Electronic Journal of Statistics
Vol. 0 (0000)
ISSN: 1935-7524
DOI: 10.1214/154957804100000000

High-Dimensional Inference for Personalized Treatment Decision

X. Jessie Jeng*

e-mail: xjjeng@ncsu.edu

Wenbin Lu†

e-mail: wlu4@ncsu.edu

and

Huimin Peng

e-mail: hpeng2@ncsu.edu

Department of Statistics, North Carolina State University, SAS Hall, 2311 Stinson Dr., Raleigh, NC 27695-8203

Abstract

Recent development in statistical methodology for personalized treatment decision has utilized high-dimensional regression to take into account a large number of patients' covariates and described personalized treatment decision through interactions between treatment and covariates. While a subset of interaction terms can be obtained by existing variable selection methods to indicate relevant covariates for making treatment decision, there often lacks statistical interpretation of the results. This paper proposes an asymptotically unbiased estimator based on Lasso solution for the interaction coefficients. We derive the limiting distribution of the estimator when baseline function of the regression model is unknown and possibly misspecified. Confidence intervals and p-values are derived to infer the effects of the patients' covariates in making treatment decision. We confirm the accuracy of the proposed method and its robustness against misspecified function in simulation and apply the method to STAR*D study for major depression disorder.

MSC 2010 subject classifications: Primary 62J05, 62F35; secondary 62P10.
Keywords and phrases: Large p Small n, Model Misspecification, Optimal Treatment Regime, Robust Regression.

1. Introduction

Precision medicine aims to design optimal treatment for each individual according to their specific conditions in the hope of lowering medical cost and improv-

*The research of X.J. Jeng is supported in part by Grant NSF-DMS 1811360.
†The research of W. Lu is supported in part by Grant NCI P01CA142538.

ing efficacy of treatments. Existing methods can be broadly partitioned into regression-based or classification-based approaches. Popular regression-based approaches include Q-learning [21, 11, 3, 6, 7, 17] and A-learning [13, 10, 8, 16, 15]. Q-learning models the conditional mean of the outcome given covariates and treatment while A-learning directly models the interaction between treatment and covariates that is sufficient for treatment decisions. Other regression-based methods have been developed in [18], [12], [19], etc. In contrast, classification-based approaches, also known as policy search or value search, estimate the marginal mean of the outcome for every treatment regime within a pre-specified class and then take the maximizer as the estimated optimal regime; see, for example, [13], [22], [23], [25], [26]. As emerging technology makes it possible to gather an extraordinarily large number of prognostic factors for each individual, such as genetic information and clinical measurements, regression-based and classification-based methods have been extended to handle high-dimensional data [12, 25, 26, 19].

This paper focuses on regression-based approaches and plans to address the current limitation that the selected interaction effects often lack statistical interpretation when the number of covariates exceeds the sample size. Consider the following semi-parametric model. Denote $\mathbf{X}$ as the $n \times p$ design matrix, where n is the number of patients and p is the number of patients' covariates. Let $\tilde{\mathbf{X}}$ be the design matrix with an additional column of 1's and $\tilde{X}_i$ the i-th row of $\tilde{\mathbf{X}}$. Denote Y_i as the outcome and $A_i = \{0, 1\}$ the treatment assignment of the i-th patient. Assume

$$Y_i = \mu(X_i) + A_i(\beta_0^T \tilde{X}_i) + \epsilon_i, \tag{1}$$

where $\mu(X_i)$ is an unspecified baseline function and β_0 is the vector of unknown coefficients for the interactions between treatment and patients' covariates. We are interested in inference on β_0 and optimal treatment decisions based on the estimated β_0.

We propose an asymptotically unbiased estimator for β_0 and derive its limiting distribution when $p > n$. Consequently, confidence intervals and p-values of the interaction coefficients can be calculated. Because the baseline function is unknown and possibly misspecified, we investigate the robustness of the estimator to misspecified $\mu(X_i)$.

The proposed estimator for β_0 and the p-value calculated from data can be utilized to make significance-based optimal treatment decision. We illustrate the efficiency of the new treatment regime in a real application to STAR*D study for major depression disorder.

2. Method and Theory

2.1. An A-learning framework

We adopt the robust regression approach of [16] and transform the interaction from $A_i(\beta_0^T \tilde{X}_i)$ to $(A_i - \pi(X_i))(\beta_0^T \tilde{X}_i)$, where $\pi(X_i) = E(A_i|X_i)$ is the propensity score. Because $E(A_i - \pi|X_i) = 0$, the transformed interaction is orthogonal

to the baseline function $\mu(X_i)$ given X_i. By doing so, we can protect the estimation of β_0 from the effect of baseline function misspecification [8, 16]. For simplicity of presentation, we consider a completely randomized study with prespecified propensity score, i.e. $A = \{0, 1\}$ and $\pi(X_i) = P(A_i = 1|X_i) = \pi$. Let $\mu_\pi(X_i) = \mu(X_i) + \pi(\beta_0^T \tilde{X}_i)$, then (1) is equivalent to

$$Y_i = \mu_\pi(X_i) + (A_i - \pi)(\beta_0^T \tilde{X}_i) + \epsilon_i.$$

The true $\mu_\pi(\cdot)$ is unknown. Let $\hat{\mu}_\pi(\cdot)$ be an estimator of $\mu_\pi(\cdot)$ that does not necessarily converge to the true $\mu_\pi(\cdot)$. Assume $\hat{\mu}_\pi(\cdot) \to \mu^*_\pi(\cdot)$ uniformly for some function $\mu^*_\pi(\cdot)$. Because $E(A_i - \pi|X_i) = 0$, then $\mu^*_\pi(X_i)$ and $(A_i - \pi)\beta_0^T \tilde{X}_i$ are orthogonal, which implies that

$$\beta_0 = \text{argmin}_\beta \left\{ E \left[Y - \mu^*_\pi(\mathbf{X}) - D(A, \pi)\tilde{\mathbf{X}}\beta \right]^2 \right\}, \tag{2}$$

where $D(A, \pi) = diag\{A_1 - \pi, \ldots, A_n - \pi\}$. We apply the Lasso to regress $Y - \hat{\mu}_\pi(\mathbf{X})$ on $D(A, \pi)\tilde{\mathbf{X}}$ and obtain

$$\hat{\beta} = \text{argmin}_\beta \left\{ \|Y - \hat{\mu}_\pi(\mathbf{X}) - D(A, \pi)\tilde{\mathbf{X}}\beta\|_2^2/n + 2\lambda_{n,p}\|\beta\|_1 \right\}. \tag{3}$$

It is well-known that the limiting distribution of Lasso solution is difficult to derive. The misspecified baseline function adds another layer of difficulty [2].

2.2. Unbiased estimation with misspecified baseline function

We propose a de-sparsified estimator for β_0 motivated by [24] and [20]:

$$\hat{b} = \hat{\beta} + \hat{\mathbf{\Theta}} \left\{ D(A, \pi)\tilde{\mathbf{X}} \right\}^T \left\{ Y - \hat{\mu}_\pi(\mathbf{X}) - D(A, \pi)\tilde{\mathbf{X}}\hat{\beta} \right\} /n, \tag{4}$$

where $\hat{\mathbf{\Theta}}$ is an estimator of the precision matrix of $D(A, \pi)\tilde{\mathbf{X}}$.

