Technical Proofs for "Mixture Cure Survival Models with Dependent Censoring"

Yi Li, Ram Tiwari and Subharup Guha

Suppose T is the survival time, e.g. time from the diagnosis of prostate cancer, and U is the potential random censoring time, e.g. study duration or death from other causes (e.g. cardiac failure), with only $X = min(T, U)$ and censoring indicator $\delta = I(X = T)$ observed in practice. Denote by $F_T(t) = P(T \leq t), F_U(t) = P(U \leq t)$ the cumulative distribution functions, and $S_T(t) = P(T > t), S_U(t) = P(U > t)$, the survival functions, for T and U, respectively. The scientific research often centers on discerning $F_T(t)$ while treating $F_U(t)$ as nuisance.

The mixture cure model assumes F_T to be an improper distribution over the entire real line and specifies

$$F_T(t) = pF_0(t) \tag{1}$$

or, equivalently,

$$S_T(t) = 1 - p + pS_0(t), \tag{2}$$

where $0 < p < 1$, $S_0(t) = 1 - F_0(t)$, and $F_0(t)$ is a proper distribution such that $\lim_{t\to\infty} F_0(t) = 1$. Models (1) and (2) consider the study population as an unobservable mixture of patients deemed susceptible (non-cured) and non-susceptible (cured) . Note that $(1-p)$ corresponds to the fraction of cured, that is, the point mass that T puts on ∞ and $F_0(t)$ is the distribution for the non-cured patients, often termed as the *latency* distribution.

We complete the model by specifying the dependence of the failure time T and censoring time U via an *strict* Archimedean copula model,

$$C(t, u) \overset{def}{=} P(T > t, U > u) = \phi^{-1}[\phi\{S_T(t)\} + \phi\{S_U(u)\}], \tag{3}$$

where $\phi : [0, 1] \to [0, \infty]$ is a nonincreasing function such that $\phi(1) = 0$ and $\phi(0) = \infty$. Examples of Φ include $\phi(t) = -\log t$, corresponding to independent censoring, the family of Clayton's models with $\phi(t) = (t^{-a} - 1)/a$ (for $a > 0$), and the Frank family with $\phi(t) = -\log((1 - \exp(-at))/(1 - \exp(-a))$ (for $a > 0$). We adopt the Archimedean copula formulation to emphasize the functional independence of the parameterizations of the marginal distribution functions, governed by S_T and S_U, and the dependence structure, governed by a class of copula functions ϕ. This formulation facilitates a derivation of the estimator for S_T, our main interest.

Suppose that we observe n i.i.d data, $(X_i, \delta_i), i = 1, \ldots, n$ and consider the counting processes $N_i(t) = I(X_i \le t, \delta_i = 1)$ and the at-risk processes $Y_i(t) = I(X_i \ge t)$. Denote by $N(t) = \sum N_i(t)$ and $Y(t) = \sum Y_i(t)$. Introduce the filtration $\mathcal{F}_t^n = \sigma\{N_i(s), Y_i(s+), 0 \le s \le t, i = 1, \ldots, n\}$, which contains the survival information up to time t for all n subjects and to which all the ensuing martingales and stopping times adapt. We denote the survival function for the observed times X_i by $\pi(t) = P(X_i > t) = C(t, t)$.

The following heuristically discusses an estimator based on (3), whose large sample properties will be considered in the next section. Denote by $\hat{S}_T$, which will be defined in (5), and $\hat{S}_U$ the estimates for S_T and S_U respectively, which are right continuous and piecewise constant functions with jumps only occurring at the observed failures and censorings, respectively. Denote by $\hat{\pi}(t)$ the empirical estimate of $\pi(t)$, which is $\hat{\pi}(t) = \sum_i I(X_i > t)/n = Y(t+)/n$.

