Bulletin of the Institute of Mathematics
Academia Sinica (New Series)
Vol. **17** (2022), No. 1, pp. 53-81
DOI: 10.21915/BIMAS.2022102

CONSISTENCY RESULT OF RECURSIVE CONDITIONAL DISTRIBUTION ESTIMATE FOR DEPENDENT DATA UNDER LEFT TRUNCATION, WITH APPLICATIONS TO THE CONDITIONAL QUANTILE

FATIMA BENZIADI[1,a] AND IMANE BOUAZZA[2,b]

[1]Laboratory of Stochastic Models, Statistics and Applications, University of Saida-Dr. Moulay Tahar, P.O. Box 138, EN-NASR, 20000, Algeria.

[a]E-mail: benziadi.fatima@univ-saida.dz

[2]Laboratory of Stochastic Models, Statistics and Applications, University of Saida-Dr. Moulay Tahar, P.O. Box 138, EN-NASR, 20000, Algeria.

[b]E-mail: imen.bouazza@univ-saida.dz

Abstract

In this paper, we discuss a question that is often asked repeatedly in the context of statistical studies, namely the presence of incomplete data in the dataset. Therefore, our goal is to study the recursive nonparametric estimation of the conditional distribution function of a vectorial response valued variable Y explained by a Hilbertian random variable $X = x$, based on the double-kernel approach. And because we are always looking for more credible methods that are in line with the research methodology, then, it is well known that the recursive methods are more efficient than its nonrecursive rival. Whereas, the variable of interest Y is left truncated by another variable T, that is, the random variables Y and T are observed if and only if $Y \geq T$; otherwise nothing is observed if $Y < T$. Under general mixing conditions, we first establish its strong uniform consistency from which we deduce the ones of the conditional quantile function estimator.

1. Introduction

In Survival Analysis, we deal generally with two popular types of data of incomplete nature which are modeled via the possible presence of right-censoring and/ or left-truncation of variables in the observed sample. In form, these two latter are very similar, however, they are quite different in

Received January 15, 2022 and in revised form January 15, 2022.

AMS Subject Classification: Primary, 62G20, Secondary 62G08, 62G35, 62E20.

Key words and phrases: Nonparametric recursive estimate, conditional distribution, left truncation, Conditional quantile, functional data, $\alpha-$mixing, uniform almost sure convergence.

nature (each type of incompleteness based on its own special properties. In other words, it is important to distinguish between these two types). In fact, many scientific domains are known to have this type of information jointly. For example, some biomedical studies suffer simultaneously of truncation and censoring, this is the situation in astronomy also. So that, the survival theory is a branch of statistics that tries to bring a solution to these different situations.

In this contribution, the case that we are focusing on is the second one. Although, in recent decades, we have observed in scientific research many published papers which make use such data because of the importance of their economic and social impacts. Thus, the development of new techniques to better model and exploit these latter whose occurrence is not known, to insert it in nonparametric statistical methods, because from an applied point of view, the truncated data are considered to be more accurate and the most expressive one. Moreover, compared with the censoring, truncation uses less information, then; the available statistical methods for use are limited. Its theoretical foundation was formally established by Woodroofe (1985)[23], whose idea dates back to Lynden-Bell (1971)[17]. As a complement to these works, the first to propose a non-parametric estimator of the truncation probability for randomly truncated data are He and Yang (1998)[12].

Literally, in the existence of these variables, various authors have provided several regression models that realize the characteristics of such data for i.i.d. and mixing cases, including O. Said and Lemdani (2006)[19], O. Said and Tatachak (2009)[20], Wang et al. (2012a)[22], Altendji et al. (2018)[1], Derrar et al. (2015)[6]. Thus, the authors Lemdani et al. (2009)[18], and later Helal and O. Said (2016)[13] have established the asymptotic properties of the conditional quantile estimates of which are robust means of forecasting.

Under weak dependence in the sense of [7], our idea in this work is to estimate recursively for the functional context, the nonparametric conditional distribution function from the case of complete data to the one of randomly truncated data on the left. So that we will establish its uniform almost sure convergence. Subsequently, using this family of estimators to derive the uniform consistency of the conditional quantile estimators. Two based models of estimators are mainly adopted; in one, a recursive kernel estimation is introduced recently by Benziadi et al. (2016)[2], the latter is studied when there are no truncation observations. The other is then presented by Hellal and O. Said (2016)[13] that extends the classical estimator of Firraty et al. (2006)[9] to the truncation setting for i.i.d. observations.

The rest of the article is divided into five sections. Section 2 specifies a unified framework for analyzing random left-truncated (RLT) type of data and the estimation procedures for our nonparametric conditional model, respectively. Note also that the latter results are used to derive some asymptotic properties of the recursive estimator of the conditional quantile function. The next section sets the assumptions used with the main results. In addition, Section 4 presents a particular case in addition to a general discussion in Section 5, while Section 6 contains all the technical proofs of Lemmas.

2. Nonparametric Estimation

2.1. Description of the functional randomly-truncated framework

As a prior knowledge that truncation does not allow the application of ordinary statistical techniques, we then begin with a reminder of some basic structures and bibliographies corresponding to this context. In the population of interest, let T denote the positive random left-truncation time. Assume that T has a support on $[0, a_F]$ and has an unknown distribution function G. This framework can be based mainly on the truncation probability defined for the two observable pairs of variables Y and T, by

$$\tau := \mathbb{P}[Y \geq T] = \int G(v)F(dv) > 0$$

and since the construction of the distribution F (resp G) of Y (resp T) has been reformulated in terms of the size n which we will talk about later, we must be aware that their joint distribution is also changeful (see Stute (1993)[21]), such that

$$H^*(y,t) = \tau^{-1} \int_{-\infty}^{y} G(t \wedge v)F(dv)$$

with the marginal ones which depend on this latter, that generate the distribution of the positive data Y and T respectively

$$F^*(y) := \tau^{-1}\int_{-\infty}^{y} G(v)F(dv) \quad \text{and} \quad G^*(t) := \tau^{-1}\int_{-\infty}^{\infty} G(t\wedge v)F(dv)$$

and thus

$$K(y) = G^*(y) - F^*(y) := \tau^{-1}G(y)\overline{F}(y)$$

with their empirical estimators defined by

$$F_n^*(y) = n^{-1}\sum_{k=1}^{n} \mathbb{I}_{(Y_k\le y)} \quad \text{and} \quad G_n^*(t) = n^{-1}\sum_{k=1}^{n} \mathbb{I}_{(T_k\le t)}$$

and the consistent estimator of $K(y)$ for $a_F \le y < +\infty$ given by

$$K_n(y) = n^{-1}\sum_{k=1}^{n} \mathbb{I}_{(T_k\le y\le Y_k)}$$

where $\mathbb{I}_A$ denotes the indicator function of the event A. For the random left-truncation model, similar to the nonparametric Kaplan-Meier estimator (NPKME) for censored data, the astrophysicist Lynden-Bell (1971)[17] has proposed the unique nonparametric estimator (NPLBE) based on maximum likelihood (ML) of the continuous functions F and G expressed as

$$F_n(y) = 1 - \prod_{s\le y}\left[1 - \frac{F_n^*(s)}{K_n(s)}\right] \quad \text{and} \quad G_n(t) = 1 - \prod_{s>t}\left[1 - \frac{G_n^*(s)}{K_n(s)}\right].$$

Note that the KME and LBE always give a valid redistribution of the upper limits, though the result may not be applicable in wider context. In addition, we will set the identifiability conditions on the support of F and G,

$$a_G \le a_F; \quad b_G \le b_F \quad \text{and} \quad \int_{a_F}^{\infty} \frac{1}{G}dF < \infty.$$

In which, the main asymptotic properties of the later estimates, including the weak and strong uniform convergence with rates of convergence, have been provided for example in the paper by the statistician Woodroofe (1985)[23], such that

$$\sup_{y \geq a_F} |F_n(y) - F(y)| \overset{P.a.s.}{\longrightarrow} 0 \quad \text{and} \quad \sup_{t \geq a_G} |G_n(t) - G(t)| \overset{P.a.s.}{\longrightarrow} 0$$

with a simpler form for the estimator of τ

$$\widehat{\tau}_n := \frac{G_n(y)\overline{F}_n(y)}{K_n(y)}$$

For this, in some references, $F_n(y)$ and $G_n(t)$ called the
Lynden-Bell-Woodroofe estimators (NPLBWE).

