Submitted to the Annals of Statistics
arXiv: arXiv:0000.0000

SUPPLEMENT TO "ESTIMATION OF A MONOTONE DENSITY IN S-SAMPLE BIASED SAMPLING MODELS"*

BY KWUN CHUEN GARY CHAN*, HOK KAN LING†, TONY SIT‡ AND SHEUNG CHI PHILLIP YAM‡

*University of Washington *, Columbia University † and The Chinese University of Hong Kong‡*

8. Appendix for Section 3.

PROOF OF PROPOSITION 3.1. For any $g \in \mathcal{G}$ with $g(T_i) = q_i$, $i = 1, \dots, n$, and $q_i \geq q_{i+1}$, $i = 1, \dots, n-1$, define a step function $\tilde{g}$ such that $\tilde{g}(x) = q_i$ for $x \in (T_{i-1}, T_i]$; otherwise $\tilde{g}(x) = 0$. Then, $\tilde{g} \leq g$, which implies that

$$\int_a^b w_i(x)\tilde{g}(x)dx \leq \int_a^b w_i(x)g(x)dx, \tag{8.1}$$

as $w_i \geq 0$. Now, define $\bar{g}(x) \triangleq \dfrac{\tilde{g}(x)}{\int_a^b \tilde{g}(x)dx}$ so that $\int_a^b \bar{g}(x)dx = 1$. By (8.1),

$$\begin{aligned}
L_n(\bar{g}) &= \frac{\prod_{i=1}^n \bar{g}(T_i)}{\prod_{i=1}^s (\int_a^b w_i(x)\bar{g}(x)dx)^{n_i}} = \frac{\prod_{i=1}^n q_i}{\prod_{i=1}^s (\int_a^b w_i(x)\tilde{g}(x)dx)^{n_i}} \\
&\geq \frac{\prod_{i=1}^n q_i}{\prod_{i=1}^s (\int_a^b w_i(x)g(x)dx)^{n_i}} = L_n(g).
\end{aligned}$$

□

PROPOSITION 8.1. *The function $\tilde{\mathcal{L}}_n$ defined in (3.3) is concave in p.*

PROOF. (i) For finite $\boldsymbol{p}$, it suffices to show that $\Omega_i(\boldsymbol{p}) \triangleq \log(\sum_{k=1}^n c_{ik}e^{p_k})$ is convex in $\boldsymbol{p}$ as sum of convex functions with non-negative weights is still a convex function. To this end, we shall show that the Hessian matrix H_i of Ω_i is positive semi-definite. Define $W_i \triangleq \sum_{k=1}^n c_{ik}e^{p_k}$ and $\boldsymbol{\gamma_i} \triangleq (c_{i1}e^{p_1}, \dots, c_{in}e^{p_n})^\top$. Note that

$$\begin{aligned}
\frac{\partial \Omega_i}{\partial p_j} &= \frac{1}{W_i} c_{ij}e^{p_j}, \\
\frac{\partial^2 \Omega_i}{\partial p_j^2} &= \frac{1}{W_i} c_{ij}e^{p_j} - \frac{1}{W_i^2}\left(c_{ij}e^{p_j}\right)^2, \\
\frac{\partial^2 \Omega_i}{\partial p_j p_l} &= -\frac{1}{W_i^2} c_{ij}c_{il}e^{p_j}e^{p_l}.
\end{aligned}$$

Hence, $H_i = \frac{1}{W_i^2}(W_i \text{diag}(\boldsymbol{\gamma_i}) - \boldsymbol{\gamma_i}\boldsymbol{\gamma_i}^\top)$. Now, for any $\boldsymbol{x} \in \mathbb{R}^n$, by Cauchy-Schwartz inequality,

$$\boldsymbol{x}^\top H_i \boldsymbol{x} \quad = \quad \frac{1}{W_i^2}\left(\sum_{j=1}^n \gamma_{ij} \sum_{j=1}^n \gamma_{ij} x_j^2 - \left(\sum_{j=1}^n \gamma_{ij} x_j \right)^2 \right) \geq 0.$$

Finally, observe that if $\boldsymbol{x} = (c, \ldots, c)^\top$ for any $c \in \mathbb{R}$, $\boldsymbol{x}^\top H_i \boldsymbol{x} = 0$. Therefore, Ω_i is only positive semi-definite.

(ii) For $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, if either one of them such that at least one, but not all, of its components equals $-\infty$, then the convex combination of $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ also has a component being equal to $-\infty$. Therefore, the concavity inequality is trivially satisfied.

□

PROPOSITION 8.2. *Suppose that $w_i > 0$ for all $i = 1, \ldots, s$. For each $n \in \mathbb{N}$, the monotone MLE $\hat{g}_n$ for the true unbiased density g_0 uniquely exists almost surely.*

PROOF. Recall that the domain of $\mathcal{L}_n(z_1, \ldots, z_n) = \prod_{i=1}^n z_i \prod_{j=1}^s (\sum_{k=1}^n c_{jk} z_k)^{-n_j}$ is $\mathcal{K}_n = \{z \in \mathbb{R}^n : z_1 \geq \ldots \geq z_n \geq 0 \text{ and } \sum_{i=1}^n z_i(T_i - T_{i-1}) = 1\}$. Note that the constraint $\sum_{i=1}^n z_i(T_i - T_{i-1}) = 1$ ensures at least one of the $z_i' s$ to be positive (non-vanishing) and hence the terms in the denominator of $\mathcal{L}_n$, $\sum_{k=1}^n z_k c_{jk} > 0$. Let $\mathcal{D}_n := \{z \in \mathcal{K}_n : \text{at least one but not all } z_i\text{'s equal } 0\}$. On $\mathcal{D}_n$, $\mathcal{L}_n$ is still well-defined and vanishes here. Therefore, $\mathcal{L}_n$ is continuous on $\mathcal{K}_n$. As $\mathcal{K}_n$ is compact almost surely, the maximum of $\mathcal{L}_n$ exists almost surely.

For uniqueness, suppose that $\boldsymbol{z}'$ and $\boldsymbol{z}''$ both maximizes (3.1) subject to (3.2); note that as the problem setting is non-trivial, none of the components in z' and z'' is vanished. Then $p_i' \triangleq \log z_i'$ and $p_i'' \triangleq \log z_i''$ both maximize (3.3) subject to the corresponding constraints with z_i replaced by e^{p_i}. As $\tilde{\mathcal{L}}_n$ is concave in $\boldsymbol{p}$, $\lambda \boldsymbol{p}' + (1-\lambda)\boldsymbol{p}''$ is also a maximizer of $\tilde{\mathcal{L}}_n$ for any $0 \leq \lambda \leq 1$. Therefore, $\tilde{\mathcal{L}}_\lambda \triangleq \tilde{\mathcal{L}}_n(\lambda \boldsymbol{p}' + (1-\lambda)\boldsymbol{p}'')$ is a constant function in λ for $0 \leq \lambda \leq 1$ and hence,

$$\frac{d^2 \tilde{\mathcal{L}}_\lambda}{d\lambda^2} = 0 \quad \text{for } \lambda \in [0, 1].$$

Since

$$\tilde{\mathcal{L}}_\lambda = \sum_{j=1}^n \left(\lambda p_j' + (1-\lambda) p_j'' \right) - \sum_{i=1}^s n_i \log \left(\sum_{k=1}^n c_{ik} \exp(\lambda p_k' + (1-\lambda) p_k'') \right),$$

we have

$$\begin{aligned}\frac{d\tilde{\mathcal{L}}_\lambda}{d\lambda} &= \sum_{j=1}^n (p'_j - p''_j) - \sum_{i=1}^s n_i \frac{\sum_{k=1}^n c_{ik}(p'_k - p''_k)\exp(\lambda p'_k + (1-\lambda)p''_k)}{\sum_{k=1}^n c_{ik}\exp(\lambda p'_k + (1-\lambda)p''_k)},\\ \frac{d^2\tilde{\mathcal{L}}_\lambda}{d\lambda^2} &= \sum_{i=1}^s n_i \frac{[\sum_{k=1}^n c_{ik}(p'_k - p''_k)\exp(\lambda p'_k + (1-\lambda)p''_k)]^2}{[\sum_{k=1}^n c_{ik}\exp(\lambda p'_k + (1-\lambda)p''_k)]^2}\\ &\quad - \sum_{i=1}^s n_i \frac{\sum_{k=1}^n c_{ik}(p'_k - p''_k)^2\exp(\lambda p'_k + (1-\lambda)p''_k)}{\sum_{k=1}^n c_{ik}\exp(\lambda p'_k + (1-\lambda)p''_k)}.\end{aligned}$$

In particular, when $\lambda = 1$, we obtain an equation:

$$\begin{aligned}\left.\frac{d^2\tilde{\mathcal{L}}_\lambda}{d\lambda^2}\right|_{\lambda=1} &= \sum_{i=1}^s n_i \frac{[\sum_{k=1}^n c_{ik}(p'_k - p''_k)e^{p'_k}]^2}{(\sum_{k=1}^n c_{ik}e^{p'_k})^2}\\ &\quad - \sum_{i=1}^s n_i \frac{\sum_{k=1}^n c_{ik}(p'_k - p''_k)^2 e^{p'_k}}{\sum_{k=1}^n c_{ik}e^{p'_k}}\\ &= 0.\end{aligned}$$

Hence, we have $\sum_{i=1}^s n_i\Delta_i = 0$, where

$$\Delta_i \triangleq \frac{(\sum_{k=1}^n c_{ik}e^{p'_k})(\sum_{k=1}^n c_{ik}(p'_k - p''_k)^2 e^{p'_k}) - (\sum_{k=1}^n c_{ik}(p'_k - p''_k)e^{p'_k})^2}{(\sum_{k=1}^n c_{ik}e^{p'_k})^2}.$$

By Cauchy-Schwartz inequality, $\Delta_i \geq 0$ for all $i = 1, \dots, s$. Note that $\sum_{i=1}^s n_i\Delta_i = 0$ and $\Delta_i \geq 0$ for all $i = 1, \dots, s$ together imply that $\Delta_i = 0$ for all $i = 1, \dots, s$. As $w_i > 0$, $c_{ik} > 0$ for all $k = 1, \dots, n$. Hence, the equality holds if and only if $p'_k - p''_k = c$ for all $k = 1, \dots, n$ and for some $c \in \mathbb{R}$. This is equivalent to $z'_k = e^c z''_k$. As $1 = \sum_{i=1}^k z'_k(T_k - T_{k-1}) = e^c \sum_{i=1}^k z''_k(T_k - T_{k-1}) = e^c$, we know that $c = 0$, and so $\boldsymbol{z}' = \boldsymbol{z}''$. □

PROOF OF PROPOSITION 3.2. The problem is to minimize ψ_n subject to $z_1 \geq \dots \geq z_n$, or equivalently $z_{i+1} - z_i \leq 0$ for $i = 1, \dots, n-1$. Note that we do not explicitly require $z_n \geq 0$ as $\mathcal{L}_n$ will not be minimized at $z_n = 0$. By the homogeneity degree of 0 of ψ_n, we also do not need to consider the constraint $\sum_{i=1}^n z_i(T_i - T_{i-1}) = 1$. Denote $\mu_0 = \mu_n \triangleq 0$. Using Karush-Kuhn-Tucker's theorem (see, for example, Chapter 11 in Luenberger and Ye (2008)), there exist μ_i for $i = 1, \dots, n-1$ such that the minimizer $\hat{\boldsymbol{z}}$ satisfies for $k = 1, \dots, n$,

$$\frac{\partial\psi_n}{\partial z_k}(\hat{\boldsymbol{z}}) + \mu_{k-1} - \mu_k = 0 \tag{8.2}$$

and

$$\text{(8.3)} \qquad \begin{cases} \mu_i(\hat{z}_{i+1} - \hat{z}_i) = & 0, \quad \text{for } i = 1, \ldots, n-1; \\ \mu_i \geq & 0, \quad \text{for } i = 1, \ldots, n-1; \\ \hat{z}_{i+1} - \hat{z}_i \leq & 0, \quad \text{for } i = 1, \ldots, n-1. \end{cases}$$

Now, summing (8.2) for $k = 1, \ldots, n$ and recursively using 8.2, we have respectively,

$$\sum_{j=1}^{n} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) = 0 \text{ and } \mu_i = \sum_{j=1}^{i} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) \geq 0, \quad \text{for } i = 1, \ldots, n-1.$$

According to Proposition 8.2, we let $\hat{\boldsymbol{z}}$ be the unique solution to (3.1)-(3.2). Define

$$\xi_K(\boldsymbol{z}) \triangleq \frac{1}{2}\left[\boldsymbol{z} - \hat{\boldsymbol{z}} + K^{-1} \bigtriangledown \psi_n(\hat{\boldsymbol{z}})\right]^{\top} K \left[\boldsymbol{z} - \hat{\boldsymbol{z}} + K^{-1} \bigtriangledown \psi_n(\hat{\boldsymbol{z}})\right],$$

where K is a positive-definite matrix to be determined. Then $\text{Hess}(\xi_K) = K$ and $\bigtriangledown(\xi_K) = K\left[\boldsymbol{z} - \hat{\boldsymbol{z}} + K^{-1} \bigtriangledown \psi_n(\hat{\boldsymbol{z}})\right]$. Hence, ξ is strictly concave with $\bigtriangledown \xi(\hat{\boldsymbol{z}}) = \bigtriangledown \psi_n(\hat{\boldsymbol{z}})$. Therefore, $\hat{z}$ satisfies the necessary and sufficient conditions for minimizing ξ subject to $z_1 \geq \ldots \geq z_n$. Also, $\hat{z}$ is the unique minimizer of ξ subject to the constraints $z_1 \geq \ldots \geq z_n$. Now, choose K to be a diagonal matrix with diagonal entry $d_i \triangleq \frac{1}{\hat{z}_i^2}$. Then K is positive-definite and $\hat{\boldsymbol{z}}$ minimizes

$$\xi(\boldsymbol{z}) = \sum_{i=1}^{n} \left[z_i - \hat{z}_i + \frac{\partial \psi_n}{\partial z_i}(\hat{z}) d_i^{-1}\right]^2 d_i = \sum_{i=1}^{n} \left[z_i - \left(\hat{z}_i - \frac{\partial \psi_n}{\partial z_i}(\hat{z}) d_i^{-1}\right)\right]^2 d_i.$$

subject to

$$z_1 \geq z_2 \geq \ldots \geq z_n,$$

and it furnishes the antitonic regression of the function

$$g(i) \triangleq \hat{z}_i - \frac{\partial \psi_n}{\partial z_i}(\hat{\boldsymbol{z}}) d_i^{-1}$$

on the ordered set $\{1, 2, \ldots, n\}$ with weight function d_i. Finally, it is well-known (see, for example, Barlow et al. (1972) or Robertson, Wright and Dykstra (1988)) that the solution is

$$(\hat{z}_1, \ldots, \hat{z}_n) \quad = \quad \text{slolcm}\left\{\sum_{j=1}^{i} d_j, \sum_{j=1}^{i} g(j) d_j\right\}_{i=0}^{n}.$$

□

9. Appendix for Section 4.

PROOF OF PROPOSITION 4.1. (i) We first state a variant of Lemma 4.5 in van de Geer (2000):

$$\sum_{i=1}^{s} h^2(\hat{f}_{i,n}, f_i) \leq \sum_{i=1}^{s} \int_a^b \frac{2\hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} d(\mathbb{F}_{i,n_i} - F_i)(x). \tag{9.1}$$

Its derivation is similar to that in van de Geer (2000) by noting the convexity of $\mathcal{G}$. Indeed, from the definition of monotone MLE, we have

$$\sum_{i=1}^{s} \int_a^b \log \hat{f}_{i,n}(x) d\mathbb{F}_{i,n_i}(x) \geq \sum_{i=1}^{s} \int_a^b \log \frac{\hat{f}_{i,n}(x) + f_i(x)}{2} d\mathbb{F}_{i,n_i}(x).$$

Hence, by noting the tangent line of the concave logarithm curve at $u = 1$, we have $\log u \leq u - 1$ for $u > 0$, and so

$$\begin{aligned}
0 &\leq \sum_{i=1}^{s} \int_a^b \log \frac{2\hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} d\mathbb{F}_{i,n_i}(x) \\
&\leq \sum_{i=1}^{s} \int_a^b \left(\frac{2\hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} - 1 \right) d\mathbb{F}_{i,n_i}(x) \\
&= \sum_{i=1}^{s} \int_a^b \left(\frac{2\hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} - 1 \right) d(\mathbb{F}_{i,n_i} - F_i)(x) \\
&\quad + \sum_{i=1}^{s} \int_a^b \left(\frac{2\hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} - 1 \right) dF_i(x) \\
&= \sum_{i=1}^{s} \int_a^b \frac{2\hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} d(\mathbb{F}_{i,n_i} - F_i)(x) \\
&\quad - \sum_{i=1}^{s} \int_a^b \frac{f_i(x) - \hat{f}_{i,n}(x)}{\hat{f}_{i,n}(x) + f_i(x)} dF_i(x).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\sum_{i=1}^{s}\int_a^b \frac{f_i(x)-\hat f_{i,n}(x)}{\hat f_{i,n}(x)+f_i(x)}dF_i(x) = \sum_{i=1}^{s}\int_a^b \frac{f_i(x)-\hat f_{i,n}(x)}{\hat f_{i,n}(x)+f_i(x)}f_i(x)dx \\
= \;& \frac{1}{2}\sum_{i=1}^{s}\int_a^b \frac{f_i(x)-\hat f_{i,n}(x)}{\hat f_{i,n}(x)+f_i(x)}(f_i(x)+\hat f_{i,n}(x))dx \\
&\quad +\frac{1}{2}\sum_{i=1}^{s}\int_a^b \frac{f_i(x)-\hat f_{i,n}(x)}{\hat f_{i,n}(x)+f_i(x)}(f_i(x)-\hat f_{i,n}(x))dx \\
= \;& \frac{1}{2}\sum_{i=1}^{s}(1-1)+\frac{1}{2}\sum_{i=1}^{s}\int_a^b \frac{(\hat f_{i,n}(x)-f_i(x))^2}{\hat f_{i,n}(x)+f_i(x)}dx \\
= \;& \frac{1}{2}\sum_{i=1}^{s}\int_a^b \left(\sqrt{\hat f_{i,n}(x)}-\sqrt{f_i(x)}\right)^2\frac{\left(\sqrt{\hat f_{i,n}(x)}+\sqrt{f_i(x)}\right)^2}{\hat f_{i,n}(x)+f_i(x)}dx \\
\geq \;& \sum_{i=1}^{s} h^2(\hat f_{i,n}, f_i).
\end{aligned}$$

