Hacettepe Journal of Mathematics and Statistics
Volume 48 (1) (2019), 290 – 311

Strong uniform consistency of a kernel conditional quantile estimator for censored and associated data

Wafaa Djelladj* and Abdelkader Tatachak†‡

Abstract

In survival or reliability studies, it is common to deal with data which are not only incomplete but weakly dependent too. Random right-censoring and random left-truncation are two common forms of such data when they are neither independent nor strongly mixing but rather associated. In this paper, we focus on kernel estimation of the conditional quantile function of a strictly stationary associated random variable T given a d-dimensional vector of covariates X, under random right-censoring. As main results, we establish a strong uniform consistency rate for the estimator. Then the finite sample performance of the estimator is illustrated on a simulation study.

Keywords: Associated data, Censored data, Convergence rate, Quantile function, Strong consistency.

2000 AMS Classification: 60G05, 62G20

Received : 23.11.2017 *Accepted* : 28.06.2018 *Doi* : 10.15672/HJMS.2018.609

1. Introduction

Let $\{T_n, n \geq 1\}$ be a strictly stationary sequence of associated random variables (rv's) of interest having an unknown absolutely continuous distribution function (df) F_T. This variable can be considered as a lifetime under biomedical studies. The major characteristic of survival time is the incompleteness.
In survival analysis especially in medical studies, we meet random censorship models which are one of the fundamental assumptions in the theory of survival analysis. Random right censoring is a well-known phenomenon which may be present when observing lifetime data. The lifetime variable may not be completely observable if the patient is

*Lab. MSTD, Department of Probability and Statistics, USTHB, BP 32, El-Alia 16111, Algeria, Email: wdjelladj@usthb.dz
†Lab. MSTD, Department of Probability and Statistics, USTHB, BP 32, El-Alia 16111, Algeria, Email: atatachak@usthb.dz
‡Corresponding Author.

still alive at the end of study or is dead for another reason or because of some departures of patients from the testing experimentation. Hence, the available data provide partial information. In this case, the variable of interest T is subject to right censoring by another non-negative rv C. In the sequel, we assume that the censoring lifetimes are independent and identically distributed (iid) and possess an unknown Lipschitz df G. We take in consideration the presence of a strictly stationary and associated covariate $\mathbf{X}$ taking values in $\mathbb{R}^d$. Under this model, the observable sequence is $\{(Y_i, \delta_i, \mathbf{X}_i), 1 \leq i \leq n\}$, with $Y_i = \min(T_i, C_i)$, $\delta_i = \mathbb{1}_{\{T_i \leq C_i\}}$ and where $\mathbb{1}_A$ denotes the indicator function of the event A.
As usual with random censoring, we assume that the censoring times $\{C_i, 1 \leq i \leq n\}$ are independent of $\{(\mathbf{X}_i, T_i), 1 \leq i \leq n\}$. This means that the censoring mechanism does not depend on the occurring event. Such a condition ensures the identifiability of the model. It is well known that the conditional df $F(\cdot|\mathbf{x})$ of $(T|\mathbf{X} = \mathbf{x})$ is defined by

$$F(t|\mathbf{x}) = \frac{1}{l(\mathbf{x})} \int_{-\infty}^{t} f(\mathbf{x}, z)dz =: \frac{F_1(\mathbf{x}, t)}{l(\mathbf{x})},$$

where $f(.,.)$ is the joint probability density function (pdf) of $(\mathbf{X}, T)$, $l(.)$ is the marginal pdf of $\mathbf{X}$ and $F_1(\mathbf{x}, .)$ is the first derivative of the joint df $F(\mathbf{x}, .)$ with respect to $\mathbf{x}$. The conditional pdf will be denoted by $f(.|\mathbf{x})$. Then, for all fixed $p \in (0,1)$, the p-th conditional quantile of T given $\mathbf{X} = \mathbf{x}$ is defined by

$$\xi_p(\mathbf{x}) := \inf\{t, F(t|\mathbf{x}) \geq p\}. \tag{1.1}$$

Hence, to get a nonparametric conditional quantile estimator, we clearly have to estimate $F_1(\mathbf{x}, t)$ by the mean of an unbiased kernel estimator and $l(\mathbf{x})$ is estimated by the famous kernel type estimator.
There has been various researches relating to the quantile estimator in view of its interesting properties. The estimator under consideration is renowned for its good description of the data (see Chaudhuri *et al.* [6]) and attracted interest of several authors.
In the complete framework, Samanta [25] established the strong convergence and the asymptotic normality of the kernel conditional quantile in the iid case. Bhattacharya and Gangopadhyay [2] gave a Bahadur-type representation of the conditional quantile and asymptotic models. Moreover, Mehra *et al.* [16] and Xiang [27] gave the almost sure convergence of a kernel type conditional quantile estimator and its asymptotic normality. Honda [12] treated the uniform convergence and asymptotic normality of the conditional quantile using local polynomial fitting approach while Abberger [1] studied quantile smoothing in financial time series.
On the same subject matter and under censoring, Dabrowska [7] established a Bahadur type representation of the quantile regression estimator. Besides, Qin and Wu [24] stated the asymptotic normality of an estimator for a conditional quantile when some auxiliary information is available using the empirical likelihood method and a linear fitting.
The strong representation of the conditional quantile estimator under right censoring and strong mixing condition was stated by Ould Saïd and Sadki [22] while Ould Saïd [20] established its strong uniform convergence rate in the iid case. Recently, Liang and de Uña-Álvarez [15] assessed its strong uniform consistency and asymptotic normality in the α-mixing setting.
Two kinds of dependency are widely used in the literature: mixing (Doukhan [8]) and association (Esary *et al.* [8]). These two concepts are not completely dissociated (see Doukhan and Louhichi [9]). In fact, we can find sequences that are associated but not mixing, associated and mixing, and mixing but not associated. The main advantage of the concept of association compared to mixing is that the conditions of limit theorems are easier to verify: indeed, a covariance is much easier to compute than a mixing coefficient.

Recall that a set of finite family of rv's $(T_1,\dots,T_n)$ are said to be associated if for all non-decreasing functions Ψ_1, Ψ_2

$$Cov(\Psi_1(T_1,\dots,T_n),\Psi_2(T_1,\dots,T_n))\geq 0,$$

whenever the covariance exists. An infinite family of rv's is associated if any finite sub-family is a set of associated rv's and any independent sequence is associated. In classical statistical inference, the observed rv's of interest are generally assumed to be iid. However, it is more common to have dependent variables in some real life situations. Dependent variables are present in several backgrounds such as medicine, biology and social sciences. Associated rv's are of considerable interest when dealing with reliability problems, percolation theory and some models in statistical mechanics.
The notion of association was firstly introduced by Esary *et al.* [11] mainly for an application in reliability. For more details on the subject we refer the reader to the monographs by Bulinski and Shashkin [3], Oliveira [19] and Prakasa Rao [23].
As far as we know, the problem of drawing nonparametric inference about the conditional quantile function under associated-censored model is not available and this motivates the study we consider here. So, the present paper deals with the almost sure uniform convergence with a rate of the estimator defined in (2.4). The paper is structured as follows: the expression of the studied estimator is presented in Section 2. Section 3 gathers the needed assumptions with some comments. A Simulation study is given in Section 4 while the last section includes the proofs of the main and some auxiliary results.

2. Notations and estimators

Recall that in the complete data case (no censoring), the traditional kernel estimator of $F(t|\mathbf{x})$ is given by

$$F_n(t|\mathbf{x})=\sum_{i=1}^{n}\omega_{in}(\mathbf{x})1\!\!1_{\{Y_i\leq t\}}, \tag{2.1}$$

where $\omega_{in}(.)$ are measurable functions. These functions called weights were introduced by Nadaraya-Watson in the context of the kernel regression and defined by

$$\omega_{in}(\mathbf{x}) \;=\; \frac{K_d\left(\dfrac{\mathbf{x}-\mathbf{X}_i}{h_{n,1}}\right)}{\displaystyle\sum_{j=1}^{n}K_d\left(\dfrac{\mathbf{x}-\mathbf{X}_j}{h_{n,1}}\right)}=\frac{\dfrac{1}{nh_{n,1}^d}K_d\left(\dfrac{\mathbf{x}-\mathbf{X}_i}{h_{n,1}}\right)}{l_n(\mathbf{x})},$$

with the convention $0|0=0$. Here K_d is a kernel function on $\mathbb{R}^d$ whereas $h_{n,1}$ is a positive sequence of bandwidths tending to 0 along with n and $l_n(.)$ is the Parzen-Rosenblatt kernel estimator of $l(.)$.
In the sequel, we will make use of the Inverse-Probability-of-Censoring Weighted (IPCW) idea of Koul *et al.* [14] to define the weights we will use after, that is

$$\omega_{in}(\mathbf{x})=\frac{1}{nh_{n,1}^d}\frac{\delta_i}{\overline{G}(Y_i)l_n(\mathbf{x})}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_{n,1}}\right). \tag{2.2}$$

It is well known that under right censoring model, the classical empirical distribution does not estimate consistently the df's F_T and G. Therefore, Kaplan and Meier [13] proposed consistent estimators $F_{T,n}$ and G_n for F_T and G, respectively, defined by

$$F_{T,n}(t)=1-\prod_{i=1}^{n}\left[1-\frac{\delta_{(i)}}{n-i+1}\right]^{1\!\!1\{Y_{(i)}\leq t\}}$$

and

$$G_n(t) = 1 - \prod_{i=1}^{n}\left[1 - \frac{1-\delta_{(i)}}{n-i+1}\right]^{\mathbb{1}\{Y_{(i)}\le t\}},$$

where $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ are the order statistics of $Y_1, Y_2, \ldots, Y_n$ and $\delta_{(i)}$ is the concomitant of $Y_{(i)}$.

The Kaplan-Meier estimator was studied in depth by many authors. For more details we refer to Stute and Wang [26] for the iid case, Cai [4] under α-mixing condition and Cai and Roussas [5] in the association setting.