By (3), at each observed time points X_i, i=1, ..., n,

$$\phi\{\hat{S}_T(X_i)\} + \phi\{\hat{S}_U(X_i)\} = \phi\{\hat{\pi}(X_i)\}.$$

Assume that $P(T = U) = 0$ (i.e. the censoring and failure cannot occur at the same time almost surely). Then at each observed failure time point X_i (such that $\delta_i = 1$), we have $\hat{S}_U(X_i^-) = \hat{S}_U(X_i)$ and

$$\begin{aligned} \phi(\hat{S}_T(X_i)) - \phi(\hat{S}_T(X_i^-)) &= \phi(\hat{\pi}(X_i)) - \phi(\hat{\pi}(X_i^-)) \\ &= \phi\left(\frac{Y(X_i)}{n} - \frac{1}{n}\right) - \phi\left(\frac{Y(X_i)}{n}\right). \end{aligned} \tag{4}$$

Applying (4) recursively, the estimator $\hat{S}_T$ can be written using the form of counting processes

$$\hat{S}_T(t) = \phi^{-1}\left[\int_0^t I(Y(s) > 0)\left\{\phi\left(\frac{Y(s)}{n} - \frac{1}{n}\right) - \phi\left(\frac{Y(s)}{n}\right)\right\} dN(s)\right], \tag{5}$$

which corresponds to the estimator derived by Rivest and Wells (2001) in the absence of cure fraction. When computing (5), we invoke the convention of $0/0 = 0$ if necessary.

To facilitate the theoretical development, introduce the crude hazard function defined by

$$d\tilde{\Lambda}(t) = \tilde{\lambda}(t)dt = P(t < T \le t + dt | T > t, U > t), \tag{6}$$

along with the martingale processes

$$M_i(t) = N_i(t) - \int_0^t Y_i(s) d\tilde{\Lambda}(t).$$

We further impose the following regularity conditions on $S_T(t)$ (or $F_T(t)$), $\pi(t)$ and the copula function ϕ.

(c.1) ϕ is strictly decreasing on $(0,1]$ and is sufficiently smooth in the following sense: the first two derivatives of $\phi(s)$ and $\psi(s) \overset{def}{=} -s\phi'(s)$ are bounded for $s \in [\epsilon, 1]$ where $\epsilon > 0$ is arbitrary. In addition, the first derivative of $\phi(s)$ is bounded away from 0 on $[0,1]$.

(c.2) $0 < \int_0^{\tau_X} \{\psi(\pi(s))\}^k d\tilde{\Lambda}(s) < \infty$ for $k = 0, 1, 2$

(c.3) $\int_0^{\tau_X} |(\psi'(\pi(s))| d\tilde{\Lambda}(s) < \infty$

(c.4) $\limsup_{t\to\tau_X} \int_t^{\tau_X} \frac{(\psi(\pi(s))^2}{\pi(s)} d\tilde{\Lambda}(s) = 0$

(c.5) $S_T(t)$ and $S_0(t)$ are continuous over $[0, \tau_X]$ if $\tau_X < \infty$. Otherwise, define $S_T(\infty) = \lim_{t\to\infty} S_T(t)$.

(c.6) $\lim_{t\to\tau_{F_0}} \frac{1-F_0(t)}{\pi(t)} < 1$.

Lemma 1 *$\phi(\hat{S}_T(t))$ converges to $\phi(S_T(t))$ uniformly on $[0, \tau_X]$. Moreover, $\hat{S}_T(t)$ converges to $S_T(t)$ uniformly on $[0, \tau_X]$ and the Nelson-Aalen estimator $\int_0^t I(Y(s) > 0) \frac{dN(s)}{Y(s)}$ converges to $\tilde{\Lambda}(t)$ in probability uniformly on $[0, \tau_X]$.*

Proof: First show that for any fixed t_0 such that $\pi(t_0) > 0$, $\sup_{t\in[0,t_0]} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \overset{pr.}{\to} 0$. Using a Taylor expansion and the regularity condition (c.1) on ϕ gives,

$$\phi(\hat{S}_T(t)) = -\int_0^t \frac{1}{n}\phi'\left\{\frac{Y(s)}{n}\right\} dN(s) + e_n$$

for $t \in [0, t_0]$, where $e_n = o_p(1)$ uniformly over $[0, t_0]$. As it can be shown that $\phi(S_T(t)) = -\int_0^t \phi'(\pi(s))\pi(s) d\tilde{\Lambda}(s)$, hence,

$$\begin{aligned} & \phi(\hat{S}_T(t)) - \phi(S_T(t)) \\ = & -\frac{1}{n}\int_0^t I(Y(s) > 0)\phi'\left\{\frac{Y(s)}{n}\right\} dM(s) + \int_0^t I(Y(s) > 0)\left[\psi\left\{\frac{Y(s)}{n}\right\} - \psi(\pi(s))\right] d\tilde{\Lambda}(s) \\ & -\int_0^t I(Y(s) = 0)\phi'(\pi(s))\pi(s) d\tilde{\Lambda}(s) + e_n \\ = & Z_1(t) + Z_2(t) + Z_3(t) + e_n, \end{aligned}$$

where $M(s) = \sum_{i=1}^n M_i(s)$ is a martingale.

