Technical Report to
"THE NPMLE OF THE JOINT DISTRIBUTION FUNCTION WITH RIGHT-CENSORED AND MASKED COMPETING RISKS DATA"

In this report, we shall give the proofs of some statements in the main paper. To be consistent with the paper, we order the sections in this report as §8 and §9. The proofs in §7 are given in §8. The other proofs are given in §9.

§8. Proofs of the lemmas in §7 related to Theorems 6.1 and 6.2.

Most proofs in §6 are given here, except for Lemma 7.1. Lemma 7.1 essentially says that the GMLE satisfies both the cdf form and the df form of the SC equation. The proof of the equivalence between the df form and the cdf form is quite similar to the proof under the interval censoring case in (Li *et al.* (1997)), and is skipped.

Proof of Lemma 7.2. Statement (1) of the lemma follows from Eq.s (7.4) and (7.5).

To prove Statement (2), it follows from (7.5) that

$$
\begin{aligned}
\mathcal{R}_F(F)(x,y) =& \int_{v<x} \frac{F(x,y)-F(v,y)}{1-F_T(v)}(1-F_T(v))dF_R(v) \\
&+ \sum_{w\in\mathcal{J}} \int_{v\le x} \frac{\sum_{i\le y, i\in w} f_{C|T}(i|v)}{\sum_{j\in w} f_{C|T}(j|v)} \sum_{c\in w} f_{C|T}(c|v) q(w|v) dF_T(v) \\
&\qquad\qquad\qquad\qquad\text{(as } \textstyle\int\int dF(t,c) = \int \sum_c f_{C|T}(c|t)dF_T(t)) \\
=& \int_{v<x} F(x,y)-F(v,y)dF_R(v) + \int_{v\le x} \sum_{w\in\mathcal{J}} \sum_{i\le y, i\in w} f_{C|T}(i|v)q(w|v)dF_T(v) \\
=& \int_{v<x} F(x,y)-F(v,y)dF_R(v) + \int_{v\le x} \sum_{i\le y} f_{C|T}(i|v) \sum_{w\in\mathcal{J}} \mathbf{1}_{(i\in w)} q(w|v)dF_T(v) \\
=& \int_{v<x}\int_{v<u\le x, c\le y} dF(u,c)dF_R(v) + \int_{v\le x} \sum_{i\le y} f_{C|T}(i|v)S_R(v-)dF_T(v) \\
&\Big(\text{as } \sum_{w\in\mathcal{J}} q(w|x)\mathbf{1}_{(y\in w)} = \sum_{w:\ y\in w}\ \sum_{h:\ w\in P_h} P(\Delta=h, R\ge x) \text{ (see A2)} \\
&\qquad\qquad = \sum_h \sum_{j:\ y\in P_{hj}} P(\Delta=h, R\ge x) = S_R(x-)\Big) \qquad (8.1) \\
=& \sum_{c\le y} f_C(c)\{\int_{v<x}\int_{v<u\le x} dF_{T|C}(u|c)dF_R(v) + \int_{v\le x}\int_{v\le r} dF_R(r)dF_{T|C}(v|c)\} \\
=& \sum_{c\le y} f_C(c)F_{T|C}(x|c) = F(x,y).
\end{aligned}
$$

It follows that $F=\mathcal{R}_F(F)=\mathcal{B}_F(Q)$ by (1). Thus statement (2) holds. Verify that in the proof of $F=\mathcal{R}_F(F)$, F is not critical, as long as $F\in\Theta_o$. Thus statement (3) holds too.

The proofs of statements (4), (5) and (6) are similar to those of (1), (2) and (3). ▫

Proof of Lemma 7.3. For each $\check F\in\Theta_o$, by Lemma 7.2, $H=\check F$ is a solution to $H=\mathcal{R}_H(\check F)$. If there exists a different solution of $H=\mathcal{R}_H(\check F)$ in Θ_o, say $H=G\ne\check F$ and $\mathcal{S}_G\supset\mathcal{S}_{\check F}$, then by Lemma 7.2, we have $G=\mathcal{R}_G(G)$ as well as $G=\mathcal{R}_G(\check F)$. That is, $H=G$ and $H=\check F$ are two different solutions to $G=\mathcal{R}_G(H)$.

On the other hand, given $H\in\Theta_o$, by Lemma 7.2, $\check F=H$ is a solution to $H=\mathcal{R}_H(\check F)$. If $\check F=G\ne H$ in Θ_o is a different solution to $H=\mathcal{R}_H(\check F)$ and $\mathcal{S}_H\supset\mathcal{S}_G$, then by Lemma 7.2, $G=\mathcal{R}_G(G)$. Thus $\check F=H$ and $\check F=G$ are two different solutions to $H=\mathcal{R}_H(\check F)$. ▫

Proof of Lemma 7.4. It is obvious that $\mathcal{B}_{\check{F}}$ is linear. Let $t_0=-\infty$, $t_{m+1}=\infty$, $\mathcal{S}_{\check{F}_T}\cap D_V=\{t_1,...,t_m\}$, where $t_1<\cdots<t_m$. It suffices to show that $\forall\ g\in\mathcal{D}_Q$, $H=\mathcal{B}_{\check{F}}(g)$ satisfies (7.2), (1) in (2.2) and $\mathcal{S}_H\subset\mathcal{S}_{\check{F}}$.

Since $\check{F}$ satisfies (2.2) and it follows from (7.4) that

$$H(x,y)=\mathcal{B}_{\check{F}}(g)(x,y)=\sum_{h=1}^{J}\mathbf{1}_{(h\le y)}\{\int_{v<x}\frac{\check{F}^s_h(x)-\check{F}^s_h(v)}{1-\check{F}_T(v)}dg(v,0,C_r)\qquad(8.2)$$

$$+\int_{v\le x,v\in\mathcal{S}_{\check{F}_T},w\in\mathcal{J}}\frac{\mathbf{1}_{(h\in w)}\check{f}_{C|T}(h|v)}{\sum_{i\in w}\check{f}_{C|T}(i|v)}dg(v,1,w)\}.$$

Since $\check{F}$ satisfies (7.2) and (1) in (2.2), in view of (8.2) H also satisfies (7.2) and (1) in (2.2). Since $||\mathcal{S}_{\check{F}}||\le(m+1)J$, in order to show $\mathcal{S}_H\subset\mathcal{S}_{\check{F}}$, it suffices to show

$$\mathcal{S}_{H_T}\subset\mathcal{S}_{\check{F}_T},\ \ \textit{i.e.}, \text{ if } t_j\le x<t_{j+1},\ j\le m, \text{ then } \mathcal{B}_{\check{F}}(g)(x,\infty)-\mathcal{B}_{\check{F}}(g)(t_j,\infty)=0\ \forall\ g\in\mathcal{D}_Q;$$
$$\text{and if } (t_j,c)\notin\mathcal{S}_{\check{F}},\ j\le m+1 \text{ and } c\in\mathcal{C}_r, \text{ then } \mu_H(\{(t_j,c)\})=0.\qquad(8.3)$$

