Electronic Journal of Statistics
Vol. 13 (2019) 5044–5087
ISSN: 1935-7524
https://doi.org/10.1214/19-EJS1645

Estimation of a bivariate conditional copula when a variable is subject to random right censoring

Taoufik Bouezmarni

Université de Sherbrooke
Centre interuniversitaire de recherche en economie quantitative (CIREQ)
e-mail: taoufik.bouezmarni@usherbrooke.ca

Félix Camirand Lemyre

Université de Sherbrooke
Centre de recherche du centre hospitalier universitaire de Sherbrooke
e-mail: felix.camirand.lemyre@usherbrooke.ca

Anouar El Ghouch

ISBA, Université catholique de Louvain, Belgique
e-mail: anouar.elghouch@uclouvain.be

Abstract: This paper concerns the dependence structure of a random pair (Y_1, Y_2) conditionally upon a covariate X in the case where the variable Y_1 is subject to random right censoring. The dependence structure is described by a conditional copula for which we propose a nonparametric estimator. We establish the asymptotic properties of the proposed estimator and investigate its finite sample behavior in a numerical study. The methodology is illustrated through a real data example featuring patients with advanced lung cancer.

MSC 2010 subject classifications: Primary 62G99; secondary 62N01.
Keywords and phrases: Conditional copula, Kendall's tau, kernel estimation, Spearman's rho, survival analysis, weak convergence.

Received September 2018.

1. Introduction

Copulas have become a popular tool to model dependence structures. Recently, many works in this field have been concerned with capturing the influence of a covariate $X \in \mathbb{R}$ on the dependence structure of a vector of interest $(Y_1, Y_2) \in \mathbb{R}^2$. An example is given in [8], where a copula function is used to illustrate how the relationship between the life expectancy of men (Y_1) and women (Y_2) varies with the gross domestic product (X). To describe this copula function, consider the conditional joint distribution of (Y_1, Y_2) given $X = x$, for a real number x, given by $F_x(y_1, y_2) = \mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2 | X = x)$. If the conditional marginal distributions of Y_1 and Y_2 given $X = x$, denoted by F_{1x} and F_{2x}, respectively, are continuous, then Sklar's theorem ensures that there exists a unique copula $\mathcal{C}_x : [0,1]^2 \to [0,1]$ such that $F_x(y_1, y_2) = \mathcal{C}_x\{F_{1x}(y_1), F_{2x}(y_2)\}$. Conversely, the copula associated to the bivariate conditional distribution F_x is given by

$$\mathcal{C}_x(u_1, u_2) = F_x\left\{F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2)\right\}. \tag{1.1}$$

The bivariate function $\mathcal{C}_x$ is called the *conditional copula* and contains all the dependence features of (Y_1, Y_2) given a fixed value x taken by the covariate.

Since the pioneering work of [13], the topic of modeling and estimating conditional copula models has recently gained momentum. For example, models specifying a relation between the covariate and a parametric copula were studied in [10] and [13]. A semiparametric estimation procedure for the conditional copula was proposed in [2] (resp. [1]) when the conditional marginal distributions are assumed to be known (resp. unknown). A nonparametric approach has been investigated in [22] and [8], and a bootstrap method suitable for this estimation procedure was developed in [12]. However, all of the previously-mentioned estimation strategies rely on the availability of a sample generated from the random vector (Y_1, Y_2, X) and lead to unsatisfactory results when the data are incomplete. Amongst others, the right censoring scheme is a source of incompleteness that frequently appears in medical studies and clinical trials.

For the censored unconditional case (i.e. censored data without covariates), the parametric and semiparametric estimation of the copula function has been studied in [14]. Nonparametric copula estimation procedures have been proposed in [9] under different censoring scenarios. In the aforementioned article, a goodness-of-fit procedure for copula models suitable for right censored data is also investigated.

In this work, we propose a nonparametric methodology to estimate the conditional copula when the random variable Y_1 is subject to random right censoring. To be more specific, we assume that the available data arise as the realizations of the random vector (T_1, Y_2, X, δ_1), with $T_1 = \min(Y_1, C_1)$ and $\delta_1 = \mathbb{I}(Y_1 \leq C_1)$, where C_1 is a censoring variable. The properties of the proposed estimator are investigated for large and finite sample sizes.

This paper is organized as follows. In Section 2, an estimator for the conditional copula in the presence of censoring is presented. This estimator is based on a new nonparametric estimator of the joint conditional distribution. In Section 3, we investigate the asymptotic properties of these estimators by providing an asymptotic representation of the conditional distribution estimator, and by identifying the weak limit of a properly re-scaled version of these estimators. A simulation study showing the performance of the conditional copula estimation procedure is presented in Section 4. In Section 5, we apply this methodology to a lung cancer dataset to illustrate the influence of age on the relationship between survival time and weight loss. All the required assumptions and conditions for the theoretical validity of the results presented in Section 3 are provided in Appendix A. The proofs are given in Appendix B, Appendix C and Appendix D. Additional technical results are provided in Appendix E and Appendix F.

2. An inverse-conditional-probability-of-censoring estimator of $\mathcal{C}_x$

As previously mentioned, the estimation of $\mathcal{C}_x$ from independent and completely observed random vectors $(Y_{11}, Y_{21}, X_1), \ldots, (Y_{1n}, Y_{2n}, X_n)$ has been investigated by [22] and [8]. Specifically, consider the estimator of the joint conditional distribution F_x given by

$$\widehat{F}_x(y_1, y_2) = \sum_{i=1}^{n} w_{ni}(x,h)\, \mathbb{I}\left(Y_{1i} \le y_1, Y_{2i} \le y_2\right),$$

where $w_{n1}(x,h), \ldots, w_{nn}(x,h)$ are kernel-based weight functions that smooth the covariate space and $h = h_n$ is a parameter called bandwidth that typically depends on the sample size. Popular choices for these weight functions include Nadaraya-Watson and local-linear weights given respectively by

$$w_{ni}^{\mathrm{NW}}(x,h) = \frac{K\left(\frac{X_i - x}{h}\right)}{S_{n,0}(x,h)} \tag{2.1}$$

and

$$w_{ni}^{\mathrm{LL}}(x,h) = \frac{K\left(\frac{X_i - x}{h}\right)\left\{S_{n,2}(x,h) - \left(\frac{X_i - x}{h}\right) S_{n,1}(x,h)\right\}}{S_{n,0}(x,h) S_{n,2}(x,h) - S_{n,1}^2(x,h)}, \tag{2.2}$$

where K is a symmetric and continuously differentiable kernel density function on $[-1,1]$ and $S_{n,j}(x,h) = \sum_{i=1}^{n} \{(X_i - x)/h\}^j K\{(X_i - x)/h\}$, for $j \in \{0,1,2\}$.

The corresponding conditional empirical marginal distributions are defined by

$$\widehat{F}_{1x}(y_1) = \sum_{i=1}^{n} w_{ni}(x,h)\, \mathbb{I}\left(Y_{1i} \le y_1\right) \quad \text{and} \quad \widehat{F}_{2x}(y_2) = \sum_{i=1}^{n} w_{ni}(x,h)\, \mathbb{I}\left(Y_{2i} \le y_2\right),$$

see [15].

In view of (1.1), a plug-in estimator of $\mathcal{C}_x$ proposed by [22] and [8] is defined by

$$\widehat{\mathcal{C}}_x(u_1, u_2) \;=\; \sum_{i=1}^{n} w_{ni}(x,h)\, \mathbb{I}\left\{Y_{1i} \le \widehat{F}_{1x}^{-1}(u_1), Y_{2i} \le \widehat{F}_{2x}^{-1}(u_2)\right\}, \tag{2.3}$$

where for $j = 1,2$, $\widehat{F}_{jx}^{-1}(u) = \inf\{y \in \mathbb{R} : \widehat{F}_{jx}(y) \ge u\}$ is the left-continuous generalized inverse of $\widehat{F}_{jx}$.

To construct our conditional copula estimator, we need first to estimate the conditional distribution function F_x. In the unconditional context (i.e., without a covariate), the nonparametric estimation of the bivariate distribution of (Y_1, Y_2) in the presence of censoring has been studied by many authors, see for example [6], [3] and [4]. However, to the best of our knowledge, the nonparametric estimation of F_x has never been studied and hence this is an original contribution of the present paper.

For the rest of the paper, a fixed design is assumed, where the outcome of the random vector (T_1, Y_2, δ_1, X) is observed at $X \in \{x_1, \ldots, x_n\}$. To be more specific, we consider independent random vectors $(T_{11}, Y_{21}, \delta_{11}, x_1)$, …, $(T_{1n}, Y_{2n}, \delta_{1n}, x_n)$, where $(T_{1i}, Y_{2i}, \delta_{1i}, x_i)$ is distributed as $(T_1, Y_2, \delta_1, X) \mid X = x_i$. Also, the censoring variable C_1 is assumed to be independent of (Y_1, Y_2) given the value taken by X.

To build our estimator for F_x, we use an inverse-probability-of-censoring weighting approach (see chapter 3.3 in [16]). Specifically, to compensate for the presence of censoring, each uncensored observation receives an extra weight equal to its inverse probability of failure. This idea is motivated by the fact that

$$F_x(y_1, y_2) = \mathrm{E}\Big\{\mathbb{I}(T_1 \le y_1, Y_2 \le y_2)\frac{\delta_1}{1 - G_x(T_1)}|X = x\Big\},$$

where $G_x(t) = \mathbb{P}(C_1 \le t \mid X = x)$ denotes the conditional distribution of the censoring variable given $X = x$. In the case where G_x is known, one could estimate F_x by

$$\widehat{F}_x^{G_x}(y_1, y_2) = \sum_{i=1}^{n} \mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2)\frac{w_{ni}(x, h)\delta_{1i}}{1 - G_x(T_{1i})}. \tag{2.4}$$

In our case G_x is unknown, so we replace it by the conditional Kaplan-Meier estimator for the censoring variable C_1:

$$\widehat{G}_x(t) = 1 - \prod_{T_{1(i)} \le t} \Big\{1 - \frac{w_{n[i]}(x, g)}{1 - \sum_{k=1}^{i-1} w_{n[k]}(x, g)}\Big\}^{1-\delta_{1[i]}}, \tag{2.5}$$

where $T_{1(1)} \le \ldots \le T_{1(n)}$ are the ordered $T_{1i}'s$, and $\delta_{1[i]}$ and $w_{n[i]}(x, h)$ are, respectively, the corresponding δ_{1i} and $w_{ni}(x, h)$. Here, $g = g_n$ is an auxiliary bandwidth parameter that may differ from h. The resulting estimator for the conditional distribution function F_x is given by

$$\widehat{F}_x^{\widehat{G}_x}(y_1, y_2) = \sum_{i=1}^{n} \mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2)\frac{w_{ni}(x, h)\delta_{1i}}{1 - \widehat{G}_x(T_{1i})}. \tag{2.6}$$

Note that when no censoring occurs, all the δ_{1i}'s are equal to 1, which implies $\widehat{G}_x = 0$ (see Equation (2.5)) and therefore $\widehat{F}_x^{\widehat{G}_x}$ reduces to $\widehat{F}_x$.

To estimate the conditional marginal distribution of Y_1 given x and taking into account the right censoring of Y_1, we consider the estimator given by

$$\widehat{F}_{1x}^{\widehat{G}_x}(y_1) = \sum_{i=1}^{n} \mathbb{I}(T_{1i} \le y_1)\frac{w_{ni}(x, h)\delta_{1i}}{1 - \widehat{G}_x(T_{1i})}.$$

When $g = h$, we show in Appendix F.1 that $\widehat{F}_{1x}^{\widehat{G}_x}$ coincides with the conditional Kaplan-Meier estimator for the survival time Y_1 introduced by [5].

Because Y_2 is completely observed, for more efficiency, we estimate F_{2x} with

$$F_{2x}(y_2) = \sum_{i=1}^{n} \mathbb{I}\left(Y_{2i} \leq y_2\right) w_{ni}(x,h).$$

So, from Equation (1.1), a plug-in estimator for the conditional copula is given by

$$\begin{aligned}\widehat{\mathcal{C}}_x^{\widehat{G}_x}(u,v) &= \widehat{F}_x^{\widehat{G}_x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v)\right\} \\ &= \sum_{i=1}^{n} \mathbb{I}\left\{T_{1i} \leq \left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), Y_{2i} \leq \widehat{F}_{2x}^{-1}(v)\right\} \frac{w_{ni}(x,h)\delta_{1i}}{1-\widehat{G}_x(T_{1i})}. \qquad (2.7)\end{aligned}$$

We can see that $\widehat{F}_x^{\widehat{G}_x} = \widehat{F}_x$ and that $\widehat{G}_x = 0$ if all the survival times are observed, hence the two estimators $\widehat{\mathcal{C}}_x^{\widehat{G}_x}$ and $\widehat{\mathcal{C}}_x$ are equal in this case. Moreover, upon setting all the weight functions $w_{ni}(x,\cdot)$ equal to n^{-1}, we retrieve an estimator that is very similar to the one proposed in [9] to estimate the unconditional copula.

3. Main theoretical results

The aim of this section is to investigate the large sample behavior of the processes

$$\mathbb{F}_x^{\widehat{G}_x} = \sqrt{nh}(\widehat{F}_x^{\widehat{G}_x} - F_x) \quad \text{and} \quad \mathbb{B}_x^{\widehat{G}_x} = \sqrt{nh}(\widehat{\mathcal{C}}_x^{\widehat{G}_x} - \mathcal{C}_x).$$

Because we estimate the conditional distribution G_x, the asymptotic results of $\mathbb{F}_x^{\widehat{G}_x}$ and $\mathbb{B}_x^{\widehat{G}_x}$ become more challenging to prove. First, we show in Theorem 3.1 that, up to a term converging in probability to zero, $\mathbb{F}_x^{\widehat{G}_x}$ can be expressed as a weighted sum of independent random functions. Note that, in the complete data case, the estimator $\widehat{F}_x$ is, by definition, a sum of independent random variables. Second, we obtain a weak convergence result for $\mathbb{F}_x^{\widehat{G}_x}$ in Corollary 3.3. Finally, the weak limit of $\mathbb{F}_x^{\widehat{G}_x}$ allows us to establish, in Proposition 3.5, the weak convergence of $\mathbb{B}_x^{\widehat{G}_x}$.

3.1. Asymptotic representation of $\mathbb{F}_x^{\widehat{G}_x}$

For any distribution function L, let τ_L be the right endpoint of its support, i.e. $\inf\{t : L(t) = 1\}$, and write $\overline{\tau}_x = \min\{\tau_{F_{1x}}, \tau_{G_x}\}$. Due to the censoring, we cannot hope to infer on the conditional distribution beyond $\overline{\tau}_x$. Before establishing the asymptotic behavior of $\mathbb{F}_x^{\widehat{G}_x}$ over any closed subset included in $[0, \overline{\tau}_x) \times \mathbb{R}$, we need the following notations: $H_{1x}(y_1) = \mathbb{P}(T_1 \leq y_1 \mid X = x)$, $H_x(y_1, y_2) = \mathbb{P}(T_1 \leq y_1, Y_2 \leq y_2 \mid X = x)$, $H_x^u(y_1, y_2) = \mathbb{P}(T_1 \leq y_1, Y_2 \leq y_2, \delta_1 = 1 \mid X = x)$, $H_{1x}^u(y_1) = \mathbb{P}(T_1 \leq y_1, \delta_1 = 1 \mid X = x)$ and $H_{1x}^c(y_1) = \mathbb{P}(T_1 \leq y_1, \delta_1 = 0 \mid X = x)$. Also, for any map $(z, y_1, y_2) \mapsto L_z(y_1, y_2)$, we let $\dot{L}_z$ and $\ddot{L}_z$ denote, respectively, its first and second partial derivative with respect to z, and we set $L_z^{[j]}(y_1, y_2) = (\partial/\partial y_j) L_z(y_1, y_2)$, for $j = 1, 2$.

For the asymptotic representation of $\mathbb{F}_x^{\widehat{G}_x}$, we introduce

$$
\begin{aligned}
\mathcal{J}_{ix}^{(1)}(y_1, y_2) =& \frac{\mathbb{I}\left(T_{1i} \le y_1, Y_{2i} \le y_2, \delta_{1i} = 1\right)}{1 - G_x(T_{1i})} - F_x(y_1, y_2), \text{ and} \\
\mathcal{J}_{ix}^{(2)}(y_1, y_2) =& \int_0^{y_1} \frac{\mathbb{I}\left(T_{1i} \le v\right) - H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \Delta_{F_x}(y_1, y_2, v) \, \mathrm{d}H_{1x}^c(v) \\
&+ \int_0^{y_1} \frac{\mathbb{I}\left(T_{1i} \le v, \delta_{1i} = 0\right) - H_{1x}^c(v)}{1 - H_{1x}(v)} \\
&\qquad \times \left\{ F_x^{[1]}(y_1, y_2) - \frac{\Delta_{F_x}(y_1, y_2, v) H_x^{u[1]}(v, y_2)}{1 - H_x(v)} \right\} \mathrm{d}v,
\end{aligned}
$$

where $\Delta_{F_x}(y_1, y_2, v) = F_x(y_1, y_2) - F_x(v, y_2)$.

In the complete data case, we can see that $\mathcal{J}_{ix}^{(1)}(y_1, y_2) = \mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2) - F_x(y_1, y_2)$ and $\mathcal{J}_{ix}^{(2)}(y_1, y_2) = 0$.

Recall that all the assumptions and conditions cited below are provided in Appendix A and that all the proofs are given in Appendix B, Appendix C and Appendix D.

Theorem 3.1. *Suppose that* $\max(g, h) \to 0$ *such that as as* $n \to \infty$ *the sequences* $nh^5(\log n)^{-1}$, $ng^5(\log n)^{-1}$, hg^{-1} *and* $\sqrt{ng}(\log n)^{-1}$ *converge to some constants as* $n \to \infty$. *Assume that Conditions* W_1–W_6 *are satisfied and that Assumptions* $(\mathcal{C}_1)$ *to* $(\mathcal{C}_6)$ *are fulfilled for* F_x, H_x, H_x^u, H_{1x}^c *and* G_x.

For any $0 < \mathfrak{t} < \overline{\tau}_x$, *write* $\mathcal{T}_{\mathfrak{t}} = [0, \mathfrak{t}] \times \mathbb{R}$. *Then, uniformly in* $(y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}$, *we have*

$$
\mathbb{F}_x^{\widehat{G}_x}(y_1, y_2) = \sqrt{nh} \sum_{i=1}^n \left\{ w_{ni}(x, h) \mathcal{J}_{ix}^{(1)}(y_1, y_2) + w_{ni}(x, g) \mathcal{J}_{ix}^{(2)}(y_1, y_2) \right\} + o_{\mathbb{P}}(1).
$$

From Theorem 3.1, it can be seen that, for a large sample size, the behaviour of $\mathbb{F}_x^{\widehat{G}_x}$ is roughly explained by the contribution of two terms. The first term, namely $\sqrt{nh} \sum_{i=1}^n w_{ni}(x, h) \mathcal{J}_{ix}^{(1)}$, provides the asymptotic representation of $\sqrt{nh}[\widehat{F}_x^{G_x}(y_1, y_2) - F_x(y_1, y_2)]$, i.e. when G_x is known. The second term is due to the estimation of the G_x.

Remark 3.2.

1. *Through the proof of Theorem 3.1, the stochastic processes* $Z_x \equiv \sum_{i=1}^n Z_{hi}$, *where*

$$
Z_{hi}(y_1, y_2, G) = \sqrt{nh}\, \mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2) w_{ni}(x, h) \frac{\delta_{1i}}{1 - G(T_{1i})},
$$

plays a central role in showing the asymptotic representation and the weak convergence of both $\mathbb{F}_x^{\widehat{G}_x}$ *and* $\mathbb{B}_x^{\widehat{G}_x}$, *see Appendix B.1. In fact,*

$$
\mathbb{F}_x^{\widehat{G}_x}(y_1, y_2) - \mathbb{F}_x^{G_x}(y_1 y_2) = Z_x(y_1, y_2, \widehat{G}_x) - Z_x(y_1, y_2, G), \text{ and}
$$

$$\mathbb{B}_x^{\widehat{G}_x}(u,v) - \mathbb{B}_x^{G_x}(u,v) = Z_x\left(\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), \widehat{G}_x\right) - Z_x\left(\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), G\right),$$

where $\mathbb{F}_x^{G_x} = \sqrt{nh}(\widehat{F}_x^{G_x} - F_x)$ *and* $\mathbb{B}_x^{G_x} = \sqrt{nh}(\widehat{\mathcal{C}}_x^{G_x} - \mathcal{C}_x)$.

