Journal of Machine Learning Research 21 (2020) 1-67 Submitted 1/20; Revised 5/20; Published 7/20

Breaking the Curse of Nonregularity with Subagging — Inference of the Mean Outcome under Optimal Treatment Regimes

Chengchun Shi c.shi7@lse.ac.uk
Department of Statistics
London School of Economics and Political Science
London, WC2A 2AE, UK

Wenbin Lu lu@stat.ncsu.edu
Rui Song rsong@ncsu.edu
Department of Statistics
North Carolina State University
Raleigh, NC 27695, USA

Editor: Xiaotong Shen

Abstract

Precision medicine is an emerging medical approach that allows physicians to select the treatment options based on individual patient information. The goal of precision medicine is to identify the optimal treatment regime (OTR) that yields the most favorable clinical outcome. Prior to adopting any OTR in clinical practice, it is crucial to know the impact of implementing such a policy. Although considerable research has been devoted to estimating the OTR in the literature, less attention has been paid to statistical inference of the OTR. Challenges arise in the nonregular cases where the OTR is not uniquely defined. To deal with nonregularity, we develop a novel inference method for the mean outcome under an OTR (the optimal value function) based on subsample aggregating (subagging). The proposed method can be applied to multi-stage studies where treatments are sequentially assigned over time.

Bootstrap aggregating (bagging) and subagging have been recognized as effective variance reduction techniques to improve unstable estimators or classifiers (Bühlmann and Yu, 2002). However, it remains unknown whether these approaches can yield valid inference results. We show the proposed confidence interval (CI) for the optimal value function achieves nominal coverage. In addition, due to the variance reduction effect of subagging, our method enjoys certain statistical optimality. Specifically, we show that the mean squared error of the proposed value estimator is strictly smaller than that based on the simple sample-splitting estimator in the nonregular cases. Moreover, under certain conditions, the length of our proposed CI is shown to be on average shorter than CIs constructed based on the existing state-of-the-art method (Luedtke and van der Laan, 2016) and the "oracle" method which works as well as if an OTR were known. Extensive numerical studies are conducted to back up our theoretical findings.

Keywords: optimal (dynamic) treatment regime; optimal value function; precision medicine; subsample aggregating; nonregularity

1. Introduction

Precision medicine is an emerging medical approach that allows physicians to select the treatment options based on individual patient information. In contrast to the "one-size-fits-all" approach, precision medicine proposes to identify the optimal treatment regime (OTR) that yields the most favorable clinical outcome. A treatment regime is a function that maps a patient's baseline covariates to the space of available treatment options. For chronic diseases such as cancer and diabetes, treatment of patients involves a series of decisions. In these applications, it is of considerable interest to estimate the optimal dynamic treatment regime (ODTR) that consists of a list of decision rules for assigning treatment based on a patient's covariates and treatment history.

In the literature, considerable research has been devoted to estimating the OTR (or ODTR). Some popular methods include Q-learning (Watkins and Dayan, 1992; Chakraborty et al., 2010), A-learning (Robins et al., 2000; Murphy, 2003; Shi et al., 2018a), policy search methods (Zhang et al., 2012, 2013a), outcome weighted learning (Zhao et al., 2012, 2015), concordance-assisted learning (Fan et al., 2017; Liang et al., 2018), maximin projection learning (Shi et al., 2018b) and decision list-based methods (Zhang et al., 2015, 2018). Prior to adopting any OTR in clinical practice, it is crucial to know the impact of implementing such a policy. This requires to evaluate the mean outcome in the population under an OTR, i.e, the optimal value function. The inference of the optimal value function helps us to evaluate whether the OTR can lead to a clinically meaningful increment value compared to fixed treatment regimes.

Despite the popularity of estimating the OTR, less attention has been paid to statistical inference of the optimal value function. This is an extremely challenging task in the non-regular cases where there is a positive probability that the interaction between treatment and covariates (i.e, the contrast function) is equal to zero. The nonregularity occurs when the treatment is neither beneficial nor harmful for a subset of patients in the population. Restricting inference to the regular cases is limited since it requires the optimal treatment to be uniquely defined for nearly all patients. The main challenge lies in that the OTR is unknown and needs to be estimated from the data. Consider the following naive method that first estimates the OTR and then evaluates its mean outcome based on the augmented inverse propensity-score weighted estimator (AIPWE) for the value function (Zhang et al., 2012, 2013a). The validity of such a procedure relies on the estimated treatment regime being consistent to a unique OTR. However, this condition is typically violated in the nonregular cases (see Section 2.2 for details).

Chakraborty et al. (2014) considered inference for the value of an estimated OTR using the m-out-of-n bootstrap. The CI based on this method is valid in the nonregular cases when m grows to infinity at a rate slower than n. However, the length of the CI shrinks at a rate of $m^{-1/2}$. As a result, such CI will be much wider than the CI of our proposed procedure which shrinks at a rate of $n^{-1/2}$. Luedtke and van der Laan (2016) proposed an online one-step estimator that is $n^{-1/2}$-consistent to the optimal value function. Their method mimics the online prediction algorithms and recursively updates the initial estimated OTR and value function using new observations. Based on the online one-step estimator, they developed a valid inference procedure. However, their procedure relies on a data ordering.

The online one-step estimator can be sensitive to the order of the data, especially when the sample size is small.

In this paper, we develop a novel inference method for the optimal value function based on subsample aggregating (subagging) and refitted cross-validation. Specifically, we estimate the OTR based on a random subsample of the data and evaluate its value based on the remaining data using AIPWE. We then iterate this procedure multiple times. Our final estimator is defined as an average of all value estimators. Bootstrap aggregating (bagging) and subagging have been recognized as effective variance reduction techniques to improve unstable estimators or classifiers (Bühlmann and Yu, 2002). However, it remains unknown whether these procedures can yield valid inference results. We show the proposed estimator is asymptotically normal even in nonregular cases. We further provide a consistent estimator for its asymptotic variance and derive a Wald-type CI for the optimal value function.

At nonregular cases, the estimated OTR might fluctuate randomly and does not converge to a fixed function, even as the sample size grows to infinity. Subagging averages over estimated OTRs computed based on different subsamples, resulting in a smoothed decision rule, yielding smaller variance and mean squared error. Due to the variance reduction effect of subagging, our method enjoys certain statistical optimality. Specifically, we prove that:

- The mean squared error of the proposed value estimator is strictly smaller than that based on the simple sample-splitting method in the nonregular cases;
- The length of our proposed CI is on average shorter than the CI constructed based on the online one-step method (Luedtke and van der Laan, 2016) in the nonregular cases;
- Our proposed CI is asymptotically narrower than the CI of the "oracle" method which works as well as if an OTR were known in the nonregular cases.

Moreover, the proposed method can be applied to multi-stage studies to evaluate the mean outcome under an ODTR. Specifically, we show that our proposed value estimator is asymptotically unbiased and our CI achieves nominal coverage, as long as the estimated contrast function at each stage satisfies certain convergence rates and the true contrast function at each stage satisfies certain margin conditions (see Appendix A). The number of treatment stages is allowed to be an arbitrary fixed integer.

The rest of the paper is organized as follows. In Section 2, we introduce our inference procedure in a point treatment study. In Section 3, we discuss the asymptotic optimality of our proposed method. Section 4 contains the extension to multi-stage studies. Simulation studies are conducted in Section 5. We apply the proposed method to a real dataset in Section 6, followed by a discussion section. All proofs are provided in the appendix.

2. Point Treatment Study

2.1. Optimal Treatment Regime and Optimal Value Function

We begin by considering a single stage study with two treatments. Let $\boldsymbol{X}_0 \in \mathbb{X}$ be a patient's baseline covariates, $A_0 \in \{0, 1\}$ denote the treatment a patient receives, and Y_0 denote a patient's clinical outcome (the larger the better by convention). The subscript 0 indicates that these are population variables. A treatment regime $d(\cdot)$ is a deterministic function that maps $\mathbb{X}$ to $\{0,1\}$. Let $Y_0^*(0)$ and $Y_0^*(1)$ be a patient's potential outcomes, representing the response he/she would get if treated by treatment 0 and 1, respectively. In addition, define the potential outcome

$$Y_0^*(d) = Y_0^*(0)\{1 - d(\boldsymbol{X}_0)\} + Y_0^*(1)d(\boldsymbol{X}_0),$$

representing the response a patient would have if treated according to a treatment regime d. Let $V(d) = \mathrm{E}\{Y_0^*(d)\}$. An OTR d^{opt} is defined as the maximizer of the expected potential outcome $V(d)$ among the set of all possible treatment regimes, i.e,

$$d^{opt} \in \arg\max_d V(d).$$

The OTR may not be unique. Let $\mathcal{D}^{opt}$ denote the set of all OTRs,

$$\mathcal{D}^{opt} = \{d_0 : V(d_0) = \max_d V(d)\}.$$

Assume the following two assumptions hold.

(A1.) Stable unit treatment value assumption (SUTVA): $Y_0 = (1 - A_0)Y_0^*(0) + A_0 Y_0^*(1)$.

(A2.) No unmeasured confounders: $Y_0^*(0), Y_0^*(1) \perp\!\!\!\perp A_0 | \boldsymbol{X}_0$.

SUTVA requires the outcome of each patient to depend on his (or her) own treatment only. In other words, there is no interference effect between patients. The no unmeasured confounders assumption automatically holds in randomized studies. These two assumptions guarantee the OTR is identifiable from the observed data. Specifically, define the conditional mean functions $h(a, \boldsymbol{x}) \equiv \mathrm{E}(Y_0 | A_0 = a, \boldsymbol{X}_0 = \boldsymbol{x})$ and the contrast function

$$\tau(\boldsymbol{x}) \equiv h(1, \boldsymbol{x}) - h(0, \boldsymbol{x}).$$

The following lemma relates OTR to the function $\tau(\cdot)$.

Lemma 1 *Let $\mathbb{X}_1 = \{\boldsymbol{x} \in \mathbb{X} : \tau(\boldsymbol{x}) > 0\}$ and $\mathbb{X}_2 = \{\boldsymbol{x} \in \mathbb{X} : \tau(\boldsymbol{x}) < 0\}$. Assume (A1), (A2) hold, and $E|\tau(\boldsymbol{X}_0)| < \infty$. Then, for any $d \in \mathcal{D}^{opt}$,*

$$Pr(\boldsymbol{X}_0 \in \mathbb{X}_1 \cap \mathbb{X}_{2,d}) = 0 \quad \text{and} \quad Pr(\boldsymbol{X}_0 \in \mathbb{X}_2 \cap \mathbb{X}_{1,d}) = 0, \tag{1}$$

where $\mathbb{X}_{1,d} = \{\boldsymbol{x} \in \mathbb{X} : d(\boldsymbol{x}) = 1\}$ and $\mathbb{X}_{2,d} = \{\boldsymbol{x} \in \mathbb{X} : d(\boldsymbol{x}) = 0\}$. Conversely, for any treatment regime d satisfying Equation 1, we have $d \in \mathcal{D}^{opt}$.

Lemma 1 implies that $d^{opt,0} \in \mathcal{D}^{opt}$ where

$$d^{opt,0}(\boldsymbol{x}) = \mathbb{I}\{\tau(\boldsymbol{x}) > 0\}, \qquad \forall \boldsymbol{x} \in \mathbb{X}, \tag{2}$$

where $\mathbb{I}(\cdot)$ stands for the indicator function. Let $V_0 = \max_d V(d) = V(d^{opt,0})$. Our objective is to construct confidence intervals (CIs) for V_0 given the observed data $\{O_i = (\boldsymbol{X}_i, A_i, Y_i), i = 1, \dots, n\}$ that are i.i.d copies of $O_0 = (\boldsymbol{X}_0, A_0, Y_0)$.

To estimate V_0, we need to estimate the OTR first. For a given subset $\mathcal{I} \subseteq \{1, \dots, n\}$, let $\widehat{d}_{\mathcal{I}}$ denote the estimated OTR based on the sub-dataset $\{O_i\}_{i \in \mathcal{I}}$. One way to derive the OTR is to consider the class of plug-in classifiers, i.e, $\widehat{d}_{\mathcal{I}}(\cdot) = \mathbb{I}\{\widehat{\tau}_{\mathcal{I}}(\cdot) > 0\}$ where $\widehat{\tau}_{\mathcal{I}}(\cdot)$ stands for the estimated contrast function based on $\{O_i\}_{i \in \mathcal{I}}$. Alternatively, we may apply some weighted classification procedure such as outcome weighted learning to directly compute $\widehat{d}_{\mathcal{I}}$. Let $\mathcal{I}^c$ denote the complement of $\mathcal{I}$ and $|\mathcal{I}|$ be its cardinality.

2.2. The Challenge in the Nonregular Cases

Before presenting the proposed method, we briefly discuss the challenge of constructing CIs for V_0. We begin by introducing a robust estimator for the value under a fixed decision rule. For any $a = 0, 1$ and any $\boldsymbol{x} \in \mathbb{X}$, let $\pi(a, \boldsymbol{x}) = \Pr(A_0 = a | \boldsymbol{X}_0 = \boldsymbol{x})$ denote the propensity score function. For $i = 0, 1, \ldots, n$ and any function $d : \mathbb{X} \to \mathbb{R}$, define

$$\begin{aligned}\psi_i(d, \pi^*, h^*) &= \frac{A_i d(\boldsymbol{X}_i) + (1 - A_i)\{1 - d(\boldsymbol{X}_i)\}}{\pi^*(A_i, \boldsymbol{X}_i)}\{Y_i - h^*(A_i, \boldsymbol{X}_i)\} \\ &+ d(\boldsymbol{X}_i) h^*(1, \boldsymbol{X}_i) + \{1 - d(\boldsymbol{X}_i)\} h^*(0, \boldsymbol{X}_i).\end{aligned}$$

Zhang et al. (2012) proposed to use the augmented inverse propensity-score weighted estimator (AIPWE) $\sum_{i=1}^n \psi_i(d, \pi^*, h^*)/n$ for $V(d)$. AIPWE is robust to the misspecification of the outcome regression model h. Specifically, it is consistent to $V(d)$ when either π^* or h^* is consistent. This is referred to as the doubly-robustness property of AIPWE (see e.g., Bang and Robins, 2005).

For illustration purposes, we assume functions π and h are known in this section and Section 2.3. In Section 2.4, we allow these functions to be estimated from the observed dataset. Let $\mathcal{I}_0 = \{1, \ldots, n\}$. Based on AIPWE, it's temping to consider the estimator $n^{-1} \sum_{i=1}^n \psi_i(\widehat{d}_{\mathcal{I}_0}, \pi, h)$. Such a plug-in estimator is consistent to V_0 under certain conditions on $\widehat{d}_{\mathcal{I}_0}$. However, it might not have a tractable limiting distribution in the nonregular cases where $\Pr\{\tau(\boldsymbol{X}_0) = 0\} > 0$. To better illustrate this, notice that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \{\psi_i(\widehat{d}_{\mathcal{I}_0}, \pi, h) - V_0\} = \underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^n \{\psi_i(d^{opt,0}, \pi, h) - V_0\}}_{\eta_1}$$

$$+ \underbrace{\sqrt{n}\{V(\widehat{d}_{\mathcal{I}_0}) - V_0\}}_{\eta_2} + \underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^n \{\psi_i(\widehat{d}_{\mathcal{I}_0}, \pi, h) - \psi_i(d^{opt,0}, \pi, h) - V(\widehat{d}_{\mathcal{I}_0}) + V_0\}}_{\eta_3}.$$

Since $\mathrm{E}\psi_i(d^{opt,0}, \pi, h) = V_0$, η_1 corresponds to a sum of i.i.d mean-zero random variables and is asymptotically normal. Under certain mild conditions, we have $\eta_2 = o_p(1)$. Assume $\eta_3 = o_p(1)$. Then it follows that $n^{-1/2} \sum_{i=1}^n \{\psi_i(\widehat{d}_{\mathcal{I}_0}, \pi, h) - V_0\}$ is asymptotically normal. However, as commented in the introduction, the condition $\eta_3 = o_p(1)$ relies on $|\widehat{d}_{\mathcal{I}_0}(\boldsymbol{X}_0) - \mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}| \stackrel{P}{\to} 0$. For any $\boldsymbol{x}$ that satisfies $\tau(\boldsymbol{x}) = 0$, the estimated OTR $\widehat{d}_{\mathcal{I}}(\boldsymbol{x})$ is "unstable" in the sense that it might recommend different treatments due to sufficiently small changes in the data. When $\Pr\{\tau(\boldsymbol{X}) = 0\} > 0$, $\widehat{d}_{\mathcal{I}_0}(\boldsymbol{X}_0)$ might not converge to a deterministic quantity. The condition $\eta_3 = o_p(1)$ is thus violated. Subagging remedies this issue by averaging across estimated OTRs from subsamples as illustrated in the next section.

Alternatively, one may apply sample-split estimation to allow for valid inference in the non-regular cases. Specifically, let $\mathcal{I}_*$ be a random subsample of $\mathcal{I}_0$ with $|\mathcal{I}_*| \approx n/2$. The estimator $|\mathcal{I}_*^c|^{-1} \sum_{i \in \mathcal{I}_*^c} \psi_i(\widehat{d}_{\mathcal{I}_*}, \pi, h)$ is consistent to V_0. Moreover, its asymptotic variance can be consistently estimated by the sampling variance estimator. However, such an estimator is not efficient since the value function is evaluated based only on samples in $\mathcal{I}_*^c$. To improve the estimation efficiency, one can also average two such value estimators, i.e,

$$\eta_4 \equiv \frac{1}{2|\mathcal{I}_*|}\sum_{i\in\mathcal{I}_*}\psi_i(\widehat{d}_{\mathcal{I}_*^c},\pi,h) + \frac{1}{2|\mathcal{I}_*|^c}\sum_{i\in\mathcal{I}_*^c}\psi_i(\widehat{d}_{\mathcal{I}_*},\pi,h).$$

However, in the nonregular cases, these two value estimators can be correlated with each other. It remains challenging to consistently estimate the asymptotic variance of η_4.

2.3. Subsample Aggregation and Sample-Split Estimation

We use a shorthand and write $\psi_i(d) = \psi_i(d,\pi,h)$ for any $i \in \{0,1,\dots,n\}$ and $d(\cdot)$. To obtain valid CI for the optimal value function, we apply sample-split estimation with subagging. More specifically, we estimate V_0 by averaging across all single sample-splitting estimators from subsamples of size s_n, i.e,

$$\widehat{V}_\infty^* = \frac{1}{\binom{n}{s_n}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}}\left(\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}})\right),$$

where s_n is some diverging sequence. Compared to the single sample-splitting estimator $|\mathcal{I}_*^c|^{-1}\sum_{i\in\mathcal{I}_*^c}\psi_i(\widehat{d}_{\mathcal{I}_*})$, such aggregation helps reduce variance in the nonregular cases (see Section 3.1 for detailed explanation).

Let $\mathcal{I}_{(-i)} = \mathcal{I}_0 - \{i\}$, we can rewrite $\widehat{V}_\infty^*$ as

$$\widehat{V}_\infty^* = \frac{1}{(n-s_n)\binom{n}{s_n}}\sum_{i=1}^n\sum_{\substack{\mathcal{I}\in\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}}\psi_i(\widehat{d}_{\mathcal{I}}) = \frac{1}{n}\sum_{i=1}^n\frac{1}{\binom{n-1}{s_n}}\sum_{\substack{\mathcal{I}\in\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}}\psi_i(\widehat{d}_{\mathcal{I}}).$$

Notice that $\psi_i(d)$ is linear in d. Consider the aggregated OTR

$$\widehat{d}_{s_n,\infty}^{(-i)}(\boldsymbol{x}) = \binom{n-1}{s_n}^{-1}\sum_{\substack{\mathcal{I}\in\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}}\widehat{d}_{\mathcal{I}}(\boldsymbol{x}). \tag{3}$$

It follows that $\widehat{V}_\infty^* = \sum_{i=1}^n\psi_i(\widehat{d}_{s_n,\infty}^{(-i)})/n$. Based on the ANOVA decomposition (see Section C.1 for details), we can show that

$$\mathrm{E}|\widehat{d}_{s_n,\infty}^{(-i)}(\boldsymbol{x}) - d_{s_n}(\boldsymbol{x})|^2 = O(s_n/n), \tag{4}$$

where the big-O term is uniform in $\boldsymbol{x}$, and $d_{s_n}(\boldsymbol{x}) = \mathrm{E}\widehat{d}_{\{1,2,\dots,s_n\}}(\boldsymbol{x})$. When $s_n = o(n)$, $\widehat{d}_{s_n,\infty}^{(-i)}$ is consistent to d_{s_n} and we might expect that

$$\widehat{V}_\infty^* \xrightarrow{P} \frac{1}{n}\sum_{i=1}^n\psi_i(d_{s_n}) \equiv \eta_5. \tag{5}$$

With a more refined analysis, we show in Theorem 2 that $\widehat{V}_\infty^* = \eta_5 + o_p(n^{-1/2})$. Notice that η_5 corresponds to the estimated value function under the smoothed decision rule $d_{s_n}(\cdot)$

	B-spline	KRR	tree-based	RF
κ_0	$\frac{2m}{(2m+1)}$	$\frac{4}{5}$	$\frac{\log(n)\log(p)}{k}$	$\frac{3}{3+4S\log(2)}$
κ^*	$\frac{m(\alpha+2)}{(2m+1)(\alpha+1)}$	$\frac{4(\alpha+2)}{5(2\alpha+2)}$	$\frac{(2\alpha+1)\log(n)\log(p)}{k(2\alpha+2)}$	$\frac{3(2\alpha+1)}{(\alpha+1)\{6+8S\log(2)\}}$
notations or conditions	$\mathbb{X}=[0,1]$, h m-th differentiable	h belongs to certain RKHS	k = minimum leaf size, p =dimension of covariates	S = number of important variables
references	Equation (8), Zhou et al. (1998)	Equations (10)-(14), Zhang et al. (2013b)	Equation (8), Wager and Walther (2015)	Corollary 6, Biau (2012)

Table 1: Values of κ_0 and κ^* for various methods.

which is a deterministic function of $\boldsymbol{x}$. As a result, $\sqrt{n}\widehat{V}_\infty^*$ is asymptotically equivalent to a sum of i.i.d mean zero random variables and can be used to construct the CI for V_0. Below, we formally establish these results. Let $\mathcal{I}_{(j)}=\{1,2,\dots,j\}$ for any j. We need the following conditions.

(A3) Assume there exists some positive constant c_0 such that $\inf_{\boldsymbol{x}\in\mathbb{X},a=0,1}\pi(a,\boldsymbol{x})\ge c_0$.

(A4) Assume there exists some constant $\kappa^*>1/2$ such that

$$\mathrm{E}V(\widehat{d}_{\mathcal{I}_{(j)}})=V_0+O(|\mathcal{I}_{(j)}|^{-\kappa^*}).$$

Condition (A4) requires the value function under an estimated OTR to satisfy certain convergence rates. When the plug-in classifier $\mathbb{I}(\widehat{\tau}_{\mathcal{I}}(\cdot)>0)$ is used, (A4) can be replaced by the following two conditions.

(A5) Assume there exist some positive constants $\bar{c}$, α and δ_0 such that

$$\Pr(0<|\tau(\boldsymbol{X}_0)|<t)\le\bar{c}t^\alpha,\qquad \forall 0<t\le\delta_0.$$

(A6) Assume there exists some constant $\kappa_0>(\alpha+2)/(2\alpha+2)$ such that

$$\mathrm{E}|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)-\tau(\boldsymbol{X}_0)|^2=O(|\mathcal{I}_{(j)}|^{-\kappa_0}).$$

Condition (A5) is very similar to the margin assumption from Audibert and Tsybakov (2007). It holds with $\alpha=1$ when $\tau(\boldsymbol{X}_0)$ has a bounded probability density function near 0. In (A6), we assume the estimated contrast function shall satisfy certain convergence rates. These rates are available for most often used nonparametric approaches including spline methods (Zhou et al., 1998), kernel ridge regression (KRR, Steinwart and Christmann, 2008; Zhang et al., 2013b), tree-based methods (Wager and Walther, 2015) and random forests (RF, Biau, 2012). See Table 1 for details. When (A5) and (A6) hold, we can show that (A4) holds with $\kappa^*=\kappa_0(1+\alpha)/(2+\alpha)$ (see Theorem 3.1 in Qian and Murphy, 2011).

We present our main results below. For any two sequences $\{a_n\},\{b_n\}$, we write $a_n\asymp b_n$ if there exist some universal constants c, $C>0$ such that $cb_n\le a_n\le Cb_n$. The notation $a_n\gg b_n$ means $\lim_n b_n/a_n=0$.

Theorem 2 *Assume (A1)-(A4) hold,* $\max_{a\in\{0,1\}}E\{Y_0^*(a)\}^2<\infty$, *and* s_n *satisfies* $s_n=o(n)$ *and* $s_n\gg n^{1/(2\kappa^*)}$. *Then, we have*

$$\widehat{V}_\infty^*=\eta_5+o_p(n^{-1/2})\ \ \textit{and}\ \ V_0=E\eta_5+o(n^{-1/2}),$$

where η_5 *is defined in Equation* (5).

Suppose we set $s_n \asymp n/\log(n)$. The condition $s_n \gg n^{1/(2\kappa^*)}$ in Theorem 2 is then automatically satisfied. Theorem 2 implies that $\sqrt{n}(\widehat{V}_\infty^* - V_0) = \sqrt{n}(\eta_5 - \mathrm{E}\eta_5) + o_p(1)$. Let $\sigma_{s_n}^2 = \mathrm{Var}\{\psi_0(s_n)\}$. Assume $\liminf_n \sigma_{s_n} > 0$. By the central limit theorem, we have

$$\frac{\sqrt{n}(\widehat{V}_\infty^* - V_0)}{\sigma_{s_n}} \stackrel{d}{\to} N(0,1).$$

For any $z_1, \dots, z_n \in \mathbb{R}$, let $\widehat{s.e}^2(\{z_i\}_{i=1}^n)$ denote the sample variance estimator, i.e, $\widehat{s.e}^2(\{z_i\}_{i=1}^n) = \sum_{i=1}^n (z_i - \bar{z})^2/(n-1)$ where $\bar{z} = \sum_{i=1}^n z_i/n$. The asymptotic variance $\sigma_{s_n}^2$ can be consistently estimated by

$$\widehat{\sigma}_\infty^{*2} = \widehat{s.e}^2\left(\left\{\psi_i(\widehat{d}_{s_n,\infty}^{(-i)})\right\}_{i=1}^n\right),$$

where $d_{s_n,\infty}^{(-i)}$ is defined in Equation (3).

Notice that it is intractable to compute $\widehat{d}_\mathcal{I}$ over all possible size s_n subsamples of the training data. In practice, we can estimate $\widehat{V}_\infty^*$ based on Monte Carlo approximations. More specifically, for a sufficiently large integer B, set

$$\widehat{V}_B^* = \frac{1}{B}\sum_{b=1}^B \frac{1}{|\mathcal{I}_b^c|}\sum_{i\in\mathcal{I}_b^c} \psi_i(\widehat{d}_{\mathcal{I}_b}), \tag{6}$$

where the subsets $\mathcal{I}_1, \dots, \mathcal{I}_B$ are drawn uniformly from the set

$$\mathcal{S}_{N_0,s_n} = \left\{\mathcal{I} \subseteq \mathcal{I}_0 : |\mathcal{I}| = s_n, N_0 \le \sum_{i\in\mathcal{I}} A_i \le s_n - N_0\right\},$$

for some positive integer N_0. Here, the constraints $N_0 \le \sum_{i\in\mathcal{I}} A_i \le s_n - N_0$ guarantee that the function $\tau(\cdot)$ is estimable based on the sub-dataset $\{O_i\}_{i\in\mathcal{I}}$. Define

$$\widehat{\sigma}_B^{*2} = \widehat{s.e}^2\left(\left\{\psi_i(\widehat{d}_{s_n,B}^{(-i)})\right\}_{i=1}^n\right), \tag{7}$$

where

$$\widehat{d}_{s_n,B}^{(-i)}(\boldsymbol{X}_i) = \frac{1}{n^{(i)}}\sum_{b:\{i\notin\mathcal{I}_b\}} \widehat{d}_{\mathcal{I}_b}(\boldsymbol{X}_i),$$

and $n^{(i)} = \sum_{b=1}^B \mathbb{I}(i \notin \mathcal{I}_b)$. The corresponding two-sided CI for V_0 is given by

$$\left[\widehat{V}_B^* - \frac{2z_{\alpha/2}\widehat{\sigma}_B^*}{\sqrt{n}}, \widehat{V}_B^* + \frac{2z_{\alpha/2}\widehat{\sigma}_B^*}{\sqrt{n}}\right].$$

2.4. Unknown propensity score and conditional mean functions

In practice, the conditional mean function $h(\cdot,\cdot)$ is unknown to us. In observational studies, the propensity score function $\pi(\cdot,\cdot)$ also needs to be estimated from the data. Parametric models are commonly used to estimate these functions. The resulting value estimator is consistent when either $h(\cdot,\cdot)$ or $\pi(\cdot,\cdot)$ is correctly specified. To avoid model misspecification and gain efficiency, we propose to estimate these function nonparametrically and use a sample-splitting method to construct AIPWE. The use of sample-splitting helps reduce the bias of AIPWE resulting from the biases of the estimated propensity score and conditional mean functions.

For any $\mathcal{I} \subseteq \mathcal{I}_0$, denote by $\widehat{h}_{\mathcal{I}}$ and $\widehat{\pi}_{\mathcal{I}}$ the corresponding estimators for h and π, based on the sub-dataset $\{O_i\}_{i\in\mathcal{I}}$. For simplicity, assume $n - s_n = 2t_n$ for some integer $t_n > 0$. We detail our procedure in the following algorithm.

Step 1. Input observations $\{O_i\}_{i=1,\dots,n}$, $0 < \alpha < 1$, and integers s_n, N_0, B.

Step 2. For $b = 1, \dots, B$,

(i) Draw a subset $\mathcal{I}_b$ from $\mathcal{S}_{N_0,s_n}$ uniformly at random.

(ii) Randomly partition $\mathcal{I}_b^c$ into 2 disjoint subsets $\mathcal{I}_b^{c(1)}$ and $\mathcal{I}_b^{c(2)}$ of equal sizes t_n.

(iii) For $j = 1, 2$, let $\mathcal{I}_b^{(j)} = \mathcal{I}_b \cup \mathcal{I}_b^{c(j)}$. Obtain the estimators $\widehat{d}_{\mathcal{I}_b}$, $\widehat{\pi}_{\mathcal{I}_b^{(1)}}$, $\widehat{\pi}_{\mathcal{I}_b^{(2)}}$, $\widehat{h}_{\mathcal{I}_b^{(1)}}$ and $\widehat{h}_{\mathcal{I}_b^{(2)}}$.

Step 3. Compute

$$\widehat{V}_B = \frac{1}{2B}\sum_{b=1}^{B}\left(\frac{1}{t_n}\sum_{i\in\mathcal{I}_b^{c(2)}}\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}}) + \frac{1}{t_n}\sum_{i\in\mathcal{I}_b^{c(1)}}\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(2)}},\widehat{h}_{\mathcal{I}_b^{(2)}})\right),$$

and

$$\widehat{\sigma}_B^2 = \widehat{s.e}^2\left(\left\{\frac{1}{n^{(i)}}\sum_{b=1}^{B}\sum_{j=1}^{2}\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(j)}},\widehat{h}_{\mathcal{I}_b^{(j)}})\mathbb{I}(i\notin\mathcal{I}_b^{(j)})\right\}_{i=1}^{n}\right),$$

where $n^{(i)} = \sum_{b=1}^{B}\mathbb{I}(i \notin \mathcal{I}_b)$.

Step 4. Output

$$\left[\widehat{V}_B - \frac{z_{\alpha/2}\widehat{\sigma}_B}{\sqrt{n}}, \widehat{V}_B + \frac{z_{\alpha/2}\widehat{\sigma}_B}{\sqrt{n}}\right]. \tag{8}$$

Notice that we apply a two-fold cross-validation procedure in Step 2 and Step 3. More generally, one can use K-fold cross-validation to construct the estimators $\widehat{V}_B$ and $\widehat{\sigma}_B^2$, for any fixed integer $K \geq 2$. The following theorem establishes the validity of the CI in Equation (8).

Theorem 3 *Assume $B \gg n$, $\liminf_n \sigma_{s_n} > 0$,*

$$Pr\left(\inf_{\mathcal{I}\in\mathcal{I}_0,\boldsymbol{x}\in\mathbb{X},a=0,1}\widehat{\pi}_{\mathcal{I}}(a,\boldsymbol{x}) \geq c^*\right) = 1, \tag{9}$$

for some constant $c^* > 0$. In addition, assume

$$\max_{a=0,1} E|\widehat{\pi}_{\mathcal{I}_{(j)}}(a, \boldsymbol{X}_0) - \pi(a, \boldsymbol{X}_0)|^2 = o(|\mathcal{I}_{(j)}|^{-1/2}), \tag{10}$$

$$\max_{a=0,1} E|\widehat{h}_{\mathcal{I}_{(j)}}(a, \boldsymbol{X}_0) - h(a, \boldsymbol{X}_0)|^2 = o(|\mathcal{I}_{(j)}|^{-1/2}). \tag{11}$$

Then, under the conditions in Theorem 2, we have

$$\frac{\sqrt{n}(\widehat{V}_B - V_0)}{\widehat{\sigma}_B} \stackrel{d}{\to} N(0,1).$$

In Equation (10) and (11), we require the estimated propensity score and conditional mean functions to satisfy certain convergence rates. These conditions guarantee that $\widehat{V}_B$ and $\widehat{\sigma}_B^2$ are asymptotically equivalent to $\widehat{V}_B^*$ and $\widehat{\sigma}_B^{*2}$, defined in Equation (6) and (7), respectively.