When $t \in [0, t_0]$,

$$\begin{aligned} 0 &< Z_3(t) \le I(Y(t)=0) \int_0^t \psi(\pi(s)) d\tilde{\Lambda}(s) \\ &< I(Y(t_0)=0) \int_0^{\tau_X} \psi(\pi(s))\pi(s) d\tilde{\Lambda}(s). \end{aligned}$$

By the strong law of large numbers $Y(t_0)/n \to \pi(t_0)(>0)$ almost surely. Hence $Y(t_0) \to \infty$ almost surely. From this, coupled with the regularity condition (c.2), we have the uniform convergence of $Z_3(t)$ over $[0, t_0]$. It remains to demonstrate the convergence of $Z_1(t)$ and $Z_2(t)$. Consider the variation process of $Z_1(t)$,

$$\begin{aligned} < Z_1, Z_1 > (t) &= \int_0^t I(Y(s)>0) \left[\phi' \left\{ \frac{Y(s)}{n} \right\}\right]^2 \frac{Y(s)}{n^2} d\tilde{\Lambda}(s) \\ &= \int_0^t \frac{I(Y(s)>0)}{Y(s)} \left[\psi \left\{ \frac{Y(s)}{n} \right\}\right]^2 d\tilde{\Lambda}(s). \end{aligned}$$

Then it follows that $Z_1^2(t) - < Z_1, Z_1 > (t)$ is a martingale. By Lenglart's inequality (see, e.g., Fleming and Harrington (1991))

$$\begin{aligned} & P(\sup_{t\in[0,t_0]} |Z_1(t)| > \epsilon) \\ < & \frac{\eta}{\epsilon^2} + P(\int_0^{t_0} \frac{I(Y(s)>0)}{Y(s)} (\psi(\frac{Y(s)}{n}))^2 d\tilde{\Lambda}(s) > \eta) \\ < & \frac{\eta}{\epsilon^2} + P\left\{ \frac{1}{Y(t_0)} \int_0^{t_0} (\psi(\frac{Y(s)}{n}))^2 d\tilde{\Lambda}(s) > \eta \right\}. \end{aligned}$$

Since the empirical process $\frac{Y(s)}{n} \to \pi(s)$ in probability uniformly on $[0, \infty)$ and because of the boundness regularity conditions on $\psi(\cdot)$ and $\psi'(\cdot)$ on $[\pi(t_0), 1]$, $\psi^2(\frac{Y(s)}{n})$ converges to $\psi^2(\pi(s))$ unformly on $[0, t_0]$. Hence $\int_0^{t_0} (\psi(\frac{Y(s)}{n}))^2 d\tilde{\Lambda}(s) \to \int_0^{t_0} (\psi(\pi(s))^2 d\tilde{\Lambda}(s) < \infty$ (by the regularity condition (c.2)). So $\frac{1}{Y(t_0)} \int_0^{t_0} (\psi(\frac{Y(s)}{n}))^2 d\tilde{\Lambda}(s) \stackrel{pr}{\to} 0$ as $Y(t_0) \stackrel{pr}{\to} \infty$. Hence, $P(\sup_{0\le t\le t_0} |Z_1(t)| > \epsilon) \to 0$ for any $\epsilon > 0$. Now consider $Z_2(t) = \int_0^t I(Y(s)>0)\psi'(\pi(s)) \left\{ \frac{Y(s)}{n} - \pi(s) \right\} d\tilde{\Lambda}(s) + e_n$ where $e_n = o_p(1/n)$ uniformly on $[0, t_0]$. Further note that $\sup |Z_2(t)| \le \sup |e_n| + \left\{ \int_0^{t_0} |\psi'(\pi(s))| d\tilde{\Lambda}(s) \right\} \sup_{0\le s\le t_0} |\frac{Y(s)}{n} - \pi(s))|$, which leads to, under the regularity condition (c.3), $\sup_{0\le t\le t_0} |Z_2(t)| \stackrel{pr.}{\to} 0$. Thus we have proved that

$$sup_{0\le t\le t_0} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \stackrel{pr.}{\to} 0$$

for any t_0 such that $\pi(t_0) > 0$.