Recall that, using the weights defined in (2.2), Ould Saïd [20] established a strong uniform consistency rate for the estimator in (2.1) in the iid case and d=1. The smoothed version of $F_n(\cdot|\cdot)$, namely

$$F_n(t|\mathbf{x}) =: \frac{F_{1,n}(\mathbf{x},t)}{l_n(\mathbf{x})} = \frac{\frac{1}{nh_{n,1}^d}\sum_{i=1}^{n}\frac{\delta_i}{\overline{G}_n(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_{n,1}}\right)H\left(\frac{t-Y_i}{h_{n,2}}\right)}{\frac{1}{nh_{n,1}^d}\sum_{i=1}^{n}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_{n,1}}\right)}, \tag{2.3}$$

was also considered and studied (strong consistency and asymptotic normality) in the iid case by Ould Saïd and Sadki [21]. Here, the bandwidth $h_{n,2}$ is not necessarily equal to $h_{n,1}$ and they will be denoted by $h_1 := h_{n,1}$ and $h_2 := h_{n,2}$.

Note that the estimator in (2.3) is an adapted version of that of Yu and Jones [28] to the censoring case. Originally, this smooth estimate for complete data (without the IPCW $\frac{\delta_i}{\overline{G}_n(Y_i)}$), was proposed and discussed by the last authors mainly to avoid the crossing problem which occurs when using an indicator function instead of a continuous df.

It follows that, in view of (2.3), a natural estimator of (1.1) can be computed by

$$\xi_{p,n}(\mathbf{x}) = \inf\{t, F_n(t|\mathbf{x}) \ge p\}. \tag{2.4}$$

To argue our main results, the following auxiliary pseudo-estimator will be of a great benefit in proving our results

$$\tilde{F}_n(t|\mathbf{x}) =: \frac{\tilde{F}_{1,n}(\mathbf{x},t)}{l_n(\mathbf{x})} = \frac{\frac{1}{nh_1^d}\sum_{i=1}^{n}\frac{\delta_i}{\overline{G}(Y_i)}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_1}\right)H\left(\frac{t-Y_i}{h_2}\right)}{\frac{1}{nh_1^d}\sum_{i=1}^{n}K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_1}\right)}. \tag{2.5}$$

Note that (2.5) can not be computed since $\overline{G}(\cdot)$ is assumed to be unknown.

3. Assumptions and main results

In the sequel, c stands for a positive constant taking different values and τ will denote a positive real number satisfying $\tau < \tau_F < \tau_G$ where, for any df W, $\tau_W := \sup\{y; W(y) < 1\}$. Define $\Omega_0 = \{\mathbf{x} \in \mathbb{R}^d / l(\mathbf{x}) \ge m_0 := \inf_x l(\mathbf{x}) > 0\}$ and let Ω and $\mathcal{C}$ be compact sets included in Ω_0 and $[0,\tau]$, respectively. The main results will be stated using the following assumptions:

A1. The bandwidths h_1 and h_2 satisfy
  (i) $h_1 \to 0$, $nh_1^{2\alpha+d(1-\alpha)} \to +\infty$ and $\frac{\log^5 n}{nh_1^d} \to 0$ as $n \to +\infty$,
  (ii) $h_2 \to 0$ and $nh_1^d h_2 \to +\infty$ as $n \to +\infty$;

A2. The kernel K_d is a bounded pdf, compactly supported and satisfies:
  (i) K_d is Hölder continuous of order $\alpha \in (0,1)$,
  (ii) $\int_{\mathbb{R}^d} u_j K_d(\mathbf{u})\mathbf{du} = 0$, for all $j = 1, ..., d$, where $\mathbf{u} = (u_1, ..., u_d)^\top$;

A3. The function H in (2.3) is of class $\mathcal{C}^1$. Furthermore, its derivative $H^{(1)}$ is assumed to be compactly supported and satisfies the properties of a second order kernel;

A4. The marginal density $l(.)$ is bounded and twice differentiable with:
$\sup_{\mathbf{x}\in\Omega}\left|\frac{\partial^k l(\mathbf{x})}{\partial x_i \partial x_j^{k-1}}(x)\right| < \infty$ for $i,j = 1,\dots,d$ and $k = 1,2$;

A5. The joint pdf $f(.,.)$ is bounded and twice continuously differentiable;

A6. The joint pdf $l_{i,j}(.,.)$ of $(\mathbf{X}_i, \mathbf{X}_j)$ is bounded;

A7. The joint pdf $f(.,.,.,.)$ of $(\mathbf{X}_i, Y_i, \mathbf{X}_j, Y_j)$ is bounded;

A8. Let us define Λ_{ij} as follows:

$$\Lambda_{ij} := \sum_{k=1}^{d}\sum_{l=1}^{d} Cov(X_i^k, X_j^l) + 2\sum_{k=1}^{d} Cov(X_i^k, Y_j) + Cov(Y_i, Y_j),$$

with X_i^k the k-th component of $\mathbf{X}_i$, such that for all $j \geq 1$ and $r > 0$

$$\sup_{i:|j-i|\geq r} \Lambda_{ij} =: \rho(r) \leq \gamma_0 e^{-\gamma r}, \text{ for all } \gamma_0, \gamma > 0;$$

A9. The function $\varsigma(\mathbf{x}) = \int_{\mathbb{R}} \frac{1}{\overline{G}(v)} f(\mathbf{x}, v)dv$ is bounded, continuously differentiable and $\sup_{\mathbf{x}\in\Omega}\left|\frac{\partial\varsigma}{\partial x_i}(\mathbf{x})\right| < \infty$ for $i = 1, ..., d$.

3.1. Remark. Assumption **A1** gives a classical choice of the bandwidths in functional estimation. For the sake of simplicity, many authors consider that $h_1 = h_2$ which is not justified in general. Note that the condition **A1** (ii) implies the first condition in **A1** (i) if $d \geq 2$. For $d = 1$, the comparison is not straightforward and depends upon the order of magnitude of h_2 with respect to h_1^α. Assumption **A2** is quite usual in kernel estimation. Assumptions **A3**-**A7** are classical in nonparametric estimation under dependency while **A8** is used for covariance calculation under association structure. Furthermore, this assumption gives a progressive trend to asymptotic independence of "past" and "future". Finally, Assumption **A9** is mainly technical.

The first result establishes the rate of convergence of the fluctuation term, that is $\left|\tilde{F}_{1,n}(\mathbf{x},t) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x},t)\right]\right|$. This will be done by applying a Bernstein-type inequality stated by Doukhan and Neumann [10] for weakly dependent rv's. The next result in Theorem 3.3 states a uniform almost sure convergence rate of $F_n(t|x)$ toward $F(t|x)$, which will be stated with the help of Theorem 3.2. Then, as an immediate result, the asymptotic behaviour of the kernel conditional quantile estimator will be deduced as presented in Corollary 3.4.

3.2. Theorem. *Suppose that assumptions **A1**-**A5** and **A7**-**A9** hold and for n large enough, we have*

$$\sup_{\boldsymbol{x}\in\Omega}\sup_{t\in\mathcal{C}}\left|\tilde{F}_{1,n}(\boldsymbol{x},t) - \mathbb{E}\left[\tilde{F}_{1,n}(\boldsymbol{x},t)\right]\right| = O\left(\sqrt{\frac{\log n}{nh_1^d}}\right), \text{ a.s.} \tag{3.1}$$

3.3. Theorem. *Under the assumptions of Theorem 3.2 and **A6**, for n large enough we have*

$$\sup_{\boldsymbol{x}\in\Omega}\sup_{t\in\mathcal{C}}|F_n(t|\boldsymbol{x}) - F(t|\boldsymbol{x})| = O\left\{(h_1^2 + h_2^2) + n^{-\theta} + \sqrt{\frac{\log n}{nh_1^d}}\right\}, \text{ a.s.}$$

with $0 < \theta < \gamma/(2\gamma + 9 + 3/2\kappa)$ for any $\kappa > 0$.

3.4. Corollary. *Under the assumptions of Theorem 3.3, and for each fixed $p \in (0,1)$ and $x \in \Omega$, if $\inf_{\boldsymbol{x}\in\Omega} f(\xi_p(\boldsymbol{x})|\boldsymbol{x}) > 0$, then for n large enough, we have*

$$\sup_{\boldsymbol{x}\in\Omega} |\xi_{p,n}(\boldsymbol{x}) - \xi_p(\boldsymbol{x})| = O\left\{(h_1^2 + h_2^2) + n^{-\theta} + \sqrt{\frac{\log n}{nh_1^d}}\right\}, \text{ a.s.}$$

3.5. Remark. The uniform positiveness condition on the conditional density in Corollary 3.4 ensures the uniform uniqueness of the conditional quantile. Hence $\forall\varepsilon > 0, \exists\beta > 0, \forall\eta_p : \Omega \to \mathbb{R}$,

$$\sup_{\mathbf{x}\in\Omega} |\xi_p(\mathbf{x}) - \eta_p(\mathbf{x})| \geq \varepsilon \Rightarrow \sup_{\mathbf{x}\in\Omega} |F(\xi_p(\mathbf{x})|\mathbf{x}) - F(\eta_p(\mathbf{x})|\mathbf{x})| \geq \beta.$$

3.6. Remark. We point out that the rate in Corollary 3.4 depends upon the parameter θ pertaining to the association dependence. In addition, remark that for γ large enough, the parameter θ approaches its upper bound (θ=1/2) and then, the covariances become negligible which in turn permits to compare our rate with those stated in the iid and strong mixing cases.

4. Simulation study

4.1. Description of the models. This part is established with the intention of giving the behaviour of the conditional quantile estimator. For this purpose, we only consider the cases of the conditional mean ($p = 1/2$) and the one dimensional covariate ($d = 1$) . The simulation is conducted for different sample sizes and censoring rates (CR). The performance of our estimator is quantified via the Global Mean Square Error (GMSE). The simulated data are obtained as follows:

- Generate $(n+1)$ iid rv's Z_i from gamma distribution ($Z_i \sim \Gamma(5, 0.5)$);
- Generate n iid rv's ε_i from normal distribution ($\varepsilon_i \sim \mathcal{N}(0, 0.01)$);
- Given Z_i, generate an n associated sequence (X_i, T_i) as follows:
  a) Linear case
  $$\begin{cases} X_i = \exp(Z_{i-1} + Z_{i-2})/2; \\ T_i = 3X_i/2 + 0.45\, \varepsilon_i. \end{cases}$$
  b) Nonlinear case
  $$\begin{cases} X_i = \exp(Z_{i-1} + Z_{i-2})/2, \\ T_i = \log(3X_i/2) + 0.45\, \varepsilon_i, \end{cases}$$
- Generate n iid rv's C_i from exponential distribution ($C_i \sim \exp(\lambda)$). The parameter λ is adjusted according to the CR's values;
- Keep the observed data $\{(Y_i := \min\{T_i, C_i\}), X_i, (\delta_i := \mathbb{1}_{\{T_i \leq C_i\}})\}$.