Remark 1. To be more precise, we note that the strong uniform consistency for the improved product limit estimator of the distribution function F over $[a_F, \infty)$ was proved under the only condition $a_F > a_G$. However, in complementary case ($a_F \leq a_G$), the desired asymptotic property does not achieved (here, the interested reader can referred directly to the paper by Chen et al. (1995) [5], who gave a comprehensive review of all the other possible cases with rich discussions that show us when we can obtain this property under some additional necessary tools).

2.2. The model and the estimate under strong mixing hypothesis and left truncation

In order to simplify and give a great flexibility for our framework and to focus on the main interest of our paper, Let us consider in the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an infinite stationary dependent random vectors $\{(X_k, Y_k), k = 1, \ldots, N\}$ drawn from the pair (X, Y), where X is the random covariate taking its values in a distanced Hilbertian space $(\mathcal{H}, d)$, Y is the interest random vectorial variable with continuous distribution function (df) F.

We consider in this model the scenario where the response variable Y assumed to be subject to truncation time T. The sample size N is fixed, but unknown. For that, in the paper's continuation, among the total number in the pooled sample N, we base only on the observed data $\{(X_k, Y_k, T_k), k = 1, \ldots, n\}$ with the conventions $n \leq N$ (n is known) and $\mathbb{P}\left[n/N \to \tau\right] = 1$.

Thus, the conditional distribution function (df) of Y given the covariate $X = x$ under the left-truncation condition exists and is often defined by

$$F_{Y/X}(y/x) = \mathbb{E}\left[\mathbb{I}_{(Y\leq y)}/X = x\right], \ \forall y \in \mathbb{R}^p.$$

Turning to the desired goal in this paper, first, a recent modification of Ferraty and al's estimator (2006)[9] in the case of non-truncated data has been introduced by Benziadi et al. (2016)[2] to estimate recursively the non-parametric conditional distribution function (cdf), they posed the following estimator

$$\ddot{F}_n^x(y) = \frac{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}(y - Y_k)\right)}{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)}. \tag{1}$$

A quick glance at the work of these latter authors shows that the estimator given above has good theoretical and practical properties when the data are assumed to be ergodic.

Now, in the same footsteps, in the case of truncated data, our purpose is to introduce the version of a recursive double kernel estimator of the model given above denoted $\widehat{F}_n^x(.)$ and defined as follow

$$\widehat{F}_n^x(y) = \frac{\sum_{k=1}^{n} G_n^{-1}(Y_k) L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)}{\sum_{k=1}^{n} G_n^{-1}(Y_k) L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)} = \frac{\widehat{\Psi}_n(x, y)}{\widehat{\Upsilon}_n(x)} \tag{2}$$

where

$$\widehat{\Psi}_n(x, y) = \frac{\widehat{\tau}_n}{n\psi_n(x, a_n)} \sum_{k=1}^{n} \frac{1}{G_n(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)$$

and

$$\widehat{\Upsilon}_n(x) = \frac{\widehat{\tau}_n}{n\psi_n(x, a_n)} \sum_{k=1}^{n} \frac{1}{G_n(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)$$

with $\Psi(\cdot,\cdot)$ is the joint probability function assumed to be bounded, $\Upsilon(\cdot)$ is the marginal one, the functions L_1 and L_2 are kernels and a_k, b_k are two positive real numbers tending to 0 as n goes to infinity.

3. Assumptions and Asymptotic Results

In order to prove that our estimate achieves the asymptotic properties, We first use the notations often introduced in many studies, $\wp_k$ the $\sigma-$field generated by $\{(X_s, Y_s);\ 1 \leq s < k\}$ and $\mathcal{B}_k$ the one generated by $\{(X_s, Y_s), (X_r), 1 \leq s < k;\ k \leq r \leq k+1\}$. Thus, let $\mathcal{S}$ and Ω be respectively two compact sets of $\mathcal{H}$ and $\mathbb{R}^p$. On the other hand, following Woodroofe (1985)[23] and He and Yang (1998)[12], let us note for the distribution function L of Z, the lower and upper boundaries of the support by

$$a_L = \inf\{z : L(z) > 0\} \quad \text{and} \quad b_L = \sup\{z : L(z) < 1\}.$$

Next, to simplify the presentation of our main results and their proofs, some important assumptions are assumed to be hold.

(**U.1**) On the hilbertian variable: there is a ball B of radius $a_k > 0$ centered at x such that

- **(i)** $\forall\ x \in \mathcal{S},\ 0 < \phi(x, a_k) \leq \mathbb{P}[X \in B(x, a_k)]$ and $\phi(x, a_k) \to 0$ as $h \to 0$;
- **(ii)** The joint density exists, is bounded and satisfies

$$0 < \sup_{k \neq l} \mathbb{P}\left[X_k \in B(x, a_k), X_l \in B(x, a_l)\right] = O\left\{\frac{(\phi(x, a_k))^{1+\gamma_1}}{n^{\gamma_1}}\right\}.$$

(**U.2**) $(X_k, Y_k)_{k \in \mathbb{N}}$ is a stationary sequence of $\alpha-$ dependent real-valued random variables whose coefficients of mixture $\alpha(n)$ satisfy the condition

$$\exists a,\ c \in \mathbb{R}_+^* :\ \forall n \in \mathbb{N},\ \alpha(n) = O(n^{-1/\gamma_1}).$$

(**U.3**) On the nonparametric model: $\forall (y_1, y_2) \in \Omega^2$, $\forall\ (x_1, x_2) \in \mathcal{N}_x^2$, $\xi^x(z)$ satisfies the Lipschitz condition

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C_1 \left(d_{\mathcal{H}}^{\nu_1}(x_1, x_2) + \|y_1 - y_2\|_{\mathbb{R}^p}^{\nu_2}\right),$$

with $C_1 > 0$, $\nu_1 > 0$, $\nu_2 > 0$.

(**U.4**) L_1 is a measurable non-negative continuous bounded kernel on its compact support $(0,1)$. Also, it is supposed to be Hölderian of order $\beta_1 = 1$ such that

$$|L_1(x_1) - L_1(x_2)| \leq C_2 d_{\mathcal{H}}(x_1, x_2)$$

(**U.5**) L_2 is an increasing, continuous and bounded distribution function satisfying:

(i) $\forall (y_1, y_2) \in \mathcal{I}^2$, $|L_2(y_1) - L_2(y_2)| \leq C_3 \|y_1 - y_2\|_{\mathbb{R}^p}$,

$$\int_{t\in\mathbb{R}^p} L_2^{(1)}(\|t\|_{\mathbb{R}^p})dt = 1, \quad \text{and} \quad \int_{t\in\mathbb{R}^p} \|t\|_{\mathbb{R}^p}^{\nu_2} L_2^{(1)}(\|t\|_{\mathbb{R}^p})dt < \infty;$$

(ii) There exists a continuous bounded function $l_\infty(\cdot)$ in the neighborhood of x such that the conditional distribution of the couple (Y_k, Y_l) knowing (X_k, X_l) exists and verifies

$$\max\left[F(y_k/x_k), F_{k,l}(y_k, y_l/x_k, x_l)\right] \leq l_\infty(x) < \infty.$$

(**U.6**) On the bandwidths: a_k and b_k satisfy the following conditions:

(i) $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$ and $\lim_{n\to\infty} n^r b_n = \infty$ for any $r > 0$;

(ii) $\sum_{k=1}^{n} \phi_k(x, a_k) = n\psi_n(x, a_n) \to \infty$ and $\lim_{n\to\infty} \frac{\log n}{n\psi_n(x, a_n)} = 0$;

(iii) $\exists\, \gamma > 0$; $\frac{1}{n^\gamma \log n} \sum_{k=1}^{n} b_k^{-1} \to 0$ as $n \to \infty$.

(**U.7**) The variables $(T_k)_{k=1,\ldots,n}$ are independent of $(Y_k)_{k=1,\ldots,n}$.

