Journal of Machine Learning Research 22 (2021) 1-51 Submitted 4/21; Revised 9/21; Published 10/21

An Online Sequential Test for Qualitative Treatment Effects

Chengchun Shi C.SHI7@LSE.AC.UK
Department of Statistics, London School of Economics and Political Science

Shikai Luo SLUO198912@163.COM
Tecent PCG

Hongtu Zhu HTZHU@EMAIL.UNC.EDU
Department of Biostatistics, University of North-Carolina

Rui Song RSONG@NCSU.EDU
Department of Statistics, North-Carolina State University

Editor: Xiaotong Shen

Abstract

Tech companies (e.g., Google or Facebook) often use randomized online experiments and/or A/B testing primarily based on the average treatment effects to compare their new product with an old one. However, it is also critically important to detect qualitative treatment effects such that the new one may significantly outperform the existing one only under some specific circumstances. The aim of this paper is to develop a powerful testing procedure to efficiently detect such qualitative treatment effects. We propose a scalable online updating algorithm to implement our test procedure. It has three novelties including adaptive randomization, sequential monitoring, and online updating with guaranteed type-I error control. We also thoroughly examine the theoretical properties of our testing procedure including the limiting distribution of test statistics and the justification of an efficient bootstrap method. Extensive empirical studies are conducted to examine the finite sample performance of our test procedure.

Keywords: A/B testing; Qualitative treatment effects; Sequential monitoring; Adaptive randomization; Online updating.

1. Introduction

Tech companies use randomized online experiments, or A/B testing to compare their new product with a well-established one. Most works in the literature focus on the average treatment effects (ATE) between the new and existing products (see Kharitonov et al., 2015; Johari et al., 2015, 2017; Yang et al., 2017; Ju et al., 2019, and the references therein). In addition to ATE, sometimes we are interested in locating the subgroup (if exists) that the new product performs significantly better than the existing one, as early as possible. Consider a ride-hailing company (e.g., Uber). Suppose some passengers are in the recession state (at a high risk of stopping using the company's app) and the company comes up with certain strategy to intervene the recession process. We would like to test if there are some subgroups that are sensitive to the strategy and pin-point these subgroups if exists. It

motivates us to consider the null hypothesis that the treatment effect is nonpositive for all passenger.

Such a null hypothesis is closely related to the notion of qualitative treatment effects in medical studies (QTE, Gail and Simon, 1985; Gunter et al., 2007, 2011; Roth and Simon, 2018; Shi et al., 2020b), and conditional moment inequalities in economics (see for example, Andrews and Shi, 2013, 2014; Chernozhukov et al., 2013; Armstrong and Chan, 2016; Chang et al., 2015; Hsu, 2017). However, these tests are computed offline and might not be suitable to implement in online settings. Moreover, it is assumed in those papers that observations are independent. In online experiment, one may wish to adaptively allocate the treatment based on the observed data stream in order to maximize the cumulative reward or to detect the alternative more efficiently. The independence assumption is thus violated. In addition, an online experiment is desired to be terminated as early as possible in order to save time and budget. Sequential testing for qualitative treatment effects has been less explored.

In the literature, there is a line of research on estimation and inference of the heterogeneous treatment effects (HTE, Athey and Imbens, 2016; Taddy et al., 2016; Wager and Athey, 2018; Yu et al., 2020). In particular, Yu et al. (2020) proposed an online test for HTE. We remark that HTE and QTE are related yet fundamentally different hypotheses. There are cases where HTE exists whereas QTE does not. See Figure 1 for an illustration. Consequently, applying their test will fail in our setting.

The contributions of this paper are summarized as follows. First, we propose a new testing procedure for treatment comparison based on the notion of QTE. When the null hypothesis is not rejected, the new product is no better than the control for any realization of covariates, and thus it is not useful at all. Otherwise, the company could implement different products according to the auxiliary covariates observed, to maximize the average reward obtained. We remark that there are plenty cases where the treatment effects are always nonpositive (see Section 5 of Chang et al., 2015; Shi et al., 2020b). A by-product of our test is that it yields a decision rule to implement personalization when the null is rejected (see Section 3.1 for details). Although we primarily focus on QTE in this paper, our procedure can be easily extended to testing ATE as well (see Section 4.1.2 for details).

Second, we propose a scalable online updating algorithm to implement our test. To allow for sequential monitoring, our procedure leverages idea from the α spending function approach (Lan and DeMets, 1983) originally designed for sequential analysis in a clinical trial (see Jennison and Turnbull, 1999, for an overview). Classical sequential tests focus on ATE. The test statistic at each interim stage is asymptotically normal and the stopping boundary can be recursively updated via numerical integration. However, the limiting distribution of the proposed test statistic does not have a tractable analytical form, making the numerical integration method difficult to apply. To resolve this issue, we propose a scalable bootstrap-assisted procedure to determine the stopping boundary.

Third, we adopt a theoretical framework that allows the maximum number of interim analyses K to diverge as the number of observations increases, since tech companies might analyze the results every few minutes (or hours) to determine whether to stop the experiment or continue collecting more data. It is ultimately different from classical sequential analysis where K is fixed. Moreover, the derivation of the asymptotic property of the proposed test is further complicated due to the adaptive randomization procedure, which makes observations dependent of each other. Despite these technical challenges, we establish a nonasymptotic

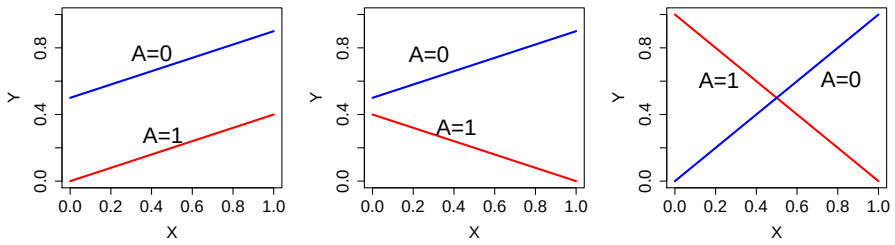

Figure 1: Plots demonstrating QTE. X denotes the observed covariates, A denotes the received treatment and Y denotes the associated reward. In the ride-hailing example, X is a feature vector describing the characteristics of a passenger, A is a binary strategy indicator and Y is the passenger's number of rides in the following two weeks. In the left panel, the treatment effect does not depend on X. Neither HTE nor QTE exists in this case. In the middle panel, HTE exists. However, the treatment effect is always negative. As such, QTE does not exist. In the right penal, both QTE and HTE exist.

upper bound on the type-I error rate by explicitly characterizing the conditions needed on randomization procedure, K and the number of samples observed at the initial decision point to ensure the validity of our test.