Therefore, (9.1) follows. Define, for each $i = 1, \ldots, s$, the classes

$$\mathcal{H}_i \triangleq \left\{ \frac{2\tilde f_i}{\tilde f_i + f_i} = \frac{\frac{2\tilde g}{\int w_i\tilde g}}{\frac{\tilde g}{\int w_i \tilde g}+\frac{g_0}{\int w_i g_0}} : \tilde g \in \mathcal{G} \text{ and } \tilde f_i \triangleq \frac{w_i\tilde g}{\int w_i \tilde g}\right\}, \tag{9.2}$$

and

$$\widetilde{\mathcal{H}}_i \triangleq \left\{ \frac{2\frac{\tilde g}{\int w_i\tilde g}\frac{g_0}{\int w_i g_0}}{\frac{\tilde g}{\int w_i \tilde g}+\frac{g_0}{\int w_i g_0}} : \tilde g \in \mathcal{G}\right\}.$$

Note that for any measurable functions h_1 and h_2 that are F_i-integrable,

$$\int |h_1(x)-h_2(x)|dF_i(x) = \int |h_1(x)-h_2(x)|\frac{g_0(x)}{\int w_i g_0}\left\|\frac{\int w_i g_0}{g_0}\right\|_{1,F_i} dQ_i(x),$$

where $dQ_i \triangleq \left\|\frac{\int w_i g_0}{g_0}\right\|_{1,F_i}^{-1} \frac{\int w_i g_0}{g_0}dF_i$. Hence,

$$H_{1,B}(\delta, \mathcal{H}_i, F_i) = H_{1,B}\left(\delta\left\|\frac{\int w_i g_0}{g_0}\right\|_{1,F_i}^{-1}, \widetilde{\mathcal{H}}_i, Q_i\right). \tag{9.3}$$

For any $\tilde g \in \mathcal{G}$, let $q_1 \triangleq \frac{\tilde g}{\int w_i \tilde g}$ and $q_0 \triangleq \frac{g_0}{\int w_i g_0}$. Note that $\frac{q_1 q_0}{q_1+q_0} = (\frac{1}{q_0}+\frac{1}{q_1})^{-1}$. Thus, $\frac{q_1 q_0}{q_1+q_0}$ is a decreasing function. Moreover, since $\frac{q_1}{q_1+q_0} \leq 1$, it is clear that $\widetilde{\mathcal{H}}_i$ is uniformly bounded by $\frac{2\sup_{x\in[a,b]}|g_0(x)|}{\int w_i g_0}$. Therefore, $\widetilde{\mathcal{H}}_i$ is a class of decreasing functions which are uniformly bounded by $\frac{2\sup_{x\in[a,b]}|g_0(x)|}{\int w_i g_0}$. Thus, by Theorem 2.7.5 in van der Vaart and Wellner (1996),

$$H_{1,B}\left(\delta\left\|\frac{\int w_i g_0}{g_0}\right\|_{1,F_i}^{-1}, \widetilde{\mathcal{H}}_i, Q_i\right) \le \frac{A\left\|\frac{\int w_i g_0}{g_0}\right\|_{1,F_i}}{\delta}, \tag{9.4}$$

for any $\delta > 0$, where A is a universal constant. Hence, $\mathcal{H}_i$ is a Glivenko-Cantelli class by Lemma 3.1 in van de Geer (2000) and so $h(\hat{f}_{i,n}, f_i) \to 0$, for $i = 1, \dots, s$, a.s. in view of (9.1) and (9.3).

(ii) By definition of $\hat{f}_{i,n} = \frac{w_i\hat{g}_n}{\int w_i\hat{g}_n}$ and $f_i = \frac{w_i g_0}{\int w_i g_0}$, since $\int \hat{g}_n = \int g_0 = 1$,

$$\begin{aligned}
&\left|\frac{1}{\int w_i\hat{g}_n} - \frac{1}{\int w_i g_0}\right| \\
=\;& \left|\int \frac{\hat{f}_{i,n}(x)}{w_i(x)}dx - \int \frac{f_i(x)}{w_i(x)}dx\right| \\
=\;& \left|\int_a^b \frac{1}{w_i}(\sqrt{\hat{f}_{i,n}(x)} - \sqrt{f_i(x)})(\sqrt{\hat{f}_{i,n}(x)} + \sqrt{f_i(x)})dx\right| \\
\le\;& \frac{1}{m}\int_a^b (\sqrt{\hat{f}_{i,n}(x)} + \sqrt{f_i(x)})\left|\sqrt{\hat{f}_{i,n}(x)} - \sqrt{f_i(x)}\right|dx \\
\le\;& \frac{1}{m}\left(\int_a^b (\sqrt{\hat{f}_{i,n}(x)} + \sqrt{f_i(x)})^2 dx\right)^{1/2}\left(\int_a^b (\sqrt{\hat{f}_{i,n}(x)} - \sqrt{f_i(x)})^2 dx\right)^{1/2} \\
\le\;& \frac{2}{m}\left(\int_a^b (\sqrt{\hat{f}_{i,n}(x)} - \sqrt{f_i(x)}dx)^2\right)^{1/2} \\
=\;& \frac{2}{m}h(\hat{f}_{i,n}, f_i) \to 0, \quad a.s.,
\end{aligned}$$

where the first inequality follows from the boundedness from below of w_i and the triangle inequality; the second inequality follows from the Cauchy-Schwartz inequality; and the third inequality follows from the inequality $(a+b)^2 \le 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$ and note that $\int_a^b \hat{f}_{i,n}(x)dx = \int_a^b f_i(x)dx = 1$.

(iii) By considering the fact that $\hat{g}_n$ is a density function, for any $0 < \delta < b - a$, $\hat{g}_n(a+\delta)\delta \le 1$ (see Figure 1 for an illustration). Hence, $\limsup_n \hat{g}_n(a+\delta) \le \frac{1}{\delta}$. To establish the second claim, denote $\phi_n(\boldsymbol{z}) \triangleq \log \mathcal{L}_n(\boldsymbol{z})$, that is:

$$\phi_n(\boldsymbol{z}) = \sum_{i=1}^n \log z_i - \sum_{j=1}^s n_j \log\left(\sum_{k=1}^n z_k c_{jk}\right).$$

Also denote $\dot{z}_i \triangleq g_0(T_i)$ and $\dot{\boldsymbol{z}} = (\dot{z}_1, \ldots, \dot{z}_n)$; note that $\hat{\boldsymbol{z}}$ is defined as the monotone MLE $(\hat{g}_n(T_1), \ldots, \hat{g}_n(T_n))$. For each $\varepsilon \in (0,1)$, $(1-\varepsilon)\hat{z}_i + \varepsilon \dot{z}_i$ is decreasing in i. Let $\Sigma_\varepsilon \triangleq \sum_{i=1}^n [(1-\varepsilon)\hat{z}_i + \varepsilon \dot{z}_i](T_i - T_{i-1})$. We have, by the definition and its globally maximizing nature of the monotone MLE $\hat{\boldsymbol{z}}$, for any $\varepsilon \in (0,1)$,

$$\{[\phi_n(((1-\varepsilon)\hat{\boldsymbol{z}} + \varepsilon \dot{\boldsymbol{z}})/\Sigma_\varepsilon) - \phi_n(\hat{\boldsymbol{z}})]\varepsilon^{-1}\} \leq 0.$$

The homogeneity of degree 0 of ϕ_n gives

$$\phi_n(((1-\varepsilon)\hat{\boldsymbol{z}} + \varepsilon \dot{\boldsymbol{z}})/\Sigma_\varepsilon) = \phi_n((1-\varepsilon)\hat{\boldsymbol{z}} + \varepsilon \dot{\boldsymbol{z}}).$$

Hence, we have

$$\lim_{\varepsilon \downarrow 0} \{[\phi_n((1-\varepsilon)\hat{\boldsymbol{z}} + \varepsilon \dot{\boldsymbol{z}}) - \phi_n(\hat{\boldsymbol{z}})]\varepsilon^{-1}\} \leq 0.$$

To evaluate this limit, note that

$$\frac{d}{d\varepsilon}\phi_n((1-\varepsilon)\hat{\boldsymbol{z}} + \varepsilon \dot{\boldsymbol{z}}) = \sum_{i=1}^n \frac{\dot{z}_i - \hat{z}_i}{\hat{z}_i + \varepsilon(\dot{z}_i - \hat{z}_i)} - \sum_{j=1}^s \frac{n_j \sum_{k=1}^n (\dot{z}_k - \hat{z}_k) c_{jk}}{\sum_{k=1}^n (\hat{z}_k + \varepsilon(\dot{z}_k - \hat{z}_k)) c_{jk}};$$

taking $\varepsilon = 0$, we deduce that

$$\lim_{\varepsilon \downarrow 0} \{[\phi_n((1-\varepsilon)\hat{\boldsymbol{z}} + \varepsilon \dot{\boldsymbol{z}}) - \phi_n(\hat{\boldsymbol{z}})]\varepsilon^{-1}\} = \sum_{i=1}^n \frac{\dot{z}_i}{\hat{z}_i} - \sum_{j=1}^s \frac{n_j \sum_{k=1}^n \dot{z}_k c_{jk}}{\sum_{k=1}^n \hat{z}_k c_{jk}};$$

or equivalently,

$$\frac{1}{n}\sum_{i=1}^n \frac{g_0(T_i)}{\hat{g}_n(T_i)} - \sum_{j=1}^s \frac{n_j}{n} \frac{\sum_{k=1}^n g_0(T_k) c_{jk}}{\sum_{k=1}^n \hat{g}_n(T_k) c_{jk}} \leq 0.$$

Rewrite the first term on the left hand side of the above inequality in terms of the empirical measures $\mathbb{F}_{i,n_i}$'s, we obtain

$$\sum_{i=1}^s \frac{n_i}{n} \int_a^b \frac{g_0(x)}{\hat{g}_n(x)} d\mathbb{F}_{i,n_i}(x) \leq \sum_{j=1}^s \frac{n_j}{n} \frac{\sum_{k=1}^n g_0(T_k) c_{jk}}{\sum_{k=1}^n \hat{g}_n(T_k) c_{jk}}.$$

Note that $\sum_{k=1}^n \hat{g}_n(T_k) c_{jk} \geq m \sum_{k=1}^n \hat{g}_n(T_k)(T_k - T_{k-1}) = m$ by the lower boundedness of w_i and the fact that $\sum_{k=1}^n \hat{g}_n(T_k)(T_k - T_{k-1}) = 1$. Moreover, $\sum_{k=1}^n g_0(T_k) c_{jk} \leq M^2(T_n - T_0) \leq M^2(b-a)$. Thus,

$$\sum_{i=1}^s \frac{n_i}{n} \int_a^b \frac{g_0(x)}{\hat{g}_n(x)} d\mathbb{F}_{i,n_i}(x) \leq \frac{M^2(b-a)}{m}.$$

Hence, for any $0 < \delta < b - a$,

$$\sum_{i=1}^{s} \frac{n_i}{n} \int_{b-\delta}^{b} \frac{g_0(x)}{\hat{g}_n(x)} d\mathbb{F}_{i,n_i}(x) \leq \frac{M^2(b-a)}{m}.$$

Note that, by the monotonicity of $\hat{g}_n$,

$$\frac{1}{\hat{g}_n(b-\delta)} \sum_{i=1}^{s} \frac{n_i}{n} \int_{b-\delta}^{b} g_0(x)\mathbb{F}_{i,n_i}(x) \leq \sum_{i=1}^{s} \frac{n_i}{n} \int_{b-\delta}^{b} \frac{g_0(x)}{\hat{g}_n(x)} d\mathbb{F}_{i,n_i}(x).$$

This implies that, for almost every ω,

$$\begin{aligned} & \limsup_{n\to\infty} \frac{1}{\hat{g}_n(b-\delta;\omega)} \\ \leq \; & \limsup_{n\to\infty} \frac{M^2(b-a)}{m\sum_{i=1}^{s} \frac{n_i}{n} \int_{b-\delta}^{b} g_0(x) d\mathbb{F}_{i,n_i}(x;\omega)} = C_2(\delta), \end{aligned}$$

where $C_2(\delta) \triangleq \frac{M^2(b-a)}{m\sum_{i=1}^{s} \lambda_i \int_{b-\delta}^{b} g_0(x) dF_i(x)} < \infty$.

(iv) Fix $\delta > 0$ such that $a + \delta < \sigma$. Fix a sample point ω such that (i),(ii) and (iii) hold. By (iii), the sequence $\hat{g}_n$ is uniformly bounded on $[a+\delta, b]$ by the constant $1/\delta$. In the rest of this proof, for the sake of simplicity, we suppress the dependence of ω. We first claim that for any converging subsequence $\hat{g}_{n_k}$ of $\hat{g}_n$ on $(a+\delta, b)$, the limit must be g_0. Then by Helly's selection principle, since $\hat{g}_n$ is monotone by definition, we can conclude the whole sequence $\hat{g}_n$ converges to g_0 on $(a+\delta, b)$.

We next prove our claim. Consider any converging subsequence g_{n_k}, by (ii), $\lim_{k\to\infty}\big(\sqrt{\hat{f}_{i,n_k}(x)} - \sqrt{f_i(x)}\big)^2 = \lim_{k\to\infty}\big(\sqrt{\frac{w_i(x)\hat{g}_{n_k}(x)}{\int w_i \hat{g}_{n_k}}} - \sqrt{f_i(x)}\big)^2$ exists. By boundedness of w_i and g_0; and the uniform boundedness of $\hat{g}_n$ on $[a+\delta, b]$, we can apply bounded convergence theorem to obtain

$$\begin{aligned} \int_{a+\delta}^{b} \lim_{k\to\infty} \big(\sqrt{\hat{f}_{i,n_k}(x)} - \sqrt{f_i(x)}\big)^2 dx &= \lim_{k\to\infty} \int_{a+\delta}^{b} \big(\sqrt{\hat{f}_{i,n_k}(x)} - \sqrt{f_i(x)}\big)^2 dx \\ &\leq \lim_{k\to\infty} 2h^2(\hat{f}_{i,n_k}, f_i) = 0, \end{aligned}$$

where the last equality follows from (i). This implies that $\lim_{k\to\infty}(\sqrt{\hat{f}_{i,n_k}(x)} - \sqrt{f_i(x)})^2 = 0$ for Lebesgue-a.e. $x \in [a+\delta, b]$. As $w_i(x) > 0$, we have

$$\frac{\hat{f}_{i,n_k}(x)}{w_i(x)} = \frac{\hat{g}_{n_k}(x)}{\int w_i \hat{g}_{n_k}} \to \frac{f_i(x)}{w_i(x)} = \frac{g_0(x)}{\int w_i g_0} \quad \text{for Lebesgue-a.e. } x \in [a+\delta, b];$$

by (ii) again, we know $\hat{g}_{n_k}(x) \to g_0(x)$ for Lebesgue-a.e. $x \in [a+\delta, b]$. As $\hat{g}_{n_k}$'s are decreasing and g_0 is continuous, we have $\hat{g}_{n_k}(x) \to g_0(x)$ for all $x \in (a+\delta, b)$. Indeed, for an arbitrary $\varepsilon > 0$ and $x_0 \in (a+\delta, b)$, choose $x' < x_0 < x''$ such that $\hat{g}_{n_k}(x') \to g_0(x')$ and $\hat{g}_{n_k}(x'') \to g_0(x'')$ as $k \to \infty$, $g_0(x') > g_0(x_0) > g_0(x'')$, and $|g_0(x'') - g_0(x')| < \varepsilon$. Then, there exists N such that for all $k \geq N$,

$$
\begin{aligned}
& |\hat{g}_{n_k}(x_0) - g_0(x_0)| \\
\leq\ & |\hat{g}_{n_k}(x_0) - \hat{g}_{n_k}(x'')| + |\hat{g}_{n_k}(x'') - g_0(x'')| + |g_0(x'') - g_0(x_0)| \\
\leq\ & |\hat{g}_{n_k}(x') - \hat{g}_{n_k}(x'')| + |\hat{g}_{n_k}(x'') - g_0(x'')| + |g_0(x'') - g_0(x_0)| \\
\leq\ & |\hat{g}_{n_k}(x') - g_0(x')| + |g_0(x') - g_0(x'')| + 2|\hat{g}_{n_k}(x'') - g_0(x'')| \\
& \quad + |g_0(x'') - g_0(x')|,
\end{aligned}
$$

that implies that $\limsup_{k\to\infty} |\hat{g}_{n_k}(x_0) - g_0(x_0)| \leq 2\varepsilon$. Thus, $\hat{g}_{n_k}(x) \to g_0(x)$ for all $x \in (a+\delta, b)$ and the claim follows.
The uniform convergence on compacta, that is, $\sup_{x\in[\sigma,\tau]} |\hat{g}_n(x) - g_0(x)| \to 0$ for $[\sigma, \tau] \subset (a+\delta, b)$, follows from the fact that $\hat{g}_n$'s, g_0 are decreasing and g_0 is continuous; also see, for example, page 1 in Resnick (2013).