3.1. Discussion of the assumptions

Generally, in nonparametric classical and/ or recursive estimation for conditional distribution functions in $\alpha-$mixing context (dependent processes), which have been adopted by Doukhan (1994)[7], all the assumptions used in this paper are necessary.

The assumption (U.1)(i) is a standard condition for functional estimate. While, (U.1)(ii) is the same as used in Ferraty et al. (2005)[8] among which the small-ball probability satisfies

$$\sup_{k\neq l} \frac{\mathbb{P}\left[X_k \in B(x,a_k), X_l \in B(x,a_l)\right]}{\mathbb{P}[X \in B(x,a_k)]} = O\left\{\left(\frac{\phi(x,a_k)}{n}\right)^{\gamma_1}\right\}. \tag{3}$$

Compared with Theorem (4.1) in Hellal and O. Said (2016)[13] for the independent framework in which they used the classical Bernstein exponential inequality for the classical kernel estimate. In the case of dependent observations, when the process (X_k, Y_k) has algebraically decreasing mixing coefficients $\alpha(n)$, we should to set the condition (U.2) in order to use the adapted Fuk-Nagaev inequality, then, to study the consistency of the estimator. While, statisticians see that the dependency structure is more complex than the previous one and has many practical applications.

On one hand, we introduce the regularity condition (U.3), defining the Hölderian property of the continuous conditional distribution which makes the proof's steps easier and enables us to obtain the rates of convergence. Moreover, the hypotheses (U.4), (U.5) are considered as classical assumptions of kernel estimation which are necessary, sufficient and always keep track of the above condition (U.3) in terms of function's class, as well as on the conditional distribution. Furthermore, (U.6) is an important technical condition on the sequences a_n and b_n, however, rather classic in recursive kernel estimation.

On the other hand, depending on the difficulty of the problem to treat properties of the proposed estimator when the sample contains truncated data. Ideally, we point out that the truncation mechanism would be examined by the assumption (U.7) which is considered as a powerful tool in nonparametric truncation estimation in the sense it gives valid solution.

Now, we are in position to state our main theoretical results.

3.2. Uniform almost sure convergence rates of the conditional distribution function

We first establish the rate of the uniform consistency, which is the object of the following theorem. Throughout the rest of the paper, $K_i, i = 1, \ldots, 7$ will be used to denote the positive constants whose value may vary, in addition to the previous constants mentioned above.

Theorem 1. *Suppose that the assumptions* (U.1)−(U.7) *hold true. For n large enough, we have*

$$\limsup_{n\to\infty} \sup_{x\in\mathcal{S}} \sup_{y\in\Omega} \frac{\left|\widehat{F}_n^x(y) - F^x(y)\right|}{\left((a_n^{\nu_1} + b_n^{\nu_2}) + \left(\dfrac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)} \leq K_1 \quad a.s.$$

The application of Theorem 1 is needed for deriving the following result.

3.3. Uniform almost sure convergence rates of the conditional quantile function

Considering that the conditional quantile estimator depends on the construction of the conditional distribution function estimator. Thus, its uniform consistency depends basically on that of the previous ones. It consists beforehand to assume that $F^x(\cdot)$ is strictly increasing and continuous in order to ensure the existence and the uniqueness of the conditional quantile function.

Our focus now is directed to the conditional quantile $q_\alpha(x)$ that naturally estimated by

$$\widehat{q}_{\alpha,n}(x) = \widehat{F}_n^{-1}(\alpha/x) = \inf\left\{y : \widehat{F}_n(y/x) \geq \alpha\right\}$$

We then have to introduce additional condition

(**U.8**) For each fixed $\alpha \in (0,1)$, the function $q_\alpha(x)$ satisfies that, for any $\epsilon > 0$ and $\eta_\alpha(x)$, there exists a $\beta > 0$ such that $\sup_{x\in\mathcal{S}} |q_\alpha(x) - \eta_\alpha(x)| \geq \epsilon$ implies that $\sup_{x\in\mathcal{S}} |F^x(q_\alpha(x)) - F^x(\eta_\alpha(x))| \geq \beta$.

Corollary 1. *Let the assumptions of Theorem* 1 *hold. In addition to* (U.8), *then, we have*

$$\limsup_{n\to\infty} \sup_{x\in\mathcal{S}} \frac{|\widehat{q}_{\alpha,n}(x) - q_\alpha(x)|}{\left((a_n^{\nu_1} + b_n^{\nu_2}) + \left(\dfrac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)} \leq K_2 \quad a.s.$$

***Proof of Theorem* 1.** The proof techniques based mainly on the following standard decomposition

$$\widehat{F}_n^x(y) - F^x(y) - \widehat{B}_n(x,y) = \frac{1}{\widehat{h}_n(x)}\Big\{\widehat{Q}_n(x,y) - \widehat{B}_n(x,y)\Big[\left(\widehat{\Upsilon}_n(x) - \widetilde{\Upsilon}_n(x)\right) + \left(\widetilde{\Upsilon}_n(x) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right)\Big]\Big\}$$

with

$$\begin{aligned}\widehat{Q}_n(x,y) \;:=\; & \left[\left(\widehat{\Psi}_n(x,y) - \widetilde{\Psi}_n(x,y)\right) + \left(\widetilde{\Psi}_n(x,y) - \mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right]\right)\right] \\ & - F^x(y)\left[\left(\widehat{\Upsilon}_n(x) - \widetilde{\Upsilon}_n(x)\right) + \left(\widetilde{\Upsilon}_n(x) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right)\right]\end{aligned}$$

and

$$\widehat{B}_n(x,y) := \frac{\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right] - F^x(y)\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]}{\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]}$$

where

$$\widetilde{\Psi}_n(x,y) = \frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^{n}\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)$$

and

$$\widetilde{\Upsilon}_n(x) = \frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^{n}\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)$$

in addition to

$$\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right] = \frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^{n}\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]$$

and

$$\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right] = \frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^{n}\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right]$$

Then, the proof of Theorem 1 is a direct consequence of Lemmas 1−5 below extending several results to the left-truncation setting.

Lemma 1. *Under the assumptions* (U.1), (U.2) *and* (U.3)−(U.6)*, we have*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\sup\frac{\left|\widetilde{\Psi}_n(x,y)-\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right]\right|}{\left(\dfrac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}}\leq K_3 \quad a.s.$$

Lemma 2. *Under the assumptions* (U.2), (U.4), (U.5) *and* (U.7) *one get*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\sup\frac{\left|\widehat{\Psi}_n(x,y)-\widetilde{\Psi}_n(x,y)\right|}{\left(n^{-1/2}\right)}\leq K_4 \quad a.s.$$

Lemma 3. *Assume that* (U.1) *and* (U.4) *hold true, for any* $x\in S$*, we have*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup\frac{\left|\widetilde{\Upsilon}_n(x)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right|}{\left(\dfrac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}}\leq K_5 \quad a.s.$$

Lemma 4. *Assume that* (U.2), (U.4) *and* (U.7) *hold true, for any* $x\in S$*, one get*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup\frac{\left|\widehat{\Upsilon}_n(x)-\widetilde{\Upsilon}_n(x)\right|}{\left(n^{-1/2}\right)}\leq K_6 \quad a.s.$$

Lemma 5. *Under the assumptions* (U.1), (U.3), (U.4) *and* (U.6)*, we have*

$$\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\left|\widehat{B}_n(x,y)\right|\leq K_7\left(a_n^{\nu_1}+b_n^{\nu_2}\right) \quad a.s.$$

***Proof of Corollary* 1.** It is easy to see that Corollary 1 can be deduced from the relation

$$\sup_{x\in\mathcal{S}}|F^x(\widehat{q}_{\alpha,n}(x))-F^x(q_\alpha(x))|\leq 2\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\left|\widehat{F}_n^x(y)-F^x(y)\right|$$

which is based primarily on the decomposition (4).

