Supplement to 'Semi-supervised D-Learning for Optimal Individual Treatment Regimes'

Xintong Li[1], Shuyi Zhang[1], and Yong Zhou[1]

[1] Key Laboratory of Advanced Theory and Application in Statistics and Data Science-MOE, School of Statistics, Academy of Statistics and Interdisciplinary Sciences, East China Normal University, Shanghai, China

1 Theoretical proofs

1.1 Proof of Lemma 1

Proof. Denote $S_i = \mathbf{X}_i^T\boldsymbol{\beta}$, $\hat{s} = \mathbf{x}^T\hat{\boldsymbol{\beta}}$ and $s^0 = \mathbf{x}^T\boldsymbol{\beta}_0$. Let $\hat{f}(s) = (nh)^{-1}\sum_{i=1}^n K_h(S_i, s)$. First we have the inequality that

$$\sup_{\mathbf{x}\in\mathcal{X}} |\hat{f}(\hat{s}) - f(s^0)| \leqslant \sup_{\mathbf{x}\in\mathcal{X}} |\hat{f}(\hat{s}) - \hat{f}(s^0)| + \sup_{\mathbf{x}\in\mathcal{X}} |\hat{f}(s^0) - f(s^0)|.$$

By the first order Taylor expansion of $\hat{f}(\hat{s})$ at $\boldsymbol{\beta}_0$, we have

$$\begin{aligned}
\hat{f}(\hat{s}) - \hat{f}(s^0) =&(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T \frac{1}{nh^2}\sum_{i=1}^n \nabla K\left(\frac{S_i^* - s^*}{h}\right)(\mathbf{x} - \mathbf{X}_i) \\
=&(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T \left\{\frac{1}{nh^2}\sum_{i=1}^n \nabla K\left(\frac{S_i^0 - s^0}{h}\right)\mathbf{x} - \frac{1}{nh^2}\sum_{i=1}^n \nabla K\left(\frac{S_i^0 - s^0}{h}\right)\mathbf{X}_i\right\} \\
&+ (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T \frac{1}{nh^2}\sum_{i=1}^n \left\{\nabla K\left(\frac{S_i^* - s^*}{h}\right) - \nabla K\left(\frac{S_i^0 - s^0}{h}\right)\right\}(\mathbf{x} - \mathbf{X}_i) \\
:=&(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T\{I_{11} - I_{12}\} + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T I_2,
\end{aligned}$$

where $S_i^* = \mathbf{X}_i^T\boldsymbol{\beta}^*$, $s^* = \mathbf{x}^T\boldsymbol{\beta}^*$ and $\boldsymbol{\beta}^*$ is a point between $\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}_0$.

By Theorem 2 in Hansen (2008), we have $\|\sup_{\mathbf{x}\in\mathcal{X}} |I_{11} - E[I_{11}]|\|_{\max} = O_p\left(\frac{\log n}{nh^3}\right)^{\frac{1}{2}}$ and $\|\sup_{\mathbf{x}\in\mathcal{X}} |I_{12} - E[I_{12}]|\|_{\max} = O_p\left(\frac{\log n}{nh^3}\right)^{\frac{1}{2}}$. For $E[I_{11}]$, we can derive that

$$\|\sup_{\mathbf{x}\in\mathcal{X}} |E[I_{11}]|\|_{\max} = \left\|\sup_{\mathbf{x}\in\mathcal{X}}\left|\mathbf{x}h^{-2}\int \nabla K\left(\frac{w - s^0}{h}\right) f(w)dw\right|\right\|_{\max}$$

$$
\begin{aligned}
&= \left\| \sup_{\mathbf{x}\in\mathcal{X}} \left| \mathbf{x} h^{-1} \int \nabla K(v) f(s^0 + hv) dv \right| \right\|_{\max} \\
&= O\left(\left\| \sup_{\mathbf{x}\in\mathcal{X}} \left| \mathbf{x} \int K(v) \nabla f(s^0 + hv) dv \right| \right\|_{\max} \right) \\
&= O(1).
\end{aligned}
$$

Note that the second '=' holds by changing of variable and the third '=' holds by integration by parts. And along similar lines $\| \sup_{\mathbf{x}\in\mathcal{X}} |E[I_{12}]| \|_{\max} = O(1)$. It is well known that $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\| = O_p(n^{-\frac{1}{2}})$, then by CauchySchwarz inequality and triangle inequality, we have

$$
\begin{aligned}
&\sup_{\mathbf{x}\in\mathcal{X}} |(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T \{I_{11} - I_{12}\}| \leqslant \|(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\| \| \sup_{\mathbf{x}\in\mathcal{X}} |I_{11} - I_{12}| \|_{\max} \\
\leqslant & \|(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\| \Big\{ \| \sup_{\mathbf{x}\in\mathcal{X}} |I_{11} - E[I_{11}]| \|_{\max} + \| \sup_{\mathbf{x}\in\mathcal{X}} |I_{12} - E[I_{12}]| \|_{\max} \\
&+ \| \sup_{\mathbf{x}\in\mathcal{X}} |E[I_{11}]| \|_{\max} + \| \sup_{\mathbf{x}\in\mathcal{X}} |E[I_{12}]| \|_{\max} \Big\} \\
= & O_p\left(\frac{\log n}{n^2 h^3}\right)^{\frac{1}{2}} + O_p\left(\frac{1}{n}\right)^{\frac{1}{2}}.
\end{aligned}
$$