□

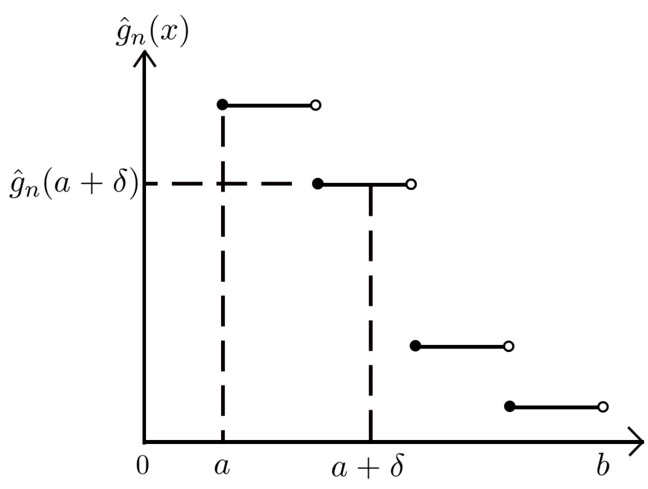

FIG 1. *A graphical illustration that $\hat{g}_(a+\delta)\delta \leq 1$ because of the fact that $\hat{g}_n$ is a density function.*

10. Appendix for Section 5.

PROOF OF LEMMA 5.1. (i) Let $U_i \triangleq F(W_i)$, for $i = 1, \ldots, n$, and thus $U_i \overset{iid}{\sim} \text{Unif}[0,1]$. Let $U_{(h)}$ be the h-th order statistic of $U_1, \ldots, U_n$. The inverse of F is well-defined; especially there is no jump of F. Note that, using the mean value theorem or otherwise,

$$
\begin{aligned}
0 &< W_{(h)} - c = F^{-1}(U_{(h)}) - F^{-1}(0) \\
&\leq \sup_{x\in[c,d]} \frac{1}{f(x)}(U_{(h)} - 0) \leq \frac{1}{m_f} U_{(h)}. \qquad (10.1)
\end{aligned}
$$

Note that we can represent $U_{(h)}$ as

$$U_{(h)} \overset{d}{=} \frac{\sum_{i=1}^{h} E_i}{\sum_{i=1}^{n+1} E_i},$$

where $E_i \overset{iid}{\sim} \text{Exp}(1)$, for $i = 1, \ldots, n$; see for example p.335 in Shorack and Wellner (2009). Hence, we have

$$\max_{h=1,\ldots,n} \frac{nU_{(h)}}{h} \overset{d}{=} \frac{n}{\sum_{i=1}^{n+1} E_i} \max_{h=1,\ldots,n} \frac{\sum_{i=1}^{h} E_i}{h}. \qquad (10.2)$$

By the strong law of large numbers, $\frac{n}{\sum_{i=1}^{n+1} E_i} \to 1$ a.s.. For the second factor on the RHS of (10.2), we first note that $\sum_{i=1}^{h} E_i \sim \text{Gamma}(h,1)$ and the fourth central moment of a $\text{Gamma}(h,1)$ random variable is $3h(2+h)$. Therefore, $\sum_{h=1}^{\infty} \mathbb{E}(\frac{\sum_{i=1}^{h} E_i}{h} - 1)^4 = \sum_{h=1}^{\infty} \frac{3h(2+h)}{h^4} < \infty$. For any $K > 1$, note that $\{\max_{h=1,\ldots,n}(\frac{\sum_{i=1}^{h} E_i}{h} - 1) > K-1\} \subset \{\max_{h=1,\ldots,n}(\frac{\sum_{i=1}^{h} E_i}{h} - 1)^4 \geq (K-1)^4\}$ as $\max_{h=1,\ldots,n}(\frac{\sum_{i=1}^{h} E_i}{h} - 1)^4 \geq (\max_{h=1,\ldots,n}(\frac{\sum_{i=1}^{h} E_i}{h} - 1))^4 \geq (K-1)^4$. Therefore, by Markov's inequality, for any n,

$$
\begin{aligned}
\mathbb{P}\left(\max_{h=1,\ldots,n} \frac{\sum_{i=1}^{h} E_i}{h} > K\right) &\leq \mathbb{P}\left(\max_{h=1,\ldots,n}\left(\frac{\sum_{i=1}^{h} E_i}{h} - 1\right)^4 > (K-1)^4\right) \\
&\leq \frac{\mathbb{E}\left(\max_{h=1,\ldots,n}\left(\frac{\sum_{i=1}^{h} E_i}{h} - 1\right)^4\right)}{(K-1)^4} \\
&\leq \frac{\sum_{h=1}^{\infty} \mathbb{E}\left(\frac{\sum_{i=1}^{h} E_i}{h} - 1\right)^4}{(K-1)^4},
\end{aligned}
$$

Hence, $\max_{h=1,\ldots,n} \frac{\sum_{i=1}^{h} E_i}{h} = O_p(1)$. Therefore, $\max_{h=1,\ldots,n} \frac{nU_{(h)}}{h} = O_p(1)$, and so is $\max_{h=1,\ldots,n} \frac{n(W_{(h)}-c)}{h}$ in view of (10.1).

(ii) Again, let $U_i \triangleq F(W_i)$, for $i = 1, \ldots, n$, and thus $U_i \overset{iid}{\sim}$ Unif$[0,1]$. Let $U_{(h)}$ be the h-th order statistic of $U_1, \ldots, U_n$. Note that in accordance to the mean value theorem or otherwise, we have

$$W_{(h)} - c = F^{-1}(U_{(h)}) - F^{-1}(0) \geq \inf_{x\in[c,d]} \frac{1}{f(x)}(U_{(h)} - 0) \geq \frac{1}{M_f} U_{(h)}.$$

Hence,

$$M_f \left(\min_{h=1,\ldots,n} \frac{nU_{(h)}}{h} \right)^{-1} \geq \left(\min_{h=1,\ldots,n} \frac{n(W_{(h)} - c)}{h} \right)^{-1}. \tag{10.3}$$

As in (a), we represent $U_{(h)}$ as $U_{(h)} \overset{d}{=} \frac{\sum_{i=1}^h E_i}{\sum_{i=1}^{n+1} E_i}$, where $E_i \overset{iid}{\sim}$ Exp(1), for $i = 1, \ldots, n$. Hence, we have

$$\min_{h=1,\ldots,n} \frac{nU_{(h)}}{h} \overset{d}{=} \frac{n}{\sum_{i=1}^{n+1} E_i} \min_{h=1,\ldots,n} \frac{\sum_{i=1}^h E_i}{h}. \tag{10.4}$$

By the strong law of large numbers, $\frac{n}{\sum_{i=1}^{n+1} E_i} \to 1$ a.s.. We first claim that $\big(\min_{h=1,\ldots,n} \frac{\sum_{i=1}^h E_i}{h}\big)^{-1} = O_p(1)$. With this claim, the lemma statement follows in view of (10.3) and (10.4). On the other hand, for the claim, for any $\varepsilon > 0$, since $\sum_{h=1}^\infty \mathbb{E}\big(\frac{\sum_{i=1}^h E_i}{h} - 1\big)^4 < \infty$, there exists N such that $\sum_{h=N}^\infty \mathbb{E}\big(\frac{\sum_{i=1}^h E_i}{h} - 1\big)^4 < \varepsilon$. Next, we choose $\delta \in (0, \frac{1}{2})$ such that $\sum_{h=1}^N \mathbb{P}\big(\frac{\sum_{i=1}^h E_i}{h} < \delta\big) < \varepsilon$, which is possible as there are only finitely many of these probabilities of continuous random variables. Note that

$$\begin{aligned}
&\mathbb{P}\left(\left(\min_{h=1,\ldots,n} \frac{\sum_{i=1}^h E_i}{h}\right)^{-1} > \frac{1}{\delta}\right) = \mathbb{P}\left(\min_{h=1,\ldots,n} \frac{\sum_{i=1}^h E_i}{h} < \delta\right) \\
= \;& \mathbb{P}\left(\bigcup_{h=1,\ldots,n} \left\{\frac{\sum_{i=1}^h E_i}{h} < \delta\right\}\right) \leq \sum_{h=1}^\infty \mathbb{P}\left(\frac{\sum_{i=1}^h E_i}{h} < \delta\right) \\
\leq \;& \varepsilon + \sum_{h=N}^\infty \mathbb{P}\left(\left(1 - \frac{\sum_{i=1}^h E_i}{h}\right)^4 > (1-\delta)^4\right) \\
\leq \;& \varepsilon + \sum_{h=N}^\infty \frac{\mathbb{E}(1 - \frac{\sum_{i=1}^h E_i}{h})^4}{(1-\delta)^4} \\
\leq \;& \varepsilon + \frac{\varepsilon}{(1-\delta)^4} \leq \varepsilon + 2^4\varepsilon = 17\varepsilon,
\end{aligned}$$

where the second inequality follows from the choice of N and note that whenever $\delta < 1$, $\{0 < \frac{\sum_{i=1}^h E_i}{h} < \delta\} = \{\frac{\sum_{i=1}^h E_i}{h} - 1 < \delta - 1 < 0\} \subset$

$\{(1 - \frac{\sum_{i=1}^{h}}{h})^4 > (1-\delta)^4\}$; the third inequality follows from Markov's inequality; the fourth inequality follows from the choice of N; and the last inequality follows as $\delta < \frac{1}{2}$.

$\square$

PROOF OF COROLLARY 5.2. (i) Let $Y_{ij} \triangleq -X_{ij}$ for $i = 1, \ldots, s, j = 1, \ldots, n_i$ so that the distribution functions of Y_{ij}'s are $F_{Y_i}(x) = 1 - F_i(-x)$ and Y_{ij}'s are supported on $[-b, -a]$. Let $S_1 \leq S_2 \leq \ldots \leq S_n$ be the order statistics of all the Y_{ij}'s, $i = 1, \ldots, s, j = 1, \ldots, n_i$. Clearly, by definition, as a mirror image, $T_n - T_{n-h} = S_{h+1} - S_1$ for all $h = 1, 2, \ldots, n-1$. Consider only those $h \in \{n_1 + 1, \ldots, n\}$, note that we have an immediate result:

$$\frac{n(S_{h+1} - S_1)}{h} \leq \frac{n(b-a)}{n_1} \to \frac{b-a}{\lambda_1}.$$

Therefore, for large enough n,

$$\begin{aligned}
\max_{h=1,\ldots,n} \frac{n(T_n - T_{n-h})}{h} &= \max_{h=1,\ldots,n} \frac{n(S_{h+1} - S_1)}{h} \\
&\leq \frac{2(b-a)}{\lambda_1} + \max_{h=1,\ldots,n_1} \frac{n(S_{h+1} - S_1)}{h} \\
&\leq \frac{2(b-a)}{\lambda_1} + \max_{h=1,\ldots,n_1} \frac{n(Y_{1(h+1)} - (-b))}{h} \\
&\leq \frac{2(b-a)}{\lambda_1} + \max_{h=1,\ldots,n_1} \frac{2n(Y_{1(h+1)} - (-b))}{h+1},
\end{aligned}$$

where $Y_{1(j)}$ is the j-th order statistics of $Y_{11}, \ldots, Y_{1n_1}$. Since the density of Y_1, $f_{Y_1}(y) = f_1(-y) \geq m$ for $y \in [-b, -a]$, the result follows by applying Lemma 5.1.

(ii) Let $U_{ij} \triangleq F_j(X_{ij})$ for $i = 1, \ldots, s, j = 1, \ldots, n_i$. Note that $U_{ij} \overset{iid}{\sim}$ Unif$[0, 1]$. Let $F_{\max}(x) \triangleq \max_{i=1,\ldots,s} F_i(x)$. Clearly, $F_{\max}$ is a distribution function on $[a, b]$ with a density function $f_{\max}$ that satisfies $0 < m \leq f_{\max} \leq M$. Note that $U_{ij} \leq F_{\max}(X_{ij})$. Hence, for any $h = 1, \ldots, n$, $U_{(h)} \leq F_{\max}(T_h)$. Therefore, $F_{\max}^{-1}(U_{(h)}) \leq T_h$. As a result, $\min_{h=1,\ldots,n} \frac{n(F_{\max}^{-1}(U_{(h)}) - a)}{h} \leq \min_{h=1,\ldots,n} \frac{n(T_h - a)}{h}$, and so

$$\left(\min_{h=1,\ldots,n} \frac{n(T_h - a)}{h}\right)^{-1} \leq \left(\min_{h=1,\ldots,n} \frac{n(F_{\max}^{-1}(U_{(h)}) - a)}{h}\right)^{-1} = O_p(1),$$

where the tightness follows from Lemma 5.1 with $W_i \triangleq F_{\max}^{-1}(U_i) \sim F_{\max}$.

$\square$

PROOF OF (5.2). Let $\Delta_i \triangleq T_i - T_{i-1}$ for $i = 1, \dots, n$ and $\mathcal{C}_j \triangleq \sum_{k=1}^{n} c_{jk} = \int_a^{T_n} w_j(x)dx$; the dependences on n of Δ_i and $\mathcal{C}_j$ are suppressed in the notation for simplicity. Suppose that $\boldsymbol{z}^*$ maximizes $\mathcal{L}_n(\boldsymbol{z})$ subject to $z_1 \geq \dots \geq z_n \geq 0$. Then, by the homogeneity of degree 0 of $\mathcal{L}_n(\boldsymbol{z})$, we can normalize $\boldsymbol{z}^*$ to obtain the maximizer of $\mathcal{L}_n(z)$ subject to $z_1 \geq \dots z_n \geq 0$ and $\sum_{i=1}^{n} z_i \Delta_i = 1$ simultaneously. Also recall that all the components of the maximizer should be positive, i.e., $\hat{\boldsymbol{z}} > 0$. Therefore, we can ignore the equality constraint and also consider the ratio $\frac{z_i}{z_n}$ for $i = 1, \dots, n-1$ instead. Then we can write $\mathcal{L}_n(\boldsymbol{z})$ as

$$\mathcal{L}_n(\boldsymbol{z}) = \frac{\frac{z_1}{z_n}\frac{z_2}{z_n}\dots\frac{z_{n-1}}{z_n}}{\prod_{j=1}^{s}(\sum_{k=1}^{n-1} c_{jk}\frac{z_k}{z_n} + c_{jn})^{n_j}} = \frac{\frac{z_1}{z_n}\frac{z_2}{z_n}\dots\frac{z_{n-1}}{z_n}}{\prod_{j=1}^{s}(\sum_{k=1}^{n-1} c_{jk}(\frac{z_k}{z_n} - 1) + \mathcal{C}_j)^{n_j}}.$$

Define $y_i \triangleq \frac{z_i}{z_n} - 1$. The original optimization problem is equivalent to maximize the following alternative objective function:

$$\mathbb{L}_n(\boldsymbol{y}) \triangleq \frac{\prod_{i=1}^{n-1}(1 + y_i)}{\prod_{j=1}^{s}(\mathcal{C}_j + \sum_{k=1}^{n-1} c_{jk} y_k)^{n_j}},$$

subject to $y_1 \geq y_2 \geq \dots \geq y_{n-1} \geq 0$, without bothering the equality constraint.