2. Background and problem formulation

We propose a potential outcome framework to formulate our problem. Suppose that we have two products including the control and the treatment. The observed data at time point t consists of a sequence of triples $\{(X_i, A_i, Y_i)\}_{i=1}^{N(t)}$, where $N(\cdot)$ is a counting process that is independent of the data stream $\{(X_i, A_i, Y_i)\}_{i=1}^{+\infty}$, A_i is a binary random variable indicating the product executed for the i-th experiment, $X_i \in \mathbb{R}^p$ denotes the associated covariates, and Y_i stands for the associated reward (the larger the better by convention). We allow A_i to depend on X_i and past observations $\{(X_j, A_j, Y_j)\}_{j<i}$ so that the randomization procedure can be adaptively changed. In addition, define $Y_i^*(0)$ and $Y_i^*(1)$ to be the potential outcome that would have been observed if the corresponding product is executed for the i-th experiment. Suppose that $\{(X_i, Y_i^*(0), Y_i^*(1))\}_{i=1}^{+\infty}$ are independently and identically distributed copies of $(X, Y^*(0), Y^*(1))$. Let $\mathbb{X}$ be the support of X and $Q_0(x, a) = \mathrm{E}\{Y^*(a)|X = x\}$ for $a = 0, 1$, we focus on testing the following hypotheses:

$$H_0 : Q_0(x, 1) \leq Q_0(x, 0), \forall x \in \mathbb{X} \quad \text{versus} \quad H_1 : Q_0(x, 1) > Q_0(x, 0), \exists x \in \mathbb{X}.$$

Notice that when there are no covariates, i.e., $\mathbb{X} = \emptyset$, the hypotheses are reduced to $H_0 : \tau_0 \leq 0$ versus $H_1 : \tau_0 > 0$, where τ_0 corresponds to ATE, i.e, $\tau_0 = \mathrm{E}\{Y^*(1) - Y^*(0)\}$. In general, we require $\mathbb{X}$ to be a compact set. We consider a large linear approximation space $\mathcal{Q}$ for the conditional mean function Q_0. Specifically, let $\mathcal{Q} = \{Q(x, a; \beta_0, \beta_1) = \varphi^\top(x)\beta_a : \beta_0, \beta_1 \in \mathbb{R}^q\}$ be the approximation space, where $\varphi(x)$ is a q-dimensional vector composed of basis functions on $\mathbb{X}$. The dimension q is allowed to diverge with the number of observations in order to alleviate the effects of model misspecification. The use of linear approximation space simplifies the computation of our testing procedure. When $Q_0(x, 0)$ and $Q_0(x, 1)$ are

well approximated by $\varphi^\top(x)\beta_0^*$ and $\varphi^\top(x)\beta_1^*$ for some β_0^* and β_1^*, it suffices to test

$$H_0 : \varphi^\top(x)(\beta_1^* - \beta_0^*) \le 0, \forall x \in \mathbb{X} \quad \text{versus} \quad H_1 : \varphi^\top(x)(\beta_1^* - \beta_0^*) > 0, \exists x \in \mathbb{X}. \tag{1}$$

We require the approximation error $\text{err} = \inf_{\beta_0,\beta_1\in\mathbb{R}^p} \sup_{x\in\mathbb{X},a\in\{0,1\}} |Q_0(x,a) - Q(x,a;\beta_0,\beta_1)|$ to decay to zero at a rate of $o\{N^{-1/2}(T)\}$ to ensure the validity of the proposed test. See Appendix A for details.

Let $\mathcal{F}_j$ denote the sub-dataset $\{(X_i, A_i, Y_i)\}_{1\le i\le j}$ for $j \ge 1$ and $\mathcal{F}_0 = \emptyset$. Throughout this paper, we assume that the following two assumptions hold.

(A1) $Y_i = A_i Y_i^*(1) + (1 - A_i)Y_i^*(0)$ for $\forall i \ge 1$.

(A2) A_i is independent of $Y_i^*(0), Y_i^*(1)$, $\{(X_k, Y_k^*(0), Y_k^*(1))\}_{k>i}$ given X_i and $\mathcal{F}_{i-1}$, for any i.

Assumption (A1) is referred to be the stable unit treatment value assumption and Assumption (A2) is the sequential randomization assumption (Zhang et al., 2013) and is automatically satisfied in a randomized study where the treatments are independently generated of the observed data. (A2) essentially assumes there is no unmeasured confounders. These assumptions guarantee that both regression coefficients (defined through potential outcomes) are estimable from the observed dataset as shown in the following lemma.

Lemma 1 *Let $\mathbb{I}(\cdot)$ denotes the indicator function. Under (A1)-(A2), we have*

$$E[\mathbb{I}(A_i = a)\{Y_i - Q_0(X_i, a)\}] = 0, \quad \forall a \in \{0,1\}, i \ge 1.$$

3. Online sequential testing for QTE

3.1 Test statistics and their limiting distribution

We first present our test statistic for testing H_0. In view of Lemma 1, we estimate β_a^* by using the ordinary least squares estimator

$$\widehat{\beta}_a(t) = \widehat{\Sigma}_a^{-1}(t) \left\{ \frac{1}{N(t)} \sum_{i=1}^{N(t)} \mathbb{I}(A_i = a)\varphi(X_i)Y_i \right\}$$

at each time point t for $a \in \{0,1\}$, where $\widehat{\Sigma}_a(t) = N^{-1}(t)\sum_{i=1}^{N(t)} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)$. A generalized inverse might be used even if $\widehat{\Sigma}_a(t)$ is not invertible. Consider the following normalized test statistic

$$S(t) = \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}]},$$

for some standard error estimator $\widehat{s.e.}[\cdot]$ whose explicit form will be presented below. The benefits of working with normalized statistics are two folds. First, there is an efficiency gain compared to the unnormalized statistics without standardization, i.e., $\sup_{x\in\mathbb{X}} \varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}$. In cases where $\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}$ is not consistently estimated for some value of x, the supremum might be attained by these values. However, due to the large estimation error, they might differ significant from the oracle maximizer, $\arg\max_x\{Q_0(x,1) - Q_0(x,0)\}$,

thus lowering the power. We conduct some simulation studies (results are not included in the paper) and find that the normalized test statistic has much better power properties compared to the unnormalized one. Second, it requires weaker assumptions than the unnormalized test statistic. We will discuss this in detail in Appendix A.

Meanwhile, we remark that the studentized supremum type statistics have been used in the economics literature. For instance, Chen and Christensen (2015) proposed to use studentization for controlling the bias term in nonparametric series regression. Belloni et al. (2015) proposed to construct uniform confidence bands in nonparametric regression based on studentized supremum type statistics. Chen and Christensen (2018) developed a uniform inference on nonlinear functionals of nonparametric instrumental variables regression using studentized supremum type statistics. The benefits of using these statistics have been discussed in these papers as well.

