Journal of Machine Learning Research 23 (2022) 1-65 Submitted 8/21; Revised 7/22; Published 7/22

Estimation and inference on high-dimensional individualized treatment rule in observational data using split-and-pooled de-correlated score

Muxuan Liang MUXUAN.LIANG@UFL.EDU
Department of Biostatistics
University of Florida
Gainesville, Florida 32611, USA

Young-Geun Choi YGCHOI@SM.AC.KR
Department of Statistics
Sookmyung Women's University
Seoul 04310, Korea

Yang Ning YN265@CORNELL.EDU
Department of Statistics and Data Science
Cornell University
Ithaca, Newyork 14853, USA

Maureen A Smith MAUREENSMITH@WISC.EDU
Departments of Population Health and Family Medicine
University of Wisconsin-Madison
Madison, Wisconsin 53706, USA

Ying-Qi Zhao YQZHAO@FREDHUTCH.ORG
Public Health Sciences Divisions
Fred Hutchinson Cancer Research Center
Seattle, Washington 98109, USA

Editor: Jie Peng

Abstract

With the increasing adoption of electronic health records, there is an increasing interest in developing individualized treatment rules, which recommend treatments according to patients' characteristics, from large observational data. However, there is a lack of valid inference procedures for such rules developed from this type of data in the presence of high-dimensional covariates. In this work, we develop a penalized doubly robust method to estimate the optimal individualized treatment rule from high-dimensional data. We propose a split-and-pooled de-correlated score to construct hypothesis tests and confidence intervals. Our proposal adopts the data splitting to conquer the slow convergence rate of nuisance parameter estimations, such as non-parametric methods for outcome regression or propensity models. We establish the limiting distributions of the split-and-pooled de-correlated score test and the corresponding one-step estimator in high-dimensional setting. Simulation and real data analysis are conducted to demonstrate the superiority of the proposed method.

Keywords: Individualized treatment rule, double-robustness, high-dimensional inference, semiparametric inference, precision medicine.

1. Introduction

An individualized treatment rule is a decision rule that maps the patient profiles $\boldsymbol{X} \in \mathcal{X}$, a subspace of $\mathbb{R}^p$, into the intervention space $A \in \mathcal{A}$, where p is the number of the covariates and $\mathcal{A}$ is the set of available interventions. Given an outcome of interest, the optimal individualized treatment rule maximizes the value function which is the mean outcome if it were applied to a target population. Understanding the driving factors of a data-driven treatment rule can help with identifying the source of the heterogeneous effects and with guiding practical applications of precision medicine.

The increasing adoption of electronic health records at healthcare centers has provided us unprecedented opportunities to understand the optimal individualized treatment rule through massive observational data. One of the difficulties in dealing with observational data is the high-dimensionality of the covariates. There have been various methods developed to estimate the optimal individualized treatment rule. For regression-based approaches, Q-learning methods (Watkins and Dayan, 1992; Chakraborty et al., 2010; Qian and Murphy, 2011; Laber et al., 2014a) pose a fully specified model assumption on the conditional mean of the outcomes given the covariates and treatments. Qian and Murphy (2011) approximates the conditional mean by a rich linear model, along with an l_1 penalty to accommodate high-dimensional data. A-learning methods (Murphy, 2003; Lu et al., 2013; Shi et al., 2016, 2018; Wu et al., 2021) pose a model assumption on the contrast function of the conditional means. With high-dimensional covariates, Shi et al. (2016, 2018) adopt penalized estimating equation or penalized regression with a linear contrast function. An alternative class of methods searches over a pre-specified class of individualized treatment rules to optimize an estimator of the mean outcome, usually called direct (Laber et al., 2014b), policy learning (Athey and Wager, 2017), value-search (Davidian et al., 2014) estimators, or C-learning (Zhang and Zhang, 2018). Especially, Zhang et al. (2012) adopts a doubly robust strategy to estimate the value under any treatment rules and directly optimize the estimated value. Their procedure can be applied to the observational data by plugging in a parametrically-estimated propensity score. Similarly, A-learning can also be extended to deal with observational data by plugging in a parametrically-estimated propensity score. Among these methods, Zhao et al. (2012) propose the outcome weighted learning approach based on an inverse probability weighted estimator of the value. Song et al. (2015) develop a variable selection method based on penalized outcome weighted learning for optimal individualized treatment selection.

Statistical inference for the optimal or estimated individualized treatment rule is particularly challenging in the presence of high-dimensional covariates. Confounding and selection bias presented in large observational data such as EHR data add one more layer of complexity. Liang et al. (2018b) propose a concordance-assisted learning algorithm in the presence of high-dimensional covariates. Nonetheless, they do not provide any inference procedures. Inference methods for A-learning approaches such as Song et al. (2017) and Jeng et al. (2018) are developed assuming the propensity score is known. Recently, Wu et al. (2021) provide an inference procedure for a high-dimensional single-index contrast function assuming a known propensity. Thus, their methods cannot be applied if data are collected from observational studies. Shi et al. (2018) derive the oracle inequalities of the proposed estimators for the parameters in a linear contrast function, but their work focuses on the selection consistency and has little discussion on the inference of the estimated rule. Their method depends on parametric assumptions on the propensity and outcome models, and thus may not be consistent when complex propensity or outcome models are expected. In practice, to avoid misspecification, flexible models may be adopted for the outcome regression or the propensity score. However, these models result in slow convergence rates for the nuisance parameters, and deteriorate the limiting distribution of the estimated decision rule. As such, it is important to propose an inference procedure for the estimated decision rule, which is valid under the high-dimensional setup and robust to flexible models for the nuisance parameters. Recent literature on the high-dimensional inference can assist with tackling this challenge. For example, van de Geer et al. (2014) propose a debiased Lasso approach for generalized linear models. Ning and Liu (2017) propose a de-correlated score test for low-dimensional parameters with the existence of the high-dimensional covariates, which is applicable for parametric models with correctly specified likelihoods. Dezeure et al. (2017) propose a bootstrap procedure for high-dimensional inference, but it is computationally intensive.

Another importance and related topic is the inference of the optimal value. The inference of the optimal value has been shown to be challenging at exceptional laws (non-regular case) where there exists a subgroup of patients for which treatment effect vanishes (Chakraborty et al., 2010; Laber et al., 2014c; Goldberg et al., 2014). To achieve the inference of the optimal value in low-dimensional setup, Chakraborty et al. (2014) propose an m-out of-n bootstrap to construct a confidence interval for the value. Luedtke and Van Der Laan (2016) propose an online one-step estimator which is the weighted average the values estimated on chunks of data increasing in size. Recently, Shi et al. (2020) use a subagging algorithm to aggregate value estimates obtained by repeated sample splittings. In both Luedtke and Van Der Laan (2016) and Shi et al. (2020), a single-split procedure is also discussed to facilitate the computation, though the resulting confidence interval might be wider. However, the value inference for high-dimensional setup is lacking.

In this work, we propose a novel penalized doubly robust approach, termed as penalized efficient augmentation and relaxation learning, to estimate the optimal individualized treatment rule in observational studies with high-dimensional covariates. We construct the decision rules by optimizing a convex relaxation of the augmented inverse probability weighted estimator of the value with penalties, which generalizes the method proposed in Zhao et al. (2019) to high-dimensional setup. The proposed procedure involves estimation of the conditional means of the outcomes and the propensity scores as nuisance parameters. As long as one of the nuisance models is correctly specified, we can consistently estimate the optimal individualized treatment rules under certain conditions. Furthermore, we propose a split-and-pooled de-correlated score test, which provides valid hypothesis testing and interval estimation procedures to identify the driving factors of the estimated decision rule. The proposed procedure generalizes the de-correlated score (Ning and Liu, 2017) to handle the potential slow convergence rates from the nuisance parameters estimation and to allow a general loss function. Sample-splitting is adopted to separate the estimation of the nuisance parameters from the construction of the de-correlated score, which is adopted in Chernozhukov et al. (2018) for inference on a low-dimensional parameter of interest in the presence of high-dimensional nuisance parameters. However, the inference on the estimated decision rule using the proposed approach requires a more sophisticated analysis due to the convex relaxation schemes. Theoretically, we show that the split-and-pooled decorrelated score is asymptotically normal even when the nuisance parameters are estimated non-parametrically with slow convergence rates. In addition, we use a single-split procedure to infer the value under the estimated decision rule.

2. Method

In this section, we propose the penalized efficient augmentation and relaxation learning and then introduce the proposed inference procedure.

2.1 Penalized Efficient Augmentation and Relaxation Learning

Let $\boldsymbol{X}$ be a p-dimensional random vector, which contains the baseline covariates capturing patient profiles. We assume that p can be much larger than the sample size n. Let $A \in \{-1, 1\}$ be the treatment assignment, and $Y \in \mathbb{R}$ be the observed outcome that higher values are preferred. Here, we adopt the framework of potential outcomes (Rubin, 1974, 2005). Denote the potential outcome under treatment $a \in \{-1, 1\}$ as $Y(a)$. Then the observed outcome is $Y = Y(a)I\{a = A\}$, where $I\{\cdot\}$ is the indicator function. An individualized treatment rule, denoted by D, is a mapping from the space of covariates $\mathcal{X} \subseteq \mathbb{R}^p$ to the space of treatments $\mathcal{A} = \{-1, 1\}$. With a slight abuse of notation, we write the observed outcome under this decision rule as $Y(D) = \sum_{a\in\{-1,1\}} Y(a)I\{a = D(\boldsymbol{X})\}$. The expectation of $Y(D)$, $V(D) = E(Y(D))$, is called the *value function* which is the average of the outcomes over the population if the decision rule were to be adopted. In order to express the value in terms of the data generative model, we assume the following conditions: 1) the stable unit treatment value assumption (Imbens and Rubin, 2015); 2) the strong ignorability $Y(-1), Y(1) \perp A \mid \boldsymbol{X}$; 3) Consistency $Y = Y(A)$. The stable unit treatment value assumption assumes that the potential outcomes for a patient do not vary with the treatments assigned to other patients. It also implies that there are no different versions of the treatment. The strong ignorability condition means that there is no unmeasured confounding between the potential outcomes and the treatment assignment mechanism. The optimal individualized treatment rule is defined as $D_{\text{opt}} = \arg\max_D\{V(D)\}$.

In this paper, due to the high-dimensional nature of the data we work with, we focus on deriving a linear decision rule of the form $D(\boldsymbol{x}) = \text{sgn}(\boldsymbol{x}^\top\boldsymbol{\beta})$, where $\boldsymbol{x} \in \mathcal{X}$ and the function $\text{sgn}(t) = 1$ if $t \geq 0$; $\text{sgn}(t) = -1$ if $t < 0$. To ensure the identifiability, we assume that the k^*-th coordinate of $\boldsymbol{\beta}$, $\boldsymbol{\beta}_{k^*} = 1$, for some k^*. The choice of k^* can be determined by the domain knowledge. Let $\pi(a; \boldsymbol{x}) = P(A = a \mid \boldsymbol{X} = \boldsymbol{x})$ and $Q(a; \boldsymbol{x}) = E(Y \mid \boldsymbol{X} = \boldsymbol{x}, A = a)$ for $a \in \{-1, 1\}$ and $\boldsymbol{x} \in \mathcal{X}$. Define the weights

$$\widehat{W}_a = W_a(Y, \boldsymbol{X}, A, \widehat{\pi}, \widehat{Q}) = \frac{YI\{A = a\}}{\widehat{\pi}(a; \boldsymbol{X})} - \frac{[I\{A = a\} - \widehat{\pi}(a; \boldsymbol{X})]\,\widehat{Q}(a; \boldsymbol{X})}{\widehat{\pi}(a; \boldsymbol{X})}$$

for $a \in \{-1, 1\}$, where $\widehat{\pi}(a; \boldsymbol{X})$ and $\widehat{Q}(a; \boldsymbol{X})$ are the estimators of $\pi(a; \boldsymbol{X})$ and $Q(a; \boldsymbol{X})$ respectively. Under the conditions above, the augmented inverse probability weighted estimator of the value function is

$$\widehat{V}(D) = E_n\left[\widehat{W}_1 I\{D(\boldsymbol{X}) = 1\} + \widehat{W}_{-1} I\{D(\boldsymbol{X}) = -1\}\right],$$

where $E_n[\cdot]$ denotes the empirical average. The estimator $\widehat{V}(D)$ enjoys the double robustness property. Assume that $\widehat{Q}(a;\boldsymbol{x})$ and $\widehat{\pi}(a;\boldsymbol{x})$ converge in probability uniformly to some deterministic limits, denoted by $Q^m(a;\boldsymbol{x})$ and $\pi^m(a;\boldsymbol{x})$, respectively. $\widehat{V}(D)$ converges to $V^m(D)$, where

$$V^m(D) = E\left[W_1^m I\left\{D(\boldsymbol{X})=1\right\} + W_{-1}^m I\left\{D(\boldsymbol{X})=-1\right\}\right].$$

Here, $W_a^m = W_a(Y,\boldsymbol{X},A,\pi^m,Q^m)$ is the limit that $\widehat{W}_a$ converges to, $a=\pm 1$. As shown in Zhao et al. (2019), if either $\pi^m(a;\boldsymbol{x}) = \pi(a;\boldsymbol{x})$ or $Q^m(a;\boldsymbol{x}) = Q(a;\boldsymbol{x})$, but not necessarily both, then $V^m(D) = V(D)$.

To avoid negative $\widehat{W}_a$, we consider its positive and negative parts separately and define $\widehat{W}_{a,+} = |\widehat{W}_a| 1\{\widehat{W}_a \geq 0\}$ and $\widehat{W}_{a,-} = |\widehat{W}_a| 1\{\widehat{W}_a \leq 0\}$. Maximizing $\widehat{V}(D)$ is equivalent to minimizing

$$E_n\left[\left(\widehat{W}_{1,+} + \widehat{W}_{-1,-}\right) I\left\{D(\boldsymbol{X}) \neq 1\right\} + \left(\widehat{W}_{1,-} + \widehat{W}_{-1,+}\right) I\left\{D(\boldsymbol{X}) \neq -1\right\}\right]. \tag{1}$$

Directly optimizing (1) is infeasible due to the indicator functions in the objective function, especially with a large number of covariates. To avoid minimizing the indicator function, we replace the indicator function with a strictly convex surrogate loss. Due to the strict convexity, the minimizer of the surrogate loss is always unique. Thus, we can relax the constraint that $\boldsymbol{\beta}_{k^*} = 1$. Furthermore, we add a sparse penalty function, which enables us to eliminate the unimportant variables from the derived rule. We denote the weight encouraging $A=1$ as $\widehat{\Omega}_+ = \widehat{W}_{1,+} + \widehat{W}_{-1,-}$ and the weight encouraging $A=-1$ as $\widehat{\Omega}_- = \widehat{W}_{1,-} + \widehat{W}_{-1,+}$. Our proposed estimator $\widehat{\boldsymbol{\beta}}$ is

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} E_n\left[\widehat{\Omega}_+ \phi\left(\boldsymbol{X}^\top \boldsymbol{\beta}\right) + \widehat{\Omega}_- \phi\left(-\boldsymbol{X}^\top \boldsymbol{\beta}\right)\right] + \lambda_n P(\boldsymbol{\beta}), \tag{2}$$

where ϕ is a convex surrogate loss, $P(\boldsymbol{\beta})$ is a sparse penalty function with respect to $\boldsymbol{\beta}$, and λ_n is a tuning parameter controlling the amount of penalization. In this paper, we focus on the l_1-lasso penalty $P(\boldsymbol{\beta}) = \|\boldsymbol{\beta}\|_1$. The framework allows a broad class of surrogate loss functions, such as logistic loss, $\phi(t) = \log\left(1+e^{-t}\right)$, see Section 3 for the detailed technical conditions on ϕ. The estimated decision rule can be subsequently obtained as $\widehat{D}(\boldsymbol{X}) = \operatorname{sgn}\left(\boldsymbol{X}^\top \widehat{\boldsymbol{\beta}}\right)$.

2.2 Split-and-pooled De-correlated Score Test

We define

$$l_\phi(\boldsymbol{\beta};\Omega_+^m,\Omega_-^m) = \Omega_+^m \phi\left(\boldsymbol{X}^\top \boldsymbol{\beta}\right) + \Omega_-^m \phi\left(-\boldsymbol{X}^\top \boldsymbol{\beta}\right),$$

and $\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta}} E\left[l_\phi(\boldsymbol{\beta};\Omega_+^m,\Omega_-^m)\right]$, where $\Omega_+^m = W_{1,+}^m + W_{-1,-}^m$ and $\Omega_-^m = W_{1,-}^m + W_{-1,+}^m$. To simplify notations, we will suppress the superscript and write them as Ω_+ and Ω_- instead. Let $X_j \in \mathbb{R}$ is the j-th covariate and $\boldsymbol{X}_{-j} \in \mathbb{R}^{p-1}$ includes the remaining covariates. Likewise, let β_j^* be the j-th coordinate of $\boldsymbol{\beta}^*$ and $\boldsymbol{\beta}_{-j}^*$ be a $p-1$ dimensional sub-vector of $\boldsymbol{\beta}^*$ without β_j^*. Without loss of generality, suppose that β_j^* is of interest. The statistical inference problem can be formulated as testing the null hypothesis $H_0: \beta_j^* = 0$ versus $H_1:$

$\beta_j^* \neq 0$, or constructing confidence intervals for β_j^*. The proposed method can be easily generalized to test any low-dimensional projection of $\boldsymbol{\beta}^*$.

Before we propose our inference procedure for $\boldsymbol{\beta}^*$, we introduce a lemma to show that under certain conditions, our inference procedure for $\boldsymbol{\beta}^*$ can provide information on $D_{\text{opt}}(\boldsymbol{x})$. In Lemma 1, we assume that $D_{\text{opt}}(\boldsymbol{x}) = \text{sgn}(\boldsymbol{x}^\top \boldsymbol{\beta}^{\text{opt}})$, which also indicates $D_{\text{opt}}(\boldsymbol{x}) = \text{sgn}(cx^\top \boldsymbol{\beta}^{\text{opt}})$ for any $c > 0$. To avoid $\boldsymbol{\beta}^{\text{opt}} = 0$ and identifiability issue, we restrict inference to regimes in which $\beta_{k^*}^{\text{opt}} = 1$. This implies that we would not infer $\beta_{k^*}^{\text{opt}}$ through $\beta_{k^*}^*$. In general, the contrast function $Q(1; \boldsymbol{x}) - Q(-1; \boldsymbol{x})$ could be a complex function of $\boldsymbol{x}$, but in many situations, the optimal rule $D_{\text{opt}}(\boldsymbol{x})$ may only depend on a linear function of $\boldsymbol{x}$ (Xu et al., 2015). In Lemma 1, we provide sufficient conditions that $\boldsymbol{\beta}^*$ satisfies $D_{\text{opt}}(\boldsymbol{X}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^*)$. In this case, the results on the sparsity pattern of $\boldsymbol{\beta}^*$ can be extended to inferring $\boldsymbol{\beta}^{\text{opt}}$.

Define two subspaces depending on $\boldsymbol{\beta}$,

$$\begin{aligned}\Delta_\phi(\boldsymbol{\beta}) &= \left\{ f(\boldsymbol{X}) \in L_2 : \text{cov}\left[f(\boldsymbol{X}), \left\{ \phi(\boldsymbol{X}^\top \boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}) \right\} \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}} \right] \geq 0 \right\}, \\ S_\phi(\boldsymbol{\beta}) &= \left\{ f(\boldsymbol{X}) \in L_2 : \text{cov}\left[f(\boldsymbol{X}), \left\{ \phi(\boldsymbol{X}^\top \boldsymbol{\beta}) + \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}) \right\} \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}} \right] \geq 0 \right\}.\end{aligned}$$

Lemma 1 *If the $D_{\text{opt}}(\boldsymbol{X})$ has a linear form, and $Q^m = Q$ or $\pi^m = \pi$ in Ω_+ and Ω_-, then $D_{\text{opt}}(\boldsymbol{X}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^*)$ if the following conditions are satisfied: (a) The contrast function $Q(1; \boldsymbol{X}) - Q(-1; \boldsymbol{X}) \in \Delta_\phi(\boldsymbol{\beta}^*)$, and the main effect $E(Y(1) + Y(-1) \mid \boldsymbol{X}) \in S_\phi(\boldsymbol{\beta}^*)$; (b) there exists a p-dimensional vector $\boldsymbol{P}$ such that $E(\boldsymbol{X} \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) = \boldsymbol{P}\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}$.*

The subspaces $\Delta_\phi(\boldsymbol{\beta})$ and $S_\phi(\boldsymbol{\beta})$ enjoy the following properties: (i) Any measurable function of $\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}$ belongs to $\Delta_\phi(\boldsymbol{\beta}) \cap S_\phi(\boldsymbol{\beta})$, $\forall \boldsymbol{\beta}$; (ii) Suppose that a function $g(\boldsymbol{X}) \in \Delta_\phi(\boldsymbol{\beta})$ (or $S_\phi(\boldsymbol{\beta})$), then the function $h(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) g(\boldsymbol{X}) \in \Delta_\phi(\boldsymbol{\beta})$ (or $S_\phi(\boldsymbol{\beta})$), where $h(\cdot)$ is an arbitrary measurable function. Thus, if $E(Y_1 \mid \boldsymbol{X})$ and $E(Y_{-1} \mid \boldsymbol{X})$ only depend on $\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}$, Condition (a) is easily satisfied. We provide examples in the Appendix B to further show that Condition (a) is satisfied by a large class of models, including data generative models that are dense and not single index models (see Example 2 in Appendix B). In particular, although method proposed in Wu et al. (2021) can deal with general single index models, it has more restriction on the sparsity level of the contrast function than our requirement.

Condition (b) on the design matrix $\boldsymbol{X}$ is common in the dimension reduction literature (Li, 1991; Zhu et al., 2006; Lin et al., 2018, 2019). It is satisfied if the distribution of $\boldsymbol{X}$ is elliptically symmetric. Li and Duan (1989); Duan and Li (1991) provide a thorough discussion on this condition in regression methods which aims to estimate a single index with an arbitrary and unknown link function. More specifically, they provide a bias bound when the elliptical symmetry is violated and show that the asymptotic bias is small when the elliptical symmetry is nearly satisfied. Further, Hall and Li (1993) shows that when the dimension of $\boldsymbol{X}$ is large, for most directions $\boldsymbol{\beta}^{\text{opt}}$ even the most nonlinear regression is still nearly linear. In addition, empirical studies by Brillinger and others suggest that quite often the bias may be negligible even for a moderate violation of condition (b) (Brillinger, 2012; Li and Duan, 1989).

Remark 2 *Alternatively, instead of assuming the conditions in Lemma 1, the desired relationship $D_{\text{opt}}(\boldsymbol{X}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^*)$ may still hold under some parametric assumptions on*

$E(Y_1 \mid \boldsymbol{X})$ and $E(Y_{-1} \mid \boldsymbol{X})$. For example, if the outcomes are non-negative and the following conditions are satisfied

$$\log\left\{E(Y_1 \mid \boldsymbol{X})/E(Y_{-1} \mid \boldsymbol{X})\right\} = \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}, \tag{3}$$

we still have $D_{\text{opt}}(\boldsymbol{X}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^*)$. Condition (3) poses a parametric assumption on $E(Y_1 \mid \boldsymbol{X})/E(Y_{-1} \mid \boldsymbol{X})$ (see Appendix B for the details). This ratio measures the relative change of the potential outcomes. Under Condition (3), hypothesis testing of $\boldsymbol{\beta}^*$ is equivalent to testing for the driving factors of the D_{opt}. Furthermore, the interval estimation of $\boldsymbol{\beta}^*$ can be interpreted through the specified model assumption in (3).

Next, we introduce our proposed inference procedure. Suppose that Ω_+ and Ω_- are known, then the estimator $\widehat{\boldsymbol{\beta}}$ is obtained by minimizing the empirical loss

$$E_n\left[l_\phi(\boldsymbol{\beta}; \Omega_+, \Omega_-)\right] + \lambda_n P(\boldsymbol{\beta}).$$

Let $\nabla l_\phi(\boldsymbol{\beta}; \Omega_+, \Omega_-) = \Omega_+ \phi'\left(\boldsymbol{X}^\top \boldsymbol{\beta}\right) - \Omega_- \phi'\left(-\boldsymbol{X}^\top \boldsymbol{\beta}\right)$. For $j \neq k^*$, the score function of $\boldsymbol{\beta}_j$ is $\mathrm{E}_n\left[\nabla l_\phi(\boldsymbol{\beta}; \Omega_+, \Omega_-) X_j\right]$. Let $\widehat{\boldsymbol{\beta}}_{\text{null(j)}}$ be a vector that equals to $\widehat{\boldsymbol{\beta}}$ with the j-th coordinate replaced by 0. In the low-dimensional setting where p is fixed, the score function with $\widehat{\boldsymbol{\beta}}_{\text{null(j)}}$, $E_n\left[\nabla l_\phi(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}; \Omega_+, \Omega_-) X_j\right]$, is asymptotically normal. Nevertheless, in a high-dimensional setting, the asymptotic normality of the score function $E_n\left[\nabla l_\phi(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}; \Omega_+, \Omega_-) X_j\right]$ is deteriorated by the high-dimensionality of $\widehat{\boldsymbol{\beta}}_{-j}$. Following Ning and Liu (2017), we use the semiparametric theory to de-couple the estimation error of $\widehat{\boldsymbol{\beta}}_{-j}$ with the score function of $\boldsymbol{\beta}_j$. A de-correlated score function is defined as $E_n\left[\nabla l_\phi(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}; \Omega_+, \Omega_-)\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)\right]$, where $\boldsymbol{w}_j^* = \left(\boldsymbol{I}^*_{-j,-j}\right)^{-1} \boldsymbol{I}^*_{-j,j}$ is chosen to reduce the uncertainty of the score function due to the estimation error of $\widehat{\boldsymbol{\beta}}_{-j}$. Denote $\nabla^2 l_\phi(\boldsymbol{\beta}; \Omega_+, \Omega_-) = \Omega_+ \phi''\left(\boldsymbol{X}^\top \boldsymbol{\beta}\right) + \Omega_- \phi''\left(-\boldsymbol{X}^\top \boldsymbol{\beta}\right)$. The $\boldsymbol{I}^*_{-j,-j}$ and $\boldsymbol{I}^*_{-j,j}$ are the corresponding partitions of $\boldsymbol{I}^* = E\left[\nabla^2 l_\phi(\boldsymbol{\beta}^*; \Omega_+, \Omega_-) \boldsymbol{X}\boldsymbol{X}^\top\right]$.

Under the null hypothesis, this de-correlated score function follows

$$n^{1/2} E_n\left[\nabla l_\phi(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}; \Omega_+, \Omega_-)\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)\right] \to N\left(0, \left(\boldsymbol{\nu}_j^*\right)^\top \boldsymbol{I} \boldsymbol{\nu}_j^*\right),$$

where $\boldsymbol{\nu}_j^*$ is a vector whose j-th coordinate is 1 and other coordinates equal to $-\boldsymbol{w}_j^*$. We propose to estimate the nuisance parameter $\boldsymbol{w}_j^*$ via

$$\min_{\boldsymbol{w}} E_n\left[\nabla^2 l_\phi\left(\widehat{\boldsymbol{\beta}}; \Omega_+, \Omega_-\right)\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}\right)^2\right] + \widetilde{\lambda}_n \|\boldsymbol{w}\|_1,$$

where $\widetilde{\lambda}_n$ is a tuning parameter. Denote the estimator for $\boldsymbol{w}_j^*$ as $\widehat{\boldsymbol{w}}_j$. A valid test for $H_0: \beta_j^* = 0$ is constructed based on

$$E_n\left[\nabla l_\phi(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}; \Omega_+, \Omega_-)\left(X_j - \boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}_j\right)\right]. \tag{4}$$

The nuisance parameters, Ω_+ and Ω_- are unknown in practice, and are estimated via modeling π and Q. To avoid misspecification, they can be estimated using flexible non-parametric or machine learning methods, which may lead to convergence rates slower than

$n^{-1/2}$. To overcome the possible slow convergence rates of $\widehat{\pi}$ and $\widehat{Q}$, we propose a split-and-pooled de-correlated score, where we consider a sample split procedure in constructing the de-correlated score function (Chernozhukov et al., 2018).