By Assumption 1 that $\|\nabla K(w_1) - \nabla K(w_2)\| \leqslant C\|w_1 - w_2\|$ for any w_1, $w_2 \in \mathbb{R}$ and some positive constant $C < \infty$, we can derive that

$$
\begin{aligned}
\| \sup_{\mathbf{x}\in\mathcal{X}} |I_2| \|_{\max} \leqslant & \frac{C}{h^3} \sup_{\mathbf{x}\in\mathcal{X}, \mathbf{X}\in\mathcal{X}} \|\mathbf{X} - \mathbf{x}\|^2 \|\boldsymbol{\beta}^* - \boldsymbol{\beta}_0\| \\
\leqslant & \frac{C}{h^3} \sup_{\mathbf{x}\in\mathcal{X}, \mathbf{X}\in\mathcal{X}} \|\mathbf{X} - \mathbf{x}\|^2 \|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\| \\
= & O_p\left(\frac{1}{n^{\frac{1}{2}} h^3}\right),
\end{aligned}
$$

thus $\sup_{\mathbf{x}\in\mathcal{X}} |(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T I_2| \leqslant \|(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\| \| \sup_{\mathbf{x}\in\mathcal{X}} |I_2| \|_{\max} = O_p\left(\frac{1}{nh^3}\right)$.

Applying Theorem 2 in Hansen (2008) again, $\sup_{\mathbf{x}\in\mathcal{X}} |\hat{f}(s^0) - E[\hat{f}(s^0)]| = O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}}$. We then derive the bias term that $\sup_{\mathbf{x}\in\mathcal{X}} |E[\hat{f}(s^0)] - f(s^0)| = O_p(h^q)$. By changing of variable, $E[\hat{f}(s^0)] = \int \frac{1}{h} K\left(\frac{w - s^0}{h}\right) f(w) dw = \int K(v) f(s^0 + hv) dv$. And then performing the q-th order Taylor expansion for $f(s^0 + hv)$ at s^0, we can derive that $\int K(v) f(s^0 + hv) dv = f(s^0) \int K(v) dv + h f^{(1)}(s^0) \int v K(v) dv + \ldots + \frac{h^q}{q!} f^{(q)}(s^0) \int v^q K(v) dv + o_p(h^q) = f(s^0) + O_p(h^q)$ by the definition of q-th order kernel function (Hall and Marron, 1987) and Assumption 1. Thus

$$
\sup_{\mathbf{x}\in\mathcal{X}} |\hat{f}(\hat{s}) - f(s^0)| = O_p\left(\frac{\log n}{n^2 h^3}\right)^{\frac{1}{2}} + O_p\left(\frac{1}{n}\right)^{\frac{1}{2}} + O_p\left(\frac{1}{nh^3}\right) + O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}} + O_p(h^q).
$$

Under the optimal bandwidth order $h_{\text{opt}} = O_p\left(n^{-\frac{1}{2q+1}}\right)$, the dominant term is $O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}}$. Let $l(s) = m(s)f(s)$ and $\hat{l}(s) = (nh)^{-1}\sum_{i=1}^{n} K_h(S_i, s) * 2A_iY_i$, then follow the similar arguments before, we have $\sup\limits_{\mathbf{x}\in\mathcal{X}} |\hat{l}(\hat{s}) - l(s^0)| = O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}}$. Finally,

$$\begin{aligned}
\sup_{\mathbf{x}\in\mathcal{X}} |\hat{m}(\hat{s}) - m(s^0)| =& \sup_{\mathbf{x}\in\mathcal{X}} \left| \frac{\hat{l}(\hat{s})}{\hat{f}(\hat{s})} - \frac{l(s^0)}{f(s^0)} \right| \\
\leqslant& \sup_{\mathbf{x}\in\mathcal{X}} \left| \frac{\hat{l}(\hat{s}) - l(s^0)}{\hat{f}(\hat{s})} \right| + \sup_{\mathbf{x}\in\mathcal{X}} \left| \frac{l(s^0)}{\hat{f}(\hat{s})f(s^0)}(\hat{f}(\hat{s}) - f(s^0)) \right| \\
=& O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}} = O_p\left(n^{-\frac{q}{2q+1}}\sqrt{\log n}\right).
\end{aligned}$$

This completes the proof. ∎

1.2 Proof of Lemma 2

Proof. Denote $S^0 = \mathbf{X}^T\boldsymbol{\beta}_0$, $S_i^0 = \mathbf{X}_i^T\boldsymbol{\beta}_0$, $\hat{S} = \mathbf{X}^T\hat{\boldsymbol{\beta}}$ and $\hat{S}_i = \mathbf{X}_i^T\hat{\boldsymbol{\beta}}$. We decompose $G(\mathbf{X}_i)$ as $G_1(\mathbf{X}_i) + G_2(\mathbf{X}_i)$ where $G_1(\mathbf{X}_i) = \mathbf{X}_i(\hat{m}(\hat{S}_i) - \hat{m}(S_i^0)) - E_{\mathbf{X}}[\mathbf{X}(\hat{m}(\hat{S}) - \hat{m}(S^0))]$ and $G_2(\mathbf{X}_1) = \mathbf{X}_i(\hat{m}(S_i^0) - m(S_i^0)) - E_{\mathbf{X}}[\mathbf{X}(\hat{m}(S^0) - m(S^0))]$.