2. *When* $g = h$, *the conditional marginal distribution estimator* $\widehat{F}_{1x}^{\widehat{G}_x}$ *is equal to the conditional Kaplan-Meier estimator introduced in [5], see Appendix F.1. Hence, the random function* $\sqrt{nh}\{\widehat{F}_{1x}^{\widehat{G}_x}(y_1) - F_{1x}(y_1)\}$, *which is equal to* $\sqrt{nh}\{\widehat{F}_x^{\widehat{G}_x}(y_1,y_2) - F_x(y_1,y_2)\}$ *as* $y_2 \to \infty$, *reduces to the conditional Kaplan-Meier process studied in [20]. As expected, it is shown, in Appendix F.2, that these two process share the same asymptotic i.i.d representation.*

3.2. Weak convergence of $\mathbb{F}_x^{\widehat{G}_x}$

In view of Theorem 3.1, the large sample behavior of $\mathbb{F}_x^{\widehat{G}_x}$ will essentially depend on the conditions imposed on the weight functions and on the bandwidths h and g. To establish its weak limit, we consider the zero mean Gaussian processes $\mathbb{J}_x^{(1)}$ and $\mathbb{J}_x^{(2)}$ with covariance function

$$\operatorname{cov}\{\mathbb{J}_x^{(1)}(y_1,y_2), \mathbb{J}_x^{(1)}(y_1',y_2')\} = K_4\left\{\int_0^{y_1\wedge y_1'} \frac{F_x^{[1]}(v, y_2\wedge y_2')}{1-G_x(v)}\,\mathrm{d}v - F_x(y_1,y_2)F_x(y_1',y_2')\right\},$$

and

$$\operatorname{cov}\{\mathbb{J}_x^{(2)}(y_1,y_2), \mathbb{J}_x^{(2)}(y_1',y_2')\} = K_4\int_0^{y_1\wedge y_2'} \frac{\Delta_{F_x}(t\wedge y_1', y_2, v)\Delta_{F_x}(t\wedge y_1', y_2', v)}{\{1-H_{1x}(v)\}^2}\,\mathrm{d}H_{1x}^c(v),$$

where the constant K_4 is defined in Assumption W_4. Moreover, let

$$b_x^{(1)}(y_1,y_2) = K_2\left\{\dot{F}_x(y_1,y_2) - \int_0^{y_1} \dot{G}_x(v)\frac{F_x^{[1]}(v,y_2) + \dot{F}_x^{[1]}(v,y_2)}{1-G_x(v)}\mathrm{d}v\right\} + \frac{K_3}{2}\left\{\ddot{F}_x(y_1,y_2) - \int_0^{y_1}\ddot{G}_x(v)\frac{F_x^{[1]}(v,y_2)}{1-G_x(v)}\,\mathrm{d}v\right\},$$

and

$$b_x^{(2)}(y_1,y_2) = K_2\int_0^{y_1} \frac{\dot{H}_{1x}(v)H_{1x}^{c[1]}(v) + \dot{H}_{1x}^{c[1]}(v)}{1-H_{1x}(v)}\Delta_{F_x}(y_1,y_2,v)\,\mathrm{d}v + \frac{K_3}{2}\int_0^{y_1} \frac{\ddot{H}_{1x}(v)H_{1x}^{c[1]}(v) + \ddot{H}_{1x}^{c[1]}(v)}{1-H_{1x}(v)}\Delta_{F_x}(y_1,y_2,v)\,\mathrm{d}v,$$

where the constants K_2 and K_3 are defined in Assumption W_2 and W_3, respectively.

Corollary 3.3. *Suppose that the assumptions of Theorem 3.1 are met. For any* $0 < \mathfrak{t} < \overline{\tau}_x$, *write* $\mathcal{T}_{\mathfrak{t}} = [0, \mathfrak{t}] \times \mathbb{R}$.
(a) If $hg^{-1} \to 0$ *and* $\sqrt{nh}g^2 \to K_5$ *as* $n \to \infty$ *for some* $K_5 > 0$, *then* $\mathbb{F}_x^{\widehat{G}_x}$ *converges weakly in* $l^\infty(\mathcal{T}_{\mathfrak{t}})$ *to* $\mathbb{J}_x^{(1)} + K_5 b_x^{(2)}$ *over* $\mathcal{T}_{\mathfrak{t}}$.
(b) If $g = h$, *and in addition if* $\sqrt{nh^5} \to K_6$ *as* $n \to \infty$ *for some* $K_6 > 0$, *then* $\mathbb{F}_x^{\widehat{G}_x}$ *converges weakly to a gaussian process having the representation* $\mathbb{J}_x \equiv \mathbb{J}_x^{(1)} + \mathbb{J}_x^{(2)} + K_6\{b_x^{(1)} + b_x^{(2)}\}$ *over* $\mathcal{T}_{\mathfrak{t}}$, *with*

$$\begin{aligned} &\operatorname{cov}\{\mathbb{J}_x^{(1)}(y_1, y_2), \mathbb{J}_x^{(2)}(y_1', y_2')\} \\ &\quad = \int_0^{y_1'} \frac{\Delta_{F_x}(y_1', y_2', v)}{\{1 - H_{1x}(v)\}^2} \{F_x(y_1 \wedge v, y_2) - H_{1x}(v) F_x(y_1, y_2)\}\, \mathrm{d}H_{1x}^c(v) \\ &\qquad - F_x(y_1, y_2) \int_0^{y_1'} \frac{H_{1x}^c(v)}{1 - H_{1x}(v)} \Big\{ F_x^{[1]}(v, y_2') - \frac{\Delta_{F_x}(y_1', y_2', v) H_{1x}^{[1]}(v)}{1 - H_{1x}(v)} \Big\}\, \mathrm{d}v\,. \end{aligned}$$

Note that $K_6 b_x^{(1)}$ represents the bias of the process $\sqrt{nh}\sum_{i=1}^n w_{ni}(x, h)\mathcal{J}_{ix}^{(1)}$ as $n \to \infty$ and $K_5 b_x^{(2)}$ is the asymptotic bias of $\sqrt{nh}\sum_{i=1}^n w_{ni}(x, g)\mathcal{J}_{ix}^{(2)}(y_1, y_2)$.

Also note that, when there is no censoring, the asymptotic bias and covariance function of $\mathbb{F}_x^{\widehat{G}_x}$ respectively reduce to $K_2 \dot{F}_x(y_1, y_2) + (K_3/2)\, \ddot{F}_x(y_1, y_2)$, and $K_4\{F_x(y_1 \wedge y_1', y_2 \wedge y_2') - F_x(y_1, y_2)F_x(y_1', y_2')\}$, which match the asymptotic bias and covariance of the process $\sqrt{nh}(\widehat{F}_x - F_x)$ in the context of complete data (see e.g. [22]).

Remark 3.4. *According to Corollary 3.3 (a), when the bandwidth* $hg^{-1} \to 0$, *i.e. g is asymptotically larger than h, the impact of estimating the conditional probability of censoring on the asymptotic distribution of* $\widehat{F}_x^{\widehat{G}_x}$ *appears only in the asymptotic bias* $K_5 b_x^{(2)}$. *Note that the bias related to the process* $\mathbb{F}_x^{G_x}$ *disappears in that case. But, as in that case* $hg^{-1} \to 0$ *and* $ng^5 < \infty$ *as* $n \to \infty$, *it follows that h must be chosen so that* $nh^5 \to 0$. *This excludes the possibility of taking* $h \sim n^{-1/5}$, *which is the usual order of the optimal bandwidth parameter in the mean squared error sense.*

3.3. Weak convergence of $\mathbb{B}_x^{\widehat{G}_x}$

The next result states the weak limit of the conditional copula estimator under random censoring. Let α_x° be a zero-mean gaussian process with covariance function $\operatorname{cov}\{\alpha_x^\circ(u), \alpha_x^\circ(u')\} = K_4(u \wedge u' - uu')$, and let $b_x^{(3)}(v) = K_2 \dot{F}_{2x}\{F_{2x}^{-1}(v)\} + 2^{-1} K_3 \ddot{F}_{2x}\{F_{2x}^{-1}(v)\}$.

Proposition 3.5. *Assume that the conditions of Theorem 3.1 as well as Condition* $(\mathcal{D})$ *are satisfied. For any* $0 < \mathfrak{t} < \overline{\tau}_x$, *let* $\widetilde{\mathcal{T}}_{\mathfrak{t}} = [0, H_{1x}(\mathfrak{t})] \times [0, 1]$.

(a) If $hg^{-1} \to 0$ and $\sqrt{nh}g^2 \to K_5$ as $n \to \infty$ for some $K_5 > 0$, then $\mathbb{B}_x^{\widehat{G}_x}$ converges weakly in $l^\infty(\widetilde{\mathcal{T}}_{\mathfrak{t}})$ to a gaussian process with the following representation:

$$\alpha_x(u,v) - \mathcal{C}_x^{[1]}(u,v)\alpha_x(u,1) - \mathcal{C}_x^{[2]}(u,v)\alpha_x^\circ(v),$$

where $\alpha_x(u,v) = \mathbb{J}_x^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\} + K_5\, b_x^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\}$ *and*

$$\operatorname{cov}\{\alpha_x(u,v), \alpha_x^\circ(v')\} = K_4\{\mathcal{C}_x(u, v\wedge v') - \mathcal{C}_x(u,v)v'\}\,.$$

(b) If $h = g$ and $\sqrt{nh^5} \to K_6$ as $n \to \infty$ for some $K_6 > 0$, then $\mathbb{B}_x^{\widehat{G}_x}$ converges weakly in $l^\infty(\widetilde{\mathcal{T}}_{\mathfrak{t}})$ to a gaussian process with the following representation

$$\beta_x(u,v) - \mathcal{C}_x^{[1]}(u,v)\beta_x(u,1) - \mathcal{C}_x^{[2]}(u,v)\{\alpha_x^\circ(v) + K_6 b_x^{(3)}(v)\},$$

where $\beta_x(u,v) = \mathbb{J}_x\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\}$, *and for* $y_1, y_2, y' \in \mathbb{R}$,

$$\begin{aligned}
&K_4^{-1}\operatorname{cov}[\mathbb{J}_x(y_1,y_2), \alpha_x^\circ\{F_{2x}(y')\}] \\
&= F_x(y_1, y_2\wedge y') - F_x(y_1,y_2)F_{2x}(y') \\
&\quad + \int_0^{y_1} \frac{H_x(z,y') - H_{1x}(z)F_{2x}(y')}{\{1-H_{1x}(z)\}^2}\Delta_{F_x}(y_1,y_2,z)\,\mathrm{d}H_{1x}^c(z) \\
&\quad\quad + \int_0^{y_1} \frac{H_x^c(z,y') - H_{1x}^c(z)F_{2x}(y')}{1-H_{1x}(z)}\left\{F_x^{[1]}(z,y_2) - \frac{\Delta_{F_x}(y_1,y_2,z)H_{1x}^{[1]}(z)}{1-H_{1x}(z)}\right\}\mathrm{d}z.
\end{aligned}$$

As shown in Appendix D, the covariance structure of the tight limit of $\mathbb{B}_x^{\widehat{G}_x}$ is a consequence of the fact that $\mathbb{B}_x^{\widehat{G}_x} = \sqrt{nh}\sum_{i=1}^n\{w_{ni}(x,h)\mathbb{j}_{ix}^{(1)} + w_{ni}(x,g)\mathbb{j}_{ix}^{(2)}\} + o_{\mathbb{P}}(1)$ on $l^\infty(\widetilde{T}_{\mathfrak{t}})$, where

$$\begin{aligned}
\mathbb{j}_{ix}^{(1)}(u,v) =& \mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\} - \mathcal{C}_x^{[1]}(u,v)\mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(1)\} \\
&- \mathcal{C}_x^{[2]}(u,v)[\mathbb{I}\{Y_{2i} \le F_{2x}^{-1}(v)\} - v],
\end{aligned}$$

and

$$\mathbb{j}_{ix}^{(2)}(u,v) = \mathcal{J}_{ix}^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\} - \mathcal{C}_x^{[1]}(u,v)\mathcal{J}_{ix}^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(1)\}.$$

When all the δ_{1i}'s are equal to one, i.e. for complete data, the term $\mathbb{J}_x^{(2)}$ reduces to 0. Hence, the covariance structure of the limit process α_x matches the one of the conditional copula process $\sqrt{nh}(\widehat{\mathcal{C}}_x - \mathcal{C}_x)$ established by [22].

4. Simulation study

The nonparametric estimation of the conditional copula involves a choice for the weight functions $w_{n1}(x,\cdot), \ldots, w_{nn}(x,\cdot)$ that fulfils the required assumptions listed in Appendix A.2. It is shown in [12] that the requirements W_1–W_5 are satisfied, among others, by the Nadaraya–Watson and the local linear weights, given respectively in (2.1) and (2.2).

The simulation results that will be reported here have been obtained using the local linear weights with the triweight function $K(y) = 35(1-y^2)^3\,\mathbb{I}(|y| \le 1)/32$. When negative weights occur, they are truncated to zero and the remaining weights are simply re-scaled so that they sum to one. As pointed out in [12], this modification is asymptotically negligible. Finally, note that all the numerical experiments were also run using the Nadaraya–Watson kernel. As the results were very similar, they are not presented here.

The primary aim of this section is to evaluate the performance of the proposed conditional copula estimator with respect to the percentage of censoring, the influence of the covariate on the dependence and the effect of the sample size. This performance is evaluated by considering the *average squared bias* (ASB) and the *average variance* (AV). To be specific, if $\widehat{\mathcal{C}}_x$ is some estimator of $\mathcal{C}_x$, then

$$\mathrm{ASB}(\widehat{\mathcal{C}}_x) = \frac{1}{K^2}\sum_{i,j=1}^{K}\Big[\mathrm{E}\{\widehat{\mathcal{C}}_x(u_i,u_j)\} - \mathcal{C}_x(u_i,u_j)\Big]^2, \text{ and}$$

$$\mathrm{AV}(\widehat{\mathcal{C}}_x) = \frac{1}{K^2}\sum_{i,j=1}^{K}\mathrm{E}\Big(\widehat{\mathcal{C}}_x^2(u_i,u_j) - \Big[\mathrm{E}\{\widehat{\mathcal{C}}_x(u_i,u_j)\}\Big]^2\Big).$$

The two criteria, ASB and AV, have been estimated from 1 000 replicates under each of the scenario with $n = 250$ and $n = 1\,000$ and $K = 15$.

The nonparametric estimation of $\mathcal{C}_x$ also requires a choice for either one or two bandwidth parameters. Indeed, an interesting aspect of Proposition 3.5 is that the limiting distribution of the copula process $\mathbb{B}_x^{\widehat{G}_x}$ differs in the case when $g = h$ and when $hg^{-1} \to 0$. Therefore, the secondary aim of this section is to evaluate the impact of using a single or both bandwidth parameters in the estimation of $\mathcal{C}_x$. In the following, let $\widehat{\mathcal{C}}_x^{\widehat{G}_x,1}$ and $\widehat{\mathcal{C}}_x^{\widehat{G}_x,2}$ denote the estimators resulting from the choices $g = h$ and $g \ne h$ respectively. Setting $g = h \times 0.25\log(n)$ for $\widehat{\mathcal{C}}_x^{\widehat{G}_x,2}$, their performance is compared for different values of h.

The covariate is generated from the standard normal distribution and the estimation of the conditional copula is evaluated at $x = 0.5$. The copula which joins the marginals is either a normal copula $\mathcal{C}_\varrho^{\mathrm{N}}$ or a Clayton copula $\mathcal{C}_\gamma^{\mathrm{CL}}$. These are defined for $-1 < \varrho < 1$ and $\gamma > 0$ by

$$\mathcal{C}_\varrho^{\mathrm{N}}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)} \varphi_\varrho(y,z)\ \mathrm{d}z\ \mathrm{d}y,$$

and

$$\mathcal{C}_\gamma^{\mathrm{CL}}(u,v) = \left(u^{-\gamma} + v^{-\gamma} - 1\right)^{-1/\gamma},$$

where φ_ϱ is the bivariate standard Normal density with correlation ϱ and Φ is the standard Normal distribution. It is convenient to quantify the dependency using Kendall's tau. For any copula $\mathcal{C}$, its associate Kendall's tau can be written as

$$\mathfrak{T}(\mathcal{C}) = 4\int_0^1\int_0^1 \mathcal{C}(u_1,u_2)\ \mathrm{d}\mathcal{C}(u_1,u_2) - 1. \tag{4.1}$$

For a given conditional copula $\mathcal{C}_x$, the conditional Kendall's tau, as suggested by [8] for complete data, is given by $\mathfrak{T}_x = \mathfrak{T}(\mathcal{C}_x)$.

In the following, both the conditional Normal and Clayton copulas are parameterized in such a way that $\mathfrak{T}_x = a_1 \times \{\Phi(x) - 0.2\}^2$. The constant a_1 is chosen such that $\mathfrak{T}_x \in \{0, .1, .25, .35\}$. From the relationships between Kendall's tau and the parameters of the Normal and Clayton copulas, this can be done by setting

$$\varrho(x) = \sin\left(\frac{\mathfrak{T}_x \times \pi}{2}\right) \qquad \text{and} \qquad \gamma(x) = \frac{2\,\mathfrak{T}_x}{1 - \mathfrak{T}_x}.$$

This means that the triplet (Y_{1i}, Y_{2i}, x_i) is obtained by considering either $\varrho(x_i)$ or $\gamma(x_i)$ combined with the corresponding copula. As discussed at the beginning of Section 3.1, the tail support of the distribution of a random variable may not be identified due to right censoring when the support of the censoring variable is included in the support of the variable of interest, i.e. when $\tau_{\mathcal{C}_x} < \tau_{F_x}$. To evaluate the impact on the estimation of $\mathcal{C}_x$, the cases where $\tau_{\mathcal{C}_x} = \tau_{F_x}$ and $\tau_{\mathcal{C}_x} < \tau_{F_x}$ are examined separately in the following two sections.

4.1. $\tau_{G_x} = \tau_{F_x}$

Here, we have considered the case where $\tau_{G_x} = \tau_{F_x} = \infty$. To do this, the conditional marginal distributions of Y_{1i} and Y_{2i} given $X = x_i$ are generated from an exponential distribution with mean given, for some constant a_2, by

$$\lambda_{x_i} = a_2 \times \{1 + \Phi(x_i) + \Phi(x_i)^2\},$$

i.e. $F_{jx_i}(y) = \mathbb{P}\{Y_{ji} \leq y \mid X = x_i\} = 1 - e^{-\frac{y}{\lambda_{x_i}}}$. The censoring variable C_{1i} is also picked as an exponential but with mean $a_3 \times \{1 + \Phi(x_i) + \Phi(x_i)^2\}$, for some constant a_3. Hence, the probability of censoring conditional on $X = x$, denoted θ thereafter, is given by $\frac{a_2}{a_2+a_3}$. The results are reported for $a_2 = 5$ and $\theta \in \{.2, .4, .6\}$ in Tables 1 and 2.

4.2. $\tau_{G_x} < \tau_{F_x}$

Here, we have considered $\tau_{F_x} = \infty$ and $\tau_{\mathcal{C}_x} < \infty$. In that case, the conditional marginal distributions of Y_{i1} and Y_{i2} given $X = x_i$ are generated from an exponential distribution with mean λ_{x_i}. The censoring variable C_{i1} is generated from a uniform distribution over $[0, a_4\lambda_{x_i}]$, for some constant a_4. We can show that in this case, the percentage of censoring is given by

$$\theta = \left(1 - e^{-a_4}\right)/a_4.$$

The constant a_4 is chosen such that $\theta \in \{0.2, 0.4\}$. We have also covered the scenario $\theta = 0$, which corresponds to the situation when all the survival times are observed. The results are reported in Table 3 and 4.

4.3. Comments on the simulations results

From the obtained results it can be seen that, globally, when the association between Y_1 and Y_2 increases, the bias increases whereas the variance decreases slightly. Another interesting finding is that there is no significant difference between the results obtained with a single bandwidth ($g = h$) and double bandwidth ($g \neq h$), except for large sample size ($n = 1000$). In this case, double bandwidth reduces the bias substantially without increasing the variance of the resulting estimator. The results from the Normal and Clayton copula are quite similar. Also, and as expected, increasing the percentage of censoring decreases the performances of the copula estimator both in terms of bias and variance. The opposite is observed regarding the effect of the sample size. A larger (resp. smaller) bandwidth is needed when censoring (resp. sample size) increases. Note that we obtain much more accurate information on the conditional copula when $\tau_{G_x} = \tau_{F_x}$. When $\tau_{G_x} < \tau_{F_x}$, one needs a large sample size to get accurate estimates otherwise the results should be interpreted with care, especially when the percentage of censoring is high. Finally, as for any kernel-based estimator, we can see that a large bandwidth, typically, leads to a larger bias and a smaller variance. This becomes clear with a large sample size (see the results for $n = 1000$).

5. Illustrative example

In this section, we illustrate our methodology by considering a dataset, analyzed by [11], on patients with advanced lung cancer from the North Central Cancer Treatment Group. Their study was originally developed to determine whether descriptive information from a patient-completed questionnaire could provide prognostic information that was independent from that already obtained by the patient's physician. Requested information, before entering the study, include age, calories intake, weight-loss in the last six months, etc. Three variables are considered in our example: the survival time between onset of lung cancer and death (Y_1), the weight lost in the last six months before entering the study (Y_2) and the age of the patient (X), which is our conditioning variable. The dataset contains information on $n = 228$ patients among whom 165 died and 63 were right-censored during the follow-up period.

We start our analysis by investigating the marginal effects of age on the survival time and weight loss. Figure 1 illustrates the local linear regression estimators of the survival time and weight loss on age. We can see that age has an effect on both variables. In fact, the survival time and the weight loss increase for young people, i.e. between 39 and 50 years. But there is no clear effect when age is higher than 50 years.

The goal here is to assess the effect of age on the relationship between the survival time and the weight loss in the last six months using the conditional

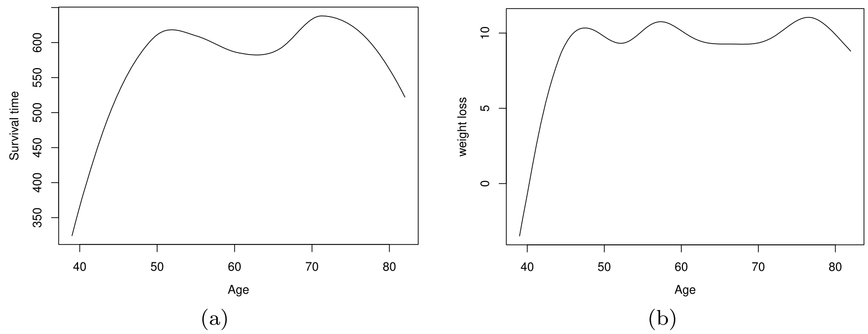

FIG 1. *(a) Nonparametric regression estimator of (a) survival time on age; (b) weight loss on age.*

Kendall's tau of (Y_1, Y_2) given $X = x$, given by

$$\mathfrak{T}(\mathcal{C}_x) = 4\int_0^1\int_0^1 \mathcal{C}_x(u_1, u_2)\ \mathrm{d}\mathcal{C}_x(u_1, u_2) - 1. \tag{5.1}$$

A natural way to estimate this coefficient is to replace the unknown quantity $\mathcal{C}_x$ in the above expression by its nonparametric estimator $\widehat{\mathcal{C}}_x^{\widehat{G}_x}$ given by (2.7). This can be expressed as

$$\mathfrak{T}(\widehat{\mathcal{C}}_x^{\widehat{G}_x}) = 4\int_0^{H_{1x}(\mathfrak{t})}\int_0^1 \widehat{\mathcal{C}}_x^{\widehat{G}_x}(u_1, u_2) d\widehat{\mathcal{C}}_x^{\widehat{G}_x}(u_1, u_2) - 1. \tag{5.2}$$

Except for the case where $H_{1x}(\mathfrak{t}) = 1$, the truncation in the integral above is needed because $\widehat{\mathcal{C}}_x^{\widehat{G}_x}$ is inconsistent outside $[0, H_{1x}(\mathfrak{t})] \times [0, 1]$, see Proposition 3.5 above. More about this subject can be found in [23]. Unfortunately, the quantity $H_{1x}(\mathfrak{t})$ is unknown and there is no obvious way to estimate it without imposing some restrictive assumptions on the data generating process. In practice one may consider (5.2) without the truncation, but then the results should be interpreted with care.