Theorem 3 shows the asymptotic normality of $\sqrt{n}(\widehat{V}_B - V_0)/\widehat{\sigma}_B$. As a result, the two-sided CI defined in Equation (8) has asymptotically nominal coverage probabilities. Moreover, it also implies that $\widehat{V}_B - z_\alpha \widehat{\sigma}_B/\sqrt{n}$ is an asymptotic $1-\alpha$ lower confidence bound for V_0.

2.5. Practical Guidance

The proposed algorithm requires selection of some hyperparameters s_n, N_0 and B. In practice, we recommend to set N_0 to a small integer such as 5 or 10. In our numerical studies, we find the resulting confidence intervals are not sensitive to this hyperparameter.

The selection of s_n is critical to our proposed method. A smaller s_n might result in a biased value estimator. However, we require $s_n = o(n)$ to guarantee its asymptotic normality. In practice, we recommend to set $s_n = \lfloor K_0 n/\log(n) \rfloor$ for some $K_0 \in [2, 5]$ to guarantee s_n is sufficiently large such that the bias of the proposed value estimator is $o(n^{-1/2})$. Here, $\lfloor z \rfloor$ denotes the largest integer smaller than or equal to z. We find such a choice of s_n works well in our simulations.

The number of subsamples B represents a trade-off. In theory, B shall be as large as possible to guarantee the validity of our CI. Yet, the computation complexity increases linearly in B. In practice, we recommend using $B = 4000$ as a good balance between computational load and inference accuracy. Such a choice of B is also being used by Efron (2014) for quantifying the uncertainty of regression parameters after model selection.

As shown in Theorem 2, the proposed value estimator corresponds to a nearly unbiased estimates of the value under the "smoothed" decision rule d_{s_n}. To implement precision medicine, we recommend to adopt the decision rule $\widehat{d}_{\mathcal{I}_0}$ learned based on all samples instead of d_{s_n}. This is because d_{s_n} might not be a deterministic rule. As such, it will cause some additional complexity in practice. Under condition (A4), the difference between values under $\widehat{d}_{\mathcal{I}_0}$ and V_0 is $o_p(n^{-1/2})$. Consequently, the proposed confidence interval is also valid for the value under $\widehat{d}_{\mathcal{I}_0}$.

3. Asymptotic Optimality

This section discusses the optimality of the proposed method. The length of the proposed CI (see (8)) is given by $L(\widehat{V}_B, \alpha) = 2z_{\alpha/2}\widehat{\sigma}_B/\sqrt{n}$. Under the given conditions in Theorem

3, the estimator $\widehat{\sigma}_B$ is consistent to σ_{s_n} and we can show

$$\sqrt{n}L(\widehat{V}_B,\alpha) = 2z_{\alpha/2}\sigma_{s_n} + o_p(1), \tag{12}$$

and

$$\sqrt{n}\mathrm{E}L(\widehat{V}_B,\alpha) = 2z_{\alpha/2}\sigma_{s_n} + o(1). \tag{13}$$

For illustration purposes, we focus on the class of plug-in classifiers $\widehat{d}_{\mathcal{I}}(\cdot) = \mathbb{I}\{\widehat{\tau}_{\mathcal{I}}(\cdot) > 0\}$ throughout this section. Before presenting our main results, we introduce the following condition.

(A7) For any $\boldsymbol{x} \in \mathbb{X}$ with $\tau(\boldsymbol{x}) = 0$, assume the following holds:

$$\Pr\left(\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) > 0\right) \to p_0(\boldsymbol{x}), \qquad \text{as } |\mathcal{I}| \to \infty,$$

for some function $p_0(\cdot)$. In addition, assume there exists some $0 < d^* \leq 1/2$ such that $d^* \leq p_0(\boldsymbol{x}) \leq 1 - d^*$ for any $\boldsymbol{x}$ that satisfies $\tau(\boldsymbol{x}) = 0$.

Suppose $\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) - \mathrm{E}\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x})$ is asymptotically normal, i.e,

$$\frac{\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) - \mathrm{E}\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x})}{\sigma^*_{|\mathcal{I}|}(\boldsymbol{x})} \stackrel{d}{\to} N(0,1), \tag{14}$$

for some sequence $\sigma^*_n(\boldsymbol{x}) \to 0$. Moreover, suppose the bias term satisfies

$$\frac{\mathrm{E}\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) - \tau(\boldsymbol{x})}{\sigma^*_{|\mathcal{I}|}(\boldsymbol{x})} \to r_0(\boldsymbol{x}), \tag{15}$$

for some bounded function $r_0(\cdot)$. Then for any $\boldsymbol{x}$ that satisfies $\tau(\boldsymbol{x}) = 0$, it follows from the definition of weak convergence that

$$\Pr\left(\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) > 0\right) = \Pr\left(\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) - \tau(\boldsymbol{x}) > 0\right) = \Pr\left(\frac{\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) - \tau(\boldsymbol{x})}{\sigma^*_{|\mathcal{I}|}(\boldsymbol{x})} > 0\right)$$
$$= \Pr\left(\frac{\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) - \mathrm{E}\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x})}{\sigma^*_{|\mathcal{I}|}(\boldsymbol{x})} > \frac{\tau(\boldsymbol{x}) - \mathrm{E}\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x})}{\sigma^*_{|\mathcal{I}|}(\boldsymbol{x})}\right) \to \Pr(N(0,1) > -r_0(\boldsymbol{x})).$$

Condition (A7) is thus satisfied.

Notice that the condition (14) holds for a wide variety of nonparametric estimators $\widehat{\tau}_{\mathcal{I}}$ computed by kernel smoothing methods (Härdle, 1990), spline methods (Zhou et al., 1998), kernel ridge regression (Zhao et al., 2016), random forests (Wager and Athey, 2018), etc. All these estimating procedures require tuning parameter selection. In Condition (15), we assume the tuning parameter is chosen such that the bias-variance ratio stabilizes. When undersmoothing is employed, we have $r_0(\boldsymbol{x}) = 0$ and hence

$$\Pr(\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}) > 0) \to \frac{1}{2}, \quad \forall \boldsymbol{x} \text{ with } \tau(\boldsymbol{x}) = 0. \tag{16}$$

For simplicity, throughout this section, we assume the following semiparametric regression model for Y_0:

$$Y_0 = h(0, \boldsymbol{X}_0) + A_0\tau(\boldsymbol{X}_0) + e_0, \tag{17}$$

where e_0 is a mean zero random error term independent of $\boldsymbol{X}_0, A_0$. Let $\sigma_0^2 = \mathrm{Var}(e_0) > 0$.

3.1. Comparison with the Single Sample-Splitting Estimator

Consider the single sample-splitting estimator,

$$\widehat{V}^{ss} \equiv \frac{1}{|\mathcal{I}_*^c|} \sum_{i \in \mathcal{I}_*^c} \psi_i(\widehat{d}_{\mathcal{I}_*}, \widehat{\pi}_{\mathcal{I}_*}, \widehat{h}_{\mathcal{I}_*}), \tag{18}$$

where $\mathcal{I}_*$ is a random subset of $\mathcal{I}_0$ with size ℓ_n. Based on such an estimator, the associated CI for V_0 is given by

$$\widehat{V}^{ss} \pm \frac{z_{\alpha/2}}{\sqrt{n-\ell_n}} \widehat{s.e}(\{\psi_i(\widehat{d}_{\mathcal{I}_*}, \widehat{\pi}_{\mathcal{I}_*}, \widehat{h}_{\mathcal{I}_*})\}_{i \in \mathcal{I}_*^c}). \tag{19}$$

In this section, we focus on comparing the mean squared error of the proposed value estimator with that of (18). Assume $\ell_n = o(n)$. According to the bias-variance decomposition, we have

$$\begin{aligned} n\text{MSE}(\widehat{V}^{ss}) &= \frac{n}{n-\ell_n} \text{EVar}[\psi_0(\widehat{d}_{\mathcal{I}_*}, \widehat{\pi}_{\mathcal{I}_*}, \widehat{h}_{\mathcal{I}_*}) | \{O_i\}_{i \in \mathcal{I}_*}] + n\{\text{E}\psi_0(\widehat{d}_{\mathcal{I}_*}, \widehat{\pi}_{\mathcal{I}_*}, \widehat{h}_{\mathcal{I}_*}) - V_0\}^2 \\ &= \text{EVar}[\psi_0(\widehat{d}_{\mathcal{I}_*}, \widehat{\pi}_{\mathcal{I}_*}, \widehat{h}_{\mathcal{I}_*}) | \{O_i\}_{i \in \mathcal{I}_*}] + n\{\text{E}\psi_0(\widehat{d}_{\mathcal{I}_*}, \widehat{\pi}_{\mathcal{I}_*}, \widehat{h}_{\mathcal{I}_*}) - V_0\}^2 + o(1). \end{aligned}$$

When $\widehat{\pi}_{\mathcal{I}_*}$ and $\widehat{h}_{\mathcal{I}_*}$ satisfy certain convergence rates, we have

$$n\text{MSE}(\widehat{V}^{ss}) = \text{EVar}[\psi_0(\widehat{d}_{\mathcal{I}_*}) | \{O_i\}_{i \in \mathcal{I}_*}] + n\{\text{E}\psi_0(\widehat{d}_{\mathcal{I}_*}) - V_0\}^2 + o(1).$$

Under (A5) and (A6), we have $\text{E}\psi_0(\widehat{d}_{\mathcal{I}_*}) = V_0 + o(n^{-1/2})$ and hence

$$n\text{MSE}(\widehat{V}^{ss}) = \text{EVar}[\psi_0(\widehat{d}_{\mathcal{I}_*}) | \{O_i\}_{i \in \mathcal{I}_*}] + o(1). \tag{20}$$

Similarly, we can show

$$n\text{MSE}(\widehat{V}_B) = \sigma_{s_n}^2 + o(1). \tag{21}$$

Theorem 4 *Assume* (17), (20), (21), *(A1)-(A3), (A5)-(A7) hold,* $s_n, \ell_n \to \infty$*. Then, we have*

$$nMSE(\widehat{V}^{ss}) - nMSE(\widehat{V}_B) \geq \frac{\sigma_0^2 (d^*)^2}{c_0^2} Pr\{\tau(\boldsymbol{X}_0) = 0\} + o(1),$$

where c_0 *is defined in (A3),* d^* *is defined in (A7) and* σ_0^2 *is defined in Equation* (17).

Theorem 4 suggests that the proposed estimator is more efficient compared to (18) in the nonregular cases. As such, the proposed confidence interval is strictly narrower than (19). In the regular cases, the two estimators are asymptotically equivalent in theory. In finite samples, however, the proposed estimator could be more efficient due to that $\widehat{V}^{ss}$ only uses $n - \ell_n$ samples to construct the value estimates.

3.2. Comparison with the Online One-Step Estimator

Let $\{\ell_n\}_n$ be a sequence of nonnegative integer with $\ell_n < n$. The online one-step estimator is defined as

$$\widehat{V}^{on} = \left(\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_{\mathcal{I}_{(j)}}^{-1}\right)^{-1} \left(\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_{\mathcal{I}_{(j)}}^{-1} \psi_{j+1}(\widehat{d}_{\mathcal{I}_{(j)}}; \widehat{\pi}_{\mathcal{I}_{(j)}}, \widehat{h}_{\mathcal{I}_{(j)}})\right),$$

where $\widetilde{\sigma}_{\mathcal{I}_{(j)}}^2$ stands for some consistent estimator of

$$\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \widehat{\pi}_{\mathcal{I}_{(j)}}, \widehat{h}_{\mathcal{I}_{(j)}}) = \mathrm{Var}\left(\psi_{j+1}(\widehat{d}_{\mathcal{I}_{(j)}}; \widehat{\pi}_{\mathcal{I}_{(j)}}, \widehat{h}_{\mathcal{I}_{(j)}}) \middle| \{O_i\}_{i\in\mathcal{I}_{(j)}}\right),$$

computed based on the observations $\{O_i\}_{i\in\mathcal{I}_{(j)}}$. Under the conditions in Theorem 2 of Luedtke and van der Laan (2016), it follows from the martingale central limit theorem that

$$\frac{\sqrt{n-\ell_n}(\widehat{V}^{on} - V_0)}{\widehat{\sigma}^{on}} \stackrel{d}{\to} N(0,1),$$

where $\widehat{\sigma}^{on} = \{\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_{\mathcal{I}_{(j)}}^{-1}/(n-\ell_n)\}^{-1}$. The corresponding two-sided CI for V_0 is given by

$$\left[\widehat{V}^{on} - z_{\alpha/2}\frac{\widehat{\sigma}^{on}}{\sqrt{n-\ell_n}}, \widehat{V}^{on} + z_{\alpha/2}\frac{\widehat{\sigma}^{on}}{\sqrt{n-\ell_n}}\right]. \tag{22}$$

Assume $\ell_n \to \infty$, under the same conditions as Theorem 3, we can show that

$$\widehat{\sigma}^{on} = \left(\frac{\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_0^{-1}(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)}{n-\ell_n}\right)^{-1} + o_p(1). \tag{23}$$

The first term on the RHS of the above expression is a random variable depending on $\{O_i\}_{i\in\mathcal{I}_{(n-1)}}$. Therefore, we focus on comparing the average length of (22) with that of our proposed CI.

When $\{\widetilde{\sigma}_{\mathcal{I}_{(j)}}\}_{j=\ell_n,\dots,n-1}$ and $\{\widetilde{\sigma}_0(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)\}_{j=\ell_n,\dots,n-1}$ are uniformly bounded from above, it follows from (23) that

$$\mathrm{E}\widehat{\sigma}^{on} = \mathrm{E}\left(\frac{\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_0^{-1}(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)}{n-\ell_n}\right)^{-1} + o(1).$$

When $\ell_n = o(n)$, the length of (22) satisfies

$$\sqrt{n}\mathrm{E}L(\widehat{V}^{on}, \alpha) = 2z_{\alpha/2}\mathrm{E}\left(\frac{\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_0^{-1}(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)}{n-\ell_n}\right)^{-1} + o(1). \tag{24}$$

The first term on the RHS of (24) is very challenging to analyze. We consider approximating it by

$$\mathrm{E}\left(\frac{\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_0^{-1}(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)}{n-\ell_n}\right)^{-1} = \left(\frac{\sum_{j=\ell_n}^{n-1} \{\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)\}^{-1/2}}{n-\ell_n}\right)^{-1} + o(1). \tag{25}$$

In Appendix D, we provide sufficient conditions on the estimated contrast function that ensure the approximation in Equation (25) is satisfied and use the kernel smoother as an example to show Equation (25) holds. Combining (24) with (25), we obtain

$$\sqrt{n}\mathrm{E}L(\widehat{V}^{on},\alpha) = 2z_{\alpha/2}\left(\frac{\sum_{j=\ell_n}^{n-1}\{\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}(j)};\pi,h)\}^{-1/2}}{n-\ell_n}\right)^{-1} + o(1). \tag{26}$$

Theorem 5 *Assume* (13)*,* (17)*,* (26)*, (A1)-(A3), (A5)-(A7) hold,* $s_n, \ell_n \to \infty$*. Then, we have*

$$\sqrt{n}EL(\widehat{V}^{on},\alpha) - \sqrt{n}EL(\widehat{V}_B,\alpha) \geq \frac{z_{\alpha/2}(d^*)^2\sigma_0^2}{c_0^2\sqrt{\tilde{c}+\sigma_0^2 c_0^{-1}}} Pr\{\tau(\boldsymbol{X}_0)=0\} + o(1),$$

where c_0 *is defined in Condition (A3),* d^* *is defined in Condition (A7), and* $\tilde{c} = Var\{h(d^{opt,0}(\boldsymbol{X}_0),\boldsymbol{X}_0)\}$*.*

Theorem 5 implies that the expected length of (22) is asymptotically larger than that of the proposed CI in the nonregular cases. The difference depends on $\Pr\{\tau(\boldsymbol{X}_0)=0\}$, which measures the degree of nonregularity. In the regular cases, we have $\widehat{V}_B = \widehat{V}^{on} + o_p(n^{-1/2})$. By Corollary 3 of Luedtke and van der Laan (2016), our proposed value estimator is asymptotically efficient.

In the following, we sketch a few lines to see why our proposed CI is narrower on average. Under the given conditions, $\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}(j)};\pi,h)$ converges to a fixed function as $j\to\infty$. Since $s_n,\ell_n\to\infty$, we have

$$\left(\frac{\sum_{j=\ell_n}^{n-1}\{\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}(j)};\pi,h)\}^{-1/2}}{n-\ell_n}\right)^{-1} = \{\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}(s_n)};\pi,h)\}^{1/2} + o(1).$$

By definition, we have $\sigma_{s_n}^2 = \widetilde{\sigma}_0^2(d_{s_n};\pi,h)$ and $d_{s_n}(\boldsymbol{x}) = \mathrm{E}\widehat{d}_{\mathcal{I}(s_n)}(\boldsymbol{x})$. The function $\widetilde{\sigma}_0^2(d;\pi,h)$ is convex in d. Therefore, it follows from Jensen's inequality that

$$\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}(s_n)};\pi,h) \geq \widetilde{\sigma}_0^2(\mathrm{E}\widehat{d}_{\mathcal{I}(s_n)};\pi,h).$$

This together with (13) and (26) yields that $\sqrt{n}\mathrm{E}L(\widehat{V}^{on},\alpha) \geq \sqrt{n}\mathrm{E}L(\widehat{V}_B,\alpha)+o(1)$.

However, it is worth mentioning that the theoretical guarantees for our proposed method make a stronger condition on the tuning parameter s_n than do the theoretical guarantees for the online one-step estimator. Specifically, we require $s_n/n\to 0$ and $s_n \gg n^{1/(2\kappa^*)}$. On the contrary, the validity of the CI in Equation 22 only requires ℓ_n/n to be bounded away from 1. As long as (A4) holds, the bias of $\widehat{V}^{on}$ is of the order $O_p(n^{-\kappa^*}) = o_p(n^{-1/2})$, which is independent of ℓ_n.

3.3. Beyond Oracle Property

In this section, we compare the proposed CI with the CI based on the oracle method. The oracle knew the set of optimal treatment regimes $\mathcal{D}^{opt}$ ahead of time and picked a single OTR d^{opt} from $\mathcal{D}^{opt}$. When functions π and h are known, we can estimate V_0 by $\sum_{i=1}^n \psi_i(d^{opt};\pi,h)/n$ for an arbitrary $d^{opt}\in\mathcal{D}^{opt}$. To deal with unknown propensity score and conditional mean functions, the oracle can construct the estimator based on the following cross-validation procedure:

Step 1 Input observations $\{O_i\}_{i\in\mathcal{I}_0}$, $0<\alpha<1$.

Step 2 Randomly partition $\mathcal{I}_0$ into 2 disjoint subsets $\mathcal{I}_1$ and $\mathcal{I}_2$ of equal sizes, assuming the sample size n is an even integer.

Step 3 Obtain the estimators $\widehat{\pi}_{\mathcal{I}_j}$ and $\widehat{h}_{\mathcal{I}_j}$ for $j=1,2$. Compute

$$\widehat{V}^{or}(d^{opt}) = \frac{1}{2|\mathcal{I}_1|}\sum_{i\in\mathcal{I}_1}\psi_i(d^{opt};\widehat{\pi}_{\mathcal{I}_2},\widehat{h}_{\mathcal{I}_2}) + \frac{1}{2|\mathcal{I}_2|}\sum_{i\in\mathcal{I}_2}\psi_i(d^{opt};\widehat{\pi}_{\mathcal{I}_1},\widehat{h}_{\mathcal{I}_1}),$$

$$\widehat{\sigma}^{or}(d^{opt}) = \left\{\frac{1}{n-1}\sum_{j=1}^{2}\sum_{i\in\mathcal{I}_j}\left(\psi_i(d^{opt};\widehat{\pi}_{\mathcal{I}_j^c},\widehat{h}_{\mathcal{I}_j^c}) - \widehat{V}^{or}(d^{opt})\right)^2\right\}^{1/2}.$$

Step 4 Output

$$\left[\widehat{V}^{or}(d^{opt}) - \frac{z_{\alpha/2}\widehat{\sigma}^{or}(d^{opt})}{\sqrt{n}}, \widehat{V}^{or}(d^{opt}) + \frac{z_{\alpha/2}\widehat{\sigma}^{or}(d^{opt})}{\sqrt{n}}\right]. \tag{27}$$

The CI in Equation (27) is valid. Under conditions (9), (10) and (11), we can show that

$$\{\widehat{\sigma}^{or}(d^{opt})\}^2 = \mathrm{Var}\{\psi_0(d^{opt})\} + o_p(1).$$

Thus, the length of (27) satisfies

$$\sqrt{n}L\{\widehat{V}^{or}(d^{opt}),\alpha\} = 2z_{\alpha/2}\sqrt{\mathrm{Var}\{\psi_0(d^{opt})\}} + o_p(1). \tag{28}$$

Theorem 6 *Assume* (12), (17), (28), *(A1)-(A3), (A5)-(A7) hold and* $s_n\to\infty$. *Assume* $\min_{a=0,1}\pi(a,\boldsymbol{x})\ge(1-d^*)^2$, $\forall\boldsymbol{x}\in\mathbb{X}$ *with* $\tau(\boldsymbol{x})=0$. *Then,*

$$\inf_{d^{opt}\in\mathcal{D}^{opt}} nL^2\{\widehat{V}^{or}(d^{opt}),\alpha\} - nL^2(\widehat{V}_B,\alpha) \ge c^{**}z_{\alpha/2}^2\sigma_0^2 Pr\{\tau(\boldsymbol{X}_0)=0\} + o(1),$$

where

$$c^{**} \equiv \inf_{\substack{a=0,1\\ \boldsymbol{x}\in\mathbb{X}:\tau(\boldsymbol{x})=0}}\left(\frac{d^*(2-d^*)}{\pi(a,\boldsymbol{x})} - \frac{(1-d^*)^2}{\pi(1-a,\boldsymbol{x})}\right)\ge 0.$$

In randomized studies, we usually have $\pi(1,\boldsymbol{x})=1-\pi(0,\boldsymbol{x})=\pi^*$ for some constant $\pi^*>0$. The condition $\min_{a=0,1}\pi(a,\boldsymbol{x})\ge(1-d^*)^2$ thus holds if $(1-d^*)^2\le\pi^*\le d^*(2-d^*)$. When (16) holds, this condition is further reduced to $1/4\le\pi^*\le 3/4$. Theorem 6 implies that the proposed CI is asymptotically narrower than (27) in the nonregular cases. As discussed in the introduction, this is due to the subagging procedure, which averages over estimated OTRs in the nonregular cases, resulting in a smoothed treatment regime $d_{s_n}(\cdot)$. To give a more formal explanation, let's assume $\tau(\boldsymbol{x})=0$ and $\pi(1,\boldsymbol{x})=\pi^*$ for any $\boldsymbol{x}$. In addition, we assume we know the true propensity score and conditional mean functions and set $\widehat{\pi}_{\mathcal{I}} = \pi^*$ and $\widehat{h}_{\mathcal{I}} = h$. Then it follows from Lemma 1 that $\mathrm{E}\widehat{V}_{\mathcal{I}_0}(d) = V_0$ for any regime d. By (17), we have

$$n\mathrm{Var}\{\widehat{V}^{or}(d)\} = n\mathrm{Var}\{\widehat{V}_{\mathcal{I}_0}(d)\} = \sigma_0^2\mathrm{E}\left(\frac{d^2(\boldsymbol{X}_0)}{\pi^*} + \frac{\{1-d(\boldsymbol{X}_0)\}^2}{1-\pi^*}\right) + \mathrm{Var}\{h(0,\boldsymbol{X}_0)\}$$
$$= \sigma_0^2\mathrm{E}\left(\frac{d(\boldsymbol{X}_0)}{\pi^*} + \frac{1-d(\boldsymbol{X}_0)}{1-\pi^*}\right) + \mathrm{Var}\{h(0,\boldsymbol{X}_0)\} \geq \sigma_0^2\min\left(\frac{1}{\pi^*}, \frac{1}{1-\pi^*}\right) + \mathrm{Var}\{h(0,\boldsymbol{X}_0)\}.$$

As for our proposed value estimator, it follows from (A7) that

$$n\mathrm{Var}\{\widehat{V}_B\} \approx \sigma_{s_n}^2 = \sigma_0^2\mathrm{E}\left(\frac{d_{s_n}^2(\boldsymbol{X}_0)}{\pi^*} + \frac{\{1-d_{s_n}(\boldsymbol{X}_0)\}^2}{1-\pi^*}\right) + \mathrm{Var}\{h(0,\boldsymbol{X}_0)\}$$
$$\leq \sigma_0^2\left(\frac{(1-d^*)^2}{\pi^*} + \frac{(1-d^*)^2}{1-\pi^*}\right) + \mathrm{Var}\{h(0,\boldsymbol{X}_0)\} + o(1).$$

When $(1-d^*)^2 \leq \pi^* \leq d^*(2-d^*)$, we have $(1-d^*)^2/\pi^* + (1-d^*)^2/(1-\pi^*) \leq \min\{1/\pi^*, 1/(1-\pi^*)\}$. This implies that our proposed estimator is more efficient than the oracle estimator.

In the regular cases where $\Pr\{\tau(\boldsymbol{X}_0) = 0\} = 0$, we can show $\widehat{V}_B = \widehat{V}^{or}(d^{opt}) + o_p(n^{-1/2})$ and $\widehat{\sigma}_B = \widehat{\sigma}^{or}(d^{opt}) + o_p(1)$ for any $d^{opt} \in \mathcal{D}^{opt}$. This means the proposed CI is asymptotically equivalent to the CI of the oracle method in the regular cases.

To summarize, we have proven that the proposed method outperforms the single sample-splitting method, the online one-step method and the "oracle" method in nonregular cases and is equivalent to these methods in regular cases. However, an unavoidable consequence of subagging is longer computational time. Specifically, the single sample-splitting method only requires to estimate the OTR once. The online on-step method requires to estimate the OTR $n - s_n$ times. The proposed method requires to estimate the OTR B many times while B shall be chosen to be much larger than n. This is a potential drawback of our method. In Section 7.2, we provide some suggestions to facilitate the computation.

4. Multiple Time Point Study

4.1. Optimal Dynamic Treatment Regime

In this section, we consider a multistage study where the treatment decisions are made at a finite number of time points $t_1, \dots, t_K$. The data for a subject can be summarized as

$$(\boldsymbol{X}_0^{(1)}, A_0^{(1)}, \boldsymbol{X}_0^{(2)}, A_0^{(2)}, \dots, \boldsymbol{X}_0^{(K)}, A_0^{(K)}, Y_0),$$

where Y_0 denotes the outcome of interest, $\boldsymbol{X}_0^{(1)}$ stands for the set of covariates obtained prior to the time point t_1, $A_0^{(1)}$ denotes the treatment received at t_1. For $k = 2, \dots, K$, $\boldsymbol{X}_0^{(k)}$ denotes some additional covariates collected between time points t_{k-1} and t_k, and $A_0^{(k)}$ denotes the treatment given at t_k. For simplicity, we assume $A_0^{(1)}, \dots, A_0^{(K)}$ are all binary treatments. For $k = 1, \dots, K$, let

$$\bar{\boldsymbol{X}}_0^{(k)} = (\boldsymbol{X}_0^{(1)}, \dots, \boldsymbol{X}_0^{(k)}) \in \bar{\mathbb{X}}^{(k)} \quad \text{and} \quad \bar{\boldsymbol{A}}_0^{(k)} = (A_0^{(1)}, \dots, A_0^{(k)}) \in \{0,1\}^k,$$

denote a patient's covariates and treatment history. For any $a_1, \dots, a_K \in \{0,1\}$, denote by $\bar{\boldsymbol{a}}_k = (a_1, \dots, a_k)$ for $k = 1, \dots, K$. The set of all potential outcomes is given by

$$\boldsymbol{W} = \left\{ \left(\boldsymbol{X}_0^{(2)*}(a_1), \boldsymbol{X}_0^{(3)*}(\bar{\boldsymbol{a}}_2), \dots, \boldsymbol{X}_0^{(K)*}(\bar{\boldsymbol{a}}_{K-1}), Y_0^*(\bar{\boldsymbol{a}}_K) \right) : \forall \bar{\boldsymbol{a}}_K \in \{0,1\}^K \right\}, \tag{29}$$

where $\boldsymbol{X}_0^{(k)*}(\bar{\boldsymbol{a}}_{k-1})$ denotes the potential time-dependent covariates of a patient that would occur between t_{k-1} and t_k assuming he/she receives treatments $(a_1, \dots, a_{k-1})$ at decision points $(t_1, \dots, t_{k-1})$ and $Y_0^*(\bar{\boldsymbol{a}}_K)$ denotes the potential outcome that would result assuming he/she receives treatments $(a_1, \dots, a_K)$.

