Supplement Material for
'Entropy Learning for Dynamic Treatment Regimes'

Binyan Jiang[1], Rui Song[2], Jialiang Li[3] and Donglin Zeng[4]

The Hong Kong Polytechnic University[1], *North Carolina State University*[2]

National University of Singapore[3] *and University of North Carolina*[4]

This Supplementary Material provides technical proofs for Proposition 1 and Theorems 1 and 2 in the paper.

S1. Technical proof

S1.1 Proof of Proposition 1

Proof. Note that each stage is a single-stage outcome weighted learning problem. By verifying that the entropy loss satisfies the two sufficient conditions given in Section 2.1, we have $d_T^*(\mathbf{S}_T) = \text{sgn}(f_T(\mathbf{X}_T))$. Using the same arguments backwards through $t = T-1, \dots, 1$, we would sequentially obtain that $d_t^*(\mathbf{S}_t) = \text{sgn}(f_t(\mathbf{X}_t))$ for $t = T-1, \dots, 1$. □

S1.2 Proof of Theorem 1

Before we proceed to prove Theorem 1, we introduce two technical lemmas.

Lemma 1. *Let Φ and ϕ be the cumulative distribution function and density function of a standard Gaussian random variable. For any $x \geq 1$ we have*

$$\frac{\phi(x)}{2x} \leq \Phi(-x) \leq \frac{\phi(x)}{x}.$$

Proof. Using integration by parts we have for $x \geq 1$:

$$\Phi(-x) = \frac{\phi(x)}{x} - \int_x^{+\infty} \frac{1}{u^2}\phi(u)du \leq \frac{\phi(x)}{x} - \Phi(-x).$$

Lemma 1 is then proved immediately from the above inequality. □

Lemma 2. *Under assumptions A1 and A2, there exist positive constants C_{T1}, C_{T1}, C_{T3} such that*

$$P(|\hat{\beta}_T - \beta_T^0|_\infty > C_{T1}\epsilon) \leq C_{T2} \exp\{-C_{T3} C_{T1}^2 n\epsilon^2\}.$$

Proof. First of all it is easy to see that $\hat{\beta}_T$ is consistent in estimating β_T. Note that for stage T,

$$\mathbf{0} = \frac{\partial l_T(\hat{\beta}_T)}{\partial \beta_T} = \frac{\partial l_T(\beta_T^0)}{\partial \beta_T} + \frac{\partial^2 l_T(\beta_T^0)}{\partial \beta_T^2}(\hat{\beta}_T - \beta_T^0) + O(|\hat{\beta}_T - \beta_T^0|_1^2),$$

where $\frac{\partial l_T(\beta_t^0)}{\partial \beta_T}$ and $\frac{\partial^2 l_T(\beta_T^0)}{\partial \beta_T^2}$ can be written as means of i.i.d. random observations. Consequently, using Bernstein's inequality (Bennett, 1962), we have there exist positive constants c_1, c_2, c_3 such that

$$P\Big(\Big|\frac{\partial^2 l_T(\beta_T^0)}{\partial \beta_T^2} - \mathrm{I}_T(\beta_T)\Big|_\infty > c_1\epsilon\Big) \leq c_2 \exp\{-c_3 c_1^2 n\epsilon^2\}.$$

Similarly there exist positive constants c_4, c_5, c_6 such that

$$P\Big(\Big|\frac{\partial l_T(\beta_T^0)}{\partial \beta_T}\Big|_\infty > c_4\epsilon\Big) \leq c_5 \exp\{-c_6 c_4^2 n\epsilon^2\}.$$

Consequently, there exists a large enough constant C_{T1} such that, when n is large enough,

$$\begin{aligned} & P(|\hat{\beta}_T - \beta_T^0|_\infty < 2C_{T1}|\mathrm{I}_T^{-1}(\beta_T^0)|_{1,\infty}\epsilon) \\ \geq\ & 1 - c_2\exp\{-c_3c_1^2n\epsilon^2\} - c_5\exp\{-c_6c_4^2n\epsilon^2\}. \end{aligned}$$

This proves the lemma. □

Proof of Theorem 1

Proof. For simplicity we use p to denote the dimension of the covariates $\mathbf{X}_t$ for all stages t. We break the proof into two steps:

(i) We show that this theorem holds for $t = T$;

(ii) Given that the theorem holds for stage $t+1, \ldots, T$, we show that it also holds for stage t;

(i) For stage T, (3.1) and (3.4) can be obtained directly from Lemma 2 and its proof. We next show that (3.2) and (3.3) hold for stage T. In what follows we use $\hat\beta_{T,-\{i\}}$ to denote the estimator obtained by leaving the ith sample $\mathbf{S}_{Ti}$ out.