We show that this estimator is asymptotically unbiased and derive its limiting distribution as follows. Decompose $\hat{b} - \beta_0$ into three terms:

$$\hat{b} - \beta_0 = \eta - \Delta_1 - \Delta_2,$$

where $\eta = \hat{\mathbf{\Theta}}\{D(A, \pi)\tilde{\boldsymbol{X}}\}^T \left(Y - \mu^*_\pi(\mathbf{X}) - D(A, \pi)\tilde{\boldsymbol{X}}\beta_0 \right) /n$, $\Delta_1 = (\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}} - \boldsymbol{I})(\hat{\beta} - \beta_0)$, $\Delta_2 = \hat{\mathbf{\Theta}}\{D(A, \pi)\tilde{\boldsymbol{X}}\}^T \left((\hat{\mu}_\pi(\mathbf{X}) - \mu^*_\pi(\mathbf{X})) \right) /n$, and $\hat{\mathbf{\Sigma}}$ is the sample covariance matrix of $D(A, \pi)\tilde{\boldsymbol{X}}$.

Consequently, for an arbitrary $q \times (p + 1)$ matrix, $\mathbf{H}$, we can decompose $\sqrt{n}\mathbf{H}(\hat{b} - \beta_0)$ into three terms. The reason to study $\sqrt{n}\mathbf{H}(\hat{b} - \beta_0)$ is to explore the asymptotic joint distribution of q arbitrary linear contrasts of $\{\sqrt{n}(\hat{b}_1 - \beta_{0,1}), \ldots, \sqrt{n}(\hat{b}_p - \beta_{0,p})\}$, which includes the limiting distribution of a single $\sqrt{n}(\hat{b}_j - \beta_{0,j})$ as a special case. We derive the limiting distribution of $\sqrt{n}\mathbf{H}(\hat{b} - \beta_0)$ by showing that $\|\sqrt{n}\mathbf{H}\Delta_1\|_\infty = o_P(1)$, $\|\sqrt{n}\mathbf{H}\Delta_2\|_\infty = o_P(1)$, and $\sqrt{n}\mathbf{H}\eta \to_d N(0, G)$.

We first show a preliminary result on the lasso solution $\hat{\beta}$ with misspecified baseline function. As $\hat{\beta} - \beta_0$ is an important component in Δ_1, this preliminary result helps to derive the asymptotic properties of $\sqrt{n}\mathbf{H}\Delta_1$. Consider the following assumptions, where C denotes a generic constant, possibly varying from place to place:

Condition 1: The random error ϵ_i of the true model (1) are independent and identically distributed with $E(\epsilon_i|\tilde{X}, A) = 0$ and $E(\epsilon_i^2|\tilde{X}, A) \leq C$.

Condition 2: There exists a $\mu_\pi^*(\cdot)$ such that for any $\tilde{X}_i \in R^{p+1}$,
(a) $\sup_{\tilde{X}_i} |\hat{\mu}_\pi(\tilde{X}_i) - \mu_\pi^*(\tilde{X}_i)| = o_p(1)$ and
(b) $E(\mu_\pi^*(\tilde{X}_i) - \mu_\pi(\tilde{X}_i))^2 \leq C$.

Condition 3: Denote $\tilde{X}^j$ as the jth column of $\tilde{\boldsymbol{X}}$. $\|\tilde{X}^j\|_\infty \leq C$ for all $1 \leq j \leq p+1$.

Condition 4: Define $\boldsymbol{\Sigma}$ as the covariance matrix of $D(A, \pi)\tilde{\boldsymbol{X}}$. $\boldsymbol{\Sigma}$ has smallest eigenvalue $\Lambda_{min}^2(\boldsymbol{\Sigma}) \geq C > 0$.

Condition 5: The number of important interaction terms satisfies that $s_0 = \|\beta_0\|_0 = o\{\sqrt{n}/\log(p)\}$.

Lemma 2.1. *Consider model (1). Assume Conditions 1 - 5. The lasso solution from (3) with $\lambda_{n,p} \asymp \sqrt{\log(p)/n}$ satisfies*

$$\|\hat{\beta} - \beta_0\|_1 = o_p(1/\sqrt{\log p}). \tag{5}$$

Next, we show that $\sqrt{n}\mathbf{H}\Delta_1$ and $\sqrt{n}\mathbf{H}\Delta_2$ are at the order of $o_p(1)$. Similar to [20], we apply lasso for nodewise regression to construct $\hat{\boldsymbol{\Theta}}$. Details of the construction are presented in Section A.1. Consider additional assumptions:

Condition 6: The prespecified matrix $\mathbf{H}_{q\times(p+1)}$ satisfies
(a) $\|\mathbf{H}\|_\infty \leq C$ and
(b) $h_t = |\{j \neq t : H_{t,j} \neq 0\}| \leq C$ for any $t \in \{1, \ldots, q\}$.

Condition 7: The precision matrix $\boldsymbol{\Theta}$ of $D(\boldsymbol{A}, \pi)\tilde{\boldsymbol{X}}$ satisfies
(a) $\|\boldsymbol{\Theta}\|_\infty \leq C$ and
(b) $s_j = |\{k \neq j : \Theta_{j,k} \neq 0\}| \leq C$ for any $j \in \{1, \ldots, p\}$.

Lemma 2.2. *Consider model (1). Assume Conditions 1 - 7. Obtain $\hat{\beta}$ by (3) with $\lambda_{n,p} \asymp \sqrt{\log(p)/n}$ and $\hat{\boldsymbol{\Theta}}$ by nodewise regression with $\lambda_j \asymp \sqrt{\log(p)/n}$ for any $j \in \{1, \ldots, p\}$. Then*

$$\|\sqrt{n}\mathbf{H}\Delta_1\|_\infty = o_P(1) \tag{6}$$

$$\|\sqrt{n}\mathbf{H}\Delta_2\|_\infty = o_P(1). \tag{7}$$

The remaining component of $\sqrt{n}\mathbf{H}(\hat{b} - \beta_0)$ is $\sqrt{n}\mathbf{H}\eta$. We show that $\sqrt{n}\mathbf{H}\eta$ converges to a multivariate normal distribution. Since η does not involve $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0$, less conditions are needed in the following lemma.

Lemma 2.3. *Consider model (1). Assume Conditions 1- 3 and 6 - 7. Obtain $\hat{\boldsymbol{\Theta}}$ by nodewise regression with $\lambda_j \asymp \sqrt{\log(p)/n}$ for any $j \in \{1, \ldots, p\}$. Then*

$$\sqrt{n}\mathbf{H}\eta \to_d N(0, \mathbf{G}), \tag{8}$$

where $\mathbf{G} = \mathbf{H\Theta V\Theta}^T\mathbf{H}^T$ *is a* $q \times q$ *nonnegative matrix,* $\mathbf{V} = E\left[(A_i - \pi)^2\sigma_\epsilon^2(A_i, \tilde{X}_i)\tilde{X}_i\tilde{X}_i^T\right] + \pi(1-\pi)E\left[(\mu_\pi(\tilde{X}_i) - \mu_\pi^*(\tilde{X}_i))^2\tilde{X}_i\tilde{X}_i^T\right]$, $\sigma_\epsilon^2(A_i, \tilde{X}_i) = Var(\epsilon_i|A_i, \tilde{X}_i)$, *and* $\|\mathbf{G}\|_\infty < \infty$.