Define $\tilde{\mu}_0 \triangleq 0$, Karush-Kuhn-Tucker (KKT) conditions for minimizing $-\log \mathbb{L}_n$ subject to $y_1 \geq y_2 \geq \dots y_{n-1} \geq 0$ imply that there exist $\tilde{\mu}_1, \dots, \tilde{\mu}_{n-1}$ such that the optimal solution $\hat{\boldsymbol{y}}$ satisfies

(10.5)
$$-\frac{1}{1 + \hat{y}_l} + \sum_{j=1}^{s} \frac{n_j c_{jl}}{\mathcal{C}_j + \sum_{k=1}^{n-1} c_{jk}\hat{y}_k} - \tilde{\mu}_l + \tilde{\mu}_{l-1} = 0, \quad \text{for } l = 1, 2, \dots, n-1,$$

such that

$$\begin{aligned} &(10.6) & \tilde{\mu}_l(\hat{y}_{l+1} - \hat{y}_l) &= 0, \quad \text{for } l = 1, 2, \dots, n-2, \\ &(10.7) & -\tilde{\mu}_{n-1}\hat{y}_{n-1} &= 0 \\ &(10.8) & \tilde{\mu}_1, \tilde{\mu}_2, \dots, \tilde{\mu}_{n-1} &\geq 0, \end{aligned}$$

where $\tilde{\mu}_1, \dots, \tilde{\mu}_{n-1}$ are the KKT multipliers. We can have two scenarios:

Case 1: $\hat{y}_1 = \dots = \hat{y}_{n-1}$. Clearly, we have $\hat{z}_1 = \dots = \hat{z}_{n-1}$ and the equality constraint $\sum_{i=1}^{n} \hat{z}_i \Delta_i = 1$ implies that

$$\hat{z}_1 = \frac{1 - \hat{z}_n \Delta_n}{\sum_{i=1}^{n-1} \Delta_i} \leq \frac{1}{T_{n-1} - a}.$$

Case 2: there exists a $j^* \in \{1, \ldots, n-1\}$ such that $\hat{y}_1 = \ldots = \hat{y}_{j^*} > \hat{y}_{j^*+1}$, where $\hat{y}_n \triangleq 0$. The strict inequality implies $\tilde{\mu}_{j^*} = 0$; next adding up the equations from (10.5) for $l = 1$ to $l = j^*$, we obtain

$$-\frac{j^*}{1+\hat{y}_1} + \sum_{j=1}^{s} \frac{n_j \sum_{l=1}^{j^*} c_{jl}}{\mathcal{C}_j + \sum_{k=1}^{n-1} c_{jk}\hat{y}_k} = 0,$$

where the first term in the LHS is obtained from the fact that $\hat{y}_1 = \ldots = \hat{y}_{j^*}$. Using the boundedness assumptions on w_j's, we obtain

$$\frac{j^*}{1+\hat{y}_1} \geq \frac{m}{M} \sum_{j=1}^{s} \frac{n_j (T_{j^*} - a)}{(T_n - a) + \sum_{k=1}^{n-1} \Delta_k \hat{y}_k}.$$

Hence, we have

$$1 + \hat{y}_1 \leq \frac{M}{m} \frac{(T_n - a) + \sum_{k=1}^{n-1} \Delta_k \hat{y}_k}{n \frac{T_{j^*} - a}{j^*}}. \tag{10.9}$$

It is noted that by what times the equality constraints are active among the components of the optimal $\hat{\boldsymbol{y}}$ may not be explicitly known. Nevertheless, we can still provide a sensible bound for $1 + \hat{y}_1$ based on (10.9), which is

$$1 + \hat{y}_1 \leq \frac{M}{m} \frac{(T_n - a) + \sum_{k=1}^{n-1} \Delta_k \hat{y}_k}{n \min_{j=1,\ldots,n} \frac{T_j - a}{j}}.$$

Combining Case 1 and Case 2, we obtain a uniform bound for $\hat{z}_1$ as

$$\hat{z}_1 \leq \hat{z}_n \frac{M}{m} \frac{(T_n - a) + \sum_{k=1}^{n-1} \Delta_k \hat{y}_k}{n \min_{j=1,\ldots,n} \frac{T_j - a}{j}} + \frac{1}{T_{n-1} - a}. \tag{10.10}$$

In view of (10.10), we shall now provide an upper bound for the term $(T_n - a) + \sum_{k=1}^{n-1} \Delta_k \hat{y}_k$. To this end, for each $h = 1, \ldots, n-1$, we consider maximizing

$$\mathbb{L}_n^h(y_1, \ldots, y_{n-h}) \triangleq \frac{\prod_{i=1}^{n-h}(1 + y_i)}{\prod_{j=1}^{s}(\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk} y_k)^{n_j}} \tag{10.11}$$

subject to $y_1 \geq \ldots y_{n-h} \geq 0$. Define $\tilde{\mu}_0^h \triangleq 0$. As in (10.5) to (10.8), there exist $\tilde{\mu}_1^h, \ldots, \tilde{\mu}_{n-h}^h$ such that the optimal solution $\hat{\boldsymbol{y}}^h$ satisfies

(10.12)

$$-\frac{1}{1+\hat{y}_l^h} + \sum_{j=1}^{s} \frac{n_j c_{jl}}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk} \hat{y}_k^h} - \tilde{\mu}_l^h + \tilde{\mu}_{l-1}^h \quad = \quad 0, \quad \text{for } l = 1, 2, \ldots, n-h,$$

such that

$$\tilde{\mu}_l^h(\hat{y}_{l+1}^h - \hat{y}_l^h) = 0, \quad \text{for } l = 1, 2, \ldots, n-h-1, \tag{10.13}$$

$$-\tilde{\mu}_{n-h}^h \hat{y}_{n-h}^h = 0 \tag{10.14}$$

$$\tilde{\mu}_1^h, \tilde{\mu}_2^h, \ldots, \tilde{\mu}_{n-h}^h \geq 0. \tag{10.15}$$

Observe that if we multiply $(1+\hat{y}_l^h)$ on both sides of (10.12) for each $l = 1, \ldots, n-h$, and add up all the resulting equations, we obtain

$$\sum_{l=1}^{n-h}\sum_{j=1}^{s} \frac{n_j c_{jl}(1+\hat{y}_l^h)}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk}\hat{y}_k^h} - (n-h) = \tilde{\mu}_{n-h}^h, \tag{10.16}$$

by using (10.13) and (10.14). Note that we can simplify the expression on the LHS of (10.16),

$$\begin{aligned}
&\sum_{l=1}^{n-h}\sum_{j=1}^{s} \frac{n_j c_{jl}(1+\hat{y}_l^h)}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk}\hat{y}_k^h} - (n-h) \\
=\ & \sum_{j=1}^{s} n_j \frac{\sum_{l=1}^{n-h} c_{jl} + \sum_{l=1}^{n-h} c_{jl}\hat{y}_l^h}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk}\hat{y}_k^h} - (n-h) \\
=\ & \sum_{j=1}^{s} n_j \left(1 - \frac{\sum_{l=n-h+1}^{n} c_{jl}}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk}\hat{y}_k^h}\right) - (n-h) \\
=\ & h - \sum_{j=1}^{s} \frac{n_j \sum_{l=n-h+1}^{n} c_{jl}}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk}\hat{y}_k^h} = \tilde{\mu}_{n-h}^h.
\end{aligned} \tag{10.17}$$

Returning to the maximizer $\hat{\boldsymbol{y}}$, we have again two scenarios:

Case 1': $\hat{y}_1 = \ldots = \hat{y}_{n-1} = 0$. Then $(T_n - a) + \sum_{k=1}^{n-1} \Delta_k \hat{y}_k = T_n - a$.

Case 2': there exists a $h \in \{1, \ldots, n-1\}$ such that $\hat{y}_{n-1} = \ldots = \hat{y}_{n-h+1} = 0$ and $\hat{y}_{n-h} > 0$. To find the optimal solution $\hat{\boldsymbol{y}}$ in this case, it suffices to maximize the reduced objective function $\mathbb{L}_n^h(y_1, \ldots, y_{n-h})$ defined in (10.11) subject to $y_1 \geq \ldots y_{n-h} \geq 0$. Note that the optimal value of $\mathbb{L}_n^h$ and that of the original objective function $\mathbb{L}_n$ are the same; meanwhile, the optimal points $\hat{y}_i^h$ for $\mathbb{L}_n^h$ are also the same as that of the original objective function $\hat{y}_i$, i.e., $\hat{y}_i^h = \hat{y}_i$ for $i = 1, \ldots, n-h$. Hence, we get from (10.17) that

$$h - \sum_{j=1}^{s} \frac{n_j \sum_{l=n-h+1}^{n} c_{jl}}{\mathcal{C}_j + \sum_{k=1}^{n-h} c_{jk}\hat{y}_k} = \tilde{\mu}_{n-h}^h = 0,$$

as $\hat{y}^h_{n-h} = \hat{y}_{n-h} > 0$ implies $\tilde{\mu}^h_{n-h} = 0$. Using the boundedness assumptions on w_j's, we have

$$\frac{M}{m}\sum_{j=1}^{s}\frac{n_j(T_n - T_{n-h})}{(T_n - a) + \sum_{k=1}^{n-h}\Delta_k \hat{y}_k} \geq h.$$

Hence, we obtain

$$(T_n - a) + \sum_{k=1}^{n-h}\Delta_k \hat{y}_k \leq \frac{M}{m}\frac{n(T_n - T_{n-h})}{h}.$$

As before, by what components of the optimal $\hat{\boldsymbol{y}}$ vanish may not be explicitly known. Nevertheless, based on the above arguments, we can still provide a bound for $(T_n - a) + \sum_{k=1}^{n-1}\Delta_k \hat{y}_k$, which is

$$(T_n - a) + \sum_{k=1}^{n-1}\Delta_k \hat{y}_k \leq \frac{M}{m}\max_{h=1,\dots,n-1}\frac{n(T_n - T_{n-h})}{h}.$$

Combining Case 1' and Case 2', we have a uniform bound

$$(T_n - a) + \sum_{k=1}^{n-1}\Delta_k \hat{y}_k \leq \frac{M}{m}\max_{h=1,\dots,n}\frac{n(T_n - T_{n-h})}{h}. \tag{10.18}$$

From (10.10) and (10.18), we establish (5.2). $\square$

PROOF OF LEMMA 5.4. (i) From equation (9.1), we know

$$\sum_{i=1}^{s} h^2(\hat{f}_{i,n}, f_i) \leq \sum_{i=1}^{s}\int_a^b \frac{2\hat{f}_{i,n}(x_i)}{\hat{f}_{i,n}(x_i) + f_i(x_i)} d(\mathbb{F}_i - F_i)(x_i).$$

and the class of functions $\frac{\hat{f}_{i,n}}{\hat{f}_{i,n}+f_{i,n}}$ is also a subset of a Donsker class in view of (9.2), (9.3) and (9.4). Hence, the RHS of the inequality is of order $O_p(n^{-1/2})$, and the claim follows.

(ii) This is a direct consequence of Lemma 5.4 (i); by using the expression in the proof in Proposition 4.1 (ii) and $\int w_i \hat{g}_n = O_p(1)$ in light of Proposition 4.1 (ii) again.

(iii) By Lemma 5.3,

$$\begin{aligned} &\int_a^b (\hat{f}_{i,n}(x) - f_i(x))^2 dx \\ = &\int_a^b (\sqrt{\hat{f}_{i,n}(x)} - \sqrt{f_i(x)})^2 (\sqrt{\hat{f}_{i,n}(x)} + \sqrt{f_i(x)})^2 dx \\ \leq\ & O_p(1)\int_a^b (\sqrt{\hat{f}_{i,n}(x)} - \sqrt{f_i(x)})^2 dx = O_p(n^{-1/2}). \end{aligned}$$

(iv) Using Lemma 5.3 again, Lemma 5.4 (ii) and (iii),

$$\begin{aligned}\int_a^b (\hat{g}_n(x) - g_0(x))^2 dx &= \int_a^b \left(\frac{\hat{f}_{i,n}(x) \int w_i \hat{g}_n}{w_i(x)} - \frac{f_i(x) \int w_i g_0}{w_i(x)} \right)^2 dx \\ \leq\ & \frac{1}{m^2} \int_a^b \left(\hat{f}_{i,n}(x) \int w_i \hat{g}_n - f_i(x) \int w_i g_0 \right)^2 dx \\ \leq\ & \frac{2}{m^2} \int_a^b (\hat{f}_{i,n}(x) - f_i(x))^2 dx \cdot \left(\int w_i \hat{g}_n \right)^2 + \\ & \frac{2}{m^2} \int_a^b f_i^2(x) dx \cdot \left(\int w_i g_0 - \int w_i \hat{g}_n \right)^2 \\ =\ & O_p(n^{-1/2}).\end{aligned}$$

$\square$

PROOF OF LEMMA 5.5. The proof follows from the Karush-Kuhn-Tucker conditions. Recall that in the proof of Proposition 3.2, we have for all $i = 1, \dots, n-1$,

$$\mu_i = \sum_{j=1}^{i} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) \geq 0, \quad \sum_{j=1}^{n} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) = 0, \quad \mu_i(\hat{z}_{i+1} - \hat{z}_i) = 0.$$

Therefore, whenever $\hat{z}_{k+1} > \hat{z}_k$ for some k, $\mu_k = \sum_{j=1}^{k} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) = 0$. Now, denote $\tau_1, \dots, \tau_m$ to be the jump points of $\hat{g}_n$ for some $m + 1 \leq n$. Define also $\tau_0 \triangleq T_1$ and $\tau_{m+1} \triangleq T_n$; note that τ_m can be equal to τ_{m+1}. We then have

$$\sum_{i:\tau_j \leq T_i < \tau_{j+1}} \frac{\partial \psi_n}{\partial z_i}(\hat{z}) = \sum_{i:T_i < \tau_{j+1}} \frac{\partial \psi_n}{\partial z_i}(\hat{z}) - \sum_{i:T_i < \tau_j} \frac{\partial \psi_n}{\partial z_i}(\hat{z}) = 0, \quad j = 0, \dots, m.$$

Hence, for any function γ and for $j = 0, \dots, m$, as there is no change of value of $\hat{g}_n$ on $[\tau_j, \tau_{j+1})$,

$$0 = \gamma(\hat{g}_n(\tau_j)) \sum_{i:\tau_j \leq T_i < \tau_{j+1}} \frac{\partial \psi_n}{\partial z_i}(\hat{z}) = \sum_{i:\tau_j \leq T_i < \tau_{j+1}} \left(\frac{\partial \psi_n}{\partial z_i}(\hat{z}) \gamma(\hat{g}_n(T_i)) \right).$$

Summing the above equation from $j = 0$ to $j = m$ completes the proof. $\square$

PROOF OF LEMMA 5.9. Let $T_j^1, j = 1, \dots, n_1$, be the order statistics from the first sample $X_{1j}, j = 1, \dots, n_1$, and $\hat{F}_1$ be the empirical distribution of F_1 from $X_{11}, \dots, X_{1n_1}$. We first claim that

$$\sup_{j=1,\dots,n} (T_j - T_{j-1}) = O_p(n^{-1/2}), \tag{10.19}$$

and

$$\int_{T_n}^{b} g_0(x)w_i(x)dx = O_p(n^{-1}). \tag{10.20}$$

Indeed, by the mean value theorem, for some $\eta_j \in (T_{j-1}^1, T_j^1)$,

$$\begin{aligned} T_j^1 - T_{j-1}^1 &= \frac{1}{F_1'(\eta_j)}\Big[F_1(T_j^1) - \hat{F}_1(T_j^1) + \hat{F}_1(T_{j-1}^1) - F_1(T_{j-1}^1) + \hat{F}_1(T_j^1) - \hat{F}_1(T_{j-1}^1)\Big] \\ &\le \frac{1}{f_1(\eta_j)}\Big[|\hat{F}_1(T_j^1) - F_1(T_j^1)| + |\hat{F}_1(T_{j-1}^1) - F_1(T_{j-1}^1)| + \frac{1}{n_1}\Big]. \end{aligned}$$

Hence, under Assumptions 2.1 (C) and by applying the Dvoretzky-Kiefer-Wolfowitz inequality,

$$\begin{aligned} &\sup_{j=1,\ldots,n}(T_j - T_{j-1}) \le \sup_{j=1,\ldots,n_1}(T_j^1 - T_{j-1}^1) \\ \le\ & 2\sup_{t\in[a,b]}\left(\frac{1}{f_1(t)}\right)\sup_{x\in[a,b]}|\hat{F}_1(x) - F_1(x)| + O(n^{-1}) \\ =\ & O_p(n^{-1/2}) + O(n^{-1}) = O_p(n^{-1/2}). \end{aligned}$$

On the other hand, for establishing (10.20), note that under Assumptions 2.1 (C), $\int_{T_n}^b g_0(x)w_i(x)dx \le M^2(b - T_n) \le M^2(b - T_{n_1}^1) = O_p(n^{-1})$; see also the proof of Lemma 5.6. Using (10.19), (10.20) and Assumptions 2.1 (B),

$$\begin{aligned} &\sum_{i=1}^{n} g_0(T_j)\int_{T_{j-1}}^{T_j} w_i(x)dx - \int_a^b g_0(x)w_i(x)dx \\ =\ & \sum_{i=1}^{n} g_0(T_j)\int_{T_{j-1}}^{T_j} w_i(x)dx - \sum_{i=1}^{n}\int_{T_{j-1}}^{T_j} g_0(x)w_i(x)dx - \int_{T_n}^{b} g_0(x)w_i(x)dx \\ =\ & \sum_{i=1}^{n}\int_{T_{j-1}}^{T_j} w_i(x)g_0'(\xi_{j,x})(T_j - x)dx + O_p(n^{-1}) \\ \le\ & \sup_{t\in(a,b)}|g_0'(t)|\sup_{j=1,\ldots,n}(T_j - T_{j-1})\int_a^b w_i(x)dx + O_p(n^{-1}) = O_p(n^{-1/2}), \end{aligned}$$

for some $\xi_{j,y} \in [T_{j-1}, T_j]$ for each $y \in [a, b]$ in light of the mean value theorem. □

PROOF OF LEMMA 5.11. Denote $J_i(t) \triangleq \sum_{j:t\le X_{ij}<t_0}\left(-\frac{1}{n_i g_0(t_0)} + \frac{\sum_{j:t\le T_j<t_0} c_{ij}}{\int w_i\hat{g}_n}\right)$. Then, for any $t < t_0$,

$$\sum_{j:t\le T_j<t_0}\left(-\frac{1}{n g_0(t_0)} + \sum_{i=1}^{s}\frac{n_i}{n}\frac{c_{ij}}{\int w_i\hat{g}_n}\right) = \sum_{i=1}^{s}\frac{n_i}{n}J_i(t).$$

We first claim that $\mathbb{P}(\sup_{t\in I_n} J_i(t)\geq 0)\leq\varepsilon$ for all $i=1,\dots,s$. The claim of our present lemma then follows as $\mathbb{P}(\sup_{t\in I_n}\sum_{i=1}^s \frac{n_i}{n}J_i(t)\geq 0)\leq\mathbb{P}(\sum_{i=1}^s\frac{n_i}{n}\sup_{t\in I_n}J_i(t)\geq 0)\leq\sum_{i=1}^s\mathbb{P}(\sup_{t\in I_n}J_i(t)\geq 0)\leq s\varepsilon$. Next, we verify the first claim. For the ease of notation, we suppress the subscript i if there is no ambiguity. Define

$$\begin{aligned}
\mathcal{A}(t) &\triangleq -\int\frac{1}{g_0(t_0)}1(t\leq x<t_0)dF(x)+\frac{\int_t^{t_0}w(x)dx}{\int wg_0},\\
\mathcal{B}(t) &\triangleq -\int\frac{1}{g_0(t_0)}1(t\leq x<t_0)d(\mathbb{F}_{n_i}-F)(x),\\
\mathcal{C}(t) &\triangleq \frac{\int_{T_L}^t w(x)dx}{\int w\hat{g}_n}-\frac{\int_{T_U}^{t_0}w(x)dx}{\int w\hat{g}_n},\\
\mathcal{D}(t) &\triangleq \frac{\int_t^{t_0}w(x)dx}{\int w\hat{g}_n}-\frac{\int_t^{t_0}w(x)dx}{\int wg_0},
\end{aligned}$$

where T_L and T_U denote the maxima of X_{ij} less than t and t_0 respectively. Telescoping the terms, we obtain $J(t)=\mathcal{A}(t)+\mathcal{B}(t)+\mathcal{C}(t)+\mathcal{D}(t)$.