Proof of (3.3) for stage T:

From Lemma 2 and the boundness of $\mathbf{X}_{Ti}$, we have that there exists a large enough constants $C > 0$ such that

$$P\left(|\mathbf{X}_{Ti}^{*\top}\hat\beta_T - \mathbf{X}_{Ti}^{*\top}\beta_T^0|_\infty > C\sqrt{\frac{\log n}{n}}\right) = o\Big(\frac{\log n}{n}\Big). \tag{S1.1}$$

On the other hand, from the boundness of R_{Ti} and $\mathbf{X}_{Ti}$, we have there exists a large enough constant C_l such that $|\hat\beta_T - \hat\beta_{T,-\{i\}}|_\infty \le C_l n^{-1}$. Consequently we have there exists a constant $B > 0$ such that $\mathrm{sgn}(\mathbf{X}_{Ti}^{*\top}\hat\beta_T) = \mathrm{sgn}(\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}})$ when $|\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}})| \ge Bn^{-1}$, and from assumption A3, we have $P(|\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}}| < Bn^{-1}) \le P(|\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}}| < Bn^{-1}, |\hat\beta_{T,-\{i\}} - \beta_T^0|_\infty < b) + P(|\hat\beta_{T,-\{i\}} - \beta_T^0|_\infty > b) = O(n^{-1})$. Consequently, by denoting $\hat d_{T,-\{i\}}(\mathbf{S}_{Ti}) = \mathrm{sgn}(\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}})$ we have

$$\begin{aligned}
& \mathbb{E}\Big|I(A_{Ti} = \hat d_T(\mathbf{S}_{Ti})) - I(A_{Ti} = d_T^*(\mathbf{S}_{Ti}))\Big| \\
= & \mathbb{E}\Big|I(A_{Ti} = \hat d_{T,-\{i\}}(\mathbf{S}_T)) - I(A_{Ti} = d_T^*(\mathbf{S}_{Ti}))\Big| + o\Big(\frac{\log n}{n}\Big) \\
= & |P(\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}} > 0, \mathbf{X}_{Ti}^{*\top}\beta_T^0 \le 0) - P(\mathbf{X}_{Ti}^{*\top}\hat\beta_{T,-\{i\}} < 0, \mathbf{X}_{Ti}^{*\top}\beta_T^0 \ge 0)| \\
& + o\Big(\frac{\log n}{n}\Big).
\end{aligned}$$

Denote $\hat{Y} = \mathbf{X}_{Ti}^{*\top}\hat{\beta}_{T,-\{i\}}$, by condition A3, we have,

$$\begin{aligned} & P(\mathbf{X}_{Ti}^{*\top}\hat{\beta}_{T,-\{i\}} > 0, \mathbf{X}_{Ti}^{*\top}\beta_T^0 \leq 0) \\ = & \int_{-C\sqrt{\frac{\log n}{n}}}^{0} P(\hat{Y} > 0 | Y = y) g_T(y) dy + o\Big(\frac{\log n}{n}\Big) \\ = & o\Big(\frac{\log n}{n}\Big). \end{aligned} \tag{S1.2}$$

Here the last step is obtained by noticing that $g(y) = o(y)$ as indicated by assumption A3. Similarly we have $P(\mathbf{X}_{Ti}^{*\top}\hat{\beta}_{T,-\{i\}} < 0, \mathbf{X}_{Ti}^{*\top}\beta_T^0 \geq 0) = o\Big(\frac{\log n}{n}\Big)$. This proves (3.3).

Before we proceed to prove (3.2) for stage T, similar to (3.3) we show that for $1 \leq i \neq j \leq n$,

$$\begin{aligned} & \mathbb{E}[I(A_{Ti} = \hat{d}_T(\mathbf{S}_{Ti})) - I(A_{Ti} = d_T^*(\mathbf{S}_{Ti}))] \\ & \times[I(A_{Tj} = \hat{d}_T(\mathbf{S}_{Tj})) - I(A_{Tj} = d_T^*(\mathbf{S}_{Tj}))] \\ = & \; o\left(\frac{\log^2 n}{n^2}\right). \end{aligned} \tag{S1.3}$$

For $k = i, j$, denote $\hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Tk}) = \mathrm{sgn}(\mathbf{X}_{Tk}^{\top}\hat{\beta}_{T,-\{i,j\}})$, where $\hat{\beta}_{T,-\{i,j\}}$ is the estimator of β_T^0 obtained by leaving the ith and jth samples. We have,

$$\begin{aligned} & \mathbb{E}\Pi_{k=i,j}[I(A_{Tk} = \hat{d}_T(\mathbf{S}_{Tk})) - I(A_{Tk} = d_T^*(\mathbf{S}_{Tk}))] \\ = & \; \mathbb{E}\Pi_{k=i,j}[(2I(A_{Tk} = \hat{d}_T(\mathbf{S}_{Tk})) - 1)I(\hat{d}_T(\mathbf{S}_{Tk}) \neq \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Tk})) \\ & + I(A_{Tk} = \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Tk})) - I(A_{Tk} = d_T^*(\mathbf{S}_{Tk}))]. \end{aligned}$$