4. The Real Valued Response Data Case

We have previously studied the strong consistency of our estimator when the random variable of interest Y is of vector nature. It remains to treat the particular case where this variable is real (i.e. $d = 1$). In this case, some current assumptions lose their usefulness and they are modified to fit the situation considered

(R.1) The conditional distribution function $F^x(\cdot)$ is such that: $\forall y = (y_1, y_2) \in \mathbb{R}$, $\exists\ \beta_1 > 0, \beta_2 > 0$ and $C_1 > 0$

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C_1 \left(d_{\mathcal{H}}^{\beta_1}(x_1, x_2) + |y_1 - y_2|^{\beta_2} \right).$$

(R.2) The kernel function L_2 will be supposed to satisfy the following conditions: $\forall (y_1, y_2) \in \mathcal{I}^2$, $|L_2(y_1) - L_2(y_2)| \leq C_3 |y_1 - y_2|$,

$$\int_{t\in\mathbb{R}} L_2^{(1)}(|t|)dt = 1 \quad \text{and} \quad \int_{t\in\mathbb{R}} |t|^{\beta_2} L_2^{(1)}(|t|)dt < \infty.$$

We will not repeat here the proofs which are the same as for the previously studied case and the result remains the same too, such that

Corollary 2. *Maintaining the same assumptions used in Theorem* 1 *and replacing* (U.3) *by* (R.1) *and* (U.5) *by* (R.2), *one have*

$$\limsup_{n\to\infty} \sup_{x\in\mathcal{S}} \sup_{y\in\Omega} \frac{\left|\widetilde{F}_n^x(y) - F^x(y)\right|}{\left(\left(a_n^{\beta_1} + b_n^{\beta_2}\right) + \left(\dfrac{\log n}{n\psi_n(x, a_n)}\right)^{1/2}\right)} \leq K_1' \quad a.s.$$

4.1. The L^1 recursive estimate

For $x \in \mathcal{H}$, the L^1 estimator of the conditional probability distribution of Y given $X = x$ is given as follows

$$\overline{F}_n^x(y) = \frac{\sum_{k=1}^{n} G_n^{-1}(Y_k) L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) \mathbb{I}_{(-\infty, y)}(Y_k)}{\sum_{k=1}^{n} G_n^{-1}(Y_k) L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)} := \frac{\overline{\Psi}_n(x, y)}{\overline{\Upsilon}_n(x)} \tag{4}$$

where $\mathbb{I}_A$ denotes the indicator function of the set A.

Theorem 2. *Under the assumptions* (U.1), (U.4), (U.6) *and* (R.1), *one have*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\frac{\left|\overline{F}_n^x(y)-F^x(y)\right|}{\left((a_n^{\nu_1})+\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)}\le M_1 \quad a.s.$$

The proof of this theorem is based on the main following results.

Lemma 6. *Let Assumptions of Theorem* 2 *hold true. Then*

(i) $$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\frac{\left|\ddot{\Psi}_n(x,y)-\mathbb{E}\left[\ddot{\Psi}_n(x,y)\right]\right|}{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}}\le M_2 \quad a.s.$$

(ii) $$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\frac{\left|\overline{\Psi}_n(x,y)-\ddot{\Psi}_n(x,y)\right|}{\left(n^{-1/2}\right)}\le M_3 \quad a.s.$$

Lemma 7. *Let the assumptions* (U.1), (U.4) *and* (U.6) *hold. Then, one have*

(i) $$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\frac{\left|\ddot{\Upsilon}_n(x)-\mathbb{E}\left[\ddot{\Upsilon}_n(x)\right]\right|}{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}}\le M_4 \quad a.s.$$

(ii) $$\lim_{n\to\infty}\sup_{x\in\mathcal{S}}\frac{\left|\overline{\Upsilon}_n(x)-\ddot{\Upsilon}_n(x)\right|}{\left(n^{-1/2}\right)}\le M_5 \quad a.s.$$

Lemma 8. *Under the same assumptions as those of Lemma* 5, *then, we have*

$$\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\left|\widetilde{B}_n(x,y)\right|\le M_6 a_n^{\nu_1} \quad a.s.$$

5. General Discussion

The principal purpose of this section is to discuss some previous results related to ours. Indeed, the sensitivity of a kernel estimator to the presence of incomplete data is a subject discussed several times in recent literature and is confirmed by several statisticians for many conditional models such as, regression, conditional density, mode, distribution and quantile functions.

We quote for this case, for a finite sample size N, the work of Guessoum and Hamrani(2014)[11] whose main interest is to compare the behavior of the relative truncated regression kernel function estimator, when the observations are supposed to be independent, $\alpha-$mixing and associated respectively and they have clearly shown on the one hand that the quality of the estimator is affected much more by the sample size $n \leq N$ than by the fixed truncation rate $\alpha = 70\%$ and that their estimator performs well in the mixing case since it is the most expressive case of the information state. On the other hand, they noted that their results are slightly better than that obtained in the independent case by Ould Saïd and Lemdani (2006)[19].

Helal and Ould Saïd(2016)[13] also obtained some important simulation results about kernel conditional quantile estimator so that they lent further support to their theoretical results and assess the performance of the estimator for discrete time processes with values in functional spaces with different truncation sizes. Therefore, they proved clearly that:

- The mean squared error is gradually decreasing (MSE= 0.20, 0.15, 0.12) with a rise of $n = 100, 300, 500$ respectively.
- The quality of the estimate deteriorates with the increase of the percentage of truncation (TR= 0%, 12%, 32%, 66%).

While, the effectiveness of the recursive estimator in such case of the presence of incomplete data (for instance the censored data) is the object of a work of Bouazza et al. (2021)[4] by considering a certain type of dependent observations. Thus, they used this robust approach to study the conditional mode function with the values of the explanatory variables X taken in a semimetric space and they dealt with the following remarks:

- The mean squared error decreases (MSE= 1.322, 1.149, 0.850) when the sample size increases ($n = 200, 400, 600$), so the quality of the estimator is better for high observed n.
- The estimator that includes incomplete data performs slightly less than the ones in complete case, when the censoring rates increases (CR= 20%, 40%, 60%).

Thus, approximately, similar computations also allow to obtain the same results. Although the recursive estimator has very excellent properties (since it is the one for which the asymptotic variance is small) compared to the classical estimator, but, it is more affected by the presence of these data.

6. Technical Proofs

In the following, let's pose $\psi_n(x,a_n) = \mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_1)\right)\right]$. Then, the proof of Theorem 1, is based essentially on the following lemmas adapted to the $\alpha-$mixing context.

Lemma 9 (Fuk-Nagaev). *Let $\{Z_k, k=1,\dots,n\}$ be a sequence of centered real random variables, of alpha-mixing coefficient $\alpha(n)=O(n^{-a})$ for $a>1$, checking for all $n\in\mathbb{N}$, $|Z_k|<\infty, k=1,\dots,n$. Then, for all $q>1$*

$$\mathbb{P}\left(\left|\sum_{k=1}^{n} Z_k\right| > 4\theta\right) \leq 4\left(1+\frac{\theta^2}{qS_n^2}\right)^{-q/2} + \frac{2nc}{q}\left(\frac{2q}{\theta}\right)^{a+1},$$

where $S_n^2 = \sum_{1\leq k,l\leq n} \left|Cov\left(Z_k,Z_l\right)\right|.$

Lemma 10 (O. Saïd and Tatachak (2009)[20]). *Under assumption* (U.2) *of mixing random variables, we have*

$$|\widehat{\tau}_n - \tau| = O\left\{n^{-1/2}(\log_2 n)^{1/2}\right\}$$

The interested reader can refer to the original article for more details and proof of this lemma, for the case of strong dependency.