Actually for each $y\in\mathcal{C}_r$, if $t_j\le x<t_{j+1}$, $j\le m$, then

$$\begin{aligned}
&\mathcal{B}_{\check{F}}(g)(x,y)-\mathcal{B}_{\check{F}}(g)(t_j,y)\\
=&\int_{v<x}\frac{\check{F}(x,y)-\check{F}(v,y)}{1-\check{F}_T(v)}dg(v,0,C_r)+\int_{v\le x,v\in\mathcal{S}_{\check{F}_T},w\in\mathcal{J}}\frac{\sum_{h\le y,h\in w}\check{f}_{C|T}(h|v)}{\sum_{i\in w}\check{f}_{C|T}(i|v)}dg(v,1,w)\\
-&\int_{v<t_j}\frac{\check{F}(t_j,y)-\check{F}(v,y)}{1-\check{F}_T(v)}dg(v,0,C_r)-\int_{v\le t_j,v\in\mathcal{S}_{\check{F}_T},w\in\mathcal{J}}\frac{\sum_{h\le y,h\in w}\check{f}_{C|T}(h|v)}{\sum_{i\in w}\check{f}_{C|T}(i|v)}dg(v,1,w)\\
=&\int_{v<t_j}\frac{\check{F}(x,y)-\check{F}(t_j,y)}{1-\check{F}_T(v)}dg(v,0,C_r)+\int_{t_j\le v<x}\frac{\check{F}(x,y)-\check{F}(v,y)}{1-\check{F}_T(v)}dg(v,0,C_r)\\
&+\int_{t_j<v\le x,v\in\mathcal{S}_{\check{F}_T},w\in\mathcal{J}}\frac{\sum_{h\le y,h\in w}\check{f}_{C|T}(h|v)}{\sum_{i\in w}\check{f}_{C|T}(i|v)}dg(v,1,w)\\
=&0\ \ (\text{as } \check{F}(t,y)=\check{F}(t_j,y) \text{ and } \check{f}_{C|T}(c|t)=0\ \forall\ t\in(t_j,t_{j+1})).
\end{aligned}$$

Moreover, if $\mu_{\check{F}}(\{(t_j,c)\})=0$, then noting $t_j\in\mathcal{S}_{\check{F}_T}$, by (8.2)

$$\begin{aligned}
\mu_H(\{(t_j,c)\})=&\int_{v<t_j}\frac{\check{F}^s_c(t_j)-\check{F}^s_c(t_j-)}{1-\check{F}_T(v)}dg(v,0,C_r)+\int_{v=t_j,w\in\mathcal{J}}\frac{\mathbf{1}_{(c\in w)}\check{f}_{C|T}(c|v)}{\sum_{i\in w}\check{f}_{C|T}(i|v)}dg(v,1,w)\}\\
=&0\ (\text{as } \check{F}^s_c(t_j)=\check{F}^s_c(t_j-) \text{ and } \check{f}_{C|T}(c|t_j)=0 \text{ due to } \mu_{\check{F}}(\{(t_j,c)\})=0).
\end{aligned}$$

Thus (8.3) holds. This completes the proof of the statement related to $\mathcal{B}_H$. Replacing dg in $\mathcal{B}_{\check{F}}$ in the foregoing arguments by $(H(\infty,J)-H(v,J))dF_R$ or $q(w|v)dH(v,c)$ in $\mathcal{R}_{\check{F}}$, we can establish a similar version of (8.3) for $\mathcal{R}_{\check{F}}$. Since $\mathcal{R}_{\check{F}}$ is also linear, it completes the proof of the lemma. ▫

Proof of Lemma 7.5. Since (1) $\mathcal{R}_{\hat{F}}(\hat{F})=\hat{F}=\mathcal{B}_{\hat{F}}(\hat{Q})$ (by Lemma 7.1), (2) $\mathcal{R}_{\hat{F}}(F)=\mathcal{B}_{\hat{F}}(Q)$ (by Lemma 7.2), $\mathcal{R}_{\hat{F}}(\hat{F})-\mathcal{R}_{\hat{F}}(F)=\mathcal{B}_{\hat{F}}(\hat{Q})-\mathcal{B}_{\hat{F}}(Q)$. Since both of the operators are linear by Lemma 7.4, $\mathcal{R}_{\hat{F}}(\hat{F}-F)=\mathcal{B}_{\hat{F}}(\hat{Q}-Q)$. ▫

Proof of Lemma 7.6. We shall give the proof in 3 steps.

Step 1 (preliminary). Recall $f = f_{T,C}$. The uniqueness of the solution to the equation $\check{F} = \mathcal{R}_{\check{F}}(F)$ is equivalent to the uniqueness of the solution to equations (7.3) (w.r.t. the measure $\mu(\cdot,\cdot)$ on $D_V \times \mathcal{C}_r$). Since we are interested in the equation $\check{F} = \mathcal{R}_{\check{F}}(H)$ rather than $\check{F} = \mathcal{R}_{\check{F}}(F)$, Eq. (7.3) becomes, for $x \in \mathcal{S}_{\check{F}_T}$ and $H \in \mathcal{D}$,

$$\check{f}(x,y) = \int_{v<x} \frac{\check{f}(x,y)(1-H_T(v))}{1-\check{F}_T(v)} dF_R(v) + \sum_{w\in\mathcal{J}} q(w|x) \frac{\mathbf{1}_{(y\in w)}\check{f}(x,y)}{\sum_{k\in w}\check{f}(x,k)} \sum_{h\in w} h(x,h); \tag{8.4}$$

$$\check{f}_T(x) = \check{f}_T(x) \int_{v<x} \frac{1-H_T(v)}{1-\check{F}_T(v)} dF_R(v) + \mathbf{1}_{(x\in\mathcal{S}_{\check{F}_T})} S_R(x-) h_T(x), \tag{8.5}$$

where (H_T, h_T) correspond to (F_T, f_T), respectively.

Step 2 (uniqueness of the solution to (8.5)). It is easy to show that $H = \check{F}$ is a solution to the aforementioned equations. In particular, $h_T = \check{f}_T$ is the solution of the SC equation (8.5). By the assumption $\mathcal{S}_H \subset \mathcal{S}_{\check{F}}$, $\mathcal{S}_{H_T} \subset \mathcal{S}_{\check{F}_T}$ which is finite. If $h_T \neq \check{f}_T$, let $t_1 < \cdots < t_m$ be all the points in $\mathcal{S}_{\check{F}_T} \cap D_V$. Let t_{i_o} be the smallest point in $\mathcal{S}_{\check{F}_T}$ such that $\check{f}_T(t_{i_o}) \neq h_T(t_{i_o})$. Since $\check{f}_T(t_i) = h_T(t_i)$ for $i < i_o$ and $H_T(-\infty) = \check{F}_T(-\infty) = 0$ by (1) in (2.2), $H_T(v) = \check{F}_T(v)$ for $v < t_{i_o}$. Letting $x = t_{i_o}$, Eq. (8.5) yields

$$\begin{aligned}
&\check{f}_T(t_{i_o}) = \check{f}_T(t_{i_o}) \int_{v<t_{i_o}} 1 dF_R(v) + S_R(t_{i_o}-) h_T(t_{i_o}),\\
\Rightarrow &\check{f}_T(t_{i_o})(1 - F_R(t_{i_o}-)) = S_R(t_{i_o}-) h_T(t_{i_o}),\\
\Rightarrow &\check{f}_T(t_{i_o}) = h_T(t_{i_o}) \text{ (as } S_R(t_{i_o}-) > 0 \text{ due to } t_{i_o} \in D_V),
\end{aligned}$$

contradicting the assumption $\check{f}_T(t_{i_o}) \neq h_T(t_{i_o})$. The contradiction implies that $h_T = \check{f}_T$ on D_V. By (7.2), $H_T = \check{F}_T$ on $[-\infty, \infty)$. Since $H_T(\infty) = \check{F}_T(\infty) = 1$, $H_T = \check{F}_T$.