Now we show that

$$sup_{0\le t\le \tau_X} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \stackrel{pr.}{\to} 0.$$

We only consider the situation when $\tau_X < \infty$ as the proof follows similarly when $\tau_X = \infty$. Fix a small $\epsilon > 0$ and consider any $t \in [\tau_X - \epsilon, \tau_X]$. With monotonicity of S_T and ϕ, it follows that

$$\begin{aligned}
& |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \\
< \; & |\phi(\hat{S}_T(\tau_X)) - \phi(\hat{S}_T(\tau_X - \epsilon))| + |\phi(\hat{S}_T(\tau_X - \epsilon)) - \phi(S_T(\tau_X - \epsilon))| \\
& + |\phi(S_T(\tau_X - \epsilon)) - \phi(S_T(\tau_X))|.
\end{aligned}$$

Also note that

$$\begin{aligned}
& \sup_{0 \le t \le \tau_X} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \\
\le \; & \sup_{0 \le t \le \tau_X - \epsilon} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| + \sup_{\tau_X - \epsilon \le t \le \tau_X} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \\
\le \; & \sup_{0 \le t \le \tau_X - \epsilon} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| + |\phi(\hat{S}_T(\tau_X)) - \phi(\hat{S}_T(\tau_X - \epsilon))| + |\phi(\hat{S}_T(\tau_X - \epsilon)) - \phi(S_T(\tau_X - \epsilon))| \\
& + |\phi(S_T(\tau_X - \epsilon)) - \phi(S_T(\tau_X))|.
\end{aligned}$$

Using the uniform convergence of $\hat{S}_T(t)$ on $[0, \tau_X - \epsilon]$ and letting $\epsilon \to 0^+$ yields $\sup_{0 \le t \le \tau_X} |\phi(\hat{S}_T(t)) - \phi(S_T(t))| \overset{pr.}{\to} 0$. As $\phi'(\cdot)$ is bounded away from 0 on $[0,1]$ [condition (c.1)], a Taylor expansion immediately yields $\sup_{0 \le t \le \tau_X} |\hat{S}_T(t) - S_T(t)| \overset{pr.}{\to} 0$.

Applying a similar argument, we may demonstrate the uniform convergence of $\int_0^t I(Y(s) > 0) \frac{dN(s)}{Y(s)}$ to $\tilde{\Lambda}(t)$ on $[0, \tau_X]$ by observing that

$$\int_0^t I(Y(s) > 0) \frac{dN(s)}{Y(s)} - \int_0^t d\tilde{\Lambda}(s) = \int_0^t I(Y(s) > 0) \frac{dM(s)}{Y(s)} - \int_0^t I(Y(s) = 0) d\tilde{\Lambda}(s).$$

□

We now consider the asymptotical normality of the proposed estimator for cure fractions. Define the stopped process

$$Z_n(t) = \sqrt{n}\{\phi(\hat{S}_T(t \wedge X^n)) - \phi(S_T(t \wedge X^n))\}. \tag{7}$$

and the covariance function

$$\begin{aligned}
\mathcal{C}(t_1, t_2) = \; & \int_0^{t_1 \wedge t_2} \pi(s)(\phi'(\pi(s)))^2 d\tilde{\Lambda}(s) + 2 \int_0^{t_1 \wedge t_2} \int_0^s \pi(s)(1 - \pi(u))\psi'(\pi(u))\psi'(\pi(s)) d\tilde{\Lambda}(u) d\tilde{\Lambda}(s) \\
& + 2 \int_0^{t_1 \wedge t_2} \int_0^s \phi'(\pi(u))\pi(s)\psi'(\pi(s)) d\tilde{\Lambda}(u) d\tilde{\Lambda}(s) \\
& + \int_{t_1 \wedge t_2}^{t_1 \vee t_2} \pi(s)\psi'(\pi(s)) d\tilde{\Lambda}(s) \int_0^{t_1 \wedge t_2} \left[\{1 - \pi(u)\}\psi'\{\pi(u)\} + \phi'\{\pi(u)\}\right] d\tilde{\Lambda}(u)
\end{aligned} \tag{8}$$

for $0 \le t_1, t_2, < \tau_X$, where $\psi(s) \overset{def}{=} -s\phi'(s)$.

