Journal of Statistical Planning and Inference 171 (2016) 10–37

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Asymptotic representation of presmoothed Kaplan–Meier integrals with covariates in a semiparametric censorship model

Gerhard Dikta [a,b], René Külheim [b,a], Jorge Mendonça [c], Jacobo de Uña-Álvarez [d,*]

[a] Fachhochschule Aachen, Abt. Jülich, Germany
[b] Department of Mathematical Sciences, University of Wisconsin, Milwaukee, USA
[c] Polytechnic Institute of Oporto, School of Engineering, Portugal
[d] Department of Statistics and Operations Research & Biomedical Research Centre (CINBIO), University of Vigo, Spain

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form 2 December 2015
Accepted 2 December 2015
Available online 14 December 2015

Keywords:
Censored data
Central limit theorem
Presmoothing
Semiparametric model
Survival Analysis

ABSTRACT

Presmoothed Kaplan–Meier integrals have been proposed as suitable estimators in semiparametric censorship models. They are based on a modification of Kaplan–Meier weights which replaces the censoring indicators by some smooth (parametric) fit to the conditional probability of uncensoring, leading to estimators with smaller variance. In this paper an asymptotic representation of these estimators as a sum of i.i.d. random variables is established. The situation in which covariates are present is considered; therefore, the present paper extends previous results in Dikta et al. (2005) to the setting with covariates. As a consequence, a CLT for presmoothed Kaplan–Meier integrals with covariates is obtained. Application to censored regression is given. The finite sample performance of the estimator is investigated through simulations.

1. Introduction

In Survival Analysis and other fields, the variable of interest Y is a lifetime which is observed under right-censoring. Therefore, rather than Y one observes (Z, δ), where $Z = \min(Y, C)$ is the recorded (possibly censored) lifetime, $\delta = 1_{\{Y \le C\}}$ is a censoring indicator, and C is the potential censoring time. Often, a p-dimensional vector of covariates X is attached to each individual. Estimation of the expectation $E[\varphi(X, Y)]$ for a general transformation φ is of interest; in particular, this allows for the estimation of regression coefficients in e.g. linear censored regression. Given a random sample (X_i, Z_i, δ_i) , $i = 1, \ldots, n$, Stute (1993) proposed the following estimator of $E[\varphi(X, Y)]$:

$$\widehat{S}_n^{\varphi} = \sum_{i=1}^{n} W_{(i)} \varphi(X_{[i]}, Z_{(i)})$$

* Correspondence to: Departamento de Estadística e Investigación Operativa, Facultad de Ciencias Económicas y Empresariales, Universidad de Vigo, Campus Universitario Lagoas-Marcosende, 36310 Vigo, Spain. Tel.: +34 986812492; fax: +34 986812401.

E-mail addresses: dikta@fh-aachen.de (G. Dikta), re.kuelheim@gmx.de (R. Külheim), jpm@isep.ipp.pt (J. Mendonça), jacobo@uvigo.es (J. de Uña-Álvarez).

http://dx.doi.org/10.1016/j.jspi.2015.12.001
0378-3758/

where $Z_{(i)}$ is the ith ordered Z-datum, $\left(X_{[i]}, \delta_{[i]}\right)$ is the ith concomitant, and

$$W_{(i)} = \frac{\delta_{[i]}}{n-i+1} \prod_{j=1}^{i-1} \left[1 - \frac{\delta_{[j]}}{n-j+1}\right]$$

is the jump of the Kaplan–Meier estimator of $F(y) = P(Y \leq y)$ at $y = Z_{(i)}$. Under the two following identifiability assumptions:

(i) Y and C are independent;
(ii) δ and X are independent conditionally on Y;

(which hold in particular if C is independent of (X, Y)), Stute (1993) established the strong consistency of the Kaplan–Meier integral $\widehat{S}_n^{\varphi}$, that is, $\widehat{S}_n^{\varphi} \to S^{\varphi}$ with probability 1, where the limit $S^{\varphi} = E\left[\varphi(X, Y)1_{\{Y \leq \tau_H\}}\right]$ may not be equal to $E\left[\varphi(X, Y)\right]$ due to censoring effects. Here, $\tau_H = \inf\{z : H(z) = 1\}$ stands for the upper bound of the support of H, the cdf of Z. See Stute (1993, 1996, 1999) for discussion on (i)–(ii), further results, and applications.

In some instances, information on the conditional probability of censoring is available. Introduce the function $m(x, z) = P(\delta = 1|X = x, Z = z)$. Without covariates, Dikta (1998) proposed a semiparametric censorship model in which the function m belongs to a certain parametric family. He introduced an estimator for F alternative to the Kaplan–Meier estimator and he proved the asymptotic superiority of the new estimator in the sense of having a smaller asymptotic variance (Dikta et al., 2005). Recently, Dikta (2014) proved that the semiparametric estimator is asymptotically efficient. In the setting with covariates, de Uña-Álvarez and Rodríguez-Campos (2004) proved the strong consistency of

$$S_n^{\varphi} = \sum_{i=1}^{n} W_{(i)}(m_n)\varphi(X_{[i]}, Z_{(i)})$$

where

$$W_{(i)}(m_n) = \frac{m_n(X_{[i]}, Z_{(i)})}{n-i+1} \prod_{j=1}^{i-1} \left[1 - \frac{m_n(X_{[j]}, Z_{(j)})}{n-j+1}\right]$$

and where m_n stands for a uniformly consistent estimator of m. These are 'presmoothed' Kaplan–Meier weights, in the sense that some preliminary smoothing of the probability of uncensoring is performed before the computation of the product-type weights.

Under Dikta (1998)'s semiparametric model, $m \in \{m(., \beta)\}$ with $\beta \in \Omega \subseteq \mathbb{R}^k$, and therefore the function m is estimated by some parametric fit $m(., \beta_n)$. In such a case, and ignoring covariates for a moment, $W_{(i)}(m_n)$ is just the jump of Dikta's semiparametric estimator of $F(y)$ at $y = Z_{(i)}$. Interestingly, application of these 'presmoothed' Kaplan–Meier weights in the presence of covariates allows for a variance reduction, similarly as for the marginal setting. This was illustrated through simulations in the context of censored linear regression (de Uña-Álvarez and Rodríguez-Campos, 2004). More recent applications in which similar features are seen include estimation of a conditional distribution function (Iglesias-Pérez and de Uña-Álvarez, 2010), or estimation of multivariate distribution functions and transition probabilities in multi-state models (de Uña-Álvarez and Amorim, 2011; Amorim et al., 2011; Moreira et al., 2013). Still, asymptotic properties of the presmoothed Kaplan–Meier integral S_n^{φ} (for a general function φ) beyond consistency are unknown. Dikta et al. (2005) established an i.i.d. representation for such integrals in absence of covariates; in this paper, we extend his results to the more general setting in which covariables are present. As a consequence, we obtain a CLT for the estimator; as an important application, we derive the asymptotic normality of regression estimators based on the semiparametric censorship model. We also prove that the asymptotic variance of $\widehat{S}_n^{\varphi}$ is larger than that of S_n^{φ}. Like in our previous papers, we assume that the cdf of Z is continuous, and we use this continuity in the proofs. The case of a discrete distribution is different; typically, one has (besides assumptions (i) and (ii) above) the additional assumption of no common jumps of the censoring and the survival distribution (Stute, 1993).

The rest of the paper is organized as follows. In Section 2 we introduce the needed assumptions and the main results. In Section 3 we include a simulation study to investigate the finite sample performance of the proposed estimator. Proofs of the main results are given in Section 4. Some auxiliary results are collected and proved in the Appendix.

Locally efficient estimation in semiparametric censorship models has been considered under the viewpoint of coarsening at random too, see e.g. Robins and Rotnitzky (1992) or Robins and Finkelstein (2000). In these papers, the approach is based on inverse probability weighted augmented estimation, which allows for the construction of doubly robust estimators. This approach depends on a model for the coarsening mechanism (i.e. the conditional distribution of the observed data (X, Z, δ) given the full data (X, Y)) as well as a model for the cumulative distribution function of the full data. The method we follow here is different in that we only model the conditional expectation of the binary indicator δ given (X, Z). Although the consistency of our approach depends on the assumed model, it may provide accurate estimators even under slight miss-specifications (see Section 3). Note besides that, since both δ and (X, Z) are observable, a reasonable model for $m(x, z) = P(\delta = 1|X = x, Z = z)$ can be postulated by using binary regression techniques (Cox and Snell, 1989; Dikta et al., 2006).

2. Main results

Under a semiparametric censorship model, we have $m \in \{m(., \beta)\}$ with $\beta \in \Omega \subseteq \mathbb{R}^k$; i.e., $m(x, z) = m(x, z, \beta_0)$ for some $\beta_0 \in \Omega$. Following Dikta (1998), the parameter β is estimated by the conditional MLE, that is, by the maximizer β_n of

$$L_n(\beta) = \prod_{i=1}^{n} m(X_i, Z_i, \beta)^{\delta_i} \left[1 - m(X_i, Z_i, \beta)\right]^{1-\delta_i}. \tag{1}$$

By repeating the arguments of the proof to Theorem 2.3 in Dikta (1998) (adapted to covariates), the asymptotic normality of $n^{1/2}(\beta_n - \beta_0)$ may be established (see our Lemma 1 in Section 4). Throughout the paper, for a given function $f(., \beta)$, $D_r f(., \beta_0)$ and $D_{r,s} f(., \beta_0)$ will denote respectively, the first-order and the second-order partial derivatives of $f(., \beta)$ with respect to β, evaluated at $\beta = \beta_0$, and $Grad(f(., \beta)) = (D_1 f(., \beta), \ldots, D_k f(., \beta))$. We will refer to the following regularity conditions.

(C1) There exists a measurable solution $\beta_n \in \Omega$ of the equation $Grad\,(l_n(\beta)) = 0$ which tends to β_0 in probability, where $l_n(\beta) = \log L_n(\beta)$.

(C2) For $i = 1, 2$, $w_i\,(x, z, \beta)$ possesses continuous partial derivatives of second order with respect to β and for each (x, z), where $w_1(x, z, \beta) = \log(m(x, z, \beta))$ and $w_2(x, z, \beta) = \log(1 - m(x, z, \beta))$. Furthermore, $D_{r,s} w_i\,(x, z, \beta)$ is measurable for each $\beta \in \Omega$, and there exist a neighborhood $V\,(\beta_0) \subset \Omega$ and a measurable function M such that, for all $\beta \in V\,(\beta_0)$ and each (x, z), $1 \le r, s \le k$,

$$\left|D_{r,s} w_1\,(x, z, \beta)\right| + \left|D_{r,s} w_2\,(x, z, \beta)\right| \le M\,(x, z) \quad \text{and} \quad E\,(M\,(X, Z)) < \infty.$$

(C3) For $1 \le r \le k$, $\left[\frac{D_r m(X,Z,\beta_0)}{m(X,Z,\beta_0)}\right]^2$ and $\left[\frac{D_r m(X,Z,\beta_0)}{1-m(X,Z,\beta_0)}\right]^2$ have finite expectations.

(C4) The matrix $I(\beta_0) = \left(a_{r,s}\right)_{1 \le r,s \le k}$, where

$$a_{r,s} = -E\left[D_{r,s} w\,(\delta, X, Z, \beta_0)\right] = E\left[\frac{D_r m\,(X, Z, \beta_0)\, D_s m\,(X, Z, \beta_0)}{m\,(X, Z, \beta_0)\,(1 - m\,(X, Z, \beta_0))}\right],$$

with

$$w\,(\delta, x, z, \beta) = \delta w_1\,(x, z, \beta) + (1 - \delta) w_2\,(x, z, \beta)$$

is positive definite.

(C5) $m\,(x, z, \beta)$ possesses continuous partial derivatives of second order with respect to β at each $\beta \in \Omega$ and for each (x, z). Furthermore $D_{r,s} m\,(x, z, \beta)$ is measurable for each $\beta \in \Omega$ and there exist a neighborhood $V\,(\beta_0) \subset \Omega$ and a measurable function M such that, for all $\beta \in V\,(\beta_0)$ and each (x, z), $1 \le r, s \le k$, $\left|D_{r,s} m\,(x, z, \beta)\right| \le M\,(x, z)$ and $E\,(M\,(X, Z)) < \infty$. Finally, $Grad\,(m\,(x, z, \beta_0))$ is bounded on $\mathbb{R}^p \times [0, T]$ with $H(T) < 1$, i.e.

$$\sup_{(x,z) \in \mathbb{R}^p \times [0,T]} \|Grad\,(m\,(x, z, \beta_0))\| < \infty.$$

(C6)

(C6.1) With $m = m(X, Z, \beta_0)$ and $D_r m = D_r m(X, Z, \beta_0)$ we have, for $1 \le r, s, t, u \le k$:

$$E\,[\mathrm{Var}(m|Z)] < \infty,$$

$$E\left[\frac{(D_s m D_t m)^4}{(m(1-m))^4}\right] < \infty,$$

$$E\left[\frac{|D_r m D_s m D_t m D_u m(-3m + 3m^2 + 1)|}{m^3(1-m)^3}\right] < \infty.$$

(C6.2) $v \in [0, T] \to E\left(\frac{D_j m(X, Z, \beta_0)}{1 - H(Z)} I_{\{Z \le v\}}\right)$ is continuous for $1 \le j \le k$ and,

for some $\varepsilon > 0$, $E\left(\left|\frac{D_j m(X, Z, \beta_0)}{1 - H(Z)}\right|^{2+\varepsilon}\right) < \infty$.

(C7) $\widetilde{m}\,(., \beta_0)$ is of bounded variation over $[0, \tau_H]$, where

$$\widetilde{m}(z, \beta_0) = E\,[m(X, Z, \beta_0)\,|Z = z\,],$$

i.e.

$$\sup\left\{\sum_{i=1}^{l} |\widetilde{m}\,(z_i, \beta_0) - \widetilde{m}\,(z_{i-1}, \beta_0)| : 0 \equiv z_0 \le z_1 \le \cdots \le z_l \le \tau_H, l \ge 1\right\} < \infty$$

where the sup is taken over all partitions $(z_i)_{i=0}^{l}$ of $[0, \tau_H]$. Besides,

$$E\left[\left|\frac{m(X,Z,\beta_0)-\widetilde{m}(Z,\beta_0)}{1-H(Z)}\right|\right]<\infty.$$

Conditions (C1)–(C4) are needed for the asymptotic normality of $n^{1/2}(\beta_n-\beta_0)$, similarly as in Dikta (1998). Condition (C5) is a technical one, which, like in Dikta et al. (2005), is needed here to control the presmoothing function $m_n(.) = m(.; \beta_n)$. Condition (C6) is a substitute for condition (A6) in Dikta et al. (2005), who used such assumption to prove the tightness of a certain process. Similarly, we use (C6) to prove tightness of certain extended processes which include the covariates. Condition (C7) is a natural adaptation of condition (A7) in Dikta et al. (2005) to our context.

Theorem 1 is our main result. It gives an extension of Theorem 2.1 in Dikta et al. (2005) to the setting in which covariables are present. Some functions appearing in the i.i.d. representation are defined now. Put

$$H_{X,Z}(x,z) = P(X\le x, Z\le z),$$

$$\widetilde{H}^{00}(x,t,\beta) = \int_{u\le x}\int_{-\infty}^{t}(1-m(u,z,\beta))\,H_{X,Z}(du,dz),$$

$$\widetilde{H}^{0}_{X,Z}(x,t) = \widetilde{H}^{00}(x,t,\beta_0),$$

$$\widetilde{H}^{11}(x,t,\beta) = \int_{u\le x}\int_{-\infty}^{t} m(u,z,\beta)\,H_{X,Z}(du,dz),\qquad \widetilde{H}^{1}_{X,Z}(x,t) = \widetilde{H}^{11}(x,t,\beta_0),$$

$$K(x,z,\delta) = \frac{\delta - m(x,z,\beta_0)}{m(x,z,\beta_0)\,(1-m(x,z,\beta_0))},$$

$$\gamma_0(z) = \exp\left(\int_0^{z^-}\frac{\widetilde{H}^0(du)}{1-H(u)}\right)\quad \text{where } \widetilde{H}^0(u) = P(Z\le u, \delta=0),$$

$$\gamma_1(z) = \frac{1}{1-H(z)}\iint 1_{\{z<t\}}\varphi(x,t)\,\gamma_0(t)\,\widetilde{H}^1_{X,Z}(dx,dt),$$

$$\gamma_2(z) = \iint \frac{1_{\{u<z,u<t\}}\varphi(x,t)\,\gamma_0(t)}{[1-H(u)]^2}\widetilde{H}^1_{X,Z}(dx,dt)\,\widetilde{H}^0_{X,Z}(ds,du),$$

$$\gamma_3(v,w) = \iint \frac{1_{\{w>s\}}\alpha(u,s,v,w)\,\varphi(x,w)\,\gamma_0(w)}{1-H(s)}\,\widetilde{H}^1_{X,Z}(dx,dw)\;H_{X,Z}(du,ds),$$

$$\gamma_4(u,s) = \iint \varphi(x,t)\,\gamma_0(t)\,\alpha(x,t,u,s)\;H_{X,Z}(dx,dt),$$

$$\alpha(u,s,v,t) = \left\langle Grad\,(m(u,s,\beta_0))\,|I^{-1}(\beta_0)\;Grad\,(m(v,t,\beta_0))\right\rangle$$

where $\langle .|.\rangle$ denotes the inner product in $\mathbb{R}^k$. Finally, some conditions on the function φ are needed. These are, with $F_{X,Y}(x,y) = P(X\le x, Y\le y)$,

(M1) $\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^2(x,u)\gamma_0(u)\,F_{X,Y}(dx,du)<\infty$

(M2) $\int_{\mathbb{R}^p}\int_0^{\tau_H}|\varphi(x,u)|\,(1-H(u))^{-1/2}\,F_{X,Y}(dx,du)<\infty$

(M3) $\int_{\mathbb{R}^p}\int_0^{\tau_H}|\varphi(x,u)|\,\gamma_0(u)\,H_{X,Z}(dx,du)<\infty$.