Recall the results in the proof of Lemma 1 that

$$\hat{f}(\hat{S}) - \hat{f}(S^0) = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T\{I_{11} - I_{12}\} + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T I_2,$$

where

$$\begin{aligned}
I_{11} &= \frac{1}{nh^2}\sum_{i=1}^{n} \nabla K\left(\frac{S_i^0 - S^0}{h}\right)\mathbf{X}, \quad I_{12} = \frac{1}{nh^2}\sum_{i=1}^{n} \nabla K\left(\frac{S_i^0 - S^0}{h}\right)\mathbf{X}_i, \\
I_2 &= \frac{1}{nh^2}\sum_{i=1}^{n} \left\{\nabla K\left(\frac{S_i^* - S^*}{h}\right) - \nabla K\left(\frac{S_i^0 - S^0}{h}\right)\right\}(\mathbf{X} - \mathbf{X}_i).
\end{aligned}$$

For $l \in \{1, \ldots, p+1\}$, $P_nX_lI_{11} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{h^2}X_l\nabla K\left(\frac{(\mathbf{X}_i - \mathbf{X}_j)^T\boldsymbol{\beta}_0}{h}\right)\mathbf{X}_j$ and $P_nX_lI_{12} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{h^2}X_l\nabla K\left(\frac{(\mathbf{X}_i - \mathbf{X}_j)^T\boldsymbol{\beta}_0}{h}\right)\mathbf{X}_i$ are both V-statistics, then by Assumption 1 and the Lemma 8.4 in Newey and McFadden (1994), we can get for any $l \in \{1, \ldots, p+1\}$ and $k \in \{1, 2\}$

$$P_n\{X_lI_{1k} - X_lE[I_{1k}]\} - E_{\mathbf{X}}\{X_lI_{1k} - X_lE[I_{1k}]\} = O_p\left(\frac{1}{nh^2}\right),$$

hence $\sqrt{n}\left[P_n\left\{\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T\{I_{1k} - E[I_{1k}]\}\right\} - E_{\mathbf{X}}\left\{\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T\{I_{1k} - E[I_{1k}]\}\right\}\right] = O_p\left(\frac{1}{nh^2}\right)$ since $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\| = O_p\left(n^{-\frac{1}{2}}\right)$.

As has been derived in the proof of Lemma 1, $\left\|\sup\limits_{\mathbf{x}\in\mathcal{X}} |E[I_{1k}]|\right\|_{\max} = O(1)$. By the boundness of $\mathbf{X}$, we further have $\left\|\sup_{\mathbf{X}\in\mathcal{X}} \|\mathbf{X}\| |E[I_{1k}]|\right\|_{\max} = O(1)$. Then there exists a constant $c_1 > 0$ such that $\left\|\sup_{\mathbf{X}\in\mathcal{X}} \|\mathbf{X}\| |E[I_{1k}]|\right\|_{\max} \leqslant c_1$, then applying the Hoeffding' s inequality we have, for any $\epsilon_1 > 0$ and any M_{ϵ_1} large enough,

$$
\begin{aligned}
&\sum_{l=1}^{p+1} \mathbb{P}\left[\left|\sqrt{n}\left\{P_n X_l E[I_{1k}]\right] - E_{\mathbf{X}}\left[X_l E[I_{1k}]\right]\right\}\right| > \frac{M_{\epsilon_1}}{\sqrt{p+1}}\right] \\
\leqslant & 2(p+1)\exp\left(-\frac{2M_{\epsilon_1}^2}{(p+1)\left\|2\sup_{\mathbf{X}\in\mathcal{X}} \|\mathbf{X}\| |E[I_{1k}]|\right\|_{\max}^2}\right) \\
\leqslant & 2(p+1)\exp\left(-\frac{M_{\epsilon_1}^2}{2(p+1)c_1^2}\right) \leqslant \epsilon_1.
\end{aligned}
$$

Hence for $l \in \{1,\ldots,p+1\}$, $\|\sqrt{n}\left\{P_n X_l E[I_{1k}]\right] - E_{\mathbf{X}}\left[X_l E[I_{1k}]\right]\}\|_{\max} = O(1)$. Therefore for $k \in \{1,2\}$

$$
\sqrt{n}\left[P_n\left\{\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T E[I_{1k}]\right\} - E_{\mathbf{X}}\left\{\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T E[I_{1k}]\right\}\right] = O_p\left(n^{-\frac{1}{2}}\right).
$$

Let $\int f(x)\mathbf{P}_n(dx) = P_n f(x)$ and $\int f(x)\mathbf{P}_x(dx) = E_x[f(x)]$. By CauchySchwarz inequality and Assumption 1 (iv), we have