We summarize the above results in the following theorem.

Theorem 2.4. *Consider model (1). Assume Conditions 1 - 7. Obtain* $\hat{\beta}$ *by (3) with* $\lambda_{n,p} \asymp \sqrt{\log(p)/n}$ *and* $\hat{\mathbf{\Theta}}$ *by nodewise regression with* $\lambda_j \asymp \sqrt{\log(p)/n}$ *for any* $j \in \{1, \ldots, p\}$. *Then,*

$$\sqrt{n}\mathbf{H}(\hat{b} - \beta_0) \to_d N(0, \mathbf{G}),$$

where $\mathbf{G}$ *is defined in (8).*

2.3. Some examples for $\hat{\mu}_\pi$

In real applications, we need to choose an estimator $\hat{\mu}_\pi$ for the baseline function. The only requirement for $\hat{\mu}_\pi$ is *Condition 2*, which is very general. Here we discuss two examples for $\hat{\mu}_\pi$.

Example 1: $\hat{\mu}_\pi(\tilde{X}_i) = constant$. In this case, *Condition 2* (a) holds trivially. *Condition 2* (b) is implied by

Condition 8: $E(Y_i^2) \le C$.

The following corollary presents the limiting distribution of $\sqrt{n}\mathbf{H}(\hat{b} - \beta_0)$ in this case. Proof of this corollary is straightforward and, thus, omitted.

Corollary 2.5. *Consider model (1). Implement* $\hat{\mu}_\pi(\tilde{X}_i) = constant$. *Assume Conditions 1, 3 - 7, and 8. Obtain* $\hat{\beta}$ *by (3) with* $\lambda_{n,p} \asymp \sqrt{\log(p)/n}$ *and* $\hat{\mathbf{\Theta}}$ *by nodewise regression with* $\lambda_j \asymp \sqrt{\log(p)/n}$ *for any* $j \in \{1, \ldots, p\}$. *Then,*

$$\sqrt{n}\mathbf{H}(\hat{b} - \beta_0) \to_d N(0, \mathbf{G}),$$

where $\mathbf{G}$ *is defined in (8).*

Example 2: $\hat{\mu}_\pi(\tilde{X}_i) = \hat{\gamma}^T\tilde{X}_i$, where $\hat{\gamma}$ is obtained by

$$(\hat{\gamma}, \hat{\beta}) = \text{argmin}_{\gamma,\beta}\left\{\|Y - \tilde{X}\gamma - D(A, \pi)\tilde{X}\beta\|_2^2/n + 2\lambda_{n,p}(\|\gamma\|_1 + \|\beta\|_1)\right\}. \quad (9)$$

Note that the solution $\hat{\beta}$ from (9) is equivalent to the solution from (3) (detail of the proof is included in the proof of Corollary 2.6). In this case, we replace *Condition 2* with

Condition 9: Define $\gamma^* = \text{argmin}_\gamma\left\{E(\mu_\pi(\tilde{X}_i) - \gamma^T\tilde{X}_i)^2\right\}$; γ^* satisfies

(a) $s_\gamma = \|\gamma^*\|_0 = o(\sqrt{n}/\log(p))$ and

(b) $E(\gamma^{*T}\tilde{X}_i - \mu_\pi(\tilde{X}_i))^2 \le C$

Corollary 2.6. *Consider model (1). Implement* $\hat{\mu}_\pi(\tilde{X}_i) = \hat{\gamma}^T\tilde{X}_i$ *and obtain* $\hat{\gamma}$ *and* $\hat{\beta}$ *from (9) with* $\lambda_{n,p} \asymp \sqrt{\log(p)/n}$. *Assume Conditions 1, 3 - 7, and 9.*

Obtain $\hat{\mathbf{\Theta}}$ *by nodewise regression with* $\lambda_j \asymp \sqrt{\log(p)/n}$ *for any* $j \in \{1, \ldots, p\}$. *Then,*

$$\sqrt{n}\mathbf{H}(\hat{b} - \beta_0) \to_d N(0, \mathbf{G}),$$

where $\mathbf{G}$ *is defined in (8).*

2.4. Interval estimation and p-value

The theoretical result on the asymptotic normality of the de-sparsified estimator $\hat{b}$ can be utilized to construct confidence intervals for the coefficients of interest and to perform hypothesis testing. The variance-covariance matrix $\mathbf{G}$ can be approximated by $\mathbf{H}\hat{\mathbf{\Theta}}\hat{\mathbf{V}}\hat{\mathbf{\Theta}}^T\mathbf{H}^T$ where

$$\hat{\mathbf{V}} = \frac{1}{n}\sum_{i=1}^{n}(A_i - \pi)^2 \left\{Y_i - \hat{\mu}_\pi(\tilde{X}_i) - (A_i - \pi)\hat{\boldsymbol{\beta}}^T\tilde{X}_i\right\}^2 \tilde{X}_i\tilde{X}_i^T.$$

Therefore, a pointwise $(1-\alpha)$ confidence interval of $\boldsymbol{\beta}_{0,j}$ can be constructed as

$$\left\{\hat{b}_j - c(\alpha, n),\ \hat{b}_j + c(\alpha, n)\right\}, \tag{10}$$

where $c(\alpha, n) = \Phi^{-1}(1-\alpha/2)\sqrt{(\mathbf{H}\hat{\mathbf{\Theta}}\hat{\mathbf{V}}\hat{\mathbf{\Theta}}^T\mathbf{H}^T)_{j,j}/n}$, and $\Phi(\cdot)$ is the c.d.f of $N(0,1)$.

One can also calculate the asymptotic p-value for testing

$$H_0 : \boldsymbol{\beta}_{0,j} = 0 \qquad \text{vs.} \qquad H_A : \boldsymbol{\beta}_{0,j} \neq 0$$

for a given $j \in \{1, \ldots, p\}$. A non-zero $\boldsymbol{\beta}_{0,j}$ means that variant X_j is relevant in making treatment decision for a patient.

3. Simulation

We consider three simulation settings with different baseline functions:

- Case 1: $Y = 1 + \boldsymbol{X}\gamma + A(\tilde{\boldsymbol{X}}\beta_0) + \epsilon$,
- Case 2: $Y = 1 + 0.5(1 + \boldsymbol{X}\gamma)^2 + A(\tilde{\boldsymbol{X}}\beta_0) + \epsilon$,
- Case 3: $Y = 1 + 1.5sin(\pi\boldsymbol{X}\gamma) + X_1^2 + A(\tilde{\boldsymbol{X}}\beta_0) + \epsilon$, where X_1 is the first covariate of $\boldsymbol{X}$.