These conditions (M1)–(M3) are a proper adaptation of conditions (M1)–(M3) in Dikta et al. (2005) in the presence of covariates.

Theorem 1. *Let Ω be a connected open subset of $\mathbb{R}^k$. If H is continuous,* (C1)–(C7)*, and* (M1)–(M3) *are satisfied, then*

$$S_n^{\varphi} = \frac{1}{n}\sum_{i=1}^{n}\xi^{\varphi}(X_i,Z_i,\delta_i;\beta_0) + o_P(n^{-1/2}), \tag{2}$$

where

$$\begin{aligned}\xi^{\varphi}(X_i,Z_i,\delta_i;\beta_0) &= \varphi(X_i,Z_i)\gamma_0(Z_i)m(X_i,Z_i,\beta_0) + [1-m(X_i,Z_i,\beta_0)]\gamma_1(Z_i)\\ &\quad -\gamma_2(Z_i) - K(X_i,Z_i,\delta_i)\gamma_3(X_i,Z_i) + K(X_i,Z_i,\delta_i)\gamma_4(X_i,Z_i).\end{aligned}$$

The fourth and fifth terms in the representation of Theorem 1 come from the estimation of β_0. These two terms have zero mean, which is easily seen by noting $E[K(X,Z,\delta)|X,Z]=0$. On the other hand, since $E[(1-m(X_i,Z_i,\beta_0))\gamma_1(Z_i)] = E[(1-\delta_i)\gamma_1(Z_i)] = E[\gamma_2(Z_i)]$ (cfr. Stute, 1996), we get that the expectation of our representation equals

$$E[\varphi(X,Z)\gamma_0(Z)m(X,Z;\beta_0)] = E[\varphi(X,Z)\,\gamma_0(Z)\delta] = S^{\varphi}.$$

Since conditions in Theorem 1 guarantee the existence of the second order moment of the leading term of the representation, and application of the CLT gives the following corollary.

Corollary 1. *Under the assumptions of Theorem 1, we have* $n^{1/2}(S_n^{\varphi} - S^{\varphi}) \to N(0, \sigma)$ *in distribution, where* $\sigma^2 = \mathrm{Var}(\xi^{\varphi}(X, Z, \delta; \beta_0))$. ■

Consider now the linear regression model $Y = X^t\theta_0 + \varepsilon$, where θ_0 is the true vector of p regression parameters and ε is an error term satisfying $E[\varepsilon|X] = 0$. Introduce the estimator θ_n of θ_0 as the minimizer of the (randomly) weighted least squares criterion

$$\theta \longmapsto \sum_{i=1}^{n} W_{(i)}(m_n)(Z_{(i)} - X_{[i]}^t\theta)^2$$

where again $m_n(.) = m(.; \beta_n)$. An application of Theorem 1 and the multivariate CLT gives the following result. See Section 4 for a proof.

Corollary 2. *Assume* $\tau_F = \tau_H$, *where* $\tau_F = \inf\{y : F(y) = 1\}$. *Assume that the matrix* $\Sigma_0 = E[XX^t]$ *is non-singular. Under the assumptions of Theorem 1 for the special* φ*-functions* $\varphi_j(x, z) = \varphi_j(x^1, \ldots, x^p, z) = x^j(z - x^t\theta_0)$ *and* $\varphi_{(i,j)}(x, z) = x^ix^j$, $1 \le i, j \le p$, *we have that* $n^{1/2}(\theta_n - \theta_0)$ *converges in distribution to a normal random variable* $N(\underline{0}, \Sigma_0^{-1}\Sigma\Sigma_0^{-1})$, *where* $\Sigma = (\sigma_{ij})_{1\le i,j\le p}$ *has entries* $\sigma_{ij} = Cov(\xi^{\varphi_i}(X, Z, \delta; \beta_0), \xi^{\varphi_j}(X, Z, \delta; \beta_0))$.

Similarly to Corollary 2.5 in Dikta et al. (2005), the next result shows that the asymptotic variance of S_n^{φ} (σ^2) is not larger than that of $\widehat{S}_n^{\varphi}$. According to Stute (1996), the latter is given by $\sigma_0^2 = \mathrm{Var}(\varphi(X, Z)\gamma_0(Z)\delta + \gamma_1(Z)(1-\delta) - \gamma_2(Z))$. Introduce $A(x, z) = \gamma_1(z) - \varphi(x, z)\gamma_0(z)$ and $B(x, z) = (\gamma_3(x, z) - \gamma_4(x, z))/(m(x, z)(1-m(x, z)))$, where we put $m(x, z) = m(x, z, \beta_0)$ for simplicity.

Corollary 3. *Under the assumptions of Theorem 1, we have*

$$\sigma_0^2 - \sigma^2 = E\left[m(X, Z)(1 - m(X, Z))(A^2(X, Z) - B^2(X, Z))\right] \ge 0$$

with equality if and only if with probability one

$$aA(X, Z)\sqrt{m(X, Z)(1 - m(X, Z))} + b\sum_{i=1}^{k} \frac{h_iD_i(m(X, Z))}{\sqrt{m(X, Z)(1 - m(X, Z))}} = 0$$

for some a, b *having* $a^2 + b^2 > 0$. *Here*

$$h = I^{-1}(\beta_0)p,$$

where $I^{-1}(\beta_0) = (\bar{\sigma}_{r,s})_{1\le r,s\le k}$ *is the inverse of* $I(\beta_0)$ *and*

$$p^t = (E(A(X, Z)D_1(m(X, Z))), \ldots, E(A(X, Z)D_k(m(X, Z)))).$$

3. Simulation study

In this section we include a simulation study to investigate the finite sample performance of the semiparametric estimator S_n^{φ}. In particular, we consider the indicator function $\varphi(u, v) = 1_{\{u\le x, v\le y\}}$ for a given (x, y) point, so S_n^{φ} reduces to an empirical bivariate distribution function $F_{X,Y,n}(x, y)$. For comparison purposes, we include the results pertaining to the Kaplan–Meier-based estimator, $\widehat{F}_{X,Y,n}(x, y) = \widehat{S}_n^{\varphi}$ (Stute, 1993).

The simulation steps are as follows. We simulate a Gaussian one-dimensional covariate with mean 0.5 and standard deviation 1.5, $X \sim Normal(0.5, 1.5)$. Given X, the lifetime Y follows a (conditional) distribution function given by

$$P(Y \le y|X = x) = 1 - \exp\{-(y/b(x))^a\}$$

where a and $b(x)$ are respectively shape and scale constants. Thus, the conditional distribution of the lifetime is of Weibull type. We take $a = 1$, and $b(x) = \exp\{\rho x\}$ with $\rho = 1$ or $\rho = 3$. We independently simulate the potential censoring time from a Weibull distribution with shape parameter $c = 1$ or $c = 3$, and scale parameter 1, leading to censoring proportions about 41.3% ($\rho = 1, c = 1$), 33.8% ($\rho = 1, c = 3$), 37.8% ($\rho = 3, c = 1$), and 34.4% ($\rho = 3, c = 3$). Since C is independent of (X, Y), assumptions (i) and (ii) are fulfilled. On the other hand, it is straightforward to see that, under our simulation scheme, the conditional probability of uncensoring is given by

$$m(X, Z) = 1/(1 + \exp(\nu(X, Z)))$$

where

$$\nu(x, z) = -\rho x + (\log(c) - \log(a)) + (c - a)\log(z).$$

We consider the semiparametric estimator $F_{X,Y,n}(x, y)$ based on the logistic model with linear predictor $\nu(x, z, \beta) = \beta_0 + \beta_1 x + \beta_2 z$. Therefore, it is clear that the parametric model is well specified whenever $c - a = 0$. Our simulation

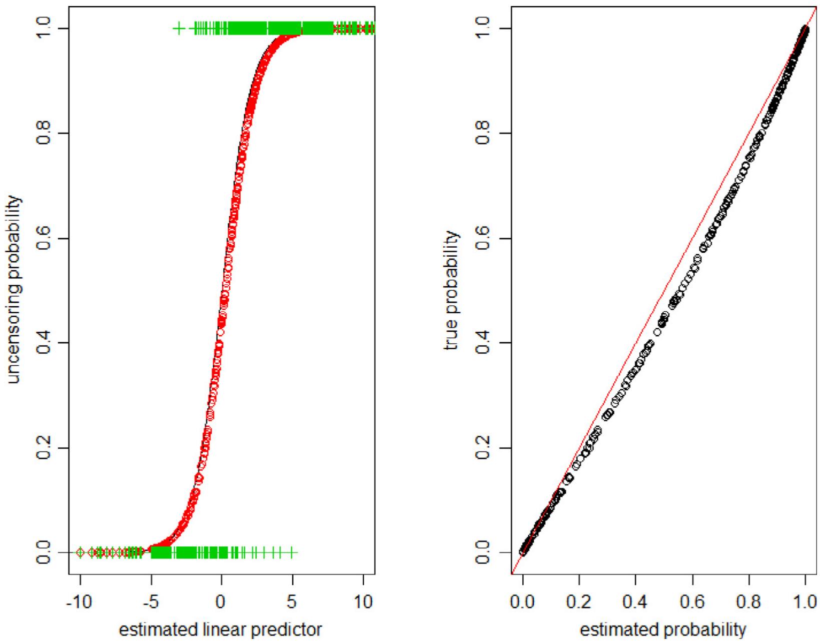

Fig. 1. Left panel: Fitted logistic model (line), true probability of uncensoring (dots), and censoring indicators (crosses) simulated from the model $c = 1$ (no miss-specification) with $\rho = 3$, $n = 500$. Right panel: True probability of uncensoring vs. estimated probability for the same simulated trial.

study includes such scenario and a situation with miss-specification of the parametric model, namely $c - a = 2$, too. In Figs. 1 and 2 we illustrate these two different scenarios from a single Monte Carlo trial of size $n = 500$ drawn from each of the two models, in the particular case $\rho = 3$. It is seen that, when $a = c$, the true values of $m(X, Z)$ are accurately estimated by the logistic model while, when $a \neq c$, the linear predictor is unable to estimate the uncensoring probability well, because it is unaware of the needed log-transformation of Z.

In Table 1 we provide the bias, the standard deviation (SD), and the mean squared error (MSE) of both the semiparametric estimator and the Kaplan–Meier-based estimator along 1000 Monte Carlo trials, for sample size $n = 500$; Table 2 reports the information corresponding to $n = 1000$. The particular (x, y)-points considered are (1.5, 0.1), (1.5, 0.4), and (1.5, 1.3), which correspond to cumulative proportions $(F_{X,Y}(x, y))$ 0.0917, 0.2746 and 0.4967 when $\rho = 1$, and 0.2272, 0.3446, and 0.4385 when $\rho = 3$. In Table 1 it is seen that, for both estimators, the bias is of smaller order compared to the SD (so the MSE is roughly the variance), and that the MSE decreases as the sample size increases. It is also seen that, when the parametric model is correctly specified, the semiparametric estimator outperforms the Kaplan–Meier-based estimator, providing a smaller MSE. In this case, the MSE improvement reaches 27.2% ($n = 500$, $\rho = 1$). On the contrary, when the model is miss-specified ($c = 3$), the semiparametric estimator is inferior in most occasions (9 out of 12) due to its systematic bias; in this case, the maximum MSE improvement offered by the purely nonparametric estimator is 15.8% ($n = 1000$, $\rho = 1$). In the setting with miss-specified model, the semiparametric estimator still has an improved MSE when $y = 1.3$ and $n = 500$ (Table 2) since, in this case, the introduced bias is compensated by the variance reduction. Overall, we conclude that much can be gained through presmoothing, provided that a suitable (or even a slightly miss-specified) parametric model is chosen.

4. Proofs

To prove Theorem 1 we follow the steps in the proof to Theorem 2.1 in Dikta et al. (2005). For this, ten lemmas will be established. The first one gives the uniform convergence with rate for the presmoothing function $m_n(.) = m(., \beta_n)$. Lemma 2 gives a suitable representation of S_n^{φ}. A Taylor expansion leads then to another representation with several terms, which are analyzed in detail along Lemmas 3–10. In Lemmas 3–10 we temporarily refer to the following support condition, which is removed at the end of the section:

(C8) There exists $T < \tau_H$ such that $\varphi(x, y) = 0$ for all $(x, y) \in \mathbb{R}^p \times (T, \infty)$.

Our Lemmas 1–10 correspond to Lemmas 4.1–4.10 in Dikta et al. (2005). To prove his Lemma 4.3, Dikta et al. (2005) used his Eq. (4.3), following results in a previous paper (Dikta, 1998). This is why we investigate the extension of the results in Dikta (1998) to the setting with covariates in order to prove our Lemma 3 (see Lemmas A.1 and A.2 in the Appendix). Roughly, the rest of the arguments behind our proofs are direct adaptations of the lines in Dikta et al. (2005) to our particular setting.

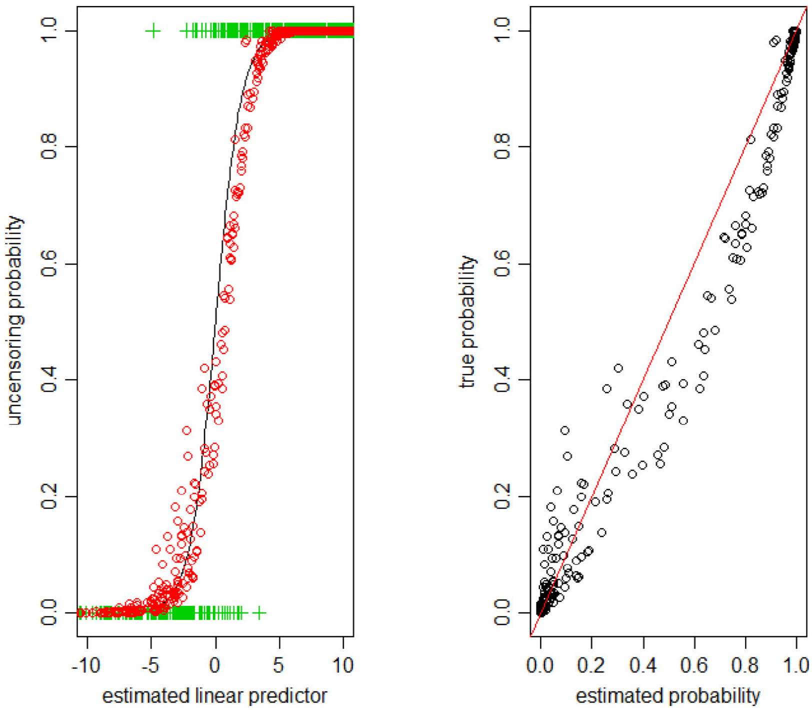

Fig. 2. Left panel: Fitted logistic model (line), true probability of uncensoring (dots), and censoring indicators (crosses) simulated from the model $c = 3$ (miss-specified model) with $\rho = 3$, $n = 500$. Right panel: True probability of uncensoring vs. estimated probability for the same simulated trial.

Table 1
Bias (multiplied by 10^4), SD and MSE (multiplied by 10^4) of the semiparametric estimator $F_{X,Y,n}(x, y)$ (SP) and of the Kaplan–Meier-based estimator $\widehat{F}_{X,Y,n}(x, y)$ (KM) along 1000 Monte Carlo trials, for $x = 1.5$ and $n = 500$. The ratio of the MSEs (R(SP,KM)) is reported too.

	KM			SP			R(SP,KM)
	Bias	SD	MSE	Bias	SD	MSE	
$c = 1, \rho = 1$							
$y = 0.1$	−0.4753	0.0138	1.8982	−0.2409	0.0118	1.3826	0.7284
$y = 0.4$	0.0140	0.0226	5.0981	2.0577	0.0207	4.2716	0.8379
$y = 1.3$	6.5446	0.0301	9.0925	−0.0438	0.0273	7.4308	0.8172
$c = 1, \rho = 3$							
$y = 0.1$	−7.0840	0.0193	3.7237	−5.4741	0.0185	3.4248	0.9197
$y = 0.4$	−0.2290	0.0227	5.1601	1.0916	0.0221	4.8823	0.9462
$y = 1.3$	3.8443	0.0267	7.1557	−2.6788	0.0255	6.5075	0.9094
$c = 3, \rho = 1$							
$y = 0.1$	−2.6770	0.0129	1.6641	−0.0045	0.0124	1.7493	1.0512
$y = 0.4$	3.5812	0.0198	3.9223	−0.0028	0.0197	3.9560	1.0086
$y = 1.3$	4.0925	0.0306	9.3718	−0.0114	0.0270	8.6129	0.9190
$c = 3, \rho = 3$							
$y = 0.1$	−3.0804	0.0187	3.4986	−0.0026	0.0186	3.5212	1.0065
$y = 0.4$	−5.8540	0.0205	4.2242	−0.0021	0.0207	4.3262	1.0242
$y = 1.3$	13.023	0.0254	6.4701	−0.0036	0.0237	5.7409	0.8873

Lemma 1. *If assumptions* (C1)–(C4) *are satisfied, then* $n^{1/2}(\beta_n - \beta_0)$ *is asymptotically normal,* $N_k(0, I(\beta_0)^{-1})$, *where* $I(\beta_0) = (a_{r,s})_{1\le r,s\le k}$ *is the Fisher information matrix defined in* (C4). *Furthermore, under* (C1)–(C5),

$$\sup_{x\in\mathbb{R}^p,0\le z\le\tau_H} |m(x, z, \beta_n) - m(x, z, \beta_0)| = O_P(n^{-1/2}).$$

Proof. The proof follows the steps in Lemma 4.1 in Dikta et al. (2005), see also Theorem 2.3 in Dikta (1998). ∎

Introduce $H_n(z) = n^{-1}\sum_{i=1}^{n} 1_{\{Z_i\le z\}}$, $H_{X,Z,n}(x, z) = n^{-1}\sum_{i=1}^{n} 1_{\{X_i\le x, Z_i\le z\}}$, which are the empirical counterparts of $H(z)$ and $H_{X,Z}(x, z)$.