$$
\begin{aligned}
&\sqrt{n}\left[P_n\left\{\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T I_2\right\} - E_{\mathbf{X}}\left\{\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T I_2\right\}\right] \\
\leqslant & \sqrt{n}\sup_{\mathbf{x}\in\mathcal{X},\mathbf{X}\in\mathcal{X}}\left\{\|\mathbf{x}\|\|\mathbf{X}-\mathbf{x}\|\right\}\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\| \\
&\times \int \frac{1}{nh^2}\sum_{i=1}^n \left|\frac{(\boldsymbol{\beta}^* - \boldsymbol{\beta}_0)^T(\mathbf{X}-\mathbf{X}_i)}{h}\right| \phi\left(\frac{\boldsymbol{\beta}_0^T(\mathbf{X}-\mathbf{X}_i)}{h}\right)(\mathbf{P}_n + \mathbf{P}_{\mathbf{X}})(d\mathbf{X}) \\
\leqslant & \sqrt{n}\sup_{\mathbf{x}\in\mathcal{X},\mathbf{X}\in\mathcal{X}}\left\{\|\mathbf{x}\|\|\mathbf{X}-\mathbf{x}\|^2\right\}\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\|^2 \int \frac{1}{nh^3}\sum_{i=1}^n \phi\left(\frac{S^0 - S_i^0}{h}\right)(\mathbf{P}_n + \mathbf{P}_{\mathbf{X}})(d\mathbf{x}) \\
\leqslant & O_p\left(n^{-\frac{1}{2}}\right)\int \frac{1}{nh^3}\sum_{i=1}^n \phi\left(\frac{S^0 - S_i^0}{h}\right)(\mathbf{P}_n + \mathbf{P}_{\mathbf{X}})(d\mathbf{x}).
\end{aligned}
$$

For the first part $P_n\frac{1}{nh^3}\sum_{i=1}^n \phi\left(\frac{S^0 - S_i^0}{h}\right) = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \frac{1}{h^3}\phi\left(\frac{S_j^0 - S_i^0}{h}\right)$ is a V-statistic and by the Lemma 8.4 in Newey and McFadden (1994) again, we have

$$
\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \frac{1}{h^3}\phi\left(\frac{S_j^0 - S_i^0}{h}\right) - \frac{1}{n}\sum_{i=1}^n \frac{1}{h^3}E_{\mathbf{X}}\left[\phi\left(\frac{S^0 - S_i^0}{h}\right)\right] - \frac{1}{n}\sum_{j=1}^n \frac{1}{h^3}E_{\mathbf{x}}\left[\phi\left(\frac{S_j^0 - s^0}{h}\right)\right] \\
&+ \frac{1}{h^3}E\left[\phi\left(\frac{S^0 - s^0}{h}\right)\right] = O_p\left(\frac{1}{nh^3}\right).
\end{aligned}
$$

And for the second part, by Assumption 1 (iv)-(v), we have

$$\begin{aligned}
&\int \frac{1}{nh^3}\sum_{i=1}^{n}\phi\left(\frac{S^0-S_i^0}{h}\right)\mathbf{P}_{\mathbf{X}}(d\mathbf{x})\\
=&\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h^3}\int\phi\left(\frac{w-S_i^0}{h}\right)f(w)dw\\
\leqslant&\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h^2}\int\phi(v_i)f(S_i^0+hv_i)dv_i\\
\leqslant&\frac{1}{h^2}\sup_{s\in\mathcal{S}}f(s)\int\phi(v)dv=O_p(h^{-2}).
\end{aligned}$$

Then we apply the Hoeffding' s inequality similarly to before, for any $\epsilon_2>0$ and any M_{ϵ_2} large enough, we have

$$\begin{aligned}
&\mathbb{P}\left\{\left|\frac{1}{n}\sum_{j=1}^{n}\frac{1}{h^3}E_{\mathbf{x}}\left[\phi\left(\frac{S_j^0-s^0}{h}\right)\right]-\frac{1}{h^3}E\left[\phi\left(\frac{S^0-s^0}{h}\right)\right]\right|>\frac{M_{\epsilon_2}}{\sqrt{n}h^2}\right\}\\
\leqslant&2\exp\left(-M_{\epsilon_2}^2O(1)\right)<\epsilon_2.
\end{aligned}$$

Thus $\int\frac{1}{nh^3}\sum_{i=1}^{n}\phi\left(\frac{S^0-S_i^0}{h}\right)(\mathbf{P}_n+\mathbf{P}_{\mathbf{X}})(d\mathbf{x})=O_p\left(\frac{1}{nh^3}\right)+O_p\left(\frac{1}{h^2}\right)+O_p\left(\frac{1}{\sqrt{n}h^2}\right)$. Under the optimal bandwidth order $h_{\text{opt}}=O_p\left(n^{-\frac{1}{2q+1}}\right)$, the dominant term is $O_p\left(\frac{1}{h^2}\right)$. Hence

$$\sqrt{n}\left\{P_n\mathbf{X}(\hat{f}(\hat{S})-\hat{f}(S^0))-E_{\mathbf{X}}[\mathbf{X}(\hat{f}(\hat{S})-\hat{f}(S^0))]\right\}=O_p\left(\frac{1}{\sqrt{n}h^2}\right)=O_p\left(n^{-\frac{2q-3}{2(2q+1)}}\right).$$

Recall that $l(S)=m(S)f(S)$ and $\hat{l}(S)=(nh)^{-1}\sum_{i=1}^{n}K_h(S_i,S)*2A_iY_i$ as defined during the proof of Lemma 1. Then by Assumption 1 and follow the similar arguments, there holds $\sqrt{n}\left\{P_n\mathbf{X}(\hat{l}(\hat{S})-\hat{l}(S^0))-E_{\mathbf{X}}[\mathbf{X}(\hat{l}(\hat{S})-\hat{l}(S^0))]\right\}=O_p\left(n^{-\frac{2q-3}{2(2q+1)}}\right)$.

