

Maximum likelihood estimation for conditional distribution single-index models under censoring

Ewa Strzalkowska-Kominiak^{1,*}

Department of Mathematics, Facultad de Informática, Universidade da Coruña, A Coruña, Spain

Ricardo Cao*

Department of Mathematics, Facultad de Informática, Universidade da Coruña, A Coruña, Spain

Abstract

A new likelihood approach is proposed for the problem of semiparametric estimation of a conditional distribution or density under censoring. Consistency and asymptotic normality for two versions of the maximum likelihood estimator of the parameter vector in the single index model are proved. The single-index model considered can be seen as a useful tool for credit scoring and estimation of the default probability in credit risk. A data-driven bandwidth selection procedure is proposed. It allows to choose the smoothing parameter involved in our approach. The finite sample performance of the estimators has been studied by simulations, where the new method has been compared with the method by Bouaziz and Lopez (2010) [1]. To the best of our knowledge this is the only existing competitor in this context. The simulation study shows the good behaviour of the proposed method.

Keywords: conditional density function, credit risk, kernel estimation, survival analysis

1. Motivation and background

The single-index model (SIM) is a flexible tool to incorporate the effect of a vector of covariates in a regression problem. By focusing on an index, the so-called “curse of dimensionality” is no longer a problem in the SIM, since multivariate nonparametric regression estimation is avoided. In other words, the SIM is as a reasonable compromise between a fully parametric and a fully nonparametric model.

Plenty of papers have dealt with the issue of estimating the parameter vector and the link function in SIM (see [6], [9], [12], [14], [15] and [23], among many other). Different methods are available for fitting the SIM: for instance, kernel smoothing, least-squares, average derivative estimation and sliced inverse regression. Parametric tests (see, for instance, [30]) and goodness-of-fit tests (see [8], [13], [27] and [32], among other) have been proposed for SIM. Generalizations of SIM has been also considered over the past few years. For instance [3], [28] and [29] studied estimation methods in partial linear single-index models. Empirical likelihood methods have been also applied for SIM (see, e.g., [33]). Finally SIM have been extended to survival analysis (as in [18]), including censored data (see [17], [19] and [20]).

In the contexts of conditional distribution function and conditional density the use of SIM has been much more limited. However, this is an important setting in applications to credit risk, where SIM’s can be used to jointly estimate the scoring and the probability of default of credits, since this probability can be expressed in terms of a conditional distribution function (see [2] for details). SIM’s have been also used for other purposes in credit risk (see, for instance, [21]).

To fix our notation, let the lifetime, Z , be a nonnegative random variable dependent on a vector of covariates $X = (X_1, \dots, X_d)'$ and $f(z|x)$ the density function of Z given $X = x$. Moreover, let

$$\theta_0 = (\theta_1, \dots, \theta_d)' \text{ where } d \geq 2$$

*Tel +34981167000; fax +34981167160; Address: Campus de Elviña, s/n 15071 A Coruña, Spain

Email address: strzalkowska@udc.es (Ewa Strzalkowska-Kominiak)

¹Corresponding author.

be a parameter vector with the property:

$$f_{\theta_0}(z|\theta'_0\mathbf{x}) = f(z|\mathbf{x}),$$

where $f_{\theta_0}(z|\theta'_0\mathbf{x})$ is the conditional density of Z given $\theta'_0X = \theta'_0\mathbf{x}$. Furthermore, let $F(z|X = \mathbf{x})$ and $F_{\theta_0}(z|\theta'_0X = \theta'_0\mathbf{x})$ be the conditional distribution functions, given X and θ'_0X , respectively. As a consequence

$$F_{\theta_0}(z|\theta'_0X = \theta'_0\mathbf{x}) = F(z|X = \mathbf{x}). \quad (1)$$

Moreover, the random variable Z may be censored from the right by $C \sim G$. Hence, we observe

$$Y = \min(Z, C) \sim H$$

together with

$$\delta = 1_{\{Z \leq C\}}.$$

For example, in credit risk applications (as in Cao, Vilar-Fernández and Devia [2]) the lifetime, Z , could be the time to default of a credit and the vector $X = (X_1, \dots, X_d)'$ would include the client covariates that are relevant for estimating the probability of default. The linear combination $\theta'_0\mathbf{x}$ is an index of the propensity to default of a credit with covariate vector $\mathbf{x}$. So, the interest is to estimate the parameter θ_0 , the conditional density, f_{θ_0} , and the distribution function, F_{θ_0} . In this context, Bouaziz and Lopez [1] have recently proposed a semiparametric procedure to estimate the conditional density, deriving consistency and asymptotic normality of some estimator of the index parameter in (1). This estimator is based on maximizing a regression-like analogue of the (uncomputable) theoretical likelihood.

In this paper a new maximum likelihood estimator is proposed. The idea is to compute some alternative observable version of the likelihood that takes into account explicitly the censoring mechanism. The rest of the paper is organized as follows. Section 2 presents the theoretical likelihood in this conditional context under censoring and some result that characterizes the true index vector as the maximizer of the expected value of this theoretical likelihood. The new likelihood version to be used in practice is presented in Section 3, where the maximum likelihood estimator is defined. Section 4 includes the main results for the maximum likelihood estimator for the index vector, namely its consistency and asymptotic normality. Section 5 is devoted to a simulation study that exhibits the good behaviour of the proposed approach, while Section 6 shows a real data example. Finally, Section 7 presents some conclusions and the proofs of the main results are collected in Appendix A.

2. Theoretical likelihood function

Suppose for a while that F_θ is known except for the value of the index vector θ . We define the theoretical likelihood function as follows:

$$\tilde{L}_n(\theta) = \prod_{i=1}^n (f_\theta(Y_i|\theta'X_i))^{\delta_i} (1 - F_\theta(Y_i|\theta'X_i))^{1-\delta_i}$$

and

$$\tilde{l}_n(\theta) = \frac{1}{n} \log \left(\tilde{L}_n(\theta) \right) = \frac{1}{n} \sum_{i=1}^n \left(\delta_i \log f_\theta(Y_i|\theta'X_i) + (1 - \delta_i) \log(1 - F_\theta(Y_i|\theta'X_i)) \right). \quad (2)$$

Set

$$\tilde{\theta}_n = \arg \max_{\theta} \tilde{l}_n(\theta) \quad (3)$$

and define the score function as the expected likelihood:

$$l(\theta) = E(\tilde{l}_n(\theta)). \quad (4)$$

Some assumptions are needed for the following result:

A1: Z is independent of C

A2: Z is conditionally independent of C given X

A3: $\mathbb{P}(Z \leq C|X, Z) = \mathbb{P}(Z \leq C|Z)$

Observe that assumption A3 allows some dependence between C and X . Moreover, under (1) and A2, we have that Z is conditionally independent of C given $\theta'_0 X$.

The following result characterizes θ_0 as the maximizer of the score function.

Theorem 1. *Under A1-A3 we have*

$$\theta_0 = \arg \max_{\theta} l(\theta) \quad (5)$$

Proof:

We will show that for every θ we have

$$E(\tilde{l}_n(\theta_0)) \geq E(\tilde{l}_n(\theta)).$$

According to (2) and since the data are i.i.d., we have

$$\begin{aligned} E(\tilde{l}_n(\theta_0)) - E(\tilde{l}_n(\theta)) &= E\left(\log\left(\frac{f_{\theta_0}(Y_1|\theta'_0 X_1)}{f_{\theta}(Y_1|\theta' X_1)}\right)^{\delta_1}\right) + E\left(\log\left(\frac{1 - F_{\theta_0}(Y_1|\theta'_0 X_1)}{1 - F_{\theta}(Y_1|\theta' X_1)}\right)^{1-\delta_1}\right) \\ &= E\left(-\log\left(\frac{f_{\theta}(Y_1|\theta' X_1)}{f_{\theta_0}(Y_1|\theta'_0 X_1)}\right)^{\delta_1}\right) + E\left(-\log\left(\frac{1 - F_{\theta}(Y_1|\theta' X_1)}{1 - F_{\theta_0}(Y_1|\theta'_0 X_1)}\right)^{1-\delta_1}\right). \end{aligned}$$

Moreover, using Jensen inequality

$$E(\tilde{l}_n(\theta_0)) - E(\tilde{l}_n(\theta)) \geq -\log(A) - \log(B),$$

where

$$A = E\left(\left(\frac{f_{\theta}(Y_1|\theta' X_1)}{f_{\theta_0}(Y_1|\theta'_0 X_1)}\right)^{\delta_1}\right)$$

and

$$B = E\left(\left(\frac{1 - F_{\theta}(Y_1|\theta' X_1)}{1 - F_{\theta_0}(Y_1|\theta'_0 X_1)}\right)^{1-\delta_1}\right).$$

Let $f_3(\mathbf{x}, c, z)$, $f_2(c, z|\mathbf{x})$ and $g(c|\mathbf{x})$ be the densities of (X, C, Z) , (C, Z) given X and of C given X , respectively. According to A2, we have that $f_3(\mathbf{x}, c, z) = f_2(c, z|\mathbf{x})f_X(\mathbf{x}) = g(c|\mathbf{x})f(z|\mathbf{x})f_X(\mathbf{x})$. Hence, by definition of θ_0 , we obtain

$$\begin{aligned} A &= \int \frac{f_{\theta}(\min(z, c)|\theta' \mathbf{x})^{1_{\{z \leq c\}}}}{f(\min(z, c)|\mathbf{x})^{1_{\{z \leq c\}}}} f_3(\mathbf{x}, c, z) d\mathbf{x} dz dc \\ &= \int \frac{f_{\theta}(\min(z, c)|\theta' \mathbf{x})^{1_{\{z \leq c\}}}}{f(\min(z, c)|\mathbf{x})^{1_{\{z \leq c\}}}} g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dz dc \\ &= \int 1_{\{z \leq c\}} f_{\theta}(z|\theta' \mathbf{x}) g(c|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dz dc + \int 1_{\{z > c\}} g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dz dc \\ &= a + b, \end{aligned}$$

where $a = \int g(c|\mathbf{x}) f_X(\mathbf{x}) F_{\theta}(c|\theta' \mathbf{x}) d\mathbf{x} dc$ and $b = \int g(c|\mathbf{x}) f_X(\mathbf{x}) (1 - F(c|\mathbf{x})) d\mathbf{x} dc$

Similarly,

$$\begin{aligned} B &= \int \frac{(1 - F_\theta(\min(z, c)|\theta' \mathbf{x}))^{1_{\{z > c\}}}}{(1 - F(\min(z, c)|\mathbf{x}))^{1_{\{z > c\}}}} g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dz dc \\ &= \int (1 - F_\theta(c|\theta' \mathbf{x})) g(c|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dc + \int F(c|\mathbf{x}) g(c|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dc = 2 - a - b. \end{aligned}$$

Hence

$$E(\tilde{l}_n(\theta_0)) - E(\tilde{l}_n(\theta)) \geq -\log((a + b)(2 - a - b)),$$

where, $a, b \in [0, 1]$. Finally, since the function $x(2 - x)$ has a global maximum at $x = 1$, we have that

$$E(\tilde{l}_n(\theta_0)) - E(\tilde{l}_n(\theta)) \geq -\log((a + b)(2 - a - b)) \geq -\log(1) = 0.$$

This completes the proof. \(\square\)

3. Maximum likelihood estimation

Since, in practice, neither f_θ nor F_θ are known, they need to be estimated. Let K be a nonnegative kernel and h_1, h_2 two positive bandwidths. Set

$$F(t) = \mathbb{P}(Z \leq t), G(t) = \mathbb{P}(C \leq t) \text{ and } H(t) = \mathbb{P}(Y \leq t)$$

together with

$$\tau_H = \inf\{y : H(y) = 1\}.$$

Furthermore, set

$$\hat{f}_\theta(y|\theta' \mathbf{x}) = \frac{\frac{1}{h_1 h_2} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) K\left(\frac{y - v}{h_2}\right) 1_{\{v \leq a_n\}} F_n^\theta(du, dv)}{\frac{1}{h_1} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) 1_{\{v \leq a_n\}} F_n^\theta(du, dv)} \quad (6)$$

and

$$1 - \hat{F}_\theta(y|\theta' \mathbf{x}) = \frac{\frac{1}{h_1} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) \mathbb{K}\left(\frac{y - v}{h_2}\right) 1_{\{v \leq a_n\}} F_n^\theta(du, dv)}{\frac{1}{h_1} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) 1_{\{v \leq a_n\}} F_n^\theta(du, dv)}, \quad (7)$$

where $\mathbb{K}(x) = \int_x^\infty K(z) dz$ and a_n is a positive sequence, used to avoid problems with the right tail of the lifetime distribution, such that $a_n \rightarrow \tau_H$ when $n \rightarrow \infty$. Moreover, $F_n^\theta(u, v)$ is the Kaplan-Meier estimator of $(\theta' X, Z)$ defined as

$$\int \varphi(u, v) F_n^\theta(du, dv) = \sum_{j=1}^n W_{jn} \varphi(\theta' X_j, Y_j),$$

where

$$W_{jn} = F_n(Y_j) - F_n(Y_{j-})$$

and $F_n(t)$ is Kaplan-Meier estimator of $F(t) = \mathbb{P}(Z \leq t)$. See, [16], [25] and [26] for details.

Moreover, it can be shown that

$$W_{jn} = \frac{\delta_j}{n(1 - G_n(Y_{j-}))},$$

where $G_n(x)$ is the Kaplan-Meier estimator of $G(t) = \mathbb{P}(C \leq t)$. See, e.g., [1] for details.

Finally, let $\hat{f}_\theta^{-i}(Y_i|\theta'X_i)$ and $1 - \hat{F}_\theta^{-i}(Y_i|\theta'X_i)$ be the estimators defined in (6) and (7), where the sums in definition of $F_n^\theta(u, v)$ runs over $j \neq i$. Set

$$\hat{l}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \left(\delta_i \log \hat{f}_\theta^{-i}(Y_i|\theta'X_i) + (1 - \delta_i) \log(1 - \hat{F}_\theta^{-i}(Y_i|\theta'X_i)) \right) 1_{\{Y_i \leq a_n, X_i \in A^{c_n}\}}, \quad (8)$$

where A^{c_n} is a set with the property $\mathbb{P}(X_i \in A^{c_n}) \rightarrow 1$ for every $i = 1, \dots, n$ when $n \rightarrow \infty$ and $c_n \rightarrow 0$, and

$$\hat{\theta}_n = \arg \max_{\theta} \hat{l}_n(\theta). \quad (9)$$

Finally, set

$$\hat{f}_\theta^{-i}(Y_i|\theta'X_i) = \frac{\hat{r}(\theta'X_i, Y_i)}{\hat{s}(\theta'X_i)} \quad (10)$$

and

$$1 - \hat{F}_\theta^{-i}(Y_i|\theta'X_i) = \frac{\hat{d}(\theta'X_i, Y_i)}{\hat{s}(\theta'X_i)}, \quad (11)$$

where

$$\hat{r}(\theta'X_i, Y_i) = \frac{1}{h_1 h_2} \sum_{j \neq i} 1_{\{Y_j \leq a_n\}} \frac{\delta_j}{n(1 - G_n(Y_j -))} K\left(\frac{\theta'X_i - \theta'X_j}{h_1}\right) K\left(\frac{Y_i - Y_j}{h_2}\right) \quad (12)$$

$$\hat{s}(\theta'X_i) = \frac{1}{h_1} \sum_{j \neq i} 1_{\{Y_j \leq a_n\}} \frac{\delta_j}{n(1 - G_n(Y_j -))} K\left(\frac{\theta'X_i - \theta'X_j}{h_1}\right) \quad (13)$$

and

$$\hat{d}(\theta'X_i, Y_i) = \frac{1}{h_1} \sum_{j \neq i} 1_{\{Y_j \leq a_n\}} \frac{\delta_j}{n(1 - G_n(Y_j -))} K\left(\frac{\theta'X_i - \theta'X_j}{h_1}\right) \mathbb{K}\left(\frac{Y_i - Y_j}{h_2}\right). \quad (14)$$

4. Main results

In this section we will study the properties of $\hat{\theta}_n$. To establish the main results we need to assume some further conditions:

A4: $E(X|\theta'_0 X, C, Z) = E(X|\theta'_0 X)$

A5: $E(XX^t) < \infty$ componentwise.

The two bandwidths h_1, h_2 and the sequence c_n should fulfill the following conditions

A6: $\sqrt{n}h_1^4 \rightarrow 0, \sqrt{n}h_2^2 \rightarrow 0, nh_1^6 \rightarrow \infty$ and $c_n n^{5/6} h_1^3 h_2 \rightarrow \infty$ when $h_1, h_2 \rightarrow 0$ and $n \rightarrow \infty$.

Remark 1. Using $h_1 = c_1 n^{-1/7}$, $h_2 = c_2 n^{-1/3}$ and $c_n = 1/\log(n)$, assumption A6 is fulfilled. These are the asymptotically optimal rates for the estimation of the first derivative of the conditional density and for the distribution function estimation, respectively.

Moreover, assume

A7: The derivatives $\frac{\partial^k}{\partial u^k} \frac{\partial^l}{\partial v^l} f_{\theta_0}(u, v)$, $\frac{d^k}{du} f_{\theta'_0 X}(u)$ and $\frac{d^k}{du} E(X|\theta'_0 X = u)$ exists for $k = 1, 2, 3$ and $l = 1, 2$.

A8: The function $h(\theta_1) = \frac{\partial}{\partial \theta_k} f_{\theta}(\theta'_1 \mathbf{x}, y)_{\theta=\theta_0}$ is continuous and $\frac{\partial^2}{\partial^2 \theta_k} f_{\theta}(\theta'_0 \mathbf{x}, y)_{\theta=\theta_0}$ exists.

A9: $\int \frac{F(dt)}{1-G(t-)} < \infty$ and $\sup_x \int \frac{f_{\theta_0}(t|x)}{1-G(t-)} dt < \infty$.

Finally, let $l^{[1]}(\theta_0) = \nabla_{\theta} l(\theta)|_{\theta=\theta_0}$ denote the gradient of $l(\theta)$ over θ evaluated in θ_0 . Further, let $l^{[2]}(\theta)$ denote the Hessian matrix of $l(\theta)$. Now we can state our first result.

Lemma 1. *Under A1-A3, A7 and A8 we have*

$$\hat{\theta}_n - \theta_0 = - \left[\hat{l}_n^{[2]}(\hat{\theta}_n^*) \right]^{-1} (\hat{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0)),$$

where $\hat{\theta}_n^*$ is between $\hat{\theta}_n$ and θ_0 .

Proof. Using (5) and (9) we have

$$l^{[1]}(\theta_0) = 0 \quad \text{and} \quad \hat{l}_n^{[1]}(\hat{\theta}_n) = 0.$$

Now a Taylor expansion gives

$$l^{[1]}(\theta_0) = 0 = \hat{l}_n^{[1]}(\hat{\theta}_n) = \hat{l}_n^{[1]}(\theta_0) + \hat{l}_n^{[2]}(\hat{\theta}_n^*)(\hat{\theta}_n - \theta_0),$$

where $\hat{\theta}_n^*$ is between $\hat{\theta}_n$ and θ_0 . This completes the proof. $\square$

Theorem 2. *Let $a_n = H^{-1}(1 - \frac{s_n}{n})$ with $s_n \approx n^{2/3}(\log(n))^{2+\varepsilon_1}$ for some $\varepsilon_1 > 0$. Set $A^{c_n} = A_0^{c_n} = \{\mathbf{x} : f_{\theta'_0 X}(\theta'_0 \mathbf{x}) > c_n\}$. Then, under A1-A9 and if $l^{[2]}(\theta^*)$ is positive definite for θ^* belonging to a neighborhood of θ_0 , we have*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow \mathcal{N}(0, \Sigma),$$

where

$$\begin{aligned} \Sigma &= \Sigma_2 \Sigma_1 \Sigma_2^t, \\ \Sigma_2 &= \left[l^{[2]}(\theta_0) \right]^{-1} \end{aligned}$$

and

$$\begin{aligned} \Sigma_1 &= \int (1 - G(z - |\mathbf{x}|)) (\nabla_{\theta} \log(f_{\theta}(z|\theta' \mathbf{x}))_{\theta=\theta_0}) (\nabla_{\theta} \log(f_{\theta}(z|\theta' \mathbf{x}))_{\theta=\theta_0})^t f(\mathbf{x}, z) d\mathbf{x} dz \\ &\quad + \int (1 - F(c|\mathbf{x})) (\nabla_{\theta} \log(1 - F_{\theta}(c|\theta' \mathbf{x}))_{\theta=\theta_0}) (\nabla_{\theta} \log(1 - F_{\theta}(c|\theta' \mathbf{x}))_{\theta=\theta_0})^t g(\mathbf{x}, c) d\mathbf{x} dc. \end{aligned}$$

Theorem 3. *Let $a_n = Y_{r_n \cdot n}$ with $r_n = n - s_n$ and $s_n \approx n^{2/3}(\log(n))^{2+\varepsilon_1}$. Set $A^{c_n} = A_n^{c_n} = \{\mathbf{x} : f_{\theta'_n X}(\theta'_n \mathbf{x}) > c_n\}$, where θ_n denotes a consistent estimator of θ_0 . Then, under A1-A9, if $\text{supp}(K) \in [-a, a]$ and $l^{[2]}(\theta^*)$ is positive definite for θ^* belonging to a neighborhood of θ_0 , we have*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow \mathcal{N}(0, \Sigma),$$

where Σ is defined as in Theorem 2.