(i) Note that, since $g_0'(t_0)<0$,

$$\begin{aligned}
\mathcal{A}(t) &= -\int_t^{t_0}\frac{g_0(x)-g_0(t_0)}{g_0(t_0)}\cdot\frac{w(t_0)}{\int wg_0}dx-\int_t^{t_0}\frac{g_0(x)-g_0(t_0)}{g_0(t_0)}\cdot\frac{w(x)-w(t_0)}{\int wg_0}dx\\
&= \left(-\int_t^{t_0}\frac{g_0'(t_0)(x-t_0)}{g_0(t_0)}\cdot\frac{w(t_0)}{\int wg_0}dx+o(|t-t_0|^2)\right)+O(|t-t_0|^3)\\
&= -\frac{w(t_0)|g_0'(t_0)|}{g_0(t_0)\int wg_0}\frac{(t-t_0)^2}{2}+o(|t-t_0|^2).
\end{aligned}$$

Hence, there exists R_0 such that for all $t\in(t_0-R_0,t_0)$, we have $\mathcal{A}(t)\leq-\alpha(t-t_0)^2$, where $\alpha\triangleq\frac{1}{4}\frac{|g_0'(t_0)|w(t_0)}{g_0(t_0)\int wg_0}>0$.

(ii) For $\mathcal{B}(t)$, by the same argument as used in the proof of Lemma 4.1 of Kim and Pollard (1990), for any $\delta>0$, there exists a tight sequence of random variables $\{M_n\}$ such that for any $t\in(t_0-R_0,t_0)$,

$$\left|-\int\frac{1}{g_0(t_0)}1(t\leq x<t_0)d(\mathbb{F}_{n_i}-F)(x)\right|\leq\delta(t_0-t)^2+n^{-2/3}M_n^2.$$

(iii) For $\mathcal{C}(t)$, it is clear that the terms $\frac{\int_{T_L}^t w(x)dx}{\int w\hat{g}_n}$ and $\frac{\int_{T_U}^{t_0}w(x)dx}{\int w\hat{g}_n}$ are of order $O_p(n^{-1})$, which are also independent of the choice of $t\in(t_0-R_0,t_0)$; see also the proof of Lemma 5.6.

(iv) For $\mathcal{D}(t)$, by Proposition 5.8,

$$\frac{\int_t^{t_0}w(x)dx}{\int w\hat{g}_n}-\frac{\int_t^{t_0}w(x)dx}{\int wg_0}=O_p(n^{-1/2})O(|t_0-t|)=O_p(n^{-1/2})|t_0-t|.$$

Now, for $t \in (t_0 - R_0, t_0)$, by choosing $\delta = \alpha/2 > 0$,

$$\begin{aligned} J(t) &\leq -\alpha(t-t_0)^2 + \delta(t_0-t)^2 + n^{-2/3}M_n^2 + O_p(n^{-1/2})|t_0 - t| \\ &= \Big(-\frac{\alpha}{2} + O_p(n^{-1/2})|t_0-t|^{-1}\Big)(t_0-t)^2 + n^{-2/3}M_n^2; \end{aligned}$$

as a consequence, with probability more than $1-\varepsilon/2$, there exists $D > 0$ such that for any $C > 0$, any $t \in (t_0 - R_0, t_0 - Cn^{-1/3}]$, and all sufficiently large n,

$$J(t) \leq n^{-2/3}\Big(-\frac{\alpha}{2}C^2 + n^{-1/6}DC + M_n^2\Big).$$

By choosing a $C_0 > 0$ large enough, we can guarantee that with probability more than $1-\frac{\varepsilon}{2}$, $-\frac{\alpha}{2}C^2 + n^{-1/6}DC + M_n^2 < 0$ for all $C \geq C_0$ as $M_n^2 = O_p(1)$. Since the above bound for $J(t)$ is independent of t, we know that there exists C_0 such that for all $C \geq C_0$, $R \leq R_0$, large enough n,

$$\mathbb{P}\Big(\sup_{t\in(t_0-R, t_0-Cn^{-1/3}]} J(t) < 0\Big) \geq 1-\varepsilon,$$

and our claim follows. $\square$

PROOF OF LEMMA 5.12. Fix $C > 0$ and define $t_n = t_0 - 2Cn^{-1/3}$. Denote $\tilde{J}_i(t) \triangleq \sum_{j:t_n\leq X_{ij}<t}\Big(-\frac{1}{n_i g_0(t_n)} + \frac{\sum_{j:t_n\leq T_j<t} c_{ij}}{\int w_i \hat{g}_n}\Big)$. Then, for $t > t_n$,

$$\sum_{j:t_n\leq T_j<t}\Big(-\frac{1}{ng_0(t_n)} + \sum_{i=1}^{s}\frac{n_i}{n}\frac{c_{ij}}{\int w_i\hat{g}_n}\Big) = \sum_{i=1}^{s}\frac{n_i}{n}\tilde{J}_i(t).$$

Following the same argument as that in Lemma 5.11, it suffices to show that $\mathbb{P}(\inf_{t\in\tilde{I}_n}\tilde{J}_i(t) \leq 0) \leq \varepsilon$. For the notational simplicity, we suppress the subscript i if there is no ambiguity. Define

$$\begin{aligned} \tilde{\mathcal{A}}(t) &\triangleq -\int \frac{1}{g_0(t_n)}1(t_n \leq x < t)dF(x) + \frac{\int_{t_n}^{t} w(x)dx}{\int wg_0}, \\ \tilde{\mathcal{B}}(t) &\triangleq -\int \frac{1}{g_0(t_n)}1(t_n \leq x < t)d(\mathbb{F}_{n_i} - F)(x), \\ \tilde{\mathcal{C}}(t) &\triangleq \frac{\int_{T_L}^{t_n} w(x)dx}{\int w\hat{g}_n} - \frac{\int_{T_U}^{t} w(x)dx}{\int w\hat{g}_n}, \\ \tilde{\mathcal{D}}(t) &\triangleq \frac{\int_{t_n}^{t} w(x)dx}{\int w\hat{g}_n} - \frac{\int_{t_n}^{t} w(x)dx}{\int wg_0}, \end{aligned}$$

where T_L and T_U denote the maxima of X_{ij} less than t_n and t respectively. Telescoping the terms, we obtain $\tilde{J}(t) = \tilde{\mathcal{A}}(t) + \tilde{\mathcal{B}}(t) + \tilde{\mathcal{C}}(t) + \tilde{\mathcal{D}}(t)$.

(i) Note that, since $g_0'(t_n) < 0$,

$$\begin{aligned}
\tilde{\mathcal{A}}(t) &= \int_{t_n}^{t} \frac{g_0(t_n) - g_0(x)}{g_0(t_n)} \cdot \frac{w(t_n)}{\int w g_0} dx \\
&\quad + \int_{t_n}^{t} \frac{g_0(t_n) - g_0(x)}{g_0(t_n)} \cdot \frac{w(x) - w(t_n)}{\int w g_0} dx \\
&= \int_{t_n}^{t} \frac{|g_0'(t_n)|(x - t_n)}{g_0(t_n)} \cdot \frac{w(t_n)}{\int w g_0} dx + o(|t - t_n|^2) \\
&\quad + O(|t - t_n|^3) \\
&= \frac{|g_0'(t_n)| w(t_n)}{g_0(t_n) \int w g_0} \frac{(t - t_n)^2}{2} + o(|t - t_n|^2).
\end{aligned}$$

Note that t_n converges to t_0 as n goes to infinity. Therefore, there exists $R_1 > 0$ such that for large enough n and $t \in [t_n, t_0 + R_1)$, $\tilde{\mathcal{A}}(t) \geq \tilde{\alpha}(t - t_n)^2$ for $\tilde{\alpha} \triangleq \frac{1}{4} \sup_{t \in [\frac{t_0}{2}, t_0]} \frac{|g_0'(t)| w(t)}{g_0(t) \int w g_0} > 0$ under Assumptions 2.1 (B) and (C).

(ii) For $\tilde{\mathcal{B}}(t)$, by the same argument as used in the proof of Lemma 4.1 of Kim and Pollard (1990), for all $\delta > 0$, there exists a tight sequence of random variables $\{\tilde{M}_n\}$ such that for all $t \in (t_n, t_0 + R_1)$,

$$\left| - \int \frac{1}{g_0(t_n)} 1(t_n \leq x < t) d(\mathbb{F} - F)(x) \right| \leq \delta (t - t_n)^2 + n^{-2/3} \tilde{M}_n^2.$$

(iii) For $\tilde{\mathcal{C}}(t)$, it is clear that the terms $\frac{\int_{T_L}^{t_n} w(x)dx}{\int w \hat{g}_n}$ and $\frac{\int_{T_U}^{t} w(x)dx}{\int w \hat{g}_n}$ are $O_p(n^{-1})$, not depending on t; see also the proof of Lemma 5.6.

(iv) For $\tilde{\mathcal{D}}(t)$, by Proposition 5.8,

$$\frac{\int_{t_n}^{t} w(x)dx}{\int w \hat{g}_n} - \frac{\int_{t_n}^{t} w(x)dx}{\int w g_0} = O_p(n^{-1/2}) |t - t_n|.$$

Now, for any $t \in [t_0 - Cn^{-1/3}, t_0 + R_1)$, by choosing $\delta = \tilde{\alpha}/2 > 0$,

$$\begin{aligned}
\tilde{J}(t) &\geq \tilde{\alpha}(t - t_n)^2 - \delta(t - t_n)^2 - n^{-2/3} \tilde{M}_n^2 + O_p(n^{-1/2}) |t - t_n| \\
&= \left(\frac{\tilde{\alpha}}{2} + O_p(n^{-1/2}) |t - t_n|^{-1} \right) (t - t_n)^2 - n^{-2/3} \tilde{M}_n^2;
\end{aligned}$$

as a consequence, with probability more than $1 - \varepsilon/2$, there exists $\tilde{D} > 0$ such that for any $C > 0$, any $t \in [t_0 - Cn^{-1/3}, t_0 + R_1)$, and all large enough n,

$$\tilde{J}(t) \geq n^{-2/3} \left(\frac{\tilde{\alpha}}{2} C^2 - \tilde{D} C n^{-1/6} - \tilde{M}_n^2 \right).$$

Hence, with probability more than $1 - \frac{\varepsilon}{2}$, there exists a sufficiently large C_1 such that $\frac{\tilde{\alpha}}{2}C^2 - \tilde{D}Cn^{-1/6} - \tilde{M}_n^2 > 0$ for all $C \geq C_1$ as $\tilde{M}_n^2 = O_p(1)$. Since the above bound for $\tilde{J}(t)$ is not depending on t, we conclude that there exists $C_1 > 0$ such that for all $C \geq C_1$, $R \leq R_1$, large enough n,

$$\mathbb{P}\left(\inf_{t \in [t_0 - Cn^{-1/3}, t_0 + R)} \tilde{J}(t) > 0\right) \geq 1 - \varepsilon,$$

and our claim follows. □

PROOF OF PROPOSITION 5.13. Fix $\varepsilon > 0$. Using the notation in Lemma 5.11 and Lemma 5.12, choose $C \geq \max\{C_0, C_1, M\}$ and $0 < R \leq \min\{R_0, R_1\}$. We first claim that

$$\mathbb{P}\big(\hat{g}_n(t_0 - Cn^{-1/3}) \leq g_0(t_0)\big) \leq 2\varepsilon, \tag{10.21}$$

and

$$\mathbb{P}\big(\hat{g}_n(t_0 - Cn^{-1/3}) \geq g_0(t_0 - 2Cn^{-1/3})\big) \leq 2\varepsilon. \tag{10.22}$$

These together imply that with the probability more than $1 - 4\varepsilon$,

$$g_0(t_0 - 2Cn^{-1/3}) > \hat{g}_n(t_0 - Cn^{-1/3}) > g_0(t_0).$$

Similarly, we have with the probability more than $1 - 4\varepsilon$,

$$g_0(t_0) > \hat{g}_n(t_0 + Cn^{-1/3}) > g_0(t_0 + 2Cn^{-1/3}).$$

Hence, with the probability more than $1 - 8\varepsilon$,

$$n^{1/3}(g_0(t_0 - 2Cn^{-1/3}) - g_0(t_0)) \geq n^{1/3}(\hat{g}_n(t_0 - Cn^{-1/3}) - g_0(t_0)) \geq 0,$$

and

$$n^{1/3}(g_0(t_0 + 2Cn^{-1/3}) - g_0(t_0)) \leq n^{1/3}(\hat{g}_n(t_0 + Cn^{-1/3}) - g_0(t_0)) \leq 0.$$

Together with the fact that $C \geq M$, we know that with the probability more than $1 - 8\varepsilon$,

$$\sup_{|h| \leq M} n^{1/3}\left|\hat{g}_n(t_0 + hn^{-1/3}) - g_0(t_0)\right| \leq \max_{K \in \{-C, C\}} n^{1/3}\left|g_0(t_0 + 2Kn^{-1/3}) - g_0(t_0)\right|.$$

The right hand side of the above inequality is eventually bounded by a positive constant in light of Assumptions 2.1 (B).

To verify (10.21), define $L_n \triangleq \{\hat{g}_n(t_0 - Cn^{-1/3}) \leq g_0(t_0)\}$. Note that for large enough n, $g_0(t_0 - R) > g_0(t_0 - Cn^{-1/3})$ as $g_0'(t_0)$ is strictly negative. By Proposition 4.1 (iv), we see that with the probability more than $1 - \varepsilon$, for large enough n, $\hat{g}_n(t_0 - R) > \hat{g}_n(t_0 - Cn^{-1/3})$, implying that $\hat{g}_n$ has a jump in $I_n = (t_0 - R, t_0 - Cn^{-1/3}]$. Define U_n to be the event that $\hat{g}_n$ has a jump point in I_n. Then, $P(U_n) \geq 1 - \varepsilon$ for large enough n. Now, fix a sample point $\omega \in L_n \cap U_n$. Denote $\tau_n \triangleq \tau_n(\omega)$ the last jump point of $\hat{g}_n$ in I_n, we have, from Karush-Kuhn-Tucker conditions,

$$0 \leq \sum_{j:0<T_j<t_0} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) = \sum_{j:0<T_j<\tau_n} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) + \sum_{j:\tau_n \leq T_j < t_0} \frac{\partial \psi_n}{\partial z_j}(\hat{z})$$

and

$$\sum_{j:0<T_j<\tau_n} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) = 0.$$

Therefore,

$$\sum_{j:\tau_n \leq T_j < t_0} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) \geq 0.$$

Note that if $t \geq \tau_n$, by the monotonicity of $\hat{g}_n$, the fact that τ_n is the last jump point and $\omega \in L_n \cap U_n$, we have $\hat{g}_n(t) \leq \hat{g}_n(\tau_n) = \hat{g}_n(t_0 - Cn^{-1/3}) \leq g_0(t_0)$. Hence,

(10.23)
$$0 \leq \sum_{j:\tau_n \leq T_j < t_0} \left(-\frac{1}{\hat{g}_n(T_j)} + \sum_{i=1}^{s} \frac{n_i c_{ij}}{\int w_i \hat{g}_n} \right) \leq \sum_{j:\tau_n \leq T_j < t_0} \left(-\frac{1}{g_0(t_0)} + \sum_{i=1}^{s} \frac{n_i c_{ij}}{\int w_i \hat{g}_n} \right).$$

Since $\tau_n \in I_n$, we have

$$0 \leq \sup_{t \in I_n} \sum_{j:t \leq T_j < t_0} \left(-\frac{1}{n g_0(t_0)} + \sum_{i=1}^{s} \frac{n_i}{n} \frac{c_{ij}}{\int w_i \hat{g}_n} \right). \tag{10.24}$$

However, from Lemma 5.11, for large enough n, we know that (10.24) happens with the probability at most ε, implying that $\mathbb{P}(L_n) = \mathbb{P}(L_n \cap U_n) + \mathbb{P}(L_n \cap U_n^c) \leq \varepsilon + \mathbb{P}(U_n^c) \leq 2\varepsilon$. The verification of (10.22) is similar and see the details in Section 9 of the supplementary materials.