Similar to the proof in (S1.2), we have there exists a large enough constant C such that $\text{sgn}(\mathbf{X}_{Ti}^{*}\hat{\beta}_T) = \text{sgn}(\mathbf{X}_{Ti}^{*\top}\hat{\beta}_{T,-\{i,j\}})$ when $|\mathbf{X}_{Ti}^{*\top}\hat{\beta}_{T,-\{i,j\}})| \geq Cn^{-1}$. Consequently,

$$\begin{aligned}
& \mathbb{E}I(\hat{d}_T(\mathbf{S}_{Ti}) \neq \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Ti})) \cdot I(\hat{d}_T(\mathbf{S}_{Tj}) \neq \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Tj})) \\
\leq \;& P\Big(|\mathbf{X}_{Ti}^{*\top}\hat{\beta}_{T,-\{i,j\}}| < Cn^{-1}, |\mathbf{X}_{Tj}^{*\top}\hat{\beta}_{T,-\{i,j\}}| < Cn^{-1}) \\
= \;& O(n^{-2}) + P(|\hat{\beta}_{T,-\{i,j\}} - \beta_T^0|_\infty > b) \\
= \;& O(n^{-2}).
\end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned}
& \mathbb{E}\big\{I(\hat{d}_T(\mathbf{S}_{Ti}) \neq \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Ti})) \\
& \times I(A_{Tj} = \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Tj})) - I(A_{Tj} = d_T^*(\mathbf{S}_{Tj}))]\big\} = o\left(\frac{\log n}{n^2}\right), \\
& \mathbb{E}\Pi_{k=i,j} I(A_{Tk} = \hat{d}_{T,-\{i,j\}}(\mathbf{S}_{Tk})) - I(A_{Tk} = d_T^*(\mathbf{S}_{Tk}))] = o\left(\frac{\log^2 n}{n^2}\right).
\end{aligned}$$

We thus conclude that (S1.3) holds for stage T.

Proof of (3.2) for stage T:

Denote

$$h_i(\beta) = \frac{R_{Ti}}{\pi(A_{Ti}, \mathbf{S}_{Ti})}\left[.5(A_{Ti}+1) - \frac{\exp(\mathbf{X}_{Ti}^{*\top}\beta)}{1+\exp(\mathbf{X}_{Ti}^{*\top}\beta)}\right]\mathbf{X}_{Ti}^{*}.$$

We have $\sum_{i=1}^n h_i(\hat{\beta}_T) = \mathbf{0}$ and $\mathbb{E}h_i(\beta_T^0) = \mathbf{0}$. Note that , $h_{ij}(\beta)$, the jth element of $h_i(\beta)$, is bounded and there exists $\beta_T^* \in [\beta_T^0, \hat{\beta}_T]$ such that

$$\sum_{i=1}^n h_{ij}(\beta_T^0) = \sum_{i=1}^n h_{ij}(\beta_T^0) - \sum_{i=1}^n h_{ij}(\hat{\beta}_T) = (\beta_T^0 - \hat{\beta}_T)^\top \sum_{i=1}^n h'_{ij}(\beta_T^*). \quad \text{(S1.4)}$$

Since $\mathbb{E}h''_{ij}(\beta_T^0)$ is finite, by Lemma 2, we have with probability larger than $1-o\Big(\frac{\log n}{n}\Big)$,

$$\left|\frac{1}{n}\sum_{i=1}^{n} h'_{ij}(\beta_T^*) - \frac{1}{n}\sum_{i=1}^{n} h'_{ij}(\beta_T^0)\right|_\infty = O\Big(\sqrt{\frac{\log n}{n}}\Big). \tag{S1.5}$$

Again, using Bernstein's inequality we have for a large enough constant C_1,

$$P\left(\Big|\frac{1}{n}\sum_{i=1}^{n} h'_{i}(\beta_T^0) - Eh'_{i}(\beta_T^0)\Big|_\infty > C_1\sqrt{\frac{\log n}{n}}\right) = o\Big(\frac{\log n}{n}\Big), \tag{S1.6}$$

where $h'_i(\beta) = (h'_{i1}(\beta), \ldots, h'_{ip}(\beta))$. Write $F := \mathbb{E}h'_i(\beta_T^0)$. From (S1.4), (S1.5) and (S1.6) we have with probability greater than $1-o\Big(\frac{\log n}{n}\Big)$,

$$\left|\hat{\beta}_T - \beta_T^0 + n^{-1}(FF^\top)^{-1}F\sum_{i=1}^{n} h_i(\beta_T^0)\right|_\infty = O\left(\frac{\log n}{n}\right).$$