Let $I_1, \ldots, I_K$ be a random partition of the observed data with approximately equal sizes, where $K \geq 2$ is a fixed pre-specified integer. We assume that $\lfloor n/K \rfloor \leq |I_k| \leq \lfloor n/K \rfloor + 1$, for all $k = 1, \ldots, K$. Let $E_n^{(k)}[\cdot]$ denote the expectation defined by the data in I_k. For each $k \in \{1, \ldots, K\}$, we repeat the following procedure. First, we obtain $\widehat{\pi}_{(-k)}$ and $\widehat{Q}_{(-k)}$ using the data excluding I_k. In the presence of high-dimensional covariates, we can use generalized linear model with penalties (van de Geer, 2008) or kernel regression after a model-free variable screening (Li et al., 2012; Cui et al., 2015) for estimating π and Q. A data-split estimator $\widehat{\boldsymbol{\beta}}^{(k)}$ is obtained by

$$\widehat{\boldsymbol{\beta}}^{(k)} = \arg\min_{\boldsymbol{\beta}} E_n^{(k)} \left[l_\phi \left(\boldsymbol{\beta}; \widehat{\Omega}_+^{(-k)}, \widehat{\Omega}_-^{(-k)} \right) \right] + \lambda_{n,k} \|\boldsymbol{\beta}\|_1, \tag{5}$$

where $\widehat{\Omega}_+^{(-k)}$ and $\widehat{\Omega}_-^{(-k)}$ are computed with $\widehat{\pi}_{(-k)}$ and $\widehat{Q}_{(-k)}$ plugged in, and $\lambda_{n,k}$ is a tuning parameter. Then, we estimate $\boldsymbol{w}_j^*$ by

$$\widehat{\boldsymbol{w}}_j^{(k)} = \arg\min_{\boldsymbol{w}} E_n^{(k)} \left[\nabla^2 l_\phi \left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(-k)}, \widehat{\Omega}_-^{(-k)} \right) \left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w} \right)^2 \right] + \widetilde{\lambda}_{n,k} \|\boldsymbol{w}\|_1, \tag{6}$$

where $\widetilde{\lambda}_{n,k}$ is a tuning parameter. Let $\left(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)} \right)$ be a vector that equals $\widehat{\boldsymbol{\beta}}^{(k)}$ except its j-th coordinate replaced by 0. Finally, we construct the data-split de-correlated score test statistic $S_j^{(k)}(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)})$ as

$$S_j^{(k)} \left(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)} \right) = E_n^{(k)} \left[\nabla l_\phi \left(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}; \widehat{\Omega}_+^{(-k)}, \widehat{\Omega}_-^{(-k)} \right) \left(X_j - \boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}_j^{(k)} \right) \right]. \tag{7}$$

Combining K data-split estimators, we obtain the pooled estimator $\widehat{\boldsymbol{\beta}} = K^{-1} \sum_{k=1}^K \widehat{\boldsymbol{\beta}}^{(k)}$. Likewise, the pooled de-correlated score test statistic is $S_j = K^{-1} \sum_{k=1}^K S_j^{(k)} \left(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)} \right)$.

As shown in Theorem 4, under null hypothesis ($\beta_j^* = 0$), we have

$$n^{1/2} S_j \to N \left(0, \left(\boldsymbol{\nu}_j^* \right)^\top \boldsymbol{I} \boldsymbol{\nu}_j^* \right),$$

uniformly holds over all j's. The detailed algorithm is provided in Algorithm 1. In this algorithm, for a fixed $1 \leq k \leq K$, $\widehat{\pi}_{(-k)}$ and $\widehat{Q}_{(-k)}$ are trained on a subset of samples of size $n(K-1)/K$.

2.3 Confidence Intervals

We use the data-split de-correlated score to construct a valid confidence interval of β_j^*. This is motivated from the fact that the data-split de-correlated score $S_j^{(k)} \left(\boldsymbol{\beta}, \widehat{\boldsymbol{w}}_j^{(k)} \right)$ is also an unbiased estimating equation for β_j^* when fixing $\boldsymbol{\beta}_{-j} = \boldsymbol{\beta}_{-j}^*$. However, directly solving this estimating equation has several drawbacks, such as the existence of multiple roots or ill-posed Hessian (Chapter 5 in van der Vaart (2000)). Ning and Liu (2017) proposed a one-step

Algorithm 1: Inference of $\boldsymbol{\beta}^*$ using a sample-split procedure

Input: A random seed; n samples; a positive integer K.

Output: $\widehat{\boldsymbol{\beta}}$ and a p-value for $\mathcal{H}_0 : \boldsymbol{\beta}_j^* = 0$.

Randomly split data into K parts $\{I_k\}_{k=1}^K$ with equal size, and set $k = 1$;

Estimate π and Q on I_k^c and denote the estimator as $\widehat{\pi}_{(-k)}$ and $\widehat{Q}_{(-k)}$;

Obtain a data-split estimator $\widehat{\boldsymbol{\beta}}^{(k)}$ on I_k by (5), where $\lambda_{n,k}$ is tuned by cross-validation ;

Obtain an estimator $\widehat{\boldsymbol{w}}_j^{(k)}$ for $\boldsymbol{w}_j^*$ by (6), where $\widetilde{\lambda}_{n,k}$ is tuned by cross-validation ;

Construct the data-split de-correlated score test statistic $S_j^{(k)}(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}, \widehat{\boldsymbol{w}}_j^{(k)})$ by Equation (7), and the estimator of the variance $\widehat{\sigma}^2_{k,j} = E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(-k)}, \widehat{\Omega}_-^{(-k)}\right)\right\}^2\left(X_j - \boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}_j^{(k)}\right)^2\right]$;

Set $k = 2, 3, \ldots, K$, and repeat Step 2 and 5. Obtain $\left\{\widehat{\boldsymbol{\beta}}^{(k)}\right\}_{k=1}^K$ and $\left\{S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}, \widehat{\boldsymbol{w}}_j^{(k)}\right)\right\}_{k=1}^K$ as well as $\left\{\widehat{\sigma}^2_{k,j}\right\}_{k=1}^K$. Aggregate them by

$$\widehat{\boldsymbol{\beta}} = K^{-1}\sum_{k=1}^K \widehat{\boldsymbol{\beta}}^{(k)}, S_j = K^{-1}\sum_{k=1}^K S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}, \widehat{\boldsymbol{w}}_j^{(k)}\right), \widehat{\sigma}_j^2 = K^{-1}\sum_{k=1}^K \widehat{\sigma}^2_{k,j}.$$

Calculate the p-value by $2\left(1 - \Phi(n^{1/2}|S_j|/\widehat{\sigma}_j)\right)$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

estimator, which solved a first order approximation of the de-correlated score. Following their procedure, we construct the data-split one-step estimator, $\widetilde{\beta}_j^{(k)}$, as the solution to,

$$S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right) + E_n^{(k)}\left[\nabla^2 l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(-k)}, \widehat{\Omega}_-^{(-k)}\right) X_j(X_j - \boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}_j^{(k)})\right](\beta_j - \widehat{\beta}_j^{(k)}) = 0.$$

Hence, we have that $\widetilde{\beta}_j^{(k)} = \widehat{\beta}_j^{(k)} - S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right)/\widehat{I}_{j|-j}^{(k)}$, where

$$\widehat{I}_{j|-j}^{(k)} = E_n^{(k)}\left[\nabla^2 l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(-k)}, \widehat{\Omega}_-^{(-k)}\right) X_j(X_j - \boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}_j^{(k)})\right].$$

Finally, the pooled one-step estimator is the aggregation of these data-split one-step estimators following $\widetilde{\beta}_j = K^{-1}\sum_{k=1}^K \widetilde{\beta}_j^{(k)}$. In Section 3, we will show the asymptotic normality of the pooled one-step estimator $\widetilde{\beta}_j$, which provides a valid confidence interval for $\boldsymbol{\beta}_j^*$. The algorithm for constructing confidence intervals is presented in Algorithm 2.

Algorithm 2: Confidence interval of β_j^* using a sample-split procedure

Input: The data-split de-correlated score $S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right)$ and $\widehat{I}_{j|-j}^{(k)}$ for $k = 1, \ldots, K$; $\widehat{\sigma}^2$ from Algorithm 1.
Output: A 95% confidence interval for β_j^*.
Construct the data-split one-step estimator by $\widetilde{\beta}_j^{(k)} = \widehat{\boldsymbol{\beta}}_j^{(k)} - S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right)/\widehat{I}_{j|-j}^{(k)}$;
Aggregate these data-split one-step estimators by $\widetilde{\beta}_j = K^{-1}\sum_{k=1}^K \widetilde{\beta}_j^{(k)}$, and calculate $\widehat{I}_{j|-j} = K^{-1}\sum_{k=1}^K \widehat{I}_{j|-j}^{(k)}$;
Construct the 95% confidence interval by $\left(\widetilde{\beta}_j - 1.96n^{-1/2}\widehat{\sigma}_j/\widehat{I}_{j|-j}, \widetilde{\beta}_j + 1.96n^{-1/2}\widehat{\sigma}_j/\widehat{I}_{j|-j}\right)$.

2.4 Inference of the Value

We adopt an analogy of the single-split procedure (Luedtke and Van Der Laan, 2016; Shi et al., 2020) to infer the value under $D^*(\boldsymbol{X})$, $V(D^*)$, where $D^*(\boldsymbol{X}) = \text{sgn}\left(\boldsymbol{X}^\top\boldsymbol{\beta}^*\right)$. The single-split procedure splits the entire data set into two parts. We use one part for training and nuisance parameter fitting, and conduct inference on the other part. When $\boldsymbol{\beta}^* \propto \boldsymbol{\beta}_{\text{opt}}$, i.e., there is a constant $c > 0$ such that $\boldsymbol{\beta}^* = c\boldsymbol{\beta}_{\text{opt}}$, our procedure provides a valid inference procedure for the optimal value. The detailed procedure for inference of the value is presented in Algorithm 3.

3. Theoretical Properties

We assume the following conditions.

(C1) Each covariate X_j's is sub-Gaussian with common proxy σ_x, $\|\boldsymbol{\beta}^*\|_1 \leq R$, and $\max_j\{\|\boldsymbol{w}_j^*\|_1\} \leq R$; $\sup_{x\in\mathcal{X}}|Q(a;\boldsymbol{X})|$ is bounded, and the conditional distribution of

Algorithm 3: Inference of the value $V(D^*)$ using a single-split procedure

Input: A random seed; n samples.
Output: A 95% confidence interval for $V(D^*)$.
Randomly split the data into two sets, $\widetilde{I}_1$ and $\widetilde{I}_2$ with sample size n_1 and n_2, and obtain $\widehat{\boldsymbol{\beta}}$ using data in $\widetilde{I}_1$ by Algorithm 1;
Estimate π and Q on $\widetilde{I}_1$ and denote the estimator as $\widehat{\pi}$ and $\widehat{Q}$;
Estimate $V(D^*)$ on $\widehat{I}_2$ and denote the estimator as $\widehat{V}$,
$\widehat{V}(\widehat{D}) = E_{n_2}^{(2)}\left[W_{\widehat{D}(\boldsymbol{X})}(Y, \boldsymbol{X}, A, \widehat{\pi}, \widehat{Q})\right]$, where $\widehat{D}(\boldsymbol{X}) = \mathrm{sgn}(\boldsymbol{X}^\top \widehat{\boldsymbol{\beta}})$;
Estimate the variance and denote the estimator as $\widehat{\sigma}_V^2$,
$\widehat{\sigma}_V^2 = \mathrm{var}_{n_2}^{(2)}\left[W_{\widehat{D}(\boldsymbol{X})}(Y, \boldsymbol{X}, A, \widehat{\pi}, \widehat{Q})\right]$, where $\mathrm{var}_{n_2}^{(2)}(\cdot)$ is the sample variance on I_2.
Construct the 95% confidence interval by
$\left(\widehat{V}(\widehat{D}) - 1.96 n_2^{-1/2}\widehat{\sigma}_V, \widehat{V}(\widehat{D}) + 1.96 n_2^{-1/2}\widehat{\sigma}_V\right)$.

$Y(a) - Q(a;\boldsymbol{X})$ given X is sub-exponential, i.e., it is either bounded or satisfies that there exists some constants $M, \nu_0 \in R$ such that

$$E\left[\exp\{|Y(a) - Q(a;\boldsymbol{X})|/M\} - 1 - |Y(a) - Q(a;\boldsymbol{X})|/M \mid \boldsymbol{X}\right] M^2 \le \nu_0/2,$$

for both $a = 1$ and $a = -1$.

(C2) There exists some constants $0 < \pi_{\min} < \pi_{\max} < 1$ such that $\pi_{\min} \le \pi(a;\boldsymbol{X}) \le \pi_{\max}$ with probability 1.

(C3) ϕ is convex, and ϕ' is bounded with $\phi'(0) < 0$; for any $t \in [-\bar{c} - \epsilon, \bar{c} + \epsilon]$ with some constant $\epsilon > 0$ and a sequence t_1 satisfying $|t_1 - t| = o(1)$, it holds that $0 < \phi''(t) \le C$ and $|\phi''(t_1) - \phi''(t)| \le C|t_1 - t|\phi''(t)$ for some constant $C > 0$.

(C4) The smallest eigenvalue of $E[\nabla^2 l_\phi(\boldsymbol{\beta}^*; \Omega_+, \Omega_-)\boldsymbol{X}\boldsymbol{X}^\top]$ is larger than κ, where κ is a positive constant.

(C5) Suppose that for some $\alpha, \beta > 0$, $\sup_{\boldsymbol{x}} |\widehat{\pi}(a;\boldsymbol{x}) - \pi(a;\boldsymbol{x})| = O_p(n^{-\alpha})$ and $\sup_{\boldsymbol{x}} \left|\widehat{Q}(a;\boldsymbol{x}) - Q(a;\boldsymbol{x})\right| = O_p(n^{-\beta})$ for $a = 1$ and -1, we require that $Rn^{-\alpha-\beta+1/2} = o(1)$. In addition, we require that

$$R \max\{s^*, s'\} \log n (\log p)^{3/2} = o(n^{1/2}), \tag{8}$$

and

$$(n^{-\alpha} + n^{-\beta}) s^* \sqrt{\log p} \to 0, \tag{9}$$

where $s^* = \|\boldsymbol{\beta}^*\|_0$ and $s' = \max_j \|\boldsymbol{w}_j^*\|_0$.

Condition (C1) on the joint distribution of $(\boldsymbol{X}, A, Y)$ is weaker than the assumption in high-dimensional inference literature (van de Geer et al., 2014; Ning and Liu, 2017). Instead of sub-gaussian design, they only consider that the design is uniformly bounded. We also assume that $Y(a) - Q(a;\boldsymbol{X})$ is sub-exponential or bounded. This condition enables a faster convergence rate of high-dimensional empirical processes involving the estimation errors of $\widehat{\pi}$ and $\widehat{Q}$. Under this condition, if $\sup_X \left|\widehat{Q}(a;\boldsymbol{X}) - Q(a;\boldsymbol{X})\right| = o_p(1)$, we have

$$\left\|E_n\left[\{Y(a) - Q(a;\boldsymbol{X})\}\left\{Q(a;X) - \widehat{Q}(a;\boldsymbol{X})\right\}\boldsymbol{X}\right]\right\|_\infty = o_p\left((\log p/n)^{1/2}\right).$$

Condition (C2) prevents the extreme values in the true propensities. Condition (C3) requires that the surrogate loss ϕ has bounded first-order and second-order derivatives. The logistic loss satisfies these conditions. Condition (C4) is standard in high-dimensional inference literature (Ning and Liu, 2017; van de Geer et al., 2014; Dezeure et al., 2017). Condition (C5) is imposed for Algorithm 1. We assume that it holds on each split data set. To simplify the notation, we do not distinguish $\widehat{\pi}$ and $\widehat{Q}$ with $\widehat{\pi}_{(-k)}$ and $\widehat{Q}_{(-k)}$ for a fixed k. First it requires that both $\widehat{\pi}$ and $\widehat{Q}$ are consistent and the convergence rates satisfy $Rn^{-\alpha-\beta+1/2} = o(1)$. This can be attained if either the convergence rate of $\widehat{\pi}$ or $\widehat{Q}$ is sufficiently fast. For example, assuming $R = O(1)$, if π is estimated by a regression spline estimator and is known to be p_π-dimensional (low dimension) by design, we have $\sup_X |\widehat{\pi}(a;\boldsymbol{X}) - \pi(a;\boldsymbol{X})| = O_p\left(n^{-1/3}\right)$, where π is assumed to belong to the Hölder class with a smoothness parameter greater than $5p_\pi$ (Newey, 1997). Then $n^{-\alpha-\beta} \ll n^{-1/2}$ is satisfied when $n^{-\beta} \ll n^{-1/6}$. Second, formula (8) in Condition (C5) requires that the number of nonzero entries of $\boldsymbol{\beta}^*$ and $\boldsymbol{w}_j^*$ is smaller than the order of $n^{1/2}/(\log p)^{3/2}$, which is slightly more restrictive than the conditions in the high-dimensional inference literature (van de Geer et al., 2014; Ning and Liu, 2017) due to the sub-gaussian design. Finally, formula (9) of Condition (C5) indicates the convergence rates of the nuisance parameter estimations cannot be too slow if s^* and p increase fast with the sample size n.

Theorem 3 *Assume that Conditions (C1)-(C5) hold. By choosing* $\lambda_{n,k} \asymp (\log p/n)^{1/2}$ *, we have* $\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_1 = O_p\left(s^*(\log p/n)^{1/2}\right)$.

Theorem 3 assumes that both the outcome and propensity score models are correctly specified, $Q^m = Q$ and $\pi^m = \pi$ (implied by Condition (C5)). Nonetheless, our proposed estimator enjoys the doubly robustness property in the sense that $\widehat{\boldsymbol{\beta}}$ is still consistent if either $Q^m = Q$ or $\pi^m = \pi$. When $Q^m \neq Q$ and $\pi^m = \pi$, we have $\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_1 = O_p\left(s^* \max\left\{(\log p/n)^{1/2}, n^{-\alpha}\right\}\right)$; when $\pi^m \neq \pi$ and $Q^m = Q$, we have $\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_1 = O_p\left(s^* \max\left\{(\log p/n)^{1/2}, n^{-\beta}\right\}\right)$. This also indicates that as long as one of the estimators $\widehat{\pi}$ and $\widehat{Q}$ has a reasonably fast rate, the estimator $\widehat{\boldsymbol{\beta}}$ is consistent.

Theorems 4 and 5 provide the uniform validity of the testing procedures in Algorithm 1 and the confidence interval constructed using the pooled one-step estimator $\widetilde{\beta}_j$'s in Algorithm 2 via sample-splitting, respectively.

Theorem 4 *Assume that Conditions (C1)-(C5) hold. For Algorithm 1, under the null hypothesis* $H_0: \beta_j^* = 0, \forall j \in \mathcal{J} \subset \{1, \cdots, p\}$*, by choosing* $\lambda_{n,k} \asymp \tilde{\lambda}_{n,k} \asymp (\log p/n)^{1/2}$ *, we have*

$$\max_{j\in\mathcal{J}} \sup_{\alpha\in(0,1)} \left|P\left(\left|\sigma_j^{-1} n^{1/2} S_j\right| \le \Phi^{-1}(1-\alpha/2)\right) - (1-\alpha)\right| = o_p(1).$$

and $\max_j |\widehat{\sigma}_j - \sigma_j^2| = o_p(1)$*, where* $\widehat{\sigma}_j^2$ *is given in Algorithm 1, and*

$$\sigma_j^2 = \left(\boldsymbol{\nu}_j^*\right)^\top \mathrm{var}\left[\nabla^2 l_\phi(\boldsymbol{\beta}^*;\Omega_+,\Omega_-)\right]\boldsymbol{\nu}_j^*.$$

Theorem 5 *Assume that Conditions (C1)-(C5) hold. The pooled one-step estimator satisfies*

$$\max_j \sup_{\alpha\in(0,1)} \left| P\left(\left| n^{1/2} \left(\widetilde{\beta}_j - \beta_j^* \right) \widehat{I}_{j|-j} / \widehat{\sigma}_j \right| \le \Phi^{-1}(1-\alpha/2) \right) - (1-\alpha) \right| = o_p(1).$$

Remark 6 *Theorems 4 and 5 assume that both the propensity and the outcome models are correctly specified and estimated. Nonetheless, when the propensity score is known by the design of the experiment, the conclusions in Theorems 4 and 5 still hold even if the outcome model is misspecified. In contrast, Q-learning requires correctly specified outcome models even when the propensity is known. In practice, an individualized treatment rule can still be linear even if the contrast function is non-linear. As such, our modeling framework is more flexible. The advantages of our methods extend to the high-dimensional setting. The outcome weighted learning approach does not involve modeling outcomes. However, the corresponding penalized estimator in the outcome weighted learning approach may have a slower convergence rate than the proposed estimator in Theorem 3 when the propensity score is estimated with a slow rate. Therefore, the de-correlated score or the one-step estimator based on the outcome weighted learning approach cannot achieve a limiting distribution with $n^{1/2}$ convergence rate as in Theorems 4 and 5.*

To derive the asymptotic property of the inference procedure for the value, we further introduce the following conditions:

(C6) There exists an increasing function ψ such that 1) $\psi(0) = 0$; 2) there exists $\zeta > 0$ and $\limsup_{t\to 0} \psi(t)/t^{\zeta} < +\infty$; 3) $|E(Y(1)-Y(-1) \mid \boldsymbol{X})| \le \psi(|\boldsymbol{X}^\top\boldsymbol{\beta}^*|)$ when $|\boldsymbol{X}^\top\boldsymbol{\beta}^*| \le t_0$, where t_0 is a constant.

(C7) There exist constants $\gamma > 0$ and $C_\gamma > 0$ such that for any t in some neighborhood of 0, we have that $P\left(0 < \left|\boldsymbol{X}^\top\boldsymbol{\beta}^*\right| \le t\right) \le C_\gamma t^\gamma$.

Theorem 7 *Assume that Y is bounded and denote the sample size of $\widetilde{I}_1$ as n_1 and $\widetilde{I}_2$ as n_2. In addition to the conditions in Theorem 3, we further assume $n_1^{-\alpha-\beta} n_2^{1/2} = o(1)$ and one of the following conditions: 1) Conditions (C6) and (C7) holds with $\left(s(\log p/n_1)^{1/2}\right)^{\zeta+\gamma} = o_p(n_2^{-1/2})$; 2) Condition (C7) holds with $P\left(\left|\boldsymbol{X}^\top\boldsymbol{\beta}^*\right| = 0\right) = 0$ and $\left(s(\log p/n_1)^{1/2}\right)^{\gamma} = o_p(n_2^{-1/2})$, then we have*

$${n_2}^{1/2}\sigma_V^{-2}(\widehat{V}(\widehat{D}) - V(D^*)) \to N(0,1),$$

where $\sigma_V^2 = \mathrm{var}\left[W_{\widehat{D}(\boldsymbol{X})}(Y,\boldsymbol{X},A,\pi,Q)\right]$.

Condition (C6) implicitly assumes that $\boldsymbol{\beta}^*$ corresponds to the optimal individualized treatment rule. When Condition (C6) fails, the inference of the value under $D^*(\boldsymbol{X})$ requires stronger assumptions (see Theorem 9 in Appendix D for details). In the simulation studies and application, we choose $n_1 = n_2 = n/2$.

4. Simulation Studies

In this section, we test our estimation and inference procedure under various simulation scenarios. Let $\Delta(\boldsymbol{X}) = \{Q(1;\boldsymbol{X}) - Q(-1;\boldsymbol{X})\}/2$ and $S(\boldsymbol{X}) = \{Q(1;\boldsymbol{X}) + Q(-1;\boldsymbol{X})\}/2$. We generate $\boldsymbol{X} \sim N(0, \boldsymbol{I}_{p\times p})$, and $Y = A\Delta(\boldsymbol{X}) + S(\boldsymbol{X}) + \epsilon$, $\epsilon \sim N(0,1)$. Let $\boldsymbol{\beta}^{\text{opt}} = (1,1,-1,-1,0,\ldots,0)^\top, \boldsymbol{\beta}^*_S = (-1,-1,1,-1,0,\ldots,0)^\top$, and $\boldsymbol{\beta}^*_\pi = (1,-1,0,\ldots,0)^\top$. The following scenarios are considered: (I) $\Delta(\boldsymbol{X}) = \xi\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}$, $S(\boldsymbol{X}) = 0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_S$, and $\pi(1;\boldsymbol{X}) = \exp(0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_\pi)/\left\{1+\exp(0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_\pi)\right\}$; (II) $\Delta(\boldsymbol{X}) = \left\{\Phi\left(\xi\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) - 0.5\right\}\times\widetilde{\Delta}(\boldsymbol{X})$, $S(\boldsymbol{X}) = \exp\left(0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_S\right)$, $\pi(1;\boldsymbol{X}) = \exp\{(X_1^2+X_2^2+X_1X_2)/4\}/\left[1+\exp\{(X_1^2+X_2^2+X_1X_2)/4\}\right]$, where $\widetilde{\Delta}(\boldsymbol{X}) = 2(\sum_{j=1}^4 X_j)^2 + 2\xi$ and $\Phi(\cdot)$ is the cdf of the standard normal distribution.

Under these settings, the magnitude of the treatment effect $\Delta(\boldsymbol{X})$ changes with ξ, which ranges from 0.1 to 1. Scenario (I) features a linear outcome model $Q(a;\boldsymbol{X})$ for both $a = 1$ and $a = -1$, and a logistic model for the propensity. Scenario (II) has a nonlinear treatment effect $\Delta(\boldsymbol{X})$, though the decision boundary is still linear. The treatment assignment mechanism is also complex. More simulation results with a mixture of both discrete and continuous covariates, as well as highly correlated design matrices and non-regular cases, can be found in Appendix A.

We compare the pooled estimator with Q-learning, a regression-based method (Qian and Murphy, 2011). With high-dimensional covariates, we fit a linear regression with a lasso penalty in Q-learning for all scenarios. The inference target of interest is $\boldsymbol{\beta}^{\text{opt}}$. However, the limits of the coefficients estimates using either proposed method or Q-learning may not be identical to $\boldsymbol{\beta}^{\text{opt}}$. In our simulation experiments, we will test and construct confidence intervals for β^*_j's, $j = 1,\ldots,8$, the j-th coordinate of $\boldsymbol{\beta}^*$, which by abuse of notations, denote the limits of estimates under either method. We generate large data sets multiple times using the same data-generating process, and empirically verify that the sparsity pattern of $\boldsymbol{\beta}^*$ matches with that of $\boldsymbol{\beta}^{\text{opt}}$. Hence, inferences on $\boldsymbol{\beta}^*$ provide insights on the true optimal decisions. We conduct the hypothesis testing for Q-learning using the de-correlated score test proposed in Ning and Liu (2017), and construct 95% confidence intervals for the coefficients of interest in the context of Q-learning. For value inference, we implement the Algorithm 3 as our proposed approach; for Q-learning, we implement the Algorithm 3 with the coefficients $\widehat{\boldsymbol{\beta}}$ estimated from Q-learning approach. The true value $V(\boldsymbol{\beta}^*)$ is approximated by the average of estimated values on a large independent data set. An R package called `ITRInference` (see `https://github.com/muxuanliang/ITRInference.git`) is coded to implement the proposed method and Q-learning approach. For the proposed method, the user can specify the method or select from a list of candidates to estimate nuisance parameters. In our implementation, we choose to estimate π and Q functions nonparametrically for all scenarios. To be more specific, we first implement a distance correlation-based variable screening procedure (Li et al., 2012). We then fit a kernel regression using the selected variables after screening. When estimating π, we set caps at 0.1 and 0.9 to trim extreme values.

In all scenarios, the sample size n and the dimension p range from 350, 500, 800, 1600, 2500 to 8000. We set the nominal significance level at 0.05, and the nominal coverage at 95%. We report the type I errors, the powers of the hypothesis tests, and the value functions under the estimated decision rules out of 500 replications. In particular, we present the type I errors for testing β^*_5 to β^*_8, and the powers for testing β^*_1 to β^*_4. For each method, we also present the coverage of the interval estimations around the limiting coefficients. We also present the bias and the length of the confidence interval for the coefficients estimations and value estimations.

Figures 1 - 3 show the simulation results for different scenarios, with the sample size n varied and the p and ξ fixed. Additional results on varying p with n and ξ fixed can be found in Appendix A. As expected, in Scenarios (I) (Figure 1) where the regression model is correctly specified for Q-learning, Q-learning yields a better value function. Conversely, the proposed method outperforms the Q-learning method in Scenario (II) (Figure 2). In terms of the type I error and power, the proposed method is comparable to the Q-learning approach in Scenario (I) (Figure 1). For Scenarios (II) (Figure 2), our method is more powerful, and the type I errors are well controlled. Figure 3 also shows that the proposed method leads to less biased point estimations and shorter confidence intervals. The excessive bias of point estimations and lengths of interval estimations for the Q-learning approach may be due to the model misspecification. The coverage of β_5^* to β_8^* are concentrated near 95%, and the coverage of the β_1^* to β_4^* gradually approach 95% for the proposed method. For the coverage of the value $V(\boldsymbol{\beta}^*)$, the inference procedure achieves a valid CI for the value under the proposed approach in both scenarios when the sample size is large enough. However, the inference for the value under the Q-learning is under-coverage due to the model misspecification. In Appendix A, we also compare the proposed value inference procedure with bootstrap methods in terms of the coverage probabilities and lengths of confidence intervals.

