The Annals of Statistics
2014, Vol. 42, No. 5, 2092–2137
DOI: 10.1214/14-AOS1256

MAXIMUM SMOOTHED LIKELIHOOD ESTIMATORS FOR THE INTERVAL CENSORING MODEL

BY PIET GROENEBOOM

Delft University

We study the maximum smoothed likelihood estimator (MSLE) for interval censoring, case 2, in the so-called separated case. Characterizations in terms of convex duality conditions are given and strong consistency is proved. Moreover, we show that, under smoothness conditions on the underlying distributions and using the usual bandwidth choice in density estimation, the local convergence rate is $n^{-2/5}$ and the limit distribution is normal, in contrast with the rate $n^{-1/3}$ of the ordinary maximum likelihood estimator.

1. Introduction. In [10], the maximum smoothed likelihood estimator (MSLE) and smoothed maximum likelihood estimator (SMLE) were studied for the current status model, the simplest interval censoring model. It is called the interval censoring, case 1, model in [5] and [12]. It was shown in [10] that, under certain regularity conditions, the MSLE and the SMLE, evaluated at a fixed interior point, converge at rate $n^{-2/5}$ to the real underlying distribution function, if one takes a bandwidth of order $n^{-1/5}$. This convergence rate is faster than the convergence rate of the nonsmoothed maximum likelihood estimator, which is $n^{-1/3}$ in this situation, as shown in [5] and [12]. Moreover, the limit distribution is normal, in contrast with the limit distribution of the nonsmoothed maximum likelihood estimator.

The interval censoring model, where there is an interval in which the relevant (unobservable) event takes place, is more common, in particular in medical statistics. It is called the interval censoring, case 2, model in [5] and [12]. A preliminary discussion of the SMLE in this situation can be found in [11], where it was shown that the development of the theory of the SMLE for this model crucially depends on a further analysis of the integral equations, studied in [2, 3] and [4]. In the present paper, we study the MSLE and prove a consistency and asymptotic normality result for this estimator. We also discuss algorithms for computing the MSLE, which is a rather complicated issue.

We recall the interval censoring, case 2, model. Let $X_1, \dots, X_n$ be a sample of unobservable random variables from an unknown distribution function F_0 on

Received March 2012; revised July 2014.
MSC2010 subject classifications. Primary 62G05, 62N01; secondary 62G20.
Key words and phrases. Interval censoring, smoothed maximum likelihood estimator, maximum smoothed likelihood estimator, consistency, asymptotic distribution, integral equations, kernel estimators.

$[0, \infty)$. Suppose that one can observe n pairs (T_i, U_i), independent of X_i, where $U_i > T_i$. Moreover,

$$\Delta_{i1} \stackrel{\text{def}}{=} 1_{\{X_i \le T_i\}}, \qquad \Delta_{i2} \stackrel{\text{def}}{=} 1_{\{T_i < X_i \le U_i\}} \quad \text{and} \quad \Delta_{i3} \stackrel{\text{def}}{=} 1 - \Delta_{i1} - \Delta_{i2},$$

provide the only information one has on the position of the random variables X_i with respect to the observation times T_i and U_i. In this set-up, one wants to estimate the unknown distribution function F_0, generating the "unobservables" X_i, on an interval $[0, M]$.

Interestingly, from a computational point of view, the MLE for the distribution function of the hidden variable in the case that one has more observation times $T_i, U_i, V_i, \dots$ "per hidden variable," can always be reduced to the case of interval censoring, case 2. This follows from the fact that at most two of the observation times of the set $\{T_i, U_i, V_i, \dots\}$ are relevant for the location of the hidden variable. If we know that the hidden variable is located between two observation times, while the other observation times for this hidden variable are either more to the right or more to the left, then these other observation times do not give extra information and can be discarded in computing the MLE. Likewise, if we know that the hidden variable lies to the right of all these observation times, all observation times smaller than the largest one do not give extra information, with a similar situation if we know that the hidden observation time lies to the left of the smallest observation time for this variable. So, in the last two cases, only one observation time gives relevant information and the other ones can be discarded. This motivates concentrating on the interval censoring, case 2, model, as an extension of the current status model.

The MSLE (maximum smoothed likelihood estimator) is defined in the following way. Let g be the joint density of the observation pairs (T_i, U_i), with first marginal g_1 and second marginal g_2. Moreover, let the densities h_{01}, h_{02} and h_0 be defined by

$$\begin{aligned} h_{01}(t) &= F_0(t) g_1(t), \\ h_{02}(u) &= \{1 - F_0(u)\} g_2(u), \\ h_0(t,u) &= \{F_0(u) - F_0(t)\} g(t,u). \end{aligned} \tag{1.1}$$

We define $\tilde{h}_{nj}$, $j = 1, 2$ and $\tilde{h}_n$ as the estimates of the densities h_{0j}, $j = 1, 2$ and the 2-dimensional density h_0, respectively, where

$$\tilde{h}_{n1}(t) = \frac{1}{n} \sum_{i=1}^n K_{b_n}(t - T_i) \Delta_{i1}, \qquad \tilde{h}_{n2}(u) = \frac{1}{n} \sum_{i=1}^n K_{b_n}(u - U_i) \Delta_{i3}, \tag{1.2}$$

$$\tilde{h}_n(t,u) = \frac{1}{n} \sum_{i=1}^n K_{b_n}(t - T_i) K_{b_n}(u - U_i) \Delta_{i2} \tag{1.3}$$

and

$$K_{b_n}(x) = \frac{1}{b_n} K\left(\frac{x}{b_n}\right),$$

for a symmetric continuously differentiable kernel K with compact support, like the triweight kernel

$$K(x) = \tfrac{35}{32}\left(1-x^2\right)^3 1_{[-1,1]}(x), \qquad x \in \mathbb{R}. \tag{1.4}$$

At boundary points, we use a boundary correction by replacing the kernel K by a linear combination of $K(u)$ and $uK(u)$. For example, if $t \in [0, b_n)$, we define

$$\tilde{h}_{n1}(t) = \alpha(t/b_n)\frac{1}{n}\sum_{i=1}^{n} K_{b_n}(t-T_i)\Delta_{i1} + \beta(t/b_n)\frac{1}{n}\sum_{i=1}^{n}\frac{t-T_i}{b_n}K_{b_n}(t-T_i)\Delta_{i1},$$

where the coefficients $\alpha(u)$ and $\beta(u)$ are defined by

$$\alpha(u)\int_{-1}^{u} K(x)\,dx + \beta(u)\int_{-1}^{u} xK(x)\,dx = 1, \qquad u \in [0,1] \tag{1.5}$$

and

$$\alpha(u)\int_{-1}^{u} xK(x)\,dx + \beta(u)\int_{-1}^{u} x^2K(x)\,dx = 0, \qquad u \in [0,1]. \tag{1.6}$$

It may happen that $\tilde{h}_{n1}(t) < 0$; in that case we put $\tilde{h}_{n1}(t) = 0$.

If $t \in (M-b_n, M]$, we similarly define

$$\begin{aligned}\tilde{h}_{n1}(t) = {} & \alpha\big((M-t)/b_n\big)\frac{1}{n}\sum_{i=1}^{n} K_{b_n}(t-T_i)\Delta_{i1}\\ & - \beta\big((M-t)/b_n\big)\frac{1}{n}\sum_{i=1}^{n}\frac{t-T_i}{b_n}K_{b_n}(t-T_i)\Delta_{i1},\end{aligned}$$

where the functions α and β are again defined by (1.5) and (1.6). The estimates $\tilde{h}_{n2}$ and $\tilde{h}_n$ are similarly defined if one or more (in the case of $\tilde{h}_n$) arguments have distance less than b_n to the boundary; for $\tilde{h}_n$ we apply this to the factors of the product of the kernels separately, in the same way as for the one-dimensional estimates $\tilde{h}_{nj}$. We finally divide $\tilde{h}_{n1}(t)$, $\tilde{h}_{n2}(t)$ and $\tilde{h}_n(t,u)$ by

$$\int_{[0,M]}\big\{\tilde{h}_{n1}(x) + \tilde{h}_{n2}(x)\big\}\,dx + \int_{[0,M]^2}\tilde{h}_n(x,y)\,dx\,dy,$$

(i.e., by a discrete approximation to this quantity) to give a total mass approximately equal to 1 to the observation density.

The MSLE $\hat{F}_n$ is now defined as the distribution function, maximizing the criterion function

$$\begin{aligned}\ell(F) = {} & \int \tilde{h}_{n1}(t)\log F(t)\,dt + \int \tilde{h}_{n2}(u)\big\{1-F(t)\big\}\,du\\ & + \int \tilde{h}_n(t,u)\log\big\{F(u)-F(t)\big\}\,dt\,du,\end{aligned} \tag{1.7}$$

as a function of the distribution function F. But in practice we discretize, and maximize

$$
\begin{aligned}
(1.8)\qquad & \sum_{i=1}^{m}\{\tilde h_{n1}(t_i)\log F(t_i)\}d_i + \sum_{i=1}^{m}\{\tilde h_{n2}(t_i)\log\{1-F(t_i)\}\}d_i \\
& + \sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\{\tilde h_n(t_i,t_j)\log\{F(t_j)-F(t_i)\}\}d_i d_j,
\end{aligned}
$$

over all distribution functions F, where $0 = t_0 < t_1, \dots, < t_m = M$ are the points of a grid and $d_i = t_i - t_{i-1}$, $i = 1, \dots, m$.

Note that $\ell(F)$ is a smoothed log likelihood for F and, therefore, the maximizing (sub)distribution function F is called the maximum smoothed likelihood estimator (MSLE). Also note that the maximization of (1.7) is the same as the minimization of the Kullback–Leibler distance

$$
\begin{aligned}
& \int \tilde h_{n1}(t)\log\frac{\tilde h_{n1}(t)}{F(t)\tilde g_{n1}(t)}\,dt + \int \tilde h_{n1}(t)\log\frac{\tilde h_{n2}(t)}{\{1-F(t)\}\tilde g_{n2}(t)}\,dt \\
& + \int \tilde h_n(t,u)\log\frac{\tilde h_n(t,u)}{\{F(u)-F(t)\}\tilde g_n(t,u)}\,dt\,du,
\end{aligned}
$$

where $\tilde g_{ni}$ and $\tilde g_n$ are kernel estimates of the densities g_i and g, computed in the same way as the estimates h_{ni} and $\tilde h_n$ (but without the indicators Δ_{ij}).

Defining the SMLE (smoothed maximum likelihood estimator) is somewhat easier. If we have computed the ordinary MLE $\hat F_n$, we simply define the SMLE $F_n^{\mathrm{SML}}(x)$ by

$$
\begin{aligned}
(1.9)\qquad & F_n^{\mathrm{SML}}(x) = \int \mathbb{K}_{b_n}(x-y)\,d\hat F_n(y), \\
& \mathbb{K}_{b_n}(u) = \int_{-\infty}^{u/b_n} K(x)\,dx,
\end{aligned}
$$

where K is of type (1.4) again. A picture of the MLE, the MSLE and the SMLE for a sample of size $n = 1000$ from an exponential distribution function is shown in Figure 1. A picture of the bivariate observation density g, with $\varepsilon = 0.1$, is shown in Figure 2.

1.1. *The SMLE and MSLE for the current status model.* Before embarking on the theory for this model, it might be instructive to recapitulate the rather different ways in which the asymptotic distributions of the SMLE and the MSLE are derived for the simpler current status model. In this case, the data are given by

$$(T_1, \Delta_1), \dots, (T_n, \Delta_n),$$

where

$$\Delta_i = 1_{\{X_i \le T_i\}},$$

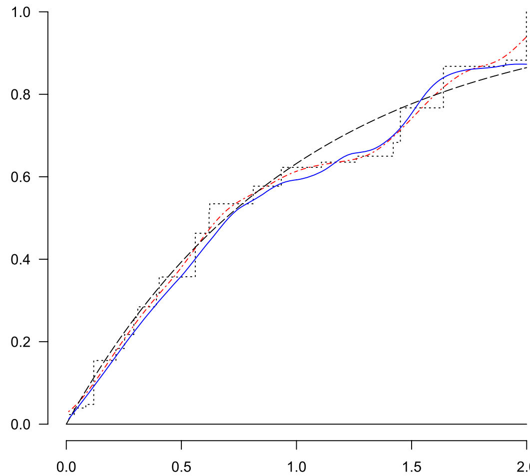

FIG. 1. *The MSLE* (*solid*), *SMLE* (*dashed-dotted*) *and MLE* (*dotted*) *on* $[0, 1]$ *for a sample of size* $n = 1000$ *from the exponential distribution function* $F_0(x) = 1 - \exp\{-x\}$ (*dashed*); *the bivariate observation density is* $g(x, y) = 6(y - x - \varepsilon)^2/\{(2 - x - \varepsilon)(2 - \varepsilon)\}^2$, $x + \varepsilon < y$ *on the triangle with vertices* $(0, \varepsilon)$, $(0, 2)$ *and* $(2 - \varepsilon, 2)$, *where* $\varepsilon = 0.1$. *The bandwidth for the computation of the MSLE was* $b_n = n^{-1/5} \approx 0.25119$.

and X_i and T_i are independent.

Let $\tilde{F}_n^{(\mathrm{SML})}$ be the SMLE for the current status model, defined by (1.9), but now using the MLE $\hat{F}_n$ in the current status model. It is shown in [10] that, under

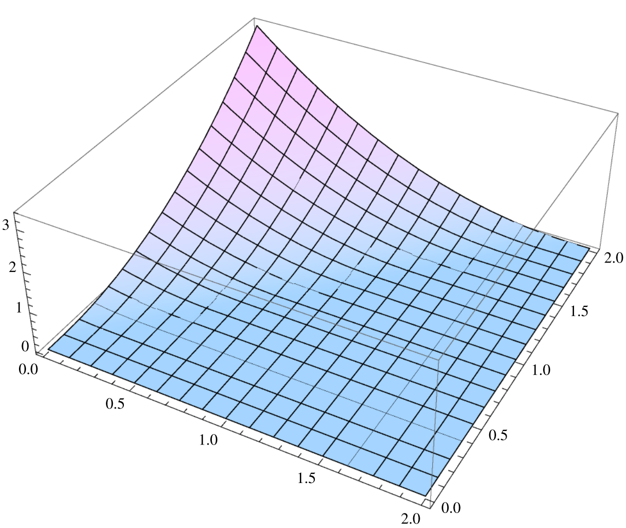

FIG. 2. *The bivariate observation density* g *on* $[0, 2]^2$, *where* $\varepsilon = 0.1$.

suitable smoothness conditions, we can write, if $b_n \asymp n^{-1/5}$,

$$\tilde F_n^{(\mathrm{SML})}(t) - \int K_{b_n}(t-u)\,dF_0(u) = \int \theta^{\mathrm{CS}}_{t,b_n,F}(u,\delta)\,d(\mathbb{Q}_n - Q_0)(u,\delta) + o_p\bigl(n^{-2/5}\bigr), \tag{1.10}$$

where

$$\theta^{\mathrm{CS}}_{t,b,F}(u,\delta) = -\frac{\delta\phi^{\mathrm{CS}}_{t,b,F}(u)}{F(u)} + \frac{(1-\delta)\phi^{\mathrm{CS}}_{t,b,F}(u)}{1-F(u)}, \qquad u \in (0,1) \tag{1.11}$$

and $\phi^{\mathrm{CS}}_{t,b,F}$ is given by

$$\phi^{\mathrm{CS}}_{t,b,F}(u) = \frac{F(u)\{1-F(u)\}}{g(u)} b^{-1} K\bigl((t-u)/b\bigr).$$

Moreover, g is the density of the (one-dimensional) observation distribution.

The solution $\phi^{\mathrm{CS}}_{t,b_n,F_0}$ gives as an approximation for $n\,\mathrm{var}(\tilde F_n(t))$:

$$\begin{aligned}\mathbb{E}\theta^{\mathrm{CS}}_{t,b_n,F_0}(T_1,\Delta_1)^2 &= \int \frac{\phi^{\mathrm{CS}}_{t,b_n,F_0}(u)^2}{F_0(u)} g(u)\,du + \int \frac{\phi^{\mathrm{CS}}_{t,b_n,F_0}(u)^2}{1-F_0(u)} g(u)\,du \\ &\sim \frac{F_0(t)\{1-F_0(t)\}}{b_n g(t)} \int K(u)^2\,du, \qquad b_n \to 0.\end{aligned}$$

Taking the bias into account, we get, if $b_n \asymp n^{-1/5}$, for the SMLE the central limit theorem

$$\sqrt{n}\left\{\tilde F_n^{\mathrm{CS}}(t) - F_0(t) - \frac{1}{2} b_n^2 f_0'(t) \int u^2 K(u)\,du\right\} \Big/ \sigma_n \xrightarrow{\mathcal{D}} N(0,1), \qquad n \to \infty, \tag{1.12}$$

where

$$\sigma_n^2 = \mathbb{E}\theta^{\mathrm{CS}}_{t,b_n,F_0}(T_1,\Delta_1)^2 \sim \frac{F_0(t)\{1-F_0(t)\}}{b_n g(t)} \int K(u)^2\,du, \qquad n \to \infty;$$

see Theorem 4.2, page 365 [10].

