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: wjelladj@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 $X$ taking values in $\mathbb{R}^d$. Under this model, the observable sequence is $\{(Y_i, \delta_i, X_i), 1 \leq i \leq n\}$, with $Y_i = \min(T_i, C_i)$, $\delta_i = 1_{\{T_i \leq C_i\}}$ and where $1_A$ 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 $\{(X_i, T_i), 1 \leq i \leq n\}$. 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 $F(\cdot | x)$ of $(T|X = x)$ is defined by

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

where $f(\cdot, \cdot)$ is the joint probability density function (pdf) of $(X, T)$, $l(\cdot)$ is the marginal pdf of $X$ and $F_1(x, \cdot)$ is the first derivative of the joint df $F(x, \cdot)$ with respect to $x$.

The conditional pdf will be denoted by $f(\cdot | x)$. Then, for all fixed $p \in (0, 1)$, the $p$-th conditional quantile of $T$ given $X = x$ is defined by

$$(1.1) \quad \xi_p(x) := \inf\{t, F(t|x) \geq p\}.$$
Recall that a set of finite family of rv’s \((T_1, \ldots, T_n)\) are said to be associated if for all non-decreasing functions \(\Psi_1, \Psi_2\)
\[
\text{Cov}(\Psi_1(T_1, \ldots, T_n), \Psi_2(T_1, \ldots, T_n)) \geq 0,
\]
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 
\(
F(t|x)
\)
is given by
\begin{equation}
F_n(t|x) = \frac{1}{n} \sum_{i=1}^{n} \omega_{in}(x) \mathbb{1}_{\{Y_i \leq t\}},
\end{equation}
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
\[
\omega_{in}(x) = \frac{K_d \left( \frac{x - X_i}{h_{n,1}} \right)}{\sum_{j=1}^{n} K_d \left( \frac{x - X_j}{h_{n,1}} \right)} = \frac{1}{nh_{n,1}} K_d \left( \frac{x - X_i}{h_{n,1}} \right) l_n(x),
\]
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
\begin{equation}
\omega_{in}(x) = \frac{1}{nh_{n,1}^d C(Y_i) l_n(x)} K_d \left( \frac{x - X_i}{h_{n,1}} \right). \tag{2.2}
\end{equation}

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
\[
F_{T,n}(t) = 1 - \prod_{i=1}^{n} \left[ 1 - \frac{\delta(i)}{n - i + 1} \right] \mathbb{1}_{\{Y_{(i)} \leq t\}}
\]
and
\[
G_n(t) = 1 - \prod_{i=1}^{n} \left( 1 - \frac{1 - \delta(i)}{n - i + 1} \right) I\{Y(i) \leq t\},
\]
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 \(\alpha\)-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 \(F_n(\cdot)\), namely
\[
F_n(t|X) = \frac{1}{n h_n} \sum_{i=1}^{n} \frac{\delta_i}{G_n(Y_i)} H\left( \frac{t - Y_i}{h_n} \right),
\]
was also considered and studied (strong consistency and asymptotic normality) in the iid case by Ould Saïd and Sadki [21]. Here, the bandwidth \(h_n\) 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 version of (2.3)), 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
\[
\xi_{p,n}(x) = \inf\{t, F_n(t|x) \geq p\}.
\]

To argue our main results, the following auxiliary pseudo-estimator will be of a great benefit in proving our results
\[
\tilde{F}_n(t|X) = \frac{1}{n h_n^2} \sum_{i=1}^{n} \frac{\delta_i}{H_n(Y_i)} H\left( \frac{t - Y_i}{h_n} \right),
\]
Note that (2.5) can not be computed since \(H_n(\cdot)\) is assumed to be unknown.

3. Assumptions and main results

In the sequel, \(\epsilon\) stands for a positive constant taking different values and \(\tau\) will denote a positive real number satisfying \(\tau < \tau_p < \tau_0\) where, for any df \(W, \tau_W := \sup\{y; W(y) < 1\}\). Define \(\Omega_0 = \{x \in \mathbb{R}^d/l(x) \geq m_0 := \inf_x l(x) > 0\}\) and let \(\Omega\) and \(\mathbb{C}\) be compact sets included in \(\Omega_0\) and \([0, \tau], \) respectively. The main results will be stated using the following assumptions:

**A1.** The bandwidths \(h_1\) and \(h_2\) satisfy
(i) \(h_1 \to 0\), \(nh_1^{2+\alpha(1-\alpha)} \to +\infty\) and \(nh_1^\alpha \to 0\) as \(n \to +\infty\),
(ii) \(h_2 \to 0\) and \(nh_2^\alpha h_1^\beta \to +\infty\) 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) \(\int_{\mathbb{R}^d} u_j K_d(u) du = 0\), for all \(j = 1, \ldots, d\), where \(u = (u_1, \ldots, u_d)^T\);
A3. The function $H$ in (2.3) is of class $C^4$. Furthermore, its derivative $H^{(1)}$ is assumed to be compactly supported and satisfies the properties of a second order kernel;

A4. The marginal density $f(.)$ is bounded and twice differentiable with:

\[
\sup_{x \in \Omega} \left| \frac{\partial^6 f(x)}{\partial x_i \partial x_j} \right| < \infty \text{ for } i, j = 1, \ldots, d \text{ and } k = 1, 2;
\]