5. Application to Complex Patients with Type-II Diabetes

In this section, we apply our proposed estimation and inference procedures to construct the optimal individualized treatment rule for complex patients with type-II diabetes. The data are collected from the electronic health records through Health Innovation Program at University of Wisconsin. The entire data set includes $n = 9101$ patients. There are 40 covariates, including socio-demographic variables, previous disease experiences, and baseline HbA1c levels, etc. The outcome is the indicator whether the patient successfully controls the HbA1c below 8% after a year. The treatment $A = 1$ if the patient received any medications, including insulin, sulphnea or OHA, and $A = -1$ otherwise. Among 9101 patients, 17.1% had a missing post-treatment HbA1c measurement, and 15.4% had the missing baseline HbA1c measurements. We impute missing values using Multivariate Imputation by Chained Equations (`MICE` package in R), which is based on the estimated conditional distributions of each covariate given other covariates (van Buuren and Groothuis-Oudshoorn, 2011). To address the possible interactions among covariates, we consider both raw covariates and all first-order interactions. We rank these covariates by their variances and select $p = 100$ covariates with top variances.

We split the data set into a training data set (80% of the entire data set) and a testing data set (20% of the entire data set). The proposed method and Q-learning are fitted on the training data set using the same strategies as described in simulation studies. To evaluate these estimated decision rules, we calculate the value function by $E_n[Y1\{A = \widehat{D}(\boldsymbol{X})\}/\widehat{\pi}_0]$, on the testing data set, where $\widehat{D}$ is the estimated decision rules on the training data set and $\widehat{\pi}_0$ is the estimated propensity scores on the testing data set. The entire procedure is

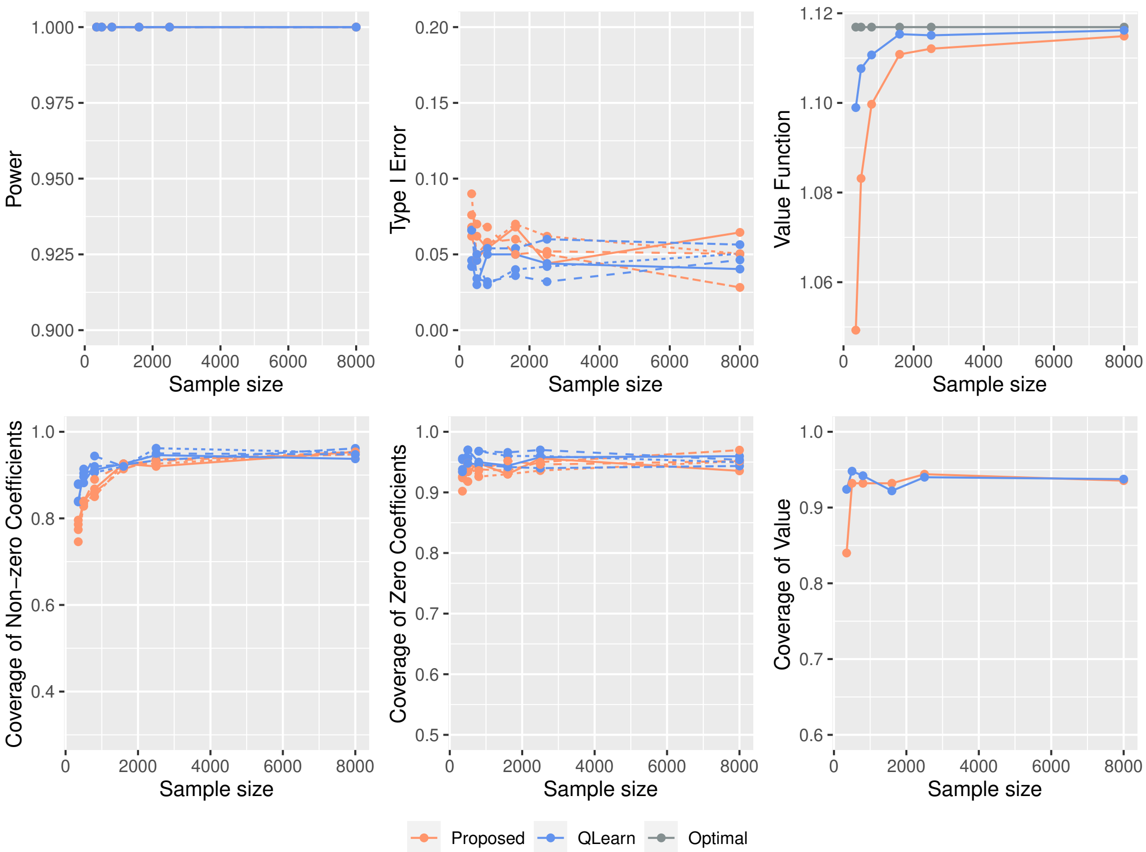

Figure 1: Simulation results for Scenario (I) with the change of sample size when $\xi = 0.7$ and $p = 2500$. Types of the line represent different coefficients.

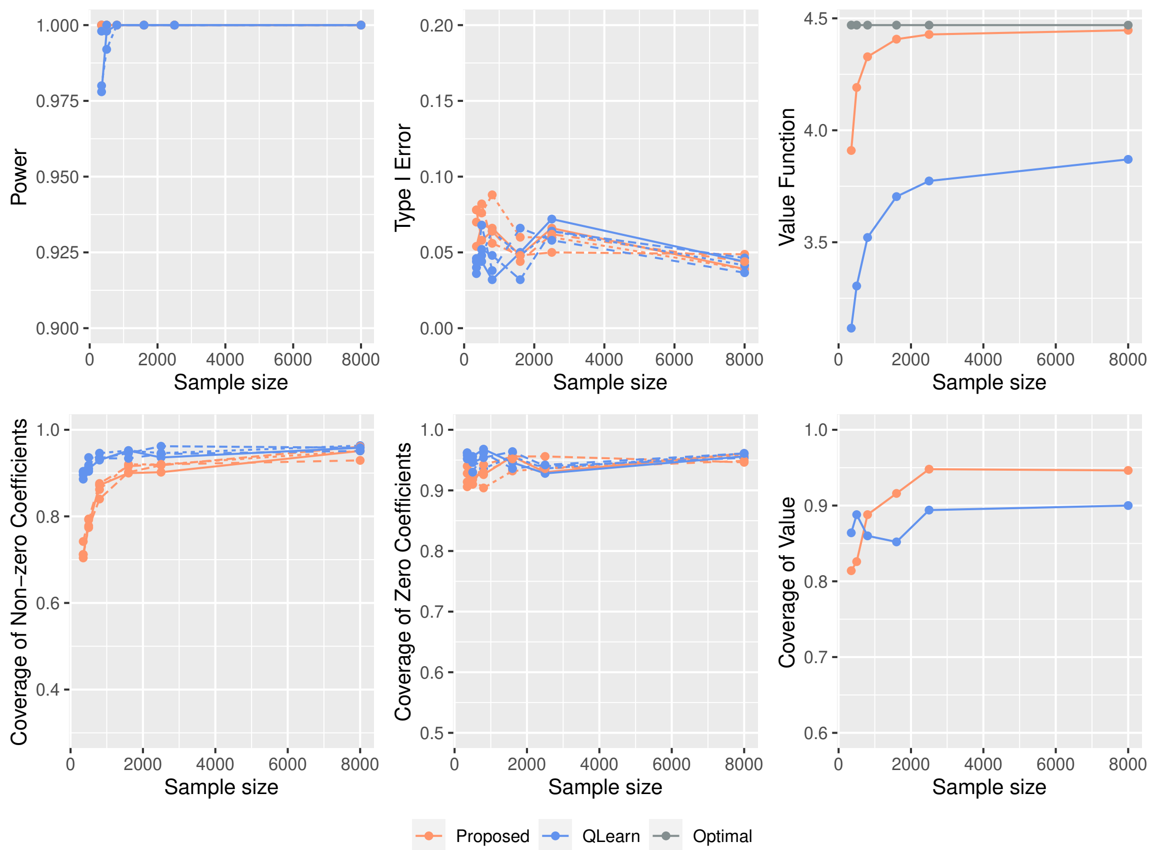

Figure 2: Simulation results for Scenario (II) with the change of sample size when $\xi = 0.8$ and $p = 2500$. Types of the line represent different coefficients.

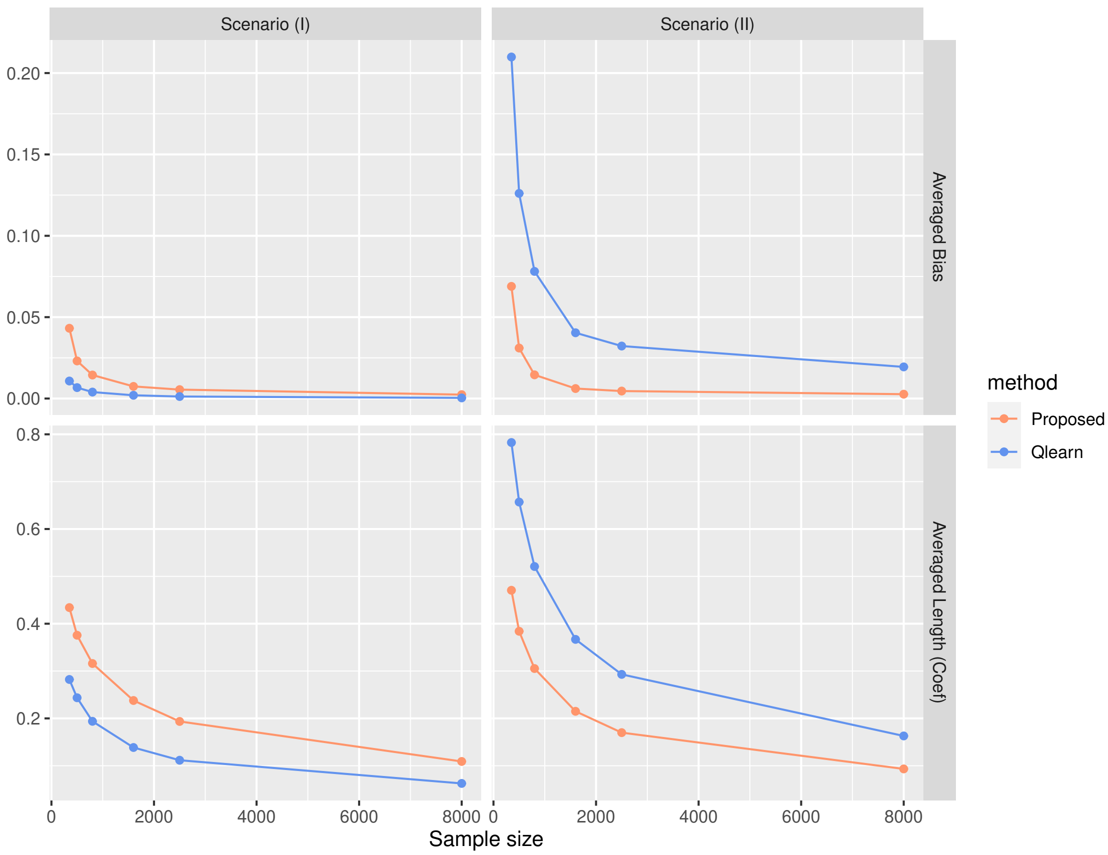

Figure 3: Simulation results in terms of the averaged bias of the coefficient estimates and the averaged length of the confidence intervals.

	Mean	Standard deviation
Observed	0.860	0.008
Proposed approach	0.877	0.015
Q-Learning	0.869	0.015

Table 1: Comparisons on value functions.

	Coef	P-value
Chronic Complications : Fluid and Electrolyte Disorders	-0.024	4.71×10^{-2}
Chronic Complications : African American	-0.027	3.58×10^{-2}
Alcohol Abuse : Entitlement Disability	-0.054	3.33×10^{-2}
HCC Community Score : Special Chronic Conditions	-0.022	2.99×10^{-2}
Hypertension : Lower Extremity Ulcer	-0.036	2.39×10^{-2}
HbA1c at Baseline : African American	0.019	2.26×10^{-2}
Entitlement Disability : Hypothyroidism	-0.024	2.25×10^{-2}
Cardiac Heart Failure : Peripheral Vascular Disease	-0.029	2.24×10^{-2}
Chronic Kidney Disease : HbA1c at Baseline	0.081	1.97×10^{-2}
Other Race : Special Chronic Conditions	0.016	1.95×10^{-2}
Liver Disease : Weight Loss	0.015	1.72×10^{-2}
Other Neurological Disorders : Female	-0.021	1.28×10^{-2}
Lower Extremity Ulcer : HbA1c at Baseline	0.039	9.60×10^{-3}
Chronic Complications : Bucketized Age	0.040	9.05×10^{-4}
HbA1c at Baseline : Female	0.044	8.47×10^{-8}

Table 2: Coefficients and p-value for the significant covariates of the estimated decision rule (CIs are included in Appendix C). Special chronic conditions refer to chronic conditions including amputation, chronic blood loss, drug abuse, lymphoma, metastatistic cancer, and peptic ulcer disease. Bucketized age refers to a variable created by bucketizing the raw age by its observed quartiles. Other Race refers to the race excluding White and Black.

repeated 100 times with random training and testing data splits. The mean and standard deviation (sd) of the value functions over these repeats are summarized in Table 1. Both the proposed and Q-learning methods construct decision rules that yield better results than the current clinical practice (sd of the difference is 0.0138 (Proposed); 0.0143 (Q-Learning)). Furthermore, our proposed method achieves a higher value function than Q-learning approach as shown in Table 1 (sd of the difference is 0.0115).

Next, we conduct the inference procedure to identify driving factors of the optimal individualized treatment rule as well as to provide an interval estimation using the entire data set. Results are presented in Table 2. After controlling for the false discovery rate below 0.05, our results indicate that a female patient with a higher HbA1c value at baseline is more likely to benefit from the treatment. The figure comparing the list of significant covariates selected by the proposed method and Q-learning can be found in Appendix C.

6. Discussion

In this paper, we consider a single-stage problem and assume a high-dimensional linear decision rule. In practice, especially in managing chronic diseases, dynamic treatment regimes are widely adopted, where sequential decision rules for individual patients adapt overtime to the evolving disease. One future direction is to develop inferential methods in the multi-decision setup. We can also extend the linear decision rule to a single index decision rule $d(\boldsymbol{X}^\top \boldsymbol{\beta}^*)$, where d is an unknown function. Throughout, we require that the surrogate loss function be differentiable. A non-differentiable surrogate loss such as the hinge loss does not have a well-defined Hessian, which hinders the construction of the de-correlated score. This can be addressed by a smoothed hinge loss or an approximation of the Hessian. We are currently working on these possible extensions.

In this work, we adopt the de-correlated score to infer the high-dimensional linear decision rule. It is also possible to use other high-dimensional influential tools developed recently. Partial penalized tests proposed in Shi et al. (2019) allow to test hypotheses involving a growing number of coefficients as the sample size increases. Ma et al. (2021) consider the global and simultaneous hypothesis testing for high-dimensional logistic regression models. Although a modified algorithm 1 can be combined with these methods, its theoretical property, especially the consequences of nuisance parameter estimation with slow rates, needs future investigations. In addition, this work can be extended to test multi-dimensional hypotheses, i.e., $\mathcal{H}_0 : \beta^*_j = 0, j \in \mathcal{G}$, where $\mathcal{G}$ is a subset of $\{1, \cdots, p\}$. For a low-dimensional sub-vector, .i.e, $|\mathcal{G}|$, the proposed de-biased estimators can be used to construct p-values or confidence regions. For a high-dimensional sub-vector, i.e. $\mathcal{G}$, the current procedure can be extended as well. However, the required relationship among α, β, p, and n needs future investigation. Thus, we will leave it as future work.

In addition, in this work, we assume that both nuisance parameters are correctly specified and estimated by nonparametric methods after the variable screening. Some recent literature in estimating the average treatment effect assumes that only one of the nuisance parameters is correctly specified (Athey et al., 2018; Ning et al., 2020; Tan, 2020; Smucler et al., 2019). In these approaches, the correct specified nuisance parameter is assumed to follow a structural model such as linear or partially linear models. It would be interesting to investigate how to extend these approaches to ITR inference.

In this work, we approximate the indicator function by a smooth convex surrogate. In addition to a smooth surrogate, many other non-smooth or non-convex loss functions such as the hinge loss (Cortes and Vapnik, 1995), the ramp loss (Collobert et al., 2006a,b) and the ψ-learning loss (Shen et al., 2003) can be considered to approximate the indicator function. Especially, the non-convex loss function such as the ramp loss and the ψ-learning loss is more robust to the presence of outliers. However, the non-convex/non-smooth surrogate loss may be hard to optimize and the non-convexity creates an additional barrier in high-dimensional inference. We would consider this extension as our future work.

Another future work is to extend the proposed approach to multiple treatment options setup. There are several possible directions. The first direction is to transform the multiple treatment problem into multiple binary decision problems. We can consider a sequential decision-making strategy (Zhou et al., 2018) by conducting a series of binary treatment selections. It is shown that such strategy is Fisher consistent. Another direction is to adopt techniques used in multi-label classification problems to estimate the optimal individualized treatment rule (Liang et al., 2018a). We can incorporate the weights based on the outcome model and propensity model into this framework and develop the corresponding inferential procedures. We are currently working on these extensions.

Acknowledgments

We would like to acknowledge support for this project from the NIH/NIDDK R01 DK108073. We also appreciate the editor, Dr. Jie Peng, and two anonymous reviewers for their helpful comments.

Appendix A.

In Appendix A, we provide additional simulation results.

Simulation Results with ξ and p Changed

In this section, we provides more simulation results. Let $\Delta(\boldsymbol{X}) = \{Q(1;\boldsymbol{X}) - Q(-1;\boldsymbol{X})\}/2$ and $S(\boldsymbol{X}) = \{Q(1;\boldsymbol{X}) + Q(-1;\boldsymbol{X})\}/2$. We generate $\boldsymbol{X} \sim N(0, \boldsymbol{I}_{p\times p})$, and $Y = A\Delta(\boldsymbol{X}) + S(\boldsymbol{X}) + \epsilon$, where $\epsilon \sim N(0,1)$ is the random error.

Denote $\boldsymbol{\beta}^{\text{opt}} = (1,1,-1,-1,0,\cdots,0)^\top$, $\boldsymbol{\beta}^*_S = (-1,-1,1,-1,0,\cdots,0)^\top$, and $\boldsymbol{\beta}^*_\pi = (1,1,1,0,-1,0,-1,0,\cdots,0)^\top$. The following scenarios are considered:
(I) $\Delta(\boldsymbol{X}) = \xi\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}$, $S(\boldsymbol{X}) = 0.8\boldsymbol{X}^\top\boldsymbol{\beta}^*_S$, and $\pi(1;\boldsymbol{X}) = 1/\left\{1+\exp(-0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_\pi)\right\}$;
(II) $\Delta(\boldsymbol{X}) = \left\{\Phi\left(\xi\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) - 0.5\right\}\times\widetilde{\Delta}(\boldsymbol{X})$, $S(\boldsymbol{X}) = \exp\left(0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_S\right)$, $\pi(1;\boldsymbol{X}) = \exp((X_1^2 + X_2^2 + X_1X_2)/4)/\left\{1+\exp((X_1^2+X_2^2+X_1X_2)/4)\right\}$, where $\widetilde{\Delta}(\boldsymbol{X}) = 2(\sum_{j=1}^4 X_j)^2 + 2\xi$ and $\Phi(\cdot)$ is the cdf of the standard normal distribution.

Under these settings, the magnitude of the treatment effect $\Delta(\boldsymbol{X})$ changes with ξ, which ranges from 0.1 to 1. In all scenarios, the sample size n and the dimension p range from 350, 500, 800, 1600, 2500, to 8000. When p changes from 350 to 2500, these settings include both low-dimensional ($p = 350$, $n = 2500$) and high-dimensional ($n = 350$, $p = 2500$) settings. We set the nominal significant level at 0.05, and the nominal coverage at 95%. We report the type I error, the power of the testing, and the value functions under the estimated decision rules out of 500 replications. For simplicity, we only test for β^*_l's, $l = 1,\cdots,8$, where β^*_l is the l-th coordinate of $\boldsymbol{\beta}^*$.

Figuress 4 and 5 show the simulation results for Scenario (I) and (II). Figure 4 shows that when we fix $n = 500$, the proposed method suffers from the high-dimensional nonparametric estimation of Q compared with Q-learning in terms of the value function as p increases. Figure 5 shows that the proposed method dominates the Q-learning in terms of the value function no matter how ξ and p changes.

Non-regular Scenarios

In this section, we provides two additional simulation scenarios. These two scenarios share the same conditional distribution of Y given $A, \boldsymbol{X}$ but have different design matrix. Both scenarios are non-regular in the sense that there exists a subgroup of patients for whom treatment is neither beneficial nor harmful. Specifically, we consider a function $z(t) = (z-0.2)^3 I\{z \geq 0.2\} + (z+0.2)^3 I\{z \leq -0.2\}$. Then we choose $\Delta(\boldsymbol{X}) = z\left(\xi\left\{4\Phi\left(\xi\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) - 2\right\}\right)$, where $\Phi(\cdot)$ is the cdf of the standard normal distribution; $S(\boldsymbol{X}) = \exp\left(0.4\boldsymbol{X}^\top\boldsymbol{\beta}^*_S\right)$; $\pi(1;\boldsymbol{X}) = \exp(0.25(X_1^2+X_2^2+X_1X_2))/\left\{1+\exp(0.25(X_1^2+X_2^2+X_1X_2))\right\}$.

In Scenario III.a, the design matrix involves only continuous variables. In Scenario III.b, the design matrix involves discrete variables and a dense correlation matrix between covariates. Specially, we consider (III.a) $\boldsymbol{X} \sim N(0, \boldsymbol{I}_{p\times p})$; (III.b) $\widetilde{\boldsymbol{X}}$ follows $N(0, \boldsymbol{\Sigma})$, and then $X_j = I\{j = 4i+3, i \in \mathbb{N}\}I\{\widetilde{X}_j > 0\} + I\{j \neq 4i+3, i \in \mathbb{N}\}\widetilde{X}_j$, where $\boldsymbol{\Sigma}$ is a $p \times p$ matrix with the (i,j)th entry $0.2^{|i-j|}$.

Figures 6 and 7 show the results under non-regular cases. Compared with Q-learning, the proposed method has higher value and comparable testing power. On the coverage of value, the proposed method is less susceptible to model mis-specification compared with

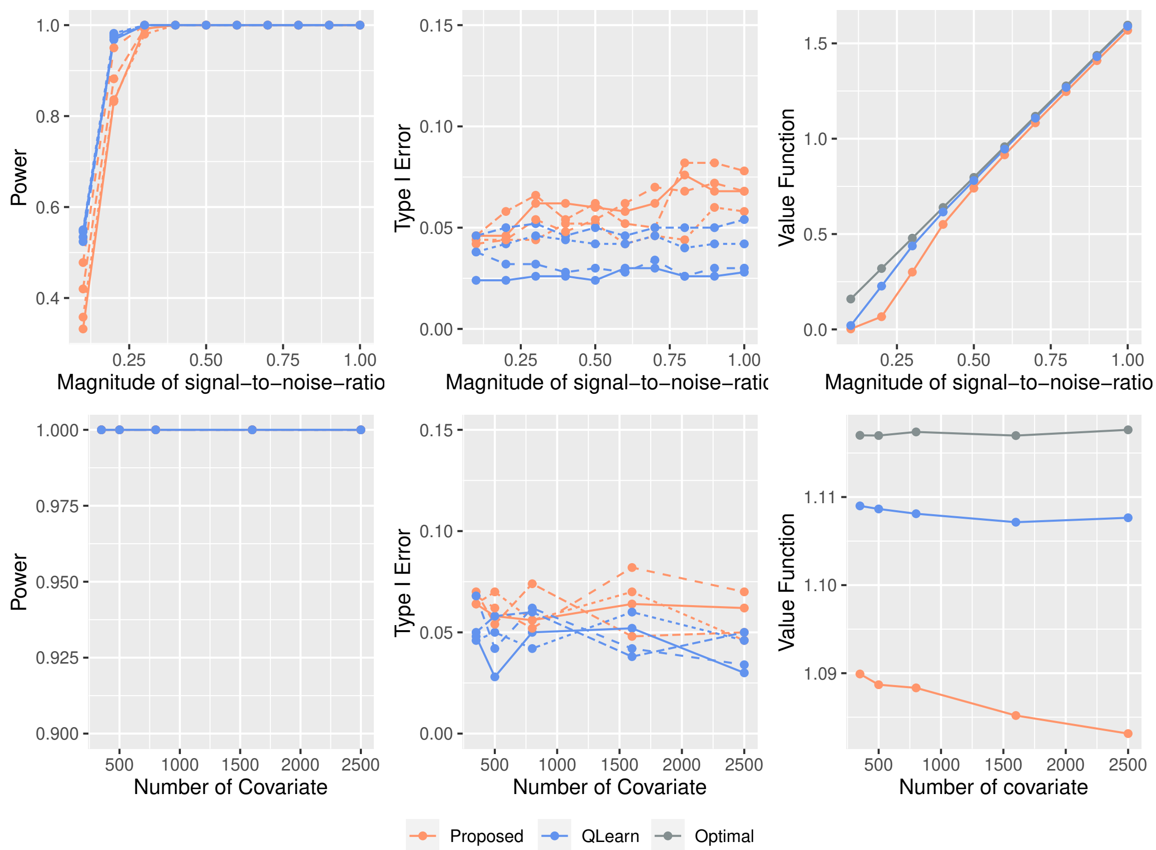

Figure 4: Simulation results for Scenario (I) with the change of ξ ($p = 2500$) and p ($\xi = 0.7$) when $n = 500$. Types of the line represent different coefficients.

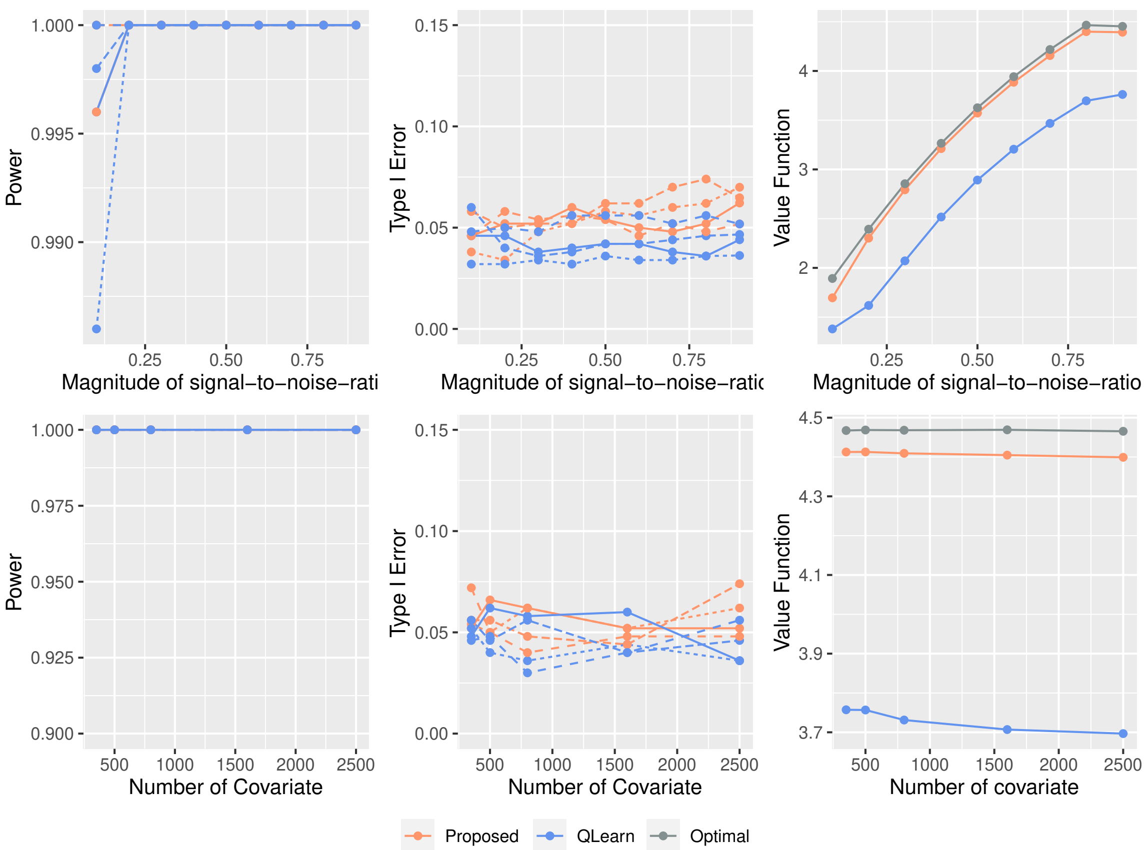

Figure 5: Simulation results for Scenario (II) with the change of ξ ($p = 2500$) and p ($\xi = 0.8$) when $n = 1600$. Types of the line represent different coefficients.