Step 3 (conclusion). Since $h_T = \check{f}_T$ by Step 2, (8.4) is equivalent to

$$\check{f}_{C|T}(y|x) = \check{f}_{C|T}(y|x) F_R(x-) + \sum_{w\in\mathcal{J}, y\in w} \sum_{j\in w} h_{C|T}(j|x) \frac{q(w|x)\check{f}_{C|T}(y|x)}{\sum_{k\in w}\check{f}_{C|T}(k|x)}, \ (x,y) \in \mathcal{S}_{\check{F}}; \tag{8.6}$$

$$\text{or } \check{f}_{C|T}(y|x) = \check{f}_{C|T}(y|x) \sum_{w\in\mathcal{J}} \frac{q(w|x)\mathbf{1}_{(y\in w)}}{S_R(x-)} \frac{\sum_{j\in w} h_{C|T}(j|x)}{\sum_{k\in w}\check{f}_{C|T}(k|x)}, \ \ (x,y) \in \mathcal{S}_{\check{F}}. \tag{8.7}$$

Verify that $h_{C|T} = \check{f}_{C|T}$ is a solution to (8.7), as the right hand side of (8.7) becomes

$$\check{f}_{C|T}(y|x) \sum_{w\in\mathcal{J}} \frac{q(w|x)\mathbf{1}_{(y\in w)}}{S_R(x-)} = \check{f}_{C|T}(y|x) \frac{S_R(x-)}{S_R(x-)} = \check{f}_{C|T}(y|x)$$

by (8.1). We shall prove that

$$\text{the solution } h_{C|T}(\cdot|x) \text{ to (8.7) is unique a.e. in } x \text{ on } D_V \text{ w.r.t. } \mu_{F_T}. \tag{8.8}$$

Consider two cases: (1) $\check{f}_{C|T}(c|x) = 0$ for some c, (2) $\check{f}_{C|T}(c|x) > 0$ for each c.

In case (1), since $\mathcal{S}_{\check F} \subset \mathcal{S}_F$ (by (7.2)), by A2 $\exists\ w = \{c\} \in \mathcal{J}$ such that $q(w|x) > 0$. Then it follows from (8.7) that if $y = c$, then (8.7) yields

$$\check f_{C|T}(c|x) = \check f_{C|T}(c|x) \sum_{w\in\mathcal{J}} [\mathbf{1}_{(w=\{c\})} + \mathbf{1}_{(w\neq\{c\})}] \frac{q(w|x)\mathbf{1}_{(c\in w)}}{S_R(x-)} \frac{\sum_{j\in w} h_{C|T}(j|x)}{\sum_{k\in w} \check f_{C|T}(k|x)}, \quad (x,c) \in \mathcal{S}_{\check F}.$$

Thus $0 = 0(\frac{\sum_{j\in\{c\}} h_{C|T}(c|x)}{0} + 0)$. It implies $h_{C|T}(c|x) = 0$ as $\frac{0}{0} \overset{def}{=} 1$. Since $\check f_{C|T}(c|x) = h_{C|T}(c|x) = 0$, it reduces to case (2) by replacing $J-1$ for J.

Now consider case (2). Then (8.7) is equivalent to

$$1 = \sum_{w\in\mathcal{J}, y\in w} \frac{q(w|x)}{S_R(x-)} \frac{\sum_{j\in w} h_{C|T}(j|x)}{\sum_{k\in w} \check f_{C|T}(k|x)},\ y \in \{1, ..., J\}. \tag{8.9}$$

Treating $\check f_{C|T}$ as given constant, $\mathbf{h}' = (h_{C|T}(1|x), ..., h_{C|T}(J|x))$ as an undetermined vector, and $\mathbf{1}$ as the vector with coordinates 1, Eq. (8.9) is actually a system of linear equations in $h_{C|T}$, say $\mathbf{1} = A\mathbf{h}$, where $A = \mathcal{K}_\phi' \mathcal{K}_{q,\check f} \mathcal{K}_\phi$, $\mathcal{K}_{q,\check f}$ is the $k_t \times k_t$ diagonal matrix with diagonal elements $\frac{q(W_j|x)}{S_R(x-)\sum_{k\in W_j} \check f_{C|T}(k|x)}$, W_1, ..., W_{k_t} are all the elements of $\mathcal{J}$ such that $q(W_j|x) > 0$, and $\mathcal{K}_\phi$ is the $k_t \times J$ matrix with the j^{th} row $(\phi(W_j))'$ (see A1, which is implied by A2). Verify that $\mathcal{K}_\phi$ is of full rank J and $\mathcal{K}_{q,\check f}$ is a diagonal matrix of full rank k_t, thus $\mathcal{K}_\phi' \mathcal{K}_{q,\check f} \mathcal{K}_\phi$ is of full rank J. Consequently, the solution of the equation $\mathbf{1} = A\mathbf{h}$ is unique and $h_{C|T} = \check f_{C|T}$, as the latter is a solution to (8.9). Thus (8.8) holds. This concludes the proof of the lemma. □

Proof of Lemma 7.7. By assumption $||\mathcal{S}_{\check F}||$ is finite, let (t_1, c_1),..., (t_m, c_m) be all the distinct points in $\mathcal{S}_{\check F} \cap (D_V \times \mathcal{C}_r)$. WLOG, we can assume that $\tau < \infty$ and $(\infty, J) \in \mathcal{S}_{\check F}$. Otherwise, the proof is simpler as $\mathcal{S}_{\check F} \subset (D_V \times \mathcal{C}_r)$. Let $(t_{m+1}, c_{m+1}) = (\infty, J)$. Let $g_i(\mathbf{x}) = \mathbf{1}_{(\mathbf{x} \geq (t_i, c_i))}$ and $H_i = \mathcal{R}_{\check F}(g_i)$, $i = 1$,..., $m+1$.

By Lemma 7.4, $\mathcal{S}_{\mathcal{R}_{\check F}} \subset \mathcal{S}_{\check F}$. We shall show that $\forall\ i \in \{1, ..., m+1\}$,

$$H_i(-\infty, -\infty) = 0,\ H_i(\infty, \infty) = 1 \text{ and } \mu_{H_i}(\{(t_k, c_k)\}) \geq 0, \text{for all possible } (i,k). \tag{8.10}$$

Now $\forall\ i$, $H_i(-\infty, -\infty) = 0$ follows from Lemma 7.4 and (1) in (2.2). Moreover,

$$\begin{aligned} H_i(\infty, \infty) = &\int_{v<\infty} \frac{\check F(\infty, \infty) - \check F(v, \infty)}{1 - \check F_T(v)} (1 - g_i(v, J)) dF_R(v) \\ &+ \sum_{w\in\mathcal{J}} \int_{v\leq\infty} \int_{c\in w} \frac{\sum_{h\leq\infty, h\in w} \check f_{C|T}(h|v)}{\sum_{j\in w} \check f_{C|T}(j|v)} q(w|v) dg_i(v,c) \\ =&\int_{v<\infty} (1 - g_i(v,J)) dF_R(v) + \sum_{w\in\mathcal{J}} \int_{v\leq\infty} \int_{c\in w} q(w|v) dg_i(v,c) \\ =&\int_{v<t_i} dF_R(v) + \sum_{w\in\mathcal{J}, c_i\in w} q(w|t_i) \qquad \text{(see A2)} \\ =&\int_{v<t_i} dF_R(v) + \sum_{w\in\mathcal{J}, c_i\in w} \sum_{h:\ w\in P_h} f_\Delta(h) S_{R|\Delta}(t_i - |h) \\ =&P(R < t_i) + S_R(t_i-) = 1. \end{aligned}$$