4.1. Remark. In computing the estimators, we use the standardized normal df and a Gaussian kernel for H and K, respectively.

In order to attenuate the boundary effect, we will use optimal local bandwidths. To do so, we first assume that $h_1 = h_2 =: h$, and this bandwidth sweeps the interval $[0.05, 0.8]$. For each model, the process above is repeated $B = 300$ times with fixed values of n and CR. Thus, we compute the conditional quantile estimator along a grid of points in

$[1.5, 4]$. At the end of the process, we keep the optimal local bandwidth which minimizes the estimating errors by means of the MSE (Mean Square Error) criterion, and then we quantify the $GMSE$. The formula calculating the $GMSE$ is

$$GMSE = \frac{1}{uB} \sum_{\ell=1}^{u} \sum_{k=1}^{B} \left[\xi_{p,n,k}(x_\ell) - \xi_p(x_\ell)\right]^2 ,$$

where $\xi_{p,n,k}(x_\ell)$ is the value of $\xi_{p,n}(x_\ell)$ at iteration k and u is the number of equidistant points x_ℓ belonging to $[1.5, 4]$.
To illustrate visually the quality of fit, we will plot the conditional quantile estimator $\xi_{p,n}(x_\ell)$ versus $\xi_p(x_\ell)$.

4.2. Simulation results.

4.2.1. ***Linear case:*** Note that under this model, the rv X follows $\Gamma(10, 0.5)$ and the conditional rv $(T|X = x)$ follows $\mathcal{N}(3x/2, 0.0045)$.
To show how is the influence of the censoring rate and the sample size on the quality of fit, we draw curves for different sample sizes $n = 50$, 100 and 300 and $CR = 40\%$, 25% and 10% as illustrated by Figures 1, 2 and 3. The corresponding errors with respect to the $GMSE$ are summarized in Table 1.

Table 1. Values of $GMSE$ for $\xi_{p,n}$ with $p = 0.5$

Linear case	$n = 50$	$n = 100$	$n = 300$
$CR = 10\%$	0.0637	0.0245	0.0069
$CR = 25\%$	0.1591	0.0586	0.0113
$CR = 40\%$	0.2465	0.1059	0.0128

4.2. Remark. From Table 1 and the graphs plotted for the linear case, we remark that the quality of fit seems to increase when the CR decreases. The curves reveal also that boundary effects on the right side tend to diminish for large values of n. Of course, the performance is quite acceptable when $n = 50$ and becomes more visible for $n = 300$. It means that the influence of the CR on the quality of fit becomes more and more insignificant along with n.

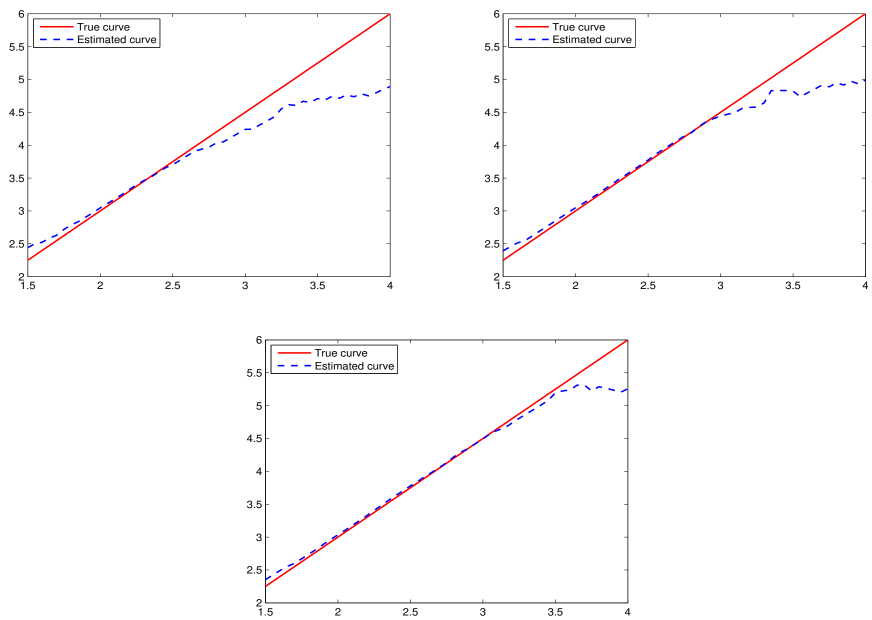

Figure 1. Linear case: $n = 50$ and $CR = 40$, 25 and 10, respectively

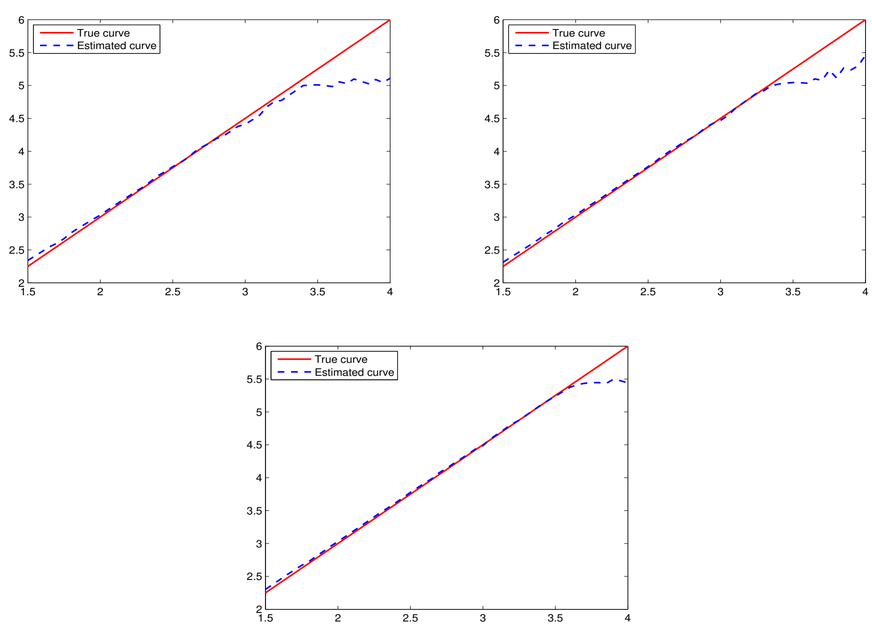

Figure 2. Linear case: $n = 100$ and $CR = 40$, 25 and 10, respectively

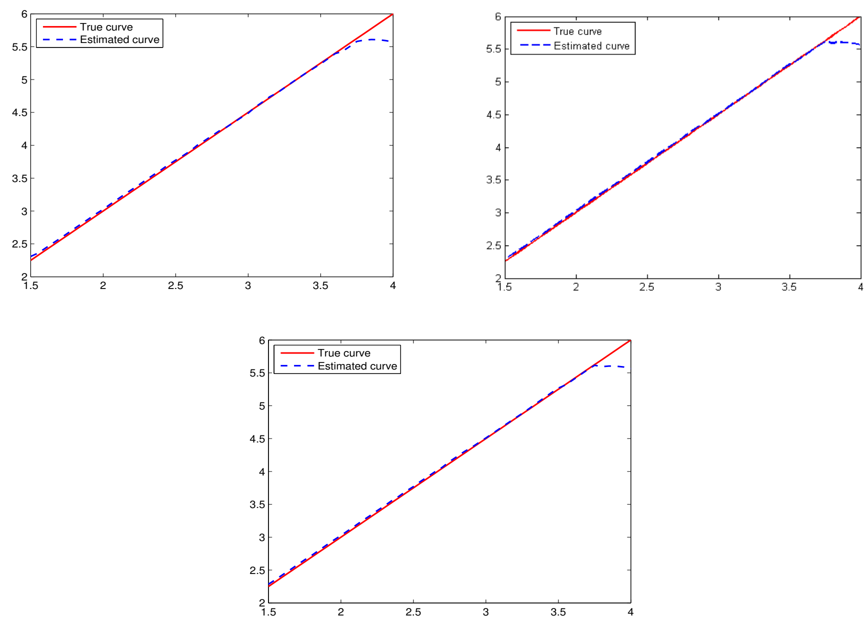

Figure 3. Linear case: $n = 300$ and $CR = 40$, 25 and 10, respectively

4.2.2. ***Non-linear case:*** Note that the rv $(T|X = x)$ follows $\mathcal{N}(\log(3x/2), 0.0045)$ and the choice of the log function permits to preserve the association property by monotonicity.
For the rest we proceed as for the linear case. The $GMSE'$s are summarized in Table 2 and the quality of fit is illustrated through Figures 4,5 and 6.

Table 2. Values of $GMSE$ for $\xi_{p,n}(.)$ with $p = 0.5$

Non-linear case	$n = 50$	$n = 100$	$n = 300$
$CR = 10\%$	24×10^{-3}	15×10^{-3}	5.54×10^{-4}
$CR = 25\%$	69×10^{-3}	25×10^{-3}	8.23×10^{-4}
$CR = 40\%$	11×10^{-2}	51×10^{-3}	16×10^{-3}

4.3. Remark. From Table 2 and the graphs, we observe that the estimator behaves similarly as for the linear case. The quality of fit becomes better along with the sample size which means that the behavior of the estimator remains correct even for large values of CR.