Theorem 4. *Under A1-A9, with $a_n = \tau_H$, we have*

$$\hat{\theta}_n \rightarrow \theta_0 \quad \text{in probability}$$

Remark 2. In Theorem 2 we assume that the quantile function $H^{-1}(u)$ and the set $A_0^{c_n}$ are known. This makes the trimming function $J(x, y) = 1_{\{y \leq a_n, x \in A_0^{c_n}\}}$ deterministic and the proofs more transparent. Nevertheless, Theorem 3 shows that we can replace the quantiles $H^{-1}(u)$ by order statistics and the set $A_0^{c_n}$ with an estimable $A_n^{c_n}$ based on preliminary estimator θ_n . The θ_n can be found by setting $1_{\{X_i \in A^{c_n}\}} \equiv 1$ and maximizing the likelihood (8). See [6] and [17], for details. In praxis, since $J(X_i, Y_i)$ goes to one in probability for every $i = 1, \dots, n$, we may take $J(X_i, Y_i) \equiv 1$. Such a choice of trimming function gives good results in the simulation study.

5. Simulations

Let us consider the model used by Bouaziz and Lopez [1]:

$$Z = \theta'_0 X + \varepsilon, \quad (15)$$

where $\theta_0 = (1, 0.5, 1.4, 0.2)$, $X = (X_1, X_2, X_3, X_4)$,

$$X_i \sim \begin{cases} \mathcal{N}(0, 1), & \text{with probability 0.2} \\ \mathcal{N}(0.25, 2), & \text{with probability 0.8} \end{cases}$$

is a normal mixture for $i = 1, 2, 3, 4$ and $\varepsilon \sim \mathcal{N}(0, |\theta'_0 X|)$. Set $C \sim \exp(\lambda)$, so that we only observe $Y = \min(Z, C)$ together with $\delta = 1_{\{Z \leq C\}}$.

The goal is to estimate θ_0 with $\hat{\theta}_n = (1, \hat{\theta}_{n1}, \hat{\theta}_{n2}, \hat{\theta}_{n3}) \equiv (1, \theta_n)$ and select the bandwidths h_1 and h_2 . For the latter we consider two possible strategies:

1. Optimizing the likelihood function over two different bandwidths h_1 and h_2 .
2. Optimizing the likelihood function over h by setting $h_1 = h\hat{\sigma}(\theta'_0 X)$ and $h_2 = h\hat{\sigma}(Y)$, where $\hat{\sigma}$ is the estimated standard deviation.

In the following sections we set $1_{\{Y_i \leq a_n, X_i \in A^{c_n}\}} \equiv 1$.

5.1. Model 1

We consider now the model from Bouaziz and Lopez [1], where λ is constant (it does not depend on X). Hence C is independent of Z and the assumption A1 is fulfilled.

Tables 1-4 show Monte Carlo approximations for the bias, the variance and the mean squared error of the Bouaziz and Lopez (BL) estimator and the new estimator with the two bandwidth choices presented above. These results are for $n = 100$ and $n = 200$ together with $\lambda = 0.3$ (25% censoring) and $\lambda = 0.85$ (40 % censoring). For the new estimator, the results are based on 500 trials.

Table 1: Estimated bias, variance and MSE for $\lambda = 0.3$, $n = 100$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	0.003	0.026	-0.006
Variance	0.031	0.080	0.023
MSE	0.1350874		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.009	0.041	-0.001
Variance	0.023	0.058	0.018
MSE	0.1014918		
Bouaziz and Lopez [1]			
Bias	0.07	0.221	0.028
Variance	0.034	0.074	0.02
MSE	0.182598		

Table 2: Estimated bias, variance and MSE for $\lambda = 0.3$ $n = 200$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	-0.007	-0.015	-0.006
Variance	0.010	0.026	0.009
MSE	0.0456373		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	-0.003	-0.009	-0.004
Variance	0.008	0.019	0.007
MSE	0.03511254		
Bouaziz and Lopez [1]			
Bias	0.034	0.107	0.014
Variance	0.007	0.011	0.0006
MSE	0.0364064		

Table 3: Estimated bias, variance and MSE for $\lambda = 0.85$ $n = 100$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	-0.008	0.008	-0.008
Variance	0.054	0.136	0.039
MSE	0.2298826		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	-0.003	0.029	0.0005
Variance	0.037	0.091	0.025
MSE	0.1542782		
Bouaziz and Lopez [1]			
Bias	0.074	0.176	0.061
Variance	0.064	0.051	0.069
MSE	0.2239023		

Table 4: Estimated bias, variance and MSE for $\lambda = 0.85$ $n = 200$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	0.005	-0.0046	-0.0047
Variance	0.021	0.043	0.0176
MSE	0.08184733		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.012	0.013	-0.0009
Variance	0.016	0.033	0.015
MSE	0.06476457		
Bouaziz and Lopez [1]			
Bias	0.043	0.14	0.021
Variance	0.018	0.022	0.014
MSE	0.07533921		

It is important to mention that the proposed estimator is mostly better than the one by Bouaziz and Lopez [1]. Furthermore, the new estimator with bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$ gives always the smallest MSE. The different performance between the BL estimator and the one proposed in here comes from the bias. It is worth mentioning that this difference increases when the sample size decreases.

5.2. Model 2

In this section we consider the same model as before, but with $\lambda = \lambda(X)$. Hence C depends on Z and assumption A1 is violated. More precisely, we take

$$\lambda(X) = \lambda_1 |\theta'_0 X|,$$

with $\lambda_1 = 0.15$ and $\lambda_1 = 0.65$, which gives, as above, 25% and 40% of censoring.

Tables 5-8 collect the results for the Monte Carlo approximations for the bias, the variance and the mean squared error of the new estimator. The sample size was set to $n = 100$ and $n = 200$, with 500 trials.

Table 5: Estimated bias, variance and MSE for $\lambda_1 = 0.15$ $n = 100$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	0.0088	0.0083	0.00002
Variance	0.041	0.087	0.030
MSE	0.1586619		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.016	0.029	0.00008
Variance	0.029	0.071	0.022
MSE	0.1247307		

Table 6: Estimated bias, variance and MSE for $\lambda_1 = 0.15$ $n = 200$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	-0.0048	-0.0044	-0.0018
Variance	0.012	0.034	0.012
MSE	0.05776507		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.0005	0.0081	0.0045
Variance	0.010	0.028	0.009
MSE	0.04728104		

Table 7: Estimated bias, variance and MSE for $\lambda_1 = 0.65$ $n = 100$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	-0.012	-0.031	-0.0048
Variance	0.168	0.693	0.081
MSE	0.942775		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.026	0.073	0.0008
Variance	0.079	0.219	0.051
MSE	0.3566779		

Table 8: Estimated bias, variance and MSE for $\lambda_1 = 0.65$ $n = 200$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	-0.004	-0.019	-0.0046
Variance	0.088	0.201	0.038
MSE	0.3281369		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.034	0.0436	0.007
Variance	0.049	0.111	0.036
MSE	0.1997059		

With 25% of censoring (Tables 5-6) the estimator is still behaving well, even if the assumption A1 is not fulfilled. A censoring of 40% causes a significant increase in MSE. Nevertheless, the estimator with $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$ still presents a reasonable behavior. It is worth mentioning that, although the estimator with a single smoothing parameter has a smaller variance, its bias tends to be larger.

5.3. Model 3

In this section we consider again model (15) with $\lambda = \lambda(X)$. As in Model 2, $Z \sim \mathcal{N}(\theta'_0 X, |\theta'_0 X|)$ and $C \sim \exp(\lambda_1 |\theta'_0 X|)$. This time, given X , C and Z are generated from the Gaussian copula with parameter 0.5. Hence C depends on Z and neither assumption A1 nor A2 is fulfilled. We take $\lambda_1 = 0.15$ and $\lambda_1 = 0.65$, which gives, as before, 25% and 40% of censoring.

Tables 9-12 collect the results for the Monte Carlo approximations for the bias, the variance and the mean squared error of the new estimator. The sample size was set to $n = 100$ and $n = 200$, with 500 trials.

Table 9: Estimated bias, variance and MSE for $\lambda_1 = 0.15$ $n = 100$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	0.005	0.022	0.003
Variance	0.036	0.088	0.029
MSE	0.1535400		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.014	0.036	0.004
Variance	0.027	0.064	0.021
MSE	0.1139752		

Table 10: Estimated bias, variance and MSE for $\lambda_1 = 0.15$ $n = 200$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	0.004	0.011	-0.002
Variance	0.016	0.036	0.012
MSE	0.0647620		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.007	0.018	0.002
Variance	0.012	0.026	0.011
MSE	0.0492625		

Table 11: Estimated bias, variance and MSE for $\lambda_1 = 0.65$ $n = 100$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	0.009	-0.041	-0.021
Variance	0.086	0.350	0.074
MSE	0.5124920		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.049	0.069	0.0001
Variance	0.068	0.196	0.052
MSE	0.3238710		

Table 12: Estimated bias, variance and MSE for $\lambda_1 = 0.65$ $n = 200$ and 500 trials.

Different bandwidths h_1 and h_2			
	$\hat{\theta}_{n1}$	$\hat{\theta}_{n2}$	$\hat{\theta}_{n3}$
Bias	-0.013	-0.012	0.004
Variance	0.071	0.160	0.054
MSE	0.2856553		
Bandwidths $h_1 = h\hat{\sigma}(\theta'X)$ and $h_2 = h\hat{\sigma}(Y)$			
Bias	0.019	0.059	0.022
Variance	0.060	0.165	0.053
MSE	0.2825447		

The results in Tables 9-12 shows that if, additionally to A1, the assumption A2 is not fulfilled the MSE stays similar to MSE in Model 2. Similarly as in Model 2, the results are worse than in Model 1 especially when censoring is heavy. Consequently, the assumption A1 seems to be crucial for our model.

5.4. Asymptotic variance

In this section we compare the asymptotic variance of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ for the estimator presented in this paper with that given in [1]. Using Theorem 2, it can be shown that the asymptotic covariance matrix, Σ , of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ can be written as

$$\begin{aligned} \Sigma = & \left[\int (1 - G(z - |\mathbf{x}|)) (\nabla_{\theta} \log(f_{\theta}(z|\theta' \mathbf{x}))_{\theta=\theta_0}) (\nabla_{\theta} \log(f_{\theta}(z|\theta' \mathbf{x}))_{\theta=\theta_0})^t f(\mathbf{x}, z) d\mathbf{x} dz \right. \\ & \left. + \int (1 - F(c|\mathbf{x})) (\nabla_{\theta} \log(1 - F_{\theta}(c|\theta' \mathbf{x}))_{\theta=\theta_0}) (\nabla_{\theta} \log(1 - F_{\theta}(c|\theta' \mathbf{x}))_{\theta=\theta_0})^t g(\mathbf{x}, c) d\mathbf{x} dc \right]^{-1}. \end{aligned} \quad (16)$$

The asymptotic covariance matrix of the estimator proposed by Bouaziz and Lopez [1] is given by

$$\Sigma_{\tau} = V_{\tau}^{-1} \Delta_{\tau}(f_1) V_{\tau}^{-1},$$

where, using the notation of the present paper, $f_1(x, y) = (f_{\theta_0}^{\tau}(y|\theta'_0 \mathbf{x}))^{-1} 1_{\{f_{\theta'_0 X}(\theta'_0 x) > c\}} \nabla_{\theta} f_{\theta}^{\tau}(y|\theta' \mathbf{x})_{\theta=\theta_0}$, f_{θ}^{τ} denotes the density of Z given $\theta' X$ and $Z \in A^{\tau}$ with $A^{\tau} \in (-\infty, \tau]$,

$$V_{\tau} = E \left(\frac{\nabla_{\theta} f_{\theta}^{\tau}(Z|\theta' X)_{\theta=\theta_0} \nabla_{\theta} f_{\theta}^{\tau}(Z|\theta' X)_{\theta=\theta_0}^t}{(f_{\theta_0}^{\tau}(Z|\theta'_0 X))^2} 1_{\{f_{\theta'_0 X}(\theta'_0 x) > c\}} 1_{\{Z \in A^{\tau}\}} \right)$$

$$\Delta_{\tau}(f_1) = Var(\psi(Y, \delta, X, f_1 1_{\{Y \in A^{\tau}\}}))$$

and

$$\psi(Y, \delta, X, f_1 1_{\{Y \in A^{\tau}\}}) = \frac{\delta f_1(X, Y) 1_{\{Y \in A^{\tau}\}}}{1 - G(Y-)} + \int \frac{\int_y^{\tau} \int f_1(x, t) 1_{\{t \in A^{\tau}\}} F(dx, dt) M^G(dy)}{1 - H(y)}.$$

Moreover, using Lemma 4, we have

$$\int_y^{\tau} \int f_1(x, t) 1_{\{t \in A^{\tau}\}} F(dx, dt) = E(f_1(X, Z) 1_{\{Z \in (y, \tau]\}}) = 0.$$

Hence

$$\Delta_{\tau}(f_1) = Var \left(\frac{\delta f_1(X, Y) 1_{\{Y \in A^{\tau}\}}}{1 - G(Y-)} \right) = E \left(\frac{\nabla_{\theta} f_{\theta}^{\tau}(Z|\theta' X)_{\theta=\theta_0} \nabla_{\theta} f_{\theta}^{\tau}(Z|\theta' X)_{\theta=\theta_0}^t}{(1 - G(Z-)) (f_{\theta_0}^{\tau}(Z|\theta'_0 X))^2} 1_{\{f_{\theta'_0 X}(\theta'_0 x) > c\}} 1_{\{Z \in A^{\tau}\}} \right),$$

where the last equation is a consequence of Lemma 4 and assumption A3 (Assumption 5 in [1]).

Note that if there is no censoring, $C = \infty$, both asymptotic covariance matrices, Σ in (16) and Σ_{τ} from [1] with $\tau = \infty$ and $c = 0$, reduce to the asymptotic covariance matrix in the complete data case

$$\Sigma_0 = \left[E \left((\nabla_{\theta} \log(f_{\theta}(Z|\theta' X))_{\theta=\theta_0}) (\nabla_{\theta} \log(f_{\theta}(Z|\theta' X))_{\theta=\theta_0})^t \right) \right]^{-1}.$$

See Theorem 2 in [5], for details.

Because of the complicated structure of the asymptotic covariance matrices, they cannot be compared in general. Nevertheless, we present a comparison based on Model (15). Tables 13-14 show the asymptotic variances $diag(\Sigma)$ and $diag(\Sigma_{\tau})$ for Model 1 for $1_{\{f_{\theta'_0 X}(\theta'_0 x) > c\}} \equiv 1$ and different values of τ . More precisely, we choose the $\tau = \tau^*$ (which differs depending on censoring) that gives the best results, as well as $\tau = F^{-1}(0.9)$ (which coincides with Table 5 in [1]) and $\tau = \infty$ (no trimming). Since there are no explicit expressions for the expectations, we approximate them via Monte Carlo based on 30000 trials.

Table 13: Asymptotic variances using $\lambda_1 = 0.3$ for Model 1.

Present estimator: $diag(\Sigma)$ given in (16)		
0.148	1.154	0.024
Bouaziz and Lopez [1]: $diag(\Sigma_\tau)$ with $\tau = \tau^* = F^{-1}(0.8)$		
0.208	1.620	0.039
Bouaziz and Lopez [1]: $diag(\Sigma_\tau)$ with $\tau = F^{-1}(0.9)$		
0.217	1.701	0.037
Bouaziz and Lopez [1]: $diag(\Sigma_\tau)$ with $\tau = \infty$		
0.581	4.553	0.093

Table 14: Asymptotic variances using $\lambda_1 = 0.85$ for Model 1.

Present estimator: $diag(\Sigma)$ given in (16)		
0.199	1.551	0.032
Bouaziz and Lopez [1]: $diag(\Sigma_\tau)$ with $\tau = \tau^* = F^{-1}(0.6)$		
0.253	1.979	0.042
Bouaziz and Lopez [1]: $diag(\Sigma_\tau)$ with $\tau = F^{-1}(0.9)$		
0.705	5.526	0.116
Bouaziz and Lopez [1]: $diag(\Sigma_\tau)$ with $\tau = \infty$		
276.5	2172	44.31

Consequently, the asymptotic variance for the present estimator is always better than that of [1] for Model 1. Moreover, the variance in [1] is strongly dependent on the choice of τ , specially when censoring is heavy. Hence, in this particular example, the estimator presented in this paper is more efficient.

6. German Credit Data

In this section we apply our model to the German Credit data set which is publicly available on the internet page [http://archive.ics.uci.edu/ml/datasets/Statlog+\(German+Credit+Data\)](http://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)). This data set includes information about 1000 credits from which 300 were classified as bad credits and 700 as good credits. In our model we use four covariates such as credit amount, checking account, time of employment and savings account. Let us denote with $X = (X_1, X_2, X_3, X_4)'$ the vector with the above mentioned covariates and with $\theta_0 = (1, \theta_2, \theta_3, \theta_4)'$ the index to be estimated.

Since some of the X_i 's are ordinal (interval) variables, in order to use our approach, we change them into numerical variables. To be more specific, X_1 is already a continuous variable denoting amount of credit in DM, $X_2 \in \{-0.05, 0.01, 0.25, 0\}$ denotes the checking account in thousands of DM, $X_3 \in \{0, 0.5, 2.5, 5.5, 8.5\}$ denotes the years of employment and $X_4 \in \{0, 0.05, 0.25, 0.75, 1.25\}$ denotes the savings account in thousands of DM.