Theorem 1 *$Z_n(t)$ converges weakly to $I[0,\tau_X)Z(t) + I\{\tau_X\}Z^\infty$ on $D[0,\tau_X]$, where $Z(t)$ is a tight Gaussian process with the covariance function $\mathcal{C}(t_1,t_2)$ and Z^∞ is a normal random variable with the variance v_0^∞ and $cov\{Z^\infty, Z(t)\} = \mathcal{C}^\infty(t)$.*

Proof: Using the same argument as in Rivest and Wells (2001), up to an $o_p(1)$ term, we have that

$$\begin{aligned} Z_n(t) &= \sqrt{n}\left(-\frac{1}{n}\int_0^{t\wedge X^n} I(Y(s)>0)\phi'\left\{\frac{Y(s)}{n}\right\}dM(s)\right. \\ &\quad \left. + \int_0^{t\wedge X^n} I(Y(s)>0)\left[\psi\left\{\frac{Y(s)}{n}\right\} - \psi(\pi(s)\right]d\tilde{\Lambda}(s)\right) \\ &= Z_{n,1}(t) + Z_{n,2}(t). \end{aligned} \tag{9}$$

Rivest and Wells (2001) showed, for any t_0 such that $\pi(t_0) > 0$, $Z_n(t)$ converges weakly to $Z(t)$ on $D[0,t_0]$. To show the weak convergence of $Z_n(t)$ on $D[0,\tau_X]$, it is sufficient to show the tightness of $Z_n(t)$ in a small (left) neighborhood of τ_X in view of Theorems 13.2 and 16.8 of Billingsley (1999). That is, it suffices to show for any $\epsilon > 0$

$$\lim_{t\to\tau_X}\limsup_n P(\sup_{s\in(t,\tau_X]}|Z_n(s) - Z_n(t)| > \epsilon) = 0; \tag{10}$$

see, also, Gill (1980).

Fix a t. Then $sup_{s\in(t,\tau_X]}|Z_n(s) - Z_n(t)| \le sup_{s\in(t,\tau_X]}|Z_{n,1}(s) - Z_{n,1}(t)| + sup_{s\in(t,\tau_X]}|Z_{n,2}(s) - Z_{n,2}(t)|$. Since X^n is a stopping time, and by the optional sampling theorem, $Z_{n,1}(s) - Z_{n,1}(t) = -\frac{1}{\sqrt{n}}\int_{t\wedge X^n}^{s\wedge X^n} I(Y(s)>0)\phi'\left\{\frac{Y(s)}{n}\right\}dM(s)$ is a local martingale and its predictable variation process is given by

$$< Z_{n,1}(s) - Z_{n,1}(t), Z_{n,1}(s) - Z_{n,1}(t) >= \int_{t\wedge X^n}^{s\wedge X^n} I(Y(s)>0)\left[\psi\left\{\frac{Y(s)}{n}\right\}\right]^2\frac{n}{Y(s)}d\tilde{\Lambda}(s),$$

hence, $(Z_{n,1}(s) - Z_{n,1}(t))^2 - < Z_{n,1}(s) - Z_{n,1}(t), Z_{n,1}(s) - Z_{n,1}(t) >$ is a martingale (again assume that t is fixed).

Therefore, by Lenglart's inequality we have

$$\begin{aligned} &P(\sup_{[t,\tau_X]}|Z_{n,1}(s) - Z_{n,1}(t)| > \epsilon) \\ &< \frac{\eta}{\epsilon^2} + P(\int_{t\wedge X^n}^{\tau_X\wedge X^n} I(Y(s)>0)\left[\psi\left\{\frac{Y(s)}{n}\right\}\right]^2\frac{n}{Y(s)}d\tilde{\Lambda}(s) > \eta) \\ &\le \frac{\eta}{\epsilon^2} + P(\int_t^{\tau_X}\left[\psi\left\{\frac{Y(s)}{n}\right\}\right]^2\frac{n}{Y(s)}d\tilde{\Lambda}(s) > \eta). \end{aligned} \tag{11}$$