Figure 2 (left side) shows the scatter plot of the observed survival times on the y axis and the weight loss in the last six months on the x axis using different symbols for censored/uncensored observations and different colors for the age of the patients. From this figure it can be seen that there is a relationship between time and weight loss: the survival time has tendency to decrease with increasing weight loss. This tendency is not very strong as the estimated unconditional (global) Kendall's tau is only of -0.002. Figure 2 (right side) shows the estimated conditional Kendall's tau between time and weight loss given the age of the patients. The dashed curve corresponds to the estimator, say $\mathfrak{T}(\widehat{\mathcal{C}}_x)$, obtained ignoring censoring, i.e. by considering all observed times as exact, and the solid curve is the estimator $\mathfrak{T}(\widehat{\mathcal{C}}_x^{\widehat{G}})$ obtained using our method that takes into account censoring. For both estimators a bandwidth $h = 8.2$ was

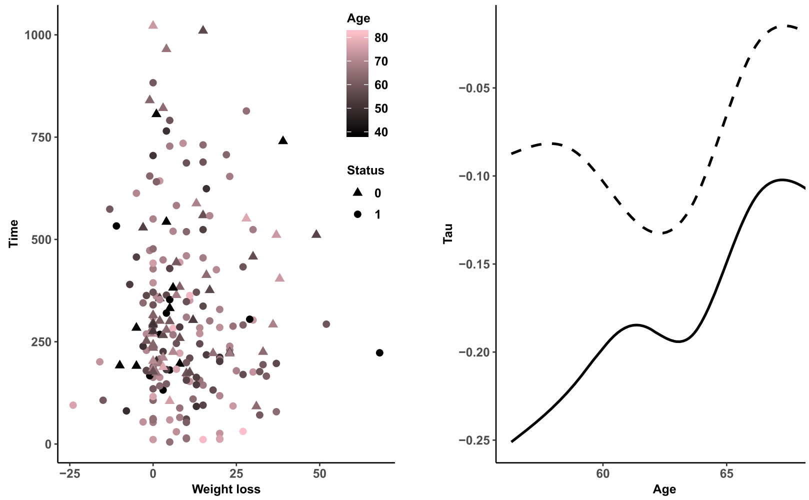

FIG 2. *Left: Scatter plot of time versus age (*$1 =$ *uncensored,* $0 =$ *censored); Right: Conditional Kendall's as a function of age estimated (i) taking into account censoring (solid curve) and (ii) considering censored observations as exact survival times (dashed curve).*

used. We can see that while the estimated conditional Kendall's tau coefficients remain negative their magnitude changes with age. When the latter increases from 55 to 60, the absolute value of $\mathfrak{T}(\widehat{\mathcal{C}}_x^{\widehat{G}})$ decreases from 0.25 to reach 0.18, and then starts increasing to reach 0.20 at the age 63 before decreasing again to reach its minimum value of 0.1 around the age 66. This figure illustrates the advantage of the conditional Kendall's tau over the global one as the former gives a more precise picture of the association between survival time and weight loss accounting for age whereas the latter measures only the "average" association. From this figure one can also see that the "uncorrected" estimator $\mathfrak{T}(\widehat{\mathcal{C}}_x)$ underestimates the strength of the association between time and weight loss.

6. Conclusion and discussion

In this work, we have studied the relationship between two random variables Y_1 and Y_2 conditional on the value taken by a covariate X in the case where the variable Y_1 is subject to random right censoring. We have proposed a kernel smoothing estimator for the conditional copula and have investigated its asymptotic and finite sample properties. To illustrate the usefulness of the proposed method, we have applied it to a dataset on advanced lung cancer.

The methodology can be extended to cover the case when censoring affects not only Y_1, as it is the case in this work, but also Y_2. More precisely, let us assume that Y_j is censored by C_j which, given $X = x$, has as a conditional survival function S_{jx}, $j = 1, 2$. It can be easily shown that

$$F_x(y_1, y_2) = E\Big\{\mathbb{I}(T_1 \leq y_1, T_2 \leq y_2)\frac{\delta_1\delta_2}{S_x^c(T_1, T_2)}|X = x\Big\},$$

where $\delta_j = I(Y_j \leq C_j)$, $j = 1, 2$, and $S_x^c(y_1, y_2) = P(C_1 \geq y_1, C_2 \geq y_2|X = x)$. Clearly, provided that a valid estimator for this quantity is available, one can estimate the copula $\mathcal{C}_x$ using exactly the same approach as we did in the above case when censoring affects only Y_1. For example, if C_1 and C_2 are known to be independent given $X = x$, then one could use $\hat{S}_x^c(y_1, y_2) = \hat{S}_{1x}(y_1)\hat{S}_{2x}(y_2)$, where $\hat{S}_{jx}$ is the Beran's estimator of S_{jx}, $j = 1, 2$. More generally, let $\mathcal{C}_x^c$ be the conditional (survival) copula of (C_1, C_2) given $X = x$. If $\mathcal{C}_x^c$ is known, then one could use $\mathcal{C}_x^c(\hat{S}_{1x}(y_1), \hat{S}_{2x}(y_2))$ to estimate $S_x^c(y_1, y_2)$. In practice $\mathcal{C}_x^c$ is unknown and needs to be estimated from the observed data $(T_{1i}, T_{2i}, \delta_{1i}, \delta_{2i}, X_i)$, $i = 1, \ldots, n$. Due to identifiability issues, this can only be done parametrically by assuming that $\mathcal{C}_x^c$ belongs to a parametric family of copulas. Once such parametric family has been chosen, one can then estimate the copula parameter(s) via maximum likelihood as done, for example, in [7].

The methodology can also be straightforwardly extended to cover the case of $k \geq 2$ variables of interest, say $Y_1, \ldots, Y_k$, where only the variable Y_1 is censored. This can be done by using the same weight functions, i.e. by using

$$\hat{F}_x^{\hat{G}_x}(y_1, y_2, ..., y_k) = \sum_{i=1}^{n} I(T_{1i} \leq y_1, Y_{2i} \leq y_2, ..., Y_{ki} \leq y_k)\frac{w_{ni}(x, h)\delta_{1i}}{1 - \hat{G}_x(T_{1i}-)},$$

as an estimator for the joint conditional distribution of $(Y_1, \ldots, Y_k)|X = x$.

Appendix A. Assumptions

Appendix A.1. (Sub-)distribution functions

Smoothness conditions over $F_x, H_x, H_x^u, H_{1x}^c$ and G_x are needed in the proof of Theorem 3.1. We formulate them for a general (sub-)distribution function L_x and for a fixed $\mathcal{T}_{\mathfrak{s}} = [0, \mathfrak{s}] \times \mathbb{R}$ with $\mathfrak{s} > \mathfrak{t}$, where $\mathfrak{t}$ is given in Theorem 3.1.

($\mathcal{C}_1$) $\dot{L}_x(y_1, y_2) = \frac{\partial}{\partial x}L_x(y_1, y_2)$ exists and is continuous over $V(x) \times \mathcal{T}_{\mathfrak{s}}$, where $V(x)$ is a neighborhood of x;

($\mathcal{C}_2$) $L_x^{[j]}(y_1, y_2) = \frac{\partial}{\partial y_j}L_x(y_1, y_2), j = 1, 2$ exist and are continuous over $V(x) \times \mathcal{T}_{\mathfrak{s}}$;

($\mathcal{C}_3$) $\ddot{L}_x(y_1, y_2) = \frac{\partial^2}{\partial x^2}L_x(y_1, y_2)$ exist and is continuous over $V(x) \times \mathcal{T}_{\mathfrak{s}}$;

($\mathcal{C}_4$) $L_x^{[i,j]}(y_1, y_2) = \frac{\partial^2}{\partial y_i \partial y_j}L_x(y_1, y_2), i, j = 1, 2$ exist and are continuous over $V(x) \times \mathcal{T}_{\mathfrak{s}}$;

($\mathcal{C}_5$) $\dot{L}_x^{[j]}(y_1, y_2) = \frac{\partial^2}{\partial x \partial y_j}L_x(y_1, y_2), j = 1, 2$ exist and are continuous over $V(x) \times \mathcal{T}_{\mathfrak{s}}$;

($\mathcal{C}_6$) $\dot{L}_x^{[1,2]}(y_1, y_2) = \frac{\partial^3}{\partial x \partial y_1 \partial y_2}L_x(y_1, y_2)$ and $\ddot{L}_x^{[1,2]}(y_1, y_2) = \frac{\partial^4}{\partial x^2 \partial y_1 \partial y_2}L_x(y_1, y_2)$ exist and are continuous over $V(x) \times \mathcal{T}_{\mathfrak{s}}$;

Table 1. *Average integrated square bias (AISB $\times 10^4$) estimated from 1 000 replicates of $\mathbb{B}_x^{\widehat{G}_x}$ with $n = 250$ and $n = 1\,000$ in the case $\tau_{\mathcal{C}_x} = \tau_{F_x} = \infty$. Upper pannel: Normal Copula. Bottom pannel: Clayton Copula.*

θ	h	$\mathfrak{T}_x = .1$				$\mathfrak{T}_x = .25$				$\mathfrak{T}_x = .35$			
		$n=250$		$n=1000$		$n=250$		$n=1000$		$n=250$		$n=1000$	
		$g=h$	$g\neq h$	$g=h$	$g\neq h$	$g=h$	$g\neq h$	$g=h$	$g\neq h$	$g=h$	$g\neq h$	$g=h$	$g\neq h$
20%	0.9	1.420	0.510	0.223	0.023	1.403	0.481	0.232	0.028	1.515	0.536	0.286	0.051
	1.2	0.356	0.167	0.014	0.009	0.361	0.178	0.022	0.028	0.471	0.264	0.057	0.057
	1.5	0.138	0.116	0.021	0.030	0.138	0.118	0.057	0.073	0.237	0.212	0.102	0.115
	1.8	0.049	0.050	0.047	0.049	0.083	0.083	0.106	0.111	0.188	0.185	0.164	0.168
	2.1	0.038	0.037	0.085	0.087	0.091	0.090	0.161	0.167	0.216	0.214	0.225	0.231
	2.4	0.049	0.048	0.095	0.096	0.114	0.114	0.189	0.195	0.254	0.251	0.265	0.270
	2.7	0.063	0.064	0.120	0.121	0.159	0.163	0.234	0.240	0.296	0.299	0.319	0.324
40%	0.9	1.707	1.464	0.532	0.050	1.525	1.282	0.535	0.045	1.553	1.304	0.609	0.068
	1.2	0.973	0.639	0.078	0.032	0.859	0.537	0.072	0.039	0.948	0.604	0.119	0.074
	1.5	0.535	0.405	0.054	0.035	0.435	0.311	0.072	0.067	0.530	0.392	0.149	0.131
	1.8	0.390	0.374	0.060	0.056	0.318	0.302	0.114	0.111	0.430	0.409	0.214	0.204
	2.1	0.329	0.321	0.079	0.077	0.300	0.286	0.145	0.151	0.457	0.432	0.270	0.265
	2.4	0.291	0.280	0.088	0.084	0.273	0.258	0.190	0.186	0.448	0.425	0.340	0.328
	2.7	0.259	0.249	0.090	0.087	0.291	0.277	0.210	0.208	0.487	0.466	0.384	0.375
60%	0.9	10.841	8.653	1.551	1.059	9.378	7.289	1.271	0.788	8.699	6.667	1.186	0.683
	1.2	5.339	4.948	0.923	0.764	4.297	3.935	0.699	0.538	3.915	3.566	0.662	0.480
	1.5	4.465	4.420	0.881	0.733	3.342	3.283	0.715	0.566	2.959	2.888	0.743	0.571
	1.8	4.441	4.230	0.752	0.758	3.329	3.130	0.618	0.590	2.968	2.762	0.723	0.662
	2.1	3.699	3.680	0.687	0.691	2.764	2.713	0.627	0.601	2.602	2.530	0.798	0.739
	2.4	3.705	3.641	0.791	0.743	2.622	2.538	0.733	0.659	2.423	2.315	0.950	0.844
	2.7	3.607	3.502	0.824	0.787	2.673	2.561	0.782	0.721	2.550	2.424	1.053	0.961
20%	0.9	1.324	0.442	0.185	0.024	1.567	0.618	0.214	0.034	1.563	0.432	0.486	0.083
	1.2	0.325	0.142	0.031	0.014	0.191	0.170	0.061	0.076	0.368	0.305	0.163	0.191
	1.5	0.141	0.085	0.047	0.049	0.161	0.135	0.174	0.180	0.323	0.325	0.318	0.347
	1.8	0.061	0.055	0.104	0.106	0.178	0.179	0.283	0.292	0.417	0.419	0.488	0.529
	2.1	0.060	0.060	0.121	0.123	0.248	0.252	0.346	0.353	0.561	0.567	0.622	0.634
	2.4	0.085	0.084	0.156	0.157	0.302	0.304	0.441	0.450	0.656	0.662	0.772	0.786
	2.7	0.097	0.096	0.223	0.226	0.368	0.370	0.551	0.560	0.792	0.797	0.915	0.930
40%	0.9	2.746	1.376	0.608	0.158	2.282	1.086	0.527	0.087	2.250	1.266	0.431	0.118
	1.2	0.864	0.600	0.072	0.039	0.910	0.606	0.130	0.085	0.925	0.506	0.228	0.191
	1.5	0.681	0.503	0.059	0.054	0.498	0.378	0.152	0.154	0.528	0.485	0.318	0.324
	1.8	0.369	0.330	0.101	0.098	0.363	0.339	0.254	0.258	0.551	0.532	0.488	0.497
	2.1	0.247	0.226	0.118	0.116	0.338	0.331	0.297	0.303	0.616	0.611	0.587	0.598
	2.4	0.295	0.284	0.131	0.130	0.402	0.393	0.372	0.378	0.716	0.710	0.714	0.724
	2.7	0.350	0.337	0.178	0.178	0.487	0.471	0.455	0.464	0.873	0.864	0.850	0.867
60%	0.9	11.450	7.530	1.866	1.264	9.481	6.093	1.555	1.000	9.294	5.838	2.023	0.767
	1.2	6.922	5.600	0.976	0.885	5.650	4.892	0.871	0.709	4.649	3.984	0.785	0.662
	1.5	4.734	4.604	0.787	0.729	4.327	3.823	0.726	0.676	3.532	3.051	0.789	0.697
	1.8	3.968	3.849	0.841	0.814	2.822	2.675	0.756	0.711	2.421	2.351	0.946	0.891
	2.1	3.569	3.428	0.955	0.972	2.428	2.359	0.877	0.823	2.281	2.173	1.072	1.026
	2.4	3.431	3.385	0.957	0.920	2.712	2.603	0.997	0.956	2.446	2.321	1.265	1.212
	2.7	4.136	4.099	0.997	0.969	3.166	3.043	1.034	0.983	2.918	2.818	1.372	1.306

TABLE 2. *Average integrated variance (AIV* $\times 10^4$*) estimated from 1 000 replicates of* $\mathbb{B}_x^{\widehat{G}_x}$ *with* $n = 250$ *and* $n = 1\,000$ *in the case* $\tau_{C_x} = \tau_{F_x} = \infty$. *Upper pannel: Normal Copula. Bottom pannel: Clayton Copula.*

θ	h	$\mathfrak{T}_x = .1$, $n = 250$, $g = h$	$\mathfrak{T}_x = .1$, $n = 250$, $g \neq h$	$\mathfrak{T}_x = .1$, $n = 1000$, $g = h$	$\mathfrak{T}_x = .1$, $n = 1000$, $g \neq h$	$\mathfrak{T}_x = .25$, $n = 250$, $g = h$	$\mathfrak{T}_x = .25$, $n = 250$, $g \neq h$	$\mathfrak{T}_x = .25$, $n = 1000$, $g = h$	$\mathfrak{T}_x = .25$, $n = 1000$, $g \neq h$	$\mathfrak{T}_x = .35$, $n = 250$, $g = h$	$\mathfrak{T}_x = .35$, $n = 250$, $g \neq h$	$\mathfrak{T}_x = .35$, $n = 1000$, $g = h$	$\mathfrak{T}_x = .35$, $n = 1000$, $g \neq h$
20%	0.9	25.033	9.210	14.408	3.398	26.750	9.235	15.656	3.515	27.752	9.117	16.448	3.559
	1.2	10.312	3.793	4.302	0.945	10.633	3.426	4.534	0.827	10.724	3.072	4.662	0.730
	1.5	5.537	4.444	1.934	0.807	5.385	4.176	1.968	0.723	5.187	3.902	1.959	0.640
	1.8	2.878	2.882	0.708	0.711	2.556	2.560	0.614	0.619	2.264	2.266	0.543	0.548
	2.1	2.691	2.693	0.645	0.648	2.378	2.382	0.569	0.571	2.069	2.075	0.497	0.499
	2.4	2.504	2.514	0.614	0.616	2.194	2.202	0.539	0.542	1.920	1.927	0.469	0.472
	2.7	2.438	2.449	0.589	0.591	2.106	2.115	0.509	0.511	1.812	1.819	0.446	0.448
40%	0.9	19.754	16.672	19.448	1.997	20.449	17.054	21.114	1.857	20.733	17.127	22.155	1.703
	1.2	12.547	6.187	5.979	1.536	12.789	5.728	6.301	1.392	12.750	5.243	6.474	1.264
	1.5	8.735	5.478	4.680	1.334	8.591	4.989	4.923	1.228	8.350	4.530	5.038	1.120
	1.8	4.810	4.798	1.172	1.166	4.347	4.339	1.052	1.051	3.932	3.923	0.952	0.956
	2.1	4.419	4.425	2.225	1.105	3.987	3.996	2.241	1.003	3.573	3.582	2.212	0.902
	2.4	4.256	4.257	1.012	1.012	3.831	3.835	0.916	0.918	3.448	3.457	0.815	0.821
	2.7	4.160	4.165	0.993	0.993	3.727	3.732	0.894	0.898	3.317	3.324	0.800	0.804
60%	0.9	39.515	30.505	15.744	6.265	40.965	30.816	16.583	6.046	41.665	30.771	17.002	5.784
	1.2	17.887	15.903	8.380	4.062	17.286	15.089	8.564	3.782	16.588	14.231	8.591	3.513
	1.5	14.052	13.855	6.879	3.610	13.212	13.017	6.934	3.323	12.337	12.156	6.879	3.052
	1.8	12.913	11.926	3.225	3.169	11.996	10.901	2.952	2.917	11.334	10.160	2.701	2.684
	2.1	11.578	11.564	3.018	2.959	10.467	10.472	2.714	2.681	9.569	9.577	2.435	2.428
	2.4	10.745	10.602	2.844	2.822	9.693	9.577	2.552	2.540	8.816	8.714	2.286	2.283
	2.7	10.262	10.160	2.665	2.641	9.278	9.225	2.373	2.360	8.480	8.454	2.124	2.124
20%	0.9	23.844	8.186	12.115	2.329	28.510	12.680	15.334	1.093	28.287	5.322	23.461	2.235
	1.2	10.258	3.809	6.478	0.967	5.766	4.588	3.260	0.844	8.048	4.297	5.865	0.754
	1.5	6.459	3.218	0.790	0.793	6.420	2.850	0.708	0.708	2.583	3.839	1.908	0.616
	1.8	4.062	2.967	0.714	0.716	2.582	2.591	0.631	0.634	2.283	2.287	1.854	0.556
	2.1	2.649	2.651	0.645	0.647	2.318	2.324	0.570	0.570	2.050	2.065	0.496	0.499
	2.4	2.499	2.500	0.617	0.618	2.138	2.146	0.527	0.530	1.873	1.872	0.459	0.459
	2.7	2.401	2.408	0.588	0.590	2.065	2.068	0.516	0.517	1.809	1.815	0.447	0.449
40%	0.9	29.534	14.413	19.150	6.339	26.838	11.294	20.597	4.213	26.139	13.168	17.852	2.914
	1.2	12.588	7.344	4.891	1.605	16.057	9.181	8.617	1.456	16.231	5.245	8.918	1.334
	1.5	9.609	5.360	3.532	1.328	8.377	4.862	1.205	1.207	5.740	4.483	1.097	1.093
	1.8	5.864	4.788	2.257	1.154	5.476	4.285	1.066	1.068	3.867	3.876	0.970	0.968
	2.1	4.466	4.469	1.079	1.078	3.979	3.963	0.973	0.973	3.525	3.534	0.885	0.889
	2.4	4.187	4.179	1.030	1.029	3.732	3.736	0.909	0.906	3.395	3.405	0.809	0.809
	2.7	4.029	4.033	0.981	0.983	3.583	3.580	0.874	0.878	3.202	3.206	0.782	0.786
60%	0.9	41.210	22.473	15.480	6.182	40.648	20.576	15.916	5.707	43.935	23.550	27.011	4.244
	1.2	23.991	15.355	6.393	4.213	20.477	15.180	8.445	3.795	19.801	14.016	7.162	3.434
	1.5	15.315	13.351	4.689	3.593	15.274	12.114	4.426	3.228	14.634	11.134	6.734	2.975
	1.8	12.811	11.739	3.189	3.106	11.655	10.466	2.904	2.884	9.899	9.875	2.664	2.652
	2.1	11.306	11.293	2.996	2.929	9.956	9.926	2.614	2.608	8.995	8.988	2.362	2.353
	2.4	10.645	10.561	2.833	2.795	9.248	9.253	2.542	2.533	8.522	8.470	2.272	2.256
	2.7	9.772	9.695	2.731	2.701	8.950	8.961	2.454	2.413	8.279	8.243	2.170	2.163