Additionally, to the explanatory variable X , we set Z , as the time to default and denote with $\delta = 1$ the bad/defaulted credits and with $\delta = 0$ the good credits. The results are presented in Table 15

Table 15: Estimated parameter $\hat{\theta}_n$			
$\hat{\theta}_{1n}$	$\hat{\theta}_{2n}$	$\hat{\theta}_{3n}$	$\hat{\theta}_{4n}$
1	3.2091	0.2312	2.1891

Figure 1 presents the estimated conditional survival function for

$$\mathbf{x}_1 = (\text{credit}, 0.1, 4, 0.25)' \text{ where } \text{credit} \in \{1, 10\}$$

$$\mathbf{x}_2 = (2, \text{account}, 4, 0.25)' \text{ where } \text{account} \in \{-0.05, 0.25\}$$

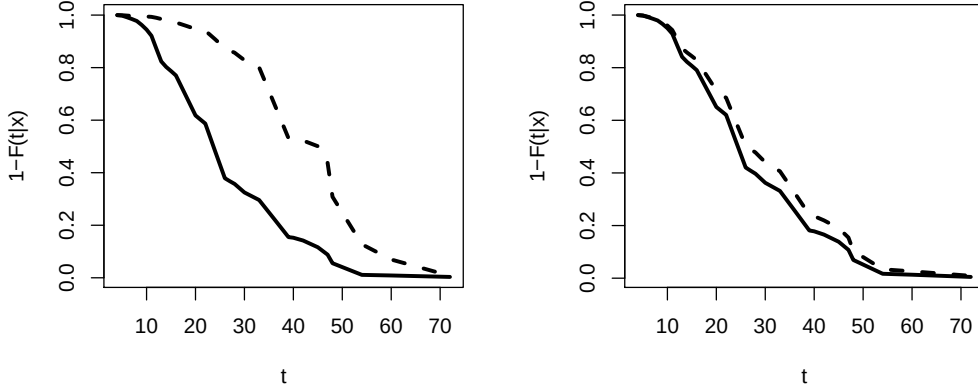


Figure 1: Estimators $1 - \hat{F}_{\hat{\theta}_n}(t|\hat{\theta}'_n \mathbf{x}_1)$ (left) for $\text{credit} = 1$ (solid line), $\text{credit} = 10$ (dashed line) and $1 - \hat{F}_{\hat{\theta}_n}(t|\hat{\theta}'_n \mathbf{x}_2)$ (right) for $\text{account} = -0.05$ (solid line), $\text{account} = 0.25$ (dashed line)

7. Conclusions

In this paper we have introduced an estimator of the parameter vector in a single-index model under censoring and proved its asymptotic properties. We compared our method with that presented by Bouaziz and Lopez [1]. The simulation study has shown that our estimator gives mostly better results and is more efficient than that presented in [1]. The main difference between the two approaches comes from the definition of the likelihood function. The theoretical likelihood in our approach, defined in (2), is commonly applied in the censored data setup. Bouaziz and Lopez [1], on the other hand, estimated the quantity $E(\log f_{\theta}^T(Z|\theta'X)J(X)1_{Z \in A_\tau})$, where f_{θ}^T is the density of Z given $\theta'X$ and $Z \in A^\tau$ while $J(X)1_{Z \in A_\tau}$ is a trimming function. This kind of likelihood, with $A^\tau = \mathbb{R}$, is generally used in the complete data setup (see, e.g., [5]). The trimming function in the estimated likelihood is also present in our model. Nevertheless, it goes to 1 and hence vanishes when the sample size increases. The next crucial difference between our method and that presented in [1] has to do with bandwidth selection. In order to estimate the conditional density, Bouaziz and Lopez [1] used one smoothing parameter h for both Z and $\theta'X$. Consequently they needed fourth order kernels in order to prove the asymptotic normality. On the contrary, our approach uses two different bandwidths, h_1 and h_2 , which allows us to use positive second order kernels. The only drawback of our method is the need of Assumption A2, which requires the independence between Z and C given the explanatory variable X . Nevertheless, both methods require Assumption A1, the unconditional independence between Z and C . Note that in many practical examples the censoring variable C is independent of X . In such a case Assumption A2 is a consequence of A1.

Acknowledgements

The authors acknowledge financial support from Ministerio de Economía y Competitividad Grants MTM2008-00166 and MTM2011-22392 (EU ERDF support included). Additionally, Ewa Strzalkowska-Kominiak acknowledges financial support from a Juan de la Cierva scholarship, from the Spanish Ministerio de Economía y Competitividad.

References

- [1] Bouaziz, O., Lopez, O. (2010). Conditional density estimation in a censored single-index model. *Bernoulli* 16, 514–542
- [2] Cao, R., Vilar J.M., Devia, A. (2009). Modeling consumer credit risk via survival analysis. *SORT* 33 (1), 3–30
- [3] Carroll, R.J., Fan, J., Gijbels, I., Wand, M.P. (1997). Generalized partially linear single-index models. *J. Amer. Statist. Assoc.* 92, 477–489.
- [4] Csörgő, S. (1996). Universal Gaussian approximations under random censorship. *The Annals of Statistics* 24, No. 6, 2744–2778
- [5] Delecroix, M., Härdle, W., Hristache, M. (2003). Efficient estimation in conditional single-index regression. *Journal of Multivariate Analysis* 86, 213–226
- [6] Delecroix, M., Hristache, M., Patilea, V. (2006). On semiparametric M-estimation in single-index regression. *Journal of Statistical Planning and Inference* 136, 730–769
- [7] Dvoretzky, A., Kiefer, J., Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Annals of Mathematical Statistics* 27, 642–669
- [8] Escanciano, J.C., Song, K. (2010). Testing single-index restrictions with a focus on average derivatives. *J. Econometrics* 156, 377–391.
- [9] Gao, J., Liang, H. (1997). Statistical inference in single-index and partially nonlinear models. *Ann. Inst. Statist. Math.* 49, 493–517.
- [10] Gascuel O., Caraux G. (1992). Bounds on expectations of order statistics via extremal dependences. *Statistics and Probability Letters* 15, 143–148
- [11] Gill, R. (1983) Large sample behaviour of the product-limit estimator on the whole real line. *The Annals of Statistics* 11, No.1, 49–58
- [12] Härdle, W., Hall, P., Ichimura, H. (1993). Optimal smoothing in single-index models. *Ann. Statist.* 21, 157–178.
- [13] Härdle, W., Mammen, E. and Proença, I. (2001). A bootstrap test for single index models. *Statistics* 35, 427–451.
- [14] Hristache, M., Juditsky, A., Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model. *Ann. Statist.* 29, 595–623.
- [15] Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *J. Econometrics* 58, 71–120.
- [16] Kaplan, E.L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53, 457–481
- [17] Lopez, O. (2009). Single-index regression models with right-censored responses. *J. Statist. Plann. Inference* 139, 1082–1097.
- [18] Lu, X., Chen, G., Singh, R.S., Song, P.X.K. (2006). A class of partially linear single-index survival models. *Canad. J. Statist.* 34, 97–112.
- [19] Lu, X., Cheng, T.L. (2007). Randomly censored partially linear single-index models. *J. Multivariate Anal.* 98, 1895–1922.
- [20] Lu, X. (2010). Asymptotic distributions of two synthetic data estimators for censored single-index models. *J. Multivariate Anal.* 101, 999–1015.
- [21] McAleer, M., da Veiga, B. (2008). Single-index and portfolio models for forecasting value-at-risk thresholds. *J. Forecast.* 27, 217–235.
- [22] Okolewski A., Rychlik T. (2001). Sharp distribution-free bounds on the bias in estimating quantiles via order statistics. *Statistics and Probability Letters* 52, 207–213
- [23] Stoker, T.M. (1986). Consistent estimation of scaled coefficients. *Econometrica* 54, 1461–1481.
- [24] Stute, W. (1994). Strong and weak representations of cumulative hazard function and Kaplan-Meier estimators on increasing sets. *Journal of Statistical Planning and Inference* 42, 315–329
- [25] Stute, W. (1995). The central limit theorem under random censorship. *The Annals of Statistics* 23, No.2, 422–439
- [26] Stute, W. (1996). Distributional convergence under random censorship when covariables are present. *Scandinavian Journal of statistics* 23, 461–471
- [27] Stute, W., Zhu, L.X. (2005). Nonparametric checks for single-index models. *Ann. Statist.* 33, 1048–1083.
- [28] Wang, J.L., Xue, L., Zhu, L., Chong, Y.S. (2010). Estimation for a partial-linear single-index model. *Ann. Statist.* 38, 246–274.
- [29] Xia, Y., Härdle, W. (2006). Semi-parametric estimation of partially linear single-index models. *J. Multivariate Anal.* 97, 1162–1184.
- [30] Zhang, R., Huang, Z., Lv, Y. (2010). Statistical inference for the index parameter in single-index models. *J. Multivariate Anal.* 101, 1026–1041.
- [31] Zhou, M. (1991). Some properties of the Kaplan-Meier estimator for independent nonidentically distributed random variables. *The Annals of Statistics* 19, No. 4, 2266–2274
- [32] Zhu, L.X., Ng, K.W. (2003). Checking the adequacy of a partial linear model. *Statist. Sinica* 13, 763–781.
- [33] Zhu, L.X. and Xue, L.G. (2006). Empirical likelihood confidence regions in a partially linear single-index model. *J. Roy. Statist. Soc. Ser. B* 68, 549–570.

Appendix A. Proofs of Theorems 2 - 4.

Proof of Theorem 2.

According to Lemma 1, we need to study the quantity:

$$\hat{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0) = \alpha_n(\theta_0) + \beta_n(\theta_0),$$

where

$$\alpha_n(\theta_0) = \hat{l}_n^{[1]}(\theta_0) - \tilde{l}_n^{[1]}(\theta_0)$$

$$\beta_n(\theta_0) = \tilde{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0).$$

For $\beta_n(\theta_0)$, according to Lemma 2, we have

$$\sqrt{n}\beta_n(\theta_0) \rightarrow \mathcal{N}(0, \Sigma_1)$$

and

$\beta_n(\theta_0) \rightarrow 0$ in probability.

Concerning $\alpha_n(\theta)$, set

$$d(\theta' X_i, Y_i) = \int_{Y_i}^{\infty} f_{\theta}(\theta' X_i, y) dy$$

Hence, by (2), (8), (10) and (11), we have

$$\begin{aligned} \hat{l}_n(\theta_0) - \tilde{l}_n(\theta_0) &= \frac{1}{n} \sum_{i=1}^n \left[\delta_i (\log(\hat{r}(\theta'_0 X_i, Y_i)) - \log(f_{\theta_0}(\theta'_0 X_i, Y_i))) + (\log(f_{\theta'_0 X}(\theta'_0 X_i)) - \log(\hat{s}(\theta'_0 X_i))) \right. \\ &\quad \left. + (1 - \delta_i) (\log(\hat{d}(\theta'_0 X_i, Y_i)) - \log(d(\theta'_0 X_i, Y_i))) \right] 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} \\ &\quad + \frac{1}{n} \sum_{i=1}^n \left[\delta_i \log f_{\theta_0}(Y_i | \theta'_0 X_i) + (1 - \delta_i) \log(1 - F_{\theta_0}(Y_i | \theta'_0 X_i)) \right] (1_{\{Y_i > a_n, X_i \in A^{c_n}\}} + 1_{\{X_i \notin A^{c_n}\}}). \end{aligned}$$

First note, that a_n and c_n are deterministic sequences such that $a_n \rightarrow \tau_H$ and $c_n \rightarrow 0$. Then, using Lemma 4, the derivative with respect to θ of the last term is of order $o_{\mathbb{P}}(n^{-1/2})$. Hence

$$\begin{aligned} \hat{l}_n^{[1]}(\theta_0) - \tilde{l}_n^{[1]}(\theta_0) &= \frac{1}{n} \sum_{i=1}^n \left[\delta_i \left(\frac{\hat{r}^{[1]}(\theta'_0 X_i, Y_i)}{\hat{r}(\theta'_0 X_i, Y_i)} - \frac{f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \right) + \left(\frac{f_{\theta'_0 X}^{[1]}(\theta'_0 X_i)}{f_{\theta'_0 X}(\theta'_0 X_i)} - \frac{\hat{s}^{[1]}(\theta'_0 X_i)}{\hat{s}(\theta'_0 X_i)} \right) \right. \\ &\quad \left. + (1 - \delta_i) \left(\frac{\hat{d}^{[1]}(\theta'_0 X_i, Y_i)}{\hat{d}(\theta'_0 X_i, Y_i)} - \frac{d^{[1]}(\theta'_0 X_i, Y_i)}{d(\theta'_0 X_i, Y_i)} \right) \right] 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} + o_{\mathbb{P}}(n^{-1/2}). \end{aligned}$$

Furthermore, by repeated use of

$$\frac{1}{\hat{r}} = \frac{1}{f_{\theta_0}} + \frac{f_{\theta_0} - \hat{r}}{f_{\theta_0} \hat{r}}$$

and similar expansions for $\hat{s}$ and $\hat{d}$, we obtain

$$\hat{l}_n^{[1]}(\theta_0) - \tilde{l}_n^{[1]}(\theta_0) = \sum_{k=1}^{12} A_{kn} + o_{\mathbb{P}}(n^{-1/2}),$$

where

$$\begin{aligned}
A_{1n} &= \frac{1}{n} \sum_{i=1}^n \left[\delta_i \frac{\hat{r}^{[1]}(\theta'_0 X_i, Y_i) - f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{2n} &= \frac{1}{n} \sum_{i=1}^n \left[\delta_i \frac{f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} (f_{\theta_0}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i)) \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{3n} &= \frac{1}{n} \sum_{i=1}^n \left[\frac{f_{\theta'_0 X}^{[1]}(\theta'_0 X_i) - \hat{s}^{[1]}(\theta'_0 X_i)}{f_{\theta'_0 X}(\theta'_0 X_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{4n} &= \frac{1}{n} \sum_{i=1}^n \left[\frac{f_{\theta'_0 X}^{[1]}(\theta'_0 X_i)}{f_{\theta'_0 X}^2(\theta'_0 X_i)} (\hat{s}(\theta'_0 X_i) - f_{\theta'_0 X}(\theta'_0 X_i)) \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{5n} &= \frac{1}{n} \sum_{i=1}^n \left[(1 - \delta_i) \frac{\hat{d}^{[1]}(\theta'_0 X_i, Y_i) - d^{[1]}(\theta'_0 X_i, Y_i)}{d(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{6n} &= \frac{1}{n} \sum_{i=1}^n \left[(1 - \delta_i) \frac{d^{[1]}(\theta'_0 X_i, Y_i)}{d^2(\theta'_0 X_i, Y_i)} (d(\theta'_0 X_i, Y_i) - \hat{d}(\theta'_0 X_i, Y_i)) \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{7n} &= \frac{1}{n} \sum_{i=1}^n \left[\delta_i \frac{(\hat{r}^{[1]}(\theta'_0 X_i, Y_i) - f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i))(f_{\theta_0}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i))}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{8n} &= \frac{1}{n} \sum_{i=1}^n \left[\delta_i \nabla_{\theta} \log(\hat{r}(\theta' X_i, Y_i))|_{\theta=\theta_0} \frac{(f_{\theta_0}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i))^2}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{9n} &= \frac{1}{n} \sum_{i=1}^n \left[\frac{(f_{\theta'_0 X}^{[1]}(\theta'_0 X_i) - \hat{s}^{[1]}(\theta'_0 X_i))(\hat{s}(\theta'_0 X_i) - f_{\theta'_0 X}(\theta'_0 X_i))}{s^2(\theta'_0 X_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{10n} &= \frac{1}{n} \sum_{i=1}^n \left[\nabla_{\theta} \log(\hat{s}(\theta' X_i))|_{\theta=\theta_0} \frac{(f_{\theta'_0 X}(\theta'_0 X_i) - \hat{s}(\theta'_0 X_i))^2}{f_{\theta'_0 X}^2(\theta'_0 X_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{11n} &= \frac{1}{n} \sum_{i=1}^n \left[(1 - \delta_i) \frac{(\hat{d}^{[1]}(\theta'_0 X_i, Y_i) - d^{[1]}(\theta'_0 X_i, Y_i))(d(\theta'_0 X_i, Y_i) - \hat{d}(\theta'_0 X_i, Y_i))}{d^2(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\
A_{12n} &= \frac{1}{n} \sum_{i=1}^n \left[(1 - \delta_i) \nabla_{\theta} \log(\hat{d}(\theta' X_i, Y_i))|_{\theta=\theta_0} \frac{(d(\theta'_0 X_i, Y_i) - \hat{d}(\theta'_0 X_i, Y_i))^2}{d^2(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}}
\end{aligned}$$

Now we need to consider each of this terms separately.

To deal with A_{in} for $i = 1, \dots, 6$, let us define the following functions:

$$\tilde{r}(\theta' \mathbf{x}, y) = \frac{1}{h_1 h_2} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) K\left(\frac{y - v}{h_2}\right) f_{\theta}(u, v) 1_{\{v \leq a_n\}} du dv$$

$$\tilde{s}(\theta' \mathbf{x}) = \frac{1}{h_1} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) f_{\theta}(u, v) 1_{\{v \leq a_n\}} du dv$$

and

$$\tilde{d}(\theta' \mathbf{x}, y) = \frac{1}{h_1} \int K\left(\frac{\theta' \mathbf{x} - u}{h_1}\right) \mathbb{K}\left(\frac{y - v}{h_2}\right) f_{\theta}(u, v) 1_{\{v \leq a_n\}} du dv.$$

Moreover let

$$A_{1n} = B_{1n} + C_{1n},$$

where

$$\begin{aligned} B_{1n} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\hat{r}^{[1]}(\theta'_0 X_i, Y_i) - \tilde{r}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} \\ C_{1n} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\tilde{r}^{[1]}(\theta'_0 X_i, Y_i) - f_{\theta}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} \end{aligned}$$

and

$$A_{2n} = B_{2n} + C_{2n},$$

where

$$\begin{aligned} B_{2n} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} (\tilde{r}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i)) 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} \\ C_{2n} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} (f_{\theta_0}(\theta'_0 X_i, Y_i) - \tilde{r}(\theta'_0 X_i, Y_i)) 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} \end{aligned}$$

Similarly $A_{in} = B_{in} + C_{in}$ for $i = 3, 4, 5, 6$. Set

$$C_n = C_{1n} + C_{2n} + C_{3n} + C_{4n} + C_{5n} + C_{6n}$$

and

$$B_n = B_{1n} + B_{2n} + B_{3n} + B_{4n} + B_{5n} + B_{6n}.$$

Now using Lemma 5,

$$\sqrt{n}C_n = o_{\mathbb{P}}(1).$$

Moreover, Lemma 13 leads to

$$\sqrt{n}B_n = o_{\mathbb{P}}(1).$$

Finally, Lemma 14 implies that $\sqrt{n} \sum_{k=7}^{12} A_{kn} = o_{\mathbb{P}}(1)$, while Lemma 15 shows that $\tilde{l}_n^{[2]}(\theta) \rightarrow l^{[2]}(\theta)$. This completes the proof. $\square$

Lemma 2. *Under the conditions in Theorem 2 we have*

$$\tilde{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0) \xrightarrow{\mathbb{P}} 0$$

and

$$\sqrt{n}(\tilde{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0)) \rightarrow \mathcal{N}(0, \Sigma_1)$$

Proof.