Under H_0, we expect $S(t)$ to be small. A large $S(t)$ can be interpreted as the evidence against H_0. As such, we reject H_0 for large $S(t)$. We remark that when H_0 is rejected, we can apply the decision rule $d(x) = \arg\max_{a\in\{0,1\}} \varphi^\top(x)\widehat{\beta}_a(t)$ for personalized recommendation.

To determine the rejection region, we next discuss the limiting distribution of $S(t)$. Under H_0,

$$\begin{aligned} S(t) &\leq \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)\{\widehat{\beta}_1(t) - \beta_1^* - \widehat{\beta}_0(t) + \beta_0^*\} + \sup_{x\in\mathbb{X}} \varphi^\top(x)(\beta_1^* - \beta_0^*)}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}]} \\ &\leq \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)\{\widehat{\beta}_1(t) - \beta_1^* - \widehat{\beta}_0(t) + \beta_0^*\}}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}]}. \end{aligned} \tag{2}$$

Both equalities hold when $\beta_0^* = \beta_1^*$. Suppose there exists some function $\pi^*(\cdot,\cdot)$ defined on $\{0,1\}\times\mathbb{X}$ that satisfies $\mathrm{E}^X|\sum_{i=1}^n n^{-1}\pi_{i-1}(a,X) - \pi^*(a,X)| \xrightarrow{P} 0, \forall a \in \{0,1\}$ as $n\to\infty$, where $\pi_n(\cdot,\cdot) = \Pr(A_n = a|X_n = x, \mathcal{F}_{n-1})$, and the expectation E^X is taken with respect to X. This condition implies that the treatment assignment mechanism cannot be arbitrary (see the discussion below Theorem 1 for details). Then we will show

$$B(t) \equiv \sqrt{N(t)}\{\widehat{\beta}_1(t) - \beta_1^* - \widehat{\beta}_0(t) + \beta_0^*\} \xrightarrow{d} N(0, \sum_{a\in\{0,1\}} \Sigma_a^{-1}\Phi_a\Sigma_a^{-1}), \text{ as } N(t)\to\infty, \tag{3}$$

where $\Sigma_a = \mathrm{E}\pi^*(a,X)\varphi(X)\varphi^\top(X)$, $\Phi_a = \mathrm{E}\pi^*(a,X)\sigma^2(a,X)\varphi(X)\varphi^\top(X)$, and $\sigma^2(a,x) = \mathrm{E}[\{Y^*(a) - Q_0(X,a)\}^2|X=x]$, for any $x\in\mathbb{X}$. Consequently, we set the denominator of $S(t)$ to

$$\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}] = \left\{\frac{1}{N(t)}\sum_{a\in\{0,1\}} \varphi^\top(x)\widehat{\Sigma}_a^{-1}(t)\widehat{\Phi}_a(t)\widehat{\Sigma}_a^{-1}(t)\right\}^{1/2},$$

where $\widehat{\Phi}_a(t)$ denotes the sandwich estimator for Φ_a computed using time points up to time t. Please refer to Algorithm 1 for a detailed definition.

In addition, according to (3), the right-hand-side (RHS) of (2) is to converge in distribution to the maximum of some Gaussian random variables. This observation forms the basis of our test.

We next discuss the sequential implementation of our test. Assume that the interim analyses are conducted at time points $t_1, t_2, \ldots, t_K \in [0, \ldots, T]$ such that $0 < t_1 < t_2 < \cdots < t_K = T$. We allow K to grow with the number of observations. In the most extreme case, one may set $t_k = \inf_t\{N(t) \geq N(t_{k-1}) + 1\}, \forall k \geq 2$. That is, we make a decision regarding the null hypothesis upon the arrival of each observation. In addition, we assume that t_1 is large so that there are enough number of samples $N(t_1)$ to guarantee the validity of the normal approximation for $B(t_1)$. We remark that in typical tech companies such as Amazon, Facebook, etc., massive data are collected even within a short time interval. Large sample approximation is validated in these applications.

To guarantee our test controls the type-I error, we reject H_0 and terminate the experiment at t_k if $\sqrt{N(t_k)}S(t_k) \geq z_k$ for some $k = 1, \ldots, K$ with some suitably chosen $z_1, \ldots, z_K > 0$ that satisfy

$$\Pr\left(\max_{k\in\{1,\ldots,K\}}\{\sqrt{N(t_k)}S(t_k) - z_k\} > 0\right) \leq \alpha + o(1)$$

for a given significance level $\alpha > 0$ under H_0. In view of (2), it suffices to find $\{z_k\}_k$ that satisfy

$$\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)B(t_k)}{\sqrt{N(t)}\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}]} - z_k\right) > 0\right\} \leq \alpha + o(1), \tag{4}$$

where the stochastic process $B(\cdot)$ is defined in (3).

To determine $\{z_k\}_k$, we need to derive the asymptotic distribution of the left-hand-side (LHS) of (4). To this end, define a mean-zero Gaussian process $G(t)$ with covariance function

$$\mathrm{cov}(G(t), G(t')) = N^{1/2}(t)N^{-1/2}(t')\sum_{a\in\{0,1\}}\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}, \quad \forall 0 < t \leq t'.$$

In the following, we show that the LHS of (4) can be uniformly approximated based on $G(\cdot)$, for any $\{z_k\}_{k=1,\ldots,K}$. To establish our theoretical results, we need some regularity conditions on $\varphi(\cdot)$. To save space, we summarize these assumptions in (A3) and put them in Appendix A. A random variable Z is said to have a sub-Gaussian tail if there exist some constants $C, \nu > 0$ such that

$$P(|Z| > z) \leq C\exp(-\nu z^2),$$

for any z.

Theorem 1 *Assume (A1)-(A3) hold. For $a = 0, 1$, assume $\inf_{x\in\mathbb{X}}\pi^*(a, x) > 0$, $E[\{Y^*(a)\}^2|X]$ is a bounded random variable, and $|Y^*(a)|$ has a sub-Gaussian tail. Assume there exists some $0 < \alpha_0 \leq 1$ such that for any sequence $\{j_n\}_n$ that satisfies $j_n^{\alpha_0}/\log^{\alpha_0} j_n \gg q^2$, the following event occurs with probability at least $1 - O(j_n^{-\alpha_0})$,*

$$\sup_{a\in\{0,1\}} E\left|\sum_{i=1}^{k}\{\pi_{i-1}(a, x) - \pi^*(a, x)\}\right| \leq O(1)qk^{1-\alpha_0}\log^{\alpha_0} k, \quad \forall k \geq j_n, \tag{5}$$

where $O(1)$ denotes some positive constant. Assume $q = O(N^{\alpha^*}(t_1))$ for some $0 \leq \alpha^* < \min(1/3, \alpha_0/2)$ and $N(t_1) \gg \log N(T)$ almost surely. Then conditional on the counting process $N(\cdot)$, there exists some constant $c > 0$ such that

$$\sup_{z_1,\ldots,z_K} \left| Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)B(t_k)}{\sqrt{N(t)}\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}]} - z_k \right) > 0 \right\} \right.$$
$$\left. - Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)G(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z_k \right) > 0 \right\} \right|$$
$$\leq c\Big[N^{-1/8}(t_1)\log^{15/8}\{KN(t_1)\} + \{N^{-\alpha_0/3}(t_1) + q^{3/2}N^{-\alpha_0}(t_1)\}\log^{(5+\alpha_0)/3}\{KN(t_1)\}$$
$$+ err\log^{1/2}\{KN(t_1)\}\Big].$$