□

PROOF OF THE CLAIM (10.22). Using the notation in Lemma 5.11 and Lemma 5.12, choose $C \geq \max\{C_0, C_1, M\}$ and $0 < R \leq \min\{R_0, R_1\}$. Define $\tilde{L}_n \triangleq \{\hat{g}_n(t_0 - Cn^{-1/3}) \geq g_0(t_0 - 2Cn^{-1/3})\}$. Note that for large enough n, $g_0(t_0 - Cn^{-1/3}) > g_0(t_0 + R)$ as $g_0'(t_0)$ is strictly negative. By Proposition 4.1 (iv), we see that with the probability more than $1 - \varepsilon$, for large enough n, $\hat{g}_n(t_0 - Cn^{-1/3}) >$

$\hat{g}_n(t_0+R)$, implying that $\hat{g}_n$ has a jump in $\tilde{I}_n = [t_0 - Cn^{-1/3}, t_0+R)$. Define $\tilde{U}_n$ to be the event that $\hat{g}_n$ has a jump in $\tilde{I}_n$. Thus, $\mathbb{P}(\tilde{U}_n) \geq 1-\varepsilon$ for large enough n. Now, fix a sample point $\omega \in \tilde{L}_n \cap \tilde{U}_n$. Denote $\tilde{\tau}_n \equiv \tilde{\tau}(\omega)$ the first jump point of $\hat{g}_n$ in $\tilde{I}_n$. Note that from Karush-Kuhn-Tucker conditions,

$$0 = \sum_{j:0<T_j<\tilde{\tau}_n} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) = \sum_{j:0<T_j<t_0-2Cn^{-1/3}} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) + \sum_{j:t_0-2Cn^{-1/3}\leq T_j<\tau_n} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}})$$

and

$$\sum_{j:0<T_j<t_0-2Cn^{-1/3}} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) \geq 0.$$

Therefore,

$$\sum_{j:t_0-2Cn^{-1/3}\leq T_j<\tilde{\tau}_n} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) \leq 0.$$

If $t \in [t_0 - 2Cn^{-1/3}, \tilde{\tau}_n)$, by the monotonicity of $\hat{g}_n$, the fact that $\tilde{\tau}_n$ is the first jump point and $\omega \in \tilde{L}_n \cap \tilde{U}_n$, we have $\hat{g}_n(t) \geq \hat{g}_n(\tilde{\tau}_n-) = \hat{g}_n(t_0 - Cn^{-1/3}) \geq g_0(t_0 - 2Cn^{-1/3})$. Therefore,

$$\begin{aligned}
&\sum_{j:t_0-2Cn^{-1/3}\leq T_j<\tilde{\tau}_n} \left(-\frac{1}{g_0(t_0-2Cn^{-1/3})} + \sum_{i=1}^{s} \frac{n_i c_{ij}}{\int w_i \hat{g}_n} \right) \\
\leq \quad &\sum_{j:t_0-2Cn^{-1/3}\leq T_j<\tilde{\tau}_n} \left(-\frac{1}{\hat{g}_n(T_j)} + \sum_{i=1}^{s} \frac{n_i c_{ij}}{\int w_i \hat{g}_n} \right) \leq 0.
\end{aligned}$$

Since $\tilde{\tau}_n \in \tilde{I}_n$, we have

$$\inf_{t\in\tilde{I}_n} \sum_{j:t_0-2Cn^{-1/3}\leq T_j<t} \left(-\frac{1}{ng_0(t_0-2Cn^{-1/3})} + \sum_{i=1}^{s} \frac{n_i}{n} \frac{c_{ij}}{\int w_i \hat{g}_n} \right) \leq 0. \tag{10.25}$$

However, from Lemma (5.12), for large enough n, we know that (10.25) happens with the probability at most ε, implying that $\mathbb{P}(\tilde{L}_n) = \mathbb{P}(\tilde{L}_n \cap \tilde{U}_n) + \mathbb{P}(\tilde{L}_n \cap \tilde{U}_n^c) \leq \varepsilon + \mathbb{P}(\tilde{U}_n^c) \leq 2\varepsilon$. □

PROOF OF LEMMA 5.14. Fix $\varepsilon > 0$ and $\tilde{C} > 0$. Denote $t_n \triangleq t_0 - 2\tilde{C}n^{-1/3}$. Define $\bar{J}_i(t) \triangleq \sum_{j:t\leq X_{ij}<t_n} \left(-\frac{1}{n_i g_0(t_n)} + \frac{\sum_{j:t\leq T_j<t_n} c_{ij}}{\int w_i \hat{g}_n} \right)$. Then, for $t < t_n$,

$$\sum_{j:t\leq T_j<t_n} \left(-\frac{1}{ng_0(t_n)} + \sum_{i=1}^{s} \frac{n_i}{n} \frac{c_{ij}}{\int w_i \hat{g}_n} \right) = \sum_{i=1}^{s} \frac{n_i}{n} \bar{J}_i(t).$$

Again, following the same argument leading to Lemma 5.11, it suffices to show that $\mathbb{P}(\sup_{t\in\bar{I}_n}\bar{J}_i(t)\geq 0)\leq\varepsilon$. For the notational simplicity, we suppress the subscript i if there is no cause of ambiguity. Define

$$\begin{aligned}
\bar{\mathcal{A}}(t) &\triangleq -\int\frac{1}{g_0(t_n)}1(t\leq x<t_n)dF(x)+\frac{\int_t^{t_n}w(x)dx}{\int wg_0},\\
\bar{\mathcal{B}}(t) &\triangleq -\int\frac{1}{g_0(t_n)}1(t\leq x<t_n)d(\mathbb{F}_{n_i}-F)(x),\\
\bar{\mathcal{C}}(t) &\triangleq \frac{\int_{T_L}^t w(x)dx}{\int w\hat{g}_n}-\frac{\int_{T_U}^{t_n}w(x)dx}{\int w\hat{g}_n},\\
\bar{\mathcal{D}}(t) &\triangleq \frac{\int_t^{t_n}w(x)dx}{\int w\hat{g}_n}-\frac{\int_t^{t_n}w(x)dx}{\int wg_0},
\end{aligned}$$

where T_L and T_U denote the maxima of X_{ij} less than t and t_n respectively. Telescoping the terms, we obtain $\bar{J}(t)=\bar{\mathcal{A}}(t)+\bar{\mathcal{B}}(t)+\bar{\mathcal{C}}(t)+\bar{\mathcal{D}}(t)$.

(i) Note that, since $g_0'(t_n)<0$,

$$\begin{aligned}
\bar{\mathcal{A}}(t) &= -\int_t^{t_n}\frac{g_0(x)-g_0(t_n)}{g_0(t_n)}\frac{w(t_n)}{\int wg_0}dx\\
&\quad -\int_t^{t_n}\frac{g_0(x)-g_0(t_n)}{g_0(t_n)}\frac{w(x)-w(t_n)}{\int wg_0}dx\\
&= -\int_t^{t_n}\frac{g_0'(t_n)(x-t_n)}{g_0(t_n)}\frac{w(t_n)}{\int wg_0}dx+o(|t_n-t|^2)+O(|t_n-t|^3)\\
&= -\frac{|g_0'(t_n)|}{g_0(t_n)}\frac{w(t_n)}{\int wg_0}\frac{(t_n-t)^2}{2}+o(|t_n-t|^2).
\end{aligned}$$

Note that t_n converges to t_0 as n goes to infinity. Therefore, there exists $\tilde{R}>0$ such that for all sufficiently large n, for all $t\in(t_0-\tilde{R},t_n)$, $\bar{\mathcal{A}}(t)\leq-\bar{\alpha}(t_n-t)^2$, for some $\bar{\alpha}>0$ by Assumptions 2.1 (B) and (C).

(ii) For $\bar{\mathcal{B}}(t)$, using argument as the proof of Lemma 4.1 in Kim and Pollard (1990), for any $\delta>0$, there exists a tight sequence of random variables $\bar{M}_n$ such that for all $t\in(t_0-\tilde{R},t_n)$,

$$\left|-\int\frac{1}{g_0(t_n)}1(t\leq x<t_n)d(\mathbb{F}_{n_i}-F)(x)\right|\leq\delta(t-t_n)^2+n^{-2/3}\bar{M}_n^2.$$

(iii) For $\bar{\mathcal{C}}(t)$, it is clear that the terms $\frac{\int_{T_L}^t w(x)dx}{\int w\hat{g}_n}$ and $\frac{\int_{T_U}^{t_n}w(x)dx}{\int w\hat{g}_n}$ are of order $O_p(n^{-1})$, which are also independent of the choice of $t\in(t_0-\tilde{R},t_n)$; see also the proof of Lemma 5.6.

(iv) For $\bar{\mathcal{D}}(t)$, by Proposition 5.8,

$$\frac{\int_t^{t_n} w(x)dx}{\int w\hat{g}_n} - \frac{\int_t^{t_n} w(x)dx}{\int wg_0} = O_p(n^{-1/2})|t - t_n|.$$

Now, for $t \in (t_0 - \tilde{R}, t_n)$, by choosing $\delta = \alpha/2$,

$$\begin{aligned} J(t) &\leq -\frac{\alpha}{2}(t - t_n)^2 + n^{-2/3}\bar{M}_n^2 + O_p(n^{-1/2})|t - t_n| \\ &= \left(-\frac{\alpha}{2} + O_p(n^{-1/2})|t - t_n|^{-1} \right)(t - t_n)^2 + n^{-2/3}\bar{M}_n^2; \end{aligned}$$

as a consequence, with probability more than $1 - \varepsilon/2$, there exists $E > 0$ such that for any $\tilde{D} > 0$, any $t \in \bar{I}_n$, and all large enough n,

$$J(t) \leq n^{-2/3}\left(-\frac{\alpha}{2}\tilde{D}^2 + E\tilde{D}n^{-1/6} + \bar{M}_n^2 \right).$$

By choosing sufficiently large $\tilde{D}$, we can guarantee that with probability more than $1 - \varepsilon$, $-\frac{\alpha}{2}\tilde{D}^2 + E\tilde{D}n^{-1/6} + \bar{M}_n^2 < 0$ as $\bar{M}_n^2 = O_p(1)$. Since the above bound for $J(t)$ is independent of $t \in \tilde{I}_n$, $\mathbb{P}(\sup_{t \in I_n} \bar{J}(t) < 0) \geq 1 - \varepsilon$ and the desired result follows. $\square$

PROOF OF LEMMA 5.15. We shall only prove that $\tau_n^- - t_0 = O_p(n^{-1/3})$ using Lemma 5.14; the proof for $\tau_n^+ - t_0 = O_p(n^{-1/3})$ is similar, and therefore we omit it. Now, in Lemma 5.14, set $\tilde{C} = \max\{C_0, C_1, K_1, K_2\}$, where C_0 and C_1 are the constants chosen in the proof of Lemma 5.13 . Define $L_n^* \triangleq \{\hat{g}_n(t_0 - (2\tilde{C} + \tilde{D})n^{-1/3}) > g_0(t_0 - 2\tilde{C}n^{-1/3})\}$. We first claim that $\mathbb{P}(L_n^*) \geq 1 - 2\varepsilon$ for large enough n. From (10.22), we know that $\mathbb{P}(\hat{g}_n(t_0 - \tilde{C}n^{-1/3}) < g_0(t_0 - 2\tilde{C}n^{-1/3})) \geq 1 - 2\varepsilon$ for large enough n. Hence, from (10.21), for large enough n, with probability more than $1 - 4\varepsilon$, we further have

$$\{\hat{g}_n(t_0 - \tilde{C}n^{-1/3}) < g_0(t_0 - 2\tilde{C}n^{-1/3}) < \hat{g}_n(t_0 - (2\tilde{C} + \tilde{D})n^{-1/3})\};$$

therefore, in particular, also by definition, with probability more than $1 - 4\varepsilon$,

$$\hat{g}_n(t_0 - K_2 n^{-1/3}) \leq \hat{g}_n(t_0 - \tilde{C}n^{-1/3}) < \hat{g}_n(t_0 - (2\tilde{C} + \tilde{D})n^{-1/3}),$$

implies that with probability more than $1 - 4\varepsilon$, for large enough n, $\hat{g}_n$ has a jump in $[t_0 - (2C + \tilde{D})n^{-1/3}, t_0 - K_2 n^{-1/3}]$ and hence $|\tau_n^- - t_0| \leq |t_0 - (t_0 - (2\tilde{C} + \tilde{D})n^{-1/3})| = (2\tilde{C} + \tilde{D})n^{-1/3}$.

We now verify the claim that $\mathbb{P}(L_n^*) \geq 1 - 2\varepsilon$ for large enough n. Denote U_n^* the event that $\hat{g}_n$ has a jump in $(t_0 - \tilde{R}, t_0 - (2\tilde{C} + \tilde{D})n^{-1/3}]$. By Proposition 4.1

(iv), we know that $\mathbb{P}(U_n^*) \geq 1 - \varepsilon$. Fix $\omega \in (L_n^*)^c \cap U_n^*$. Let $\tau_n \triangleq \tau_n(\omega)$ denote the last jump point on $(t_0 - \tilde{R}, t_0 - (2\tilde{C} + \tilde{D})n^{-1/3}]$. From Karush-Kuhn-Tucker condition, we know that

$$0 \leq \sum_{j:0<T_j<t_0-2\tilde{C}n^{-1/3}} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) = \sum_{j:0<T_j<\tau_n} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) + \sum_{j:\tau_n \leq T_j<t_0-2\tilde{C}n^{-1/3}} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}})$$

and

$$\sum_{j:0<T_j<\tau_n} \frac{\partial \psi_n}{\partial z_j}(\hat{\boldsymbol{z}}) = 0.$$

Therefore,

$$\sum_{j:\tau_n \leq T_j<t_0-2\tilde{C}n^{-1/3}} \frac{\partial \psi_n}{\partial z_j}(\hat{z}) \geq 0.$$

If $t \geq \tau_n$, then $\hat{g}_n(t) \leq \hat{g}_n(\tau_n) = \hat{g}_n(t_0 - (2\tilde{C} + \tilde{D})n^{-1/3}) \leq g_0(t_0 - 2\tilde{C}n^{-1/3})$. Therefore,

$$\begin{aligned} 0 &\leq \sum_{j:\tau_n \leq T_j<t_0-2\tilde{C}n^{-1/3}} \left(-\frac{1}{\hat{g}_n(T_j)} + \sum_{i=1}^{s} \frac{n_i}{n} \frac{c_{ij}}{\int w_i \hat{g}_n} \right) \\ &\leq \sum_{j:\tau_n \leq T_j<t_0-2\tilde{C}n^{-1/3}} \left(-\frac{1}{g_0(t_0 - 2Cn^{-1/3})} + \sum_{i=1}^{s} \frac{n_i}{n} \frac{c_{ij}}{\int w_i \hat{g}_n} \right). \end{aligned}$$

Since $\tau_n \in I_n$, as is defined in Lemma 5.14, we have

$$\sup_{t \in I_n} \sum_{j:t \leq T_j<t_0-2\tilde{C}n^{-1/3}} \left(-\frac{1}{ng_0(t_0 - 2\tilde{C}n^{-1/3})} + \sum_{i=1}^{s} \frac{n_i}{n} \frac{c_{ij}}{\int w_i \hat{g}_n} \right) \geq 0. \tag{10.26}$$

For large enough n, by Lemma 5.14, the event that the inequality (10.26) holds has a probability less than ε, thus $1 - \mathbb{P}(L_n^*) = \mathbb{P}((L_n^*)^c) = \mathbb{P}((L_n^*)^c \cap U_n^*) + \mathbb{P}((U_n^*)^c) \leq 2\varepsilon$. $\square$

11. Appendix for Section 6.

PROOF OF LEMMA 6.1. We only write the proof for the case $0 \leq t \leq K$, as the proof for the case $-K \leq t \leq 0$ is similar. Let $A_n \triangleq (t_0, t_0 + tn^{-1/3}]$. Note that

$$\tilde{G}_{n,g_0}(t) - \tilde{G}_{n,\hat{g}_n}(t) = \frac{n^{1/3}}{\lambda} \frac{1}{n} \sum_{j=1}^{n} \left(\frac{1}{g_0^2(T_j)} - \frac{1}{\hat{g}_n^2(T_j)} \right) 1_{A_n}(T_j).$$

For $T_j \in A_n$, if $\hat{g}_n(T_j) \geq g_0(T_j)$, by the monotonicity of $\hat{g}_n$ and g_0, we have $0 \leq \hat{g}_n(T_j) - g_0(T_j) \leq \hat{g}_n(t_0) - g_0(t_0 + Kn^{-1/3})$. Similarly, if $\hat{g}_n(T_j) < g_0(T_j)$, $\hat{g}_n(t_0+Kn^{-1/3}) - g_0(t_0) \leq \hat{g}_n(T_j) - g_0(T_j) < 0$. Hence, by also using Proposition 5.13,

$$\begin{aligned}
& |\hat{g}_n(T_j) - g_0(T_j)| \\
\leq\ & \max\{|\hat{g}_n(t_0) - g_0(t_0 + Kn^{-1/3})|, |\hat{g}_n(t_0 + Kn^{-1/3}) - g_0(t_0)|\} \\
\leq\ & |\hat{g}_n(t_0) - g_0(t_0)| + |g_0(t_0) - g_0(t_0 + Kn^{-1/3})| \\
& +|\hat{g}_n(t_0 + Kn^{-1/3}) - g_0(t_0)| = O_p(n^{-1/3}).
\end{aligned}$$

Note also that

$$\begin{aligned}
& \frac{1}{n}\sum_{j=1}^{n} 1_{A_n}(T_j) = \sum_{i=1}^{s} \frac{n_i}{n}\frac{1}{n_i}\sum_{j=1}^{n_i} 1_{A_n}(X_{ij}) \\
=\ & \sum_{i=1}^{s} \frac{n_i}{n}\int 1_{A_n}(x)d(\mathbb{F}_{i,n_i} - F_i)(x) + \sum_{i=1}^{s}\frac{n_i}{n}\int 1_{A_n}(x)dF_i(x) \\
=\ & O_p(n^{-1/2}) + O_p(n^{-1/3}) = O_p(n^{-1/3}).
\end{aligned}$$

Now, note that $t_0 \in (a,b)$. For every $\delta_1 \in (0, \frac{b-t_0}{2})$, if $T_j \in A_n$, then $T_j \in (t_0, b-\delta_1)$ for large enough n and therefore Proposition 4.1 (iii) implies that eventually $\frac{1}{\hat{g}_n^2(T_j)} \leq \frac{1}{\hat{g}_n^2(b-\delta_1)} = O_p(1)$. Similarly, Proposition 4.1 (iii) with $\delta_2 = \frac{t_0-a}{2}$ implies that eventually $\hat{g}_n^2(T_j) \leq \hat{g}_n^2(a+\delta_2) = O_p(1)$. These facts together with the boundedness from below and above of g_0 give

$$\begin{aligned}
& |\tilde{G}_{n,g_0}(t) - \tilde{G}_{n,\hat{g}_n}(t)| \\
\leq\ & \frac{n^{1/3}}{\lambda}\frac{1}{n}\sum_{j=1}^{n}\left|\hat{g}_n(T_j) - g_0(T_j)\right|\frac{\hat{g}_n(T_j) + g_0(T_j)}{g_0^2(T_j)\hat{g}_n^2(T_j)}1_{A_n}(T_j) \\
=\ & \frac{n^{1/3}}{\lambda}\frac{1}{n}\sum_{j=1}^{n} O_p(n^{-1/3})O_p(1)1_{A_n}(T_j) \\
=\ & \frac{n^{1/3}}{\lambda}O_p(n^{-1/3})O_p(1)\frac{1}{n}\sum_{j=1}^{n}1_{A_n}(T_j) = O_p(n^{-1/3}).
\end{aligned}$$