According to (2) and since $l^{[1]}(\theta_0) = E(\tilde{l}_n^{[1]}(\theta_0)) = 0$, we have

$$\sqrt{n}(\tilde{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\delta_i \frac{f_\theta^{[1]}(Y_i|\theta'X_i)_{\theta=\theta_0}}{f_{\theta_0}(Y_i|\theta'_0X_i)} - (1 - \delta_i) \frac{F_\theta^{[1]}(Y_i|\theta'X_i)_{\theta=\theta_0}}{1 - F_{\theta_0}(Y_i|\theta'_0X_i)} \right].$$

Then, $\tilde{l}_n^{[1]}(\theta) - l^{[1]}(\theta)$ is a sum of i.i.d. random vectors and the Law of Large Numbers gives the convergence to zero in probability. Moreover, by the Central Limit Theorem we have

$$\sqrt{n}(\tilde{l}_n^{[1]}(\theta_0) - l^{[1]}(\theta_0)) \rightarrow \mathcal{N}(0, \Sigma_1)$$

To compute Σ_1 , let $f_3(\mathbf{x}, z, c)$ be the density of (X, Z, C) , while $f(\mathbf{x}, z)$ and $g(\mathbf{x}, z)$ are the densities of (X, Z) and (X, C) . Recall that $Y = \min(Z, C)$ while $\delta = 1_{\{Z \leq C\}}$. Hence, using A1, A2 and A3, since $\delta(1 - \delta) = 0$, we obtain

$$\begin{aligned} \Sigma_1 &= \lim_{n \rightarrow \infty} E \left[\left(\delta_1 \frac{f_\theta^{[1]}(Y_1|\theta'X_1)_{\theta=\theta_0}}{f_{\theta_0}(Y_1|\theta'_0X_1)} - (1 - \delta_1) \frac{F_\theta^{[1]}(Y_1|\theta'X_1)_{\theta=\theta_0}}{1 - F_{\theta_0}(Y_1|\theta'_0X_1)} \right) \right. \\ &\quad \left. \left(\delta_1 \frac{f_\theta^{[1]}(Y_1|\theta'X_1)_{\theta=\theta_0}}{f_{\theta_0}(Y_1|\theta'_0X_1)} - (1 - \delta_1) \frac{F_\theta^{[1]}(Y_1|\theta'X_1)_{\theta=\theta_0}}{1 - F_{\theta_0}(Y_1|\theta'_0X_1)} \right)^t \right] \\ &= \int (1 - G(z - |\mathbf{x}|)) (\nabla_\theta \log(f_\theta(z|\theta'\mathbf{x}))_{\theta=\theta_0}) (\nabla_\theta \log(f_\theta(z|\theta'\mathbf{x}))_{\theta=\theta_0})^t f(\mathbf{x}, z) d\mathbf{x} dz \\ &\quad + \int (1 - F(c|\mathbf{x})) (\nabla_\theta \log(1 - F_\theta(c|\theta'\mathbf{x}))_{\theta=\theta_0}) (\nabla_\theta \log(1 - F_\theta(c|\theta'\mathbf{x}))_{\theta=\theta_0})^t g(\mathbf{x}, c) d\mathbf{x} dc. \end{aligned}$$

Observe that the last equation is a consequence of Lemma 4. \(\square\)

Next we consider the gradients of f_θ . We have following results:

Lemma 3. *Under A8, we have*

a)

$$\frac{\partial}{\partial \theta_k} f_\theta(\theta'\mathbf{x}, y)_{\theta=\theta_0} = \frac{\partial}{\partial t} \left(f_{\theta_0}(t, y) \left[\mathbf{x}_k - E(X_k|\theta'_0X = t, Y = y) \right] \right)_{t=\theta'_0\mathbf{x}}.$$

b)

$$\frac{\partial}{\partial \theta_k} f_{\theta'X}(\theta'\mathbf{x})_{\theta=\theta_0} = \frac{\partial}{\partial t} \left(f_{\theta'_0X}(t) \left[\mathbf{x}_k - E(X_k|\theta'_0X = t) \right] \right)_{t=\theta'_0\mathbf{x}}$$

Proof. Let

$$\theta_0 = (\theta_{01}, \dots, \theta_{0d})'$$

and

$$\theta^* = \theta_0 + (0, \dots, h, \dots, 0)', \text{ where } h \text{ is in position } k$$

Then, for every $\mathbf{x} = (x_1, \dots, x_d)'$ and $y \in \mathbb{R}$, we have

$$\frac{\partial}{\partial \theta_k} f_\theta(\theta'\mathbf{x}, y)_{\theta=\theta_0} = A + B + C,$$

where

$$\begin{aligned} A &= \lim_{h \rightarrow 0} \frac{f_{\theta^*}(\theta'_0 \mathbf{x}, y) - f_{\theta_0}(\theta'_0 \mathbf{x}, y)}{h} \\ B &= \lim_{h \rightarrow 0} \frac{f_{\theta_0}(\theta^{*'} \mathbf{x}, y) - f_{\theta_0}(\theta'_0 \mathbf{x}, y)}{h} \\ C &= \lim_{h \rightarrow 0} \left[\frac{f_{\theta^*}(\theta^{*'} \mathbf{x}, y) - f_{\theta_0}(\theta^{*'} \mathbf{x}, y)}{h} - \frac{f_{\theta^*}(\theta'_0 \mathbf{x}, y) - f_{\theta_0}(\theta'_0 \mathbf{x}, y)}{h} \right]. \end{aligned}$$

As to C , using a Taylor expansion in θ^* , we obtain

$$C = \lim_{h \rightarrow 0} \left[\frac{\partial}{\partial \theta_k} f_{\theta}(\theta^{*'} \mathbf{x}, y)_{\theta=\theta^*} - \frac{\partial}{\partial \theta_k} f_{\theta}(\theta'_0 \mathbf{x}, y)_{\theta=\theta^*} \right] + \lim_{h \rightarrow 0} \frac{h}{2} \left[\frac{\partial^2}{\partial \theta_k^2} f_{\theta}(\theta^{*'} \mathbf{x}, y)_{\theta=\tilde{\theta}} - \frac{\partial^2}{\partial \theta_k^2} f_{\theta}(\theta'_0 \mathbf{x}, y)_{\theta=\tilde{\theta}} \right],$$

where $\tilde{\theta} = \theta_0 + (0, \dots, \tilde{h}, \dots, 0)'$, $\tilde{\theta} = \theta_0 + (0, \dots, \tilde{h}, \dots, 0)'$ and $\tilde{h}, \tilde{h} \in (0, h)$. Hence, using A8, $C = 0$.

Further, it is easy to prove that

$$B = \frac{\partial}{\partial \theta_k} f_{\theta_0}(\theta' \mathbf{x}, y)_{\theta=\theta_0} = x_k \frac{\partial}{\partial t} f_{\theta_0}(t, y)_{t=\theta'_0 \mathbf{x}}.$$

Finally, we have

$$A = \frac{\partial}{\partial \theta_k} f_{\theta}(\theta'_0 \mathbf{x}, y)_{\theta=\theta_0} = \frac{\partial}{\partial t} \left(\frac{\partial}{\partial \tilde{y}} \left(\frac{\partial}{\partial \theta_k} P(\theta' X \leq t, Y \leq \tilde{y})_{\theta=\theta_0} \right)_{\tilde{y}=y} \right)_{t=\theta'_0 \mathbf{x}}.$$

Set $\theta = (\theta_{01}, \dots, \theta_k, \dots, \theta_{0d})'$ so that $\theta_i = \theta_{0i}$ for $i \neq k$ is a fixed value and hence $\theta' X$ is a function of only one variable θ_k . Hence, for the inner derivative in A , we have

$$\theta' X = \theta_{01} X_1 + \dots + \theta_k X_k + \dots + \theta_{0d} X_d = \theta'_0 X + \theta_k X_k - \theta_{0k} X_k.$$

Let $f^{X_k, \theta'_0 X, Y}(x_k, u, \tilde{y})$ denote the density of $(X_k, \theta'_0 X, Y)$. Hence

$$\begin{aligned} \frac{\partial}{\partial \theta_k} P(\theta' X \leq t, Y \leq \tilde{y}) &= \frac{\partial}{\partial \theta_k} P(\theta'_0 X + \theta_k X_k - \theta_{0k} X_k \leq t, Y \leq \tilde{y}) \\ &= \frac{\partial}{\partial \theta_k} \int_{-\infty}^{\tilde{y}} \int \int 1_{\{u + \theta_k x_k - \theta_{0k} x_k \leq t\}} f^{X_k, \theta'_0 X, Y}(x_k, u, z) du dx_k dz \\ &= \int_{-\infty}^{\tilde{y}} \int \left(\frac{\partial}{\partial \theta_k} \int_{-\infty}^{t - \theta_k x_k + \theta_{0k} x_k} f^{X_k, \theta'_0 X, Y}(x_k, u, z) du \right) dx_k dz \\ &= \int_{-\infty}^{\tilde{y}} \int f^{X_k, \theta'_0 X, Y}(x_k, t - \theta_k x_k + \theta_{0k} x_k, z) (-x_k) dx_k dz \end{aligned}$$

Thus,

$$\frac{\partial}{\partial \theta_k} P(\theta' X \leq t, Y \leq \tilde{y})_{\theta=\theta_0} = \int_{-\infty}^{\tilde{y}} \int f^{X_k, \theta'_0 X, Y}(x_k, t, z) (-x_k) dx_k dz.$$

Finally,

$$\begin{aligned}
A &= \frac{\partial}{\partial t} \frac{\partial}{\partial \tilde{y}} \left(\frac{\partial}{\partial \theta_k} P(\theta' X \leq t, Y \leq \tilde{y})_{\theta=\theta_0} \right)_{\tilde{y}=y, t=\theta'_0 \mathbf{x}} = \frac{\partial}{\partial t} \left(\int f^{X_k, \theta'_0 X, Y}(x_k, t, y)(-x_k) dx_k \right)_{t=\theta'_0 \mathbf{x}} \\
&= \frac{\partial}{\partial t} \left(f_{\theta_0}(t, y) \int f^{X_k | \theta'_0 X, Y}(x_k | t, y)(-x_k) dx_k \right)_{t=\theta'_0 \mathbf{x}} = \frac{\partial}{\partial t} (-f_{\theta_0}(t, y) E(X_k | \theta'_0 X = t, Y = y))_{t=\theta'_0 \mathbf{x}}
\end{aligned}$$

and the proof is finished. $\square$

Lemma 4. *Under A4, we have*

a)

$$\nabla_{\theta} f_{\theta}(y | \theta' \mathbf{x})_{\theta=\theta_0} = \left[\mathbf{x} - E(X | \theta'_0 X = \theta'_0 \mathbf{x}) \right] \frac{\partial}{\partial t} f_{\theta_0}(y | t)_{t=\theta'_0 \mathbf{x}}.$$

b)

$$E(\nabla_{\theta} [\log f_{\theta}(Z_i | \theta' X_i)]_{\theta=\theta_0} | Z_i, C_i, \theta' X_i) = 0$$

and

$$E(\nabla_{\theta} [\log(1 - F_{\theta}(C_i | \theta' X_i))]_{\theta=\theta_0} | Z_i, C_i, \theta' X_i) = 0.$$

Proof.

Using A4, part a) is an immediate consequence of Lemma 3. As to b), by A4, we have

$$E \left([X_i - E(X | \theta'_0 X = \theta'_0 X_i)] | Z_i, C_i, \theta' X_i \right) = E(X | \theta'_0 X = \theta'_0 X_i) - E(X | \theta'_0 X = \theta'_0 X_i) = 0.$$

This completes the proof. $\square$

Lemma 5. *Under A2, A4, A6 and A7, we have following results:*

$$\sqrt{n}C_n = O_{\mathbb{P}}(\sqrt{n}h_1^4 + \sqrt{n}h_2^2) = o_{\mathbb{P}}(1).$$

Proof.

First, using a change of variable, we have

$$\begin{aligned}
\tilde{r}(\theta'_0 X_i, Y_i) 1_{\{Y_i \leq a_n\}} &= \frac{1}{h_1 h_2} \int K \left(\frac{\theta'_0 X_i - u}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) f_{\theta_0}(u, v) 1_{\{v \leq a_n\}} du dv 1_{\{Y_i \leq a_n\}} \\
&= \int K(z_1) K(z_2) f_{\theta_0}(\theta'_0 X_i - h_1 z_1, Y_i - h_2 z_2) 1_{\{Y_i - h_2 z_2 \leq a_n\}} dz_1 dz_2 1_{\{Y_i \leq a_n\}}
\end{aligned}$$

and

$$\begin{aligned}
\tilde{r}^{[1]}(\theta'_0 X_i, Y_i) 1_{\{Y_i \leq a_n\}} &= \frac{1}{h_1^2 h_2} \int (X_i - E(X | \theta'_0 X = u)) K' \left(\frac{\theta'_0 X_i - u}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \\
&\quad f_{\theta_0}(u, v) 1_{\{v \leq a_n\}} du dv 1_{\{Y_i \leq a_n\}} \\
&= \int (X_i - E(X | \theta'_0 X = \theta'_0 X_i - h_1 z_1)) K'(z_1) K(z_2) \\
&\quad f_{\theta_0}(\theta'_0 X_i - h_1 z_1, Y_i - h_2 z_2) 1_{\{Y_i - h_2 z_2 \leq a_n\}} dz_1 dz_2 1_{\{Y_i \leq a_n\}}
\end{aligned}$$

Then, observe that $\int z^a K(z) dz = 0$ for a odd, $\int K(z) dz = 1$ and $b_K = \int z^2 K(z) dz < \infty$, together with $\int z^b K'(z) dz = 0$ for b even, $\int z K'(z) dz = -1$ and $\int z^3 K'(z) dz = -3 \int z^2 K(z) dz$.

Moreover, using A7, the function $f_\theta(u, v) 1_{\{v \leq a_n\}} 1_{\{Y_i \leq a_n\}}$ is three times differentiable on $[v, Y_i]$ if $v \leq Y_i$ and on $[Y_i, v]$ if $v > Y_i$. Hence, using a Taylor expansion, we have

$$\begin{aligned} \tilde{r}(\theta'_0 X_i, Y_i) 1_{\{Y_i \leq a_n\}} &= f_{\theta_0}(\theta'_0 X_i, Y_i) 1_{\{Y_i \leq a_n\}} + \frac{b_K h_1^2}{2} \frac{\partial^2}{\partial^2 u} f_{\theta_0}(u, Y_i)_{u=\theta'_0 X_i} 1_{\{Y_i \leq a_n\}} \\ &\quad + \frac{b_K h_2^2}{2} \frac{\partial^2}{\partial^2 v} (f_{\theta_0}(\theta'_0 X_i, v) 1_{\{Y_i \leq a_n\}})_{v=Y_i} 1_{\{Y_i \leq a_n\}} + O_{\mathbb{P}}(h_1^4 + h_1^2 h_2^2 + h_2^4). \end{aligned}$$

Similarly, Lemma 3 additionally implies that

$$\begin{aligned} \tilde{r}^{[1]}(\theta'_0 X_i, Y_i) 1_{\{Y_i \leq a_n\}} &= f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i) 1_{\{Y_i \leq a_n\}} \\ &\quad + \frac{b_K h_1^2}{2} \frac{\partial^3}{\partial^3 u} [(X_i - E(X | \theta'_0 X = u)) f_{\theta_0}(u, Y_i)]_{u=\theta'_0 X_i} 1_{\{Y_i \leq a_n\}} \\ &\quad + \frac{b_K h_2^2}{2} \frac{\partial}{\partial u} \frac{\partial^2}{\partial^2 v} [(X_i - E(X | \theta'_0 X = u)) f_{\theta_0}(u, v) 1_{\{v \leq a_n\}}]_{u=\theta'_0 X_i, v=Y_i} 1_{\{Y_i \leq a_n\}} \\ &\quad + O_{\mathbb{P}}(h_1^4 + h_1^2 h_2^2 + h_2^4). \end{aligned}$$

Finally, repeating the same steps for $\tilde{s}(\theta'_0 X_i)$ and $\tilde{d}(\theta'_0 X_i, Y_i)$ we obtain

$$C_n = c_{1n} + c_{2n} + \tilde{c}_{1n} + O_{\mathbb{P}}(h_1^4 + h_2^2)$$

where

$$c_{1n} = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} h_1^2 b_K c_{1i} 1_{\{X_i \in A^{c_n}\}},$$

$$c_{2n} = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} h_1^2 b_K c_{2i} 1_{\{X_i \in A^{c_n}\}},$$

$$\tilde{c}_{1n} = -\frac{1}{n} \sum_{i=1}^n \frac{1}{2} h_1^2 b_K (c_{1i} + c_{2i}) 1_{\{X_i \in A^{c_n}, Y_i > a_n\}},$$

$$\begin{aligned} c_{1i} &= \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \frac{d^3}{du^3} \left([X_i - E(X | \theta'_0 X = u)] f_{\theta_0}(u, Y_i) \right)_{u=\theta'_0 X_i} \\ &\quad - \frac{1}{f_{\theta'_0 X}(\theta'_0 X_i)} \frac{d^3}{du^3} \left([X_i - E(X | \theta'_0 X = u)] f_{\theta'_0 X}(u) \right)_{u=\theta'_0 X_i} \\ &\quad + \frac{1 - \delta_i}{d(Y_i, \theta'_0 X_i)} \int_{(Y_i, \infty)} \frac{d^3}{du^3} \left([X_i - E(X | \theta'_0 X = u)] f_{\theta_0}(u, v) \right)_{u=\theta'_0 X_i} dv \end{aligned}$$

and

$$\begin{aligned}
c_{2i} &= -\delta_i \frac{\nabla_{\theta} f_{\theta}(\theta' X_i, Y_i)_{\theta=\theta_0}}{f_{\theta_0}^2(Y_i, \theta'_0 X_i)} \frac{d^2}{du^2} f_{\theta_0}(u, Y_i)|_{u=\theta'_0 X_i} \frac{\nabla_{\theta} f_{\theta' X}(\theta' X_i)_{\theta=\theta_0}}{f_{\theta'_0 X}^2(\theta'_0 X_i)} f_{\theta'_0 X}''(\theta'_0 X_i) \\
&\quad - (1 - \delta_i) \frac{\nabla_{\theta} d(\theta' X_i, Y_i)_{\theta=\theta_0}}{d^2(\theta'_0 X_i, Y_i)} \int_{(Y_i, \infty)} \frac{d^2}{du^2} f_{\theta_0}(u, v)_{u=\theta'_0 X_i} dv.
\end{aligned}$$

First we deal with $\tilde{c}_{1n}$. Lemma 2.3 in [24] leads to

$$a_n = H^{-1} \left(1 - \frac{s_n}{n}\right) \geq Y_{n-\tilde{s}_n:n},$$

where $\tilde{s}_n = \lfloor s_n/2 \rfloor$. Hence

$$\sqrt{n}|\tilde{c}_{1n}| \leq \frac{1}{\sqrt{n}} \sum_{i=n-\tilde{s}_n}^n \frac{1}{2} h_1^2 b_K |c_{1i} + c_{2i}| 1_{\{X_i \in A^{c_n}\}} = O_{\mathbb{P}}(n^{1/6} h_1^2 (\log(n))^{2+\varepsilon_1}) = o_{\mathbb{P}}(1).$$

As to c_{1n} and c_{2n} , we will show that $E(c_{kn}) = 0$, $Var(c_{kn}) = o(n^{-1})$ for $k = 1, 2$. For this set $f_3(\mathbf{x}, z, c)$ the density of (X, Z, C) and $g(c|\mathbf{x})$ the conditional density of C given X . By A2 we have

$$f_3(\mathbf{x}, z, c) = g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}).$$

Moreover, since $f_{\theta_0}(z|\theta'_0 \mathbf{x}) = f(z|\mathbf{x})$, we have that $f(\mathbf{x}, z) = f_{\theta_0}(\theta'_0 \mathbf{x}, z) f_X(\mathbf{x}) / f_{\theta'_0 X}(\theta'_0 \mathbf{x})$ and $\int 1_{\{c \leq z\}} f(z|\mathbf{x}) dz = d(\theta'_0 \mathbf{x}, c) / f_{\theta'_0 X}(\theta'_0 \mathbf{x})$. Hence

$$\begin{aligned}
E(c_{1i} 1_{\{X_i \in A^{c_n}\}}) &= \int \frac{1_{\{z \leq c, \mathbf{x} \in A^{c_n}\}}}{f_{\theta_0}(\theta'_0 \mathbf{x}, z)} \frac{d^3}{du^3} \left([\mathbf{x} - E(X|\theta'_0 X = u)] f_{\theta_0}(u, z) \right)_{u=\theta'_0 \mathbf{x}} g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dz dc \\
&\quad - \int \frac{1_{\{\mathbf{x} \in A^{c_n}\}}}{f_{\theta'_0 X}(\theta'_0 \mathbf{x})} \frac{d^3}{du^3} \left([\mathbf{x} - E(X|\theta'_0 X = u)] f_{\theta'_0 X}(u) \right)_{u=\theta'_0 \mathbf{x}} f_X(\mathbf{x}) d\mathbf{x} \\
&\quad + \int \frac{1_{\{z > c, \mathbf{x} \in A^{c_n}\}}}{d(\theta'_0 \mathbf{x}, c)} \int_{(c, \infty)} \frac{d^3}{du^3} \left([\mathbf{x} - E(X|\theta'_0 X = u)] f_{\theta_0}(u, v) \right)_{u=\theta'_0 \mathbf{x}} dv \\
&\quad g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} dz dc = 0,
\end{aligned}$$

where the last equation is consequence of change v to z in last coefficient and the fact that $\int g(c|\mathbf{x}) dc = 1$. Moreover c_{1i} are i.i.d. and, since $h_1^2 \rightarrow 0$, $Var(\sqrt{n}c_{1n}) \rightarrow 0$. Hence $\sqrt{n}c_{1n} \rightarrow 0$ in probability.