On the other hand, for the MSLE in the current status model it is first shown that the MSLE corresponds to the slope of greatest convex minorant of the *continuous* cusum diagram

$$\begin{aligned}&\left(\int \mathbb{K}_b(t-x)\,d\mathbb{G}_n(x), \int \delta\mathbb{K}_b(t-x)\,d\mathbb{P}_n(x,\delta)\right), \\ &\mathbb{K}_b(y) = \int_{-\infty}^{y/b} K(u)\,du, \qquad t \ge 0,\end{aligned} \tag{1.13}$$

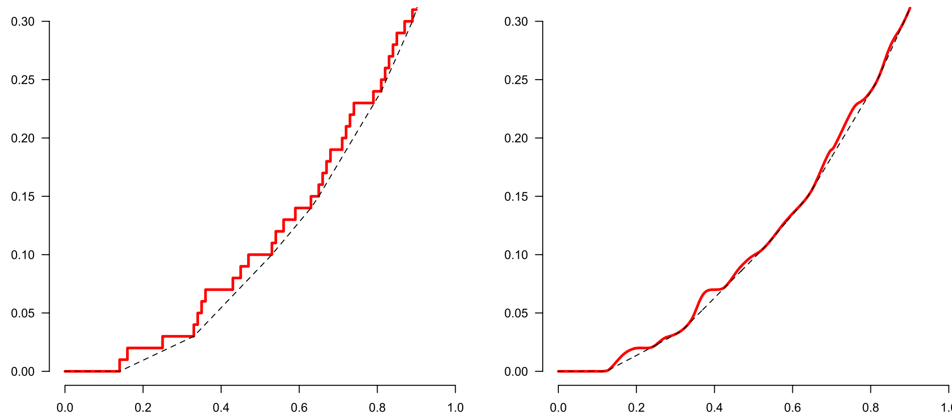

FIG. 3. *Unsmoothed and smoothed cusum diagram.*

where $\mathbb{G}_n$ is the empirical distribution function of the T_i and $\mathbb{P}_n$ the empirical distribution function of the pairs (T_i, Δ_i), analogously to the way the MLE corresponds to the slope of greatest convex minorant of the cusum diagram

$$\left(\int_{[0,t]} d\mathbb{G}_n(u), \int_{[0,t]} \delta\, d\mathbb{P}_n(u,\delta)\right), \qquad t \ge 0.$$

A picture of the cusum diagram for the MLE and the SMLE for the same sample is shown in Figure 3.

Next, it is shown that the MSLE is at interior points asymptotically equivalent to the ratio of kernel estimators

$$\frac{g^{\delta}_{n,b_n}(t)}{g_{n,b_n}(t)}, \tag{1.14}$$

where

$$g_{n,b_n}(t) = \int K_{b_n}(t-u)\, d\mathbb{G}_n(u), \qquad g^{\delta}_{n,b_n}(t) = \int \delta K_{b_n}(t-u)\, d\mathbb{P}_n(u,\delta).$$

This leads to the following central limit theorem for the MSLE, if $b_n \asymp n^{-1/5}$:

$$\begin{aligned} &\sqrt{n}\left\{\tilde F_n^{\mathrm{MSLE}}(t) - F_0(t) - \frac{1}{2}b_n^2\left\{f_0'(t) + \frac{2f_0(t)g'(t)}{g(t)}\right\}\int u^2K(u)\,du\right\}\Big/\sigma_n \\ &\quad \xrightarrow{\mathcal{D}} N(0,1), \end{aligned} \tag{1.15}$$

as $n \to \infty$, where σ_n is defined as in (1.12). Note that (1.12) and (1.15) only differ in the bias term $b_n^2 f_0(t)g'(t)/g(t)$.

1.2. *The SMLE and MSLE for the interval censoring*, *case* 2, *model*. For interval censoring, case 2, we cannot rely on explicit representations, as in the current status model. For the SMLE, we only have a representation of type (1.10) via the solution ϕ^{IC} of an integral equation, and we have to follow arguments analogous to the arguments in [3, 6] and [4].

In the separated case (specified by Condition 1.1 below), the integral equation (in $\phi = \phi^{\mathrm{IC}}$) is given by

$$\phi(u) = d_F(u)\biggl\{b^{-1}K\bigl((t-u)/b\bigr) + \int_{v>u} \frac{\phi(v)-\phi(u)}{F(v)-F(u)} g(u,v)\,dv - \int_{v<u} \frac{\phi(u)-\phi(v)}{F(u)-F(v)} g(v,u)\,dv\biggr\}, \tag{1.16}$$

where we take either $F = \hat F_n$ or $F = F_0$, and where

$$d_F(u) = \frac{F(u)\{1-F(u)\}}{g_1(u)\{1-F(u)\} + g_2(u)F(u)}. \tag{1.17}$$

Moreover, let the function $\theta^{\mathrm{IC}}_{t,b,F}$ be defined by

$$\theta^{\mathrm{IC}}_{t,b,F}(u,v,\delta_1,\delta_2) = -\frac{\delta_1\phi^{\mathrm{IC}}_{t,b,F}(u)}{F(u)} - \frac{\delta_2\{\phi^{\mathrm{IC}}_{t,b,F}(v)-\phi^{\mathrm{IC}}_{t,b,F}(u)\}}{F(v)-F(u)} + \frac{\delta_3\phi^{\mathrm{IC}}_{t,b,F}(v)}{1-F(v)}, \tag{1.18}$$

where $u < v$. Then, as in [4], we have the representation

$$\begin{aligned}
&\int \mathbb{K}\bigl((t-u)/b\bigr)\,d(\hat F_n - F_0)(u)\\
&\quad= \int \theta^{\mathrm{IC}}_{t,b,\hat F_n}(u,v,\delta_1,\delta_2)\,dP_0(u,v,\delta_1 d_2)\\
&\quad= \int \frac{\phi^{\mathrm{IC}}_{t,b,\hat F_n}(u)}{\hat F_n(u)} F_0(u)g_1(u)\,du - \int \frac{\phi^{\mathrm{IC}}_{t,b,\hat F_n}(v)}{1-\hat F_n(v)}\bigl\{1-F_0(v)\bigr\}g_2(v)\,dv\\
&\qquad+ \int \frac{\phi^{\mathrm{IC}}_{t,b,\hat F_n}(v)-\phi^{\mathrm{IC}}_{t,b,\hat F_n}(u)}{\hat F_n(v)-\hat F_n(u)}\bigl\{F_0(v)-F_0(u)\bigr\}g(u,v)\,du\,dv.
\end{aligned}$$

Using the theory in [4] again, we get that $\phi^{\mathrm{IC}}_{t,b,F_0}$ gives as an approximation for $n\,\mathrm{var}(\tilde F_n(t))$:

$$\begin{aligned}
&E\theta^{\mathrm{IC}}_{t,b,F_0}(T_1,U_1,\Delta_{11},\Delta_{12})^2\\
&\quad= \int \frac{\phi^{\mathrm{IC}}_{t,b,F_0}(u)^2}{F_0(u)} g_1(u)\,du + \int \frac{\{\phi^{\mathrm{IC}}_{t,b,F_0}(v)-\phi^{\mathrm{IC}}_{t,b,F_0}(u)\}^2}{F_0(v)-F_0(u)} h(u,v)\,du\,dv\\
&\qquad+ \int \frac{\phi^{\mathrm{IC}}_{t,b,F_0}(v)^2}{1-F_0(v)} g_2(v)\,dv.
\end{aligned}$$

Taking $b_n \asymp n^{-1/5}$ and defining

$$\sigma_n^2 = E\theta_{t,b_n,F_0}^{\mathrm{IC}}(T_1, U_1, \Delta_{11}, \Delta_{12})^2,$$

we get

$$\begin{aligned}\lim_{b_n \downarrow 0} b_n \sigma_n^2 = d_{F_0}(t)\biggl\{1 + d_{F_0}(t)\int_{v>t} \frac{g(t,v)}{F_0(v)-F_0(t)}\,dv \\ + d_{F_0}(t)\int_{v<t} \frac{g(v,t)}{F_0(t)-F_0(v)}\,dv\biggr\}^{-1} \int K(u)^2\,du,\end{aligned} \tag{1.19}$$

where d_{F_0} is defined by (1.17). This means, as we shall show below, that the limit variance for the SMLE is again (as in the current status model) equal to the limit variance of the MSLE. This leads to Conjecture 11.15 in [9]:

$$\begin{aligned}\sqrt{n}\biggl\{\tilde F_n^{\mathrm{SML}}(t) - F_0(t) - \frac{1}{2}b_n^2 f_0'(t)\int u^2 K(u)\,du\biggr\}\Big/\sigma_n \xrightarrow{\mathcal{D}} N(0,1), \\ n \to \infty,\end{aligned} \tag{1.20}$$

under the conditions given in [9]. This also means that the asymptotic bias is of the same form as for the SMLE in the current status model, which is much simpler than the bias of the MSLE.

Throughout this paper, we will assume that the following conditions are satisfied, which were also assumed in [3] and [4].

CONDITION 1.1. (S1) *g_1 and g_2 are continuous, with $g_1(x) + g_2(x) > 0$ for all $x \in [0, M]$.*

(S2) *$\mathbb{P}\{V - U < \varepsilon\} = 0$ for some ε with $0 < \varepsilon \le 1/2M$, so g does not have mass close to the diagonal; this is called the separated case.*

(S3) *$(u, v) \mapsto g(u, v)$ is continuous on $\{(x, y) : 0 \le x < y < M\}$ and is zero outside this set. Moreover, $g(u, v) = 0$ if $v - u < \varepsilon$.*

(S4) *F is a continuous distribution function with support $[0, M]$; F satisfies*

$$F(u) - F(t) \ge c > 0, \qquad \text{if } u - t \ge \varepsilon.$$

(S5) *The partial derivatives $\partial_1 g(t, u)$ and $\partial_2 g(t, u)$ exist, except for at most a countable number of points, where left and right derivatives exist. The derivatives are bounded, uniformly over t and u.*

(S6) *If both G_1 and G_2 put zero mass on some set A, then F has zero mass on A as well, so $F \ll H_1 + H_2$. This means that F does not have mass on sets in which no observations can occur.*

Note that (S1) implies that d_F, defined by (1.17), is bounded. Conditions (S2) and (S4) are needed to avoid singularity in the integral equation: if $F(x) - F(t)$

becomes very small, we have $g(t, x) = 0$. A picture of an observation density, satisfying the above conditions, is shown in Figure 2; g is defined by

$$g(x, y) = 6(y - x - \varepsilon)^2/\{(2 - x - \varepsilon)(2 - \varepsilon)\}^2, \qquad x + \varepsilon < y, \tag{1.21}$$

on the triangle with vertices $(0, \varepsilon)$, $(0, 2)$ and $(2 - \varepsilon, 2)$, where $\varepsilon = 0.1$.

We use the following conditions for the kernel estimators.

CONDITION 1.2 (Conditions on the kernel estimators). *We assume that $\tilde{h}_{nj}$ and $\tilde{h}_n$ are kernel estimators of h_{0j} and h_0, respectively, defined by* (1.2) *and* (1.3), *respectively, for a symmetric continuously differentiable kernel K of type* (1.4), *with compact support. Moreover, for points near the boundary, boundary kernels are used, with coefficients $\alpha(t)$ and $\beta(t)$, defined by* (1.5) *and* (1.6), *respectively, where the functions α, β, and its derivatives α' and β' are bounded on* $[0, 1]$. *We assume*:

$$0 = \inf\left\{t \in [0, M] : \tilde{h}_{n1}(t) \vee \int_{u=0}^{t} \tilde{h}_n(u, t)\, du > 0\right\} \tag{1.22}$$

and

$$M = \sup\left\{t \in [0, M] : \tilde{h}_{n2}(t) \vee \int_{u=t}^{M} \tilde{h}_n(t, u)\, du > 0\right\}. \tag{1.23}$$

An example of a kernel estimate, satisfying the conditions of Condition 1.2, is given by kernel estimates which use the triweight kernel, defined by (1.4). For this kernel, the weight functions α and β, used in constructing the boundary kernel, are decreasing on [0, 1], and the derivatives are bounded on [0, 1]. Using Condition 1.1, we give a characterization in terms of necessary and sufficient (duality) conditions for the MSLE in Section 2. In that section, we also prove consistency of the MSLE, using techniques, similar to the method, used in [12], Part II, Section 4.

In Section 3, we discuss algorithms for computing the MSLE: the EM algorithm and an iterative convex minorant algorithm. The iterative convex minorant algorithm is an adapted version of the algorithm, introduced in [5] and (again in) [12]. It turns out that the latter algorithm performs best in our experiments. The EM algorithm is very slow and, therefore, not suitable for larger sample sizes or simulation purposes.

In Section 4, we will prove asymptotic normality of the MSLE at a fixed interior point of the domain of definition (Theorem 4.1). In this paper, we concentrate on the "separated case," where $U_i - T_i \geq \varepsilon$ for some $\varepsilon > 0$, as in [3] and [4]. This case seems to be the most important case, and also to be the usual situation in medical statistics. The nonseparated case is rather different and has its own specific difficulties. The behavior of the MLE and SMLE in this situation is discussed in [2, 6] and [11], but the theory is still rather incomplete, even for the MLE. There is a conjecture for its asymptotic distribution, put forward in [5] and [12], but this conjecture has not been proved up till now, although a simulation study, supporting the conjecture is given in [11]. The theory for the MSLE in this situation has still not been developed.

2. Characterization of the MSLE and consistency. Let, for an estimate $\tilde h_n$ of h_0, satisfying

$$\tilde h_n(t,u)=0,\qquad u-t<\varepsilon, \tag{2.1}$$

for some $\varepsilon>0$, the nabla function ∇_F be defined by

$$\begin{aligned}\nabla_F(u)&=\frac{\tilde h_{n1}(u)}{F(u)}-\frac{\tilde h_{n2}(u)}{1-F(u)}+\int_{v=0}^{u}\frac{\tilde h_n(v,u)}{F(u)-F(v)}\,dv\\&\quad-\int_{v=u}^{M}\frac{\tilde h_n(u,v)}{F(v)-F(u)}\,dv,\end{aligned} \tag{2.2}$$

if $0<F(u)<1$. If $F(u)=0$ or $F(u)=1$, we define $\nabla_F(u)=0$.

Then, similarly to the ordinary MLE, the MSLE can be characterized by the so-called Fenchel duality conditions.

LEMMA 2.1. *Let $\tilde h_n$ satisfy* (2.1), *for some $\varepsilon>0$. Then the distribution function $\hat F_n$ maximizes* (1.7) *if and only if $\hat F_n$ is continuous on $[0,M]$ and satisfies the conditions*

$$\int_{v=t}^{M}\nabla_{\hat F_n}(v)\,dv\le 0,\qquad t\in[0,M) \tag{2.3}$$

and

$$\int_0^M\nabla_{\hat F_n}(v)\hat F_n(v)\,dv=0. \tag{2.4}$$

Moreover, if $t\in[0,M)$ is a point of increase of $\hat F_n$, that is,

$$\begin{cases}\hat F_n(u)-\hat F_n(u')>0,\\ \qquad \text{for all } u,u'\in[0,M] \text{ such that } u'<t<u, & \text{if } t>0,\\ \hat F_n(u)>0,\\ \qquad u\in(0,M], & \text{if } t=0,\end{cases} \tag{2.5}$$

we have

$$\nabla_{\hat F_n}(t)=0\quad\text{and}\quad\int_t^M\nabla_{\hat F_n}(v)\,dv=0. \tag{2.6}$$

The proof of this lemma is given in the Appendix.

Note that if $\nabla_F(t)=0$ for all $t\in(0,M)$, where ∇_F is defined by (2.2), the conditions of Lemma 2.1 are satisfied for F, and hence F would be the MSLE if it also would be a distribution function. But unfortunately, the function F satisfying $\nabla_F(t)=0$ for all $t\in(0,M)$ need not be monotone. We will call a function $\tilde F_n$, satisfying $\nabla_{\tilde F_n}(t)=0$, $t\in(0,M)$, a *plug-in estimator* or *naive estimator* (as in [10]). This plug-in estimator will be further studied in Section 4 in the proof of the local asymptotic normality of the MSLE, where it will be shown that the MSLE is indeed locally asymptotically equivalent to this plug-in estimator.

COROLLARY 2.1. *Let $\tilde{h}_n$ satisfy* (2.1), *for some $\varepsilon > 0$. Then the distribution function $\hat{F}_n$ maximizes* (1.7) *if and only if $\hat{F}_n$ is continuous on $[0, M)$, $\hat{F}_n(M) > 0$, and if $\hat{F}_n$ satisfies the conditions*

$$\int_0^t \nabla_{\hat{F}_n}(v)\, dv \geq 0, \qquad t \in (0, M) \tag{2.7}$$

and

$$\int_0^M \nabla_{\hat{F}_n}(v)\, dv = 0. \tag{2.8}$$

Moreover, if $t \in [0, M)$ is a point of increase of $\hat{F}_n$, that is, satisfies condition (2.5) *of Lemma* 2.1, *then*

$$\nabla_{\hat{F}_n}(t) = 0 \quad \textit{and} \quad \int_0^t \nabla_{\hat{F}_n}(v)\, dv = 0. \tag{2.9}$$

PROOF. Suppose $\hat{F}_n$ maximizes $\ell(F)$. Defining

$$F_\delta(t) = \{1 - (1+\delta)(1 - \hat{F}_n(t))\} \vee 0, \qquad t \in [0, M],$$

we find:

$$\lim_{\delta \to 0} \frac{\ell(F_\delta) - \ell(\hat{F}_n)}{\delta} = -\int_0^M \nabla_{\hat{F}_n}(u)\{1 - \hat{F}_n(u)\}\, du = 0. \tag{2.10}$$

So if $\hat{F}_n$ maximizes $\ell(F)$, (2.8) follows from (2.10) and (2.4) of Lemma 2.1.

$$\int_0^M \nabla_{\hat{F}_n}(u)\, du = 0.$$

This implies

$$\int_0^t \nabla_{\hat{F}_n}(v)\, dv = -\int_t^M \nabla_{\hat{F}_n}(v)\, dv,$$

and condition (2.9) now also follows.