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

A6. The joint pdf $h_1,\ldots,(\ldots)$ of $(X_i, X_j)$ is bounded;

A7. The joint pdf $f(\ldots, \ldots, \ldots)$ of $(X_i, X_j, Y_j)$ is bounded;

A8. Let us define $\Lambda_{ij}$ as follows:

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

with $X_i^k$ the $k$-th component of $X_i$, 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 $\varsigma(\mathbf{x}) = \int_{\Omega} \frac{1}{\gamma(x,v)} f(\mathbf{x},v)dv$ is bounded, continuously differentiable and

\[
\sup_{x \in \Omega} \left| \frac{\partial \varsigma}{\partial x_i} \right| < \infty \text{ for } i = 1, \ldots, d.
\]

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 $h_2^\gamma$. 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|. \text{ 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|\mathbf{x}) \text{ toward } F(t|\mathbf{x}), \text{ 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

\[
(3.1) \quad \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathbb{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[ \tilde{F}_{1,n}(\mathbf{x}, t) \right] \right| = O \left( \sqrt{\frac{\log n}{nh_1^2}} \right), \text{ a.s.}
\]

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

\[
\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathbb{C}} \left| F_n(t|\mathbf{x}) - F(t|\mathbf{x}) \right| = O \left( \left( h_1^2 + h_2^2 \right) + n^{-\theta} \right), \text{ a.s.}
\]

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 \( \inf_{\mathbf{x} \in \Omega} f(\xi_p(\mathbf{x})|\mathbf{w}) > 0 \), then for \( n \) large enough, we have

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

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

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

3.6. Remark. We point out that the rate in Corollary 3.4 depends upon the parameter \( \theta \) pertaining to the association dependence. In addition, remark that for \( \gamma \) large enough, the parameter \( \theta \) approaches its upper bound \( (\theta-1/2) \) 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 \( \varepsilon_i \) from normal distribution \( (\varepsilon_i \sim N(0, 0.01)) \);
- Given \( Z_i \), generate an \( n \) associated sequence \((X_i, T_i)\) as follows:
  a) Linear case
  \[
  \begin{align*}
  X_i &= \exp(Z_{i-1} + Z_{i-2})/2; \\
  T_i &= 3X_i/2 + 0.45 \varepsilon_i.
  \end{align*}
  \]
  b) Nonlinear case
  \[
  \begin{align*}
  X_i &= \exp(Z_{i-1} + Z_{i-2})/2; \\
  T_i &= \log(3X_i/2) + 0.45 \varepsilon_i,
  \end{align*}
  \]
- Generate \( n \) iid rv’s \( C_i \) from exponential distribution \( (C_i \sim \exp(\lambda)) \). The parameter \( \lambda \) is adjusted according to the \( CR \)'s values;
- Keep the observed data \( \{(Y_i := \min(T_i, C_i)), X_i, (\delta_i := 1\{T_i \leq C_i\})\} \).

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
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

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

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 $x_{\ell}$ 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 $N(3x/2, 0.0045)$. 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>
<thead>
<tr>
<th>Linear case</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 300$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CR = 10%$</td>
<td>0.0637</td>
<td>0.0245</td>
<td>0.0069</td>
</tr>
<tr>
<td>$CR = 25%$</td>
<td>0.1591</td>
<td>0.0586</td>
<td>0.0113</td>
</tr>
<tr>
<td>$CR = 40%$</td>
<td>0.2465</td>
<td>0.1059</td>
<td>0.0128</td>
</tr>
</tbody>
</table>

#### 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
4.2.2. Non-linear case: Note that the rv $(T|X = x)$ follows $N(\log(3x/2), 0.0045)$ and the choice of the log function permits to preserve the association property by monotonicty.

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>
<thead>
<tr>
<th>CR</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 300$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>$24 \times 10^{-3}$</td>
<td>$15 \times 10^{-3}$</td>
<td>$5.54 \times 10^{-4}$</td>
</tr>
<tr>
<td>25%</td>
<td>$69 \times 10^{-3}$</td>
<td>$25 \times 10^{-3}$</td>
<td>$8.23 \times 10^{-4}$</td>
</tr>
<tr>
<td>40%</td>
<td>$11 \times 10^{-2}$</td>
<td>$51 \times 10^{-3}$</td>
<td>$16 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

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
\[
\Delta_i(x, t) = \frac{\delta_i}{G(Y_i)} K_d \left( \frac{x - X_i}{h_1} \right) H \left( \frac{t - Y_i}{h_2} \right) - \mathbb{E} \left[ \frac{\delta_1}{G(Y_1)} K_d \left( \frac{x - X_1}{h_1} \right) H \left( \frac{t - Y_1}{h_2} \right) \right],
\]
for all $i = 1, ..., n$. It is easily seen that
\[
\tilde{F}_{1,n}(x, t) - \mathbb{E} \left[ \tilde{F}_{1,n}(x, t) \right] = \frac{1}{nh_1} \sum_{i=1}^{n} \Delta_i(x, t).
\]

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 $\Delta_1(x, t), \Delta_2(x, t), ..., \Delta_n(x, t)$ 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
\[
\begin{align*}
&\text{a) } \text{Cov} \left( \prod_{i=s_1}^{s_p} \Delta_i(x, t), \prod_{j=v_1}^{v_q} \Delta_j(x, t) \right) \leq c^{s_1+q}h_1^d h_2^q \frac{2^{2r}}{pq} \Upsilon(v_1 - s_p), \\
&\text{b) } \sum_{s=0}^{\infty} (s+1)^{\lambda_0} \Upsilon(s) \leq L_1 L_2^B (k_0!)^\mu, \forall k_0 \geq 0,
\end{align*}
\]
c) $\mathbb{E} [\left| \Delta_i(x,t) \right|^k] \leq (k_0!)^c M^{k_0}$.