Table 3. *Average integrated square bias (AISB $\times 10^4$) estimated from 1 000 replicates of $\mathbb{B}_x^{\widehat{G}_x}$ with $n = 250$ and $n = 1\,000$ in the case $\tau_C < \tau_{F_x}$. Upper pannel: Normal Copula. Bottom pannel: Clayton Copula.*

θ	h	$\mathfrak{T}_x = .1$				$\mathfrak{T}_x = .25$				$\mathfrak{T}_x = .35$			
		$n = 250$		$n = 1000$		$n = 250$		$n = 1000$		$n = 250$		$n = 1000$	
		$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$
0%	0.9	0.230	0.230	0.005	0.005	0.228	0.228	0.013	0.013	0.284	0.284	0.032	0.032
	1.2	0.089	0.089	0.013	0.013	0.116	0.116	0.035	0.035	0.198	0.198	0.059	0.059
	1.5	0.047	0.047	0.043	0.043	0.073	0.073	0.087	0.087	0.158	0.158	0.117	0.117
	1.8	0.030	0.030	0.075	0.075	0.083	0.083	0.143	0.143	0.167	0.167	0.183	0.183
	2.1	0.029	0.029	0.126	0.126	0.105	0.105	0.209	0.209	0.200	0.200	0.251	0.251
	2.4	0.046	0.046	0.148	0.148	0.130	0.130	0.254	0.254	0.235	0.235	0.300	0.300
	2.7	0.082	0.082	0.189	0.189	0.195	0.195	0.314	0.314	0.301	0.301	0.358	0.358
20%	0.9	1.454	0.729	0.608	0.036	1.437	0.696	0.644	0.039	1.543	0.761	0.739	0.065
	1.2	0.464	0.184	0.091	0.010	0.466	0.192	0.085	0.027	0.584	0.276	0.138	0.057
	1.5	0.150	0.125	0.022	0.030	0.147	0.126	0.058	0.072	0.245	0.221	0.103	0.114
	1.8	0.061	0.060	0.049	0.051	0.093	0.092	0.108	0.113	0.197	0.194	0.168	0.172
	2.1	0.049	0.047	0.086	0.087	0.103	0.100	0.159	0.163	0.228	0.224	0.225	0.229
	2.4	0.059	0.058	0.097	0.098	0.121	0.120	0.190	0.194	0.264	0.261	0.269	0.272
	2.7	0.072	0.073	0.121	0.122	0.167	0.169	0.233	0.238	0.305	0.306	0.320	0.324
40%	0.9	7.057	3.481	6.590	0.773	6.437	2.904	6.698	0.605	6.289	2.718	6.979	0.551
	1.2	4.061	2.231	1.035	0.642	3.590	1.823	0.909	0.515	3.547	1.744	0.931	0.491
	1.5	2.029	1.718	0.798	0.597	1.608	1.339	0.740	0.526	1.548	1.284	0.824	0.559
	1.8	1.829	1.701	0.603	0.615	1.461	1.357	0.596	0.574	1.445	1.343	0.722	0.669
	2.1	1.630	1.539	0.605	0.600	1.357	1.272	0.641	0.610	1.439	1.347	0.799	0.743
	2.4	1.689	1.563	0.642	0.611	1.374	1.260	0.712	0.658	1.448	1.332	0.933	0.852
	2.7	1.538	1.438	0.643	0.626	1.321	1.237	0.741	0.701	1.440	1.354	0.999	0.929
0%	0.9	0.212	0.212	0.008	0.008	0.215	0.215	0.035	0.035	0.298	0.298	0.088	0.088
	1.2	0.065	0.065	0.022	0.022	0.115	0.115	0.096	0.096	0.263	0.263	0.210	0.210
	1.5	0.036	0.036	0.070	0.070	0.133	0.133	0.207	0.207	0.335	0.335	0.374	0.374
	1.8	0.049	0.049	0.142	0.142	0.215	0.215	0.345	0.345	0.464	0.464	0.576	0.576
	2.1	0.064	0.064	0.180	0.180	0.297	0.297	0.429	0.429	0.617	0.617	0.697	0.697
	2.4	0.114	0.114	0.226	0.226	0.392	0.392	0.551	0.551	0.744	0.744	0.867	0.867
	2.7	0.130	0.130	0.305	0.305	0.468	0.468	0.674	0.674	0.893	0.893	1.024	1.024
20%	0.9	1.610	0.533	0.354	0.047	1.841	0.452	0.632	0.040	2.768	0.560	0.364	0.087
	1.2	0.421	0.160	0.108	0.016	0.255	0.181	0.062	0.078	0.466	0.329	0.166	0.191
	1.5	0.129	0.090	0.049	0.050	0.162	0.144	0.169	0.173	0.343	0.329	0.301	0.346
	1.8	0.081	0.065	0.107	0.109	0.180	0.187	0.280	0.287	0.424	0.425	0.510	0.522
	2.1	0.070	0.068	0.127	0.128	0.252	0.252	0.342	0.349	0.566	0.570	0.611	0.623
	2.4	0.097	0.096	0.157	0.158	0.316	0.319	0.442	0.451	0.663	0.665	0.765	0.778
	2.7	0.107	0.107	0.222	0.224	0.380	0.382	0.546	0.556	0.798	0.802	0.903	0.916
40%	0.9	8.009	4.098	5.463	1.269	8.886	3.196	4.665	0.943	8.132	2.789	4.685	0.759
	1.2	3.588	2.221	1.990	0.706	3.255	1.710	1.865	0.691	2.935	1.569	1.852	0.690
	1.5	2.634	1.901	0.875	0.671	2.095	1.520	0.840	0.684	1.720	1.367	0.976	0.805
	1.8	1.888	1.535	0.750	0.723	1.517	1.384	0.869	0.824	1.541	1.406	1.038	0.989
	2.1	1.444	1.332	0.760	0.749	1.341	1.277	0.885	0.849	1.522	1.440	1.151	1.108
	2.4	1.649	1.531	0.756	0.752	1.485	1.400	0.998	0.974	1.634	1.548	1.288	1.250
	2.7	1.776	1.674	0.840	0.812	1.649	1.571	1.038	1.014	1.872	1.791	1.403	1.357

TABLE 4. *Average integrated variance (AIV $\times 10^4$) estimated from 1 000 replicates of $\mathbb{B}_x^{\widehat{G}_x}$ with $n = 250$ and $n = 1\,000$ in the case $\tau_C < \tau_{F_x}$. Upper pannel: Normal Copula. Bottom pannel: Clayton Copula.*

θ	h	$\tau_x = .1$				$\tau_x = .25$				$\tau_x = .35$			
		$n = 250$		$n = 1000$		$n = 250$		$n = 1000$		$n = 250$		$n = 1000$	
		$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$	$g = h$	$g \neq h$
0%	0.9	3.581	3.581	0.849	0.849	3.164	3.164	0.750	0.750	2.771	2.771	0.649	0.649
	1.2	2.703	2.703	0.691	0.691	2.387	2.387	0.589	0.589	2.078	2.078	0.500	0.500
	1.5	2.442	2.442	0.583	0.583	2.109	2.109	0.509	0.509	1.790	1.790	0.436	0.436
	1.8	2.124	2.124	0.511	0.511	1.839	1.839	0.434	0.434	1.558	1.558	0.369	0.369
	2.1	1.933	1.933	0.467	0.467	1.659	1.659	0.398	0.398	1.394	1.394	0.334	0.334
	2.4	1.844	1.844	0.455	0.455	1.565	1.565	0.388	0.388	1.337	1.337	0.327	0.327
	2.7	1.820	1.820	0.423	0.423	1.542	1.542	0.354	0.354	1.282	1.282	0.300	0.300
20%	0.9	24.896	13.329	24.072	4.462	26.604	13.807	26.315	4.695	27.605	13.997	27.771	4.815
	1.2	12.322	3.652	10.927	0.918	12.878	3.294	11.853	0.800	13.125	2.950	12.432	0.700
	1.5	5.445	4.353	1.911	0.787	5.294	4.083	1.946	0.704	5.099	3.811	1.936	0.620
	1.8	2.832	2.831	0.682	0.685	2.498	2.500	0.590	0.594	2.197	2.200	0.520	0.524
	2.1	2.631	2.641	0.628	0.630	2.316	2.326	0.553	0.555	2.013	2.021	0.481	0.484
	2.4	2.438	2.436	0.599	0.601	2.125	2.126	0.526	0.528	1.855	1.858	0.457	0.460
	2.7	2.390	2.391	0.574	0.576	2.055	2.057	0.495	0.498	1.765	1.767	0.431	0.433
40%	0.9	44.135	19.743	63.494	4.405	47.572	20.341	69.854	4.422	49.689	20.560	74.010	4.372
	1.2	26.871	9.619	13.628	1.718	28.587	9.369	14.703	1.524	29.528	9.043	15.368	1.364
	1.5	8.899	6.801	9.151	1.501	8.618	6.285	9.786	1.333	8.284	5.806	10.156	1.188
	1.8	5.181	5.136	1.325	1.293	4.574	4.536	1.142	1.123	4.068	4.038	1.010	1.000
	2.1	4.700	4.698	1.278	1.252	4.118	4.120	1.105	1.088	3.630	3.641	0.963	0.953
	2.4	4.587	4.577	1.189	1.173	3.998	3.999	1.017	1.013	3.528	3.533	0.876	0.876
	2.7	4.440	4.411	1.178	1.156	3.860	3.842	0.998	0.986	3.378	3.366	0.856	0.851
0%	0.9	3.605	3.605	0.884	0.884	3.197	3.197	0.769	0.769	2.773	2.773	0.658	0.658
	1.2	2.790	2.790	0.708	0.708	2.417	2.417	0.611	0.611	2.084	2.084	0.523	0.523
	1.5	2.332	2.332	0.577	0.577	1.996	1.996	0.496	0.496	1.717	1.717	0.425	0.425
	1.8	2.154	2.154	0.513	0.513	1.827	1.827	0.438	0.438	1.550	1.550	0.372	0.372
	2.1	1.926	1.926	0.467	0.467	1.626	1.626	0.398	0.398	1.382	1.382	0.336	0.336
	2.4	1.798	1.798	0.446	0.446	1.476	1.476	0.370	0.370	1.246	1.246	0.311	0.311
	2.7	1.764	1.764	0.419	0.419	1.463	1.463	0.354	0.354	1.244	1.244	0.301	0.301
20%	0.9	26.717	9.075	17.445	4.482	31.666	7.874	26.904	3.450	42.212	8.880	19.740	3.477
	1.2	12.285	3.736	11.903	0.937	8.020	4.481	6.845	0.823	11.665	5.456	7.110	0.731
	1.5	5.279	3.113	0.774	0.775	5.148	2.772	0.682	0.683	5.023	2.496	3.182	0.602
	1.8	5.078	2.892	0.688	0.690	3.727	2.527	0.613	0.615	2.192	2.203	0.535	0.537
	2.1	2.557	2.559	0.626	0.629	2.263	2.268	0.545	0.547	1.973	1.975	0.475	0.478
	2.4	2.430	2.432	0.598	0.598	2.073	2.072	0.516	0.516	1.805	1.807	0.445	0.447
	2.7	2.344	2.347	0.568	0.571	1.997	2.003	0.496	0.497	1.739	1.743	0.433	0.434
40%	0.9	47.333	22.684	52.828	10.644	56.433	19.019	53.310	8.996	55.804	15.543	56.266	5.505
	1.2	21.724	7.633	27.012	1.795	25.893	6.950	29.195	2.762	25.443	6.473	29.539	1.450
	1.5	13.852	5.634	9.070	1.513	12.944	4.986	9.597	1.323	10.548	4.486	9.952	1.181
	1.8	9.272	5.075	2.425	1.300	5.661	4.508	3.542	1.148	5.270	4.026	2.299	1.029
	2.1	4.781	4.750	2.330	1.219	4.180	4.170	1.096	1.085	3.714	3.688	0.971	0.964
	2.4	4.441	4.421	1.207	1.188	3.934	3.912	1.029	1.016	3.564	3.569	0.887	0.879
	2.7	4.291	4.294	1.154	1.136	3.681	3.683	0.978	0.971	3.246	3.235	0.854	0.850

The following assumption is needed to guarantee the weak convergence of $\mathbb{B}_x^{\widehat{G}_x}$.

($\mathcal{D}$) The partial derivatives $\mathcal{C}_x^{[1]}(u,v) = (\partial/\partial u)\mathcal{C}_x(u,v)$ and $\mathcal{C}_x^{[2]}(u,v) = (\partial/\partial v)\mathcal{C}_x(u,v)$ exist and are continuous on $(0,1)\times[0,1]$ and $[0,1]\times(0,1)$, respectively.

Appendix A.2. Weight functions

Assumptions W_1–W_6 below are required to establish the theoretical results presented in Section 3.

W_1. $\sqrt{nh}\max_{1\le i\le n}|w_{ni}(x,h)| = o(1)$;

W_2. $\sqrt{nh}\Big|\sum_{i=1}^n w_{ni}(x,h)(x_i-x) - h^2 K_2\Big| = o(1)$ for some $K_2 = K_2(x)\in[0,\infty)$;

W_3. $\sqrt{nh}\Big|\sum_{i=1}^n w_{ni}(x,h)(x_i-x)^2 - h^2 K_3\Big| = o(1)$ for some $K_3 = K_3(x)\in(0,\infty)$;

W_4. $nh\sum_{i=1}^n\{w_{ni}(x,h)\}^2 - K_4 = o(1)$ for some $K_4 = K_4(x)\in(0,\infty)$;

W_5. $\max_{i\in I_{nx}} x_i - \min_{i\in I_{nx}} x_i = o(1)$, where $I_{nx} = \{i : w_{ni}(x,h) > 0\}$;

W_6. $\sum_{i=1}^n w_{ni}(x,h) - 1 = O\left(\frac{1}{\sqrt{nh}}\right)$.

In what follows, all the expectations of the form $\mathrm{E}\{\psi(T_{1i},Y_{2i},C_i)\}$ have to be understood as taken conditional upon $X = x_i$. Formally, for any $1\le i\le n$ and any real function ψ,

$$\mathrm{E}\{\psi(T_{1i},Y_{2i},C_i)\} = \int\int \psi(y_1,y_2,c)\,\mathrm{d}F_{x_i}(y_1,y_2)\,\mathrm{d}G_{x_i}(c),$$

whenever the left-hand side of the integral exists.

Appendix B. Proof of Theorem 3.1

We start by observing that

$$\widehat{F}_x^{\widehat{G}_x}(y_1,y_2) = \sum_{i=1}^n \delta_{1i}\phi(y_1,y_2,T_{1i},Y_{2i},\widehat{G}_x)w_{ni}(x,h)\,,$$

where for any $(y_1,y_2)\in\mathcal{T}_{\mathrm{t}}$, $(v,v')\in\mathbb{R}^2$ and for any function $G:\mathbb{R}\to[0,1)$:

$$\phi(y_1,y_2,v,v',G) = \frac{\mathbb{I}(v\le y_1, v'\le y_2)}{1-G(v)}.$$

To prove the desired result, we apply the ideas of [18]. To this end, let first define the operator $\mathbb{E}(\cdot)$ by

$$\mathbb{E}\left\{\delta_{1i}L(y_1,y_2,T_{1i},Y_{2i},G)\right\} = \int L(y_1,y_2,v,v',G)\,\mathrm{d}H^u_{x_i}(v,v'), \tag{B.1}$$

whenever the right-hand side of the integral exists. In this definition, L, y_1, y_2 and G may be any stochastic elements. In other words, $\mathbb{E}$ should be understood as the conditional expectation with respect to $(\delta_{1i},T_{1i},Y_{2i})$ only. Observe that when L, y_1,y_2 and G are fixed (non random), $\mathbb{E}\left\{\delta_{1i}L(y_1,y_2,T_{1i},Y_{2i},G)\right\} = \mathrm{E}\left\{\delta_{1i}L(y_1,y_2,T_{1i},Y_{2i},G)\right\}$.

Let start with the following decomposition:

$$\begin{aligned}\sqrt{nh}(\widehat{F}_x^{\widehat{G}_x}(y_1,y_2) - F_x(y_1,y_2)) = \{A_x(y_1,y_2) - \mathbb{E}\left[A_x(y_1,y_2)\right]\} \\ + B_x(y_1,y_2) + \mathbb{E}\left[A_x(y_1,y_2)\right],\end{aligned}$$

where, recalling the definition of $\widehat{F}_x^{G_x}$ at Equation (2.4),

$$A_x(y_1,y_2) = \sqrt{nh}\left\{\widehat{F}_x^{\widehat{G}_x}(y_1,y_2) - \widehat{F}_x^{G_x}(y_1,y_2)\right\} \tag{B.2}$$

and

$$\begin{aligned}B_x(y_1,y_2) &= \sqrt{nh}\Big\{\sum_{i=1}^n \mathbb{I}(T_{1i}\le y_1, Y_{2i}\le y_2)w_{ni}(x,h)\frac{\delta_{1i}}{1-G_x(T_{1i})} - F_x(y_1,y_2)\Big\} \\ &= \sqrt{nh}\sum_{i=1}^n w_{ni}(x,h)\mathcal{J}^{(1)}_{ix}(y_1,y_2).\end{aligned}$$

As B_x is already in the required form, we derive the asymptotic representation of $\mathbb{F}_x^{\widehat{G}_x}$ by showing in Appendix B.1 that $A_x - \mathbb{E}\left(A_x\right)$ is asymptotically negligible, and then by proving in Appendix B.2 that

$$\mathbb{E}\left(A_x(y_1,y_2)\right) = \sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\mathcal{J}^{(2)}_{ix}(y_1,y_2) + o_{\mathrm{a.s}}(1)\,. \tag{B.3}$$

Appendix B.1. Asymptotic negligibility of $A_x - \mathbb{E}(A_x)$

Hereafter, it is important to remember that the X_i's are assumed to be fixed by design. Therefore, the w_{ni}'s are treated as (non-stochastic) constants satisfying assumptions (W_1)–(W_6) above. Also, recall from Theorem 3.1 statement that $\mathcal{T}_\mathfrak{t} = [0,\mathfrak{t}]\times\mathbb{R}$, where $0<\mathfrak{t}<\bar{\tau}_x$. Let $\epsilon>0$ such that $G_x(\mathfrak{t})+\epsilon<1$, and set $\mathfrak{g}_\epsilon \equiv G_x(\mathfrak{t})+\epsilon$. Also, let $\mathcal{G}_{t_\epsilon} \equiv \{G:\mathbb{R}\to[0,1]$ nondecreasing and $G(\mathfrak{t})<\mathfrak{g}_\epsilon\}$, and for any $(y_1,y_2,G)\in\mathcal{T}_\mathfrak{t}\times\mathcal{G}_{\mathfrak{g}_\epsilon}$, define the stochastic processes $Z_x\equiv\sum_{i=1}^n Z_{hi}$, where

$$Z_{hi}(y_1,y_2,G) = \sqrt{nh}\,\mathbb{I}(T_{1i}\le y_1, Y_{2i}\le y_2)w_{ni}(x,h)\frac{\delta_{1i}}{1-G(T_{1i})}. \tag{B.4}$$

We note that Z_x can be viewed as a weighted process indexed by the set of functions from $\mathbb{R}^2 \times \{0,1\} \to \mathbb{R}$ given by

$$\mathcal{F} = \Big\{(v, v', w) \mapsto \mathbb{I}(v \leq y_1, v' \leq y_2)\frac{w}{1-G(v)}, (y_1, y_2) \in \mathcal{T}_{\mathfrak{t}},\, G \in \mathcal{G}_{\mathfrak{g}_\epsilon}\Big\}.$$

Hence, each function $f \in \mathcal{F}$ may be formally identified by a triplet (y_1, y_2, G). The definition of the weighted process Z_x is motivated by the fact that $A_x(y_1, y_2) = Z_x(y_1, y_2, \widehat{G}_x) - Z_x(y_1, y_2, G_x)$. While the ϵ*–enlargement* in the definition of the class $\mathcal{G}_{\mathfrak{g}_\epsilon}$ might appear overdone, it is however required to guarantee that $\widehat{G}_x$ asymptotically fits into $\mathcal{G}_{\mathfrak{g}_\epsilon}$.

Finally, we equip the index set $\mathcal{F}$ with the semimetric ρ_{F_x} defined for $f = (y_1, y_2, G)$ and $f' = (y_1', y_2', G')$ as

$$\rho_{F_x}(f, f') = |F_{1x}(y_1) - F_{1x}(y_1')| + |F_{2x}(y_2) - F_{2x}(y_2')| + \sup_{z\in[0,\mathfrak{t}]} |G(z) - G'(z)|.$$

Note that $(\mathcal{F}, \rho)$ is totally bounded for ρ_{F_x}, as

$$\mathbb{I}\{v \leq y_1, v' \leq y_2\}w\{1-G(v)\}^{-1} \leq \mathbb{I}\{v \leq \mathfrak{t}\}\{1-G(v)\}^{-1} \leq (1-\mathfrak{g}_\epsilon)^{-1} < \infty\,.$$

Moreover, $\mathfrak{F} \equiv \frac{1}{1-\mathfrak{g}_\epsilon}$ is an envelope function of $\mathcal{F}$ (i.e for all $f \in \mathcal{F}$, $f \leq \mathfrak{F}$).