***Proof of Lemma* 1.** We start by noting for all couple $(x,y)\in\mathcal{S}\times\Omega$ that

$$\begin{aligned}
&\widetilde{\Psi}_n(x,y) - \mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right] \\
&=\frac{1}{n}\sum_{k=1}^{n}\frac{\tau}{\psi_n(x,a_n)}\left\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right.\\
&\quad\left.-\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\right\}\\
&=\frac{1}{n}\sum_{k=1}^{n}Z_{k,n}(x,y).
\end{aligned}$$

with

$$Z_{k,n}(x,y) = \frac{\tau}{\psi_n(x,a_n)}\left\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right.$$
$$\left. -\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\right\}.$$

The use of the boundedness of L_1 from the assumption (U.4) ensures that for two existent constants $(m_1,m_2)\in\mathbb{R}_+^2$, we would have

$$0 < m_1\phi(x,a_k) \leq \mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right] \leq m_2\phi(x,a_k). \tag{5}$$

In view of the following quantity, since the condition $\mathbb{I}_{(T_k\leq Y_k)}=1$ is always validated in the left truncated model by definition of the probability τ. Then, by condition (5) and applying the assumption (U.5)(ii), we can write it in its simplest form, thus

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\right|\\
&=\left|\mathbb{E}\left[\mathbb{E}\Big[\frac{1}{G(Y_k)}L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\mathbb{I}_{(T_k\leq Y_k)}/X_k\Big]L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right]\right|\\
&\leq \mathbb{E}\left[\frac{1}{G(Y_k)}\mathbb{E}\left[L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)/X_k\right]\mathbb{P}[T_k\leq Y_k]L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right]\\
&\leq l_\infty(x)\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right]\\
&\leq l_\infty(x)m_2\phi(x,a_k).
\end{aligned}$$

and so, we employ the decomposition

$$\begin{aligned}
\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\left|\sum_{k=1}^n Z_{k,n}(x,y)\right| \leq{}& \underbrace{\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\left|\sum_{k=1}^n Z^*_{k,n}(x_i,y)\right|}_{A_1} + \underbrace{\max_{j\in\{1,\ldots,r_n\}}\sup_{x\in\mathcal{S}}\left|\sum_{k=1}^n \widetilde{Z}_{k,n}(x,y)\right|}_{A_2}\\
&+\underbrace{\max_{i\in\{1,\ldots,h_n\}}\max_{j\in\{1,\ldots,r_n\}}\left|\sum_{k=1}^n Z_{k,n}(x_i,y_j)\right|}_{A_3}.
\end{aligned}\tag{6}$$

Therefore, the compactness property of the two subsets Ω and $\mathcal{S}$ help us to write for any $y_1, y_2, \ldots, y_{r_n}$ and $x_1, x_2, \ldots, x_{h_n}$,

$$\Omega \subset \bigcup_{j=1}^{r_n} \mathcal{B}(y_j, s_n) \quad \text{and} \quad \mathcal{S} \subset \bigcup_{i=1}^{h_n} \mathcal{B}(x_i, s_n).$$

Thus, we can take for a constant M, $s_n \leq Mn^{-\beta}$ with $(\beta > 0)$ and $j(y) = \arg \min_{j\in\{1,\ldots,r_n\}} \|y - y_j\|_{\mathbb{R}^p}$ and $h(x) = \arg \min_{i\in\{1,\ldots,h_n\}} d_{\mathcal{H}}(x, x_i)$.
For the first term of the decomposition (6), we have for any $(x, y) \in \mathcal{S} \times \Omega$

$$\begin{aligned}\left|\sum_{k=1}^{n} Z^*_{k,n}(x_i, y)\right| \leq& \left|\widetilde{\Psi}_n(x, y) - \widetilde{\Psi}_n(x_i, y)\right| + \left|\mathbb{E}\left[\widetilde{\Psi}_n(x_i, y)\right] - \mathbb{E}\left[\widetilde{\Psi}_n(x, y)\right]\right| \\ \leq& \frac{\tau}{\psi_n(x, a_n)} \sum_{k=1}^{n} \frac{L_2\left(b_k^{-1}\|y_- Y_k\|_{\mathbb{R}^p}\right)}{G(Y_k)} \Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) \\ &- L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right)\Big| \\ &+ \frac{\tau}{\psi_n(x, a_n)} \sum_{k=1}^{n} \mathbb{E}\Big[\frac{L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)}{G(Y_k)} \Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) \\ &- L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)\Big|\Big].\end{aligned}$$

Using the fact that the kernel L_1 is of Lipschitz class. Then, one get

$$\begin{aligned}A_1 \leq\ & C_2 \frac{\tau}{G(a_F)\psi_n(x, a_n)} \sup_{x\in\mathcal{S}} \sup_{y\in\Omega} \sum_{k=1}^{n} \frac{d_{\mathcal{H}}(x, x_i)}{a_k} \left|L_2\left(b_k^{-1}\|y_- Y_k\|_{\mathbb{R}^p}\right)\right| \\ &+ C_2 \frac{\tau}{G(a_F)\psi_n(x, a_n)} \sup_{x\in\mathcal{S}} \sup_{y\in\Omega} \sum_{k=1}^{n} \frac{d_{\mathcal{H}}(x_i, x)}{a_k} \mathbb{E}\left[L_2\left(b_k^{-1}\|y_- Y_k\|_{\mathbb{R}^p}\right)\right]\end{aligned}$$

and since for all $s_n = n^{-\beta}$, it follows that $A_1 \longrightarrow 0 \;\; as \;\; n \to \infty$. Then, for the study of A_2, we first write the following decomposition which leads to

$$\begin{aligned}A_2 =& \max_{j\in\{1,\ldots,r_n\}} \sup_{x\in\mathcal{S}} \left|\widetilde{\Psi}_n(x_i, y) - \widetilde{\Psi}_n(x_i, y_j)\right| \\ &+ \max_{j\in\{1,\ldots,r_n\}} \sup_{x\in\mathcal{S}} \left|\mathbb{E}\left[\widetilde{\Psi}_n(x_i, y_j)\right] - \mathbb{E}\left[\widetilde{\Psi}_n(x_i, y)\right]\right| \\ \leq& \max_{j\in\{1,\ldots,r_n\}} \sup_{x\in\mathcal{S}} \sum_{k=1}^{n} \frac{\tau}{\psi_n(x, a_n)} \left\{ \frac{1}{G(Y_k)} \Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)\right.\end{aligned}$$

$$
\begin{aligned}
& - L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\Bigg|\Bigg\} \\
+ & \max_{j\in\{1,\dots,r_n\}} \sup_{x\in\mathcal{S}} \sum_{k=1}^{n} \frac{\tau}{\psi_n(x,a_n)} \Bigg\{\mathbb{E}\Bigg[\frac{1}{G(Y_k)}\Bigg| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right) \\
& - L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)\Bigg|\Bigg]\Bigg\} \\
= & A_2^{(1)} + A_2^{(2)}
\end{aligned}
$$

so that under the assumptions (U.5) and (U.6)(iii), we have

$$
\begin{aligned}
A_2^{(1)} &\leq C_3 \frac{\|y - y_j\|_{\mathbb{R}^p}}{\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{\tau}{b_k G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) \\
&\leq C_3 \frac{\tau s_n}{G(a_F)\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right)}{b_k} \\
&\leq \frac{M_1 n^{-\gamma}}{\psi_n(x,a_n)} \sum_{k=1}^{n} b_k^{-1} \\
&\leq \frac{M_1 \log n}{\psi_n(x,a_n)} \frac{1}{n^{\gamma}\log n} \sum_{k=1}^{n} b_k^{-1} \longrightarrow 0 \quad as \quad n\to\infty
\end{aligned} \tag{7}
$$

for the second term of the decomposition, the same arguments as for $A_2^{(1)}$ with condition (5) lead as n goes to infinity to

$$
A_2^{(2)} \leq C_3 \Big[\psi_n(x,a_n)\Big]^{-1} \sum_{k=1}^{n} \frac{\|y_j - y\|_{\mathbb{R}^p}}{b_k} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right)\right] \longrightarrow 0.
$$

We move now to the next term

$$
A_3 = \max_{i\in\{1,\dots,h_n\}} \max_{j\in\{1,\dots,r_n\}} \Big|\sum_{k=1}^{n} Z_{k,n}(x_i,y_j)\Big|
$$

Let's calculate first

$$
S_n^2 = \underbrace{\sum_{k\neq l}\Big|Cov\Big(Z_{k,n}(x_i,y_j), Z_{l,n}(x_i,y_j)\Big)\Big|}_{S_n^{Cov}} + \underbrace{\sum_{k=l}\Big|Cov\Big(Z_{k,n}(x_i,y_j), Z_{k,n}(x_i,y_j)\Big)\Big|}_{S_n^{Var}}. \tag{8}
$$