A dynamic treatment regime $d = \{d_k\}_{k=1}^K$ is a set of decision rules that treats a patient over time. For $k = 1, \dots, K$, $d_k = d_k(\bar{\boldsymbol{a}}_{k-1}, \bar{\boldsymbol{x}}_k)$ corresponds to the kth decision rule that takes as input a patient's realized covariate and treatment history and outputs a treatment option $a_k \in \{0,1\}$. Let $\bar{d}_k = \{d_j\}_{j=1}^k$ for $k = 1, \dots, K-1$, the potential outcome associated with d is given by

$$\left(\boldsymbol{X}_0^{(2)*}(d_1), \boldsymbol{X}_0^{(3)*}(\bar{d}_2), \dots, \boldsymbol{X}_0^{(K)*}(\bar{d}_{K-1}), Y_0^*(d) \right),$$

where $\boldsymbol{X}_0^{(k)*}(\bar{d}_{k-1})$ stands for the potential covariates of a patient between t_{k-1} and t_k assuming he/she receives the treatments sequentially according to the decision rules $(d_1, \dots, d_{k-1})$ and $Y_0^*(d)$ stands for the potential outcome assuming the treatments he/she receives are determined by the treatment regime d. An optimal dynamic treatment regime d^{opt} is defined to maximize the average potential outcome, i.e,

$$d^{opt} = \arg\max_d \mathrm{E} Y_0^*(d).$$

For any $\bar{\boldsymbol{a}}_K \in \{0,1\}^K$ and $\bar{\boldsymbol{x}}_K \in \bar{\mathbb{X}}^{(K)}$, let $h_K(\bar{\boldsymbol{a}}_K, \bar{\boldsymbol{x}}_K) = \mathrm{E}(Y_0 | \bar{\boldsymbol{X}}_0^{(K)} = \bar{\boldsymbol{x}}_K, \bar{\boldsymbol{A}}_0^{(K)} = \bar{\boldsymbol{a}}_K)$ and $\tau_K(\bar{\boldsymbol{a}}_{K-1}, \bar{\boldsymbol{x}}_K) = h_K\{(\bar{\boldsymbol{a}}_{K-1}, 1), \bar{\boldsymbol{x}}_K\} - h_K\{(\bar{\boldsymbol{a}}_{K-1}, 0), \bar{\boldsymbol{x}}_K\}$. In addition, for $k = K-1, \cdots, 2$, we sequentially define

$$h_k(\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{x}}_k) = \mathrm{E} \left(\arg\max_{a_{k+1} \in \{0,1\}} h_{k+1}\{(\bar{\boldsymbol{a}}_k, a_{k+1}), \bar{\boldsymbol{X}}_0^{(k+1)}\} \middle| \bar{\boldsymbol{X}}_0^{(k)} = \bar{\boldsymbol{x}}_k, \bar{\boldsymbol{A}}_0^{(k)} = \bar{\boldsymbol{a}}_k \right),$$

and $\tau_k(\bar{\boldsymbol{a}}_{k-1}, \bar{\boldsymbol{x}}_k) = h_k\{(\bar{\boldsymbol{a}}_{k-1}, 1), \bar{\boldsymbol{x}}_k\} - h_k\{(\bar{\boldsymbol{a}}_{k-1}, 0), \bar{\boldsymbol{x}}_k\}$, for any $\bar{\boldsymbol{a}}_k \in \{0,1\}^k$ and $\bar{\boldsymbol{x}}_k \in \bar{\mathbb{X}}^{(k)}$. For $k = 1$, let

$$h_1(a_1, \boldsymbol{x}_1) = \mathrm{E} \left(\arg\max_{a_2 \in \{0,1\}} h_2\{(a_1, a_2), \bar{\boldsymbol{X}}_0^{(2)}\} \middle| \boldsymbol{X}_0^{(1)} = \boldsymbol{x}_1, A_0^{(1)} = a_1 \right)$$

and $\tau_1(\boldsymbol{x}_1) = h_1(1, \boldsymbol{x}_1) - h_1(0, \boldsymbol{x}_1)$ for any $a_1 \in \{0,1\}, \boldsymbol{x}_1 \in \bar{\mathbb{X}}_1$. Under the following two conditions,

(C1.) $\boldsymbol{X}_0^{(k)} = \sum_{\bar{\boldsymbol{a}}_{k-1} \in \{0,1\}^{k-1}} \boldsymbol{X}_0^{(k)*}(\bar{\boldsymbol{a}}_{k-1}) \mathbb{I}(\bar{\boldsymbol{A}}_0^{(k-1)} = \bar{\boldsymbol{a}}_{k-1})$ and
$Y_0 = \sum_{\bar{\boldsymbol{a}}_K \in \{0,1\}^K} Y_0^*(\bar{\boldsymbol{a}}_K) \mathbb{I}(\bar{\boldsymbol{A}}_0^{(K)} = \bar{\boldsymbol{a}}_K)$, $\forall k = 2, \dots, K$ and $\bar{\boldsymbol{a}}_K \in \{0,1\}$,

(C2.) $A_0^{(k)} \perp\!\!\!\perp \boldsymbol{W} | \bar{\boldsymbol{X}}_0^{(k)}, \bar{\boldsymbol{A}}_0^{(k-1)}, \forall k = 1, \dots, K$ where $\boldsymbol{W}$ is defined in Equation (29), we can show

$$h(\bar{\boldsymbol{a}}_K, \bar{\boldsymbol{x}}_K) = \mathrm{E}\{Y_0^*(\bar{\boldsymbol{a}}_K)|\bar{\boldsymbol{X}}_0^{(K)*}(\bar{\boldsymbol{a}}_{K-1}) = \bar{\boldsymbol{x}}_K\}, \tag{30}$$

and for $2 \le k \le K-1$,

$$h(\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{x}}_k) = \mathrm{E}[V_0^{(k+1)}\{\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{X}}_0^{(k+1)*}(\bar{\boldsymbol{a}}_k)\}|\bar{\boldsymbol{X}}_0^{(k)*}(\bar{\boldsymbol{a}}_{k-1}) = \bar{\boldsymbol{x}}_k], \tag{31}$$

and

$$h(a_1, \boldsymbol{x}_1) = \mathrm{E}[V_0^{(2)}\{a_1, \bar{\boldsymbol{X}}_0^{(2)*}(a_1)\}|\boldsymbol{X}_0^{(1)} = \boldsymbol{x}_1], \tag{32}$$

where

$$\begin{aligned}
&V_0^{(K)}(\bar{\boldsymbol{a}}_{K-1}, \bar{\boldsymbol{x}}_K) = \max_{a_K \in \{0,1\}} \mathrm{E}\{Y_0^*(\bar{\boldsymbol{a}}_K)|\bar{\boldsymbol{X}}_0^{(K)*}(\bar{\boldsymbol{a}}_{K-1}) = \bar{\boldsymbol{x}}_K\},\\
&V_0^{(k)}(\bar{\boldsymbol{a}}_{k-1}, \bar{\boldsymbol{x}}_k) = \max_{a_k \in \{0,1\}} \mathrm{E}[V_0^{(k+1)}\{\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{X}}_0^{(k+1)*}(\bar{\boldsymbol{a}}_k)\}|\bar{\boldsymbol{X}}_0^{(k)*}(\bar{\boldsymbol{a}}_{k-1}) = \bar{\boldsymbol{x}}_k],\\
&\bar{\boldsymbol{X}}_0^{(k)*}(\bar{\boldsymbol{a}}_{k-1}) = \{\boldsymbol{X}_0^{(1)}, \boldsymbol{X}_0^{(2)*}(a_1), \dots, \boldsymbol{X}_0^{(k)*}(\bar{\boldsymbol{a}}_{k-1})\}.
\end{aligned}$$

Here, Condition (C2) automatically holds in sequentially randomized studies (Murphy, 2005).

Define the set of dynamic treatment regimes $\mathcal{D}^{opt}$ such that any $d = \{d_k\}_{k=1}^K \in \mathcal{D}^{opt}$ shall satisfy

$$d_K(\bar{\boldsymbol{a}}_{k-1}, \bar{\boldsymbol{x}}_k) \in \arg\max_{a \in \{0,1\}} a\tau_k(\bar{\boldsymbol{a}}_{k-1}, \bar{\boldsymbol{x}}_k), k = 2, \dots, K, \tag{33}$$
$$\text{and} \quad d_1(\boldsymbol{x}_1) \in \arg\max_{a \in \{0,1\}} a\tau_1(\boldsymbol{x}_1),$$

for any $\bar{\boldsymbol{x}}_K \in \bar{\mathbb{X}}^{(K)}, \dots, \bar{\boldsymbol{x}}_2 \in \bar{\mathbb{X}}^{(2)}, \boldsymbol{x}_1 \in \bar{\mathbb{X}}^{(1)}$ and $\bar{\boldsymbol{a}}_{K-1} \in \{0,1\}^{K-1}, \dots, \bar{\boldsymbol{a}}_2 \in \{0,1\}^2, a_1 \in \{0,1\}$. By (30)-(32) and backward induction, we can show that

$$\mathcal{D}^{opt} \subseteq \arg\max_d \mathrm{E} Y_0^*(d).$$

Notice that the argmax in Equation 33 is not unique when $\tau_k(\bar{\boldsymbol{a}}_{k-1}, \bar{\boldsymbol{x}}_k) = 0$ or $\tau_1(\boldsymbol{x}_1) = 0$. Therefore, the optimal dynamic treatment regime may not be unique.

4.2. Confidence Interval for the Optimal Value Function

In this section, we focus on constructing CIs for the optimal value function $V_0 = \max_d \mathrm{E} Y_0^*(d)$, based on the observed dataset:

$$\left\{O_i = \left(\boldsymbol{X}_i^{(1)}, A_i^{(1)}, \boldsymbol{X}_i^{(2)}, A_i^{(2)}, \dots, \boldsymbol{X}_i^{(K)}, A_i^{(K)}, Y_i\right) : i = 1, \dots, n\right\}.$$

For $k = 1, \dots, K, i = 0, 1, \dots, n$, let

$$\bar{\boldsymbol{X}}_i^{(k)} = (\boldsymbol{X}_i^{(1)}, \dots, \boldsymbol{X}_i^{(k)}) \quad \text{and} \quad \bar{\boldsymbol{A}}_i^{(k)} = (A_i^{(1)}, \dots, A_i^{(k)}).$$

Define the propensity score function $\pi_k(\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{x}}_k) = \Pr(A_0^{(k)} = a_k | \bar{\boldsymbol{X}}_0^{(k)} = \bar{\boldsymbol{x}}_k, \bar{\boldsymbol{A}}_0^{(k-1)} = \bar{\boldsymbol{a}}_{k-1})$ for $k = 2, \dots, K$ and $\pi_1(a_1, \boldsymbol{x}_1) = \Pr(A_0^{(1)} = a_k | \boldsymbol{X}_0^{(1)} = \boldsymbol{x}_1)$. For any dynamic treatment regime $d = \{d_k\}_{k=1}^K$, let $\widehat{V}_i^{(K+1)} = Y_i$ and recursively define

$$\begin{aligned}\widehat{V}_i^{(k)}(d; \pi^*, h^*) &= \frac{\mathrm{g}\{A_i^{(k)}, d_k(\bar{\boldsymbol{A}}_i^{(k-1)}, \bar{\boldsymbol{X}}_i^{(k)})\}}{\pi_k^*(\bar{\boldsymbol{A}}_i^{(k)}, \bar{\boldsymbol{X}}_i^{(k)})}\{\widehat{V}_i^{(k+1)}(d; \pi^*, h^*) - h_k^*(\bar{\boldsymbol{A}}_i^{(k)}, \bar{\boldsymbol{X}}_i^{(k)})\} \\ &+ h_k^*[\{\bar{\boldsymbol{A}}_i^{(k-1)}, d_k(\bar{\boldsymbol{A}}_i^{(k-1)}, \bar{\boldsymbol{X}}_i^{(k)})\}, \bar{\boldsymbol{X}}_i^{(k)}],\end{aligned}$$

for $k = K-1, \dots, 2, 1$, $i = 0, 1, \dots, n$, where $\bar{\boldsymbol{A}}_i^{(0)} = \emptyset$, $\pi^* \equiv \{\pi_k^*\}_{k=1}^K$ and $h^* \equiv \{h_k^*\}_{k=1}^K$ denote the estimated propensity score and conditional mean functions, and the function $\mathrm{g}(a, z) = az + (1-a)(1-z)$ for $a, z \in \mathbb{R}$. For any $\mathcal{I} \subseteq \mathcal{I}_0$, consider the following augmented inverse propensity-score weighted estimator for $\mathrm{E}Y_0^*(d)$,

$$\frac{1}{|\mathcal{I}|}\sum_{i \in \mathcal{I}} \widehat{V}_i^{(1)}(d; \pi^*, h^*).$$

Notice that it is unbiased when either $\pi^* = \pi$ or $h^* = h$.

For any $\mathcal{I} \subseteq \mathcal{I}_0$, let $\{\widehat{h}_{\mathcal{I},k}\}_{k=1}^K$, $\{\widehat{\pi}_{\mathcal{I},k}\}_{k=1}^K$ denote some consistent estimators for $\{h_k\}_{k=1}^K$ and $\{\pi_k\}_{k=1}^K$, computed based on the sub-dataset $\{O_i\}_{i \in \mathcal{I}}$. Consider the estimated treatment regime $\widehat{d}_{\mathcal{I},k}(\cdot, \cdot)$, $k = 2, \dots, K$ and $\widehat{d}_{\mathcal{I},1}(\cdot)$. Define the set

$$\mathcal{S}_{N_0, s_n} = \left\{\mathcal{I} \subseteq \mathcal{I}_0 : |\mathcal{I}| = s_n, \min_{a_1, \dots, a_K \in \{0,1\}} \sum_{i \in \mathcal{I}} \mathbb{I}(A_i^{(1)} = a_1, \dots, A_i^{(K)} = a_K) \geq N_0\right\},$$

for some integers $s_n > N_0 > 0$. We summarize our procedure in the following algorithm.

Step 1 Input observations $\{O_i\}_{i \in \mathcal{I}_0}$, $0 < \alpha < 1$ and integers s_n, N_0 and B.

Step 2 For $b = 1, \dots, B$,

(i) Draw a subset $\mathcal{I}_b$ uniformly from $\mathcal{S}_{N_0, s_n}$.

(ii) Randomly partition $\mathcal{I}_b^c$ into 2 disjoint subsets $\mathcal{I}_b^{c(1)}$ and $\mathcal{I}_b^{c(2)}$ of equal sizes t_n.

(iii) For $j = 1, 2$, let $\mathcal{I}_b^{(j)} = \mathcal{I}_b \cup \mathcal{I}_b^{c(j)}$. Obtain the estimators $\widehat{d}_{\mathcal{I}_b} = \{\widehat{d}_{\mathcal{I}_b,k}\}_{k=1}^K$, $\widehat{\pi}_{\mathcal{I}_b^{(1)}} = \{\widehat{\pi}_{\mathcal{I}_b^{(1)},k}\}_{k=1}^K$, $\widehat{\pi}_{\mathcal{I}_b^{(2)}} = \{\widehat{\pi}_{\mathcal{I}_b^{(2)},k}\}_{k=1}^K$, $\widehat{h}_{\mathcal{I}_b^{(1)}} = \{\widehat{h}_{\mathcal{I}_b^{(1)},k}\}_{k=1}^K$ and $\widehat{h}_{\mathcal{I}_b^{(2)}} = \{\widehat{\pi}_{\mathcal{I}_b^{(2)},k}\}_{k=1}^K$.

Step 3 Compute

$$\widehat{V}_B = \frac{1}{2Bt_n}\sum_{b=1}^B \left(\sum_{i \in \mathcal{I}_b^{c(2)}} \widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b}; \widehat{\pi}_{\mathcal{I}_b^{(1)}}, \widehat{h}_{\mathcal{I}_b^{(1)}}) + \sum_{i \in \mathcal{I}_b^{c(1)}} \widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b}; \widehat{\pi}_{\mathcal{I}_b^{(2)}}, \widehat{h}_{\mathcal{I}_b^{(2)}})\right),$$

and

$$\widehat{\sigma}_B^2 = \widehat{s.e}^2\left(\left\{\frac{1}{n^{(i)}}\sum_{j=1}^2\sum_{b=1}^B \widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b}; \widehat{\pi}_{\mathcal{I}_b^{(j)}}, \widehat{h}_{\mathcal{I}_b^{(j)}})\mathbb{I}(i \notin \mathcal{I}_b^{(j)})\right\}_{i=1}^n\right),$$

where $n^{(i)} = \sum_{b=1}^B \mathbb{I}(i \notin \mathcal{I}_b)$.

	(A)	(B)	(C)	(D)	(E)	(F)
$\Phi(x_1, x_2)$	0.3	0.3	x_2^2	x_2^2	x_2^2	x_2^2
$\tau(x_1, x_2)$	$0.4\mathbb{I}(x_1 = 0)$	0.4	$x_1 x_2^2$	$x_2^2 - 4/3$	$2x_1 \cos(\pi x_2/4)$	$2\cos(\pi x_2/4) - 4/\pi$
V_0	0.5	0.7	2	1.85	1.97	1.60

Table 2: Simulation setting

Step 4 Output

$$\left[\widehat{V}_B - \frac{z_{\alpha/2}\widehat{\sigma}_B}{\sqrt{n}}, \widehat{V}_B + \frac{z_{\alpha/2}\widehat{\sigma}_B}{\sqrt{n}}\right]. \tag{34}$$

In Appendix A, we prove the CI in Equation (34) achieves nominal coverage.

5. Simulations

5.1. Point Treatment Study

We consider simulation studies based on the following model:

$$Y_0 = \Phi(X_{0,1}, X_{0,2}) + A_0\tau(X_{0,1}, X_{0,2}) + e_0,$$

where the covariate $X_{0,1}$ and the treatment A_0 are generated from Ber(0.5) and Ber($0.5 + 0.1X_{0,1}$), respectively, where Ber(p_0) stands for the Bernoulli distribution with probability of success p_0. The random error term e_0 satisfies $\mathrm{E}(e_0|A_0, X_{0,1}, X_{0,2}) = 0$. We consider six scenarios. In Scenario (A) and (B), $X_{0,2}$ is generated from Ber(0.5), and

$$e_0 \sim \mathrm{Ber}\{\Phi(X_{0,1}, X_{0,2}) + A_0\tau(X_{0,1}, X_{0,2})\} - \Phi(X_{0,1}, X_{0,2}) - A_0\tau(X_{0,1}, X_{0,2}).$$

In Scenario (C)-(F), $X_{0,2}$ follows a uniform distribution on the interval $[-2, 2]$, and $e_0 \sim N(0, 0.25)$ is independent of A_0, $X_{0,1}$ and $X_{0,2}$. In addition, $X_{0,1}$ and $X_{0,2}$ are independently generated in all scenarios. Table 2 summarizes the information of the baseline function, the contrast function and the optimal value V_0 under different scenarios. In all scenarios, V_0 can be explicitly calculated. The OTR is not uniquely defined in Scenario (A), (C) and (E), since the contrast functions in these scenarios satisfy

$$\Pr\{\tau(X_{0,1}, X_{0,2}) = 0\} = \Pr(X_{0,1} = 0) = 0.5.$$

On the contrary, we have $\Pr\{\tau(X_{0,1}, X_{0,2}) = 0\} = 0$ in the remaining three scenarios. For each scenario, we further consider two different sample sizes, $n = 500$ and $n = 1000$. This yields a total of 12 settings.

Comparison is made among the following four methods:

(i) The proposed CI in Equation (8).
(ii) The CI constructed by the online one-step method in Equation (22).
(iii) The CI constructed by the oracle method in Equation 27 with $d^{opt} = d^{opt,0}$ (see Equation 2). (Notice that $d^{opt,0}$ is unknown in practice, we implement this method for comparison purposes only.)

(A)	Proposed $s_n = 3n/\log(n)$		Online		Oracle		SSS	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
500	93.6 (0.8)	11.2 (0.02)	94.1 (0.7)	12.8 (0.02)	94.0 (0.8)	13.1 (0.02)	93.7 (0.8)	17.1 (0.03)
1000	93.7 (0.8)	7.8 (0.01)	93.9 (0.8)	8.8 (0.01)	94.1 (0.7)	9.0 (0.01)	94.2 (0.7)	11.4 (0.02)
(B)	Proposed $s_n = 3n/\log(n)$		Online		Oracle		SSS	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
500	95.3 (0.7)	11.1 (0.01)	93.9 (0.8)	11.5 (0.02)	95.4 (0.7)	11.2 (0.01)	94.7 (0.7)	15.4 (0.03)
1000	95.3 (0.7)	7.8 (0.01)	95.5 (0.7)	7.9 (0.01)	95.3 (0.7)	7.8 (0.01)	95.1 (0.7)	10.3 (0.01)
(C)	Proposed $s_n = 3n/\log(n)$		Online		Oracle		SSS	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
500	94.7 (0.7)	36.8 (0.04)	92.7 (0.8)	41.1 (0.06)	94.3 (0.7)	38.0 (0.08)	93.9 (0.8)	52.9 (0.09)
1000	95.0 (0.7)	25.9 (0.02)	93.4 (0.8)	27.4 (0.03)	94.5 (0.7)	26.3 (0.02)	95.3 (0.7)	34.9 (0.04)
(D)	Proposed $s_n = 3n/\log(n)$		Online		Oracle		SSS	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
500	95.3 (0.7)	36.6 (0.04)	93.2 (0.8)	40.6 (0.05)	93.7 (0.8)	38.0 (0.20)	94.1 (0.7)	51.4 (0.09)
1000	94.3 (0.7)	25.7 (0.02)	93.1 (0.8)	27.1 (0.02)	94.2 (0.7)	25.9 (0.02)	95.1 (0.7)	34.3 (0.03)
(E)	Proposed $s_n = 3n/\log(n)$		Online		Oracle		SSS	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
500	94.4 (0.7)	22.7 (0.02)	88.2 (1.0)	25.8 (0.03)	94.0 (0.8)	24.6 (0.09)	95.2 (0.7)	33.3 (0.12)
1000	95.4 (0.7)	15.9 (0.01)	91.9 (0.9)	17.2 (0.01)	95.3 (0.7)	16.7 (0.02)	95.0 (0.7)	21.9 (0.02)
(F)	Proposed $s_n = 3n/\log(n)$		Online		Oracle		SSS	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
500	92.9 (0.8)	21.4 (0.04)	87.5 (1.0)	23.3 (0.03)	93.8 (0.8)	24.3 (0.36)	94.5 (0.7)	31.1 (0.14)
1000	94.4 (0.7)	14.9 (0.01)	90.8 (0.9)	15.6 (0.01)	94.1 (0.7)	15.3 (0.03)	93.4 (0.8)	20.0 (0.03)

Table 3: ECP and AL of the CIs with standard errors in parenthesis

(iv) The CI constructed by the single sample-splitting method in Equation (19) (Denote by SSS).

All these methods require estimation of the propensity score and conditional mean functions. For scenario (A) and (B), we use the nonparametric maximum likelihood estimator to estimate these functions. For scenario (C)-(F), we estimate these functions using cubic B-splines. More specifically, for $a = 0, 1$, define

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{\pi,a} = \arg\min_{\boldsymbol{\xi}} \sum_{i\in\mathcal{I}} \left(A_i - \sum_{j=1}^{K+4} N_j(X_{i,2})\xi_j \right)^2 \mathbb{I}(X_{i,1} = a),$$

and

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_1,a} = \arg\min_{\boldsymbol{\xi}} \sum_{i\in\mathcal{I}} \left(Y_i - \sum_{j=1}^{K+4} N_j(X_{i,2})\xi_j \right)^2 \mathbb{I}(A_i = 1, X_{i,1} = a),$$

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_0,a} = \arg\min_{\boldsymbol{\xi}} \sum_{i\in\mathcal{I}} \left(Y_i - \sum_{j=1}^{K+4} N_j(X_{i,2})\xi_j \right)^2 \mathbb{I}(A_i = 0, X_{i,1} = a),$$

where $N_1(\cdot), \ldots, N_{K+4}(\cdot)$ stand for the cubic B-spline basis, and K denotes the number of interior knots. Given K, the interior knots are placed at equally spaced sample quantiles of

$\{X_{i,2}\}_{i\in\mathcal{I}_0}$. The hyperparameter K is selected via 5-fold cross-validation. After computing $\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{\pi,a}$, $\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_1,a}$ and $\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_0,a}$, we set

$$\begin{aligned}
\widehat{\pi}_{\mathcal{I}}(1,\boldsymbol{x}_1) &= \min\Big(\sum_{\substack{a=\{0,1\}\\1\le j\le K+4}} \mathbb{I}(x_{1,1}=a)N_j(x_{1,2})\widehat{\xi}_{\mathcal{I},j}^{\pi,a}, 0.05\Big), && (35)\\
\widehat{\pi}_{\mathcal{I}}(0,\boldsymbol{x}_1) &= \min\{1-\widehat{\pi}_{\mathcal{I}}(1,\boldsymbol{x}_1), 0.05\}, && (36)\\
\widehat{h}_{\mathcal{I}}(1,\boldsymbol{x}_1) &= \sum_{a=0,1}\mathbb{I}(x_{1,1}=a)\sum_{j=1}^{K+4}N_j(x_{1,2})\widehat{\xi}_{\mathcal{I},j}^{h_1,a},\\
\widehat{h}_{\mathcal{I}}(0,\boldsymbol{x}_1) &= \sum_{a=0,1}\mathbb{I}(x_{1,1}=a)\sum_{j=1}^{K+4}N_j(x_{1,2})\widehat{\xi}_{\mathcal{I},j}^{h_0,a},
\end{aligned}$$

where $\boldsymbol{x}_1=(x_{1,1},x_{1,2})$. Truncation is used in Equations (35) and (36) to avoid extreme weights, resulting in a more "stabilized" value estimator. The estimated contrast function is defined as

$$\widehat{\tau}_{\mathcal{I}}(\boldsymbol{x}_1)=\widehat{h}_{\mathcal{I}}(1,\boldsymbol{x}_1)-\widehat{h}_{\mathcal{I}}(0,\boldsymbol{x}_1).$$

To calculate the CI in Equation (8), we set $s_n=\lfloor K_0 n/\log n\rfloor$ with $K_0=3$. To implement the online one-step and single sample-splitting method, we need to specify ℓ_n. In general, the lengths of CIs in Equations (22) and (19) increase as ℓ_n increases. Nonetheless, ℓ_n should be large enough to guarantee that biases of the resulting value estimates are negligible. For the single sample-splitting method, we set $\ell_n=\lfloor 3n/\log n\rfloor$. For the online one-step method, in Scenarios (A) and (B), we set $\ell_n=50$. In Scenarios (C)-(F), we find that when $\ell_n=50$, the resulting CIs have very poor coverage probabilities. Therefore, we set $\ell_n=100$ in these scenarios. The variance estimator $\widetilde{\sigma}^2_{\mathcal{I}_{(j)}}$ is computed by

$$\widetilde{\sigma}^2_{\mathcal{I}_{(j)}}=\widehat{s.e}^2\left(\left\{\psi_i(\widehat{d}_{\mathcal{I}_{(j)}};\widehat{\pi}_{\mathcal{I}_{(j)}},\widehat{h}_{\mathcal{I}_{(j)}})\right\}_{i=1}^{j}\right).$$

We implement the simulation program in R. Some subroutines are written in C with the GNU Scientific Library (Galassi et al., 2015) to facilitate the computation.

Reported in Table 3 are the empirical coverage probability (ECP) and average length (AL) of the CIs in (i)-(iv). Results are aggregated over 1000 replications. It can be seen that all four CIs achieve nominal coverage in Scenario (A)-(D). However, the CIs based on the online one-step method, the oracle method and the single sample-splitting method are wider than the proposed CIs in all cases. Take Scenario (A) as an example. ALs of our proposed method are at least 13% smaller than other competing methods. In Scenarios (E) and (F), ECPs of the online one-step method are smaller than 90% when $n=500$. In contrast, ECPs of the proposed CIs are close to the nominal level in all cases. In addition, the proposed CIs achieve smaller ALs in these scenarios.

Notice that in Scenarios (B), (D) and (F), the contrast function is almost surely nonzero. In theory, when $\ell_n=o(n)$, the lengths of all four CIs should be asymptotically the same. However, it can been seen from Table 3 that in finite samples, ALs of the CIs based on our proposed method are always smaller than other competing methods.

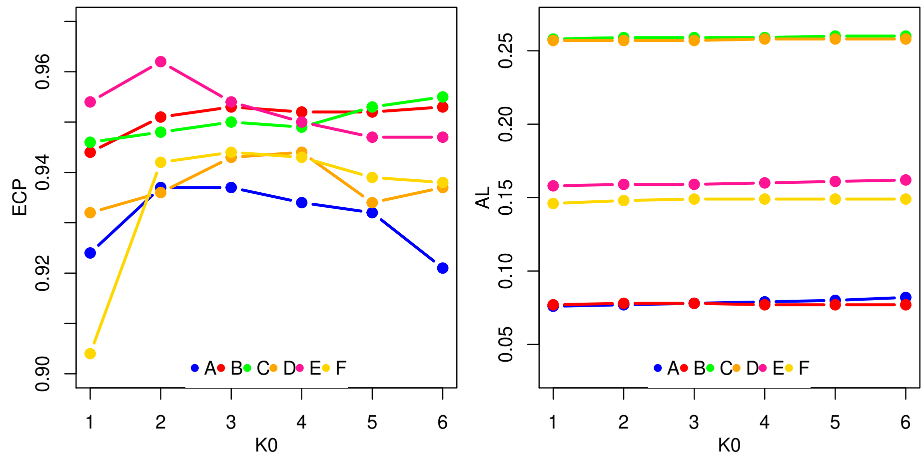

Figure 1: ECP and AL of the proposed confidence interval with difference choices of K_0 under Settings (A)-(F). $n = 1000$ in all cases.

	(G)	(H)	(I)
$\Phi(x_{1,1}, x_{1,2}, a_1, x_2)$	$x_{1,1}^2 - a_1(0.25 + x_{1,1}^2)$	x_2^2	0
$\tau(x_{1,1}, x_{1,2}, a_1, x_2)$	$a_1 x_2^2$	0	x_2^2
V_0	1.33	1.58	1.58

Table 4: Simulation setting

We next investigate the robustness of our method to the choice of K_0. We fix $n = 1000$. Figure 1 depicts the empirical coverage probabilities (ECP) and average lengths (AL) of the proposed confidence intervals with different choices of K_0 under Settings (A)-(F). It can be seen from Figure 1 that the ECPs of the proposed CIs are close to the nominal level for any K_0 under most settings. In addition, for each setting, ALs of our CIs are roughly the same across different K_0. This shows the robustness of our procedure to the choice of K_0.

5.2. Multiple Time Point Study

Consider the following model:

$$Y_0 = \Phi(X_{0,1}^{(1)}, X_{0,2}^{(1)}, A_0^{(1)}, X_0^{(2)}) + A_0^{(2)} \tau(X_{0,1}^{(1)}, X_{0,2}^{(1)}, A_0^{(1)}, X_0^{(2)}) + e_0^{(2)}, \quad X_0^{(2)} = A_0^{(1)} X_{0,1}^{(1)} + e_0^{(1)},$$

where $X_{0,1}^{(1)}$ and $X_{0,2}^{(1)}$ are the baseline covariates, $A_0^{(1)}$ and $A_0^{(2)}$ denote the first and second treatment a patient receives at t_1 and t_2, $X_0^{(2)}$ stands for the intermediate covariate collected between t_1 and t_2. Variables $A_0^{(1)}$, $A_0^{(2)}$, $X_{0,1}^{(1)}$, $X_{0,2}^{(1)}$, $e_0^{(1)}$ and $e_0^{(2)}$ are all independent. In addition, we assume $A_0^{(1)}, A_0^{(2)} \sim \text{Ber}(0.5)$, $e_0^{(1)}, e_0^{(2)} \sim N(0, 0.25)$ and $X_{0,1}^{(1)}, X_{0,2}^{(1)} \sim \text{Unif}[-2, 2]$ where $\text{Unif}[a, b]$ denotes the uniform distribution on the interval $[a, b]$.

We consider three scenarios. The functional forms of Φ and τ and the optimal value function V_0 under these scenarios are reported in Table 4.

In Scenario (G), we have

$$h_1(a_1, \boldsymbol{x}_1) = \mathrm{E}\{A_0^{(1)} X_0^{(2)} + \Phi(X_{0,1}^{(1)}, X_{0,2}^{(1)}, A_0^{(1)}, X_0^{(2)}) | X_{0,1}^{(1)} = x_{1,1}, X_{0,2}^{(1)} = x_{1,2}, A_0^{(1)} = a_1\} = a_1(0.25 + x_{1,1}^2) + x_{1,1}^2 - a_1(0.25 + x_{1,1}^2) = x_{1,1}^2,$$

where $\boldsymbol{x}_1 = (x_{1,1}, x_{1,2})$. Therefore, the first stage contrast function $\tau_1(\cdot)$ equals zero.

In Scenario (H), the second stage contrast function $\tau_2(\cdot, \cdot)$ equals zero. Hence, the ODTR is not unique in these two scenarios.

In the last scenario, we have

$$h_2(\bar{\boldsymbol{a}}_2, \bar{\boldsymbol{x}}_2) = x_2^2, h_1(a_1, \boldsymbol{x}_1) = \mathrm{E}\{(X_{0,2}^{(2)})^2 | \bar{\boldsymbol{X}}_0^{(1)} = \boldsymbol{x}_1, A_0^{(1)} = a_1\} = a_1 x_{1,1}^2 + 0.25,$$

where $\bar{\boldsymbol{x}}_2 = (x_{1,1}, x_{1,2}, x_2)$. In this scenario, the ODTR is uniquely defined and we have $d_1^{opt}(\boldsymbol{x}_1) = 1$, $d_2^{opt}(a_1, \bar{\boldsymbol{x}}_2) = 1$.

We compare our proposed CI (see (34)) with the CI based on the online one-step method, defined as $[\widehat{V}^{on} - z_{\alpha/2}\widehat{\sigma}^{on}/\sqrt{n - \ell_n}, \widehat{V}^{on} + z_{\alpha/2}\widehat{\sigma}^{on}/\sqrt{n - \ell_n}]$ where

$$\begin{aligned} \widehat{V}^{on} &= \left(\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_{\mathcal{I}_{(j)}}^{-1} \right)^{-1} \left(\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_{\mathcal{I}_{(j)}}^{-1} \psi_{j+1}(\widehat{d}_{\mathcal{I}_{(j)}}; \widehat{\pi}_{\mathcal{I}_{(j)}}, \widehat{h}_{\mathcal{I}_{(j)}}) \right), \\ \widetilde{\sigma}_{\mathcal{I}_{(j)}}^2 &= \widehat{s.e}^2 \left(\left\{ \psi_i(\widehat{d}_{\mathcal{I}_{(j)}}; \widehat{\pi}_{\mathcal{I}_{(j)}}, \widehat{h}_{\mathcal{I}_{(j)}}) \right\}_{i=1}^{j} \right), \\ \widehat{\sigma}^{on} &= \left(\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_{\mathcal{I}_{(j)}}^{-1} / (n - \ell_n) \right)^{-1}, \end{aligned}$$

for some divergent sequence ℓ_n. Notice that both methods require to calculate $\widehat{h}_{\mathcal{I}} = \{\widehat{h}_{\mathcal{I},k}\}_{k=1}^2$, $\widehat{\pi}_{\mathcal{I}} = \{\widehat{\pi}_{\mathcal{I},k}\}_{k=1}^2$, $\widehat{d}_{\mathcal{I}} = \{\widehat{d}_{\mathcal{I},k}\}_{k=1}^2$. These estimators are computed based on cubic B-spline methods. To save space, we present the detailed estimating procedure in Appendix B.

We consider two sample sizes, $n = 600$ and $n = 1200$. In Table 5, we report the ECP and AL of the proposed CI with $s_n = 3n/\log(n)$, and the CI based on online one-step method with $\ell_n = 200$ and $\ell_n = 400$. In Appendix E, we report the ECP and AL of the proposed CI with $s_n = 3.5n/\log(n)$ and $s_n = 4n/\log(n)$. It can be seen that the proposed CIs are not sensitive to the choice of K_0. ECPs of our CIs are close to the nominal level in almost all cases. In contrast, ECPs of the CIs based on the online one-step method are well below the nominal level in Scenarios (G) and (H). Moreover, CIs based on our proposed method are much shorter than those based on the online one-step method.