Because of the uniform convergence of $\frac{Y(s)}{n}$ to $\pi(s)$ on $[0,\tau_X]$,

$$\int_t^{\tau_X}\left[\psi\left\{\frac{Y(s)}{n}\right\}\right]^2\frac{n}{Y(s)}d\tilde{\Lambda}(s)\overset{pr.}{\to}\int_t^{\tau_X}\frac{(\psi(\pi(s))^2}{\pi(s)}d\tilde{\Lambda}(s),$$

for any $t<\tau_X$. Hence by the regularity condition (c.4), the second term in (11) converges to 0 for any $\eta>0$ as $t\to\tau_X$. Hence, we have

$$\lim_{t\to\tau_X}\limsup_n P(\sup_{[t,\tau_X]}|Z_{n,1}(s)-Z_{n,1}(t)|>\epsilon)=0. \tag{12}$$

Now we turn to show that

$$\lim_{t\to\tau_X}\limsup_n P(\sup_{[t,\tau_X]}|Z_{n,2}(s)-Z_{n,2}(t)|>\epsilon)=0. \tag{13}$$

As $Z_{n,2}(s)-Z_{n,2}(t)=\int_{t\wedge X_n}^{s\wedge X_n}I(Y(s)>0)\psi'(\pi(s))\sqrt{n}\left\{\frac{Y(s)}{n}-\pi(s)\right\}d\tilde{\Lambda}(s)+o_p(1)$, it follows that (13) holds as $\sqrt{n}(\frac{Y(s)}{n}-\pi(s))$ converges weakly to a tight Gaussian process over $[0,\infty)$. Combining (12) and (13) gives (10). Hence the desired result follows.

We now compute the covariance function for the limiting process $Z(t)$. The derivation of this covariance function, which is not given in Rivest and Wells (2001), is involved as $Z(t)$ is not an independent increment process. For any $t<\tau_X$, as $X^n\to\tau_X$ almost surely and following Rivest and Wells (2001), we can show that (9) is asymptotically equal to (up to an $o_p(1)$ term) $W_n(t)=\frac{1}{\sqrt{n}}\int_0^t-\phi'(\pi(u))dM(u)+\int_0^t X_n(s)\psi'(\pi(s))d\tilde{\Lambda}(s)=W_{n,1}(t)+W_{n,2}(t)$, where $X_n(s)=\sqrt{n}\left\{\frac{Y_n(s)}{n}-\pi(s)\right\}$. Hence we only need to compute the limiting covariance function for $W_n(t)$.

Consider $0\le t_1\le t_2\le\tau_X$. Then

$$\begin{aligned}\text{cov}\{W_n(t_1),W_n(t_2)\} &= E\{W_{n,1}(t_1)W_{n,1}(t_2)\}+E\{W_{n,2}(t_1)W_{n,2}(t_2)\}\\ &\quad +E\{W_{n,1}(t_1)W_{n,2}(t_2)\}+E\{W_{n,1}(t_2)W_{n,2}(t_1)\}.\end{aligned}$$

Since $W_{n,1}(\cdot)$ is a square integrable martingale,

$$E\{W_{n,1}(t_1)W_{n,1}(t_2)\}=\frac{1}{n}E\left\{\int_0^{t_1}[\phi'\{\pi(s)\}]^2Y(s)d\tilde{\Lambda}(s)\right\}=\int_0^{t_1}[\phi'\{\pi(s)\}]^2\pi(s)d\tilde{\Lambda}(s).$$

Also

$$\begin{aligned}&E\{W_{n,2}(t_1)W_{n,2}(t_2)\}\\ =\ & E\int_0^{t_2}\int_0^{t_1}(Y_1(u)-\pi(u))(Y_1(s)-\pi(s))\psi'(\pi(u))d\tilde{\Lambda}(u)\psi'(\pi(s))d\tilde{\Lambda}(s)\end{aligned}$$

$$
\begin{aligned}
&= E\int_0^{t_1}\int_0^{t_1}(Y_1(u)-\pi(u))(Y_1(s)-\pi(s))\psi'(\pi(u))d\tilde{\Lambda}(u)\psi'(\pi(s))d\tilde{\Lambda}(s)\\
&\quad +E\int_{t_1}^{t_2}\int_0^{t_1}(Y_1(u)-\pi(u))(Y_1(s)-\pi(s))\psi'(\pi(u))d\tilde{\Lambda}(u)\psi'(\pi(s))d\tilde{\Lambda}(s)\\
&= 2\int_0^{t_1}\int_0^{s}\pi(s)(1-\pi(u))\psi'(\pi(u))d\tilde{\Lambda}(u)\psi'(\pi(s))d\tilde{\Lambda}(s)\\
&\quad +\int_{t_1}^{t_2}\pi(s)\psi'(\pi(s))d\tilde{\Lambda}(s)\int_0^{t_1}(1-\pi(u))\psi'(\pi(u))d\tilde{\Lambda}(u),
\end{aligned}
$$

where the calculation of $E\int_0^{t_1}\int_0^{t_1}$ comes from Rivest and Wells (2001) [after correcting a typographic error in their formula - $\pi(u)-\pi(s)\pi(u)$ on line 8 of p.151 of Rivest and Wells (2001) should read $\pi(s)-\pi(s)\pi(u)$].