Assumptions W_1, W_4, W_5, W_6 and $\mathcal{C}_1$ together with Lemma E.1 imply that the process $\overline{Z}_x \equiv Z_x - \mathrm{E}(Z_x)$ indexed by $(\mathcal{F}, \rho_{F_x})$ is asymptotically ρ_{F_x}–equicontinuous in probability. This implies that for any $\eta > 0$ and $\eta' > 0$, there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}^\star\Big\{\sup_{\rho_{F_x}(f,f')<\delta} |\overline{Z}_x(f') - \overline{Z}_x(f)| > \eta\Big\} < \eta'. \tag{B.5}$$

From there, to show the asymptotic negligibility of $A_x - \mathbb{E}(A_x)$, it suffices to prove that for sufficiently large n, (a) $\widehat{G}_x \in \mathcal{G}_{\mathfrak{g}_\epsilon}$, and (b) with probability 1, we have $\rho_{F_x}\{(y_1, y_2, \widehat{G}_x), (y_1, y_2, G_x)\} < \delta$. Indeed, if (a) and (b) hold, we obtain

$$\sup_{(y_1,y_2)\in\mathcal{T}_{\mathfrak{t}}} \left|\sqrt{nh}\,[A_x(y_1, y_2) - \mathbb{E}\{A_x(y_1, y_2)\}]\,\right| \leq \sup_{\rho_{F_x}(f,f')<\delta} |\overline{Z}_x(f') - \overline{Z}_x(f)|,$$

and the desired result follows from Equation (B.5).

As for (a), since it is assumed that the w_{ni}'s satisfy Assumptions W_1 to W_5, one can use Lemma 3 in [12] to derive an exponential inequality for the two following random quantities $\sup_{y_2}\sum_{i=1}^n w_{ni}(x, g)\{\mathbb{I}(T_{1i} \leq y_2) - H_{1x_i}(y_2)\}$ and $\sup_{y_2}\sum_{i=1}^n w_{ni}(x, g)\{\mathbb{I}(T_{1i} \leq y_2, \delta_{1i} = 0) - H^c_{x_i}(y_2)\}$, and then, from standard arguments, show that $\sum_{i=1}^n w_{ni}(x, g)\{H_{1x_i}(y_2) - H_{1x}(y_2)\}$ and $\sum_{i=1}^n w_{ni}(x, g)$ $\{H^c_{x_i}(y_2) - H^c_x(y_2)\}$ are $O(g^2)$. Using theses and following a similar arguments in [19], one can prove that

$$\sup_{t\in[0,\mathfrak{t}]} |\widehat{G}_x(t) - G_x(t)| = O\{(ng)^{-1/2}(\log n)^{1/2}\} \qquad \text{a.s.}, \tag{B.6}$$

which implies (a).

(b) follows from (B.6) as $\rho_{F_x}\{(y_1, y_2, \widehat{G}_x), (y_1, y_2, G_x)\} = \sup_{t\in[0,\mathfrak{t}]} |\widehat{G}_x(t) - G_x(t)| = o_{\text{a.s}}(1)$. This concludes the proof of the asymptotic negligibility of $A_x - \mathbb{E}(A_x)$.

Appendix B.2. Asymptotic representation of $\mathbb{E}(A_x)$

Using Equation (B.1), the random function $D_x = \mathbb{E}(A_x)$ can be expressed as

$$\begin{aligned} D_x(y_1, y_2) &= \sqrt{nh} \int_0^{y_1} \int_0^{y_2} \frac{\widehat{G}_x(v) - G_x(v)}{\{1 - G_x(v)\}\{1 - \widehat{G}_x(v)\}} \sum_{i=1}^n w_{ni}(x, h) \, \mathrm{d}H^u_{x_i}(v, v') \\ &= \sqrt{nh} \int_0^{y_1} \frac{\widehat{G}_x(v) - G_x(v)}{\{1 - G_x(v)\}\{1 - \widehat{G}_x(v)\}} \sum_{i=1}^n w_{ni}(x, h) \int_0^{y_2} \mathrm{d}H^u_{x_i}(v, v') \\ &= \sqrt{hg^{-1}} \int_0^{y_1} \frac{\mathbb{G}_x(v)}{\{1 - G_x(v)\}\{1 - \widehat{G}_x(v)\}} \sum_{i=1}^n w_{ni}(x, h) H^{u[1]}_{x_i}(v, y_2) \, \mathrm{d}v, \end{aligned}$$

where $\mathbb{G}_x = \sqrt{ng}(\widehat{G}_x - G_x)$. Next, let

$$\widetilde{D}_x(y_1, y_2) = \sqrt{hg^{-1}} \int_0^{y_1} \frac{\mathbb{G}_x(v)}{\{1 - G_x(v)\}^2} \sum_{i=1}^n w_{ni}(x, h) H^{u[1]}_{x_i}(v, y_2) \, \mathrm{d}v,$$

and, for any function $\chi \in l^\infty([0, \mathfrak{t}])$, let $\Lambda_x : l^\infty([0, \mathfrak{t}]) \to l^\infty(\mathcal{T}_\mathfrak{t})$ be defined as

$$\begin{aligned} \Lambda_x(\chi)(y_1, y_2) &= \int_0^{y_1} \frac{\chi(v)}{\{1 - G_x(v)\}^2} H^{u[1]}_x(v, y_2) \, \mathrm{d}v \\ &= \int_0^{y_1} \frac{\chi(v)}{1 - G_x(v)} F^{[1]}_x(v, y_2) \, \mathrm{d}v, \end{aligned} \tag{B.7}$$

where we used the identity $H^u_x(v, y_2) = \int_0^v \{1 - G_x(z)\} F^{[1]}_x(z, y_2) \, \mathrm{d}z$. We derive below the asymptotic representation of D_x by showing that: (a) $\sqrt{hg^{-1}} \Lambda_x(\mathbb{G}_x) = \sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \mathcal{J}^{(2)}_{ix} + o_{\mathrm{a.s}}(1)$, (b) $\widetilde{D}_x - \sqrt{hg^{-1}} \Lambda_x(\mathbb{G}_x) = o_{\mathrm{a.s}}(1)$ and (c) $D_x - \widetilde{D}_x = o_{\mathrm{a.s}}(1)$. Note that (a), (b) and (c) together yields Equation (B.3).

To show (a), let

$$\begin{aligned} \chi_{ix}(y_1) = &\{1 - G_x(y_1)\} \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \le v) - H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \, \mathrm{d}H^c_{1x}(v) \\ &- \{1 - G_x(y_1)\} \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \le v, \delta_{1i} = 0) - H^c_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \, \mathrm{d}H_{1x}(v) \\ &+ \{1 - G_x(y_1)\} \frac{\mathbb{I}(T_{1i} \le y_1, \delta_{1i} = 0) - H^c_{1x}(y_1)}{1 - H_{1x}(y_1)}. \end{aligned} \tag{B.8}$$

It is shown in Theorem 2.1 of [20] that if G_x and H^c_{1x} both satisfy Conditions $(\mathcal{C}_1)$ to $(\mathcal{C}_5)$, and if $ng^5(\log n)^{-1} < \infty$ as $n \to \infty$, then, for a particular choice of w_{ni}'s, we have uniformly in $t \in [0, \mathfrak{t}]$, that $\mathbb{G}_x = \sqrt{ng} \sum_{i=1}^n w_{ni}(x, g) \chi_{ix} + o_{\mathrm{a.s}}(1)$. But, we can show that their result is true in the case where the w_{ni}'s are only required to satisfy Assumptions W_1 to W_5. From this, we can conclude that, uniformly in $t \in [0,\mathfrak{t}]$, $\sqrt{hg^{-1}}\mathbb{G}_x = \sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\chi_{ix} + o_{\text{a.s}}(1)$. So, by continuous mapping theorem,

$$\sqrt{hg^{-1}}\Lambda_x\left(\mathbb{G}_x\right) = \sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\Lambda_x(\chi_{ix}) + o_{\text{a.s}}(1).$$

This proves (a), since by switching the order of integration, it is straightforward to show that $\Lambda_x(\chi_{ix})(y_1,y_2) = \mathcal{J}_{ix}^{(2)}(y_1,y_2)$.

To show (b), we start by proving that $\widetilde{D}_x - \sqrt{hg^{-1}}\Lambda_x(\mathbb{G}_x)$ is asymptotically negligible. For this, we observe that for any $(y_1,y_2) \in \mathcal{T}_\mathfrak{t}$,

$$\begin{aligned}|\widetilde{D}_x(y_1,y_2) - \sqrt{hg^{-1}}\Lambda_x(\mathbb{G}_x)(y_1,y_2)| &= \Big|\int_0^{y_1} \frac{\widehat{G}_x(v) - G_x(v)}{\{1-G_x(v)\}^2} R_x(v,y_2)\,\mathrm{d}v\Big| \\ &\leq \{1-G_x(\mathfrak{t})\}^{-2} \sup_{v\in[0,\mathfrak{t}]} |\widehat{G}_x(v) - G_x(v)| \int_0^{y_1} |R_x(v,y_2)|\,\mathrm{d}v, \qquad \text{(B.9)}\end{aligned}$$

where $R_x(v,y_2) = \sqrt{nh}\sum_{i=1}^n w_{ni}(x,h)\{H_x^{u[1]}(v,y_2) - H_{x_i}^{u[1]}(v,y_2)\}$.

Now condition $(\mathcal{C}_6)$ together with W_5 allows the Taylor expansion

$$\begin{aligned}R_x(v,y_2) &= \dot{H}_x^{u[1]}(v,y_2)\sqrt{nh}\sum_{i=1}^n (x-x_i)w_{ni}(x,h) \\ &\quad -2^{-1}\sqrt{nh}\sum_{i=1}^n (x-x_i)^2 w_{ni}(x,h)\,\ddot{H}_{z_i}^{u[1]}(v,y_2),\end{aligned}$$

where z_i lies between x_i and x. Using Assumptions W_2, W_3, W_5 and $(\mathcal{C}_6)$, we obtain that $\int_0^{y_1} |R_x(u,v)|\,\mathrm{d}v = O(1)$. This together with Equations (B.6) and (B.9) implies that, uniformly in $(y_1,y_2) \in \mathcal{T}_\mathfrak{t}$, $\widetilde{D}_x - \sqrt{hg^{-1}}\Lambda_x(\mathbb{G}_x) = o_{\text{a.s}}(1)$, which proves (b).

To show (c), note that

$$\begin{aligned}&|D_x(y_1,y_2) - \widetilde{D}_x(y_1,y_2)| \\ &\quad = \left|\sqrt{nh}\int_0^{y_1} \frac{\{\widehat{G}_x(v) - G_x(v)\}^2}{\{1-G_x(v)\}^2\{1-\widehat{G}_x(v)\}} \sum_{i=1}^n w_{ni}(x,h) H_{x_i}^{u[1]}(v,y_2)\,\mathrm{d}v\right| \\ &\quad \leq \sqrt{hg^{-1}}(ng)^{-1/2} \frac{\sup_{v\in[0,\mathfrak{t}]}\{\mathbb{G}_x(v)\}^2}{\{1-G_x(\mathfrak{t})\}^2\{1-\widehat{G}_x(\mathfrak{t})\}} \\ &\qquad\quad \times \int_0^{y_1} \left|\sum_{i=1}^n w_{ni}(x,h) H_{x_i}^{u[1]}(v,y_2)\right| \mathrm{d}v.\end{aligned}$$

Hence, using Equation (B.6), we deduce that the latter is $o_{\text{a.s}}(1)$, provided that $hg^{-1} < \infty$ and that $\sqrt{ng}(\log n)^{-1} \to \infty$ as $n \to \infty$. This proves that (c) holds and hence concludes the proof of Equation (B.3) and of Theorem 3.1.

Appendix C. Proof of Corollary 3.3

Appendix C.1. Weak convergence of $\mathbb{J}_{xh}^{(1)} \equiv \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)\mathcal{J}_{ix}^{(1)}$

In this section, our goal is to prove that $\mathbb{J}_{xh}^{(1)}$ converges in $l^\infty(\mathcal{T}_\mathfrak{t})$ to a Gaussian process whose representation is $\mathbb{J}_x^{(1)} + K_6 b_x^{(1)}$, where K_6 is a constant such that $\sqrt{nh}h^2 \to K_6$ as $n \to \infty$. For this, we first show that $\mathrm{E}\left(\mathbb{J}_{xh}^{(1)}\right) \to K_6 b_x^{(1)}$ as $n \to \infty$ and that the covariance function of $\mathbb{J}_{xh}^{(1)}$ asymptotically matches the one of $\mathbb{J}_x^{(1)}$. From there, the result will follow from an application of Theorem 2.11.1 of [17].

We now prove that $\mathrm{E}(\mathbb{J}_{xh}^{(1)}) = K_6 b_x^{(1)} + o(1)$. Using

$$\mathrm{E}\left\{\frac{\mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2)\delta_{1i}}{1-G_x(T_{1i})}\right\} = \mathrm{E}\left\{\mathrm{E}(\delta_{1i} \mid Y_{1i}, Y_{2i})\frac{\mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2)}{1-G_x(Y_{1i})}\right\}$$
$$= \mathrm{E}\left\{\frac{1-G_{x_i}(Y_{1i})}{1-G_x(Y_{1i})}\mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2)\right\},$$

Assumptions $\mathcal{C}_1$ and $\mathcal{C}_3$, and Taylor expansion of the map $z \mapsto G_z$ (under Condition W_5), we get

$$\begin{aligned}
&\mathrm{E}\{\mathbb{J}_{xh}^{(1)}(y_1,y_2)\} \\
&\quad = \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)\{F_{x_i}(y_1,y_2) - F_x(y_1,y_2)\} \\
&\qquad + \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)(x_i - x)\mathrm{E}\left\{\frac{-\dot{G}_x(Y_{1i})}{1-G_x(Y_{1i})}\mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2)\right\} \\
&\qquad + \frac{1}{2}\sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)(x_i - x)^2\mathrm{E}\Big\{\frac{-\ddot{G}_{z_i}(Y_{1i})}{1-G_x(Y_{1i})}\mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2)\Big\} \\
&\quad = I_{n,1} + I_{n,2} + I_{n,3},
\end{aligned} \tag{C.1}$$

where z_i lies between x_i and x. A Taylor expansion of $z \mapsto F_z$ leads to

$$\begin{aligned}
I_{n,1} = \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)(x_i - x)\dot{F}_x(y_1,y_2) \\
+ \frac{1}{2}\ddot{F}_x(y_1,y_2)\sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)(x_i - x)^2 + o(1),
\end{aligned}$$

which, by Assumptions W_2, W_3, W_5 and $\mathcal{C}_1$ and the fact that $\sqrt{nh}h^2 \to K_6$, implies that

$$I_{n,1} = K_6\left\{K_2\dot{F}_x(y_1,y_2) + \frac{K_3}{2}\ddot{F}_x(y_1,y_2) + o(1)\right\}, \tag{C.2}$$

where the constants K_2 and K_3 are given in Assumptions W_2 and W_3.

In view of Assumptions $\mathcal{C}_1$ and $\mathcal{C}_5$, we have that the maps $z \mapsto \int_0^{y_1} \dot{G}_x(v) \{1 - G_x(v)\}^{-1} F_z^{[1]}(v, y_2)\, \mathrm{d}v$, $z \mapsto \int_0^{y_1} \dot{G}_x(v)\{1 - G_x(v)\}^{-1} \dot{F}_z^{[1]}(v, y_2)\, \mathrm{d}v$ and $z \mapsto \ddot{G}_z$ are uniformly continuous on $V(x)$. Hence, by proceeding similarly as for $I_{n,1}$, we deduce that

$$
\begin{aligned}
I_{n,2} &= K_6 \int_0^{y_1} \dot{G}_x(v)\{1 - G_x(v)\}^{-1}\{K_2 F_x^{[1]}(v, y_2) + K_3 \dot{F}_x^{[1]}(v, y_2)\}\, \mathrm{d}v \, + o(1) \\
I_{n,3} &= K_6 K_3 \int_0^{y_1} \ddot{G}_x(v)\{1 - G_x(v)\}^{-1} F_x^{[1]}(v, y_2)\, \mathrm{d}v \, + o(1).
\end{aligned} \tag{C.3}
$$

From Equations (C.1) to (C.3), we conclude that $\mathrm{E}\big(\mathbb{J}_{xh}^{(1)}\big) = K_6 b_x^{(1)} + o(1)$.

Now, we calculate the covariance function of $\mathbb{J}_{xh}^{(1)}$. From Assumption W_4 and W_5, can we show that

$$
\begin{aligned}
&\mathrm{cov}\{\mathbb{J}_{xh}^{(1)}(y_1, y_2), \mathbb{J}_{xh}^{(1)}(y_1', y_2')\} \\
&\qquad = nh \sum_{i=1}^n w_{ni}(x, h)^2 \Bigg(\mathrm{E}\Big[\frac{\mathbb{I}(T_{1i} \le y_1 \wedge y_1', Y_{2i} \le y_2 \wedge y_2')\delta_{1i}}{\{1 - G_x(T_{1i})\}^2} \Big] \\
&\qquad\quad - \ \mathrm{E}\left\{ \frac{\mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2)\delta_{1i}}{1 - G_x(T_{1i})} \right\} \mathrm{E}\left\{ \frac{\mathbb{I}(T_{1i} \le y_1', Y_{2i} \le y_2')\delta_{1i}}{1 - G_x(T_{1i})} \right\} \Bigg) \\
&\qquad = K_4 \left\{ \int_0^{y_1 \wedge y_1'} \frac{F_x^{[1]}(v, y_2 \wedge y_2')}{1 - G_x(v)}\, \mathrm{d}v - F_x(y_1, y_2) F_x(y_1', y_2') \right\} + o(1).
\end{aligned}
$$

We are now ready to show that $\mathbb{J}_{xh}^{(1)}$ is asymptotically Gaussian. To do so, note that from the definition of $\overline{Z}_x$, we can write $\mathbb{J}_{xh}^{(1)}(y_1, y_2) = \overline{Z}_x(y_1, y_2, G_x) + \mathrm{E}\{\mathbb{J}_{xh}^{(1)}(y_1, y_2)\}$. As Assumptions W_1, W_4, W_5 and $\mathcal{C}_1$ are satisfied, and since $\{(v, v', w) \mapsto \mathbb{I}(v \le y_1, v' \le y_2) \frac{w}{1 - G_x(v)}, (y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}, \} \subset \mathcal{F}$ (where $\mathcal{F}$ is defined at page 5065), we deduce from the proof of Lemma E.1 that the process $\{\overline{Z}_x(y_1, y_2, G_x)\}_{(y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}}$ fulfils the requirements of Theorem 2.11.1 of [17]. Now, because

$$
\begin{aligned}
\mathrm{cov}\{\mathbb{J}_{xh}^{(1)}(y_1, y_2), \mathbb{J}_{xh}^{(1)}(y_1', y_2')\} &= \mathrm{cov}\{\overline{Z}_x(y_1, y_2, G_x), \overline{Z}_x(y_1', y_2', G_x)\} \\
&\to \mathrm{cov}\{\mathbb{J}_x^{(1)}(y_1, y_2), \mathbb{J}_x^{(1)}(y_1', y_2')\} \quad \text{as } n \to \infty,
\end{aligned}
$$

we conclude that $\{\overline{Z}_x(y_1, y_2, G_x)\}_{(y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}}$ converges weakly in $l^\infty(\mathcal{T}_{\mathfrak{t}})$ to a Gaussian process whose representation matches the one of $\mathbb{J}_x^{(1)}$. The weak convergence of $\mathbb{J}_{xh}^{(1)}$ to $\mathbb{J}_x^{(1)} + K b_x^{(1)}$ follows from the fact that $\mathrm{E}(\mathbb{J}_{xh}^{(1)}) = K b_x^{(1)} + o(1)$.

Appendix C.2. Asymptotic behavior of $\sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \mathcal{J}_{ix}^{(2)}$

We start by computing the mean and the covariance function of $\mathcal{J}_{ix}^{(2)}$. For this, recall the definition of Λ_x at Equation (B.7) and the definition of χ_{ix}

at Equation (B.8). Using the fact that $\Lambda_x(\chi_{ix})(y_1,y_2) = \mathcal{J}^{(2)}_{ix}(y_1,y_2)$ (see Appendix B.1), and since Λ_x is a linear functional, we have

$$\sum_{i=1}^n w_{ni}(x,g)\mathcal{J}^{(2)}_{ix} = \sum_{i=1}^n w_{ni}(x,g)\Lambda_x(\chi_{ix}) = \Lambda_x\left\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}\right\}.$$

From there, using the linearity of Λ_x, we can write

$$\mathrm{E}\left[\Lambda_x\left\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}\right\}\right] = \Lambda_x\left[\mathrm{E}\left\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}\right\}\right].$$

As Assumptions $\mathcal{C}_1$ and $\mathcal{C}_3$ are satisfied, we mimic the proof of Equation (3.3) in [20] to deduce that as long as the w_{ni}'s fulfil Conditions W_2, W_3 and W_5, we have

$$\begin{aligned}\mathrm{E}\Big\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}(y_1)\Big\} &\\ = g^2\{1-G_x(y_1)\}\int_0^{y_1}\frac{K_2\dot{H}_{1x}(s)+\frac{K_3}{2}\ddot{H}_{1x}(s)}{\{1-H_{1x}(s)\}^2}\,\mathrm{d}H^c_{1x}(s)&\\ +\ g^2\{1-G_x(y_1)\}\int_0^{y_1}\frac{\mathrm{d}\left\{K_2\dot{H}^c_x(s)+\frac{K_3}{2}\ddot{H}^c_x(s)\right\}}{\{1-H_{1x}(s)\}} + o(g^2).&\end{aligned}$$

Then, we use integrations by parts to show that

$$\Lambda_x\left[\mathrm{E}\left\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}\right\}\right](y_1,y_2) = g^2\mathrm{E}\left\{\mathbb{J}^{(2)}_x(y_1,y_2)\right\} + o(g^2). \tag{C.4}$$

Focusing now on the covariance function of $\mathcal{J}^{(2)}_{ix}$, by adapting the proof of Lemma A2 in [21] to the case where the w_{ni}'s are only required to satisfy Conditions W_1 to W_5, we obtain using Assumptions W_4 and W_5 that for any $y_1, y_1' \in [0,\mathfrak{t}]$:

$$\begin{aligned}&ng\,\mathrm{cov}\Big\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}(y_1), \sum_{i=1}^n w_{ni}(x,g)\chi_{ix}(y_1')\Big\}\\ &= K_4\{1-G_x(y_1)\}\{1-G_x(y_1')\}\int_0^{y_1\wedge y_1'}\{1-H_{1x}(s)\}^{-2}\,\mathrm{d}H^c_{1x}(s)+o(1).\end{aligned} \tag{C.5}$$

Then, recalling that $\Delta_{F_x}(y_1,y_2,v) = F_x(y_1,y_2) - F_x(v,y_2)$, we have

$$\begin{aligned}&ng\,\mathrm{cov}\Big[\Lambda_x\Big\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}\Big\}(y_1,y_2), \Lambda_x\Big\{\sum_{i=1}^n w_{ni}(x,g)\chi_{ix}\Big\}(y_1',y_2')\Big]\\ &= ng\int_0^{y_1}\int_0^{y_1'}\frac{\sum_{i=1}^n w^2_{ni}(x,g)\mathrm{cov}\{\chi_{ix}(v),\chi_{ix}(v')\}}{\{1-G_x(v)\}\{1-G_x(v')\}}F^{[1]}_x(v,y_2)F^{[1]}_x(v',y_2')\,\mathrm{d}v\,\mathrm{d}v'\end{aligned}$$

$$= K_4 \int_0^{y_1 \wedge y_1'} \int_s^{y_1} \int_s^{y_1'} F_x^{[1]}(v, y_2) F_x^{[1]}(v', y_2')\, \mathrm{d}v\, \mathrm{d}v' \frac{\mathrm{d}H_{1x}^c(s)}{\{1 - H_{1x}(s)\}^2} + o(1)$$
$$= K_4 \int_0^{y_1 \wedge y_1'} \Delta_{F_x}(y_1, y_2, s) \Delta_{F_x}(y_1', y_2', s) \frac{\mathrm{d}H_{1x}^c(s)}{\{1 - H_{1x}(s)\}^2} + o(1)$$
$$= \mathrm{cov}\{\mathbb{J}_x^{(2)}(y_1, y_2), \mathbb{J}_x^{(2)}(y_1', y_2')\} + o(1). \tag{C.6}$$

We next discuss the large sample behavior of $\sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\mathcal{J}_{ix}^{(2)}$. To this end, we start by showing that $\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\mathcal{J}_{ix}^{(2)} - \mathrm{E}(\mathcal{J}_{ix}^{(2)})\}$ converges in $l^\infty(\mathcal{T}_\mathfrak{t})$ to a Gaussian process. From there, we derive the weak limit of $\sqrt{nh}$ $\sum_{i=1}^n w_{ni}(x,g)\mathcal{J}_{ix}^{(2)}$ according to whether $hg^{-1} \to 0$ or $g = h$.