6. Real Data Analysis

In this section, we apply the proposed method to a data from AIDS Clinical Trials Group Protocol 175 (ACTG175). We focus on a subset of the data which consists of 1046 patients that were treated with either ZDV + zalcitabine (zal) ($A = 0$) or ZDV + didanosine (ddI)

Setting (G)	Proposed $s_n = 3n/\log(n)$		Online ($\ell_n = 200$)		Online ($\ell_n = 400$)	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
600	93.6 (0.8)	27.7 (0.15)	82.1 (1.2)	38.3 (0.12)	90.8 (0.9)	54.0 (0.18)
1200	93.3 (0.8)	18.3 (0.05)	87.9 (1.0)	24.2 (0.06)	90.7 (0.9)	27.0 (0.07)
Setting (H)	Proposed $s_n = 3n/\log(n)$		Online ($\ell_n = 200$)		Online ($\ell_n = 400$)	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
600	92.7 (0.8)	37.3 (0.09)	84.1 (1.2)	45.3 (0.10)	91.4 (0.9)	64.0 (0.13)
1200	93.8 (0.8)	25.5 (0.04)	89.5 (1.0)	28.6 (0.05)	92.8 (0.8)	32.0 (0.05)
Setting (I)	Proposed $s_n = 3n/\log(n)$		Online ($\ell_n = 200$)		Online ($\ell_n = 400$)	
n	ECP(%)	AL*100	ECP(%)	AL*100	ECP(%)	AL*100
600	93.5 (0.8)	39.4 (0.14)	84.5 (1.2)	44.5 (0.10)	90.1 (0.9)	63.1 (0.13)
1200	94.2 (0.7)	26.4 (0.04)	90.5 (0.9)	28.4 (0.05)	92.5 (0.8)	31.8 (0.05)

Table 5: ECP and AL of the CIs with standard errors in parenthesis

Method	Estimated value function	95% CI	Length of CI
Proposed ($K_0 = 3$)	399.6	$[387.9, 411.3]$	23.4
Proposed ($K_0 = 3.5$)	399.6	$[387.8, 411.4]$	23.6
Proposed ($K_0 = 4$)	399.5	$[387.6, 411.4]$	23.8
Online ($\ell_n = 50$)	399.2	$[385.6, 412.7]$	27.1
Online ($\ell_n = 100$)	398.3	$[384.4, 412.2]$	27.7
Online ($\ell_n = 200$)	403.9	$[389.5, 418.4]$	28.9

Table 6: Estimated value functions and confidence intervals

($A = 1$). The outcome of interests were CD4 count (cells/mm^3) at 20 ± 5 weeks after receiving the treatment. Fan et al. (2017) found that patient's age is the only variable that has significant interaction with the treatment. Therefore, in the following, we use age to construct the OTR. Since ACTG175 is a randomized trial, the no unmeasured confounders assumption (A2) automatically holds.

In Table 6, we report the estimated optimal value function and its 95% CI based on our proposed method and the online one-step method with $\ell_n = 50, 100$ and 200. To construct these CIs, we set $\widehat{\pi}_{\mathcal{I}} = 0.5$ for any $\mathcal{I} \subseteq \mathcal{I}_0$. The conditional mean functions are estimated using cubic B-splines. The detailed estimating procedure is very similar to that in Section 5.1 and is hence omitted for brevity. In addition, we set $s_n = \lfloor K_0 n/\log(n) \rfloor$ with $K_0 \in \{3, 3.5, 4\}$.

It can be seen from Table 6 that all methods yield similar estimated optimal value functions. These estimated values are larger than those based on linear decision rules (see Section 4 in Fan et al., 2017). Besides, we notice that our proposed CI is at least 14% shorter compared to those based on the online one-step method. Such phenomenon is consistent with our theoretical findings and simulation results.

7. Discussion

7.1. Inference via Subagging and Refitted Cross-Validation

In this paper, we propose to construct the confidence interval for the optimal value function based on subsample aggregating and refitted cross-validation. Such an inference method can be applied to some other non-regular problems as well. Variation of this approach has been used by Wang et al. (2020) for inference of the treatment effect in high-dimensional models.

7.2. Inference with Moderate or High-Dimensional Covariates

In this paper, we investigate the empirical performance of our method under settings with only a few covariates. Modern machine learning algorithms are well-suited to estimating the outcome regression functions in moderate or high-dimensions. Using these plug-in decision rules, the proposed method can handle high-dimensional covariates in theory. However, the major challenge lies in efficiently applying these machine learning algorithms B many times. Next, we provide some suggestions to facilitate the computation.

First, we note that most machine learning procedures use cross-validation for hyperparameter selection. Cross-validation could be very computationally expensive in practice. To compute the proposed confidence interval, there is no need to implement cross-validations on all B subsamples. It suffices to apply cross-validation for tuning parameter selection to the data set in the first subsample and then use the same tuning parameters in the rest $B-1$ subsamples. This could greatly simplify the computation. Second, we note that Step 2 of our algorithm can be naturally implemented in parallel. This could further reduce the computational cost.

Acknowledgments

The authors wish to thank the Action Editor and anonymous referees for their constructive comments, which lead to significant improvement of this work. Dr. Lu is partially supported by National Cancer Institute grant P01 CA142538, Dr. Song is partially supported by the National Science Foundation grant DMS 1555244 and National Cancer Institute grant P01 CA142538.

Appendix A. Validity of the Proposed Confidence Interval in (4.6)

To provide more interpretable conditions, we focus on the class of plug-in estimators $\widehat{d}_{\mathcal{I},k}(\cdot,\cdot) = \mathbb{I}\{\widehat{\tau}_{\mathcal{I},k}(\cdot,\cdot) > 0\}$, $k = 2,\dots,K$ and $\widehat{d}_{\mathcal{I},1}(\cdot) = \mathbb{I}\{\widehat{\tau}_{\mathcal{I},1}(\cdot) > 0\}$, where $\widehat{\tau}_{\mathcal{I},k}$ denotes the estimated contrast function at the kth stage. We introduce the following conditions.

(C3) Assume there exists some positive constant c_0 such that

$$\min_{k=2,\dots,K} \inf_{\substack{\bar{\boldsymbol{x}}_k \in \bar{\mathbb{X}}^{(k)} \\ \bar{\boldsymbol{a}}_k \in \{0,1\}^k}} \pi_k(\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{x}}_k) \geq c_0 \quad \text{and} \quad \inf_{\substack{\boldsymbol{x}_1 \in \bar{\mathbb{X}}^{(1)} \\ a_1 = 0,1}} \pi_1(a_1, \boldsymbol{x}_1) \geq c_0.$$

(C4) Assume there exist some positive constants $\bar{c}$, α and δ_0 such that

$$\max_{k=2,\dots,K} \Pr(0 < |\tau_k(\bar{\boldsymbol{A}}_0^{(k-1)}, \bar{\boldsymbol{X}}_0^{(k)})| < t) \le \bar{c}t^\alpha \text{ and } \Pr(0 < |\tau_1(\boldsymbol{X}_0^{(1)})| < t) \le \bar{c}t^\alpha,$$

for any $0 < t \le \delta_0$.
(C5) Assume there exists some constant $\kappa_0 > (\alpha+2)/(2\alpha+2)$ such that

$$\max_{k=2,\dots,K} \mathrm{E}|\widehat{\tau}_{\mathcal{I},k}(\bar{\boldsymbol{A}}_0^{(k-1)}, \bar{\boldsymbol{X}}_0^{(k)}) - \tau_k(\bar{\boldsymbol{A}}_0^{(k-1)}, \bar{\boldsymbol{X}}_0^{(k)})|^2 = O(|\mathcal{I}|^{-\kappa_0}),$$
$$\mathrm{E}|\widehat{\tau}_{\mathcal{I},1}(\boldsymbol{X}_0^{(1)}) - \tau_1(\boldsymbol{X}_0^{(1)})|^2 = O(|\mathcal{I}|^{-\kappa_0}),$$

for any $\mathcal{I} \subseteq \mathcal{I}_0$.

Notice that Conditions (C3)-(C5) are very similar to (A3), (A5) and (A6) in single-stage studies. In (C4), we assume the contrast function at each stage satisfies the margin assumption. In (C5), the estimated contrast function at each stage is required to satisfy certain convergence rates.

For any $\mathcal{I} \subseteq \mathcal{I}_0$ with $|\mathcal{I}| = s_n$, define

$$\sigma_{s_n}^2 = \mathrm{Var}\left\{\mathrm{E}\left(\frac{\mathrm{g}\{A_0^{(1)}, \widehat{d}_{\mathcal{I},1}(\bar{\boldsymbol{X}}_0^{(1)})\}}{\pi_1(A_0^{(1)}, \bar{\boldsymbol{X}}_0^{(1)})}\{\widehat{V}_0^{(2)}(\widehat{d}_{\mathcal{I}};\pi,h) - h_1(A_0^{(1)}, \bar{\boldsymbol{X}}_0^{(1)})\} + h_1\{\widehat{d}_{\mathcal{I},1}(\bar{\boldsymbol{X}}_0^{(1)}), \bar{\boldsymbol{X}}_0^{(1)}\} \mid O_0\right)\right\},$$

where $O_0 = (\bar{\boldsymbol{A}}_0^{(K)}, \bar{\boldsymbol{X}}_0^{(K)}, Y_0)$. We have the following results.

Theorem 7 *Assume (C1)-(C6) hold, s_n satisfies $s_n \gg n^{(2\alpha+2)/\{(\alpha+2)\kappa_0\}}$, and $\max_{\bar{\boldsymbol{a}}_K\in\{0,1\}^K} E\{Y_0^*(\bar{\boldsymbol{a}}_K)\}^2 < +\infty$. Assume $B \gg n$, $\liminf_n \sigma_{s_n} > 0$,*

$$Pr\Big(\min_{\substack{k=2,\dots,K\\ \mathcal{I}\subseteq\mathcal{I}_0}} \inf_{\substack{\bar{\boldsymbol{x}}_k\in\bar{\mathbb{X}}^{(k)}\\ \bar{\boldsymbol{a}}_k\in\{0,1\}^k}} \widehat{\pi}_{\mathcal{I},k}(\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{x}}_k) \ge c^*\Big) = 1, \tag{37}$$

$$Pr\Big(\inf_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\ \boldsymbol{x}_1\in\bar{\mathbb{X}}^{(1)}, a_1=0,1}} \widehat{\pi}_{\mathcal{I},1}(a_1, \boldsymbol{x}_1) \ge c^*\Big) = 1, \tag{38}$$

for some constant $c^* > 0$. In addition, assume

$$E|\widehat{\pi}_{\mathcal{I}_{(j)},k}\{(\bar{\boldsymbol{A}}_0^{(k-1)}, a), \bar{\boldsymbol{X}}_0^{(k)}\} - \pi_k\{(\bar{\boldsymbol{A}}_0^{(k-1)}, a), \bar{\boldsymbol{X}}_0^{(k)}\}|^2 = o(|\mathcal{I}_{(j)}|^{-1/2}), \tag{39}$$
$$E|\widehat{h}_{\mathcal{I}_{(j)},k}\{(\bar{\boldsymbol{A}}_0^{(k-1)}, a), \bar{\boldsymbol{X}}_0^{(k)}\} - h_k\{(\bar{\boldsymbol{A}}_0^{(k-1)}, a), \bar{\boldsymbol{X}}_0^{(k)}\}|^2 = o(|\mathcal{I}_{(j)}|^{-1/2}), \tag{40}$$

for any $a = 0, 1$, $k = 1, \dots, K$, where $\bar{\boldsymbol{A}}_0^{(0)} = \emptyset$. Then, we have

$$\frac{\sqrt{n}(\widehat{V}_B - V_0)}{\widehat{\sigma}_B} \xrightarrow{d} N(0,1).$$

Appendix B. Detailed Estimating Procedure in Section 5.2

For any $a_1 \in \{0, 1\}$, we calculate

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{\pi_2,a_1} = \operatorname*{arg\,min}_{\boldsymbol{\xi} \in \mathbb{R}^{3(K+4)}} \sum_{i \in \mathcal{I}} \left(A_i^{(2)} - \sum_{j=1}^{K+4} N_j(X_i^{(2)}) \xi_j - \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(X_{i,k}^{(1)}) \xi_{j+k(K+4)} \right)^2 \times \mathbb{I}(A_i^{(1)} = a_1),$$

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_{2,1},a_1} = \operatorname*{arg\,min}_{\boldsymbol{\xi} \in \mathbb{R}^{3(K+4)}} \sum_{i \in \mathcal{I}} \left(Y_i - \sum_{j=1}^{K+4} N_j(X_i^{(2)}) \xi_j - \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(X_{i,k}^{(1)}) \xi_{j+k(K+4)} \right)^2 \times \mathbb{I}(A_i^{(2)} = 1, A_i^{(1)} = a_1),$$

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_{2,0},a_1} = \operatorname*{arg\,min}_{\boldsymbol{\xi} \in \mathbb{R}^{3(K+4)}} \sum_{i \in \mathcal{I}} \left(Y_i - \sum_{j=1}^{K+4} N_j(X_i^{(2)}) \xi_j - \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(X_{i,k}^{(1)}) \xi_{j+k(K+4)} \right)^2 \times \mathbb{I}(A_i^{(2)} = 0, A_i^{(1)} = a_1),$$

and compute

$$\widehat{\pi}_{\mathcal{I},2}((a_1, 1), \bar{\boldsymbol{x}}_2) = \min \left(\sum_{j=1}^{K+4} N_j(x_2) \widehat{\xi}_{\mathcal{I},j}^{\pi_2,1} + \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(x_{1,k}) \widehat{\xi}_{\mathcal{I},j+k(K+4)}^{\pi_2,1}, 0.05 \right) \times \mathbb{I}(A_i^{(1)} = a_1),$$

$$\widehat{\pi}_{\mathcal{I},2}((a_1, 0), \bar{\boldsymbol{x}}_2) = \min\{1 - \widehat{\pi}_{\mathcal{I},2}((a_1, 1), \bar{\boldsymbol{x}}_2), 0.05\},$$

$$\widehat{h}_{\mathcal{I},2}((a_1, a_2), \bar{\boldsymbol{x}}_2) = \left(\sum_{j=1}^{K+4} N_j(x_2) \widehat{\xi}_{\mathcal{I},j}^{h_{2,a_2},a_1} + \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(x_{1,k}) \widehat{\xi}_{\mathcal{I},j+k(K+4)}^{h_{2,a_2},a_1} \right) \times \mathbb{I}(A_i^{(1)} = a_1),$$

$$\widehat{d}_{\mathcal{I},2}(a_1, \bar{\boldsymbol{x}}_2) = \mathbb{I}[\widehat{h}_{\mathcal{I},2}\{(a_1, 1), \bar{\boldsymbol{x}}_2\} > \widehat{h}_{\mathcal{I},2}\{(a_1, 0), \bar{\boldsymbol{x}}_2\}].$$

Then we construct the pseudo outcome

$$\widehat{V}_{i,\mathcal{I}} = \frac{\mathrm{g}\{A_i^{(2)}, \widehat{d}_{\mathcal{I},2}(A_i^{(1)}, \bar{\boldsymbol{X}}_i^{(2)})\}}{\widehat{\pi}_{\mathcal{I},2}(\bar{\boldsymbol{A}}_i^{(2)}, \bar{\boldsymbol{X}}_i^{(2)})} \{Y_i - \widehat{h}_{\mathcal{I},2}(\bar{\boldsymbol{A}}_i^{(2)}, \bar{\boldsymbol{X}}_i^{(2)})\} + \widehat{h}_{\mathcal{I},2}[\{A_i^{(1)}, \widehat{d}_{\mathcal{I},2}(A_i^{(1)}, \bar{\boldsymbol{X}}_i^{(2)})\}, \bar{\boldsymbol{X}}_i^{(2)}],$$

for any $i \in \mathcal{I}_0$, and compute

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{\pi_1} = \operatorname*{arg\,min}_{\boldsymbol{\xi} \in \mathbb{R}^{2(K+4)}} \sum_{i \in \mathcal{I}} \left(A_i^{(1)} - \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(X_{i,k}^{(1)}) \xi_{j+(k-1)(K+4)} \right)^2,$$

$$\widehat{\boldsymbol{\xi}}_{\mathcal{I}}^{h_{a_1}} = \operatorname*{arg\,min}_{\boldsymbol{\xi} \in \mathbb{R}^{2(K+4)}} \sum_{i \in \mathcal{I}} \left(\widehat{V}_{i,\mathcal{I}} - \sum_{j=1}^{K+4} \sum_{k=1}^{2} N_j(X_{i,k}^{(1)}) \xi_{j+(k-1)(K+4)} \right)^2 \mathbb{I}(A_i^{(1)} = a_1),$$

for $a_1 = \{0, 1\}$. Finally, we set

$$\widehat{\pi}_{\mathcal{I}}(1, \boldsymbol{x}_1) = \min\left(\sum_{j=1}^{K+4}\sum_{k=1}^{2} N_j(x_{1,k})\widehat{\xi}^{\pi_1}_{\mathcal{I},j+(k-1)(K+4)}, 0.05\right),$$
$$\widehat{\pi}_{\mathcal{I}}(0, \boldsymbol{x}_1) = \min\{1 - \widehat{\pi}_{\mathcal{I}}(1, \boldsymbol{x}_1), 0.05\},$$

and

$$\begin{aligned}
\widehat{h}_{\mathcal{I}}(1, \boldsymbol{x}_1) &= \sum_{j=1}^{K+4}\sum_{k=1}^{2} N_j(x_{1,k})\widehat{\xi}^{h_1}_{\mathcal{I},j+(k-1)(K+4)}, \\
\widehat{h}_{\mathcal{I}}(0, \boldsymbol{x}_1) &= \sum_{j=1}^{K+4}\sum_{k=1}^{2} N_j(x_{1,k})\widehat{\xi}^{h_0}_{\mathcal{I},j+(k-1)(K+4)}, \\
\widehat{d}_{\mathcal{I}}(\boldsymbol{x}_1) &= \mathbb{I}\{\widehat{h}_{\mathcal{I}}(1, \boldsymbol{x}_1) > \widehat{h}_{\mathcal{I}}(0, \boldsymbol{x}_1)\}.
\end{aligned}$$

Similar to Section 5.1, we select the number of interior knots K via 5-folded cross-validation.

Appendix C. Proofs

C.1. Proof of Theorem 2

For any $\mathcal{I} = \{i_1, i_2, \dots, i_s\} \subseteq \mathcal{I}_0$, the estimated treatment regime $|\widehat{d}_{\mathcal{I}}(\cdot)|$ is upper bounded by 1. It follows from the ANOVA decomposition of Efron and Stein (1981) that

$$\begin{aligned}
\widehat{d}_{\mathcal{I}}(\boldsymbol{x}) = d_s(\boldsymbol{x}) + \sum_{i\in\mathcal{I}} d_{s,1}(O_i; \boldsymbol{x}) + \sum_{\substack{i,j\in\mathcal{I}\\ i\neq j}} d_{s,2}(O_i, O_j; \boldsymbol{x}) \qquad (41)\\
+ \sum_{\substack{i,j,k\in\mathcal{I}\\ i\neq j, i\neq k, j\neq k}} d_{s,3}(O_i, O_j, O_k; \boldsymbol{x}) + \cdots + d_{s,s}(O_{i_1}, O_{i_2}, \dots, O_{i_s}; \boldsymbol{x}), \quad \forall \boldsymbol{x},
\end{aligned}$$

where $d_s(x) = \mathrm{E}\widehat{d}_{\mathcal{I}}(\boldsymbol{x}) = \Pr(\widehat{d}_{\mathcal{I}}(\boldsymbol{x}) = 1)$, is the grand mean; $d_{s,1}(o; \boldsymbol{x}) = \mathrm{E}\{\widehat{d}_{\mathcal{I}}(\boldsymbol{x})|O_{i_1} = o\} - d_s(\boldsymbol{x})$, is the main effect;

$$\begin{aligned}
& d_{s,2}(o_1, o_2; \boldsymbol{x}) = \mathrm{E}\{\widehat{d}_{\mathcal{I}}(\boldsymbol{x})|O_{i_1} = o_1, O_{i_2} = o_2\} \\
- \;& \mathrm{E}\{\widehat{d}_{\mathcal{I}}(\boldsymbol{x})|O_{i_1} = o_1\} - \mathrm{E}\{\widehat{d}_{\mathcal{I}}(\boldsymbol{x})|O_{i_2} = o_2\} + d_s(\boldsymbol{x}),
\end{aligned}$$

is the second-order interaction; etc.

All the 2^s random variables on the right-hand side (RHS) of (41) are uncorrelated. Therefore,

$$\sum_{k=1}^{s}\binom{s}{k}\mathrm{E}d^2_{s,k}(O_{i_1}, O_{i_2}, \dots, O_{i_k}; \boldsymbol{x}) = \mathrm{Var}\{\widehat{d}_{\mathcal{I}}(\boldsymbol{x})\} \leq \mathrm{E}\widehat{d}^2_{\mathcal{I}}(\boldsymbol{x}) \leq 1. \qquad (42)$$

Let $\eta_6 = \widehat{V}_\infty^* - \eta_5$. In the following, we show $\eta_6 = o_p(n^{-1/2})$. With some calculations, we can show that

$$\eta_6 = \underbrace{\frac{1}{(n-s_n)\binom{n}{s_n}} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \frac{(2A_i-1)R_{\mathcal{I}}(\boldsymbol{X}_i)}{\pi(A_i,\boldsymbol{X}_i)}\{Y_i - h(A_i,\boldsymbol{X}_i)\}}_{\eta_7} + \underbrace{\frac{1}{(n-s_n)\binom{n}{s_n}} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \tau(\boldsymbol{X}_i)R_{\mathcal{I}}(\boldsymbol{X}_i)}_{\eta_8},$$

where $R_{\mathcal{I}}(\boldsymbol{x}) = \widehat{d}_{\mathcal{I}}(\boldsymbol{x}) - d_{s_n}(\boldsymbol{x})$.

Below, we break the proof into two steps. In the first step, we show $\eta_7 = o_p(n^{-1/2})$. In the second step, we prove $\eta_8 = o_p(n^{-1/2})$.

Step 1: For $i = 0, 1, \ldots, n$, let $\omega_{0,i} = (1-A_i)\{Y_i - h(0,\boldsymbol{X}_i)\}/\pi(0,\boldsymbol{X}_i)$ and $\omega_{1,i} = A_i\{Y_i - h(1,\boldsymbol{X}_i)\}/\pi(1,\boldsymbol{X}_i)$. We have

$$\eta_7 = -\frac{1}{(n-s_n)\binom{n}{s_n}} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \omega_{0,i}R_{\mathcal{I}}(\boldsymbol{X}_i) + \frac{1}{(n-s_n)\binom{n}{s_n}} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \omega_{1,i}R_{\mathcal{I}}(\boldsymbol{X}_i).$$

Below, we show

$$\eta_7^{(1)} \equiv \frac{1}{(n-s_n)\binom{n}{s_n}} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \omega_{0,i}R_{\mathcal{I}}(\boldsymbol{X}_i) = o_p(n^{-1/2}). \tag{43}$$

It follows from Equation 41 that

$$\eta_7^{(1)} = \frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \omega_{0,i} \left(\sum_{k=1}^{s_n} \sum_{\{j_1,\ldots,j_k\}\subseteq\mathcal{I}} d_{s_n,k}(O_{j_1},\ldots,O_{j_k};\boldsymbol{X}_i) \right).$$

Notice that $(n-s_n)\binom{n}{s_n} = (n-s_n)\binom{n}{n-s_n} = n\binom{n-1}{n-s_n-1} = n\binom{n-1}{s_n}$. With some calculations, we have

$$\eta_7^{(1)} \quad = \quad \frac{1}{n}\sum_{i=1}^{n} \omega_{0,i} \sum_{k=1}^{s_n} \frac{\binom{n-1-k}{s_n-k}}{\binom{n-1}{s_n}} \sum_{\{j_1,\ldots,j_k\}\subseteq\mathcal{I}_{(-i)}} d_{s_n,k}(O_{j_1},\ldots,O_{j_k};\boldsymbol{X}_i).$$

By (A1) and (A2), we have for any $i = 1, \ldots, n$,

$$\mathrm{E}(\omega_{0,i}|\boldsymbol{X}_i) = \mathrm{E}\left(\frac{1-A_i}{\pi(0,\boldsymbol{X}_i)}\{Y_i^*(0) - h(0,\boldsymbol{X}_i)\}|\boldsymbol{X}_i \right) = \mathrm{E}[\{Y_i^*(0) - h(0,\boldsymbol{X}_i)\}|\boldsymbol{X}_i] = 0.$$

Moreover, $d_{s_n,k_1}(O_{j_1^{(1)}},\cdots,O_{j_{k_1}^{(1)}};\boldsymbol{x})$ and $d_{s_n,k_2}(O_{j_1^{(2)}},\cdots,O_{j_{k_2}^{(2)}};\boldsymbol{x})$ are uncorrelated for any $\{j_1^{(1)},\ldots,j_{k_1}^{(1)}\} \neq \{j_1^{(2)},\ldots,j_{k_2}^{(2)}\}$. Therefore, for any $1 \le i_1, i_2 \le n$, $\{j_1^{(1)},\ldots,j_k^{(1)}\} \in \mathcal{I}_{(-i_1)}$, $\{j_1^{(2)},\ldots,j_k^{(2)}\} \in \mathcal{I}_{(-i_2)}$,

$$\mathrm{E}\omega_{0,i_1}\omega_{0,i_2}d_{s_n,k_1}(O_{j_1^{(1)}},\cdots,O_{j_{k_1}^{(1)}};\boldsymbol{X}_{i_1})d_{s_n,k_2}(O_{j_1^{(2)}},\cdots,O_{j_{k_2}^{(2)}};\boldsymbol{X}_{i_2}) \neq 0,$$

only when $k_1 = k_2$ and $\{i_1, j_1^{(1)}, \ldots, j_k^{(1)}\} = \{i_2, j_1^{(2)}, \ldots, j_k^{(2)}\}$. This occurs when either $i_1 = i_2$, $\{j_1^{(1)}, \ldots, j_k^{(1)}\} = \{j_1^{(2)}, \ldots, j_k^{(2)}\}$ or $i_2 = j_{k_0}^{(1)}$ for some $1 \le k_0 \le k$ and $\{i_1, j_1^{(1)}, \ldots, j_k^{(1)}\} - \{j_{k_0}^{(1)}\} = \{j_1^{(2)}, \ldots, j_k^{(2)}\}$. As a result, we have by Cauchy-Schwarz inequality and (42) that

$$\begin{aligned}
n\mathrm{E}(\eta_7^{(1)})^2 &= \frac{1}{n}\sum_{k=1}^{s_n} \frac{\binom{n-1-k}{s_n-k}^2}{\binom{n-1}{s_n}^2} \sum_{\substack{\{i_1, j_1^{(1)}, \ldots, j_k^{(1)}\} \\ = \{i_2, j_1^{(2)}, \ldots, j_k^{(2)}\}}} \mathrm{E}\prod_{a\in\{1,2\}} \{\omega_{0,i_a} d_{s_n,k}(O_{j_1^{(a)}}, \cdots, O_{j_k^{(a)}}; \boldsymbol{X}_{i_a})\} \\
&\le \frac{2}{n}\sum_{k=1}^{s_n} \frac{\binom{n-1-k}{s_n-k}^2}{\binom{n-1}{s_n}^2} \sum_{\substack{i\in\{1,\ldots,n\} \\ \{j_1,\ldots,j_k\}\subseteq \mathcal{I}_{(-i)}}} (k+1)\mathrm{E}\omega_{0,i}^2 d_{s_n,k}^2(O_{j_1}, \cdots, O_{j_k}; \boldsymbol{X}_i) \\
&= 2\sum_{k=1}^{s_n} \frac{\binom{n-1-k}{s_n-k}^2}{\binom{n-1}{s_n}^2} \binom{n-1}{k}(k+1)\mathrm{E}\omega_{0,0}^2 d_{s_n,k}^2(O_1, \cdots, O_k; \boldsymbol{X}_0) \\
&= 2\sum_{k=1}^{s_n} \frac{\binom{n-1-k}{s_n-k}^2}{\binom{n-1}{s_n}^2} \binom{n-1}{k}(k+1)\mathrm{E}[\{\mathrm{E}^{\boldsymbol{X}_0}\omega_{0,0}^2\}\{\mathrm{E}^{\boldsymbol{X}_0} d_{s_n,k}^2(O_1, \cdots, O_k; \boldsymbol{X}_0)\}] \\
&\le 2\max_{k\in\{1,\ldots,s_n\}} \frac{\binom{n-1-k}{s_n-k}^2}{\binom{n-1}{s_n}^2} \binom{n-1}{k} \frac{k+1}{\binom{s_n}{k}} \mathrm{E}\{\mathrm{E}^{\boldsymbol{X}_0}\omega_{0,0}^2\},
\end{aligned}$$

where the expectation $\mathrm{E}^{\boldsymbol{X}_0}$ is taken conditional on the covariates $\boldsymbol{X}_0$. With some calculations, we have for any $1 \le k \le s_n$,

$$\frac{\binom{n-1-k}{s_n-k}}{\binom{n-1}{s_n}} = \frac{\binom{s_n}{k}}{\binom{n-1}{k}} \quad \text{and} \quad \frac{\binom{n-1-k}{s_n-k}}{\binom{n-1}{s_n}} \le \frac{s_n^k}{(n-1)^k}.$$

It follows that

$$n\mathrm{E}(\eta_7^{(1)})^2 \le 2\max_{k\in\{1,\ldots,s_n\}} \frac{(k+1)s_n^k}{(n-1)^k}\mathrm{E}\omega_{0,0}^2. \tag{44}$$

Since $s_n/(n-1) \to 0$, we have for sufficiently large n that

$$\frac{(k+2)s_n^{k+1}}{(n-1)^{k+1}} \le \frac{(k+1)s_n^k}{(n-1)^k},$$

for any $k \ge 1$. Therefore,

$$\max_{k\in\{1,\ldots,s_n\}} \frac{(k+1)s_n^k}{(n-1)^k} \le \frac{s_n}{n-1} \to 0. \tag{45}$$

By (A3) and the condition $\max_{a\in\{0,1\}} \mathrm{E}\{Y_0^*(a)\}^2 < +\infty$, we have

$$\begin{aligned}
\mathrm{E}\omega_{0,0}^2 &\le \mathrm{E}\frac{\{Y_0^*(0) - h(0, \boldsymbol{X}_0)\}^2}{\pi^2(0, \boldsymbol{X}_0)} \le \frac{1}{c_0^2}\mathrm{E}\{Y_0^*(0) - h(0, \boldsymbol{X}_0)\}^2 \\
&= \frac{1}{c_0^2}\mathrm{Var}\{Y_0^*(0)\}^2 \le \frac{1}{c_0^2}\mathrm{E}\{Y_0^*(0)\}^2 = O(1). \qquad (46)
\end{aligned}$$

Combining this together with (44) and (45) yields $n\mathrm{E}(\eta_7^{(1)})^2 = o(1)$. Consequently, we obtain $\eta_7^{(1)} = o_p(n^{-1/2})$, by Markov's inequality. Similarly, we can show

$$\frac{1}{(n-s_n)\binom{n}{s_n}} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}} \sum_{i\in\mathcal{I}^c} \omega_{1,i} R_{\mathcal{I}}(\boldsymbol{X}_i) = o_p(n^{-1/2}).$$

Thus, we obtain $\eta_7 = o_p(n^{-1/2})$.

Step 2: Notice that

$$\begin{aligned}
\eta_8 &= \frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \tau(\boldsymbol{X}_i)[\widehat{d}_{\mathcal{I}}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)] \\
&= \underbrace{\frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \tau(\boldsymbol{X}_i)[\widehat{d}_{\mathcal{I}}(\boldsymbol{X}_i) - \mathbb{I}\{\tau(\boldsymbol{X}_i) > 0\}]}_{\eta_8^{(1)}} \\
&- \underbrace{\frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \tau(\boldsymbol{X}_i)[d_{s_n}(\boldsymbol{X}_i) - \mathbb{I}\{\tau(\boldsymbol{X}_i) > 0\}]}_{\eta_8^{(2)}}.
\end{aligned}$$

To prove $\eta_8 = o_p(n^{-1/2})$, it suffices to show $\eta_8^{(1)}, \eta_8^{(2)} = o_p(n^{-1/2})$.