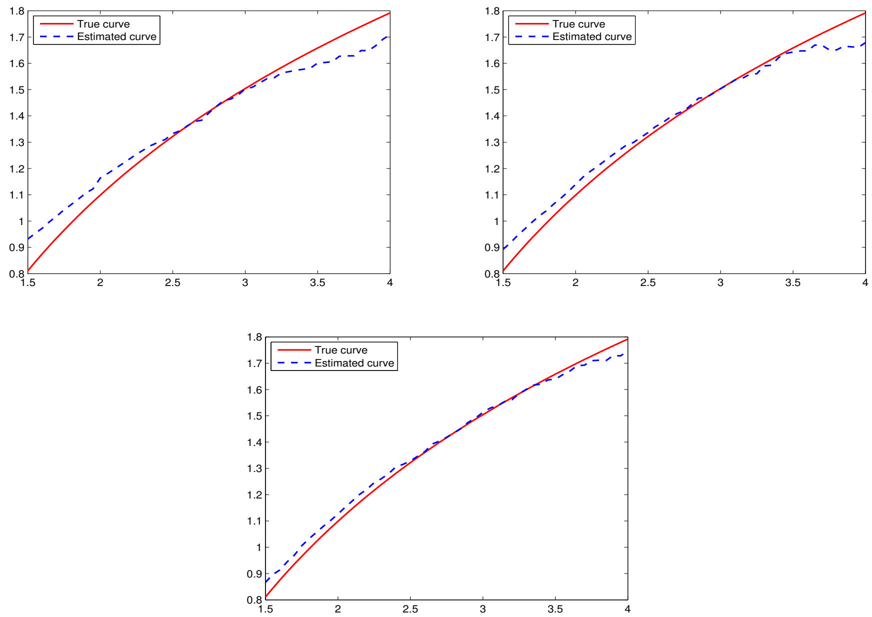

Figure 4. Non linear case: $n = 50$ and $CR = 40$, 25 and 10, respectively

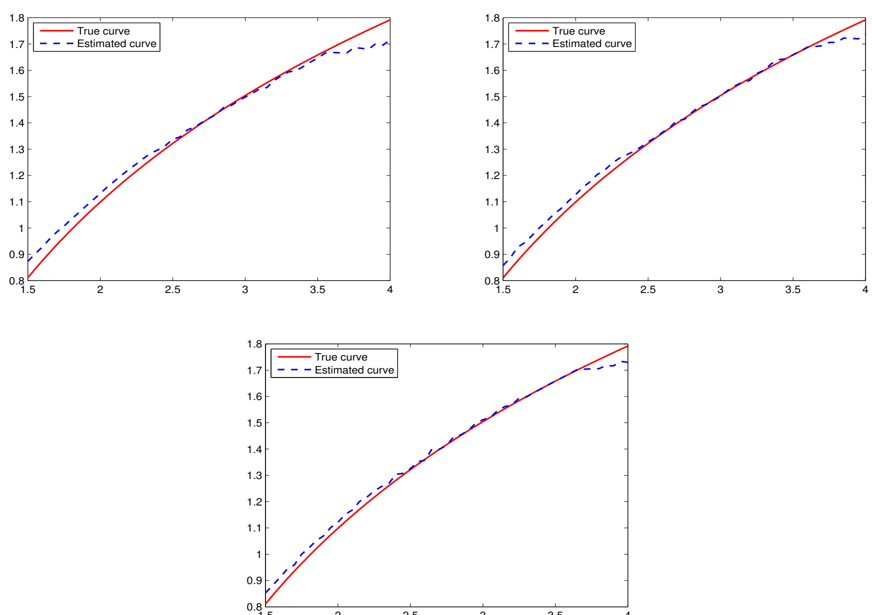

Figure 5. Non linear case: $n = 100$ and $CR = 40$, 25 and 10, respectively

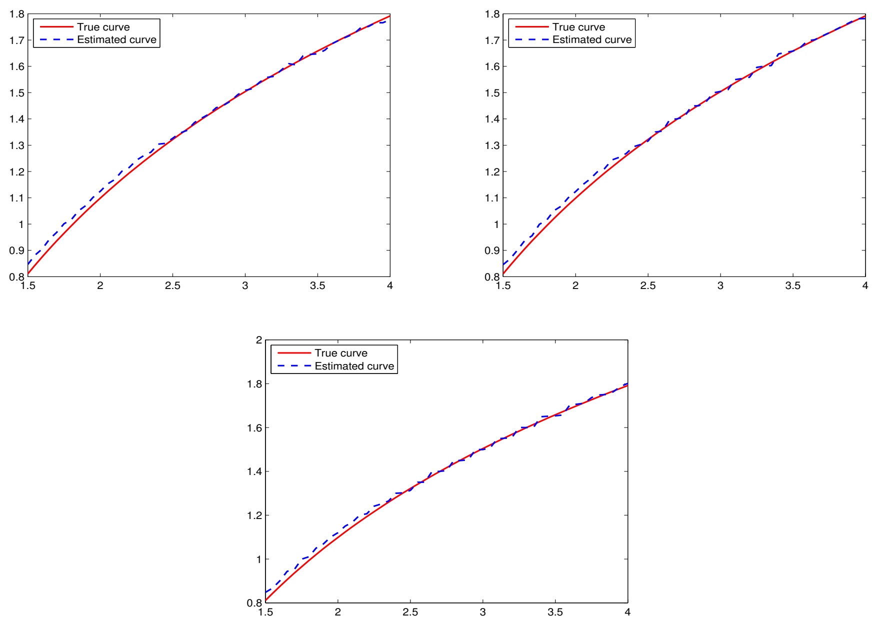

Figure 6. Non linear case: $n = 300$ and $CR = 40$, 25 and 10, respectively

5. Auxiliary results and proofs

For notational convenience, let us define

$$\begin{aligned}\Delta_i(\mathbf{x}, t) &= \frac{\delta_i}{\overline{G}(Y_i)} K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_1}\right) H\left(\frac{t-Y_i}{h_2}\right) \\ &- \mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)} K_d\left(\frac{\mathbf{x}-\mathbf{X}_1}{h_1}\right) H\left(\frac{t-Y_1}{h_2}\right)\right],\end{aligned}$$

for all $i = 1, ..., n$. It is easily seen that

$$\tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}, t)\right] = \frac{1}{nh_1^d} \sum_{i=1}^{n} \Delta_i(\mathbf{x}, t). \tag{5.1}$$

The items of the following proposition are similar to the conditions of Theorem 1 in Doukhan and Neumann [10]. Once the conditions are met, it becomes possible to use an exponential inequality to prove Theorem 3.2 related to the fluctuation term.

5.1. Proposition. *Let $\Delta_1(\boldsymbol{x}, t)$, $\Delta_2(\boldsymbol{x}, t)$,..., $\Delta_n(\boldsymbol{x}, t)$ be defined as above. Then, there exist constants M, L_1, L_2, $\mu \geq 0$, $\lambda \geq 0$ and a non-decreasing sequence of real coefficients $(\Upsilon(n))_{n\geq 0}$ so that for all p-tuples $(s_1, ..., s_p)$ and all q-tuples $(v_1, ..., v_q)$ with $1 \leq s_1 \leq ... \leq s_p \leq v_1 \leq ... \leq v_q \leq n$, we have*

a) $Cov\left(\prod_{i=s_1}^{s_p} \Delta_i(\boldsymbol{x}, t), \prod_{j=v_1}^{v_q} \Delta_j(\boldsymbol{x}, t)\right) \leq c^{p+q} h_1^d h_2^{\frac{2}{d+1}} pq \Upsilon(v_1 - s_p)$,

b) $\sum_{s=0}^{\infty} (s+1)^{k_0} \Upsilon(s) \leq L_1 L_2^{k_0} (k_0!)^{\mu}, \forall k_0 \geq 0$,

c) $\mathbb{E}\left[|\Delta_i(\boldsymbol{x},t)|^{k_0}\right] \leq (k_0!)^{\lambda} M^{k_0}$.

Proof. **Proof of Proposition 5.1** To prove the first item of Proposition 5.1, we need the following lemma:

5.2. Lemma. *Under assumptions **A2**, **A5**, **A7** and **A8**, we have*

i) $Cov\left(\prod_{i=s_1}^{s_p} \Delta_i(\boldsymbol{x},t), \prod_{j=v_1}^{v_q} \Delta_j(\boldsymbol{x},t)\right) =: C_1 \leq c^{p+q} h_1^{-2} h_2^{-2} pq\rho(v_1 - s_p)$,

ii) $Cov\left(\prod_{i=s_1}^{s_p} \Delta_i(\boldsymbol{x},t), \prod_{j=v_1}^{v_q} \Delta_j(\boldsymbol{x},t)\right) =: C_2 \leq c^{p+q} h_1^{2d} h_2^{2}$.

Proof. Exploiting the definition 5.1, p.88 in Bulinski & Shashkin [3], we recall that the partial Lipschitz constants are defined as follows

$$Lip_i(\Phi_m) = \sup_{\substack{z_1,\ldots,z_m \\ z_i \neq z_i', z_i' \in \mathbb{R}}} \frac{|\Phi_m(z_1,\ldots,z_{i-1},z_i,z_{i+1},\ldots,z_m) - \Phi_m(z_1,\ldots,z_{i-1},z_i',z_{i+1},\ldots,z_m)|}{|z_i - z_i'|}, \tag{5.2}$$

where $\Phi_m : \mathbb{R}^m \to \mathbb{R}$ and $Lip(\Phi_m)$ denotes the Lipschitz continuity modulus of Φ_m, viz

$$Lip(\Phi_m) = \sup_{\mathbf{x} \neq \mathbf{y}} \frac{|\Phi_m(\mathbf{x}) - \Phi_m(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_1},$$

with $\|(z_1,\ldots,z_n)\|_1 = |z_1| + \ldots + |z_n|$.
To prove part (i) in Lemma 5.2, we use Theorem 5.3, p.89 in (Bulinski and Shashkin [3]). Firstly, we set

$$\Phi_p =: \prod_{i=s_1}^{s_p} \Delta_i \quad \text{and} \quad \Phi_q =: \prod_{j=v_1}^{v_q} \Delta_j.$$

Then, using the fact that K_d, H and G are Lipschitz functions, we have

$$Cov(\Phi_p, \Phi_q) \leq \sum_{i=s_1}^{s_p} \sum_{j=v_1}^{v_q} Lip_i(\Phi_p) Lip_j(\Phi_q) \Lambda_{ij},$$