Table 2
Bias (multiplied by 10^4), SD and MSE (multiplied by 10^4) of the semiparametric estimator $F_{X,Y,n}(x, y)$ (SP) and of the Kaplan–Meier-based estimator $\widehat{F}_{X,Y,n}(x, y)$ (KM) along 1000 Monte Carlo trials, for $x = 1.5$ and $n = 1000$. The ratio of the MSEs (R(SP,KM)) is reported too.

	KM			SP			R(SP,KM)
	Bias	SD	MSE	Bias	SD	MSE	
$c = 1, \rho = 1$							
$y = 0.1$	−0.4293	0.0093	0.8666	−0.1515	0.0082	0.6710	0.7743
$y = 0.4$	1.3732	0.0152	2.3036	0.1091	0.0141	1.9833	0.8610
$y = 1.3$	0.7717	0.0211	4.4576	−5.1175	0.0188	3.5439	0.7950
$c = 1, \rho = 3$							
$y = 0.1$	−6.1212	0.0138	1.9112	−5.2986	0.0134	1.7995	0.9415
$y = 0.4$	−3.4667	0.0160	2.5573	−1.8701	0.0154	2.3807	0.9309
$y = 1.3$	0.5980	0.0190	3.5933	−2.1852	0.0177	3.1461	0.8756
$c = 3, \rho = 1$							
$y = 0.1$	−2.0069	0.0091	0.8232	−0.0045	0.0088	0.9783	1.1883
$y = 0.4$	2.9310	0.0144	2.0757	−0.0029	0.0144	2.1643	1.0427
$y = 1.3$	3.2151	0.0218	4.7506	−0.0120	0.0195	5.2392	1.1029
$c = 3, \rho = 3$							
$y = 0.1$	4.9824	0.0131	1.7261	−0.0019	0.0130	1.7382	1.0070
$y = 0.4$	5.1819	0.0145	2.1166	−0.0011	0.0146	2.1416	1.0118
$y = 1.3$	10.998	0.0183	3.3651	−0.0038	0.0171	3.0723	0.9130

Lemma 2. *For a continuous H we have*

$$S_n^{\varphi} = \int_{\mathbb{R}^p} \int \varphi(x, w) m(x, w, \beta_n) \exp\left\{ n \int_{\mathbb{R}^p} \int_0^{w^-} \log\left[1 + \frac{1 - m(u, v, \beta_n)}{n(1 - H_n(v))}\right] H_{X,Z,n}(du, dv) \right\} H_{X,Z,n}(dx, dw).$$

Proof. We have:

$$\begin{aligned}
S_n^{\varphi} &= \sum_{i=1}^{n} \frac{\varphi\left(X_{[i]}, Z_{(i)}\right) m\left(X_{[i]}, Z_{(i)}, \beta_n\right)}{n - i + 1} \prod_{j=1}^{i-1} \left(1 - \frac{m\left(X_{[j]}, Z_{(j)}, \beta_n\right)}{n - j + 1}\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi\left(X_{[i]}, Z_{(i)}\right) m\left(X_{[i]}, Z_{(i)}, \beta_n\right) \prod_{j=1}^{i-1} \left(1 + \frac{1 - m\left(X_{[j]}, Z_{(j)}, \beta_n\right)}{n - j}\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi\left(X_{[i]}, Z_{(i)}\right) m\left(X_{[i]}, Z_{(i)}, \beta_n\right) \prod_{j=1}^{i-1} \left(1 + \frac{1 - m\left(X_{[j]}, Z_{(j)}, \beta_n\right)}{n\left(1 - H_n\left(Z_{(j)}\right)\right)}\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \varphi\left(X_{[i]}, Z_{(i)}\right) m\left(X_{[i]}, Z_{(i)}, \beta_n\right) \exp\left(\sum_{j=1}^{n} 1_{\{Z_{(j)} < Z_{(i)}\}} \log\left(1 + \frac{1 - m\left(X_{[j]}, Z_{(j)}, \beta_n\right)}{n\left(1 - H_n\left(Z_{(j)}\right)\right)}\right)\right) \\
&= \int_{\mathbb{R}^p} \int \varphi(x, w)\, m(x, w, \beta_n) \exp\left\{ n \int_{\mathbb{R}^p} \int_0^{w^-} \log\left[1 + \frac{1 - m(u, v, \beta_n)}{n(1 - H_n(v))}\right] H_{X,Z,n}(du, dv) \right\} H_{X,Z,n}(dx, dw). \quad \blacksquare
\end{aligned}$$

Now, expand the exponential term from expression in Lemma 2 at the points

$$A_i = \int_{\mathbb{R}^p} \int_0^{Z_i^-} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z}(du, dv)$$

to get

$$\begin{aligned}
&\exp\left\{ n \int_{\mathbb{R}^p} \int_0^{Z_i^-} \log\left[1 + \frac{1 - m(u, v, \beta_n)}{n(1 - H_n(v))}\right] H_{X,Z,n}(du, dv) \right\} \\
&= \exp(A_i)\left[1 + n \int_{\mathbb{R}^p} \int_0^{Z_i^-} \log\left[1 + \frac{1 - m(u, v, \beta_n)}{n(1 - H_n(v))}\right] H_{X,Z,n}(du, dv) - A_i\right] \\
&\quad + \frac{\exp(\Delta_i)}{2}\left[n \int_{\mathbb{R}^p} \int_0^{Z_i^-} \log\left[1 + \frac{1 - m(u, v, \beta_n)}{n(1 - H_n(v))}\right] H_{X,Z,n}(du, dv) - A_i\right]^2
\end{aligned}$$

where Δ_i is between the two terms in brackets. Note that $\exp(A_i) = \gamma_0(Z_i)$. With these expansions, we have

$$\begin{aligned} S_n^{\varphi} = & \frac{1}{n}\sum_{i=1}^{n} \varphi\left(X_i, Z_i\right) m\left(X_i, Z_i, \beta_0\right) \gamma_0(Z_i) \left(1 + B_{i,n}(\beta_n) + C_{i,n}(\beta_n)\right) \\ & + \frac{1}{n}\sum_{i=1}^{n} \varphi\left(X_i, Z_i\right) \left(m\left(X_i, Z_i, \beta_n\right) - m\left(X_i, Z_i, \beta_0\right)\right) \times \gamma_0(Z_i) \left(1 + B_{i,n}(\beta_n) + C_{i,n}(\beta_n)\right) \\ & + \frac{1}{2n}\sum_{i=1}^{n} \varphi\left(X_i, Z_i\right) \left(m\left(X_i, Z_i, \beta_n\right)\right) \exp(\Delta_i) \left(B_{i,n}(\beta_n) + C_{i,n}(\beta_n)\right)^2, \end{aligned} \tag{3}$$

where

$$B_{i,n}(\beta) = n\int_{\mathbb{R}^p}\int_0^{Z_i^-} \log\left(1 + \frac{1 - m\left(u, v, \beta_n\right)}{n\left(1 - H_n(v)\right)}\right) H_{X,Z,n}(du, dv) - \int_{\mathbb{R}^p}\int_0^{Z_i^-} \frac{1 - m(u, v, \beta)}{1 - H_n(v)} H_{X,Z,n}(du, dv)$$

and

$$C_{i,n}(\beta) = \int_{\mathbb{R}^p}\int_0^{Z_i^-} \frac{1 - m\left(u, v, \beta\right)}{1 - H_n(v)} H_{X,Z,n}(du, dv) - \int_{\mathbb{R}^p}\int_0^{Z_i^-} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z}(du, dv).$$

In Lemmas 3–7 we obtain an i.i.d. representation for

$$\frac{1}{n}\sum_{i=1}^{n} \varphi(X_i, Z_i) m(X_i, Z_i; \beta_0) \gamma_0(Z_i) C_{i,n}(\beta_n)$$

under the extra assumption (C8).

Lemma 3. *If H is continuous, $E[|\varphi(X, Y)|] < \infty$, and* (C1)–(C5) *and* (C8) *are satisfied, then, as $n \to \infty$,*

$$\frac{1}{n}\sum_{i=1}^{n} \varphi(X_i, Z_i) m(X_i, Z_i, \beta_0) \gamma_0(Z_i) C_{i,n}(\beta_n) = \int_{\mathbb{R}^p}\int_0^{\infty} \varphi(s, t) m(s, t, \beta_0) \gamma_0(t) C_n(\beta_0, t) H_{X,Z,n}(ds, dt) + o_P(n^{-1/2})$$

where

$$\begin{aligned} C_n(\beta_0, t) = & \int_{\mathbb{R}^p}\int_0^{t^-} \frac{1 - m\left(u, v, \beta_0\right)}{1 - H(v)} H_{X,Z,n}(du, dv) - \int_{\mathbb{R}^p}\int_0^{t^-} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z}(du, dv) \\ & + \int_{\mathbb{R}^p}\int_0^{t^-} \frac{\left(1 - m(u, v, \beta_0)\right)\left(H_n(v^-) - H(v)\right)}{\left(1 - H(v)\right)^2} H_{X,Z}(du, dv) \\ & - n^{-1}\sum_{i=1}^{n} K\left(X_i, Z_i, \delta_i\right) \int_{\mathbb{R}^p}\int_0^{t} \frac{\alpha(u, v, X_i, Z_i)}{1 - H(v)} H_{X,Z}(du, dv). \end{aligned}$$

Proof. Note that $C_{i,n}(\beta_n) = \Lambda_n^0(Z_i) - \Lambda^0(Z_i)$ where

$$\Lambda_n^0(Z_i) = \int_{\mathbb{R}^p}\int_0^{Z_i^-} \frac{1 - m\left(u, v, \beta_n\right)}{1 - H_n(v)} H_{X,Z,n}(du, dv)$$

and

$$\Lambda^0(Z_i) = \int_{\mathbb{R}^p}\int_0^{Z_i} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z}(du, dv).$$

Similarly as in Dikta et al. (2005, Lemma 4.3), we have

$$\max_{Z_i \le T} \left|C_{i,n}(\beta_n) - C_n(\beta_0, Z_i)\right| \le \sup_{0 \le v \le T} \left|\Lambda_n^0(v) - \Lambda^0(v) - C_n(\beta_0, v)\right|. \tag{4}$$

In the setting without covariates, this quantity was studied in detail in Dikta (1998), see his Lemma 3.8. In our setting, to conclude we need to prove that the sup at the right-hand side of (4) is $o_P(n^{-1/2})$. This is done in the Appendix, Lemma A.1. ■

Lemmas 4–6 show that the integrating measure $H_{X,Z,n}$ appearing at the right-hand side of the representation in Lemma 3 can be replaced by $H_{X,Z}$.

Lemma 4. *If H is continuous, $E[\varphi(X,Y)^2]<\infty$, and (C8) is satisfied, then, as $n\to\infty$,*

$$\begin{aligned}&\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1-m(u,v,\beta_0)}{1-H(v)}d\left(H_{X,Z,n}(u,v)-H_{X,Z}(u,v)\right)dH_{X,Z,n}(s,t)\\&\quad=\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1-m(u,v,\beta_0)}{1-H(v)}d\left(H_{X,Z,n}(u,v)-H_{X,Z}(u,v)\right)dH_{X,Z}(s,t)+O_P(n^{-1}).\end{aligned}$$

Proof. Denote the left-hand side of the above equation by V_n. Since

$$\int_{\mathbb{R}^p\times[0,t]}\frac{1-m(u,v,\beta_0)}{1-H(v)}dH_{X,Z}(du,dv)=-\log(1-G(t))$$

we have

$$V_n=\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+}h(s,t,u,v)H_{X,Z,n}(du,dv)H_{X,Z,n}(ds,dt)$$

where

$$h(s,t,u,v)=\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\left(1_{\{v\le t\}}\frac{1-m(u,v,\beta_0)}{1-H(v)}+\log\left(1-G(t)\right)\right),$$

where G stands for the censoring distribution. The integrability assumption on φ, (C8) and Lemma 5.7.3 in Serfling (1980), guarantees that $V_n=U_n+O_P(n^{-1})$ where U_n denotes the associated U-statistic. Observe that $h\in L_2(H_{X,Z}\otimes H_{X,Z})$ and apply Theorem 5.3.2 in Serfling (1980) to get

$$\begin{aligned}V_n&=\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+}h(s,t,u,v)H_{X,Z,n}(du,dv)H_{X,Z}(ds,dt)+\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+}h(s,t,u,v)H_{X,Z}(du,dv)H_{X,Z,n}(ds,dt)\\&\quad-\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+}h(s,t,u,v)H_{X,Z}(du,dv)H_{X,Z}(ds,dt)+O_P(n^{-1}).\end{aligned}$$

Note that the first integral on the right-hand side is identical to the integral on the right-hand side of the lemma. One may easily see that $E\left[h(s,t,X,Z)\right]=0$ and thus $E\left[h(X_1,Z_1,X_2,Z_2)\right]=0$; hence, the second and third integral on the right-hand side vanish, which completes the proof. ∎

By using a similar argumentation, the following lemma can be established.

Lemma 5. *If H is continuous, $E[\varphi(X,Y)^2]<\infty$, and (C8) is satisfied, then, as $n\to\infty$,*

$$\begin{aligned}&\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{H_n(v^-)-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)dH_{X,Z,n}(s,t)\\&\quad=\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{H_n(v^-)-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)dH_{X,Z}(s,t)+O_P(n^{-1}).\end{aligned}$$

Proof. Denote the left-hand side of the above equation by V_n, and observe that

$$\begin{aligned}V_n&=\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{H_n(v^-)-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)dH_{X,Z,n}(s,t)\\&=\frac{1}{n}\sum_{j=1}^n\int_{\mathbb{R}^p\times[0,Z_j]}\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\frac{\frac{1}{n}\sum_{i=1}^n\left(1_{\{Z_i<v\}}-H(v)\right)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)\\&=\frac{1}{n^2}\sum_{j=1}^n\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\sum_{i=1}^n\int_{\mathbb{R}^p\times[0,Z_j]}\frac{1_{\{Z_i<v\}}-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)\\&=\frac{1}{n(n-1)}U_n+R_n\end{aligned}$$

where

$$\begin{aligned}U_n&=\sum_{1\le i\ne j\le n}\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\int_{\mathbb{R}^p}\int_0^{Z_j}\frac{1_{\{Z_i<v\}}-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)\\&\equiv\sum_{1\le i\ne j\le n}h(X_i,Z_i,X_j,Z_j)\end{aligned}$$

and

$$\begin{aligned}R_n &= -\frac{1}{n}U_n + \frac{1}{n^2}\sum_{j=1}^{n}\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\int_{\mathbb{R}^p}\int_0^{Z_j}\frac{1_{\{Z_j<v\}}-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)\\ &= -\frac{1}{n}U_n - \frac{1}{n^2}\sum_{j=1}^{n}\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\int_{\mathbb{R}^p}\int_0^{Z_j}\frac{H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v).\end{aligned}$$

Now, $E[U_n|X_k,Z_k] = \sum_{1\le i\ne j\le n}E(h(X_i,Z_i,X_j,Z_j)|X_k,Z_k)$. For a fixed k, we have $E[U_n|X_k,Z_k]=0$ if $(i\ne k, j\ne k)$. Therefore,

$$\begin{aligned}E[U_n|X_k,Z_k] &= \sum_{1\le i\ne j\le n}E(h(X_i,Z_i,X_j,Z_j)|X_k,Z_k)\\ &= \sum_{j\ne k}E(h(X_k,Z_k,X_j,Z_j)|X_k,Z_k) + \sum_{i\ne k}E(h(X_i,Z_i,X_k,Z_k)|X_k,Z_k)\\ &= \sum_{j\ne k}E\left[\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\int_{\mathbb{R}^p}\int_0^{Z_j}\frac{1_{\{Z_k\le v\}}-H(v)}{(1-H(v))^2}dH^0_{X,Z}(u,v)|X_k,Z_k\right]\\ &\quad + \sum_{i\ne k}E\left[\varphi(X_k,Z_k)m(X_k,Z_k,\beta_0)\gamma_0(Z_k)E\left[\int_{\mathbb{R}^p}\int_0^{Z_k}\frac{1_{\{Z_i\le v\}}-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)|X_k,Z_k\right]\right].\end{aligned}$$

The expected value in the second term becomes, with $j\ne i$ and $j\ne k$,

$$\begin{aligned}E\left[\int_{\mathbb{R}^p}\int_0^{Z_k}\frac{1_{\{Z_i<v\}}-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)|X_k,Z_k\right] &= E\left[\frac{1_{\{Z_i<Z_j\}}-H(Z_j)}{(1-H(Z_j))^2}(1-\delta_j)1_{\{Z_j\le Z_k\}}|X_k,Z_k\right]\\ &= E\left[E\left[\frac{1_{\{Z_i<Z_j\}}-H(Z_j)}{(1-H(Z_j))^2}(1-\delta_j)1_{\{Z_j\le Z_k\}}|Z_j,\delta_j,X_k,Z_k\right]\Bigg|X_k,Z_k\right]\\ &= E\left[\frac{H(Z_j)-H(Z_j)}{(1-H(Z_j))^2}(1-\delta_j)1_{\{Z_j\le Z_k\}}|X_k,Z_k\right]=0.\end{aligned}$$