Notice that $\eta_8^{(1)}$ and $\eta_8^{(2)}$ are non-positive, we have

$$\begin{aligned}
\mathrm{E}|\eta_8^{(1)}| &= \frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \mathrm{E}\tau(\boldsymbol{X}_i)[\mathbb{I}\{\tau(\boldsymbol{X}_i) > 0\} - \widehat{d}_{\mathcal{I}}(\boldsymbol{X}_i)], \\
\mathrm{E}|\eta_8^{(2)}| &= \frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \mathrm{E}\tau(\boldsymbol{X}_i)[\mathbb{I}\{\tau(\boldsymbol{X}_i) > 0\} - d_{s_n}(\boldsymbol{X}_i)].
\end{aligned}$$

Moreover, since $\mathrm{E}(\widehat{d}_{\mathcal{I}}(\boldsymbol{X}_i)|\boldsymbol{X}_i) = d_{s_n}(\boldsymbol{X}_i)$, we obtain that $\mathrm{E}|\eta_8^{(1)}| = \mathrm{E}|\eta_8^{(2)}|$. By Markov's inequality, it suffices to show $\mathrm{E}|\eta_8^{(1)}| = o(n^{-1/2})$. This is immediate to see by noting that

$$\begin{aligned}
&\frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \mathrm{E}\tau(\boldsymbol{X}_i)[\mathbb{I}\{\tau(\boldsymbol{X}_i) > 0\} - \widehat{d}_{\mathcal{I}}(\boldsymbol{X}_i)] \\
= &\frac{1}{\binom{n}{s_n}(n-s_n)} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \sum_{i\in\mathcal{I}^c} \{V(d^{opt,0}) - \mathrm{E}V(\widehat{d}_{\mathcal{I}})\} = V(d^{opt,0}) - \mathrm{E}V(\widehat{d}_{\mathcal{I}}) = O(s_n^{-\kappa^*}) = o(n^{-1/2}),
\end{aligned}$$

where the last two equalities are due to (A4) and the condition that $s_n \gg n^{-1/(2\kappa^*)}$. This proves $\eta_8 = o_p(n^{-1/2})$.

To summarize, we've shown $\eta_6 = o_p(n^{-1/2})$. Next, we show $V_0 = \mathrm{E}\eta_5 + o(n^{-1/2})$. It follows from the definitions of V_0, ψ_i and η_5 that

$$\begin{aligned} & V_0 - \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\psi_i(d_{s_n}) = V_0 - \mathrm{E}\psi_0(d_{s_n}) \qquad (47)\\ = \;& V_0 - \mathrm{E}[d_{s_n}(\boldsymbol{X}_0)h(1,\boldsymbol{X}_0) + \{1-d_{s_n}(\boldsymbol{X}_0)\}h(0,\boldsymbol{X}_0)]\\ = \;& \mathrm{E}[h(0,\boldsymbol{X}_0) + \tau(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}] - \mathrm{E}[d_{s_n}(\boldsymbol{X}_0)h(1,\boldsymbol{X}_0)\\ + \;& \{1-d_{s_n}(\boldsymbol{X}_0)\}h(0,\boldsymbol{X}_0)] = \mathrm{E}\tau(\boldsymbol{X}_0)[\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\} - d_{s_n}(\boldsymbol{X}_0)]. \end{aligned}$$

Using similar arguments in bounding $\mathrm{E}|\eta_8^{(1)}|$, we can show

$$\mathrm{E}|\tau(\boldsymbol{X}_0)||d_{s_n}(\boldsymbol{X}_0) - \mathbb{I}\{\tau(\boldsymbol{X}_0>0)\}| = o(n^{-1/2}).$$

In view of (47), this yields $V_0 = \mathrm{E}\eta_5 + o(n^{-1/2})$. The proof is hence completed. ■

C.2. Proof of Theorem 3

For any functions $d(\cdot)$, $\pi^*(\cdot,\cdot)$ and $h^*(\cdot,\cdot)$, define $V(d;\pi^*,h^*) = \mathrm{E}\psi_0(d,\pi^*,h^*)$. It is immediate to see that $V_0 = V(d^{opt};\pi,h)$ for any $d^{opt}\in\mathcal{D}^{opt}$. Let $n_{\mathcal{S}} = |\mathcal{S}_{N_0,s_n}|$. Before proving Theorem 3, we present the following lemma whose proof is given in Section C.7.

Lemma 8 *Under the conditions in Theorem 3, there exist some constants* $c_1,c_2,c_3>0$ *and* $0<p_*<1$, $\beta_0>0$ *such that*

$$Pr\left(\frac{\binom{n}{s_n} - n_{\mathcal{S}}}{\binom{n}{s_n}} \le c_1 p_*^{c_2 n^{\beta_0}}\right) \ge 1 - 2\exp(-c_3 n). \tag{48}$$

In addition, we have

$$\max_{i\in\{1,\dots,n\}} Pr\left(\left|\frac{n^{(i)}}{B} - \frac{n-s_n}{n}\right| \le \frac{\sqrt{\log n}}{\sqrt{n}}\right) \ge 1 - 4\exp(-c_4 n) - \frac{2}{B}, \tag{49}$$

for some constant $c_4>0$*, where* $n^{(i)} = \sum_{b=1}^{B}\mathbb{I}(i\in\mathcal{I}_b^c)$.

Theorem 2 implies that $\widehat{V}_\infty^* - V_0 = \eta_5 - \mathrm{E}\eta_5 + o_p(n^{-1/2})$. Under (A3) and the conditions $\max_{a\in\{0,1\}}\mathrm{E}\{Y_0^*(a)\}^2 = O(1)$, $\liminf_n \sigma_{s_n} > 0$, it follows from the central limit theorem that

$$\frac{\sqrt{n}(\widehat{V}_\infty^* - V_0)}{\sigma_{s_n}} \xrightarrow{d} N(0,1). \tag{50}$$

Assume for now, we've shown

$$\widehat{V}_B = \widehat{V}_\infty^* + o_p(n^{-1/2}), \tag{51}$$

and

$$\widehat{\sigma}_B^2 = \sigma_{s_n}^2 + o_p(1). \tag{52}$$

In view of (50), we have

$$\frac{\sqrt{n}(\widehat{V}_B - V_0)}{\widehat{\sigma}_B} \xrightarrow{d} N(0,1),$$

by Slutsky's theorem and the condition that $\liminf_n \sigma_{s_n} > 0$. Therefore, it suffices to show (51) and (52).

In the following, we break the proof into three steps. In the first step, we show $\widehat{V}_B = \widehat{V}_B^* + o_p(n^{-1/2})$ where

$$\widehat{V}_B^* \equiv \frac{1}{B}\sum_{b=1}^{B}\frac{1}{|\mathcal{I}_b^c|}\sum_{i\in\mathcal{I}_b^c}\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h) = \frac{1}{2B}\Big\{\frac{1}{t_n}\sum_{i\in\mathcal{I}_b^{c(2)}}\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h) + \frac{1}{t_n}\sum_{i\in\mathcal{I}_b^{c(1)}}\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\Big\}. \tag{53}$$

Next, we show $\widehat{V}_B^* = \widehat{V}_\infty^* + o_p(n^{-1/2})$. In the last step, we show (52) hold.

Step 1: Recall that $\widehat{V}_B$ is defined as

$$\widehat{V}_B = \frac{1}{2B}\sum_{b=1}^{B}\Big\{\frac{1}{t_n}\sum_{i\in\mathcal{I}_b^{c(2)}}\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}}) + \frac{1}{t_n}\sum_{i\in\mathcal{I}_b^{c(1)}}\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(2)}},\widehat{h}_{\mathcal{I}_b}^{(2)})\Big\}.$$

In view of (53), it suffices to show

$$\frac{1}{Bt_n}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\Big(\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\Big) = o_p(n^{-1/2}), \tag{54}$$

$$\frac{1}{Bt_n}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(1)}}\Big(\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(2)}},\widehat{h}_{\mathcal{I}_b^{(2)}}) - \psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\Big) = o_p(n^{-1/2}). \tag{55}$$

In the following, we prove (54). Let $\mathcal{A}_0$ denote the event defined in Equation 48. On the set $\mathcal{A}_0$, we have

$$n_S \geq \frac{1}{2}\binom{n}{s_n}, \tag{56}$$

for sufficiently large n. Notice that $\mathcal{A}_0$ depends only on the data subset $\{O_i\}_{i\in\mathcal{I}_0}$. By Lemma 8, we have $\Pr(\mathcal{A}_0) \to 1$.

For any $\mathcal{I} \subseteq \mathcal{I}_0$ with $|\mathcal{I}| = s_n$, let $\mathcal{P}(\mathcal{I})$ denote the set of partitions, i.e,

$$\mathcal{P}(\mathcal{I}) \equiv \Big\{(\mathcal{I}^{c(1)},\mathcal{I}^{c(2)}) : \mathcal{I}^{c(1)} \cup \mathcal{I}^{c(2)} = \mathcal{I}^c, \mathcal{I}^{c(1)} \cap \mathcal{I}^{c(2)} = \emptyset, |\mathcal{I}^{c(1)}| = |\mathcal{I}^{c(2)}| = t_n\Big\}.$$

Notice that $|\mathcal{P}(\mathcal{I}_1)| = |\mathcal{P}(\mathcal{I}_2)|$ for any subsets $\mathcal{I}_1, \mathcal{I}_2$ such that $|\mathcal{I}_1| = |\mathcal{I}_2|$. Define $\mathcal{P}_0 = |\mathcal{P}(\mathcal{I})|$ for any $\mathcal{I} \subseteq \mathcal{I}_0$ such that $|\mathcal{I}| = s_n$. For $j = 1, 2$, let $\mathcal{I}^{(j)} = \mathcal{I} \cup \mathcal{I}^{c(j)}$, we have

$$
\begin{aligned}
& \mathrm{E}\left|\frac{1}{Bt_n}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0) \\
\leq\ & \frac{1}{Bt_n}\sum_{b=1}^{B}\mathrm{E}\left|\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0) \\
=\ & \frac{1}{t_n}\mathrm{E}\left|\sum_{i\in\mathcal{I}_1^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_1};\widehat{\pi}_{\mathcal{I}_1^{(1)}},\widehat{h}_{\mathcal{I}_1^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}_1};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0) \\
=\ & \mathrm{E}\frac{1}{t_n n_S \mathcal{P}_0}\sum_{\substack{\mathcal{I}\in\mathcal{S}_{N_0,s_n}\\ (\mathcal{I}^{c(1)},\mathcal{I}^{c(2)})\in\mathcal{P}(\mathcal{I})}}\left|\sum_{i\in\mathcal{I}^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}};\widehat{\pi}_{\mathcal{I}^{(1)}},\widehat{h}_{\mathcal{I}^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0),
\end{aligned}
$$

where the first equality is due to the fact that $(\mathcal{I}_1, \mathcal{I}_1^{c(1)}, \mathcal{I}_1^{c(2)}), \ldots, (\mathcal{I}_B, \mathcal{I}_B^{c(1)}, \mathcal{I}_B^{c(2)})$ are independent and identically distributed conditional on $\{O_i\}_{i\in\mathcal{I}_0}$, the second equality is due to the fact that

$$
\begin{aligned}
& \mathrm{E}^{\{O_i\}_{i\in\mathcal{I}_1}}\left|\sum_{i\in\mathcal{I}_1^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_1};\widehat{\pi}_{\mathcal{I}^{(1)}},\widehat{h}_{\mathcal{I}_1^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}_1};\pi,h)\}\right| \\
=\ & \frac{1}{n_S\mathcal{P}_0}\sum_{\substack{\mathcal{I}\in\mathcal{S}_{N_0,s_n}\\ (\mathcal{I}^{c(1)},\mathcal{I}^{c(2)})\in\mathcal{P}(\mathcal{I})}}\left|\sum_{i\in\mathcal{I}^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}};\widehat{\pi}_{\mathcal{I}^{(1)}},\widehat{h}_{\mathcal{I}^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\}\right|,
\end{aligned}
$$

where $\mathrm{E}^{\{O_i\}_{i\in\mathcal{I}_1}}$ denotes the conditional expectation given $\{O_i\}_{i\in\mathcal{I}_1}$. It follows from (56) that

$$
\begin{aligned}
& \mathrm{E}\left|\frac{1}{Bt_n}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0) \qquad (57) \\
\leq\ & \frac{2}{t_n\binom{n}{s_n}\mathcal{P}_0}\mathrm{E}\sum_{\substack{\mathcal{I}\in\mathcal{S}_{N_0,s_n}\\ (\mathcal{I}^{c(1)},\mathcal{I}^{c(2)})\in\mathcal{P}(\mathcal{I})}}\left|\sum_{i\in\mathcal{I}^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}};\widehat{\pi}_{\mathcal{I}^{(1)}},\widehat{h}_{\mathcal{I}^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\}\right| \\
\leq\ & \frac{2}{t_n\binom{n}{s_n}\mathcal{P}_0}\mathrm{E}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n\\ (\mathcal{I}^{c(1)},\mathcal{I}^{c(2)})\in\mathcal{P}(\mathcal{I})}}\left|\sum_{i\in\mathcal{I}^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}};\widehat{\pi}_{\mathcal{I}^{(1)}},\widehat{h}_{\mathcal{I}^{(1)}}) - \psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\}\right|.
\end{aligned}
$$

Let $\mathcal{I}_*$ be a random subset uniformly sampled from $\{\mathcal{I} \subseteq \mathcal{I}_0 : |\mathcal{I}| = s_n\}$, independent of $\{O_i\}_{i\in\mathcal{I}_0}$. Given $\mathcal{I}_*$, let $\mathcal{I}_*^{c(1)}$ and $\mathcal{I}_*^{c(2)}$ denote the random partition of $\mathcal{I}_*^c$ generated by the algorithm in Section 2.4. Notice that $(\mathcal{I}_*, \mathcal{I}_*^{c(1)}, \mathcal{I}_*^{c(2)})$ is independent of $\{O_i\}_{i\in\mathcal{I}_0}$. So far, we have shown

$$\begin{aligned}
&\frac{1}{2}\mathrm{E}\left|\frac{1}{t_nB}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0)\\
\leq\;& \frac{1}{t_n}\mathrm{E}\left|\sum_{i\in\mathcal{I}_*^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})-\psi_i(\widehat{d}_{\mathcal{I}_*};\pi,h)\}\right|.
\end{aligned}$$

It follows from triangle inequality that

$$\begin{aligned}
&\frac{1}{2}\mathrm{E}\left|\frac{1}{t_nB}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0) \qquad (58)\\
&\leq \underbrace{\mathrm{E}\left|\frac{1}{t_n}\sum_{i\in\mathcal{I}_*^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_*};\pi,h)-\psi_i(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\pi,h)+V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\}\right|}_{\eta_9}\\
&\qquad + \underbrace{\mathrm{E}\left|V(\widehat{d}_{\mathcal{I}_*};\pi,h)-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\right|}_{\eta_{10}}.
\end{aligned}$$

Below, we prove $\eta_9, \eta_{10} = o(n^{-1/2})$. This implies for any $\varepsilon > 0$,

$$\begin{aligned}
&\Pr\left(\left|\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|>\frac{Bt_n\varepsilon}{\sqrt{n}}\right) \qquad (59)\\
&\leq \Pr\left(\left\{\left|\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|>\frac{Bt_n\varepsilon}{\sqrt{n}}\right\}\cap\mathcal{A}_0\right)+\Pr(\mathcal{A}_0^c)\\
&\leq \frac{\sqrt{n}}{\varepsilon}\mathrm{E}\left|\frac{1}{Bt_n}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\}\right|\mathbb{I}(\mathcal{A}_0)+\Pr(\mathcal{A}_0^c)=o(1).
\end{aligned}$$

Hence, (54) is proven.

By Cauchy-Schwarz inequality, we have

$$\eta_9^2 \leq \mathrm{E}\left|\frac{1}{t_n}\sum_{i\in\mathcal{I}_*^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_*};\pi,h)-\psi_i(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\pi,h)+V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\}\right|^2.$$

Conditional on $\{O_i\}_{i\in\mathcal{I}_*^{(1)}}$, $\mathcal{I}_*$ and $\mathcal{I}_*^{(1)}$,

$$\sum_{i\in\mathcal{I}_*^{c(2)}}\{\psi_i(\widehat{d}_{\mathcal{I}_*};\pi,h)-\psi_i(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\pi,h)+V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\}$$

corresponds to a sum of i.i.d mean zero random variables. Therefore, we have

$$\eta_9^2\leq\frac{1}{t_n}\mathrm{EVar}\left(\psi_0(\widehat{d}_{\mathcal{I}_*};\pi,h)-\psi_0(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\right|\{O_i\}_{i\in\mathcal{I}_*}^{(1)},\mathcal{I}_*,\mathcal{I}_*^{(1)}\bigg).\tag{60}$$

Notice that $t_n=(n-s_n)/2$. Under the condition that $s_n=o(n)$, we have

$$|t_n|\asymp n.\tag{61}$$

It thus follows from (60) that

$$\eta_9^2\leq\frac{C}{n}\mathrm{EVar}\left(\psi_0(\widehat{d}_{\mathcal{I}_*};\pi,h)-\psi_0(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\right|\{O_i\}_{i\in\mathcal{I}_*}^{(1)},\mathcal{I}_*,\mathcal{I}_*^{(1)}\bigg),$$

for some constant $C>0$. For any random variable $\mathbb{Z}$, we have $\mathrm{Var}(\mathbb{Z})\leq\mathrm{E}\mathbb{Z}^2$. Hence, we have

$$C^{-1}n\eta_9^2\asymp\mathrm{E}\left|\psi_0(\widehat{d}_{\mathcal{I}_*};\pi,h)-\psi_0(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\right|^2.\tag{62}$$

By Cauchy-Schwarz inequality, the right-hand side (RHS) of (62) can be upper bounded by

$$\underbrace{3\,\mathrm{E}\left|\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)}Y_0-\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\pi(A_0,\boldsymbol{X}_0)}Y_0\right|^2}_{\eta_9^{(1)}}\tag{63}$$

$$+\underbrace{3\,\mathrm{E}\left|\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)}\widehat{h}_{\mathcal{I}_*^{c(1)}}(A_0,\boldsymbol{X}_0)-\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\pi(A_0,\boldsymbol{X}_0)}h(A_0,\boldsymbol{X}_0)\right|^2}_{\eta_9^{(2)}}$$

$$+\underbrace{3\,\mathrm{E}|\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\{\widehat{h}_{\mathcal{I}_*^{(1)}}(1,\boldsymbol{X}_0)-h(1,\boldsymbol{X}_0)\}+\{1-\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}\{\widehat{h}_{\mathcal{I}_*^{c(1)}}(0,\boldsymbol{X}_0)-h(0,\boldsymbol{X}_0)\}|^2}_{\eta_9^{(3)}}.$$

In the following, we show $\eta_9^{(1)},\eta_9^{(2)},\eta_9^{(3)}=o(1)$. This together with (62) implies $\eta_9=o(n^{-1/2})$.

By Condition (A1), we have $|Y_0|^2\leq A_0|Y_0^*(1)|^2+(1-A_0)|Y_0^*(0)|^2$. Since $\max_{a\in\{0,1\}}\mathrm{E}|Y_0^*(a)|^2<\infty$, we obtain $\mathrm{E}|Y_0|^2<\infty$. As a result,

$$\mathrm{E}|Y_0|^2\mathbb{I}(|Y_0|>n^{1/4})\to 0.\tag{64}$$

Notice that

$$\begin{aligned}\eta_9^{(1)} &= \mathrm{E}\frac{Y_0^2|\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)-\pi(A_0,\boldsymbol{X}_0)|^2}{\widehat{\pi}^2_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)\pi^2(A_0,\boldsymbol{X}_0)} \le \frac{1}{c_0^2(c^*)^2}\mathrm{E}Y_0^2|\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)-\pi(A_0,\boldsymbol{X}_0)|^2 \\ &\le \frac{1}{c_0^2(c^*)^2}\mathrm{E}Y_0^2|\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)-\pi(A_0,\boldsymbol{X}_0)|^2\mathbb{I}(|Y_0>n^{-1/4}|) \\ &+ \frac{n^{1/2}}{c_0^2(c^*)^2}\mathrm{E}|\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)-\pi(A_0,\boldsymbol{X}_0)|^2 = o(1),\end{aligned}$$

where the first inequality is due to (9) and (A3), the last equality is due to (10) and (64).

By Cauchy-Schwarz inequality, we have

$$\begin{aligned}\eta_9^{(2)} &\le \underbrace{2\,\mathrm{E}\left|\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)}h(A_0,\boldsymbol{X}_0)-\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\pi(A_0,\boldsymbol{X}_0)}h(A_0,\boldsymbol{X}_0)\right|^2}_{\eta_9^{(4)}} \\ &+ \underbrace{2\,\mathrm{E}\left|\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)}\widehat{h}_{\mathcal{I}_*^{c(1)}}(A_0,\boldsymbol{X}_0)-\frac{\mathbb{I}\{A_0=\widehat{d}_{\mathcal{I}_*}(\boldsymbol{X}_0)\}}{\widehat{\pi}_{\mathcal{I}_*^{(1)}}(A_0,\boldsymbol{X}_0)}h(A_0,\boldsymbol{X}_0)\right|^2}_{\eta_9^{(5)}}.\end{aligned}$$

Using similar arguments in showing $\eta_9^{(1)}=o(1)$, we can show $\eta_9^{(4)}=o(1)$. Besides, under the conditions in Equations 9 and 11, we have

$$\eta_9^{(5)} \le \frac{1}{c^*}\mathrm{E}|\widehat{h}_{\mathcal{I}_*^{c(1)}}(A_0,\boldsymbol{X}_0)-h(A_0,\boldsymbol{X}_0)|^2 = o(1).$$

This shows $\eta_9^{(2)}=o(1)$. Under the condition that $\max_{a=0,1}\mathrm{E}|\widehat{h}_{\mathcal{I}}(a,\boldsymbol{X}_0)-h(a,\boldsymbol{X}_0)|^2 = o(|\mathcal{I}|^{-1/2})$, we have $\eta_9^{(3)}=o(1)$. In view of Equations (62) and (63), we've shown

$$\eta_9 = o(n^{-1/2}). \tag{65}$$

We next show $\eta_{10}=o(n^{-1/2})$. Note that for any regime d and functions π^*, h^*,

$$\begin{aligned}V(d;\pi^*,h^*) &= \mathrm{E}\left(\frac{\mathbb{I}\{A_0=d(\boldsymbol{X}_0)\}}{\pi^*(A_0,\boldsymbol{X}_0)}\{Y-h^*(A_0,\boldsymbol{X}_0)\}+h^*(d(\boldsymbol{X}_0),\boldsymbol{X}_0)\right) \\ &= \mathrm{E}h(d(\boldsymbol{X}_0),\boldsymbol{X}_0)+\mathrm{E}\left(\frac{\pi(1,\boldsymbol{X}_0)}{\pi^*(1,\boldsymbol{X}_0)}-1\right)d(\boldsymbol{X}_0)\{h(1,\boldsymbol{X}_0)-h^*(1,\boldsymbol{X}_0)\} \\ &+ \mathrm{E}\left(\frac{\pi(0,\boldsymbol{X}_0)}{\pi^*(0,\boldsymbol{X}_0)}-1\right)\{1-d(\boldsymbol{X}_0)\}\{h(0,\boldsymbol{X}_0)-h^*(0,\boldsymbol{X}_0)\}.\end{aligned}$$

Therefore,

$$\begin{aligned}|V(d;\pi^*,h^*)-V(d;\pi,h)| &\le \sum_{a=0,1}\mathrm{E}\left|\left(\frac{\pi(a,\boldsymbol{X}_0)}{\pi^*(a,\boldsymbol{X}_0)}-1\right)\{h(a,\boldsymbol{X}_0)-h^*(a,\boldsymbol{X}_0)\}\right| \\ &\le \frac{1}{\inf_{a=0,1,\boldsymbol{x}\in\mathbb{X}}\pi^*(a,\boldsymbol{x})}\sum_{a=0,1}\mathrm{E}|\pi(a,\boldsymbol{X}_0)-\pi^*(a,\boldsymbol{X}_0)||h(a,\boldsymbol{X}_0)-h^*(a,\boldsymbol{X}_0)| \\ &\le \frac{1}{\inf_{a=0,1,\boldsymbol{x}\in\mathbb{X}}\pi^*(a,\boldsymbol{x})}\sum_{a=0,1}\frac{1}{2}\mathrm{E}\left(|\pi(a,\boldsymbol{X}_0)-\pi^*(a,\boldsymbol{X}_0)|^2+|h(a,\boldsymbol{X}_0)-h^*(a,\boldsymbol{X}_0)|^2\right),\end{aligned}$$

where the last inequality follows by Cauchy-Schwarz inequality. Hence, under the conditions in Equations (9), (10) and (11), we have

$$\begin{aligned}
&\eta_{10} = \mathrm{E}|V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}}) - V(\widehat{d}_{\mathcal{I}_*};\pi,h)| \\
\leq\ & \frac{1}{2c^*}\sum_{a=0,1}\mathrm{E}\left(|\pi(a,\boldsymbol{X}_0) - \widehat{\pi}_{\mathcal{I}_*^{(1)}}(a,\boldsymbol{X}_0)|^2 + |h(a,\boldsymbol{X}_0) - \widehat{h}_{\mathcal{I}_*^{(1)}}(a,\boldsymbol{X}_0)|^2\right) \\
\leq\ & \frac{1}{c^*}\max_{a=0,1}\mathrm{E}\left(|\pi(a,\boldsymbol{X}_0) - \widehat{\pi}_{\mathcal{I}_*^{(1)}}(a,\boldsymbol{X}_0)|^2 + |h(a,\boldsymbol{X}_0) - \widehat{h}_{\mathcal{I}_*^{(1)}}(a,\boldsymbol{X}_0)|^2\right) = o(|\mathcal{I}_*^{(1)}|^{-1/2}).
\end{aligned}$$

Besides, similar to Equation 61, we can show $|\mathcal{I}_*^{(1)}| \asymp n$. Hence, we obtain $\eta_{10} = o(n^{-1/2})$. By Markov's inequality, this together with (65) yields (54). Similarly, we can show (55) holds. Therefore, we have $\widehat{V}_B = \widehat{V}_B^* + o_p(n^{-1/2})$.

Step 2: Recall that $\widehat{V}_B^*$ is defined as

$$\widehat{V}_B^* = \frac{1}{B}\sum_{b=1}^{B}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}_b^c}\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h).$$

The expectation and variance of $\widehat{V}_B^*$ conditional on $\{O_i\}_{i\in\mathcal{I}_0}$ are given by

$$\mathrm{E}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0}) = \frac{1}{n_{\mathcal{S}}}\sum_{\mathcal{I}\in\mathcal{S}_{N_0,s_n}}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h),$$

$$\mathrm{Var}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0}) = \frac{n_{\mathcal{S}}-1}{n_{\mathcal{S}}B}\widehat{s.e}^2\left(\left\{\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right\}_{\mathcal{I}\in\mathcal{S}_{N_0,s_n}}\right).$$

For any $\varepsilon > 0$, we have

$$\begin{aligned}
&\Pr(\sqrt{n}|\widehat{V}_B^* - \widehat{V}_\infty^*| > 2\varepsilon) \leq \Pr\left(\sqrt{n}|\mathrm{E}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0}) - \widehat{V}_\infty^*| > \varepsilon\right) \\
+\ & \Pr\left(\sqrt{n}|\widehat{V}_B^* - \mathrm{E}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0})| > \varepsilon\right) \leq \Pr\left(\left\{\sqrt{n}|\mathrm{E}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0}) - \widehat{V}_\infty^*| > \varepsilon\right\}\cap\mathcal{A}_0\right) \\
+\ & \Pr\left(\sqrt{n}|\widehat{V}_B^* - \mathrm{E}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0})| > \varepsilon\right) + \Pr(\mathcal{A}_0^c) \leq \frac{\sqrt{n}}{\varepsilon}\underbrace{\mathrm{E}|\mathrm{E}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0}) - \widehat{V}_\infty^*|\mathbb{I}(\mathcal{A}_0)}_{\zeta_1} \\
+\ & \frac{n}{\varepsilon^2}\underbrace{\mathrm{E}\mathrm{Var}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}_0})}_{\zeta_2} + \Pr(\mathcal{A}_0^c),
\end{aligned}$$

where the first inequality follows by Bonferroni's inequality and the last inequality is due to Markov's inequality.

By Lemma 8, to prove $\widehat{V}_B^* - \widehat{V}_\infty^* = o_p(n^{-1/2})$, it suffices to show $\zeta_1 = o(n^{-1/2})$ and $\zeta_2 = o(n^{-1})$. We first show $\zeta_1 = o(n^{-1/2})$. It follows from triangle inequality that

$$\zeta_1 \leq \underbrace{\mathrm{E}\left|\frac{1}{n_{\mathcal{S}}}\sum_{\mathcal{I}\in\mathcal{S}_{N_0,s_n}}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h) - \frac{1}{n_{\mathcal{S}}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right|\mathbb{I}(\mathcal{A}_0)}_{\zeta_1^{(1)}}$$

$$+ \underbrace{\mathrm{E}\left|\frac{1}{n_{\mathcal{S}}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h) - \frac{1}{\binom{n}{s_n}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right|\mathbb{I}(\mathcal{A}_0)}_{\zeta_1^{(2)}}$$

Recall that for any regime d and any $\mathcal{I}\subseteq\mathcal{I}_0$, $\sum_{i\in\mathcal{I}^c}\psi_i(d;\pi,h)/(n-s_n)$ is defined as

$$\frac{1}{|\mathcal{I}^c|}\sum_{i\in\mathcal{I}^c}\left(\frac{\mathbb{I}\{A_i = d(\boldsymbol{X}_i)\}}{\pi(A_i,\boldsymbol{X}_i)}\{Y_i - h(A_i,\boldsymbol{X}_i)\} + h(d(\boldsymbol{X}_i),\boldsymbol{X}_i)\right).$$

By Condition (A1) and (A3), we have

$$\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(d;\pi,h)\right| \leq \frac{1+(c^*)^{-1}}{|\mathcal{I}^c|}\sum_{i=1}^n\sum_{a\in\{0,1\}}\{|Y_i^*(a)| + |h(a,\boldsymbol{X}_i)|\}. \tag{66}$$

Notice that $\mathrm{E}Y_i^*(a) = h(a,\boldsymbol{X}_i)$. By Jensen's inequality, Cauchy-Schwarz inequality and the conditions $s_n = o(n)$, $\max_{a\in\{0,1\}}\mathrm{E}\{Y^*(a)\}^2 < +\infty$, we obtain

$$\mathrm{E}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(d;\pi,h)\right| \leq \frac{2n+2n(c^*)^{-1}}{n-s_n}\max_{a\in\{0,1\}}\mathrm{E}\{|Y_0^*(a)| + |h(a,\boldsymbol{X}_0)|\}$$
$$\leq \frac{4n+4n(c^*)^{-1}}{n-s_n}\max_{a\in\{0,1\}}\mathrm{E}\{|Y_0^*(a)|\} \leq \frac{4n+4n(c^*)^{-1}}{n-s_n}\sqrt{\max_{a\in\{0,1\}}\mathrm{E}^2\{|Y_0^*(a)|\}} = O(1).$$

Combining this together with (56) and the definition of $\mathcal{A}_0$, we have

$$\zeta_1^{(1)} = \mathrm{E}\left|\frac{1}{n_{\mathcal{S}}}\sum_{\substack{\mathcal{I}\in\mathcal{S}^c_{N_0,s_n}\\|\mathcal{I}|=s_n}}\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right|\mathbb{I}(\mathcal{A}_0)$$
$$\leq \frac{2\left(\binom{n}{s_n} - |\mathcal{S}_{N_0,s_n}|\right)}{\binom{n}{s_n}}\mathrm{E}\max_{\substack{\mathcal{I}\subseteq\mathcal{I}_0^c\\|\mathcal{I}|=s_n}}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right|\mathbb{I}(\mathcal{A}_0)$$
$$\leq 2c_1 p_*^{c_2 n^{\beta_0}}\mathrm{E}\max_{\substack{\mathcal{I}\subseteq\mathcal{I}_0^c\\|\mathcal{I}|=s_n}}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right| = O\left((p^*)^{c_2 n^{\beta_0}}\right).$$

Since $0 < p_* < 1$, $\beta_0 > 0$ and $c_2 > 0$, we have $\zeta_1^{(1)} = o(n^{-1/2})$. Similarly, we can show $\zeta_1^{(2)} = o(n^{-1/2})$. Therefore, we have $\zeta_1 = o(n^{-1/2})$.

Next we show $\zeta_2 = o(n^{-1})$. Recall that

$$\mathrm{Var}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}}) = \frac{n_{\mathcal{S}}-1}{n_{\mathcal{S}}B}\widehat{s.e}^2\left(\left\{\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right\}_{\mathcal{I}\in\mathcal{S}_{N_0,s_n}}\right),$$

we have

$$\mathrm{Var}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}}) \le \frac{1}{n_{\mathcal{S}}B}\sum_{\mathcal{I}\in\mathcal{S}_{N_0,s_n}}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right|^2. \tag{67}$$

Besides, it follows from (66) and Cauchy-Schwarz inequality that

$$\begin{aligned}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(d;\pi,h)\right|^2 &\le \frac{\{1+(c^*)^{-1}\}^2}{(n-s_n)^2}\left(\sum_{i=1}^n\sum_{a\in\{0,1\}}\{Y_i^*(a)+|h(a,\boldsymbol{X}_i)|\}\right)^2\\ &\le \frac{4n\{1+(c^*)^{-1}\}^2}{(n-s_n)^2}\sum_{i=1}^n\sum_{a\in\{0,1\}}\{|Y_i^*(a)|^2+|h(a,\boldsymbol{X}_i)|^2\}.\end{aligned}$$

By Jensen's inequality and the conditions $s_n = o(n)$, $\max_{a\in\{0,1\}}\mathrm{E}\{Y^*(a)\}^2 < +\infty$, we have

$$\mathrm{E}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(d;\pi,h)\right|^2 \le \frac{16n^2\{1+(c^*)^{-1}\}^2}{(n-s_n)^2}\mathrm{E}|Y_0^*(a)|^2 = O(1).$$

Combining this together with (67) yields

$$\zeta_2 = \mathrm{E}\mathrm{Var}(\widehat{V}_B^*|\{O_i\}_{i\in\mathcal{I}}) \le \mathrm{E}\frac{1}{n_{\mathcal{S}}B}\sum_{\mathcal{I}\in\mathcal{S}_{N_0,s_n}}\left|\frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\right|^2 = O(B^{-1}).$$

Notice that we require $B \gg n$. It follows that

$$\zeta_2 = O(B^{-1}) = o(n^{-1}).$$

Therefore, we've shown $\widehat{V}_B^* - \widehat{V}_\infty^* = o_p(n^{-1/2})$.