For c_{2n} ,

$$\begin{aligned}
E(c_{2i} 1_{\{X_i \in A^{c_n}\}}) &= - \int 1_{\{z \leq c\}} \frac{\nabla_{\theta} f_{\theta}(\theta' \mathbf{x}, z)_{\theta=\theta_0}}{f_{\theta_0}^2(\theta'_0 \mathbf{x}, z)} \frac{d^2}{du^2} f_{\theta_0}(u, z)|_{u=\theta'_0 \mathbf{x}} g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) 1_{\{\mathbf{x} \in A^{c_n}\}} d\mathbf{x} dz dc \\
&\quad + \int \frac{\nabla_{\theta} f_{\theta' X}(\theta' \mathbf{x})_{\theta=\theta_0}}{f_{\theta'_0 X}^2(\theta'_0 \mathbf{x})} f_{\theta'_0 X}''(\theta'_0 \mathbf{x}) f_X(\mathbf{x}) 1_{\{\mathbf{x} \in A^{c_n}\}} d\mathbf{x} \\
&\quad - \int 1_{\{z > c\}} \frac{\nabla_{\theta} d(\theta' \mathbf{x}, c)_{\theta=\theta_0}}{d^2(\theta'_0 \mathbf{x}, c)} \left(\int_{(c, \infty)} \frac{d^2}{du^2} f_{\theta_0}(u, v)_{u=\theta'_0 \mathbf{x}} dv \right) g(c|\mathbf{x}) f(z|\mathbf{x}) f_X(\mathbf{x}) \\
&\quad 1_{\{\mathbf{x} \in A^{c_n}\}} d\mathbf{x} dz dc \\
&= - \int 1_{\{z \leq c\}} \frac{\nabla_{\theta} f_{\theta}(\theta' \mathbf{x}, z)_{\theta=\theta_0}}{f_{\theta_0}(\theta'_0 \mathbf{x}, z)} \frac{d^2}{du^2} f_{\theta_0}(u, z)|_{u=\theta'_0 \mathbf{x}} g(c|\mathbf{x}) \frac{f_X(\mathbf{x})}{f_{\theta'_0 X}(\theta'_0 \mathbf{x})} 1_{\{\mathbf{x} \in A^{c_n}\}} d\mathbf{x} dz dc \\
&\quad + \int \frac{\nabla_{\theta} f_{\theta' X}(\theta' \mathbf{x})_{\theta=\theta_0}}{f_{\theta'_0 X}(\theta'_0 \mathbf{x})} \int \frac{d^2}{du^2} \left(\int f_{\theta_0}(u, z) dz \right)_{u=\theta'_0 \mathbf{x}} \frac{f_X(\mathbf{x})}{f_{\theta'_0 X}(\theta'_0 \mathbf{x})}
\end{aligned}$$

$$\begin{aligned}
& \int (1_{\{z \leq c\}} + 1_{\{z > c\}}) g(c|\mathbf{x}) 1_{\{\mathbf{x} \in A^{c_n}\}} dcd\mathbf{x} \\
& - \int 1_{\{v > c\}} \frac{\nabla_\theta d(\theta' \mathbf{x}, c)_{\theta=\theta_0}}{d(\theta'_0 \mathbf{x}, c)} \frac{d^2}{du^2} f_{\theta_0}(u, v)_{u=\theta'_0 \mathbf{x}} g(c|\mathbf{x}) \frac{f_X(\mathbf{x})}{f_{\theta'_0 X}(\theta'_0 \mathbf{x})} 1_{\{\mathbf{x} \in A^{c_n}\}} d\mathbf{x} dv dc \\
& = - \int \left[1_{\{z \leq c\}} \nabla_\theta \log f_\theta(z|\theta' \mathbf{x})_{\theta=\theta_0} + 1_{\{z > c\}} \nabla_\theta \log(1 - F_\theta(c|\theta' \mathbf{x}))_{\theta=\theta_0} \right] \\
& \quad \frac{d^2}{du^2} \left(\int f_{\theta_0}(u, z) dz \right)_{u=\theta'_0 \mathbf{x}} g(c|\mathbf{x}) \frac{f_X(\mathbf{x})}{f_{\theta'_0 X}(\theta'_0 \mathbf{x})} \frac{f_3(\mathbf{x}, z, c)}{f_3(\mathbf{x}, z, c)} 1_{\{\mathbf{x} \in A^{c_n}\}} dcd\mathbf{x} \\
& = -E \left(\left[\delta_1 \nabla_\theta \log f_\theta(Z_1|\theta' X_1)_{\theta=\theta_0} + (1 - \delta_1) \nabla_\theta \log(1 - F_\theta(C_1|\theta' X_1))_{\theta=\theta_0} \right] \right. \\
& \quad \left. \frac{d^2}{du^2} (f_{\theta_0}(u, Z_1))_{u=\theta'_0 X_1} g(C_1|X_1) \frac{f_X(X_1)}{f_{\theta'_0 X}(\theta'_0 X_1)} \frac{1}{f_3(X_1, Z_1, C_1)} 1_{\{X_1 \in A^{c_n}\}} \right)
\end{aligned}$$

Moreover, $1_{\{X_1 \in A^{c_n}\}} = 1_{\{f_{\theta'_0 X}(\theta'_0 X_1) > c_n\}}$ and is measurable with respect to $\theta'_0 X_1$. Additionally,

$$g(C_1|X_1) \frac{f_X(X_1)}{f_{\theta'_0 X}(\theta'_0 X_1)} \frac{1}{f_3(X_1, Z_1, C_1)} = \frac{1}{f_{\theta_0}(\theta'_0 X_1, Z_1)}.$$

Hence

$$\begin{aligned}
E(c_{2i} 1_{\{X_i \in A^{c_n}\}}) &= -E \left(\frac{\frac{d^2}{du^2} (f_{\theta_0}(u, Z_1))_{u=\theta'_0 X_1}}{f_{\theta_0}(\theta'_0 X_1, Z_1)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_1) > c_n\}} E \left[\delta_1 \nabla_\theta \log f_\theta(Z_1|\theta' X_1)_{\theta=\theta_0} \right. \right. \\
&\quad \left. \left. + (1 - \delta_1) \nabla_\theta \log(1 - F_\theta(C_1|\theta' X_1))_{\theta=\theta_0} | \theta'_0 X_1, C_1, Z_1 \right] \right) = 0,
\end{aligned}$$

where the last equation is a consequence of Lemma 4.

Finally, $\sqrt{n}c_{2n} \rightarrow 0$ in probability, which completes the proof. $\square$

Now we need some properties of the Kaplan-Meier estimator.

Lemma 6. *Under A9 we have*

$$\max_{i=1, \dots, n} \frac{1 - G(Y_i -)}{1 - G_n(Y_i -)} = O_{\mathbb{P}}(1). \quad (\text{A.1})$$

Moreover, set $H_0(s) = \mathbb{P}(Y \leq s, \delta = 0)$. Then we have

$$G_n(t) - G(t) = (1 - G(t)) \frac{1}{n} \sum_{k=1}^n \psi(Y_k, \delta_k, t) + R_n(t), \quad (\text{A.2})$$

where

$$\psi(Y_k, \delta_k, t) = \frac{1 - \delta_k}{1 - H(Y_k -)} 1_{\{Y_k \leq t\}} - \int_{(-\infty, t]} \frac{H_0(ds)}{1 - H(s-)} + \int_{(-\infty, t]} \frac{1_{\{Y_k \leq s\}} - H(s)}{(1 - H(s-))^2} H_0(ds)$$

and for $r_n/n \uparrow$, $(n - r_n)/\log(n) \rightarrow \infty$ and each $\varepsilon > 0$

$$\sup_{t \leq Y_{r_n:n}} |R_n| = O \left(\frac{(\log(n))^{2+\varepsilon}}{2n - r_{2n}} \right) \text{ with probability } 1.$$

Additionally,

$$\sup_{t \leq Y_{r_n:n}} \left| \frac{G_n(t) - G(t)}{1 - G(t)} \right| = \begin{cases} O\left(\sqrt{\frac{\log(n)}{2n-r_{2n}}}\right) & \text{with probability } 1 \\ O_{\mathbb{P}}\left(\frac{1}{\sqrt{n-r_n}}\right) & . \end{cases} \quad (\text{A.3})$$

Proof.

The proof of (A.1) can be found, e.g., in Zhou [31]. Equation (A.2) is a consequence of Theorems 1.1 and 1.4 from Stute [24], while (A.3) follows from Theorem 2 Csörgő [4]. Both rewritten for case where $C \sim G$ is a variable of interest observed if $1 - \delta = 1$. □

Lemma 7. *Under the conditions in Theorem 2 we have*

$$B_{1n} = \frac{1}{n} \sum_{j=1}^n \bar{H}_{1j} + \frac{1}{n} \sum_{k=1}^n \bar{N}_{1k} + o_{\mathbb{P}}(n^{-1/2}),$$

where

$$\begin{aligned} \bar{H}_{1j} = & -\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j-)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j)(1 - G_{\theta_0}(Y_j - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta_0 X_j} \\ & + E\left(\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j-)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j)(1 - G_{\theta_0}(Y_j - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta_0 X_j}\right), \end{aligned}$$

and

$$\bar{N}_{1k} = - \int \frac{d}{dt} [(E(X | \theta'_0 X = t) - u)(1 - G_{\theta_0}(v - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta'_0 \mathbf{u}} \frac{\psi(Y_k, \delta_k, v-)}{1 - G(v-)} f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv.$$

Proof.

Let $H_1(x, y) = \mathbb{P}(X \leq x, Y \leq y, \delta = 1)$ and note that, under A1 and A3, we have

$$H_1(dx, dy) = (1 - G(y-))F(dx, dy).$$

Hence

$$\begin{aligned} B_{1n} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\hat{r}^{[1]}(\theta'_i X_i, Y_i)_{\theta=\theta_0} - \tilde{r}^{[1]}(\theta'_i X_i, Y_i)_{\theta=\theta_0}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1^2 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G_n(Y_j-)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\ &\quad \left. - \frac{1}{h_1^2 h_2} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{v \leq a_n\}}}{1 - G(v-)} H_1(d\mathbf{u}, dv) \right] 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}. \end{aligned}$$

where $K'(t)$ denotes the derivatives of K with respect to t .

Finally, using

$$\frac{1}{1 - G_n(v-)} = \frac{1}{1 - G(v-)} + \frac{G_n(v-) - G(v-)}{(1 - G(v-))(1 - G_n(v-))}$$

we obtain the following decomposition

$$B_{1n} = b_{1n} + b_{2n} + b_{3n},$$

where

$$\begin{aligned} b_{1n} &= \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1^2 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\ &\quad \left. - \frac{1}{h_1^2 h_2} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{v \leq a_n\}}}{1 - G(v -)} H_1(d\mathbf{u}, dv) \right] 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}} \\ b_{2n} &= \frac{1}{h_1^2 h_2 n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{G_n(Y_j -) - G(Y_j -)}{(1 - G(Y_j -))^2} 1_{\{Y_j \leq a_n\}} \\ &\quad 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}} \\ b_{3n} &= \frac{1}{h_1^2 h_2 n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{(G_n(Y_j -) - G(Y_j -))^2}{(1 - G(Y_j -))^2 (1 - G_n(Y_j -))} \\ &\quad 1_{\{Y_j \leq a_n\}} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}. \end{aligned}$$

Using (A.1) and (A.3) it is easy to prove that, choosing $r_n = n - s_n$ and $s_n = cn^{2/3}(\log(n))^{2+\varepsilon_1}$, we obtain $\sqrt{n}b_{3n} = o_{\mathbb{P}}(1)$. Moreover, we have

$$b_{2n} = b_{21n} + b_{22n},$$

where

$$\begin{aligned} b_{21n} &= \frac{1}{h_1^2 h_2 n^3} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq i, j} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{\psi(Y_k, \delta_k, Y_j -)}{1 - G(Y_j -)} \\ &\quad 1_{\{Y_j \leq a_n\}} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}} \\ b_{22n} &= \frac{1}{h_1^2 h_2 n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{R_n(Y_j -) 1_{\{Y_j \leq a_n\}}}{(1 - G(Y_j -))^2} \\ &\quad 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}. \end{aligned}$$

As to b_{22n} , according to Csörgő [4], for n large enough and with high probability

$$a_n = H^{-1}\left(1 - \frac{s_n}{n}\right) \leq Y_{r_n:n},$$

where $r_n = n - \lfloor s_n/5 \rfloor + 1$ and $s_n \approx n^{2/3}(\log(n))^{2+\varepsilon_1}$. Hence, using Lemma 6 with $\varepsilon < \varepsilon_1$, we obtain

$$\sup_{t \leq a_n} |R_n(t)| \leq \sup_{t \leq Y_{r_n:n}} |R_n(t)| = o(n^{-2/3})$$

and

$$\sqrt{n}b_{22n} = o_{\mathbb{P}}(n^{-1/6}h_1^{-1}) = o_{\mathbb{P}}(1).$$

Concerning b_{1n} , set

$$\begin{aligned}
H_{ij} &= H(\theta'_0 X_i, Y_i, \delta_i, \theta'_0 X_j, Y_j, \delta_j) \\
&= \frac{1}{h_1^2 h_2} \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{\delta_j}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) 1_{\{Y_j \leq a_n\}} \right. \\
&\quad \left. - \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{v \leq a_n\}}}{1 - G(v -)} H_1(d\mathbf{u}, dv) \right] 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}
\end{aligned}$$

and

$$\bar{H}_j = E(H_{ij} | X_j, Y_j, \delta_j).$$

Then

$$\sqrt{n} b_{1n} = \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} (H_{ij} - \bar{H}_j) + \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} \bar{H}_j = \sqrt{n} b_{11n} + \sqrt{n} b_{12n}. \quad (\text{A.4})$$

For the first term, observe that

$$E(H_{ij} - \bar{H}_j) = E(H_{ij} - \bar{H}_j | X_i, Y_i, \delta_i) = E(H_{ij} - \bar{H}_j | X_j, Y_j, \delta_j) = 0.$$

Since the goal is to show that the vector $\sqrt{n} b_{11n} \in \mathbb{R}^d$ goes to zero in probability, it is enough to prove that component-wise. Let denote by $(v)_m$, for $m = 1, \dots, d$, the m th component of vector $v \in \mathbb{R}^d$. Hence

$$\begin{aligned}
E \left(\frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} (H_{ij} - \bar{H}_j)_m \right)^2 &= \frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i} \sum_{k=1}^n \sum_{l \neq k} E((H_{ij} - \bar{H}_j)_m (H_{kl} - \bar{H}_l)_m) \\
&= \frac{1}{n} [E(H_{12} - \bar{H}_2)_m^2 + E(H_{21} - \bar{H}_1)_m (H_{12} - \bar{H}_2)_m].
\end{aligned}$$

Let consider first $E(H_{12} - \bar{H}_2)_m^2$. Using A3,

$$E \left(\frac{\delta_2}{(1 - G(Y_2 -))^2} | Z_2, X_2 \right) = \frac{1}{1 - G(Z_2 -)}$$

and, since $Z_i \geq 0$, we have componentwise

$$\begin{aligned}
\frac{1}{n} E(H_{12} - \bar{H}_2)_m^2 &\leq \frac{1}{n} E((H_{12})_m^2) \leq \frac{1}{h_1^4 h_2^2} E \left(\frac{\delta_1 1_{\{f_{\theta'_0 X}(\theta'_0 X_1) > c_n, 0 \leq Y_1 \leq a_n, 0 \leq Y_2 \leq a_n\}}}{f_{\theta'_0}^2(\theta'_0 X_1, Y_1)} \frac{\delta_2}{(1 - G(Y_2 -))^2} \right. \\
&\quad \left. (X_1 - X_2)_m^2 (K')^2 \left(\frac{\theta'_0 X_1 - \theta'_0 X_2}{h_1} \right) K^2 \left(\frac{Y_1 - Y_2}{h_2} \right) \right) \\
&= \frac{1}{h_1^3 h_2} \int \frac{f(\mathbf{u}, v) (1 - G(v + h_2 z_2 -)) 1_{\{f_{\theta'_0 X}(\theta'_0 \mathbf{u} + h_1 z_1) > c_n\}}}{f_{\theta_0}(\theta'_0 \mathbf{u} + h_1 z_1, v + h_2 z_2) (1 - G(v -))} (K')^2(z_1) K^2(z_2) \\
&\quad E[(X_1 - \mathbf{u})_m^2 | \theta'_0 X_1 = \theta'_0 \mathbf{u} + h_1 z_1] 1_{\{0 \leq v \leq a_n, 0 \leq v + h_2 z_2 \leq a_n\}} dz_1 dz_2 d\mathbf{u} dv.
\end{aligned}$$

Moreover, since $\frac{f(\mathbf{u}, v)}{f_{\theta_0}(\theta'_0 \mathbf{u}, v)} = \frac{f_X(\mathbf{u})}{f_{\theta'_0 X}(\theta'_0 \mathbf{u})}$, the first term of a Taylor expansion is just

$$\frac{a_n}{h_1^3 h_2} \int E[(X_1 - \mathbf{u})_m^2 | \theta'_0 X_1 = \theta'_0 \mathbf{u}] 1_{\{f_{\theta'_0 X}(\theta'_0 \mathbf{u}) > c_n\}} \frac{f_X(\mathbf{u})}{f_{\theta'_0 X}(\theta'_0 \mathbf{u})} d\mathbf{u} \int (K')^2(z_1) K^2(z_2) dz_1 dz_2.$$

Hence, using A5, we obtain

$$\frac{1}{n}E((H_{12})_m^2) \leq \frac{a_n}{c_n n h_1^3 h_2} C,$$

where C is a constant.

Moreover, according to inequality (1) in [22], we have for $p = j/n$

$$H^{-1}(p) \leq E(Y_{j:n}) + \sigma \frac{H_{j:n}(p)}{\sqrt{p(1-p)}},$$

where $\sigma = \text{Var}(Y) < \infty$.

Furthermore, according to the inequality in Proposition 1 in [10],

$$E(Y_{j:n}) = O(\sqrt{(j-1)/(n-j-1)}) = O(\sqrt{(p-1/n)/(1-p-1/n)}).$$

Hence

$$a_n \leq H^{-1}\left(\frac{n - [n^{2/3}]}{n}\right) = O(n^{1/6})$$

and, using A6, we obtain $\frac{1}{n}E((H_{12})_m^2) = o(1)$ for $m = 1, \dots, d$. A similar property holds for $E(H_{21} - \bar{H}_1)_m(H_{12} - \bar{H}_2)_m$. So that

$$\sqrt{n}b_{11n} \xrightarrow{\mathbb{P}} 0.$$

For $\sqrt{n}b_{12n}$, using A2, we have

$$\begin{aligned} \bar{H}_j &= E\left(\frac{1}{h_1^2 h_2} \int \frac{1_{\{z \leq c\}}}{f_{\theta_0}(x, z)} \frac{\delta_j}{1 - G(Y_j -)} (E(X_i | \theta_0 X_i = x) - X_j) K' \left(\frac{x - \theta'_0 X_j}{h_1} \right) K \left(\frac{z - Y_j}{h_2} \right) \right. \\ &\quad \left. f_{\theta_0}(x, z) g_{\theta_0}(c|x) 1_{\{f_{\theta'_0 X}(x) > c_n, Y_j \leq a_n, z \leq a_n\}} dx dz \right. \\ &\quad \left. - \frac{1}{h_1^2 h_2} \int \int \frac{1_{\{z \leq c\}}}{f_{\theta_0}(x, z)} (E(X_i | \theta_0 X_i = x) - \mathbf{u}) K' \left(\frac{x - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{z - v}{h_2} \right) f_{\theta_0}(x, z) g_{\theta_0}(c|x) \right. \\ &\quad \left. f(\mathbf{u}, v) 1_{\{f_{\theta'_0 X}(x) > c_n, v \leq a_n, z \leq a_n\}} dx dz dv d\mathbf{u} | X_j, Y_j, \delta_j \right) \end{aligned}$$

Finally, using a change of variable and a Taylor expansion, we obtain:

$$\bar{H}_j = \bar{H}_{1j} + h_{1j},$$

where

$$\begin{aligned} \bar{H}_{1j} &= -\frac{\delta_j}{1 - G(Y_j -)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j)(1 - G_{\theta_0}(Y_j - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta_0 X_j} 1_{\{Y_j \leq a_n\}} \\ &\quad + E\left(\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j)(1 - G_{\theta_0}(Y_j - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta_0 X_j} \right) \end{aligned}$$

and

$$h_{1j} = o_{\mathbb{P}}(1).$$

Moreover, $h_{1j} = \bar{H}_j - \bar{H}_{1j}$. Hence h_{1j} are i.i.d. with expectation 0. Hence $\frac{1}{\sqrt{n}} \sum_{j=1}^n h_{1j} \rightarrow 0$ in probability.