On the other hands, since $P_n\mathbf{X}\hat{f}(S)=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{h}\mathbf{X}_iK\left(\frac{(\mathbf{X}_j-\mathbf{X}_i)^T\boldsymbol{\beta}}{h}\right)$ is a V-statistic, then by Assumption 1 and the Lemma 8.4 in Newey and McFadden (1994), we can get

$$P_n\{\mathbf{X}\hat{f}(S^0)-\mathbf{X}E[\hat{f}(S^0)]\}-E_{\mathbf{X}}\{\mathbf{X}\hat{f}(S^0)-\mathbf{X}E[\hat{f}(S^0)]\}=O_p\left(\frac{1}{nh}\right).$$

As derived in the proof of Lemma 1, the bias term $\sup_{\mathbf{x}\in\mathcal{X}}|E[\hat{f}(s^0)]-f(s^0)|=O_p(h^q)$. And since $\mathbf{X}$ is bounded, there holds $\sup_{\mathbf{x}\in\mathcal{X}}|||\mathbf{X}|||E[\hat{f}(s^0)]-f(s^0)||=O_p(h^q)$. Hence there exists a constant $c_2>0$ such that $c_2\sup_{\mathbf{x}\in\mathcal{X}}|||\mathbf{X}|||E[\hat{f}(s^0)]-f(s^0)||\leqslant h^q$. Then applying the Hoeffding' s inequality, for any $\epsilon_3 > 0$ and any M_{ϵ_3} large enough we have

$$\sum_{l=1}^{p+1} \mathbb{P}\left[\left|\sqrt{n}\left\{P_n X_l[E[\hat{f}(S^0)] - f(S^0)] - E_{\mathbf{X}}\left[X_l[E[\hat{f}(S^0)] - f(S^0)]\right]\right\}\right| > \frac{M_{\epsilon_3} h^q}{\sqrt{p+1}}\right]$$
$$\leqslant 2(p+1)\exp\left(-\frac{2M_{\epsilon_3}^2 h^{2q}}{(p+1)(2\sup_{\mathbf{x}\in\mathcal{X}}|E[\hat{f}(s^0)] - f(s^0)|)^2}\right).$$

And then by the property of probability that $\mathbb{P}\left(\sum_{i=1}^k a_i > b\right) \leqslant \sum_{i=1}^k \mathbb{P}\left(a_i > \frac{b}{k}\right)$ for some positive integer k, we can derive that

$$\mathbb{P}\left[\left\|\sqrt{n}\left\{P_n \mathbf{X}[E[\hat{f}(S^0)] - f(S^0)] - E_{\mathbf{X}}\left[\mathbf{X}[E[\hat{f}(S^0)] - f(S^0)]\right]\right\}\right\| > M_{\epsilon_3} h^q\right]$$
$$\leqslant 2(p+1)\exp\left(-\frac{M_{\epsilon_3}^2 c_2^2}{2(p+1)}\right) \leqslant \epsilon_3.$$

Hence $\sqrt{n}\left\{P_n\mathbf{X}(\hat{f}(S^0) - f(S^0)) - E_{\mathbf{X}}[\mathbf{X}(\hat{f}(S^0) - f(S^0))]\right\} = O_p\left(\frac{1}{\sqrt{n}h}\right) + O_p\left(h^q\right)$.

Then by Assumption 1 and follow the similar arguments, there holds

$$\sqrt{n}\left\{P_n\mathbf{X}(\hat{l}(S^0) - l(S^0)) - E_{\mathbf{X}}[\mathbf{X}(\hat{l}(S^0) - l(S^0))]\right\} = O_p\left(\frac{1}{\sqrt{n}h}\right) + O_p\left(h^q\right).$$

For notation simplicity, we next write $\hat{h}(\hat{S})$ as $\hat{h}$, $\hat{h}(S^0)$ as $\hat{h}_0$ and $h(S^0)$ as h for function $h(\cdot) \in \{m(\cdot), f(\cdot), l(\cdot)\}$. To derive the order of $\sqrt{n}P_nG_1(\mathbf{X})$, we decompose $\hat{m} - \hat{m}_0$ as

$$\begin{aligned}
\hat{m} - \hat{m}_0 =& \frac{\hat{l}}{\hat{f}} - \frac{\hat{l}_0}{\hat{f}_0} \\
=& \frac{\hat{l} - \hat{l}_0}{f} + \frac{\hat{l}_0(\hat{f} - \hat{f}_0)}{f^2} - \frac{(\hat{l} - \hat{l}_0)(\hat{f} - \hat{f}_0)}{\hat{f}\hat{f}_0} - \frac{\hat{l}_0(\hat{f} - \hat{f}_0)^2}{\hat{f}f^2} \\
&- \frac{(\hat{l} - \hat{l}_0)(\hat{f}_0 - f)}{\hat{f}_0 f} + \frac{\hat{l}_0(\hat{f}_0 + f)(\hat{f} - \hat{f}_0)(\hat{f}_0 - f)}{\hat{f}\hat{f}_0 f^2} \\
=& \frac{\hat{l} - \hat{l}_0}{f} + \frac{\hat{l}_0(\hat{f} - \hat{f}_0)}{f^2} + O_p(a_{n,1}^2) + O_p(a_{n,1}a_{n,2}),
\end{aligned}$$