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) $\text{Cov} \left( \prod_{i=1}^{s_p} \Delta_i(x,t), \prod_{j=1}^{s_q} \Delta_j(x,t) \right) =: C_1 \leq c^{p+q} h_1^{-2} h_2^{-2} p q \rho (1 - s_p)$,

ii) $\text{Cov} \left( \prod_{i=1}^{s_p} \Delta_i(x,t), \prod_{j=1}^{s_q} \Delta_j(x,t) \right) =: C_2 \leq c^{p+q} h_1^{2d} h_2^2$.

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

\begin{equation}
\text{Lip}_i(\Phi_m) = \sup_{z_1, \ldots, z_m \in \mathbb{R}} \frac{\left| \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) \right|}{|z_i - z_i'|},
\end{equation}

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

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

with $\|(z_1, \ldots, z_m)\|_1 = |z_1| + \ldots + |z_n|$.

To prove part (i) in Lemma 5.2, we use Theorem 5.3, p.89 in (Bulinski and Shashkin [3]). Firstly, we set

$\Phi_p := \prod_{i=1}^{s_p} \Delta_i$ and $\Phi_q := \prod_{j=1}^{s_q} \Delta_j$.

Then, using the fact that $K_2$, $H$ and $G$ are Lipschitz functions, we have

$\text{Cov}(\Phi_p, \Phi_q) \leq \sum_{i=1}^{s_p} \sum_{j=1}^{s_q} \text{Lip}_i(\Phi_p) \text{Lip}_j(\Phi_q) A_{ij}$,

The definition in (5.2) leads to

$\text{Lip}_i(\Phi_p) \leq \frac{M_0}{h_1 h_2} \left( \frac{2}{G(\tau)} \right)^p \|K_2\|_{\infty}^{p-1}$

and

$\text{Lip}_j(\Phi_q) \leq \frac{M_0}{h_1 h_2} \left( \frac{2}{G(\tau)} \right)^q \|K_2\|_{\infty}^{q-1}$,

where $M_0 = \max \left\{ h_2 \text{Lip}(K) \|K\|_{\infty}^{d-1}, h_1 \left( \text{Lip}(H) + h_2 \frac{\text{Lip}(G)}{G(\tau)} \right) \|K_2\|_{\infty} \right\}$.

Note that the partial Lipschitz constants are obtained as follows

$\text{Lip}_i(\Phi_p) \leq \frac{M_0}{h_1 h_2} \left( \frac{2}{G(\tau)} \right)^{p-1} \|K_2\|_{\infty}^{p-1} \frac{1}{G(\tau)}$

$\leq \frac{M_0}{h_1 h_2} \left( \frac{2}{G(\tau)} \right)^p \|K_2\|_{\infty}^{p-1}$.
If Assumption A8 holds, by stationarity we get

\[
\text{Cov}(\Phi_p, \Phi_q) \leq \frac{M^2}{h_1^2 h_2^2} \left( \frac{2}{G(\tau)} \right)^{p+q} \| K_d \|_{\text{op}}^{p+q-2} \sum_{i=1}^{s_1} \sum_{j=1}^{s_2} \Lambda_{ij}
\]

\[
\leq \frac{c^{p+q}}{h_1^2 h_2^2} \rho q(v_1 - s_p). 
\]

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

\[
\mathbb{E}[\delta_i \delta_j | T_i, T_j] = \mathbb{E} \left[ \mathbb{1}_{(T_i \leq C_1)} \mathbb{1}_{(T_j \leq C_2)} | T_i, T_j \right] = \overline{G}(T_i) \overline{G}(T_j).
\]

We also use the following simplified notations

\[
K_{d,x,i} := K_d \left( \frac{x - X_i}{h_1} \right) \quad \text{and} \quad H_{t,i} := H \left( \frac{t - Y_i}{h_2} \right).
\]

Indeed, we have

\[
\text{Cov}(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))
\]

\[
= \mathbb{E} \left[ \frac{\delta_i \delta_j}{G(Y_i) G(Y_j)} K_{d,x,i} H_{t,i} K_{d,x,j} H_{t,j} \right] 
\]

\[- \mathbb{E} \left[ \frac{\delta_i}{G(Y_i)} K_{d,x,i} H_{t,i} \right] \times \mathbb{E} \left[ \frac{\delta_j}{G(Y_j)} K_{d,x,j} H_{t,j} \right] 
\]

\[
= \mathbb{E} \left[ K_{d,x,i} K_{d,x,j} \mathbb{E} \left[ \frac{\delta_i}{G(Y_i)} H_{t,i} | \mathbf{X}_i, \mathbf{X}_j \right] \right] 
\]

\[- \mathbb{E} \left[ K_{d,x,i} \mathbb{E} \left[ \frac{\delta_i}{G(Y_i)} H_{t,i} \right] \right] \times \mathbb{E} \left[ K_{d,x,j} \mathbb{E} \left[ \frac{\delta_j}{G(Y_j)} H_{t,j} | \mathbf{X}_i, \mathbf{X}_j \right] \right] 
\]