The definition in (5.2) leads to

$$Lip_i(\Phi_p) \leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)}\right)^p \|K_d\|_\infty^{p-1}$$

and

$$Lip_j(\Phi_q) \leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)}\right)^q \|K_d\|_\infty^{q-1},$$

where $M_0 = \max\left\{h_2 Lip(K) \|K\|_\infty^{d-1}, h_1\left(Lip(H) + h_2 \frac{Lip(\overline{G})}{\overline{G}(\tau)}\right) \|K_d\|_\infty\right\}$.
Note that the partial Lipschitz constants are obtained as follows

$$\begin{aligned} Lip_i(\Phi_p) &\leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)}\right)^{p-1} \|K_d\|_\infty^{p-1} \frac{1}{\overline{G}(\tau)} \\ &\leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)}\right)^{p} \|K_d\|_\infty^{p-1}. \end{aligned}$$

If Assumption **A8** holds, by stationarity we get

$$\begin{aligned} Cov(\Phi_p, \Phi_q) &\leq \frac{M_0^2}{h_1^2 h_2^2} \left(\frac{2}{\overline{G}(\tau)} \right)^{p+q} \|K_d\|_\infty^{p+q-2} \sum_{i=s_1}^{s_p} \sum_{j=v_1}^{v_q} \Lambda_{ij} \\ &\leq \frac{c^{p+q}}{h_1^2 h_2^2} \, pq \, \rho(v_1 - s_p). \end{aligned}$$

This achieves the proof of (i). In order to prove the second part of Lemma 5.2, we need to calculate the covariance term as shown hereafter by using the fact that

$$\mathbb{E}\left[\delta_i \delta_j | T_i, T_j\right] = \mathbb{E}\left[\mathbb{1}_{\{T_i \leq C_i\}} \mathbb{1}_{\{T_j \leq C_j\}} | T_i, T_j \right] = \overline{G}(T_i)\overline{G}(T_j).$$

We also use the following simplified notations

$$K_{d,\mathbf{x},i} := K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1}\right) \quad \text{and} \quad H_{t,i} := H\left(\frac{t - Y_i}{h_2}\right).$$

Indeed, we have

$$\begin{aligned} &Cov(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t)) \\ &= \mathbb{E}\left[\frac{\delta_i \delta_j}{\overline{G}(Y_i)\overline{G}(Y_j)} K_{d,\mathbf{x},i} H_{t,i} K_{d,\mathbf{x},j} H_{t,j}\right] \\ &- \mathbb{E}\left[\frac{\delta_i}{\overline{G}(Y_i)} K_{d,\mathbf{x},i} H_{t,i}\right] \times \mathbb{E}\left[\frac{\delta_j}{\overline{G}(Y_j)} K_{d,\mathbf{x},j} H_{t,j}\right] \\ &= \mathbb{E}\left[K_{d,\mathbf{x},i} K_{d,\mathbf{x},j} \ \mathbb{E}\left(\frac{\delta_i \delta_j}{\overline{G}(Y_i)\overline{G}(Y_j)} H_{t,i} H_{t,j} | \mathbf{X}_i, \mathbf{X}_j\right)\right] \\ &- \mathbb{E}\left[K_{d,\mathbf{x},i} \ \mathbb{E}\left(\frac{\delta_i}{\overline{G}(Y_i)} H_{t,i} | \mathbf{X}_i\right)\right] \times \mathbb{E}\left[K_{d,\mathbf{x},j} \ \mathbb{E}\left(\frac{\delta_j}{\overline{G}(Y_j)} H_{t,j} | \mathbf{X}_j\right)\right] \\ &= \mathbb{E}\left[K_{d,\mathbf{x},i} K_{d,\mathbf{x},j} \ \mathbb{E}\left(H_{t,i} H_{t,j} \frac{\mathbb{E}\left[\delta_i \delta_j | T_i, T_j\right]}{\overline{G}(T_i)\overline{G}(T_j)} | \mathbf{X}_i, \mathbf{X}_j\right)\right] \\ &- \mathbb{E}\left[K_{d,\mathbf{x},i} \ \mathbb{E}\left(H_{t,i} \frac{\mathbb{E}\left[\delta_i | T_i\right] | \mathbf{X}_i}{\overline{G}(T_i)}\right)\right] \mathbb{E}\left[K_{d,\mathbf{x},j} \ \mathbb{E}\left(H_{t,j} \frac{\mathbb{E}\left[\delta_j | T_j\right]}{\overline{G}(T_j)} | \mathbf{X}_j\right)\right]. \end{aligned}$$

Then, we get

$$\begin{aligned} &|Cov(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))| \\ &\leq \left| \int_{\mathbb{R}^{2d} \times \mathbb{R}^2} K_d\left(\frac{\mathbf{x} - \mathbf{u}}{h_1}\right) H\left(\frac{t - s}{h_2}\right) K_d\left(\frac{\mathbf{x} - \mathbf{r}}{h_1}\right) H\left(\frac{t - v}{h_2}\right) \right. \\ &\times f(\mathbf{u}, s, \mathbf{r}, v) \mathbf{du} ds \mathbf{dr} dv| \\ &+ \left| \int_{\mathbb{R}^{d+1}} K_d\left(\frac{\mathbf{x} - \mathbf{u}}{h_1}\right) H\left(\frac{t - s}{h_2}\right) f(\mathbf{u}, s) \mathbf{du} ds \right. \\ &\times \left. \int_{\mathbb{R}^{d+1}} K_d\left(\frac{\mathbf{x} - \mathbf{r}}{h_1}\right) H\left(\frac{t - v}{h_2}\right) f(\mathbf{r}, v) \mathbf{dr} dv \right|. \end{aligned}$$

Moreover, under assumptions **A2**, **A5** and **A7**, using a change of variables we get

$$|Cov(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))| = O(h_1^{2d} h_2^2). \tag{5.3}$$

Finally, the second part of Lemma 5.2 follows by simple algebra. □

We need some auxiliary notations to set up the proof of Proposition 5.1. Impose $\Upsilon(.) = \rho^{\frac{d}{2d+2}}(.)$ and use the upper bounds of Lemma 5.2, namely

$$C_1^{\frac{d}{2d+2}} \leq c^{\frac{(p+q)d}{2d+2}} h_1^{\frac{-2d}{2d+2}} h_2^{\frac{-2d}{2d+2}} (pq)^{\frac{d}{2d+2}} \rho^{\frac{d}{2d+2}}(v_1 - s_p), \tag{5.4}$$

$$C_2^{\frac{d+2}{2d+2}} \leq c^{(p+q)\frac{d+2}{2d+2}} h_1^{\frac{2d(d+2)}{2d+2}} h_2^{\frac{2(d+2)}{2d+2}}. \tag{5.5}$$

Combining (5.4) and (5.5), we get

$$\begin{aligned} C_1^{\frac{d}{2d+2}} C_2^{\frac{d+2}{2d+2}} &\leq c^{p+q} h_1^d h_2^{\frac{2}{d+1}} (pq)^{\frac{d}{2d+2}} \rho^{\frac{d}{2d+2}}(v_1 - s_p) \\ &\leq c^{p+q} h_1^d h_2^{\frac{2}{d+1}} pq\, \Upsilon(v_1 - s_p). \end{aligned}$$

This inequality concludes the proof of part (a) of Proposition 5.1. Next, under Assumption **A8** and choosing $\lambda = 0$, $\mu = 1$, $L_1 = L_2 = \frac{1}{1-e^{\frac{-\gamma d}{2d+2}}}$, the proofs of the results in (b) and (c) are similar to those used in proving Proposition 8 in (Doukhan and Neumann [10]), then we omit them. The proof of Proposition 5.1 is complete. □

Proof. **Proof of Theorem 3.2** In order to set up the uniform asymptotic expression of the fluctuation term $\left|\tilde{F}_{1,n}(\mathbf{x},t) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x},t)\right]\right|$, we apply the triangular inequality and classical techniques to cover compacts. So, Ω can be covered by a finite number $d_{x,n}$ of balls $B_k(\mathbf{x}_k, a_n^d)$ centred at $\mathbf{x}_k = (x_{k,1}, ..., x_{k,d})$ and $\mathcal{C}$ is split into $d_{t,n}$ subintervals $J_1, ..., J_{d_{t,n}}$ of lengths b_n, centred at t_ℓ. In other words, for all $\mathbf{x} \in \Omega$, $t \in \mathcal{C}$, there exist integers $k \in \{1, ..., d_{x,n}\}$ and $\ell \in \{1, ..., d_{t,n}\}$ such that $\|\mathbf{x} - \mathbf{x}_k\| \leq a_n^d$ and $|t - t_\ell| \leq b_n$, with $a_n^d = \left(n^{-1} h_1^{2\alpha+d}\right)^{1/2\alpha}$ and $b_n = \left(n h_1^d\right)^{-1/2} h_2$. Then, as Ω and $\mathcal{C}$ are bounded, let m_1 and m_2 be positive constants satisfying $d_{x,n} a_n^d \leq m_1$ and $d_{t,n} b_n \leq m_2$.

5.3. Remark. In proving our results we will use Lemma 5.4 stated in Menni and Tatachak [17] (see their Lemma 3) which governs a strong uniform consistency rate of the kernel estimator $l_n(.)$. We recall it hereinafter without proof.