Step 3: Recall that $\widehat{\sigma}_B^2$ is defined as

$$\widehat{\sigma}_B^2 = \frac{1}{n-1}\sum_{i=1}^n\{\psi^{(i)}\}^2 - \frac{n}{n-1}(\bar{\psi})^2,$$

where

$$\psi^{(i)} = \frac{1}{n^{(i)}}\sum_{b=1}^B\left(\psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})\mathbb{I}(i\notin\mathcal{I}_b^{(1)}) + \psi_i(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(2)}},\widehat{h}_{\mathcal{I}_b^{(2)}})\mathbb{I}(i\notin\mathcal{I}_b^{(2)})\right),$$

and $\bar{\psi} = \sum_{i=1}^n \psi^{(i)}/n$. Let $\mathcal{A}^{(i)}$ denote the event defined in Equation 49. Since $s_n = o(n)$, when $\mathcal{A}^{(i)}$ holds, we have for sufficiently large n,

$$n^{(i)} \geq \frac{B}{2}. \tag{68}$$

For any $i \in \mathcal{I}_0$, define

$$\psi^{*(i)} = \frac{1}{n^{(i)}} \sum_{b=1}^{B} \psi_i(\widehat{d}_{\mathcal{I}_b}; \pi, h) \mathbb{I}(i \notin \mathcal{I}_b).$$

Using similar arguments in Step 1 of the proof, we can show

$$\max_{i \in \mathcal{I}_0} \mathrm{E}|\psi^{(i)} - \psi^{*(i)}|^2 \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) = o(1). \tag{69}$$

In addition, using similar arguments in bounding ζ_2 in Step 2 of the proof, we can show

$$\max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{*(i)}|^2 = O(1). \tag{70}$$

This together with (69) yields

$$\begin{aligned}\max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{(i)} + \psi^{*(i)}|^2 \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) &= \max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{(i)} - \psi^{*(i)} + 2\psi^{*(i)}|^2 \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) \\ &\leq 2 \max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{(i)} - \psi^{*(i)}|^2 \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) + 8 \max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{*(i)}|^2 = O(1). \end{aligned} \tag{71}$$

In view of (69) and (71), it follows from Cauchy-Schwarz inequality that

$$\begin{aligned}\max_{i \in \mathbb{I}_0} \mathrm{E}|(\psi^{(i)})^2 - (\psi^{*(i)})^2| \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) &= \max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{(i)} - \psi^{*(i)}||\psi^{(i)} + \psi^{*(i)}| \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) \\ &\leq \sqrt{\max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{(i)} + \psi^{*(i)}|^2 \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) \max_{i \in \mathbb{I}_0} \mathrm{E}|\psi^{(i)} - \psi^{*(i)}|^2 \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0)} = o(1).\end{aligned}$$

Hence,

$$\begin{aligned}&\mathrm{E}\left|\frac{1}{n-1}\sum_{i=1}^n \{\psi^{(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^n \{\psi^{*(i)}\}^2\right| \mathbb{I}(\mathcal{A}_0)\mathbb{I}(\cap_{i=1}^n \mathcal{A}^{(i)}) \\ \leq\ & \frac{n}{n-1} \max_{i \in \mathbb{I}_0} \mathrm{E}|(\psi^{(i)})^2 - (\psi^{*(i)})^2| \mathbb{I}(\mathcal{A}^{(i)} \cap \mathcal{A}_0) = o(1).\end{aligned} \tag{72}$$

Notice that $B \gg n$. By Lemma 8 and Bonferroni's inequality, we have

$$\begin{aligned}\Pr\left\{\mathcal{A}_0^c \cup \left(\cup_{i=1}^n \mathcal{A}^{(i)^c}\right)\right\} &\leq \Pr(\mathcal{A}_0^c) + \sum_{i=1}^n \Pr(\mathcal{A}^{(i)^c}) \\ &\leq \frac{n}{B} + 4n\exp(-c_4 n) + 2\exp(-c_3 n) \to 0.\end{aligned} \tag{73}$$

This together with (72) implies that

$$\begin{aligned}&\Pr\left(\left|\frac{1}{n-1}\sum_{i=1}^n \{\psi^{(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^n \{\psi^{*(i)}\}^2\right| > \varepsilon\right) \leq \Pr\left\{\mathcal{A}_0^c \cup \left(\cup_{i=1}^n \mathcal{A}^{(i)^c}\right)\right\} \\ +\ & \frac{1}{\varepsilon}\mathrm{E}\left|\frac{1}{n-1}\sum_{i=1}^n \{\psi^{(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^n \{\psi^{*(i)}\}^2\right| \mathbb{I}(\mathcal{A}_0)\mathbb{I}(\cap_{i=1}^n \mathcal{A}^{(i)}) \to 0,\end{aligned}$$

for any $\varepsilon > 0$. Therefore, we've shown

$$\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{(i)}\}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{*(i)}\}^2 + o_p(1). \tag{74}$$

Conditional on the event defined in $\mathcal{A}^{(i)}$, we have

$$\underbrace{\frac{n-s_n}{n} - \frac{\sqrt{\log n}}{\sqrt{n}}}_{p_L} \le \frac{n^{(i)}}{B} \le \underbrace{\frac{n-s_n}{n} + \frac{\sqrt{\log n}}{\sqrt{n}}}_{p_U}.$$

Let

$$\psi_L^{*(i)} = \frac{1}{Bp_U}\sum_{b=1}^{B}\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\mathbb{I}(i\notin\mathcal{I}_b) \text{ and } \psi_U^{*(i)} = \frac{1}{Bp_L}\sum_{b=1}^{B}\psi_i(\widehat{d}_{\mathcal{I}_b};\pi,h)\mathbb{I}(i\notin\mathcal{I}_b),$$

we have

$$\{\psi_L^{*(i)}\}^2 \le \{\psi^{*(i)}\}^2 \le \{\psi_U^{*(i)}\}^2.$$

Define

$$\psi_\infty^{*(i)} = \frac{1}{\binom{n-1}{s_n}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\|\mathcal{I}|=s_n}}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h)\mathbb{I}(i\notin\mathcal{I}) = \frac{1}{\binom{n-1}{s_n}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}}\psi_i(\widehat{d}_{\mathcal{I}};\pi,h).$$

We now claim

$$\max_{i\in\mathcal{I}_0}\mathrm{E}|\{\psi_\infty^{*(i)}\}^2 - \{\psi^{*(i)}\}^2|\mathbb{I}(\mathcal{A}_0\cap\mathcal{A}^{(i)}) = o(1). \tag{75}$$

To prove this, it suffices to show

$$\max_{i\in\mathcal{I}_0}\mathrm{E}|\{\psi_\infty^{*(i)}\}^2 - \{\psi_L^{*(i)}\}^2|\mathbb{I}(\mathcal{A}_0\cap\mathcal{A}^{(i)}) = o(1), \tag{76}$$

and

$$\max_{i\in\mathcal{I}_0}\mathrm{E}|\{\psi_\infty^{*(i)}\}^2 - \{\psi_U^{*(i)}\}^2|\mathbb{I}(\mathcal{A}_0\cap\mathcal{A}^{(i)}) = o(1). \tag{77}$$

(76) and (77) can be similarly proven as (74). We omit the technical details for brevity. Under (75), we have

$$\mathrm{E}\left|\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{*(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^{n}\{\psi_\infty^{*(i)}\}^2\right|\mathbb{I}\left(\mathcal{A}_0\cap(\cap_{j=1}^{n}\mathcal{A}^{(j)})\right) = o(1).$$

It thus follows from (73) and Markov's inequality that

$$\Pr\left(\left|\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{*(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^{n}\{\psi_\infty^{*(i)}\}^2\right| > \varepsilon\right) \le \Pr\left(\mathcal{A}_0^c\cup(\cup_{j=1}^{n}\mathcal{A}^{(i)^c})\right)$$
$$+\mathrm{E}\left|\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{*(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^{n}\{\psi_\infty^{*(i)}\}^2\right|\mathbb{I}\left(\mathcal{A}_0\cap(\cap_{j=1}^{n}\mathcal{A}^{(j)})\right) \to 0,$$

for any $\varepsilon > 0$. This implies

$$\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{*(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^{n}\{\psi_{\infty}^{*(i)}\}^2 = o_p(1). \tag{78}$$

Notice that we have $\psi_{\infty}^{*(i)} = \psi_i(\widehat{d}_{s_n}^{(-i)}; \pi, h)$. Based on the ANOVA decomposition (see (41)), we have

$$\widehat{d}_{s_n}^{(-i)}(\cdot) = d_{s_n}(\cdot) + \sum_{k=1}^{s_n} \frac{\binom{n-1-k}{s_n-k}}{\binom{n-1}{s_n}} \sum_{\{j_1,\dots,j_k\}\subseteq \mathcal{I}_{(-i)}} d_{s_n,k}(O_{j_1},\dots,O_{j_k};\cdot).$$

Using similar arguments in bounding $\eta_3^{(2)}$ and $\eta_3^{(3)}$ in the proof of Theorem 2, we can show

$$\max_{i\in\mathcal{I}_0} \mathrm{E}|\widehat{d}_{s_n}^{(-i)}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)|^2 = o(1), \tag{79}$$

and hence $|\widehat{d}_{s_n}^{(-i)}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)|^2 = o_p(1)$. By dominated convergence theorem, we obtain

$$\begin{aligned}
\mathrm{E}|\psi_i(\widehat{d}_{s_n}^{(-i)}; \pi, h) - \psi_i(d_{s_n}; \pi, h)|^2 &\le 4\mathrm{E}\frac{A_i|\widehat{d}_{s_n}^{(-i)}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)|^2}{\pi^2(1,\boldsymbol{X}_i)}|Y_i^*(1) - h(1,\boldsymbol{X}_i)|^2 \\
+ \quad & 4\mathrm{E}\frac{(1-A_i)|\widehat{d}_{s_n}^{(-i)}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)|^2}{\pi^2(0,\boldsymbol{X}_i)}|Y_i^*(0) - h(0,\boldsymbol{X}_i)|^2 \\
+ \quad & 4\mathrm{E}h^2(0,\boldsymbol{X}_i)|\widehat{d}_{s_n}^{(-i)}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)|^2 + 4\mathrm{E}h^2(1,\boldsymbol{X}_i)|\widehat{d}_{s_n}^{(-i)}(\boldsymbol{X}_i) - d_{s_n}(\boldsymbol{X}_i)|^2 = o(1).
\end{aligned}$$

Since $\psi_1(\widehat{d}_{s_n}^{(-i)}; \pi, h)$, $\psi_2(\widehat{d}_{s_n}^{(-i)}; \pi, h)$, $\cdots$, $\psi_n(\widehat{d}_{s_n}^{(-n)}; \pi, h)$ are exchangeable, we obtain

$$\max_{i\in\mathcal{I}_0} \mathrm{E}|\psi_i(\widehat{d}_{s_n}^{(-i)}; \pi, h) - \psi_i(d_{s_n}; \pi, h)|^2 = o(1). \tag{80}$$

By Markov's inequality, we obtain

$$\frac{1}{n-1}\sum_{i=1}^{n}\{\psi_{\infty}^{*(i)}\}^2 - \frac{1}{n-1}\sum_{i=1}^{n}\psi_i^2(d_{s_n}; \pi, h) = o_p(1).$$

Combining this with (74) and (78), we've shown

$$\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{(i)}\}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\psi_i^2(d_{s_n}; \pi, h) + o_p(1). \tag{81}$$

Similarly, we can show

$$\left(\frac{1}{n}\sum_{i=1}^{n}\psi^{(i)}\right)^2 = \left(\frac{1}{n}\sum_{i=1}^{n}\psi_i(d_{s_n}; \pi, h)\right)^2 + o_p(1) = \eta_5^2 + o_p(1).$$

Combining this together with (81), we have

$$\widehat{\sigma}_B^2 = \frac{1}{n-1}\sum_{i=1}^{n}\psi_i^2(d_{s_n}; \pi, h) - \frac{n}{n-1}\eta_5^2 + o_p(1).$$

Under the given conditions, it follows from law of larger numbers that

$$\frac{1}{n}\sum_{i=1}^{n}\psi_i^2(d_{s_n};\pi,h)=\mathrm{E}\psi_0^2(d_{s_n};\pi,h)+o_p(1),$$

and $\eta_5=\{\mathrm{E}\psi_0(d_{s_n};\pi,h)\}^2+o_p(1)$. Therefore, we have $\widehat{\sigma}_B^2=\sigma_{s_n}^2+o_p(1)$. The proof is hence completed. ■

C.3. Proof of Theorem 5

We first study the asymptotic property of $\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}};\pi,h)$. Next, we study the asymptotic property of $\sigma_{s_n}^2=\widetilde{\sigma}_0^2(d_{s_n};\pi,h)$. Finally, we bound the difference between the average lengths of the two CIs.

Let $\nu_0=\Pr\{\tau(\boldsymbol{X}_0)=0\}$ and $d_0=d(\boldsymbol{X}_0)$ for any function d. Since e_0 is independent of A_0 and $\boldsymbol{X}_0$, we have

$$\begin{aligned}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}};\pi,h)&=\mathrm{Var}\left(\frac{\mathrm{g}(A_0,\widehat{d}_{\mathcal{I}_{(j)},0})}{\pi(A_0,\boldsymbol{X}_0)}e_0+h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\\&=\mathrm{Var}\left(\frac{\mathrm{g}(A_0,\widehat{d}_{\mathcal{I}_{(j)},0})}{\pi(A_0,\boldsymbol{X}_0)}e_0\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)+\mathrm{Var}\{h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)|\{O_i\}_{i\in\mathcal{I}_{(j)}}\}.\end{aligned}\tag{82}$$

By definition, we have

$$h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)=\widehat{d}_{\mathcal{I}_{(j)},0}h(1,\boldsymbol{X}_0)+(1-\widehat{d}_{\mathcal{I}_{(j)},0})h(0,\boldsymbol{X}_0)=h(0,\boldsymbol{X}_0)+\tau(\boldsymbol{X}_0)\mathbb{I}\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)>0\}.$$

For any $d^{opt}\in\mathcal{D}^{opt}$ and $\varepsilon>0$, it follows from Markov's inequality that

$$\begin{aligned}&\Pr\left\{\mathrm{E}\left(|h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)-h(d_0^{opt},\boldsymbol{X}_0)|^2|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)>\varepsilon\right\}\\\leq\;&\frac{1}{\varepsilon^2}\mathrm{E}|h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)-h(d_0^{opt},\boldsymbol{X}_0)|^2=\frac{1}{\varepsilon^2}\mathrm{E}\tau^2(\boldsymbol{X}_0)|\mathbb{I}\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)>0\}-\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}|.\end{aligned}$$

Here, $h(d_0^{opt},\boldsymbol{X}_0)=h(0,\boldsymbol{X}_0)+\max\{\tau(\boldsymbol{X}_0),0\}$ is independent of d^{opt}.

Since $|\mathbb{I}\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)>0\}-\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}|\leq\mathbb{I}\{|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)-\tau(\boldsymbol{X}_0)|\geq\tau(\boldsymbol{X}_0)\}$, by Condition (A6) and Markov's inequality, we have

$$\begin{aligned}\mathrm{E}\tau^2(\boldsymbol{X}_0)|\mathbb{I}\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)>0\}-\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}|&\leq\mathrm{E}\mathbb{I}\{|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)-\tau(\boldsymbol{X}_0)|\geq\tau(\boldsymbol{X}_0)\}\\&\leq\mathrm{E}|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)-\tau(\boldsymbol{X}_0)|^2\to0,\end{aligned}$$

as $j\to\infty$. This implies that as $j\to\infty$,

$$\mathrm{E}\left(|h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)-h(d_0^{opt},\boldsymbol{X}_0)|^2|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\xrightarrow{P}0.\tag{83}$$

Under the given conditions, we can show as $j\to\infty$ that,

$$\mathrm{E}\left|\mathrm{Var}\{h(\widehat{d}_{\mathcal{I}_{(j)},0},\boldsymbol{X}_0)|\{O_i\}_{i\in\mathcal{I}_{(j)}}\}-\mathrm{Var}\{h(d_0^{opt},\boldsymbol{X}_0)\}\right|\to0.\tag{84}$$

Consequently, the second term on the second line of (82) is equivalent to $\text{Var}\{h(d_0^{opt}, \boldsymbol{X}_0)\}$.

Consider the first term. With some calculations, we have

$$\text{Var}\left(\frac{\text{g}(A_0, \widehat{d}_{\mathcal{I}_{(j)},0})}{\pi(A_0, \boldsymbol{X}_0)} e_0 \middle| \{O_i\}_{i\in\mathcal{I}_{(j)}} \right) = \sigma_0^2 \text{E}\left(\frac{\text{g}^2\{A_0, \widehat{d}_{\mathcal{I}_{(j)},0}\}}{\pi^2(A_0, \boldsymbol{X}_0)} \middle| \{O_i\}_{i\in\mathcal{I}_{(j)}} \right)$$

$$= \sigma_0^2 \text{E}\left(\frac{\widehat{d}^2_{\mathcal{I}_{(j)},0}}{\pi(1, \boldsymbol{X}_0)} + \frac{(1-\widehat{d}_{\mathcal{I}_{(j)},0})^2}{\pi(0, \boldsymbol{X}_0)} \middle| \{O_i\}_{i\in\mathcal{I}_{(j)}} \right) = \sigma_0^2 \text{E}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(1, \boldsymbol{X}_0)} + \frac{1-\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(0, \boldsymbol{X}_0)} \middle| \{O_i\}_{i\in\mathcal{I}_{(j)}} \right), \tag{85}$$

where the last equality is due to that $\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0) \in \{0, 1\}$.

In the following, we show

$$\lim_j \text{E}\left| \frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} - \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} \right| = 0. \tag{86}$$

Notice that

$$\begin{aligned} & \text{E}\left| \frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} - \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} \right| \\ \leq\ & \underbrace{\text{E}\left| \frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} - \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} \right| \mathbb{I}\{0 < \tau(\boldsymbol{X}_0) \leq j^{-1/4}\}}_{\zeta_3} \\ +\ & \underbrace{\text{E}\left| \frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} - \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} \right| \mathbb{I}\{\tau(\boldsymbol{X}_0) > j^{-1/4}\}}_{\zeta_4}. \end{aligned} \tag{87}$$

It follows from Condition (A3) and (A5) that

$$\zeta_3 \leq \frac{1}{c_0} \text{E}\mathbb{I}\{0 < \tau(\boldsymbol{X}_0) \leq j^{-1/4}\} \leq \frac{\bar{c}}{c_0} j^{-1/(4\alpha)} \to 0, \qquad \text{as } j \to \infty.$$

In addition, notice that $\widehat{d}_{\mathcal{I}_j}(\boldsymbol{x}) \neq \mathbb{I}\{\tau(\boldsymbol{x}) > 0\}$ only when $|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}) - \tau(\boldsymbol{x})| \geq |\tau(\boldsymbol{x})|$. It follows that

$$\begin{aligned} \zeta_4 &\leq \frac{1}{c_0} \text{E}|\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0) - \mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}|\mathbb{I}\{\tau(\boldsymbol{X}_0) > j^{-1/4}\} \\ &\leq \frac{1}{c_0} \text{E}\frac{|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0) - \tau(\boldsymbol{X}_0)|^2}{|\tau(\boldsymbol{X}_0)|^2} \mathbb{I}\{\tau(\boldsymbol{X}_0) > j^{-1/4}\} \\ &\leq \frac{j^{1/2}}{c_0} \text{E}|\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0) - \tau(\boldsymbol{X}_0)|^2 = O(j^{-\kappa_0+1/2}) \to 0, \qquad \text{as } j \to \infty, \end{aligned}$$

the last inequality is due to the relation that $\kappa_0 > (\alpha+2)/(2\alpha+2) > 1/2$. By (87), we've shown (86) holds. By Jensen's inequality, this further implies

$$\lim_j \text{E}\left| \text{E}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} - \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0) > 0\}}{\pi(1, \boldsymbol{X}_0)} \middle| \{O_i\}_{i\in\mathcal{I}_j} \right) \right| = 0. \tag{88}$$

Similarly, we can show

$$\lim_j \mathrm{E}\left|\mathrm{E}\left(\frac{\{1-\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\}\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}-\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\right|=0,$$
$$\lim_j \mathrm{E}\left|\mathrm{E}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(1,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\right|=0,$$
$$\lim_j \mathrm{E}\left|\mathrm{E}\left(\frac{\{1-\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{X}_0)\}\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(0,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\right|=0.$$

Combining these together with (88) yields

$$\begin{aligned}\mathrm{E}\Bigg|\mathrm{E}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(1,\boldsymbol{X}_0)}+\frac{1-\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(0,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)-\mathrm{E}\left(\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)}+\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}\right)\\ -\mathrm{E}\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(1,\boldsymbol{X}_0)}+\frac{1-\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(0,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\Bigg|=o(1),\qquad \text{as}\quad j\to\infty.\end{aligned}$$

Since $\ell_n\to\infty$, we have

$$\begin{aligned}&\max_{j\ge\ell_n}\mathrm{E}\Bigg|\mathrm{E}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(1,\boldsymbol{X}_0)}+\frac{1-\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(0,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)-\mathrm{E}\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)}\\ -\quad&\mathrm{E}\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}-\mathrm{E}\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(1,\boldsymbol{X}_0)}+\frac{1-\widehat{d}_{\mathcal{I}_{(j)},0}}{\pi(0,\boldsymbol{X}_0)}\middle|\{O_i\}_{i\in\mathcal{I}_{(j)}}\right)\Bigg|=o(1).\end{aligned}$$

This together with (84) and (85) yields

$$\sup_{j\ge\ell_n}\mathrm{E}\left|\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}};\pi,h)-\nu_1-\sigma_0^2\underbrace{\int_{\boldsymbol{x}\in\mathbb{X}}\left(\frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x})}{\pi(1,\boldsymbol{x})}+\frac{1-\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x})}{\pi(0,\boldsymbol{x})}\right)dF_X(\boldsymbol{x})}_{\kappa_j}\right|=o(1),\tag{89}$$

where $\mathbb{X}=\{\boldsymbol{x}\in\mathbb{X}:\tau(\boldsymbol{x})=0\}$, $F_X(\cdot)$ denotes the cumulative distribution function of $\boldsymbol{X}_0$, and

$$\nu_1=\mathrm{Var}\{h(d_0^{opt},\boldsymbol{X}_0)\}+\sigma_0^2\mathrm{E}\left(\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)}+\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}\right).$$

This further implies

$$\sup_{j\ge\ell_n}|\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}};\pi,h)-\nu_1-\sigma_0^2\mathrm{E}\kappa_j|=o(1).\tag{90}$$

Consider the expectation $\mathrm{E}\kappa_j$. Let

$$\kappa_j(\boldsymbol{x})=\frac{\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x})}{\pi(1,\boldsymbol{x})}+\frac{1-\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x})}{\pi(0,\boldsymbol{x})},$$

we have

$$\mathrm{E}\kappa_j = \int_{\boldsymbol{x}\in\mathbb{X}} \mathrm{E}\kappa_j(\boldsymbol{x}) dF_X(\boldsymbol{x}).$$

Under the conditions in (A7), we have

$$\mathrm{E}\kappa_j(\boldsymbol{x}) = \frac{\Pr\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}) > 0\}}{\pi(1,\boldsymbol{x})} + \frac{1-\Pr\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}) > 0\}}{\pi(0,\boldsymbol{x})} \to \frac{p_0(\boldsymbol{x})}{\pi(1,\boldsymbol{x})} + \frac{1-p_0(\boldsymbol{x})}{\pi(0,\boldsymbol{x})}.$$

By (A3), $|\kappa_j(\boldsymbol{x})|$ is uniformly bounded for any j and $\boldsymbol{x}$. It follows from dominated convergence theorem that

$$\mathrm{E}\kappa_j \to \mathrm{E}\left(\frac{p_0(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-p_0(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right)\mathbb{I}(\boldsymbol{X}_0 \in \mathbb{X}), \quad \text{as } j \to \infty,$$

and hence

$$\sup_{j\geq \ell_n}\left|\mathrm{E}\kappa_j - \mathrm{E}\left(\frac{p_0(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-p_0(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right)\mathbb{I}(\boldsymbol{X}_0 \in \mathbb{X})\right| = o(1).$$

Let

$$\nu_2 = \nu_1 + \sigma_0^2\mathrm{E}\left(\frac{p_0(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-p_0(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right)\mathbb{I}(\boldsymbol{X}_0 \in \mathbb{X}).$$

This together with (90) yields that

$$\sup_{j\geq \ell_n}|\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}};\pi,h) - \nu_2| = o(1). \tag{91}$$

By Condition (A3) and that $\max_{a\in\{0,1\}}\mathrm{E}|Y^*(a)|^2 < +\infty$, ν_2 is bounded away from 0 and ∞. It follows from (91) that

$$\frac{n-l_n}{\sum_{j=l_n}^{n-1}\{\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}};\pi,h)\}^{-1/2}} = \sqrt{\nu_2} + o(1). \tag{92}$$

This establishes the asymptotic length of the CI obtained by the one-step online estimator.

Now, let's consider

$$\sigma_{s_n}^2 = \widetilde{\sigma}_0^2(d_{s_n};\pi,h) = \mathrm{Var}\left(\frac{\mathrm{g}(A_0,d_{s_n,0})}{\pi(A_0,\boldsymbol{X}_0)}e_0\right) + \mathrm{Var}\{h(d_{s_n,0},\boldsymbol{X}_0)\}. \tag{93}$$

Similar to (84), we can show

$$\mathrm{Var}\{h(d_{s_n},\boldsymbol{X}_0)\} = \mathrm{Var}\{h(d_0^{opt},\boldsymbol{X}_0)\} + o(1). \tag{94}$$

With some calculations, we have

$$\begin{aligned}\mathrm{Var}\left(\frac{\mathrm{g}(A_0,d_{s_n,0})}{\pi(A_0,\boldsymbol{X}_0)}e_0\right) = \sigma_0^2\mathrm{E}\frac{\mathrm{g}^2(A_0,d_{s_n,0})}{\pi^2(A_0,\boldsymbol{X}_0)} &= \sigma_0^2\mathrm{E}\left(\frac{A_0 d_{s_n}^2(\boldsymbol{X}_0)}{\pi^2(1,\boldsymbol{X}_0)} + \frac{(1-A_0)\{1-d_{s_n}(\boldsymbol{X}_0)\}^2}{\pi^2(0,\boldsymbol{X}_0)}\right)\\ &= \sigma_0^2\mathrm{E}\left(\frac{d_{s_n}^2(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{\{1-d_{s_n}(\boldsymbol{X}_0)\}^2}{\pi(0,\boldsymbol{X}_0)}\right).\end{aligned}$$

Using similar arguments in Equation 90, we can show

$$\mathrm{E}\frac{d_{s_n}^2(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)} = \mathrm{E}\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)} + o(1),$$
$$\mathrm{E}\frac{\{1-d_{s_n}(\boldsymbol{X}_0)\}^2\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)} = \mathrm{E}\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)} + o(1),$$
$$\mathrm{E}\frac{d_{s_n}^2(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(1,\boldsymbol{X}_0)} = o(1) \quad \text{and} \quad \mathrm{E}\frac{\{1-d_{s_n}(\boldsymbol{X}_0)\}^2\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)} = o(1).$$

In addition, by Condition (A7), we have

$$\mathrm{E}\frac{d_{s_n}^2(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(1,\boldsymbol{X}_0)} = \mathrm{E}\frac{p_0^2(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(1,\boldsymbol{X}_0)} + o(1),$$
$$\mathrm{E}\frac{\{1-d_{s_n}(\boldsymbol{X}_0)\}^2\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(0,\boldsymbol{X}_0)} = \mathrm{E}\frac{\{1-p_0(\boldsymbol{X}_0)\}^2\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(0,\boldsymbol{X}_0)} + o(1).$$

Therefore,

$$\mathrm{Var}\left(\frac{\mathrm{g}(A_0,p_{s_n,0})}{\pi(A_0,\boldsymbol{X}_0)}e_0\right) = \sigma_0^2\mathrm{E}\left(\frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)>0\}}{\pi(1,\boldsymbol{X}_0)} + \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}\right)$$
$$+ \quad \sigma_0^2\mathrm{E}\frac{p_0^2(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(1,\boldsymbol{X}_0)} + \sigma_0^2\mathrm{E}\frac{\{1-p_0(\boldsymbol{X}_0)\}^2\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(0,\boldsymbol{X}_0)} + o(1).$$

Combining this together with (93) and (94) yields

$$\widetilde{\sigma}_0^2(d_{s_n};\pi,h) \to \nu_1 + \sigma_0^2\mathrm{E}\left(\frac{p_0^2(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)} + \frac{\{1-p_0(\boldsymbol{X}_0)\}^2}{\pi(1,\boldsymbol{X}_0)}\right)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} \equiv \nu_3. \tag{95}$$

Notice that

$$\nu_2 - \nu_3 = \sigma_0^2\mathrm{E}\left(\frac{p_0(\boldsymbol{X}_0)-p_0^2(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)} + \frac{1-p_0(\boldsymbol{X}_0)-\{1-p_0(\boldsymbol{X}_0)\}^2}{\pi(1,\boldsymbol{X}_0)}\right)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} \tag{96}$$
$$= \sigma_0^2\mathrm{E}\frac{p_0(\boldsymbol{X}_0)\{1-p_0(\boldsymbol{X}_0)\}}{\pi(0,\boldsymbol{X}_0)\pi(1,\boldsymbol{X}_0)}\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} \geq \frac{\sigma_0^2(d^*)^2}{c_0^2}\mathrm{E}\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} = \frac{\sigma_0^2(d^*)^2\nu_0}{c_0^2},$$

where the last inequality is due to Condition (A3) and (A7). It follows that

$$\sqrt{\nu_2} - \sqrt{\nu_3} = \frac{\nu_2-\nu_3}{\sqrt{\nu_2}+\sqrt{\nu_3}} \geq \frac{\sigma_0^2(d^*)^2\nu_0}{2\sqrt{\nu_2}c_0^2}. \tag{97}$$

By the definition of ν_2 and Condition (A3), we have

$$\nu_2 \leq \tilde{c} + \sigma_0^2 c_0^{-1}\mathrm{E}[\mathbb{I}\{\tau(\boldsymbol{X}_0)\neq 0\} + \mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}] = \tilde{c} + \sigma_0^2 c_0^{-1}.$$

This together with (97) yields

$$\sqrt{\nu_2} - \sqrt{\nu_3} \geq \frac{\sigma_0^2(d^*)^2\nu_0}{2\sqrt{\tilde{c}+c_0^{-1}\sigma_0^2}c_0^2}.$$

In view of (13) and (26), we've shown

$$\sqrt{n}\mathrm{E}L(\widehat{V}^{on},\alpha) - \sqrt{n}\mathrm{E}L(\widehat{V}_B,\alpha) \geq \frac{z_{\alpha/2}\sigma_0^2(d^*)^2\nu_0}{c_0^2\sqrt{\tilde{c}+c_0^{-1}\sigma_0^2}} + o(1).$$

This completes the proof. ■

C.4. Proof of Theorem 4

Similar to (91), we can show that

$$\mathrm{E}\mathrm{Var}[\psi_0(\widehat{d}_{\mathcal{I}_*})|\{O_i\}_{i\in\mathcal{I}_*}] = \nu_2 + o(1),$$

where the definition of ν_2 is given in the proof of Theorem 5. This together with (20) yields that

$$\mathrm{MSE}(\widehat{V}^{ss}) = \nu_2 + o(1).$$

By (95) and (21), we obtain

$$\mathrm{MSE}(\widehat{V}_B) = \nu_3 + o(1),$$

where the definition of ν_3 is given in the proof of Theorem 5. In view of (96), it is immediate to see the assertion in Theorem 4 holds. ■

C.5. Proof of Theorem 6

Recall that $\nu_0 = \Pr\{\tau(\boldsymbol{X}_0) = 0\}$. By (13), (17) and (28), it suffices to show

$$\inf_{d^{opt}\in\mathcal{D}^{opt}} \mathrm{Var}\{\psi_0(d^{opt})\} \geq \mathrm{Var}\{\psi_0(d_{s_n})\} + c^{**}\sigma_0^2\nu_0 + o(1). \tag{98}$$

Similar to (82) and (85), we have

$$\begin{aligned}\mathrm{Var}\{\psi_0(d^{opt})\} &= \mathrm{Var}\left(\frac{\mathrm{g}\{A_0, d^{opt}(\boldsymbol{X}_0)\}}{\pi(A_0,\boldsymbol{X}_0)}e_0\right) + \mathrm{Var}\{h(d^{opt}(\boldsymbol{X}_0),\boldsymbol{X}_0)\} \\ &= \mathrm{Var}\{h(d^{opt}(\boldsymbol{X}_0),\boldsymbol{X}_0)\} + \sigma_0^2\mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-d^{opt}(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right). \end{aligned} \tag{99}$$

By Lemma 1, we have

$$\begin{aligned}\mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-d^{opt}(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right) &= \mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-d^{opt}(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} \\ &+\mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)\mathbb{I}\{\tau(\boldsymbol{X}_0>0)\}}{\pi(1,\boldsymbol{X}_0)} + \frac{\{1-d^{opt}(\boldsymbol{X}_0)\}\mathbb{I}\{\tau(\boldsymbol{X}_0)<0\}}{\pi(0,\boldsymbol{X}_0)}\right).\end{aligned}$$

This together with (99) gives

$$\mathrm{Var}\{\psi_0(d^{opt})\} = \nu_1 + \sigma_0^2\mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-d^{opt}(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)}\right)\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}$$

In view of (95) and Condition (A7), we have

$$\sigma_{s_n}^2 \leq \nu_1 + \sigma_0^2 \mathrm{E}\left(\frac{(1-d^*)^2}{\pi(0,\boldsymbol{X}_0)} + \frac{(1-d^*)^2}{\pi(1,\boldsymbol{X}_0)} \right) \mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}.$$

To prove (98), it suffices to show

$$\inf_{d^{opt}\in\mathcal{D}^{opt}} \mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-d^{opt}(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)} \right) \mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} \geq (1-d^*)^2 \mathrm{E} \sum_{a=0,1} \frac{\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\}}{\pi(a,\boldsymbol{X}_0)} + c^{**}\nu_0.$$

Notice that $d^{opt}(\boldsymbol{X}_0)\in\{0,1\}$. For any $\boldsymbol{x}\in\mathbb{X}$, we have

$$\begin{aligned} & \frac{d^{opt}(\boldsymbol{x})}{\pi(1,\boldsymbol{x})} + \frac{1-d^{opt}(\boldsymbol{x})}{\pi(0,\boldsymbol{x})} - (1-d^*)^2 \left(\frac{1}{\pi(0,\boldsymbol{x})} + \frac{1}{\pi(1,\boldsymbol{x})} \right) \\ \geq \quad & \min_{a=0,1} \frac{1}{\pi(a,\boldsymbol{x})} - (1-d^*)^2 \left(\frac{1}{\pi(0,\boldsymbol{x})} + \frac{1}{\pi(1,\boldsymbol{x})} \right) \geq c^{**}. \end{aligned}$$

Thus, for any $d^{opt}\in\mathcal{D}^{opt}$, we have

$$\mathrm{E}\left(\frac{d^{opt}(\boldsymbol{X}_0)}{\pi(1,\boldsymbol{X}_0)} + \frac{1-d^{opt}(\boldsymbol{X}_0)}{\pi(0,\boldsymbol{X}_0)} - \sum_{a=0,1} \frac{(1-d^*)^2}{\pi(a,\boldsymbol{X}_0)} \right) \mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} \geq c^{**}\mathrm{E}\mathbb{I}\{\tau(\boldsymbol{X}_0)=0\} = c^{**}\nu_0.$$

The proof is hence completed. ■

C.6. Proof of Theorem 7

For notational convenience, we use a shorthand and write $\widehat{V}_i^{(k)}(d;\pi,h)$ as $\widehat{V}_i^{(k)}(d)$ for any $k=2,\dots,K, i=0,1,\dots,n$ and any dynamic treatment regime d. In addition, for any $d=\{d_k\}_{k=1}^K$, $\pi^*=\{\pi_k^*\}_{k=1}^K$, $h^*=\{h_k^*\}_{k=1}^K$ and $i=0,1,\dots,n$, let $d_{k,i}=d_k(\bar{\boldsymbol{A}}_i^{(k-1)},\bar{\boldsymbol{X}}_i^{(k)})$, $\pi_{k,i}^*=\pi_k^*(\bar{\boldsymbol{A}}_i^{(k)},\bar{\boldsymbol{X}}_i^{(k)})$, $h_{k,i}^*=h_k^*(\bar{\boldsymbol{A}}_i^{(k)},\bar{\boldsymbol{X}}_i^{(k)})$, $\forall k=2,\dots,K$ and $d_{1,i}=d_1(\boldsymbol{X}_i^{(1)})$, $\pi_{1,i}^*=\pi_1^*(\boldsymbol{X}_i^{(1)})$, $h_{1,i}^*=h_1^*(\boldsymbol{X}_i^{(i)})$. When $\pi^*=\pi$ and $h^*=h$, we write $\widehat{V}_i^{(k)}(d;\pi,h)$ as $\widehat{V}_i^{(k)}(d)$ for any dynamic treatment regime d.