While the expected value in the first term becomes, with $i\ne j$ and $i\ne k$

$$E\left[\varphi(X_j,Z_j)m(X_j,Z_j,\beta_0)\gamma_0(Z_j)\frac{1_{\{Z_k<Z_i\}}-H(Z_i)}{(1-H(Z_i))^2}(1-\delta_i)1_{\{Z_i\le Z_j\}}|X_k,Z_k\right]=\psi(Z_k).$$

Hence, $E\left[U_n|X_k,Z_k\right]=\sum_{j\ne k}\psi(Z_k)=(n-1)\psi(Z_k)$. Put

$$\hat{U}_n = \sum_{k=1}^{n}E\left[U_n|X_k,Z_k\right]-(n-1)\theta = (n-1)\sum_{k=1}^{n}\psi(Z_k)$$

where $\theta = E\left[U_n\right] = E\left[E\left[U_n|X_k,Z_k\right]\right] = (n-1)E\left[\psi(Z_k)\right]=0$. The statistic $\frac{1}{n(n-1)}\hat{U}_n = \frac{1}{n}\sum_{k=1}^{n}\psi(Z_k)$ is the Hajek projection of the U-statistic $\frac{1}{n(n-1)}U_n$. Since

$$\frac{1}{n}\sum_{i=1}^{n}\psi(Z_k) = \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{H_n(v^-)-H(v)}{(1-H(v))^2}d\widetilde{H}^0_{X,Z}(u,v)dH_{X,Z}(s,t)$$

and $\frac{1}{n(n-1)}U_n - \frac{1}{n(n-1)}\hat{U}_n = O_P(n^{-1})$, this completes the proof. ■

Lemma 6. *If H is continuous, $E[\varphi(X,Y)^2]<\infty$, and (C3)–(C5) and (C8) are satisfied, then, as $n\to\infty$,*

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)H_{X,Z,n}(ds,dt)\\ &\quad = \frac{1}{n}\sum_{i=1}^{n}K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)H_{X,Z}(ds,dt)+O_P(n^{-1}).\end{aligned}$$

Proof. Denote the left-hand side of the equation in the lemma by V_n, and note that

$$V_n = \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_1(s,t,u,v)\widetilde{H}^1_{X,Z,n}(du,dv)H_{X,Z,n}(ds,dt) - \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_0(s,t,u,v)\widetilde{H}^0_{X,Z,n}(du,dv)H_{X,Z,n}(ds,dt) = V_n^1 - V_n^0$$

where

$$h_1(s,t,u,v) = \varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1}{m(u,v,\beta_0)}\int_{\mathbb{R}^p\times[0,t]}\frac{\alpha(x,y,u,v)}{1-H(y)}H_{X,Z}(dx,dy),$$

$$h_0(s,t,u,v) = \varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1}{1-m(u,v,\beta_0)}\int_{\mathbb{R}^p\times[0,t]}\frac{\alpha(x,y,u,v)}{1-H(y)}H_{X,Z}(dx,dy),$$

$$\widetilde{H}^1_{X,Z,n}(u,v) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i\le u, Z_i\le v, \delta_i=1\}}$$

and

$$\widetilde{H}^0_{X,Z,n}(u,v) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i\le u, Z_i\le v, \delta_i=0\}}.$$

According to (C5), $Grad(m(.,\beta_0))$ is bounded and therefore we get for an appropriate constant $c>0$,

$$\left|\frac{1}{m(X_j,Z_j,\beta_0)}\int_{\mathbb{R}^p\times[0,t]}\frac{\alpha(u,v,X_j,Z_j)}{1-H(v)}H_{X,Z}(du,dv)\right| \le \frac{c}{1-H(t)}\sum_{s=1}^k\frac{|D_s m(X_j.Z_j,\beta_0)|}{m(X_j.Z_j,\beta_0)}.$$

The integrability assumption on φ, (C8), (C3) and Lemma 5.7.3 in Serfling (1980) guarantees that $V_n^1 = U_n^1 + O_P(n^{-1})$ where U_n^1 denotes the associated U-statistic. Note that

$$\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_1^2(s,t,u,v)\widetilde{H}^1_{X,Z}(du,dv)H_{X,Z}(ds,dt) \le \int_{\mathbb{R}^p\times\mathbb{R}^+}\left(\frac{\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)}{1-H(t)}\right)^2 H_{X,Z}(ds,dt)c^2\int_{\mathbb{R}^p\times\mathbb{R}^+}\left(\sum_{s=1}^k\frac{|D_s m(u,v,\beta_0)|}{m(u,v,\beta_0)}\right)^2 H_{X,Z}(du,dv) < \infty$$

and apply Theorem 5.3.2 in Serfling (1980) to get

$$V_n^1 = \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_1(s,t,u,v)\widetilde{H}^1_{X,Z,n}(du,dv)H_{X,Z}(ds,dt) + \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_1(s,t,u,v)\widetilde{H}^1_{X,Z}(du,dv)H_{X,Z,n}(ds,dt) - \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_1(s,t,u,v)\widetilde{H}^1_{X,Z}(du,dv)H_{X,Z}(ds,dt) + O_P(n^{-1}).$$

The same argumentation is valid for V_n^0. In particular:

$$V_n^0 = \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_0(s,t,u,v)\widetilde{H}^0_{X,Z,n}(du,dv)H_{X,Z}(ds,dt) + \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_0(s,t,u,v)\widetilde{H}^0_{X,Z}(du,dv)H_{X,Z,n}(ds,dt) - \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times\mathbb{R}^+} h_0(s,t,u,v)\widetilde{H}^0_{X,Z}(du,dv)H_{X,Z}(ds,dt) + O_P(n^{-1}).$$

Since

$$\begin{aligned}
&\int_{\mathbb{R}^p\times\mathbb{R}^+} h_1(s,t,u,v)\widetilde{H}^1_{X,Z}(du,dv) - \int_{\mathbb{R}^p\times\mathbb{R}^+} h_0(s,t,u,v)\widetilde{H}^0_{X,Z}(du,dv)\\
&= \int_{\mathbb{R}^p\times\mathbb{R}^+}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1}{m(u,v,\beta_0)}\left[\int_{\mathbb{R}^p\times[0,t]}\frac{\alpha(x,y,u,v)}{1-H(y)}H_{X,Z}(dx,dy)\right]m(u,v,\beta_0)H_{X,Z}(du,dv)\\
&\quad - \int_{\mathbb{R}^p\times\mathbb{R}^+}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1}{1-m(u,v,\beta_0)}\\
&\quad \times\left[\int_{\mathbb{R}^p\times[0,t]}\frac{\alpha(x,y,u,v)}{1-H(y)}H_{X,Z}(dx,dy)\right](1-m(u,v,\beta_0))H_{X,Z}(du,dv) = 0
\end{aligned}$$

we get

$$V_n = n^{-1}\sum_{i=1}^{n} K(X_i, Z_i, \delta_i) \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]} \varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)} H_{X,Z}(du,dv)H_{X,Z}(ds,dt) + O_P(n^{-1})$$

which completes the proof of the lemma. ■

Now combine last four lemmas and note that some terms cancel out to get the following result.

Lemma 7. *If H is continuous, $E[\varphi(X,Y)^2] < \infty$, and (C1)–(C5) and (C8) are satisfied, then, as $n \to \infty$,*

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\varphi(X_i,Z_i)m(X_i,Z_i,\beta_0)\gamma_0(Z_i)C_{i,n}(\beta_n)\\
&\quad= \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1}{1-H(v)}H^0_{X,Z,n}(du,dv)H_{X,Z}(ds,dt)\\
&\quad\quad - \int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{1-H_n(v)}{(1-H(v))^2}\widetilde{H}^0_{X,Z}(du,dv)H_{X,Z}(ds,dt)\\
&\quad\quad - n^{-1}\sum_{i=1}^{n}K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p\times\mathbb{R}^+}\int_{\mathbb{R}^p\times[0,t]}\varphi(s,t)m(s,t,\beta_0)\gamma_0(t)\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)H_{X,Z}(ds,dt) + o_P(n^{-1/2}),
\end{aligned}$$

where $H^0_{X,Z,n}(u,v) = \frac{1}{n}\sum_{i=1}^{n}(1-m(X_i,Z_i,\beta_0))1_{\{X_i\le u, Z_i\le v\}}$. ■

For the next term in Eq. (3), namely

$$\frac{1}{n}\sum_{i=1}^{n}\varphi(X_i,Z_i)m(X_i,Z_i,\beta_0)\gamma_0(Z_i)B_{i,n}(\beta_n),$$

where recall that

$$B_{i,n}(\beta) = n\int_{\mathbb{R}^p}\int_0^{Z_i^-}\log\left(1+\frac{1-m(u,v,\beta_n)}{n(1-H_n(v))}\right)H_{X,Z,n}(du,dv) - \int_{\mathbb{R}^p}\int_0^{Z_i^-}\frac{1-m(u,v,\beta)}{1-H_n(v)}H_{X,Z,n}(du,dv)$$

we note that

$$x - \frac{x^2}{2} \le \log(1+x) \le x \quad \text{for } x \ge 0;$$

therefore,

$$\begin{aligned}
\frac{1-m(u,v,\beta_n)}{n(1-H_n(v))} - \frac{1}{2}\left(\frac{1-m(u,v,\beta_n)}{n(1-H_n(v))}\right)^2 &\le \log\left(1+\frac{1-m(u,v,\beta_n)}{n(1-H_n(v))}\right)\\
&\le \frac{1-m(u,v,\beta_n)}{n(1-H_n(v))}
\end{aligned}$$

which implies

$$B_{i,n}(\beta_n) \le n\int_{\mathbb{R}^p}\int_0^{Z_i^-}\frac{1-m(u,v,\beta_n)}{n(1-H_n(v))}H_{X,Z,n}(du,dv) - \int_{\mathbb{R}^p}\int_0^{Z_i^-}\frac{1-m(u,v,\beta_n)}{1-H_n(v)}H_{X,Z,n}(du,dv) = 0$$

and we get

$$-\frac{n}{2}\int_{\mathbb{R}^p}\int_0^{Z_i^-}\frac{(1-m(u,v,\beta_n))^2}{n^2(1-H_n(v))^2}H_{X,Z,n}(du,dv) \le B_{i,n}(\beta_n) \le 0. \tag{5}$$

The SLLN, Glivenko–Cantelli and (C8) then yield the following result.

Lemma 8. *If H is continuous, $E[|\varphi(X,Y)|] < \infty$, and (C8) is satisfied, then, as $n \to \infty$,*

$$\frac{1}{n}\sum_{i=1}^{n}\varphi(X_i,Z_i)m(X_i,Z_i,\beta_0)\gamma_0(Z_i)B_{i,n}(\beta_n) = O(n^{-1})$$

with probability 1. ■

The next result gives a representation of the second term in Eq. (3) as a sum of i.i.d. random variables.

Lemma 9. *If H is continuous, $E[\varphi(X,Y)^2] < \infty$, and* (C1)–(C6) *and* (C8) *are satisfied, then, as $n \to \infty$,*

$$\frac{1}{n}\sum_{i=1}^{n}\varphi(X_i,Z_i)(m(X_i,Z_i,\beta_n)-m(X_i,Z_i,\beta_0))\gamma_0(Z_i)(1+B_{i,n}(\beta_n)+C_{i,n}(\beta_n))$$
$$= n^{-1}\sum_{i=1}^{n}K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p\times\mathbb{R}^+}\varphi(s,t)\gamma_0(t)\alpha(s,t,X_i,Z_i)H_{X,Z}(ds,dt)+o_P(n^{-1/2}).$$

Proof. By arguments similar to those in Lemma 3.5 in Dikta (1998) extended to covariates, we have by a Taylor expansion:

$$n^{1/2}\left(m(x,z,\beta_n)-m(x,z,\beta_0)\right) = n^{1/2}\left\langle Grad\left(m(x,z,\beta_0)\right)|\beta_n-\beta_0\right\rangle$$
$$+\frac{n^{1/2}}{2}\sum_{1\le r,s\le k}D_{r,s}(m(x,z,\beta^*))(\beta_{nr}-\beta_{0r})(\beta_{ns}-\beta_{0s}),$$

where $\beta^* \in \Omega$ in the interior of the line segment connecting β_n and β_0. Since $n^{1/2}(\beta_n-\beta_0)$ is asymptotically normal, (C1) and (C5) yield

$$\frac{n^{1/2}}{2}\left|\sum_{1\le r,s\le k}D_{r,s}(m(x,z,\beta^*))(\beta_{nr}-\beta_{0r})(\beta_{ns}-\beta_{0s})\right| \le M(x,z).O_P(n^{-1/2})+o_P(1)$$

where $o_P(1)$ is not depending on (x,z). Now apply (C5) and (C4) and proceed similarly to Dikta (1998, Theorem 2.3), to get for $0 \le z \le T$

$$n^{1/2}\left\langle Grad\left(m(x,z,\beta_0)\right)|\beta_n-\beta_0\right\rangle = n^{-1/2}\sum_{i=1}^{n}\left\langle Grad(m(x,z,\beta_0))|I^{-1}(\beta_0)Grad(w(\delta_i,X_i,Z_i,\beta_0))\right\rangle + o_P(1)$$
$$= n^{-1/2}\sum_{i=1}^{n}K(X_i,Z_i,\delta_i)\alpha(x,z,X_i,Z_i)+o_P(1).$$

So, proceeding as for term J_5 in Lemma A.1 in the Appendix, we observe that uniformly on $\mathbb{R}^p \times [0,T]$,

$$m(x,z,\beta_n)-m(x,z,\beta_0) = n^{-1}\sum_{i=1}^{n}K(X_i,Z_i,\delta_i)\alpha(x,z,X_i,Z_i)+O_P(n^{-1}). \tag{6}$$

Therefore, (C8) and the SLLN yield

$$n^{-1}\sum_{i=1}^{n}\varphi\left(X_i,Z_i\right)\left(m\left(X_i,Z_i,\beta_n\right)-m\left(X_i,Z_i,\beta_0\right)\right)\gamma_0(Z_i)$$
$$= n^{-1}\sum_{j=1}^{n}K(X_j,Z_j,\delta_j)\int_{\mathbb{R}^p\times\mathbb{R}^+}\varphi(s,t)\gamma_0(t)\alpha(s,t,X_j,Z_j)H_{X,Z,n}(ds,dt)+O_P(n^{-1})$$
$$= n^{-1}\sum_{j=1}^{n}K(X_j,Z_j,\delta_j)\int_{\mathbb{R}^p\times\mathbb{R}^+}\varphi(s,t)\gamma_0(t)\alpha(s,t,X_j,Z_j)H_{X,Z}(ds,dt)+O_P(n^{-1})$$

where the last step follows by the same argument used in the proof of Lemma 6. Now, (C8) and inequality (5), Lemma 1, and the SLLN imply that

$$n^{-1}\sum_{i=1}^{n}\varphi(X_i,Z_i)(m(X_i,Z_i,\beta_n)-m(X_i,Z_i,\beta_0))\gamma_0(Z_i)B_{i,n}(\beta_n) = O_P(n^{-3/2}).$$

For the last term, use Eq. (4) and Lemma A.1 to get

$$n^{-1}\sum_{i=1}^{n}\varphi(X_i,Z_i)(m(X_i,Z_i,\beta_n)-m(X_i,Z_i,\beta_0))\gamma_0(Z_i)C_{i,n}(\beta_n)$$
$$= n^{-1}\sum_{i=1}^{n}\varphi(X_i,Z_i)(m(X_i,Z_i,\beta_n)-m(X_i,Z_i,\beta_0))\gamma_0(Z_i)C_n(\beta_0,Z_i)+o_P(n^{-1/2})$$

where $C_n(\beta_0, v)$ is given in Lemma 3. Now, we have:

$$\sup_{0\le t\le T}|C_n(\beta_0,t)| = O_P(n^{-1/2}). \tag{7}$$

See Appendix, Lemma A.2. Hence, from Lemma 1, the SLLN and (C8) we get the result. ■

The following lemma shows that the third term at the right-hand side of representation (3) is negligible, and thus we obtain the i.i.d. representation of Theorem 1 under condition (C8).

Lemma 10. *If H is continuous, $E[\varphi(X,Y)^2]<\infty$, and (C1)–(C6) and (C8) are satisfied, then, as $n\to\infty$,*

$$\frac{1}{2n}\sum_{i=1}^{n}\varphi(X_i,Z_i)m(X_i,Z_i,\beta_n)\exp(\Delta_i)(B_{i,n}(\beta_n)+C_{i,n}(\beta_n))^2=O_P(n^{-1}).$$

Proof. Recall from representation (3) that Δ_i lies between the terms

$$B_{i,n}(\beta_n)+\int_{\mathbb{R}^p}\int_0^{Z_i^-}\frac{1-m(u,v,\beta_n)}{1-H_n(v)}H_{X,Z,n}(du,dv)$$

and

$$\int_{\mathbb{R}^p}\int_0^{Z_i}\frac{1-m(u,v,\beta_0)}{1-H(v)}H_{X,Z}(du,dv).$$

Therefore, we get due to (C8), Glivenko–Cantelli, and inequality (5)

$$\sup_{i:Z_i\le T}\{\Delta_i\}\le c\quad\text{for an appropriate constant } c.$$

Now the term on the left-hand side in our lemma is bounded by

$$c\iint|\varphi(u,v)|H_{X,Z,n}(du,dv)\times\sup_{i:Z_i\le T}\left\{\left(B_{i,n}(\beta_n)+C_{i,n}(\beta_n)\right)^2\right\}.$$

According to the SLLN and the bounds given under (4), Lemma A.1, (5) and (7), this term is $O_P(n^{-1})$. This concludes the proof. ■

In summary, we have seen that representation in Theorem 1 is valid if H is continuous, assumptions (C1)–(C6) and (C8) hold, and $E[\varphi(X,Y)^2]<\infty$. In the following, we prove that the support condition (C8) is not needed as long as assumptions (M1)–(M3) are fulfilled.