where $a_{n,1} = O_p\left(\frac{\log n}{n^2h^3}\right)^{\frac{1}{2}} + O_p\left(\frac{1}{n}\right)^{\frac{1}{2}} + O_p\left(\frac{1}{nh^3}\right)$ and $a_{n,2} = O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}} + O_p(h^q)$ as derived in the proof of Lemma 1. Under the optimal bandwidth order $h_{\text{opt}} = O_p\left(n^{-\frac{1}{2q+1}}\right)$, the dominant terms are $O_p\left(\frac{1}{nh^3}\right)$ for $a_{n,1}$ and $O_p\left(\frac{\log n}{nh}\right)^{\frac{1}{2}}$ for $a_{n,2}$. Thus $\sqrt{n}P_nG_1(\mathbf{X}) = O_p\left(\frac{1}{\sqrt{n}h^2}\right) + O_p(\sqrt{n}a_{n,1}^2) + O_p(\sqrt{n}a_{n,1}a_{n,2})$ which dominant term is $O_p\left(\frac{1}{\sqrt{n}h^2}\right) = O_p\left(n^{-\frac{2q-3}{2(2q+1)}}\right)$ under the optimal bandwidth order.

To derive the order of $\sqrt{n}P_nG_2(\mathbf{X})$, we decompose $\hat{m}_0 - m$ as

$$\begin{aligned}\hat{m}_0 - m =& \frac{\hat{l}_0}{\hat{f}_0} - \frac{l}{f}\\ =& \frac{\hat{l}_0 - l}{f} - \frac{l(\hat{f}_0 - f)}{f^2} - \frac{(\hat{l}_0 - l)(\hat{f}_0 - f)}{\hat{f}_0 f} + \frac{l(\hat{f}_0 - f)^2}{\hat{f}_0 f^2}\\ =& \frac{\hat{l}_0 - l}{f} - \frac{l(\hat{f}_0 - f)}{f^2} + O_p(a_{n,2}^2).\end{aligned}$$

Hence, $\sqrt{n}P_nG_2(\mathbf{X}) = O_p\left(\frac{1}{\sqrt{nh}}\right) + O_p\left(h^q\right) + O_p\left(\sqrt{n}a_{n,2}^2\right)$, in which the dominant term is $O_p\left(\sqrt{n}a_{n,2}^2\right) = O_p\left(\frac{\log n}{\sqrt{nh}}\right) = O_p\left(n^{-\frac{2q-1}{2(2q+1)}}\log n\right)$ under the optimal bandwidth order.

Therefore, under the optimal bandwidth order,

$$\sqrt{n}P_nG(\mathbf{X}) = O_p\left(n^{-\frac{2q-3}{2(2q+1)}}\right) + O_p\left(n^{-\frac{2q-1}{2(2q+1)}}\log n\right),$$

which dominant term is $O_p\left(n^{-\frac{2q-3}{2(2q+1)}}\right)$.

This completes the proof. ∎

1.3 Proof of Theorem 1

Proof. Denote $\Gamma = E[\mathbf{X}\mathbf{X}^T]$ and $\Gamma_n = P_n\mathbf{X}\mathbf{X}^T$. We have

$$\begin{aligned}\Gamma_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) =& P_n\mathbf{X}(2AY - \hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}}) - \mathbf{X}^T\hat{\boldsymbol{\theta}} + \mathbf{X}^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0))\\ =& P_n\mathbf{X}(2AY - m(\mathbf{X}^T\boldsymbol{\beta}_0) - \mathbf{X}^T\boldsymbol{\theta}_0) + P_n\mathbf{X}(m(\mathbf{X}^T\boldsymbol{\beta}_0) - \hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}})).\end{aligned}$$

Then by Lemma 1, $P_n\mathbf{X}(m(\mathbf{X}^T\boldsymbol{\beta}_0) - \hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}})) \leqslant \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\||\hat{m}(\mathbf{x}^T\hat{\boldsymbol{\beta}}) - m(\mathbf{x}^T\boldsymbol{\beta}_0)| = O_p\left(a_n\right)$. And by CLT, $\Gamma_n = \Gamma + O_p\left(n^{-\frac{1}{2}}\right)$ and $P_n\mathbf{X}(2AY - m(\mathbf{X}^T\boldsymbol{\beta}_0) - \mathbf{X}^T\boldsymbol{\theta}_0) = O_p\left(n^{-\frac{1}{2}}\right)$ since $E[\mathbf{X}(2AY - m(\mathbf{X}^T\boldsymbol{\beta}_0) - \mathbf{X}^T\boldsymbol{\theta}_0)] = 0$. Moreover, since Γ_n is invertible a.s., $\Gamma_n^{-1} = \Gamma^{-1} + O_p\left(n^{-\frac{1}{2}}\right)$. Let $\varphi(\mathbf{Z}_i) = \Gamma^{-1}\{\mathbf{X}_i[2A_iY_i - \nu(\mathbf{X}_i; \boldsymbol{\beta}_0, \boldsymbol{\theta}_0)]\}$, hence

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = P_n\varphi(\mathbf{Z}) + \Gamma^{-1}P_n\mathbf{X}(m(\mathbf{X}^T\boldsymbol{\beta}_0) - \hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}})) + O_p\left(n^{-1}\right) + O_p\left(n^{-\frac{1}{2}}a_n\right).$$