Conversely, if the conditions of the corollary hold, we get

$$\begin{aligned} &\int_0^M \hat{F}_n(u) \nabla_{\hat{F}_n}(u)\, du \\ &\quad = \hat{F}_n(M) \int_0^M \nabla_{\hat{F}_n}(v)\, dv + \int_{t=0}^M \int_{v=0}^u \nabla_{\hat{F}_n}(v)\, dv\, d\hat{F}_n(u) \\ &\quad = \hat{F}_n(M) \int_0^M \nabla_{\hat{F}_n}(v)\, dv = 0, \end{aligned}$$

implying condition (2.4) of Lemma 2.1. The other conditions of Lemma 2.1 follow similarly. □

We now simplify the conditions somewhat, in view of the iterative convex minorant algorithm, to be discussed in Section 3. Multiplying ∇_F by $F(1-F)$ yields the function

$$\overline{\nabla}_F(u) = \tilde{h}_{n1}(u)\{1 - F(u)\} - \tilde{h}_{n2}(u)F(u) + F(u)\{1 - F(u)\} \times \left\{\int_{v=0}^{u} \frac{\tilde{h}_n(v,u)}{F(u) - F(v)}\,dv - \int_{v=u}^{M} \frac{\tilde{h}_n(u,v)}{F(v) - F(u)}\,dv\right\}. \tag{2.11}$$

COROLLARY 2.2. *Let $\tilde{h}_n$ satisfy* (2.1), *for some $\varepsilon > 0$ and let the function $\overline{\nabla}_F$ be defined by* (2.11). *Then the distribution function $\hat{F}_n$ maximizes* (1.7) *if and only if $\hat{F}_n(M) > 0$, and $\hat{F}_n$ is continuous on $[0, M)$ and satisfies the conditions*

$$\int_0^t \overline{\nabla}_{\hat{F}_n}(v)\,dv \geq 0, \qquad t \in [0, M] \tag{2.12}$$

and

$$\int_0^M \nabla_{\hat{F}_n}(v)\,dv = 0. \tag{2.13}$$

Moreover, if $t \in [0, M)$ is a point of increase of $\hat{F}_n$, that is, satisfies condition (2.5) *of Lemma* 2.1, *then*

$$\overline{\nabla}_{\hat{F}_n}(t) = 0 \quad \textit{and} \quad \int_0^t \overline{\nabla}_{\hat{F}_n}(v)\,dv = 0. \tag{2.14}$$

PROOF. We have, for $t \in (0, M)$,

$$\int_0^t \overline{\nabla}_{\hat{F}_n}(v)\,dv = \int_0^t \hat{F}_n(v)\{1 - \hat{F}_n(v)\}\nabla_{\hat{F}_n}(v)\,dv.$$

Furthermore,

$$\begin{aligned}\int_a^t \hat{F}_n(v)&\{1 - \hat{F}_n(v)\}\nabla_{\hat{F}_n}(v)\,dv \\ &= \hat{F}_n(t)\{1 - \hat{F}_n(t)\}\int_{u=0}^t \nabla_{\hat{F}_n}(u)\,du \\ &\quad - \int_0^t \{1 - 2\hat{F}_n(u)\}\int_{v=0}^u \nabla_{\hat{F}_n}(v)\,dv\,d\hat{F}_n(u) \\ &= \hat{F}_n(t)\{1 - \hat{F}_n(t)\}\int_{u=0}^t \nabla_{\hat{F}_n}(u)\,du.\end{aligned}$$

Hence, condition (2.12) is equivalent to condition (2.7) of Corollary 2.1. Relation (2.14) follows similarly, and (2.13) is the same as (2.8). □

The preceding results imply the consistency of the MSLE. The proof, which is given in the Appendix, is somewhat analogous to the proof of the consistency of the MLE in [12].

THEOREM 2.1 (Consistency of the MSLE). *Let Condition 1.1 be satisfied on* $[0, M]$ *for the distribution function* F_0 *and the observation density* g. *Moreover, let* $\tilde{h}_{nj}$ *and* $\tilde{h}_n$ *be kernel estimators of* h_{0j} *and* h_0, *respectively, of the type defined in Condition 1.2. Finally, let* $\hat{F}_n$ *be the MSLE of* F_0. *Then, with probability one,*

$$\lim_{n\to\infty} \hat{F}_n(t) = F_0(t),$$

for each $t \in [0, M)$. *The convergence is uniform on each subinterval* $[a, b]$ *of* $(0, M)$.

The proof of this theorem is given in the Appendix.

3. Algorithms. We explained in Section 1 that the MSLE can be computed for current status data via a continuous cusum diagram. In the present case we do not have a similar algorithm, which computes the MSLE in one step. The EM algorithm is based on the following "self-consistency equations"

$$\hat{f}_n(t) = \left\{ \int_t^M \frac{\tilde{h}_{n1}(v)}{\hat{F}_n(v)}\, dv + \int_0^t \frac{\tilde{h}_{n2}(v)}{1 - \hat{F}_n(v)}\, dv + \int_{v<t<u} \frac{\tilde{h}_n(v,u)}{\hat{F}_n(u) - \hat{F}_n(v)}\, dv\, du \right\} \hat{f}_n(t),$$

where $\hat{f}_n(t) = \hat{F}_n'(t)$. This yields the iteration steps

$$f^{(k+1)}(t) = \left\{ \int_t^M \frac{\tilde{h}_{n1}(v)}{F^{(k)}(v)}\, dv + \int_0^t \frac{\tilde{h}_{n2}(v)}{1 - F^{(k)}(v)}\, dv + \int_{v<t<u} \frac{\tilde{h}_n(v,u)}{F^{(k)}(u) - F^{(k)}(v)}\, dv\, du \right\} f^{(k)}(t). \tag{3.1}$$

One can indeed use a discretized version of (3.1) to compute the MSLE, but the EM algorithm is (as is usual for this type of problem with many parameters) very slow. Simply enhancing the EM algorithm by a Newton step is also not helpful because of the many constraints the solution has to satisfy, leading to very small "feasible steps." For this reason, a Newton-improved EM algorithm does not improve very much on the EM algorithm itself.

In our experience, the fastest algorithm is a combination of the EM algorithm with a version of the iterative convex minorant (ICM) algorithm, introduced in [5] and [12]. We use a sequence of cusum diagrams

$$\left(W_n^{(k)}(t), V_n^{(k)}(t)\right), \qquad t \in [0, M], k = 0, 1, 2, \ldots, \tag{3.2}$$

for which we compute the greatest convex minorants at each kth step. We alternate this with an EM-step (the combination is sometimes called the "hybrid algorithm"). The cumulative weight function $W_n^{(k)}$ is of the form

$$W_n^{(k)} = \int_0^t w_n^{(k)}(u)\,du, \qquad t \ge 0,$$

for suitably (but somewhat arbitrarily) chosen weights $w_n^{(k)}$, and the cusum function $V_n^{(k)}$ is of the form:

$$V_n^{(k)}(t) = \int_0^t F^{(k)}(u) w_n^{(k)}(u)\,du + \int_0^t \overline{\nabla}_{F^{(k)}}(u)\,du, \qquad t \ge 0,$$

where, for a distribution function F, $\overline{\nabla}_F$ is the function, defined by (2.11), evaluated at $F = F^{(k)}$. The idea is that the iterations force the Fenchel duality conditions (2.12) and (2.13) to be satisfied at the end of the iterations.

The following weight function, chosen by taking the diagonal elements of the Hessian matrix, corresponding to the function $\overline{\nabla}_F$, gave good convergence results in our simulation study of the MSLE:

$$\begin{aligned} w_n^{(k)}(t) = {} & \tilde{h}_{n1}(t) + \tilde{h}_{n2}(t) \\ & - \{1 - 2F_n^{(k)}(t)\} \\ & \times \left\{ \int_{u=0}^t \frac{\tilde{h}_n(u,t)}{F^{(k)}(t) - F^{(k)}(u)}\,du - \int_{u=t}^M \frac{\tilde{h}_n(t,u)}{F^{(k)}(u) - F^{(k)}(t)}\,du \right\} \\ & + F^{(k)}(t)\{1 - F^{(k)}(t)\} \int_{u=0}^t \frac{\tilde{h}_n(u,t)}{\{F^{(k)}(t) - F^{(k)}(u)\}^2}\,du \\ & + \int_{u=t}^M \frac{\tilde{h}_n(t,u)}{\{F^{(k)}(u) - F^{(k)}(t)\}^2}\,du. \end{aligned}$$

To prevent divergence of the algorithm, Armijo's line search method, as implemented in [13], was used for determining the step size at each iteration. The integrals were computed by a discrete approximation, using Riemann sums.

Note that, in the case of current status data, the function $\overline{\nabla}_F$ is just given by

$$\overline{\nabla}_F = \tilde{h}_n(t) - \tilde{g}_n(t) F(t), \tag{3.3}$$

from which we can compute $\hat{F}_n$ in one step.

4. Asymptotic distribution.

4.1. *Main result and road map.* We will prove the following theorem.

THEOREM 4.1. *Let condition* (1.1) *be satisfied. Moreover, let F_0 be twice differentiable, with a bounded continuous derivative f_0 on the interior of* $[0, M]$, *which is bounded away from zero on* $[0, M]$, *with a finite positive right limit at* 0 *and a positive left limit at M. Also, let f_0 have a bounded continuous* *derivative on* $(0, M)$ *and let* g_1 *and* g_2 *be twice differentiable on the interior of their supports* S_1 *and* S_2, *respectively. Furthermore, let the joint density* g *of the pair of observation times* (T_i, U_i) *have a bounded (total) second derivative on* $\{(x, y): 0 < x < y < M\}$. *Suppose that* X_i *is independent of* (T_i, U_i), *and let* d_{F_0} *be defined by*

$$d_{F_0}(v) = \frac{F_0(v)\{1 - F_0(v)\}}{g_1(v)\{1 - F_0(v)\} + F_0(v)g_2(v)}.$$

Then, if $b_n \asymp n^{-1/5}$, *we have for each* $v \in (0, M)$,

$$\sqrt{nb_n}\{\hat{F}_n(v) - F_0(v) - \beta(v)b_n^2\} \xrightarrow{\mathcal{D}} N(0, \sigma(v)^2),$$

where $N(0, \sigma(v)^2)$ *is a normal distribution with first moment zero and variance* $\sigma(v)^2$, *and where, defining*

$$\sigma_1(v) = 1 + d_{F_0}(v)\left\{\int_{t<v} \frac{g(t, v)}{F_0(v) - F_0(t)}\, dt + \int_{w>v} \frac{g(v, w)}{F_0(w) - F_0(v)}\, dw\right\}, \tag{4.1}$$

the variance $\sigma(v)^2$ *is given by*

$$\sigma(v)^2 = \frac{d_{F_0}(v)}{\sigma_1(v)} \int K(u)^2\, du. \tag{4.2}$$

Defining

$$\begin{aligned} \beta_1(v) = \frac{1}{2\sigma_1(v)}\Bigg\{&\frac{\{1 - F_0(v)\}h_1''(v) - F_0(v)h_2''(v)}{g_1(v)\{1 - F_0(v)\} + F_0(v)g_2(v)} \\ &+ d_{F_0}(v)\left\{\int_{t=0}^{v} \frac{(\partial^2/\partial v^2)h_0(t, v)}{F_0(v) - F_0(t)}\, dt \right. \\ &\left. - \int_{u=v}^{M} \frac{(\partial^2/\partial v^2)h_0(v, u)}{F_0(u) - F_0(v)}\, du\right\}\Bigg\} \int u^2 K(u)\, du, \end{aligned} \tag{4.3}$$

the bias $\beta(v)$ *is given by*

$$\beta(v) = \beta_1(v) + \frac{d_{F_0}(v)}{\sigma_1(v)}\left\{\int_{u=0}^{v} \frac{g(u, v)\beta_1(u)}{F_0(v) - F_0(u)}\, du + \int_{u=v}^{M} \frac{g(v, u)\beta_1(u)}{F_0(u) - F_0(v)}\, du\right\}.$$

REMARK 4.1. The asymptotic bias of the MSLE is of a very complicated form, certainly compared to the asymptotic bias of the SMLE, which is just

$$\frac{1}{2} f_0'(t)b_n^2 \int u^2 K(u)\, du;$$

see (1.20). It would be nice if some simplification could be found. Note however, that also in the current status model the asymptotic bias of the MSLE is more complicated than the asymptotic bias of the SMLE, since the derivatives of the estimates of the observation density come into play.

We now first give a "road map" of the proof of Theorem 4.1. Our starting point is given by the duality conditions (2.12) and (2.13). It is clear that if we would have equality in (2.12) instead of inequality, we would get the following relation by differentiating w.r.t. t:

$$
\begin{aligned}
\overline{\nabla}_F(t) &= \tilde h_{n1}(t)\{1-F(t)\} - \tilde h_{n2}(t)F(t)\\
&\quad + F(t)\{1-F(t)\}\left\{\int_{v=0}^{t}\frac{\tilde h_n(v,t)}{F(t)-F(v)}\,dv - \int_{v=t}^{M}\frac{\tilde h_n(t,v)}{F(v)-F(t)}\,dv\right\}\\
&= 0.
\end{aligned}
\tag{4.4}
$$

Conversely, if F solves (4.4) for each $t \in (0, M)$ and F is a distribution function such that $F(t) \in (0, 1)$, for each $t \in (0, M)$, F also satisfies (2.12) and (2.13) and is therefore the MSLE.

The solution of equation (4.4) takes the role of the plugin estimator (1.14) in the current status model. In the proof of the central limit theorem for the MSLE for the current status model, it was shown that the solution (in F) of (4.4) is a distribution function on a subinterval (a, b) of $[0, M]$ for large n with high probability, where we can take $a > 0$ arbitrarily close to 0 and $b < M$ arbitrarily close to M. In the present case, we prove the stronger fact that the solution of (4.4) is a (sub)distribution function on $[0, M]$ itself. This implies that the MSLE $\hat F_n$ coincides with the solution of (4.4) on the interval $(0, M)$ for large n with high probability.

To show that the solution (in F) of (4.4) is with high probability a (sub)distribution function on $[0, M]$ for large n, we first show that the solution is close to the solution of the *linear* integral equation

$$
\begin{aligned}
&F(t)-F_0(t)+d_{F_0}(t)\left\{\int_{u=0}^{t}\frac{g(u,t)\{F(t)-F_0(t)-F(u)+F_0(u)\}}{F_0(t)-F_0(u)}\,du\right.\\
&\qquad\qquad\left. - \int_{u=t}^{M}\frac{g(t,u)\{F(u)-F_0(u)-F(t)+F_0(t)\}}{F_0(u)-F_0(t)}\,du\right\}\\
&\quad = \frac{\tilde h_{n1}(t)\{1-F_0(t)\}-\tilde h_{n2}(t)F_0(t)}{\{1-F_0(t)\}g_1(t)+F_0(t)g_2(t)}\\
&\qquad + d_{F_0}(t)\left\{\int_{u<t}\frac{\tilde h_n(u,t)}{F_0(t)-F_0(u)}\,du - \int_{u>v}\frac{\tilde h_n(t,u)}{F_0(u)-F_0(t)}\,du\right\},
\end{aligned}
\tag{4.5}
$$

where d_{F_0} is defined by (1.17), with $F = F_0$.

We next show that the "toy estimator," solving the equation

$$
\begin{aligned}
&\{F(t)-F_0(t)\}\left\{1+d_{F_0}(t)\left\{\int_{u=0}^{t}\frac{g(u,t)}{F_0(t)-F_0(u)}\,du+\int_{u=t}^{M}\frac{g(t,u)}{F_0(u)-F_0(t)}\,du\right\}\right\}\\
&\quad = \frac{\tilde h_{n1}(t)\{1-F_0(t)\}-\tilde h_{n2}(t)F_0(t)}{\{1-F_0(t)\}g_1(t)+F_0(t)g_2(t)}\\
&\qquad + d_{F_0}(t)\left\{\int_{u<t}\frac{\tilde h_n(u,t)}{F_0(t)-F_0(u)}\,du - \int_{u>v}\frac{\tilde h_n(t,u)}{F_0(u)-F_0(t)}\,du\right\},
\end{aligned}
\tag{4.6}
$$

where the off-diagonal terms

$$
\begin{aligned}
-d_{F_0}(t)\biggl\{&\int_{u=0}^{t}\frac{g(u,t)\{F(u)-F_0(u)\}}{F_0(t)-F_0(u)}\,du \\
&+\int_{u=t}^{M}\frac{g(t,u)\{F(u)-F_0(u)\}}{F_0(u)-F_0(t)}\,du\biggr\}
\end{aligned}
\tag{4.7}
$$

on the left-hand side of (4.5) are omitted, also solves (4.5) to the right order, apart from a deterministic shift term. This last step is somewhat similar to a part of the proof of the asymptotic distribution of the MLE for interval censoring under the separation condition in [6]. However, in the latter case a corresponding "off-diagonal" term (4.7) plays no role asymptotically, since for the MLE the contribution to the bias is of lower order. In this way, we have reduced the proof to the asymptotic equivalence of the MSLE with the solution of (4.6) on $[0, M]$.

A comparison of the MSLE and the toy estimator, solving (4.6), is shown in part (a) of Figure 4 for bandwidth $b_n = n^{-1/5}$. One can see that for this bandwidth isotonization is still needed (the MSLE has derivative zero on a piece in the middle of the interval). Note that the toy estimator is not monotone, but has a very small distance to the MSLE. If we take $b_n = 2n^{-1/5}$, as in part (b) of Figure 4, which seems a better choice in this case, isotonization is not needed, except at the very last end of the interval (where the MSLE has derivative zero). Note that, as $n \to \infty$, the bandwidth will become smaller than $\varepsilon/2$, where ε is the separation distance in (1.21), but that this is still not the case in Figures 4.

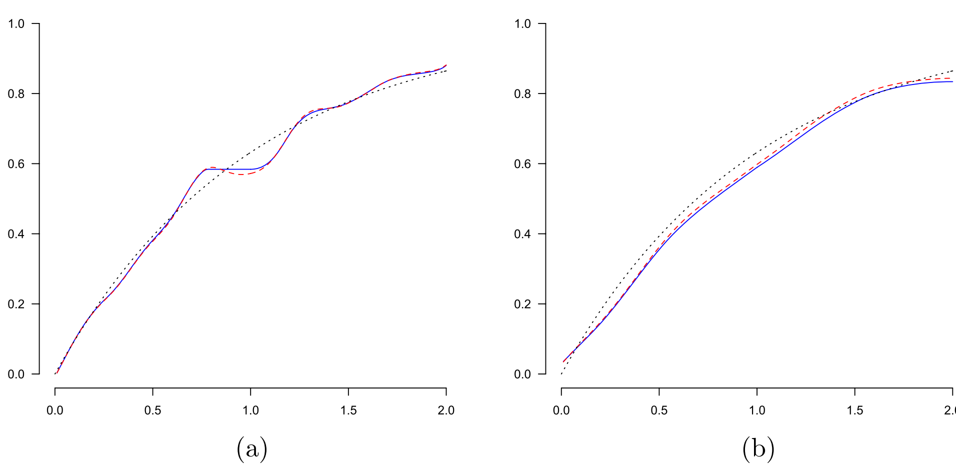

FIG. 4. (a) *The MSLE* (*solid*) *and the toy estimator, solving equation* (4.6) (*dashed*), *for a sample of size* $n = 1000$ *from the exponential distribution function* $F_0(x) = 1 - \exp\{-x\}$ (*dotted*), *using bandwidth* $b_n = n^{-1/5} \approx 0.25119$. *The bivariate observation density* g *is defined by* (1.21), *where* $\varepsilon = 0.1$. (b) *The same, but now using the bandwidth* $b_n = 2n^{-1/5} \approx 0.36411$.