To prove Theorem 7, we break the proof into four steps. In the first step, we show $\widehat{V}_B=\widehat{V}_B^*+o_p(n^{-1/2})$ where

$$\widehat{V}_B^* = \frac{1}{B}\sum_{b=1}^B \frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c} \widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}}).$$

In the second step, we show $\widehat{V}_B^*=\widehat{V}_\infty^*+o_p(n^{-1/2})$ where

$$\widehat{V}_\infty^* = \frac{1}{\binom{n}{s_n}} \sum_{\mathcal{I}\subseteq\mathcal{I}_0,|\mathcal{I}|=s_n} \frac{1}{n-s_n}\sum_{i\in\mathcal{I}^c} \widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}}).$$

In the third step, we show $\sqrt{n}(\widehat{V}_\infty^*-V_0)/\sigma_{s_n} \xrightarrow{d} N(0,1)$. In the last step, we show $\widehat{\sigma}_B^2 = \sigma_{s_n}^2+o_p(1)$. The proof is hence completed.

Step 1: By the definitions of $\widehat{V}_B$ and $\widehat{V}_B^*$, we need to show

$$\frac{1}{t_n B}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(2)}}\left(\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b})\right)=o_p(n^{-1/2}),$$

$$\frac{1}{t_n B}\sum_{b=1}^{B}\sum_{i\in\mathcal{I}_b^{c(1)}}\left(\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b};\widehat{\pi}_{\mathcal{I}_b^{(1)}},\widehat{h}_{\mathcal{I}_b^{(1)}})-\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b})\right)=o_p(n^{-1/2}).$$

For any $d=\{d_k\}_{k=1}^K$, $\pi^*=\{\pi_k^*\}_{k=1}^K$ and $h^*=\{h_k^*\}_{k=1}^K$, define $V(d;\pi^*,h^*)=\mathrm{E}\widehat{V}_0^{(1)}(d;\pi^*,h^*)$. Using similar arguments in Equations 57—59, it suffices to show $\eta_{11},\eta_{12}=o(n^{-1/2})$ where

$$\eta_{11}=\mathrm{E}\left|\frac{1}{t_n}\sum_{i\in\mathcal{I}_*^{c(2)}}\{\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_*};\pi,h)-\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\pi,h)+V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\}\right|,$$

$$\eta_{12}=\mathrm{E}\left|V(\widehat{d}_{\mathcal{I}_*};\pi,h)-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}},\widehat{h}_{\mathcal{I}_*^{(1)}})\right|,$$

where $\mathcal{I}_*$ denotes a random subset uniformly sampled from the set $\{\mathcal{I}\subseteq\mathcal{I}_0:|\mathcal{I}|=s_n\}$, $\mathcal{I}_*^{c(1)}$ and $\mathcal{I}_*^{c(2)}$ correspond to a random partition of $\mathcal{I}_*^c$ with $|\mathcal{I}_*^{c(1)}|=|\mathcal{I}_*^{c(2)}|=t_n=(n-s_n)/2$, and $\mathcal{I}_*^{(j)}=\mathcal{I}_*^{c(j)}\cup\mathcal{I}_*$ for $j=1,2$.

Define the functions $\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l)}=\{\widehat{\pi}_{\mathcal{I}_*^{(1)},k}^{(l)}\}_{k=1}^K$, $\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l)}=\{\widehat{h}_{\mathcal{I}_*^{(1)},k}^{(l)}\}_{k=1}^K$ as follows:

$$\widehat{\pi}_{\mathcal{I}_*^{(1)},k}^{(l)}=\pi_k\mathbb{I}(l<k)+\widehat{\pi}_{\mathcal{I}_*^{(1)},k}\mathbb{I}(l\ge k)\text{ and }\widehat{h}_{\mathcal{I}_*^{(1)},k}^{(l)}=h_k\mathbb{I}(l<k)+\widehat{h}_{\mathcal{I}_*^{(1)},k}\mathbb{I}(l\ge k),$$

for any $k=1,\dots,K$, $l=0,\dots,K$. Notice that for $l=0,1,2,\dots,K-1$, $\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l+1)},\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l+1)})-\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l)},\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l)})$ equals

$$\begin{aligned}&\prod_{j=1}^{l}\frac{\mathrm{g}(A_i^{(j)},\widehat{d}_{\mathcal{I}_*,j,i})}{\pi_{j,i}}\left(\frac{\mathrm{g}(A_i^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,i})}{\widehat{\pi}_{\mathcal{I}_*,l+1,i}}\{\widehat{V}_i^{(l+2)}(\widehat{d}_{\mathcal{I}_*})-\widehat{h}_{\mathcal{I}_*^{(1)},l+1,i}\}\right.\\&+\ \widehat{h}_{\mathcal{I}_*^{(1)},l+1}\{(\bar{\boldsymbol{A}}_i^{(l)},\widehat{d}_{\mathcal{I}_*,l+1,i}),\bar{\boldsymbol{X}}_i^{(l+1)}\}-\frac{\mathrm{g}(A_i^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,i})}{\pi_{l+1,i}}\{\widehat{V}_i^{(l+2)}(\widehat{d}_{\mathcal{I}_*})-h_{l+1,i}\}\\&-\ \left.h_{l+1}\{(\bar{\boldsymbol{A}}_i^{(l)},\widehat{d}_{\mathcal{I}_*,l+1,i}),\bar{\boldsymbol{X}}_i^{(l+1)}\}\right),\end{aligned}$$

where $\widehat{V}_i^{(K+1)}(\widehat{d}_{\mathcal{I}_*};\pi,h)=Y_i$ and $\bar{\boldsymbol{A}}_i^{(0)}=\emptyset$, for $i=0,1,\dots,n$.

Using similar arguments in bounding η_9 in the proof of Theorem 3, we can show

$$\begin{aligned}\max_{l=0,\dots,K-1}\mathrm{E}\Bigg|\frac{1}{t_n}\sum_{i\in\mathcal{I}_*^{c(2)}}&\{\widehat{V}_i(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l+1)},\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l+1)})-\widehat{V}_i(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l)},\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l)})\\&-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l+1)},\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l+1)})+V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}_{\mathcal{I}_*^{(1)}}^{(l)},\widehat{h}_{\mathcal{I}_*^{(1)}}^{(l)})\}\Bigg|=o(n^{-1/2}),\end{aligned}$$

for any l. An application of triangle inequality yields $\eta_{11} = o(n^{-1/2})$.

Next we show $\eta_{12} = o(n^{-1/2})$. Notice that

$$\mathrm{E}\{\widehat{V}_0^{(K)}(\widehat{d}_{\mathcal{I}})|\widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(K)}, \bar{\boldsymbol{X}}_0^{(K)}\} = h_K\{(\bar{\boldsymbol{A}}_0^{(K-1)}, \widehat{d}_{\mathcal{I},K,0}), \bar{\boldsymbol{X}}_0^{(K)}\},$$

for any $\mathcal{I} \subseteq \mathcal{I}_0$ with $|\mathcal{I}| = s_n$. Therefore, we have for any $d^{opt} \in \mathcal{D}^{opt}$,

$$\begin{aligned} & \mathrm{E}\left|\mathrm{E}\left(\widehat{V}_0^{(K)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(K)}(d^{opt})\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(K)}, \bar{\boldsymbol{X}}_0^{(K)}\right)\Big| \\ = \ & \mathrm{E}\left|\tau_K(\bar{\boldsymbol{A}}_0^{(K-1)}, \bar{\boldsymbol{X}}_0^{(K)})\right| \left|\widehat{d}_{\mathcal{I},K,0} - \mathbb{I}\{\tau_K(\bar{\boldsymbol{A}}_0^{(K-1)}, \bar{\boldsymbol{X}}_0^{(K)}) > 0\}\right|. \end{aligned}$$

Under Condition (C4) and (C5), using similar arguments in bounding $\eta_8^{(1)}$ in the proof of Theorem 2, we have

$$\mathrm{E}\left|\mathrm{E}\left(\widehat{V}_0^{(K)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(K)}(d^{opt})\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(K)}, \bar{\boldsymbol{X}}_0^{(K)}\right)\Big| = o(n^{-1/2}). \tag{100}$$

Assume for now, we've shown

$$\mathrm{E}\left|\mathrm{E}\left(\widehat{V}_0^{(k+1)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k+1)}(d^{opt})\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k+1)}, \bar{\boldsymbol{X}}_0^{(k+1)}\right)\Big| = o(n^{-1/2}). \tag{101}$$

We aim to show the above expression holds with the superscript $(k+1)$ replaced by (k) as well. By the definition of $\widehat{V}_0^{(k)}(d)$, we have

$$\widehat{V}_0^{(k)}(d) = \frac{\mathrm{g}(A_0^{(k)}, d_{k,0})}{\pi_{k,0}}\{\widehat{V}_0^{(k+1)}(d) - h_{k,0}\} + h_k\{(\bar{\boldsymbol{A}}_0^{(k-1)}, d_{k,0}), \bar{\boldsymbol{X}}_0^{(k)}\}.$$

Therefore,

$$\begin{aligned} & \mathrm{E}\left|\mathrm{E}\left(\widehat{V}_0^{(k)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k)}(d^{opt})\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k)}, \bar{\boldsymbol{X}}_0^{(k)}\right)\Big| \\ \le \ & \underbrace{\mathrm{E}\left|\mathrm{E}\left\{\frac{\mathrm{g}(A_0^{(k)}, \widehat{d}_{\mathcal{I},k,0})}{\pi_{k,0}}\left(\widehat{V}_0^{(k+1)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k+1)}(d^{opt})\right)\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k)}, \bar{\boldsymbol{X}}_0^{(k)}\right\}\Big|}_{\eta_{13}} \\ + \ & \underbrace{\mathrm{E}\left|h_k\{(\bar{\boldsymbol{A}}_0^{(k-1)}, \widehat{d}_{\mathcal{I},k,0}), \bar{\boldsymbol{X}}_0^{(k)}, \} - h_k\{(\bar{\boldsymbol{A}}_0^{(k-1)}, d_{k,0}^{opt}), \bar{\boldsymbol{X}}_0^{(k)}\}\right|}_{\eta_{14}}. \end{aligned}$$

By Condition (C3) and (101), we have

$$\begin{aligned} \eta_{13} \ \le \ & \mathrm{E}\left|\mathrm{E}\left\{\frac{\mathrm{g}(A_0^{(k)}, \widehat{d}_{\mathcal{I},k,0})}{\pi_{k,0}}\left(\widehat{V}_0^{(k+1)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k+1)}(d^{opt})\right)\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k+1)}, \bar{\boldsymbol{X}}_0^{(k+1)}\right\}\Big| \\ = \ & \mathrm{E}\frac{\mathrm{g}(A_0^{(k)}, \widehat{d}_{\mathcal{I},k,0})}{\pi_{k,0}}\left|\mathrm{E}\left\{\left(\widehat{V}_0^{(k+1)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k+1)}(d^{opt})\right)\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k+1)}, \bar{\boldsymbol{X}}_0^{(k+1)}\right\}\Big| \\ \le \ & \frac{1}{c_0}\mathrm{E}\left|\mathrm{E}\left\{\left(\widehat{V}_0^{(k+1)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k+1)}(d^{opt})\right)\right| \widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k+1)}, \bar{\boldsymbol{X}}_0^{(k+1)}\right\}\Big| = o(n^{-1/2}). \end{aligned}$$

Under Condition (C4) and (C5), using similar arguments in bounding $\eta_8^{(1)}$ in the proof of Theorem 2, we can show

$$\eta_{14} = o(n^{-1/2}). \tag{102}$$

Thus, we've shown

$$\mathrm{E}\left|\mathrm{E}\left(\widehat{V}_0^{(k)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k)}(d^{opt})\middle|\widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k)}, \bar{\boldsymbol{X}}_0^{(k)}\right)\right| = o(n^{-1/2}).$$

Since K is a fixed constant, we have

$$\max_{k=2,\dots,K}\mathrm{E}\left|\mathrm{E}\left(\widehat{V}_0^{(k)}(\widehat{d}_{\mathcal{I}}) - \widehat{V}_0^{(k)}(d^{opt})\middle|\widehat{d}_{\mathcal{I}}, \bar{\boldsymbol{A}}_0^{(k)}, \bar{\boldsymbol{X}}_0^{(k)}\right)\right| = o(n^{-1/2}). \tag{103}$$

Let $\mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}}$ denote the conditional expectation given $\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}$, we have

$$\begin{aligned}
& V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l+1)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l+1)}_{\mathcal{I}_*^{(1)}}) - V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l)}_{\mathcal{I}_*^{(1)}}) \\
= \; & \mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}}\prod_{j=1}^{l}\frac{\mathrm{g}(A_0^{(j)},\widehat{d}_{\mathcal{I}_*,j,0})}{\pi_{j,0}}\left(\widehat{V}_0^{(l+1)}(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l+1)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l+1)}_{\mathcal{I}_*^{(1)}}) - \widehat{V}_0^{(l+1)}(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l)}_{\mathcal{I}_*^{(1)}})\right) \\
= \; & \mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}}\prod_{j=1}^{l}\frac{\mathrm{g}(A_0^{(j)},\widehat{d}_{\mathcal{I}_*,j,0})}{\pi_{j,0}}\left(\frac{\mathrm{g}(A_0^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,0})}{\widehat{\pi}_{\mathcal{I}_*^{(1)},l+1,0}}\{\widehat{V}_0^{(l+2)}(\widehat{d}_{\mathcal{I}_*}) - \widehat{h}_{\mathcal{I}_*^{(1)},l+1,0}\}\right. \\
+ \; & \widehat{h}_{\mathcal{I}_*^{(1)},l+1}\{(\bar{\boldsymbol{A}}_0^{(l)},\widehat{d}_{\mathcal{I}_*,l+1,0}),\bar{\boldsymbol{X}}_0^{(l+1)}\} - \frac{\mathrm{g}(A_0^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,0})}{\pi_{l+1,0}}\{\widehat{V}_0^{(l+2)}(\widehat{d}_{\mathcal{I}_*}) - h_{l+1,0}\} \\
- \; & \left. h_{l+1}\{(\bar{\boldsymbol{A}}_0^{(l)},\widehat{d}_{\mathcal{I}_*,l+1,0}),\bar{\boldsymbol{X}}_0^{(l+1)}\}\right).
\end{aligned}$$

By Condition (C3), (37) and (103), we have

$$\begin{aligned}
& \mathrm{E}\left|\mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}}\prod_{j=1}^{l}\frac{\mathrm{g}(A_0^{(j)},\widehat{d}_{\mathcal{I}_*,j,0})}{\pi_{j,0}}\frac{\mathrm{g}(A_0^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,0})}{\widehat{\pi}_{\mathcal{I}_*^{(1)},l+1,0}}\{\widehat{V}_0^{(l+2)}(\widehat{d}_{\mathcal{I}_*}) - \widehat{V}_0^{(l+2)}(d^{opt})\}\right| \\
\le \; & \mathrm{E}\left|\mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}},\bar{\boldsymbol{A}}_0^{(l+2)},\bar{\boldsymbol{X}}_0^{(l+2)}}\prod_{j=1}^{l}\frac{\mathrm{g}(A_0^{(j)},\widehat{d}_{\mathcal{I}_*,j,0})}{\pi_{j,0}}\frac{\mathrm{g}(A_0^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,0})}{\widehat{\pi}_{\mathcal{I}_*^{(1)},l+1,0}}\right. \\
\times \; & \left.\{\widehat{V}_0^{(l+2)}(\widehat{d}_{\mathcal{I}_*}) - \widehat{V}_0^{(l+2)}(d^{opt})\}\right| \le \frac{1}{c^* c_0^l}\mathrm{E}\left|\mathrm{E}^{\mathcal{I}_*,\widehat{d}_{\mathcal{I}_*},\bar{\boldsymbol{A}}_0^{(l+2)},\bar{\boldsymbol{X}}_0^{(l+2)}}\{\widehat{V}_0^{(l+2)}(\widehat{d}_{\mathcal{I}_*}) - \widehat{V}_0^{(l+2)}(d^{opt})\}\right| \\
= \; & o(n^{-1/2}),
\end{aligned} \tag{104}$$

where $\mathrm{E}^{\mathcal{I}_*,\widehat{d}_{\mathcal{I}_*},\bar{\boldsymbol{A}}_0^{(l+2)},\bar{\boldsymbol{X}}_0^{(l+2)}}$ denotes the expectation conditional on $\mathcal{I}_*,\widehat{d}_{\mathcal{I}_*},\bar{\boldsymbol{A}}_0^{(l+2)},\bar{\boldsymbol{X}}_0^{(l+2)}$.

Similarly, we can show

$$\mathrm{E}\left|\mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}}\prod_{j=1}^{l+1}\frac{\mathrm{g}(A_0^{(j)},\widehat{d}_{\mathcal{I}_*,j,0})}{\pi_{j,0}}\{\widehat{V}_0^{(l+2)}(\widehat{d}_{\mathcal{I}_*}) - \widehat{V}_0^{(l+2)}(d^{opt})\}\right| = o(n^{-1/2}). \tag{105}$$

Besides, using similar arguments in bounding η_{10} in the proof of Theorem 3, we have

$$\begin{aligned}
&\mathrm{E}\Bigg|\mathrm{E}^{\mathcal{I}_*^{(1)},\mathcal{I}_*,\{O_i\}_{i\in\mathcal{I}_*^{(1)}}}\prod_{j=1}^{l}\frac{\mathrm{g}(A_0^{(j)},\widehat{d}_{\mathcal{I}_*,j,0})}{\pi_{j,0}}\Bigg(\frac{\mathrm{g}(A_0^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,0})}{\widehat{\pi}_{\mathcal{I}_*^{(1)},l+1,0}}\{\widehat{V}_0^{(l+2)}(d^{opt})-\widehat{h}_{\mathcal{I}_*^{(1)},l+1,0}\}\\
+\ &\widehat{h}_{\mathcal{I}_*^{(1)},l+1}\{(\bar{\boldsymbol{A}}_0^{(l)},\widehat{d}_{\mathcal{I}_*,l+1,0}),\bar{\boldsymbol{X}}_0^{(l+1)}\}-\frac{\mathrm{g}(A_0^{(l+1)},\widehat{d}_{\mathcal{I}_*,l+1,0})}{\pi_{l+1,0}}\{\widehat{V}_0^{(l+2)}(d^{opt})-h_{l+1,0}\}\\
-\ &h_{l+1}\{(\bar{\boldsymbol{A}}_0^{(l)},\widehat{d}_{\mathcal{I}_*,l+1,0}),\bar{\boldsymbol{X}}_0^{(l+1)}\}\Bigg)\Bigg|=o(n^{-1/2}).
\end{aligned}$$

Combining this together with (104) and (105) yields

$$\mathrm{E}\left|V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l+1)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l+1)}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l)}_{\mathcal{I}_*^{(1)}})\right|=o(n^{-1/2}),$$

for all $l=0,\dots,K-1$. Since K is a fixed integer, we have

$$\max_{l=0,\dots,K-1}\mathrm{E}\left|V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l+1)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l+1)}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l)}_{\mathcal{I}_*^{(1)}})\right|=o(n^{-1/2}).$$

By triangle inequality, we obtain

$$\begin{aligned}
\eta_8 &\le \sum_{l=0,\dots,K-1}\mathrm{E}\left|V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l+1)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l+1)}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l)}_{\mathcal{I}_*^{(1)}})\right|\\
&\le\ K\max_{l=0,\dots,K-1}\mathrm{E}\left|V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l+1)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l+1)}_{\mathcal{I}_*^{(1)}})-V(\widehat{d}_{\mathcal{I}_*};\widehat{\pi}^{(l)}_{\mathcal{I}_*^{(1)}},\widehat{h}^{(l)}_{\mathcal{I}_*^{(1)}})\right|=o(n^{-1/2}).
\end{aligned}$$

This implies $\widehat{V}_B=\widehat{V}_B^*+o_p(n^{-1/2})$.

Step 2: The assertion $\widehat{V}_B^*=\widehat{V}_\infty^*+o_p(n^{-1/2})$ can be proven using similar arguments in the second step of the proof of Theorem 3. We omit the details for brevity.

Step 3: For any $i\in\mathcal{I}_0$, $\mathcal{I}\subseteq\mathcal{I}_{(-i)}$ with $|\mathcal{I}|=s_n$, define

$$Q_i=\mathrm{E}\left(\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}})|O_i\right).$$

Notice that $Q_1,\dots,Q_n$ are i.i.d random variables with $\mathrm{Var}(Q_i)=\sigma_{s_n}^2$. We first show

$$\widehat{V}_\infty^*=\frac{1}{n}\sum_{i=1}^{n}Q_i+o_p(n^{-1/2}).\tag{106}$$

Recall that

$$\widehat{V}_i^{(K)}(\widehat{d}_{\mathcal{I}})=\frac{\mathrm{g}(A_i^{(K)},\widehat{d}_{\mathcal{I},K,i})}{\pi_{K,i}}(Y_i-h_{K,i})+h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)},\widehat{d}_{\mathcal{I},K,i}),\bar{\boldsymbol{X}}_i^{(K)}\},\tag{107}$$

for any $i\in\mathcal{I}_0$, $\mathcal{I}\subseteq\mathcal{I}_{(-i)}$ with $|\mathcal{I}|=s_n$. Let

$$T_i^{(K)}(\mathcal{I})=(Y_i-h_{K,i})\prod_{k=1}^{K}\frac{\mathrm{g}(A_i^{(k)},\widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}}\ \ \text{and}\ \ T_i^{(K)}=\mathrm{E}\{T_i^{(K)}(\mathcal{I})|O_i\}.$$

Using similar arguments in bounding $|\eta_7|$ in the proof of Theorem 2, we can show that

$$\frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \left(T_i^{(K)}(\mathcal{I}) - T_i^{(K)}\right) = o_p(n^{-1/2}). \tag{108}$$

Besides, by Condition (C3) and (100), we have

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} [h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, \widehat{d}_{\mathcal{I},K,i}), \bar{\boldsymbol{X}}_i^{(K)}\} - h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, d_{K,i}^{opt}), \bar{\boldsymbol{X}}_i^{(K)}\}]\right| \\
&\qquad \le \frac{1}{c_0^{K-1}} \max_{\substack{\mathcal{I}\subseteq\mathcal{I}_0\\ |\mathcal{I}|=s_n}} \mathrm{E}\left| h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, \widehat{d}_{\mathcal{I},K,i}), \bar{\boldsymbol{X}}_i^{(K)}\} - h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, d_{K,i}^{opt}), \bar{\boldsymbol{X}}_i^{(K)}\}\right| = o(n^{-1/2}),
\end{aligned}$$

for any $d^{opt} \in \mathcal{D}^{opt}$. By Markov's inequality, we obtain

$$\begin{aligned}
& \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, \widehat{d}_{\mathcal{I},K,i}), \bar{\boldsymbol{X}}_i^{(K)}\} \\
= & \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, d_{K,i}^{opt}), \bar{\boldsymbol{X}}_i^{(K)}\} + o_p(n^{-1/2}).
\end{aligned}$$

Combining this together with (107) and (108) yields

$$\begin{aligned}
& \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} \widehat{V}_i^{(K)}(\widehat{d}_{\mathcal{I}}) \qquad (109) \\
= & \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \left(T_i^{(K)}(\mathcal{I}) + \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, d_{K,i}^{opt}), \bar{\boldsymbol{X}}_i^{(K)}\} \right) \\
+ & o_p(n^{-1/2}) = \frac{1}{n} \sum_{i=1}^{n} T_i^{(K)} + o_p(n^{-1/2}) \\
+ & \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}} \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} h_K\{(\bar{\boldsymbol{A}}_i^{(K-1)}, d_{K,i}^{opt}), \bar{\boldsymbol{X}}_i^{(K)}\}.
\end{aligned}$$

Let $\varepsilon_i^{(j)} = h_j\{(\bar{\boldsymbol{A}}_i^{(j-1)}, d_{j,i}^{opt}), \bar{\boldsymbol{X}}_i^{(j)}\} - h_{j-1,i}$. Similarly, we can show for all $j = 2, \dots, K-1$,

$$\frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}} \prod_{k=1}^{j-2} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} \left(\frac{\mathrm{g}(A_i^{(j-1)}, \widehat{d}_{\mathcal{I},j-1,i})}{\pi_{j-1,i}} \varepsilon_i^{(j)} + h_{j-1}\{(\bar{\boldsymbol{A}}_i^{(j-2)}, \widehat{d}_{\mathcal{I},j-1,i}), \bar{\boldsymbol{X}}_i^{(j-1)}\} \right)$$

$$= \frac{1}{n}\sum_{i=1}^n T_i^{(j)} + \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}} \prod_{k=1}^{j-2} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} h_{j-1}\{(\bar{\boldsymbol{A}}_i^{(j-2)}, d_{j-1,i}^{opt}), \bar{\boldsymbol{X}}_i^{(j-1)}\} + o_p(n^{-1/2}),$$

where

$$T_i^{(j)} = \mathrm{E}\left(\prod_{k=1}^{j-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} \varepsilon_i^{(j)} \middle| O_i \right),$$

for any $\mathcal{I} \subseteq \mathcal{I}_{(-i)}$ with $|\mathcal{I}| = s_n$. This together with (109) yields

$$\widehat{V}_\infty^* = \frac{1}{n\binom{n-1}{s_n}} \sum_{i\in\mathcal{I}_0} \sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\|\mathcal{I}|=s_n}} \left\{ \prod_{k=1}^{K-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k_i})}{\pi_{k,i}} \widehat{V}_i^{(K)}(\widehat{d}_\mathcal{I}) \right. \tag{110}$$
$$\left. - \sum_{j=1}^{K-1} \prod_{k=1}^{j-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k_i})}{\pi_{k,i}} \left(\frac{\mathrm{g}(A_i^{(j)}, \widehat{d}_{\mathcal{I},j,i})}{\pi_{j,i}} h_{j,i} - h_{j-1}\{(\bar{\boldsymbol{A}}_i^{(j-2)}, \widehat{d}_{\mathcal{I},j-1,i}), \bar{\boldsymbol{X}}_i^{(j-1)}\} \right) \right\}$$
$$= \frac{1}{n}\sum_{k=2}^K T_i^{(k)} + \frac{1}{n}\sum_{i=1}^n h_1(d_{1,i}^{opt}, \boldsymbol{X}_i^{(1)}) + o_p(n^{-1/2}).$$

Define

$$\bar{T}_i^{(j)} = \mathrm{E}\left(\prod_{k=1}^{j-1} \frac{\mathrm{g}(A_i^{(k)}, \widehat{d}_{\mathcal{I},k,i})}{\pi_{k,i}} \{h_j\{(\bar{\boldsymbol{A}}_i^{(j-1)}, \widehat{d}_{\mathcal{I},j,i}), \bar{\boldsymbol{X}}_i^{(j)}\} - h_{j-1,i}\} \middle| O_i \right).$$

By Condition (C3) and (102), we have

$$\mathrm{E}\left| T_i^{(j)} - \bar{T}_i^{(j)} \right| \le \frac{1}{c_0^{j-1}} \mathrm{E}\left| h_j\{(\bar{\boldsymbol{A}}_i^{(j-1)}, \widehat{d}_{\mathcal{I},j,i}), \bar{\boldsymbol{X}}_i^{(j)}\} - h_j\{(\bar{\boldsymbol{A}}_i^{(j-1)}, d_{j,i}^{opt}), \bar{\boldsymbol{X}}_i^{(j)}\} \right| = o(n^{-\frac{1}{2}}) \tag{111}$$

In addition, let

$$\bar{T}_i^{(1)} = \mathrm{E}\left(h_1(\widehat{d}_{\mathcal{I},1,i}, \boldsymbol{X}_i^{(1)}) \middle| O_i \right),$$

for any $\mathcal{I} \subseteq \mathcal{I}_{(-i)}$ with $|\mathcal{I}| = s_n$. Similar to the proof of Theorem (2), we can show

$$\max_{i\in\mathcal{I}_0} \mathrm{E}\left| h_1(d_{1,i}^{opt}, \bar{\boldsymbol{X}}_i) - \bar{T}_i^{(1)} \right| = o(n^{-1/2}). \tag{112}$$

Combining this together with (111) and (110), we have

$$\widehat{V}_{\infty}^{*} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K}\bar{T}_i^{(j)} + o_p(n^{-1/2}).$$

Notice that $Q_i = \sum_{j=1}^{K}\bar{T}_i^{(j)}, \forall i \in \mathcal{I}_0$. Thus, we've shown (106).