Proof to Theorem 1. Assume $\varphi\ge0$ without loss of generality. Define for the given φ and $T<\tau_H$, $\varphi_T(u,v)=\varphi(u,v)\,1_{\{v\le T\}}$. Then we get with $\varphi^T=\varphi-\varphi_T$

$$\begin{aligned}
S_n^{\varphi}&=S_n^{\varphi_T}+S_n^{\varphi^T}\\
&=S_n(\varphi_T)+S_n^{\varphi^T}+o_P\left(n^{-1/2}\right)\\
&=S_n(\varphi)-\left(S_n(\varphi)-S_n(\varphi_T)-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T dF_{X,Y}\right)+\left(S_n^{\varphi^T}-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T dF_{X,Y}\right)+o_P\left(n^{1/2}\right)\\
&\equiv S_n(\varphi)+P_{1,n}(T)+P_{2,n}(T)+o_P\left(n^{1/2}\right)
\end{aligned}$$

where $S_n(\varphi)=\frac{1}{n}\sum_{i=1}^n\xi^{\varphi}(X_i,Z_i,\delta_i;\beta_0)$ is the leading term in representation (2). Consider first the $P_{1,n}(T)$ term to get explicitly

$$\begin{aligned}
n^{1/2}P_{1,n}(T)=\;&n^{-1/2}\sum_{i=1}^n\left(\varphi^T(X_i,Z_i)\,\gamma_0(Z_i)\,m(X_i,Z_i,\beta_0)-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)F_{X,Y}(du,dv)\right)\\
&+n^{-1/2}\sum_{i=1}^n\left(\frac{1-m(X_i,Z_i,\beta_0)}{1-H(Z_i)}\int_{\mathbb{R}^p}\int_{Z_i}^{\tau_H}\varphi^T(u,v)F_{X,Y}(du,dv)\right.\\
&\left.-\int_{\mathbb{R}^p}\int_0^{\tau_H}\int_{\mathbb{R}^p}\int_z^{\tau_H}\frac{\varphi^T(u,v)}{1-H(z)}F_{X,Y}(du,dv)\tilde{H}^0_{X,Z}(dx,dz)\right)\\
&-n^{-1/2}\sum_{i=1}^n\left(\left(\int_{\mathbb{R}^p}\int_0^{Z_i}\int_{\mathbb{R}^p}\int_s^{\tau_H}\frac{\varphi^T(u,v)}{(1-H(s))^2}F_{X,Y}(du,dv)\tilde{H}^0_{X,Z}(dx,ds)\right.\right.\\
&\left.\left.-\int_{\mathbb{R}^p}\int_0^{\tau_H}\int_{\mathbb{R}^p}\int_z^{\tau_H}\frac{\varphi^T(u,v)}{1-H(z)}F_{X,Y}(du,dv)\tilde{H}^0_{X,Z}(dx,dz)\right)\right.\\
&-n^{-1/2}\sum_{i=1}^n K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{\tau_H}\int_{\mathbb{R}^p}\int_s^{\tau_H}\frac{\alpha(u,s,X_i,Z_i)\,\varphi^T(\tilde{u},\tilde{v})}{1-H(s)}F_{X,Y}(d\tilde{u},d\tilde{v})H_{X,Z}(du,ds)
\end{aligned}$$

$$+\,n^{-1/2}\sum_{i=1}^{n}K\left(X_i,Z_i,\delta_i\right)\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)\gamma_0(v)\alpha(u,v,X_i,Z_i)H_{X,Z}(du,dv)$$
$$\equiv P_{1,1,n}\left(T\right)+P_{1,2,n}(T)-P_{1,3,n}(T)-P_{1,4,n}(T)+P_{1,5,n}(T).$$

Note that each of these terms consists of a centered i.i.d. sum. Therefore, it is sufficient to show that the variances can be made arbitrarily small choosing T large enough.

For $P_{1,1,n}\left(T\right)$ we get:

$$\begin{aligned}
Var\left[P_{1,1,n}\left(T\right)\right] &= Var\left[n^{-1/2}\sum_{i=1}^{n}\left(\varphi^T\left(X_i,Z_i\right)\gamma_0\left(Z_i\right)m\left(X_i,Z_i,\beta_0\right)-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)F_{X,Y}(du,dv)\right)\right]\\
&= n^{-1/2}\sum_{i=1}^{n}\left(\varphi^T\left(X_i,Z_i\right)\gamma_0\left(Z_i\right)m\left(X_i,Z_i,\beta_0\right)-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)F_{X,Y}(du,dv)\right)\\
&= Var\left[\left(\varphi^T\left(X_1,Z_1\right)\gamma_0\left(Z_1\right)m\left(X_1,Z_1,\beta_0\right)\right)\right]\\
&\le E\left[\left(\varphi^T\left(X_1,Z_1\right)^2\gamma_0^2\left(Z_1\right)m\left(X_1,Z_1,\beta_0\right)\right)^2\right]\\
&\le E\left[\varphi^T\left(X_1,Z_1\right)^2\gamma_0^2\left(Z_1\right)m\left(X_1,Z_1,\beta_0\right)\right]\quad(\text{since } 0\le m\le 1)\\
&= \int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)^2\gamma_0^2(v)m\left(u,v,\beta_0\right)H_{X,Z}(du,dv)\\
&= \int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)^2\gamma_0(v)F_{X,Y}(du,dv).
\end{aligned}$$

According to condition (M1) this bound can be made arbitrarily small choosing T large enough. The variance of the terms $P_{1,j,n}(T)$, $2\le j\le 5$, can be analyzed similarly; condition (M2) is needed for the cases $2\le j\le 4$, while (M3) serves to prove $Var[P_{1,5,n}(T)]\to 0$ as $T\to\tau_H$ (the lengthy details are omitted).

Now, consider $P_{2,n}\left(T\right)$. Applying Lemma 2 we get

$$\begin{aligned}
n^{1/2}P_{2,n}(T) &= n^{-1/2}\sum_{i=1}^{n}\left(\varphi^T\left(X_i,Z_i\right)m\left(X_i,Z_i,\beta_n\right)\gamma_0\left(Z_i\right)\exp\left(B_{i,n}\left(\beta_n\right)+C_{i,n}(\beta_n)\right)-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)F_{X,Y}(du,dv)\right)\\
&= n^{1/2}\left(\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)m\left(u,v,\beta_0\right)\gamma_0(v)H_{X,Z,n}(du,dv)-\int_{\mathbb{R}^p}\int_0^{\tau_H}\varphi^T(u,v)F_{X,Y}(du,dv)\right)\\
&\quad+n^{-1/2}\sum_{i=1}^{n}\varphi^T(X_i,Z_i)m\left(X_i,Z_i,\beta_0\right)\gamma_0(Z_i)\left(\exp\left(B_{i,n}(\beta_n)\right)+C_{i,n}\left(\beta_n\right)-1\right)\\
&\quad+n^{-1/2}\sum_{i=1}^{n}\varphi^T\left(X_i,Z_i\right)\gamma_0\left(Z_i\right)\left(m\left(X_i,Z_i,\beta_n\right)-m\left(X_i,Z_i,\beta_0\right)\right)\exp\left(B_{i,n}(\beta_n)+C_{i,n}(\beta_n)\right)\\
&\equiv P_{2,1,n}\left(T\right)+P_{2,2,n}\left(T\right)+P_{2,3,n}\left(T\right).
\end{aligned}$$

Since $P_{2,1,n}\left(T\right)=P_{1,1,n}\left(T\right)$ this term can be neglected. For the second and third term, we first observe that, due to (5), we get $\sup_{1\le i\le n}\left|B_{i,n}\left(\beta_n\right)\right|\le 1$. Thus, Lemma A.3 in the Appendix gives

$$\sup_{1\le i\le n}\exp\left(\left|B_{i,n}\left(\beta_n\right)\right|+\left|C_{i,n}\left(\beta_n\right)\right|\right)=O_P\left(1\right). \tag{8}$$

Observe that

$$\left|P_{2,2,n}\left(T\right)\right|\le n^{-1/2}\sum_{i=1}^{n}\left|\varphi^T\left(X_i,Z_i\right)\right|m\left(X_i,Z_i,\beta_0\right)\gamma_0\left(Z_i\right)\left(\left|B_{i,n}\left(\beta_n\right)\right|+\left|C_{i,n}\left(\beta_n\right)\right|\right)\exp\left(\left|B_{i,n}\left(\beta_n\right)\right|+\left|C_{i,n}\left(\beta_n\right)\right|\right).$$

It remains to consider

$$n^{-1/2}\sum_{i=1}^{n}\left|\varphi^T\left(X_i,Z_i\right)\right|m\left(X_i,Z_i,\beta_0\right)\gamma_0\left(Z_i\right)\left(\left|B_{i,n}\left(\beta_n\right)\right|+\left|C_{i,n}\left(\beta_n\right)\right|\right).$$

But the term

$$n^{-1/2}\sum_{i=1}^{n}\left|\varphi^T\left(X_i,Z_i\right)\right|m\left(X_i,Z_i,\beta_0\right)\gamma_0\left(Z_i\right)\left|B_{i,n}\left(\beta_n\right)\right|$$

can be neglected, in the sense that it is $O_P(1)$, see Stute (1995, p. 436). We have to consider finally

$$A_n(T) \equiv n^{-1/2}\sum_{i=1}^{n}\left|\varphi^T(X_i,Z_i)\right| m(X_i,Z_i,\beta_0)\,\gamma_0(Z_i)\left|C_{i,n}(\beta_n)\right|.$$

Recall the definition of $C_{i,n}(\beta_n)$ to bound $A_n(T)$ by

$$\begin{aligned}
A_n(T) \le\ & n^{-1/2}\sum_{i=1}^{n}\left|\varphi^T(X_i,Z_i)\right| m(X_i,Z_i,\beta_0)\,\gamma_0(Z_i)\\
&\times\left|\int_{\mathbb{R}^p}\int_0^{Z_i-}\frac{1-m(u,v,\beta_0)}{1-H_n(v)}H_{X,Z,n}(du,dv)-\int_{\mathbb{R}^p}\int\frac{1-m(u,v,\beta_0)}{1-H(v)}H_{X,Z}(du,dv)\right|\\
&+n^{-1/2}\sum_{i=1}^{n}\left|\varphi^T(X_i,Z_i)\right| m(X_i,Z_i,\beta_0)\,\gamma_0(Z_i)\left|\int_{\mathbb{R}^p}\int_0^{Z_i-}\frac{m(u,v,\beta_n)-m(u,v,\beta_0)}{1-H_n(v)}H_{X,Z,n}(du,dv)\right|\\
\equiv\ & P_{2,2,1,n}(T)+P_{2,2,2,n}(T).
\end{aligned}$$

Similar arguments to those in Stute (1995, p. 437), show that the term $P_{2,2,1,n}(T)$ is negligible. For the second term we get

$$\begin{aligned}
P_{2,2,2,n}(T) \le\ & n^{-1}\sum_{i=1}^{n}\left|\varphi^T(X_i,Z_i)\right| m(X_i,Z_i,\beta_0)\,\gamma_0(Z_i)\,(1-H(Z_i))^{-1/2}\sup_{\substack{0\le v\le\tau_H\\ u\in\mathbb{R}^p}}\left|n^{1/2}(m(u,v,\beta_n)-m(u,v,\beta_0))\right|\\
&\times\sup_{0\le v<Z_{(n)}}\left(\frac{1-H(v)}{1-H_n(v)}\right)\int_{\mathbb{R}^p}\int_0^{\tau_H}\frac{1}{(1-H(v))^{1/2}}H_{X,Z,n}(du,dv).
\end{aligned}$$

Due to Lemma 1,

$$\sup_{\substack{0\le v\le\tau_H\\ u\in\mathbb{R}^p}}\left|n^{1/2}(m(u,v,\beta_n)-m(u,v,\beta_0))\right|\quad\text{is stochastically bounded.}$$

The same holds for $\sup_{0\le v<Z_{(n)}}\left(\frac{1-H(v)}{1-H_n(v)}\right)$, see Shorack and Wellner (1986, p. 415). According to SLLN and (M2), this term $P_{2,2,2,n}(T)$ can be neglected. Finally the term $P_{2,3,n}(T)$ can be bounded by

$$P_{2,3,n}(T) \le n^{-1}\sum_{i=1}^{n}\left|\varphi^T(X_i,Z_i)\right|\gamma_0(Z_i)\sup_{\substack{0\le v\le\tau_H\\ u\in\mathbb{R}^p}}\left|n^{1/2}(m(u,v,\beta_n)-m(u,v,\beta_0))\right|\sup_{1\le i\le n}\exp\left(\left|B_{i,n}(\beta_n)\right|+\left|C_{i,n}(\beta_n)\right|\right).$$

The term can be neglected, according to Lemma 1, (8), and (M3). This completes the proof of the theorem. ■

Proof to Corollary 2. Since the estimator θ_n is the minimizer of

$$\theta \longmapsto \sum_{i=1}^{n} W_{(i)}(m_n)(Z_{(i)}-X_{[i]}^t\theta)^2,$$

we have $\theta_n = M_{1n}^{-1}M_{2n}\widetilde{Z}_n$, where $\widetilde{Z}_n = (Z_{(1)},\ldots,Z_{(n)})^t$ and the matrices M_{1n} and M_{2n} have entries

$$M_{1n}(i,j) = \sum_{k=1}^{n} W_{(k)}(m_n)X_{[k]}^iX_{[k]}^j,\quad 1\le i,j\le p,$$

and

$$M_{2n}(i,s) = W_{(s)}(m_n)X_{[s]}^i,\quad 1\le i\le p,\ 1\le s\le n,$$

respectively, and where $X_{[s]}^i$ denotes the ith coordinate of $X_{[s]}$. Introduce $\underline{S}_n^{\varphi} = (S_n^{\varphi_1},\ldots,S_n^{\varphi_p})^t$. Note that, since $\tau_F=\tau_H$, we have $S^{\varphi_i}=0$, $1\le i\le p$. From Theorem 1 and the multivariate CLT we thus have

$$n^{1/2}\underline{S}_n^{\varphi}\to N(\underline{0},\Sigma)$$

in distribution, where Σ is the variance–covariance matrix defined in Corollary 2. Since the functions $\varphi_{(i,j)}$ satisfy (M1)–(M3) we also have from Theorem 1 and the WLLN $M_{1n}\to\Sigma_0=E[XX^t]$ in probability. Therefore, $n^{1/2}(\theta_n-\theta_0)=n^{1/2}M_{1n}^{-1}\underline{S}_n^{\varphi}\to N(\underline{0},\Sigma_0^{-1}\Sigma\Sigma_0^{-1})$ in distribution and the proof is finished. ■

Proof to Corollary 3. The proof follows exactly the lines of the proof to Corollary 2.5 in Dikta et al. (2005). ■

Acknowledgment

Work supported by the Grant MTM2011-23204 (FEDER support included) of the Spanish Ministerio de Ciencia e Innovación, and by the Grant MTM2014-55966-P of the Spanish Ministerio de Economía y Competitividad.

Appendix. Auxiliary results

In this section we give some auxiliary results needed in the proofs of the main lemmas of Section 4. When needed, $C[0,T]$ and $D[0,T]$ will denote, respectively, the space of the continuous functions and the space of the right-continuous functions with left-hand limits on $[0,T]$.