\[- \mathbb{E} \left[ K_{d,x,i} H_{t,i} \mathbb{E} \left[ \frac{\delta_i}{G(Y_i)} | \mathbf{X}_i \right] \right] \times \mathbb{E} \left[ K_{d,x,j} \mathbb{E} \left[ \frac{\delta_j}{G(Y_j)} H_{t,j} | \mathbf{X}_i, \mathbf{X}_j \right] \right] 
\]

\[- \mathbb{E} \left[ K_{d,x,i} H_{t,i} \mathbb{E} \left[ \frac{\delta_i}{G(Y_i)} H_{t,i} | \mathbf{X}_i \right] \right] \times \mathbb{E} \left[ K_{d,x,j} \mathbb{E} \left[ \frac{\delta_j}{G(Y_j)} | \mathbf{X}_i, \mathbf{X}_j \right] \right] 
\]

Then, we get

\[
|\text{Cov}(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))|
\]

\[
\leq \left| \int_{\mathbb{R}^{2d+2}} K_d \left( \frac{x - u}{h_1} \right) H \left( \frac{t - s}{h_2} \right) K_d \left( \frac{x - r}{h_1} \right) H \left( \frac{t - v}{h_2} \right) f(u, s, r, v) \, du \, ds \, dr \, dv \right|
\]

\[
+ \left| \int_{\mathbb{R}^{2d+1}} K_d \left( \frac{x - u}{h_1} \right) H \left( \frac{t - s}{h_2} \right) f(u, s) \, du \, ds \right|
\]

\[
\times \left| \int_{\mathbb{R}^{2d+1}} K_d \left( \frac{x - r}{h_1} \right) H \left( \frac{t - v}{h_2} \right) f(r, v) \, dr \, dv \right|.
\]

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

\[
(5.3) \quad |\text{Cov}(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))| = O(h_1^{2d} h_2^2).
\]

Finally, the second part of Lemma 5.2 follows by simple algebra. \qed
We need some auxiliary notations to set up the proof of Proposition 5.1. Impose $\Upsilon(.) = \rho^{\frac{d}{2(d+2)}}(.)$ and use the upper bounds of Lemma 5.2, namely
\begin{align}
C_1^{\frac{d}{2(d+2)}} & \leq c^{\frac{d}{2(d+2)}} h_1^{\frac{2d}{2(d+2)}} h_2^{\frac{2d}{2(d+2)}} (pq)^{\frac{d}{2(d+2)}} \rho^{\frac{d}{2(d+2)}} (v_1 - s_p), \\
C_2^{\frac{d+2}{2(d+2)}} & \leq c^{\frac{(d+2)}{2(d+2)}} h_1^{\frac{2(d+2)}{2(d+2)}} h_2^{\frac{2(d+2)}{2(d+2)}}.
\end{align}
Combining (5.4) and (5.5), we get
\begin{align*}
C_1^{\frac{d}{2(d+2)}} C_2^{\frac{d+2}{2(d+2)}} & \leq c^{\frac{d}{2(d+2)}} h_1^{\frac{d}{2(d+2)}} (pq)^{\frac{d}{2(d+2)}} \rho^{\frac{d}{2(d+2)}} (v_1 - s_p) \\
& \leq c^{\frac{d}{2(d+2)}} h_1^{\frac{d}{2(d+2)}} (pq)^{\frac{d}{2(d+2)}} \Upsilon(v_1 - s_p).
\end{align*}
This inequality concludes the proof of part (a) of Proposition 5.1. Next, under Assumption A8 and choosing $\lambda = 0$, $\mu = 1$, $L_1 = L_2 = \frac{1}{1 - \epsilon^{2/3}}$, 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 $\hat{F}_{1,n}(x,t) - \mathbb{E}\left[\hat{F}_{1,n}(x,t)\right]$, we apply the triangular inequality and classical techniques to cover compacts. So, $\Omega$ can be covered by a finite number $d_{\epsilon,n}$ of balls $B_k(x_k, a_{d_k})$ centred at $x_k = (x_{k,1}, ..., x_{k,d})$ and $\mathcal{C}$ is split into $d_{\epsilon,n}$ subintervals $J_1, ..., J_{d_{\epsilon,n}}$ of lengths $b_n$, centred at $t_k$. In other words, for all $x \in \Omega$, $t \in \mathcal{C}$, there exist integers $k \in \{1, ..., d_{\epsilon,n}\}$ and $\ell \in \{1, ..., d_{\epsilon,n}\}$ such that $|x - x_k| \leq a_{d_k}$ and $|t - t_k| \leq b_n$, with $a_{d} = (n^{-1}h_2^2 + d_{\epsilon,n})^{-1/2}$ and $b_n = (nh_1^2)^{-1/2}$. Then, as $\Omega$ and $\mathcal{C}$ are bounded, let $m_1$ and $m_2$ be positive constants satisfying $d_{\epsilon,n}a_{d_k} \leq m_1$ and $d_{\epsilon,n}b_n \leq m_2$.