Given $\mathbf{X}_T^*$, we denote

$$W_T^2 = \mathbb{V}ar\{\sqrt{n}\mathbf{X}_T^{*\top}(\beta_T^0 - \hat{\beta}_T)\} = \mathbb{V}ar\{\mathbf{X}_T^{*\top}(FF^\top)^{-1}Fh_i(\beta_T^0)\} + O\Big(\frac{\log n}{n}\Big).$$

So far the order terms in the above derivation are obtained from Bernstein's inequality as in the proof of Lemma 2, and depends on the bounds of $R_T, \mathbf{X}_T$ and C only. From Lemma 1 and classical Cramer-Petrov type large deviation results (see for example Lin and Lu (2013), Petrov (1996)), we have for $x = o(\sqrt{n})$ and $x > 1$, as $n \to \infty$, for any $j = 1, \ldots, p$,

$$\begin{aligned} & P\left(\frac{\sum_{i=1}^{n}\mathbf{X}_T^{*\top}(FF^\top)^{-1}Fh_i(\beta_T^0)}{\sqrt{n}W_T} \geq x\right) \\ = & \left\{1 + O\Big(\frac{x^3}{\sqrt{n}}\Big)\right\}[1-\Phi(x)] + O\left(\sqrt{\frac{\log n}{n}}\right), \end{aligned} \tag{S1.7}$$

where $\Phi(x)$ is the CDF of the standard normal distribution. Consequently we have when n is large enough,

$$\begin{aligned}
& P\left(\mathbf{X}_T^{*\top}(\beta_T^0 - \hat{\beta}_T) > \frac{xW_T}{\sqrt{n}}\right) \\
= \; & P\left(\Big[\frac{\sum_{i=1}^n \mathbf{X}_T^{*\top}(FF^\top)^{-1}Fh_i(\beta_T^0)}{n} + O\Big(\frac{\log n}{n}\Big)\Big] > \frac{xW_T}{\sqrt{n}}\right) + o\Big(\frac{\log n}{n}\Big) \\
= \; & P\left(\frac{\sum_{i=1}^n \mathbf{X}_T^{*\top}(FF^\top)^{-1}Fh_i(\beta_T^0)}{\sqrt{n}W_T} \geq x + O\Big(\frac{\log n}{\sqrt{n}}\Big)\right) + o\Big(\frac{\log n}{n}\Big) \\
= \; & \left\{1 + O\Big(\frac{x^3}{\sqrt{n}}\Big)\right\}[1 - \Phi(x)] + O\Big(\frac{\log n}{\sqrt{n}}\Big),
\end{aligned}$$

Similarly, we have

$$P\left(\mathbf{X}_T^{*\top}(\beta_T^0 - \hat{\beta}_T) < -\frac{xW_T}{\sqrt{n}}\right) = \left\{1 + O\Big(\frac{x^3}{\sqrt{n}}\Big)\right\}[1 - \Phi(x)] + O\Big(\frac{\log n}{\sqrt{n}}\Big).$$

This proves (3.2) for stage T.

(ii) Now suppose Theorem 1 holds for stage $t+1, \dots, T$ and (S1.3) hold for stage $t+1$ in that for any $1 \leq i \neq l \leq n$,

$$\begin{aligned}
& \mathbb{E}\Pi_{k=i,l}\left[\prod_{j=t+1}^{T} I(A_{jk} = \hat{d}_j(\mathbf{S}_{jk})) - \prod_{j=t+1}^{T} I(A_{jk} = d_j^*(\mathbf{S}_{jk}))\right] \\
= \; & o\left(\frac{\log^2 n}{n^2}\right). \qquad \text{(S1.8)}
\end{aligned}$$

We complete the proof of this theorem by showing that (3.4), (3.1), (3.2) (3.3) and (S1.8) hold for stage t respectively.

Note that for stage t,

$$\mathbf{0} = \frac{\partial l_t(\hat{\beta}_t)}{\partial \beta_t} = \frac{\partial l_t(\beta_t^0)}{\partial \beta_t} + \frac{\partial^2 l(\beta_t^0)}{\partial \beta_t^2}(\hat{\beta}_t - \beta_t^0) + O(|\hat{\beta}_t - \beta_t^0|_1^2),$$