Appendix A

Lemma A.1. *Under the conditions in Lemma 3 we have*

$$\sup_{0\le v\le T}\left|\Lambda_n^0(v)-\Lambda^0(v)-C_n(\beta_0,v)\right|=o_P(n^{-1/2}).$$

Proof. We start by following lines similar to those in the proof to Lemma 3.7 in Dikta (1998). Write

$$\begin{aligned}\frac{1-m(u,v,\beta_n)}{1-H_n(v)}=&\frac{1-m(u,v,\beta_0)}{1-H(v)}+\frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}\\&+\frac{m(u,v,\beta_0)-m(u,v,\beta_n)}{1-H(v)}+\frac{(1-m(u,v,\beta_n))(H_n(v)-H(v))^2}{(1-H_n(v))(1-H(v))^2}\\&+\frac{(m(u,v,\beta_0)-m(u,v,\beta_n))\,(H_n(v)-H(v))}{(1-H(v))^2}.\end{aligned}$$

The difference becomes

$$\begin{aligned}&\Lambda_n^0(t)-\Lambda^0(t)-C_n(\beta_0,t)\\&=\int_{\mathbb{R}^p}\int_0^{t^-}\frac{1-m(u,v,\beta_n)}{1-H_n(v)}H_{X,Z,n}(du,dv)-\int_{\mathbb{R}^p}\int_0^{t}\frac{1-m(u,v,\beta_0)}{1-H(v)}H_{X,Z}(du,dv)\\&\quad-\int_{\mathbb{R}^p}\int_0^{t^-}\frac{1-m\,(u,v,\beta_0)}{1-H(v)}H_{X,Z,n}(du,dv)+\int_{\mathbb{R}^p}\int_0^{t}\frac{1-m(u,v,\beta_0)}{1-H(v)}H_{X,Z}(du,dv)\\&\quad-\int_{\mathbb{R}^p}\int_0^{t}\frac{(1-m(u,v,\beta_0))\left(H_n(v^-)-H(v)\right)}{(1-H(v))^2}H_{X,Z}(du,dv)\\&\quad+\frac{1}{n}\sum_{i=1}^n K\,(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{t}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)\\&=\int_{\mathbb{R}^p}\int_0^{t^-}\frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z,n}(du,dv)+\int_{\mathbb{R}^p}\int_0^{t^-}\frac{m(u,v,\beta_0)-m(u,v,\beta_n)}{1-H(v)}H_{X,Z,n}(du,dv)\\&\quad+\int_{\mathbb{R}^p}\int_0^{t^-}\frac{(1-m(u,v,\beta_n))(H_n(v)-H(v))^2}{(1-H_n(v))(1-H(v))^2}H_{X,Z,n}(du,dv)\\&\quad+\int_{\mathbb{R}^p}\int_0^{t^-}\frac{(m(u,v,\beta_0)-m(u,v,\beta_n))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z,n}(du,dv)\\&\quad-\int_{\mathbb{R}^p}\int_0^{t}\frac{(1-m(u,v,\beta_0))(H_n(v^-)-H(v))}{(1-H(v))^2}H_{X,Z}(du,dv)\\&\quad+\frac{1}{n}\sum_{i=1}^n K\,(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{t}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)\\&=J_2+J_3+J_4+J_5+J_6+n^{-1}\sum_{i=1}^n K\,(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{t}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv).\end{aligned}$$

Since $\left|H_n(v^-) - H_n(v)\right| \le 1/n$, in J_6 the term $H_n(v^-)$ may be replaced by $H_n(v)$ in the sense that the difference is $O(n^{-1})$ w.p.1 uniformly on $0 \le t \le T$. With this in mind, the sum $J_2 + J_6$ equals

$$\int_{\mathbb{R}^p}\int_0^{t^-} \frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}\left(H_{X,Z,n}(du,dv) - H_{X,Z}(du,dv)\right) + O(n^{-1})$$

uniformly in $0 \le t \le T$. Following lines similar to those in Lemma 3.8 in Dikta (1998) we have, as $n \to \infty$,

$$n^{1/2}\sup_{0\le t\le T}\left|\int_{\mathbb{R}^p}\int_0^{t^-} \frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}\left(H_{X,Z,n}(du,dv) - H_{X,Z}(du,dv)\right)\right| \to 0$$

in probability. See Lemmas B.1 and B.2.

For the term J_5 we get

$$\begin{aligned}
&n^{1/2}\sup_{0\le t\le T}\left|\int_{\mathbb{R}^p}\int_0^{t^-} \frac{(m(u,v,\beta_0)-m(u,v,\beta_n))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z,n}(du,dv)\right| \\
&\le n^{1/2}\frac{\|H_n - H\|}{(1-H(T))^2} \times \int_{\mathbb{R}^p}\int_0^{T} |m(u,v,\beta_0)-m(u,v,\beta_n)|\, H_{X,Z,n}(du,dv).
\end{aligned}$$

Since $n^{1/2}\,\|H_n - H\|$ is bounded in probability, Lemma 1 together with the SLLN guarantees that

$$n^{1/2}\sup_{0\le t\le T}\left|\int_{\mathbb{R}^p}\int_0^{t^-} \frac{(m(u,v,\beta_0)-m(u,v,\beta_n))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z,n}(du,dv)\right| = O_P(n^{-1/2}).$$

For the term J_4, take $\epsilon > 0$ such that $H(T) + \epsilon < 1$. Then we get with probability one for large n

$$n^{1/2}\sup_{0\le t\le T}\left|\int_{\mathbb{R}^p}\int_0^{t^-} \frac{(1-m(u,v,\beta_n))(H_n(v)-H(v))^2}{(1-H_n(v))(1-H(v))^2}H_{X,Z,n}(du,dv)\right| \le n^{1/2}\frac{\|H_n - H\|^2}{(1-H(T)-\epsilon)^3}.$$

Since $n^{1/2}\,\|H_n - H\|$ is bounded, the right hand side tends to zero in probability according to Glivenko–Cantelli theorem.

For the term J_3, we follow steps similar to those in Lemma 3.5 in Dikta (1998) to get

$$\begin{aligned}
&\sup_{0\le t\le T}\left|n^{1/2}\int_{\mathbb{R}^p}\int_0^{t^-} \frac{m(u,v,\beta_n)-m(u,v,\beta_0)}{1-H(v)}H_{X,Z,n}(du,dv)\right. \\
&\qquad\left. - n^{-1/2}\sum_{i=1}^n K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{t^-}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z,n}(du,dv)\right| \\
&\le \int_{\mathbb{R}^p}\int_0^{T}\frac{|R_n(u,v)|}{1-H(v)}H_{X,Z,n}(du,dv)
\end{aligned}$$

where $R_n(u,v)$ is a remainder term such that $|R_n(u,v)| \le M(u,v)O_P(n^{-1/2}) + o_P(1)$ uniformly on $u \in \mathbb{R}^p$, $0 \le v \le T$, $O_P(1)$ and $o_P(1)$ not depending on u and v, and $M(u,v)$ is given in (C5). Hence,

$$\begin{aligned}
&\int_{\mathbb{R}^p}\int_0^{T}\frac{|R_n(u,v)|}{1-H(v)}H_{X,Z,n}(du,dv) \\
&\le O_P(n^{-1/2})\int_{\mathbb{R}^p}\int_0^{T}\frac{|M(u,v)|}{1-H(v)}H_{X,Z,n}(du,dv) + o_P(1)\int_{\mathbb{R}^p}\int_0^{T}\frac{H_{X,Z,n}(du,dv)}{1-H(v)} \\
&\le O_P(n^{-1/2})c\int_{\mathbb{R}^p}\int_0^{T}|M(u,v)|\,H_{X,Z,n}(du,dv) + o_P(1)c\int_{\mathbb{R}^p}\int_0^{T}H_{X,Z,n}(du,dv) \\
&= o_P(1).
\end{aligned}$$

It remains to prove that the difference

$$n^{-1}\sum_{i=1}^n K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{t}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z,n}(du,dv) - n^{-1}\sum_{i=1}^n K(X_i,Z_i,\delta_i)\int_{\mathbb{R}^p}\int_0^{t}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)$$

tends to 0 in probability at rate $n^{-1/2}$ uniformly on $t \in [0,T]$. This is done with the help of the next two lemmas, which extend results in Dikta (1998) to the setting with covariates. ■

Lemma A.1.1. *If H is continuous, and* (C1)–(C5) *are satisfied, then, with probability one, as $n \to \infty$*

$$n^{-1/2}\left(\sum_{i=1}^{n} K_i \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u,v,X_i,Z_i)}{1-H(v)} H_{X,Z,n}(du,dv) - \sum_{i=1}^{n} K_i \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u,v,X_i,Z_i)}{1-H(v)} H_{X,Z}(du,dv)\right)$$
$$= n^{1/2}U_n(t) + O(n^{-1/2})$$

uniformly on $[0,T]$, *where*

$$U_n(t) = \frac{1}{n(n-1)} \sum_{1\le i\ne j\le n} K_i\left(\frac{\alpha(X_j,Z_j,X_i,Z_i)}{1-H(Z_j)} 1_{\{Z_j\le t\}} - \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u,v,X_i,Z_i)}{1-H(v)} H_{X,Z}(du,dv)\right)$$

and

$$K_i = \frac{\delta_i - m(X_i,Z_i,\beta_0)}{m(X_i,Z_i,\beta_0)(1-m(X_i,Z_i,\beta_0))}.$$

Proof. Write $I_i = 1_{\{Z_i\le t\}}$, $A_i = \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u,v,X_i,Z_i)}{1-H(v)} H_{X,Z}(du,dv)$, $\alpha_{j,i} = \alpha(X_j,Z_j,X_i,Z_i)$, $m_i = m(X_i,Z_i,\beta_0)$ and $H_j = H(Z_j)$. Then we have:

$$\begin{aligned}
&n^{-1/2}\left(\sum_{i=1}^{n} K_i\left(\int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u,v,X_i,Z_i)}{1-H(v)} H_{X,Z,n}(du,dv) - A_i\right)\right)\\
&= n^{-1/2}\left(\sum_{i=1}^{n} K_i\left(\frac{1}{n}\sum_{j=1}^{n}\left(\frac{\alpha_{j,i}}{1-H_j} I_j - A_i\right)\right)\right)\\
&= n^{-1/2}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} K_i \frac{1}{n}\frac{\alpha_{j,i}}{1-H_j} I_j - \sum_{i=1}^{n}\sum_{j=1}^{n} K_i \frac{1}{n} A_i\right)\\
&= n^{-1/2}\left(\frac{1}{n}\sum_{1\le i\ne j\le n} \frac{\delta_i\alpha_{j,i} - m_i\alpha_{j,i}}{m_i(1-m_i)(1-H_j)} I_j + \frac{1}{n}\sum_{i=1}^{n} K_i \frac{\alpha_{i,i}}{1-H_i} I_i - \frac{1}{n}\sum_{1\le i\ne j\le n} K_iA_i - \frac{1}{n}\sum_{i=1}^{n} K_iA_i\right). \qquad (9)
\end{aligned}$$

Define $R_n(t) = \frac{1}{n}\sum_{i=1}^{n} K_i \frac{\alpha_{i,i}}{1-H_i} I_i - \frac{1}{n}\sum_{i=1}^{n} K_iA_i$. Then:

$$\begin{aligned}
(9) &= n^{-1/2}\left(\sum_{1\le i\ne j\le n} \frac{\delta_i\alpha_{j,i} - m_i\alpha_{j,i}}{m_i(1-m_i)(1-H_j)}\left(\frac{n}{(n-1)n} - \frac{1}{n(n-1)}\right) I_j\right.\\
&\quad\left. - \sum_{1\le i\ne j\le n} K_iA_i\left(\frac{n}{(n-1)n} - \frac{1}{n(n-1)}\right) + R_n(t)\right)\\
&= n^{-1/2}\left(\frac{1}{n-1}\sum_{1\le i\ne j\le n} \frac{\delta_i\alpha_{j,i} - m_i\alpha_{j,i}}{m_i(1-m_i)(1-H_j)} I_j - \frac{1}{n(n-1)}\sum_{1\le i\ne j\le n} \frac{\delta_i\alpha_{j,i} - m_i\alpha_{j,i}}{m_i(1-m_i)(1-H_j)} I_j\right.\\
&\quad\left. - \frac{1}{n-1}\sum_{1\le i\ne j\le n} K_iA_i + \frac{1}{n(n-1)}\sum_{1\le i\ne j\le n} K_iA_i + R_n(t)\right)\\
&= n^{-1/2}\left(\frac{1}{n-1}\left(\sum_{1\le i\ne j\le n}\left(\frac{\delta_i\alpha_{j,i} - m_i\alpha_{j,i}}{m_i(1-m_i)(1-H_j)} I_j - K_iA_i\right)\right)\right.\\
&\quad\left. - \frac{1}{n(n-1)}\left(\sum_{1\le i\ne j\le n}\left(\frac{\delta_i\alpha_{j,i} - m_i\alpha_{j,i}}{m_i(1-m_i)(1-H_j)} I_j - K_iA_i\right)\right) + R_n(t)\right)\\
&= n^{-1/2}(nU_n(t) - U_n(t) + R_n(t)).
\end{aligned}$$

Since

$$n^{-1/2}(nU_n(t) - U_n(t) + R_n(t)) = n^{1/2}U_n(t) - n^{-1/2}U_n(t) + n^{-1/2}R_n(t)$$

it remains to show that, as $n \to \infty$, $\sup_{0\le t\le T} |U_n(t)| = O(1)$ and $\sup_{0\le t\le T} |R_n(t)| = O(1)$. Now, according to the Cauchy–Schwarz inequality,

$$\sup_{x\in\mathbb{R}^p, 0\le z\le T} |\alpha(x,z,X_i,Z_i)| \le L\cdot M\cdot \sum_{s=1}^{k} |D_s(m(X_i,Z_i,\beta_0))|,$$

where L and M are upper bounds for the norm of $Grad(m(x,z,\beta_0))$ and $I^{-1}(\beta_0)$ respectively, and so we have

$$\begin{aligned}\sup_{0\le t\le T}|U_n(t)| &= \sup_{0\le t\le T}\left|\frac{1}{n(n-1)}\sum_{1\le i\ne j\le n}K_i\left(\frac{\alpha(X_j,Z_j,X_i,Z_i)}{1-H(Z_j)}I_j-\int_0^t\int_{\mathbb{R}^p}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv)\right)\right|\\ &\le \frac{1}{n(n-1)}\sum_{1\le i\ne j\le n}\frac{1}{m_i(1-m_i)}\left(\frac{LM\sum_{s=1}^k|D_s(m(X_i,Z_i,\beta_0))|}{1-H(T)}+\frac{LM\sum_{s=1}^k|D_s(m(X_i,Z_i,\beta_0))|}{1-H(T)}\right)\\ &= \frac{2L\cdot M}{1-H(T)}\sum_{s=1}^k\frac{1}{n}\sum_{i=1}^n\frac{|D_sm(X_i,Z_i,\beta_0)|}{m(X_i,Z_i,\beta_0)(1-m(X_i,Z_i,\beta_0))}=O(1)\end{aligned}$$

by the SLLN, according to (C3). A similar calculation serves to prove $\sup_{0\le t\le T}|R_n(t)|=O(1)$, which proves the lemma. ∎

Lemma A.1.2. *Under the conditions of Lemma A.1.1, provided that* (C6) *holds, the process $\zeta_n(t)=n^{1/2}U_n(t)$ satisfies $\sup_{0\le t\le T}|\zeta_n(t)|\to 0$ in probability, where $U_n(t)$ is defined in Lemma A.1.1.*

Proof. The process $\zeta_n(t)$ is a degenerate U-statistic process. It can be shown that its finite dimensional distributions converge to zero, and that it is tight, and hence $\zeta_n(t)\to 0$ in probability in $D[0,T]$. To show that the finite dimensional distributions of $\zeta_n(t)$ converge to zero, it is enough to prove $Cov(\zeta_n(s),\zeta_n(t))\to 0$ as $n\to\infty$ for $0\le s\le t\le T$, which follows by direct calculations and is also established in Lemma B.3. ∎

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

$$\sup_{0\le t\le T}|C_n(\beta_0,t)|=O_P(n^{-1/2}).$$

Proof. According to the definition given in Lemma 3, the process $n^{1/2}C_n(\beta_0,t)$ is a sum of three centered processes defined on the space $D[0,T]$, namely $n^{1/2}C_n(\beta_0,t)=\alpha_n(t)+\beta_n(t)-\gamma_n(t)$ where

$$\begin{aligned}\alpha_n(t)&=n^{1/2}\int_0^t\int_{\mathbb{R}^p}\frac{1-m(u,v,\beta_0)}{1-H(v)}\left(H_{X,Z,n}(du,dv)-H_{X,Z}(du,dv)\right)\\ \beta_n(t)&=n^{1/2}\int_0^t\int_{\mathbb{R}^p}\frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z}(du,dv)\\ \gamma_n(t)&=n^{-1/2}\sum_{i=1}^n\frac{\delta_i-m(X_i,Z_i,\beta_0)}{m(X_i,Z_i,\beta_0)(1-m(X_i,Z_i,\beta_0))}\int_0^t\int_{\mathbb{R}^p}\frac{\alpha(u,v,X_i,Z_i)}{1-H(v)}H_{X,Z}(du,dv).\end{aligned}$$

Therefore, the convergence of the finite dimensional distributions of $n^{1/2}C_n(\beta_0,t)$ to a multivariate normal distribution is guaranteed by the multivariate CLT. Besides, each of the three processes above is tight. For $\alpha_n(t)$ and $\beta_n(t)$ tightness easily follows from Billingsley (1968, Theorem 16.4), and the continuous mapping theorem (the first integrability assumption in (C6.1) is needed for $\alpha_n(t)$). For $\gamma_n(t)$ one may prove tightness under (C6.2) as follows. Note that (C6.2) assures that $\gamma_n\in C[0,T]$, and that tightness in $C[0,T]$ implies tightness in $D[0,T]$. Write $K_i=\frac{\delta_i-m(X_i,Z_i,\beta_0)}{m(X_i,Z_i,\beta_0)(1-m(X_i,Z_i,\beta_0))}$. Then

$$\gamma_n(t)=n^{-1/2}\sum_{i=1}^n\sum_{1\le j,l\le k}K_i\bar{\sigma}_{j,l}D_lm(X_i,Z_i,\beta_0)\iint\frac{D_jm(u,v,\beta_0)}{1-H(v)}1_{\{v\le t\}}H_{X,Z}(du,dv).$$

Define

$$\gamma_n^{j,l}(t)=n^{-1/2}\sum_{i=1}^nK_i\bar{\sigma}_{j,l}D_lm(X_i,Z_i,\beta_0)\iint\frac{D_jm(u,v,\beta_0)}{1-H(v)}1_{\{v\le t\}}H_{X,Z}(du,dv).$$

Note that $\gamma_n(t)=\sum_{j,l}\gamma_n^{j,l}(t)$. Now, for tightness in $C[0,T]$, it suffices to show, for $s\le t$

$$E[(\gamma_n^{j,l}(t)-\gamma_n^{j,l}(s))^2]\le(Q(t)-Q(s))^{\frac{2+2\varepsilon}{2+\varepsilon}}$$

for an nondecreasing function Q. We have:

$$E[(\gamma_n^{j,l}(t)-\gamma_n^{j,l}(s))^2]=E\left[\left(n^{-1/2}\sum_{i=1}^nK_i\bar{\sigma}_{j,l}D_lm(X_i,Z_i,\beta_0)\right)^2\right]\left(\iint\frac{D_jm(u,v,\beta_0)}{1-H(v)}1_{\{v\le t\}}H_{X,Z}(du,dv)\right)^2.$$