The definition of the probability τ and because of the boundedness of the kernels L_1 and L_2, one can show that $Z_{k,n}$ really satisfies the condition $|Z_{k,n}(x_i, y_j)| < \infty$. In particular, it is bounded $\forall k \in \mathbb{N}$, such that

$$\begin{aligned}|Z_{k,n}(x_i, y_j)| \leq & \frac{\tau}{G(a_F)\psi_n(x, a_n)}\Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right) \\ & - \mathbb{E}\Big[L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\Big]\Big| \\ \leq & C\frac{\tau}{G(a_F)\psi_n(x, a_n)} \\ = & O\Big(\frac{1}{\phi(x, a_k)}\Big).\end{aligned}$$

Then, the linearity of the expectation with the standard Jensen inequality lead directly to

$$\begin{aligned}& |\mathbb{E}[Z_{k,n}(x_i, y_j)]| \\ \leq \ & 2\mathbb{E}\Big[\frac{\tau}{G(Y_k)\psi_n(x, a_n)}\Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\Big|\Big] \\ \leq \ & 2\frac{\tau}{G(a_F)}\frac{1}{\psi_n(x, a_n)}\mathbb{E}\Big[\Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\Big|\Big] \\ = \ & O(1). \qquad (9)\end{aligned}$$

and

$$\begin{aligned}& \mathbb{E}[Z_{k,n}^2(x_i, y_j)] \\ \leq \ & C\frac{1}{G(a_F)\psi_n^2(x, a_n)}\mathbb{E}\Big[\frac{\tau}{G(Y_k)} L_1^2\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2^2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\Big] \\ \leq \ & C\frac{1}{G(a_F)\psi_n^2(x, a_n)}\mathbb{E}\Big[L_1^2\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right)\mathbb{E}\Big(L_2^2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)/X_k\Big)\Big] \\ \leq \ & l_\infty(x)\frac{C}{G(a_F)}\frac{1}{\phi(x, a_k)} \leq O\Big(\frac{1}{\phi(x, a_k)}\Big). \qquad (10)\end{aligned}$$

Furthermore,

$$\begin{aligned}& |\mathbb{E}[Z_{k,n}(x_i, y_j) \cdot Z_{l,n}(x_i, y_j)]| \\ = & \Big|\mathbb{E}\Big[\frac{\tau}{G(Y_k)G(Y_l)\psi_n^2(x, a_n)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right) \\ & \times L_1\left(a_l^{-1} d_{\mathcal{H}}(x_i, X_l)\right) L_2\left(b_l^{-1}\|y_j - Y_l\|_{\mathbb{R}^p}\right)\Big]\Big| \\ \leq & C\frac{\tau}{G^2(a_F)\psi_n^2(x, a_n)}\mathbb{E}\Big[\left(L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_1\left(a_l^{-1} d_{\mathcal{H}}(x_i, X_l)\right)\right)\end{aligned}$$

$$\times \left|\mathbb{E}[L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right) L_2\left(b_l^{-1}\|y_j - Y_l\|_{\mathbb{R}^p}\right)/X_k, X_l]\right|\Big]$$
$$\leq C\frac{\tau}{G^2(a_F)\psi_n^2(x,a_n)} l_\infty(x)\mathbb{E}[L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right) L_1\left(a_l^{-1}d_{\mathcal{H}}(x_i,X_l)\right)]$$

then, by assumptions (U.1)(ii), (U.4) and condition (3), one get

$$\begin{aligned}\left|\mathbb{E}\Big[Z_{k,n}(x_i,y_j)\cdot Z_{l,n}(x_i,y_j)\Big]\right| &\leq C\frac{\tau}{G^2(a_F)\psi_n^2(x,a_n)} l_\infty(x)\left(\frac{(\phi(x,a_k))^{1+\gamma_1}}{n^{\gamma_1}}\right)\\ &\leq C\frac{\tau}{G^2(a_F)} l_\infty(x)\left[\frac{\phi(x,a_k)}{n}\right]^{\gamma_1}\cdot\frac{1}{\phi(x,a_k)} \quad (11)\end{aligned}$$

in which we deduce on the one hand from (9) and (11)

$$\begin{aligned}&\Big|Cov\left(Z_{k,n}(x_i,y_j), Z_{l,n}(x_i,y_j)\right)\Big|\\ &\leq \left|\mathbb{E}\Big[Z_{k,n}(x_i,y_j)\cdot Z_{l,n}(x_i,y_j)\Big]\right| + \Big(\mathbb{E}\big(Z_{k,n}(x_i,y_j)\big)\Big)^2\\ &= C\Big\{\Big(\frac{\phi(x,a_k)}{n}\Big)^{\gamma_1}\cdot\frac{1}{\phi(x,a_k)}\Big\}+1. \quad (12)\end{aligned}$$

In the other hand, applying the usual modified Davydov-Rio's covariance inequality for the mixing processes. $\forall k \neq l$, we have

$$\begin{aligned}|Cov\left(Z_{k,n}(x_i,y_j), Z_{l,n}(x_i,y_j)\right)| &\leq 4\|Z_{k,n}(x_i,y_j)\|\|Z_{l,n}(x_i,y_j)\|\\ &\leq C\,|k-l|^{-1/\gamma_1}. \quad (13)\end{aligned}$$

It is therefore useful to set from now on the two subsets

$$J_1 = \{(k,l)\,; 0 < |k-l| \leq \mu_n\} \quad \text{and} \quad J_2 = \{(k,l)\,; \mu_n < |k-l| \leq n-1\}$$

a combination between (12) and (13) simply allows us to have the following

$$\begin{aligned}S_n^{Cov} &= \sum\sum_{J_1} |Cov\left(Z_{k,n}(x_i,y_j), Z_{l,n}(x_i,y_j)\right)| + \sum\sum_{J_2} |Cov\left(Z_{k,n}(x_i,y_j), Z_{l,n}(x_i,y_j)\right)|\\ &\leq Cn\mu_n\left\{\left(\frac{\phi(x,a_k)}{n}\right)^{\gamma_1}\cdot\frac{1}{\phi(x,a_k)}+1\right\} + Cn^2\mu_n^{-1/\gamma_1}. \quad (14)\end{aligned}$$

For the variance term, we apply the general definition and we get

$$|Var\left[Z_{k,n}(x_i,y_j)\right]| \leq \mathbb{E}\left[Z_{k,n}^2(x_i,y_j)\right] + \mathbb{E}\left[Z_{k,n}(x_i,y_j)\right]^2$$

$$= O\Big\{\frac{1}{\phi(x,a_k)}\Big\}$$

thus,

$$S_n^{Var} = O\Big\{\frac{n}{\phi(x,a_k)}\Big\}. \tag{15}$$

It follows from (14) and (15) that,

$$S_n^2 = O\Bigg\{n\mu_n\bigg(\Big(\frac{\phi(x,a_k)}{n}\Big)^{\gamma_1}\cdot\frac{1}{\phi(x,a_k)}+1\bigg)+n^2\mu_n^{-1/\gamma_1}\Bigg\}+O\Big\{\frac{n}{\phi(x,a_k)}\Big\}. \tag{16}$$

The complementarity of this proof depends primarily on the choice of the sequence μ_n. We put $\mu_n = \Big(\frac{\phi(x,a_k)}{n}\Big)^{-\gamma_1}$, thus, we will have $S_n^2 = O\Big\{\frac{n}{\phi(x,a_k)}\Big\}$. At this stage of the proof, we use the Fuk-Nagaev inequality adapted to the $\alpha-$mixing context for $\theta = \theta_0\Big(\frac{\log n}{n\psi_n(x,a_n)}\Big)^{1/2}$ yields

$$\begin{aligned}
&\mathbb{P}\Big\{\Big|\sum_{k=1}^{n} Z_{k,n}(x_i,y_j)\Big| > 4\Big(\frac{n\theta}{4}\Big)\Big\}\\
&\leq 4\left(1+\frac{\theta^2}{qS_n^2}\right)^{-q/2} + \frac{2nc}{q}\left(\frac{2q}{\theta}\right)^{a+1}\\
&\leq 4\left(1+\frac{\frac{n^2\theta_0^2\log n}{16n\phi(x,a_k)}}{q\frac{n}{\phi(x,a_k)}}\right)^{-q/2} + \frac{2nc}{q}\left(\frac{8q\left(\frac{\log n}{n\phi(x,a_k)}\right)^{-1/2}}{n\theta_0}\right)^{a+1}\\
&\leq 4\left(1+\frac{\theta_0^2\log n}{16q}\right)^{-q/2} + \frac{2nc}{q}\left(\frac{8q\,(\phi(x,a_k))^{1/2}}{\theta_0\sqrt{n\log n}}\right)^{a+1}\\
&= I + II.
\end{aligned}$$