where

$$\begin{aligned}
&\frac{\partial l_t(\beta_t^0)}{\partial \beta_t} \qquad\qquad (S1.9)\\
= \quad & -\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(\sum_{j=t}^{T} R_{ji})\prod_{j=t+1}^{T} I(A_{ji}=d_j^*(\mathbf{S}_{ji}))}{\prod_{j=t}^{T}\pi(A_{ji},\mathbf{S}_{ji})}\right.\\
&\left.\times\left[.5(A_{ti}+1)-\frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}{1+\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}\right]\right\}\mathbf{X}_{ti}^*\\
&-\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(\sum_{j=t}^{T} R_{ji})\Delta_{ti}(d^*,\hat{d})}{\prod_{j=t}^{T}\pi(A_{ji},\mathbf{S}_{ji})}\left[.5(A_{ti}+1)-\frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}{1+\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}\right]\right\}\mathbf{X}_{ti}^*,
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial^2 l(\beta_t^0)}{\partial \beta_t^2} \qquad\qquad (S1.10)\\
= \quad & \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{ti}^*\left\{\frac{(\sum_{j=t}^{T} R_{ji})\prod_{j=t+1}^{T} I(A_{ji}=d_j^*(\mathbf{S}_{ji}))}{\prod_{j=t}^{T}\pi(A_{ji},\mathbf{S}_{ji})}\right.\\
&\left.\times\frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t^0)}{[1+\exp(\mathbf{X}_{ti}^{*\top}\beta_t^0)]^2}\right\}\mathbf{X}_{ti}^{*\top}\\
&+\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{ti}^*\left\{\frac{(\sum_{j=t}^{T} R_{ji})\Delta_{ti}(d^*,\hat{d})}{\prod_{j=t}^{T}\pi(A_{ji},\mathbf{S}_{ji})}\cdot\frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t^0)}{[1+\exp(\mathbf{X}_{ti}^{*\top}\beta_t^0)]^2}\right\}\mathbf{X}_{ti}^{*\top},
\end{aligned}$$

where $\Delta_{ti}(d^*,\hat{d})=\prod_{j=t+1}^{T} I(A_{ji}=\hat{d}_j(\mathbf{S}_{ji}))-\prod_{j=t+1}^{T} I(A_{ji}=d_j^*(\mathbf{S}_{ji}))$.

Proof of (3.4) for stage t

By Slutsky's theorem, it suffices to show that

(ii.1) $\sqrt{n}\frac{\partial l_t(\beta_t^0)}{\partial\beta_t}$ converges in distribution to $N(\mathbf{0},\Gamma_t)$.

(ii.2) $\frac{\partial^2 l(\beta_t^0)}{\partial\beta_t^2}$ converges to $\mathbf{I}_t(\beta_t^0)$ in probability.

We look at (ii.1) first:

Note that the first term in the right hand side of (S1.9) is a mean of independent random variables. By Slutsky's theorem, it suffices to show that the second term in the right hand side of (S1.9) is $o_p(n^{-\frac{1}{2}})$. By the boundness of of R_{ji} and $\mathbf{X}_{ji}$ and Markov's inequality, it suffices to show that for $1 \le l, k \le n$,

$$\mathbb{E}\Delta_{tl}^2(d^*, \hat{d}) = o(1), \quad \mathbb{E}[\Delta_{tl}(d^*, \hat{d}) \cdot \Delta_{tk}(d^*, \hat{d})] = o(n^{-1}). \tag{S1.11}$$

On the other hand, by the assumption that (3.3) and (S1.8) hold for stage $t+1$, we immediately have $\mathbb{E}\Delta_{tl}^2(d^*, \hat{d}) = \mathbb{E}|\Delta_{tl}(d^*, \hat{d})| = o\left(\frac{\log n}{n}\right)$ and $\mathbb{E}[\Delta_{tl}(d^*, \hat{d}) \cdot \Delta_{tk}(d^*, \hat{d})] = O\left(\frac{\log^2 n}{n^2}\right) = o(n^{-1})$. This proves (ii.1). The proof for (ii.2) is similar to that for (ii.1): the first term on the right hand side of (S1.10) tends to $\mathbf{I}_t(\beta_t^0)$ almost surely by the law of large numbers, and the second term is $o_p\left(n^{-\frac{1}{2}}\right)$.

Proof of (3.1) for stage t

Let e_i be the ith column of the $p \times p$ identity matrix and denote

$$\begin{aligned}
\nu_{tn} &= -\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(\sum_{j=t}^{T} R_{ji})\prod_{j=t+1}^{T} I(A_{ji} = d_j^*(\mathbf{S}_{ji}))}{\prod_{j=t}^{T}\pi(A_{ji}, \mathbf{S}_{ji})}\right. \\
&\quad \left.\times\left[.5(A_{ti}+1) - \frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}{1+\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}\right]\right\}\mathbf{X}_{ti}^*, \\
\mu_{tn} &= -\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(\sum_{j=t}^{T} R_{ji})\Delta_{ti}(d^*, \hat{d})}{\prod_{j=t}^{T}\pi(A_{ji}, \mathbf{S}_{ji})}\right. \\
&\quad \times \left.\left[.5(A_{ti}+1) - \frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}{1+\exp(\mathbf{X}_{ti}^{*\top}\beta_t)}\right]\right\}\mathbf{X}_{ti}^*.
\end{aligned}$$