We simulate the random error ϵ from $N(0,1)$ and the design matrix $\boldsymbol{X}$ from multivariate normal $MN(0, \mathbf{\Sigma})$, where $\mathbf{\Sigma} = AR(0.5)$. The parameters in the baseline functions is set as $\gamma = (1, 1, 0, \cdots, 0)^T$, and the parameter of interest β_0 has s nonzero elements with values 1 or 1.5. The treatment main effect is set at zero.

We apply our de-sparsified estimator $\hat{b}$ in (4) with $\hat{\mu}_\pi(\tilde{X}_i) = \hat{\gamma}^T\tilde{X}_i$ and derive confidence intervals by (10). To evaluate the finite-sample performance of our method, we report the following measures. Denote CI_j as the 95% confidence interval for $\beta_{0,j}$ and $S_{true} = \{j \in \{2, \ldots p+1\} : \beta_{0,j} \neq 0\}$.

- MAB($\hat{b}$): the mean absolute bias of $\hat{b}$ for relevant variables, calculated as $s^{-1}\sum_{j\in S_{true}}|\hat{b}_j-\beta_{0,j}|$.
- MAB($\hat{\beta}$): the mean absolute bias of lasso solution for relevant variables, calculated as $s^{-1}\sum_{j\in S_{true}}|\hat{\beta}_j-\beta_{0,j}|$.
- Coverage for noise variables: the empirical value of $(p-s)^{-1}\sum_{j\in S^c_{true}} P[0\in \mathrm{CI}_j]$.
- Coverage for relevant variables: the empirical value of $s^{-1}\sum_{j\in S_{true}} P[\beta_{0,j}\in \mathrm{CI}_j]$.
- Length of CI for noise variables: the empirical value of $(p-s)^{-1}\sum_{j\in S^c_{true}}\mathrm{length}(\mathrm{CI}_j)$.
- Length of CI for relevant variables: the empirical value of $s^{-1}\sum_{j\in S_{true}}\mathrm{length}(\mathrm{CI}_j)$.

Tables 1-2 summarize the performance measures for case 1-3 from 100 simulations with $n=200$, $p=300$ and $s=5$ or 10. It is shown that the mean absolute bias for relevant variables of our estimator $\hat{b}$ is significantly smaller than that of the lasso solution in all settings. The coverages of the confidence intervals are fairly consistent for all cases 1-3, concurring with the theoretical results on the robustness of the method in Corollary 2.6. We notice that the coverage of confidence intervals for noise variables is very close to the nominal level of 0.95, but that for the relevant variants is lower than 0.95. Similar phenomena have been observed for the original de-sparsified estimator in [20] and [4]. This phenomenon indicates that even though the bias of the original Lasso estimator has been corrected asymptotically by the de-sparsifying procedure, the non-zero coefficients can still be under-estimated with finite sample. Further, given that the implemented $\hat{\mu}_\pi(\tilde{X}_i)=\hat{\gamma}^T\tilde{X}_i$, it agrees with our expectation that the lengths of the confidence intervals are shortest for Case 1 when the true baseline function is linear and increases in Case 2 and 3 with nonlinear baseline functions.

TABLE 1
Coverage of confidence interval and mean absolute bias (MAB) of $\hat{b}$ with $p=300$ and $s=5$. Standard errors are in parenthesis.

	Intensity=1			Intensity=1.5		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
MAB($\hat{b}$)	0.17(0.06)	0.40(0.14)	0.30(0.10)	0.17(0.06)	0.41(0.15)	0.30(0.11)
MAB($\hat{\beta}$)	0.23(0.06)	0.50(0.14)	0.39(0.11)	0.23(0.06)	0.53(0.18)	0.40(0.13)
Coverage for noise variables	0.95(0.01)	0.95(0.01)	0.95(0.03)	0.95(0.01)	0.94(0.05)	0.95(0.01)
Coverage for relevant variables	0.88(0.13)	0.89(0.15)	0.85(0.18)	0.88(0.13)	0.86(0.20)	0.86(0.17)
Length of CI for noise variables	0.65(0.04)	1.50(0.17)	1.21(0.91)	0.65(0.04)	1.50(0.17)	1.12(0.12)
Length of CI for relevant variables	0.65(0.04)	1.49(0.16)	1.20(0.92)	0.65(0.04)	1.49(0.16)	1.11(0.11)

TABLE 2
Coverage of confidence interval and mean absolute bias (MAB) of $\hat{b}$ with $p = 300$ and $s = 10$.

	Intensity=1			Intensity=1.5		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
MAB($\hat{b}$)	0.19(0.04)	0.42(0.10)	0.33(0.08)	0.19(0.04)	0.46(0.27)	0.43(0.96)
MAB($\hat{\beta}$)	0.24(0.06)	0.51(0.10)	0.40(0.09)	0.24(0.06)	0.57(0.27)	0.52(0.96)
Coverage for noise variables	0.95(0.03)	0.94(0.04)	0.95(0.03)	0.94(0.05)	0.94(0.04)	0.94(0.05)
Coverage for relevant variables	0.86(0.14)	0.86(0.16)	0.84(0.14)	0.84(0.18)	0.84(0.16)	0.82(0.18)
Length of CI for noise variables	0.69(0.05)	1.58(0.18)	1.19(0.13)	0.69(0.05)	1.64(0.46)	1.23(0.47)
Length of CI for relevant variables	0.69(0.05)	1.58(0.18)	1.18(0.13)	0.69(0.05)	1.63(0.45)	1.22(0.47)

4. Application to STAR*D study

We consider the dataset from multi-site, randomized, multi-step STAR*D (Sequenced Treatment Alternatives to Relieve Depression) study for Major depressive disorder (MDD), which is a chronic and recurrent common disorder [5, 14]. In STAR*D study, patients received citalopram (CIT) which is a selective serotonin reuptake inhibitor antidepressant at Level 1. Patients who have had unsatisfactory outcome at level 1 are included in level 2 to randomly receive one of the two treatment switch options: sertraline (SER) and bupropion (BUP). Patients who received treatment at level 2 but have not showed sufficient improvement will be randomized at level 2A.

Our data contains 319 patients who received treatment switch options at level 2. Among them, 48% of the patients received BUP and 52% received SER. Except for treatment indicator, 308 covariate variables of the patients are included. The outcome of interest is the Quick Inventory of Depressive Symptomatology-Self-report QIDS-SR_{16}. Similar to existing studies on STAR*D [15], we transform the outcome to be the negative of the original outcome. We are interested in making optimal treatment decision between SER and BUR to maximize the mean outcome.

We apply the de-sparsified estimator $\hat{b}$ in (4) with $\hat{\mu}_\pi(\tilde{X}_i) = \hat{\gamma}^T \tilde{X}_i$ and derive the p-value based on the limiting distribution in Corollary 2.6. The top-ranked p-values that are less than 0.05 are presented in Table 3 with the corresponding covariates.

The meanings of the notations in Table 3 are as follows. Qccur_r_rate : QIDS-C score changing rates. URNONE: no symptoms in patients' urination category. NVTRM: tremors. IMPWR: indicating whether patients thought they have special powers. hWL: hRS Weight loss. DSMTD: recurrent thoughts of death, recurrent suicidal ideation, or suicide attempt. hMNIN: hRS Middle insomnia. EMSTU: did you worry a lot that you might do something to make people think that you were stupid or foolish?