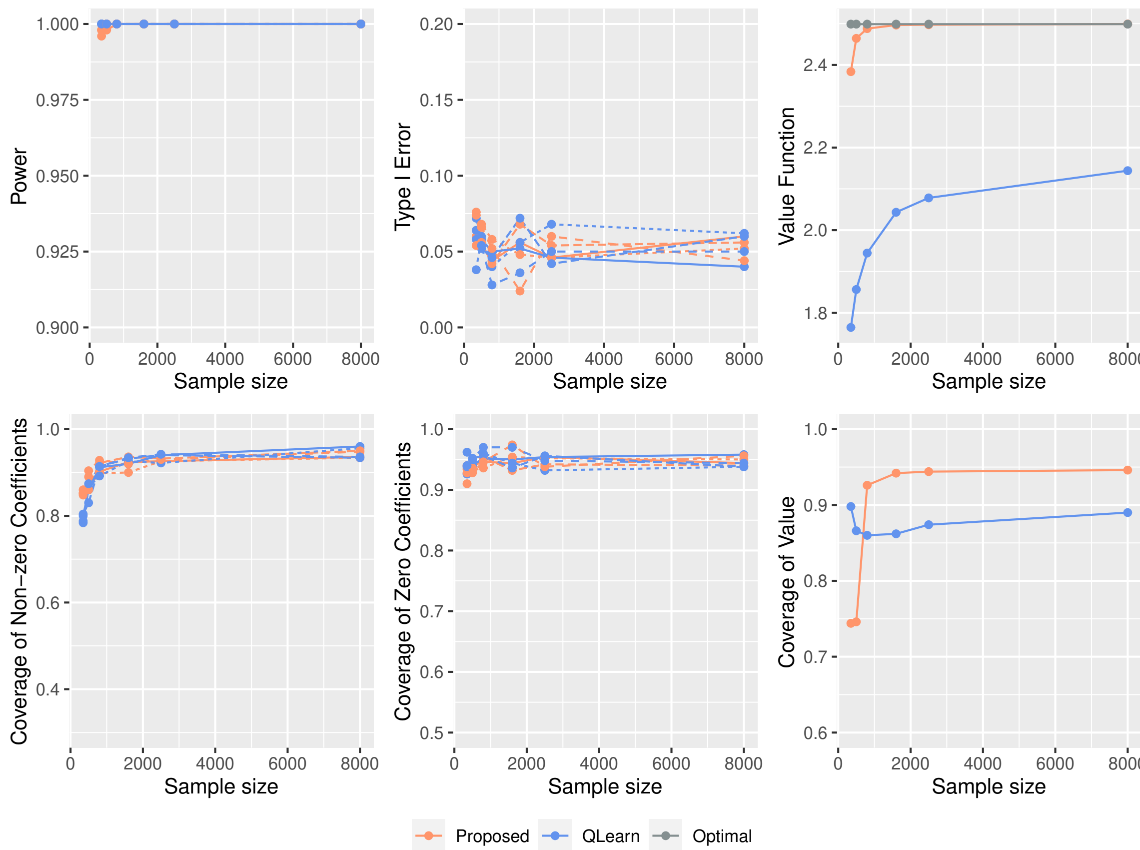

Figure 6: Simulation results for Scenario (III.a) with the change of sample size when $\xi = 0.8$ and $p = 2500$. Types of the line represent different coefficients.

the Q-learning approach. Figures 8 and 9 show the results with changing ξ and p. When ξ increases, the region where treatment effect vanishes gets smaller. Especially, when $\xi = 0.1$, there is no treatment effect for any patients. We can see that the proposed approach has higher value function no matter how ξ and p changes. The Q-learning has slightly worse lower when the non-regularity is severe.

Comparison of Inference Methods for the Optimal Value

In this section, we compare different methods to infer the optimal value. We consider the proposed method (direct), standard bootstrap procedure (bootstrap), and the weighted bootstrap procedure (weightedBootstrap). For the weighted bootstrap procedure, for each bootstrap, we only bootstrap the sampling weighted from a random variable with a mean 1 and unit variance. We do $B = 1000$ bootstraps for each bootstrap-based procedure. For each inference method, we use the PEARL and Q-learning to estimate the optimal decision rules. We follow the data generation procedure in Scenario (I), (II), (III.a), and (III.b) and report the coverages of the optimal value and the lengths of the confidence interval.

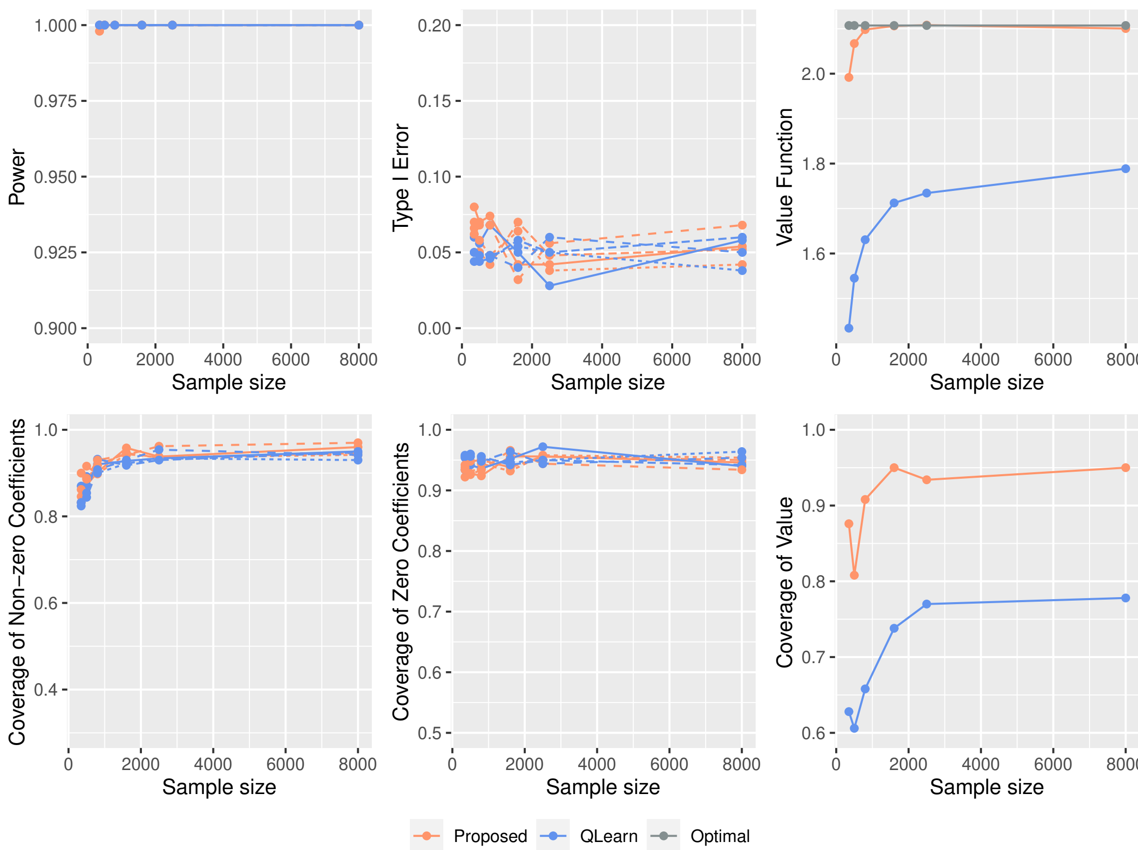

Figure 7: Simulation results for Scenario (III.b) with the change of sample size when $\xi = 0.8$ and $p = 2500$. Types of the line represent different coefficients.

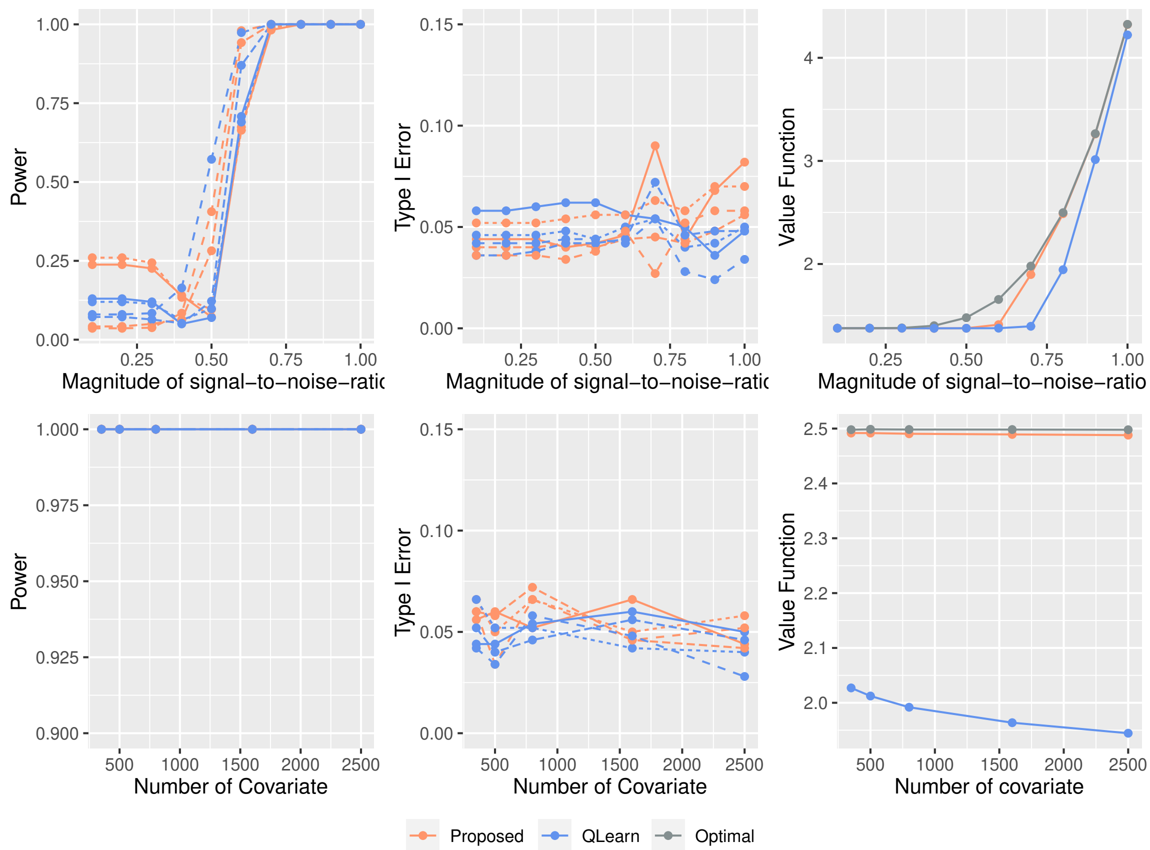

Figure 8: Simulation results for Scenario (III.a) with the change of ξ ($p = 2500$) and p ($\xi = 0.8$) when $n = 800$. Types of the line represent different coefficients.

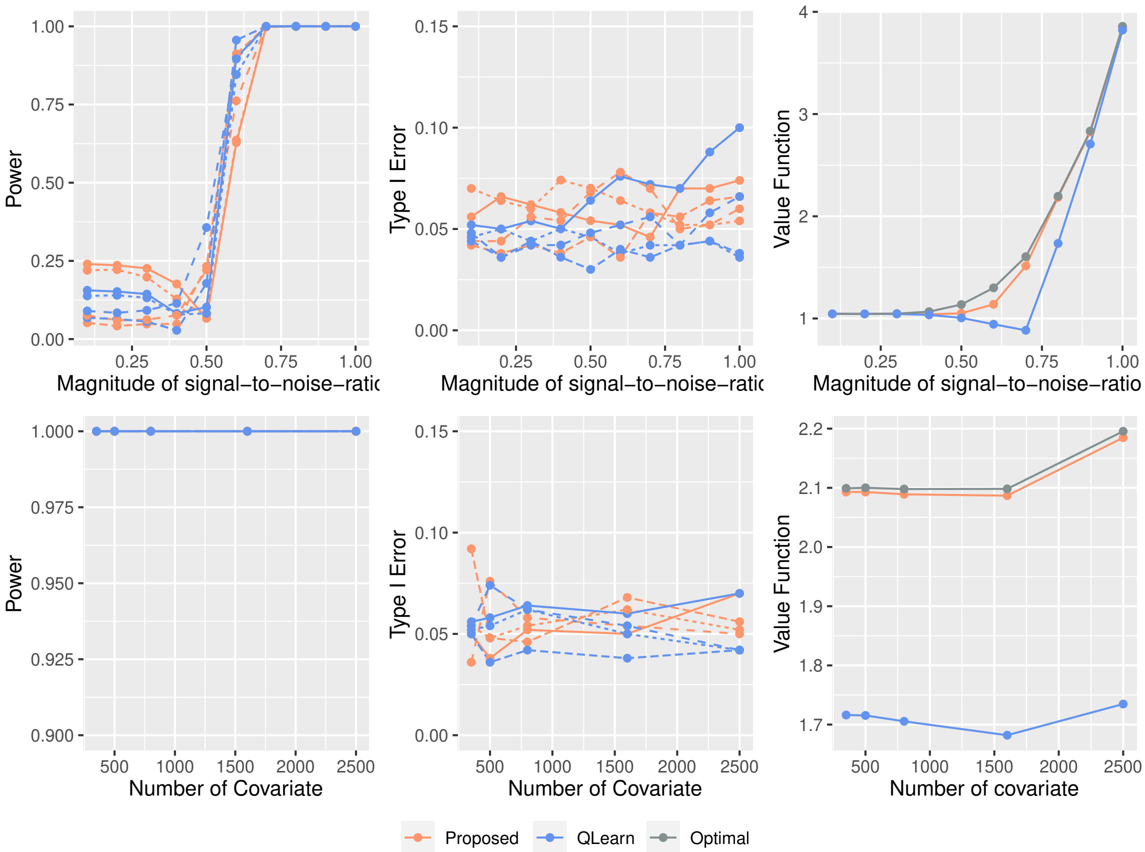

Figure 9: Simulation results for Scenario (III.b) with the change of ξ ($p = 2500$) and p ($\xi = 0.8$) when $n = 800$. Types of the line represent different coefficients.

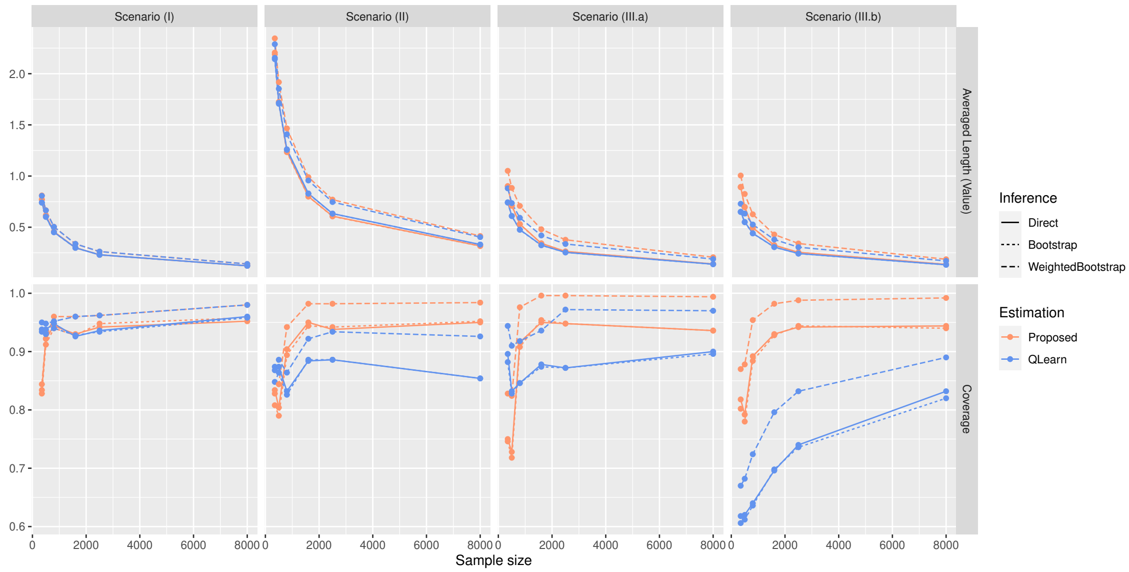

Figure 10: Simulation results for value inference.

From Figure 10, the proposed method has a similar performance to the standard bootstrap method in terms of the coverages and lengths of the CIs. However, the weighted bootstrap leads to slightly over-coverage, especially for non-regular settings, i.e., Scenario (III.a) and (III.b).

Illustration of Double Robustness

In this section, we check the double robustness of the proposed method under either misspecified propensity model or outcome models. We generate the data following Scenario (II) and consider alternative methods to estimate the propensity and outcome models. Specifically, we consider 1) the proposed method with non-parametrically estimated propensity and outcome models (Both Correct); 2) the proposed method with a non-parametrically estimated propensity and outcome models estimated by linear regressions (Outcome Missed); 3) the proposed method with non-parametrically estimated outcome models and a propensity estimated by logistic regression (Propensity Missed). Figure 11 shows that as the sample size increases, the values of the decision rules derived by all these approaches achieve the optimal value. This implies that our proposed method is consistent if either the outcome models or the propensity is correctly specified.

Appendix B.

In Appendix B, we provide a sufficient condition and examples for $\boldsymbol{\beta}^* \propto \boldsymbol{\beta}^{\text{opt}}$. Lemma 1 provides sufficient conditions that $\boldsymbol{\beta}^*$ satisfies $D_{\text{opt}}(\boldsymbol{X}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) = \text{sgn}(\boldsymbol{X}^\top \boldsymbol{\beta}^*)$, which indicates the inference of $\boldsymbol{\beta}^*$ is equivalent to that of $\boldsymbol{\beta}^{\text{opt}}$. First, we show the proof of Lemma 1.

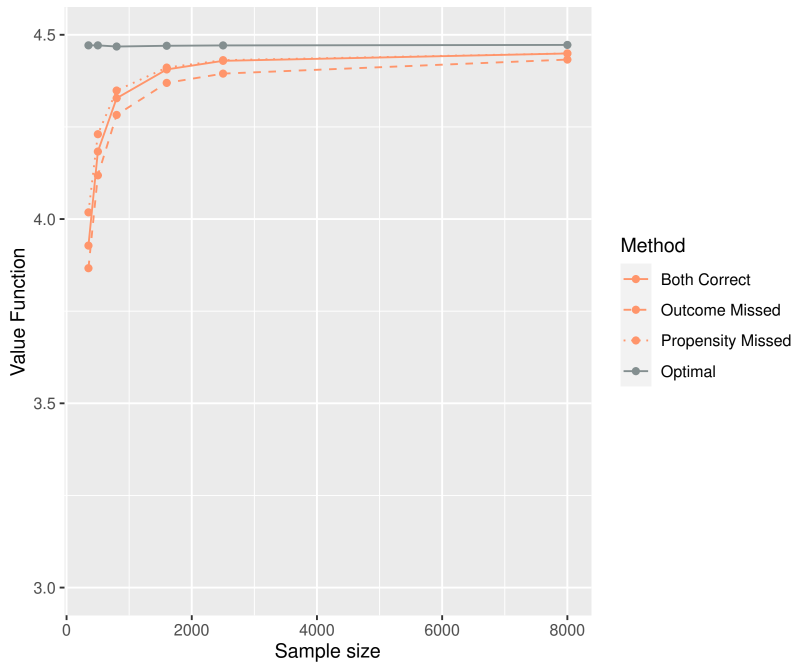

Figure 11: Simulation results for double robustness.

Proof. First, we provide a proof for Lemma 1. Under the Lemma 1, we show that $\boldsymbol{\beta}^* \propto \boldsymbol{\beta}^{\text{opt}}$. Without loss of generality, we assume that W_a's are non-negative.

For $\beta = \boldsymbol{\beta}^*$, we have the following

$$
\begin{aligned}
& E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{X}^\top\boldsymbol{\beta}\right) + E(\Omega_- \mid \boldsymbol{X})\phi(-\boldsymbol{X}^\top\boldsymbol{\beta})\right] \\
= \; & E\left[\Delta/2\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top\boldsymbol{\beta})\right\} + S/2\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta})\right\}\right] \\
= \; & E\left[E\left[E[\Delta/2\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top\boldsymbol{\beta})\right\} + S/2\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta})\right\} \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right]\right] \\
\geq \; & E\left[E(\Delta(\boldsymbol{X})/2 \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})E\left[\phi(\boldsymbol{X}^\top\boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right]\right. \\
& \left. + E\left[S(\boldsymbol{X})/2 \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}]E\left[\phi(\boldsymbol{X}^\top\boldsymbol{\beta}) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right]\right)\right] \\
= \; & E\left[E(\Omega_+ \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})E\left[\phi\left(\boldsymbol{X}^\top\boldsymbol{\beta}\right) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right]\right. \\
& \left. + E(\Omega_- \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})E\left[\phi\left(-\boldsymbol{X}^\top\boldsymbol{\beta}\right) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right]\right] \\
\geq \; & E\left[E(\Omega_+ \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right] \\
= \; & E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + \mathrm{E}(\Omega_- \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right].
\end{aligned}
$$

The first inequality comes from Condition (a). The last equality comes from the convexity of $\phi(\cdot)$ and Condition (b). Next, we notice that

$$\begin{aligned}
& E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right] \\
& -E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right] \\
= \ & E\left[\{E(\Omega_+ \mid \boldsymbol{X}) - E(\Omega_- \mid \boldsymbol{X})\}\left\{\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) - \phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right\}\right] \\
= \ & E\left[\{Q(1;\boldsymbol{X}) - Q(-1;\boldsymbol{X})\}\left\{\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) - \phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right\}\right]
\end{aligned}$$

If $\boldsymbol{P}^\top\boldsymbol{\beta} > 0$, we have

$$\begin{aligned}
& E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right] \\
\leq \ & E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right].
\end{aligned}$$

If $\boldsymbol{P}^\top\boldsymbol{\beta} \leq 0$, we have

$$\begin{aligned}
& E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right] \\
\geq \ & E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(-\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(\boldsymbol{P}^\top\boldsymbol{\beta}\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right].
\end{aligned}$$

Combining this inequality with the inequality above, we have for $\boldsymbol{\beta} = \boldsymbol{\beta}^*$,

$$\begin{aligned}
& E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{X}^\top\boldsymbol{\beta}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(-\boldsymbol{X}^\top\boldsymbol{\beta}\right)\right] \\
\geq \ & E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(|\boldsymbol{P}^\top\boldsymbol{\beta}|\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(-|\boldsymbol{P}^\top\boldsymbol{\beta}|\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right)\right]
\end{aligned}$$

Notice the fact that $\boldsymbol{\beta}^*$ minimizes

$$E\left[E(\Omega_+ \mid \boldsymbol{X})\phi\left(\boldsymbol{X}^\top\boldsymbol{\beta}\right) + E(\Omega_- \mid \boldsymbol{X})\phi\left(-\boldsymbol{X}^\top\boldsymbol{\beta}\right)\right].$$

By the strict convexity of ϕ, we have $\boldsymbol{\beta}^* = |\boldsymbol{P}^\top\boldsymbol{\beta}^*|\boldsymbol{\beta}^{\text{opt}}$. This concludes the proof.

Next, we provide a proof for our remark. Notice that $\boldsymbol{\beta}^*$ solves

$$E\left[\left\{E(\Omega_+ \mid \boldsymbol{X})\phi^{'}\left(\boldsymbol{X}^\top\boldsymbol{\beta}\right) - E(\Omega_- \mid \boldsymbol{X})\phi^{'}\left(-\boldsymbol{X}^\top\boldsymbol{\beta}\right)\right\}\boldsymbol{X}\right] = 0.$$

Take the logistic loss as an example where $\phi^{'}(t) = -\exp(-t)/(1+\exp(-t))$. We will show that

$$E[\Omega_+ \mid \boldsymbol{X}]\phi^{'}\left\{\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right\} - E[\Omega_- \mid \boldsymbol{X}]\phi^{'}\left\{-\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right\} = 0.$$

From the equation in the remark, we have that

$$\begin{aligned}
& (Q(1;\boldsymbol{X}) - Q(-1;\boldsymbol{X}))/(Q(1;\boldsymbol{X}) + Q(-1;\boldsymbol{X})) \\
= \ & (\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}) - \phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}))/(\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}) + \phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})).
\end{aligned}$$

Notice that $Q(1;\boldsymbol{X}) - Q(-1;\boldsymbol{X}) = E(\Omega_+ \mid \boldsymbol{X}) - E(\Omega_- \mid \boldsymbol{X})$ and $Q(1;\boldsymbol{X}) + Q(-1;\boldsymbol{X}) = E(\Omega_+ \mid \boldsymbol{X}) + E(\Omega_- \mid \boldsymbol{X})$. As such, we have

$$E(\Omega_+ \mid \boldsymbol{X})\phi'\left(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right) - E(\Omega_- \mid \boldsymbol{X})\phi'\left(-\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right) = 0,$$

and

$$E\left[\left\{E(\Omega_+ \mid \boldsymbol{X})\phi'\left(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right) - E(\Omega_- \mid \boldsymbol{X})\phi'\left(-\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right)\right\}\boldsymbol{X}\right] = 0.$$

By the strict convexity of ϕ, we have $\boldsymbol{\beta}^* = \boldsymbol{\beta}^{\mathrm{opt}}$.

At last, we claim that when $Y(a)$'s are bounded, we can shift $Y(a)$ by a constant such that $E(|W_a| \mid \boldsymbol{X}) = Q(a;\boldsymbol{X})$. Take $a = 1$ as an example.

$$\begin{aligned}
& E(|W_1| \mid \boldsymbol{X}) \\
= \; & E\left[\pi^{-1}(1;\boldsymbol{X})\left|YI\{A=1\} - [I\{A=1\} - \pi(1;\boldsymbol{X})]\,Q(1;\boldsymbol{X})\right| \mid \boldsymbol{X}\right] \\
= \; & E\left[|(Y(1) - Q(1;\boldsymbol{X}))I\{A=1\}/\pi(1;\boldsymbol{X}) + Q(1;\boldsymbol{X})| \mid \boldsymbol{X}\right] \\
= \; & E\left[|Q(1;\boldsymbol{X})| \mid \boldsymbol{X}\right]\pi(-1;\boldsymbol{X}) + E\left[|\{Y(1) - Q(1;X)\}/\pi(1;\boldsymbol{X}) + Q(1;\boldsymbol{X})| \mid \boldsymbol{X}\right]\pi(1;\boldsymbol{X})
\end{aligned}$$

Now, we shift $Y(a)$ by a constant and notice that $Y(1) - Q(1;\boldsymbol{X})$ does not change under any constant shift. With a sufficiently large shift, we can guarantee that $Q(1;\boldsymbol{X}) \geq 0$ and $\{Y(1) - Q(1;\boldsymbol{X})\}/\pi(1;\boldsymbol{X}) + Q(1;\boldsymbol{X}) \geq 0$. Thus, with a sufficiently large shift on the outcome, we can achieve $E(|W_a| \mid \boldsymbol{X}) = Q(a;\boldsymbol{X})$. ☐

In the following, we show two non-trivial examples where Condition (a) in Lemma 1 is satisfied.

Example 1 *If X_j's are independent with mean 0, let $\mathcal{A} = \left\{j \in \mathbb{N} : \beta_j^{\mathrm{opt}} \neq 0\right\}$. Consider the following model*

$$E(Y(a) \mid \boldsymbol{X}) = f_a(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}) + g(\boldsymbol{X}_{\mathcal{A}^c}).$$

Define

$$\mathcal{G} = \{g(\boldsymbol{X}_{\mathcal{A}^c}) : E(g(\boldsymbol{X}_{\mathcal{A}^c})X_j) = 0, E(g(\boldsymbol{X}_{\mathcal{A}^c})) = 0, \forall j \in \mathcal{A}^c\}.$$

Then if $g(\boldsymbol{X}_{\mathcal{A}^c}) \in \mathcal{G}$, then Condition (a) is satisfied for the model above.