We see here that the preferred choice of q is $\log^2 n$, such that the first term in the right hand side is thus increased by

$$I \leq cn^{-\frac{\theta_0^2}{32}} \longrightarrow 0$$

moreover, for the second term and by the same choice of q we also find as

$n \to \infty$ the following

$$II \leq cn^{\dfrac{-1}{2(1-a)}(-4a^2+a+1)} \longrightarrow 0.$$

Hence, both (I) and (II) fall on the following result

$$\mathbb{P}\Big[A_3 > 4\Big(\frac{n\theta}{4}\Big)\Big] \leq r_n h_n \max_{i\in\{1,\ldots,h_n\}} \max_{j\in\{1,\ldots,r_n\}} \mathbb{P}\Big[\Big|\sum_{k=1}^{n} Z_{k,n}(x_i,y_j)\Big| > 4\Big(\frac{n\theta}{4}\Big)\Big]$$

which give us for an appropriate choice of θ_0

$$\sum_{n\geq 1} \mathbb{P}\Big\{\max_{i\in\{1,\ldots,h_n\}} \max_{j\in\{1,\ldots,r_n\}} \Big|\sum_{k=1}^{n} Z_{k,n}(x_i,y_j)\Big| > n\theta\Big\} < \infty. \qquad (17)$$

Proof of Lemma 2. We have form the definition of the estimators $\widehat{\Psi}$ and $\widetilde{\Psi}$

$$\begin{aligned}
&\widehat{\Psi}_n(x,y) - \widetilde{\Psi}_n(x,y) \\
&\quad = \frac{\widehat{\tau}_n}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{1}{G_n(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \\
&\qquad - \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right).
\end{aligned}$$

such that

$$\begin{aligned}
&\left|\widehat{\Psi}_n(x,y) - \widetilde{\Psi}_n(x,y)\right| \\
&\leq \frac{1}{n\psi_n(x,a_n)} \Bigg\{ |\widehat{\tau}_n - \tau| \sum_{k=1}^{n} \frac{1}{G_n(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \\
&\quad + \tau \sum_{k=1}^{n} \left|\frac{G_n(Y_k)-G(Y_k)}{G(Y_k)G_n(Y_k)}\right| L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \Bigg\} \\
&\leq \frac{1}{n\psi_n(x,a_n)} \Bigg\{ \frac{|\widehat{\tau}_n - \tau|}{G_n(a_F)} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \\
&\quad + \frac{\tau}{G_n(a_F)} \left|\frac{G_n(y)-G(y)}{G(a_F)}\right| \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \Bigg\}.
\end{aligned}$$

Thus,

$$\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}\left|\widehat{\Psi}_n(x,y)-\widetilde{\Psi}_n(x,y)\right| \leq \frac{1}{G_n(a_F)}\left\{|\widehat{\tau}_n-\tau|+\tau\frac{\sup_{y\geq a_F}|G_n(y)-G(y)|}{G(a_F)}\right\}$$
$$\sup_{x\in\mathcal{S}}\sup_{y\in\Omega}|\Psi_n^*(x,y)|$$

with

$$\Psi_n^*(x,y)=\frac{1}{n\psi_n(x,a_n)}\sum_{k=1}^{n}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)$$

Recall the Lemma 10 in O. Saïd and Tatachak (2009)[20], when the process (X_k,Y_k) has a decreasing mixing coefficients $\alpha(n)$, such that

$$|\widehat{\tau}_n-\tau|=O_{a.s.}\left\{n^{-1/2}(\log_2 n)^{1/2}\right\}$$

and the direct application of Remark 6 in Woodroofe (1985)[23] which gives

$$|G_n(a_F)-G(a_F)|=O_{a.s.}\left\{n^{-1/2}\right\}.$$

Thus, the result is an immediate consequence of what has already been mentioned.

***Proof of Lemma* 3.** Keeping the same conditions concerning the compactness of $\mathcal{S}$, almost certainly identical as in Lemma 1 and we decompose the studied quantity as follows

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right|>3\eta\right\}\\
&\quad\leq\underbrace{\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x)-\widetilde{\Upsilon}_n(x_i)\right|>\eta\right\}}_{I_1}+\underbrace{\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x_i)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right|>\eta\right\}}_{I_2}\\
&\qquad+\underbrace{\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]-\mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right|>\eta\right\}}_{I_3} \qquad (18)
\end{aligned}$$

For the first term of the decomposition (18), L_1 being a Lipschitzian kernel.

In addition, $s_n = O\{n^{-\beta}\}$; implies that

$$\begin{aligned}
&\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x)-\widetilde{\Upsilon}_n(x_i)\right| \\
&\leq \sup_{x\in\mathcal{S}}\sum_{k=1}^{n}\frac{\tau}{\psi_n(x,a_n)}\frac{1}{G(Y_k)}\left|L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)-L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right| \\
&\leq \frac{\tau}{G(a_F)\psi_n(x,a_n)}\sum_{k=1}^{n}\frac{d_{\mathcal{H}}(x,x_i)}{a_k} \\
&\leq M\frac{s_n}{\psi_n(x,a_n)}\sum_{k=1}^{n}\frac{1}{a_k}\longrightarrow 0 \quad \text{as} \quad n\longrightarrow\infty
\end{aligned}$$

and the same for I_3 : $I_3 \to 0$ as $n \to \infty$ and therefore we deal with

$$I_1 \overset{a.s.}{=} O\left\{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right\} \quad \text{and} \quad I_3 \overset{a.s.}{=} O\left\{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right\}$$

Drawing attention now to the second term I_2, we appeal again to the Bernstein's type inequality adapted to this context of dependence by taking $\eta = \eta_0\left(\dfrac{\log n}{n\psi_n(x,a_n)}\right)^{1/2} > 0$

$$\begin{aligned}
I_2 &= \mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x_i)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right|>\eta\right\} \\
&\leq \sum_{i=1}^{h_n}\mathbb{P}\left\{\left|\widetilde{\Upsilon}_n(x_i)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right|>\eta\right\} \\
&\leq h_n\max_{1\leq i\leq h_n}\mathbb{P}\left\{|\widetilde{\Upsilon}_n(x_i)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]|>\eta\right\}
\end{aligned}$$

where, we have for all $k\in\mathbb{N}$,

$$\begin{aligned}
\widetilde{\Upsilon}_n(x_i)-\mathbb{E}[\widetilde{\Upsilon}_n(x_i)] &= \frac{1}{n}\sum_{k=1}^{n}\frac{\tau}{\psi_n(x,a_n)}\left\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right. \\
&\qquad \left.-\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right]\right\} \\
&= \frac{1}{n}\sum_{k=1}^{n}\Lambda_{k,n}(x_i)
\end{aligned}$$

with

$$\Lambda_{k,n}(x_i) = \frac{\tau}{\psi_n(x,a_n)}\Big\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right) - \mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right]\Big\}.$$

For the reminder of this proof, the same steps as term (A_3) in the proof of Lemma 1 are followed, in which under assumption (U.4), one can check that $\Lambda_{k,n}$ fulfills the condition of Lemma 9, such that

$$|\Lambda_{k,n}(x_i)| \leq M\frac{\tau}{G(a_F)\phi(x,a_k)} = O\left\{\frac{1}{\phi(x,a_k)}\right\}.$$

Furthermore, we deduce that

$$|Cov(\Lambda_{k,n}(x_i),\Lambda_{l,n}(x_i))| \leq |\mathbb{E}\left[\Lambda_{k,n}(x_i)\Lambda_{l,n}(x_i)\right]| = O\left\{\frac{(\phi(x,a_k))^{\gamma_1-1}}{n^{\gamma_1}}\right\} \tag{19}$$

and that

$$|Var\left[\Lambda_{k,n}(x_i)\right]| \leq \left|\mathbb{E}\left[\Lambda_{k,n}^2(x_i)\right]\right| = O\left\{\frac{1}{\phi(x,a_k)}\right\}. \tag{20}$$

Finally, combining (19) with (20) and following some additional classical calculations, we get $S_n^2 = O\Big\{\dfrac{n}{\phi(x,a_k)}\Big\}$. Therefore, a direct application of Fuk-Nagaev exponential inequality makes it possible to deduce the proof.