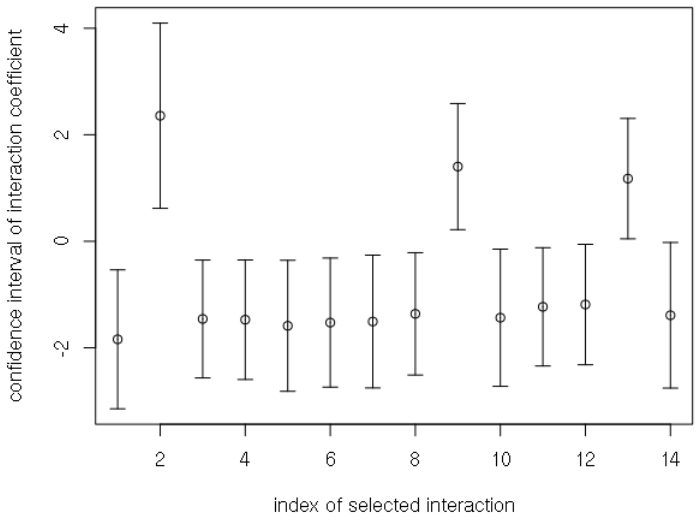
We have also derived the 95% confidence intervals by (10) for the interaction effects between all covariates and treatment options. The confidence intervals

TABLE 3
Top-ranked p-values and the corresponding covariates that are most relevant for making treatment decision.

Index	Covariate	p-value	$\hat{b}$	se of $\hat{b}$
1	qccur_r_rate	0.0056	-1.84	0.66
2	URNONE	0.0079	2.35	0.88
3	NVTRM	0.0097	-1.45	0.56
4	IMPWR	0.010	-1.47	0.57
5	hWL	0.011	-1.58	0.62
6	GLT2W	0.013	-1.52	0.61
7	DSMTD	0.017	-1.50	0.63
8	IMSPY	0.019	-1.36	0.58
9	hMNIN	0.020	1.40	0.60
10	EARNG	0.028	-1.43	0.65
11	URPN	0.029	-1.23	0.56
12	PETLK	0.039	-1.18	0.57
13	NVCRD	0.041	1.17	0.57
14	EMSTU	0.046	-1.39	0.69

for the top 14 interaction effects are presented in Figure 1.

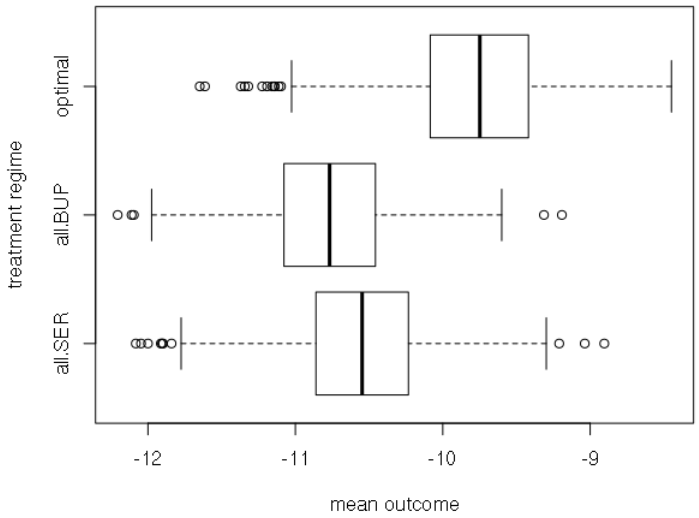
FIG 1. *The 95% confidence intervals of the top 14 interaction effects*



We also apply cross validation to evaluate the effects of the selected variables in making optimal treatment decision. Specifically, we randomly divide the data in half into the training dataset and the testing dataset. On the training dataset, we use the 14 selected variables to compute the least squares estimator for interaction coefficients and formulate the estimated optimal treatment regime. On

the corresponding testing dataset, we use the estimated value function derived by the inverse probability weighted estimator [23] to measure the performance of the estimated optimal treatment regime. For comparison, we also compute the estimated value functions for the regimes of assigning only SER and only BUP. We conducted 1000 data splittings and present the boxplot of the estimated value functions of all the treatment regimes in Figure 2. It shows that our estimated optimal treatment regime generally gives larger estimated value functions than those of assigning only SER and only BUP.

FIG 2. *Boxplot of the estimated value functions from cross validation testing samples. 'optimal' is for the estimated optimal treatment regime; 'all.BUP' is for the regime of assigning only BUP; 'all.SER' is for the regime of assigning only SER.*



5. Discussion

In this paper, we have considered randomized studies with a pre-specified propensity score. The proposed method can be extended to observation studies with high-dimensional covariates, where the propensity score model needs to be correctly specified from data. Related work on propensity score estimation can be found in [16]. However, the derivation of the limiting distribution of our desparsified estimator for the treatment-covariates interaction coefficients would be much more involved and requires further investigation.

Appendix A: Appendix

A.1. Construction of $\hat{\mathbf{\Theta}}$

We apply lasso for nodewise regression to obtain a matrix $\hat{\mathbf{\Theta}}$ such that $\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}}$ is close to $\mathbf{I}$ as in [9]. Let $\tilde{\mathbf{X}}_{-j}$ denote the matrix obtained by removing the jth column of $D(A,\pi)\tilde{\mathbf{X}}$. For each $j = 1,\ldots,p+1$, let

$$\hat{\gamma}_j = \underset{\gamma\in\mathbb{R}^p}{\operatorname{argmin}} \left(n^{-1} \left\| D(A,\pi)\tilde{\mathbf{X}}_j - \tilde{\mathbf{X}}_{-j}\gamma \right\|_2^2 + 2\lambda_j \|\gamma\|_1 \right)$$

with components $\hat{\gamma}_{j,k}, k = 1,\ldots,p+1$ and $k \neq j$. Further, define

$$\hat{\tau}_j^2 = n^{-1} \left\| D(A,\pi)\tilde{\mathbf{X}}_j - \tilde{\mathbf{X}}_{-j}\hat{\gamma}_j \right\|_2^2 + 2\lambda_j \|\hat{\gamma}_j\|_1$$

and

$$\hat{\mathbf{\Theta}} = \operatorname{diag}\left(\hat{\tau}_1^{-2},\cdots,\hat{\tau}_{p+1}^{-2}\right) \begin{pmatrix} 1 & -\hat{\gamma}_{1,2} & \cdots & -\hat{\gamma}_{1,p+1} \\ -\hat{\gamma}_{2,1} & 1 & \cdots & -\hat{\gamma}_{2,p+1} \\ \vdots & \vdots & \vdots & \vdots \\ -\hat{\gamma}_{p+1,1} & -\hat{\gamma}_{p+1,2} & \cdots & 1 \end{pmatrix}.$$

A.2. A preliminary lemma

Define

$$\epsilon_{\hat{\mu}} = \epsilon - [\hat{\mu}_\pi(\tilde{\mathbf{X}}) - \mu_\pi(\tilde{\mathbf{X}})]. \tag{11}$$