Hence we have $e_i^\top \frac{\partial l_t(\beta_t^0)}{\partial \beta_t} = e_i^\top \nu_{tn} + e_i^\top \mu_{tn}$. Due to the boundness of R_{ji} and $\mathbf{X}_{ji}$ and Markov's inequality, and the fact that (3.3), (S1.8) holds for stage

$t+1$, we have

$$P\left(\left|e_i^\top \mu_{tn}\right| > \sqrt{\frac{\log n}{n}}\right) \leq \frac{E\{(e_i^\top \mu_{tn})^2\}}{\frac{\log n}{n}} = o\left(\frac{\log n}{n}\right). \tag{S1.12}$$

On the other hand, by Bernstein's inequality, there exist positive constants c_1, c_2, c_3, depending on the bounds of R_j and $\mathbf{X}_j, j = t, \ldots, T$ only, such that,

$$P\Big(\Big|e_i^\top \nu_{tn}\Big| > c_1\epsilon\Big) \leq c_2 \exp\{-c_3 c_1^2 n\epsilon^2\}. \tag{S1.13}$$

Consequently, by choosing c_1 to be large enough, we have

$$P\Big(\Big|e_i^\top \frac{\partial l_t(\beta_t^0)}{\partial \beta_t}\Big| > (1+c_1)\sqrt{\frac{\log n}{n}}\Big) = o\left(\frac{\log n}{n}\right). \tag{S1.14}$$

Similarly it can be shown that for some large enough constant c_4,

$$P\left(\Big|\frac{\partial^2 l_t(\beta_t^0)}{\partial \beta_t^2} - \mathbf{I}_t(\beta_t)\Big|_\infty > (1+c_4)\sqrt{\frac{\log n}{n}}\right) = o\left(\frac{\log n}{n}\right). \tag{S1.15}$$

Similar to the proof of Lemma 2, from (S1.14) and (S1.15) we have (3.1) holds for stage t.

Proof of (3.2) for stage t

Using the same arguments as in the proof for stage T, we have for $x = o(\sqrt{n})$ and $x > 1$,:

$$\begin{aligned} & P\left(\frac{\sum_{i=1}^n \mathbf{X}_t^{*\top}(F_t^\top F_t)^{-1}F_t^\top h'_{t,i}(\beta_t^0)}{\sqrt{n}W_t} \geq x\right) \\ = & \left\{1 + O\Big(\frac{x^3}{\sqrt{n}}\Big)\right\}[1-\Phi(x)] + O\left(\frac{\log n}{n}\right), \end{aligned} \tag{S1.16}$$

where

$$W_t^2 = \mathbb{V}ar\{\mathbf{X}_t^{*\top}(\hat{\beta}_t - \beta_t^0)\} = \mathbb{V}ar\{\mathbf{X}_t^{*\top}(F_t^\top F_t)^{-1}F_t^\top h'_{t,i}(\beta_t^0)\} + O(\frac{\log n}{n}),$$

and $F_t = \mathbb{E}h'_{t,i}(\beta_t^0)$, with

$$\begin{aligned} & h'_{t,i}(\beta_t^0) \\ = \;& \mathbf{X}_{ti}^* \left\{ \frac{(\sum_{j=t}^T R_{ji}) \prod_{j=t+1}^T I(A_{ji} = d_j^*(\mathbf{S}_{ji}))}{\prod_{j=t}^T \pi(A_{ji}, \mathbf{S}_{ji})} \cdot \frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t^0)}{[1+\exp(X_{ti}^{*\top}\beta_t^0)]^2} \right\} \mathbf{X}_{ti}^{*\top} \\ & + \mathbf{X}_{ti}^* \left\{ \frac{(\sum_{j=t}^T R_{ji}) \Delta_{ti}(d^*, \hat{d})}{\prod_{j=t}^T \pi(A_{ji}, \mathbf{S}_{ji})} \cdot \frac{\exp(\mathbf{X}_{ti}^{*\top}\beta_t^0)}{[1+\exp(X_{ti}^{*\top}\beta_t^0)]^2} \right\} \mathbf{X}_{ti}^{*\top}. \end{aligned}$$

The rest of the proof is the same as that in the proof for stage T.