Replacing $\tilde{h}_{nj}$ by h_{0j} in (4.4), $j = 1, 2$, and $\tilde{h}_n$ by h_0, we obtain the equation

$$h_{01}(t)\{1 - F(t)\} - h_{02}(t)F(t) + F(t)\{1 - F(t)\}\left\{\int_{v=0}^{t} \frac{h_0(v,t)}{F(t) - F(v)}\,dv - \int_{u=t}^{M} \frac{h_0(t,u)}{F(u) - F(t)}\,du\right\} = 0,$$

which, using the definition of h_{0j} and h_0, turns into

$$\begin{aligned} &g_1(t)F_0(t)\{1 - F(t)\} - g_2(t)\{1 - F_0(t)\}F(t) \\ &\quad + F(t)\{1 - F(t)\} \\ &\quad \times \left\{\int_{v=0}^{t} \frac{g(v,t)\{F_0(t) - F_0(u)\}}{F(t) - F(v)}\,dv - \int_{u=t}^{M} \frac{g(t,u)\{F_0(u) - F_0(t)\}}{F(u) - F(t)}\,du\right\} \\ &\quad = 0. \end{aligned}$$

This equation is clearly solved by F_0 itself.

This motivates us to consider the equation

$$\phi(t; h_1, h_2, h, F) = 0, \qquad t \in [0, M],$$

where

$$\begin{aligned} &\phi(t; h_1, h_2, h, F) \\ &\quad = h_1(t)\{1 - F(t)\} - h_2(t)F(t) \\ &\qquad + F(t)\{1 - F(t)\}\left\{\int_{v=0}^{t} \frac{h(v,t)}{F(t) - F(v)}\,dv - \int_{u=t}^{M} \frac{h(t,u)}{F(u) - F(t)}\,du\right\}. \end{aligned} \tag{4.8}$$

The functions h belong to a closed subset of the Banach space $C(S)$, where S is given by

$$S = \{(x, y) : 0 \le x \le x + \varepsilon_0 \le y \le M\},$$

for a fixed $\varepsilon_0 > 0$. We further define

$$S_1 = [0, M - \varepsilon_0], \qquad S_2 = [\varepsilon_0, M].$$

We now define the space E by

$$E = C[S_1] \times C[S_2] \times C(S) \times C[0, M], \tag{4.9}$$

and put the following norm on E:

$$\|(h_1, h_2, h, F)\| = \max\{\|h_1\|, \|h_2\|, \|h\|, \|F\|\}, \tag{4.10}$$

where the norms on the right-hand side denote the supremum norm, which we also denote by $\|\cdot\|$. Note that E is a Banach space for the norm (4.10).

We will also need another norm on $C[S]$, defined by

$$\|h\|_S = \sup_{t \in [0,M]} \left\{\int_{u:(u,t)\in S} |h(u,t)|\,du + \int_{u:(t,u)\in S} |h(t,u)|\,du\right\}. \tag{4.11}$$

Note that this is indeed a norm on $C[S]$, since $\|h\|_S = 0$ implies $h = 0$ and since the triangle inequality and homogeneity property for scalars are obviously satisfied.

The proof of Theorem 4.1 now proceeds via a sequence of lemmas. The proofs of these lemmas are given in the Appendix. The first lemma tells us that we can apply the implicit function theorem in Banach spaces to ensure that, locally, using the norms just introduced, there is a unique solution F to the equation $\phi(t; h_1, h_2, h, F) = 0$.

LEMMA 4.1. *Let F_0, h_{01}, h_{02} and h_0 satisfy the conditions of Theorem 4.1. Furthermore, let the function ϕ be defined by (4.8). Then there exists for all small $\eta > 0$ an open set U in the Banach space $C[S_1] \times C[S_2] \times C(S)$, endowed with the norm*

$$\|(h_1, h_2, h)\| = \max\{\|h_1\|, \|h_2\|, \|h\|\},$$

such that, if $(h_1, h_2, h) \in U$, the equation

$$\phi(t; h_1, h_2, h, F) = 0, \qquad t \in [0, M],$$

where ϕ is defined by (4.8), has a unique solution F in the open ball $B(F_0, \eta) \subset C[0, M]$ with midpoint F_0.

Having established the existence of a solution, we also consider the derivative of the solution.

LEMMA 4.2. *Let, under the conditions of Lemma 4.1, for a small $\eta > 0$, $F \in B(F_0, \eta)$ be the solution of*

$$\phi(t; h_1, h_2, h, F) = 0, \qquad t \in [0, M],$$

where ϕ is defined by (4.8), and where h_j has a bounded continuous derivative on the interior of S_j, having finite limits approaching the boundary of S_j, for $j = 1, 2$. Similarly, we suppose that h is differentiable on the interior of its support S and has finite limits approaching the boundary of S. Then, if $(h_1, h_2, h) \in U_\delta$, where U_δ is defined by (A.8), the solution F has a continuous and bounded derivative for sufficiently small η and δ.

The following lemma will be used to show that, with probability tending to one, $\tilde{F}_n$ belongs to the allowed class, for all large n, and is a consistent estimate of F_0.

LEMMA 4.3. *Let, under the conditions of Lemma 4.1, $F^{(n)} \in B(F_0, \eta)$ be the solution of*

$$\phi(t; h_1^{(n)}, h_2^{(n)}, h^{(n)}, F) = 0, \qquad t \in [0, M],$$

where ϕ is defined by (4.8), and where $h^{(n)} \in C[S]$, $h_1^{(n)} \in C[S_1]$ an $h_2^{(n)} \in C[S_2]$ are nonnegative functions which have bounded continuous derivatives on the supports S, S_1 and S_2, respectively, with finite limits approaching the boundary, respectively. Furthermore, let

$$\|h_j^{(n)} - h_{0j}\| \longrightarrow 0 \quad \textit{and} \quad \|(h_j^{(n)})' - h_{0j}'\| \longrightarrow 0, \qquad j = 1, 2,$$

where, as before, $\|\cdot\|$ *denotes the supremum norm on* $C[S_j]$. *Finally, let*

$$\|h^{(n)} - h_0\|_S \longrightarrow 0 \quad \text{and} \quad \|\partial_j h^{(n)} - \partial_j h_0\|_S \longrightarrow 0, \qquad j = 1, 2, \tag{4.12}$$

where $\|\cdot\|_S$ *is defined by* (4.11). *Then* $F^{(n)} \to F_0$ *in the supremum metric, as* $n \to \infty$, *and* $F^{(n)}$ *is strictly increasing on* $[0, M]$ *and satisfies* $F^{(n)}(t) \in [0, 1]$, $t \in [0, M]$, *for all large* n.

We still need to show that the estimates $\tilde h_{nj}$ of h_{0j} and $\tilde h_n$ of h_0 have the properties of h_j and h, as defined in Lemma 4.3.

LEMMA 4.4. *Let the conditions of Theorem* 4.1 *be satisfied and let the estimates* $\tilde h_{nj}$ *of and* $\tilde h_n$ *satisfy Condition* 1.2. *Then*

$$\|\tilde h_{nj} - h_{0j}\| \xrightarrow{p} 0 \quad \text{and} \quad \|\tilde h'_{nj} - h'_{0j}\| \xrightarrow{p} 0, \qquad j = 1, 2.$$

Moreover,

$$\|\tilde h_n - h_0\|_S \xrightarrow{p} 0 \quad \text{and} \quad \|\partial_j \tilde h_n - \partial_j h_0\|_S \xrightarrow{p} 0, \qquad j = 1, 2. \tag{4.13}$$

We now get the following result.

LEMMA 4.5. *Let the conditions of Theorem* 4.1 *be satisfied and let, for small* $\eta > 0$, $F = \tilde F_n \in B(F_0, \eta)$ *be the solution of the equation*

$$\phi(t; \tilde h_{n1}, \tilde h_{n2}, \tilde h_n, F) = 0, \qquad t \in [0, M],$$

where ϕ *is defined by* (4.8). *Moreover, let* $\|\cdot\|$ *denote the supremum norm. Then:*

(i) *With probability tending to one,* $\tilde F_n$ *is strictly increasing on* $[0, M]$, *and satisfies* $\tilde F_n(t) \in [0, 1]$, $t \in [0, M]$, *for all large* n. *Hence, with probability tending to one,* $\tilde F_n$ *coincides with the MSLE for large* n *and*

$$\|\tilde F_n - F_0\| \xrightarrow{p} 0, \qquad n \to \infty.$$

(ii)

$$\|\tilde F_n - F_0\| = O_p\big(n^{-2/5}\sqrt{\log n}\big), \qquad n \to \infty.$$

(iii)

$$\|\tilde F_n - \bar F_n\| = O_p\big(n^{-4/5}\log n\big),$$

where $\bar F_n$ *is the solution in* F *of the linear integral equation* (4.5).

We did now in principle solve our problem, since we have shown that $\tilde F_n$ is locally asymptotically equivalent to the solution $\bar F_n$ of a *linear* integral equation. Since $\tilde F_n$ coincides with the MSLE for large n, the MSLE is also locally asymptotically equivalent with $\bar F_n$. However, to get an explicit expression for the bias and variance of the MSLE, we now study a still simpler "toy estimator," which turns out also to be locally asymptotically equivalent to the MSLE.

LEMMA 4.6. *Let the toy estimator $F = F_n^{\text{toy}}$ be defined as the solution of the equation*

$$\begin{aligned}
&\{F(t) - F_0(t)\} \\
&\quad \times \left\{1 + d_{F_0}(t)\left\{\int_{u<t} \frac{g(u,t)}{F_0(t) - F_0(u)}\,dt + \int_{u>v} \frac{g(t,u)}{F_0(u) - F_0(t)}\,du\right\}\right\} \\
&\quad = \frac{\tilde{h}_{n1}(t)\{1 - F_0(t)\} - \tilde{h}_{n2}(t)F_0(t)}{\{1 - F_0(t)\}g_1(t) + F_0(t)g_2(t)} \\
&\qquad + d_{F_0}(t)\left\{\int_{u<t} \frac{\tilde{h}_n(u,t)}{F_0(t) - F_0(u)}\,du - \int_{u>v} \frac{\tilde{h}_n(t,u)}{F_0(u) - F_0(t)}\,du\right\}.
\end{aligned} \tag{4.14}$$

Then, under the conditions of Theorem 4.1,

$$\sqrt{nb_n}\left\{F_n^{\text{toy}}(v) - F_0(v) - \frac{\beta_1(v)b_n^2}{2\sigma_1(v)}\right\} \xrightarrow{\mathcal{D}} N\big(0, \sigma(v)^2\big),$$

where $b_n \asymp n^{-1/5}$ and $\beta_1(v)$, $\sigma_1(v)$ and $\sigma(v)$ are defined as in Theorem 4.1.

REMARK 4.2. In Lemma 4.6, a toy estimator is introduced, which plays a similar role as the toy estimator in the study of the ordinary MLE for interval censoring, introduced in [5] and [12] (the term “toy estimator” was coined by Jon Wellner). It is called a toy estimator because we cannot use it in an actual sample, since F_0 is unknown (and is in fact the object we want to estimate). Actually, the solution $\bar{F}_n$ of the linear integral equation (4.5) in part (iii) of Lemma 4.5 is also a toy estimator in this sense (but does not produce explicit expressions for the expectation and variance of the asymptotic distribution).

Lemma 4.7 shows that the solution of the linear integral equation (4.5) is equivalent in first order to the toy estimator of Lemma 4.6, apart from a deterministic bias term.

LEMMA 4.7. *Let, under the conditions of Theorem* 4.1, F_n^{toy} *solve equation* (4.14) *of Lemma* 4.6 *and let $\bar{F}_n$ be the solution of the linear integral equation* (4.5). *Then*

$$\begin{aligned}
\bar{F}_n(t) &= F_n^{\text{toy}}(t) + d_{F_0}(t)\left\{\int_{u=0}^{t} \frac{\gamma_n(u)g(u,t)}{F_0(t) - F_0(u)}\,du + \int_{u=t}^{M} \frac{\gamma_n(u)g(t,u)}{F_0(u) - F_0(t)}\,du\right\} \\
&\qquad + O_p\big(n^{-1/2}\big),
\end{aligned}$$

where

$$\gamma_n(u) = \frac{\beta_1(u)b_n^2}{\sigma_1(u)}$$

and $\beta_1(u)$ and $\sigma_1(u)$ are defined as in Theorem 4.1.

Theorem 4.1 now follows from Lemma 4.7 and the asymptotic equivalence of $\bar{F}_n$ with the MSLE.

5. Concluding remarks and open problems. In the preceding, it was shown that, under the so-called separation hypothesis, the MSLE locally converges to the underlying distribution function at rate $n^{-2/5}$, if we use bandwidths $b_n \asymp n^{-1/5}$ in the estimates $\tilde{h}_{nj}$ and $\tilde{h}_n$. The asymptotic (normal) distribution was also determined. The results can be used to construct a two-sample likelihood ratio test, of the same type as the test, discussed in [7] for the current status model, but this is not done in the present paper. It is also possible to use the results to construct pointwise bootstrap confidence intervals, as is done in [8] and [9] for the current status model. In that case, it might be advisable to use *undersmoothing*, and work with bandwidths of order $n^{-\alpha}$, where $1/3 < \alpha < 1/5$, as is done in [8] and [9]. In this way one gets rid of the bias and it is expected that the SMLE and MSLE will then be very similar, since their asymptotic variances are the same, which implies that their asymptotic (normal) limits will also be the same.

If the separation hypothesis does not hold, which means that we can have arbitrarily small observation intervals, the asymptotic behavior of the MSLE is still unknown. In this situation the local asymptotic limit distribution for the ordinary MLE is also still unknown, although it is conjectured that the rate $n^{-1/3}$, holding under the separation hypothesis, is improved to the rate $(n \log n)^{-1/3}$ in this case. There even exists a conjectured limit distribution in this case, put forward in [5] (see also [9] and [12]). Supporting evidence for this conjecture is given in a simulation study in [11], but a proof is still missing. The latter paper also gives simulation results for the SMLE, and the asymptotic variance of the SMLE is the same as that of the MSLE, but the asymptotic bias is different, just as in the current status model. The bias of the SMLE is considerably simpler than the bias of the MSLE. The asymptotic behavior of the SMLE again has to be deduced from an associated integral equation; this is further discussed in [9] and [11].

It is possible to extend the theory to the situation that there are more than two observation times than just T_i and U_i or to the so-called mixed case (see, e.g., [14]), where there are a random number of observation times per unobservable event X_i. However, since this leads to further complications in defining the integral equations, we did not do this in the present paper.

APPENDIX

PROOF OF LEMMA 2.1. First suppose that the conditions (2.3) and (2.4) are satisfied. Then we cannot have $\hat{F}_n(t) = 0$ for t in an interval where $\tilde{h}_{n1}(t) > 0$ or $\int_{u=0}^{t} \tilde{h}_n(u,t) > 0$, since otherwise $\ell(\hat{F}_n) = -\infty$. Similarly, we cannot have $\hat{F}_n(t) = 1$ for t in an interval where $\tilde{h}_{n2}(t) > 0$ or $\int_{u=t}^{M} \tilde{h}_n(t,u) > 0$.

Since the criterion function $F \mapsto \ell(F)$ is concave in F, we get

$$\ell(F) - \ell(\hat{F}_n) \le \int_0^M \nabla_{\hat{F}_n}(u)\{F(u) - \hat{F}_n(u)\}\,du, \tag{A.1}$$

where we use the facts that the integrals defining $\ell(F)$ are all nonpositive. Note that this is similar to the relation (1.11) in [12].

By (2.4),

$$\int_0^M \nabla_{\hat{F}_n}(u)\hat{F}_n(u)\,du = 0,$$

and hence

$$\int_0^M \nabla_{\hat{F}_n}(u)\{F(u) - \hat{F}_n(u)\}\,du = \int_0^M \nabla_{\hat{F}_n}(u)F(u)\,du.$$

If $F = 1_{[t,\infty)}$ for some $t \in [0, M)$, we get by (2.3),

$$\int_0^M \nabla_{\hat{F}_n}(u)F(u)\,du = \int_t^M \nabla_{\hat{F}_n}(u)\,du \le 0.$$

So we also get, for subdistribution functions of the type

$$F = \sum_{i=1}^k \alpha_i 1_{[t_i,\infty)}, \qquad 0 \le t_1 < \cdots, t_k \le M, \alpha_i \in (0,1), \sum_{i=1}^k \alpha_i \le 1,$$

that

$$\int_0^M \nabla_{\hat{F}_n}(u)F(u)\,du = \sum_{i=1}^k \alpha_i \int_{t_i}^M \nabla_{\hat{F}_n}(u)\,du \le 0.$$

Since we can approximate any subdistribution function F on $[0, M)$ by subdistribution functions of this type, this implies $\ell(F) \le \ell(\hat{F}_n)$, for all subdistribution functions F.

Conversely, suppose that $\hat{F}_n$ maximizes $\ell(F)$. Then we must have, if $t \in (0, M)$, $F = 1_{[t,\infty)}$ and $\delta \in (0, 1)$,

$$\int_{v=t}^M \nabla_{\hat{F}_n}(v)\,dv = \lim_{\delta \downarrow 0} \delta^{-1}\{\ell((1-\delta)\hat{F}_n + \delta F) - \ell(\hat{F}_n)\} \le 0$$

(using the concavity of ℓ for the existence of the limit), and hence (2.3) has to be satisfied for $\hat{F}_n$. Moreover, defining F_δ by

$$F_\delta(t) = (1+\delta)\hat{F}_n(t) \wedge 1, \qquad t \in [0, M],$$

we find

$$\lim_{\delta \to 0} \frac{\ell(F_\delta) - \ell(\hat{F}_n)}{\delta} = 0,$$

since the limit has to be nonpositive, if we let δ tend to zero, either from above or from below.