□

PROOF OF LEMMA 6.2. We shall only consider the case when $0 \leq t \leq K$, as the case $-K \leq t \leq 0$ can be done similarly. Again, we use the notation $A_n = (t_0, t_0 + tn^{-1/3}]$. With notation and derivation (11.1) as in Section 11 of the supplementary material, we first have

$$
\begin{aligned}
(11.1)\qquad & \lambda(\tilde{U}_{n,g_0}(t) - \tilde{U}_{n,\hat{g}_n}(t)) \\
= \;& n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\psi'_{n,j} - \psi'_{0,j})1_{A_n}(T_j) + n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\psi'''_{0,j} - \psi'''_{n,j})1_{A_n}(T_j) \\
& -\Bigg[n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\hat{g}_n(T_j) - g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\int w_i\hat{g}_n)^2}1_{A_n}(T_j) \\
& -n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(g_0(T_j) - g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\sum_{k=1}^{n} g_0(T_k)c_{ik})^2}1_{A_n}(T_j)\Bigg].
\end{aligned}
$$

First note that for j such that $T_j \in A_n$, $\hat{g}_n(T_j) \le \hat{g}_n(0+) = O_p(1)$ by Lemma 5.3 and $g_0(T_j) \le g_0(t_0) \le M$. Also, from the proof of Lemma 5.9, we have $\sup_j c_{ij} \le M \sup_j(T_j - T_{j-1}) = O_p(n^{-1/2})$ so that the order of c_{ij} is $O_p(n^{-1/2})$ uniformly for all j at a time. For the first term in (11.1),

$$
\begin{aligned}
& n^{2/3}\frac{1}{n}\sum_{j=1}^{n}((\psi'_{n,j} - \psi'_{0,j})1_{A_n}(T_j)) \\
= \;& n^{2/3}\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{s} n_i c_{ij}\Bigg[\frac{1}{\int w_i\hat{g}_n + g_0(t_0)c_{ij} - \hat{g}_n(T_j)c_{ij}} \\
& \cdot\frac{1}{\sum_{k=1}^{n} g_0(T_k)c_{ik} + g_0(t_0)c_{ij} - g_0(T_j)c_{ij}} \\
& \cdot\Bigg(\sum_{k=1}^{n} g_0(T_k)c_{ik} - \int w_i\hat{g}_n - g_0(T_j)c_{ij} + \hat{g}_n(T_j)c_{ij}\Bigg)1_{A_n}(T_j)\Bigg].
\end{aligned}
$$

Note that the denominators in the two fractions in the above equation satisfy

$$
\begin{aligned}
(11.2)\qquad & \left|\int w_i\hat{g}_n + g_0(t_0)c_{ij} - \hat{g}_n(T_j)c_{ij} - \int w_i g_0\right| \\
\le \;& \left|\int w_i\hat{g}_n - \int w_i g_0\right| + g_0(t_0)c_{ij} + \hat{g}_n(T_j)c_{ij} \\
= \;& O_p(n^{-1/2}) + O_p(n^{-1/2}) + O_p(1)O_p(n^{-1/2}) = O_p(n^{-1/2}),
\end{aligned}
$$

by Proposition 5.8, where the order does not depend on j; in addition,

$$
\begin{aligned}
(11.3) \qquad & |\sum_{k=1}^{n} g_0(T_k)c_{ik} + g_0(t_0)c_{ij} - g_0(T_j)c_{ij} - \int w_i g_0| \\
\leq \quad & |\sum_{k=1}^{n} g_0(T_k)c_{ik} - \int w_i g_0| + g_0(t_0)c_{ij} + g_0(T_j)c_{ij} \\
= \quad & O_p(n^{-1/2}) + O_p(n^{-1/2}) + O(1)O_p(n^{-1/2}) = O_p(n^{-1/2}),
\end{aligned}
$$

by Lemma 5.9, where the order also does not depend on j. Thirdly,

$$
\begin{aligned}
& \left|\sum_{k=1}^{n} g_0(T_k)c_{ik} - \int w_i \hat{g}_n - g_0(T_j)c_{ij} + \hat{g}_n(T_j)c_{ij}\right| \\
\leq \quad & \left|\sum_{k=1}^{n} g_0(T_k)c_{ik} - \int w_i \hat{g}_n\right| + g_0(T_j)c_{ij} + \hat{g}_n(T_j)c_{ij} \\
= \quad & O_p(n^{-1/2}) + O_p(n^{-1/2}) + O_p(1)O_p(n^{-1/2}) = O_p(n^{-1/2}),
\end{aligned}
$$

by Corollary 5.10, independent of j. Hence,

$$
\begin{aligned}
& n^{2/3}\frac{1}{n}\sum_{j=1}^{n}((\psi'_{n,j} - \psi'_{0,j})1_{A_n}(T_j)) \\
= \quad & n^{2/3}\left[\sum_{i=1}^{s}\frac{n_i}{n}\left(\sum_{j=1}^{n} c_{ij}1_{A_n}(T_j)\right)\right]O_p(1)O_p(1)O_p(n^{-1/2}) \\
= \quad & n^{2/3}\sum_{i=1}^{s}\frac{n_i}{n}O_p(n^{-1/3})O_p(n^{-1/2}) = O_p(n^{-1/6}),
\end{aligned}
$$

where the second equality holds as

$$
\begin{aligned}
& \sum_{j=1}^{n} c_{ij}1_{A_n}(T_j) = \sum_{j=1}^{n}\int_{T_{i-1}}^{T_i} w_i(x)1_{(t_0, t_0+n^{-1/3}]}(T_i)dx \\
\leq \quad & \sum_{j=1}^{n}\int_{T_{i-1}}^{T_i} w_i(x)1_{(t_0-O_p(n^{-1/2}), t_0+O_p(n^{-1/2})+n^{-1/3}]}(x)dx \\
\leq \quad & \int_a^b w_i(x)1_{(t_0-O_p(n^{-1/2}), t_0+O_p(n^{-1/2})+n^{-1/3}]}(x)dx = O_p(n^{-1/3}).
\end{aligned}
$$

For the second term in (11.1), note that for j such that $T_j \in A_n$,

$$\begin{aligned}
\psi'''_{0,j} &= \left(-\frac{1}{g_0^*(T_j)^3} + \sum_{i=1}^{s} \frac{n_i c_{ij}^3}{(\sum_{i=1}^n g_0(T_k)c_{ik} + g_0^*(T_j)c_{ij} - g_0(T_j)c_{ij})^3} \right) \\
&\quad \cdot (g_0(T_j) - g_0(t_0))^2 \\
&= (O(1) + O(n)O_p(n^{-3/2})O_p(1))O(n^{-2/3}) = O_p(n^{-2/3}),
\end{aligned}$$

where the second equality follows from the facts that $\frac{1}{g_0^*(T_j)} \le \frac{1}{g_0(t_0+Kn^{-1/3})}$, $\sup_j c_{ij} = O_p(n^{-1/2})$, $(g_0(T_j) - g_0(t_0))^2 \le (g_0(t_0 + Kn^{-1/3}) - g_0(t_0))^2 = O(n^{-2/3})$, and the denominator in the fraction satisfies

$$\left| \sum_{i=1}^{n} g_0(T_k)c_{ik} + g_0^*(T_j)c_{ij} - g_0(T_j)c_{ij} - \int w_i g_0 \right| = O_p(n^{-1/2}),$$

where the order is independent of j and this follows from the similar arguments used in deriving (11.3). Hence, the order $O_p(n^{-2/3})$ of $\psi'''_{0,j}$ does not depend on j. Similarly,

$$\begin{aligned}
\psi'''_{n,j} &= \left(-\frac{1}{\hat{g}_n^*(T_j)^3} + \sum_{i=1}^{s} \frac{n_i c_{ij}^3}{(\sum_{i=1}^n \hat{g}_n(T_k)c_{ik} + \hat{g}_n^*(T_j)c_{ij} - \hat{g}_n(T_j)c_{ij})^3} \right) \\
&\quad \cdot (\hat{g}_n(T_j) - g_0(t_0))^2 \\
&= (O_p(1) + O(n)O_p(n^{-3/2})O_p(1))O_p(n^{-2/3}) = O_p(n^{-2/3}),
\end{aligned}$$

where second equality follows as Proposition 4.1 (iii) implies that $\frac{1}{\hat{g}_n^*(T_j)} \le \frac{1}{\hat{g}_n(T_j)} \le \frac{1}{\hat{g}_n(t_0+Kn^{-1/3})} = O_p(1)$, $\sup_j c_{ij} = O_p(n^{-1/2})$, Proposition 5.13 gives $(\hat{g}_n(T_j) - g_0(t_0))^2 = O_p(n^{-2/3})$, and the denominator in the fraction satisfies

$$\left| \sum_{i=1}^{n} \hat{g}_n(T_k)c_{ik} + \hat{g}_n^*(T_j)c_{ij} - \hat{g}_n(T_j)c_{ij} - \int w_i g_0 \right| = O_p(n^{-1/2}),$$

where the order is independent of j and this follows from the similar arguments used in deriving (11.2). Hence, the order $O_p(n^{-2/3})$ of $\psi'''_{n,j}$ does not depend on j as well. Therefore,

$$n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\psi'''_{0,j} - \psi'''_{n,j})1_{A_n}(T_j) = n^{2/3}O_p(n^{-2/3})\frac{1}{n}\sum_{j=1}^{n}1_{A_n}(T_j) = O_p(n^{-1/3}),$$

since $\frac{1}{n}\sum_{j=1}^{n} 1_{A_n}(T_j) = O_p(n^{-1/3})$ as shown in the proof of Lemma 6.1.

Finally, using the arguments as above, the last term in (11.1) can be shown similarly to be of order $O_p(n^{-1/2})$. Indeed, note that

$$\begin{aligned}
&\left| n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\hat{g}_n(T_j)-g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\int w_i\hat{g}_n)^2}1_{A_n}(T_j)\right| \\
= \;& n^{2/3}\frac{1}{n}\sum_{j=1}^{n}O_p(n^{-1/3})\sum_{i=1}^{s}n_i O_p(n^{-1/2})O_p(1)c_{ij}1_{A_n}(T_j) \\
= \;& O_p(n^{-1/6})\sum_{i=1}^{s}\frac{n_i}{n}\sum_{j=1}^{n}c_{ij}1_{A_n}(T_j) = O_p(n^{-1/6})O_p(n^{-1/3}) = O_p(n^{-1/2}),
\end{aligned}$$

since $\sum_{j=1}^{n} c_{ij}1_{A_n}(T_j) = O_p(n^{-1/3})$; essentially the same, we also have,

$$\begin{aligned}
&\left| n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(g_0(T_j)-g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\sum_{k=1}^{n} g_0(T_k)c_{ik})^2}1_{A_n}(T_j)\right| \\
= \;& n^{2/3}\frac{1}{n}\sum_{j=1}^{n}O(n^{-1/3})\sum_{i=1}^{s}n_i O_p(n^{-1/2})O_p(1)c_{ij}1_{A_n}(T_j) \\
= \;& O_p(n^{-1/6})\sum_{i=1}^{s}\frac{n_i}{n}\sum_{j=1}^{n}c_{ij}1_{A_n}(T_j) = O_p(n^{-1/6})O_p(n^{-1/3}) = O_p(n^{-1/2}).
\end{aligned}$$

$\square$

PROOF OF (11.1). By direct applications of Taylor's theorem and simple rearrangements, we have

$$\begin{aligned}
-\psi_n' + \psi_n''' &= (\hat{g}_n(T_j)-g_0(t_0))\frac{\partial^2\psi_n}{\partial z_j^2}(\hat{z}) - \frac{\partial\psi_n}{\partial z_j}(\hat{z}); \\
-\psi_0' + \psi_0''' &= (g_0(T_j)-g_0(t_0))\frac{\partial^2\psi_n}{\partial z_j^2}(z_0) - \frac{\partial\psi_n}{\partial z_j}(z_0),
\end{aligned}$$

where

$$\begin{aligned}
\psi_0' &\triangleq \frac{\partial\psi_n}{\partial z_j}(g_0(T_1),\dots,g_0(T_{j-1}),g_0(t_0),g_0(T_{j+1}),\dots,g_0(T_n)), \\
\psi_n' &\triangleq \frac{\partial\psi_n}{\partial z_j}(\hat{g}_n(T_1),\dots,\hat{g}_n(T_{j-1}),g_0(t_0),\hat{g}_n(T_{j+1}),\dots,\hat{g}_n(T_n)), \\
\psi_0''' &\triangleq \frac{1}{2}\frac{\partial^3\psi_n}{\partial z_j^3}(g_0(T_1),\dots,g_0(T_{j-1}),g_0^*(T_j),g_0(T_{j+1}),\dots,g_0(T_n))(g_0(T_j)-g_0(t_0))^2, \\
\psi_n''' &\triangleq \frac{1}{2}\frac{\partial^3\psi_n}{\partial z_j^3}(\hat{g}_n(T_1),\dots,\hat{g}_n(T_{j-1}),\hat{g}_n^*(T_j),\hat{g}_n(T_{j+1}),\dots,\hat{g}_n(T_n))(\hat{g}_n(T_j)-g_0(t_0))^2,
\end{aligned}$$

in which $g_0^*(T_j)$ is lying between $g_0(T_j)$ and $g_0(t_0)$, $\hat{g}_n^*(T_j)$ is lying between $\hat{g}_n(T_j)$ and $g_0(t_0)$, and

$$\begin{aligned}
\frac{\partial\psi_n}{\partial z_j}(z_0) &\triangleq \frac{\partial\psi_n}{\partial z_j}(g_0(T_1),\ldots,g_0(T_{j-1}),g_0(T_j),g_0(T_{j+1}),\ldots,g_0(T_n)),\\
\frac{\partial\psi_n}{\partial z_j}(\hat{z}) &\triangleq \frac{\partial\psi_n}{\partial z_j}(\hat{g}_n(T_1),\ldots,\hat{g}_n(T_{j-1}),\hat{g}_n(T_j),\hat{g}_n(T_{j+1}),\ldots,\hat{g}_n(T_n)),\\
\frac{\partial^2\psi_n}{\partial z_j^2}(z_0) &\triangleq \frac{\partial^2\psi_n}{\partial z_j^2}(g_0(T_1),\ldots,g_0(T_{j-1}),g_0(T_j),g_0(T_{j+1}),\ldots,g_0(T_n)),\\
\frac{\partial^2\psi_n}{\partial z_j^2}(\hat{z}) &\triangleq \frac{\partial^2\psi_n}{\partial z_j^2}(\hat{g}_n(T_1),\ldots,\hat{g}_n(T_{j-1}),\hat{g}_n(T_j),\hat{g}_n(T_{j+1}),\ldots,\hat{g}_n(T_n)).
\end{aligned}$$

Hence, we can write $\tilde{U}_{n,\hat{g}_n}(t)$ as:

$$\begin{aligned}
\lambda\tilde{U}_{n,\hat{g}_n}(t) &= n^{2/3}\frac{1}{n}\sum_{j=1}^{n}\left((\hat{g}_n(T_j)-g_0(t_0))\frac{\partial^2\psi_n}{\partial z_j^2}(\hat{z})-\frac{\partial\psi_n}{\partial z_j}(\hat{z})\right)1_{A_n}(T_j)\\
&\quad +n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\hat{g}_n(T_j)-g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\int w_i\hat{g}_n)^2}1_{A_n}(T_j)\\
&= n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(-\psi_n'+\psi_n''')1_{A_n}(T_j)\\
&\quad +n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(\hat{g}_n(T_j)-g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\int w_i\hat{g}_n)^2}1_{A_n}(T_j). \qquad (11.4)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\lambda\tilde{U}_{n,g_0}(t) &= n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(-\psi_0'+\psi_0''')1_{A_n}(T_j)\\
&\quad +n^{2/3}\frac{1}{n}\sum_{j=1}^{n}(g_0(T_j)-g_0(t_0))\sum_{i=1}^{s}\frac{n_i c_{ij}^2}{(\sum_{k=1}^{n}g_0(T_k)c_{ik})^2}1_{A_n}(T_j). \qquad (11.5)
\end{aligned}$$

By subtracting (11.4) from (11.5), (11.1) follows. □

PROOF OF LEMMA 6.3. Consider the case of $0 \le t \le K$. Define $A_n \triangleq$

$(t_0, t_0 + tn^{-1/3}]$. Note that

$$\begin{aligned}
\lambda \tilde{G}_{n,g_0}(t) &= \sum_{i=1}^{s} \frac{n_i}{n} n^{1/3} \int \frac{1}{g_0^2(x)} 1_{A_n}(x) d(\mathbb{F}_{i,n_i} - F_i)(x) \\
&\quad + \sum_{i=1}^{s} \frac{n_i}{n} n^{1/3} \int \frac{1}{g_0^2(x)} 1_{A_n}(x) dF_i(x) \\
&= n^{1/3} n^{-1/6} O_p(n^{-1/2}) + \sum_{i=1}^{s} \frac{n_i}{n} n^{1/3} \int_{t_0}^{t_0 + tn^{-1/3}} \frac{f_i(x)}{g_0^2(x)} dx \\
&= O_p(n^{-1/3}) + \sum_{i=1}^{s} \frac{n_i}{n} \frac{\int_0^{t_0 + tn^{-1/3}} \frac{f_i(x)}{g_0^2(x)} dx - \int_0^{t_0} \frac{f_i(x)}{g_0^2(x)} dx}{tn^{-1/3}} t.
\end{aligned}$$