Thus the first 2 equations in (8.10) hold. Notice that

$$\begin{aligned}
H_i(t_k,c_k) =& \int_{v<t_k} \frac{\check F(t_k,c_k)-\check F(v,c_k)}{1-\check F_T(v)}(1-g_i(v,J))dF_R(v) \\
&+\sum_{w\in\mathcal{J}}\int_{v\le t_k}\int_{c\in w}\frac{\sum_{h\le c_k,h\in w}\check f_{C|T}(h|v)}{\sum_{j\in w}\check f_{C|T}(j|v)}q(w|v)dg_i(v,c)\\
=&\sum_{h=1}^{J}\mathbf{1}_{(h\le c_k)}\{\int_{v<t_k}(\check F^s_h(t_k)-\check F^s_h(v))\frac{1-g_i(v,J)}{1-\check F_T(v)}dF_R(v)\\
&+\sum_{w\in\mathcal{J}}\int_{v\le t_k}\int_{c\in w}\frac{\mathbf{1}_{(h\in w)}\check f_{C|T}(h|v)}{\sum_{j\in w}\check f_{C|T}(j|v)}q(w|v)dg_i(v,c)\}
\end{aligned}$$

$$\begin{aligned}
&\mu_{H_i}(\{(t_k,c_k)\})\\
=&\mathbf{1}_{(h=c_k)}\{\int_{v<t_k}(\check F^s_h(t_k)-\check F^s_h(t_k-))\frac{1-g_i(v,J)}{1-\check F_T(v)}dF_R(v)\\
&+\sum_{w\in\mathcal{J}}\int_{v=t_k}\int_{c\in w}\frac{\mathbf{1}_{(h\in w)}\check f_{C|T}(h|v)}{\sum_{j\in w}\check f_{C|T}(j|v)}q(w|v)dg_i(v,c)\}\\
=&\int_{v<t_k}\mu_{\check F}(\{(t_k,c_k)\})\frac{1-1(v\ge t_i)}{1-\check F_T(v)}dF_R(v)+\sum_{w\in\mathcal{J}}\mathbf{1}_{(c_i,c_k\in w)}\frac{\check f_{C|T}(c_k|t_k)}{\sum_{j\in w}\check f_{C|T}(j|t_k)}q(w|t_k)\mathbf{1}_{(t_k=t_i)}\\
\ge&0,
\end{aligned}$$

Thus (8.10) holds. (8.10) implies that $H_i\in\Theta_o$.

Notice that $(g_1, ..., g_{m+1})$ is a base of the linear space $\mathcal{D}_{\check F}$ with $g_i(\infty,J)=1$ and $g_i\in\Theta_o$, thus each H in $\Theta_o\cap\mathcal{D}_{\check F}$ satisfies $H=\sum_{i=1}^{m+1}a_ig_i$, where $\sum_i a_i=1$ and $a_i\ge 0$. It follows that $\mathcal{R}_{\check F}(H)\in\Theta_o$. □

Proof of Lemma 7.8. Since $||\mathcal{S}_{\check F}||$ is finite, by Lemma 7.4, $\mathcal{R}_{\check F}$ is a linear map from $\mathcal{D}_{\check F}$ to $\mathcal{D}_{\check F}$. WLOG, we can assume that (∞,J) belongs to $\mathcal{S}_{\check F}$ $(=\{(t_1,c_1),...,(t_{m+1},c_{m+1})\})$, where $(t_{m+1},c_{m+1})=(\infty,J)$. In order to show that $\mathcal{R}^{-1}_{\check F}$ exists on $\mathcal{D}_{\check F}\cap\mathcal{D}_i$, it suffices to show that (1) $\mathcal{R}_{\check F}$ is 1-1 on $\mathcal{D}_{\check F}\cap\mathcal{D}_i$ and (2) $\mathcal{R}_{\check F}(H)(\infty,J)=i$ if $H\in\mathcal{D}_{\check F}\cap\mathcal{D}_i$.

Now suppose that $\mathcal{R}_{\check F}(H)\equiv 0$ and $H\in\mathcal{D}_{\check F}\cap\mathcal{D}_0$. Since $\mathcal{R}_{\check F}(\check F)=\check F$ by Lemma 7.2 $\check F+0=\mathcal{R}_{\check F}(\check F)+\mathcal{R}_{\check F}(H)=\mathcal{R}_{\check F}(\check F+H)$. Write $H_1=\check F+H$, then $H_1(\infty,J)=1$, $H_1\in\mathcal{D}_{\check F}$ and $\mathcal{R}_{\check F}(H_1)=\check F$. Thus by Lemma 7.6, $H_1=\check F$, that is, $\check F+H=\check F$. Hence $H\equiv 0$, *i.e.*, $\mathcal{R}_{\check F}$ is 1-1 on $D_{\check F}\cap\mathcal{D}_0$. Thus (1) holds for $i=0$.

If $H\in\mathcal{D}_{\check F}\cap\mathcal{D}_0$, then $H(x,y)=\sum_{i=1}^{m+1}d_i\mathbf{1}_{((x,y)\ge(t_i,c_i))}$, where $\sum_{i=1}^{m+1}d_i=0$ and $m+1=||\mathcal{S}_{\check F}||$. It follows that $H=H_+-H_-$, where $H_+(x,y)=\sum_{i=1}^{m+1}d_i\mathbf{1}_{((x,y)\ge(t_i,c_i),d_i\ge 0)}$ and $H_-(x,y)=-\sum_{i=1}^{m+1}d_i\mathbf{1}_{((x,y)\ge(t_i,c_i),d_i<0)}$. Let $d=\sum_{i=1}^{m+1}d_i\mathbf{1}_{(d_i\ge 0)}$. Then H_+/d and H_-/d both belong to Θ_o. By Lemma 7.7, $\mathcal{R}_{\check F}(H_+/d)\in\Theta_o$ and $\mathcal{R}_{\check F}(H_-/d)\in\Theta_o$. Thus $\mathcal{R}_{\check F}(H)\in\mathcal{D}_0$. That is, (2) holds for $i=0$. It follows that $\mathcal{R}^{-1}_{\check F}$ exists on $\mathcal{D}_{\check F}\cap\mathcal{D}_0$.

If H and G are in $\mathcal{D}_{\check F}\cap\mathcal{D}_1$ and $\mathcal{R}_{\check F}(H)=\mathcal{R}_{\check F}(G)$, then $0=\mathcal{R}_{\check F}(H-G)$. Notice that $H-G\in\mathcal{D}_0\cap\mathcal{D}_{\check F}$, thus $H-G=\mathcal{R}^{-1}_{\check F}(0)=0$, that is $H=G$. As a consequence, $\mathcal{R}_{\check F}$ is 1-1 on $\mathcal{D}_1\cap\mathcal{D}_{\check F}$. By Lemma 7.7, $\mathcal{R}_{\check F}(H)\in\Theta_o$ if $H\in\Theta_o$, It follows that $\mathcal{R}^{-1}_{\check F}$ exists on $\mathcal{D}_{\check F}\cap\mathcal{D}_1$.