For the first factor, the expected value of the squared term is 0 for some $i \neq j$ upon conditioning on (X_i, Z_i). Hence, $n^{-1}\bar{\sigma}^2_{j,l} nE\left[\left(\frac{\delta_i - m(X,Z,\beta_0)}{m(X,Z,\beta_0)(1-m(X,Z,\beta_0))} D_l m(X, Z, \beta_0)\right)^2\right]$ remains. Next,

$$\left(\iint \frac{D_j m(u, v, \beta_0)}{1 - H(v)} 1_{\{s\le v\le t\}} H_{X,Z}(du, dv)\right)^2 \le \left(\iint \left|\frac{D_j m(u, v, \beta_0)}{1 - H(v)}\right| 1_{\{s\le v\le t\}} H_{X,Z}(du, dv)\right)^2$$
$$= \left(E\left(\left|\frac{D_j m(X, Z, \beta_0)}{1 - H(Z)}\right| 1_{\{s\le Z\le t\}}\right)\right)^2.$$

Apply Hölder's inequality with $p = 2 + \varepsilon, q = \frac{2+\varepsilon}{1+\varepsilon}$ to see that the last term is less than or equal to

$$\left(E\left(\left|\frac{D_j m(X, Z, \beta_0)}{1 - H(Z)}\right|^{2+\varepsilon}\right)^{\frac{1}{2+\varepsilon}} E\left(1_{\{s\le Z\le t\}}\right)^{\frac{1+\varepsilon}{2+\varepsilon}}\right)^2 \le E\left(\left|\frac{D_j m(X, Z, \beta_0)}{1 - H(Z)}\right|^{2+\varepsilon}\right)^{\frac{2}{2+\varepsilon}} (H(t) - H(s))^{\frac{2+2\varepsilon}{2+\epsilon}}.$$

Thus,

$$E[(\gamma_n^{j,l}(t) - \gamma_n^{j,l}(s))^2] \le K(H(t) - H(s))^{\frac{2+2\varepsilon}{2+\epsilon}}$$

where $K = \bar{\sigma}^2_{j,l} E\left[\left(\frac{\delta_i - m(X,Z,\beta_0)}{m(X,Z,\beta_0)(1-m(X,Z,\beta_0))} D_l m(X, Z, \beta_0)\right)^2\right] E\left(\left|\frac{D_j m(X,Z,\beta_0)}{1-H(Z)}\right|^{2+\varepsilon}\right)^{\frac{2}{2+\varepsilon}}$ and $\frac{2+2\varepsilon}{2+\epsilon} > 1$.

Therefore, $n^{1/2} C_n(\beta_0, t)$ tends weakly to a centered Gaussian process which is concentrated on $D[0, T]$, and the result follows. ■

Lemma A.3. *Under the conditions in Theorem 1 we have*

$$\max_{1\le i\le n} \left|C_{i,n}(\beta_n)\right| = O_P(1).$$

Proof. Write

$$\int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_n)}{1 - H_n(v)} H_{X,Z,n}(du, dv)$$
$$= \int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_0)}{1 - H_n(v)} H_{X,Z,n}(du, dv) + \int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{m(u, v, \beta_0) - m(u, v, \beta_n)}{1 - H_n(v)} H_{X,Z,n}(du, dv).$$

By Lemma 1, the $\max_{1\le i\le n}$ of the second term is $O_P(n^{-1/2} \sum_{j=1}^{n-1} \frac{1}{n-j}) = O_P(1)$. Therefore, uniformly on $i = 1, \ldots, n$,

$$C_{i,n}(\beta_n) = \int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_0)}{1 - H_n(v)} H_{X,Z,n}(du, dv) - \int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z}(du, dv) + O_P(1).$$

By reasoning as in the proof to Lemma 3.1 in Dikta (2001), we get (take $r_n = n - 1$ in Lemma 3.1)

$$\int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_0)}{1 - H_n(v)} H_{X,Z,n}(du, dv) = \int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z,n}(du, dv)$$
$$+ \int_{\mathbb{R}^p} \int_0^{Z_i-} \frac{1 - m(u, v, \beta_0)(H_n(v) - H(v))}{(1 - H(v))^2} H_{X,Z}(du, dv) + O_P(1)$$

uniformly on the Z_i's (note that the upper integration limit at the right-hand side may be replaced by the preceding Z-statistic in the ordered sample). This shows that

$$\max_{1\le i\le n} \left|C_{i,n}(\beta_n)\right|$$
$$= \max_{1\le i\le n} \left|\int_{\mathbb{R}^p} \int_0^{Z_{(i-1)}} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z,n}(du, dv) - \int_{\mathbb{R}^p} \int_0^{Z_{(i)}} \frac{(1 - m(u, v, \beta_0))(1 - H_n(v))}{(1 - H(v))^2} H_{X,Z}(du, dv)\right|$$
$$+ O_P(1)$$
$$\le \max_{1\le i\le n} \left|\int_{\mathbb{R}^p} \int_0^{Z_{(i-1)}} \frac{1 - m(u, v, \beta_0)}{1 - H(v)} H_{X,Z,n}(du, dv) - \int_{\mathbb{R}^p} \int_0^{Z_{(i-1)}} \frac{(1 - m(u, v, \beta_0))(1 - H_n(v))}{(1 - H(v))^2} H_{X,Z}(du, dv)\right|$$

$$+\max_{1\le i\le n}\left|\int_{\mathbb{R}^p}\int_{Z_{(i-1)}}^{Z_{(i)}}\frac{(1-m(u,v,\beta_0))(1-H_n(v))}{(1-H(v))^2}H_{X,Z}(du,dv)\right|+O_P(1)$$
$$\equiv I+II+O_P(1).$$

For proving that the second term II is bounded in probability, note that

$$\max_{1\le i\le n}\left|\int_{\mathbb{R}^p}\int_{Z_{(i-1)}}^{Z_{(i)}}\frac{(1-m(u,v,\beta_0))(1-H_n(v))}{(1-H(v))^2}H_{X,Z}(du,dv)\right|$$
$$\le\max_{1\le i\le n}\frac{1-H_n(Z_i)}{1-H(Z_i)}\max_{1\le i\le n}\left|\int_{\mathbb{R}^p}\int_{Z_{(i-1)}}^{Z_{(i)}}\frac{1-m(u,v,\beta_0)}{1-H(v)}H_{X,Z}(du,dv)\right|.$$

Since Daniel's theorem (cfr. Shorack and Wellner, 1986, p. 345) implies that $\max_{1\le i\le n}(1-H_n(Z_i))/(1-H(Z_i))=O_P(1)$, one may proceed exactly as in the proof to Lemma 3.2 in Dikta (2001) to conclude $II=O_P(1)$. For the first term I, we proceed similarly to pages 406–408 in Dikta (2001), but taking the presence of covariates into account. To be specific, one goes back to the one-dimensional setting by adding and subtracting

$$\int_0^{Z_{(i-1)}}\frac{1-\widetilde{m}(v,\beta_0)}{1-H(v)}H_n(dv)$$

where (recall) $\widetilde{m}(v,\beta_0)=E[m(X,Z,\beta_0)|Z=v]$ and H_n stands for the empirical cdf of the Z's. In this manner, a new term appears (due to the presence of covariates):

$$\max_{1\le i\le n-1}\left|\frac{1}{n}\sum_{j=1}^{i}\frac{m(X_{[j]},Z_{(j)},\beta_0)-\widetilde{m}(Z_{(j)},\beta_0)}{1-H(Z_{(j)})}\right|.$$

To show that this new term is also $O_P(1)$, one may note that it is bounded by

$$\frac{1}{n}\sum_{j=1}^{n}\frac{\left|m(X_j,Z_j,\beta_0)-\widetilde{m}(Z_j,\beta_0)\right|}{1-H(Z_j)}.$$

Under the integrability condition in (C7) the SLLN applies to this bound, which completes the proof. ■

Appendix B

Lemma B.1. *If $H(T)<1$, then, as $n\to\infty$, in probability*

$$n^{1/2}\sup_{0\le t\le T}\left|\int_0^t\int_{\mathbb{R}^p}\frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z,n}(du,dv)\right.$$
$$\left.-\int_0^t\int_{\mathbb{R}^p}\frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z}(du,dv)\right|\to 0. \quad (10)$$

Proof. First note that we can write

$$\int_0^t\int_{\mathbb{R}^p}\frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2}H_{X,Z,n}(du,dv)=U_n(t)-\frac{1}{n}U_n(t)+\frac{1}{n^2}\sum_{i=1}^{n}\frac{(1-m(X_i,Z_i,\beta_0))}{1-H(Z_i)}1_{\{Z_i\le t\}}$$

where

$$U_n(t)=\frac{1}{n(n-1)}\sum_{1\le i\ne j\le n}\frac{(1-m(X_i,Z_i,\beta_0))\left(1_{\{Z_j\le Z_i\}}-H(Z_i)\right)}{(1-H(Z_i))^2}1_{\{Z_i\le t\}}.$$

We have, since $0\le m(x,z)\le 1,\ \forall(x,z)\in\mathbb{R}^{p+1}$:

$$\sup_{0\le t\le T}\frac{1}{n^2}\sum_{i=1}^{n}\frac{(1-m(X_i,Z_i,\beta_0))}{1-H(Z_i)}1_{\{Z_i\le t\}}\le\frac{1}{n^2}\frac{n}{1-H(T)}=O(n^{-1})$$

and

$$\sup_{0\le t\le T}\left|\frac{1}{n}U_n(t)\right|\le\frac{1}{n}\frac{1}{n(n-1)}n(n-1)\frac{1}{(1-H(T))^2}=O(n^{-1}).$$

Define

$$\tilde{U}_n(t) = \frac{1}{n(n-1)} \sum_{1\le i\ne j\le n} \frac{(1-\widetilde{m}(Z_i,\beta_0))\left(1_{\{Z_j\le Z_i\}} - H(Z_i)\right)}{(1-H(Z_i))^2} 1_{\{Z_i\le t\}}.$$

The process $\tilde{U}_n(t)$ was studied in Dikta (1998, Lemma 3.8). Now,

$$U_n(t) - \tilde{U}_n(t) = \frac{1}{n(n-1)} \sum_{1\le i\ne j\le n} \frac{(\widetilde{m}(Z_i,\beta_0) - m(X_i,Z_i,\beta_0))\left(1_{\{Z_j\le Z_i\}} - H(Z_i)\right)}{(1-H(Z_i))^2} 1_{\{Z_i\le t\}}.$$

Put

$$X_n(t) := \frac{1}{n(n-1)} \sum_{1\le i\ne j\le n} \frac{(\widetilde{m}(Z_i,\beta_0) - m(X_i,Z_i,\beta_0))\left(1_{\{Z_j\le Z_i\}} - H(Z_i)\right)}{(1-H(Z_i))^2} 1_{\{Z_i\le t\}}$$

and put $n^{1/2}X_n(t) = \kappa_n(t)$, where $\kappa_n(t)$ is studied in Lemma B.2. As a consequence of Lemma B.2, we get that $\sup_{0\le t\le T}|\kappa_n(t)| \to 0$ in probability. Thus, the LHS of (10) is always less than or equal to

$$\begin{aligned}
&n^{1/2} \sup_{0\le t\le T}\left| U_n(t) - \tilde{U}_n(t) + \tilde{U}_n(t) - \int_0^t \int_{\mathbb{R}^p} \frac{(1-m(u,v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2} H_{X,Z}(du,dv) + O(n^{-1})\right| \\
&\le n^{1/2} \sup_{0\le t\le T}\left|X_n(t)\right| + n^{1/2} \sup_{0\le t\le T}\left|\tilde{U}_n(t) - \int_0^t \frac{(1-\widetilde{m}(v,\beta_0))(H_n(v)-H(v))}{(1-H(v))^2} H(dv)\right| + O(n^{-1/2}) \\
&\to 0
\end{aligned}$$

in probability. ■

Lemma B.2. (a) *Under the first integrability condition in* (C6.1) *we have for* $0 \le s \le t \le T$,

$$Cov(\kappa_n(s),\kappa_n(t)) \to 0 \quad \text{as } n \to \infty.$$

(b) *The finite dimensional distributions of* $\kappa_n(t)$ *converge to the degenerate distribution of* $X \equiv 0$.
(c) *For* $0 \le s \le t \le T$, *we have for some constant* C

$$E\left[(\kappa_n(s)-\kappa_n(r))^2(\kappa_n(t)-\kappa_n(s))^2\right] \le C.[H(t)-H(r)]^2.$$

Proof. For notational simplicity, we prove the result skipping the denominator $(1-H(Z_i))^2$ in the definition of $\kappa_n(t)$. Since this term is multiplied by $1_{\{Z_i\le t\}}$, and since $t \le T$ and $H(T) < 1$, the given arguments will be enough to conclude. We replace $m(x,z,\beta_0)$ and $\widetilde{m}(z,\beta_0)$ by $m(x,z)$ and $\widetilde{m}(z)$ for notational simplicity too.

Proof to (a). We have

$$\begin{aligned}
&Cov(\kappa_n(s),\kappa_n(t)) \\
&= \frac{1}{n(n-1)^2} \sum_{i\ne j,k\ne l} E\Big((\widetilde{m}(Z_i) - m(X_i,Z_i))(1_{\{Z_j\le Z_i\}} - H(Z_i))1_{\{Z_i\le s\}}(\widetilde{m}(Z_k) - m(X_k,Z_k))(1_{\{Z_l\le Z_k\}} - H(Z_k))1_{\{Z_k\le t\}}\Big).
\end{aligned}$$

There are two cases to consider. For the first one, assume that one index is distinct from all the others. W.l.o.g., by symmetry, we may assume that this index is i or j. If i is distinct from all others indices (say $i = 1$), we have

$$E\Big[(\widetilde{m}(Z_1) - m(X_1,Z_1))(1_{\{Z_j\le Z_1\}} - H(Z_1))1_{\{Z_1\le s\}}(\widetilde{m}(Z_k) - m(X_k,Z_k))(1_{\{Z_l\le Z_k\}} - H(Z_k))1_{\{Z_k\le t\}}\Big].$$

Condition w.r.t. $(Z_1, Z_l, Z_j, X_k, Z_j)$ and by independence:

$$E\left[m(X_1,Z_1)|Z_1,Z_l,Z_j,X_k,Z_k\right] = E\left[m(X_1,Z_1)|Z_1\right] = \widetilde{m}(Z_1).$$

Thus, the expected value is 0. If j is distinct from all other indices (say $j = 1$), we have

$$E\Big[(\widetilde{m}(Z_i) - m(X_i,Z_i))(1_{\{Z_1\le Z_i\}} - H(Z_i))1_{\{Z_i\le s\}}(\widetilde{m}(Z_k) - m(X_k,Z_k))(1_{\{Z_l\le Z_k\}} - H(Z_k))1_{\{Z_k\le t\}}\Big].$$

Condition w.r.t. $(X_i, Z_i, X_k, Z_k, Z_l)$ to get

$$E\left[(1_{\{Z_1\le Z_i\}} - H(Z_i))|X_i,Z_i,X_k,Z_k,Z_l\right] = 0.$$

If two indices are equal, then either ($i = k$ and $j = l$) or ($i = l$ and $j = k$). In the first case, say $i = k = 1$ and $j = l = 2$, we have

$$E\Big[(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_2 \le Z_1\}} - H(Z_1))1_{\{Z_1 \le s\}}(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_2 \le Z_1\}} - H(Z_1))1_{\{Z_1 \le t\}}\Big]$$
$$= E\left[(\widetilde{m}(Z_1) - m(X_1, Z_1))^2(1_{\{Z_2 \le Z_1\}} - H(Z_1))^2 1_{\{Z_1 \le \min(s,t)\}}\right].$$

Condition with respect to (Z_1, Z_2) to get that this equals

$$E\left[\mathrm{Var}(m(X_1, Z_1)|Z_1)\underbrace{(1_{\{Z_2 \le Z_1\}} - H(Z_1))^2}_{\le 1}\underbrace{1_{\{Z_1 \le \min(s,t)\}}}_{\le 1}\right].$$

If $i = l = 1$ and $j = k = 2$, then

$$E\Big[(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_2 \le Z_1\}} - H(Z_1))1_{\{Z_1 \le s\}}(\widetilde{m}(Z_2) - m(X_2, Z_2))(1_{\{Z_1 \le Z_2\}} - H(Z_2))1_{\{Z_2 \le t\}}\Big].$$

Condition w.r.t. (Z_1, Z_2, X_2) to see that this expected value is 0.

This covers all possible cases, and we see that

$$Cov(\kappa_n(s), \kappa_n(t)) \le \frac{n(n-1)}{n(n-1)^2}E\left[\mathrm{Var}(m(X, Z)|Z)\right] \to 0$$

as $n \to \infty$.

Proof to (b). (a) implies that, for $0 \le t \le T$,

$$\mathrm{Var}(\kappa_n(t)) \to 0 \quad \text{as } n \to \infty. \tag{11}$$

Fix $0 \le t \le T$. According to (11) and Chebyshev's inequality, we get for $\varepsilon > 0$:

$$P(\kappa_n(t) > \varepsilon) \le \frac{\mathrm{Var}(\kappa_n(t))}{\varepsilon^2} \to 0$$

as $n \to \infty$ and we conclude.