We have

$$0 = \lim_{\delta \to 0} \frac{\ell(F_\delta) - \ell(\hat{F}_n)}{\delta} = \int_0^M \nabla_{\hat{F}_n}(u)\hat{F}_n(u)\,du,$$

so (2.4) must hold.

Suppose $\hat{F}_n$ has a jump at $t \in (0, M)$ and suppose $\nabla_{\hat{F}_n}(t-) > 0$. Define

$$F_\delta(u) = \begin{cases} \hat{F}_n(u), & u < t - \delta, \\ \hat{F}_n(t), & u \in [t - \delta, t), \\ \hat{F}_n(u), & u \in [t, M]. \end{cases}$$

Then

$$\int \nabla_{\hat{F}_n}(u)\{F_\delta(u) - \hat{F}_n(u)\}\,du > 0,$$

for small $\delta > 0$, a contradiction. Hence, we must have: $\nabla_{\hat{F}_n}(t-) \le 0$. If $\nabla_{\hat{F}_n}(t) < 0$, we define

$$F_\delta(u) = \begin{cases} \hat{F}_n(u), & u < t, \\ \hat{F}_n(t-), & u \in [t, t + \delta), \\ \hat{F}_n(u), & u \in [t + \delta, M], \end{cases}$$

and then again:

$$\int \nabla_{\hat{F}_n}(u)\{F_\delta(u) - \hat{F}_n(u)\}\,du > 0,$$

for small $\delta > 0$, a contradiction, so we must have: $\nabla_{\hat{F}_n}(t) \ge 0$, implying we must have

$$\nabla_{\hat{F}_n}(t-) \le 0 \le \nabla_{\hat{F}_n}(t). \tag{A.2}$$

On the other hand, we have by the continuity of $\tilde{h}_{nj}$, $j = 1, 2$ and $\tilde{h}_n$:

$$\begin{aligned}
&\nabla_{\hat{F}_n}(t) - \nabla_{\hat{F}_n}(t-) \\
&\quad = \frac{\tilde{h}_{n1}(t)}{\hat{F}_n(t)} - \frac{\tilde{h}_{n1}(t)}{\hat{F}_n(t-)} - \frac{\tilde{h}_{n2}(t)}{1 - \hat{F}_n(t)} + \frac{\tilde{h}_{n2}(t)}{1 - \hat{F}_n(t-)} \\
&\qquad + \int_{v=0}^{u} \frac{\tilde{h}_n(v, t)}{\hat{F}_n(t) - \hat{F}_n(v)}\,dv - \int_{v=0}^{u} \frac{\tilde{h}_n(v, t)}{\hat{F}_n(t-) - \hat{F}_n(v)}\,dv \\
&\qquad - \int_{v=t}^{M} \frac{\tilde{h}_n(t, v)}{\hat{F}_n(v) - \hat{F}_n(t)}\,dv + \int_{v=t}^{M} \frac{\tilde{h}_n(t, v)}{\hat{F}_n(v) - \hat{F}_n(t-)}\,dv \\
&\quad = \frac{\tilde{h}_{n1}(t)}{\hat{F}_n(t)} - \frac{\tilde{h}_{n1}(t)}{\hat{F}_n(t-)} - \frac{\tilde{h}_{n2}(t)}{1 - \hat{F}_n(t)} + \frac{\tilde{h}_{n2}(t)}{1 - \hat{F}_n(t-)} \\
&\quad = \tilde{h}_{n1}(t)\frac{\hat{F}_n(t-) - \hat{F}_n(t)}{\hat{F}_n(t-)\hat{F}_n(t)} + \tilde{h}_{n2}(t)\frac{\hat{F}_n(t-) - \hat{F}_n(t)}{\{1 - \hat{F}_n(t-)\}\{1 - \hat{F}_n(t)\}} < 0,
\end{aligned}$$

contradicting (A.2). The conclusion is that we must have: $\hat{F}_n(t-) = \hat{F}_n(t)$, for $t \in (0, M)$.

Finally, suppose (2.5) is satisfied for a point $t \in (0, M)$ and suppose $\nabla_{\hat{F}_n}(t) > 0$. Then, by the continuity of $\hat{F}_n$, there also exists a neighborhood of t such that $\nabla_{\hat{F}_n}(u) > 0$ for u in this neighborhood. We now define a perturbation F_δ of $\hat{F}_n$ by

$$F_\delta(u) = \begin{cases} \hat{F}_n(u), & u < t - \delta, \\ \hat{F}_n(t + \delta), & u \in [t - \delta, t + \delta), \\ \hat{F}_n(u), & u \in [t + \delta, M]. \end{cases}$$

Then we have for sufficiently small $\delta > 0$:

$$\int \nabla_{\hat{F}_n}(u)\{F_\delta(u) - \hat{F}_n(u)\}\,du > 0,$$

contradicting

$$\int \nabla_{\hat{F}_n}(u)\{F_\delta(u) - \hat{F}_n(u)\}\,du \le 0.$$

If (2.5) is satisfied for a point $t \in (0, M)$ and $\nabla_{\hat{F}_n}(t) < 0$, we define the perturbation F_δ of $\hat{F}_n$ by

$$F_\delta(u) = \begin{cases} \hat{F}_n(u), & u < t - \delta, \\ \hat{F}_n(t - \delta), & u \in [t - \delta, t + \delta), \\ \hat{F}_n(u), & u \in [t + \delta, M] \end{cases}$$

and get a contradiction in the same way. So, if (2.5) is satisfied for a point $t \in (0, M)$, we must have:

$$\nabla_{\hat{F}_n}(t) = 0.$$

This proves the left-hand side of (2.6).

Furthermore,

$$\begin{aligned} &\int_0^M \nabla_{\hat{F}_n}(u)\hat{F}_n(u)\,du \\ &\quad = \left[-\hat{F}_n(t)\int_{u=t}^M \nabla_{\hat{F}_n}(u)\,du\right]_{t=0}^M + \int_{t=0}^M \int_{u=t}^M \nabla_{\hat{F}_n}(u)\,du\,d\hat{F}_n(t) \\ &\quad = \hat{F}_n(0)\int_{u=0}^M \nabla_{\hat{F}_n}(u)\,du + \int_{t=0}^M \int_{u=t}^M \nabla_{\hat{F}_n}(u)\,du\,d\hat{F}_n(t) \\ &\quad = \int_{t=0}^M \int_{u=t}^M \nabla_{\hat{F}_n}(u)\,du\,d\hat{F}_n(t), \end{aligned} \tag{A.3}$$

implying by (2.3) that

$$\int_{u=t}^M \nabla_{\hat{F}_n}(u)\,du = 0,$$

for points t satisfying (2.5), since otherwise the right-hand side of (A.3) would be strictly negative. $\square$

PROOF OF THEOREM 2.1. By the assumption on the kernel estimates and the observation density g, we may assume that $\tilde{h}_n$ satisfies (2.1), for some $\varepsilon > 0$, and all large n. Let the function ψ be defined by

$$\begin{aligned}\psi(F; h_1, h_2, h) = \int h_1(t)\log F(t)\,dt + \int h_2(t)\log\{1 - F(t)\}\,dt \\ + \int h(t,u)\log\{F(u) - F(t)\}\,dt\,du.\end{aligned} \tag{A.4}$$

Then we must have, if $h_j = \tilde{h}_{nj}$, $j = 1, 2$ and $h = \tilde{h}_n$,

$$\lim_{\varepsilon\downarrow 0}\varepsilon^{-1}\{\psi\big((1-\varepsilon)\hat{F}_n + \varepsilon F_0; h_1, h_2, h\big) - \psi(\hat{F}_n; h_1, h_2, h)\} \le 0.$$

This implies (see also (4.20) in [12]):

$$\begin{aligned}\int \frac{F_0(t)}{\hat{F}_n(t)}\tilde{h}_{n1}(t)\,dt + \int \frac{1 - F_0(t)}{1 - \hat{F}_n(t)}\tilde{h}_{n2}(t)\,dt \\ + \int \frac{F_0(u) - F_0(t)}{\hat{F}_n(u) - \hat{F}_n(t)}\tilde{h}_n(t,u)\,dt\,du \le 1.\end{aligned} \tag{A.5}$$

Fix a small $\delta \in [0, M/2]$ and let the intervals A_δ and B_δ be defined by

$$A_\delta = [\delta, M], \qquad B_\delta = [0, M - \delta].$$

Then, arguing as in [12], Part II, Chapter 4, we find that there exists an $M > 0$ such that for all n,

$$\sup_{t\in A_\delta} 1 \Big/ \hat{F}_n(t) + \sup_{t\in B_\delta} 1 \Big/ \{1 - \hat{F}_n(t)\} \le M.$$

We also cannot have that $\hat{F}_n(u_{k_n}) - \hat{F}_n(t_{k_n}) \to 0$, for a sequence of points $(t_{k_n}, u_{k_n}) \in C_\delta$. For suppose, if necessary by taking a subsequence, that $t_{k_n} \to t_0$ and $u_{k_n} \to u_0$. The $u_0 - t_0 \ge \varepsilon + \delta$. By the vague convergence of $\hat{F}_n$ to F, there are continuity points t_1 and u_1 such that $t_0 < t_1 < u_1 < u_0$, $u_1 - t_1 \ge \frac{1}{2}\delta + \varepsilon$, and $\hat{F}_n(t_1) \to F(t_1)$ and $\hat{F}_n(u_1) \to F(u_1)$. Moreover, since

$$\hat{F}_n(u_1) - \hat{F}_n(t_1) \le \hat{F}_n(u_{k_n}) - \hat{F}_n(t_{k_n}),$$

for large n, we must have: $F(u_1) - F(t_1) = 0$. We then would get that there exists a rectangle $[t_1, t_2] \times [u_2, u_1]$ such that $u_2 - t_2 > \varepsilon$ and

$$\begin{aligned}&\liminf_{n\to\infty}\int_{[t_1,t_2]\times[u_2,u_1]} \frac{F_0(u) - F_0(t)}{\hat{F}_n(u) - \hat{F}_n(t)}\tilde{h}_n(t,u)\,dt\,du \\ &\qquad \ge K\{F_0(u_2) - F_0(t_2)\}\int_{[t_1,t_2]\times[u_2,u_1]} h_0(t,u)\,dt\,du,\end{aligned}$$

for any $K > 0$, contradicting (A.5). So, we may also assume that

$$\inf_{(t,u)\in C_\delta}\{\hat F_n(u) - \hat F_n(t)\} \ge \frac{1}{M},$$

for all n.

As in [12], Part II, Chapter 4, we have by the Helly compactness theorem that the exists a set of probability one, such that for each ω in this set the sequence $(\hat F_n(\cdot;\omega))$ has a subsequence $\hat F_{n_k}(\cdot;\omega)$ which converges vaguely to a subdistribution function $F = F(\cdot;\omega)$. By the vague convergence of $\hat F_{n_k}(\cdot;\omega)$ to F, we now get

$$\begin{aligned}
&\int_{A_\delta} \frac{F_0(t)}{\hat F_{n_k}(t;\omega)}\tilde h_{n1}(t)\,dt + \int_{B_\delta} \frac{1-F_0(t)}{1-\hat F_{n_k}(t;\omega)}\tilde h_{n2}(t)\,dt \\
&\quad + \int_{C_\delta} \frac{F_0(u)-F_0(t)}{\hat F_{n_k}(u;\omega)-\hat F_{n_k}(t;\omega)}\tilde h_n(t,u)\,dt\,du \\
&\quad\quad \to \int_{A_\delta} \frac{F_0(t)}{F(t)} h_{01}(t)\,dt + \int_{B_\delta}\frac{1-F_0(t)}{1-F(t)} h_{02}(t)\,dt \\
&\quad\quad\quad + \int_{C_\delta}\frac{F_0(u)-F_0(t)}{F(u)-F(t)} h_0(t,u)\,dt\,du, \qquad n\to\infty.
\end{aligned}$$

By monotone convergence, we now also have

$$\begin{aligned}
&\int_{[0,M]} \frac{F_0(t)}{F(t)} h_{01}(t)\,dt + \int_{[0,M]}\frac{1-F_0(t)}{1-F(t)} h_{02}(t)\,dt \\
&\quad + \int_{[0,M]^2}\frac{F_0(u)-F_0(t)}{F(u)-F(t)} h_0(t,u)\,dt\,du \\
&\quad = \lim_{\delta\downarrow 0}\biggl\{\int_{A_\delta} \frac{F_0(t)}{F(t)} h_{01}(t)\,dt + \int_{B_\delta}\frac{1-F_0(t)}{1-F(t)} h_{02}(t)\,dt \\
&\qquad\qquad + \int_{C_\delta}\frac{F_0(u)-F_0(t)}{F(u)-F(t)} h_0(t,u)\,dt\,du\biggr\} \le 1.
\end{aligned} \tag{A.6}$$

Suppose $F(t) \ne F_0(t)$ for some $t\in[0, M/2]$. Then there exist a $u\in(t,M)$ such that $h_0(t,u) > 0$ and

$$\frac{F_0(t)^2}{F(t)} + \frac{\{1-F_0(u)\}^2}{1-F(u)} + \frac{\{F_0(u)-F_0(t)\}^2}{F(u)-F(t)} > 1,$$

since

$$\frac{F_0(t)^2}{x} + \frac{\{1-F_0(u)\}^2}{1-y} + \frac{\{F_0(u)-F_0(t)\}^2}{y-x}\begin{cases} = 1, & F_0(t)=x,\ F_0(u)=y,\\ > 1, & \text{otherwise}\end{cases}$$

(see also (4.27) in [12]). By the continuity of F_0 and the monotonicity and right continuity of F there exist therefore also $h > 0$ such that

$$\frac{F_0(t')^2}{F(t')} + \frac{\{1 - F_0(u')\}^2}{1 - F(u')} + \frac{\{F_0(u') - F_0(t')\}^2}{F(u') - F(t')} > 1,$$

if $t' \in [t, t + h]$ and $u' \in [u, u + h]$. This implies

$$\begin{aligned}
&\int_{[0,M]} \frac{F_0(t)}{F(t)} h_{01}(t)\,dt + \int_{[0,M]} \frac{1 - F_0(t)}{1 - F(t)} h_{02}(t)\,dt \\
&\quad + \int_{[0,M]^2} \frac{F_0(u) - F_0(t)}{F(u) - F(t)} h_0(t, u)\,dt\,du \\
&\quad = \int_{[0,M]} \frac{F_0(t)^2}{F(t)} g_1(t)\,dt + \int_{[0,M]} \frac{\{1 - F_0(t)\}^2}{1 - F(t)} g_2(t)\,dt \\
&\qquad + \int_{[0,M]^2} \frac{\{F_0(u) - F_0(t)\}^2}{F(u) - F(t)} g(t, u)\,dt\,du \\
&\quad = \int_{[0,M]^2} \left\{ \frac{F_0(t)^2}{F(t)} + \frac{\{1 - F_0(t)\}^2}{1 - F(t)} + \frac{\{F_0(u) - F_0(t)\}^2}{F(u) - F(t)} \right\} g(t, u)\,dt\,du > 1,
\end{aligned}$$

in contradiction with (A.6). So, we must have $F(t) = F_0(t)$ if $t \in [0, M/2]$. A similar argument yields $F(t) = F_0(t)$ if $t \in [M/2, M)$.

So, for each ω outside a set of probability zero, the sequence $(\hat{F}_n(\cdot; \omega))$ has a subsequence which converges weakly to F_0. This implies that $\hat{F}_n(t)$ converges almost surely to $F_0(t)$ for each $t \in [0, M)$. The uniformity of the convergence on subintervals follows from the continuity of F_0. □

PROOF OF LEMMA 4.1. We will use the line of argument of the proof of the implicit function theorem 10.2.1 in [1]. We define the function $\bar{\phi}$ by

$$[\bar{\phi}(h_1, h_2, h, F)](t) = \phi(t, h_1, h_2, h, F), \qquad t \in [0, M], \tag{A.7}$$

so $\bar{\phi}$ maps E to $C[0, M]$. The derivative of $\bar{\phi}$ w.r.t. F, is given by the function

$$\begin{aligned}
&[[\partial_4 \bar{\phi}(h_1, h_2, h, F)](A)](t) \\
&\quad \stackrel{\text{def}}{=} -\{h_1(t) + h_2(t)\} A(t) \\
&\qquad + \{1 - 2F(t)\} \left\{ \int_{v=0}^{t} \frac{h(v, t)}{F(t) - F(v)}\,dv - \int_{u=t}^{M} \frac{h(t, u)}{F(u) - F(t)}\,du \right\} A(t) \\
&\qquad - F(t)\{1 - F(t)\} \\
&\qquad\quad \times \left\{ \int_{u=0}^{t} \frac{h(u, t)\{A(t) - A(u)\}}{\{F(t) - F(u)\}^2}\,du + \int_{u=t}^{M} \frac{h(t, u)\{A(t) - A(u)\}}{\{F(u) - F(t)\}^2}\,du \right\},
\end{aligned}$$

where $A \in C[0, M]$. Note that the right-hand side is well defined for $F \in B(F_0, \eta)$ and small $\eta > 0$, since $h(t, u) = 0$ if $u - t < \varepsilon$, and since F_0 has a nonvanishing derivative on $[0, M]$, implying that $F_0(u) - F_0(t)$ stays away from zero if $u - t \geq \varepsilon$.