Note that the second equality holds because the class of functions $\{\frac{1}{g_0^2} 1_{A_n}(\cdot)\}$ is a subset of $\{\frac{1}{g_0^2}(h_1 - h_2) : h_1, h_2$ are increasing functions on $\mathbb{R}$ and$||h||_\infty \le 1\}$, which has a bracketing entropy of the order $1/\delta$ (this follows from a similar argument as in Lemma 3.8 in van de Geer (2000)). Hence, this class possesses a finite bracketing integral and is a Donsker class. Finally, note that the second term converges uniformly on $0 \le t \le K$ to λt. For the case of $-K \le t \le 0$, the proof is similar, and we omit it. □

PROOF OF LEMMA 6.4. We need to show the desired convergence on $l^\infty[-K, K]$ for any $K > 0$. In particular, we only have to show the validity of the three items in Condition (2.11.21) on P.220 and the entropy integral condition in Theorem 2.11.22 in van der Vaart and Wellner (1996). Then, by this theorem, each summand of $A_1(t)$, namely $n^{1/2} \int q_{n,t}(x) d(\mathbb{F}_{i,n_i} - F_i)(x)$ for $i = 1, \ldots, s$, is asymptotically tight in $l^\infty[-K, K]$ and converges in distribution to a Gaussian process with covariance function $mK_i(u, t) = \lim_{n\to\infty}(\mathbb{E}_{F_i}(q_{n,u} \cdot q_{n,t}) - \mathbb{E}_{F_i}(q_{n,u})\mathbb{E}_{F_i}(q_{n,t}))$. For each pair u, t such that $ut > 0$, i.e., they are of the same sign,

$$\begin{aligned}
K_i(u, t) &= \lim_{n\to\infty} \left(n^{1/3} \int_{t_0 - (u_- \wedge v_-) n^{-1/3}}^{t_0 + (u_+ \wedge t_+) n^{-1/3}} \left(\frac{1}{g_0(x)} + \frac{g_0(x) - g_0(t_0)}{g_0^2(x)} \right)^2 f_i(x) dx \right) \\
&\quad - \lim_{n\to\infty} O\left(\frac{f_i(t_0)}{g_i(t_0)} n^{-1/6} \right) \\
&= \frac{f_i(t_0)}{g_0^2(t_0)} (|u| \wedge |t|).
\end{aligned}$$

Otherwise, such as when u and t are of opposite signs, we must have that $K_i(u, t) = 0$. Now, note that K_i has the same form as the covariance function of the Gaussian process $\sqrt{f_i(t_0)g_0^{-2}(t_0)}W(t)$. Hence, by the independence of different samples, as an independent sum, $\sum_{i=1}^{s}\frac{n_i}{n}n^{1/2}\int q_{n,t}(x)d(\mathbb{F}_{i,n_i}-F_i)(x)$ converges in $l^{\infty}[-K,K]$ to the Gaussian process $\sqrt{\sum_{i=1}^{s}\lambda_i f_i(t_0)g_0^{-2}(t_0)}W(t)$. It is clear that

$$q_{n,t}(x)\leq \frac{M}{m^2}n^{1/6}1(t_0-Kn^{-1/3}\leq x\leq t_0+Kn^{-1/3}),$$

under Assumptions 2.1 (C). Hence, we can choose $Q_n(x)\triangleq \frac{M}{m^2}n^{1/6}1(x\in[t_0-Kn^{-1/3},t_0+Kn^{-1/3}])$ as the envelope function of the classes of functions $\mathcal{Q}_n\triangleq\{q_{n,t}:t\in[-K,K]\}$. We now verify that each item in Condition (2.11.21) in van der Vaart and Wellner (1996) are satisfied in order:

(i) Clearly,

$$\mathbb{E}_{F_i}(Q_n^2)=\int_{t_0-Kn^{-1/3}}^{t_0+Kn^{-1/3}}\frac{M^2}{m^4}n^{1/3}f_i(x)dx=O(1).$$

(ii) Next, we clearly have that $\mathbb{E}_{F_i}(Q_n^2 1\{Q_n>\eta\sqrt{n}\})\to 0$ as $n\to\infty$ for every $\eta>0$; indeed, for large enough n, $1\{Q_n>\eta\sqrt{n}\}=1\{\frac{M}{m^2}>\eta n^{1/3}\}=0$.

(iii) Now, we also have, for any diminishing sequence $\delta_n\downarrow 0$,

$$\sup_{|u-t|<\delta_n,-K\leq u,t\leq K}\mathbb{E}_{F_i}(q_{n,u}-q_{n,t})^2\to 0;$$

indeed, we illustrate the details for the case $-K<t<0<u<K$ and $u-t<\delta_n$ as other cases can be handled similarly. Note that

$$\begin{aligned}\mathbb{E}_{F_i}(q_{n,u}-q_{n,t})^2 &= \int(q_{n,u}-q_{n,t})^2 f_i(x)dx\\ \leq\ n^{1/3}\frac{M^2}{m^4}&\int_{t_0+tn^{-1/3}}^{t_0+un^{-1/3}} f_i(x)dx\leq\frac{M^4}{m^6(b-a)}(u-t)<\frac{M^4}{m^6(b-a)}\delta_n.\end{aligned}$$

Finally, for the verification of the entropy integral condition, the argument is similar to the proof of Lemma 2.3 in Banerjee (2007a), and we omit it. $\square$

PROOF OF LEMMA 6.5. Since $\lambda\tilde{U}_{n,g_0}(t)=A_1(t)+A_2(t)+A_3(t)$, in light of Lemma 6.4, we understood the asymptotic behavior of $A_1(t)$, it remains to show that $A_2(t)$ converges uniformly to $\lambda\frac{g_0'(t_0)}{2}t^2$ on $[-K,K]$ and $A_3(t)=o_p(1)$, where the $o_p(1)$ is uniform on $[-K,K]$. Indeed, we have the following:

(i) For any $0 \le t \le K$, for each $i = 1, \dots, s$, each summand in $A_2(t)$,

$$\begin{aligned}
& n^{2/3} \int_{t_0}^{t_0+tn^{-1/3}} \frac{g_0(x) - g_0(t_0)}{g_0^2(x)} dF_i(x) \\
= \; & n^{1/3} \int_0^t (g_0(t_0 + un^{-1/3}) - g_0(t_0)) \frac{f_i(t_0 + un^{-1/3})}{g_0^2(t_0 + un^{-1/3})} du \\
= \; & \int_0^t u g_0'(t_0) \frac{f_i(t_0)}{g_0^2(t_i)} du + \int_0^t g_0'(t_0) u \Big(\frac{f_i(t_0 + un^{-1/3})}{g_i^2(t_0 + un^{-1/3})} - \frac{f_i(t_0)}{g_0^2(t_0)} \Big) du \\
& + \int_0^t [n^{1/3}(g_0(t_0 + un^{-1/3}) - g_0(t_0)) - g_0'(t_0) u] \frac{f_i(t_0 + un^{-1/3})}{g_0^2(t_0 + un^{-1/3})} du \\
= \; & f_i(t_0) g_0^{-2}(t_0) \frac{g_0'(t_0)}{2} t^2 + o(1),
\end{aligned}$$

where the $o(1)$ is uniform on $0 \le t \le K$ since

$$\sup_{0 < u \le K} \left| \frac{g_0(t_0 + un^{-1/3}) - g_0(t_0)}{un^{-1/3}} - g_0'(t_0) \right| \to 0,$$

and

$$\sup_{0 \le u \le K} \left| \frac{f_i(t_0 + un^{-1/3})}{g_i^2(t_0 + un^{-1/3})} - \frac{f_i(t_0)}{g_0^2(t_0)} \right| \to 0,$$

as $n \to \infty$. For $-K \le t \le 0$, the argument is similar. Hence, $A_2(t)$ converges uniformly to $-\lambda \frac{|g_0'(t_0)|}{2} t^2$ on $[-K, K]$.

(ii) For $0 \le t \le K$, let T_U and T_L denote the maxima of X_{ij} less than $t_0 + tn^{-1/3}$ and t_0 respectively. Then,

$$\begin{aligned}
|A_3(t)| &= \Big| \sum_{i=1}^s \frac{n_i}{n} n^{2/3} \Big(\int_{t_0}^{t_0+tn^{-1/3}} \frac{w_i(x)}{\int w_i g_0} dx - \int_{t_0}^{t_0+tn^{-1/3}} \frac{w_i(x)}{\sum_{k=1}^n c_{ik} g_0(T_k)} dx \\
&\quad + \int_{T_U}^{t_0+tn^{-1/3}} \frac{w_i(x)}{\sum_{k=1}^n c_{ik} g_0(T_k)} dx - \int_{T_L}^{t_0} \frac{w_i(x)}{\sum_{k=1}^n c_{ik} g_0(T_k)} dx \Big) \Big| \\
&\le \sum_{i=1}^s \frac{n_i}{n} n^{2/3} \Big(\frac{\int_{t_0}^{t_0+Kn^{-1/3}} w_i(x) dx}{\int w_i g_0 \sum_{k=1}^n c_{ik} g_0(T_k)} \Big| \sum_{k=1}^n c_{ik} g_0(T_k) - \int w_i g_0 \Big| + O_p(n^{-1}) \Big) \\
&= n^{2/3} [O(n^{-1/3}) O_p(1) O_p(n^{-1/2}) + O_p(n^{-1})] = O_p(n^{-1/6}),
\end{aligned}$$

which is independent of $t \in [0, K]$. Here, in the first inequality, it follows by considering the orders of third and fourth terms in the line above because both $\int_{T_U}^{t_0+tn^{-1/3}} \frac{w_i(x)}{\sum_{k=1}^n c_{ik} g_0(T_k)} dx$ and $\int_{T_L}^{t_0} \frac{w_i(x)}{\sum_{k=1}^n c_{ik} g_0(T_k)} dx$ are of order $O_p(n^{-1})$

(see also the proof of Lemma 5.6), which are also independent of $t \in [0, K]$. The second last equality follows from Lemma 5.9. Similar argument holds for the case of $-K \leq t \leq 0$.

□

The following lemma is adapted from Prakasa Rao (1969) and is used for proving Theorem 6.6.

LEMMA 11.1. *Suppose that $\{\boldsymbol{V}_{n\varepsilon}\}, \{\boldsymbol{V}_n\}$ and $\{\boldsymbol{V}_\varepsilon\}$ are three sets of random vectors such that*

(i) $\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \mathbb{P}(\boldsymbol{V}_{n\varepsilon} \neq \boldsymbol{V}_n) = 0$*;*
(ii) $\lim_{\varepsilon \to 0} \mathbb{P}(\boldsymbol{V}_\varepsilon \neq \boldsymbol{V}) = 0$*; and*
(iii) For every $\varepsilon > 0$*,* $\boldsymbol{V}_{n\varepsilon} \stackrel{\mathbb{D}}{\to} \boldsymbol{V}_\varepsilon$*, as* $n \to \infty$*.*

Then $\boldsymbol{V}_n \stackrel{\mathbb{D}}{\to} \boldsymbol{V}$*, as* $n \to \infty$*.*

PROOF OF THEOREM 6.6. Clearly, by definition and Proposition 3.2, we have

$$\{\hat{g}_n(T_i) - g_0(t_0)\}_{i=1}^n = \text{slolcm}\{G_{n,\hat{g}_n}(T_i), U_{n,\hat{g}_n}(T_i) - g_0(t_0)G_{n,\hat{g}_n}(T_i)\}_{i=0}^n.$$

From the construction, it is clear that $\{\mathbb{X}_n(t_i) : i = 1, \ldots, k\}$ are equal to the slopes of the LCM of $\{(\tilde{G}_{n,\hat{g}_n}(t), \tilde{U}_{n,\hat{g}_n}(t)) : t \in \mathbb{R}\}$ evaluated at points $\tilde{G}_{n,\hat{g}_n}(t_i)$, $i = 1, \ldots, k$. Fix $C > 0$. We denote $\mathbb{X}_{n,C}(\omega)$ and $g_{a^*,b^*,C}(w)$ the respective slopes of the LCM of $\{(\tilde{G}_{n,\hat{g}_n}(t), \tilde{U}_{n,\hat{g}_n}(t)) : t \in [-C, C]\}$ evaluated at $\tilde{G}_{n,\hat{g}_n}(w)$ and that of$\{X_{a^*,b^*}(h) : h \in [-C, C]\}$ evaluated at w, for $|w| < C$. Now, fix $M > 0$ such that the points $t_1, \ldots, t_k$ are in the interior of the set $[-M, M]$. For any $\varepsilon > 0$, by Lemma 5.15, there exists $M_\varepsilon > 0$ such that with probability more than $1 - \varepsilon$, eventually, $-M_\varepsilon < \tilde{\tau}_n^- \leq -M < M \leq \tilde{\tau}_n^+ < M_\varepsilon$, where $\tilde{\tau}_n^- \triangleq n^{1/3}(\tau_n^- - t_0)$ and $\tilde{\tau}_n^+ \triangleq n^{1/3}(\tau_n^+ - t_0)$. To apply Lemma 11.1, define

$$\begin{aligned}
\boldsymbol{V}_{n,\varepsilon} &\triangleq \{\mathbb{X}_{n,M_\varepsilon}(t_i), i = 1, \ldots, k\}, \\
\boldsymbol{V}_n &\triangleq \{\mathbb{X}_n(t_i), i = 1, \ldots, k\}, \\
\boldsymbol{V}_\varepsilon &\triangleq \{g_{a^*,b^*,M_\varepsilon}(t_i), i = 1, \ldots, k\}, \\
\boldsymbol{V} &\triangleq \{g_{a^*,b^*}(t_i), i = 1, \ldots, k\}.
\end{aligned}$$

By Lemmas 6.1, 6.2, 6.3 and 6.5, the verification of Conditions (i)-(iii) in Lemma 11.1 is similar to that in Theorem 2.1 in Banerjee (2007b), and therefore is omitted. By Lemma 11.1, $X_n(t) \stackrel{\mathbb{D}}{\to} g_{a^*,b^*}(t)$. Finally, the convergence of $\mathbb{X}_n$ to g_{a^*,b^*} in $\mathcal{L}^2_{\text{loc}}(\mathbb{R})$ follows from the corresponding finite dimensional convergence and the monotonicity of $\mathbb{X}_n$; see Corollary 2 in Huang and Zhang (1994). □

PROOF OF THEOREM 1.1. From Theorem 6.6, we have $\mathbb{X}_n(0) = n^{1/3}(\hat{g}_n(t_0) - g_0(t_0)) \xrightarrow{\mathbb{D}} g_{a^*,b^*}(0)$. As in Equations (6.7)-(6.9) in Banerjee and Wellner (2001), it is easy to see that

$$g_{a^*,b^*}(t) \stackrel{\mathbb{D}}{=} a^*(b^*/a^*)^{1/3} \cdot g_{1,1}((b^*/a^*)^{2/3}t),$$

as a processes indexed by $t \in \mathbb{R}$. Using switch relationship, we also know that $g_{1,1}(0) \stackrel{\mathbb{D}}{=} 2\mathbb{Y}$. Hence, we obtain $n^{1/3}(\hat{g}_n(t_0) - g_0(t_0)) \xrightarrow{\mathbb{D}} 2(a^*)^{2/3}(b^*)^{1/3}\mathbb{Y}$. The claim follows by rearranging the terms and using the definitions of a^* and b^*. □

References.

BANERJEE, M. (2007a). Likelihood based inference for monotone response models (A more detailed version of the paper). *http://dept.stat.lsa.umich.edu/ moulib/wilks3.pdf*.

BANERJEE, M. (2007b). Likelihood based inference for monotone response models. *The Annals of Statistics* **35** 931–956.

BANERJEE, M. and WELLNER, J. A. (2001). Likelihood ratio tests for monotone functions. *The Annals of Statistics* **29** 1699–1731.

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical inference under order restrictions*. Wiley.

HUANG, Y. and ZHANG, C.-H. (1994). Estimating a monotone density from censored observations. *The Annals of Statistics* **22** 1256–1274.

KIM, J. and POLLARD, D. (1990). Cube root asymptotics. *The Annals of Statistics* 191-219.

LUENBERGER, D. G. and YE, Y. (2008). *Linear and nonlinear programming*. Springer.

PRAKASA RAO, B. L. S. (1969). Estimation of a unimodal density. *Sankhya Ser. A* **31** 23—36.

RESNICK, S. I. (2013). *Extreme values, regular variation and point processes*. Springer.

ROBERTSON, T., WRIGHT, F. T. and DYKSTRA, R. L. (1988). *Order Restricted Statistical Inference*. Wiley.

SHORACK, G. R. and WELLNER, J. A. (2009). *Empirical processes with applications to statistics*. Siam.

VAN DE GEER, S. (2000). *Empirical Processes in M-Estimation*. Cambridge University Press.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag New York Inc.

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WA 98195
USA
E-MAIL: kcgchan@uw.edu

DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
1255 AMSTERDAM AVENUE
NEW YORK, NEW YORK 10027
USA
E-MAIL: hokkan.ling@columbia.edu

DEPARTMENT OF STATISTICS
CHINESE UNIVERSITY OF HONG KONG
SHATIN, NT
HONG KONG SAR
E-MAIL: tonysit@sta.cuhk.edu.hk,
scpyam@sta.cuhk.eud.hk