Proof of (3.3) and (S1.8) for stage t

We look at (3.3) first. For simplicity we use (a, b) where $a, b \in \{0, 1\}$ to denote the event that $\{\prod_{j=t+1}^T I(A_{ji} = \hat{d}_j(\mathbf{S}_{ji})) = a, \prod_{j=t+1}^T I(A_{ji} = d_j^*(\mathbf{S}_{ji})) = b\}$. Since (3.3) holds for stage $t+1$ we have

$$P((1,0)) = o\Big(\frac{\log n}{n}\Big), P((0,1)) = o\Big(\frac{\log n}{n}\Big).$$

Note that

$$\begin{aligned}
& \mathbb{E}\Big|\prod_{j=t}^{T} I(A_{ji}=\hat{d}_j(\mathbf{S}_{ji})) - \prod_{j=t}^{T} I(A_{ji}=d_j^*(\mathbf{S}_{ji}))\Big| \\
= \;& P\Big(I(A_{ti}=\hat{d}_t(\mathbf{S}_{ti}))=1\Big|(1,0)\Big)P((1,0)) \\
& +P\Big(I(A_{ti}=d_t^*(\mathbf{S}_{ti}))=1\Big|(0,1)\Big)P((0,1)) \\
& +P\Big(I(A_{ti}=\hat{d}_t(\mathbf{S}_{ti}))=1, I(A_{ti}=d_t^*(\mathbf{S}_{ti}))=0\Big|(1,1)\Big)P((1,1)) \\
& +P\Big(I(A_{ti}=\hat{d}_t(\mathbf{S}_{ti}))=0, I(A_{ti}=d_t^*(\mathbf{S}_{ti}))=1\Big|(1,1)\Big)P((1,1)) \\
\leq \;& o\Big(\frac{\log n}{n}\Big) + P\Big(I(A_{ti}=\hat{d}_t(\mathbf{S}_{ti}))=1, I(A_{ti}=d_t^*(\mathbf{S}_{ti}))=0\Big) \\
& +P\Big(I(A_{ti}=\hat{d}_t(\mathbf{S}_{ti}))=0, I(A_{ti}=d_t^*(\mathbf{S}_{ti}))=1\Big) \\
= \;& o\Big(\frac{\log n}{n}\Big).
\end{aligned}$$

Here the last step can be obtained using (3.1) and similar derivations as in (S1.2).

For (S1.8), note that

$$\begin{aligned}
& \mathbb{E}\Pi_{k=i,l}\left[\prod_{j=t}^{T} I(A_{jk}=\hat{d}_j(\mathbf{S}_{jk})) - \prod_{j=t}^{T} I(A_{jk}=d_j^*(\mathbf{S}_{jk}))\right] \\
= \;& \mathbb{E}\Pi_{k=i,l}\Bigg[\prod_{j=t}^{T} I(A_{jk}=\hat{d}_j(\mathbf{S}_{jk})) - \prod_{j=t}^{T} I(A_{jk}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jk})) \\
& +\prod_{j=t}^{T} I(A_{jk}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jk})) - \prod_{j=t}^{T} I(A_{jk}=d_j^*(\mathbf{S}_{jk}))\Bigg]. \qquad \text{(S1.17)}
\end{aligned}$$

Using the same argument as in the proof of (3.1) we can also show that $P(|\hat{\beta}_t - \beta_t^0| > b) = o\Big(\frac{\log^2 n}{n^2}\Big)$. Consequently, for any $t \leq j_1, j_2 \leq T$, there exists a constant C large enough such that

$$\begin{aligned}
& \mathbb{E}\Big|I(A_{j_1i} = \hat{d}_{j_1}(\mathbf{S}_{j_1i})) - I(A_{j_1i} = \hat{d}_{j_1,-\{i,l\}}(\mathbf{S}_{j_1i}))\Big| \\
& \quad \times \Big|(A_{j_2} = \hat{d}_{j_2}(\mathbf{S}_{j_2l})) - I(A_{j_2l} = \hat{d}_{j_2,-\{i,l\}}(\mathbf{S}_{j_2l}))\Big| \\
\leq \ & P(|\mathbf{X}_{j_1i}^\top \hat{\beta}_{j_1,-\{i,l\}}| < Cn^{-1}, |\hat{\beta}_{j_1,-\{i,l\}} - \beta_{j_1}^0|_\infty < b, \\
& |\mathbf{X}_{j_2l}^\top \hat{\beta}_{j_2,-\{i,l\}}| < Cn^{-1}, |\hat{\beta}_{j_2,-\{i,l\}} - \beta_{j_2}^0|_\infty < b) + o\Big(\frac{\log^2 n}{n^2}\Big) \\
= \ & o\Big(\frac{\log^2 n}{n^2}\Big).
\end{aligned}$$