For b_{21n} , equation (A.1) leads to

$$\begin{aligned} \sqrt{n}b_{21n} &= \frac{\sqrt{n}}{h_1^2 h_2 n^3} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq i, j} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{\psi(Y_k, \delta_k, Y_j -)}{1 - G(Y_j -)} \\ &\quad 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n, Y_j \leq a_n\}} + o_{\mathbb{P}}(1) \end{aligned}$$

Set

$$N_{ijk} = \frac{1}{h_1^2 h_2} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{\psi(Y_k, \delta_k, Y_j -)}{1 - G(Y_j -)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}.$$

Since $E(\psi(Y_k, \delta_k, t)) = 0$, we have

$$E(N_{ijk}|Y_i, \delta_i, X_i) = E(N_{ijk}|Y_j, \delta_j, X_j) = 0.$$

Set

$$\bar{N}_k = E(N_{ijk}|Y_k, \delta_k).$$

Hence

$$\sqrt{n}b_{2n} = \frac{\sqrt{n}}{n^3} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq i, j} (N_{ijk} - \bar{N}_k) + \frac{1}{\sqrt{n}} \sum_{k=1}^n \bar{N}_k + o_{\mathbb{P}}(1).$$

Moreover, since $E(\psi(Y_k, \delta_k, t-)) = 0$ we have, for every component $m = 1, \dots, d$,

$$\begin{aligned} E \left(\frac{\sqrt{n}}{n^3} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq i, j} (N_{ijk} - \bar{N}_k)_m \right)^2 &= n^{-2} E(N_{123} - \bar{N}_3)_m^2 + n^{-1} E(N_{123} - \bar{N}_3)_m (N_{423} - \bar{N}_3)_m \\ &\quad + n^{-1} E(N_{123} - \bar{N}_3)_m (N_{143} - \bar{N}_3)_m + n^{-1} E(N_{123} - \bar{N}_3)_m (N_{124} - \bar{N}_4)_m. \end{aligned}$$

Furthermore, we can show that

$$E(\psi^2(Y_k, \delta_k, t-)) \leq 4 \frac{1 - F(t-)}{(1 - H(t-))^2}.$$

Moreover, since $\int \frac{F(dt)}{1 - G(t-)} < \infty$, there exists some constant K_1 , so that

$$1 - F(t-) \leq K_1 (1 - G(t-))$$

and

$$1 - F(t-) \leq \sqrt{K_1} \sqrt{1 - H(t-)}.$$

Hence

$$E(\psi^2(Y_k, \delta_k, t-)) \leq \frac{4\sqrt{K_1}}{(1 - H(t-))^{3/2}}$$

Additionally,

$$(1 - H(a_n -))^{-3/2} = \left(\frac{n}{s_n} \right)^{3/2} = o(n^{1/2}).$$

Finally, for every $m = 1, \dots, d$, we obtain

$$n^{-2}E((N_{123} - \bar{N}_3)_m^2) = o_{\mathbb{P}}(n^{-2}h_1^{-3}h_2^{-1}n^{1/2}) = o_{\mathbb{P}}(1)$$

$$n^{-1}E((N_{123} - \bar{N}_3)_m(N_{423} - \bar{N}_3)_m) = o_{\mathbb{P}}(n^{-1}h_1^{-2}n^{1/2}) = o_{\mathbb{P}}(1)$$

$$n^{-1}E((N_{123} - \bar{N}_3)_m(N_{143} - \bar{N}_3)_m) = o_{\mathbb{P}}(n^{-1}h_1^{-2}n^{1/2}) = o_{\mathbb{P}}(1)$$

and, since $E(\psi(Y_k, \delta_k, t-)) = 0$,

$$n^{-1}E((N_{123} - \bar{N}_3)_m(N_{124} - \bar{N}_4)_m) = 0.$$

Hence, as before, the first term goes to zero in probability. As to the second, set

$$\begin{aligned} \bar{N}_k &= \frac{1}{h_1^2 h_2} \int \frac{1_{\{z \leq c\}}}{f_{\theta_0}(x, z)} (E(X|\theta'_0 X = x) - \mathbf{u}) K' \left(\frac{x - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{z - v}{h_2} \right) \psi(Y_k, \delta_k, v-) \\ &\quad f(\mathbf{u}, v) f_{3\theta_0}(x, z, c) 1_{\{f_{\theta'_0 X}(x) > c_n, v \leq a_n, z \leq a_n\}} dx dz dc d\mathbf{u} dv. \end{aligned}$$

Using Taylor expansions we can show that $\bar{N}_k = \bar{N}_{1k} + n_{1k}$, where

$$\bar{N}_{1k} = - \int \frac{d}{dt} [(E(X|\theta'_0 X = t) - \mathbf{u})(1 - G_{\theta_0}(v - |t|) 1_{\{f_{\theta'_0 X}(t) > c_n\}})]_{t=\theta'_0 \mathbf{u}} \psi(Y_k, \delta_k, v-) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv$$

and $n_{1k} = o_{\mathbb{P}}(1)$. Finally, n_{1k} are i.i.d. random variables with expectation 0. Hence $\frac{1}{\sqrt{n}} \sum_{k=1}^n n_{1k} \rightarrow 0$ in probability. This completes the proof. $\square$

Now we state some similar lemmas.

Lemma 8. *Under the conditions in Theorem 2 we have*

$$B_{3n} = \frac{1}{n} \sum_{j=1}^n \bar{H}_{3j} + \frac{1}{n} \sum_{k=1}^n \bar{N}_{3k} + o_{\mathbb{P}}(n^{-1/2}).$$

where

$$\begin{aligned} \bar{H}_{3j} &= \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} \frac{d}{dt} ((E(X|\theta'_0 X = t) - X_j) 1_{\{f_{\theta'_0 X}(t) > c_n\}})_{t=\theta'_0 X_j} \\ &\quad - E \left(\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} \frac{d}{dt} ((E(X|\theta'_0 X = t) - X_j) 1_{\{f_{\theta'_0 X}(t) > c_n\}})_{t=\theta'_0 X_j} \right), \end{aligned}$$

and

$$\bar{N}_{3k} = \int \frac{d}{dt} ([E(X|\theta'_0 X = t) - \mathbf{u}] 1_{\{f_{\theta'_0 X}(t) > c_n\}})_{t=\theta'_0 \mathbf{u}} \psi(Y_k, \delta_k, v-) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv.$$

Lemma 9. *Under the conditions in Theorem 2 we have*

$$B_{5n} = \frac{1}{n} \sum_{j=1}^n \bar{H}_{5j} + \frac{1}{n} \sum_{k=1}^n \bar{N}_{5k} + o_{\mathbb{P}}(n^{-1/2}).$$

where

$$\begin{aligned}\bar{H}_{5j} &= -\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j) G_{\theta_0}(Y_j - |t) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta'_0 X_j} \\ &\quad + E\left(\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j) G_{\theta_0}(Y_j - |t) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta'_0 X_j}\right),\end{aligned}$$

and

$$\bar{N}_{5k} = - \int \frac{d}{dt} (E(X | \theta'_0 X = t) - \mathbf{u}) G_{\theta_0}(v - |t) 1_{\{f_{\theta'_0 X}(t) > c_n\}})_{t=\theta'_0 \mathbf{u}} \psi(Y_k, \delta_k, v-) 1_{\{v \leq a_n\}} f(\mathbf{u}, v) d\mathbf{u} dv.$$

Lemma 10. Under the conditions in Theorem 2 we have

$$B_{2n} = \frac{1}{n} \sum_{j=1}^n \bar{H}_{2j} + \frac{1}{n} \sum_{k=1}^n \bar{N}_{2k} + o_{\mathbb{P}}(n^{-1/2}).$$

where

$$\begin{aligned}\bar{H}_{2j} &= E\left(E(\nabla_{\theta} \log f_{\theta' X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = \theta'_0 X_j) (1 - G_{\theta_0}(Y_j - | \theta'_0 X_j)) \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}}\right) \\ &\quad - E(\nabla_{\theta} \log f_{\theta' X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = \theta'_0 X_j) (1 - G_{\theta_0}(Y_j - | \theta'_0 X_j)) \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}},\end{aligned}$$

and

$$\bar{N}_{2k} = - \int E(\nabla_{\theta} f_{\theta}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = u) (1 - G_{\theta_0}(v - | \mathbf{u})) f_{\theta_0}(u, v) \psi(Y_k, \delta_k, v-) 1_{\{f_{\theta'_0 X}(u) > c_n, v \leq a_n\}} d\mathbf{u} dv.$$

Proof.

The proof is similar to the proof of Lemma 7. Observe only, that using Lemma 4, we have

$$\begin{aligned}E(\nabla_{\theta} \log(f_{\theta}(\theta' X_i, Y_i))_{\theta=\theta_0} | \theta'_0 X_i, Y_i) &= E(\nabla_{\theta} \log(f_{\theta}(\theta' X_i | Y_i))_{\theta=\theta_0} | \theta'_0 X_i, Y_i) \\ &\quad + E(\nabla_{\theta} \log(f_{\theta' X}(\theta' X_i))_{\theta=\theta_0} | \theta'_0 X_i) = E(\nabla_{\theta} \log(f_{\theta' X}(\theta' X_i))_{\theta=\theta_0} | \theta'_0 X_i)\end{aligned}$$

□

Lemma 11. Under the conditions in Theorem 2 we have

$$B_{4n} = \frac{1}{n} \sum_{j=1}^n \bar{H}_{4j} + \frac{1}{n} \sum_{k=1}^n \bar{N}_{4k} + o_{\mathbb{P}}(n^{-1/2}).$$

where

$$\begin{aligned}\bar{H}_{4j} &= \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} E(\nabla_{\theta} \log f_{\theta' X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = \theta'_0 X_j) 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}} \\ &\quad - E\left(\frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} E(\nabla_{\theta} \log f_{\theta' X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = \theta'_0 X_j) 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}}\right),\end{aligned}$$

and

$$\bar{N}_{4k} = \int E(\nabla_{\theta} \log f_{\theta' X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = u) \psi(Y_k, \delta_k, v-) 1_{\{v \leq a_n\}} f_{\theta_0}(u, v) 1_{\{f_{\theta'_0 X}(u) > c_n\}} d\mathbf{u} dv.$$

Lemma 12. *Under the conditions in Theorem 2 we have*

$$B_{6n} = \frac{1}{n} \sum_{j=1}^n \bar{H}_{6j} + \frac{1}{n} \sum_{k=1}^n \bar{N}_{6k} + o_{\mathbb{P}}(n^{-1/2}).$$

where

$$\begin{aligned} \bar{H}_{6j} = & -G_{\theta_0}(Y_j - |\theta'_0 X_j|) \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} E(\nabla_{\theta} \log f_{\theta'X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = \theta'_0 X_j) 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}} \\ & + E\left(G_{\theta_0}(Y_j - |\theta'_0 X_j|) \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} E(\nabla_{\theta} \log f_{\theta'X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = \theta'_0 X_j) 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}}\right), \end{aligned}$$

and

$$\begin{aligned} \bar{N}_{6k} = & - \int E(\nabla_{\theta} \log f_{\theta'X}(\theta' X_i)_{\theta=\theta_0} | \theta'_0 X_i = u) G_{\theta_0}(v - |u) \psi(Y_k, \delta_k, v -) G_{\theta_0}(v - |u) \\ & f_{\theta_0}(u, v) 1_{\{f_{\theta'_0 X}(u) > c_n, v \leq a_n\}} dx du dv. \end{aligned}$$

Lemma 13. *Under the conditions in Theorem 2 we have*

$$B_{1n} + B_{3n} + B_{5n} + B_{2n} + B_{4n} + B_{6n} = o_{\mathbb{P}}(n^{-1/2}).$$

Proof.

Using Lemmas 7-12, we can show that

$$\bar{H}_{1j} + \bar{H}_{3j} + \bar{H}_{5j} = \bar{H}_{2j} + \bar{H}_{4j} + \bar{H}_{6j} = 0$$

and

$$\bar{N}_{1k} + \bar{N}_{3k} + \bar{N}_{5k} = \bar{N}_{2k} + \bar{N}_{4k} + \bar{N}_{6k} = 0.$$

⊗

Lemma 14. *Under the conditions in Theorem 2 we have*

$$A_{7n} + A_{8n} + A_{9n} + A_{10n} + A_{11n} + A_{12n} = o_{\mathbb{P}}(n^{-1/2}).$$

Proof. See Appendix Appendix B.

⊗

Lemma 15. *Under the conditions in Theorem 2 we have*

$$\hat{l}_n^{[2]}(\theta) \rightarrow l^{[2]}(\theta)$$

Proof.

We have

$$\begin{aligned} \hat{l}_n^{[2]}(\theta) = & \frac{1}{n} \sum_{i=1}^n \left[\delta_i \left(\frac{\hat{r}^{[2]}(\theta' X_i, Y_i)}{\hat{r}(\theta' X_i, Y_i)} - \frac{\hat{r}^{[1]}(\theta' X_i, Y_i)(\hat{r}^{[1]}(\theta' X_i, Y_i))^t}{\hat{r}^2(\theta' X_i, Y_i)} - \frac{\hat{s}^{[2]}(\theta' X_i)}{\hat{s}(\theta' X_i)} + \frac{\hat{s}^{[1]}(\theta' X_i)(\hat{s}^{[1]}(\theta' X_i))^t}{\hat{s}^2(\theta' X_i)} \right) \right. \\ & + (1 - \delta_i) \left(\frac{\hat{d}^{[2]}(\theta' X_i, Y_i)}{\hat{d}(\theta' X_i, Y_i)} - \frac{\hat{d}^{[1]}(\theta' X_i, Y_i)(\hat{d}^{[1]}(\theta' X_i, Y_i))^t}{\hat{d}^2(\theta' X_i, Y_i)} - \frac{\hat{s}^{[2]}(\theta' X_i)}{\hat{s}(\theta' X_i)} + \frac{\hat{s}^{[1]}(\theta' X_i)(\hat{s}^{[1]}(\theta' X_i))^t}{\hat{s}^2(\theta' X_i)} \right) \Big] \\ & 1_{\{Y_i \leq a_n, X_i \in A^{c_n}\}} \end{aligned}$$

Recall,

$$\hat{r}^{[2]}(y, \theta'x) = \frac{1}{nh_1^3 h_2} \sum_{j=1}^n 1_{\{Y_j \leq a_n\}} \frac{\delta_j}{1 - G(Y_j -)} (x - X_j)(x - X_j)^t K''\left(\frac{\theta'x - \theta'X_j}{h_1}\right) K\left(\frac{y - Y_j}{h_2}\right).$$

Since $a_n \rightarrow \tau_H$ and $c_n \rightarrow 0$, it is easy to show, that

$$\hat{r}^{[2]}(\theta'x, y) \rightarrow \frac{d^2}{dt^2} \{f_\theta(t, y) E((\mathbf{x} - X)(\mathbf{x} - X)^t | \theta'X = t)\}_{t=\theta'x} \text{ in probability.}$$

Moreover, using Lemma 3, it can be proved that

$$\frac{d^2}{dt^2} \{f_\theta(y, t) E((\mathbf{x} - X)(\mathbf{x} - X)^t | \theta'X = t)\}_{t=\theta'x} = f_\theta^{[2]}(\theta'x, y).$$

Similarly, it can be shown that $\hat{s}^{[2]}(\theta'x) \rightarrow f_{\theta'X}^{[2]}(\theta'x)$ and $\hat{d}^2(\theta'x, y) \rightarrow d_\theta^{[2]}(\theta'x, y)$. This completes the proof. $\square$

Proof of Theorem 4.

In view of Lemma 2, It remains to show that

$$\sum_{k=1}^{12} A_{kn} \xrightarrow{\mathbb{P}} 0.$$

For this, note that for every $\varepsilon > 0$ there exist a small c and a sequence $b_n \rightarrow \tau_H$ so that

$$1 - H(b_n) = \frac{c}{n} \text{ and } \mathbb{P}(Y_i \leq b_n \text{ for } i = 1, \dots, n) \geq 1 - \varepsilon.$$

Additionally, observe that the remainder $R_n(Y_j)$ in Lemma 6 can be written as

$$R_n(Y_j) = R_n(Y_j)1_{\{Y_j \leq Y_{n-\tilde{s}_n:n}\}} + R_n(Y_j)1_{\{Y_j > Y_{n-\tilde{s}_n:n}\}}$$

with $\tilde{s}_n \approx n^{1/6}(\log(n))^{2+\varepsilon_1}$. Hence

$$R_n(Y_j) = o_{\mathbb{P}}(n^{-1/6}) + O_{\mathbb{P}}(1)1_{\{Y_j > Y_{n-\tilde{s}_n:n}\}}. \quad (\text{A.5})$$

The proof goes similarly to proofs of Lemmas 7-13 with a_n replaced by b_n and using (A.5) when necessary. $\square$

Proof of Theorem 3.

According to Proposition A.9 from [17] we may replace $A^{c_n} = A_n^{c_n} = \{\mathbf{x} : f_{\theta'_n X}(\theta'_n \mathbf{x}) > c_n\}$ by $A^{c_n} = A_0^{c_n} = \{\mathbf{x} : f_{\theta'_0 X}(\theta'_0 \mathbf{x}) > c_n\}$. The rest of the proof is similar to the proof of Theorem 2 but it is much longer. Since the proof is similar to the proof of Theorem 2 but it is much longer, we will present only parts of it in Appendix C. $\square$

Appendix B. Proof of Lemma 14.

Let first consider the terms in A_{7n} and A_{8n} . Recall

$$A_{7n} = \frac{1}{n} \sum_{i=1}^n \left[\delta_i \frac{(\hat{r}^{[1]}(\theta'_0 X_i, Y_i) - f_{\theta'_0}^{[1]}(\theta'_0 X_i, Y_i))(f_{\theta'_0}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i))}{f_{\theta'_0}^2(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{c_n}, Y_i \leq a_n\}}$$

and

$$A_{8n} = \frac{1}{n} \sum_{i=1}^n \left[\delta_i \nabla_{\theta} \log(\hat{r}(\theta' X_i, Y_i))|_{\theta=\theta_0} \frac{(f_{\theta_0}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i))^2}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} \right] 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}}.$$

Set

$$\begin{aligned} M_n(X_i, Y_i) &= \frac{\hat{r}^{[1]}(\theta'_0 X_i, Y_i) - f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{cn}\}} \\ N_n(X_i, Y_i) &= \frac{f_{\theta_0}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}}. \end{aligned}$$

Hence

$$\begin{aligned} A_{7n} &= \frac{1}{n} \sum_{i=1}^n \delta_i M_n(X_i, Y_i) N_n(X_i, Y_i) \\ A_{8n} &= \frac{1}{n} \sum_{i=1}^n \delta_i \nabla_{\theta} \log(\hat{r}(\theta' X_i, Y_i))|_{\theta=\theta_0} N_n^2(X_i, Y_i) \end{aligned}$$

Moreover,

$$M_n(X_i, Y_i) = M_{1n}(X_i, Y_i) + M_{2n}(X_i, Y_i),$$

where

$$\begin{aligned} M_{1n}(X_i, Y_i) &= \frac{\hat{r}^{[1]}(\theta'_0 X_i, Y_i) - \tilde{r}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}} \\ M_{2n}(X_i, Y_i) &= \frac{\tilde{r}^{[1]}(\theta'_0 X_i, Y_i) - f_{\theta_0}^{[1]}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{cn}, Y_i \leq a_n\}}. \end{aligned}$$

Further $\sup_{x,y} |M_{2n}(x, y)| = O_{\mathbb{P}}(h_1^2 + h_2^2)$. As to M_{1n} , recall

$$\begin{aligned} M_{1n}(X_i, Y_i) &= \frac{1}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1^2 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G_n(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\ &\quad \left. - \frac{1}{h_1^2 h_2} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv \right] \\ &\quad 1_{\{Y_i \leq a_n\}} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}} \end{aligned}$$

Hence

$$M_{1n}(X_i, Y_i) = M_{11n}(X_i, Y_i) + M_{12n}(X_i, Y_i),$$

where

$$\begin{aligned}
M_{11n}(X_i, Y_i) &= \frac{1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1^2 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\
&\quad \left. - \frac{1}{h_1^2 h_2} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv \right] 1_{\{Y_i \leq a_n\}} \\
M_{12n}(X_i, Y_i) &= \frac{1}{h_1^2 h_2 n} \sum_{j \neq i} \frac{\delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \\
&\quad \frac{G_n(Y_j -) - G(Y_j -)}{(1 - G(Y_j -))(1 - G_n(Y_j -))} 1_{\{Y_j \leq a_n\}} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}
\end{aligned}$$

Moreover, note that, by (A.3),

$$\sup_{t \leq Y_{r_n:n}} \left| \frac{G_n(t) - G(t)}{1 - G(t)} \right| = O(n^{-1/3}).$$

Hence, by choosing $r_n = n - \lfloor s_n/5 \rfloor + 1$ and since $a_n \leq Y_{r_n:n}$ with high probability, we have

$$\sup_{x,y} |M_{12n}(x, y)| = O_{\mathbb{P}}(h_1^{-1} n^{-1/3}).$$

and using Proposition 6 in [1],

$$\sup_{x,y} |M_{11n}(x, y)| = O_{\mathbb{P}}(n^{-1/2} h_1^{-3/2} h_2^{-1/2} (\log(n))^{1/2})$$