Theorem 1 implies that the approximation error depends on the number of observations obtained up to the first decision point $N(t_1)$, the maximum number of interim analyses K, the total number of basis functions q, err, and α_0, which characterizes the convergence rate of the treatment assignment mechanism $\sum_{i=1}^n n^{-1}\pi_{i-1}$. Clearly, the error will decay to zero when the followings hold with probability tending to 1,

$$\log(K) \ll \min\{N^{1/15-2\alpha^*/5}(t_1), N^{(\alpha_0-3\alpha^*)/(5+\alpha_0)}(t_1)\}. \tag{6}$$

In Section 3.3, we show that $\alpha_0 = 1/2$, when an ϵ-greedy strategy is used for randomization to balance the trade-off between exploration and exploitation. Condition (6) is satisfied when K grows polynomially fast with respect to $N(t_1)$. In addition to ϵ-greedy, other adaptive allocation procedures (e.g., upper confidence bound or Thompson sampling) could be applied as well.

As discussed in the introduction, the derivation of Theorem 1 is nontrivial. One way to obtain the magnitude of the approximation error is to apply the strong approximation theorem for multidimensional martingales (see Morrow and Philipp, 1982; Zhang, 2004). However, the rate of approximation typically depends on the dimension and decays fast as the dimension increases. To derive Theorem 1, we view $\{\varphi^\top(x)B(t_K)\}_{x\in\mathbb{X},k\in\{1,\ldots,\kappa\}}$ as a high-dimensional martingale and adopt the Gaussian approximation techniques that have been recently developed by Belloni and Oliveira (2018). In view of (2), an application of Theorem 1 yields the following result.

Theorem 2 *Assume that the conditions of Theorem 1 hold,* (6) *holds with probability tending to* 1*. Then for any $z_1, \ldots, z_k$ that satisfy*

$$Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)G(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z_k \right) > 0 \right\} = \alpha + o(1), \tag{7}$$

as $N(t_1)$ diverges to infinity, we have under H_0,

$$Pr\left(\max_{k\in\{1,\ldots,K\}} \{S(t_k) - z_k\} > 0 \right) \leq \alpha + o(1).$$

The above equality holds when $\beta_0^* = \beta_1^*$.

Theorem 2 suggests that the type-I error rate of the proposed test can be well controlled. It remains to find critical values $\{z_k\}_{1\le k\le K}$ that satisfy (7). In the next section, we propose a bootstrap-assisted procedure to determine these critical values.

3.2 Bootstrap stopping boundary

We first outline a method based on the wild bootstrap (Wu, 1986) to approximate the limiting distribution of $\{S(t_k)\}_k$. Then we discuss its limitation and present our proposal, a scalable bootstrap algorithm to determine the stopping boundary.

The idea is to generate bootstrap samples $\{\widehat{\beta}_a^{\mathrm{MB}}(t_k)\}_{a,k}$ that have asymptotically the same joint distribution as $\{\widehat{\beta}_a(t_k)-\beta_a^*\}_{a,k}$. Then the joint distribution of $\{S(t_k)\}_k$ can be well-approximated by the conditional distribution of $\{\widehat{S}^{\mathrm{MB}}(t_k)\}_k$ given the data, where $\widehat{S}^{\mathrm{MB}}(t)=\sup_{x\in\mathbb{X}}\varphi^\top(x)\{\widehat{\beta}_1^{\mathrm{MB}}(t)-\widehat{\beta}_0^{\mathrm{MB}}(t)\}$ for any t. Specifically, let $\{\xi_i\}_{i=1}^{+\infty}$ be a sequence of i.i.d. standard normal random variables independent of $\{(X_i,A_i,Y_i)\}_{i=1}^{+\infty}$. For $a\in\{0,1\}$, define

$$\widehat{\beta}_a^{\mathrm{MB}}(t)=\widehat{\Sigma}_a^{-1}(t)\left[\frac{1}{N(t)}\sum_{i=1}^{N(t)}\mathbb{I}(A_i=a)\varphi(X_i)\{Y_i-\varphi^\top(X_i)\widehat{\beta}(t)\}\xi_i\right],\quad \forall a\in\{0,1\}.$$

Both the asymptotic means of $\sqrt{N(t)}\widehat{\beta}_a^{\mathrm{MB}}(t)$ and $\sqrt{N(t)}(\widehat{\beta}_a(t)-\beta_a^*)$ are zero. In addition, their covariance functions are asymptotically the same. By design, $\{\widehat{\beta}_a^{\mathrm{MB}}(t_k)\}_{a,k}$ is multivariate normal. Similar to (3), we can show $\{\widehat{\beta}_a(t_k)-\beta_a^*\}_{a,k}$ is asymptotically multivariate normal. Consequently, the limiting distributions of $\{\widehat{\beta}_a^{\mathrm{MB}}(t_k)\}_{a,k}$ and $\{\widehat{\beta}_a(t_k)-\beta_a^*\}_{a,k}$ are asymptotically equivalent. As such, the bootstrap approximation is valid.

However, calculating $\widehat{\beta}_a^{\mathrm{MB}}(t_k)$ requires $O(N(t_k)q+q^3)$ operations. The time complexity of the resulting bootstrap algorithm is $O(BN(t_k)q+q^3)$ up to the k-th interim stage, where B is the total number of bootstrap samples. This can be time consuming when $\{N(t_k)-N(t_{k-1})\}_{k=1}^K$ are large. To facilitate the computation, we observe that in the calculation of $\widehat{\beta}_a^{\mathrm{MB}}$, the random noise is generated upon the arrival of each observation. This is unnecessary as we aim to approximate the distribution of $\widehat{\beta}_a(\cdot)$ only at finitely many time points.