5.4. Lemma. *Under assumptions **A1**, **A2**, **A4**, **A6** and **A8**, for n large enough we have*

$$\sup_{\boldsymbol{x} \in \Omega} |l_n(\boldsymbol{x}) - l(\boldsymbol{x})| = O\left(\max\left\{\sqrt{\frac{\log n}{n h_1^d}}, h_1^2\right\}\right) \quad \text{a.s.}$$

Next, using basic arguments, the left hand side in (3.1) is upper bounded as follows

$$\sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} \left|\tilde{F}_{1,n}(\mathbf{x},t) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x},t)\right]\right| \leq I_{1n} + I'_{1n} + I_{2n} + I'_{2n} + I_{3n},$$

with

$$I_{1n} = \max_{1\leq k\leq d_{x,n}} \sup_{\mathbf{x}\in B_k} \sup_{t\in\mathcal{C}} \left|\tilde{F}_{1,n}(\mathbf{x},t) - \tilde{F}_{1,n}(\mathbf{x}_k,t)\right|,$$

$$I'_{1n} = \max_{1\leq k\leq d_{x,n}} \sup_{\mathbf{x}\in B_k} \sup_{t\in\mathcal{C}} \left|\mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}_k,t)\right] - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x},t)\right]\right|,$$

$$I_{2n} = \max_{1\leq k\leq d_{x,n}} \max_{1\leq \ell\leq d_{t,n}} \sup_{t\in J_\ell} \left|\tilde{F}_{1,n}(\mathbf{x}_k,t) - \tilde{F}_{1,n}(\mathbf{x}_k,t_\ell)\right|,$$

$$I'_{2n} = \max_{1\leq k\leq d_{x,n}} \max_{1\leq \ell\leq d_{t,n}} \sup_{t\in J_\ell} \left|\mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}_k,t_\ell)\right] - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}_k,t)\right]\right|,$$

$$I_{3n} = \max_{1\leq k\leq d_{x,n}} \max_{1\leq \ell\leq d_{t,n}} \left|\tilde{F}_{1,n}(\mathbf{x}_k,t_\ell) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}_k,t_\ell)\right]\right.$$

Concerning I_{1n} and I'_{1n}, we apply the SLLN for associated sequences (see Newman [18]) and Assumption **A2(i)**. We obtain

$$\begin{aligned}
&\left|\tilde{F}_{1,n}(\mathbf{x},t)-\tilde{F}_{1,n}(\mathbf{x}_k,t)\right| \\
&=\left|\frac{1}{nh_1^d}\sum_{i=1}^{n}\frac{\delta_i}{\overline{G}(Y_i)}H\left(\frac{t-Y_i}{h_2}\right)\left[K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_1}\right)-K_d\left(\frac{\mathbf{x}_k-\mathbf{X}_i}{h_1}\right)\right]\right| \\
&\leq \frac{c}{h_1^d\,\overline{G}(\tau)}\frac{\|\mathbf{x}-\mathbf{x}_k\|^\alpha}{h_1^\alpha}\frac{1}{n}\sum_{i=1}^{n}\delta_i \\
&\leq \frac{c}{\overline{G}(\tau)}\frac{a_n^{d\alpha}}{h_1^{d+\alpha}}\frac{1}{n}\sum_{i=1}^{n}\delta_i \\
&= O\left(\frac{1}{\sqrt{nh_1^d}}\right). \qquad (5.6)
\end{aligned}$$

To treat the terms I_{2n} and I'_{2n}, we use Assumption **A3** and Lemma 5.4. We get

$$\begin{aligned}
&\left|\tilde{F}_{1,n}(\mathbf{x}_k,t)-\tilde{F}_{1,n}(\mathbf{x}_k,t_\ell)\right| \\
&=\left|\frac{1}{nh_1^d}\sum_{i=1}^{n}\frac{\delta_i}{\overline{G}(Y_i)}K_d\left(\frac{\mathbf{x}_k-\mathbf{X}_i}{h_1}\right)\left[H\left(\frac{t-Y_i}{h_2}\right)-H\left(\frac{t_\ell-Y_i}{h_2}\right)\right]\right| \\
&\leq \frac{c}{\overline{G}(\tau)}\frac{|t-t_\ell|}{h_2}\frac{1}{nh_1^d}\sum_{i=1}^{n}K_d\left(\frac{\mathbf{x}_k-\mathbf{X}_i}{h_1}\right) \\
&\leq \frac{c}{\overline{G}(\tau)}\frac{b_n}{h_2}l_n(\mathbf{x}_k) \\
&= O\left(\frac{1}{\sqrt{nh_1^d}}\right). \qquad (5.7)
\end{aligned}$$

We can focus now on upper bounding the term I_{3n}. To do so, we apply an exponential inequality adapted to associated sequences (see, Theorem 1, p.19 in Doukhan and Neumann [10]). Indeed, for all $\varepsilon > 0$, we have

$$\mathbb{P}\left(\sum_{i=1}^{n}\Delta_i(\mathbf{x}_k,t_\ell)\geq\varepsilon\right)\leq\exp\left(-\frac{\varepsilon^2/2}{A_n+B_n^{1/(\mu+\lambda+2)}\varepsilon^{(2\mu+2\lambda+3)/(\mu+\lambda+2)}}\right), \qquad (5.8)$$

where A_n is any number greater than σ_n^2 and

$$\sigma_n^2 := \left(\sum_{i=1}^{n}\Delta_i(\mathbf{x}_k,t_\ell)\right),$$

$$B_n = 2cL_2\max\left(\frac{2^{4+\mu+\lambda}cnh_1^dh_2^{\frac{2}{d+1}}L_1}{A_n},1\right). \qquad (5.9)$$

Firstly, we have to calculate σ_n^2. Indeed $\sigma_n^2 = (nh_1^d)^2 \mathrm{Var}\left(\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell)\right)$.
On the other hand, we have

$$
\begin{aligned}
&(nh_1^d)^2 \mathrm{Var}\left(\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell)\right) = n\mathrm{Var}\left(\frac{\delta_1}{\overline{G}(Y_1)} K_{d,\mathbf{x}_k,1} H_{t_l,1}\left(\frac{t_\ell - Y_1}{h_2}\right)\right) \\
&+ \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} Cov\left(\frac{\delta_i}{\overline{G}(Y_i)} K_{d,\mathbf{x}_k,i} H_{t_l,i}, \frac{\delta_j}{\overline{G}(Y_j)} K_{d,\mathbf{x}_k,j} H_{t_l,j}\right) \\
&=: V + S.
\end{aligned}
$$

Firstly, let us focus on V.

$$
\begin{aligned}
V &= n\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)^2} K_d^2\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1}\right) H^2\left(\frac{t_\ell - Y_1}{h_2}\right)\right] \\
&- n\mathbb{E}^2\left[\frac{\delta_1}{\overline{G}(Y_1)} K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1}\right) H\left(\frac{t_\ell - Y_1}{h_2}\right)\right] \\
&=: n(D_1 - D_2).
\end{aligned}
$$

Concerning D_1, we use classical conditional expectation techniques. So, under assumptions **A1(i)**, **A2** and **A9**, a change of variable and a Taylor expansion around $\mathbf{x}_k$, we get

$$
\begin{aligned}
D_1 &= \mathbb{E}\left[K_d^2\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1}\right)\mathbb{E}\left[\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)^2} H^2\left(\frac{t_\ell - Y_1}{h_2}\right)|T_1\right]|\mathbf{X}_1\right]\right] \\
&= \int_{\mathbb{R}^d} K_d^2\left(\frac{\mathbf{x}_k - \mathbf{u}}{h_1}\right)\mathbb{E}\left[H^2\left(\frac{t_\ell - T_1}{h_2}\right)\frac{1}{\overline{G}(T_1)}|\mathbf{X}_1 = \mathbf{u}\right] l(\mathbf{u})d\mathbf{u} \\
&\leq \int_{\mathbb{R}^d} K_d^2\left(\frac{\mathbf{x}_k - \mathbf{u}}{h_1}\right)\int_{\mathbb{R}} \frac{1}{\overline{G}(v)} f(\mathbf{u}, v)dv d\mathbf{u}, \quad \text{because H(.) is a df;} \\
&= h_1^d \int_{\mathbb{R}^d} K_d^2(\mathbf{z})\varsigma(\mathbf{x}_k - \mathbf{z}h_1)d\mathbf{z} \\
&= h_1^d \int_{\mathbb{R}^d} \varsigma(\mathbf{x}_k)K_d^2(\mathbf{z})d\mathbf{z} - h_1^{d+1}\int_{\mathbb{R}^d} K_d^2(\mathbf{z})\left[z_1\frac{\partial\varsigma(\mathbf{x}_k^*)}{\partial x_{k,1}} + \cdots + z_d\frac{\partial\varsigma(\mathbf{x}_k^*)}{\partial x_{k,d}}\right]d\mathbf{z} \\
&= O(h_1^d).
\end{aligned}
$$

Here $\mathbf{x}_k^*$ is between $\mathbf{x}_k - \mathbf{z}h_1$ and $\mathbf{x}_k$. Again, to upper bound D_2 we work similarly as before. Indeed, using a change of variable, Taylor expansion and assumptions **A1(i)**, **A2** and **A4**, we get

$$
\begin{aligned}
D_2 &= \mathbb{E}^2\left[K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1}\right)\mathbb{E}\left[\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)} H\left(\frac{t_\ell - Y_1}{h_2}\right)|T_1\right]|\mathbf{X}_1\right]\right] \\
&= O(h_1^{2d}).
\end{aligned}
$$

Consequently, we get

$$V = O(nh_1^d).$$

Secondly, to evaluate S, we first define

$$B_1 = \{(i,j); 1 \leq |i-j| \leq \eta_n\} \text{ and } B_2 = \{(i,j); \eta_n + 1 \leq |i-j| \leq n-1\},$$

where $\eta_n = o(n)$. Then, we have

$$\begin{aligned} S &= \sum_{i=1}^{n}\sum_{B_1} Cov\left(\frac{\delta_i}{\overline{G}(Y_i)}K_{d,\mathbf{x}_k,i}H_{t_l,i}, \frac{\delta_j}{\overline{G}(Y_j)}K_{d,\mathbf{x}_k,j}H_{t_l,j}\right) \\ &+ \sum_{i=1}^{n}\sum_{B_2} Cov\left(\frac{\delta_i}{\overline{G}(Y_i)}K_{d,\mathbf{x}_k,i}H_{t_l,i}, \frac{\delta_j}{\overline{G}(Y_j)}K_{d,\mathbf{x}_k,j}H_{t_l,j}\right) \\ &= : S_1 + S_2. \end{aligned}$$

From (5.3), it is clear that

$$S_1 = n\eta_n O(h_1^{2d}h_2^2) = O(n\eta_n h_1^{2d}h_2^2). \tag{5.10}$$

Next, the term S_2 will be upper bounded by remaking that result $a)$ in Proposition 5.1 and Assumption **A8** permit to write

$$\begin{aligned} S_2 &\leq \sum_{i=1}^{n}\sum_{B_2} c^2 h_1^d h_2^{\frac{2}{d+1}} \rho^{\frac{d}{2d+2}}(|i-j|) \\ &\leq nc^2 h_1^d h_2^{\frac{2}{d+1}} \sum_{B_2} \gamma_0^{\frac{d}{2d+2}} e^{-\frac{\gamma|i-j|d}{2d+2}} \\ &\leq nc^2 h_1^d h_2^{\frac{2}{d+1}} \int_{\eta_n}^{n} e^{-\frac{\gamma u d}{2d+2}} du \\ &= O\left(n h_1^d h_2^{\frac{2}{d+1}} e^{-\frac{\gamma \eta_n d}{2d+2}}\right). \end{aligned} \tag{5.11}$$

So, under Assumption **A1** and taking $\eta_n = O(h_1^{\nu_1 - d}h_2^{\nu_2 - 1})$ with $0 < \nu_1 < d$ and $0 < \nu_2 < 1$, the bounds in (5.10) and (5.11) become of order $o(nh_1^d h_2)$ and $o(nh_1^d h_2^{\frac{2}{d+1}})$, respectively. Consequently

$$\sigma_n^2 = V + S = O(nh_1^d) + o(nh_1^d h_2^{\frac{2}{d+1}}) = O(nh_1^d).$$

Thereby, we get $A_n = O(nh_1^d)$. Next, from (5.9) and choosing μ, λ, L_1 and L_2 as those in the proof of Proposition 5.1, we get $B_n = O(1)$.