For N_n , we have

$$N_n(X_i, Y_i) = N_{1n}(X_i, Y_i) + N_{2n}(X_i, Y_i),$$

where

$$\begin{aligned}
N_{1n}(X_i, Y_i) &= \frac{\tilde{r}(\theta'_0 X_i, Y_i) - \hat{r}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{c_n}\}} \\
N_{2n}(X_i, Y_i) &= \frac{f_{\theta_0}(\theta'_0 X_i, Y_i) - \tilde{r}(\theta'_0 X_i, Y_i)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} 1_{\{X_i \in A^{c_n}\}}.
\end{aligned}$$

Furthermore, $\sup_{x,y} |N_{2n}(x, y)| = O_{\mathbb{P}}(h_1^2 + h_2^2)$. As to N_{1n} , recall

$$\begin{aligned}
N_{1n}(X_i, Y_i) &= \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1 h_2} \int K \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv \right. \\
&\quad \left. - \frac{1}{h_1 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G_n(Y_j -)} K \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right] 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}
\end{aligned}$$

Hence

$$N_{1n}(X_i, Y_i) = N_{11n}(X_i, Y_i) + N_{12n}(X_i, Y_i),$$

where

$$\begin{aligned}
N_{11n}(X_i, Y_i) &= \frac{1}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1 h_2} \int K\left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1}\right) K\left(\frac{Y_i - v}{h_2}\right) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv \right. \\
&\quad \left. - \frac{1}{h_1 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} K\left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1}\right) K\left(\frac{Y_i - Y_j}{h_2}\right) \right] 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}} \\
N_{12n}(X_i, Y_i) &= -\frac{1}{h_1 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq a_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K\left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1}\right) K\left(\frac{Y_i - Y_j}{h_2}\right) \\
&\quad \frac{G_n(Y_j -) - G(Y_j -)}{(1 - G(Y_j -))(1 - G_n(Y_j -))} 1_{\{Y_j \leq a_n\}}
\end{aligned}$$

Moreover, using Proposition 6 in [1],

$$\sup_{x, y} |N_{11n}(x, y)| = O_{\mathbb{P}}(n^{-1/2} h_1^{-1/2} h_2^{-1/2} (\log(n))^{1/2}).$$

and, by (A.3),

$$\sup_{x, y} |N_{12n}(x, y)| = O_{\mathbb{P}}(n^{-1/3}).$$

Finally, it can be shown that

$$\sqrt{n} A_{8n} \xrightarrow{\mathbb{P}} 0.$$

For $\sqrt{n} A_{7n}$ we need to show that

$$\sqrt{n} \frac{1}{n} \sum_{i=1}^n \delta_i M_{11n}(X_i, Y_i) N_{12n}(X_i, Y_i) \xrightarrow{\mathbb{P}} 0 \quad (\text{B.1})$$

$$\sqrt{n} \frac{1}{n} \sum_{i=1}^n \delta_i M_{12n}(X_i, Y_i) N_{11n}(X_i, Y_i) \xrightarrow{\mathbb{P}} 0 \quad (\text{B.2})$$

$$\sqrt{n} \frac{1}{n} \sum_{i=1}^n \delta_i M_{11n}(X_i, Y_i) N_{11n}(X_i, Y_i) \xrightarrow{\mathbb{P}} 0 \quad (\text{B.3})$$

$$\sqrt{n} \frac{1}{n} \sum_{i=1}^n \delta_i M_{11n}(X_i, Y_i) N_{2n}(X_i, Y_i) \xrightarrow{\mathbb{P}} 0 \quad (\text{B.4})$$

and

$$\sqrt{n} \frac{1}{n} \sum_{i=1}^n \delta_i M_{12n}(X_i, Y_i) N_{2n}(X_i, Y_i) \xrightarrow{\mathbb{P}} 0 \quad (\text{B.5})$$

As to (B.1) and (B.2), we choose $b_n = H^{-1}(1 - \frac{\tilde{s}_n}{n})$ with $\tilde{s}_n \approx n^{5/6} (\log(n))^{2+\varepsilon_1}$ and write $1_{\{Y_j \leq a_n\}} = 1_{\{Y_j \leq b_n\}} + 1_{\{b_n < Y_j \leq a_n\}}$. The proof follows from (A.3) and uniform bounds for M_{11n} and N_{11n} .

For (B.3), set

$$\begin{aligned}\psi_1(X_i, Y_i, X_j, Y_j, \delta_j) &= \int K\left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1}\right) K\left(\frac{Y_i - v}{h_2}\right) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv \\ &\quad - \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} K\left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1}\right) K\left(\frac{Y_i - Y_j}{h_2}\right)\end{aligned}$$

and

$$\begin{aligned}\psi_2(X_i, Y_i, X_j, Y_j, \delta_j) &= \int (X_i - \mathbf{u}) K'\left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1}\right) K\left(\frac{Y_i - v}{h_2}\right) f(\mathbf{u}, v) 1_{\{v \leq a_n\}} d\mathbf{u} dv \\ &\quad - \frac{\delta_j 1_{\{Y_j \leq a_n\}}}{1 - G(Y_j -)} (X_i - X_j) K'\left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1}\right) K\left(\frac{Y_i - Y_j}{h_2}\right)\end{aligned}$$

Then the term (B.3) is equal to

$$\frac{\sqrt{n}}{n^3 h_1^3 h_2^2} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq i} \frac{\delta_i}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} \psi_1(X_i, Y_i, X_j, Y_j, \delta_j) \psi_2(X_i, Y_i, X_k, Y_k, \delta_k).$$

Moreover, for $m = 1, 2$

$$E(\psi_m(X_i, Y_i, X_j, Y_j, \delta_j) | X_i, Y_i) = 0$$

and

$$E|\psi_m(X_i, Y_i, X_j, Y_j, \delta_j)| = O(h_1 h_2).$$

Hence (B.3) is fulfilled. For (B.4) and (B.5), observe that

$$N_{2n}(X_i, Y_i) = \left[-h_2^2 \frac{1}{2} \frac{d^2}{dy^2} f_\theta(\theta' X_i, y)|_{y=Y_i} - h_1^2 \frac{1}{2} \frac{d^2}{dx^2} f_\theta(x, Y_i)|_{x=\theta' X_i} \right] \int z^2 K(z) dz + O_{\mathbb{P}}(h_1^4 + h_2^4).$$

Hence, it remains to prove that

$$\frac{\sqrt{n}}{n^2 h_1^2 h_2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i \psi_2(X_i, Y_i, X_j, Y_j, \delta_j)}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} \left[-h_2^2 \frac{1}{2} \frac{d^2}{dy^2} f_\theta(\theta' X_i, y)|_{y=Y_i} - h_1^2 \frac{1}{2} \frac{d^2}{dx^2} f_\theta(x, Y_i)|_{x=\theta' X_i} \right] \xrightarrow{\mathbb{P}} 0$$

and, using Lemma 6, that

$$\begin{aligned}&\frac{\sqrt{n}}{n^3 h_1^2 h_2} \sum_{i=1}^n \sum_{j \neq i} \sum_{k \neq i} \delta_i \frac{\delta_j (X_i - X_j)}{f_{\theta_0}^2(\theta'_0 X_i, Y_i)} K'\left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1}\right) K\left(\frac{Y_i - Y_j}{h_2}\right) \frac{\psi(Y_k, \delta_k, Y_j -)}{1 - G(Y_j -)} \\ &\quad 1_{\{Y_j \leq a_n\}} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}} \left[-h_2^2 \frac{1}{2} \frac{d^2}{dy^2} f_\theta(\theta' X_i, y)|_{y=Y_i} - h_1^2 \frac{1}{2} \frac{d^2}{dx^2} f_\theta(x, Y_i)|_{x=\theta' X_i} \right] \xrightarrow{\mathbb{P}} 0.\end{aligned}$$

The proof is similar to that of Lemma 10, with the difference that we have an additional term which goes to zero when $n \rightarrow \infty$.

The remaining terms A_{kn} , for $k = 9, \dots, 12$, can be handled similarly.

□

Appendix C. Proof of Theorem 3.

The proof follows the lines of that of Theorem 2. However the deterministic sequence a_n has to be replaced by $a_n = Y_{r_n:n}$. For this reason only three chosen auxiliary results are shown here.

Lemma 16. *Under the conditions in Theorem 3 we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\delta_i \nabla_{\theta} \log f_{\theta}(Y_i | \theta' X_i)_{\theta=\theta_0} + (1 - \delta_i) \nabla_{\theta} \log(1 - F_{\theta}(Y_i | \theta' X_i))_{\theta=\theta_0} \right] 1_{\{Y_i > Y_{r_n:n}, X_i \in A^{c_n}\}} = o_{\mathbb{P}}(1)$$

Proof.

Set

$$A_i = \delta_i \nabla_{\theta} \log f_{\theta}(Y_i | \theta' X_i)_{\theta=\theta_0} 1_{\{Y_i > Y_{r_n:n}, X_i \in A^{c_n}\}}.$$

We will show that, for every component $m = 1, \dots, d$ of the vector $A_i \in \mathbb{R}^d$, we have

$$E\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n A_i\right)_m^2 \rightarrow 0.$$

For this, let us consider

$$\begin{aligned} E\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n A_i\right)_m^2 &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n E((A_i A_j)_m) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n E\left(\delta_i \delta_j \frac{\partial}{\partial \theta_m} \log f_{\theta}(Y_i | \theta' X_i)_{\theta=\theta_0} \right. \\ &\quad \left. \frac{\partial}{\partial \theta_m} \log f_{\theta}(Y_j | \theta' X_j)_{\theta=\theta_0} 1_{\{H_n(Y_i) > \frac{r_n}{n}, H_n(Y_j) > \frac{r_n}{n}\}} 1_{\{X_i, X_j \in A^{c_n}\}}\right) \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j \neq i}^n E\left(\delta_i \delta_j \frac{\partial}{\partial \theta_m} \log f_{\theta}(Y_i | \theta' X_i)_{\theta=\theta_0} \frac{\partial}{\partial \theta_m} \log f_{\theta}(Y_j | \theta' X_j)_{\theta=\theta_0} \right. \\ &\quad \left. 1_{\{H_n(Y_i) > \frac{r_n}{n}, H_n(Y_j) > \frac{r_n}{n}\}} 1_{\{X_i, X_j \in A^{c_n}\}}\right) + O(n^{-1/3}(\log(n))^{2+\varepsilon_1}) \end{aligned}$$

Further, setting $H_n^{-(i)}(v) = \frac{1}{n} \sum_{k \neq i} 1_{\{Y_k \leq v\}}$, we have

$$\begin{aligned} 1_{\{H_n(Y_i) > \frac{r_n}{n}, H_n(Y_j) > \frac{r_n}{n}\}} &= 1_{\{H_n^{-(i)}(Y_i) > \frac{r_n-1}{n}, H_n^{-(j)}(Y_j) > \frac{r_n-1}{n}\}} = 1_{\{H_n^{-(i,j)}(Y_i) > \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_j) > \frac{r_n-1}{n}\}} \\ &\quad + 1_{\{H_n^{-(i,j)}(Y_i) > \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_j) = \frac{r_n-1}{n}, Y_i < Y_j\}} \\ &\quad + 1_{\{H_n^{-(i,j)}(Y_j) > \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_i) = \frac{r_n-1}{n}, Y_j \leq Y_i\}} \\ &= 1_{\{H_n^{-(i,j)}(Y_i) > \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_j) > \frac{r_n-1}{n}\}} \end{aligned}$$

Additionally, using Lemma 3 and since $1_{\{X_i \in A^{c_n}\}} = 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}$ is known given $\theta'_0 X_i$, we have

$$E\left(\delta_i \nabla_{\theta} \log f_{\theta}(Y_i | \theta' X_i)_{\theta=\theta_0} 1_{\{H_n^{-(i,j)}(Y_i) > \frac{r_n-1}{n}\}} | C_i, Z_i, \theta'_0 X_i, Y_s \text{ for } s \neq i\right) = 0.$$

This completes the proof. \(\square\)

Next we will deal with a version of Lemma 7, but for the choice $a_n = Y_{r_n:n}$.

Lemma 17. *Under the conditions in Theorem 3 we have*

$$B_{1n} = \frac{1}{n} \sum_{j=2}^n \bar{H}_{1j} + \frac{1}{n} \sum_{k=3}^n \bar{N}_{1k} + o_{\mathbb{P}}(n^{-1/2}),$$

where

$$\begin{aligned} \bar{H}_{1j} = & -\frac{\delta_j 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j-)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j)(1 - G_{\theta_0}(Y_j - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta_0 X_j} \\ & + E \left(\frac{\delta_j 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j-)} \frac{d}{dt} [(E(X_i | \theta_0 X_i = t) - X_j)(1 - G_{\theta_0}(Y_j - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta_0 X_j} | Y_s, s \neq j \right), \end{aligned}$$

and

$$\begin{aligned} \bar{N}_{1k} = & - \int \frac{d}{dt} [(E(X | \theta'_0 X = t) - u)(1 - G_{\theta_0}(v - |t)) 1_{\{f_{\theta'_0 X}(t) > c_n\}}]_{t=\theta'_0 \mathbf{u}} \frac{\psi(Y_k, \delta_k, v-)}{1 - G(v-)} \\ & 1_{\{H_n^{-(1,2,k)}(v) \leq \frac{r_{n-3}}{n}\}} f(\mathbf{u}, v) d\mathbf{u} dv. \end{aligned}$$

Proof.

As before

$$B_{1n} = b_{1n} + b_{2n} + b_{3n},$$

where

$$\begin{aligned} b_{1n} &= \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{1}{h_1^2 h_2 n} \sum_{j \neq i} \frac{\delta_j 1_{\{Y_j \leq Y_{r_n:n}\}}}{1 - G(Y_j-)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\ & \quad \left. - \frac{1}{h_1^2 h_2} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{v \leq Y_{r_n:n}\}}}{1 - G(v-)} H_1(d\mathbf{u}, dv) \right] 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq Y_{r_n:n}\}} \\ b_{2n} &= \frac{1}{h_1^2 h_2 n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{G_n(Y_j-) - G(Y_j-)}{(1 - G(Y_j-))^2} 1_{\{Y_j \leq Y_{r_n:n}\}} \\ & \quad 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq Y_{r_n:n}\}} \\ b_{3n} &= \frac{1}{h_1^2 h_2 n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i \delta_j (X_i - X_j)}{f_{\theta_0}(\theta'_0 X_i, Y_i)} K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \frac{(G_n(Y_j-) - G(Y_j-))^2}{(1 - G(Y_j-))^2 (1 - G_n(Y_j-))} \\ & \quad 1_{\{Y_j \leq Y_{r_n:n}\}} 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq Y_{r_n:n}\}}. \end{aligned}$$

We have

$$\sqrt{n} b_{1n} := \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} H_{ij},$$

where

$$\begin{aligned}
H_{ij} &= H(\theta'_0 X_i, Y_i, \delta_i, \theta'_0 X_j, Y_j, \delta_j) \\
&= \frac{1}{h_1^2 h_2} \frac{\delta_i}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{\delta_j}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) 1_{\{Y_j \leq Y_{r_n:n}\}} \right. \\
&\quad \left. - \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{v \leq Y_{r_n:n}\}}}{1 - G(v-)} H_1(d\mathbf{u}, dv) \right] \\
&\quad 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n, Y_i \leq Y_{r_n:n}\}}.
\end{aligned}$$

Additionally, we define

$$\begin{aligned}
H_{ij}^* &= \frac{1}{h_1^2 h_2} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{\delta_j 1_{\{H_n^{-(i,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\
&\quad \left. - \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}\}}}{1 - G(v-)} H_1(d\mathbf{u}, dv) \right] 1_{\{H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}}.
\end{aligned}$$

As first we show that

$$\frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} (H_{ij} - H_{ij}^*) = o_{\mathbb{P}}(1). \quad (\text{C.1})$$

For this we consider following indicators

$$\begin{aligned}
1_{\{Y_j \leq Y_{r_n:n}\}} 1_{\{Y_i \leq Y_{r_n:n}\}} &= 1_{\{Y_i \leq Y_j\}} 1_{\{H_n^{-(i,j)}(Y_j) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\
&\quad + 1_{\{Y_i > Y_j\}} 1_{\{H_n^{-(i,j)}(Y_j) \leq \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} \\
&= 1_{\{H_n^{-(i,j)}(Y_j) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}}
\end{aligned}$$

and

$$\begin{aligned}
1_{\{v \leq Y_{r_n:n}\}} 1_{\{Y_i \leq Y_{r_n:n}\}} &= 1_{\{Y_i \leq v\}} 1_{\{H_n^{-(i)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}} + 1_{\{Y_i > v\}} 1_{\{H_n^{-(i)}(v) \leq \frac{r_n}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\
&= 1_{\{H_n^{-(i)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
1_{\{H_n^{-(i)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}} &= 1_{\{Y_i \leq v\}} 1_{\{H_n^{-(i)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}} + 1_{\{Y_i > v\}} 1_{\{H_n^{-(i)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\
&= 1_{\{Y_j \leq Y_i \leq v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} + 1_{\{Y_i < Y_j \leq v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\
&\quad + 1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} + 1_{\{Y_j > Y_i > v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_n}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\
&\quad + 1_{\{Y_i \geq Y_j > v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_n}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} + 1_{\{Y_i > v \geq Y_j\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} \\
&= 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} + 1_{\{Y_i < Y_j \leq v\}} 1_{\{H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}\}} \\
&\quad + 1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} + 1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}\}} \\
&\quad + 1_{\{Y_j > Y_i > v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}\}} + 1_{\{Y_j > Y_i > v\}} 1_{\{H_n^{-(i,j)}(v) \in [\frac{r_{n-1}}{n}, \frac{r_n}{n}], H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\
&\quad + 1_{\{Y_i \geq Y_j > v\}} 1_{\{H_n^{-(i,j)}(v) \in [\frac{r_{n-1}}{n}, \frac{r_n}{n}], H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} + 1_{\{Y_i > v \geq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}}
\end{aligned}$$

Observe that

$$1_{\{Y_i < Y_j \leq v\}} 1_{\{H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}\}} = 0$$

$$1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}\}} = 0$$

$$1_{\{Y_j > Y_i > v\}} 1_{\{H_n^{-(i,j)}(v) \in [\frac{r_{n-1}}{n}, \frac{r_n}{n}], H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} = 0$$

$$1_{\{Y_i \geq Y_j > v\}} 1_{\{H_n^{-(i,j)}(v) \in [\frac{r_{n-1}}{n}, \frac{r_n}{n}], H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} = 0$$

and

$$1_{\{Y_i > v \geq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} = 0.$$

Hence

$$\begin{aligned} 1_{\{H_n^{-(i)}(v) \leq \frac{r_{n-1}}{n}, H_n^{-(i)}(Y_i) \leq \frac{r_{n-1}}{n}\}} &= 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} \\ &\quad + 1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \\ &\quad + 1_{\{Y_j > Y_i > v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}\}} \end{aligned}$$

Finally, we obtain

$$H_{ij} = H_{ij}^* + T_{ij},$$

where

$$\begin{aligned} T_{ij} &= -\frac{1}{h_1^2 h_2} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \\ &\quad \left(1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \right. \\ &\quad \left. + 1_{\{Y_j > Y_i > v\}} 1_{\{H_n^{-(i,j)}(v) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) = \frac{r_{n-1}}{n}\}} \right) \frac{1}{1 - G(v-)} H_1(d\mathbf{u}, dv). \end{aligned}$$

Hence, the quantity in (C.1) equals

$$\frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} T_{ij} = T_{1n} + O_{\mathbb{P}} \left(\frac{n^{1/2}}{h_1^2 h_2} \frac{n - r_n}{n^2} \right),$$

where

$$\begin{aligned} T_{1n} &= \frac{1}{h_1^2 h_2} \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \\ &\quad 1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_{n-1}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-1}}{n}\}} \frac{1}{1 - G(v-)} H_1(d\mathbf{u}, dv) \end{aligned}$$

and, since $n - r_n \approx n^{2/3} (\log(n))^{2+\varepsilon_1}$

$$O_{\mathbb{P}}\left(\frac{n^{1/2}}{h_1^2 h_2} \frac{n - r_n}{n^2}\right) = O_{\mathbb{P}}\left(\frac{(\log(n))^{2+\varepsilon_1}}{n^{5/6} h_1^2 h_2}\right) = o_{\mathbb{P}}(1).$$