We next present our proposal. Let $\{e_{i,a}\}_{i=1,\ldots,K,a=0,1}$ be a sequence of i.i.d $N(0,I_q)$ random vectors independent of the observed data, where I_q denotes the $q\times q$ identity matrix. At the k-th interim stage, we compute

$$\widehat{S}^{\mathrm{MB}*}(t_k)=\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)\{\widehat{\beta}_1^{\mathrm{MB}*}(t_k)-\widehat{\beta}_0^{\mathrm{MB}*}(t_k)\}}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t)-\widehat{\beta}_0(t)\}]},$$

where $\widehat{\beta}_a^{\mathrm{MB}*}(t_k)$ equals

$$\frac{1}{N(t_k)}\sum_{j=1}^{k}\left(\sum_{i=N(t_{j-1})+1}^{N(t_j)}\widehat{\Sigma}_a^{-1}(t_j)\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi(X_i)^\top\widehat{\beta}_a(t_j)\}^2\widehat{\Sigma}_a^{-1}(t_j)\right)^{1/2}e_{j,a}.$$

For any k_1 and k_2, the conditional covariance of $\sqrt{N(t_{k_1})}\{\widehat{\beta}_1^{\mathrm{MB}*}(t_{k_1}) - \widehat{\beta}_0^{\mathrm{MB}*}(t_{k_1})\}$ and $\sqrt{N(t_{k_2})}\{\widehat{\beta}_1^{\mathrm{MB}*}(t_{k_2}) - \widehat{\beta}_0^{\mathrm{MB}*}(t_{k_2})\}$ equals

$$\frac{1}{\sqrt{N(t_{k_1})N(t_{k_2})}} \sum_{a=0}^{1} \sum_{j=1}^{k_1} \sum_{i=N(t_{j-1})+1}^{N(t_j)} \widehat{\Sigma}_a^{-1}(t_j)\mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\widehat{\Sigma}_a^{-1}(t_j).$$

Under the given conditions in Theorem 1, it is to converge to

$$\frac{\sqrt{N(t_{k_1})}}{\sqrt{N(t_{k_2})}} \sum_{a=0}^{1} \Sigma_a^{-1}\Phi_a\Sigma_a^{-1} = \mathrm{cov}(G(t_{k_1}), G(t_{k_2})).$$

This means $\{\sqrt{N(t_k)}(\widehat{\beta}_1^{\mathrm{MB}*}(t_k) - \widehat{\beta}_0^{\mathrm{MB}*}(t_k))\}_k$ and $\{G(t_k)\}_k$ have the same asymptotic distribution. Consequently, $\{\sqrt{N(t_k)}\widehat{S}^{\mathrm{MB}*}(t_k)\}_{k=1}^K$ can be used to approximate the joint distribution of $\{\sup_{x\in\mathbb{X}} \varphi^\top(x)G(t_k)/\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}\}_{k=1}^K$.

To choose $\{z_k\}_k$ that satisfies (7), we adopt the α-spending approach that allocates the total allowable type I error at each interim stage according to an error-spending function. This guarantees our test controls the type-I error. We begin by specifying an α spending function $\alpha(t)$ that is non-increasing and satisfies $\alpha(0) = 0$, $\alpha(T) = \alpha$. Popular choices of $\alpha(\cdot)$ include

$$\begin{aligned} &\alpha_1(t) = \alpha \log\left(1 + (e-1)\frac{t}{T}\right), && \alpha_2(t) = 2 - 2\Phi\left(\frac{\Phi^{-1}(1-\alpha/2)\sqrt{T}}{\sqrt{t}}\right), \\ &\alpha_3(t) = \alpha\left(\frac{t}{T}\right)^\theta, \ \text{ for } \ \theta > 0, && \alpha_4(t) = \alpha\frac{1-\exp(-\gamma t/T)}{1-\exp(-\gamma)}, \ \text{ for } \ \gamma \neq 0, \end{aligned} \tag{8}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard normal variable and $\Phi^{-1}(\cdot)$ is its quantile function.

Based on $\alpha(\cdot)$, we iteratively calculate $\widehat{z}_k, k = 1, \ldots, K$ as the solution of

$$\Pr^*\left\{\max_{j\in\{1,\ldots,k-1\}}\left(\sqrt{N(t_j)}\widehat{S}^{\mathrm{MB}*}(t_j) - \widehat{z}_j\right) \le 0, \sqrt{N(t_k)}\widehat{S}^{\mathrm{MB}*}(t_k) > \widehat{z}_k\right\} = \alpha(t_k) - \alpha(t_{k-1}), \tag{9}$$

and reject H_0 when $\sqrt{N(t_k)}S(t_k) > \widehat{z}_k$ holds for some k.

The validity of the bootstrap test is summarized in Theorems 3 and 4 below.

Theorem 3 *Assume the conditions in Theorem 1 hold. Then conditional on the counting process $N(\cdot)$, we have*

$$\begin{aligned} &\sup_{z_1,\ldots,z_K}\left|Pr^*\left\{\max_{k\in\{1,\ldots,K\}}\left(\sqrt{N(t_k)}\widehat{S}^{MB*}(t_k) - z_k\right) > 0\right\}\right. \\ &\left. - \ Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)G(t_k)}{\sqrt{\sum_a \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z_k\right) > 0\right\}\right| \\ &\le c\left[q^{6/5}N^{-1/6}(t_1)\log^{11/6}\{KN(t_1)\} + q^{5/3}N^{-\alpha_0/3}(t_1)\log^{(5+\alpha_0)/3}\{KN(t_1)\}\right] \end{aligned}$$

for some constant $c > 0$ with probability at least $1 - O(N^{-\alpha_0}(t_1))$, where $Pr^*(\cdot)$ denotes the probability measure conditional on the data stream $\{X_i, A_i, Y_i\}_{i=1}^{+\infty}$.

Theorem 4 *Assume the conditions in Theorem 3 hold. Then conditional on $N(\cdot)$, the critical values $\{\widehat{z}_k\}_k$ satisfy*

$$\begin{aligned}&\left|Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)G(t_k)}{\sqrt{\sum_a\varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}}-\widehat{z}_k\right)>0\right\}-\alpha\right| \\ \leq& c\left[q^{6/5}N^{-1/6}(t_1)\log^{11/6}\{KN(t_1)\}+q^{5/3}N^{-\alpha_0/3}(t_1)\log^{(5+\alpha_0)/3}\{KN(t_1)\}\right],\end{aligned} \tag{10}$$

for some constant $c>0$.

When the RHS of (10) is $o_p(1)$, it follows from Theorems 2 and 4 that our test is valid. The conditional distribution in (9) can be approximated by the empirical distribution of Bootstrap samples.

Next, we study the power property and the stopping time of the proposed test. Let sig$=\sup_x\{Q(x,1)-Q(x,0)\}$ denote the qualitative treatment effect signal. We will show that the power of the proposed test approaches to one as long as $\sqrt{N(t_K)}\text{sig}\gg\sqrt{q\log\{N(t_1)\}}$. In addition, the stopping time depends crucially on sig. Specifically, we will show that the proposed test will reject the null at the kth interim stage as long as $\sqrt{N(t_k)}\text{sig}\gg\sqrt{q\log\{N(t_1)\}}$.