Consequently, we have

$$\begin{aligned}
& \mathbb{E}\Pi_{k=i,l}\left[\prod_{j=t}^{T} I(A_{jk} = \hat{d}_j(\mathbf{S}_{jk})) - \prod_{j=t}^{T} I(A_{jk} = \hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jk}))\right] \\
\leq \ & \mathbb{E}\left[\sum_{j=t}^{T}\Big|I(A_{ji} = \hat{d}_j(\mathbf{S}_{ji})) - I(A_{ji} = \hat{d}_{j,-\{i,l\}}(\mathbf{S}_{ji}))\Big|\right] \\
& \times \left[\sum_{j=t}^{T}\Big|I(A_{jl} = \hat{d}_j(\mathbf{S}_{jl})) - I(A_{jl} = \hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jl}))\Big|\right] \\
= \ & o\Big(\frac{\log^2 n}{n^2}\Big). \qquad \text{(S1.18)}
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathbb{E}\Big\{\Big[\prod_{j=t}^{T} I(A_{ji}=\hat{d}_j(\mathbf{S}_{ji}))-\prod_{j=t}^{T} I(A_{ji}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{ji}))\Big] \\
&\quad\times\Big[\prod_{j=t}^{T} I(A_{jl}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jl}))-\prod_{j=t}^{T} I(A_{jl}=d_j^*(\mathbf{S}_{jl}))\Big]\Big\} \\
\leq\ &\mathbb{E}\left[\sum_{j=t}^{T}\Big|I(A_{ji}=\hat{d}_j(\mathbf{S}_{ji}))-I(A_{ji}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{ji}))\Big|\right] \\
&\times\left[\sum_{j=t}^{T}\Big|I(A_{jl}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jl}))-I(A_{jl}=d_j^*(\mathbf{S}_{jl}))\Big|\right] \\
=\ &o\Big(\frac{\log^2 n}{n^2}\Big). \qquad\qquad (S1.19)
\end{aligned}$$

And

$$\begin{aligned}
&\mathbb{E}\Pi_{k=i,l}\Big[\prod_{j=t}^{T} I(A_{jk}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jk}))-\prod_{j=t}^{T} I(A_{jk}=d_j^*(\mathbf{S}_{jk}))\Big] \\
\leq\ &\mathbb{E}\left[\sum_{j=t}^{T}\Big|I(A_{ji}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{ji}))-I(A_{ji}=d_j^*(\mathbf{S}_{ji}))\Big|\right] \\
&\times\left[\sum_{j=t}^{T}\Big|I(A_{jl}=\hat{d}_{j,-\{i,l\}}(\mathbf{S}_{jl}))-I(A_{jl}=d_j^*(\mathbf{S}_{jl}))\Big|\right] \\
=\ &o\Big(\frac{\log^2 n}{n^2}\Big). \qquad\qquad (S1.20)
\end{aligned}$$

(S1.8) is then proved by combing (S1.17), (S1.18), (S1.19) and (S1.20). □

S1.3 Proof of Theorem 2

Proof. For simplicity, we only prove stage $t = T$. Proofs for stage $t = T-1,\ldots,1$ are similar to the proofs of (3.4). Note that

$$\begin{aligned}
& \sqrt{n}\hat{V}_T \\
= \quad & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{I(A_{Ti}=\hat{d}(\mathbf{X}_{Ti}))R_{Ti}}{(A_{Ti}\pi_T+(1-A_{Ti})/2)} \\
= \quad & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{I(A_{Ti}=d_T(\mathbf{X}_{Ti}))R_{Ti}}{(A_{Ti}\pi_T+(1-A_{Ti})/2)} \\
& +\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{[I(A_{Ti}=\hat{d}_T(\mathbf{X}_{Ti}))-I(A_{Ti}=d_T(\mathbf{X}_{Ti}))]R_{Ti}}{(A_{Ti}\pi_T+(1-A_{Ti})/2)}
\end{aligned}$$

Using central limit theorem we immediately have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{I(A_{Ti}=d_T(\mathbf{X}_{Ti}))R_{Ti}}{(A_{Ti}\pi_T+(1-A_{Ti})/2)} \to N(\sqrt{n}V_T,\Sigma_{V_T}).$$

On the other hand, by Markov's inequality and (3.3) we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{I(\mathrm{sgn}(\mathbf{X}_{Ti}^{*\top}\beta_T^0)\neq\mathrm{sgn}(\mathbf{X}_{Ti}^{*\top}\hat{\beta}_T))R_i}{(A_{Ti}\pi_T+(1-A_{Ti})/2)} \to 0,$$

in probability. Theorem 2 is then proved by Slutsky's theorem. □

References

BENNETT, G. (1962). Probability inequalities for the sum of independent random variables. *J. Am. Statist. Assoc.* **57**, 33–45.

LIN, Z. & LU, Z. (2013). *Strong Limit Theorems*, Volume 4. Springer Science & Business Media.

PETROV, V. (1966). A generalization of cramer's limit theorem. *Uspekhi. Mat. Nauk.* 9.