5.3. Remark. In proving our results we will use Lemma 5.4 stated in Menni and Tatchak [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
\[ \sup_{x \in \Omega} |l_n(x) - l(x)| = O \left( \max \left\{ \frac{\log n}{nh_1^2}, h_1^2 \right\} \right) \text{ a.s.} \]

Next, using basic arguments, the left hand side in (3.1) is upper bounded as follows
\[ \sup_{x \in \Omega, t \in \mathcal{C}} \left| \hat{F}_{1,n}(x,t) - \mathbb{E}\left[\hat{F}_{1,n}(x,t)\right]\right| \leq I_{1n} + I_{1n} + I_{2n} + I_{2n} + I_{3n}, \]
with
\begin{align*}
I_{1n} &= \max_{1 \leq k \leq d_{\epsilon,n}} \sup_{x \in B_k(t_\mathcal{C})} \left| \hat{F}_{1,n}(x,t) - \hat{F}_{1,n}(x_k,t) \right|, \\
I_{1n} &= \max_{1 \leq k \leq d_{\epsilon,n}} \sup_{x \in B_k(t_\mathcal{C})} \left| \mathbb{E}\left[\hat{F}_{1,n}(x_k,t)\right] - \mathbb{E}\left[\hat{F}_{1,n}(x,t)\right] \right|, \\
I_{2n} &= \max_{1 \leq k \leq d_{\epsilon,n}} \max_{1 \leq \ell \leq d_{\epsilon,n}} \sup_{t \in \mathcal{C}} \left| \hat{F}_{1,n}(x_k,t) - \hat{F}_{1,n}(x_k,t_\ell) \right|, \\
I_{2n} &= \max_{1 \leq k \leq d_{\epsilon,n}} \max_{1 \leq \ell \leq d_{\epsilon,n}} \sup_{t \in \mathcal{C}} \left| \mathbb{E}\left[\hat{F}_{1,n}(x_k,t_\ell)\right] - \mathbb{E}\left[\hat{F}_{1,n}(x_k,t)\right] \right|, \\
I_{3n} &= \max_{1 \leq k \leq d_{\epsilon,n}} \max_{1 \leq \ell \leq d_{\epsilon,n}} \sup_{t \in \mathcal{C}} \left| \mathbb{E}\left[\hat{F}_{1,n}(x_k,t_\ell)\right] - \mathbb{E}\left[\hat{F}_{1,n}(x_k,t)\right] \right|.
\end{align*}
Concerning $I_{1n}$ and $I'_{1n}$, we apply the SLLN for associated sequences (see Newman [18]) and Assumption A2(i). We obtain

$$\left| \tilde{F}_{1,n}(x,t) - \tilde{F}_{1,n}(x_k,t) \right|$$

$$= \left| \frac{1}{nh_1^d} \sum_{i=1}^{n} \delta_i \frac{H(t-Y_i)}{h_2} \left[ K_d \left( \frac{x-X_i}{h_1} \right) - K_d \left( \frac{x_k-X_i}{h_1} \right) \right] \right|$$

$$\leq \frac{c}{G(\tau)} \frac{\|x-x_k\|}{h_1^\alpha} \frac{1}{n} \sum_{i=1}^{n} \delta_i$$

$$\leq \frac{c}{G(\tau)} \frac{\|x-x_k\|}{h_1^\alpha} \frac{1}{n} \sum_{i=1}^{n} \delta_i$$

$$= O \left( \frac{1}{\sqrt{nh_1^d}} \right).$$

(5.6)

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

$$\left| \tilde{F}_{1,n}(x_k,t) - \tilde{F}_{1,n}(x_k,t_\ell) \right|$$

$$= \left| \frac{1}{nh_1^d} \sum_{i=1}^{n} \delta_i \frac{H(t-Y_i)}{h_2} \left[ K_d \left( \frac{x_k-X_i}{h_1} \right) - K_d \left( \frac{x_k-X_i}{h_1} \right) \right] \right|$$

$$\leq \frac{c}{G(\tau)} \frac{\|t-t_\ell\|}{h_2} \frac{1}{nh_1^d} \sum_{i=1}^{n} \delta_i$$

$$\leq \frac{c}{G(\tau)} \frac{h_n}{h_2} I_n(x_k)$$

$$= O \left( \frac{1}{\sqrt{nh_1^d}} \right).$$

(5.7)

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

$$P \left( \sum_{i=1}^{n} \Delta_i(x_k,t_\ell) \geq \varepsilon \right) \leq \exp \left( - \frac{\varepsilon^2 / 2}{A_n + B_n^{1/(\mu+\lambda+2)} \varepsilon^{(2\mu+2\lambda+4)/(\mu+\lambda+2)}} \right),$$

(5.8)

where $A_n$ is any number greater than $\sigma_n^2$ and

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

$$B_n = 2\varepsilon L_2 \max \left( \frac{2^{4+\mu+\lambda} \cosh h_1^d h_2^2 L_1}{A_n}, 1 \right).$$

(5.9)
Firstly, we have to calculate \( \sigma_n^2 \). Indeed \( \sigma_n^2 = (nh_1^d)^2 \text{Var}\left( \tilde{F}_{1,n}(x_k, t_i) \right) \). On the other hand, we have

\[
(nh_1^d)^2 \text{Var}\left( \tilde{F}_{1,n}(x_k, t_i) \right) = n \text{Var}\left( \frac{\delta_i}{G(Y_i)} K_{d,x_k,i} H_{t_i,1}(\frac{t_i - Y_i}{h_2}) \right) \\
+ \sum_{i=1}^n \sum_{j=1, j \neq i}^n \text{Cov}\left( \frac{\delta_i}{G(Y_i)} K_{d,x_k,i} H_{t_i,j}, \frac{\delta_j}{G(Y_j)} K_{d,x_k,j} H_{t_j,i} \right) \\
=: V + S.
\]