Regarding I_{3n}, in view of (5.1), (5.8) and letting $\varepsilon = \varepsilon_0 \sqrt{\dfrac{\log n}{nh_1^d}}$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\max_{1\le k\le d_{x,n}} \max_{1\le \ell\le d_{t,n}} \left| \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell)\right]\right| \ge \varepsilon \right) \\
&= \mathbb{P}\left(\max_{1\le k\le d_{x,n}} \max_{1\le \ell\le d_{t,n}} \left| \sum_{i=1}^{n} \Delta_i(\mathbf{x}_k, t_\ell)\right| \ge nh_1^d \varepsilon \right) \\
&\le \sum_{k=1}^{d_{x,n}} \sum_{\ell=1}^{d_{t,n}} \mathbb{P}\left(\left| \sum_{i=1}^{n} \Delta_i(\mathbf{x}_k, t_\ell)\right| \ge nh_1^d \varepsilon \right) \\
&\le 2 d_{x,n} d_{t,n} \exp\left(\frac{-(nh_1^d)^2 \frac{\varepsilon_0^2}{2} \frac{\log n}{nh_1^d}}{cnh_1^d + \varepsilon_0^{5/3} (nh_1^d)^{5/3} \left(\frac{\log n}{nh_1^d}\right)^{5/6}} \right) \\
&\le 2 \frac{m_1}{a_n^d} \frac{m_2}{b_n} \exp\left(\frac{-\frac{\varepsilon_0^2}{2} \log n}{c + \varepsilon_0^{5/3} \left(\frac{\log n^5}{nh_1^d}\right)^{1/6}} \right) \\
&\le c \left(n^{-1} h_1^{2\alpha+d}\right)^{-1/2\alpha} \left(nh_1^d\right)^{1/2} h_2^{-1} n^{-c\varepsilon_0^2}
\end{aligned}
$$

$$
(5.12) \qquad = c\left(nh_1^{2\alpha+d(1-\alpha)}\right)^{-1/2\alpha} (nh_2)^{-1} n^{-c\varepsilon_0^2 + \frac{1}{\alpha} + \frac{3}{2}}.
$$

So, under Assumption **A1** and taking $\varepsilon_0^2 > \frac{1}{c}\left(\frac{1}{\alpha} + \frac{5}{2}\right)$, the term in (5.12) is the general term of a convergent series. Then, we have

$$
\sum_{n\ge 1} \mathbb{P}\left(\max_{1\le k\le d_{x,n}} \max_{1\le \ell\le d_{t,n}} \left| \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell)\right]\right| \ge \varepsilon_0 \sqrt{\frac{\log n}{nh_1^d}} \right) < \infty.
$$

Applying the lemma of Borel-Cantelli, we obtain that

$$
(5.13) \qquad I_{3n} = O\left(\sqrt{\frac{\log n}{nh_1^d}} \right).
$$

Involving (5.6), (5.7) and (5.13), we deduce that

$$
\sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}, t)\right]\right| = O\left(\sqrt{\frac{\log n}{nh_1^d}} \right).
$$

The proof of Theorem 3.2 is achieved. □

Proof. **Proof of Theorem 3.3** First, observe that

$$
\begin{aligned}
\sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})| \quad \le \quad & \frac{1}{\inf_{\mathbf{x}\in\Omega} (l_n(\mathbf{x}))} \left\{ \sup_{x\in\Omega} \sup_{t\in\mathcal{C}} \left| \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}, t)\right] - F_1(\mathbf{x}, t)\right| \right. \\
& + \sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} \left| F_{1,n}(\mathbf{x}, t) - \tilde{F}_{1,n}(\mathbf{x}, t)\right| \\
& + \sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}, t)\right]\right| \\
& \left. + m_0^{-1} \sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} F_1(\mathbf{x}, t) \sup_{\mathbf{x}\in\Omega} |l_n(\mathbf{x}) - l(\mathbf{x})| \right\}.
\end{aligned}
$$

As $m_0 := \inf_x l(\mathbf{x})$, it is easily seen that

$$\begin{aligned}\frac{1}{l_n(\mathbf{x})} &\leq \frac{1}{l(\mathbf{x}) - |l_n(\mathbf{x}) - l(\mathbf{x})|} \\ &\leq \frac{1}{m_0 - \sup_{\mathbf{x}\in\Omega} |l_n(\mathbf{x}) - l(\mathbf{x})|}.\end{aligned}$$

This allows to write

$$\begin{aligned}&\sup_{\mathbf{x}\in\Omega}\sup_{t\in\mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})| \\ &\leq \frac{1}{m_0 - \sup_{\mathbf{x}\in\Omega} |l_n(\mathbf{x}) - l(\mathbf{x})|}\left\{\vartheta_1 + \vartheta_2 + \vartheta_3 + \vartheta_4 m_0^{-1} \sup_{\mathbf{x}\in\Omega}\sup_{t\in\mathcal{C}} F_1(\mathbf{x}, t)\right\}.\end{aligned}$$

As for the term ϑ_3, it has been bounded in Theorem 3.2. The following lemmas establish respectively the result of ϑ_1, ϑ_2. Then we apply Lemma 5.4 for ϑ_4.
The bias term ϑ_1 will be stated in Lemma 5.5 by using conditional expectation techniques and a Taylor expansion up to order 2 while Lemma 5.6 deals with bounding ϑ_2.

5.5. Lemma. *Under assumptions **A1**, **A2**, **A3** and **A5**, for n large enough we have*

$$\sup_{\boldsymbol{x}\in\Omega}\sup_{t\in\mathcal{C}} \left| \mathbb{E}\left[\tilde{F}_{1,n}(\boldsymbol{x}, t)\right] - F_1(\boldsymbol{x}, t)\right| = O(h_1^2 + h_2^2), \text{ a.s.}$$

Proof. The following proof does not depend on the dependence structure.

$$\begin{aligned}\mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}, t)\right] &= \frac{1}{h_1^d}\,\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)} K_d\left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1}\right) H\left(\frac{t - Y_1}{h_2}\right)\right] \\ &= \frac{1}{h_1^d}\,\mathbb{E}\left[K_d\left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1}\right)\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)} H\left(\frac{t - Y_1}{h_2}\right)|\mathbf{X}_1\right]\right].\end{aligned}$$

We use integration by parts, a change of variable and Assumption **A3**, then we have

$$\begin{aligned}\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)} H\left(\frac{t - Y_1}{h_2}\right)|\mathbf{X}_1\right] &= \mathbb{E}\left[\mathbb{E}\left[\frac{\delta_1}{\overline{G}(Y_1)} H\left(\frac{t - Y_1}{h_2}\right)|T_1\right]|\mathbf{X}_1\right] \\ &= \mathbb{E}\left[\frac{1}{\overline{G}(T_1)} H\left(\frac{t - T_1}{h_2}\right)\mathbb{E}\left[\mathbb{1}_{\{C_1 \geq T_1\}}\right]|\mathbf{X}_1\right] \\ &= \mathbb{E}\left[H\left(\frac{t - T_1}{h_2}\right)|\mathbf{X}_1\right] \\ &= \int_{\mathbb{R}} H\left(\frac{t - y}{h_2}\right) f(y|\mathbf{X}_1)dy \\ &= \int_{\mathbb{R}} H^{(1)}(z)F(t - zh_2|\mathbf{X}_1)dz.\end{aligned}$$

Again, by a change of variable we get

$$\begin{aligned}\mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x}, t)\right] &= \frac{1}{h_1^d}\,\mathbb{E}\left[K_d\left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1}\right)\int_{\mathbb{R}} H^{(1)}(z)F(t - zh_2|\mathbf{X}_1)dz\right] \\ &= \int_{\mathbb{R}^d} \frac{1}{h_1^d} K_d\left(\frac{\mathbf{x} - \mathbf{u}}{h_1}\right)\int_{\mathbb{R}} H^{(1)}(z)F(t - zh_2|\mathbf{X}_1 = \mathbf{u})l(\mathbf{u})\mathbf{du}dz \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}} \frac{1}{h_1^d} K_d\left(\frac{\mathbf{x} - \mathbf{u}}{h_1}\right) H^{(1)}(z)F_1(\mathbf{u}, t - zh_2)\mathbf{du}dz \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}} K_d(\mathbf{r})H^{(1)}(z)F_1(\mathbf{x} - \mathbf{r}h_1, t - zh_2)\mathbf{dr}dz.\end{aligned}$$