With slight abuse of notation, let $P_Nf(x) = \frac{1}{N}\sum_{j=n+1}^{n+N} f(x_j)$, and denote $\Gamma_N = P_N\mathbf{X}\mathbf{X}^T$. We can similarly derive

$$\begin{aligned}\Gamma_N(\hat{\boldsymbol{\beta}}_{sp} - \boldsymbol{\beta}_0) =& P_N\mathbf{X}(\hat{\nu}(\mathbf{X}; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}) - \mathbf{X}^T\hat{\boldsymbol{\beta}}_{sp} + \mathbf{X}^T(\hat{\boldsymbol{\beta}}_{sp} - \boldsymbol{\beta}_0))\\ =& P_N\mathbf{X}(\nu(\mathbf{X}; \boldsymbol{\beta}_0, \boldsymbol{\theta}_0) - \mathbf{X}^T\boldsymbol{\beta}_0) + P_N\mathbf{X}(\hat{\nu}(\mathbf{X}; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}) - \nu(\mathbf{X}; \boldsymbol{\beta}_0, \boldsymbol{\theta}_0))\\ =& P_N\mathbf{X}(\nu(\mathbf{X}; \boldsymbol{\beta}_0, \boldsymbol{\theta}_0) - \mathbf{X}^T\boldsymbol{\beta}_0) + P_N\mathbf{X}(\hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}}) - m(\mathbf{X}^T\boldsymbol{\beta}_0)) + \Gamma_N(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\end{aligned}$$

$$=\Gamma_N(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)+E_{\mathbf{X}}[\mathbf{X}(\hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}})-m(\mathbf{X}^T\boldsymbol{\beta}_0))]+O_p\left(N^{-\frac{1}{2}}\right).$$

Hence by Lemma 2

$$\begin{aligned}n^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_{sp}-\boldsymbol{\beta}_0)=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi(\mathbf{Z}_i)-\Gamma^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}G(\mathbf{X}_i)+O_p\left(n^{-\frac{1}{2}}\right)+O_p\left(a_n\right)+O_p\left(\frac{n}{N}\right)^{\frac{1}{2}}\\=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi(\mathbf{Z}_i)+O_p\left(n^{-\frac{1}{2}}\right)+O_p\left(\frac{n}{N}\right)^{\frac{1}{2}}+O_p\left(a_n\right)+O_p\left(b_n\right).\end{aligned}$$

Note that under the optimal bandwidth order, the dominant term of the remainder is $O_p\left(\frac{n}{N}\right)^{\frac{1}{2}}+O_p\left(b_n\right):=r_{n,N}$. Thus $r_{n,N}=o(1)$ as $n,N\to\infty$ with $N\gg n$.

This completes the proof. ∎

1.4 Proof of Theorem 2

Proof. As the decomposition derived in the proof of Theorem 1, we have

$$\Gamma_n(\hat{\boldsymbol{\theta}}_{\mathbb{K}}-\boldsymbol{\theta}_0)=P_n\mathbf{X}(2AY-m(\mathbf{X}^T\boldsymbol{\beta}_0)-\mathbf{X}^T\boldsymbol{\theta}_0)+\frac{1}{n}\sum_{k=1}^{\mathbb{K}}\sum_{i\in\mathcal{I}_k}\mathbf{X}_i(m(\mathbf{X}_i^T\boldsymbol{\beta}_0)-\hat{m}_k(\mathbf{X}_i^T\hat{\boldsymbol{\beta}}_k)).$$

And by $\|\hat{\boldsymbol{\beta}}_k-\boldsymbol{\beta}_0\|=O_p\left(n_{\mathbb{K}}^{-}\right)^{-\frac{1}{2}}$, we have $\sup_{\mathbf{X}\in\mathcal{X}}|\hat{m}_k(\mathbf{X}^T\hat{\boldsymbol{\beta}}_k)-m(\mathbf{X}^T\boldsymbol{\beta}_0)|=O_p(a_{n_{\mathbb{K}}^{-}})$ similar to Lemma 1. Hence there also holds

$$\hat{\boldsymbol{\theta}}_{\mathbb{K}}-\boldsymbol{\theta}_0=P_n\varphi(\mathbf{Z})+\Gamma^{-1}\frac{1}{n}\sum_{k=1}^{\mathbb{K}}\sum_{i\in\mathcal{I}_k}\mathbf{X}_i(m(\mathbf{X}_i^T\boldsymbol{\beta}_0)-\hat{m}_k(\mathbf{X}_i^T\hat{\boldsymbol{\beta}}_k))+O_p\left(n^{-1}\right)+O_p\left(n^{-\frac{1}{2}}a_{n_{\mathbb{K}}^{-}}\right).$$

Then we can derive

$$\Gamma_N(\hat{\boldsymbol{\beta}}_{sp,\mathbb{K}}-\boldsymbol{\beta}_0)=\Gamma_N(\hat{\boldsymbol{\theta}}_{\mathbb{K}}-\boldsymbol{\theta}_0)+\frac{1}{\mathbb{K}}\sum_{k=1}^{\mathbb{K}}E_{\mathbf{X}}[\mathbf{X}(\hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}}_k)-m(\mathbf{X}^T\boldsymbol{\beta}_0))]+O_p\left(N^{-\frac{1}{2}}\right).$$