Proof. First, we have

$$E(Y(1) \mid \boldsymbol{X}) - E(Y(-1) \mid \boldsymbol{X}) = f_1(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}) - f_{-1}(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}) \in \Delta_\phi(\boldsymbol{\beta}),$$

for any $\boldsymbol{\beta}$. Next, we will show that $E(Y(1) \mid \boldsymbol{X}) + E(Y(-1) \mid \boldsymbol{X}) \in S_\phi(\boldsymbol{\beta}^*)$. Because $f_1(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}) + f_{-1}(\boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}) \in S_\phi(\boldsymbol{\beta})$ for all β, we just need to verify

$$\begin{aligned}
& E\left[g(\boldsymbol{X}_{\mathcal{A}^c})\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right] \\
- \; & E\left[g(\boldsymbol{X}_{\mathcal{A}^c}) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right]E\left[\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right] = 0.
\end{aligned}$$

Notice that $E\left[g(\boldsymbol{X}_{\mathcal{A}^c}) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right] = E\left[g(\boldsymbol{X}_{\mathcal{A}^c})\right] = 0$, we just need to show that

$$E\left[g(\boldsymbol{X}_{\mathcal{A}^c})\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\mathrm{opt}}\right] = 0.$$

Define $\mathcal{A}^* = \left\{ j \in \mathbb{N} : \beta_j^* \neq 0 \right\}$. This is equivalent to show that

$$\begin{aligned} 0 &= E\left[g(\boldsymbol{X}_{\mathcal{A}^c}) \left\{ \phi(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}} \right] \\ &= E\left[g(\boldsymbol{X}_{\mathcal{A}^c}) \left\{ \phi(\boldsymbol{X}_{\mathcal{A}^*}^\top \boldsymbol{\beta}_{\mathcal{A}^*}^*) + \phi(-\boldsymbol{X}_{\mathcal{A}^*}^\top \boldsymbol{\beta}_{\mathcal{A}^*}^*) \right\} \mid \boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^{\text{opt}} \right]. \end{aligned}$$

This is satisfied if $\mathcal{A}^* \subset \mathcal{A}$ due to $E[g(\boldsymbol{X}_{\mathcal{A}^c})] = 0$. To show this, assume W_1 and W_{-1} are positive, we consider the optimization problem

$$E[E(Y(1) \mid \boldsymbol{X})\phi(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}) + E(Y(-1) \mid \boldsymbol{X})\phi(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}})].$$

Due to the strictly convexity, we can find the unique minimizer of this optimization problem and denote it as $\boldsymbol{\beta}_{\mathcal{A}}^*$. The $\boldsymbol{\beta}_{\mathcal{A}}^*$ satisfies that

$$E[\{E(Y(1) \mid \boldsymbol{X})\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - E(Y(-1) \mid \boldsymbol{X})\phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)\}X_j],$$

for all $j \in \mathcal{A}$ due to the first-order condition. Now, we claim that $\boldsymbol{\beta}^* = (\boldsymbol{\beta}_{\mathcal{A}}^*, 0)^\top$. By first order condition, we just need to show that

$$E[\{E(Y(1) \mid \boldsymbol{X})\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - E(Y(-1) \mid \boldsymbol{X})\phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)\}X_j] = 0,$$

for all $j \in \mathcal{A}^c$. Notice that

$$\begin{aligned} & E[\{E(Y(1) \mid \boldsymbol{X})\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - E(Y(-1) \mid \boldsymbol{X})\phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)\}X_j] \\ =\; & E[\{f_1(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - f_{-1}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})\phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)\}X_j] \\ & + E[\{g(\boldsymbol{X}_{\mathcal{A}^c})\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - g(\boldsymbol{X}_{\mathcal{A}^c})\phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)\}X_j] \\ =\; & E[f_1(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - f_{-1}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})\phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)]E[X_j] \\ & + E[\phi^{'}(\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*) - \phi^{'}(-\boldsymbol{X}_{\mathcal{A}}^\top \boldsymbol{\beta}_{\mathcal{A}}^*)]E[g(\boldsymbol{X}_{\mathcal{A}^c})X_j] \\ =\; & 0. \end{aligned}$$

Thus, we have that $\boldsymbol{\beta}^* = (\boldsymbol{\beta}_{\mathcal{A}}^*, 0)^\top$, which implies that $\mathcal{A}^* \subset \mathcal{A}$. □

Example 2 *If $\boldsymbol{\beta}^{\text{opt}} = e_1$, $\boldsymbol{X} \sim N(0, \boldsymbol{I})$, and ϕ is a logistics loss, define*

$$\mathcal{A} = \left\{ j \in \mathbb{N} : \beta_j^{\text{opt}} \neq 0 \right\},$$

and

$$\mathcal{F} = \left\{ f(\boldsymbol{X}_{\mathcal{A}^c}) \in R_+ : \operatorname{cov}\left(f(\boldsymbol{X}_{\mathcal{A}^c}), \boldsymbol{X} \mid \boldsymbol{X}_1\right) = 0 \text{ and } E(f(\boldsymbol{X}_{\mathcal{A}^c})\boldsymbol{X}_{\mathcal{A}^c}) = 0 \right\}.$$

Consider the following model

$$E(Y(a) \mid \boldsymbol{X}) = f_a(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) f(\boldsymbol{X}_{\mathcal{A}^c}) + g(\boldsymbol{X}_{\mathcal{A}^c}).$$

then if $g(\boldsymbol{X}_{\mathcal{A}^c}) \in \mathcal{G}$ and $f(\boldsymbol{X}_{\mathcal{A}^c}) \in \mathcal{F}$, then Condition (a) is satisfied for the model above. Especially, all even polynomials $B_l(\boldsymbol{X}_{\mathcal{A}^c}) = \prod_{j \in \boldsymbol{X}_{\mathcal{A}^c}} X_j^{2k_j}$ belongs to $\mathcal{F}$, where $k_j \in N$.

Proof. Let

$$\begin{aligned}
& C(\boldsymbol{\beta}; f) \\
= \;& \text{cov}\left[E(Y(1) \mid \boldsymbol{X}) - E(Y(-1) \mid \boldsymbol{X}), \phi(\boldsymbol{X}^\top \boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}) \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}\right] \\
= \;& \text{cov}\left[(f_1(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) - f_{-1}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}))f(\boldsymbol{X}_{\mathcal{A}^c}), \phi(\boldsymbol{X}^\top \boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}) \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}\right] \\
= \;& (f_1(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}) - f_{-1}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}))\text{cov}\left[f(\boldsymbol{X}_{\mathcal{A}^c}), \phi(\boldsymbol{X}^\top \boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}) \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}\right].
\end{aligned}$$

We just need to show that

$$\text{cov}\left[f(\boldsymbol{X}_{\mathcal{A}^c}), \phi(\boldsymbol{X}^\top \boldsymbol{\beta}) - \phi(-\boldsymbol{X}^\top \boldsymbol{\beta}) \mid \boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}}\right] = 0.$$

The gradient of $C(\boldsymbol{\beta}; f)$ is proportional to $\text{cov}\left(f(\boldsymbol{X}_{\mathcal{A}^c}), \boldsymbol{X} \mid X_1\right)$. Because

$$\text{cov}\left(f(\boldsymbol{X}_{\mathcal{A}^c}), \boldsymbol{X} \mid X_1\right) = 0,$$

we can then conclude that $C(\boldsymbol{\beta}; f) = 0$ for all $\boldsymbol{\beta}$. Thus, we have that

$$E(Y(1) \mid \boldsymbol{X}) - E(Y(-1) \mid \boldsymbol{X}) \in \Delta_\phi(\boldsymbol{\beta}^*).$$

Next, we will verify that

$$E(Y(1) \mid \boldsymbol{X}) + E(Y(-1) \mid \boldsymbol{X}) \in S_\phi(\boldsymbol{\beta}^*).$$

Similar to Example 1, we will show that $\mathcal{A}^* \subset \mathcal{A}$. To show this, assume W_1 and W_{-1} are positive, we consider the optimization problem

$$E[E(Y(1) \mid \boldsymbol{X})\phi(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}) + E(Y(-1) \mid \boldsymbol{X})\phi(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A})].$$

Due to the strictly convexity, we can find the unique minimizer of this optimization problem and denote it as $\boldsymbol{\beta}_\mathcal{A}^*$. The $\boldsymbol{\beta}_\mathcal{A}^*$ satisfies that

$$E[\{E(Y(1) \mid \boldsymbol{X})\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - E(Y(-1) \mid \boldsymbol{X})\phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)\}X_j],$$

for all $j \in \mathcal{A}$ due to the first-order condition. Now, we claim that $\boldsymbol{\beta}^* = (\boldsymbol{\beta}_\mathcal{A}^*, 0)^\top$. By first order condition, we just need to show that

$$E[\{E(Y(1) \mid \boldsymbol{X})\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - E(Y(-1) \mid \boldsymbol{X})\phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)\}X_j] = 0,$$

for all $j \in \mathcal{A}^c$. Notice that

$$\begin{aligned}
& E[\{E(Y(1) \mid \boldsymbol{X})\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - E(Y(-1) \mid \boldsymbol{X})\phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)\}X_j] \\
= \;& E[\{f_1(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})f(\boldsymbol{X}_{\mathcal{A}^c})\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - f_{-1}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})f(\boldsymbol{X}_{\mathcal{A}^c})\phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)\}X_j] \\
& +E[\{g(\boldsymbol{X}_{\mathcal{A}^c})\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - g(\boldsymbol{X}_{\mathcal{A}^c})\phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)\}X_j] \\
= \;& E[f_1(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - f_{-1}(\boldsymbol{X}^\top \boldsymbol{\beta}^{\text{opt}})\phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)]E[f(\boldsymbol{X}_{\mathcal{A}^c})X_j] \\
& +E[\phi^{'}(\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*) - \phi^{'}(-\boldsymbol{X}_\mathcal{A}^\top \boldsymbol{\beta}_\mathcal{A}^*)]E[g(\boldsymbol{X}_{\mathcal{A}^c})X_j] \\
= \;& 0.
\end{aligned}$$

Thus, we have that $\boldsymbol{\beta}^* = (\boldsymbol{\beta}^*_{\mathcal{A}}, 0)^\top$, which implies that $\mathcal{A}^* \subset \mathcal{A}$. Thus, we have that

$$\begin{aligned} & E\left[g(\boldsymbol{X}_{\mathcal{A}^c})\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right] \\ - \quad & E(g(\boldsymbol{X}_{\mathcal{A}^c}) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}})E(\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}) = 0, \end{aligned}$$

and

$$\begin{aligned} & E\left[(f_1(\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}) + f_{-1}(\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}))f(\boldsymbol{X}_{\mathcal{A}^c})\left\{\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}\right] \\ - \quad & E[(f_1(\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}) + f_{-1}(\boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}))f(\boldsymbol{X}_{\mathcal{A}^c}) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}] \\ \times \quad & E(\phi(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \phi(-\boldsymbol{X}^\top\boldsymbol{\beta}^*) \mid \boldsymbol{X}^\top\boldsymbol{\beta}^{\text{opt}}) = 0. \end{aligned}$$

Thus, we have that

$$E(Y(1) \mid \boldsymbol{X}) + E(Y(-1) \mid \boldsymbol{X}) \in S_\phi(\boldsymbol{\beta}^*).$$

At the end, we will verify that $\text{cov}\,(B_l(\boldsymbol{X}_{\mathcal{A}^c}), \boldsymbol{X} \mid \boldsymbol{X}_1) = 0$ and $E(f(\boldsymbol{X}_{\mathcal{A}^c})\boldsymbol{X}_{\mathcal{A}^c}) = 0$. When $j \neq 1$, we have

$$\text{cov}\,(B_l(\boldsymbol{X}_{\mathcal{A}^c}), X_j \mid \boldsymbol{X}_1) = E(B_l(\boldsymbol{X}_{\mathcal{A}^c})X_j \mid \boldsymbol{X}_1) = 0,$$

by $E(X_j \mid X_1) = E(X_j) = 0$ and the distribution of X_j is symmetric. When $j = 1$, we also have that $\text{cov}\,(B_l(\boldsymbol{X}_{\mathcal{A}^c}), X_1 \mid X_1) = 0$ by the definition of conditional expectation. This concludes that $B_l(\boldsymbol{X}_{\mathcal{A}^c}) = \prod_{j\in\mathcal{A}^c} X_j^{2k_j} \in \mathcal{F}$. □

Appendix C.

In Appendix C, we compared the list of significant covariates selected by the proposed method and Q-learning. The proposed approach identified most of the covariates selected by Q-learning. In addition, the proposed approach also identifies 10 new driving factors, which provide additional insights for further investigations. The figure 12 shows the Venn plot of the selected covariates. The 95%-confidence interval of the selected covariates by the proposed method is reported in Table 3.

Appendix D.

In Appendix D, we study the limiting property of the proposed method. We show the proof of Theorems 3 - 7. In addition, we also show the validity of the Algorithm 4, where the nuisance parameters are estimated nonparametrically and there is no sample splitting procedure.

Theorem 8 shows that Algorithm 2, where there is no sample-splitting procedure, is valid when the nuisance parameters are estimated parametrically.

Theorem 8 *Assume that $\boldsymbol{X}$ is bounded. Suppose that $\pi(a; \boldsymbol{X})$ and $Q(a; \boldsymbol{X})$ are known to follow parametric models $\pi(a; \boldsymbol{X}, \boldsymbol{\beta}_\pi)$ and $Q(a; \boldsymbol{X}, \boldsymbol{\beta}_Q)$ with true parameters $\boldsymbol{\beta}^*_\pi$ and $\boldsymbol{\beta}^*_Q$ respectively. Assume $\pi(a; \boldsymbol{X}, \boldsymbol{\beta}_\pi)$ and $Q(a; \boldsymbol{X}, \boldsymbol{\beta}_Q)$ are second order continuously differentiable, and $\|\nabla_{\boldsymbol{\beta}_\pi}\pi(a; \boldsymbol{X}, \boldsymbol{\beta}^*_\pi)\|_\infty$ and $\left\|\nabla_{\boldsymbol{\beta}_Q}Q(a; \boldsymbol{X}, \boldsymbol{\beta}^*_Q)\right\|_\infty$ are bounded. Further, there*

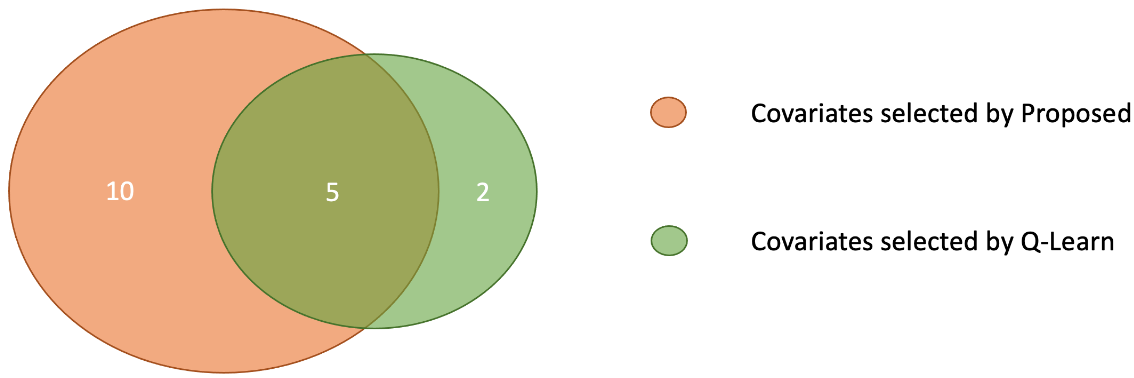

Figure 12: Variable selection using the proposed method and Q-learning.

Algorithm 4: Inference of $\boldsymbol{\beta}^*$ with parametric propensity and outcome model estimations

---

Input: n samples.

Output: $\widehat{\beta}$ and a p-value for $\mathcal{H}_0 : \boldsymbol{\beta}_1^* = 0$.

Use all data to fit a parametric regression model with a lasso penalty and obtain an estimator $\widehat{\pi}$ for the propensity and an estimator $\widehat{Q}$ for the outcome model;

Obtain the proposed estimator $\widehat{\beta}$ by $\min_\beta E_n\left[l_\phi\left(\beta;\widehat{\Omega}_+,\widehat{\Omega}_-\right)\right]+\lambda_n\|\beta\|_1$, where $\widehat{\Omega}_+$ and $\widehat{\Omega}_-$ are computed with $\widehat{\pi}$ and $\widehat{Q}$ plugged in, and λ_n is tuned by cross-validation;

Obtain an estimator $\widehat{w}$ for w^* by $\min_w E_n\left[\nabla^2 l_\phi\left(\widehat{\beta};\widehat{\Omega}_+,\widehat{\Omega}_-\right)\left(X_1-X_{-1}^\top w\right)^2\right]+\widetilde{\lambda}_n\|w\|_1$, where $\widetilde{\lambda}_n$ is tuned by cross-validation;

Let $\left(\widehat{\beta}_{\text{null}}\right)^\top=\left(0,\left(\widehat{\beta}_{-1}\right)^\top\right)$, where $\widehat{\beta}_{-1}$ is a $p-1$ dimensional sub-vector of $\widehat{\beta}$ without $\widehat{\beta}_1$. Construct the de-correlated score test statistic $S(\widehat{\beta}_{\text{null}},\widehat{w})$ as $S\left(\widehat{\beta}_{\text{null}},\widehat{w}\right)=E_n\left[\nabla l_\phi\left(\widehat{\beta}_{\text{null}};\widehat{\Omega}_+,\widehat{\Omega}_-\right)\left(X_1-X_{-1}^\top\widehat{w}\right)\right]$, and the estimator of the variance $\widehat{\sigma}^2=E_n\left[\left\{\nabla l_\phi\left(\widehat{\beta}_{\text{null}};\widehat{\Omega}_+,\widehat{\Omega}_-\right)\right\}^2\left(X_1-X_{-1}^\top\widehat{w}\right)^2\right]$;

Calculate the p-value by $2\left(1-\Phi(|S|/\widehat{\sigma})\right)$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

	Coef	95% - CI
Chronic Complications : Fluid and Electrolyte Disorders	-0.024	[-0.047,-0.001]
Chronic Complications : African American	-0.027	[-0.052,-0.001]
Alcohol Abuse : Entitlement Disability	-0.054	[-0.104,-0.004]
HCC Community Score : Special Chronic Conditions	-0.022	[-0.042,-0.002]
Hypertension : Lower Extremity Ulcer	-0.036	[-0.068,-0.005]
HbA1c at Baseline : African American	0.019	[0.003,0.036]
Entitlement Disability : Hypothyroidism	-0.024	[-0.045,-0.003]
Cardiac Heart Failure : Peripheral Vascular Disease	-0.029	[-0.057, -0.001]
Chronic Kidney Disease : HbA1c at Baseline	0.081	[0.014, 0.149]
Other Race : Special Chronic Conditions	0.016	[0.003,0.029]
Liver Disease : Weight Loss	0.015	[0.003,0.027]
Other Neurological Disorders : Female	-0.021	[-0.038,-0.005]
Lower Extremity Ulcer : HbA1c at Baseline	0.039	[0.010,0.069]
Chronic Complications : Bucketized Age	0.040	[0.016,0.063]
HbA1c at Baseline : Female	0.044	[0.028,0.061]

Table 3: Coefficients and CI for the significant covariates of the estimated decision rule. Special chronic conditions refer to chronic conditions including amputation, chronic blood loss, drug abuse, lymphoma, metastatistic cancer, and peptic ulcer disease. Bucketized age refers to a variable created by bucketizing the raw age by its observed quartiles. Other Race refers to the race excluding White and Black.

exist constants C_π and C_Q such that $\nabla^2_{\boldsymbol{\beta}_\pi}\pi(a;\boldsymbol{X},\boldsymbol{\beta}_\pi) \prec C_\pi \boldsymbol{X}\boldsymbol{X}^\top$ and $\nabla^2_{\boldsymbol{\beta}_Q} Q(a;\boldsymbol{X},\boldsymbol{\beta}_Q) \prec C_Q\boldsymbol{X}\boldsymbol{X}^\top$, where for two matrices A and B, $A \prec B$ implies that $B-A$ is positive semi-definite. In addition, suppose that $\|\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*_\pi\|_1 = O_p(n^{-\alpha})$ and $\|\widehat{\boldsymbol{\beta}}_Q - \boldsymbol{\beta}^*_Q\|_1 = O_p(n^{-\beta})$ for some $\alpha,\beta > 0$, we require that $\alpha+\beta > 1/2$. In addition, we require that

$$R\max\{s^*, s'\}\log n(\log p)^{3/2} = o(n^{1/2})$$

and

$$(n^{-\alpha}+n^{-\beta})R \to 0,$$

where $s^* = \|\boldsymbol{\beta}^*\|_0$ and $s' = \max_j \|\boldsymbol{w}^*_j\|_0$.

Assume that Conditions (C1)-(C4) hold. For Algorithm 4, under the null hypothesis $H_0 : \beta^*_j = 0$, by choosing $\lambda_n \asymp \tilde{\lambda}_n \asymp (\log p/n)^{1/2}$, we have

$$n^{1/2}S_j \to N(0,\sigma^2),$$

and $\widehat{\sigma}^2_j \to \sigma^2_j$, where $\widehat{\sigma}^2_j$ is given in Algorithm 2, and $\sigma^2_j = \left(\boldsymbol{\nu}^*_j\right)^\top \operatorname{var}\left[\nabla^2 l_\phi(\boldsymbol{\beta}^*;\Omega_+,\Omega_-)\right]\boldsymbol{\nu}^*_j$.

For the value inference, we also have the following theorem.

Theorem 9 *Assume that* Y *is bounded and denote the sample size of* $\widetilde{I}_1$ *as* n_1 *and* $\widetilde{I}_2$ *as* n_2*. In addition to the conditions in Theorem 3, we further assume* $n_1^{-\alpha-\beta}n_2^{1/2} = o(1)$ *and one of the following conditions: 1) Conditions (C6) and (C7) holds with* $\left(s(\log p/n_1)^{1/2}\right)^{\zeta+\gamma} = o_p(n_2^{-1/2})$*; 2) Condition (C7) holds with* $P\left(\left|\boldsymbol{X}^\top\boldsymbol{\beta}^*\right| = 0\right) = 0$ *and* $\left(s(\log p/n_1)^{1/2}\right)^{\gamma} = o_p(n_2^{-1/2})$*, then we have*

$$n_2{}^{1/2}\sigma_V^{-2}(\widehat{V}(\widehat{D}) - V(D^*)) \to N(0,1),$$

where $\sigma_V^2 = \operatorname{var}\left[W_{\widehat{D}(\boldsymbol{X})}(Y, \boldsymbol{X}, A, \pi, Q)\right]$.

Under Conditions (C6) and (C7) with $\left(s(\log p/n_1)^{1/2}\right)^{\zeta+\gamma} = o_p(n_2^{-1/2})$, Theorem 7 holds for both regular and non-regular cases. Condition (C6) implicitly assumes that $\boldsymbol{\beta}^*$ corresponds to the optimal individualized treatment rule. When Condition (C6) fails, the inference of the value under $D^*(\boldsymbol{X})$ is challenging but possible if Condition (C7) holds with $P\left(\left|\boldsymbol{X}^\top\boldsymbol{\beta}^*\right| = 0\right) = 0$ and $\left(s(\log p/n_1)^{1/2}\right)^{\gamma} = o_p(n_2^{-1/2})$.

Proof of Theorems and Corollary.

The proof of Thoerem 3 can be found for the proof of Lemma 10. The following provides the proof of Theorem 4.

Proof of Thoerem 4. Let's compute the following

$$
\begin{aligned}
& (n/K)^{1/2} S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}, \widehat{\boldsymbol{w}}_j^{(k)}\right) && (10)\\
= \;& (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}})\right\}(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)})\right] && (11)\\
= \;& (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}})\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] && (12)\\
& + (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}})\right\}(\boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\\
= \;& (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\\
& + (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)\right.\\
& \left.(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\\
& + (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\\
& + o_p(1), && (13)
\end{aligned}
$$

uniformly over $j \in \{1, \cdots, J\}$. The first line from (10) to (11) is the definition of

$$S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}, \widehat{\boldsymbol{w}}_j^{(k)}\right).$$

From Line (12) to (13), we use that

$$
\begin{aligned}
& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
= \;& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
& + E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null,j}}) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null,j}})\right\}\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)\right. \\
& \left.(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
= \;& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
& + E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
& + E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\left[\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \phi''(\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}})\right]\right.\right. \\
& \left.\left. + \widehat{\Omega}_-^{(k)}\left[\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null,j}})\right]\right\}\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right],
\end{aligned}
$$

$$
\begin{aligned}
& \left|E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\left[\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \phi''(\boldsymbol{X}^\top\beta_{\text{null}})\right] + \widehat{\Omega}_-^{(k)}\left[\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null}})\right]\right\}\right.\right. \\
& \left.\left.\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| \\
\leq \;& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\left|\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \phi''(\boldsymbol{X}^\top\beta_{\text{null}})\right| + \widehat{\Omega}_-^{(k)}\left|\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \phi''(-\boldsymbol{X}^\top\beta_{\text{null}})\right|\right\}\right. \\
& \left.|\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)||X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*|\right] \\
\leq \;& CE_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}|\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)|^2|X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*|\right].
\end{aligned}
$$

By the sub-gaussian of $X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*$ and $\sup_j\|\boldsymbol{w}_j^*\|_1 \leq R$, there exists a constant σ_x such that

$$
P\left\{\max_{j=1,\cdots,p}\max_{1\leq i\leq n}|X_{i,j} - \boldsymbol{X}_{i,-j}^\top\boldsymbol{w}_j^*| \geq 2\sigma_x R\sqrt{\log(np)}\right\} \leq 2\exp\{-\log(np)\}.
$$

One the event that $\max_{j=1,\cdots,p}\max_{1\leq i\leq n}|X_{i,j} - X_{i,-j}^\top\boldsymbol{w}_j^*| \leq 2\sigma_w R\sqrt{\log(np)}$, we have

$$
\begin{aligned}
& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}|\boldsymbol{X}_{-j}^\top(\widehat{\beta}_{-1}^{(k)} - \beta_{-1}^*)|^2|X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*|\right] \\
\leq \;& 2\sigma_w R\sqrt{\log(np)}E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}|\boldsymbol{X}_{-j}^\top(\widehat{\beta}_{-1}^{(k)} - \beta_{-1}^*)|^2\right].
\end{aligned}
$$

By Lemma 10, we have that

$$
\begin{aligned}
& \sup_{j=1,\cdots,p} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}|\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)|^2\right] \\
= \;& O_p(s^*\log p/n).
\end{aligned}
$$

Thus, we have that

$$
\begin{aligned}
& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}|\boldsymbol{X}_{-j}^\top(\widehat{\beta}_{-1}^{(k)} - \beta_{-1}^*)|^2|X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*|\right] \\
= \;& O_p(R\sqrt{\log(np)}s^*\log p/n)
\end{aligned}
$$

Under the condition $Rs^*(\log p)^{3/2}/\sqrt{n} \to 0$, we have Line (12) to (13).

Now, we focus on a bound for (13). For the first term in (13), by Claim 13, we have that

$$\begin{aligned}
& n^{1/2} E_n^{(k)} \left[\left\{ \widehat{\Omega}_+^{(k)} \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)} \phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \right] \\
= \;& n^{1/2} E_n^{(k)} \left[\left\{ \Omega_+ \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \Omega_- \phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \right] \\
& + n^{1/2} E_n^{(k)} \left[\left\{ \left[\widehat{\Omega}_+^{(k)} - \Omega_+ \right] \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \left[\widehat{\Omega}_-^{(k)} - \Omega_- \right] \phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \right] \\
= \;& n^{1/2} E_n^{(k)} \left[\left\{ \Omega_+ \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \Omega_- \phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \right] + \\
& O_p \left(R n^{-\alpha-\beta} + R(n^{-\alpha} + n^{-\beta}) \sqrt{\log p / n} \right),
\end{aligned}$$

uniformaly holds in j. By the conditions, we have

$$R n^{-\alpha-\beta+1/2} \to 0, \quad R(n^{-\alpha} + n^{-\beta}) \sqrt{\log p} \to 0.$$

Thus, we have that the first term in (13) is equivalent to

$$(n/K)^{1/2} E_n^{(k)} \left[\left\{ \Omega_+ \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \Omega_- \phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \right]$$

uniformly.

The second term in (13) can be bounded by the following

$$\begin{aligned}
& \max_{j \in \mathcal{J}} \left| (n/K)^{1/2} E_n^{(k)} \left[\left\{ \widehat{\Omega}_+^{(k)} \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)} \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} \right.\right. \\
& \left.\left. \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*)(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \right] \right| \\
\leq \;& (n/K)^{1/2} \max_{j \in \mathcal{J}} \left\| E_n^{(k)} \left[\left\{ \widehat{\Omega}_+^{(k)} \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)} \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \right] \right\|_\infty \\
& \left\| \widehat{\boldsymbol{\beta}}^{(k)} - \boldsymbol{\beta}^* \right\|_1 \\
\leq \;& (n/K)^{1/2} \max_{j \in \mathcal{J}} \left\| E_n^{(k)} \left[\left\{ \Omega_+ \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \Omega_- \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \right] \right\|_\infty \\
& \left\| \widehat{\boldsymbol{\beta}}^{(k)} - \boldsymbol{\beta}^* \right\|_1 + O_p \left(\left(R n^{-\alpha-\beta+1/2} + R(n^{-\alpha} + n^{-\beta}) \sqrt{\log p} \right) s^* (\log p / n)^{1/2} \right).
\end{aligned}$$

The second equality comes from Claim 13 and Lemma 10. To bound

$$\left\| E_n^{(k)} \left[\left\{ \Omega_+ \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \Omega_- \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \right] \right\|_\infty,$$

we consider the following decomposition

$$\begin{aligned}
& \left\| E_n^{(k)} \left[\left\{ \Omega_+ \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \Omega_- \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \right] \right\|_\infty \\
\leq \;& \left\| E_n^{(k)} \left[\left\{ \frac{1\{A = 1\}}{\pi_1} (Y - Q(1; \boldsymbol{X})) \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \right.\right.\right. \\
& \left.\left.\left. \frac{1\{A = -1\}}{\pi_{-1}} (Y - Q(-1; \boldsymbol{X})) \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \right] \right\|_\infty \\
& + \left\| E_n^{(k)} \left[\left\{ Q(1; \boldsymbol{X}) \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + Q(1; \boldsymbol{X}) \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \right] \right\|_\infty .
\end{aligned}$$

By the condition on $Y - Q(a;\boldsymbol{X})$ on $A = a$, by the proof of Claim 13, the first term can be bounded

$$\begin{aligned} & n^{1/2} \max_{j\in\mathcal{J}} \Bigg\| E_n^{(k)} \Bigg[\Bigg\{ \frac{1\{A=1\}}{\pi_1} \left(Y - Q(1;\boldsymbol{X})\right) \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \\ & \frac{1\{A=-1\}}{\pi_{-1}} \left(Y - Q(-1;\boldsymbol{X})\right) \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Bigg\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \Bigg] \Bigg\|_\infty \\ = \; & O_p(R\sqrt{\log p/n}). \end{aligned}$$

For the second term, by the boundedness of $Q(a;\boldsymbol{X})$'s and $\phi''(\cdot)$, by Lemma 14 in Loh and Wainwright (2015), if $\log p = O(n)$, we have

$$\begin{aligned} & \max_{j\in\mathcal{J}} \Big\| E_n^{(k)} \Big[\Big\{ Q(1;\boldsymbol{X})\phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + Q(1;\boldsymbol{X})\phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Big\} (X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \boldsymbol{X}_{-j} \Big] \Big\|_\infty \\ = \; & O_p(R\sqrt{\log p/n}). \end{aligned}$$

Because our condition that

$$Rn^{-1/2} s^* \log p \to 0, \quad Rn^{-\alpha-\beta+1/2} \to 0, \quad \text{and } R(n^{-\alpha} + n^{-\beta})\sqrt{\log p} \to 0,$$

the second term in (13) is $o_p(1)$.