From Lemma E.2, we deduce that $\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\chi_{ix} - \mathrm{E}(\chi_{ix})\}$ is asymptotically tight in $l^\infty([0,\mathfrak{t}])$. Moreover, in view of Equation (C.5), we can use similar arguments as those used at the end of Appendix C.1 to conclude that $\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\chi_{ix} - \mathrm{E}(\chi_{ix})\}$ converges weakly in $l^\infty([0,\mathfrak{t}])$ to a Gaussian process.

As Λ_x is linear and continuous, given Lemma 3.9.8 in [17], we can say that the sequence $\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\mathcal{J}_{ix}^{(2)} - \mathrm{E}(\mathcal{J}_{ix}^{(2)})\} = \Lambda_x[\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\chi_{ix} - \mathrm{E}(\chi_{ix})\}]$ is asymptotically tight in $l^\infty(\mathcal{T}_\mathfrak{t})$ and, in view of Equations (C.6) its tight limit is Gaussian with mean 0 and covariance function matching the one of $\mathbb{J}_x^{(2)}$.

Thus, in the case where $hg^{-1} \to 0$ and $\sqrt{nh}g^2 \to K_5$ as $n \to \infty$ (part (a) of the corollary), we obtain, using Equation (C.4), that

$$\begin{aligned}
\sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\mathcal{J}_{ix}^{(2)} &= \sqrt{hg^{-1}}\Big[\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\left\{\mathcal{J}_{ix}^{(2)} - \mathrm{E}\big(\mathcal{J}_{ix}^{(2)}\big)\right\}\Big] \\
&\quad + \sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\mathrm{E}\big(\mathcal{J}_{ix}^{(2)}\big) \\
&= O_\mathbb{P}(\sqrt{hg^{-1}}) + K_5 b_x^{(2)} + o(1) \\
&= K_5 b_x^{(2)} + o_\mathbb{P}(1) + o(1),
\end{aligned}$$

while in the case where $g = h$ and $\sqrt{nh}h^2 = \sqrt{nh}g^2 \to K_6$ as $n \to \infty$ (Part (b) of the corollary), we obtain that $\sqrt{nh}\sum_{i=1}^n w_{ni}(x,g)\mathcal{J}_{ix}^{(2)}$ converges to a Gaussian process with representation $\mathbb{J}_x^{(2)} + K_6 b_x^{(2)}$.

Appendix C.3. Asymptotic normality of $\sqrt{nh}\sum_{i=1}^n w_{ni}(x,h)(\mathcal{J}_{ix}^{(1)} + \mathcal{J}_{ix}^{(2)})$

Let $\mathcal{J}_{xh} \equiv \sqrt{nh}\sum_{i=1}^n w_{ni}(x,h)(\mathcal{J}_{ix}^{(1)} + \mathcal{J}_{ix}^{(2)})$. It is easy to see that

$$\mathrm{cov}\{\mathcal{J}_{xh}(y_1,y_2), \mathcal{J}_{xh}(y_1',y_2')\} = nh\sum_{i=1}^n w_{ni}(x,h)^2 \mathrm{cov}\{\mathcal{J}_{ix}^{(1)}(y_1,y_2), \mathcal{J}_{ix}^{(1)}(y_1',y_2')\}$$

$$
\begin{aligned}
&+ nh\sum_{i=1}^{n} w_{ni}(x,h)^2 \mathrm{cov}\{\mathcal{J}_{ix}^{(1)}(y_1,y_2), \mathcal{J}_{ix}^{(2)}(y_1',y_2')\}\\
&+ nh\sum_{i=1}^{n} w_{ni}(x,h)^2 \mathrm{cov}\{\mathcal{J}_{ix}^{(1)}(y_1',y_2'), \mathcal{J}_{ix}^{(2)}(y_1,y_2)\}\\
&+ nh\sum_{i=1}^{n} w_{ni}(x,h)^2 \mathrm{cov}\{\mathcal{J}_{ix}^{(2)}(y_1,y_2), \mathcal{J}_{ix}^{(2)}(y_1',y_2')\}\\
=&I_{n,1}^* + I_{n,2}^* + I_{n,3}^* + I_{n,4}^*.
\end{aligned}
$$

Now,

$$
\begin{aligned}
&I_{n,2}^* = o(1)\\
&- K_4\{F_{2x}(y_2') - F_x(y_1',y_2')\}\int_0^{y_1'} \frac{F_x(y_1\wedge v, y_2) - H_{1x}(v)F_x(y_1,y_2)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}}\,\mathrm{d}H_{1x}^c(v)\\
&- K_4F_x(y_1,y_2)\int_0^{y_1'} \frac{1-H_{1x}^u(v)}{1-H_{1x}(v)}F_x^{[1]}(v,y_2')\,\mathrm{d}v - F_x(y_1,y_2)F_x(y_1',y_2')\\
&+ K_4F_x(y_1,y_2)\{F_{2x}(y_2') - F_x(y_1',y_2')\}\int_0^{y_1'} \frac{H_{1x}^u(v)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}}\,\mathrm{d}H_{1x}(v).
\end{aligned}
$$

Hence, $\mathrm{cov}\{\mathcal{J}_{xh}(y_1,y_2), \mathcal{J}_{xh}(y_1',y_2')\} = \mathrm{cov}\{\mathcal{J}_x(y_1,y_2), \mathcal{J}_x(y_1',y_2')\} + o(1)$. The proof of the weak convergence of $\mathcal{J}_{xh}$ follows from similar arguments given in the end of Appendix C.1.

Appendix D. Proof of Proposition 3.5

This section follows the lines of Theorem 1 in [22]. First, we decompose

$$
\sqrt{nh}(\widehat{\mathcal{C}}_x^{\widehat{G}_x} - \mathcal{C}_x) = \{\widetilde{A}_x - \mathbb{E}\,(\widetilde{A}_x)\} + \{\widetilde{A}_x' - \mathbb{E}(\widetilde{A}_x')\} + \widetilde{B}_x + \mathbb{E}\,(\widetilde{A}_x) + \mathbb{E}\,(\widetilde{A}_x'), \quad \text{(D.1)}
$$

where, recalling the definition of A_x at Equation (B.2) and the one of $\widehat{F}_x^{G_x}$ at Equation (2.4),

$$
\begin{aligned}
\widetilde{A}_x(u,v) &= A_x\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v)\right\},\\
\widetilde{A}_x'(u,v) &= \sqrt{nh}\left[\widehat{F}_x^{G_x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v)\right\} - \widehat{F}_x^{G_x}\left\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\right\}\right]
\end{aligned}
$$

and

$$
\widetilde{B}_x(u,v) = \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)\mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\}.
$$

We show next that both $\{\widetilde{A}_x - \mathbb{E}\,(\widetilde{A}_x)\}$ and $\{\widetilde{A}_x' - \mathbb{E}\,(\widetilde{A}_x')\}$ are asymptotically negligible. Then, we prove that

$$
\mathbb{E}\,\{\widetilde{A}_x(u,v)\} = \sqrt{nh}\sum_{i=1}^{n} w_{ni}(g,h)\mathcal{J}_{ix}^{(2)}\left\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\right\} + o_{\mathbb{P}}(1). \quad \text{(D.2)}
$$

Finally, we derive the asymptotic representation of $\mathbb{E}(\widetilde{A}_x')$.

Appendix D.1. Asymptotic negligibility of $\sqrt{nh}\{\widetilde{A}_x - \mathbb{E}\,(\widetilde{A}_x)\}$ *and* $\sqrt{nh}\{\widetilde{A}'_x - \mathbb{E}\,(\widetilde{A}'_x)\}$

First, using Corollary 3.3, we have uniformly in $(y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}$ that

$$\widehat{F}_{1x}^{\widehat{G}_x}(y_1) - F_{1x}(y_1) = o_{\mathbb{P}}(1) \quad \text{and} \quad \widehat{F}_{2x}(y_2) - F_{2x}(y_2) = o_{\mathbb{P}}(1).$$

Secondly, it follows from Assumption W_1 that, for any sufficiently small $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left\{F_{1x}^{-1}(u-\epsilon) \le \left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u) \le F_{1x}^{-1}(u+\epsilon), u \in [0, H_{1x}(\mathfrak{t})]\right\} = 1,$$
$$\lim_{n\to\infty} \mathbb{P}\left\{F_{2x}^{-1}(v-\epsilon) \le \widehat{F}_{2x}^{-1}(v) \le F_{1x}^{-1}(v+\epsilon), v \in [0,1]\right\} = 1. \tag{D.3}$$

Now, from the definition of Z_x above Equation (B.4), we note that

$$\widetilde{A}_x(u,v) = Z_x\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), \widehat{G}_x\right\} - Z_x\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), G_x\right\}$$

and

$$\widetilde{A}'_x(u,v) = Z_x\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), G_x\right\} - Z_x\left\{F_{1x}^{-1}(u), F_{2x}^{-1}(v), G_x\right\}.$$

It follows from Equations (B.6) and (D.3) that

$$\rho_{F_x}(f_1, f_2) = o_{\mathbb{P}}(1) \quad \text{and} \quad \rho_{F_x}(f_2, f_3) = o_{\mathbb{P}}(1) \qquad \text{as } n \to \infty,$$

where

$$f_1 = \left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), \widehat{G}_x\right\}, f_2 = \left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v), G_x\right\},$$

and $f_3 = \{F_{1x}^{-1}(u), F_{2x}^{-1}(v), G_x\}$. Hence, the negligibility of $\widetilde{A}_x - \mathbb{E}(A_x)$ and $\widetilde{A}'_x - \mathbb{E}(A'_x)$ is then ensured by Lemma E.1 and the arguments used in the end of Appendix B.1.

Appendix D.2. Asymptotic representation of $\mathbb{E}(\widetilde{A}_x)$

Let $\epsilon > 0$ be such that $H_{1x}(\mathfrak{t}+\epsilon) < 1$ and $\mathfrak{t}+\epsilon < \mathfrak{s}$, where $\mathfrak{s}$ is given in Appendix A.1. From Equation (D.3), we obtain with probability going to 1 that, $\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u) \in [0, \mathfrak{t}+\epsilon]$, uniformly in $u \in [0, H_{1x}(\mathfrak{t})]$. Thus, we get from Appendix B.2 that, uniformly in $(u,v) \in \widetilde{\mathcal{T}}_{\mathfrak{t}}$,

$$\mathbb{E}\{\widetilde{A}_x(u,v)\} = \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,g)\mathcal{J}_{ix}^{(2)}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v)\right\} + o_{\mathbb{P}}(1). \tag{D.4}$$

Now let $y_{1x} = \left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)$, $y'_{1x} = F_{1x}^{-1}(u)$, and define $y_{2x} = \widehat{F}_{2x}^{-1}(v)$ and $y'_{2x} = F_{2x}^{-1}(v)$. Recall, from Appendix B.2, that

$$\mathcal{J}_{ix}^{(2)}(y_{1x}, y_{2x}) = \Lambda_x(\chi_{ix})(y_{1x}, y_{2x}) = \int_0^{y_{1x}} \chi_{ix}(v)\{1 - G_x(v)\}^{-1} F_x^{[1]}(v, y_{2x}) \mathrm{d}v\,.$$

Using the mean value theorem, we obtain that

$$\mathcal{J}_{ix}^{(2)}(y_{1x}, y_{2x}) - \mathcal{J}_{ix}^{(2)}(y'_{1x}, y_{2x}) = \chi_{ix}(y^\star_{1x}) \frac{F_x^{[1]}(y^\star_{1x}, y_{2x})(y_{1x} - y'_{1x})}{1 - G_x(y^\star_{1x})},$$

for some $y^\star_{1x}$ between y_{1x} and y'_{1x}. Hence,

$$\sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \left\{ \mathcal{J}_{ix}^{(2)}(y_{1x}, y_{2x}) - \mathcal{J}_{ix}^{(2)}(y'_{1x}, y_{2x}) \right\} = \left\{ \sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \chi_{ix}(y^\star_{1x}) \right\} \frac{F_x^{[1]}(y^\star_{1x}, y_{2x})(y_{1x} - y'_{1x})}{1 - G_x(y^\star_{1x})}. \quad \text{(D.5)}$$

Now note that we proved in Lemma E.2 that $\sqrt{ng} \sum_{i=1}^n w_{ni}(x, g)\{\chi_{ix} - \mathrm{E}(\chi_{ix})\}$ is asymptotically tight in $l^\infty([0, \mathfrak{t}])$. With ϵ as above, a meticulous inspection of the arguments used there shows that this result is still true when $\mathfrak{t}$ there is replaced by $\mathfrak{t} + \epsilon$, leading us to conclude $\sqrt{ng} \sum_{i=1}^n w_{ni}(x, g)\{\chi_{ix} - \mathrm{E}(\chi_{ix})\}$ is asymptotically tight in $l^\infty([0, \mathfrak{t}+\epsilon])$. Consequently, since $y_{1x}, y'_{1x} \in [0, \mathfrak{t}+\epsilon]$ with probability going to 1 (see above Equation (D.4)) implies that $y^\star_{1x} \in [0, \mathfrak{t} + \epsilon]$ with probability going to 1, we have

$$\begin{aligned} &\left| \sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \chi_{ix}(y^\star_{1x}) \{1 - G_x(y^\star_{1x})\}^{-1} F_x^{[1]}(y^\star_{1x}, y_{2x}) \right| \\ &\quad \le \frac{\sup_{y_1 \in [0, \mathfrak{t}+\epsilon]} \left| \sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \chi_{ix}(y_1) \right|}{1 - G_x(\mathfrak{t} + \epsilon)} \sup_{(y_1, y_2) \in [0, \mathfrak{t}+\epsilon] \times \mathbb{R}} F_x^{[1]}(y_1, y_2) \\ &\quad = O_{\mathbb{P}}(1)\,, \end{aligned}$$

where, to deduce the last equality, we used Condition $\mathcal{C}_2$. Thus, Equations (D.3) and (D.5) imply that

$$\sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \left\{ \mathcal{J}_{ix}^{(2)}(y_{1x}, y_{2x}) - \mathcal{J}_{ix}^{(2)}(y'_{1x}, y_{2x}) \right\} = o_{\mathbb{P}}(1).$$

Similarly, we have

$$\sqrt{nh} \sum_{i=1}^n w_{ni}(x, g) \left\{ \mathcal{J}_{ix}^{(2)}(y'_{1x}, y_{2x}) - \mathcal{J}_{ix}^{(2)}(y'_{1x}, y'_{2x}) \right\} = o_{\mathbb{P}}(1).$$

Therefore, Equation (D.2) can be deduced from the two last equations combined with Equation (D.4).

Appendix D.3. Asymptotic representation of $\mathbb{E}(\widetilde{A}'_x)$

From Equation (B.1), we can write

$$\mathbb{E}(\widetilde{A}'_x)(u,v) = \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)\left[F_{x_i}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u), \widehat{F}_{2x}^{-1}(v)\right\} - F_{x_i}\left\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\right\}\right].$$

Conditions $\mathcal{C}_1$, $\mathcal{C}_3$, W_2, W_3, W_5 and W_6 allow to mimic the proof of Theorem 1 in [22] to obtain, uniformly in $(u,v) \in \widetilde{\mathcal{T}}_{\mathfrak{t}}$, that

$$\begin{aligned}&\mathbb{E}(\widetilde{A}'_x)(u,v)\\ &= \sqrt{nh}\left(\mathcal{C}_x\left[F_{1x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\}, F_{2x}\left\{\widehat{F}_{2x}^{-1}(v)\right\}\right] - \mathcal{C}_x(u,v)\right) + o_{\mathbb{P}}(1).\end{aligned} \tag{D.6}$$

We next show that

$$\begin{aligned}\xi_{xh} &\equiv \sqrt{nh}\left[F_{1x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\} - u\right]\\ &= -\sqrt{nh}\left[\widehat{F}_{1x}^{\widehat{G}_x}\left\{F_{1x}^{-1}(u)\right\} - u\right] + o_{\mathbb{P}}(1).\end{aligned} \tag{D.7}$$

For this, note that from Equation (D.1) and Appendix D.1, we have that uniformly on $\widetilde{\mathcal{T}}_{\mathfrak{t}}$,

$$\sqrt{nh}(\mathcal{C}_x^{\widehat{G}_x} - \mathcal{C}_x) = \widetilde{B}_x + \mathbb{E}\left(\widetilde{A}_x\right) + \mathbb{E}\left(\widetilde{A}'_x\right) + o_{\mathbb{P}}(1). \tag{D.8}$$

As the fact that $Y_{2i} \leq \widehat{F}_{2x}^{(-1)}(1)$ for all $i \in \{1,\dots,n\}$ implies

$$\mathcal{C}_x^{\widehat{G}_x}(u,1) = \widehat{F}_{1x}^{\widehat{G}_x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\},$$

we deduce from (D.8) that

$$\begin{aligned}\xi_{xh} &= \sqrt{nh}\{\mathcal{C}_x^{\widehat{G}_x}(u,1) - \mathcal{C}_x(u,1)\}\\ &= \widetilde{B}_x(u,1) + \mathbb{E}\left\{\widetilde{A}_x(u,1)\right\} + \mathbb{E}\left\{\widetilde{A}'_x(u,1)\right\} + o_{\mathbb{P}}(1)\\ &= \sqrt{nh}\sum_{i=1}^{n}\left[w_{ni}(x,h)\mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(1)\} + w_{ni}(x,g)\mathcal{J}_{ix}^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(1)\}\right]\\ &\quad + \sqrt{nh}\left(\mathcal{C}_x\left[F_{1x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\}, F_{2x}\left\{\widehat{F}_{2x}^{-1}(1)\right\}\right] - u\right) + o_{\mathbb{P}}(1)\\ &= \sqrt{nh}\left[\widehat{F}_{1x}^{\widehat{G}_x}\left\{F_{1x}^{-1}(u)\right\} - u\right]\\ &\quad + \sqrt{nh}\left(\mathcal{C}_x\left[F_{1x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\}, F_{2x}\left\{\widehat{F}_{2x}^{-1}(1)\right\}\right] - u\right) + o_{\mathbb{P}}(1).\end{aligned} \tag{D.9}$$

To obtain the third equality, we used the definition of $\widetilde{B}_x$, Equations (D.2) and (D.6), while the last equation follows from Theorem 3.1.

Hence, to show Equation (D.7), we need to prove that

$$\sqrt{nh}\left[\widehat{F}_{1x}^{\widehat{G}_x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\}-u\right]=o_{\mathbb{P}}(1),\quad \text{uniformly in}\quad u\in[0,H_{1x}(\mathfrak{t})], \tag{D.10}$$

and

$$\sqrt{nh}\big(\mathcal{C}_x\left[z,F_{2x}\{\widehat{F}_{2x}^{-1}(1)\}\right]-z\}\big)=o_{\mathbb{P}}(1),\quad \text{uniformly in}\quad z\in[0,1]. \tag{D.11}$$

First recall from Appendix D.2 that for any $\epsilon>0$ satisfying $H_{1x}(\mathfrak{t}+\epsilon)<1$, we have with probability going to 1 that $\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\in[0,\mathfrak{t}+\epsilon]$ uniformly in $u\in[0,H_{1x}(\mathfrak{t})]$. Also, note that Equation (B.6) implies that, for sufficiently large n, $1-\widehat{G}_x(\mathfrak{t}+\epsilon)>1/2\{1-G_x(\mathfrak{t}+\epsilon)\}$ almost surely. Consequently, for sufficiently large n, we have

$$\begin{aligned}\sqrt{nh}\left|\widehat{F}_{1x}^{\widehat{G}_x}\left\{\left(\widehat{F}_{1x}^{\widehat{G}_x}\right)^{-1}(u)\right\}-u\right| &\le \sqrt{nh}\frac{\max_i w_{ni}(x,h)\mathbb{I}(T_{1i}\le\mathfrak{t}+\epsilon)}{1-\widehat{G}_x(T_{1i})}\\ &\le 2\sqrt{nh}\frac{\max_i w_{ni}(x,h)}{1-G_x(\mathfrak{t}+\epsilon)}\quad \text{a.s.}\end{aligned}$$

From there, (D.10) follows from Assumption W_1.