Firstly, let us focus on \( V \).

\[
V = n \mathbb{E}\left[ \frac{\delta_i}{G(Y_i)^2} K_d^2 \left( \frac{x_k - X_1}{h_1} \right) H^2 \left( \frac{t_i - Y_i}{h_2} \right) \right] \\
- n \mathbb{E}^2\left[ \frac{\delta_i}{G(Y_i)} K_d \left( \frac{x_k - X_1}{h_1} \right) H \left( \frac{t_i - Y_i}{h_2} \right) \right] \\
= n(D_1 - D_2).
\]

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 \( x_k \), we get

\[
D_1 = \mathbb{E}\left[ K_d^2 \left( \frac{x_k - X_1}{h_1} \right) \mathbb{E}\left[ \frac{\delta_i}{G(Y_i)^2} H^2 \left( \frac{t_i - Y_i}{h_2} \right) | T_1 \right] | X_1 \right] \\
= \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h_1} \right) \mathbb{E}\left[ H^2 \left( \frac{t_i - T_1}{h_2} \right) \right] \frac{1}{G(T_1)} l(u) du \\
\leq \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h_1} \right) \int_{\mathbb{R}} \frac{1}{G(v)} f(u,v) dv du, \quad \text{because } H(.) \text{ is a df;}
\]

\[
D_1 = \int_{\mathbb{R}^d} K_d^2 \left( \frac{x_k - u}{h_1} \right) \int_{\mathbb{R}} \frac{1}{G(v)} f(u,v) dv du, \quad \text{because } H(.) \text{ is a df;}
\]

\[
D_1 = h_1^d \int_{\mathbb{R}^d} K_d^2(z) \left( \frac{x_k - z h_1}{h_1} \right) dz \\
= h_1^d \int_{\mathbb{R}^d} \zeta(x_k) K_d^2(z) dz - h_1^{d+1} \int_{\mathbb{R}^d} K_d^2(z) \left[ z_1 \frac{\partial \zeta(x_k)}{\partial x_{k,1}} + \cdots + z_d \frac{\partial \zeta(x_k)}{\partial x_{k,d}} \right] dz \\
= O(h_1^{2d}).
\]

Here \( \zeta(x_k) \) is between \( x_k - z h_1 \) and \( x_k \). 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

\[
D_2 = \mathbb{E}^2\left[ K_d \left( \frac{x_k - X_1}{h_1} \right) \mathbb{E}\left[ \frac{\delta_i}{G(Y_i)} H \left( \frac{t_i - Y_i}{h_2} \right) | T_1 \right] | X_1 \right] \\
= O(h_1^{2d}).
\]

Consequently, we get

\[
V = O(nh_1^d).
\]

Secondly, to evaluate \( S \), we first define

\[
B_1 = \{(i,j); 1 \leq |i - j| \leq \eta_n \} \quad \text{and} \quad B_2 = \{(i,j); \eta_n + 1 \leq |i - j| \leq n - 1 \},
\]
where $\eta_n = o(n)$. Then, we have

$$S = \sum_{i=1}^{n} \sum_{B_1} \text{Cov} \left( \frac{\delta_i}{G(Y_i)} K_{d,x_{k,i}} H_{t_i,i}, \frac{\delta_j}{G(Y_j)} K_{d,x_{k,j}} H_{t_i,j} \right)$$

$$+ \sum_{i=1}^{n} \sum_{B_2} \text{Cov} \left( \frac{\delta_i}{G(Y_i)} K_{d,x_{k,i}} H_{t_i,i}, \frac{\delta_j}{G(Y_j)} K_{d,x_{k,j}} H_{t_i,j} \right)$$

$$= : S_1 + S_2.$$ 

From (5.3), it is clear that

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

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

$$S_2 \leq \sum_{i=1}^{n} \sum_{B_2} c^2 h_1^d h_2^{2\tau} \rho^\frac{d}{\nu_2} (|i - j|)$$

$$\leq \frac{nc^2 h_1^d h_2^{2\tau}}{\nu_2} \sum_{B_2} \gamma_{d, i-j} e^{-\gamma_{d, i-j}}$$

$$\leq \frac{nc^2 h_1^d h_2^{2\tau}}{\nu_2} \int_{\eta_n}^{n} e^{-\frac{\gamma_{d, i-j}}{\nu_2}} du$$

$$= O \left( nh_1^d h_2^{2\tau} e^{-\frac{\gamma_{d, i-j}}{\nu_2}} \right).$$

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 $o(nh_1^d h_2^{\frac{2\tau}{\nu_2}})$ respectively. Consequently

$$\sigma_n^2 = V + S = O(nh_1^d) + o(nh_1^d h_2^{\frac{2\tau}{\nu_2}}) = O(nh_1^d).$$