Further, we can rewrite the T_{1n} as follow

$$\begin{aligned} T_{1n} &= \frac{1}{h_1^2 h_2} \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \int (X_i - E(X|\theta'_0 X = u)) K'\left(\frac{\theta'_0 X_i - u}{h_1}\right) K\left(\frac{Y_i - v}{h_2}\right) \\ &\quad 1_{\{Y_i \leq v \leq Y_j\}} 1_{\{H_n^{-(i,j)}(v) = \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_n-1}{n}\}} f_{\theta_0}(u, v) du dv \end{aligned}$$

Moreover, by change of variables and Taylor expansion, we obtain

$$\begin{aligned} T_{1n} &= \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \int \frac{d}{dt} [(X_i - E(X|\theta'_0 X = t)) f_{\theta_0}(t, Y_i)]_{t=\theta'_0 X_i} K(z_2) \\ &\quad 1_{\{Y_i \leq Y_i + h_2 z_2 \leq Y_j\}} 1_{\{H_n^{-(i,j)}(Y_i + h_2 z_2) = \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_n-1}{n}\}} dz_2 \\ &\quad + O_{\mathbb{P}}(h_1^2 n^{1/2} (\log(n))^{2+\varepsilon_1} n^{-1/3}) + O_{\mathbb{P}}(h_2 n^{1/2} (\log(n))^{2+\varepsilon_1} n^{-1/3}). \end{aligned}$$

The last two parts are $o_{\mathbb{P}}(1)$. As to the first one, note that

$$\begin{aligned} \{H_n^{-(i,j)}(Y_i + h_2 z_2) = \frac{r_n-1}{n}\} &\cap \{Y_i \leq Y_i + h_2 z_2 \leq Y_j\} \cap \{H_n^{-(i,j)}(Y_i) \leq \frac{r_n-1}{n}\} \\ &\in \{H_n(Y_i + h_2 z_2) = \frac{r_n}{n}\} \cap \{H_n(Y_i) \leq \frac{r_n}{n}\} = \{Y_i \in [Y_{r_n:n} + h_2 z_2, Y_{r_n:n}]\} \end{aligned}$$

Further, if $\text{supp}(K) \in [-a, a]$ and since $Y_j \geq Y_{r_n:n}$, we can show that

$$\begin{aligned} &\frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \int \frac{d}{dt} [(X_i - E(X|\theta'_0 X = t)) f_{\theta_0}(t, Y_i)]_{t=\theta'_0 X_i} K(z_2) \\ &1_{\{Y_i \leq Y_i + h_2 z_2 \leq Y_j\}} 1_{\{H_n^{-(i,j)}(Y_i + h_2 z_2) = \frac{r_n-1}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_n-1}{n}\}} dz_2 = O_{\mathbb{P}}(h_2 n^{1/2} (\log(n))^{2+\varepsilon_1} n^{-1/3}) \\ &= o_{\mathbb{P}}(1). \end{aligned}$$

Next, set

$$\bar{H}_j^* = E(H_{ij}^* | Y_j, \delta_j, X_j, Y_s \text{ for } s \neq i).$$

Hence

$$\sqrt{n} b_{1n} := \frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i} (H_{ij}^* - \bar{H}_j^*) + \frac{1}{\sqrt{n}} \sum_{j=2}^n \bar{H}_j^* + o_{\mathbb{P}}(1).$$

As for the second part, we have

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=2}^n \bar{H}_j^* &= \frac{1}{h_1^2 h_2} \frac{1}{\sqrt{n}} \sum_{j=2}^n \left[\int \frac{(1 - G(y - |x|)) 1_{\{f_{\theta'_0 X}(x) > c_n\}}}{f_{\theta_0}(x, y)} \left[\frac{\delta_j}{1 - G(Y_j -)} (E(X | \theta'_0 X = x) - X_j) \right. \right. \\
&\quad \left. K' \left(\frac{x - \theta'_0 X_j}{h_1} \right) K \left(\frac{y - Y_j}{h_2} \right) 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}} \right. \\
&\quad \left. - \int (E(X | \theta'_0 X = x) - \mathbf{u}) K' \left(\frac{x - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{y - v}{h_2} \right) \frac{1_{\{H_n^{-(1,j)}(v) \leq \frac{r_{n-2}}{n}\}}}{1 - G(v -)} H_1(d\mathbf{u}, dv) \right] \\
&\quad \left. 1_{\{H_n^{-(1,j)}(y) \leq \frac{r_{n-2}}{n}\}} f_{\theta_0}(x, y) dx dy \right] \\
&= \frac{1}{\sqrt{n}} \sum_{j=2}^n \left[\frac{1}{h_1} \int (1 - G(Y_j + h_2 z_2 - |\theta'_0 X_j + h_1 z_1|)) 1_{\{f_{\theta'_0 X}(\theta'_0 X_j + h_1 z_1) > c_n\}} \frac{\delta_j 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j -)} \right. \\
&\quad (E(X | \theta'_0 X = \theta'_0 X_j + h_1 z_1) - X_j) K'(z_1) K(z_2) 1_{\{H_n^{-(1,j)}(Y_j + h_2 z_2) \leq \frac{r_{n-2}}{n}\}} dz_1 dz_2 \\
&\quad - \frac{1}{h_1} \int (1 - G(v + h_2 z_2 - |\theta'_0 u + h_1 z_1|)) 1_{\{f_{\theta'_0 X}(\theta'_0 u + h_1 z_1) > c_n\}} \int (E(X | \theta'_0 X = \theta'_0 u + h_1 z_1) - \mathbf{u}) \\
&\quad \left. K'(z_1) K(z_2) 1_{\{H_n^{-(1,j)}(v) \leq \frac{r_{n-2}}{n}\}} 1_{\{H_n^{-(1,j)}(v + h_2 z_2) \leq \frac{r_{n-2}}{n}\}} f(u, v) dz_1 dz_2 d\mathbf{u} dv \right] \\
&= \frac{1}{\sqrt{n}} \sum_{j=2}^n \left[\frac{d}{dt} \left(\bar{G}(Y_j - |t|) (E(X | \theta'_0 X = t) - X_j) \right) \right]_{t=\theta'_0 X_j} \frac{\delta_j 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j -)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}} \\
&\quad - \int \frac{d}{dt} \left(\bar{G}(v - |t|) (E(X | \theta'_0 X = t) - \mathbf{u}) \right) \right]_{t=\theta'_0 \mathbf{u}} 1_{\{f_{\theta'_0 X}(\theta'_0 u) > c_n\}} 1_{\{H_n^{-(1,j)}(v) \leq \frac{r_{n-2}}{n}\}} f(u, v) d\mathbf{u} dv \\
&\quad + \frac{1}{\sqrt{n}} \sum_{j=2}^n h_{1j}^*,
\end{aligned}$$

Set

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=2}^n \bar{H}_{1j}^* &= \frac{1}{\sqrt{n}} \sum_{j=2}^n \left[\frac{d}{dt} \left[\bar{G}(Y_j - |t|) (E(X | \theta'_0 X = t) - X_j) \right] \right]_{t=\theta'_0 X_j} \frac{\delta_j 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j -)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}} \\
&\quad - \int \frac{d}{dt} \left[\bar{G}(v - |t|) (E(X | \theta'_0 X = t) - \mathbf{u}) \right] \right]_{t=\theta'_0 \mathbf{u}} 1_{\{f_{\theta'_0 X}(\theta'_0 u) > c_n\}} 1_{\{H_n^{-(1,j)}(v) \leq \frac{r_{n-2}}{n}\}} f(u, v) d\mathbf{u} dv.
\end{aligned}$$

Then

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=2}^n h_{1j}^* &= \frac{1}{\sqrt{n}} \sum_{j=2}^n h_1^2 \left[\frac{d^3}{dt^3} \left[\bar{G}(Y_j - |t|) (E(X | \theta'_0 X = t) - X_j) \right] \right]_{t=\theta'_0 X_j} \frac{\delta_j 1_{\{H_n^{-(1,j)}(Y_j) \leq \frac{r_{n-2}}{n}\}}}{1 - G(Y_j -)} 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}} \\
&\quad - \int \frac{d^3}{dt^3} \left[\bar{G}(v - |t|) (E(X | \theta'_0 X = t) - \mathbf{u}) \right] \right]_{t=\theta'_0 \mathbf{u}} 1_{\{f_{\theta'_0 X}(\theta'_0 u) > c_n\}} 1_{\{H_n^{-(1,j)}(v) \leq \frac{r_{n-2}}{n}\}} f(u, v) d\mathbf{u} dv \\
&\quad + o_{\mathbb{P}}(1)
\end{aligned}$$

Moreover, by Lemma 18

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=2}^n h_{1j}^* &= \frac{1}{\sqrt{n}} \sum_{j=2}^n h_1^2 \left[\frac{d^3}{dt^3} \left[\bar{G}(Y_j - |t|)(E(X|\theta'_0 X = t) - X_j) \right]_{t=\theta'_0 X_j} 1_{\{f_{\theta'_0 X}(\theta'_0 X_j) > c_n\}} \frac{\delta_j 1_{\{Y_j \leq \tilde{a}_n\}}}{1 - G(Y_j -)} \right. \\
&\quad \left. - \int \frac{d^3}{dt^3} \left[\bar{G}(v - |t|)(E(X|\theta'_0 X = t) - \mathbf{u}) \right]_{t=\theta'_0 \mathbf{u}} 1_{\{f_{\theta'_0 X}(\theta'_0 \mathbf{u}) > c_n\}} 1_{\{v \leq \tilde{a}_n\}} f(u, v) d\mathbf{u} dv \right] \\
&\quad + O_{\mathbb{P}}(n^{1/2-\beta/2} h_1^2) + o_{\mathbb{P}}(1),
\end{aligned}$$

where $\tilde{a}_n = H^{-1}(\frac{r_n-2}{n})$. Finally, since the first part is a sum of i.i.d. random variables with variance going to zero as $n \rightarrow \infty$, we have

$$\frac{1}{\sqrt{n}} \sum_{j=2}^n h_{1j}^* = o_{\mathbb{P}}(1).$$

Next we show that

$$\frac{\sqrt{n}}{n^2} \sum_{i=1}^n \sum_{j \neq i}^n (H_{ij}^* - \bar{H}_j^*) = o_{\mathbb{P}}(1).$$

For this let us define

$$\begin{aligned}
H_{ij(kl)}^* &= \frac{1}{h_1^2 h_2} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{\delta_j 1_{\{H_n^{-(i,j,k,l)}(Y_j) \leq \frac{r_n-4}{n}\}}}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\
&\quad \left. - \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) \frac{1_{\{H_n^{-(i,j,k,l)}(v) \leq \frac{r_n-4}{n}\}}}{1 - G(v -)} H_1(d\mathbf{u}, dv) \right] 1_{\{H_n^{-(i,j,k,l)}(Y_i) \leq \frac{r_n-4}{n}\}}
\end{aligned}$$

and

$$\bar{H}_{j(kl)}^* = E(H_{ij(kl)}^* | Y_j, \delta_j, X_j, Y_s \text{ for } s \neq i).$$

We have

$$E[(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*) | Y_j, \delta_j, X_j, Y_s \text{ for } s \neq i] = E[(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*) | Y_i, \delta_i, X_i, Y_s \text{ for } s \neq j] = 0$$

and given Y_s for $s \neq i, j, k, l$, we have that $(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)$ is independent of $(H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)$. Hence, similarly as in the proof of Lemma 7, we can show that for every component $m = 1, \dots, d$

$$\frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i}^n \sum_{k=1}^n \sum_{l \neq k}^n E[(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m] = o(1). \quad (\text{C.2})$$

Therefore, it remains to prove that

$$D = \frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i}^n \sum_{k=1}^n \sum_{l \neq k}^n \left\{ E[(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m (H_{kl}^* - \bar{H}_l^*)_m] \right\} = o(1) \quad (\text{C.3})$$

The term in (C.3) can be handled in the following way

$$\begin{aligned}
D &= \frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i}^n \sum_{k=1}^n \sum_{l \neq k}^n \left\{ E[(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m (H_{kl}^* - \bar{H}_l^*)_m] \right\} \\
&= \frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i}^n \sum_{k=1}^n \sum_{l \neq k}^n \left\{ E[\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m \right. \\
&\quad \left. + (H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m \{(H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m - (H_{kl}^* - \bar{H}_l^*)_m\} \right. \\
&\quad \left. + \{(H_{ij}^* - \bar{H}_j^*)_m - (H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m\} \{(H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m - (H_{kl}^* - \bar{H}_l^*)_m\} \right\}
\end{aligned}$$

Let us consider the first part

$$\frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i}^n \sum_{k=1}^n \sum_{l \neq k}^n E[\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m]$$

First observe that, componentwise,

$$E[(H_{ij}^* - \bar{H}_j^*)|Y_j, \delta_j, X_j, Y_s \text{ for } s \neq i] = E[(H_{ij}^* - \bar{H}_j^*)|Y_i, \delta_i, X_i, Y_s \text{ for } s \neq j] = 0$$

- If $i = k \neq j = l$, then

$$\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} = 0$$

- If $i \neq k \neq j \neq l$, then

$$\begin{aligned}
&E[\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m] = E[E(\dots | X_k, \delta_k, X_l, \delta_l, Y_s \text{ for } s \neq i, j)] \\
&= E[(H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m E(\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} | X_k, \delta_k, X_l, \delta_l, Y_s \text{ for } s \neq i, j)]
\end{aligned}$$

Moreover,

$$\begin{aligned}
&E(\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} | X_k, \delta_k, X_l, \delta_l, Y_s \text{ for } s \neq i, j) \\
&= E(E(\{(H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} | X_s, \delta_s, Y_s \text{ for } s \neq i) | X_k, \delta_k, X_l, \delta_l, Y_s \text{ for } s \neq i, j) = 0
\end{aligned}$$

- If $i = k \neq j \neq l$, we have similarly

$$\begin{aligned}
&E[\{(H_{ij(il)}^* - \bar{H}_{j(il)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} (H_{il(ij)}^* - \bar{H}_{l(ij)}^*)_m] = E[E(\dots | X_i, \delta_i, X_l, \delta_l, Y_s \text{ for } s \neq j)] \\
&= E[(H_{il(ij)}^* - \bar{H}_{l(ij)}^*)_m E(\{(H_{ij(il)}^* - \bar{H}_{j(il)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} | X_i, \delta_i, X_l, \delta_l, Y_s \text{ for } s \neq j)]
\end{aligned}$$

and

$$\begin{aligned}
&E(\{(H_{ij(il)}^* - \bar{H}_{j(il)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} | X_i, \delta_i, X_l, \delta_l, Y_s \text{ for } s \neq j) \\
&= E(E(\{(H_{ij(il)}^* - \bar{H}_{j(il)}^*)_m - (H_{ij}^* - \bar{H}_j^*)_m\} | X_s, \delta_s, Y_s \text{ for } s \neq j) | X_i, \delta_i, X_l, \delta_l, Y_s \text{ for } s \neq j) = 0
\end{aligned}$$

- If $i \neq k \neq j = l$, we do the same as above only under $s \neq i$.

Since the second part can be proven similarly to the first one, we can conclude

$$D = \frac{1}{n^3} \sum_{i=1}^n \sum_{j \neq i} \sum_{k=1}^n \sum_{l \neq k} E \left[(H_{ij}^* - \bar{H}_j^*)_m - (H_{ij(kl)}^* - \bar{H}_{j(kl)}^*)_m \right] \left\{ (H_{kl(ij)}^* - \bar{H}_{l(ij)}^*)_m - (H_{kl}^* - \bar{H}_l^*)_m \right\}$$

Now we can prove that

$$\begin{aligned} & 1_{\{H_n^{-(i,j)}(Y_j) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j)}(Y_i) \leq \frac{r_{n-2}}{n}\}} = 1_{\{H_n^{-(i,j,k,l)}(Y_j) \leq \frac{r_{n-4}}{n}, H_n^{-(i,j,k,l)}(Y_i) \leq \frac{r_{n-4}}{n}\}} \\ & + 1_{\{Y_k, Y_l > Y_i\}} 1_{\{Y_k, Y_l > Y_j\}} \left(1_{\{H_n^{-(i,j,k,l)}(Y_j) \leq \frac{r_{n-4}}{n}, H_n^{-(i,j,k,l)}(Y_i) \in [\frac{r_{n-3}}{n}, \frac{r_{n-2}}{n}]\}} \right. \\ & \left. + 1_{\{H_n^{-(i,j,k,l)}(Y_i) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j,k,l)}(Y_j) \in [\frac{r_{n-3}}{n}, \frac{r_{n-2}}{n}]\}} \right) \end{aligned}$$

Hence

$$H_{ij}^* = H_{ij(kl)}^* + T_{ijkl},$$

where

$$\begin{aligned} T_{ijkl} &= \frac{1}{h_1^2 h_2} \frac{\delta_i 1_{\{f_{\theta'_0 X}(\theta'_0 X_i) > c_n\}}}{f_{\theta_0}(\theta'_0 X_i, Y_i)} \left[\frac{\delta_j}{1 - G(Y_j -)} (X_i - X_j) K' \left(\frac{\theta'_0 X_i - \theta'_0 X_j}{h_1} \right) K \left(\frac{Y_i - Y_j}{h_2} \right) \right. \\ & 1_{\{Y_k, Y_l > Y_i\}} 1_{\{Y_k, Y_l > Y_j\}} \left(1_{\{H_n^{-(i,j,k,l)}(Y_j) \leq \frac{r_{n-4}}{n}, H_n^{-(i,j,k,l)}(Y_i) \in [\frac{r_{n-3}}{n}, \frac{r_{n-2}}{n}]\}} \right. \\ & \left. + 1_{\{H_n^{-(i,j,k,l)}(Y_i) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j,k,l)}(Y_j) \in [\frac{r_{n-3}}{n}, \frac{r_{n-2}}{n}]\}} \right) \\ & - \int (X_i - \mathbf{u}) K' \left(\frac{\theta'_0 X_i - \theta'_0 \mathbf{u}}{h_1} \right) K \left(\frac{Y_i - v}{h_2} \right) 1_{\{Y_k, Y_l > Y_i\}} 1_{\{Y_k, Y_l > v\}} \\ & \left(1_{\{H_n^{-(i,j,k,l)}(v) \leq \frac{r_{n-4}}{n}, H_n^{-(i,j,k,l)}(Y_i) \in [\frac{r_{n-3}}{n}, \frac{r_{n-2}}{n}]\}} + 1_{\{H_n^{-(i,j,k,l)}(Y_i) \leq \frac{r_{n-2}}{n}, H_n^{-(i,j,k,l)}(v) \in [\frac{r_{n-3}}{n}, \frac{r_{n-2}}{n}]\}} \right) \\ & \left. \frac{1}{1 - G(v-)} H_1(d\mathbf{u}, dv) \right] \end{aligned}$$

Since, $(T_{ijkl})_m (T_{kl ij})_m = 0$ for every $m = 1, \dots, d$, the proof of (C.3) is completed.

Finally, the term b_{2n} can be handled, by using Lemma 6, similarly to b_{1n} . The proof for b_{3n} is an immediate consequence of Lemma 6. Hence the proof is finished. $\square$

Lemma 18. *We have*

$$|1_{\{H_n(Y_i) \leq \frac{r_n}{n}\}} - 1_{\{H(Y_i) \leq \frac{r_n}{n}\}}| \leq 1_{A^{\delta_n}(Y_i)} + 1_{(\delta_n, \infty)}(Z_n), \quad (\text{C.4})$$

where

$$A^{\delta_n}(y) = \{y : |H(y) - \frac{r_n}{n}| \leq \delta_n\}$$

and

$$Z_n = \max_{i=1, \dots, n} |H_n(Y_i) - H(Y_i)|.$$

Set, $\delta_n = n^{-\beta/2}$ and $\beta < 1$. Hence

$$1_{A^{\delta_n}}(Y_i) = O_{\mathbb{P}}(n^{-\beta/2})$$

and

$$1_{(\delta_n, \infty)}(Z_n) = o_{\mathbb{P}}(n^{-\alpha}). \tag{C.5}$$

Proof.

The proof is based on Lemma B.2 in [6]. Observe that, $Z_n \leq \sup_{t \in \mathbb{R}} |H_n(t) - H(t)|$ and, using the Dvoretzky-Kiefer-Wolfowitz inequality (see [7]) we have

$$n^\alpha E \left(1_{(\delta_n, \infty)}(Z_n) \right) \leq n^\alpha \mathbb{P}(D_n \geq \delta_n) \leq cn^\alpha \exp(-2n\delta_n^2) \rightarrow 0.$$

□