Secondly, from the proof of Theorem 1 in [22], we have, uniformly in $v\in[0,1]$, that

$$\sqrt{nh}\left[F_{2x}\left\{\widehat{F}_{2x}^{-1}(v)\right\}-v\right]=-\sqrt{nh}\left[\widehat{F}_{2x}\left\{F_{2x}^{-1}(v)\right\}-v\right]+o_{\mathbb{P}}(1). \tag{D.12}$$

Thus,

$$\sqrt{nh}\left[F_{2x}\left\{\widehat{F}_{2x}^{-1}(1)\right\}-1\right]=-\sqrt{nh}\left[\widehat{F}_{2x}\left\{F_{2x}^{-1}(1)\right\}-1\right]+o_{\mathbb{P}}(1)\ =o_{\mathbb{P}}(1).$$

Using the inequality $|\mathcal{C}_x(u,v)-\mathcal{C}_x(u',v')|\le|u-u'|+|v-v'|$ combined with the fact that for any $z\in[0,1]$ we have $\mathcal{C}_x(z,1)=z$, we found that

$$\sqrt{nh}\big|\mathcal{C}_x\left[z,F_{2x}\{\widehat{F}_{2x}^{-1}(1)\}\right]-z\big|\le\sqrt{nh}\big|F_{2x}\{\widehat{F}_{2x}^{-1}(1)\}-z\big|=o_{\mathbb{P}}(1),$$

which proves Equation (D.11) and therefore Equation (D.7) holds.

From there, following the end of the proof of Theorem 1 in [22], as assumption $(\mathcal{D})$ is satisfied, and in view of (D.7) and (D.12), we obtain that uniformly in $(u,v)\in\widetilde{\mathcal{T}}_{\mathfrak{t}}$:

$$\begin{aligned}\mathbb{E}(A_x')(u,v)=&-\sqrt{nh}\mathcal{C}_x^{[1]}(u,v)\left[\widehat{F}_{1x}^{\widehat{G}_x}\left\{F_{1x}^{-1}(u)\right\}-u\right]\\ &-\sqrt{nh}\mathcal{C}_x^{[2]}(u,v)\left[\widehat{F}_{2x}\left\{F_{2x}^{-1}(v)\right\}-v\right]+o_{\mathbb{P}}(1).\end{aligned}$$

Appendix D.4. Conclusion of the proof

Using Theorem 3.1, Equation (D.1) and collecting the results obtained in Appendix D.1 to Appendix D.3, we deduce that uniformly on $l^\infty(\widetilde{\mathcal{T}}_{\mathfrak{t}})$:

$$\sqrt{nh}(\widehat{\mathcal{C}}_x^{\hat{G}_x} - \mathcal{C}_x) = \sqrt{nh}\sum_{i=1}^{n}\{w_{ni}(x,h)\mathbb{j}_{ix}^{(1)} + w_{ni}(x,g)\mathbb{j}_{ix}^{(2)}\} + o_{\mathbb{P}}(1),$$

where

$$\begin{aligned}\mathbb{j}_{ix}^{(1)}(u,v) = \mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\} - \mathcal{C}_x^{[1]}(u,v)\mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(1)\}\\ - \mathcal{C}_x^{[2]}(u,v)[\mathbb{I}\{Y_{2i} \le F_{2x}^{-1}(v)\} - v]\end{aligned}$$

and

$$\mathbb{j}_{ix}^{(2)}(u,v) = \mathcal{J}_{ix}^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\} - \mathcal{C}_x^{[1]}(u,v)\mathcal{J}_{ix}^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(1)\}.$$

Recalling, from Appendix A.2, that $I_{nx} = \{i : w_{ni}(x,h) > 0\}$, we have, for any $i \in I_{nx}$, $(y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}$ and $y \in \mathbb{R}$, that

$$\begin{aligned}\mathrm{cov}[\mathcal{J}_{ix}^{(1)}(y_1,y_2), \mathbb{I}\{Y_{2i} \le y'\}] &= F_{x_i}(y_1, y \wedge y') - F_{x_i}(y_1,y_2)F_{2x}(y')\\ &= F_x(y_1, y \wedge y') - F_x(y_1,y_2)F_{2x}(y') + o(1),\end{aligned}$$

and

$$\begin{aligned}&\mathrm{cov}[\mathcal{J}_{ix}^{(2)}(y_1,y_2), \mathbb{I}\{Y_{2i} \le y'\}]\\ &\qquad = \int_0^{y_1} \frac{H_x(z,y') - H_{1x}(z)F_{2x}(y')}{\{1-H_{1x}(z)\}^2}\Delta_{F_x}(y_1,y_2,z)\,\mathrm{d}H_{1x}^c(z)\\ &\qquad\quad + \int_0^{y_1} \frac{H_x^c(z,y') - H_{1x}^c(z)F_{2x}(y')}{1-H_{1x}(z)}F_x^{[1]}(z,y_2)\,\mathrm{d}z\\ &\qquad\quad - \int_0^{y_1} \frac{H_x^c(z,y') - H_{1x}^c(z)F_{2x}(y')}{\{1-H_{1x}(z)\}^2}\Delta_{F_x}(y_1,y_2,z)\,\mathrm{d}H_{1x}(z) \; + o(1).\end{aligned}$$

As such, we have, for any $i \in I_{nx}$, $(u,v) \in \widetilde{\mathcal{T}}_{\mathfrak{t}}$ and $v' \in [0,1]$, that

$$\mathrm{cov}[\mathcal{J}_{ix}^{(1)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\}, \mathbb{I}\{Y_{2i} \le F_{2x}^{-1}(v)\}] = \mathcal{C}_x(u, v\wedge v') - \mathcal{C}_x(u,v)v' + o(1),$$

and using that $\Delta_{F_x}(y_1,y_2,z) = \mathcal{C}_x\{F_{1x}(y_1), F_{2x}(y_2)\} - \mathcal{C}_x\{F_{1x}(z), F_{2x}(y_2)\}$,

$$\begin{aligned}&\mathrm{cov}[\mathcal{J}_{ix}^{(2)}\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\}, \mathbb{I}\{Y_{2i} \le F_{2x}^{-1}(v')\}]\\ &= \int_0^{F_{1x}^{-1}(u)} \frac{H_x\{z, F_{2x}^{-1}(v')\} - H_{1x}(z)v'}{\{1-H_{1x}(z)\}^2}\left[\mathcal{C}_x(u,v) - \mathcal{C}_x\{F_{1x}(z), v\}\right]\mathrm{d}H_{1x}^c(z)\\ &\quad + \int_0^{F_{1x}^{-1}(u)} \frac{H_x^c\{z, F_{2x}^{-1}(v')\} - H_{1x}^c(z)v'}{1-H_{1x}(z)}F_x^{[1]}\{z, F_{2x}^{-1}(v)\}\,\mathrm{d}z\end{aligned}$$

$$- \int_0^{F_{1x}^{-1}(u)} \frac{H_x^c\{z, F_{2x}^{-1}(v')\} - H_{1x}^c(z)v'}{\{1 - H_{1x}(z)\}^2} \left[\mathcal{C}_x(u, v) - \mathcal{C}_x\{F_{1x}(z), v\}\right] \mathrm{d}H_{1x}(z) + o(1).$$

From there, the conclusion of the Proposition is obtained by using similar arguments as those exposed at the end of each of Appendix C.1 to Appendix C.3.

Appendix E. Auxiliary lemmas

The following lemma is required in Appendix B.1 to establish the asymptotic representation of $\mathbb{F}_x^{\widehat{G}_x}$.

Lemma E.1. *Recall the definition of $Z_x, \mathcal{F}$ and ρ_{F_x} at the beginning of Appendix B.1. Suppose Assumptions W_1, W_4 and W_5 are satisfied, and that the maps $z \mapsto F_{1z}$ and $z \mapsto F_{2z}$ are uniformly continuous for all z in a neighborhood of x. Then, process $\overline{Z}_x \equiv Z_x - \mathrm{E}Z_x$ indexed by $(\mathcal{F}, \rho_{F_x})$ is asymptotically ρ_{F_x}-equicontinuous in probability.*

Proof of Lemma E.1. Let $\|\cdot\|_{\mathcal{F}}$ stand for the supremum norm over $\mathcal{F}$. From Theorem 2.11.1 of [17], we can conclude that Z_x is tight if the following requirements hold:

- $\mathcal{R}_1$: $\sum_{i=1}^n \mathrm{E}(\|Z_{hi}\|_{\mathcal{F}}^2)\mathbb{I}\{\|Z_{hi}\|_{\mathcal{F}} > \eta\} \to 0$ for any $\eta > 0$;
- $\mathcal{R}_2$: $\sup_{\rho_{F_x}(f,f')<\delta_n} \sum_{i=1}^n \mathrm{E}[\{Z_{hi}(f) - Z_{hi}(f')\}^2] \to 0$ for every $\delta_n \downarrow 0$.
- $\mathcal{R}_3$: $\int_0^{\delta_n} \{\log N(\epsilon, \mathcal{F}, d_n)\}^{1/2}\, \mathrm{d}\epsilon \xrightarrow{\mathbb{P}} 0$ for every $\delta_n \downarrow 0$, where $N(\epsilon, \mathcal{F}, d_n)$ is the covering number of the set $\mathcal{F}$ with respect to the random semi-metric $d_n^2(f, f') = \sum_{i=1}^n \{Z_{hi}(f) - Z_{hi}(f')\}^2$.

Recall from Page 5065 that each $f \in \mathcal{F}$ can be formally identified by a triplet (y_1, y_2, G) with $(y_1, y_2) \in \mathcal{T}_{\mathfrak{t}}$ and $G \in \mathcal{G}_{\mathfrak{g}_\epsilon}$, where

$$\mathcal{G}_{t_\epsilon} = \{G : \mathbb{R} \to [0, 1] \text{ nondecreasing and } G(\mathfrak{t}) < \mathfrak{g}_\epsilon\}$$

and $\mathfrak{g}_\epsilon = G_x(\mathfrak{t}) + \epsilon < 1$. As such, throughout this proof, let $f = (y_1, y_2, G)$ and $f' = (y_1', y_2', G')$.

Let's prove that $\mathcal{R}_1$ to $\mathcal{R}_3$ hold.

Starting with $\mathcal{R}_1$, because $Z_{hi}(f) = 0$ whenever $Y_{1i} > t$ and G is nondecreasing, we have

$$Z_{hi}(f) \le \sqrt{nh} w_{ni}(x, h)\{1 - G(\mathfrak{t})\}^{-1} \le \sqrt{nh} w_{ni}(x, h)(1 - \mathfrak{g}_\epsilon)^{-1}.$$

In view of Assumption W_1, the latter is $o(1)$. Hence, for any $\eta > 0$, there exists a constant $N_\eta \ge 1$ such that $\max_{1\le i\le n} Z_{hi}(f) < \eta$ for any $n \ge N_\eta$, which proves that requirement $\mathcal{R}_1$ is fulfilled.

To show that $\mathcal{R}_2$ holds, we assume wlog that $y_1 \le y_1'$ and $y_2 \le y_2'$. Since $\mathbb{I}(T_{1i} \le y_1)\delta_{1i} \le \mathbb{I}(Y_{1i} \le y_1)$, we have

$$Z_{hi}(f') - Z_{hi}(f)$$

$$= \sqrt{nh}\left\{\mathbb{I}(T_{1i} \le y_1', Y_{2i} \le y_2') - \mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2)\right\} \frac{\delta_{1i} w_{ni}(x,h)}{1 - G'(T_{1i})}$$
$$+ \sqrt{nh}\mathbb{I}(T_{1i} \le y_1, Y_{2i} \le y_2)\delta_{1i} w_{ni}(x,h) \frac{G(T_{1i}) - G'(T_{1i})}{\{1 - G(T_{1i})\}\{1 - G'(T_{1i})\}}$$
$$\le \sqrt{nh}\frac{w_{ni}(x,h)}{(1-\mathfrak{g}_\epsilon)^2}\Big[\{\mathbb{I}(Y_{1i} \le y_1', Y_{2i} \le y_2') - \mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2)\} + \|G - G'\|_{[0,\mathfrak{t}]}\Big].$$

From there, and using the inequality $(a+b)^2 \le 2(a^2+b^2)$, we deduce that

$$(nh)^{-1}\sum_{i=1}^n \mathrm{E}\left[\{Z_{hi}(f') - Z_{hi}(f)\}^2\right]$$
$$\le 2\sum_{i=1}^n \frac{w_{ni}(x,h)^2}{(1-\mathfrak{g}_\epsilon)^4}\mathrm{E}\Big[\{\mathbb{I}(Y_{1i} \le y_1', Y_{2i} \le y_2') - \mathbb{I}(Y_{1i} \le y_1, Y_{2i} \le y_2)\}^2 + \|G - G'\|^2_{[0,\mathfrak{t}]}\Big]$$
$$\le 2\sum_{i=1}^n \frac{w_{ni}(x,h)^2}{(1-\mathfrak{g}_\epsilon)^4}\Big\{|F_{1x_i}(y_1') - F_{1x_i}(y_1)| + |F_{2x_i}(y_2') - F_{2x_i}(y_2)| + \|G - G'\|^2_{[0,\mathfrak{t}]}\Big\}.$$

As Assumption W_5 holds, and since the maps $z \mapsto F_{1z}$ and $z \mapsto F_{2z}$ are assumed uniformly continuous, we deduce that $\max_{i \in I_{nx}} |F_{jx_i}(y_j') - F_{jx_i}(y_j)| = |F_{jx}(y_j') - F_{jx}(y_j)| + o(1)$, $j = 1, 2$. Hence, using the above inequality, we obtain that

$$\sup_{\rho_{F_x}(f,f') < \delta_n} \sum_{i=1}^n \mathrm{E}\left[\{Z_{hi}(f') - Z_{hi}(f)\}^2\right] \le 2\Big\{\sum_{i=1}^n \frac{w_{ni}(x,h)^2}{(1-\mathfrak{g}_\epsilon)^4}\Big\}\{\delta_n + o(1)\}.$$

From Assumption W_4, we deduce that the latter display is bounded by $O(1)\{\delta_n + o(1)\}$. Hence, requirement $\mathcal{R}_2$ is fulfilled as $\delta_n \to 0$.

To show $\mathcal{R}_3$, our goal is to apply Lemma 2.11.6 of [17]. First, we rewrite

$$\{Z_{hi}(f') - Z_{hi}(f)\}^2 = \int \left\{\frac{\mathbb{I}(v \le y_1, v' \le y_2)}{1 - G(v)} - \frac{\mathbb{I}(v \le y_1', v' \le y_2')}{1 - G'(v)}\right\}^2 \mathrm{d}\mu_{ni},$$

where $\mu_{ni} = nhw_{ni}(x,h)^2\delta_{1i}\mathbb{I}(v = Y_{1i}, v' = Y_{2i})$. Hence, the process Z_n is *measurelike* with respect to the random measure μ_{ni} (see [17], Section 2.11). Secondly, as Assumption W_4 holds, $\sum_{i=1}^n \int \frac{1}{1-\mathfrak{g}_\epsilon}\mathrm{d}\mu_{ni} = nh\sum_{i=1}^n \frac{w_{ni}(x,h)^2}{1-\mathfrak{g}_\epsilon} = O(1)$. Thirdly, it is required to show that the class $\mathcal{F}$ satisfies the uniform entropy condition (2.11.5) of [17]. To prove that it is the case, let

$$\mathcal{F}_1 = \{(v, v', w) \mapsto \mathbb{I}\{v \le y_1, v' \le y\}, (y_1, y_2) \in \mathcal{T}_\mathfrak{t}, w \in \{0, 1\}\}$$

and let $\mathcal{F}_2$ be the class of monotone and bounded functions over $[0, \frac{1}{1-\mathfrak{g}_\epsilon}]$. Now we observe that $\mathcal{F} \subset \mathcal{F}_1\mathcal{F}_2 = \{f = f_1 f_2, f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2\}$. As $\mathcal{F}_1$ is a VC-class and $\mathcal{F}_2$ is a VC-hull class for sets with envelope functions respectively $\mathfrak{F}_1 = 1$ and $\mathfrak{F}_2 = \frac{1}{1-\mathfrak{g}_\epsilon}$, an application Lemma 2.6.20 of [17] allows to conclude that $\mathcal{F}_1\mathcal{F}_2$ is VC-hull class for sets with envelope function $\mathfrak{F}_1 \times \mathfrak{F}_2 = \mathfrak{F}$. Therefore, using Corollary 2.6.12 of [17], we conclude that the uniform entropy condition is fulfilled. As a result, Lemma 2.11.6 of [17] applies which proves the requirement $\mathcal{R}_3$, which concludes the proof of Lemma E.1. □

From Theorem 1 in [21], the process

$$\{\sqrt{ng}\sum_{i=1}^{n} w_{ni}(x,g)[\chi_{ix}(y_1) - \mathrm{E}\{\chi_{ix}(y_1)\}], y_1 \in [0,\mathfrak{t}]\}$$

is asymptotically tight in $D[0,\mathfrak{t}]$, the space of right continuous functions with left hand limits, endowed with Skorohod topology. The next lemma states that the same result holds in $l^\infty([0,\mathfrak{t}])$ equipped with the uniform topology.

Lemma E.2. *Assume that Conditions W_1 to W_5 and $(\mathcal{C}_1)$ to $(\mathcal{C}_3)$ are satisfied. Then, with χ_{ix} as in Equation (B.8), the process $\mathrm{g}_x \equiv \sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\chi_{ix} - \mathrm{E}(\chi_{ix})\}$ is asymptotically tight in $l^\infty([0,\mathfrak{t}])$.*

Proof of Lemma E.2. From the definition of χ_{ix} at Equation (B.8), we have

$$\mathrm{g}_x(y_1) - \mathrm{E}\{\mathrm{g}_x(y_1)\} = \{1 - G_x(y_1)\}\{\widetilde{\mathrm{g}}_x^{(1)}(y_1) - \widetilde{\mathrm{g}}_x^{(2)}(y_1) + \widetilde{\mathrm{g}}_x^{(3)}(y_1)\},$$

where for $j = 1,2,3$, $\widetilde{\mathrm{g}}_x^{(j)} = \sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\widetilde{g}_{ix}^{(j)}$ with

$$\widetilde{g}_{ix}^{(1)}(y_1) = \int_0^{y_1} \{\mathbb{I}(T_{1i} \le v) - H_{1x_i}(v)\}\{1 - H_{1x}(v)\}^{-2}\,\mathrm{d}H_{1x}^c(v),$$

$$\widetilde{g}_{ix}^{(2)}(y_1) = \int_0^{y_1} \{\mathbb{I}(T_{1i} \le v, \delta_{1i} = 0) - H_{1x_i}^c(v)\}\{1 - H_{1x}(v)\}^{-2}\,\mathrm{d}H_{1x}(v),$$

$$\widetilde{g}_{ix}^{(3)}(y_1) = \{\mathbb{I}(T_{1i} \le y_1, \delta_{1i} = 0) - H_{1x_i}^c(y_1)\}\{1 - H_{1x}(y_1)\}^{-1}.$$

Because G_x is bounded, we conclude that $\mathrm{g}_x - \mathrm{E}(\mathrm{g}_x)$ is asymptotically tight by showing that $\widetilde{\mathrm{g}}_x^{(j)}$ is asymptotically tight for $j = 1,2,3$.

We start with $\widetilde{\mathrm{g}}_x^{(3)}$. Mimicking the proof of Lemma 2 in [22], and in view of the connection between asymptotic tightness and asymptotic equicontinuity in $l^\infty([0,\mathfrak{t}])$ (see Theorem 1.5.7 in [17]), we can show that the process $\sqrt{ng}\sum_{i=1}^n w_{ni}(x,g)\{\mathbb{I}(T_{1i} \le y_1, \delta_{1i} = 0) - H_{1x_i}^c(y_1)\}$ is asymptotically tight in $l^\infty([0,\mathfrak{t}])$. Therefore, the asymptotic tightness of $\widetilde{\mathrm{g}}_x^{(3)}$ in $l^\infty([0,\mathfrak{t}])$ is ensured by the fact that the map $y_1 \mapsto \{1 - H_{1x}(y_1)\}^{-1}$ is bounded and continuous on $[0,\mathfrak{t}]$.

To prove the asymptotic tightness of $\widetilde{\mathrm{g}}_x^{(1)}$, we introduce the set

$$\mathcal{H} \equiv \left\{\mathbb{f}_{y_1} : \mathbb{R} \mapsto \mathbb{R} : \mathbb{f}_{y_1}(y_1^\star) \equiv \int_0^{y_1} \frac{\mathbb{I}(y_1^\star \le v)}{\{1 - H_{1x}(v)\}^2}\,\mathrm{d}H_{1x}^c(v), y_1 \in [0,\mathfrak{t}]\right\},$$

which allows us to write $\widetilde{\mathrm{g}}_x^{(1)}(y_1) = \sum_{i=1}^n [z_{hi}(\mathbb{f}_{y_1}) - \mathrm{E}\{z_{ni}(\mathbb{f}_{y_1})\}]$, where $z_{hi}(\mathbb{f}) = \sqrt{ng}w_{ni}(x,g)\mathbb{f}(T_{1i})$ for any $\mathbb{f} \in \mathcal{H}$. Now, let us equip $\mathcal{H}$ with a semimetric ρ defined as

$$\rho(\mathbb{f}_{y_1}, \mathbb{f}_{y_1'}) = |y_1 - y_1'| \quad \text{for} \quad \mathbb{f}_{y_1}, \mathbb{f}_{y_1'} \in \mathcal{H}.$$

Next verify that the sequence $z_{h1}, \ldots, z_{hn}$ indexed by $(\mathcal{H}, \rho)$ satisfies

(a) $\sum_{i=1}^n \mathrm{E}\left(\|z_{hi}\|^2_{\mathcal{H}}\right) \mathbb{I}\left\{\|z_{hi}\|_{\mathcal{H}} > \eta\right\} \to 0$ for any $\eta > 0$;
(b) $\sup_{|y_1 - y_1'| < \delta_n} \sum_{i=1}^n \mathrm{E}[\{z_{hi}(\mathbb{f}_{y_1}) - z_{hi}(\mathbb{f}_{ty_1})\}^2] \to 0$ for every $\delta_n \downarrow 0$
(c) $\int_0^{\delta_n} \{\log N(\epsilon, \mathcal{H}, d_n)\}^{1/2}\, \mathrm{d}\epsilon \xrightarrow{\mathbb{P}} 0$ for every $\delta_n \downarrow 0$, where $N(\epsilon, \mathcal{H}, d_n)$ is the covering number of the set $\mathcal{H}$ with respect to the random semi-metric $d_n^2(\mathbb{f}_{y_1}, \mathbb{f}_{y_1'}) = \sum_{i=1}^n \{z_{hi}(\mathbb{f}_{y_1}) - z_{hi}(\mathbb{f}_{y_1'})\}^2$.