The third term in (13) can also be bounded following a similar steps to the first term in (13).

$$\begin{aligned} & \Big| (n/K)^{1/2} E_n^{(k)} \Big[\Big\{ \widehat{\Omega}_+^{(k)} \phi'(\boldsymbol{X}^\top \widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}) - \widehat{\Omega}_-^{(k)} \phi'(-\boldsymbol{X}^\top \widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}) \Big\} (\boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \Big] \Big| \\ \le \; & \Big| (n/K)^{1/2} E_n^{(k)} \Big[\Big\{ \widehat{\Omega}_+^{(k)} \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)} \phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Big\} (\boldsymbol{X}_{-j}^\top \widehat{\boldsymbol{w}}^{(k)} - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*) \Big] \Big| \\ & + \Big| (n/K)^{1/2} E_n^{(k)} \Big[\Big\{ \widehat{\Omega}_+^{(k)} \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)} \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Big\} \\ & \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*) \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*) \Big] \Big| + o_p(1), \end{aligned}$$

uniformly in j. To bound the second term above, by Lemma 10 and 11, we have that

$$\begin{aligned} & \max_{j\in\mathcal{J}} \Big| (n/K)^{1/2} E_n^{(k)} \Big[\Big\{ \widehat{\Omega}_+^{(k)} \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)} \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Big\} \\ & \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*) \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*) \Big] \Big| \\ \le \; & (n/K)^{1/2} \Bigg\{ \max_{j\in\mathcal{J}} E_n^{(k)} \Bigg[\Big\{ \widehat{\Omega}_+^{(k)} \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)} \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Big\} \Big| \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*) \Big|^2 \Bigg] \\ & \max_{j\in\mathcal{J}} E_n^{(k)} \Bigg[\Big\{ \widehat{\Omega}_+^{(k)} \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)} \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \Big\} \Big| \boldsymbol{X}_{-j}^\top (\widehat{\boldsymbol{\beta}}_{-j}^{(k)} - \boldsymbol{\beta}_{-j}^*) \Big|^2 \Bigg] \Bigg\}^{1/2} \\ = \; & o_p(1). \end{aligned}$$

To bound the first term, we have

$$\begin{aligned}
& \max_{j\in\mathcal{J}} \left|(n/K)^{1/2} E_n^{(k)} \left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(\boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| \\
\leq\ & \max_{j\in\mathcal{J}} \left\|(n/K)^{1/2} E_n^{(k)} \left[\left\{\widehat{\Omega}_+^{(k)}\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-^{(k)}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}_{-j}\right]\right\|_\infty \left\|\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*)\right\|_1 \\
\leq\ & \left\|(n/K)^{1/2} E_n^{(k)} \left[\left\{\Omega_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \Omega_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right]\right\|_\infty \max_{j\in\mathcal{J}}\left\|\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*)\right\|_1 \\
& + \left\|(n/K)^{1/2} E_n^{(k)} \left[\left\{\left(\widehat{\Omega}_+^{(k)} - \Omega_+\right)\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \left(\widehat{\Omega}_-^{(k)} - \Omega_-\right)\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right]\right\|_\infty \\
& \max_{j\in\mathcal{J}}\left\|\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*)\right\|_1
\end{aligned}$$

For the second term, by Lemma 11 and Claim 13, we know that the second term is negligible. For the first term, we decompose it into two terms, i.e.,

$$\begin{aligned}
& n^{1/2}E_n^{(k)}\left[\left\{\Omega_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \Omega_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right] \\
=\ & n^{1/2}E_n^{(k)}\left[\left\{\frac{1\{A=1\}}{\pi_1}(Y - Q(1;\boldsymbol{X}))\,\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right.\right. \\
& \left.\left. -\frac{1\{A=-1\}}{\pi_{-1}}(Y - Q(-1;\boldsymbol{X}))\,\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right] \\
& + n^{1/2}E_n^{(k)}\left[\left\{Q(1;\boldsymbol{X})\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - Q(-1;\boldsymbol{X})\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right].
\end{aligned}$$

Because $Q(a;\boldsymbol{X})$'s are bounded and $\phi^{'}(\cdot)$ is bounded, by the sub-Gaussian condition on $\boldsymbol{X}$, under the null hypothesis, we have that

$$\begin{aligned}
& \max_{j\in\mathcal{J}}\left\|E_n^{(k)}\left[\left\{Q(1;\boldsymbol{X})\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - Q(-1;\boldsymbol{X})\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right]\right\|_\infty \\
=\ & O_p(\sqrt{\log p/n}).
\end{aligned}$$

By the condition of $Y - Q(a;\boldsymbol{X})$ on $A = 1$ and the sub-Gaussian condition on $\boldsymbol{X}$, by the proof of Claim 13, if $\log p = O(n)$, we have

$$\begin{aligned}
& \max_{j\in\mathcal{J}}\left|E_n^{(k)}\left[\frac{1\{A=1\}}{\pi_1}(Y - Q(1;\boldsymbol{X}))\,\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| \\
=\ & O_p(\sqrt{\log p/n}).
\end{aligned}$$

Similarly, we can show that

$$\begin{aligned}
& \max_{j\in\mathcal{J}}\left\|E_n^{(k)}\left[\frac{1\{A=-1\}}{\pi_{-1}}(Y - Q(-1;\boldsymbol{X}))\,\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\boldsymbol{X}\right]\right\|_\infty \\
=\ & O_p(\sqrt{\log p/n}).
\end{aligned}$$

Thus, the first term is $O_p(n^{1/2}\sqrt{\log p/n}\max\{s^*, s'\}\sqrt{\log p/n})$, which is negligiable.

In conclusion, we have that

$$\begin{aligned}
& \max_{j\in\mathcal{J}}\left|(n/K)^{1/2}S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right) - (n/K)^{1/2}E_n^{(k)}\left[\left\{\Omega_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right.\right.\right. \\
& \left.\left.\left. -\Omega_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| = o_p(1).
\end{aligned}$$

Define

$$\sigma_j^2 = E\left[\left\{\Omega_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}) - \Omega_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta})\right\}^2 (X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2\right].$$

Applying the Berry-Esseen bound for CLT, there exists some universal constant such that

$$\begin{aligned}\max_{j\in\mathcal{J}}\sup_{\alpha\in(0,1)}&\Big|P\Big(\Big|\sigma_j^{-1}(n/K)^{1/2}E_n^{(k)}\Big[\Big\{\Omega_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\\ &-\Omega_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\Big\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\Big]\Big| \le \Phi^{-1}(1-\alpha/2)\Big) - (1-\alpha)\Big|\\ \le\quad & \frac{c_0}{\sqrt{n}}\max_j E[|M_j|^3],\end{aligned}$$

where

$$M_j = \left\{\Omega_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \Omega_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*).$$

From the sub-Gaussian conditions, we have that $\max_j E[|M_j|^3] = O(R^3)$. Thus, if $R^3/\sqrt{n} \to 0$, we have that $\frac{c_0}{\sqrt{n}}\max_j E[|M_j|^3] \to 0$.

Combining the inequalities above, we have

$$\max_{j\in\mathcal{J}}\sup_{\alpha\in(0,1)}\left|P\left(\left|\sigma_j^{-1}(n/K)^{1/2}S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}, \widehat{\boldsymbol{w}}_j^{(k)}\right)\right| \le \Phi^{-1}(1-\alpha/2)\right) - (1-\alpha)\right| = o_p(1).$$

Averaging all the k's, we can conclude the proof.

□

Proof of Theorem 8. We can observe that if Claim 13 holds with

$$n^{1/2}\left\|E_n\left[\left(\widehat{\Omega}_a^{(k)} - \Omega_a^{(k)}\right)h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty = o_p(1),$$

then the proof of Theorem 4 is applicable to Algorithm 2. However, the proof of Claim 13 uses the fact that the data used to train $\widehat{\pi}^{(-k)}$ and $\widehat{Q}^{(-k)}$ is independent with those used to fit the proposed method and form the score test statistic. Here, we provide a separate proof of Claim 13 utilizing the parametric structure. Without loss of generality, we assume that π is estimated parametrically. To start with, denote the parametric model of π as $\pi(A;\boldsymbol{X},\boldsymbol{\beta})$. Under this notation, we rewrite that $\widehat{\pi}(A;\boldsymbol{X}) = \pi\left(A;\boldsymbol{X},\widehat{\boldsymbol{\beta}}_\pi\right)$ and $\pi(A;\boldsymbol{X}) = \pi\left(A;\boldsymbol{X},\boldsymbol{\beta}_\pi^*\right)$.

We want to show that for any function h with $\|h\|_\infty \le C$, we have for $a = 1$ or -1,

$$\left\|E_n\left[\left(\widehat{\Omega}_a^{(k)} - \Omega_a^{(k)}\right)h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty = O_p\left(n^{-\alpha-\beta} + (n^{-\alpha} + n^{-\beta})(\log p/n)^{1/2}\right).$$

We consider $a = 1$. Note that

$$\begin{aligned}&\left\|E_n\left[\left(\widehat{\Omega}_a^{(k)} - \Omega_a\right)h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty\\ \le&\left\|E_n\left[I\{A=1\}\left(\widehat{\pi}^{-1}(1;\boldsymbol{X}) - \pi^{-1}(1;\boldsymbol{X})\right)\left(\widehat{Q}^{(-k)}(1;\boldsymbol{X}) - Q(1;\boldsymbol{X})\right)h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty\\ &+\left\|E_n\left[I\{A=1\}\left(\widehat{\pi}^{-1}(1;\boldsymbol{X}) - \pi^{-1}(1;\boldsymbol{X})\right)(Y_1 - Q(1;\boldsymbol{X}))h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty\\ &+\left\|E_n\left[(I\{A=1\}/\pi(1;\boldsymbol{X}) - 1)\left(\widehat{Q}^{(-k)}(1;\boldsymbol{X}) - Q(1;\boldsymbol{X})\right)h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty\\ =&I_1 + I_2 + I_3\end{aligned}$$

For I_1, we can show that $I_1 \lesssim n^{-\beta} \left\|\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*_\pi\right\|_1$ following the proof of Claim 13 and $\left(\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*_\pi\right)^\top E_n[\boldsymbol{X}\boldsymbol{X}^\top]\left(\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*_\pi\right) \lesssim n^{-1/2}$. Likewise, from the proof of Claim 13, we can conclude that $I_3 \lesssim O_p\left(n^{-\beta}(\log p/n)^{1/2}\right)$. For I_2, we have

$$\begin{aligned}
& \left\|E_n\left[I\{A=1\}\left(\widehat{\pi}^{-1}(1;\boldsymbol{X}) - \pi^{1/2}(1;\boldsymbol{X})\right)(Y_1 - Q(1;\boldsymbol{X}))\,h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty \\
= & \left\|E_n\left[I\{A=1\}\left(\pi^{-1}\left(A;\boldsymbol{X},\widehat{\boldsymbol{\beta}}_\pi\right) - \pi^{-1}\left(A;\boldsymbol{X},\boldsymbol{\beta}^*_\pi\right)\right)(Y_1 - Q(1;\boldsymbol{X}))\,h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty \\
\lesssim & \left\|E_n\left[-I\{A=1\}\pi^{-2}\left(A;\boldsymbol{X},\boldsymbol{\beta}^*_\pi\right)\nabla_{\boldsymbol{\beta}}\pi(A;\boldsymbol{X},\boldsymbol{\beta}^*)\left(\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*\right)^\top (Y_1 - Q(1;\boldsymbol{X}))\right.\right. \\
& \left.\left. h(\boldsymbol{X})\boldsymbol{X}\right]\right\|_\infty \\
\leq & \left\|E_n\left[-\boldsymbol{X}I\{A=1\}\pi^{-2}\left(A;\boldsymbol{X},\boldsymbol{\beta}^*_\pi\right)\{\nabla_{\boldsymbol{\beta}}\pi(A;\boldsymbol{X},\boldsymbol{\beta}^*)\}^\top (Y_1 - Q(1;\boldsymbol{X}))\,h(\boldsymbol{X})\right]\right\|_\infty \\
& \times \left\|\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*\right\|_1 \\
= & \left\|\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*\right\|_1 (\log p/n)^{1/2}
\end{aligned}$$

When π is estimated by linear or logistic regression with lasso penalty, we have

$$\left\|\widehat{\boldsymbol{\beta}}_\pi - \boldsymbol{\beta}^*\right\|_1 \lesssim n^{-\alpha}.$$

Thus, the claim holds. □

Proof of Theorem 5. First, we will show that

$$(n/K)^{1/2}\max_j\left|\left(\widetilde{\beta}_j^{(k)} - \beta_j^*\right)\widehat{I}^{(k)}_{j|-j} + E_n^{(k)}\left[\{\nabla l_\phi\left(\boldsymbol{\beta}^*;\Omega_+,\Omega_-\right)\}\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)\right]\right| = o_p(1).$$

By the definition of $\widetilde{\beta}_j^{(k)}$,

$$\begin{aligned}
& (n/K)^{1/2}\left|\left(\widetilde{\beta}_j^{(k)} - \beta_j^*\right)\widehat{I}^{(k)}_{j|-j} + E_n^{(k)}\left[\{\nabla l_\phi\left(\boldsymbol{\beta}^*;\Omega_+,\Omega_-\right)\}\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)\right]\right| \\
= & (n/K)^{1/2}\left|\left(\widetilde{\beta}_j^{(k)} - \beta_j^*\right)\widehat{I}^{(k)}_{j|-j} - \left\{S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)},\widehat{\boldsymbol{w}}_j^{(k)}\right) - E_n^{(k)}\left[\{\nabla l_\phi\left(\boldsymbol{\beta}^*;\Omega_+,\Omega_-\right)\}\right.\right.\right. \\
& \left.\left.\left.\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)\right]\right\}\right| \\
\leq & (n/K)^{1/2}\left|\left(\widetilde{\beta}_j^{(k)} - \beta_j^*\right)\widehat{I}^{(k)}_{j|-j} - \left\{S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)},\widehat{\boldsymbol{w}}_j^{(k)}\right) - S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}},\widehat{\boldsymbol{w}}_j^{(k)}\right)\right\}\right| \\
& (n/K)^{1/2}\left|S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}},\widehat{\boldsymbol{w}}_j^{(k)}\right) - E_n^{(k)}\left[\{\nabla l_\phi\left(\boldsymbol{\beta}^*;\Omega_+,\Omega_-\right)\}\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)\right]\right| \\
\leq & I_{1j} + I_{2j},
\end{aligned}$$

where $\widehat{\boldsymbol{\beta}}^{(k)}_{\text{null(j)}}$ equals to $\widehat{\boldsymbol{\beta}}^{(k)}$ except its jth coefficient replaced by β_j^*.

By the proof of Theorem 4, we have

$$\max_j I_{2j} = o_p(1).$$

To bound I_{1j} uniformly, we consider

$$
\begin{aligned}
& (n/K)^{1/2} S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right) - S_j^{(k)}\left(\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}, \widehat{\boldsymbol{w}}_j^{(k)}\right) \\
= \; & (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
& + (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}}) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}})\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)\right. \\
& \left.\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol{w}}_j^{(k)} - \boldsymbol{w}_j^*)\right] \\
& + (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\left\{\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}}) - \phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} + \right.\right. \\
& \left.\left.\widehat{\Omega}_-^{(k)}\left\{\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}}) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*\right],
\end{aligned}
$$

where $\boldsymbol{\beta}_{\text{null(j)}}$ is in between $\widehat{\boldsymbol{\beta}}_{\text{null(j)}}^{(k)}$ and $\widehat{\boldsymbol{\beta}}_j^{(k)}$. By Claim 13 and the proof of Theorem 4, we can see that

$$
\begin{aligned}
& (n/K)^{1/2} E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
= \; & (n/K)^{1/2}\left(\widehat{\beta}_j^{(k)} - \beta_j^*\right) I_{j|-j} + o_p(1),
\end{aligned}
$$

uniformly holds over $j = 1, \cdots, J$. For the third term, we have

$$
\begin{aligned}
& (n/K)^{1/2}\left|E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\left\{\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}}) - \phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} + \right.\right.\right. \\
& \left.\left.\left.\widehat{\Omega}_-^{(k)}\left\{\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}_{\text{null(j)}}) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| \\
\leq \; & (n/K)^{1/2} C E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\right. \\
& \left.\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}^{(k)} - \boldsymbol{\beta}^*) X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right].
\end{aligned}
$$

By the proof of Thoerem 4, we can show that

$$
\begin{aligned}
& \max j E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\right. \\
& \left.\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}^{(k)} - \boldsymbol{\beta}^*) X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
= \; & O_p(Rs^*(\log p)^{3/2}/n).
\end{aligned}
$$

Thus, the second term is negligible due to $Rs^*(\log p)^{3/2}/\sqrt{n} \to 0$. Similarly, we can derive the second term is negligible.

Combining these results, we have that

$$
\begin{aligned}
& (n/K)^{1/2}\max_j\left|\left(\widetilde{\beta}_j^{(k)} - \beta_j^*\right)\widehat{I}_{j|-j}^{(k)} + E_n^{(k)}\left[\left\{\nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right\}\left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*\right)\right]\right| \\
= \; & (n/K)^{1/2}\max_j\left|\left(\widehat{\beta}_j^{(k)} - \beta_j^*\right)\left(\widehat{I}_{j|-j}^{(k)} - I_{j|-j}\right)\right| + o_p(1).
\end{aligned}
$$

Notice that

$$
\begin{aligned}
& \left(\widehat{\beta}_j^{(k)} - \beta_j^*\right)\left(\widehat{I}_{j|-j}^{(k)} - I_{j|-j}^{(k)}\right) \\
= \;& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)})\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)})\right] \\
& - E\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right].
\end{aligned}
$$

By Lemma 10, 11 and the proof of Thoerem 4, we have

$$
\begin{aligned}
& (n/K)^{1/2}E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)}) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}^{(k)})\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)\right. \\
& \left.(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)})\right] \\
= \;& (n/K)^{1/2}E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(\widehat{\beta}_j^{(k)} - \beta_j^*)(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] \\
& + O_p(Rs^*(\log p)^{3/2}/\sqrt{n})
\end{aligned}
$$

uniformly over j. Thus,

$$
\begin{aligned}
& \max_j\left|\left(\widehat{\beta}_j^{(k)} - \beta_j^*\right)\left(\widehat{I}_{j|-j}^{(k)} - I_{j|-j}^{(k)}\right)\right| \\
\leq \;& \max_j\left|E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^{(k)}) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^{(k)})\right\} X_j(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] - \right. \\
& \left. E\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| \|\widehat{\boldsymbol{\beta}}^{(k)} - \boldsymbol{\beta}^*\|_2.
\end{aligned}
$$

By Claim 13, we have

$$
\begin{aligned}
& \max_j\left|E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-^{(k)}\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right] - \right. \\
& \left. E\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\} X_j(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right]\right| \\
= \;& O_p(Rn^{-\alpha-\beta} + R(n^{-\alpha} + n^{-\beta})\sqrt{\log p/n}).
\end{aligned}
$$

Thus, we have $(n/K)^{1/2}\max_j\left|\left(\widehat{\beta}_j^{(k)} - \beta_j^*\right)\left(\widehat{I}_{j|-j}^{(k)} - I_{j|-j}\right)\right| = o_p(1)$.

Next, define

$$
\left(\widehat{\sigma}_j^{(k)}\right)^2 = E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(k)}, \widehat{\Omega}_-^{(k)}\right)\right\}^2\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j^{(k)}\right)^2\right].
$$

and

$$
\sigma_j^2 = E\left[\{\nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\}^2\left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*\right)^2\right].
$$

We will show that $\max_j \left|\left(\widehat{\sigma}_j^{(k)}\right)^2 - \sigma_j^2\right| = o_p(1)$. To show this,

$$\begin{aligned}
& E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol\beta}^{(k)};\widehat{\Omega}_+^{(k)},\widehat{\Omega}_-^{(k)}\right)\right\}^2\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2\right] \\
& -E\left[\left\{\nabla l_\phi\left(\boldsymbol\beta^*;\Omega_+,\Omega_-\right)\right\}^2\left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right] \\
= \;& E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol\beta}^{(k)};\widehat{\Omega}_+^{(k)},\widehat{\Omega}_-^{(k)}\right)\right\}^2\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right] \\
& +E_n^{(k)}\left[\left\{\left\{\nabla l_\phi\left(\widehat{\boldsymbol\beta}^{(k)};\widehat{\Omega}_+^{(k)},\widehat{\Omega}_-^{(k)}\right)\right\}^2 - \left\{\nabla l_\phi\left(\boldsymbol\beta^*;\Omega_+,\Omega_-\right)\right\}^2\right\}\left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right] \\
& +\left(E_n^{(k)} - E\right)\left[\left\{\nabla l_\phi\left(\boldsymbol\beta^*;\Omega_+,\Omega_-\right)\right\}^2\left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right] \\
= \;& I_1 + I_2 + I_3.
\end{aligned}$$

For I_1, we have

$$\begin{aligned}
& E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol\beta}^{(k)};\widehat{\Omega}_+^{(k)},\widehat{\Omega}_-^{(k)}\right)\right\}^2\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right] \\
\leq \;& CE_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\right\}^2\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right] \\
& +CE_n^{(k)}\left[\left\{\widehat{\Omega}_-^{(k)}\right\}^2\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right]
\end{aligned}$$

By Condition (C5), we have

$$\begin{aligned}
& E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\right\}^2\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right] \\
= \;& E_n^{(k)}\left[\Omega_+\widehat{\Omega}_+^{(k)}\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right] \\
& +(n^{-\alpha} + n^{-\beta})E_n^{(k)}\left[\widehat{\Omega}_+^{(k)}\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right].
\end{aligned}$$

Notice that

$$\begin{aligned}
& \left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2 \\
= \;& \boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol w}_j^{(k)} - \boldsymbol w_j^*)\left(2X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)} - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right) \\
= \;& -\left\{\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol w}_j^{(k)} - \boldsymbol w_j^*)\right\}^2 + 2\boldsymbol{X}_{-j}^\top(\widehat{\boldsymbol w}_j^{(k)} - \boldsymbol w_j^*)\left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right).
\end{aligned}$$

By the condition on $Y - Q(a;\boldsymbol X)$, $\boldsymbol X$ and $X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*$, we have

$$\begin{aligned}
& \max_j E_n^{(k)}\left[\left\{\widehat{\Omega}_+^{(k)}\right\}^2\left\{\left(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol w}_j^{(k)}\right)^2 - \left(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol w_j^*\right)^2\right\}\right] \\
= \;& O_p(R\log n\log(np)\max\{s^*, s'\}\sqrt{\log p/n}).
\end{aligned}$$

For I_2, we have

$$
\begin{aligned}
& I_2 \\
= \; & E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(k)}, \widehat{\Omega}_-^{(k)}\right) - \nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right\}^2 \left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)^2\right] \\
& -2E_n^{(k)}\left[\left\{\nabla l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(k)}, \widehat{\Omega}_-^{(k)}\right) - \nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right\} \nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right) \right. \\
& \left. \left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)^2\right].
\end{aligned}
$$

Under the sub-Gaussian conditions, we have

$$
\begin{aligned}
& \max_i \left|\nabla l_\phi\left(\widehat{\boldsymbol{\beta}}^{(k)}; \widehat{\Omega}_+^{(k)}, \widehat{\Omega}_-^{(k)}\right) - \nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right| \\
= \; & O_p\left((n^{-\alpha} + n^{-\beta}) + R\sqrt{\log(np)} s^* \sqrt{\log p/n}\right).
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
& I_2 \\
\leq \; & O_p\left((n^{-\alpha} + n^{-\beta}) + R\sqrt{\log(np)} s^* \sqrt{\log p/n}\right) E_n^{(k)}\left[\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)^2\right] \\
& + O_p\left((n^{-\alpha} + n^{-\beta}) + R\sqrt{\log(np)} s^* \sqrt{\log p/n}\right) \\
& \times E_n^{(k)}\left[\left|\nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right| \left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)^2\right].
\end{aligned}
$$

By the condition on $Y - Q(a; \boldsymbol{X})$, we have

$$
\max_i \left|\nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right| = O_p(\sqrt{\log n}).
$$

By Lemma 14 in Loh and Wainwright (2015), we have

$$
\begin{aligned}
& \max_j \left|E_n^{(k)}\left[\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)^2\right]\right| \\
= \; & \max_j E\left[\left(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*\right)^2\right] + O_p(\sqrt{\log p/n}).
\end{aligned}
$$

Thus, we have $\max_j I_2 = o_p(1)$. For I_3, similar to I_2, we can derive that $\max_j I_3 = o_p(1)$. Thus, we have

$$
\max_j \left|\left(\widehat{\sigma}_j^{(k)}\right)^2 - \sigma_j^2\right| = o_p(1).
$$

Now, we show that $\min_j \sigma_j^2$ is bounded away from 0. To see this,

$$
\begin{aligned}
& \sigma_j^2 \\
\geq \; & E\left[(X_j - \boldsymbol{X}_{-j}^\top \boldsymbol{w}_j^*)^2\right] \left\{E\left[\left\{\nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right\}^2\right]\right\}^{-1} \\
\geq \; & \lambda_{\min} \left\{E\left[\left\{\nabla l_\phi\left(\boldsymbol{\beta}^*; \Omega_+, \Omega_-\right)\right\}^2\right]\right\}^{-1} > 0.
\end{aligned}
$$

Combining these results, we have

$$(n/K)^{1/2}\max_j\Big|\left(\widetilde{\beta}_j^{(k)}-\beta_j^*\right)\widehat{I}_{j|-j}^{(k)}/\widehat{\sigma}_j^{(k)} \\ +\sigma_j^{-1}E_n^{(k)}\left[\left\{\nabla l_\phi\left(\boldsymbol{\beta}^*;\Omega_+,\Omega_-\right)\right\}\left(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*\right)\right]\Big|=o_p(1).$$

By the Berry-Esseen bound for CLT, we have

$$\max_j\sup_{\alpha\in(0,1)}\left|P\left(\left|\left(\widetilde{\beta}_j^{(k)}-\beta_j^*\right)\widehat{I}_{j|-j}^{(k)}/\widehat{\sigma}_j^{(k)}\right|\le\Phi^{-1}(1-\alpha/2)\right)-(1-\alpha)\right|=o_p(1).$$

□

Proof of Theorem 7. Let the two split data set be $\mathcal{I}_1$ and $\mathcal{I}_2$ with sample size n_1 and n_2. We rearrange the data as the order of index in $\mathcal{I}_1$ and $\mathcal{I}_2$. Define the sigma-field generated by first j samples as $\mathcal{F}_j$.

$$\begin{aligned}\widehat{V}(\widehat{D})-V(D^*) &= (\widehat{V}-V)(\widehat{D}) \\ &\quad +V(\widehat{D})-V(D^*) \\ &= I_1+I_2.\end{aligned}$$

For I_1,we have

$$\begin{aligned}I_1 &= (E_{n_2}^{(2)}-E)\left(\widehat{W}_{\widehat{D}(\boldsymbol{X})}(Y,\boldsymbol{X},A,\widehat{\pi}_{(-2)},\widehat{Q}_{(-2)})\right) \\ &\quad +E\left[\widehat{W}_{\widehat{D}(\boldsymbol{X})}(Y,\boldsymbol{X},A,\widehat{\pi}_{(-2)},\widehat{Q}_{(-2)})-W_{\widehat{D}(\boldsymbol{X})}(Y,\boldsymbol{X},A,\pi,Q)\right] \\ &= (E_{n_2}^{(2)}-E)\left(\widehat{W}_{\widehat{D}(\boldsymbol{X})}(Y,\boldsymbol{X},A,\widehat{\pi}_{(-2)},\widehat{Q}_{(-2)})\right)+O_p(n_1^{-\alpha-\beta}).\end{aligned}$$

The second equality is due to Condition (C5). Let $Z_{n,i}=\widehat{W}_{\widehat{D}(\boldsymbol{X}_i)}(Y_i,\boldsymbol{X}_i,A_i,\widehat{\pi}_{(-2)},\widehat{Q}_{(-2)})-E\left[\widehat{W}_{\widehat{D}(\boldsymbol{X}_i)}(Y_i,\boldsymbol{X}_i,A_i,\widehat{\pi}_{(-2)},\widehat{Q}_{(-2)})\right]$. Consider $M_{n,i}=n_2^{-1/2}\widehat{\sigma}_V^{-1}Z_{n,i}$. We have

$$E[M_{n,i}|\mathcal{F}_{i-1}]=0$$

for any $i\in\mathcal{I}_2$. We also have that $\sum_{i\in\mathcal{I}_2}E[M_{n,i}^2|\mathcal{F}_{i-1}]=\sigma_V^2/\widehat{\sigma}_V^2\to 1$. Because Y is bounded, $\widehat{\pi}$ and $\widehat{Q}$ are consistent, and $\pi(a;\boldsymbol{X})$ is bounded away from 0 and 1, we have that $Z_{n,i}$ is bounded. Because $\widehat{\sigma}_V^2\to\sigma_V^2>0$, the conditional Linderberg condition (Condition C2) in Luedtke and Van Der Laan (2016)) holds, the martingle central limit theorem for triangular arrays [see, e.g., Theorem 2 in] shows that $n_2^{1/2}\widehat{\sigma}_V^{-1}I_1=\sum_{i\in\mathcal{I}_2}M_{n,i}\to N(0,1)$.