It is easy to show that $\mathcal{R}^{-1}_{\check F}$ is also linear. □

Proof of Lemma 7.9. By (7.5), it is obvious that statement (1) holds, as

$$\mathcal{R}_{\check F}(\mathcal{U}_o)(x,y) = \int_{v<x} \frac{\check F(x,y)-\check F(v,y)}{1-\check F_T(v)}(0-0)dF_R(v) + \sum_{w\in\mathcal{J}} \int_{\substack{v\le x\\ v\in\mathcal{S}_{\check F_T}}} \int_{c\in w} \frac{\sum_{i\le y, i\in w}\check f_{C|T}(i|v)}{\sum_{j\in w}\check f_{C|T}(j|v)} q(w|v)d0 = 0 = \mathcal{U}_o.$$

Moreover, if $x<\check t_o$, then $\mathcal{R}_{\check F}(\mathcal{U}_{\check F})(x,y)=\mathcal{U}_{\check F}(x,y)=0$. If $x\ge \check t_o$,then $\mathcal{R}_{\check F}(\mathcal{U}_{\check F})(x,y)$

$$\begin{aligned}
&= \int_{v<x} \frac{\check F(x,y)-\check F(v,y)}{1-\check F_T(v)}(\mathcal{U}_{\check F}(\infty,J)-\mathcal{U}_{\check F}(v,J))dF_R(v) \\
&\quad + \sum_{w\in\mathcal{J}} \int_{\substack{v\le x\\ v\in\mathcal{S}_{\check F_T}}} \int_{c\in w} \frac{\sum_{i\le y, i\in w}\check f_{C|T}(i|v)}{\sum_{j\in w}\check f_{C|T}(j|v)} q(w|v)d\mathcal{U}_{\check F}(v,c) \\
&= \int_{\inf \mathcal{S}_{F_R}\le v<\check t_o} \check F(x,y)dF_R(v) + \sum_{w\in\mathcal{J}} \frac{\sum_{i\le y, i\in w}\check f_{C|T}(i|v)}{\sum_{j\in w}\check f_{C|T}(j|v)} \sum_{c\in w} \check f_{C|T}(c|\check t_o)q(w|v)|_{v=\check t_o} \\
&= \check F(x,y)\int_{\inf \mathcal{S}_{F_R}\le v<\check t_o} dF_R(v) + \sum_{i\le y}\check f_{C|T}(i|v)\sum_{w\in\mathcal{J}}\mathbf{1}_{(i\in w)}q(w|v)|_{v=\check t_o} \\
&= \check F(x,y)\int_{\inf \mathcal{S}_{F_R}\le v<\check t_o} dF_R(v) + \sum_{i\le y}\check f_{C|T}(i|v)S_R(v-)|_{v=\check t_o}.
\end{aligned} \tag{8.11}$$

In particular, if $\check F = F$, then $\mathcal{R}_F(\mathcal{U}_F)(x,y) = \sum_{i\le y} f_{C|T}(i|v)S_R(v-)|_{v=t_o}\mathbf{1}_{(x\ge t_o)} = \mathcal{U}_F(x,y)$. Moreover, (8.11) yields statement (3) of the lemma by A3. ▫

Proof of Lemma 7.10. For $\check F\in\Theta_o$ with $\check t_o=t_o$, let C_k be the collection of all the distinct points among $c_{k,j}$s, where $c_{k,j}=\inf\{x:\ \check F_T(x)\ge j/2^k\}$, $j=0,\ ...,\ 2^k$, $k\ge 1$. Let F_k be a function in Θ_o such that $F_k(t,c)=\check F(t,c)\ \forall\ (t,c)\in C_k\times\mathcal{C}_r$ and its F^s_{kc} is a step function with its discontinuity points in C_k. Denote $\mathcal{W}_k$ the subclass of $\mathcal{D}$ such that each member H satisfies that H^s_c is a step function and the collection of discontinuity points is a subset of $\mathcal{S}_{F^s_{kc}}$.

Set $i=1$ first. For each $g\in\mathcal{D}_{\check F}\cap\mathcal{D}_i$ and $k\ge1$, let $g_k\in\mathcal{W}_k\cap\mathcal{D}_i$ be such that $g_k(x,c)=g(x,c)$ if $(x,c)\in C_k\times\mathcal{C}_r$. Then $||g_k-g||\to 0$, since $\mathcal{S}_{g_k}\in\mathcal{S}_{\check F},\mathcal{S}_g\in\mathcal{S}_{\check F}$ and the set $\cup_k C_k$ is dense in $\mathcal{S}_{\check F}$. By Lemma 7.8, $\mathcal{R}^{-1}_{F_k}$ exists, so $\exists!\ H_k\in\mathcal{W}_k\cap\mathcal{D}_i$ such that $g_k=\mathcal{R}_{F_k}(H_k)$. $\forall\ K>k$ and $\forall$ $H\in\mathcal{D}_{\check F}\cap\mathcal{D}_i$, $\mathcal{W}_k\subset\mathcal{W}_K$, $||F_k-F_K||\le 1/2^k$ and $||\mathcal{R}_{F_k}(H)-\mathcal{R}_{F_K}(H)||\to 0$ as $k\to\infty$ by the BCT. Lemma 7.9 says that $\mathcal{U}_o$ and $\mathcal{U}_F$ play the role of 0 and 1 in some sense, respectively, and if k is large, one can write

$$(\mathcal{R}_{F_k}-\mathcal{R}_{F_K})(H)=o(1)\mathcal{U}_{F_K} \text{ and } (\mathcal{R}_{F_k}-\mathcal{R}_{F_K})(H)=o(1)\mathcal{R}_{F_K}(\mathcal{U}_{F_K}), \tag{8.12}$$

as $||\mathcal{R}_{F_k}(\mathcal{U}_{F_k})-\mathcal{U}_{F_k}||\to 0$ by Lemma 7.9. Let $G_k\in\mathcal{D}_{H_k}\cap\mathcal{D}_i$ such that $G_k=\mathcal{R}^{-1}_{F_K}(g_k)$.

$$\begin{aligned}
\mathcal{R}_{F_K}(G_k-H_k) &= \mathcal{R}_{F_K}(G_k)-\mathcal{R}_{F_k}(H_k)+\mathcal{R}_{F_k}(H_k)-\mathcal{R}_{F_K}(H_k) \\
&= g_k-g_k+\mathcal{R}_{F_k}(H_k)-\mathcal{R}_{F_K}(H_k) \\
&= (\mathcal{R}_{F_k}-\mathcal{R}_{F_K})(H_k) \qquad\qquad (\text{as } H_k=\mathcal{R}^{-1}_{F_k}(g_k)).
\end{aligned}$$

$$(\mathcal{R}^{-1}_{F_K}-\mathcal{R}^{-1}_{F_k})(g_k)=G_k-H_k = \mathcal{R}^{-1}_{F_K}((\mathcal{R}_{F_k}-\mathcal{R}_{F_K})(H_k)) \text{ (by Lemma 7.8,}$$

as $(\mathcal{R}_{F_k} - \mathcal{R}_{F_K})(H_k) \in \mathcal{D}_{F_K} \cap \mathcal{D}_0)$.

$$
\begin{aligned}
||(\mathcal{R}^{-1}_{F_K} - \mathcal{R}^{-1}_{F_k})(g_k)|| &= ||\mathcal{R}^{-1}_{F_K}((\mathcal{R}_{F_k} - \mathcal{R}_{F_K})(H_k))|| \\
&= ||\mathcal{R}^{-1}_{F_K}(o(1)\mathcal{R}_{F_K}(\mathcal{U}_{F_K}))|| && \text{(by (8.12))} \\
&= ||o(1)\mathcal{R}^{-1}_{F_K}(\mathcal{R}_{F_K}(\mathcal{U}_{F_K}))|| && \text{(by Lemma 7.9)} \\
&= ||o(1)|| \cdot ||\mathcal{U}_{F_K}|| \to 0 \cdot 1 \text{ (by Lemma 7.9).} && (8.13)
\end{aligned}
$$

$$
\begin{aligned}
||H_k - H_K|| \le& ||\mathcal{R}^{-1}_{F_k}(g_k) - \mathcal{R}^{-1}_{F_K}(g_k)|| + ||\mathcal{R}^{-1}_{F_K}(g_k) - \mathcal{R}^{-1}_{F_K}(g_K)|| \\
=& ||\mathcal{R}^{-1}_{F_k}(g_k) - \mathcal{R}^{-1}_{F_K}(g_k)|| + ||\mathcal{R}^{-1}_{F_K}(o(1)\mathcal{U}_{F_K})|| \\
& \text{(by Lemma 7.9 as } ||g_k - g|| \to 0 \text{ and thus } ||g_k - g_K|| = o(1)\mathcal{U}_{F_K}\text{)} \\
=& ||\mathcal{R}^{-1}_{F_k}(g_k) - \mathcal{R}^{-1}_{F_K}(g_k)|| + ||\mathcal{R}^{-1}_{F_K}(o(1)\mathcal{R}_{F_K}(\mathcal{U}_{F_K}))|| \text{ (by Lemma 7.9)} \\
\to& 0 + 0 \text{ as } k \to \infty \text{ (by (8.13)).}
\end{aligned}
$$