Denote $G_k(\mathbf{X}_i)=\mathbf{X}_i(\hat{m}_k(\mathbf{X}_i^T\hat{\boldsymbol{\beta}}_k))-m(\mathbf{X}_i^T\boldsymbol{\beta}_0))-E_{\mathbf{X}}[\mathbf{X}(\hat{m}(\mathbf{X}^T\hat{\boldsymbol{\beta}}_k)-m(\mathbf{X}^T\boldsymbol{\beta}_0))]$, $\mathbb{G}_k=\frac{1}{\sqrt{n_{\mathbb{K}}}}\sum_{i\in\mathcal{I}_k}G_k(\mathbf{X}_i)$ and $\mathbb{G}_{n,\mathbb{K}}=\frac{1}{\sqrt{\mathbb{K}}}\sum_{k=1}^{\mathbb{K}}\mathbb{G}_k$, we have

$$n^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_{sp,\mathbb{K}}-\boldsymbol{\beta}_0)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi(\mathbf{Z}_i)-\Gamma^{-1}\mathbb{G}_{n,\mathbb{K}}+O_p\left(n^{-\frac{1}{2}}\right)+O_p\left(a_{n_{\mathbb{K}}^{-}}\right)+O_p\left(\frac{n}{N}\right)^{\frac{1}{2}}.$$

By the similar technique as the proof of Lemma 2, we have for any $\epsilon>0$, there exists M_ϵ large enough such that

$$\mathbb{P}(\|\mathbb{G}_{n,\mathbb{K}}\|>M_\epsilon a_{n_{\mathbb{K}}^{-}})$$

$$
\begin{aligned}
&\leqslant \sum_{k=1}^{\mathbb{K}} \sum_{l=1}^{p+1} \mathbb{P}\left(|\mathbb{G}_{k[l]}| > \frac{M_\epsilon a_{n_{\mathbb{K}}^-}}{\sqrt{\mathbb{K}(p+1)}} \right) \\
&\leqslant 2\mathbb{K}(p+1) \exp\left(- \frac{2M_\epsilon^2 a_{n_{\mathbb{K}}^-}^2}{\mathbb{K}(p+1)(2\sup_{\mathbf{X}\in\mathcal{X}} \|\mathbf{X}\|_{max} |\hat{m}_k(\mathbf{X}^T\hat{\boldsymbol{\beta}}_k) - m(\mathbf{X}^T\boldsymbol{\beta}_0)|)^2} \right) \\
&< \epsilon.
\end{aligned}
$$

Note that the first inequality holds by CauchySchwarz inequality and the property of probability, the second inequality holds by the Hoeffding' s inequality, and the last inequality holds by the boundedness of $\mathbf{X}$ and $\sup_{\mathbf{X}\in\mathcal{X}} |\hat{m}_k(\mathbf{X}^T\hat{\boldsymbol{\beta}}_k) - m(\mathbf{X}^T\boldsymbol{\beta}_0)| = O_p(a_{n_{\mathbb{K}}^-})$. Hence $\mathbb{G}_{n,\mathbb{K}} = O_p\left(a_{n_{\mathbb{K}}^-}\right)$.

This completes the proof. ∎

2 Additional results for real data application

In the main analysis, we set the propensity score to 0.5 for simplicity, as our primary focus was on the semi-supervised D-learning framework rather than propensity score estimation. However, to assess the sensitivity of our results to this choice, we conducted additional analyses comparing treatment recommendations under three different propensity score specifications:

1. Fixed propensity score ($\pi(A, \mathbf{X}) = 0.5$) (as in the main manuscript).
2. Estimated $\pi(A, \mathbf{X})$ via logistic regression model (results presented in Table S1).
3. Estimated $\pi(A, \mathbf{X})$ via Probit model (results presented in Table S2).

Table S1: Treatment recommendations ($\pi(A, \mathbf{X})$ estimated via logistic regression)

Treatment	Methods					
	SUP	NP	SP	SP.CV	KRLS	KRLS.CV
A=-1: IV Fluid Resuscitation	4999	4999	4999	5041	5028	5040
A=1: Vasopressors	2808	2808	2808	2766	2779	2767

Table S2: Treatment recommendations ($\pi(A, \mathbf{X})$ estimated via Probit model)

Treatment	Methods					
	SUP	NP	SP	SP.CV	KRLS	KRLS.CV
A=-1: IV Fluid Resuscitation	4994	4422	4995	5036	5026	5031
A=1: Vasopressors	2813	3385	2812	2771	2781	2776

The results demonstrate that the treatment recommendations are highly consistent across all three approaches. This strongly supports our decision to use a fixed propensity score of

0.5 in the main analysis, as this choice does not materially affect the conclusions of this real data application.

While propensity score misspecification can influence treatment recommendations in observational studies, our results suggest minimal impact in this specific analysis. Nevertheless, we acknowledge that our method, like traditional D-learning approaches, may be sensitive to misspecified propensity scores in general settings (Song et al., 2017; Qi and Liu, 2018). Future work will explore robust semi-supervised methods to address this limitation.

References

Hall, P. and Marron, J. S. (1987). Estimation of integrated squared density derivatives. *Statistics & Probability Letters*, 6(2):109–115.

Hansen, B. E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory*, 24(3):726–748.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245.

Qi, Z. and Liu, Y. (2018). D-learning to estimate optimal individual treatment rules. *Electronic Journal of Statistics*, 12(2):3601–3638.

Song, R., Luo, S., Zeng, D., Zhang, H. H., Lu, W., and Li, Z. (2017). Semiparametric single-index model for estimating optimal individualized treatment strategy. *Electronic journal of statistics*, 11(1):364.