Next, we will show that $I_2=o_p(n_2^{-1/2})$.

$$\begin{aligned}V(\widehat{D}^*)-V(D^*) &= E\left[|\Delta|1\left\{\widehat{D}^*\neq D^*\right\}\right] \\ &\le E\left[|\Delta|1\left\{|\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)|\ge|\boldsymbol{X}^\top\boldsymbol{\beta}^*|\right\}\right] \\ &\le E\left[|\Delta|1\left\{\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*\|_1\ge|\boldsymbol{X}^\top\boldsymbol{\beta}^*|\right\}\right]\end{aligned}$$

If $|\Delta|\le\psi\left(|\boldsymbol{X}^\top\boldsymbol{\beta}^*|\right)$ when $\boldsymbol{X}^\top\boldsymbol{\beta}^*$ is at a neighborhood of 0, then $V(\widehat{D}^*)-V(D^*)\lesssim\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*\|_1^{\zeta+\gamma}$. If we do not have condition, then $V(\widehat{D}^*)-V(D^*)\lesssim\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*\|_1^{\gamma}$. Under conditions, we have $I_2=o_p(n_2^{-1/2})$. This concludes the proof. □

Proof of Lemmas and Claims.

The proofs of theorems use lemmas and claims below. For the simplicity of the notation, we omit the superscript indicating split data set. For example, $\widehat{\beta}^{(k)}$ is written as $\widehat{\beta}$.

Lemma 10 *Assuming conditions in Theorem 4, we have*

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_1 = O_p\left(s^*(\log p/n)^{1/2}\right),$$

where $s^* = \|\boldsymbol{\beta}^*\|_0$. *Further,*

$$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top H_X(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \lesssim s^* \log p/n,$$

and

$$\max_j(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-1})^\top H_{-j}(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j}) \lesssim s^* \log p/n,$$

where $H_{\boldsymbol{X}}$ *and* H_{-j} *are defined in the proof.*

Lemma 11 *We have*

$$\max_j \|\widehat{\boldsymbol{w}}_j - \boldsymbol{w}_j^*\|_1 = O_p\left(R \max\left\{s^*, s'\right\}(\log p/n)^{1/2}\right),$$

where $s' = \max \|\boldsymbol{w}_j^*\|_0$.

The following claims are useful in the proofs of lemmas. It essentially takes advantage of the subgaussian tail of $Y_a - Q(a; \boldsymbol{X})$.

Claim 12 *Let* $Y = (Y_1, \cdots, Y_n)^T$ *be the* n *dimensional independent random vector and* $a \in \mathrm{R}^n$. *Then*

a. If Y_i*'s are bounded in* $[c, d]$ *for some* $c, d \in \mathrm{R}$*., then for any* $t \in (0, +\infty)$

$$P\left(\left|a^\top Y - a^\top E[Y]\right| > t\right) \leq 2\exp\left\{-t^2/\|a\|_2^2(d-c)^2\right\}.$$

b. If Y_i*'s are unbounded and there exists some* $M, \nu_0 \in \mathrm{R}$ *such that*

$$\max_{i=1,\cdots,n} E\left\{\exp\left[|Y_i - E(Y_i)|/M\right] - 1 - |Y_i - E(Y_i)|/M\right\} M^2 \leq \nu_0/2,$$

then for any $t \in (0, +\infty)$

$$P\left(\left|a^\top Y - a^\top E(Y)\right| > t\right) \leq 2\exp\left\{-t^2/2(\|a\|_2^2\nu_0 + \|a\|_\infty Mt)\right\}.$$

Proof of Claim 12. The proof is the same as Proposition 4 in Fan and Lv (2011). The results from Hoeffding's and Bernstein's inequality, respectively. □

Claim 13 *For any function vector $\boldsymbol{h}:\mathbb{R}^p\to\mathbb{R}^q$ with $\boldsymbol{h}$ is sub-Gaussian with a proxy of the order $O(R)$, $\|\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top\|_{\max}\le O_p(a_n)$ and $\|E[\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top]\|_\infty = O(R)$ and $\max_{j_1,j_2} E\left[h_{j_1}^4(\boldsymbol{X})\right]E\left[X_{j_2}^4\right]\le R^4$, if $a_nR^{-1}\sqrt{\log p/n}\to 0$, we have for $a=1$ or -1,*

$$\left\|E_n[(\widehat{\Omega}_a-\Omega_a)\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top]\right\|_{\max} = O_p\left(Rn^{-\alpha-\beta}+R(n^{-\alpha}+n^{-\beta})\sqrt{\log p/n}\right),$$

where $\|\boldsymbol{V}\|_{\max}$ denotes the maximum of the absolute values of the entries if $\boldsymbol{V}$ is a matrix; if $\boldsymbol{V}$ is a vector, $\|\boldsymbol{V}\|_{\max}=\|\boldsymbol{V}\|_\infty$.

Proof of Claim 13. For simplicity, we will prove when $a=1$. The proof can also be applied to $a=-1$. Note that

$$\begin{aligned}
&\|E_n[(\widehat{\Omega}_a-\Omega_a)\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top]\|_{\max}\\
\le&\left\|E_n\left[I\{A=1\}\left(\widehat{\pi}^{-1}(1;\boldsymbol{X})-\pi^{-1}(1;\boldsymbol{X})\right)\left(\widehat{Q}(1;\boldsymbol{X})-Q(1;\boldsymbol{X})\right)\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top\right]\right\|_{\max}\\
&+\left\|E_n\left[I\{A=1\}\left(\widehat{\pi}^{-1}(1;\boldsymbol{X})-\pi^{-1}(1;\boldsymbol{X})\right)(Y-Q(1;\boldsymbol{X}))\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top\right]\right\|_{\max}\\
&+\left\|E_n\left[\left(I\{A=1\}\pi^{-1}(1;\boldsymbol{X})-1\right)\left(\widehat{Q}(1;\boldsymbol{X})-Q(1;\boldsymbol{X})\right)\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top\right]\right\|_{\max}\\
=&I_1+I_2+I_3
\end{aligned}$$

Define

$$\begin{aligned}
\Omega_n(c) &= \{\|\widehat{Q}-Q\|_\infty\le cn^{\beta}\}\cap\{\|\widehat{\pi}-\pi\|_\infty\le cn^{-\alpha}\}\cap\{\|\boldsymbol{h}(\boldsymbol{X})\boldsymbol{X}^\top\|_{\max}\le ca_n\}\\
\Theta_n &= \{\|E_n[\boldsymbol{h}(\boldsymbol{X})\boldsymbol{h}^\top(\boldsymbol{X})]-E[\boldsymbol{h}(\boldsymbol{X})\boldsymbol{h}^\top(\boldsymbol{X})]\|_{\max}\le cR\sqrt{\log p/n}\}\\
&\quad\cup\{\|E_n[\boldsymbol{X}\boldsymbol{X}^\top]-E[\boldsymbol{X}\boldsymbol{X}^\top]\|_{\max}\le cR\sqrt{\log p/n}\}
\end{aligned}$$

For I_1, we have

$$P(|I_1|>t)\le P(|I_1|>t\mid\Omega_n(c)\cap\Theta_n)+P(\Omega_n^c(c))+P(\Theta_n^c).$$

By Condition (C5) and sub-Gaussian conditions, we have

$$P(\Omega_n^c(c))\to 0,\quad P(\Theta_n^c)\to 0.$$

On $\Omega_n(c)\cap\Theta_n$, we have

$$\begin{aligned}
|I_1| &\le cn^{-\alpha}cn^{-\beta}\max_{j_1,j_2}E_n[|h_{j_1}(\boldsymbol{X})X_{j_2}|]\\
&\le c^2n^{-\alpha-\beta}\max_{j_1,j_2}\left\{(E_n-E)[|h_{j_1}(\boldsymbol{X})X_{j_2}|]+E[|h_{j_1}(\boldsymbol{X})X_{j_2}|]\right\}.
\end{aligned}$$

By Lemma 14 in Loh and Wainwright (2011), we have

$$P(|(E_n-E)[|h_{j_1}(\boldsymbol{X})X_{j_2}|]|\ge t)\le 6pq\exp\left\{-cn\min\left(t^2/\sigma^2,t/\sigma\right)\right\}$$

where σ is the multiplication of the proxy of X_{j_2} and $h_{j_1}(\boldsymbol{X})$. Thus, we know that

$$|I_1|\le CRn^{-\alpha-\beta}$$

on $\Omega_n(c) \cap \Theta_n$.

Now we focus on I_2. For any $t, c > 0$, we have

$$P\left(I_2 \geq t\right) \leq P\left(I_2 \geq t \mid \Omega_n(c) \cap \Theta_n\right) + P(\Omega_n^c(c)) + P(\Theta_n^c).$$

Notice that by Claim 12 and the independence of $\widehat{\pi}$. Let

$$\mathcal{E}_t = \left\{\left|E_n\left[I\{A=1\}\left(\widehat{\pi}^{-1}(1;\boldsymbol{X}) - \pi^{-1}(1;\boldsymbol{X})\right)\left(Y_1 - Q(1;\boldsymbol{X})\right)h_{j_1}(\boldsymbol{X})X_{j_2}\right]\right| > t\right\}.$$

For any $j_1 \in \{1, \cdots, q\}$ and $j_2 \in \{1, \cdots, p\}$, we have

$$\begin{aligned}
& P\left\{\mathcal{E}_t \mid A=1, \Omega_n \cap \Theta_n\right\} \\
\leq\ & 2\exp\left\{-\frac{1}{2}nt^2\left\{c^2 n^{-2\alpha}\nu_0 E_n\left[h_{j_1}^2(\boldsymbol{X})X_{j_2}^2\right] + \max_{i=1,\cdots,n}|h_{j_1}(\boldsymbol{X}_i)X_{i,j_2}|cn^{-\alpha}Mt\right\}\right\} \\
\leq\ & 2\exp\left\{-\frac{1}{2}nt^2\left\{c^2 n^{-2\alpha}\nu_0\sqrt{E_n\left[h_{j_1}^4(\boldsymbol{X})\right]E_n\left[X_{j_2}^4\right]}\right.\right. \\
& \left.\left.+\max_{i=1,\cdots,n}|h_{j_1}(\boldsymbol{X}_i)X_{i,j_2}|cn^{-\alpha}Mt\right\}\right\} \\
\leq\ & 2\exp\left\{-\frac{1}{2}nt^2\left\{c^4 n^{-2\alpha}\nu_0\max_{j_1,j_2}\sqrt{E\left[h_{j_1}^4(\boldsymbol{X})\right]E\left[X_{j_2}^4\right]} + c^2 a_n n^{-\alpha}Mt\right\}\right\}.
\end{aligned}$$

The last inequality uses the concentration inequality for polynomials of sub-Gaussian random variables Adamczak and Wolff (2015), .i.e. there exist C,C_d, and c_1, c_2 such that

$$\begin{aligned}
P\left(|(E_n - E)[h_j^4(\boldsymbol{X})]| > t\right) &\leq 2C\exp\left\{-\frac{1}{C_D}\min\left(nt^2/(c_1R^2), n^{1/2}t^{1/2}/(c_2R^2)\right)\right\}, \\
P\left(|(E_n - E)[X_j^4]| > t\right) &\leq 2C\exp\left\{-\frac{1}{C_D}\min\left(nt^2/c_1, n^{1/2}t^{1/2}/c_2\right)\right\},
\end{aligned}$$

and the assumption that $\log p/\sqrt{n} \to 0$.

Thus, under $\log p/\sqrt{n} \to 0$, we have

$$\begin{aligned}
& P\left(I_2 \geq t \mid \Omega_n(c) \cap \Theta_n\right) \\
\leq & 2p^2\exp\left\{-nt^2/2\left\{Cc^4R^2n^{-2\alpha}\nu_0 + c^2a_nn^{-\alpha}Mt\right\}\right\}.
\end{aligned}$$

Let $t = 8(Cc^4\nu_0 + c^2M)Rc^2n^{-\alpha}\sqrt{\log p/n}$, we have

$$P\left\{I_2 \geq 8(Cc^4\nu_0 + c^2M)Rc^2n^{-\alpha}\sqrt{\log p/n} \mid \Omega_n(c) \cap \Theta_n\right\} \to 0.$$

For I_3, notice that π is bounded and apply (a) in Claim 12, we can similarly conclude $I_3 \lesssim O_p(Rn^{-\beta}\sqrt{\log pn})$. Thus, we con conclude the claim. □

Proof of Lemma 10. To simplify the notation, we omite the superscript (k)l instead, we assume that the nuisance parameter estimations used in constructing $\widehat{\Omega}_+$ and $\widehat{\Omega}_-$ are independent from the observed samples. Let S denote the support of $\boldsymbol{\beta}^*$. Denote

$$\begin{aligned}
& D(\boldsymbol{\beta}, \boldsymbol{\beta}^*) \\
=\ & E_n\left[\left\{\widehat{\Omega}_+[\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}) - \phi'(\boldsymbol{X}^\top\beta^*)] - \widehat{\Omega}_-[\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}) - \phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)]\right\}\boldsymbol{X}^\top(\boldsymbol{\beta} - \boldsymbol{\beta}^*)\right].
\end{aligned}$$

Let $\widehat{\boldsymbol{\Delta}} = \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*$. Thus,

$$\begin{aligned} & D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*) \\ = \quad & E_n\left[\left\{\widehat{\Omega}_+\phi'(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \widehat{\Omega}_-\phi'(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}\boldsymbol{X}_S^\top\widehat{\boldsymbol{\Delta}}_S\right] \\ & +E_n\left[\left\{\widehat{\Omega}_+\phi'(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \widehat{\Omega}_-\phi'(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}\boldsymbol{X}_S^\top\widehat{\boldsymbol{\beta}}_{\bar{S}}\right] \\ & -E_n\left[\left\{\widehat{\Omega}_+\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{\Omega}_-\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}^\top\widehat{\boldsymbol{\Delta}}\right] \\ = \quad & (I) + (II) + (III) \end{aligned}$$

By KKT condition,

$$(I) \le \lambda_n\|\widehat{\boldsymbol{\Delta}}_S\|_1, (II) = -\lambda_n\|\widehat{\boldsymbol{\Delta}}_{\bar{S}}\|.$$

Assuming that $\widehat{W}_1$, $\widehat{W}_{-1}$, W_1 and W_{-1} are positive, we have

$$\begin{aligned} & (III) \\ = & -E_n\left[\left\{\widehat{W}_1\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - \widehat{W}_{-1}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}^\top\widehat{\boldsymbol{\Delta}}\right] \\ = & -E_n\left[\left\{W_1\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - W_{-1}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}^\top\widehat{\boldsymbol{\Delta}}\right] \\ & -E_n\left[\left\{[\widehat{W}_1 - W_1]\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - [\widehat{W}_{-1} - W_{-1}]\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}^\top\widehat{\boldsymbol{\Delta}}\right]. \end{aligned}$$

Due to the sub-Gaussian condition on X, following the proof of Theorem 4, we can show that

$$\begin{aligned} & P\left(\left\|E_n\left[\left\{W_1\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - W_{-1}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right]\right\|_\infty \ge C\sqrt{\log p/n}\right) \\ \le \quad & c_0\exp(-c_1\log p), \end{aligned}$$

where C, c_0 and c_1 are som constants. By (the proof of) Claim 13, the second term can be bounded by $O_p\left\{[(n^{-\alpha} + n^{-\beta})(\log p/n)^{1/2} + n^{-\alpha-\beta}]\|\widehat{\Delta}\|_1\right\}$ when both model are correct. Thus, $(III) \le C(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}\|_1$ with a large enough C.

Let $\lambda_n = 2C(\log p/n)^{1/2}$, on the event,

$$\left\{\left\|E_n\left[\left\{W_1\phi'(\boldsymbol{X}^\top\boldsymbol{\beta}^*) - W_{-1}\phi'(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}\right]\right\|_\infty \le C\sqrt{\log p/n}\right\},$$

we have

$$D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*) \le C(\log p/n)^{1/2}(3\|\widehat{\boldsymbol{\Delta}}_S\|_1 - \|\widehat{\boldsymbol{\Delta}}_{\bar{S}}\|_1).$$

Since ϕ is convex, we have $D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*) \ge 0$ and $\|\widehat{\boldsymbol{\Delta}}_{\bar{S}}\|_1 \le 3\|\widehat{\boldsymbol{\Delta}}_S\|_1$. In addition, $D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*)$ can be rewritten as

$$\begin{aligned} & (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top E_n[\nabla l_\phi(\widehat{\boldsymbol{\beta}}; \widehat{\Omega}_+, \widehat{\Omega}_-) - \nabla l_\phi(\boldsymbol{\beta}^*; \widehat{\Omega}_+, \widehat{\Omega}_-)X] \\ = & E_n\left\{[\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widetilde{\boldsymbol{\beta}}) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widetilde{\boldsymbol{\beta}})](\boldsymbol{X}^\top\widehat{\boldsymbol{\Delta}})^2\right\} \\ = & E_n\left\{\left[\Omega_+\phi''(\boldsymbol{X}^\top\widetilde{\boldsymbol{\beta}}) + \Omega_-\phi''(-\boldsymbol{X}^\top\widetilde{\boldsymbol{\beta}})\right](\boldsymbol{X}^\top\widehat{\Delta})^2\right\} \\ & + E_n\left\{\left[\left(\widehat{\Omega}_+ - \Omega_+\right)\phi''(\boldsymbol{X}^\top\widetilde{\boldsymbol{\beta}}) + \left(\widehat{\Omega}_- - \Omega_-\right)\phi''(-\boldsymbol{X}^\top\widetilde{\boldsymbol{\beta}})\right](\boldsymbol{X}^\top\widehat{\boldsymbol{\Delta}})^2\right\} \end{aligned}$$

The first term satisfies the RSC condition following the Proposition 1 in Loh and Wainwright (2015) when X's are zero-mean sub-Gaussians with probablity at least $1 - c_1 \exp(-c_2 n)$. The second term has the following bound

$$\begin{aligned}
&\left| E_n \left\{ \left[\left(\widehat{\Omega}_+ - \Omega_+ \right) \phi''(\boldsymbol{X}^\top \widetilde{\boldsymbol{\beta}}) + \left(\widehat{\Omega}_- - \Omega_- \right) \phi''(-\boldsymbol{X}^\top \widetilde{\boldsymbol{\beta}}) \right] (\boldsymbol{X}^\top \widehat{\boldsymbol{\Delta}})^2 \right\} \right| \\
\leq &\|\widehat{\boldsymbol{\Delta}}\|_1^2 \left\| E_n \left[\left\{ \left(\widehat{\Omega}_+ - \Omega_+ \right) \phi''(\boldsymbol{X}^\top \widetilde{\boldsymbol{\beta}}) + \left(\widehat{\Omega}_- - \Omega_- \right) \phi''(-\boldsymbol{X}^\top \widetilde{\boldsymbol{\beta}}) \right\} \boldsymbol{X}\boldsymbol{X}^\top \right] \right\|_{\max} \\
\lesssim &(n^{-\alpha-\beta} + (n^{-\beta} + n^{-\beta})\sqrt{\log p/n}) s^* \|\widehat{\boldsymbol{\Delta}}\|_2^2.
\end{aligned}$$

The last inequality uses the fact that $\|\widehat{\boldsymbol{\Delta}}\|_1 \leq 4\|\widehat{\boldsymbol{\Delta}}_S\|_1 \leq 4(s^*)^{1/2}\|\widehat{\boldsymbol{\Delta}}_S\|_2 \leq 4(s^*)^{1/2}\|\widehat{\boldsymbol{\Delta}}\|_2$ and the concentration inequality for the polynomial functions of independent sub-Gaussian (Adamczak and Wolff, 2015). Thus, we have $D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*) \geq \kappa \|\widehat{\boldsymbol{\Delta}}\|_2^2 - \tau\sqrt{\log p/n}\|\widehat{\boldsymbol{\Delta}}\|_1\|\widehat{\boldsymbol{\Delta}}\|_2$ on the event

$$\begin{aligned}
&\left\{ \left\| E_n \left[\left\{ W_1 \phi'(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - W_{-1}\phi'(-\boldsymbol{X}^\top \boldsymbol{\beta}^*) \right\} \boldsymbol{X} \right] \right\|_\infty \leq C\sqrt{\log p/n} \right\} \cup \\
&\left\{ \|\widehat{Q} - Q\|_\infty \leq cn^\beta \right\} \cap \left\{ \|\widehat{\pi} - \pi\|_\infty \leq cn^{-\alpha} \right\} \cup \{\text{RSC condition holds}\}.
\end{aligned}$$

Notice that

$$\tau\sqrt{\log p/n}\|\widehat{\boldsymbol{\Delta}}\|_1\|\widehat{\boldsymbol{\Delta}}\|_2 \leq \kappa/2\|\widehat{\boldsymbol{\Delta}}\|_2^2 + \frac{\tau^2}{2\kappa}\frac{\log p}{n}\|\widehat{\boldsymbol{\Delta}}\|_1^2$$

Combining the upper bound of $D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*)$ derived above, given $\|\boldsymbol{\beta}^*\|_1 \leq \sqrt{n/\log p}$, we have

$$\kappa/2\|\widehat{\boldsymbol{\Delta}}\|_2^2 \leq C(\log p/n)^{1/2}(3\|\widehat{\boldsymbol{\Delta}}_S\|_1 - \|\widehat{\boldsymbol{\Delta}}_{\bar{S}}\|_1) + \frac{\tau^2}{2\kappa}\sqrt{\frac{\log p}{n}}\|\widehat{\boldsymbol{\Delta}}\|_1.$$

Thus, we have

$$\begin{aligned}
\|\widehat{\boldsymbol{\Delta}}\|_2 &\leq \left\{ \frac{24C}{\kappa} + \frac{4\tau^2}{\kappa^2} \right\} \sqrt{\frac{s^* \log p}{n}}, \\
\|\widehat{\boldsymbol{\Delta}}\|_1 &\leq \left\{ \frac{96C}{\kappa} + \frac{16\tau^2}{\kappa^2} \right\} s^* \sqrt{\frac{\log p}{n}}, \\
D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*) &\leq 3C \left\{ \frac{96C}{\kappa} + \frac{16\tau^2}{\kappa^2} \right\} s^* \frac{\log p}{n}.
\end{aligned}$$

Let $H_{\boldsymbol{X}} = E_n \left\{ [\widehat{\Omega}_+ \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) + \widehat{\Omega}_- \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*)](\boldsymbol{X}^\top \boldsymbol{X}) \right\}$. Note that

$$\begin{aligned}
&|D(\widehat{\beta}, \boldsymbol{\beta}^*) - (\widehat{\beta} - \boldsymbol{\beta}^*)^\top H_X(\widehat{\beta} - \boldsymbol{\beta}^*)| \\
= &\left| E_n \left[\widehat{\Omega}_+ \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}) + \widehat{\Omega}_- \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}) - \widehat{\Omega}_+ \phi''(\boldsymbol{X}^\top \boldsymbol{\beta}^*) - \widehat{\Omega}_- \phi''(-\boldsymbol{X}^\top \boldsymbol{\beta}^*)(\boldsymbol{X}^\top \widehat{\boldsymbol{\Delta}})^2 \right] \right| \\
\lesssim &C\sqrt{\log(np)} s^* (\log p/n)^{1/2} \left| (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top H_{\boldsymbol{X}} (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \right|.
\end{aligned}$$

Combining with the upper bound on $D(\widehat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*)$, by $s^* \log p/\sqrt{n} \to 0$, we know that

$$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top H_{\boldsymbol{X}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \leq \frac{3C}{1+C}\left\{\frac{96C}{\kappa} + \frac{16\tau^2}{\kappa^2}\right\} s^* \frac{\log p}{n}.$$

Let $H_{-j} = E_n\left[\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\}\boldsymbol{X}_{-j}\boldsymbol{X}_{-j}^\top\right]$. We have

$$\begin{aligned}
&(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j})^\top H_{-j}(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j})\\
\leq &2(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top H_{\boldsymbol{X}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)\\
&+ 2(\widehat{\beta}_j - \beta_j^*)^2 E_n\left[[\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)]X_j^2\right].
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\max_j(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j})^\top H_{-j}(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j})\\
\leq\ & 2(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)^\top H_{\boldsymbol{X}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)\\
&+2\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_2^2 \max_j E_n\left[[\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)]X_j^2\right].
\end{aligned}$$

By the sub-Gaussian assumption, we have

$$\begin{aligned}
&\max_j \Big| E_n\left[[\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)]X_j^2\right]\\
&-E\left[[\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)]X_j^2\right]\Big| = o_p(1).
\end{aligned}$$

By $\max_j E[X_j^2]$ is bounded, we have that

$$\max_j(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j})^\top H_{-j}(\widehat{\boldsymbol{\beta}}_{-j} - \boldsymbol{\beta}^*_{-j}) = O_p\left(s^*\frac{\log p}{n}\right).$$

□

Before the formal proof of Lemma 11, we establish the following claim. The following claim plays the same role as the RSC in the proof of Lemma 10. Let S' be the support of w^*.

Claim 14 *Denote $\widehat{F}(\boldsymbol{\beta}) = E_n[\widehat{U}(\boldsymbol{\beta})\boldsymbol{X}\boldsymbol{X}^\top]$, where $\widehat{U}(\boldsymbol{\beta}) = \widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta})$. And*

$$\kappa_D(s') = \min\left\{(s')^{1/2}(\boldsymbol{v}^\top\widehat{F}(\widehat{\boldsymbol{\beta}})\boldsymbol{v})^{1/2}/\|\boldsymbol{v}_{S'}\|_1 : \boldsymbol{v} \in R^p\backslash\{0\}, \|\boldsymbol{v}_{\bar{S}'}\|_1 \leq \xi\|\boldsymbol{v}_{S'}\|_1\right\},$$

where ξ is a positive constant. Assuming assumptions in Theorem 4, with probability tending to one, $\kappa_D(s') \geq \kappa/\sqrt{6}$.