We now define the open set $U = U_\delta$ of functions (h_1, h_2, h) by

$$
U_\delta = \{(h_1, h_2, h) \in C[S_1] \times C[S_2] \times C[S] : \max\{\|h_1 - h_{01}\|, \|h_2 - h_{02}\|, \|h - h_0\|_S\} < \delta\}. \tag{A.8}
$$

There exists a $\delta > 0$ such that for $(h_1, h_2, h) \in U_\delta$,

$$
\begin{aligned}
&\|\bar{\phi}(h_1, h_2, h, F_1) - \bar{\phi}(h_1, h_2, h, F_2) - [\partial_4\bar{\phi}(h_{01}, h_{02}, h_0, F_0)](F_1 - F_2)\| \\
&\quad \leq \varepsilon' \|F_1 - F_2\|,
\end{aligned}
$$

if $F_1, F_2 \in B(F_0, \eta)$, where $\varepsilon' > 0$ can be made arbitrarily small by making δ small, using the definition of differentiability in Banach spaces.

The equation

$$
[\partial_4\bar{\phi}(F_0; h_{01}, h_{02}, h_0, F_0)](A) = 0
$$

only has the trivial solution $A \equiv 0$ in $C[0, M]$. This is seen in the following way. Suppose there exists a solution in $A \in C[0, M]$ such that $[\partial_4\bar{\phi}(F_0; h_{01}, h_{02}, h_0, F_0)](A) = 0$ and $A(s) > 0$ for some $s \in [0, M]$. Then also $\max_{s \in [0, M]} A(s) > 0$. Suppose the maximum is attained at $t \in [0, M]$. Then

$$
\begin{aligned}
&[[\partial_4\bar{\phi}(h_{01}, h_{02}, h_0, F_0)](A)](t) \\
&\quad = -\{h_{01}(t) + h_{02}(t)\}A(t) \\
&\qquad + \{1 - 2F_0(t)\}\left\{\int_{v=0}^{t} \frac{h_0(v, t)}{F_0(t) - F_0(v)}\,dv - \int_{u=t}^{M} \frac{h_0(t, u)}{F_0(u) - F_0(t)}\,du\right\}A(t) \\
&\qquad - F_0(t)\{1 - F_0(t)\} \\
&\qquad\quad \times \left\{\int_{u=0}^{t} \frac{h_0(u, t)\{A(t) - A(u)\}}{\{F_0(t) - F_0(u)\}^2}\,du + \int_{u=t}^{M} \frac{h_0(t, u)\{A(t) - A(u)\}}{\{F_0(u) - F_0(t)\}^2}\,du\right\} \\
&\quad = -\{g_1(t)\{1 - F_0(t)\} + g_2(t)F_0(t)\}A(t) \\
&\qquad - F_0(t)\{1 - F_0(t)\} \\
&\qquad\quad \times \left\{\int_{u=0}^{t} \frac{h_0(u, t)\{A(t) - A(u)\}}{\{F_0(t) - F_0(u)\}^2}\,du + \int_{u=t}^{M} \frac{h_0(t, u)\{A(t) - A(u)\}}{\{F_0(u) - F_0(t)\}^2}\,du\right\} \\
&\quad \leq -\{g_1(t)\{1 - F_0(t)\} + g_2(t)F_0(t)\}A(t) < 0,
\end{aligned}
$$

using $g_1(t)\{1 - F_0(t)\} + g_2(t)F_0(t) > 0$, in contradiction with the assumption

$$
[\partial_4\bar{\phi}(h_{01}, h_{02}, h_0, F_0)](A) = 0.
$$

We similarly get a contradiction if we assume that $A(t) < 0$ for some $t \in [0, M]$ (similar arguments were used for the integral equation, studied in [3]). This shows that $\partial_4\bar{\phi}(h_{01}, h_{02}, h_0, F)$ is a linear homeomorphism of $C[0, M]$ onto $C[0, M]$ and that we can in fact use arguments of the type used in the proof of the implicit function theorem in Banach spaces, as given, for example, in [1], Theorem 10.2.1.

Denoting (as in the proof of Theorem 10.2.1 of [1]) the linear mapping $\partial_4\bar{\phi}(h_{01}, h_{02}, h_0, F_0)$ by T_0 and its inverse by T_0^{-1}, we find that

$$\begin{aligned}\|T_0^{-1}\cdot\{\bar{\phi}(h_1, h_2, h, F_1) - \bar{\phi}(h_1, h_2, h, F_2)\} - (F_1 - F_2)\| &\le \varepsilon'\|T_0^{-1}\|\,\|F_1 - F_2\|\\ &\le \tfrac{1}{2}\|F_1 - F_2\|,\end{aligned}$$

so we have a *contraction*, and this implies that the equation

$$F = F - T_0^{-1}\cdot\bar{\phi}(\cdot, F; h_1, h_2, h)$$

has a unique solution $F \in B(F_0, \eta)$ which can be obtained by successive approximations, if we take the balls around h_{0j} and h_0, to which h_j and h belong, respectively, sufficiently small, using a result like 10.1.1 in [1]. This, in turn, implies that the equation

$$\bar{\phi}(h_1, h_2, h, F) = 0$$

has a unique solution in $F \in B(F_0, \eta)$, for $(h_1, h_2, h) \in U_\delta$ and small δ. $\square$

PROOF OF LEMMA 4.2. If $\phi(\cdot; h_1, h_2, h, F) = 0$, we have

$$\begin{aligned}&\{1 - F(t)\}h_1(t) - F(t)h_2(t)\\ &\quad + F(t)\{1 - F(t)\}\left\{\int_{u:(u,t)\in S}\frac{h(u,t)}{F(t) - F(u)}\,du - \int_{u:(t,u)\in S}\frac{h(t,u)}{F(u) - F(t)}\,du\right\}\\ &\quad = 0.\end{aligned}$$

Note that the differentiability proerties of h, h_1 and h_2 and the fact that F solves the integral equation imply that we can differentiate F too. Differentiation w.r.t. t, and defining $f = F'$, yields:

$$\begin{aligned}&\{1 - F(t)\}h_1'(t) - F(t)h_2'(t) - f(t)\{h_1(t) + h_2(t)\}\\ &\quad + \{1 - 2F(t)\}\\ &\qquad \times f(t)\left\{\int_{u:(u,t)\in S}\frac{h(u,t)}{F(t) - F(u)}\,du - \int_{u:(t,u)\in S}\frac{h(t,u)}{F(u) - F(t)}\,du\right\}\\ &\quad + F(t)\{1 - F(t)\}\\ &\qquad \times \left\{\int_{u:(u,t)\in S}\frac{\partial_2 h(u,t)}{F(t) - F(u)}\,du - \int_{u:(t,u)\in S}\frac{\partial_1 h(t,u)}{F(u) - F(t)}\,du\right\}\\ &\quad - f(t)F(t)\{1 - F(t)\}\\ &\qquad \times \left\{\int_{u:(u,t)\in S}\frac{h(u,t)}{\{F(t) - F(u)\}^2}\,du + \int_{u:(t,u)\in S}\frac{h(t,u)}{\{F(u) - F(t)\}^2}\,du\right\}\\ &\quad = 0.\end{aligned}\tag{A.9}$$

Temporarily replacing F by F_0, and h_j and h by h_{0j} and h_0, respectively, we would obtain

$$\begin{aligned}
&\{1-F_0(t)\}h_{01}'(t)-F_0(t)h_{02}'(t)-f(t)\{\{1-F_0(t)\}g_1(t)+F_0(t)g_2(t)\}\\
&\quad+F_0(t)\{1-F_0(t)\}\\
&\qquad\times\left\{\int_{u:(u,t)\in S}\frac{\partial_2 h_0(u,t)}{F_0(t)-F_0(u)}\,du-\int_{u:(t,u)\in S}\frac{\partial_1 h_0(t,u)}{F_0(u)-F_0(t)}\,du\right\}\\
&\quad-f(t)F_0(t)\{1-F_0(t)\}\\
&\qquad\times\left\{\int_{u:(u,t)\in S}\frac{h_0(u,t)}{\{F_0(t)-F_0(u)\}^2}\,du+\int_{u:(t,u)\in S}\frac{h_0(t,u)}{\{F_0(u)-F_0(t)\}^2}\,du\right\}\\
&\quad=0,
\end{aligned}$$

that is,

$$\begin{aligned}
&f_0(t)\Bigg\{\{\{1-F_0(t)\}g_1(t)1_{S_1}(t)+F_0(t)g_2(t)1_{S_2}(t)\}\\
&\qquad+F_0(t)\{1-F_0(t)\}\left\{\int_{u:(u,t)\in S}\frac{h_0(u,t)}{\{F_0(t)-F_0(u)\}^2}\,du\right.\\
&\qquad\qquad\left.+\int_{u:(t,u)\in S}\frac{h_0(t,u)}{\{F_0(u)-F_0(t)\}^2}\,du\right\}\Bigg\}\\
&\quad=\{1-F_0(t)\}h_{01}'(t)1_{S_1}(t)-F_0(t)h_{02}'(t)1_{S_2}(t)\\
&\qquad+F_0(t)\{1-F_0(t)\}\\
&\qquad\times\left\{\int_{u:(u,t)\in S}\frac{\partial_2 h_0(u,t)}{F_0(t)-F_0(u)}\,du-\int_{u:(t,u)\in S}\frac{\partial_1 h_0(t,u)}{F_0(u)-F_0(t)}\,du\right\}.
\end{aligned}$$

This means that the coefficient of $f(t)$ in equation (A.9) stays away from zero if F belongs to a sufficiently small ball $B(F_0,\eta)$ around F_0, and $(h_1,h_2,h)\in U_\delta$ for small $\delta>0$. Denoting this coefficient by $c(t,h_1,h_2,h,F)$, we get the equation

$$\begin{aligned}
f(t)&=\frac{1}{c(t,h_1,h_2,h,F)}\\
&\quad\times\Bigg\{\{1-F(t)\}h_1'(t)-F(t)h_2'(t)\\
&\qquad+F(t)\{1-F(t)\}\\
&\qquad\times\left\{\int_{u=0}^{t}\frac{\partial_2 h(u,t)}{F(t)-F(u)}\,du-\int_{u=t}^{M}\frac{\partial_1 h(t,u)}{F(u)-F(t)}\,du\right\}\Bigg\}.
\end{aligned}\tag{A.10}$$

The statement of the lemma now follows. □

PROOF OF LEMMA 4.3. By Lemma 4.1, we have that $F^{(n)}$ tends to F_0 in the supremum norm on $C[0, M]$, since for any η we can choose a $\delta > 0$ such that

$$\|F^{(n)} - F_0\| < \eta,$$

if $\|h_j^{(n)} - h_{0j}\| < \delta$ and $\|h^{(n)} - h_0\| < \delta$.

Using Lemma 4.2, we get that, under the conditions of the lemma, that $F^{(n)}$ is differentiable with a bounded derivative $f^{(n)}$; see (A.10). Specifically, (A.10) yields

$$\begin{aligned} f^{(n)}(t) &\stackrel{\text{def}}{=} \big(F^{(n)}\big)'(t) \\ &= \Bigg\{F^{(n)}(t)\{1 - F^{(n)}(t)\} \\ &\qquad \times \left\{\int_{u=0}^{t} \frac{\partial_2 h^{(n)}(u,t)}{F^{(n)}(t) - F^{(n)}(u)}\,du - \int_{u=t}^{M} \frac{\partial_1 h^{(n)}(t,u)}{F^{(n)}(u) - F^{(n)}(t)}\,du\right\} \\ &\qquad\qquad + \{1 - F^{(n)}(t)\}\big(h_1^{(n)}\big)'(t) - F^{(n)}(t)\big(h_2^{(n)}\big)'(t)\Bigg\} \\ &\qquad \Big/ c\big(t, h_1^{(n)}, h_2^{(n)}, h^{(n)}, F^{(n)}\big), \end{aligned}$$

where $c(t, h_1^{(n)}, h_2^{(n)}, h^{(n)}, F^{(n)})$ stays away from zero, as $n \to \infty$. The corresponding density f_0 of the underlying model similarly has the representation

$$\begin{aligned} f_0(t) = &\Bigg\{F_0(t)\{1 - F_0(t)\} \\ &\times \left\{\int_{u\,:\,(u,t)\in S} \frac{\partial_2 h_0(u,t)}{F_0(t) - F_0(u)}\,du - \int_{u\,:\,(t,u)\in S} \frac{\partial_1 h_0(t,u)}{F_0(u) - F_0(t)}\,du\right\} \\ &\qquad + \{1 - F_0(t)\}h_{01}'(t) - F_0(t)h_{02}'(t)\Bigg\} \\ &\Big/ c_0(t), \end{aligned} \tag{A.11}$$

where $c_0(t)$ is given by

$$c_0(t) = g_1(t)\{1 - F_0(t)\} + g_2(t)F_0(t).$$

By

$$\|h_j^{(n)} - h_{0j}\| \to 0, \qquad \|h^{(n)} - h_0\| \to 0, \qquad \|F^{(n)} - F_0\| \to 0,$$

and (4.12), we now get

$$\sup_{t\in[0,M]} \big|c\big(t, h_1^{(n)}, h_2^{(n)}, h^{(n)}, F^{(n)}\big) - c_0(t)\big| \to 0.$$

Again using (4.12), we also get

$$\|f^{(n)} - f_0\| \to 0,$$

that is, $f^{(n)}$ converges to f_0 in the supremum norm. Since f_0 stays away from zero on $[0, M]$, this means that $F^{(n)}$ is strictly increasing on $[0, M]$ for all sufficiently large n.

Furthermore, since $\phi(t, h_1^{(n)}, h_2^{(n)}, h^{(n)}, F^{(n)}) = 0$, we get for large n, and t in a right neighborhood of 0,

$$\begin{aligned} &F^{(n)}(t) \\ &\quad = \frac{h_1^{(n)}(t)}{h_1^{(n)}(t) + h_2^{(n)}(t) + \{1 - F^{(n)}(t)\} \int_{u=t}^{M} (h^{(n)}(t,u)/(F^{(n)}(u) - F^{(n)}(t)))\,du} \\ &\quad \ge 0, \end{aligned}$$

since, by the convergence of $F^{(n)}$ to F_0, we may assume $1 - F^{(n)}(t) > 0$ for t in a neighborhood of 0, and since $h_j^{(n)}$ and $h^{(n)}$ are nonnegative.

Likewise, if $F_0(M) = 1$, we have, for t in a small left neighborhood of M,

$$\begin{aligned} &F^{(n)}(t)\{1 - F^{(n)}(t)\} \int_{v=0}^{t} \frac{h^{(n)}(v,t)}{F^{(n)}(t) - F^{(n)}(v)}\,dv \\ &\qquad = \{h_1^{(n)}(t) + h_2^{(n)}(t)\} F^{(n)}(t) - h_1^{(n)}(t) \\ &\qquad\quad + F^{(n)}(t)\{1 - F^{(n)}(t)\} \int_{v=t}^{M} \frac{h^{(n)}(v,t)}{F^{(n)}(t) - F^{(n)}(v)}\,dv \\ &\qquad = h_2^{(n)}(t) F^{(n)}(t) - h_1^{(n)}(t)\{1 - F^{(n)}(t)\}, \end{aligned}$$

and hence, for t in a small left neighborhood of M,

$$\{1 - F^{(n)}(t)\}\left\{F^{(n)}(t) \int_{v=0}^{t} \frac{h^{(n)}(v,t)}{F^{(n)}(t) - F^{(n)}(v)}\,dv + h_1^{(n)}(t)\right\} = h_2^{(n)}(t) F^{(n)}(t),$$

implying that, for all large n, $1 - F^{(n)}(t) \ge 0$ for t in a neighborhood of M. This will a fortiori hold if $F_0(M) < 1$. This shows that, for all large n and all $t \in [0, M]$, $F^{(n)}(t) \in [0, 1]$. $\square$

PROOF OF LEMMA 4.4. Since we use boundary kernels near the boundary of $[0, M]$, $\tilde{h}_{nj}(t)$ is a consistent estimate of $h_{0j}(t)$ for each $t \in S_j$. For if $t \in [b_n, M - b_n] \cap S_j$ we just have

$$\mathbb{E}\tilde{h}_{n1}(t) = \mathbb{E}\Delta_{11} K_{b_n}(t - T_1) = \int K_{b_n}(t - u) h_{01}(u)\,du = h_{01}(t) + O(n^{-2/5}),$$

where the remainder term is uniform in $t \in [b_n, M - b_n] \cap S_j$. Since $b_n \downarrow 0$, we have $b_n < \varepsilon$ for all large n, where ε is the "separation parameter" of Condition 1.1, and hence the boundary kernels are only relevant for $\tilde h_{n1}$ in a neighborhood of 0 and for $\tilde h_{n2}$ in a neighborhood of M.

If $t \in [0, b_n]$, we have

$$\begin{aligned}
\mathbb{E}\tilde h_{n1}(t) &= \alpha(t/b_n)E\Delta_{11}K_{b_n}(t-T_1) + \beta(t/b_n)\mathbb{E}\left\{\Delta_{11}\frac{t-T_1}{b_n}K_{b_n}(t-T_1)\right\}\\
&= \alpha(t/b_n)\int_{u=0}^{M} K_{b_n}(t-u)h_{01}(u)\,du\\
&\quad + \beta(t/b_n)\int_{u=0}^{M}\frac{t-u}{b_n}K_{b_n}(t-u)h_{01}(u)\,du\\
&= h_{01}(t)\int_{u=-1}^{t/b_n}\{\alpha(t/b_n)K(u)+\beta(t/b_n)K(u)\}\,du + O(n^{-2/5})\\
&= h_{01}(t) + O(n^{-2/5}),
\end{aligned}$$

again uniformly for $t \in [0, b_n]$. A similar computation can be made for $\tilde h_{n2}$ if $t \in [M-b_n, M]$. Since

$$\sup_{t\in S_1}\left|\tilde h_{n1}(t) - \mathbb{E}\tilde h_{n1}(t)\right| = O_p\left(n^{-2/5}\sqrt{\log n}\right),$$

we now get the uniform convergence in probability of $\tilde h_{n1}$ to h_{01} on S_1, and similarly we have uniform convergence in probability of $\tilde h_{n2}$ to h_{02} on S_2.