Thus $\{H_k\}_{k\ge1}$ is a Cauchy sequence. Since $\mathcal{D}_{\check F}$ is a Banach space, $\exists\ H_o \in \mathcal{D}_{\check F}$ such that $||H_k - H_o|| \to 0$. By the BCT, $g = \lim_{k\to\infty} \mathcal{R}_{F_k}(H_k) = \mathcal{R}_{\check F}(H_o)$. Define $H_o = \mathcal{R}^{-1}_{\check F}(g)$.

Verify that the foregoing arguements hold also for $i = 0$. ▫

Proof of Lemma 7.13. Since $\mathcal{S}_H = \mathcal{S}_F, \mathcal{D}_H = \mathcal{D} = \mathcal{D}_F$ by (7.2). Given $H \in \Theta_o$, let $G \in \Theta_o$ be an arbitrary solution to $H = \mathcal{R}_H(G)$. Since $H = \mathcal{R}_H(H)$ by (3) of Lemma 7.2, $0 = H - H = \mathcal{R}_H(G) - \mathcal{R}_H(H) = \mathcal{R}_H(G - H)$ by Lemma 7.4. That is, $0 = \mathcal{R}_H(G - H)$. Since $\mathcal{R}^{-1}_H$ exists on $\mathcal{D}_H \cap \mathcal{D}_0$ by Lemma 7.10 and $G - H \in \mathcal{D}_H \cap \mathcal{D}_0$, $0 = \mathcal{R}^{-1}_H(0) = G - H$. That is, $G = H$. In other words, the solution G in Θ_o to the equation $H = \mathcal{R}_H(G)$ is unique. Hence the second statement of Lemma 7.3 holds. Then the solution H to $\mathcal{R}_H(F) = H$ is unique in Θ_o by Lemma 7.3, and $H = F$ by Lemma 7.2. Since $\mathcal{B}_H(Q) = \mathcal{R}_H(F)$ for each $H \in \Theta_o$ by Lemma 7.2, the solution H in Θ_o to $H = \mathcal{B}_H(Q)$ is unique and is $H = F$.

Now consider $H = \mathcal{B}_{H,f_{C|T,\cdot}}(Q)$. Notice that $H \in \Theta_o$ and thus $h_{C|T,w}$ are constant in $w \in \mathcal{J}$. By (7.6) and (7.7)

$$
\begin{aligned}
&\mathcal{B}_{H,f_{C|T,\cdot}}(Q)(x,y) \quad (= \mathcal{B}^o_H(Q)(x,y)) \\
=& \sum_{w\in\mathcal{J}} \Big[\int_{v\le x, v\in D^c_T, v\in\mathcal{S}_{H_T}} \frac{\sum_{i\le y, i\in w} f_{C|T}(i|v)}{\sum_{j\in w} f_{C|T}(j|v)} dQ(v,1,w) \\
&+ \int_{v\le x, v\in D_T, v\in\mathcal{S}_{H_T}} \frac{\sum_{i\le y, i\in w} h_{C|T}(i|v)}{\sum_{j\in w} h_{C|T}(j|v)} dQ(v,1,w)\Big] + \int_{v<x} \frac{H(x,y) - H(v,y)}{1 - H_T(v)} dQ(v,0,\mathcal{C}_r) \\
=& \sum_{w\in\mathcal{J}} \int_{v\le x, v\in\mathcal{S}_{H_T}} \int_{c\in w} \Big[\mathbf{1}_{(v\in D^c_T)} \frac{\sum_{i\le y, i\in w} f_{C|T}(i|v)}{\sum_{j\in w} f_{C|T}(j|v)} \\
&+ \mathbf{1}_{(v\in D_T)} \frac{\sum_{i\le y, i\in w} h_{C|T}(i|v)}{\sum_{j\in w} h_{C|T}(j|v)}\Big] q(w|v) dF(v,c) + \int_{v<x} \frac{H(x,y) - H(v,y)}{1 - H_T(v)} (1 - F_T(v)) dF_R(v) \\
=& \mathcal{R}^o_H(F) \text{ (by (7.8), as } F \in \Theta_o).
\end{aligned}
$$

Now verify that replacing $(\mathcal{B}^o_H, \mathcal{R}^o_H)$ for $(\mathcal{B}_H, \mathcal{R}_H)$ in Lemmas 7.2 through 7.9, except for Lemma 7.5, one can show that those modified lemmas hold too. Then the proof of the lemma regarding $H = \mathcal{B}_{H,f_{C|T,\cdot}}(Q)$ follows almost line by line as the proof in the first paragraph making use of the modified lemmas. In particular, Eq. (8.5) and Eq. (8.6) become

$$
\check f_T(x) = \check f_T(x) \int_{v<x} \frac{1 - H_T(x)}{1 - \check F_T(x)} dF_R(v) + S_R(x-)[\mathbf{1}_{(x\in D^c_T)} \tilde f_T(x) + \mathbf{1}_{(x\in D_T)} h_T(x)],
$$

$$\check{f}_{C|T}(y|x)S_R(x-) = \sum_{w\in\mathcal{J},y\in w} q(w|x)[\frac{\sum_{j\in w} h_{C|T}(j|x)}{\sum_{k\in w}\check{f}_{C|T}(k|x)}\mathbf{1}_{(x\in D_T)} + \frac{\sum_{j\in w}\tilde{f}_{C|T}(j|x)}{\sum_{k\in w}\check{f}_{C|T}(k|x)}\mathbf{1}_{(x\in\mathcal{D}_T^c)}],$$

$\forall\ (x,y)\in\mathcal{S}_{\check{F}}$ and $H\in\mathcal{D}$. This completes the proof of the lemma. ▫

Proof of Lemma 7.15. **Step 1** (preliminary). By definition, $\tilde{F}^*$ is the GMLE of the modified data $(V_{mi},\delta_i,\mathcal{M}_i)$s, thus it maximizes the modified likelihood function

$$\Lambda_2 = \prod_{i=1}^n \mu_{\check{F}}(I_{mi}),\ \text{where } I_{mi} = \begin{cases}(V_i,\infty)\times\mathcal{M}_i & \text{if } \delta_i = 0\\ \{V_{mi}\}\times\mathcal{M}_i & \text{otherwise,}\end{cases}\ \text{and } \check{F} \text{ is a cdf.}$$

It can be shown as in Lemma 7.1 that $H=\tilde{F}^*$ is a solution to the equation $H=\mathcal{B}_H(\hat{Q}_m)$, where $\hat{Q}_m$ is the empirical cdf of Q, based on $(V_{mi},\delta_i,\mathcal{M}_i)$s, and $\mathcal{B}_H$ is defined in (7.4). Let Ω_6 be the event that $\hat{Q}\to Q$. By the SLLN $P(\Omega_6)=1$. Let Ω_7 be the event that

$$\sup_{t\in[-\infty,\infty],u\in\{0,1\},\mathbf{w}\in\mathcal{J}} |\hat{Q}_m(t,u,\mathbf{w}) - Q(t,u,\mathbf{w})| \to 0. \tag{8.14}$$