Lemma A.1. *Assume conditions 1 - 3, we have*

$$\|2\epsilon_{\hat{\mu}}^T D(A,\pi)\tilde{\mathbf{X}}/n\|_\infty = O_p(\sqrt{\log p}/\sqrt{n}).$$

Proof of Lemma A.1:

$$\begin{aligned} \epsilon_{\hat{\mu}}^T D(A,\pi)\tilde{\mathbf{X}} &= (\epsilon - [\mu_\pi^*(\tilde{\mathbf{X}}) - \mu_\pi(\tilde{\mathbf{X}})] - [\hat{\mu}_\pi(\tilde{\mathbf{X}}) - \mu_\pi^*(\tilde{\mathbf{X}})])^T D(A,\pi)\tilde{\mathbf{X}} \\ &= \epsilon^T D(A,\pi)\tilde{\mathbf{X}} - [\mu_\pi^*(\tilde{\mathbf{X}}) - \mu_\pi(\tilde{\mathbf{X}})]^T D(A,\pi)\tilde{\mathbf{X}} - [\hat{\mu}_\pi(\tilde{\mathbf{X}}) - \mu_\pi^*(\tilde{\mathbf{X}})]^T D(A,\pi)\tilde{\mathbf{X}} \end{aligned} \tag{12}$$

For the first term of (12), it is clear that by condition 1,

$$E(\epsilon^T D(A,\pi)\tilde{\mathbf{X}}) = EE(\epsilon^T D(A,\pi)\tilde{\mathbf{X}}|A,\tilde{\mathbf{X}}) = 0.$$

Further, by the moment inequality in Chapter 6.2.2 of [1],

$$E(\|\epsilon^T D(A,\pi)\tilde{\mathbf{X}}\|_\infty^2) \leq 8\log(2p)\sum_{i=1}^n (A_i-\pi)^2 (\max_{1\leq j\leq p}|X_i^j|)^2 E(\epsilon_i^2) \leq Cn\log(p),$$

where the second inequality is by conditions 1 and 3. Then, by Markov inequality,

$$\|2\epsilon^T D(A,\pi)\tilde{\mathbf{X}}/n\|_\infty = O_p(\sqrt{\log(p)}/\sqrt{n}). \tag{13}$$

Next, consider the second term of (12). It is easy to see that since $E(D(A,\pi)|\tilde{\mathbf{X}}) = 0$,

$$E([\mu_\pi^*(\tilde{\mathbf{X}}) - \mu_\pi(\tilde{\mathbf{X}})]^T D(A,\pi)\tilde{\mathbf{X}}) = EE([\mu_\pi^*(\tilde{\mathbf{X}}) - \mu_\pi(\tilde{\mathbf{X}})]^T D(A,\pi)\tilde{\mathbf{X}}|\tilde{\mathbf{X}}) = 0.$$

Then, similar arguments as those leading to (13) combined with conditions 2 (b) and 3 gives

$$\|2[\mu_\pi^*(\tilde{\mathbf{X}}) - \mu_\pi(\tilde{\mathbf{X}})]^T D(A,\pi)\tilde{\mathbf{X}}/n\|_\infty = O_p(\sqrt{\log(p)}/\sqrt{n}). \tag{14}$$

Finally, by condition 2 (a), the third term of (12)

$$\|2[\hat{\mu}_\pi(\tilde{\mathbf{X}}) - \mu_\pi^*(\tilde{\mathbf{X}})]^T D(A,\pi)\tilde{\mathbf{X}}/n\|_\infty = o_p(\sqrt{\log(p)}/\sqrt{n}). \tag{15}$$

Combining (13) - (15) gives the claim in Lemma A.1.

A.3. Proof of Lemma 2.1

By the construction of $\hat{\beta}$, we have

$$\|Y - \hat{\mu}_\pi(\tilde{\mathbf{X}}) - D(A,\pi)\tilde{\mathbf{X}}\hat{\beta}\|_2^2/n + 2\lambda_{n,p}\|\hat{\beta}\|_1 \le \|Y - \hat{\mu}_\pi(\tilde{\mathbf{X}}) - D(A,\pi)\tilde{\mathbf{X}}\beta_0\|_2^2/n + 2\lambda_{n,p}\|\beta_0\|_1.$$

Recall the definition of $\epsilon_{\hat{\mu}}$ in (11), then

$$\|D(A,\pi)\tilde{\mathbf{X}}(\hat{\beta} - \beta_0)\|_2^2/n + 2\lambda_{n,p}\|\hat{\beta}\|_1 \le 2\epsilon_{\hat{\mu}}^T D(A,\pi)\tilde{\mathbf{X}}(\hat{\beta} - \beta_0)/n + 2\lambda_{n,p}\|\beta_0\|_1.$$

The first term of the righ-hand side

$$2\epsilon_{\hat{\mu}}^T D(A,\pi)\tilde{\mathbf{X}}(\hat{\beta} - \beta_0)/n \le \|2\epsilon_{\hat{\mu}}^T D(A,\pi)\tilde{\mathbf{X}}/n\|_\infty \|\hat{\beta} - \beta_0\|_1,$$

where the order of $\|2\epsilon_{\hat{\mu}}^T D(A,\pi)\tilde{\mathbf{X}}/n\|_\infty$ is derived in Lemma A.1.

The rest of the proof is similar to the proof of the second claim of Lemma 2 in [2], where it is shown that given conditions 3 - 5, the compatibility condition holds with probability tending to 1. Then using Lemma A.1 and similar arguments in Section 6.2.2 of [1], the inequality in (5) holds.

A.4. Proof of Lemma 2.2

Consider (6) first. Recall

$$\sqrt{n}\mathbf{H}\Delta_1 = \frac{1}{\sqrt{n}}\mathbf{H}\hat{\mathbf{\Theta}}\{D(A,\pi)\tilde{\boldsymbol{X}}\}^T\{\hat{\mu}_\pi(\tilde{\mathbf{X}}) - \mu_\pi^*(\tilde{\mathbf{X}})\}.$$

Rewrite $\sqrt{n}\mathbf{H}\boldsymbol{\Delta}_1$ as

$$\begin{aligned}\sqrt{n}\mathbf{H}\Delta_1 &= \mathbf{H}\hat{\mathbf{\Theta}}\frac{1}{\sqrt{n}}\sum_{i=1}^n (A_i - \pi)\tilde{X}_i\{\hat{\mu}_\pi(\tilde{X}_i) - \mu_\pi^*(\tilde{X}_i)\} \\ &\le \sup_{\tilde{X}_i}|\hat{\mu}_\pi(\tilde{X}_i) - \mu_\pi^*(\tilde{X}_i)|\mathbf{H}\hat{\mathbf{\Theta}}\frac{1}{\sqrt{n}}\sum_{i=1}^n (A_i - \pi)\tilde{X}_i.\end{aligned}$$

Since $E[(A_i-\pi)\tilde{X}_i]=0$ and $\|E[(A_i-\pi)^2\tilde{X}_i\tilde{X}_i^T]\|_\infty<\infty$ by condition 3, by Multivariate Central Limit Theorem, the $(p+1)$-vector

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(A_i-\pi)\tilde{X}_i=O_p(1).$$

On the other hand, condition 7 combined with similar arguments as in the proof of the first claim of Lemma 2 in [2] gives $\|\hat{\Theta}^j-\Theta^j\|_1=o_p(1/\sqrt{\log p})$. Summarizing the above with condition 6 and Slutsky's Theorem gives

$$\|\mathbf{H}\hat{\mathbf{\Theta}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(A_i-\pi)\tilde{X}_i\|_\infty=O_p(1). \tag{16}$$

(6) follows from (16) and condition 2 (a).