Proof of Claim 14. By the definition of $\kappa_D(s')$ and the fact that $\|\boldsymbol{v}_{S'}\|_1 \leq (s')^{1/2}\|\boldsymbol{v}_{S'}\|_2 \leq (s')^{1/2}\|\boldsymbol{v}\|_2$, we only need to show that

$$\kappa_D^2(s') = \min\left\{\boldsymbol{v}^\top \widehat{F}(\widehat{\boldsymbol{\beta}})\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 : \boldsymbol{v} \in \mathrm{R}^p\backslash\{0\}, \|\boldsymbol{v}_{\bar{S}'}\| \leq \xi\|\boldsymbol{v}_{S'}\|_1\right\}.$$

Let $F(\boldsymbol{\beta}) = E_n[U(\boldsymbol{\beta})\boldsymbol{X}\boldsymbol{X}^\top]$, where $U(\boldsymbol{\beta}) = \Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta})$. We have

$$\begin{aligned}
&\boldsymbol{v}^\top \widehat{F}(\widehat{\boldsymbol{\beta}})\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 \\
=&\boldsymbol{v}^\top F(\boldsymbol{\beta}^*)\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 \\
&+ E_n\left[\left\{(\widehat{\Omega}_+ - \Omega_+)\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + (\widehat{\Omega}_- - \Omega_-)\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(\boldsymbol{X}^\top\boldsymbol{v})^2/\|\boldsymbol{v}\|_2^2\right] \\
&+ E_n\left[\left\{\Omega_+(\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)) + \Omega_-(\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}(\boldsymbol{X}^\top\boldsymbol{v})^2/\|\boldsymbol{v}\|_2^2\right] \\
&+ E_n\left[\left\{(\widehat{\Omega}_+ - \Omega_+)(\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}(\boldsymbol{X}^\top\boldsymbol{v})^2/\|\boldsymbol{v}\|_2^2\right] \\
&+ E_n\left[\left\{(\widehat{\Omega}_- - \Omega_-)(\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}(\boldsymbol{X}^\top\boldsymbol{v})^2/\|\boldsymbol{v}\|_2^2\right] \\
=&\boldsymbol{v}^\top F(\boldsymbol{\beta}^*)\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 + I_1 + I_2 + I_3 + I_4
\end{aligned}$$

By (the proof of) Lemma 10, I_1 can be bounded by $O_p\left(n^{-\alpha} + n^{-\beta}\right)$. For I_2, we have

$$\begin{aligned}
I_2 &\leq CE_n\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)(\boldsymbol{X}^\top\boldsymbol{v})^2/\|\boldsymbol{v}\|_2^2\right] \\
&\leq Cs^*\log(np)/\sqrt{n}\boldsymbol{v}^\top F(\boldsymbol{\beta}^*)\boldsymbol{v}/\|\boldsymbol{v}\|_2^2.
\end{aligned}$$

By $s^*\log(np)/\sqrt{n} \to 0$, we know that

$$I_2 \leq \boldsymbol{v}^\top F(\boldsymbol{\beta}^*)\boldsymbol{v}/2\|\boldsymbol{v}\|_2^2.$$

For I_3, we have the following.

$$\begin{aligned}
I_3 \leq& E_n\left[\left\{|\widehat{\Omega}_+ - \Omega_+|\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)|\boldsymbol{X}^\top(\widehat{\beta} - \boldsymbol{\beta}^*)|\right\}(\boldsymbol{X}_{-j}^\top v)^2/\|v\|_2^2\right] \\
\leq& Cs^*\log(np)/\sqrt{n}E_n\left[\left\{|\widehat{\Omega}_+ - \Omega_+|\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(\boldsymbol{X}_{-j}^\top v)^2/\|v\|_2^2\right] \\
\lesssim& (n^{-\alpha} + n^{-\beta})s^*\log(np)/\sqrt{n}.
\end{aligned}$$

Similarly, $I_4 \lesssim (n^{-\alpha} + n^{-\beta})s^*\log(np)/\sqrt{n}$. Hence, $\boldsymbol{v}^\top \widehat{F}(\widehat{\boldsymbol{\beta}})\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 \geq (1/4)\boldsymbol{v}^\top F(\boldsymbol{\beta}^*)\boldsymbol{v}/\|\boldsymbol{v}\|_2^2$, with probability tending to 1. Note that

$$\begin{aligned}
\boldsymbol{v}^\top \widehat{F}(\widehat{\boldsymbol{\beta}})\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 &\geq (1/4)\boldsymbol{v}^\top F(\boldsymbol{\beta}^*)\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 \\
&= 1/4\left\{\boldsymbol{v}^\top \boldsymbol{I}^*\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 + \boldsymbol{v}^\top\left(F(\boldsymbol{\beta}^*) - \boldsymbol{I}^*\right)\boldsymbol{v}/\|\boldsymbol{v}\|_2^2\right\} \\
&\geq 1/4\left\{\lambda_{\min}\left(\boldsymbol{I}^*\right) - \left|\boldsymbol{v}^\top\left(F(\boldsymbol{\beta}^*) - \boldsymbol{I}^*\right)\boldsymbol{v}\right|/\|\boldsymbol{v}\|_2^2\right\} \\
&\geq 1/4\left\{\kappa^2 - \|\boldsymbol{v}\|_1^2\|F(\boldsymbol{\beta}^*) - \boldsymbol{I}^*\|_{\max}/\|\boldsymbol{v}\|_2^2\right\}.
\end{aligned}$$

By $\|\boldsymbol{v}\|^2 \leq (\xi+1)^2\|\boldsymbol{v}_{S'}\|_1^2 \leq s'(\xi+1)^2\|\boldsymbol{v}\|_2^2$, we have

$$\boldsymbol{v}^\top \widehat{F}(\widehat{\boldsymbol{\beta}})\boldsymbol{v}/\|\boldsymbol{v}\|_2^2 \geq 1/4\left\{\kappa^2 - (\xi+1)^2 s'\|F(\boldsymbol{\beta}^*) - \boldsymbol{I}^*\|_{\max}\right\}.$$

By sub-gaussian condition, we know that $\|F(\boldsymbol{\beta}^*) - \boldsymbol{I}^*\|_{\max} = O_P((\log p/n)^{1/2})$. By

$$s'(\log p/n)^{1/2} \to 0,$$

we have probability tending to 1 such that $s'\|F(\boldsymbol{\beta}^*) - \boldsymbol{I}^*\|_{\max} \le \kappa^2/[3(\xi+1)^2]$. Thus, we can conclude the claim. □

Proof of Lemma 11. Let $\widehat{\Delta}_j = \widehat{\boldsymbol{w}}_j - \boldsymbol{w}_j^*$. By definition, we have

$$\begin{aligned}
&E_n\left[\left\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(X_j - \boldsymbol{X}_{-j}^\top\widehat{\boldsymbol{w}}_j)^2\right] + \lambda_n'\|\widehat{\boldsymbol{w}}_j\|_1 \\
\le &E_n\left[\left\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2\right] + \lambda_n'\|\boldsymbol{w}_j^*\|_1.
\end{aligned}$$

By rearranging terms, we have equivalently

$$\begin{aligned}
&E_n\left[\left\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j})^2\right] \\
\le &2E_n\left[\left\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) + \widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j}\right] \\
&+ \lambda_n'\|\boldsymbol{w}_j^*\|_1 - \lambda_n'\|\widehat{\boldsymbol{w}}_j\|_1 \\
\le &2I_1 + \lambda_n'\|\boldsymbol{w}_j^*\|_1 - \lambda_n'\|\widehat{\boldsymbol{w}}_j\|_1.
\end{aligned}$$

Notice that

$$\lambda_n'\|\boldsymbol{w}_j^*\|_1 - \lambda_n'\|\widehat{\boldsymbol{w}}_j\|_1 = \lambda_n'\|\boldsymbol{w}_{j,S'}^*\|_1 - \lambda_n'\|\widehat{\boldsymbol{w}}_{j,S'}\|_1 - \lambda_n'\|\widehat{\boldsymbol{w}}_{j,\bar{S}'}\|_1 \le \lambda_n'\|\widehat{\boldsymbol{\Delta}}_{j,S'}\|_1 - \lambda_n'\|\widehat{\boldsymbol{\Delta}}_{j,\bar{S}'}\|_1.$$

We just need to bound I_1.

$$\begin{aligned}
&I_1 \\
=&E_n\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}^*)\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j}\right] \\
&+ E_n\left[\left\{(\widehat{\Omega}_+ - \Omega_+)\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) + (\widehat{\Omega}_- - \Omega_-)\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j}\right] \\
&+ E_n\left[\left\{\Omega_+(\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)) + \Omega_-(\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}}) - \phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}\right. \\
&\left.(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j}\right] \\
=&I_{11} + I_{12} + I_{13}.
\end{aligned}$$

By the proof of Theorem 4, on the event

$$\begin{aligned}
&\left\{\max_j\left\|E_n\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}^*)\boldsymbol{X}_{-j}\right]\right\|_\infty\right. \\
&\left.\le CR(\log p/n)^{1/2}\right\},
\end{aligned}$$

we have that

$$\begin{aligned}
|I_{11}| &\le \left\|E_n\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*) + \Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j - \boldsymbol{X}_{-j}^\top\boldsymbol{w}^*)\boldsymbol{X}_{-j}\right]\right\|_\infty \|\widehat{\boldsymbol{\Delta}}_j\|_1 \\
&\le CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_j\|_1.
\end{aligned}$$

For I_{12}, we have

$$\begin{aligned}&|I_{12}|\\ \leq&\|E_n\left[\left\{(\widehat{\Omega}_+-\Omega_+)\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})+(\widehat{\Omega}_--\Omega_-)\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\boldsymbol{X}_{-j}\right]\|_\infty\|\widehat{\boldsymbol{\Delta}}_j\|_1.\end{aligned}$$

and

$$\begin{aligned}&\|E_n\left[\left\{(\widehat{\Omega}_+-\Omega_+)\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})+(\widehat{\Omega}_--\Omega_-)\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\boldsymbol{X}_{-j}\right]\|_\infty\\ \leq&\|E_n\left[\left\{(\widehat{\Omega}_+-\Omega_+)\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)+(\widehat{\Omega}_--\Omega_-)\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}^*)\boldsymbol{X}_{-j}\right]\|_\infty\\ &+\|E_n\left[\left\{(\widehat{\Omega}_+-\Omega_+)(\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})-\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}^*)\boldsymbol{X}_{-j}\right]\|_\infty\\ &+\|E_n\left[\left\{(\widehat{\Omega}_--\Omega_-)(\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}^*)\boldsymbol{X}_{-j}\right]\|_\infty\\ =&I_{21}+I_{22}+I_{23}.\end{aligned}$$

First, by (the proof of) Claim 13, $\max_j I_{21}$ can be bounded by $O_p[R(n^{-\alpha}+n^{-\beta})(\log p/n)^{1/2}+Rn^{-\alpha-\beta}]$. Second, by the sub-Gaussian of $\boldsymbol{X}$, Lemma 10, and Claim 13, we have

$$\begin{aligned}\max_j I_{22}\leq&C\max_j\left\|E_n\left[\left\{|\widehat{\Omega}_+-\Omega_+|\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}_{-j}|X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}^*||\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)|\right]\right\|_\infty\\ \leq&C's^*\log p/\sqrt{n}\max_j\left\|E_n\left[\left\{|\widehat{\Omega}_+-\Omega_+|\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}_{-j}|X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}^*|\right]\right\|_\infty\\ \lesssim&R(n^{-\alpha}+n^{-\beta})s^*\log p/\sqrt{n}.\end{aligned}$$

Similarly, $\max I_{23}\lesssim R(n^{-\alpha}+n^{-\beta})s^*\log p/\sqrt{n}$. Thus,

$$|I_{12}|\lesssim o\left(R(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_j\|_1\right)$$

uniformly holds for all j's.

Now, we will bound I_{13}. Let $U(\boldsymbol{\beta})=\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta})+\Omega_+\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta})$.

$$\begin{aligned}&|I_{13}|\\ \leq&E_n\left[\left|(U(\widehat{\boldsymbol{\beta}})-U(\boldsymbol{\beta}^*))/(U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)))\right|\left|U^{1/2}(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)\right|\right.\\ &\left.\times\left|U^{1/2}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}^\top\boldsymbol{X}_{-j}\right|\right]\\ \leq&C\left(E_n[U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))^2(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2]\right)^{1/2}\left(E_n[U(\boldsymbol{\beta}^*)(\widehat{\boldsymbol{\Delta}}^\top\boldsymbol{X}_{-j})^2]\right)^{1/2}\\ =&C(\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j)^{1/2}\left(E_n[U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))^2(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2]\right)^{1/2},\end{aligned}$$

where $F_{-j}(\boldsymbol{\beta})=E_n[U(\boldsymbol{\beta})\boldsymbol{X}_{-j}\boldsymbol{X}_{-j}^\top]$. To bound

$$\max_j E_n[U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))^2(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2],$$

we have

$$\begin{aligned}
& \max_j E_n[U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))^2(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2] \\
\leq\ & \max_j \left|(E_n-E)[U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))^2(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2]\right| \\
& +\max_j E[U(\boldsymbol{\beta}^*)(\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))^2(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2] \\
\leq\ & \|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*\|_1^2\|(E_n-E)[U(\boldsymbol{\beta}^*)(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2]\boldsymbol{X}\boldsymbol{X}^\top\|_{\max} \\
& +C\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*\|_2^2\max_j E[(X_j-\boldsymbol{X}_{-j}^\top\boldsymbol{w}_j^*)^2]
\end{aligned}$$

By Lemma 10, we know that there exists a constant C such that

$$|I_{13}|\leq CR(\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j)^{1/2}(s^*\log p/n)^{1/2}$$

uniformly holds for all j's.

Taking $\lambda_n'\asymp R(\log p/n)^{1/2}$ and combining bound on I_1 and $\lambda_n'\|\boldsymbol{w}_j^*\|_1-\lambda_n'\|\widehat{\boldsymbol{w}}_j\|_1$, we have

$$\begin{aligned}
& P\left(\cap_j\left\{E_n\left[\left\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})+\widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j})^2\right]\right.\right. \\
& \leq CR(\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j)^{1/2}(s^*\log p/n)^{1/2} \\
& \left.\left.+3RC(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{j,S'}\|_1-CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{j,\bar{S}'}\|_1\right\}\right)\to 1.
\end{aligned}$$

Let $\widehat{F}_{-j}(\boldsymbol{\beta})=E_n[\widehat{U}(\boldsymbol{\beta})\boldsymbol{X}_{-j}\boldsymbol{X}_{-j}^\top]$, where $\widehat{U}(\boldsymbol{\beta})=\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta})+\widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta})$. To make the above inequality useful, we further link $\widehat{\boldsymbol{\Delta}}_j^\top\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j$ with $\widehat{\boldsymbol{\Delta}}_j^\top F(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j$. Consider

$$\begin{aligned}
& \left|\widehat{\boldsymbol{\Delta}}_j^\top\left[\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})-F_{-j}(\boldsymbol{\beta}^*)\right]\widehat{\boldsymbol{\Delta}}_j\right| \\
= & \left|E_n\left[\left\{\widehat{\Omega}_+\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})+\widehat{\Omega}_-\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})-\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)-\Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}(\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j})^2\right]\right| \\
\leq & \left|E_n\left[\left\{(\widehat{\Omega}_+-\Omega_+)\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})+(\widehat{\Omega}_--\Omega_-)\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}(\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j})^2\right]\right| \\
& +\left|E_n\left[\left\{\Omega_+(\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})-\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*))+\Omega_-(\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*))\right\}(\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j})^2\right]\right| \\
\leq & \left\|\widehat{\boldsymbol{\Delta}}_j\right\|_1^2\left\|E_n\left[\left\{(\widehat{\Omega}_+-\Omega_+)\phi''(\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})+(\widehat{\Omega}_--\Omega_-)\phi''(-\boldsymbol{X}^\top\widehat{\boldsymbol{\beta}})\right\}\boldsymbol{X}\boldsymbol{X}^\top\right]\right\|_{\max} \\
& +C\left|E_n\left[\left\{\Omega_+\phi''(\boldsymbol{X}^\top\boldsymbol{\beta}^*)+\Omega_-\phi''(-\boldsymbol{X}^\top\boldsymbol{\beta}^*)\right\}\boldsymbol{X}^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)(\widehat{\boldsymbol{\Delta}}_j^\top\boldsymbol{X}_{-j})^2\right]\right| \\
\leq & C(n^{-\alpha}+n^{-\beta})\|\widehat{\boldsymbol{\Delta}}_j\|_1^2+C|\widehat{\boldsymbol{\Delta}}_j^\top F(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j|s^*\log(np)/n^{1/2}.
\end{aligned}$$

Thus, for some constant C'', we have

$$\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j\leq\left(1+C''s^*\log(np)/n^{1/2}\right)\widehat{\boldsymbol{\Delta}}_j^\top\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j+C''(n^{-\alpha}+n^{-\beta})\|\widehat{\boldsymbol{\Delta}}_j\|_1^2,$$

uniformly hold over all j's. Combining with the inequality above, we have

$$\begin{aligned}
& \widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j \\
\leq\ & CR(\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j)^{1/2}(s^*\log p/n)^{1/2} \\
& +3CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 - CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1 \\
\leq\ & CR\sqrt{\frac{s^*\log p}{n}}\left(\left(1+C''s^*\log(np)/n^{1/2}\right)\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j + C''(n^{-\alpha}+n^{-\beta})\|\widehat{\boldsymbol{\Delta}}_j\|_1^2\right)^{1/2} \\
& +3CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 - CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1,
\end{aligned}$$

uniformly holds over all j's.

Notice that $(n^{-\alpha}+n^{-\beta})s^* \to 0$ and $s^*\log(np)/\sqrt{n} \to 0$, we have

$$\begin{aligned}
& \widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j \leq C'R(s^*\log p/n)^{1/2}(\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j)^{1/2} \\
& +3CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 - CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1, \qquad (14)
\end{aligned}$$

uniformly holds over all j's with a sufficient large C' and C.

If $\left(\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j\right)^{1/2} \leq C'R(s^*\log p/n)^{1/2}$, Inequality (14) holds trivially. If

$$\left(\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j\right)^{1/2} > C'R(s^*\log p/n)^{1/2},$$

this implies that $3\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 \geq \|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1$. Due to the claim with $\xi = 3$, we conclude that $\|\widehat{\boldsymbol{\Delta}}_{s',j}\|_1 \leq C(s')^{1/2}(\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j)^{1/2}$. Combining with Inequality (14), we have $(\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j)^{1/2} \lesssim R(\max\{s^*, s'\}\log p/n)^{1/2}$. Thus,

$$\widehat{\boldsymbol{\Delta}}_j^\top \widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j \lesssim R^2\max\{s^*, s'\}\log p/n$$

uniformly holds over all j's.

Now, we consider to bound $\|\widehat{\boldsymbol{\Delta}}_j\|_1$. First, if $6\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 \geq \|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1$, then we have $\|\widehat{\boldsymbol{\Delta}}_j\|_1 \leq 7\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 \lesssim (s')^{1/2}(\widehat{\boldsymbol{\Delta}}_j\widehat{F}_{-j}(\widehat{\beta})\widehat{\boldsymbol{\Delta}}_j)^{1/2}$ by claim with $\xi = 6$. Therefore, we obtain that

$$\|\widehat{\boldsymbol{\Delta}}_j\|_1 \lesssim R\max\{s^*, s'\}(\log p/n)^{1/2}.$$

Otherwise, we have $6\|\widehat{\boldsymbol{\Delta}}_{S',j}\|_1 \leq \|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1$. Then Inequality (14) implies that

$$\widehat{\boldsymbol{\Delta}}_j\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j \leq C'R(s^*\log p/n)^{1/2}(\widehat{\boldsymbol{\Delta}}_j\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j)^{1/2} - CR(\log p/n)^{1/2}\|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1/2.$$

Hence,

$$\|\widehat{\boldsymbol{\Delta}}_j\|_1 \leq 7/6\|\widehat{\boldsymbol{\Delta}}_{\bar{S}',j}\|_1 \lesssim R(s^*)^{1/2}(\widehat{\boldsymbol{\Delta}}_j\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\Delta}}_j)^{1/2} \lesssim R\max\{s^*, s'\}(\log p/n)^{1/2},$$

uniformly holds for all j's.

Because

$$\left|\widehat{\boldsymbol{\Delta}}_j^\top\left[\widehat{F}_{-j}(\widehat{\boldsymbol{\beta}}) - F(\boldsymbol{\beta}^*)\right]\widehat{\boldsymbol{\Delta}}_j\right| \lesssim (n^{-\alpha}+n^{-\beta})\|\widehat{\boldsymbol{\Delta}}_j\|_1^2 + |\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j|s^*\log p/\sqrt{n}$$

uniformly holds over all j's, we have $\widehat{\boldsymbol{\Delta}}_j^\top F_{-j}(\boldsymbol{\beta}^*)\widehat{\boldsymbol{\Delta}}_j \lesssim R\max\{s^*, s'\}\log p/n$ holds uniformly over all j's. □

References

Radosław Adamczak and Paweł Wolff. Concentration inequalities for non-lipschitz functions with bounded derivatives of higher order. *Probab Theory Relat Fields*, 162(3):531–586, 2015.

Susan Athey and Stefan Wager. Efficient Policy Learning. Research Papers 3506, Stanford University, Graduate School of Business, February 2017.

Susan Athey, Guido W. Imbens, and Stefan Wager. Approximate residual balancing: debiased inference of average treatment effects in high dimensions. *J. R. Statist. Soc. B*, 80 (4):597–623, 2018.

David R Brillinger. A generalized linear model with "gaussian" regressor variables. In *Selected Works of David Brillinger*, pages 589–606. Springer, 2012.

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.

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

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *Econom. J.*, 21(1):C1–C68, 2018.

Ronan Collobert, Fabian Sinz, Jason Weston, and Léon Bottou. Trading convexity for scalability. In *International Conference on Machine Learning*, pages 201–208, 2006a.

Ronan Collobert, Fabian Sinz, Jason Weston, Léon Bottou, and Thorsten Joachims. Large scale transductive svms. *J Mach Learn Res*, 7(8), 2006b.

Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Mach Learn*, 20(3):273–297, 1995.

Hengjian Cui, Runze Li, and Wei Zhong. Model-free feature screening for ultrahigh dimensional discriminant analysis. *J. Am. Stat. Assoc.*, 110(510):630–641, 2015.

M. Davidian, A.A. Tsiatis, and E.B. Laber. *Value search estimators.* Springer, 2014.

Ruben Dezeure, Peter Bühlmann, and Cun-Hui Zhang. High-dimensional simultaneous inference with the bootstrap. *TEST*, 26(4):685–719, 2017.

Naihua Duan and Ker-Chau Li. Slicing regression: a link-free regression method. *Ann. Statist.*, pages 505–530, 1991.

Jianqing Fan and Jinchi Lv. Nonconcave penalized likelihood with np-dimensionality. *IEEE Trans. Inf. Theory*, 57(8):5467–5484, 2011.

Yair Goldberg, Rui Song, Donglin Zeng, and Michael R Kosorok. Comment on "dynamic treatment regimes: Technical challenges and applications". *Electron. J. Statist.*, 8(1): 1290, 2014.

Peter Hall and Ker-Chau Li. On almost linearity of low dimensional projections from high dimensional data. *Ann. Statist.*, pages 867–889, 1993.

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press, 2015.

X. Jessie Jeng, Wenbin Lu, and Huimin Peng. High-dimensional inference for personalized treatment decision. *Electron. J. Statist.*, 12(1):2074–2089, 2018.

Eric B. Laber, Kristin A. Linn, and Leonard A. Stefanski. Interactive model building for Q-learning. *Biometrika*, 101(4):831–847, 10 2014a.

Eric B Laber, Daniel J Lizotte, Min Qian, William E Pelham, and Susan A Murphy. Dynamic treatment regimes: technical challenges and applications. *Electron. J. Statist.*, 8(1):1225, 2014b.

Eric B. Laber, Daniel J. Lizotte, Min Qian, William E. Pelham, and Susan A. Murphy. Rejoinder of "Dynamic treatment regimes: Technical challenges and applications". *Electron. J. Statist.*, 8(1):1312 – 1321, 2014c.

Ker-Chau Li. Sliced inverse regression for dimension reduction. *J. Am. Stat. Assoc.*, 86 (414):316–327, 1991.

Ker-Chau Li and Naihua Duan. Regression analysis under link violation. *Ann. Statist.*, 17 (3):1009–1052, 1989.

Runze Li, Wei Zhong, and Liping Zhu. Feature screening via distance correlation learning. *J. Am. Stat. Assoc.*, 107(499):1129–1139, 2012.

Muxuan Liang, Ting Ye, and Haoda Fu. Estimating individualized optimal combination therapies through outcome weighted deep learning algorithms. *Stat. Med.*, 37(27):3869–3886, 2018a.

Shuhan Liang, Wenbin Lu, Rui Song, and Lan Wang. Sparse concordance-assisted learning for optimal treatment decision. *J. Mach. Learn. Res.*, 18(202):1–26, 2018b.

Qian Lin, Zhigen Zhao, Jun S Liu, et al. On consistency and sparsity for sliced inverse regression in high dimensions. *Ann. Statist.*, 46(2):580–610, 2018.

Qian Lin, Zhigen Zhao, and Jun S Liu. Sparse sliced inverse regression via lasso. *J. Am. Stat. Assoc.*, 114(528):1726–1739, 2019.

Po-Ling Loh and Martin J Wainwright. High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. In *Advances in Neural Information Processing Systems*, volume 24, 2011.

Po-Ling Loh and Martin J. Wainwright. Regularized m-estimators with nonconvexity: Statistical and algorithmic theory for local optima. *J. Mach. Learn. Res.*, 16(19):559–616, 2015.

Wenbin Lu, Hao Helen Zhang, and Donglin Zeng. Variable selection for optimal treatment decision. *Stat. Methods Med. Res.*, 22(5):493–504, 2013.

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, 2016.

Rong Ma, T Tony Cai, and Hongzhe Li. Global and simultaneous hypothesis testing for high-dimensional logistic regression models. *J. Am. Stat. Assoc.*, 116(534):984–998, 2021.

Susan A Murphy. Optimal dynamic treatment regimes. *J. R. Statist. Soc. B*, 65(2):331–355, 2003.

Whitney K. Newey. Convergence rates and asymptotic normality for series estimators. *J. Econom.*, 79(1):147–168, 1997.

Yang Ning and Han Liu. A general theory of hypothesis tests and confidence regions for sparse high dimensional models. *Ann. Statist.*, 45(1):158–195, 2017.

Yang Ning, Peng Sida, and Kosuke Imai. Robust estimation of causal effects via a high-dimensional covariate balancing propensity score. *Biometrika*, 107(3):533–554, 2020.

M. Qian and S. A. Murphy. Performance guarantees for individualized treatment rules. *Ann. Statist.*, 39(2):1180, 2011.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.*, 66(5):688, 1974.

Donald B Rubin. Causal inference using potential outcomes. *J. Am. Stat. Assoc.*, 100(469): 322–331, 2005.

Xiaotong Shen, George C Tseng, Xuegong Zhang, and Wing Hung Wong. On ψ-learning. *J. Am. Stat. Assoc.*, 98(463):724–734, 2003.

Chengchun Shi, Rui Song, and Wenbin Lu. Robust learning for optimal treatment decision with np-dimensionality. *Electron. J. Statist.*, 10(2):2894–2921, 2016.

Chengchun Shi, Ailin Fan, Rui Song, and Wenbin Lu. High-dimensional a-learning for optimal dynamic treatment regimes. *Ann. Statist.*, 46(3):925–957, 06 2018.

Chengchun Shi, Rui Song, Zhao Chen, Runze Li, et al. Linear hypothesis testing for high dimensional generalized linear models. *Ann. Statist.*, 47(5):2671–2703, 2019.

Chengchun Shi, Wenbin Lu, and Rui Song. Breaking the curse of nonregularity with subagging: inference of the mean outcome under optimal treatment regimes. *J. Mach. Learn. Res.*, 21(176):1–67, 2020.

E Smucler, A Rotnitzky, and JM Robins. A unifying approach for doubly-robust ℓ_1 regularized estimation of causal contrasts. *arXiv preprint arXiv:1904.03737*, 2019.

Rui Song, Michael Kosorok, Donglin Zeng, Yingqi Zhao, Eric Laber, and Ming Yuan. On sparse representation for optimal individualized treatment selection with penalized outcome weighted learning. *Stat*, 4(1):59–68, 2015.

Rui Song, Shikai Luo, Donglin Zeng, Hao Helen Zhang, Wenbin Lu, and Zhiguo Li. Semiparametric single-index model for estimating optimal individualized treatment strategy. *Electron. J. Statist.*, 11(1):364–384, 2017.

Zhiqiang Tan. Model-assisted inference for treatment effects using regularized calibrated estimation with high-dimensional data. *Ann. Statist.*, 48(2):811–837, 2020.

Stef van Buuren and Karin Groothuis-Oudshoorn. Mice: Multivariate imputation by chained equations in r. *J. Stat. Softw.*, 45(3):1–67, 2011.

Sara van de Geer, Peter Bühlmann, Ya'acov Ritov, and Ruben Dezeure. On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.*, 42(3):1166–1202, 06 2014.

Sara A. van de Geer. High-dimensional generalized linear models and the lasso. *Ann. Statist.*, 36(2):614–645, 04 2008.

Aad W van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Mach. Learn.*, 8(3):279–292, 1992.

Yunan Wu, Lan Wang, and Haoda Fu. Model-assisted uniformly honest inference for optimal treatment regimes in high dimension. *J. Am. Stat. Assoc.*, in print, 2021.

Yaoyao Xu, Menggang Yu, Ying-Qi Zhao, Quefeng Li, Sijian Wang, and Jun Shao. Regularized outcome weighted subgroup identification for differential treatment effects. *Biometrics*, 71(3):645–653, 2015.

Baqun Zhang and Min Zhang. C-learning: A new classification framework to estimate optimal dynamic treatment regimes. *Biometrics*, 74(3):891–899, 2018.

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.

Ying-Qi Zhao, Eric B Laber, Yang Ning, Sumona Saha, and Bruce E Sands. Efficient augmentation and relaxation learning for individualized treatment rules using observational data. *J. Mach. Learn. Res.*, 20(48):1–23, 2019.

Yingqi Zhao, Donglin Zeng, A. John Rush, and Michael R. Kosorok. Estimating individualized treatment rules using outcome weighted learning. *J. Am. Stat. Assoc.*, 107(499):1106–1118, 2012.

Xuan Zhou, Yuanjia Wang, and Donglin Zeng. Outcome-weighted learning for personalized medicine with multiple treatment options. In *IEEE 5th international conference on data science and advanced analytics*, pages 565–574. IEEE, 2018.

Lixing Zhu, Baiqi Miao, and Heng Peng. On sliced inverse regression with high-dimensional covariates. *J. Am. Stat. Assoc.*, 101(474):630–643, 2006.