Theorem 5 *Assume the conditions in Theorem 2 hold, there exists some sufficiently large constant $C>0$ such that $\min\{\alpha(t_1),\min_k\alpha(t_k)-\alpha(t_{k-1})\}\gg N^{-C}(t_1)$, and that $\sqrt{N(t_K)}sig\gg\sqrt{q\log\{N(t_1)\}}$. Then the power of the proposed test approaches one as $N(t_1)$ diverges to infinity. In addition, for any k such that $\sqrt{N(t_k)}sig\gg\sqrt{q\log\{N(t_1)\}}$, the stopping time will be smaller than or equal to t_k, with probability tending to 1.*

Finally, we remark that our test can be online updated as batches of observations arrive at the end of each interim stage. A pseudocode summarizing our procedure is given in Algorithm 1. In Algorithm 1, we use O_{p+1} to denote a $(p+1)\times(p+1)$ zero matrix and 0_{p+1} to denote a $(p+1)$-dimensional zero vector. The spatial complexity of the proposed algorithm is $O(Bq+q^2)$, where B is the number of bootstrap samples. The time complexity is $O(Bkq^2+N(t_k)q^2)$ up to the k-th interim stage. Suppose $N(t_j)-N(t_{j-1})=n$ for any $1\leq j\leq K$, we have $Bkq^2+N(t_k)q^2=(B+n)kq^2\ll Bnkq=BN(t_k)q$ when $Bn\gg(B+n)q$, or equivalently, $\min(B,n)\gg q$. Hence, our procedure is much faster compared to the standard wild bootstrap as long as the number of bootstrap samples and the number of observations per batch are much large than the number of basis functions.

3.3 Adaptive randomization

In practice, the company might want to allocate more traffic to a better treatment based on the observed data stream. The ϵ-greedy strategy is commonly used to balance the trade-off between exploration and exploitation. For a given $0<\varepsilon_0<1$, consider the following randomization procedure: for some integer $N_0>0$ and any $j\geq N_0$, $a\in\{0,1\}$, $x\in\mathbb{X}$, we set

$$\pi_{j-1}(a,x)=(1-\varepsilon_0)a\mathbb{I}\{\varphi^\top(x)(\widehat{\beta}_{1,j-1}-\widehat{\beta}_{0,j-1})>0\}+\varepsilon_0(1-a)\mathbb{I}\{\varphi^\top(x)(\widehat{\beta}_{1,j-1}-\widehat{\beta}_{0,j-1})\leq 0\},$$

Input: Number of bootstrap samples B, an α spending function $\alpha(\cdot)$.
Initialize: $n=0$, $\widehat{\Sigma}_0=\widehat{\Sigma}_1=O_{p+1}$, $\widehat{\gamma}_0=\widehat{\gamma}_1=0_{p+1}$, $\widehat{\beta}_{0,b}=\widehat{\beta}_{1,b}=0_{p+1}$, and a set $\mathcal{I}=\{1,\ldots,B\}$.
For $k=1$ to K **do**
  Initialize: $m=0$ and $\widehat{\Phi}_0=\widehat{\Phi}_1=O_{p+1}$.
  Step 1: Online update of $\widehat{\beta}_a$
  For $i=N(t_{k-1})+1$ to $N(t_k)$ **do**
    $n=n+1$ and $m=m+1$;
    $\widehat{\Sigma}_a=(1-n^{-1})\widehat{\Sigma}_a+n^{-1}\varphi(X_i)\varphi^\top(X_i)\mathbb{I}(A_i=a), a=0,1$;
    $\widehat{\gamma}_a=(1-n^{-1})\widehat{\gamma}_a+n^{-1}\varphi(X_i)Y_i\mathbb{I}(A_i=a), a=0,1$;
  Compute $\widehat{\beta}_a=\widehat{\Sigma}_a^{-1}\widehat{\gamma}_a$ for $a\in\{0,1\}$;
  Step 2: Bootstrap
  For $i=N(t_{k-1})+1$ to $N(t_k)$ **do**
    $\widehat{\Phi}_a=\widehat{\Phi}_a+\widehat{\Sigma}_a^{-1}\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\widehat{\beta}_a\}^2\widehat{\Sigma}_a^{-1}\mathbb{I}(A_i=a), a=0,1$;
  Compute $S=\sup_{x\in\mathbb{X}}[\varphi^\top(x)(\widehat{\beta}_1-\widehat{\beta}_0)/\sqrt{\sum_a\varphi^\top(x)\widehat{\Sigma}_a^{-1}\widehat{\Phi}_a\widehat{\Sigma}_a^{-1}}]$;
  For $b=1,\ldots,B$ **do**
    Generate two independent $N(0,I_{p+1})$ Gaussian vectors e_0, e_1;
    $\widehat{\beta}_{a,b}=(1-mn^{-1})\widehat{\beta}_{a,b}+n^{-1}\widehat{\Phi}_a^{1/2}e_a, a=0,1$;
    Compute $\widehat{S}_b=\sup_{x\in\mathbb{X}}[\varphi^\top(x)(\widehat{\beta}_{1,b}-\widehat{\beta}_{0,b})/\sqrt{\sum_a\varphi^\top(x)\widehat{\Sigma}_a^{-1}\widehat{\Phi}_a\widehat{\Sigma}_a^{-1}}]$;
  Step 3: Reject or not
  Set z to be the upper $\{\alpha(t)-|\mathcal{I}|^c/B|\}/(1-|\mathcal{I}^c|/B)$-th percentile of $\{\widehat{S}_b\}_{b\in\mathcal{I}}$;
  Update $\mathcal{I}$ as $\mathcal{I}\leftarrow\{b\in\mathcal{I}:\widehat{S}_b\le z\}$.
  If $S>z$: Reject H_0 and terminate the experiment.

Algorithm 1: the Pseudocode that summarizing the online bootstrap testing procedure.

where

$$\widehat{\beta}_{a,j}=\widehat{\Sigma}_{a,j}^{-1}\frac{1}{j}\sum_{i=1}^{j}\{\mathbb{I}(A_i=a)\varphi(X_i)Y_i\}\ \text{ and }\ \widehat{\Sigma}_{a,j}=\frac{1}{j}\sum_{i=1}^{j}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i).$$

It is immediate to see that $\widehat{\Sigma}_a(t)=\widehat{\Sigma}_{a,n(t)}$ and $\widehat{\beta}_a(t)=\widehat{\beta}_{a,n(t)}$. Define

$$\pi^*(a,x)=(1-\varepsilon_0)a\mathbb{I}\{\varphi^\top(x)(\beta_1-\beta_0)>0\}+\varepsilon_0(1-a)\mathbb{I}\{\varphi^\top(x)(\beta_1-\beta_0)\le 0\}$$

for any $a\in\{0,1\}$ and $x\in\mathbb{X}$.