Step 2 (to show $\Omega_6\subset\Omega_7$). Fix $\omega\in\Omega_6$. Since $V_{mi}=V_i$ if $\delta_i=0$, for each $\mathbf{w}\in\mathcal{J}$, $\sup_{t,u}|\hat{Q}_m(t,u,\mathbf{w}) - \hat{Q}(t,u,\mathbf{w})| = \sup_t|\hat{Q}_m(t,1,\mathbf{w}) - \hat{Q}(t,1,\mathbf{w})|$ $(\stackrel{def}{=} d_{nt})$. Then $d_{nt} = \frac{1}{n}|\sum_{i=1}^n d_i|$, where $d_i = \mathbf{1}_{(V_{mi}\le t,\delta_i=1,\mathcal{M}_i=\mathbf{w})} - \mathbf{1}_{(V_i\le t,\delta_i=1,\mathcal{M}_i=\mathbf{w})}$. Verify that $d_{nt}=0$ if $t=0$ (by the definition of V_{mi}s). For $t>0$, $\exists\ j$ such that $t\in(s_{n,j-1},s_{n,j}]$. Then due to the definition of $s_{n,j}$, $d_i = -1_{(\delta_i=1,V_i\in(s_{n,j-1},t],\ t<s_{n,j})}$.

$$d_{nt} = |\sum_{i=1}^n \frac{1}{n}\mathbf{1}_{(V_i\in(s_{n,j-1},t],\ t<s_{n,j},\delta_i=1,\mathcal{M}_i=\mathbf{w})}| \le \mu_{\hat{F}_T}((s_{n,j-1},t])\mathbf{1}_{(t\in s_{n-j},s_j])} \le 1/\sqrt{n}\to 0.$$

That is, $\sup_{t,u}|\hat{Q}_m(t,u,\mathbf{w}) - \hat{Q}(t,u,\mathbf{w})|\to 0$. Since $\omega\in\Omega_6$, $\sup_{t,u}|\hat{Q}(t,u,\mathbf{w}) - Q(t,u,\mathbf{w})|\to 0$ by assumption on Ω_6. Thus it yields that for $(u,\mathbf{w})\in\{0,1\}\times\mathcal{J}$,

$$\sup_t|\hat{Q}_m(t,u,\mathbf{w}) - Q(t,u,\mathbf{w})| \le \sup_t\{|\hat{Q}_m(t,u,\mathbf{w}) - \hat{Q}(t,u,\mathbf{w})| + |\hat{Q}(t,u,\mathbf{w}) - Q(t,u,\mathbf{w})|\}\to 0.$$

Hence (8.14) holds. That is, $\omega\in\Omega_7$. Thus $\Omega_6\subset\Omega_7$.

Step 3 (to derive $\lim\tilde{f}_{C|T}$). For each $j\in\mathcal{C}_r$, $\tilde{F}_j^{s*}$ is monotone and bounded by 0 and 1, each given subsequence of $\tilde{F}_j^{s*}$ has a further convergent subsequence. Thus each subsequence of $(\tilde{F}_1^{s*},...,\tilde{F}_J^{s*})$, has a further convergent subsequence, as $J<\infty$, and so is $\tilde{F}^*(\cdot,c)$, $c\in\mathcal{C}_r$. Denote the limiting point by H^* together with H_c^{s*}, $c\in\mathcal{C}_r$. For the given $\omega\in\Omega_6$ and for each n, one can define $\tilde{f}_{C|T}(c|t)$ for each $(c,t)\in\mathcal{C}_r\times\{V_i:\ i=1,...,n\}$. Notice also that $\tilde{F}_c^{s*}(x) = \int_{t\le x}\tilde{f}^*_{C|T}(c|t)d\tilde{F}_T^*(t) = \int_{t\le x}\tilde{f}_{C|T}(c|t)d\tilde{F}_T^*(t)$ (see step 2.4)). Since $\tilde{f}_{C|T}$ is bounded and $\mathcal{C}_r\times\{V_i:\ i=1,...,n\}$, is countable, by Helly's selection theorem there is a convergent subsequence of $\tilde{f}_{C|T}$ with limiting point $h_{C|T}$ such that $\tilde{f}_{C|T}(c|t)\to h_{C|T}(c|t)$ $\forall\ (c,t)\in\mathcal{C}_r\times\{V_i:\ i=1,...,n\}$. By taking further subsequence, without loss of generality, one can assume that $\{\tilde{F}^*_{n_h},\tilde{f}_{C|T,n_h}\}_{h\ge1}$ converges. Now by the BCT, $H_c^{s*}(x) = \int_{t\le x}h_{C|T}(c|t)dH_T^*(t)$ $\forall\ x$. Thus for each $c\in\mathcal{C}_r$, $h_{C|T}(c|\cdot) = h^*_{C|T}(c|\cdot)$ (induced by H^*) (except for a set of zero $\mu_{H_T^*}$ measure) is a limiting point of $\tilde{f}_{C|T}$.

Step 4 (conclusion). Verify that $\sup_{x\in D_V}|\tilde{F}_T^*(x) - F_T(x)|\to 0$. Under A1, for each t, if n is large enough, there are $\mathcal{M}_{k_j}$s (where $V_{k_j}, t\in(s_{n,j-1},s_{n,j}]$ and $\delta_{k_j}=1$) such that $\forall\ c\in\mathcal{C}_r\ \exists$

constants b_{cj} satisfying $\phi(\{c\}) = \sum_j b_{cj}\phi(\mathcal{M}_{k_j})$. It follows that $\tilde f^*_{C|T}(\cdot|t)$ does not based on only one $\mathbf{w} \in \mathcal{J}_t$, at least for large enough n. In fact, if n is large, each $\mathbf{w}$ in $\{W \in \mathcal{J}_t : t \in (s_{n,j-1}, s_{n,j}]\}$ would be observed and thus $\tilde f^*_{C|T}(\cdot|t)$ is based on all these $\mathbf{w}$'s. Then for the convergent subsequence $\{\tilde F^*_{n_h}\}_{h\ge1}$, the sequences $\{\tilde f_{C|T,n_h}\}_{h\ge1}$ induced by $\{\tilde F^*_{n_h}\}_{h\ge1}$ converges to the limiting functions $h_{C|T}$ on the set $\{V_i : i = 1, ..., n\}$, which contains the support of $\tilde F^*_T$. Notice that each of the $||\mathcal{J}||+1$ integrands of $\mathcal{B}_{\tilde F^*}$ (see (7.4)), namely, $\frac{\tilde F^*_{n_h}(x,y)-\tilde F^*_{n_h}(v,y)}{1-\tilde F^*_{T,n_h}(v)}$ and $\frac{\sum_{i\le y, i\in\mathbf{W}} \tilde f_{C|T,\mathbf{W},n_h}(i|v)}{\sum_{j\in\mathbf{W}} \tilde f_{C|T,\mathbf{W},n_h}(j|v)}$ where $\mathbf{w} \in \mathcal{J}_t$, are bounded by 1 and converge as $h \to \infty$, for each $x \in D_V$, $y \in \mathcal{C}_r$ and $v \in \{V_i : i = 1, ..., n\}$. Thus by the BCT, the limiting equation of $\tilde F^*(x,y) = \mathcal{B}_{\tilde F^*}(\hat Q)(x,y)$ is $H^*(x,y) = \mathcal{B}_{H^*}(Q)(x,y)$ for all $(x,y) \in D_V \times \mathcal{C}_r$. It can be shown that $\mathcal{S}_{H^*} = \mathcal{S}_F$ and $H^* \in \Theta_o$. By Lemma 7.13, the solution to $\mathcal{B}_H(Q) = H$, $H \in \mathcal{F}$, is unique in the sense specified there. It follows that each limiting point $h_{C|T}$ of $\tilde f_{C|T}$ satisfies that $h_{C|T}(\cdot|t) = f_{C|T}(\cdot|t)$ for each $(t,c) \in \{V_i : i = 1, ..., n\} \times \mathcal{C}_r$, and $\tilde F^* \to F$. Since $\tilde F^s_c(x) = \int_{t\le x} \tilde f_{C|T}(c|t)d\hat F_T(t)$, by the BCT, $\tilde F^s_c(t) \to F^s_c(t)$ on D_V and $\tilde F \to F$ for the given ω. Since ω is arbitrary in Ω_6 and $P(\Omega_6) = 1$, the lemma is proved. □

§9. Proofs in Sections 1 and 2.