Next, we consider the derivative of $\tilde h_{n1}(t)$. If $t \in [b_n, M-b_n] \cap S_1$, we just have

$$\begin{aligned}
\mathbb{E}\tilde h'_{n1}(t) &= \frac{d}{dt}\mathbb{E}\Delta_{11}K'_{b_n}(t-T_1)\\
&= \int \frac{d}{dt}K_{b_n}(t-u)h_{01}(u)\,du\\
&= b_n^{-1}\int K'(u)h_{01}(t-b_n u)\,du\\
&= b_n^{-1}\int K'(u)\left\{h_{01}(t) - b_n u h'_{01}(t) + \frac{1}{2}b_n^2u^2h'_{01}(t) - \frac{1}{6}b_n^3u^3h''_{01}(t)\right\}du\\
&\quad + o(b_n^2)\\
&= h'_{01}(t) + O(b_n) = h'_{01}(t) + O(n^{-1/5}),
\end{aligned}$$

again uniformly in t. Since

$$\sup_{t\in[b_n, M-b_n]\cap S_1}\left|\tilde h'_{n1}(t) - \mathbb{E}\tilde h'_{n1}(t)\right| = O_p\left(n^{-1/5}\sqrt{\log n}\right),$$

we only have to consider what happens near the boundary.

In treating the boundary kernels, we denote for simplicity b_n by b. If $t \in [0, b]$, we have

$$\mathbb{E}\tilde{h}_{n1}(t) = \alpha\left(\frac{t}{b}\right)\int_{x=0}^{t+b} K_b(t-x)h_{01}(x)\,dx + \beta\left(\frac{t}{b}\right)\int_{x=0}^{t+b} \frac{t-x}{b} K_b(t-x)h_{01}(x)\,dx.$$

This can also be written

$$\mathbb{E}\tilde{h}_{n1}(t) = \int_{-1}^{t/b}\left\{\alpha\left(\frac{t}{b}\right)K(u) + \beta\left(\frac{t}{b}\right)uK(u)\right\}h_{01}(t-bu)\,du.$$

We write this in the form

$$\begin{aligned}\mathbb{E}\tilde{h}_{n1}(t) &= h_{01}(t)\\ &\quad + \int_{-1}^{t/b}\left\{\alpha\left(\frac{t}{b}\right)K(u) + \beta\left(\frac{t}{b}\right)uK(u)\right\}\int_{t-bu}^{t}(w-t+bu)h_{01}''(w)\,dw\,du,\end{aligned}$$

using a second-order Taylor development of h_{01} with the integral remainder term. Hence,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\tilde{h}_{n1}(t) &= h_{01}'(t) + \frac{1}{b}\left\{\alpha\left(\frac{t}{b}\right)K\left(\frac{t}{b}\right) + \beta\left(\frac{t}{b}\right)\frac{t}{b}K\left(\frac{t}{b}\right)\right\}\int_0^t wh_{01}''(w)\,dw\\ &\quad + \frac{1}{b}\int_{-1}^{t/b}\left\{\alpha'\left(\frac{t}{b}\right)K(u) + \beta'\left(\frac{t}{b}\right)uK(u)\right\}\\ &\quad \times \int_{t-bu}^{t}(w-t+bu)h_{01}''(w)\,dw\,du\\ &\quad - \int_{-1}^{t/b}\left\{\alpha\left(\frac{t}{b}\right)K(u) + \beta\left(\frac{t}{b}\right)uK(u)\right\}\int_{t-bu}^{t}h_{01}''(w)\,dw\,du\\ &\quad + \int_{-1}^{t/b}\left\{\alpha\left(\frac{t}{b}\right)K(u) + \beta\left(\frac{t}{b}\right)uK(u)\right\}buh_{01}''(t)\,du\\ &= h_{01}'(t) + O(b), \qquad b \downarrow 0.\end{aligned}$$

Note that, by Condition 1.2, the functions α, β, α' and β' are bounded on $[0, 1]$. We also have

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\tilde{h}_{n1}(t) &= \frac{1}{b}\left\{\alpha'\left(\frac{t}{b}\right)\int_{x=0}^{t+b} K_b(t-x)h_{01}(x)\,dx\right.\\ &\qquad \left. + \beta'\left(\frac{t}{b}\right)\int_{x=0}^{t+b}\frac{t-x}{b}K_b(t-x)h_{01}(x)\,dx\right\}\\ &\quad + \alpha\left(\frac{t}{b}\right)\int_{x=0}^{t+b}\frac{d}{dt}K_b(t-x)h_{01}(x)\,dx\end{aligned}$$

$$+\beta\left(\frac{t}{b}\right)\int_{x=0}^{t+b}\frac{t-x}{b}\frac{d}{dt}K_b'(t-x)h_{01}(x)\,dx$$
$$+\frac{1}{b}\beta\left(\frac{t}{b}\right)\int_{x=0}^{t+b}K_b(t-x)h_{01}(x)\,dx,$$

so

$$\mathbb{E}\tilde{h}_{n1}'(t)=\frac{d}{dt}\mathbb{E}\tilde{h}_{n1}(t)=h_{01}'(t)+O(b)=h_{01}'(t)+O\big(n^{-1/5}\big).$$

Since we have

$$\sup_{t\in[0,b_n]}\big|\tilde{h}_{n1}'(t)-\mathbb{E}\tilde{h}_{n1}'(t)\big|=O_p\big(n^{-1/5}\big),$$

we now also get that

$$\sup_{u\in S_1}\big|\tilde{h}_{n1}'(t)-h_{01}'(t)\big|=o_p(1).$$

The other cases can be treated in a similar way. □

PROOF OF LEMMA 4.5. Part (i) is an immediate consequence of Lemma 4.3.

(ii) We get, again using the approach of the implicit function theorem 10.2.1 in Banach spaces of [1], denoting the derivative w.r.t. (h_1, h_2, h) by D_1 and the derivative w.r.t. F by D_2:

$$\begin{aligned}\|\tilde{F}_n-F_0\| &= \big\|\big[D_2\bar{\phi}(h_{01},h_{02},h_0,F_0)^{-1}\circ D_1\bar{\phi}(h_{01},h_{02},h_0,F_0)\big]\\ &\qquad\times(\tilde{h}_{n1}-h_{01},\tilde{h}_{n2}-h_{02},\tilde{h}_n-h_0)\big\|\\ &\quad+o_p\big(\big\|(\tilde{h}_{n1}-h_{01},\tilde{h}_{n2}-h_{02},\tilde{h}_n-h_0)\big\|\big),\end{aligned}\tag{A.12}$$

where the norm $\|\cdot\|$ on the left-hand side and the first norm on the right-hand side denote the supremum norm on $C[0, M]$ and the norm in the o_p-term denotes the norm

$$\big\|(h_1,h_2,h)\big\|=\max\{\|h_1\|,\|h_2\|,\|h\|_S\},$$

where the first two norms denote again the supremum norm and the third norm $\|\cdot\|_S$ is defined by (4.11).

By well-known results in density estimation, we have, if $b_n \asymp n^{-1/5}$,

$$\max\big(\|\tilde{h}_{n1}-h_{01}\|,\|\tilde{h}_{n2}-h_{02}\|\big)=O_p\big(n^{-2/5}\sqrt{\log n}\big).$$

The boundary kernels ensure that the rates are not spoiled by what happens at the boundary. So, we have to determine the rate of convergence of $\|\tilde{h}_n - h_0\|_S$. We get

$$\begin{aligned}&\int_{u\,:\,(u,t)\in S}\big|\tilde{h}_n(u,t)-h_0(u,t)\big|\,du\\ &\qquad\le M^{1/2}\left\{\int_{u\,:\,(u,t)\in S}\big\{\tilde{h}_n(u,t)-h_0(u,t)\big\}^2\,du\right\}^{1/2}\end{aligned}$$

and

$$\left\{\int_{u\,:\,(u,t)\in S}\{\tilde h_n(u,t)-E\tilde h_n(u,t)\}^2\,du\right\}^{1/2}=O_p\big(n^{-2/5}\sqrt{\log n}\big),$$

uniformly in t. For the bias we get, if $b_n<u<u+\varepsilon\le t<M-b_n$,

$$\begin{aligned}
&\mathbb{E}\tilde h_n(u,t)-h_0(u,t)\\
&\quad=\mathbb{E}K_{b_n}(u-T_1)K_{b_n}(t-U_1)\Delta_{12}-h_0(u,t)\\
&\quad=\int K_{b_n}(u-t')K_{b_n}(t-u')h_0(t',u')\,dt'\,du'-h_0(u,t)\\
&\quad=\int K(v)K(w)h_0(u-b_nw,t-b_nv)\,dv\,dw-h_0(u,t)\\
&\quad=O\big(b_n^2\big).
\end{aligned}$$

The use of the boundary kernels ensures that the bias is also of order $O(b_n^2)$ is $u<b_n$ or $t>M-b_n$. The conclusion is

$$\|(\tilde h_{n1}-h_{01},\tilde h_{n2}-h_{02},\tilde h_n-h_0)\|=O_p\big(n^{-2/5}\sqrt{\log n}\big). \tag{A.13}$$

The derivative D_2 was computed in the proof of Lemma 4.1 (denoted by ∂_4 there) and the derivative D_1 is given by

$$\begin{aligned}
&\big[\big[D_1\bar\phi(h_{01},h_{02},h_0,F_0)\big](A)\big](t)\\
&\quad=B_1(t)\{1-F_0(t)\}-B_2(t)F_0(t)\\
&\qquad+F_0(t)\{1-F_0(t)\}\left\{\int_{v=0}^t\frac{B(v,t)}{F_0(t)-F_0(v)}\,dv-\int_{u=t}^M\frac{B(t,u)}{F_0(u)-F_0(t)}\,du\right\},
\end{aligned}$$

where B_1, B_2 and B are of the form

$$B_1=h_1-h_{01},\qquad B_2=h_2-h_{02},$$
$$B=h-h_0.$$

Hence, defining $\bar F_n$ by

$$\begin{aligned}
\bar F_n-F_0&=-\big[D_2\bar\phi(h_{01},h_{02},h_0,F_0)^{-1}\circ D_1\bar\phi(h_{01},h_{02},h_0,F_0)\big]\\
&\qquad\times(\tilde h_{n1}-h_{01},\tilde h_{n2}-h_{02},\tilde h_n-h_0),
\end{aligned}\tag{A.14}$$

we get that $F=\bar F_n$ is the solution of the linear integral equation

$$\begin{aligned}
&\big[D_2\bar\phi(h_{01},h_{02},h_0,F_0)\big](F-F_0)\\
&\quad=-\big[D_1\bar\phi(h_{01},h_{02},h_0,F_0)\big](\tilde h_{n1}-h_{01},\tilde h_{n2}-h_{02},\tilde h_n-h_0),
\end{aligned}$$

which, letting $A = F - F_0$ and $(B_1, B_2, B) = (\tilde{h}_{n1} - h_{01}, \tilde{h}_{n2} - h_{02}, \tilde{h}_n - h_0)$, boils down to the equation

$$\begin{aligned}
&\{h_{01}(t) + h_{02}(t)\}A(t) \\
&\quad - \{1 - 2F_0(t)\}\left\{\int_{v=0}^{t} \frac{h_0(v,t)}{F_0(t) - F_0(v)}\,dv - \int_{u=t}^{M} \frac{h_0(t,u)}{F_0(u) - F_0(t)}\,du\right\}A(t) \\
&\quad + F_0(t)\{1 - F_0(t)\} \\
&\quad \times \left\{\int_{u=0}^{t} \frac{h_0(u,t)\{A(t) - A(u)\}}{\{F_0(t) - F_0(u)\}^2}\,du + \int_{u=t}^{M} \frac{h_0(t,u)\{A(t) - A(u)\}}{\{F_0(u) - F_0(t)\}^2}\,du\right\} \\
&\quad = B_1(t)\{1 - F_0(t)\} - B_2(t)F_0(t) \\
&\qquad + F_0(t)\{1 - F_0(t)\} \\
&\qquad \times \left\{\int_{v=0}^{t} \frac{B(v,t)}{F_0(t) - F_0(v)}\,dv - \int_{u=t}^{M} \frac{B(t,u)}{F_0(u) - F_0(t)}\,du\right\}.
\end{aligned} \tag{A.15}$$

We have

$$\begin{aligned}
&\{h_{01}(t) + h_{02}(t)\}A(t) \\
&\quad - \{1 - 2F_0(t)\}\left\{\int_{v=0}^{t} \frac{h_0(v,t)}{F_0(t) - F_0(v)}\,dv - \int_{u=t}^{M} \frac{h_0(t,u)}{F_0(u) - F_0(t)}\,du\right\}A(t) \\
&\quad = \{g_1(t)F_0(t)(t) + g_2(t)\{1 - F_0(t)\}\}A(t) \\
&\qquad - \{1 - 2F_0(t)\}\{g_2(t) - g_1(t)\}A(t) \\
&\quad = \{\{1 - F_0(t)\}g_1(t) + F_0(t)g_2(t)\}A(t).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\int_{u=0}^{t} \frac{h_0(u,t)\{A(t) - A(u)\}}{\{F_0(t) - F_0(u)\}^2}\,du + \int_{u=t}^{M} \frac{h_0(t,u)\{A(t) - A(u)\}}{\{F_0(u) - F_0(t)\}^2}\,du \\
&\qquad = \int_{u=0}^{t} \frac{g(u,t)\{A(t) - A(u)\}}{F_0(t) - F_0(u)}\,du + \int_{u=t}^{M} \frac{g(t,u)\{A(t) - A(u)\}}{F_0(u) - F_0(t)}\,du.
\end{aligned}$$

Finally,

$$\begin{aligned}
&h_{01}(t)\{1 - F_0(t)\} - h_{02}(t)F_0(t) \\
&\quad + F_0(t)\{1 - F_0(t)\}\left\{\int_{v=0}^{t} \frac{h_0(v,t)}{F_0(t) - F_0(v)}\,dv - \int_{u=t}^{M} \frac{h_0(t,u)}{F_0(u) - F_0(t)}\,du\right\} = 0.
\end{aligned}$$

So, we obtain the linear integral equation (4.5) by dividing both sides of (A.15) by $\{1 - F_0(t)\}g_1(t) + F_0(t)g_2(t)$.

Hence, $\bar{F}_n$ is the solution of the linear integral equation (4.5), and by (A.12) and (A.13),

$$\|\tilde{F}_n - F_0\| = \|\bar{F}_n - F_0\| + o_p\big(n^{-2/5}\sqrt{\log n}\big).$$

Note that we have

$$\|\bar{F}_n - F_0\| = O_p\big(n^{-2/5}\sqrt{\log n}\big)$$

by the fact that $D_1\bar{\phi}(h_{01}, h_{02}, h_0, F_0)$ and $D_2\bar{\phi}(h_{01}, h_{02}, h_0, F_0)^{-1}$ are bounded linear mappings, and hence

$$\|\bar{F}_n - F_0\| = O_p\big(\|(\tilde{h}_{n1} - h_{01}, \tilde{h}_{n2} - h_{02}, \tilde{h}_n - h_0)\|\big) = O_p\big(n^{-2/5}\sqrt{\log n}\big).$$

(iii) The function $\tilde{F}_n$ satisfies the equation $\bar{\phi}(\tilde{h}_{n1}, \tilde{h}_{n2}, \tilde{h}_n, \tilde{F}_n) = 0$, where $\bar{\phi}$ is defined by (A.7). Hence,

$$\begin{aligned}&\tilde{h}_{n1}(t)\big\{1 - \tilde{F}_n(t)\big\} - \tilde{h}_{n2}(t)\tilde{F}_n(t)\\&\qquad + F(t)\big\{1 - F(t)\big\}\left\{\int_{v=0}^{t} \frac{\tilde{h}_n(v,t)}{\tilde{F}_n(t) - \tilde{F}_n(v)}\,dv - \int_{u=t}^{M} \frac{\tilde{h}_n(t,u)}{\tilde{F}_n(u) - \tilde{F}_n(t)}\,du\right\} = 0.\end{aligned}$$

By (ii), we have $\|\tilde{F}_n - F_0\| = O_p(n^{-2/5}\sqrt{\log n})$, and hence

$$\begin{aligned}&\int_{v=0}^{t} \frac{\tilde{h}_n(v,t)}{\tilde{F}_n(t) - \tilde{F}_n(v)}\,dv\\&\qquad = \int_{v=0}^{t} \frac{\tilde{h}_n(v,t)}{F_0(t) - F_0(v)}\,dv\\&\qquad\quad - \int_{v=0}^{t} \frac{g(u,t)\{\tilde{F}_n(t) - F_0(t) - \tilde{F}_n(u) + F_0(u)\}}{F_0(t) - F_0(u)}\,du + O_p\big(n^{-4/5}\log n\big)\end{aligned}$$

and similarly

$$\begin{aligned}&\int_{u=t}^{M} \frac{\tilde{h}_n(t,u)}{\tilde{F}_n(u) - \tilde{F}_n(t)}\,du\\&\qquad = \int_{u=t}^{M} \frac{\tilde{h}_n(t,u)}{F_0(u) - F_0(t)}\,du\\&\qquad\quad - \int_{u=t}^{M} \frac{g(t,u)\{\tilde{F}_n(u) - F_0(u) - \tilde{F}_n(t) + F_0(t)\}}{F_0(u) - F_0(t)}\,du + O_p\big(n^{-4/5}\log n\big).\end{aligned}$$

Hence, we get

$$\begin{aligned}&\tilde{F}_n(t) - F_0(t) + d_{F_0}(t)\left\{\int_{u=0}^{t} \frac{g(u,t)\{\tilde{F}_n(t) - F_0(t) - \tilde{F}_n(u) + F_0(u)\}}{F_0(t) - F_0(u)}\,du\right.\\&\qquad\qquad\qquad\qquad\left. - \int_{u=t}^{M} \frac{g(t,u)\{\tilde{F}_n(u) - F_0(u) - \tilde{F}_n(t) + F_0(t)\}}{F_0(u) - F_0(t)}\,du\right\}\\&\qquad = \frac{\tilde{h}_{n1}(t)\{1 - F_0(t)\} - \tilde{h}_{n2}(t)F_0(t)}{\{1 - F_0(t)\}g_1(t) + F_0(t)g_2(t)}\end{aligned}$$

$$+ d_{F_0}(t)\left\{\int_{u=0}^{t} \frac{\tilde h_n(u,t)}{F_0(t)-F_0(u)}\,du - \int_{u=t}^{M} \frac{\tilde h_n(t,u)}{F_0(u)-F_0(t)}\,du\right\}$$
$$+ O_p\bigl(n^{-4/5}\log n\bigr),$$

uniformly for $t \in [0, M]$, implying

$$\begin{aligned}\tilde F_n &= -\bigl[D_2\bar\phi(h_{01}, h_{02}, h_0, F_0)^{-1} \circ D_1\bar\phi(h_{01}, h_{02}, h_0, F_0)\bigr]\\ &\qquad \times (\tilde h_{n1} - h_{01}, \tilde h_{n2} - h_{02}, \tilde h_n - h_0)\\ &\qquad + O_p\bigl(n^{-4/5}\log n\bigr)\\ &= \bar F_n + O_p\bigl(n^{-4/5}\log n\bigr). \qquad \square\end{aligned}$$