Thereby, we get $A_n = O(nh_1^d)$. Next, from (5.9) and choosing $\mu, \lambda, L_1$ 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 $\varepsilon = \varepsilon_0 \sqrt{\frac{\log n}{n h_1^d}}$, we have

$$
P \left( \max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \left| \tilde{F}_{1,n}(x_k, t_\ell) - \mathbb{E} \left[ \tilde{F}_{1,n}(x_k, t_\ell) \right] \right| \geq \varepsilon \right)
\leq \mathbb{P} \left( \sum_{k=1}^{d_{x,n}} \sum_{\ell=1}^{d_{t,n}} P \left( \sum_{i=1}^{n} \Delta_i(x_k, t_\ell) \geq n h_1^d \varepsilon \right) \right)
\leq 2 d_{x,n} d_{t,n} \exp \left( \frac{- \left( n h_1^d \right)^2 \varepsilon_0^2}{2 \log n} \right)
\leq 2^{m_1} m_2 \frac{a_2}{b_n} \exp \left( \frac{- \frac{\varepsilon_0^2}{2} \log n}{c + \varepsilon_0 \left( \log \frac{n b_n}{n h_1^d} \right)^{1/6}} \right)
\leq c \left( n^{-1} h_1^{2a+d} \right)^{-1/2} \left( n h_1^d \right)^{1/2} h_2^{-1} n^{-c_0^2}
\leq c \left( n h_1^{2a+d(1 - \alpha)} \right)^{-1/2} (n h_2)^{-1} n^{-c_0^2} (\frac{1}{\alpha} + \frac{1}{2}) (5.12)
$$

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

$$
\sum_{n \geq 1} P \left( \max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \left| \tilde{F}_{1,n}(x_k, t_\ell) - \mathbb{E} \left[ \tilde{F}_{1,n}(x_k, t_\ell) \right] \right| \geq \varepsilon_0 \sqrt{\frac{\log n}{n h_1^d}} \right) < \infty.
$$

Applying the lemma of Borel-Cantelli, we obtain that

$$
I_{3n} = O \left( \left. \sqrt{\frac{\log n}{n h_1^d}} \right. \right) . (5.13)
$$

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

$$
\sup_{x \in \mathbb{U}} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(x, t) - \mathbb{E} \left[ \tilde{F}_{1,n}(x, t) \right] \right| = O \left( \sqrt{\frac{\log n}{n h_1^d}} \right) .
$$

The proof of Theorem 3.2 is achieved. \(\square\)

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

$$
\sup_{x \in \mathbb{U}} \sup_{t \in \mathcal{C}} \left| F_n(t|x) - F(t|x) \right| \leq \inf_{x \in \mathbb{U}} \left( \sup_{x \in \mathbb{U}} \sup_{t \in \mathcal{C}} \left| \mathbb{E} \left[ \tilde{F}_{1,n}(x, t) \right] - F_1(x, t) \right| \right)
+ \sup_{x \in \mathbb{U}} \sup_{t \in \mathcal{C}} \left| F_{1,n}(x, t) - \tilde{F}_{1,n}(x, t) \right|
+ \sup_{x \in \mathbb{U}} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(x, t) - \mathbb{E} \left[ \tilde{F}_{1,n}(x, t) \right] \right|
+ m_0^{-1} \sup_{x \in \mathbb{U}} \sup_{t \in \mathcal{C}} \left| l_n(x) - l(x) \right| .
$$


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

\[
\frac{1}{l_n(x)} \leq \frac{1}{l(x) - |l_n(x) - l(x)|} \leq \frac{1}{m_0 - \sup_{x \in \Omega} |l_n(x) - l(x)|}.
\]

This allows to write

\[
\sup_{x \in \Omega} \sup_{t \in C} |F_{t,n}(t|x) - F(t|x)| \leq \frac{1}{m_0 - \sup_{x \in \Omega} |l_n(x) - l(x)|} \left( \vartheta_1 + \vartheta_2 + \vartheta_3 + \vartheta_4 m_0^{-1} \sup_{x \in \Omega} \sup_{t \in C} F_1(x,t) \right).
\]

As for the term \( \vartheta_3 \), it has been bounded in Theorem 3.2. The following lemmas establish respectively the result of \( \vartheta_1, \vartheta_2 \). Then we apply Lemma 5.4 for \( \vartheta_3 \).

The bias term \( \vartheta_1 \) 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 \( \vartheta_2 \).

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

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

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

\[
E \left[ \tilde{F}_{1,n}(x,t) \right] = \frac{1}{h_1^2} E \left[ \frac{\delta_1}{\overline{G}(Y_1)} K_d \left( \frac{x - X_1}{h_1} \right) H \left( \frac{t - Y_1}{h_2} \right) \right] = \frac{1}{h_1^2} E \left[ K_d \left( \frac{x - X_1}{h_1} \right) E \left[ \frac{\delta_1}{\overline{G}(Y_1)} H \left( \frac{t - Y_1}{h_2} \right) | X_1 \right] \right].
\]

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

\[
E \left[ \frac{\delta_1}{\overline{G}(Y_1)} H \left( \frac{t - Y_1}{h_2} \right) | X_1 \right] = E \left[ \frac{\delta_1}{\overline{G}(Y_1)} H \left( \frac{t - Y_1}{h_2} \right) | T_1 \right] E \left[ \frac{\delta_1}{\overline{G}(Y_1)} H \left( \frac{t - T_1}{h_2} \right) | X_1 \right] = E \left[ H \left( \frac{t - T_1}{h_2} \right) | X_1 \right] = \int_{\mathbb{R}} H \left( \frac{t - y}{h_2} \right) f(y|X_1) dy = \int_{\mathbb{R}} H^{(1)}(z) F(t - z h_2 | X_1) dz.
\]