Note that (a) is immediate, as for any $y_1 \in [0, \mathfrak{t}]$ we have $0 \le \mathbb{f}_{y_1} \le \mathbb{f}_{\mathfrak{t}} < \infty$, and therefore from Assumption W_1, we have $|z_{hi}| \le o(1)$.

To prove (b), we note that for $y_1, y_1' \in [0, \mathfrak{t}]$ with $y_1 \le y_1'$, we have

$$\begin{aligned}\sum_{i=1}^n \mathrm{E}\left[\{z_{hi}\ (\mathbb{f}_{y_1}) - z_{hi}(\mathbb{f}_{y_1'})\}^2\right] \\ &= ng \sum_{i=1}^n w_{ni}^2(x, g) \mathrm{E}\left(\left[\int_{y_1}^{y_1'} \mathbb{I}(T_{1i} \le v)\{1 - H_{1x}(v)\}^{-2}\, \mathrm{d}H_{1x}^c(v)\right]^2\right) \\ &\le (y_1 - y_1')^2 ng \sum_{i=1}^n w_{ni}^2(x, g) \times \kappa_{\mathfrak{t}},\end{aligned}$$

where $\kappa_{\mathfrak{t}} = 2\{1 - H_{1x}(\mathfrak{t})\}^{-2} \sup_{v \in [0,\mathfrak{t}]}\{H_{1x}^{c[1]}(v)\}$. Therefore, from Assumption W_4, the fact that $\sup_{v \in [0,\mathfrak{t}]}\{H_{1x}^{c[1]}(v)\} < \infty$ (since H_{1x}^c fulfils Condition $(\mathcal{C}_2)$), we conclude that (b) holds.

To prove (c), we show that for sufficiently large n, $N(\epsilon, \mathcal{H}, d_n) \le \sqrt{2K_4\kappa_{\mathfrak{t}}}\mathfrak{t} \times \epsilon^{-1}$. Indeed, from Assumption W_4 we have for sufficiently large n,

$$ng \sum_{i=1}^n w_{ni}^2(x, g) \le 2K_4\,.$$

Now, following similar calculation in the proof of (b), we have, for any $y_1 \le y_1'$ and sufficiently large n, that

$$\begin{aligned}d_n^2(\mathbb{f}_{y_1}, \mathbb{f}_{y_1'}) &= \sum_{i=1}^n \{z_{hi}(\mathbb{f}_{y_1}) - z_{hi}(\mathbb{f}_{y_1'})\}^2 \\ &\le (y_1 - y_1')^2 ng \sum_{i=1}^n w_{ni}^2(x, g)\kappa_{\mathfrak{t}} \le (y_1 - y_1')^2 2K_4\kappa_{\mathfrak{t}}\,.\end{aligned}$$

Then, we have $N(\epsilon, \mathcal{H}, d_n) \le N\{\epsilon/\sqrt{2K_4\kappa_{\mathfrak{t}}}, [0, \mathfrak{t}], |\cdot|\} = \sqrt{2K_4\kappa_{\mathfrak{t}}}\mathfrak{t}\epsilon^{-1}$. Therefore, we get

$$\int_0^{\delta_n} \{\log N(\epsilon, \mathcal{H}, d_n)\}^{1/2}\, \mathrm{d}\epsilon \le \int_0^{\delta_n} \{\log(\sqrt{2K_4\kappa_{\mathfrak{t}}}\mathfrak{t}\epsilon^{-1})\}^{1/2} \to 0 \quad \text{as } n \to \infty.$$

This concludes the proof of the asymptotic tightness of $\widetilde{\mathbb{g}}_x^{(1)}$ in $l^\infty([0, \mathfrak{t}])$. □

Similarly we can show that $\widetilde{\mathbb{g}}_x^{(2)}$ is also asymptotically tight in $l^\infty([0, \mathfrak{t}])$. This concludes the proof of Lemma E.2.

Appendix F. Auxiliary computations

Appendix F.1. Equivalence between the conditional Kaplan-Meier estimator of F_{1x} and $\widehat{F}_{1x}^{\widehat{G}_x}$ when $g = h$

In this section, we show that in the case where the bandwidths g and h in the definition of $\widehat{F}_{1x}^{\widehat{G}_x}$ are equal, then $\widehat{F}_{1x}^{\widehat{G}_x}$ coincide with the conditional Kaplan-Meier estimator of F_{1x} defined as

$$\begin{aligned} F_{1xh}^{(\mathrm{KM})}(y_1) =& :1 - \left[\prod_{T_{(i)} \le y_1} \left\{ 1 - \frac{w_{n[i]}(x,h)}{1 - \widehat{F}_{1x}(T_{(i)}-)} \right\}^{\delta_{[i]}} \right] \\ =& 1 - \left[\prod_{T_{(i)} \le y_1} \left\{ 1 - \frac{w_{n[i]}(x,h)}{1 - \sum_{k=1}^{i-1} w_{n[k]}(x,h)} \right\}^{\delta_{[i]}} \right], \end{aligned} \tag{F.1}$$

where we recall that $T_{(1)} < \ldots < T_{(n)}$ are the ordered T_{1i}'s, and $w_{n[i]}(x,h) = w_{nj}(x,h)$ and $\delta_{[i]} = \delta_j$ when j satisfies $T_j = T_{(i)}$.

Set $T_{(0)} = 0$ and $T_{n+1} = \infty$. In oder to prove that $F_{1xh}^{(\mathrm{KM})} = \widehat{F}_{1x}^{\widehat{G}_x}$, it suffices to show that for any $0 \le k \le n+1$, $\widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)}) = F_{1xh}^{(\mathrm{KM})}(T_{(k)})$. For this, we use the following induction argument.

Basis: Trivially, $\widehat{F}_{1x}^{\widehat{G}_x}(0) = F_{1xh}^{(\mathrm{KM})}(0) = 0$. Moreover,

$$\widehat{F}_{1x}^{\widehat{G}_x}(T_{(1)}) = \frac{\delta_{[1]} w_{n[1]}(x,h)}{\left\{1 - w_{n[1]}(x,h)\right\}^{1-\delta_{[1]}}} = \delta_{[1]} w_{n[1]}(x,h)$$

and

$$F_{1xh}^{(\mathrm{KM})}(T_{(1)}) = 1 - \left\{1 - w_{n[1]}(x,h)\right\}^{\delta_{[1]}} = \delta_{[1]} w_{n[1]}(x,h).$$

Hence, the Basis step is verified.

Induction hypothesis: The equality $\widehat{F}_{1x}^{\widehat{G}_x}(y_1) = F_{1xh}^{(\mathrm{KM})}(y_1)$ holds for $y_1 = T_{(0)}$ up to $y_1 = T_{(k)}$.

Induction step: Let's now show that the equality is verified for $y_1 = T_{(k+1)}$ assuming the induction hypothesis. From direct computations,

$$\begin{aligned} F_{1xh}^{(\mathrm{KM})}(T_{(k+1)}) &= 1 - \prod_{i=1}^{k} \left\{ 1 - \frac{w_{n[i]}(x,h)}{1 - \widehat{F}_{1x}(T_{(i-1)})} \right\}^{\delta_{[i]}} \left\{ 1 - \frac{w_{n[k+1]}(x,h)}{1 - \widehat{F}_{1x}(T_{(k)})} \right\}^{\delta_{[k+1]}} \\ &= 1 + \{\widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)}) - 1\} \left(1 - \frac{w_{n[k+1]}(x,h)}{1 - \widehat{F}_{1x}(T_{(k)})} \right)^{\delta_{[k+1]}}, \end{aligned}$$

where the latter equation follows from the induction hypothesis. If $\delta_{[k+1]} = 0$, we use the induction hypothesis to obtain

$$F_{1xh}^{(\mathrm{KM})}(T_{(k+1)}) = F_{1xh}^{(\mathrm{KM})}(T_{(k)}) = \widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)}) = \widehat{F}_{1x}^{\widehat{G}_x}(T_{(k+1)}).$$

Otherwise, if $\delta_{[k+1]} = 1$, then

$$\begin{aligned} F_{1xh}^{(\mathrm{KM})}(T_{(k+1)}) &= \widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)}) + \{1 - \widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)})\} \left(\frac{w_{n[k+1]}(x,h)}{1 - \widehat{F}_{1x}(T_{(k+1)})} \right) \\ &= \widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)}) + \{1 - F_{1xh}^{(\mathrm{KM})}(T_{(k)})\} \left(\frac{w_{n[k+1]}(x,h)}{1 - \widehat{F}_{1x}(T_{(k+1)})} \right). \end{aligned}$$

From the identity $\{1 - F_{1xh}^{(\mathrm{KM})}(T_{(k)})\}\{1 - \widehat{G}_{1x}(T_{(k)})\} = 1 - \widehat{F}_{1x}(T_{(k)})$, we deduce that

$$F_{1xh}^{(\mathrm{KM})}(T_{(k+1)}) = \widehat{F}_{1x}^{\widehat{G}_x}(T_{(k)}) + \frac{w_{n[k+1]}(x,h)}{1 - \widehat{G}_{1x}(T_{(k)})}.$$

The proof follows from the fact that $\widehat{G}_{1x}(T_{(k)}) = \widehat{G}_{1x}(T_{(k+1)})$ since $\delta_{[k+1]} = 1$.

Appendix F.2. Equivalence between the asymptotic representation of the conditional Kaplan-Meier estimator and that of $\widehat{F}_{1x}^{\widehat{G}_x}$ when $g = h$

We next show that the asymptotic representation of $\lim_{y\to\infty} \mathbb{F}_x^{\widehat{G}_x}(y_1, y_2)$ in the case $g = h$ as stated in Theorem 3.1 is the same as the one of $\sqrt{nh}(F_{1xh}^{(KM)} - F_{1x})$ derived in Theorem 2.1 of [20]. In fact, [20] stated that, $\sqrt{nh}(F_{1xh}^{(KM)} - F_{1x}) = \sqrt{nh}\sum_{j=1}^n w_{ni}(x,h)\{\widetilde{\mathcal{J}}_{ix} - \widetilde{\mathcal{J}}_x\} + o_{\mathrm{a.s}}(1)$, where

$$\begin{aligned} \widetilde{\mathcal{J}}_{ix}(y_1) = \{1 - F_{1x}(y_1)\}\times &\\ \Bigg[\int_0^{y_1} \frac{\mathbb{I}(T_{1i} \le v)\,\mathrm{d}H_{1x}^u(v)}{\{1 - H_{1x}(v)\}^2} &+ \frac{\mathbb{I}(T_{1i} \le y_1, \delta_{1i} = 1)}{1 - H_{1x}(y_1)} \\ &- \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \le v, \delta_{1i} = 1)\,\mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \Bigg], \end{aligned}$$

and

$$\begin{aligned} \widetilde{\mathcal{J}}_x(y_1) = \{1 - F_{1x}(y_1)\} &\\ \times \Bigg[\int_0^{y_1} \frac{H_{1x}(v)\,\mathrm{d}H_{1x}^u(v)}{\{1 - H_{1x}(v)\}^2} &+ \frac{H_{1x}^u(y_1)}{1 - H_{1x}(y_1)} \\ &- \int_0^{y_1} \frac{H_{1x}^u(v)\,\mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \Bigg]. \end{aligned}$$

To show the result, we write $\lim_{y\to\infty}\{\mathcal{J}_{ix}^{(1)}(y_1,y_2)+\mathcal{J}_{ix}^{(2)}(y_1,y_2)\}=\check{\mathcal{J}}_{ix}(y_1)+\check{\mathcal{J}}_{x}(y_1)$, where

$$\begin{aligned}\check{\mathcal{J}}_{ix}(y_1) = \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\Delta_{F_x}(y_1,\infty,v)}{\{1-H_{1x}(v)\}^2}\,\mathrm{d}H_{1x}^c(v) \\ + \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v,\delta_{1i}=0)\,\mathrm{d}H_{1x}^u(v)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}} \\ - \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v,\delta_{1i}=0)\Delta_{F_x}(y_1,\infty,v)}{\{1-H_{1x}(v)\}^2}\mathrm{d}H_{1x}(v) \\ + \frac{\mathbb{I}\{T_{1i}\le y_1,\delta_{1i}=1\}}{1-G_x(T_{1i})},\end{aligned} \tag{F.2}$$

and

$$\begin{aligned}\check{\mathcal{J}}_{x}(y_1) = \int_0^{y_1} \frac{H_{1x}^u(v)\Delta_{F_x}(y_1,\infty,v)}{\{1-H_{1x}(v)\}^2}\,\mathrm{d}H_{1x}^c(v) \\ + \int_0^{y_1} \frac{H_{1x}^c(v)\,\mathrm{d}H_{1x}^u(v)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}} \\ - \int_0^{y_1} \frac{H_{1x}^c(v)\Delta_{F_x}(y_1,\infty,v)}{\{1-H_{1x}(v)\}^2}\,\mathrm{d}H_{1x}(v)\ + F_{1x}(y_1).\end{aligned}$$

We start by showing that $\check{\mathcal{J}}_{ix}=\widetilde{\mathcal{J}}_{ix}$. First, as $1-H_{1x}(v)=\{1-F_{1x}(v)\}\{1-G_x(v)\}$, and since $H_{1x}=H_{1x}^u+H_{1x}^c$, we obtain that

$$\begin{aligned}&\int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\Delta_{F_x}(y_1,\infty,v)}{\{1-H_{1x}(v)\}^2}\,\mathrm{d}H_{1x}^c(v) \\ &\quad = -\{1-F_{1x}(y_1)\}\int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}^c(v)}{\{1-H_{1x}(v)\}^2} \\ &\qquad + \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\{1-F_{1x}(v)\}\,\mathrm{d}H_{1x}^c(v)}{\{1-H_{1x}(v)\}^2} \\ &\quad = -\{1-F_{1x}(y_1)\}\int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}^c(v)}{\{1-H_{1x}(v)\}^2} \\ &\qquad + \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}^c(v)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}} \\ &\quad = -\{1-F_{1x}(y_1)\}\int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}(v)}{\{1-H_{1x}(v)\}^2} \\ &\qquad + \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}(v)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}} \\ &\qquad + \{1-F_{1x}(y_1)\}\int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}^u(v)}{\{1-H_{1x}(v)\}^2} \\ &\qquad + \int_0^{y_1} \frac{\mathbb{I}(T_{1i}\le v)\,\mathrm{d}H_{1x}^u(v)}{\{1-H_{1x}(v)\}\{1-G_x(v)\}}\,.\end{aligned}$$

Second, using the fact that $\mathbb{I}(T_{1i} \leq v, \delta_{1i} = 0) = \mathbb{I}(T_{1i} \leq v) - \mathbb{I}(T_{1i} \leq v, \delta_{1i} = 1)$, $H^u_{1x}(v) = \int_0^v \{1 - G_x(t-)\}\, \mathrm{d}F_{1x}(t)$, and $1 - H_{1x}(v) = \{1 - F_{1x}(v)\}\{1 - G_x(v)\}$, we obtain by integrating by parts that,

$$
\begin{aligned}
&\int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v, \delta_{1i} = 0)\, \mathrm{d}H^u_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}} \\
&\quad = \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v) - \mathbb{I}(T_{1i} \leq v, \delta_{1i} = 1)\, \mathrm{d}H^u_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}} \\
&\quad = \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v)\, \mathrm{d}H^u_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}} + \frac{\mathbb{I}\{T_{1i} \leq y_1, \delta_{1i} = 1\}}{1 - G_x(y_1)} \\
&\qquad\quad - \frac{\mathbb{I}\{T_{1i} \leq y_1, \delta_{1i} = 1\}}{1 - G_x(T_{1i})} - \int_0^{y_1} \frac{\mathbb{I}\{T_{1i} \leq v, \delta_{1i} = 1\}\, \mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \\
&\quad = \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v)\, \mathrm{d}H^u_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}} + \{1 - F_{1x}(y_1)\} \frac{\mathbb{I}\{T_{1i} \leq y_1, \delta_{1i} = 1\}}{1 - H_{1x}(y_1)} \\
&\qquad\quad - \frac{\mathbb{I}\{T_{1i} \leq y_1, \delta_{1i} = 1\}}{1 - G_x(T_{1i})} - \int_0^{y_1} \frac{\mathbb{I}\{T_{1i} \leq v, \delta_{1i} = 1\}\, \mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}}.
\end{aligned}
$$

Third, using similar arguments, with the third term on the right hand side of Equation (F.2), we found

$$
\begin{aligned}
&\int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v, \delta_{1i} = 0) \Delta_{F_x}(y_1, \infty, v)}{\{1 - H_{1x}(v)\}^2} \mathrm{d}H_{1x}(v) \\
&= \{1 - F_{1x}(y_1)\} \left[\int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v)\, \mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} - \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v, \delta_{1i} = 1)\, \mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}^2} \right] \\
&\qquad\qquad - \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v)\, \mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}} \\
&\qquad\qquad\qquad + \int_0^{y_1} \frac{\mathbb{I}(T_{1i} \leq v, \delta_{1i} = 1)\, \mathrm{d}H_{1x}(v)}{\{1 - H_{1x}(v)\}\{1 - G_x(v)\}}.
\end{aligned}
$$

Plugging these new expressions into Equation (F.2) demonstrates that $\check{\mathcal{J}}_{ix} = \widetilde{\mathcal{J}}_{ix}$. The proof that $\check{\mathcal{J}}_x = \widetilde{\mathcal{J}}_x$ is similar.

Acknowledgements

This work has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC 402521-2013) and Le Fond de recherche du Québec (Nature et technologies) FRQNT 190250?-2016. This research was also undertaken thanks in part to funding from the Canada First Research Excellence Fund and to funding from Australian Research Council Discovery Project DP140100125.

References

[1] Abegaz, F., Gijbels, I. and Veraverbeke, N. (2012). Semiparametric estimation of conditional copulas. *J. Multivariate Anal.* **110** 43–73. MR2927509

[2] Acar, E. F., Craiu, R. V. and Yao, F. (2011). Dependence calibration in conditional copulas: a nonparametric approach. *Biometrics* **67** 445–453. MR2829013

[3] Akritas, M. G. (1994). Nearest neighbor estimation of a bivariate distribution under random censoring. *Ann. Statist.* **22** 1299–1327. MR1311977

[4] Akritas, M. G. and Van Keilegom, I. (2003). Estimation of bivariate and marginal distributions with censored data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65** 457–471. MR1983758

[5] Beran, R. (1981). Nonparametric regression with randomly censored survival data Technical Report, Univ. California, Berkeley.

[6] Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *Ann. Statist.* **16** 1475–1489. MR0964934

[7] Geerdens, C., Acar, E. F. and Janssen, P. (2017). Conditional copula models for right-censored clustered event time data. *Biostatistics* **19** 247–262. MR3799615

[8] Gijbels, I., Veraverbeke, N. and Omelka, M. (2011). Conditional copulas, association measures and their applications. *Comput. Statist. Data Anal.* **55** 1919–1932. MR2765054

[9] Gribkova, S. and Lopez, O. (2015). Non-parametric copula estimation under bivariate censoring. *Scand. J. Stat.* **42** 925–946. MR3426303

[10] Jondeau, E. and Rockinger, M. (2006). The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance* **25** 827–853.

[11] Loprinzi, C. L., Laurie, J. A., Wieand, H. S., Krook, J. E., Novotny, P. J., Kugler, J. W., Bartel, J., Law, M., Bateman, M. and Klatt, N. E. (1994). Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. *Journal of Clinical Oncology* **12** 601–607.

[12] Omelka, M., Veraverbeke, N. and Gijbels, I. (2013). Bootstrapping the conditional copula. *J. Statist. Plann. Inference* **143** 1–23. MR2969007

[13] Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *Internat. Econom. Rev.* **47** 527–556. MR2216591

[14] Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51** 1384–1399. MR1381050

[15] Stone, C. J. (1977). Consistent nonparametric regression. *Ann. Statist.* **5** 595–645. With discussion and a reply by the author. MR0443204

[16] van der Laan, M. J. and Robins, J. M. (2003). *Unified methods for censored longitudinal data and causality. Springer Series in Statistics.* Springer-Verlag, New York. MR1958123

[17] van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence*

and empirical processes. *Springer Series in Statistics*. Springer-Verlag, New York With applications to statistics. MR1385671

[18] van der Vaart, A. W. and Wellner, J. A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: particles, processes and inverse problems*. *IMS Lecture Notes Monogr. Ser.* **55** 234–252. Inst. Math. Statist., Beachwood, OH. MR2459942

[19] Van Keilegom, I. and Veraverbeke, N. (1996). Uniform strong convergence results for the conditional Kaplan-Meier estimator and its quantiles. *Comm. Statist. Theory Methods* **25** 2251–2265. MR1411556

[20] Van Keilegom, I. and Veraverbeke, N. (1997a). Estimation and bootstrap with censored data in fixed design nonparametric regression. *Ann. Inst. Statist. Math.* **49** 467–491. MR1482368

[21] Van Keilegom, I. and Veraverbeke, N. (1997b). Weak convergence of the bootstrapped conditional Kaplan-Meier process and its quantile process. *Comm. Statist. Theory Methods* **26** 853–869. MR1436082

[22] Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Stat.* **38** 766–780. MR2859749

[23] Wang, W. and Wells, M. T. (2000). Estimation of Kendall's tau under censoring. *Statist. Sinica* **10** 1199–1215. MR1804550