Next, consider (7).

$$\|\sqrt{n}\mathbf{H}\Delta_2\|_\infty=\|\sqrt{n}\mathbf{H}(\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}}-\boldsymbol{I})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\|_\infty\le\sqrt{n}\|\mathbf{H}(\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}}-\boldsymbol{I})\|_\infty\|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\|_1.$$

By condition 6, $\|\mathbf{H}(\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}}-\boldsymbol{I})\|_\infty\le C\|\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}}-\boldsymbol{I}\|_\infty$. Similar arguments as those leading to inequality (10) in [20], combined with conditions 3 and 7, result in $\|\hat{\mathbf{\Theta}}\hat{\mathbf{\Sigma}}-\boldsymbol{I}\|_\infty=O_p(\sqrt{\log p}/\sqrt{n})$. Combing the above with Lemma 2.2 gives (7).

A.5. Proof of Lemma 2.3

We first show $E(\sqrt{n}\mathbf{H}\eta)=0$. By definition of η, we have

$$\begin{aligned}E(\sqrt{n}\mathbf{H}\eta) &= E\left[\frac{1}{\sqrt{n}}\mathbf{H}\hat{\mathbf{\Theta}}\{D(A,\pi)\tilde{\boldsymbol{X}}\}^T\left(Y-\mu_\pi^*\tilde{\mathbf{X}}-D(A,\pi)\tilde{\boldsymbol{X}}\boldsymbol{\beta_0}\right)\right]\\ &= \frac{1}{\sqrt{n}}\mathbf{H}E\left[\hat{\mathbf{\Theta}}\{D(A,\pi)\tilde{\boldsymbol{X}}\}^T\epsilon\right]+\frac{1}{\sqrt{n}}\mathbf{H}E\left[\hat{\mathbf{\Theta}}\{D(A,\pi)\tilde{\boldsymbol{X}}\}^T\left(\mu_\pi(\tilde{\mathbf{X}})-\mu_\pi^*(\tilde{\mathbf{X}})\right)\right]\end{aligned}$$

The first term of the above is equal to 0 because $E(\epsilon|\tilde{\mathbf{X}},A)=0$ by condition 1. The second term also equals to 0 since $E(D(A,\pi)|\tilde{\mathbf{X}})=0$.

Next, rewrite $\sqrt{n}\mathbf{H}\eta$ as

$$\sqrt{n}\mathbf{H}\eta=\mathbf{H}\hat{\Theta}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}W_i,$$

where

$$W_i=(A_i-\pi)(y_i-\mu_\pi^*(\tilde{X}_i)-(A_i-\pi)\beta_0^T\tilde{X}_i)\tilde{X}_i$$

The second moment of W_i is

$$E\left[(A_i-\pi)^2(\epsilon_i+\mu_\pi(\tilde{X}_i)-\mu_\pi^*(\tilde{X}_i))^2\tilde{X}_i\tilde{X}_i^T\right]=E\left[(A_i-\pi)^2\epsilon_i^2\tilde{X}_i\tilde{X}_i^T\right]+E\left[(A_i-\pi)^2(\mu_\pi(\tilde{X}_i)-\mu_\pi^*(\tilde{X}_i))^2\tilde{X}_i\tilde{X}_i^T\right.$$

The first term

$$E\left[(A_i-\pi)^2\epsilon_i^2\tilde{X}_i\tilde{X}_i^T\right] = E\left[(A_i-\pi)^2E(\epsilon_i^2|\tilde{X},A)\tilde{X}_i\tilde{X}_i^T\right]=E\left[(A_i-\pi)^2\sigma_\epsilon^2(A_i,\tilde{X}_i)\tilde{X}_i\tilde{X}_i^T\right]$$

The second term

$$\begin{aligned} E\left[(A_i-\pi)^2(\mu_\pi(\tilde{X}_i)-\mu_\pi^*(\tilde{X}_i))^2\tilde{X}_i\tilde{X}_i^T\right] &= E\left[E((A_i-\pi)^2|\tilde{X}_i)(\mu_\pi(\tilde{X}_i)-\mu_\pi^*(\tilde{X}_i))^2\tilde{X}_i\tilde{X}_i^T\right] \\ &= \pi(1-\pi)E\left[(\mu_\pi(\tilde{X}_i)-\mu_\pi^*(\tilde{X}_i))^2\tilde{X}_i\tilde{X}_i^T\right]. \end{aligned}$$

Note that $\sigma_\epsilon^2(A_i,\tilde{X}_i)<\infty$ by condition 1. This combined with conditions 2 (b) and 3 gives $\|\mathbf{V}\|_\infty<\infty$. Consequently, $\|\mathbf{H\Theta V\Theta}^T\mathbf{H}^T\|_\infty<\infty$ is implied by conditions 6 - 7.

On the other hand, condition 7 combined with similar arguments as in the proof of the first claim of Lemma 2 in [2] gives $\|\hat{\Theta}^j-\Theta^j\|_1=o_p(1/\sqrt{\log p})$. By Multivariate Central Limit Theorem and Slutsky's Theorem,

$$\mathbf{H}\hat{\mathbf{\Theta}}\frac{1}{\sqrt{n}}\sum_{i=1}^n W_i \to_d N(0,\mathbf{H\Theta V\Theta}^T\mathbf{H}^T).$$

Result in (8) follows.

A.6. Proof of Corollary 2.6

After obtaining $\hat{\gamma}$ and $\hat{\beta}$ from (9), implement $\hat{\mu}_\pi(\tilde{X}_i)=\hat{\gamma}^T\tilde{X}_i$. Define

$$\beta^*=\text{argmin}_\beta\left\{\|Y-\tilde{\mathbf{X}}\hat{\gamma}-D(A,\pi)\tilde{\mathbf{X}}\beta\|_2^2/n+2\lambda_{n,p}\|\beta\|_1\right\}.$$

First, we show that $\hat{\beta}=\beta^*$. By the construction of β^*,

$$\|Y-\tilde{\mathbf{X}}\hat{\gamma}-D(A,\pi)\tilde{\mathbf{X}}\beta^*\|_2^2/n+2\lambda_{n,p}\|\beta^*\|_1\le\|Y-\tilde{\mathbf{X}}\hat{\gamma}-D(A,\pi)\tilde{\mathbf{X}}\hat{\beta}\|_2^2/n+2\lambda_{n,p}\|\hat{\beta}\|_1. \tag{17}$$