PROOF OF LEMMA 4.6. We have

$$\tilde h_{n1}(v) = \frac{1}{n}\sum_{i=1}^{n} K_{b_n}(v - T_i)\Delta_{i1}$$

and hence

$$\begin{aligned}\operatorname{var}\bigl(\tilde h_{n1}(v)\bigr) &= \frac{1}{n}\operatorname{var}\bigl(K_{b_n}(v - T_1)\Delta_{11}\bigr)\\ &= \frac{1}{n} E K_{b_n}(v - T_1)^2\bigl(\Delta_{11} - F_0(T_1)\bigr)^2\\ &\sim \frac{F_0(v)\{1 - F_0(v)\}g_1(v)}{nb_n}\int K(u)^2\,du.\end{aligned}$$

Likewise,

$$\operatorname{var}\bigl(\tilde h_{n2}(v)\bigr) \sim \frac{F_0(v)\{1 - F_0(v)\}g_2(v)}{nb_n}\int K(u)^2\,du.$$

Furthermore,

$$\begin{aligned}&\operatorname{covar}\bigl(\tilde h_{n1}(v), \tilde h_{n2}(v)\bigr)\\ &\quad = \frac{1}{n}\mathbb{E}K_{b_n}(v - T_1)K_{b_n}(v - U_1)\bigl(\Delta_{11} - F_0(T_1)\bigr)\bigl(\Delta_{13} - F_0(U_1)\bigr)\\ &\quad = -\frac{1}{n}\mathbb{E}K_{b_n}(v - T_1)K_{b_n}(v - U_1)F_0(T_1)F_0(U_1)\\ &\quad \sim -\frac{F_0(v)^2 g(v,v)}{nb_n}\int K(u)^2\,du = 0,\end{aligned}$$

using the "separation condition" $g(v, v) = 0$. So, we obtain

$$\begin{aligned}&\operatorname{var}\left(\frac{\{1 - F_0(v)\}\tilde h_{n1}(v) - F_0(v)\tilde h_{n2}(v)}{g_1(v)\{1 - F_0(v)\} + F_0(v)g_2(v)}\right)\\ &\qquad \sim \frac{F_0(v)\{1 - F_0(v)\}(\{1 - F_0(v)\}^2 g_1(v) + F_0(v)^2 g_2(v))}{nb_n\{g_1(v)\{1 - F_0(v)\} + F_0(v)g_2(v)\}^2}.\end{aligned}$$

Furthermore,

$$\int_{t<v} \frac{\tilde{h}_n(t,v)}{F_0(v)-F_0(t)}\,dt = n^{-1}\sum_{i=1}^{n} K_{b_n}(v-U_i)\Delta_{i2}\int_{t<v}\frac{K_{b_n}(t-T_i)}{F_0(v)-F_0(t)}\,dt$$

and hence

$$\begin{aligned}
\operatorname{var}&\biggl(\int_{t<v} \frac{\tilde{h}_n(t,v)}{F_0(v)-F_0(t)}\,dt\biggr)\\
&= n^{-1}\operatorname{var}\biggl(K_{b_n}(v-U_1)\Delta_{12}\int_{t<v}\frac{K_{b_n}(t-T_1)}{F_0(v)-F_0(t)}\,dt\biggr)\\
&= n^{-1}\mathbb{E}K_{b_n}(v-U_1)^2\bigl(\Delta_{12}-F_0(U_1)+F_0(T_1)\bigr)^2\biggl\{\int_{t<v}\frac{K_{b_n}(t-T_1)}{F_0(v)-F_0(t)}\,dt\biggr)^2\\
&= n^{-1}\mathbb{E}K_{b_n}(v-U_1)^2\{F_0(U_1)-F_0(T_1)\}\{1-F_0(U_1)+F_0(T_1)\}\\
&\qquad\times\biggl\{\int_{t<v}\frac{K_{b_n}(t-T_1)}{F_0(v)-F_0(t)}\,dt\biggr)^2\\
&\sim \frac{1}{nb_n}\int_{t<v}\frac{1-F_0(v)+F_0(t)}{F_0(v)-F_0(t)}g(t,v)\,dt\int K(u)^2\,du.
\end{aligned}$$

Likewise,

$$\begin{aligned}
\operatorname{var}&\biggl(\int_{w>v} \frac{\tilde{h}_n(v,w)}{F_0(w)-F_0(v)}\,dw\biggr)\\
&\sim \frac{1}{nb_n}\int_{w>v}\frac{1-F_0(w)+F_0(v)}{F_0(w)-F_0(v)}g(v,w)\,dw\int K(u)^2\,du.
\end{aligned}$$

Finally,

$$\begin{aligned}
\operatorname{covar}&\biggl(\tilde{h}_{n2}(v),\int_{t<v} \frac{\tilde{h}_n(t,v)}{F_0(v)-F_0(t)}\,dt\biggr)\\
&= \frac{1}{n}\mathbb{E}K_{b_n}(v-U_1)^2\bigl(\Delta_{13}-\{1-F_0(U_1)\}\bigr)\bigl(\Delta_{12}-F_0(U_1)+F_0(T_1)\bigr)\\
&\qquad\times\int_{t<v}\frac{K_{b_n}(t-T_1)}{F_0(v)-F_0(t)}\,dt\\
&= -\frac{1}{n}\mathbb{E}K_{b_n}(v-U_1)^2\{1-F_0(U_1)\}\{F_0(U_1)-F_0(T_1)\}\\
&\qquad\times\int_{t<v}\frac{K_{b_n}(t-T_1)}{F_0(v)-F_0(t)}\,dt\\
&\sim -\frac{\{1-F_0(v)\}g_2(v)}{nb_n}\int K(u)^2\,du
\end{aligned}$$

and similarly

$$\operatorname{covar}\left(\tilde{h}_{n1}(v), \int_{w>v} \frac{\tilde{h}_n(v,w)}{F_0(w)-F_0(v)}\,dw\right) \sim -\frac{F_0(v)g_1(v)}{nb_n}\int K(u)^2\,du.$$

Combining these facts, we obtain that the variance of the right-hand side of (4.14) is given by

$$\begin{aligned}
&\frac{F_0(v)\{1-F_0(v)\}(\{1-F_0(v)\}^2 g_1(v)+F_0(v)^2 g_2(v))}{nb_n\{g_1(v)\{1-F_0(v)\}+F_0(v)g_2(v)\}^2}\int K(u)^2\,du\\
&\quad+\frac{d_{F_0}(v)^2}{nb_n}\\
&\qquad\times\left\{\int_{t<v}\frac{g(t,v)}{F_0(v)-F_0(t)}\,dt+\int_{w>v}\frac{g(v,w)}{F_0(w)-F_0(v)}\,dw-g_1(v)-g_2(v)\right\}\\
&\qquad\times\int K(u)^2\,du+\frac{2d_{F_0}(v)^2}{nb_n}\{g_1(v)+g_2(v)\}\int K(u)^2\,du\\
&\qquad=\frac{F_0(v)\{1-F_0(v)\}(\{1-F_0(v)\}^2 g_1(v)+F_0(v)^2 g_2(v))}{nb_n\{g_1(v)\{1-F_0(v)\}+F_0(v)g_2(v)\}^2}\int K(u)^2\,du\\
&\qquad\quad+\frac{d_{F_0}(v)^2}{nb_n}\left\{\int_{t<v}\frac{g(t,v)}{F_0(v)-F_0(t)}\,dt\right.\\
&\qquad\qquad\left.+\int_{w>v}\frac{g(v,w)}{F_0(w)-F_0(v)}\,dw+g_1(v)+g_2(v)\right\}\int K(u)^2\,du\\
&\qquad=\frac{d_{F_0}(v)}{nb_n}\left\{1+d_{F_0}(v)\left\{\int_{t<v}\frac{g(t,v)}{F_0(v)-F_0(t)}\,dt+\int_{w>v}\frac{g(v,w)}{F_0(w)-F_0(v)}\,dw\right\}\right\}\\
&\qquad\quad\times\int K(u)^2\,du.
\end{aligned}$$

Hence, we get that the asymptotic variance at a fixed interior point v of the solution F of the equation (4.14) is given by

$$\frac{d_{F_0}(v)}{nb_n\sigma_1}\int K(u)^2\,du,$$

where σ_1 is defined by (4.1).

We still have to compute the bias. We have

$$\begin{aligned}
&\mathbb{E}\frac{\{1-F_0(v)\}\{\tilde{h}_1(v)-h_1(v)\}-F_0(v)\{\tilde{h}_2(v)-h_2(v)\}}{g_1(v)\{1-F_0(v)\}+F_0(v)g_2(v)}\\
&\quad+d_{F_0}(v)\mathbb{E}\int_{t<v}\frac{\tilde{h}(t,v)-h(t,v)}{F_0(v)-F_0(t)}\,dt-d_{F_0}(v)\mathbb{E}\int_{u>v}\frac{\tilde{h}_n(v,u)-h(v,u)}{F_0(u)-F_0(v)}\,du
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\{1-F_0(v)\}h_1''(v) - F_0(v)h_2''(v)}{2\{g_1(v)\{1-F_0(v)\}+F_0(v)g_2(v)\}} b_n^2 \int u^2 K(u)\,du \\
&\quad + \frac{1}{2} b_n^2 d_{F_0}(v) \left\{ \int_{t<v} \frac{(\partial^2/\partial v^2)h(t,v)}{F_0(v)-F_0(t)}\,dt - \int_{u>v} \frac{(\partial^2/\partial v^2)h(v,u)}{F_0(u)-F_0(v)}\,du \right\} \\
&\quad \times \int u^2 K(u)\,du + o(b_n^2),
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde h_1(t) &= \int K_{b_n}(t-u)F_0(u)g_1(u)\,du, \\
\tilde h_2(t) &= \int K_{b_n}(t-u)\{1-F_0(u)\}g_2(u)\,du
\end{aligned}
\tag{A.16}
$$

and

$$
\tilde h(t,u) = \int K_{b_n}(t-v)K_{b_n}(u-w)\{F_0(w)-F_0(v)\}g(v,w)\,dv\,dw. \tag{A.17}
$$

Moreover,

$$
\begin{aligned}
&\{1-F_0(v)\}h_1(v) - F_0(v)h_2(v) \\
&\quad + F_0(v)\{1-F_0(v)\}\left\{ \int_{t<v} \frac{h(t,v)}{F_0(v)-F_0(t)}\,dt - \int_{u>v} \frac{h(v,u)}{F_0(u)-F_0(v)}\,du \right\} \\
&\quad = F_0(v)\{1-F_0(v)\}\{g_1(v)-g_2(v)\} - F_0(v)\{1-F_0(v)\}\{g_1(v)-g_2(v)\} \\
&\quad = 0.
\end{aligned}
$$

This yields the result of the lemma, since we can use the central limit theorem for i.i.d. random variables on the right-hand side of (4.14), using

$$
\begin{aligned}
&\frac{\tilde h_{n1}(t)\{1-F_0(t)\} - \tilde h_{n2}(t)F_0(t)}{\{1-F_0(t)\}g_1(t)+F_0(t)g_2(t)} \\
&\qquad + d_{F_0}(t)\left\{ \int_{u<t} \frac{\tilde h_n(u,t)}{F_0(t)-F_0(u)}\,du - \int_{u>t} \frac{\tilde h_n(t,u)}{F_0(u)-F_0(t)}\,du \right\} \\
&\qquad = n^{-1}\sum_{i=1}^n \left\{ \frac{\{1-F_0(t)\}K_{b_n}(t-T_i)\Delta_{i1} - F_0(t)K_{b_n}(t-U_i)\Delta_{i3}}{\{1-F_0(t)\}g_1(t)+F_0(t)g_2(t)} \right. \\
&\qquad\qquad + d_{F_0}(t)K_{b_n}(t-U_i)\Delta_{i2}\int_{u<t} \frac{K_{b_n}(u-T_i)}{F_0(t)-F_0(u)}\,du \\
&\qquad\qquad \left. - d_{F_0}(t)K_{b_n}(t-T_i)\Delta_{i2}\int_{u>t} \frac{K_{b_n}(u-U_i)}{F_0(u)-F_0(t)}\,du \right\}. \qquad \square
\end{aligned}
\tag{A.18}
$$

PROOF OF LEMMA 4.7. The difference between the equation defining the toy estimator and the solution of the integral equation (4.5) resides in the term

$$
-d_{F_0}(t)\left\{ \int_{u=0}^{t} \frac{g(u,t)\{F(u)-F_0(u)\}}{F_0(t)-F_0(u)}\,du + \int_{u=t}^{M} \frac{g(t,u)\{F(u)-F_0(u)\}}{F_0(u)-F_0(t)}\,du \right\};
$$

see (4.5) and (4.6). Take $F = F_n^{\text{toy}}$, and consider the integral within the brackets. An easy computation yields

$$\int_{u=0}^{t} \frac{\{F_n^{\text{toy}}(u) - \mathbb{E}F_n^{\text{toy}}(u)\}g(u,t)}{F_0(t) - F_0(u)}\,du = O_p(n^{-1/2}).$$

So, we get

$$\begin{aligned}
d_{F_0}(t)&\left\{\int_{u=0}^{t} \frac{g(u,t)\{F_n^{\text{toy}}(u) - F_0(u)\}}{F_0(t) - F_0(u)}\,du + \int_{u=t}^{M} \frac{g(t,u)\{F_n^{\text{toy}}(u) - F_0(u)\}}{F_0(u) - F_0(t)}\,du\right\} \\
&= d_{F_0}(t)\left\{\int_{u=0}^{t} \frac{g(u,t)\{\mathbb{E}F_n^{\text{toy}}(u) - F_0(u)\}}{F_0(t) - F_0(u)}\,du \right.\\
&\qquad \left. + \int_{u=t}^{M} \frac{g(t,u)\{\mathbb{E}F_n^{\text{toy}}(u) - F_0(u)\}}{F_0(u) - F_0(t)}\,du\right\} + O_p(n^{-1/2}).
\end{aligned}$$

This yields the result of the lemma, since the bias of the toy estimator is given by

$$\frac{\beta_1(u)b_n^2}{\sigma_1(u)} + o(b_n^2),$$

which implies, by the preceding, that

$$F_n^{\text{toy}}(t) + d_{F_0}(t)\left\{\int_{u=0}^{t} \frac{\gamma_n(u)g(u,t)}{F_0(t) - F_0(u)}\,du + \int_{u=t}^{M} \frac{\gamma_n(u)g(t,u)}{F_0(u) - F_0(t)}\,du\right\}$$

satisfies (4.5), apart from a term of order $O_p(n^{-1/2})$. $\square$

REFERENCES

[1] DIEUDONNÉ, J. (1969). *Foundations of Modern Analysis*. Academic Press, New York. Enlarged and corrected printing, *Pure and Applied Mathematics* **10-I**. MR0349288

[2] GESKUS, R. and GROENEBOOM, P. (1999). Asymptotically optimal estimation of smooth functionals for interval censoring, case 2. *Ann. Statist.* **27** 627–674. MR1714713

[3] GESKUS, R. B. and GROENEBOOM, P. (1996). Asymptotically optimal estimation of smooth functionals for interval censoring. I. *Stat. Neerl.* **50** 69–88. MR1381209

[4] GESKUS, R. B. and GROENEBOOM, P. (1997). Asymptotically optimal estimation of smooth functionals for interval censoring. II. *Stat. Neerl.* **51** 201–219. MR1466426

[5] GROENEBOOM, P. (1991). Nonparametric maximum likelihood estimators for interval censoring and deconvolution. Technical Report 378, Statistics Dept., Stanford Univ.

[6] GROENEBOOM, P. (1996). Lectures on inverse problems. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1994). *Lecture Notes in Math.* **1648** 67–164. Springer, Berlin. MR1600884

[7] GROENEBOOM, P. (2012). Likelihood ratio type two-sample tests for current status data. *Scand. J. Stat.* **39** 645–662. MR3000840

[8] GROENEBOOM, P. and JONGBLOED, G. (2014). Nonparametric confidence intervals for monotone functions. Unpublished manuscript.

[9] GROENEBOOM, P. and JONGBLOED, G. (2014). *Nonparametric Estimation Under Shape Constraints*. Cambridge Univ. Press, Cambridge.

[10] GROENEBOOM, P., JONGBLOED, G. and WITTE, B. I. (2010). Maximum smoothed likelihood estimation and smoothed maximum likelihood estimation in the current status model. *Ann. Statist.* **38** 352–387. MR2589325

[11] GROENEBOOM, P. and KETELAARS, T. (2011). Estimators for the interval censoring problem. *Electron. J. Stat.* **5** 1797–1845. MR2870151

[12] GROENEBOOM, P. and WELLNER, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. *DMV Seminar* **19**. Birkhäuser, Basel. MR1180321

[13] JONGBLOED, G. (1998). The iterative convex minorant algorithm for nonparametric estimation. *J. Comput. Graph. Statist.* **7** 310–321. MR1646718

[14] SCHICK, A. and YU, Q. (2000). Consistency of the GMLE with mixed case interval-censored data. *Scand. J. Stat.* **27** 45–55. MR1774042

DELFT INSTITUTE OF APPLIED MATHEMATICS
DELFT UNIVERSITY OF TECHNOLOGY
MEKELWEG 4, 2628 CD DELFT
THE NETHERLANDS
E-MAIL: P.Groeneboom@tudelft.nl
URL: http://dutiosc.twi.tudelft.nl/~pietg/