Lemma 2 *Assume (A1)-(A3) hold. Assume* $\inf_{x\in\mathbb{X}}\pi^*(a,x)>0$ *and* $|Y^*(a)|$ *is bounded almost surely, for* $a\in\{0,1\}$. *Assume* $Pr(|Q_0(X,1)-Q_0(X,0)|\le\epsilon)\le L_0\epsilon$, *for some constant* $L_0>0$ *and any* $\epsilon>0$. *Then for any* $\{j_n\}_n$ *that satisfies* $\sqrt{j_n}/\sqrt{\log j_n}\gg q^2$, *the following event occurs with probability at least* $1-O(j_n^{-1})$,

$$\sum_{a\in\{0,1\}}E^{\mathcal{F}_{i-1}}\left|\sum_{i=1}^{k}\{\pi_{i-1}(a,X)-\pi^*(a,X)\}\right|\preceq q\sqrt{k\log k},\quad\forall k\ge j_n. \tag{11}$$

We make a few remarks. First, Lemma 2 implies that Condition (5) in Theorem 1 automatically holds with $\alpha_0 = 1/2$, when the epsilon-greedy strategy is used. Nonetheless, Condition (5) is weaker than (11), as it only requires the average estimated policy aggregated over different interim stages to converge at certain rate. Second, by setting $\epsilon = 0$, the assumption $\Pr(|Q_0(X,1) - Q_0(X,0)| \leq \epsilon) \leq L_0\epsilon$ requires $\Pr(Q_0(X,1) = Q_0(X,0)) = 0$, almost surely. It essentially requires that the difference between the two Q-functions is nonzero, almost surely. This implies that the optimal decision rule is uniquely defined and is thus identifiable. This condition is necessary to guarantee that the estimated policy converges to the oracle optimal policy. Without this identifiability condition, the estimated optimal decision rule could fluctuate randomly and will not stabilize (Luedtke and van der Laan, 2016). Third, our condition also requires $\Pr(0 < |Q_0(X,1) - Q_0(X,0))| \leq \epsilon) \leq L_0\epsilon$ for any $\epsilon > 0$. The latter condition is well-known as the margin condition that is commonly imposed in the literature to bound the difference between the expect return under the estimated and the optimal decision rule (Qian and Murphy, 2011; Luedtke and van der Laan, 2016; Shi et al., 2020a). The identifiability assumption is not needed to establish the rate of convergence of the expected return under the estimated decision rule.

4. Numerical studies

4.1 Simulation studies

4.1.1 Testing QTE

In this section, we conduct Monte Carlo simulations to examine the finite sample properties of the proposed test. We generated the potential outcomes as $Y_i^*(a) = 1 + (X_{i1} - X_{i2})/2 + a\tau(X_i) + \varepsilon_i$, where ε_i's are i.i.d $N(0, 0.5^2)$. The covariates $X_i = (X_{i1}, X_{i2}, X_{i3})^\top$ were generated as follows. We first generated $X_i^* = (X_{i1}^*, X_{i2}^*, X_{i3}^*)^\top$ from a multivariate normal distribution with zero mean and covariance matrix equal to $\{0.5^{|i-j|}\}_{i,j}$. Then we set $X_{ij} = X_{ij}^*\mathbb{I}(X_{ij}^*| \leq 2) + 2\text{sgn}(X_{ij}^*)\mathbb{I}(X_{ij}^*| > 2)$. We consider two randomization designs. In the first design, the treatment assignment is nondynamic and completely random. Specifically we set $\pi_i(a,x) = 0.5$, for any a, x and i. In the second design, we use an ϵ-greedy strategy to generate the treatment with $\varepsilon = 0.3$. In addition, we set $N(T_1) = 2000$ and $N(T_j) - N(T_{j-1}) = 2n$ for $2 \leq j \leq K$ and some $n > 0$. We consider two combinations of (n, K), corresponding to $(n, K) = (200, 5)$ and $(20, 50)$.

We set the significance level $\alpha = 0.05$ and choose $B = 10000$. We set $\tau(X_i) = \phi_\delta\{(X_{i1} + X_{i2})/\sqrt{2}\}X_{i3}^2$ for some function ϕ_δ parameterized by some $\delta \geq 0$. We consider two scenarios for ϕ_δ. Specifically, we set $\phi_\delta(x) = \delta x^2/3$ in Scenario 1 and $\phi_\delta = \delta\cos(\pi x)$ in Scenario 2. For each setting, we further consider four cases by setting $\delta = 0, 0.05, 0.10$, and 0.15. When $\delta = 0$, H_0 holds. Otherwise, H_1 holds. For all settings, we construct the basis function $\varphi(\cdot)$ using additive cubic splines. For each univariate spline, we set the number of internal knots to be 4. These knots are equally spaced between $[-2, 2]$.

We denote our test by BAT, short for bootstrap-assisted test. We run our experiments on a single computer instance with 40 Intel(R) Xeon(R) 2.20GHz CPUs. It takes 1-2 seconds on average to compute each test. In Figure 2, we plot the rejection probabilities of our tests and the average stopping times (defined as the average number of samples consumed when the experiment is terminated), aggregated over 400 simulations when $\alpha_1(\cdot)$ is chosen as the

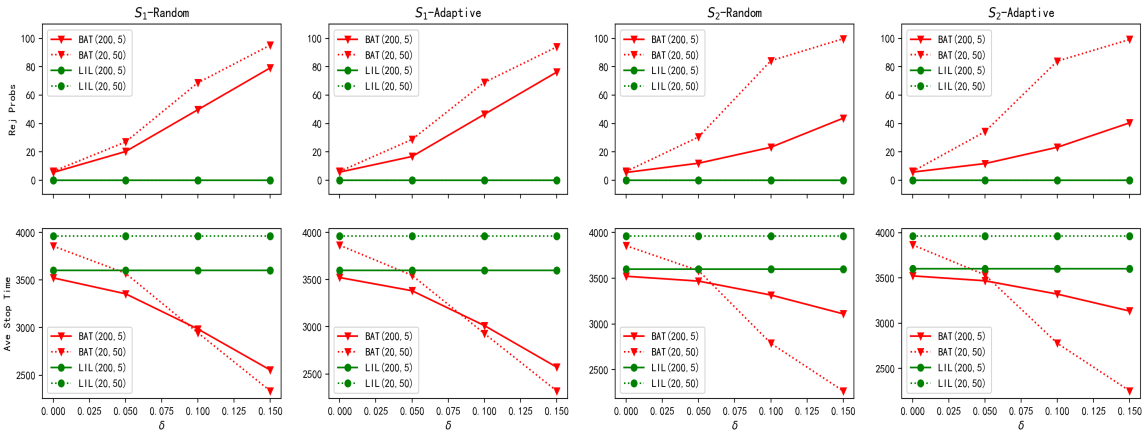

Figure 2: Rejection probabilities and average stopping times of the proposed test when $\alpha_1(\cdot)$ is chosen as the spending function. From left to right: Scenario 1 with random design, Scenario 1 with ϵ-greedy design, Scenario 2 with random design and Scenario 2 with ϵ-greedy design.

spending function. The detailed values of these rejection probabilities and average stopping times can be found in Table 1. It can be seen that the type-I error rates are close to the nominal level in all cases. The power of our test increases as δ increases, demonstrating its consistency. In addition, when $\delta > 0$, our experiments are stopped early in all cases.