§9.1. Remark 9.1. Dinse (1982, p.426) provides a data set with $J = 2$. Dinse comments that the GMLE of $f_{C|T}$ is "extremely erratic", and partitions the observations into several equal-sized groups and derives a new smoothed estimator of $f_{C|T}$ based on the grouped data. The new estimator assigns positive weights to both $(V_i, 1)$ and $(V_i, 2)$, even if the MI is not $\{V_i\} \times \{1, 2\}$. Thus the new estimator is not a GMLE based on the original data. Moreover, most discrete GMLEs of continuous density functions are not consistent, thus most of them are erratic. However, the cdfs of based on the inconsistent GMLE of the densities are often consistent, just like the case we are studying.

§9.2. Proofs in Example 2.1. Under given assumptions, the log likelihood function is

$$\mathcal{L} = n_1\ln(p_1 + p_2) + n_2\ln(1 - p_1 - p_2) + n_3\ln(p_1 + p_3) + n_4\ln(1 - p_1 - p_3),$$

where $n_i = \sum_{j=1}^n \mathbf{1}_{(\mathcal{M}_j = W_i)}$, $p_i = f(1, i)$ and $n = n_1 + \cdots + n_4$. The normal equations are $\frac{n_1}{p_1+p_2} - \frac{n_2}{1-p_1-p_2} + \frac{n_3}{p_1+p_3} - \frac{n_4}{1-p_1-p_3} = 0$, $\frac{n_1}{p_1+p_2} - \frac{n_2}{1-p_1-p_2} = 0$, $\frac{n_3}{p_1+p_3} - \frac{n_4}{1-p_1-p_3} = 0$, which reduce to $\frac{n_1}{p_1+p_2} - \frac{n_2}{1-p_1-p_2} = 0$, $\frac{n_3}{p_1+p_3} - \frac{n_4}{1-p_1-p_3} = 0$. Solving these two equations leads to the GMLE in Example 2.1: $\hat p_2 = r_1 - \hat p_1$, $\hat p_3 = r_2 - \hat p_1$, and $\hat p_4 = 1 - \hat p_1 - \hat p_2 - \hat p_3$, where $\hat p_1$ is arbitrary in $[\max\{0, r_1 + r_2 - 1), \min\{r_1, r_2\}]$, $r_1 = \frac{n_1}{(n_1+n_2)}$ and $r_2 = \frac{n_3}{(n_3+n_4)}$. □

§9.3. Proofs in Example 2.2 (existence of both inconsistent GMLE and consistent GMLE). Suppose that $J = 2$; partitions $P_0 = \{\{1\}, \{2\}\}$ and $P_1 = \{\mathcal{C}_r\}$; F^s_js satisfy $\frac{\partial}{\partial t}F^s_1(t) = p_1\mathbf{1}_{(t\in(1,2))}$ and $\frac{\partial}{\partial t}F^s_2(t) = \frac{p_2}{3}\mathbf{1}_{(t\in(0,3))}$, where $p_1 + p_2 = 1$ and $p_i \ge 0$; $f_\Delta(h) = 1/2$; there is no censoring; $(T, C) \perp (\Delta, R)$ and A1 holds. Verify that

(a) all n $(V_i, \delta_i, \mathcal{M}_i)$s are of the forms (1) $(V_i, 1, \{j\})$, $j \in \{1, 2\}$, or (2) $(V_i, 1, \mathcal{C}_r)$;
(b) they are all distinct and thus each of them is an MI induced by these n observations.

Thus the GMLE $\hat F_1$ assigns weight $\frac{1}{n}$ to $\{V_i\}\times\{j\}$ if $\mathcal{M}_i = \{j\}$, and assigns weight $\frac{1}{2n}$ to $\{V_i\}\times\{j\}$ if $\mathcal{M}_i = \mathcal{C}_r$ and $j \in \mathcal{C}_r$. Now $\hat F_1(t, 1) = \frac{1}{2n}\sum_{i=1}^n \mathbf{1}_{(V_i\le t, \mathcal{M}_i=\mathcal{C}_r)} + \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{(V_i\le t, \mathcal{M}_i=\{1\})} \to \frac{1}{2}\frac{p_2 t}{3} f_\Delta(2) = p_2t/12 \ne F(t, 1)$. Thus $\hat F_1$ is not consistent on $(0, 1)$. In the aforementioned example, since the data are of the form either (1) or (2), there is a consistent GMLE with a closed form solution: $\hat F_3(t, 1) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{(V_i\le t)}\tilde f_{C|T}(1|V_i)$, where $\tilde f_{C|T}(1|t) = \begin{cases} n_{1t}/n_t & \text{if } n_t \ne 0 \\ 1/2 & \text{otherwise} \end{cases}$, $n_t = \sum_{i=1}^n \mathbf{1}_{(|V_i-t|\le\frac{1}{\sqrt n},\ \delta_i=1,\ \mathcal{M}_i=\{1\}\ or\ \{2\})}$ and $n_{1t} = \sum_{i=1}^n \mathbf{1}_{(|V_i-t|\le\frac{1}{\sqrt n},\ \delta_i=1,\ \mathcal{M}_i=\{1\})}$. □

§9.4. Proofs in Example 6.1. Suppose $T \equiv 1$. $J = 3$, $P_2 = \{\{1, 2\}, \{3\}\}$, $P_3 = \{\{1, 3\}, \{2\}\}$, $f_\Delta(2) = f_\Delta(3) = 1/2$. $R \equiv 2$. The possible observations are of the forms $(1, 1, W_i)$, $i = 1$, ..., 4, where $W_1 = \{1, 2\}$, $W_2 = \{2\}$, $W_3 = \{3\}$ and $W_4 = \{1, 3\}$. Let $N_j = \sum_{i=1}^n \mathbf{1}_{\mathcal{M}_i=W_j}$,

$j = 1, ..., 4$. Then the GMLEs are $\hat{f}_{C|T}(2|1) = \frac{N_2}{N_2+N_4}$ and $\hat{f}_{C|T}(3|1) = \frac{N_3}{N_1+N_3}$, provided that $\hat{\theta} \leq 1$, where $\hat{\theta} = \frac{N_2}{N_2+N_4} + \frac{N_3}{N_1+N_3}$. However, if $f_C(2) = f_C(3) = 1/2$, then $\sqrt{n}(\hat{\theta} - 1)$ converges in distribution to $N(0, \sigma^2)$, where $\sigma > 0$. Thus $P\{\hat{\theta} > 1\} \to 1/2$. It follows that the GMLE $\hat{F}(1,1) = \hat{f}_{C|T}(1|1) = \begin{cases} 1 - \hat{\theta} & \text{if } \hat{\theta} < 1 \\ 0 & \text{otherwise.} \end{cases}$ Thus it is not asymptotically normally distributed.