***Proof of Lemma* 4.** Similarly to the proof of Lemma 2, one may follows the same lines and arguments, such that

$$\begin{aligned}\sup_{x\in\mathcal{S}}\left|\widehat{\Upsilon}_n(x) - \widetilde{\Upsilon}_n(x)\right| &\leq \frac{1}{n\psi_n(x,a_n)}\sup_{x\in\mathcal{S}}\left\{\widehat{\tau}_n\sum_{k=1}^{n}\frac{1}{G_n(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right. \\ &\left. -\tau\sum_{k=1}^{n}\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right\} \\ &\leq \frac{1}{G_n(a_F)}\left\{|\widehat{\tau}_n-\tau| + \tau\frac{\sup_{y\geq a_F}|G_n(y)-G(y)|}{G(a_F)}\right\}\sup_{x\in\mathcal{S}}|\Upsilon_n^*(x)|.\end{aligned}$$

with

$$\Upsilon_n^*(x) = \frac{1}{n\psi_n(x,a_n)} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right).$$

Again, a direct application of Lemma 10 with Remark 6 in Woodroofe (1985)[23] complete the proof.

***Proof of Lemma* 5.** According to the definition of the bias term $\widehat{B}_n(x,y)$ above, which is not affected by the dependence condition. We use the fact that $\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]$ is bounded. Then, we can rewrite it in the following form

$$\widehat{B}_n(x,y) = \frac{\mathbb{E}\left[\mathbb{E}\left(\widetilde{\Psi}_n(x,y) - F^x(y)/X_k\right)\widetilde{\Upsilon}_n(x)\right]}{\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]}.$$

As we have already proved in previous steps the fact that this writing is satisfied by definition of the probability of trancature

$$\begin{aligned}
\mathbb{E}[\widetilde{\Psi}_n(x,y)] &= \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[\frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)\right] \\
&= \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)\right]
\end{aligned}$$

Therefore, by a conditioning to X_k we have

$$\begin{aligned}
\left|\mathbb{E}\left(\widetilde{\Psi}_n(x,y) - F^x(y)/X_k = u\right)\right| \leq & \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\Bigg[L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) \\
& \times \left|\mathbb{E}\left[L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right) - F^x(y)/X_k = u\right]\right|\Bigg].
\end{aligned}$$

Next, an integration by parts, a change of variable and because of condition (U.3), for any $u \in \mathcal{B}(x, a_k)$, we get

$$\begin{aligned}
&\left|\mathbb{E}\left[L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right) - F^x(y)/X_k = u\right]\right| \\
&\leq \int_{\mathbb{R}^p} L_2^{(1)}(\|t\|_{\mathbb{R}^p}) \left|F^{(u)}(y - b_k t) - F^{(x)}(y)\right| dt \\
&\leq \int_{\mathbb{R}^p} L_2^{(1)}(\|t\|_{\mathbb{R}^p}) \left|F^{(u)}(y - b_k t) - F^{(u)}(y)\right| dt + \left|F^{(u)}(y) - F^{(x)}(y)\right|
\end{aligned}$$

$$\leq C_1 \int_{\mathbb{R}^p} L_2^{(1)}(\|t\|_{\mathbb{R}^p}) \left(a_k^{\nu_1} + |b_k|^{\nu_2} \|t\|_{\mathbb{R}^p}^{\nu_2}\right) dt$$

which finishes the proof of this Lemma.

Acknowledgments

The authors would like to thank both the anonymous referees and the Editor in Chief for their valuable comments and constructive suggestions that have led to improve the presentation of paper substantially.

References

1. B. Altendji, J. Demongeot, A. Laksaci and M. Rachdi, Functional data analysis: Estimation of the relative error in functional regression under random left-truncation model, *Journal of Nonparametric Statistics*, **30** (2018), No.2, 1048-5252.

2. F. Benziadi, A. Laksaci and F. Tebboune, Recursive kernel estimate of the conditional quantile for functional ergodic data, *Comm in Statist-Theory and Methods*, **45** (2016), No.11, 3097-3113.

3. A. Berlinet, A. Gannoun and E. Matzner-Lober, Asymptotic normality of convergent estimates of conditional quantiles, *Statistics*, **35** (2001), 139-168.

4. I. Bouazza, F. Benziadi and T. Guendouzi, Nonparametric recursive estimate for right-censored conditional mode function with ergodic functional data, *J Indian Soc Probab Stat*, **22** (2021), 389-415. https://doi.org/10.1007/s41096-021-00110-5.

5. K. Chen, M. T. Chao and S. H. Lo, On strong uniform consistency of the Lynden-Bell estimator for truncated data. *The Annals of Statistics*, **23** (1995), No.2, 440-449.

6. S. Derrar, A. Laksaci, and E. O. Said, On the nonparametric estimation of the functional ψ-regression for a random left-truncation model, *J. Stat. Theory Pract*, **9** (2015), 823-849.

7. P. Doukhan, Mixing: Properties and examples, *Lecture Notes in Statistics*, **85**, Berlin: Springer, 1994.

8. F. Ferraty, A. Rabhi and P. Vieu., Conditional quantiles for functional dependant data with application to the climatic El Nio phenomenon, *Sankhy*, **67**(2005), 378-398.

9. F. Ferraty, A. Laksaci and P. Vieu, Estimating some characteristics of the conditional distribution in nonparametric functional models, *Statist. Inf. Stoch. Processes*, **9** (2006), 47-76.

10. F. Ferraty, A. Laksaci, A. Tadj and P. Vieu, Rate of uniform consistency for nonparametric estimates with functional variables, *J. of Statist. Plann. and Inference*, **140** (2010), 335-352.

11. Z. Guessoum and F. Hamrani., Simulation de l'estimateur noyau de la fonction de régression dans un modèle de troncature à gauche, MSS'2014 *Book of Abstracts.* (*Colloque International Modélisation Stochastique et Statistique*).

12. S. He and G. Yang, Estimation of the truncation probability in the random truncation model, *Annals of Statistics*, **26** (1998), 1011-1027.

13. N. Helal and E. O. Said, Kernel conditional quantile estimator under left truncation for functional regressors, *Opuscula Math*, **36** (2016), 25-48.

14. W. Horrigue and E. O. Said, Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors, Random *Oper. Stoch. Equ*, **19** (2011), 131-156.

15. W. Horrigue and E. O. Said, Nonparametric regression quantile estimation for dependant functional data under random censorship: Asymptotic normality, *Comm. in Statistics. Theory and Methods*, **44** (2014), No.20, 4307-4332.

16. J. P. Klein and M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*, Springer, second edition, New-York, 2003.

17. D. Lynden-Bell, A method of allowing for known observational selection in small samples applied to 3CR quasars, *Monthly Notices of the Royal Astronomical Society*, **155** (1971), 95-118.

18. M. Lemdani, E. O. Said and N. Poulin, Asymptotic properties of a conditional quantile estimator with randomly truncated data, *J. Multivariate Anal*, **100** (2009), 546-559.

19. E. O. Said and M. Lemdani, Asymptotic properties of a nonparametric regression function estimator with randomly truncated data, *Ann. Inst. Statist. Math*, **58** (2006), 357-378.

20. E. O. Said and A. Tatachak, Strong Consistency Rate for the Kernel Mode Estimator Under Strong Mixing Hypothesis and Left Truncation, *Communications in Statistics - Theory and Methods*, **38** (2009), No.8, 1154-1169.

21. W. Stute, Almost sure representations of the product-limit estimator for truncated data, *Ann. Statist*, **21** (1993), 146-156.

22. J. Wang, H. Y. Liang, and G. Fan, Asymptotic properties for an M-estimator of the regression function with truncation and dependent data, *Journal of the Korean Statistical Society*, **41** (2012a), 35-367.

23. M. Woodroofe, Estimating a distribution function with truncated data. *The Annals of Statistics*, **13** (1985), 163-177.