On the other hand, the construction of $\hat{\beta}$ implies

$$\begin{aligned} &\|Y-\tilde{\mathbf{X}}\hat{\gamma}-D(A,\pi)\tilde{\mathbf{X}}\hat{\beta}\|_2^2/n+2\lambda_{n,p}(\|\hat{\gamma}\|_1+\|\hat{\beta}\|_1) \\ \le\ &\|Y-\tilde{\mathbf{X}}\hat{\gamma}-D(A,\pi)\tilde{\mathbf{X}}\beta^*\|_2^2/n+2\lambda_{n,p}(\|\hat{\gamma}\|_1+\|\beta^*\|_1) \\ \le\ &\|Y-\tilde{\mathbf{X}}\hat{\gamma}-D(A,\pi)\tilde{\mathbf{X}}\hat{\beta}\|_2^2/n+2\lambda_{n,p}(\|\hat{\gamma}\|_1+\|\hat{\beta}\|_1), \end{aligned}$$

where the second inequality is by (17). Therefore, by the uniqueness of the solution of convex optimization, $\hat{\beta}=\beta^*$.

Secondly, we show that

$$\|\hat{\gamma}-\gamma^*\|_1+\|\hat{\beta}-\beta_0\|_1=o_p(1/\sqrt{\log p}). \tag{18}$$

Given the fact that $\mu_\pi(\tilde{X}_i)$ and $(A_i-\pi)\beta_0^T\tilde{X}_i$ are orthogonal and $E(\epsilon_i|\tilde{\mathbf{X}})=0$ from condition 1, definition of γ^* implies

$$(\gamma^*,\beta_0)=\text{argmin}_{\gamma,\beta}\left\{E(Y_i-\gamma^TX_i-(A_i-\pi)\beta^TX_i)^2\right\}.$$

Define $\xi_i = Y_i - \gamma^{*T} X_i - (A_i - \pi)\beta_0^T X_i = \epsilon_i + \mu_\pi(\tilde{X}_i) - \gamma^{*T} X_i$. Then conditions 1 and 9 (b) imply

$$E(\xi_i)^2 \leq C. \tag{19}$$

By the second claim of Lemma 2 of [2], (19) combined with conditions 3 - 5 and 9 (a) gives (18).

Next, it is easy to see that condition 2 (a) is implied by (18) and condition 3, and condition 2 (b) holds trivially given condition 9 (b). Therefore, condition 2 is satisfied and the rest of the proof is the same as the proof of Theorem 2.4.

References

[1] Bühlmann, P. and S. van de Geer (2011). *Statistics for high-dimensional data - methods, theory and applications*. Springer.

[2] Bühlmann, P. and S. van de Geer (2015). High-dimensional inference in misspecified linear models. *Electronic Journal of Statistics 9*, 1449–1473.

[3] Chakraborty, B., S. A. Murphy, and V. J. Strecher (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research 19*(3), 317–343.

[4] Dezeure, R., P. Bühlmann, L. Meier, and N. Meinshausen (2015). High-dimensional inference: Confidence intervals, p-values and r-software hdi. *Statistical Science 30*(4), 533–558.

[5] Fava, M., A. J. Rush, M. H. Trivedi, A. A. Nierenberg, M. E. Thase, H. A. Sackeim, F. M. Quitkin, S. Wisniewski, P. W. Lavori, J. F. Rosenbaum, et al. (2003). Background and rationale for the sequenced treatment alternatives to relieve depression (star* d) study. *Psychiatric Clinics of North America 26*(2), 457–494.

[6] Goldberg, Y., R. Song, and M. R. Kosorok (2013). Adaptive q-learning. *Institute of Mathematical Statistics Collections 9*, 150–162.

[7] Laber, E. B., K. A. Linn, and L. A. Stefanski (2014). Interactive model building for q-learning. *Biometrika 101*, 831–847.

[8] Lu, W., H. H. Zhang, and D. Zeng (2013). Variable selection for optimal treatment decision. *Statistical Methods in Medical Research 22*(5), 493–504.

[9] Meinshausen, N. and P. Bühlmann (2006). High-dimensional graphs and variable selection with the Lasso. *The Annals of Statistics 34*(3), 1436–1462.

[10] Murphy, S. (2003). Optimal dynamic treatment regimes. *Journal of The Royal Statistical Society: Series B 65*, 331–366.

[11] Murphy, S. (2005). A generalization error for q-learning. *Journal of Machine Learning Research 6*, 1073–1097.

[12] Qian, M. and S. A. Murphy (2011). Performance guarantees for individualized treatment rules. *The Annals of Statistics 39*, 1180–1210.

[13] Robins, J., M. A. Hernan, and B. Brumback (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology 11*, 550–560.

[14] Rush, A. J., M. Fava, S. R. Wisniewski, P. W. Lavori, M. H. Trivedi, H. A. Sackeim, M. E. Thase, A. A. Nierenberg, F. M. Quitkin, T. M. Kashner,

et al. (2004). Sequenced treatment alternatives to relieve depression (star* d): rationale and design. *Controlled Clinical Trials 25*(1), 119–142.

[15] Shi, C., A. Fan, R. Song, and W. Lu (2018). High-dimensional a-learning for optimal dynamic treatment regimes. *The Annals of Statistics 46(3)*, 925–957.

[16] Shi, C., R. Song, and W. Lu (2016). Robust learning for optimal treatment decision with np-dimensionality. *Electronic Journal of Statistics 10*, 2894–2921.

[17] Song, R., W. Wang, D. Zeng, and M. R. Kosorok (2015). Penalized q-learning for dynamic treatment regimens. *Statistica Sinica 25*, 901–920.

[18] Su, X., C.-L. Tsai, H. Wang, D. M. Nickerson, and B. Li (2009). Subgroup analysis via recursive partitioning. *Journal of Machine Learning Research 10*, 141–158.

[19] Tian, L., A. A. Alizadeh, A. J. Gentles, and R. Tibshirani (2014). A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association 109*, 1517–1532.

[20] van de Geer, S., P. Bühlmann, Y. Ritov, and R. Dezeure (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *The Annals of Statistics 42*(3), 1166–1202.

[21] Watkins, C. J. C. H. and P. Dayan (1992). Q-learning. *Machine Learning 8*, 279–292.

[22] Zhang, B., A. A. Tsiatis, M. Davidian, M. Zhang, and E. B. Laber (2012). Estimating optimal treatment regimes from a classification perspective. *Stat 1*, 103–104.

[23] Zhang, B., A. A. Tsiatis, E. B. Laber, and M. Davidian (2012). A robust method for estimating optimal treatment regimes. *Biometrics 68*, 1010–1018.

[24] Zhang, C. and S. S. Zhang (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *Journal of the Royal Statistical Society. Series B 76*(1), 217–242.

[25] Zhao, Y., D. Zeng, A. J. Rush, and M. R. Kosorok (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association 107*, 1106–1118.

[26] Zhao, Y. Q., D. Zeng, E. B. Laber, R. Song, M. Yuan, and M. R. Kosorok (2015). Doubly robust learning for estimating individualized treatment with censored data. *Biometrika 102*, 151–168.