Introduce $A(s)=-\phi'(\pi(s))$ and $B(s)ds=\psi'(\pi(s))d\tilde{\Lambda}(s)$. Applying the result $E(M_1(u)Y_1(s))=-\pi(s)\tilde{\Lambda}(u\wedge s)$ and integration by parts, we have

$$
\begin{aligned}
&cov\{W_{n,2}(t_1),W_{n,1}(t_2)\}\\
&= E\left\{\int_0^{t_1}A(u)dM_1(u)\int_0^{t_2}(Y_1(s)-\pi(s))B(s)ds\right\}\\
&= E\left\{\int_0^{t_2}A(t_1)M_1(t_1)Y_1(s)B(s)ds\right\} \qquad (14)\\
&\quad +E\left\{\int_0^{t_2}\int_0^{t_1}-M_1(u)Y_1(s)dA(u)B(s)ds\right\} \qquad (15)
\end{aligned}
$$

Using $\int_0^{t_2}=\int_0^{t_1}+\int_{t_1}^{t_2}$, (14) is

$$
-\tilde{\Lambda}(t_1)\int_0^{t_1}\pi(s)\tilde{\Lambda}(s)B(s)ds-\tilde{\Lambda}(t_1)A(t_1)\int_{t_1}^{t_2}\pi(s)B(s)ds \qquad (16)
$$

while using $\int_0^{t_2}\int_0^{t_1}=\int_0^{t_1}\int_0^{t_1}+\int_{t_1}^{t_2}\int_0^{t_1}$, (15) is

$$
\int_0^{t_1}\int_0^{t_1}\pi(s)\tilde{\Lambda}(u\wedge s)dA(u)B(s)ds+\int_{t_1}^{t_2}\int_0^{t_1}\pi(s)\tilde{\Lambda}(u)dA(u)B(s)ds. \qquad (17)
$$

Adding the first term of (16) and the first term of (17) gives $-\int_0^{t_1}\int_0^{s}A(u)d\tilde{\Lambda}(u)\pi(s)B(s)ds$ following Rivest and Wells (2001) [, though a minus sign is missing in the statement between (20) and (21) in their article). Integration by parts with respect to $dA(u)$ in the second term of (17) gives the summation of the second term in (16) and the second term in (17) is $-\int_{t_1}^{t_2}\pi(s)B(s)ds\int_0^{t_1}A(u)d\tilde{\Lambda}(u)$ So,

$$
cov\{W_{n,2}(t_1),W_{n,1}(t_2)\}=-\int_0^{t_1}\int_0^{s}A(u)d\tilde{\Lambda}(u)\pi(s)B(s)ds-\int_{t_1}^{t_2}\pi(s)B(s)ds\int_0^{t_1}A(u)d\tilde{\Lambda}(u). \qquad (18)
$$

Similarly we obtain

$$cov\{W_{n,1}(t_2), W_{n,2}(t_1)\} = -\int_0^{t_1}\int_0^s A(u)d\tilde{\Lambda}(u)\pi(s)B(s)ds. \tag{19}$$

Plugging back $A(u) = -\phi'(\pi(s))$ and $B(s)ds = \psi'(\pi(s))d\tilde{\Lambda}(s)$ in (18) and (19) and using the weak convergence of a tight process W_n to $Z(t)$, we have thus obtained the covariance function $\mathcal{C}(t_1, t_2)$ as stated in the theorem. □

Reference

Fleming, T. R. and Harrington, D. P. (1991) *Counting processes and survival analysis*, New York: John Wiley & Sons.

Gill, R. D. (1980) "Censoring and stochastic integrals," CWI, Math. Centrum.

Rivest, L.P. and Wells, M.T. (2001) "A martingale approach to the copula-graphic estimator for the survival function under dependent censoring," *J. of Multivariate Analysis*, 79, 138-155.