Moreover, it follows from (111) and (112) that

$$\max_{i\in\mathcal{I}_0} \mathrm{E}\left| Q_i - \sum_{j=2}^{K} T_i^{(j)} - h_1(d_{1,i}^{opt}, \boldsymbol{X}_i^{(1)}) \right| = o(n^{-1/2}).$$

Therefore, we have

$$\mathrm{E}Q_i = \mathrm{E}\left(\sum_{j=2}^{K} T_i^{(j)} + h_1(d_{1,i}^{opt}, \boldsymbol{X}_i^{(1)})\right) + o(n^{-1/2}) = \mathrm{E}h_1(d_{1,i}^{opt}, \boldsymbol{X}_i^{(1)}) + o(n^{-1/2}). \quad (113)$$

Notice that $\mathrm{E}h_1(d_{1,i}^{opt}, \boldsymbol{X}_i^{(1)}) = V_0$. Under the condition that $\liminf_n \sigma_n > 0$, it follows from (106) and (113) that

$$\frac{\sqrt{n}}{\sigma_{s_n}}(\widehat{V}_{\infty}^{*} - V_0) \xrightarrow{d} N(0,1).$$

Step 4: For $i = 0, 1, \ldots, n$, define

$$\psi^{(i)} = \frac{1}{n^{(i)}}\sum_{j=1}^{2}\sum_{b=1}^{B}\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}_b}; \widehat{\pi}_{\mathcal{I}_b^{(j)}}, \widehat{h}_{\mathcal{I}_b^{(j)}})\mathbb{I}(i \notin \mathcal{I}_b^{(j)}).$$

Using similar arguments in Step 3 of the proof of Theorem 3, we can show

$$\frac{1}{n-1}\sum_{i=1}^{n}\{\psi^{(i)}\}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\{\psi_{\infty}^{*(i)}\}^2 + o_p(1),$$

and

$$\left(\frac{1}{n}\sum_{i=1}^{n}\psi^{(i)}\right)^2 = \left(\frac{1}{n}\sum_{i=1}^{n}\psi_{\infty}^{*(i)}\right)^2 + o_p(1),$$

where

$$\psi_{\infty}^{*(i)} = \frac{1}{\binom{n-1}{s_n}}\sum_{\substack{\mathcal{I}\subseteq\mathcal{I}_{(-i)}\\ |\mathcal{I}|=s_n}}\widehat{V}_i^{(1)}(\widehat{d}_{\mathcal{I}}).$$

This implies that

$$\widehat{\sigma}_B^2 = \frac{1}{n-1}\sum_{i=1}^{n}\{\psi_{\infty}^{*(i)}\}^2 - \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}\psi_{\infty}^{*(i)}\right)^2 + o_p(1). \quad (114)$$

Since $s_n = o(n)$, by the ANOVA decomposition (Efron and Stein, 1981), we have

$$\max_{i \in \mathcal{I}_0} \mathrm{E} \left| \psi_{\infty}^{*(i)} - Q_i \right|^2 = o(1).$$

In addition, under the condition that $\max_{\bar{\boldsymbol{a}}_K \in \{0,1\}^K} \mathrm{E}\{Y_0^*(\bar{\boldsymbol{a}}_K)\}^2 < +\infty$, we can show

$$\max_{i \in \mathcal{I}_0} \mathrm{E} \left| \psi_{\infty}^{*(i)} + Q_i \right|^2 = O(1).$$

By Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} & \mathrm{E} \left| \frac{1}{n-1} \sum_{i=1}^{n} \{\psi_{\infty}^{*(i)}\}^2 - \frac{1}{n-1} \sum_{i=1}^{n} Q_i^2 \right| \le \max_{i \in \mathcal{I}_0} \mathrm{E} \left| \{\psi_{\infty}^{*(i)}\}^2 - Q_i^2 \right| \\ \le & \sqrt{\max_{i \in \mathcal{I}_0} \mathrm{E} \left| \psi_{\infty}^{*(i)} - Q_i \right|^2 \max_{i \in \mathcal{I}_0} \mathrm{E} \left| \psi_{\infty}^{*(i)} + Q_i \right|^2} = o(1). \end{aligned}$$

It follows from Markov's inequality that

$$\frac{1}{n-1} \sum_{i=1}^{n} \{\psi_{\infty}^{*(i)}\}^2 - \frac{1}{n-1} \sum_{i=1}^{n} Q_i^2 = o_p(1).$$

Similarly, we can show

$$\left(\frac{1}{n} \sum_{i=1}^{n} \psi_{\infty}^{*(i)} \right)^2 = \left(\frac{1}{n} \sum_{i=1}^{n} Q_i \right)^2 + o_p(1).$$

In view of (114), we've shown

$$\widehat{\sigma}_B^2 = \frac{1}{n-1} \sum_{i=1}^{n} Q_i^2 - \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^{n} Q_i \right)^2 + o_p(1).$$

In addition, it follows from the law of large numbers that

$$\frac{1}{n-1} \sum_{i=1}^{n} Q_i^2 - \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^{n} Q_i \right)^2 = \sigma_{s_n}^2 + o_p(1).$$

Thus, we have $\widehat{\sigma}_B^2 = \sigma_{s_n}^2 + o_p(1)$. The proof is hence completed. ■

C.7. Proof of Lemma 8

Let $p_0 = \Pr(A_0 = 1)$. By Condition (A3), we have

$$0 < c_0 \le p_0 \le 1 - c_0 < 1. \tag{115}$$

Consider the event

$$\mathcal{A}_* = \{c_0 n/2 < n_A < (1 - c_0/2) n\},$$

where $n_A = \sum_{i=1}^n A_i$. It follows from Hoeffding's inequality (Hoeffding, 1963) that

$$\Pr(\mathcal{A}_*^c) \le \Pr\left(|n_A - np_0| \le c_0 n/3\right) \le 2\exp\left(-\frac{18n}{c_0^2}\right) \to 0. \tag{116}$$

Note that the random variable $n_{\mathcal{S}}$ is completely determined by n_A. For $s_n < n_A < n - s_n$, we have

$$\frac{\binom{n}{s_n} - n_{\mathcal{S}}}{\binom{n}{s_n}} = \frac{\sum_{k=0}^{N_0-1}\binom{n_A}{s_n-k}\binom{n-n_A}{k} + \sum_{k=0}^{N_0-1}\binom{n_A}{k}\binom{n-n_A}{s_n-k}}{\binom{n}{s_n}}. \tag{117}$$

Let $m^{(s)} = m(m-1)\cdots(m-s+1)$ for any integers $m \ge s > 0$, we have for any $0 \le k \le N_0 - 1 \le s_n$,

$$\frac{\binom{n_A}{s_n-k}\binom{n-n_A}{k}}{\binom{n}{s_n}} = \binom{s_n}{k}\frac{n_A^{(s_n-k)}(n-n_A)^{(k)}}{n^{(s_n)}} \le \frac{n_A^{(s_n-k)}(n-n_A)^{(k)}}{n^{(s_n)}} \le \frac{n_A^{s_n-k}(n-n_A)^k}{(n-s_n+1)^{s_n}}.$$

Since $s_n = o(n)$, for sufficiently large n, we have $n - s_n + 1 \ge (1 - c_0/3)n$. Thus, under the event defined in $\mathcal{A}_*$, we have

$$\frac{\binom{n_A}{s_n-k}\binom{n-n_A}{k}}{\binom{n}{s_n}} \le \left(\frac{1-c_0/2}{1-c_0/3}\right)^{s_n}, \qquad \forall 0 \le k \le N_0 - 1 \le s_n.$$

Similarly, we can show

$$\frac{\binom{n_A}{k}\binom{n-n_A}{s_n-k}}{\binom{n}{s_n}} \le \left(\frac{1-c_0/2}{1-c_0/3}\right)^{s_n}, \qquad \forall 0 \le k \le N_0 - 1 \le s_n,$$

under the event defined in $\mathcal{A}_*$.

By (117), we obtain

$$\frac{\binom{n}{s_n} - n_{\mathcal{S}}}{\binom{n}{s_n}} \le 2N_0 p_*^{s_n},$$

under the event defined in $\mathcal{A}_*$, where $p_* = (1 - c_0/2)/(1 - c_0/3)$. Notice that N_0 is a fixed constant. Under the given conditions, we have $s_n \gg n^{1/(2\kappa^*)}$. Set $c_3 = 18c_0^{-2}$ and $\beta_0 = 1/(2\kappa^*)$, it follows from (116) that

$$\Pr\left(\frac{\binom{n}{s_n} - n_{\mathcal{S}}}{\binom{n}{s_n}} \le c_1 p_*^{c_2 n^{\beta_0}}\right) \ge \Pr(\mathcal{A}_*) \ge 1 - 2\exp(-c_3 n) \to 1,$$

for some constants $c_1, c_2 > 0$. This completes the proof of (48).

For any $i \in \{1, \ldots, n\}$, define $\mathcal{S}_{N_0,s_n}^{(i)} = \{\mathcal{I} \in \mathcal{S}_{N_0,s_n} : i \notin \mathcal{I}\}$ and $n_{\mathcal{S}}^{(i)} = |\mathcal{S}_{N_0,s_n}^{(i)}|$. Similar to (48), there exist some constants $c_1^*, c_2^*, c_3^* > 0$ and $0 < p_{**} < 1$ such that

$$\Pr\left(\frac{\binom{n-1}{s_n} - n_{\mathcal{S}}^{(i)}}{\binom{n-1}{s_n}} \le c_1^* p_{**}^{c_2^* n^{\beta_0}}\right) \ge 1 - 2\exp(-c_3^* n). \tag{118}$$

Let $\mathcal{A}^{(i)}$ be the event defined in Equation 118. Set $c_4 = \min(c_3^*, c_3)$, it follows from Bonferroni's inequality that

$$\Pr(\mathcal{A}^{(i)} \cap \mathcal{A}_*) \geq 1 - \Pr(\mathcal{A}^{(i)}) - \Pr(\mathcal{A}_*) \geq 1 - 4\exp(-c_4 n).$$

Under the events defined in $\mathcal{A}^{(i)}$ and $\mathcal{A}_*$, we have

$$\begin{aligned}
\left| \frac{n_{\mathcal{S}}^{(i)}}{n_{\mathcal{S}}} - \frac{\binom{n-1}{s_n}}{\binom{n}{s_n}} \right| &= \frac{\binom{n-1}{s_n}}{\binom{n}{s_n}} \left| \frac{n_{\mathcal{S}}^{(i)}/\binom{n-1}{s_n} - n_{\mathcal{S}}/\binom{n}{s_n}}{n_{\mathcal{S}}/\binom{n}{s_n}} \right| = \frac{n-s_n}{n} \left| \frac{n_{\mathcal{S}}^{(i)}/\binom{n-1}{s_n} - n_{\mathcal{S}}/\binom{n}{s_n}}{n_{\mathcal{S}}/\binom{n}{s_n}} \right| \\
&\leq \frac{2(n-s_n)}{n} \left| \frac{n_{\mathcal{S}}^{(i)}}{\binom{n-1}{s_n}} - \frac{n_{\mathcal{S}}}{\binom{n}{s_n}} \right| \leq \frac{2(n-s_n)}{n} \left(\left| \frac{n_{\mathcal{S}}^{(i)}}{\binom{n-1}{s_n}} - 1 \right| + \left| \frac{n_{\mathcal{S}}}{\binom{n}{s_n}} - 1 \right| \right) \\
&\leq 2(c_1 p_*^{c_2 n^{\beta_0}} + c_1^* p_{**}^{c_2^* n^{\beta_0}}) \ll \frac{\sqrt{\log n}}{\sqrt{n}},
\end{aligned}$$

where the first inequality is due to (56), the third inequality follows by the definitions of $\mathcal{A}^{(i)}$ and $\mathcal{A}_*$.

Conditional on $\{O_i\}_{i\in\mathcal{I}_0}$, the random variables $\mathbb{I}(i \notin \mathcal{I}_1), \ldots, \mathbb{I}(i \notin \mathcal{I}_B)$ are independent Bernoulli random variables with mean $\Pr(i \notin \mathcal{I}_b | \{O_i\}_{i\in\mathcal{I}_0}) = n_{\mathcal{S}}^{(i)}/n_{\mathcal{S}}$. Hence, it follows from Hoeffding's inequality that

$$\Pr\left(\left| \frac{n^{(i)}}{B} - \frac{n_{\mathcal{S}}^{(i)}}{n_{\mathcal{S}}} \right| \leq \frac{\sqrt{\log B}}{\sqrt{2B}} \middle| \{O_i\}_{i\in\mathcal{I}_0} \right) \geq 1 - 2\exp\left(-2\log B/2\right) = 1 - \frac{2}{B}.$$

Therefore, we have

$$\Pr\left(\left| \frac{n^{(i)}}{B} - \frac{n_{\mathcal{S}}^{(i)}}{n_{\mathcal{S}}} \right| \leq \frac{\sqrt{\log B}}{\sqrt{2B}} \right) \geq 1 - \frac{2}{B}. \tag{119}$$

Since $B \gg n$, we have $\sqrt{\log B}/\sqrt{2B} \leq \sqrt{\log n}/\sqrt{2n}$. Under the events defined (119), $\mathcal{A}^{(i)}$ and $\mathcal{A}_*$, we have

$$\left| \frac{n^{(i)}}{B} - \frac{n-s_n}{n} \right| = \left| \frac{n^{(i)}}{B} - \frac{\binom{n-1}{s_n}}{\binom{n}{s_n}} \right| \leq \left| \frac{n^{(i)}}{B} - \frac{n_{\mathcal{S}}^{(i)}}{n_{\mathcal{S}}} \right| + \left| \frac{n_{\mathcal{S}}^{(i)}}{n_{\mathcal{S}}} - \frac{\binom{n-1}{s_n}}{\binom{n}{s_n}} \right| \leq \frac{\sqrt{\log n}}{\sqrt{n}}.$$

The proof is hence completed by noting that

$$\begin{aligned}
\Pr\left(\left| \frac{n^{(i)}}{B} - \frac{n-s_n}{n} \right| > \frac{\sqrt{\log n}}{\sqrt{n}} \right) &\leq \Pr\left(\left| \frac{n^{(i)}}{B} - \frac{n_{\mathcal{S}}^{(i)}}{n_{\mathcal{S}}} \right| \leq \frac{\sqrt{\log B}}{\sqrt{2B}} \right) + \Pr\left(\mathcal{A}_*^c \cup (\mathcal{A}^{(i)})^c \right) \\
&\leq \frac{2}{B} + 4\exp(-c_4 n).
\end{aligned}$$

■

Appendix D. Additional Details Regarding (25)

In this section, we show the approximation in Equation 25 holds if kernel smoothers is used to estimate the contrast function. For simplicity, we assume all covariates are continuous on $\prod_{j=1}^{p}[a_j, b_j]$ for some $0 < a_j < b_j < +\infty$, with a strictly positive density function $f_X(\cdot)$. More generally, we can show (25) holds when at least one of the covariates is continuous.

Although $\widehat{d}_{\mathcal{I}_{(j)}}$ might not converge to a deterministic function, we will show that $\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)$ converges as $j \to \infty$. Under the given conditions, we can show there exists some constant $\bar{C} \geq 1$ such that the following holds with probability 1,

$$\bar{C}^{-1} \leq \widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h) \leq \bar{C}, \quad \forall j \geq 1. \tag{120}$$

For any two positive sequences $\{y_i\}_i, \{z_i\}_i$, it follows from a first-order Taylor expansion that

$$\frac{m}{\sum_{i=1}^{m} y_i^{-1/2}} = \frac{m}{\sum_{i=1}^{m} z_i^{-1/2}} + \frac{m \sum_{i=1}^{m} (z_i^*)^{-3/2}(y_i - z_i)}{\{\sum_{i=1}^{m} (z_i^*)^{-1/2}\}^2}.$$

Set $y_i = \widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(i+\ell_n-1)}}; \pi, h)$ and $z_i = \mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(i+\ell_n-1)}}; \pi, h)$, it follows from (120) that

$$\left| \frac{n-\ell_n}{\sum_{j=\ell_n}^{n-1} \widetilde{\sigma}_0^{-1}(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)} - \frac{n-\ell_n}{\sum_{j=\ell_n}^{n-1} \{\mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)\}^{-1/2}} \right| \leq \frac{\bar{C}^{2/5}}{n-\ell_n} \sum_{j=\ell_n}^{n-1} |\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h) - \mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)|.$$

Thus, the approximation in Equation 25 holds as long as

$$\mathrm{E}|\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h) - \mathrm{E}\widetilde{\sigma}_0^2(\widehat{d}_{\mathcal{I}_{(j)}}; \pi, h)| \to 0.$$

By (89), it suffices to show $\mathrm{E}|\kappa_j - \mathrm{E}\kappa_j| \to 0$, or $\mathrm{Var}(\kappa_j) \to 0$, as $j \to \infty$, where κ_j is defined in the proof of Theorem 5.

With some calculations, we have

$$\begin{aligned}\mathrm{Var}(\kappa_j) = \int_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}} & \frac{\{\pi(0, \boldsymbol{x}_1) - \pi(1, \boldsymbol{x}_1)\}\{\pi(0, \boldsymbol{x}_2) - \pi(1, \boldsymbol{x}_2)\}}{\pi(0, \boldsymbol{x}_1)\pi(1, \boldsymbol{x}_1)\pi(0, \boldsymbol{x}_2)\pi(1, \boldsymbol{x}_2)} \\ & \mathrm{cov}(\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1), \widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)) dF_X(\boldsymbol{x}_1) dF_X(\boldsymbol{x}_2).\end{aligned}$$

It follows from Condition (C3) that

$$\mathrm{Var}(\kappa_j) \leq \frac{1}{c_0^4} \int_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}} \mathrm{cov}(\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1), \widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)) dF_X(\boldsymbol{x}_1) dF_X(\boldsymbol{x}_2).$$

Under the conditions in (A7), the estimator $\{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}) - \tau(\boldsymbol{x}_0)\}/\sigma_j^*(\boldsymbol{x})$ is asymptotically normal for each $\boldsymbol{x}$. Suppose for now, there exists a non-increasing sequence $\{h_j\}_j$ that satisfies $h_j > 0$, $h_j \to 0$ as $j \to \infty$ such that

$$\begin{aligned} & \forall \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty > h_j, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}, \\ \left(\frac{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1) - \mathrm{E}\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1)}{\sigma_j^*(\boldsymbol{x}_1)}, \right. & \left. \frac{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2) - \mathrm{E}\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)}{\sigma_j^*(\boldsymbol{x}_2)} \right) \stackrel{d}{\to} N(0, \boldsymbol{I}_2),\end{aligned} \tag{121}$$

where $\boldsymbol{I}_2$ denotes a 2×2 identity matrix.

Condition (121) essentially requires $\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1)$ and $\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)$ to be asymptotically independent for any $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty > h_j$. As we will see below, h_j can be set as the bandwidth parameter when kernel smoothers are used to estimate the contrast.

Notice that

$$c_0^4 \mathrm{Var}(\kappa_j) \le \underbrace{\int_{\substack{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X} \\ \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty \le h_j}} \mathrm{cov}(\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1), \widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)) dF_X(\boldsymbol{x}_1) dF_X(\boldsymbol{x}_2)}_{\zeta_5}$$
$$+ \underbrace{\int_{\substack{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X} \\ \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty > h_j}} \mathrm{cov}(\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1), \widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)) dF_X(\boldsymbol{x}_1) dF_X(\boldsymbol{x}_2)}_{\zeta_6}.$$

Suppose $f_X(\cdot)$ is uniformly bounded. Since $|\mathrm{cov}(\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1), \widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2))| \le 1$, we have

$$\zeta_5 \le \int_{\substack{\boldsymbol{x}_1 \in \mathbb{X} \\ \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty \le h_j}} dF_X(\boldsymbol{x}_1) f_X(\boldsymbol{x}_2) d\boldsymbol{x}_2 = O(h_j^p) = o(1). \tag{122}$$

In addition, it follows from (121) that

$$\left|\mathrm{cov}(\widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1), \widehat{d}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2))\right| = \left|\Pr\left(\frac{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1) - \mathrm{E}\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1)}{\sigma_j^*(\boldsymbol{x}_1)} > 0, \frac{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2) - \mathrm{E}\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)}{\sigma_j^*(\boldsymbol{x}_2)} > 0\right)\right.$$
$$\left. - \Pr\left(\frac{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1) - \mathrm{E}\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_1)}{\sigma_j^*(\boldsymbol{x}_1)} > 0\right) \Pr\left(\frac{\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2) - \mathrm{E}\widehat{\tau}_{\mathcal{I}_{(j)}}(\boldsymbol{x}_2)}{\sigma_j^*(\boldsymbol{x}_2)} > 0\right)\right| = o(1).$$

By the dominated convergence theorem, we obtains $\zeta_6 = o(1)$. This together with (122) yields that $\mathrm{Var}(\kappa_j) \to 0$, as $j \to \infty$.

In the following, we show (25) holds when kernel smoothers are used. Let $K(\cdot)$ be a p-dimensional multivariate kernel with bounded support $[-1/2, 1/2]^p$. Given a bandwidth parameter $h_j > 0$, consider the following nonparametric estimator for $m(\boldsymbol{x}) = \tau(\boldsymbol{x}) f(\boldsymbol{x})$:

$$\widehat{m}_j(\boldsymbol{x}) = \frac{1}{j h_j^p} \sum_{i=1}^{j} \left(\frac{A_i}{\pi(1, \boldsymbol{X}_i)} - \frac{1 - A_i}{\pi(0, \boldsymbol{X}_i)}\right) Y_i K\left(\frac{\boldsymbol{x} - \boldsymbol{X}_i}{h_j}\right). \tag{123}$$

It is consistent when the bandwidth satisfies $h_j \to 0$ and $j h_j^p \to \infty$. An augmented version of (123) can be similarly derived.

Since $\mathbb{I}\{m(\boldsymbol{x}) > 0\} = \mathbb{I}\{\tau(\boldsymbol{x}) > 0\}$, we may set $\widehat{d}_{\mathcal{I}_j}(\boldsymbol{x}) = \mathbb{I}\{\widehat{m}_j(\boldsymbol{x}) > 0\}$. Let $\widehat{m}_j(\cdot) = \widehat{m}_{\mathcal{I}_{(j)}}(\cdot)$. For any $\boldsymbol{x}_1, \boldsymbol{x}_2$ with $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty > h_j$, we have

$$K\left(\frac{\boldsymbol{x}_1 - \boldsymbol{X}_i}{h_j}\right) K\left(\frac{\boldsymbol{x}_2 - \boldsymbol{X}_i}{h_j}\right) = 0.$$

As a result, for any $\boldsymbol{x}_1, \boldsymbol{x}_2$ that satisfy $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_\infty > h_j$ and $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}$,

$$\mathrm{cov}(\widehat{m}_j(\boldsymbol{x}_1), \widehat{m}_j(\boldsymbol{x}_2)) = \mathrm{E}\widehat{m}_j(\boldsymbol{x}_1) \mathrm{E}\widehat{m}_j(\boldsymbol{x}_2) = o(1),$$

where the last equality is due to that the bias $\mathrm{E}\widehat{m}_j(\boldsymbol{x}) - m_j(\boldsymbol{x}) = \mathrm{E}\widehat{m}_j(\boldsymbol{x})$ will converge to zero for any $\boldsymbol{x} \in \mathbb{X}$. Now we can similarly show $\mathrm{Var}(\kappa_j) \to 0$ as $j \to \infty$.

Setting (G)	$s_n = 3.5n/\log(n)$		$s_n = 4n/\log(n)$	
n	ECP(%)	AL*100	ECP(%)	AL*100
600	93.3 (0.8)	28.1 (0.17)	93.2 (8.0)	28.5 (0.18)
1200	92.8 (0.8)	18.6 (0.05)	92.3 (0.8)	18.8 (0.06)
Setting (H)	$s_n = 3.5n/\log(n)$		$s_n = 4n/\log(n)$	
n	ECP(%)	AL*100	ECP(%)	AL*100
600	92.5 (0.8)	37.4 (0.09)	92.8 (0.8)	37.6 (0.10)
1200	93.8 (0.8)	25.6 (0.04)	93.6 (0.9)	25.6 (0.04)
Setting (I)	$s_n = 3.5n/\log(n)$		$s_n = 4n/\log(n)$	
n	ECP(%)	AL*100	ECP(%)	AL*100
600	93.8 (0.8)	39.4 (0.14)	94.0 (0.8)	39.4 (0.15)
1200	94.4 (0.7)	26.5 (0.04)	94.2 (0.7)	26.5 (0.04)

Table 7: ECP and AL of the CIs with standard errors in parenthesis

Appendix E. Additional Tables

In this section, we attach a table that reports the performance of our method in Scenarios (G)-(I) where K_0 is set to 3.5 or 4.

References

Jean-Yves Audibert and Alexandre B. Tsybakov. Fast learning rates for plug-in classifiers. *Ann. Statist.*, 35(2):608–633, 2007. ISSN 0090-5364. doi: 10.1214/009053606000001217. URL http://dx.doi.org/10.1214/009053606000001217.

Heejung Bang and James M. Robins. Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4):962–972, 2005. ISSN 0006-341X. doi: 10.1111/j.1541-0420.2005.00377.x.

Gérard Biau. Analysis of a random forests model. *J. Mach. Learn. Res.*, 13:1063–1095, 2012. ISSN 1532-4435.

Peter Bühlmann and Bin Yu. Analyzing bagging. *Ann. Statist.*, 30(4):927–961, 2002. ISSN 0090-5364. URL https://doi.org/10.1214/aos/1031689014.

Bibhas Chakraborty, Susan Murphy, and Victor Strecher. Inference for non-regular parameters in optimal dynamic treatment regimes. *Stat. Methods Med. Res.*, 19(3):317–343, 2010. ISSN 0962-2802. doi: 10.1177/0962280209105013. URL http://dx.doi.org/10.1177/0962280209105013.

Bibhas Chakraborty, Eric B Laber, and Ying-Qi Zhao. Inference about the expected performance of a data-driven dynamic treatment regime. *Clinical Trials*, 11(4):408–417, 2014.

B. Efron and C. Stein. The jackknife estimate of variance. *Ann. Statist.*, 9(3):586–596, 1981. ISSN 0090-5364.

Bradley Efron. Estimation and accuracy after model selection. *J. Amer. Statist. Assoc.*, 109(507):991–1007, 2014. ISSN 0162-1459. doi: 10.1080/01621459.2013.823775. URL `https://doi.org/10.1080/01621459.2013.823775`.

Caiyun Fan, Wenbin Lu, Rui Song, and Yong Zhou. Concordance-assisted learning for estimating optimal individualized treatment regimes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 79(5):1565–1582, 2017. ISSN 1369-7412.

Mark Galassi, Jim Davies, James Theiler, Brian Gough, Gerard Jungman, Patrick Alken, Michael Booth, Fabrice Rossi, and Rhys Ulerich. *GNU Scientific Library Reference Manual (Version 2.1)*, 2015. URL `http://www.gnu.org/software/gsl/`.

Wolfgang Härdle. *Applied nonparametric regression*, volume 19 of *Econometric Society Monographs*. Cambridge University Press, Cambridge, 1990. ISBN 0-521-38248-3. doi: 10.1017/CCOL0521382483. URL `https://doi.org/10.1017/CCOL0521382483`.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963. ISSN 0162-1459. URL `http://links.jstor.org/sici?sici=0162-1459(196303)58:301<13:PIFSOB>2.0.CO;2-D&origin=MSN`.

Shuhan Liang, Wenbin Lu, Rui Song, and Lan Wang. Sparse concordance-assisted learning for optimal treatment decision. *J. Mach. Learn. Res.*, 18:1–26, 2018.

Alexander R. Luedtke and Mark J. van der Laan. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *Ann. Statist.*, 44(2):713–742, 2016. ISSN 0090-5364. doi: 10.1214/15-AOS1384. URL `http://dx.doi.org/10.1214/15-AOS1384`.

S. A. Murphy. Optimal dynamic treatment regimes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 65(2):331–366, 2003. ISSN 1369-7412. doi: 10.1111/1467-9868.00389. URL `http://dx.doi.org/10.1111/1467-9868.00389`.

S. A. Murphy. An experimental design for the development of adaptive treatment strategies. *Stat. Med.*, 24(10):1455–1481, 2005. ISSN 0277-6715. URL `https://doi.org/10.1002/sim.2022`.

Min Qian and Susan A. Murphy. Performance guarantees for individualized treatment rules. *Ann. Statist.*, 39(2):1180–1210, 2011. ISSN 0090-5364. doi: 10.1214/10-AOS864. URL `http://dx.doi.org/10.1214/10-AOS864`.

J.M. Robins, M.A. Hernan, and B. Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiol.*, 11:550–560, 2000.

Chengchun Shi, Ailin Fan, Rui Song, and Wenbin Lu. High-dimensional A-learning for optimal dynamic treatment regimes. *Ann. Statist.*, 46(3):925–957, 2018a. ISSN 0090-5364. doi: 10.1214/17-AOS1570.

Chengchun Shi, Rui Song, Wenbin Lu, and Bo Fu. Maximin projection learning for optimal treatment decision with heterogeneous individualized treatment effects. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 80(4):681–702, 2018b. ISSN 1369-7412. doi: 10.1111/rssb.12273.

Ingo Steinwart and Andreas Christmann. *Support vector machines*. Information Science and Statistics. Springer, New York, 2008. ISBN 978-0-387-77241-7.

Stefan Wager and Susan Athey. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *J. Amer. Statist. Assoc.*, 113(523):1228–1242, 2018. ISSN 0162-1459. doi: 10.1080/01621459.2017.1319839. URL `https://doi.org/10.1080/01621459.2017.1319839`.

Stefan Wager and Guenther Walther. Adaptive concentration of regression trees, with application to random forests. *arXiv preprint arXiv:1503.06388*, 2015.

Jingshen Wang, Xuming He, and Gongjun Xu. Debiased Inference on Treatment Effect in a High-Dimensional Model. *J. Amer. Statist. Assoc.*, 115(529):442–454, 2020. ISSN 0162-1459. doi: 10.1080/01621459.2018.1558062.

C.J.C.H Watkins and P. Dayan. Q-learning. *Mach. Learn.*, 8:279–292, 1992.

Baqun Zhang, Anastasios A. Tsiatis, Eric B. Laber, and Marie Davidian. A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018, 2012. ISSN 0006-341X. doi: 10.1111/j.1541-0420.2012.01763.x. URL `http://dx.doi.org/10.1111/j.1541-0420.2012.01763.x`.

Baqun Zhang, Anastasios A. Tsiatis, Eric B. Laber, and Marie Davidian. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika*, 100 (3):681–694, 2013a. ISSN 0006-3444. doi: 10.1093/biomet/ast014. URL `http://dx.doi.org/10.1093/biomet/ast014`.

Yichi Zhang, Eric B. Laber, Anastasios Tsiatis, and Marie Davidian. Using decision lists to construct interpretable and parsimonious treatment regimes. *Biometrics*, 71(4):895–904, 2015. ISSN 0006-341X. doi: 10.1111/biom.12354. URL `http://dx.doi.org/10.1111/biom.12354`.

Yichi Zhang, Eric B. Laber, Marie Davidian, and Anastasios A. Tsiatis. Estimation of optimal treatment regimes using lists. *J. Amer. Statist. Assoc.*, 113(524):1541–1549, 2018. ISSN 0162-1459. doi: 10.1080/01621459.2017.1345743.

Yuchen Zhang, John Duchi, and Martin Wainwright. Divide and conquer kernel ridge regression. In *Conference on Learning Theory*, pages 592–617, 2013b.

Tianqi Zhao, Guang Cheng, and Han Liu. A partially linear framework for massive heterogeneous data. *Ann. Statist.*, 44(4):1400–1437, 2016. ISSN 0090-5364. URL `https://doi.org/10.1214/15-AOS1410`.

Ying-Qi Zhao, Donglin Zeng, Eric B. Laber, and Michael R. Kosorok. New statistical learning methods for estimating optimal dynamic treatment regimes. *J. Amer. Statist. Assoc.*, 110(510):583–598, 2015. ISSN 0162-1459. doi: 10.1080/01621459.2014.937488. URL `http://dx.doi.org/10.1080/01621459.2014.937488`.

Yingqi Zhao, Donglin Zeng, A. John Rush, and Michael R. Kosorok. Estimating individualized treatment rules using outcome weighted learning. *J. Amer. Statist. Assoc.*, 107(499):1106–1118, 2012. ISSN 0162-1459. doi: 10.1080/01621459.2012.695674. URL http://dx.doi.org/10.1080/01621459.2012.695674.

S. Zhou, X. Shen, and D. A. Wolfe. Local asymptotics for regression splines and confidence regions. *Ann. Statist.*, 26(5):1760–1782, 1998. ISSN 0090-5364. doi: 10.1214/aos/1024691356. URL https://doi.org/10.1214/aos/1024691356.