Proof to (c). Note that $\kappa_n(t) = n^{-1/2}\sum_{i=1}^n \eta_i 1_{\{Z_i \le t\}}$, where

$$\eta_i = \frac{1}{n-1}\sum_{1 \le j \ne i \le n}(\widetilde{m}(Z_i) - m(X_i, Z_i))(1_{\{Z_j \le Z_i\}} - H(Z_i)).$$

Thus,

$$E\left[(\kappa_n(s) - \kappa_n(r))^2(\kappa_n(t) - \kappa_n(s))^2\right] = \frac{1}{n^2}\sum_{1 \le i,j,k,l \le n}E\left[\eta_i\eta_j\eta_k\eta_l 1_{\{r \le Z_i \le s\}}1_{\{r \le Z_j \le s\}}1_{\{s \le Z_k \le t\}}1_{\{s \le Z_l \le t\}}\right].$$

Define $I_a^b(i) := 1_{\{a \le Z_i \le b\}}$ and $L_i(j) := (\widetilde{m}(Z_i) - m(X_i, Z_i))(1_{\{Z_j \le Z_i\}} - H(Z_i))$, $\alpha_i := \eta_i I_r^s(i)$ and $\beta_i := \eta_i I_s^t(i)$. Thus,

$$E\left[(\kappa_n(s) - \kappa_n(r))^2(\kappa_n(t) - \kappa_n(s))^2\right] = \frac{1}{n^2(n-1)^4}\sum_{1 \le i,j,k,l \le n}E\left[\alpha_i\alpha_j\alpha_k\alpha_l\right].$$

We analyze all possible cases separately.

1. $E\left[\alpha_1^2\beta_1^2\right] = 0$, since Z_1 cannot be in two disjoint intervals. The same applies to $E[\alpha_1^2\beta_1\beta_2]$, $E[\alpha_1\alpha_2\beta_1^2]$, $E[\alpha_1\alpha_2\beta_1\beta_2]$, and $E[\alpha_1\alpha_2\beta_1\beta_3]$.
2. $E\left[\alpha_1\alpha_2\beta_3\beta_4\right]$

$$E\left[\alpha_1\alpha_2\beta_3\beta_4\right] = E\left[\left(\sum_{j_1 \ne 1}L_1(j_1)\right)\left(\sum_{j_2 \ne 2}L_2(j_2)\right)\left(\sum_{j_3 \ne 3}L_3(j_3)\right)\left(\sum_{j_4 \ne 4}L_4(j_4)\right)I_r^s(1)I_r^s(2)I_s^t(3)I_s^t(4)\right]$$
$$= \sum_{\substack{j_1 \ne 1, j_2 \ne 2 \\ j_3 \ne 3, j_4 \ne 4}}E\left[L_1(j_1)L_2(j_2)L_3(j_3)L_4(j_4)I\right]$$

where $I := I_r^s(1)I_r^s(2)I_s^t(3)I_s^t(4)$, and so

$$E\left[L_1(j_1)L_2(j_2)L_3(j_3)L_4(j_4)I\right] = E\Big[(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_{j_1} \le Z_1\}} - H(Z_1))(\widetilde{m}(Z_2) - m(X_2, Z_2))(1_{\{Z_{j_2} \le Z_2\}} - H(Z_2))$$
$$\times(\widetilde{m}(Z_3) - m(X_3, Z_3))(1_{\{Z_{j_3} \le Z_3\}} - H(Z_3))(\widetilde{m}(Z_4) - m(X_4, Z_4))(1_{\{Z_{j_4} \le Z_4\}} - H(Z_4))I\Big].$$

Regardless of the values for j_1, j_2, j_3 and j_4, we condition w.r.t. $(Z_1, X_2, Z_2, \ldots, X_n, Z_n)$. Then we get, for a proper function $T = T(X_1, X_2, Z_2, \ldots, X_n, Z_n)$:

$$E\left[L_1(j_1)L_2(j_2)L_3(j_3)L_4(j_4)I\right] = E(T(Z_1, X_2, Z_2, \ldots, X_n, Z_n)E(\widetilde{m}(Z_1) - m(X_1, Z_1)|Z_1)).$$

But since $E(m(X_1, Z_1)|Z_1) = \widetilde{m}(Z_1)$, the expected value is 0.

3. The expected value of $\alpha_1^2\beta_2^2$ equals:

$$E\left[\alpha_1^2\beta_2^2\right] = \sum_{\substack{j_1\neq 1, j_2\neq 1\\ j_3\neq 2, j_4\neq 2}} E\left[L_1(j_1)L_1(j_2)L_2(j_3)L_2(j_4)I_r^s(1)I_s^t(2)\right]$$

and

$$\begin{aligned} &E\left[L_1(j_1)L_1(j_2)L_2(j_3)L_2(j_4)I_r^s(1)I_s^t(2)\right] \\ &= E\Big[(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_{j_1}\le Z_1\}} - H(Z_1))(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_{j_2}\le Z_1\}} - H(Z_1)) \\ &\quad \times (\widetilde{m}(Z_2) - m(X_2, Z_2))(1_{\{Z_{j_3}\le Z_2\}} - H(Z_2))(\widetilde{m}(Z_2) - m(X_2, Z_2))(1_{\{Z_{j_4}\le Z_2\}} - H(Z_2))I_r^s(1)I_s^t(2)\Big] \\ &\le E\left[1 I_r^t(1)I_r^t(2)\right] = E\left[I_r^t(1)\right]E\left[I_r^t(2)\right] \\ &\le (H(t) - H(r))^2. \end{aligned}$$

4. For $E(\alpha_1^2\beta_2\beta_3)$, we have to consider

$$\begin{aligned} &E\left[L_1(j_1)L_1(j_2)L_2(j_3)L_3(j_4)I_r^s(1)I_s^t(2)I_s^t(3)\right] \\ &= E\Big[(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_{j_1}\le Z_1\}} - H(Z_1))(\widetilde{m}(Z_1) - m(X_1, Z_1))(1_{\{Z_{j_2}\le Z_1\}} - H(Z_1)) \\ &\quad \times (\widetilde{m}(Z_2) - m(X_2, Z_2))(1_{\{Z_{j_3}\le Z_2\}} - H(Z_2))(\widetilde{m}(Z_3) - m(X_3, Z_3))(1_{\{Z_{j_4}\le Z_3\}} - H(Z_3))I_r^s(1)I_s^t(2)\Big]. \end{aligned}$$

Condition w.r.t. $X_1, Z_1, X_2, Z_2, Z_3, X_4, Z_3, \ldots, X_n, Z_n$ shows that this expected value is 0.

This covers all cases. We get

$$\begin{aligned} E\left[(\kappa_n(s) - \kappa_n(r))^2(\kappa_n(t) - \kappa_n(s))^2\right] &\le \frac{n(n-1)(n-1)^4}{n^2(n-1)^4}(H(t) - H(r))^2 \\ &\le (H(t) - H(r))^2. \quad \blacksquare \end{aligned}$$

Lemma B.3. (a) *Under* (C3)–(C5) *(resp.* (C4), (C5) *and* (C6.1)*) we have for* $0 \le s \le t \le T$,

$$Cov(\zeta_n(s), \zeta_n(t)) \to 0 \quad \textit{as } n \to \infty.$$

(b) *The finite dimensional distributions of* $\zeta_n(t)$ *converge to the degenerate distribution of* $X \equiv 0$.

(c) *Under* (C4), (C5) *and* (C6.1) *the process* $\zeta_n(t)$ *is tight.*

Proof. Proof to (a). Let K_i be as in Lemma A.1.1. Put $m(x, z) = m(x, z, \beta_0)$. Write

$$\begin{aligned} Cov(\zeta_n(s), \zeta_n(t)) &= E(\zeta_n(s)\zeta_n(t)) \\ &= \frac{1}{n(n-1)^2}\sum_{1\le i\neq j\le n}\sum_{1\le k\neq l\le n} E\Bigg[K_i\left(\frac{\alpha(X_j, Z_j, X_i, Z_i)}{1 - H(Z_j)}1_{\{Z_j\le s\}} - \int_0^s\int_{\mathbb{R}^p}\frac{\alpha(u, v, X_i, Z_i)}{1 - H(v)}H_{X,Z}(du, dv)\right) \\ &\quad \times K_k\left(\frac{\alpha(X_l, Z_l, X_k, Z_k)}{1 - H(Z_l)}1_{\{Z_l\le t\}} - \int_0^t\int_{\mathbb{R}^p}\frac{\alpha(u, v, X_k, Z_k)}{1 - H(v)}H_{X,Z}(du, dv)\right)\Bigg]. \end{aligned}$$

There are two cases to consider. For the first one, assume that one index is distinct from all the others. W.l.o.g., by symmetry, we may assume that this index is i or j. If i is distinct from all other indices (say $i = 1$), we have

$$\begin{aligned} &E\Bigg[K_1\left(\frac{\alpha(X_j, Z_j, X_1, Z_1)}{1 - H(Z_j)}1_{\{Z_j\le s\}} - \int_0^s\int_{\mathbb{R}^p}\frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)}H_{X,Z}(du, dv)\right) \\ &\quad \times K_k\left(\frac{\alpha(X_l, Z_l, X_k, Z_k)}{1 - H(Z_l)}1_{\{Z_l\le t\}} - \int_0^t\int_{\mathbb{R}^p}\frac{\alpha(u, v, X_k, Z_k)}{1 - H(v)}H_{X,Z}(du, dv)\right)\Bigg]. \end{aligned}$$

Condition w.r.t. $(X_1, Z_1, X_j, Z_j, X_k, Z_k, \delta_k, X_l, Z_l)$, and since $E\left[K_1|X_1, Z_1\right] = 0$, the expected value is 0.

If j is distinct from all other indices (say $j = 1$), then we have

$$E\left[K_i\left(\frac{\alpha(X_1, Z_1, X_i, Z_i)}{1 - H(Z_1)} 1_{\{Z_1 \le s\}} - \int_0^s \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_i, Z_i)}{1 - H(v)} H_{X,Z}(du, dv)\right)\right.$$
$$\left.\times K_k\left(\frac{\alpha(X_l, Z_l, X_k, Z_k)}{1 - H(Z_l)} 1_{\{Z_l \le t\}} - \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_k, Z_k)}{1 - H(v)} H_{X,Z}(du, dv)\right)\right].$$

Condition w.r.t. $(X_i, Z_i, \delta_i, X_k, Z_k, \delta_k)$, and since

$$E\left(\frac{\alpha(X_1, Z_1, X_i, Z_i)}{1 - H(Z_1)} 1_{\{Z_1 \le s\}} \middle| X_i, Z_i\right) - \int_0^s \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv) = 0$$

the expected value is 0.

If two indices are equal, then either ($i = k$ and $j = l$) or ($i = l$ and $j = k$). In the first case, say $i = k = 1$ and $j = l = 2$, we have

$$E\left[K_1^2\left(\frac{\alpha(X_2, Z_2, X_1, Z_1)}{1 - H(Z_2)} 1_{\{Z_2 \le s\}} - \int_0^s \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv)\right)\right.$$
$$\left.\times \left(\frac{\alpha(X_2, Z_2, X_1, Z_1)}{1 - H(Z_2)} 1_{\{Z_2 \le t\}} - \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv)\right)\right].$$

First, condition w.r.t. (X_1, Z_1, Z_2, X_2) and note that $E\left[K_1^2 | X_1, Z_1\right] = \frac{1}{m(X_1,Z_1)(1-m(X_1,Z_1))}$. Upon multiplying and taking absolute value, we get

$$\begin{aligned}
&\le E\left[\frac{\alpha^2(X_2, Z_2, X_1, Z_1)}{m(X_1, Z_1)(1 - m(X_1, Z_1))} 1_{\{Z_2 \le t\}}\right] \\
&\quad + E\left[\frac{\alpha(X_2, Z_2, X_1, Z_1)}{(1 - H(Z_2))m(X_1, Z_1)(1 - m(X_1, Z_1))} 1_{\{Z_2 \le s\}} \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv)\right] \\
&\quad + E\left[\frac{\alpha(X_2, Z_2, X_1, Z_1)}{(1 - H(Z_2))m(X_1, Z_1)(1 - m(X_1, Z_1))} 1_{\{Z_2 \le t\}} \int_0^s \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv)\right] \\
&\quad + E\left[\frac{1}{m(X_1, Z_1)(1 - m(X_1, Z_1))} \int_0^s \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv) \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv)\right].
\end{aligned} \tag{12}$$

It is easily seen that, according to our assumptions, all these expected values are finite.

If $i = l = 1$ and $j = k = 2$, then

$$E\left[K_1\left(\frac{\alpha(X_2, Z_2, X_1, Z_1)}{1 - H(Z_2)} 1_{\{Z_2 \le s\}} - \int_0^s \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_1, Z_1)}{1 - H(v)} H_{X,Z}(du, dv)\right)\right.$$
$$\left.\times K_2\left(\frac{\alpha(X_1, Z_1, X_2, Z_2)}{1 - H(Z_1)} 1_{\{Z_1 \le t\}} - \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_2, Z_2)}{1 - H(v)} H_{X,Z}(du, dv)\right)\right].$$

Condition w.r.t. (X_1, Z_1, X_2, Z_2) to see that this expected value is 0. Thus,

$$Cov(\zeta_n(s), \zeta_n(t)) \le \frac{1}{n(n-1)^2} n(n-1)C \to 0 \quad \text{as } n \to \infty$$

where C is an upper bound for the expected value (12).

Proof to (b). The proof follows from (a) similarly as in Lemma B.2(b).

Proof to (c). The lengthy proof, which we omit, starts by decomposing $\zeta_n(t)$ as a sum of two processes, namely $\zeta_n(t) = \zeta_n^{(1)}(t) + \zeta_n^{(2)}(t)$ where

$$\zeta_n^{(1)}(t) = n^{-1/2} \sum_{j=1}^n \left(\frac{1}{n-1} \sum_{i \ne j} K_i \frac{\alpha(X_j, Z_j, X_i, Z_i)}{1 - H(Z_j)} 1_{\{Z_j \le t\}}\right)$$

$$\zeta_n^{(2)}(t) = n^{-1/2} \sum_{j=1}^n \left(\frac{1}{n-1} \sum_{i \ne j} K_i \int_0^t \int_{\mathbb{R}^p} \frac{\alpha(u, v, X_i, Z_i)}{1 - H(v)} H_{X,Z}(du, dv)\right).$$

For the tightness of $\zeta_n^{(1)}(t)$, conditions (C4), (C5) and (C6.1) are needed, while (C6.1) suffices for the tightness of $\zeta_n^{(2)}(t)$. More explicitly, under the respective conditions it can be seen that, for $0 \le r \le s \le t \le T$,

$$E\left[(\zeta_n^{(i)}(s) - \zeta_n^{(i)}(r))^2 (\zeta_n^{(i)}(t) - \zeta_n^{(i)}(s))^2\right] \le C_i.(H(t) - H(r))^{3/2},$$

$i = 1, 2$, for some constants C_1 and C_2, which is enough to conclude. ■

References

Amorim, A.P., de Uña-Álvarez, J., Meira-Machado, L., 2011. Presmoothing the transition probabilities in the illness-death model. Statist. Probab. Lett. 81, 797–806.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Cox, D.R., Snell, E.J., 1989. Analysis of Binary Data. Chapman and Hall, London.

de Uña-Álvarez, J., Amorim, A.P., 2011. A semiparametric estimator of the bivariate distribution function for censored gap times. Biom. J. 53, 113–127.

de Uña-Álvarez, J., Rodríguez-Campos, C., 2004. Strong consistency of presmoothed Kaplan–Meier integrals when covariables are present. Statistics 38, 483–496.

Dikta, G., 1998. On semiparametric random censorship models. J. Statist. Plann. Inference 66, 253–279.

Dikta, G., 2001. Weak representation of the cumulative hazard function under semiparametric random censorship models. Statistics 35, 395–409.

Dikta, G., 2014. Asymptotically efficient estimation under semi-parametric random censorship models. J. Multivariate Anal. 124, 10–24.

Dikta, G., Ghorai, J., Schmidt, C., 2005. The central limit theorem under semiparametric random censorship models. J. Statist. Plann. Inference 127, 23–51.

Dikta, G., Kvesic, M., Schmidt, C., 2006. Bootstrap approximations in model checks for binary data. J. Amer. Statist. Assoc. 101, 521–530.

Iglesias-Pérez, M.C., de Uña-Álvarez, J., 2010. Nonparametric estimation of the conditional distribution function in a semiparametric censorship model. J. Statist. Plann. Inference 138, 3044–3058.

Moreira, A., de Uña-Álvarez, J., Machado, L., 2013. Presmoothing the Aalen-Johansen estimator in the illness-death model. Electron. J. Stat. 7, 1491–1516.

Robins, J.M., Finkelstein, D., 2000. Correcting for non-compliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 56, 779–788.

Robins, J.M., Rotnitzky, A., 1992. Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell, N., Dietz, K., Farewell, V. (Eds.), AIDS Epidemiology—Methodological Issues. Birkhäuser, Boston, MA, pp. 297–331.

Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. John Wiley and Sons, New York.

Shorack, G.R., Wellner, J.A., 1986. Empirical Processes with Applications to Statistics. Wiley, New York.

Stute, W., 1993. Consistent estimation under random censorship when covariables are present. J. Multivariate Anal. 45, 89–103.

Stute, W., 1995. The central limit theorem under random censorship. Ann. Statist. 23, 422–439.

Stute, W., 1996. Distributional convergence under random censorship when covariables are present. Scand. J. Statist. 23, 461–471.

Stute, W., 1999. Nonlinear censored regression. Statist. Sinica 9, 1089–1102.