Then, expanding $F_1(\mathbf{x}-\mathbf{r}h_1, t-zh_2)$ up to order 2 around $(\mathbf{x},t)$ gives

$$\begin{aligned}
&F_1(\mathbf{x}-\mathbf{r}h_1, t-zh_2) = F_1(\mathbf{x},t)\\
&-h_1\left[r_1\frac{\partial F_1(\mathbf{x},t)}{\partial x_1}+\cdots+r_d\frac{\partial F_1(\mathbf{x},t)}{\partial x_d}\right]-h_2\left[z\frac{\partial F_1(\mathbf{x},t)}{\partial t}\right]\\
&+\frac{h_1^2}{2}\left[r_1^2\frac{\partial^2 F_1(\mathbf{x}^*,t^*)}{\partial x_1^2}+\cdots+r_d^2\frac{\partial^2 F_1(\mathbf{x}^*,t^*)}{\partial x_d^2}+2\sum_{i,j;i\neq j} r_i r_j\frac{\partial^2 F_1(\mathbf{x}^*,t^*)}{\partial x_i\partial x_j}\right]\\
&+\frac{h_2^2}{2}\left[z^2\frac{\partial^2 F_1(\mathbf{x}^*,t^*)}{\partial t^2}\right]+h_1h_2\left[r_1 z\frac{\partial^2 F_1(\mathbf{x}^*,t^*)}{\partial x_1\partial t}+\cdots+r_d z\frac{\partial^2 F_1(\mathbf{x}^*,t^*)}{\partial x_d\partial t}\right].
\end{aligned}$$

Here, $(\mathbf{x}^*,t^*)$ lies between $(\mathbf{x}-\mathbf{r}h_1, t-zh_2)$ and $(\mathbf{x},t)$. Finally, assumptions **A1**, **A2**, **A3** and **A5** entail

$$\sup_{\mathbf{x}\in\Omega}\sup_{t\in\mathcal{C}}\left|\,\mathbb{E}\left[\tilde{F}_{1,n}(\mathbf{x},t)\right]-F_1(\mathbf{x},t)\right|\leq c(h_1^2+h_2^2).$$

This provides the desired result. □

5.6. Lemma. *Under assumptions **A2**, **A4** and **A8**, for n large enough, we have*

$$\sup_{\boldsymbol{x}\in\Omega}\sup_{t\in\mathcal{C}}\left|F_{1,n}(\boldsymbol{x},t)-\tilde{F}_{1,n}(\boldsymbol{x},t)\right|=o\left(n^{-\theta}\right),\ \text{a.s.}$$

Proof. Firstly, we have

$$\begin{aligned}
&|F_{1,n}(\mathbf{x},t)-\tilde{F}_{1,n}(\mathbf{x},t)|\\
&=\left|\frac{1}{nh_1^d}\sum_{i=1}^n\delta_i K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_1}\right)H\left(\frac{t-Y_i}{h_2}\right)\left(\frac{1}{\overline{G}_n(Y_i)}-\frac{1}{\overline{G}(Y_i)}\right)\right|\\
&\leq\frac{1}{nh_1^d}\sum_{i=1}^n K_d\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_1}\right)H\left(\frac{t-Y_i}{h_2}\right)\left|\frac{1}{\overline{G}_n(Y_i)}-\frac{1}{\overline{G}(Y_i)}\right|\\
&\leq\frac{l_n(\mathbf{x})}{\overline{G}_n(\tau)\overline{G}(\tau)}\sup_{t\in\mathcal{C}}\left|\overline{G}_n(t)-\overline{G}(t)\right|.
\end{aligned}$$

Then, following Theorem 1.4 in Cai and Roussas [5] and for n large enough, we easily get

$$\sup_{t\in\mathcal{C}}\left|\overline{G}_n(t)-\overline{G}(t)\right|=o\left(n^{-\theta}\right),\ \text{a.s.}\tag{5.14}$$

Recall that $0<\theta<\gamma/(2\gamma+9+3/2\kappa)$ for any real $\kappa>0$ and γ is referred in Assumption **A8**. Hence, Lemma 5.4 and (5.14) end the proof of Lemma 5.6. □

To end the proof of Theorem 3.3, it suffices to apply Lemma 5.4 for ϑ_4. □

The last step consists in proving the result on the behavior of the conditional quantile estimator $\xi_{p,n}(\mathbf{x})$, stated in Corollary 3.4.

Proof. **Proof of Corollary 3.4** It suffices to use the following triangular inequality jointly with basic arguments. Let $\mathbf{x}\in\Omega$, then we have

$$\begin{aligned}
|F(\xi_{p,n}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})| &\leq |F_n(\xi_{p,n}(\mathbf{x})|\mathbf{x})-F(\xi_{p,n}(\mathbf{x})|\mathbf{x})|\\
&+\ |F_n(\xi_{p,n}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})|\\
&\leq 2\sup_{t\in\mathcal{C}}|F_n(t|\mathbf{x})-F(t|\mathbf{x})|.
\end{aligned}\tag{5.15}$$

So, the first part of Corollary 3.4 is straightforwardly deduced from Theorem 3.3. And, a Taylor expansion of $F(\xi_{p,n}|\mathbf{x})$ in the neighborhood of $\xi_p(\mathbf{x})$ permits to get

$$F(\xi_{p,n}(\mathbf{x})|\mathbf{x}) - F(\xi_p(\mathbf{x})|\mathbf{x}) = (\xi_{p,n}(\mathbf{x}) - \xi_p(\mathbf{x}))f(\xi_p^*(\mathbf{x})|\mathbf{x}),$$

where $\xi_p^*(\cdot)$ is between $\xi_p(\cdot)$ and $\xi_{p,n}(\cdot)$. Thus from (5.15), we obtain

$$\sup_{\mathbf{x}\in\Omega} |\xi_{p,n}(\mathbf{x}) - \xi_p(\mathbf{x})| f(\xi_p^*(\mathbf{x})|\mathbf{x}) \leq 2 \sup_{\mathbf{x}\in\Omega} \sup_{t\in\mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})|.$$

Note that if the condition in Corollary 3.4 is not checked, one has to consider a higher order-Taylor expansion. Finally, the desired result follows immediately from Assumption **A5** and Theorem 3.3. □

References

[1] Abberger, K. *Quantile smoothing in financial time series*, Statist. Papers **38**, 125–148, 1997.

[2] Bhattacharya, P.K. and Gangopadhyay, A. *Kernel and nearest neighbour estimation of conditional quantile*, Ann. Statist. **18**, 1400–1415, 1990.

[3] Bulinski, A. and Shashkin, A. *Limit theorems for associated random fields and related systems*, Vol **10**, Advanced series on statistical science & applied probability, World Scientific, Singapore, 2007.

[4] Cai, Z. *Asymptotic properties of Kaplan-Meier estimator for censored dependent data*, Statist. Probab. Lett. **37**, 381–389, 1998.

[5] Cai, Z. and Roussas, G.G. *Kaplan-Meier estimator under association*, J. Multivariate Anal. **67**, 318–348, 1998.

[6] Chaudhuri, P., Doksum, K. and Samarov, A. *On average derivative quantile regression*, Ann. Statist. **25**, 715–744, 1997.

[7] Dabrowska, D. *Nonparametric quantile regression with censored data*, Sankhya A **54**, 252–259, 1992.

[8] Doukhan, P. *Mixing: Properties and Examples*, Lect. Notes in Statist. **61**, Springer Verlag, 1994.

[9] Doukhan, P. and Louhichi, S. *A new weak dependence condition and applications to moment inequalities*, Stochastic Process. Appl. **84**, 313–342, 1999.

[10] Doukhan, P. and Neumann, M. *Probability and moment inequalities for sums of weakly dependent random variables, with applications*, Stochastic Process. Appl. **117**, 878–903, 2007.

[11] Esary, J., Proschan, F. and Walkup, D. *Association of random variables with applications*, Ann. Math. Statist. **38**, 1466–1476, 1967.

[12] Honda, T. *Nonparametric estimation of a conditional quantile for strong mixing processes*, Ann. Inst. Statist. Math. **52**, 459–470, 2000.

[13] Kaplan, E.M. and Meier, P. *Nonparametric estimation from incomplete observations*, J. Amer. Statist. Assoc. **53**, 457–481, 1958.

[14] Koul, H., Susarla, V. and Van Ryzin, J. *Regression analysis with randomly right censored data*, Ann. Statist. **9**, 1276–1288, 1981.

[15] Liang, H.-Y. and de Uña-Álvarez, J. *Asymptotic properties of conditional quantile estimator for censored dependent observations*, Ann. Inst. Statist. Math. **63**, 267–289, 2011.

[16] Mehra, K.L., Rao M.S. and Upadrasta S.P. *A smooth conditional quantile estimator and related applications of conditional empirical processes*, J. Multivariate Anal. **37**, 151–179, 1991.

[17] Menni, N. and Tatachak, A. *A note on estimating the conditional expectation under censoring and association: strong uniform consistency*, Statist. Papers, DOI:10.1007/s00362-016-0801-8, 2016.

[18] Newman, C.M. *Asymptotic independence and limit theorems for positively and negatively dependent random variables*, in: Tong, Y.L. (Ed.) Inequalities in statistics and probability, IMS Lecture Notes-Monograph Series, vol. **5**, Hayward, CA, 127–140, 1984.

[19] Oliveira, P.E. *Asymptotics for Associated Random Variables*, Springer Verlag, 2012.

[20] Ould Saïd, E. *A strong uniform convergence rate of kernel conditional quantile estimator under random censorship*, Statist. Probab. Lett. **76**, 579–586, 2006.

[21] Ould Saïd, E. and Sadki, O. *Prediction via the conditional quantile for right censorship model*, Far East Journal of Theoretical Statistics **25**, 145–179, 2008.

[22] Ould Saïd, E. and Sadki, O. *Strong approximation of quantile function for strong mixing and censored processes*, Comm. Statist. Theory Methods **34**, 1449–1459, 2005.

[23] Prakasa Rao, B.L.S. *Associated Sequences, Demimartingales and Nonparametric Inference*, Probability and its Applications, Springer Basel AG, 2012.

[24] Qin, Y.S. and Wu, Y. *An estimator of a conditional quantile in the presence of auxiliary information*, J. Statist. Plann. Inference **99**, 59–70, 2001.

[25] Samanta, M. *Nonparametric estimation of conditional quantiles*, Statist. Probab. Lett. **7**, 407–412, 1989.

[26] Stute, W. and Wang, J.L. *The strong law under random censorship*, Ann. Statist. **21**, 1591–1607, 1993.

[27] Xiang, X. *A kernel estimator of a conditional quantile*, J. Multivariate Anal. **59**, 206-216, 1996.

[28] Yu, K. and Jones, M.C. *Local linear quantile regression*, J. Amer. Statist. Assoc. **93**, 228–238, 1998.