Again, by a change of variable we get

\[
E \left[ \tilde{F}_{1,n}(x,t) \right] = \frac{1}{h_1^2} E \left[ K_d \left( \frac{x - X_1}{h_1} \right) \int_{\mathbb{R}} H^{(1)}(z) F(t - z h_2 | X_1) dz \right] = \int_{\mathbb{R}^d} \frac{1}{h_1^2} K_d \left( \frac{x - u}{h_1} \right) \int_{\mathbb{R}} H^{(1)}(z) F(t - z h_2 | X_1 = u) f(u) du dz = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{1}{h_1^2} K_d \left( \frac{x - u}{h_1} \right) H^{(1)}(z) F_1(u, t - z h_2) du dz = \int_{\mathbb{R}^d} \int_{\mathbb{R}} K_d(r) H^{(1)}(z) F_1(x - r h_1, t - z h_2) dr dz.
\]
Then, expanding $F_1(x - rh_1, t - z h_2)$ up to order 2 around $(x, t)$ gives

$$F_1(x - rh_1, t - z h_2) = F_1(x, t) - h_1 \left[ \frac{\partial F_1(x, t)}{\partial x_1} + \ldots + r_d \frac{\partial F_1(x, t)}{\partial x_d} \right] - h_2 \left[ z \frac{\partial F_1(x, t)}{\partial t} \right]$$

$$+ \frac{h_1^2}{2} \left[ r_1 \frac{\partial^2 F_1(x^*, t^*)}{\partial x_1^2} + \ldots + r_d^2 \frac{\partial^2 F_1(x^*, t^*)}{\partial x_d^2} + 2 \sum_{i,j=1}^d r_i r_j \frac{\partial^2 F_1(x^*, t^*)}{\partial x_i \partial x_j} \right]$$

$$+ \frac{h_2^2}{2} \left[ z \frac{\partial^2 F_1(x^*, t^*)}{\partial t^2} \right] + h_1 h_2 \left[ r_1 z \frac{\partial^2 F_1(x^*, t^*)}{\partial x_1 \partial t} + \ldots + r_d z \frac{\partial^2 F_1(x^*, t^*)}{\partial x_d \partial t} \right].$$

Here, $(x^*, t^*)$ lies between $(x - rh_1, t - z h_2)$ and $(x, t)$. Finally, assumptions $A_1$, $A_2$, $A_3$ and $A_5$ entail

$$\sup_{x \in I} \sup_{t \in C} \left| \mathbb{E} \left[ F_{1,n}(x, t) \right] - F_1(x, t) \right| \leq c(h_1^2 + h_2^2).$$

This provides the desired result. □

5.6. Lemma. Under assumptions $A_2$, $A_4$ and $A_8$, for $n$ large enough, we have

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

Proof. Firstly, we have

$$\left| F_{1,n}(x, t) - \tilde{F}_{1,n}(x, t) \right| = \left| \frac{1}{nh_1^n} \sum_{i=1}^n \tilde{h}_d \left( \frac{x - X_i}{h_1} \right) H \left( \frac{t - Y_i}{h_2} \right) \left( \frac{1}{G_n(Y_i)} - \frac{1}{\overline{G}(Y_i)} \right) \right|$$

$$\leq \frac{1}{nh_1^n} \sum_{i=1}^n K_d \left( \frac{x - X_i}{h_1} \right) H \left( \frac{t - Y_i}{h_2} \right) \left( \frac{1}{G_n(Y_i)} - \frac{1}{\overline{G}(Y_i)} \right)$$

$$\leq \frac{\overline{f}_n(x)}{G_{\alpha}(\tau) G(\tau)} \sup_{t \in C} \left| G_n(t) - \overline{G}(t) \right|.$$

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

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

Recall that $0 < \theta < \gamma/(2\gamma + 9 + 3/2\kappa)$ for any real $\kappa > 0$ and $\gamma$ is referred in Assumption $A_8$. 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 $\partial_x$. □

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

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

$$\left| F(\xi_{p,n}(x)|x) - F(\xi_p(x)|x) \right| \leq \left| F_n(\xi_{p,n}(x)|x) - F(\xi_{p,n}(x)|x) \right| + \left| F_n(\xi_{p,n}(x)|x) - F(\xi_p(x)|x) \right|$$

$$+ \left| F(\xi_p(x)|x) - F(\xi_p(x)|x) \right|$$

$$\leq 2 \sup_{t \in C} \left| F_n(t|x) - F(t|x) \right|.$$
So, the first part of Corollary 3.4 is straightforwardly deduced from Theorem 3.3. And, a Taylor expansion of \( F(x) \) in the neighborhood of \( \xi_p(x) \) permits to get
\[
F(\xi_{p,n}(x)|x) - F(\xi_p(x)|x) = (\xi_{p,n}(x) - \xi_p(x)) f(\xi_p(x)|x),
\]
where \( \xi_{p,n}(\cdot) \) is between \( \xi_p(\cdot) \) and \( \xi_{p,n}(\cdot) \). Thus from (5.15), we obtain
\[
\sup_{s \in \Omega} |\xi_{p,n}(x) - \xi_p(x)| f(\xi_p(x)|x) \leq 2 \sup_{s \in \Omega} |F_n(t|x) - F(t|x)|.
\]
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. \( \square \)

References