To further evaluate our method, we compare it with a test based on the law of iterated logarithm (denoted by LIL). LIL determines the decision boundary based on an always valid finite error bound (see Appendix C for details about the competing method). It can be seen from Figure 2 that our method has good power properties, whereas LIL fails to detect the alternative and does not have any power at all.

4.1.2 Testing ATE

We extend our proposal to testing ATE in this section. Specifically, we focus on testing the following hypothesis,

$$H_0 : \mathrm{E}Y_i^*(1) \leq \mathrm{E}Y_i^*(0) \quad \text{versus} \quad H_1 : \mathrm{E}Y_i^*(1) > \mathrm{E}Y_i^*(0).$$

Under (A1) and (A2), it suffices to test

$$H_0 : \mathrm{E}Q(X_i, 1) \leq \mathrm{E}Q(X_i, 0) \quad \text{versus} \quad H_1 : \mathrm{E}Q(X_i, 1) > \mathrm{E}Q(X_i, 0).$$

We similarly use basis approximations to model the Q-function. The proposed method is very similar to that in Section 3. The main difference lies in that instead of employing a supremum type statistics, we aggregate the estimated treatment effect and set $S(t) = N^{-1} \sum_{i=1}^{N} \varphi^\top(X_i)\{\widehat{\beta}_1(t) - \widehat{\beta}_0(t)\}$. Its asymptotic distribution can be approximated by the corresponding bootstrap statistic $N^{-1} \sum_{i=1}^{N} \varphi^\top(X_i)\{\widehat{\beta}_1^{\mathrm{MB}*}(t) - \widehat{\beta}_0^{\mathrm{MB}*}(t)\}$. The proposed algorithm can then be applied to determine the rejection boundary. To save space, we summarize our proposal in the following algorithm. We next conduct simulation studies to evaluate this algorithm.

Input: Number of bootstrap samples B, an α spending function $\alpha(\cdot)$.
Initialize: $n = 0$, $\widehat{\Sigma}_0 = \widehat{\Sigma}_1 = O_{p+1}$, $\widehat{\gamma}_0 = \widehat{\gamma}_1 = 0_{p+1}$, $\widehat{\beta}_{0,b} = \widehat{\beta}_{1,b} = 0_{p+1}$, $\bar{\varphi} = 0$ and a set $\mathcal{I} = \{1, \ldots, B\}$.
For $k = 1$ to K **do**
Initialize: $m = 0$, $\widehat{\phi} = 0$ and $\widehat{\Phi}_0 = \widehat{\Phi}_1 = O_{p+1}$.

For $i = N(t_{k-1}) + 1$ to $N(t_k)$ **do**
$n = n + 1$, $m = m + 1$ and $\bar{\varphi} = n^{-1}(n-1)\bar{\varphi} + n^{-1}\varphi(X_i)$;
$\widehat{\Sigma}_a = (1 - n^{-1})\widehat{\Sigma}_a + n^{-1}\varphi(X_i)\varphi^\top(X_i)\mathbb{I}(A_i = a), a = 0, 1$;
$\widehat{\gamma}_a = (1 - n^{-1})\widehat{\gamma}_a + n^{-1}\varphi(X_i)Y_i\mathbb{I}(A_i = a), a = 0, 1$;
Compute $\widehat{\beta}_a = \widehat{\Sigma}_a^{-1}\widehat{\gamma}_a$ for $a \in \{0, 1\}$ and $S = \bar{\varphi}^\top(\widehat{\beta}_1 - \widehat{\beta}_0)$;

For $i = N(t_{k-1}) + 1$ to $N(t_k)$ **do**
$\widehat{\phi} = \widehat{\phi} + [\{\varphi(X_i) - \bar{\varphi}\}^\top(\widehat{\beta}_1 - \widehat{\beta}_0)]^2$.
$\widehat{\Phi}_a = \widehat{\Phi}_a + \widehat{\Sigma}_a^{-1}\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_a\}^2\widehat{\Sigma}_a^{-1}\mathbb{I}(A_i = a), a = 0, 1$;
For $b = 1, \ldots, B$ **do**
Generate two independent $N(0, I_{p+1})$ Gaussian vectors e_0, e_1, $N(0, 1)$ random variable e_2;
$\widehat{\beta}_{a,b} = (1 - mn^{-1})\widehat{\beta}_{a,b} + n^{-1}\widehat{\Phi}_a^{1/2}e_a + n^{-1}\widehat{\phi}^{1/2}e_2, a = 0, 1$;
Compute $\widehat{S}_b = \bar{\varphi}^\top(\widehat{\beta}_{1,b} - \widehat{\beta}_{0,b})$;

Set z to be the upper $\{\alpha(t) - |\mathcal{I}|^c/B|\}/(1 - |\mathcal{I}^c|/B)$-th percentile of $\{\widehat{S}_b\}_{b\in\mathcal{I}}$;
Update $\mathcal{I}$ as $\mathcal{I} \leftarrow \{b \in \mathcal{I} : \widehat{S}_b \le z\}$.
If $S > z$:
Reject H_0 and terminate the experiment;

We compare our procedure with the always valid test (AVT, Johari et al., 2017) that extends the two-sample t-test for sequential monitoring. AVT requires to impose a parametric likelihood model assumption (in addition to the conditional mean model) to construct a likelihood-ratio-based "always valid p-value". To implement the test, we follow the proposal detailed in Section 4.3 of Johari et al. (2017), assume the responses are normal with constant means and known variances, and compute the mixture sequential probability ratio test statistic accordingly. Please refer to Appendix C for details. We remark that the validity of the resulting test requires no confounding variables exist, as the test statistic is derived without adjusting for confounders.

We generate the potential outcomes with the same model, except that ε_i's are i.i.d $N(0, 1)$. However, we set $N(T_1) = 1000$ and $N(T_j) - N(T_{j-1}) = 2n$ for $2 \le j \le K$ and some $n > 0$. We consider two combinations of (n, K), corresponding to $(n, K) = (100, 5)$ and $(10, 50)$. For all settings, we use a linear function to approximate Q.

In Table 2 and Figure 3, we show the rejection probabilities and average stopping times of the proposed test aggregated over 400 simulations, when $\alpha_1(\cdot)$ is chosen as the spending function. It can be seen that our method behaves better than the always valid test when the effect size is small, and comparable when the effect size is large. The always valid test fails in the adaptive randomization settings, as the type-I error rates are around 50%

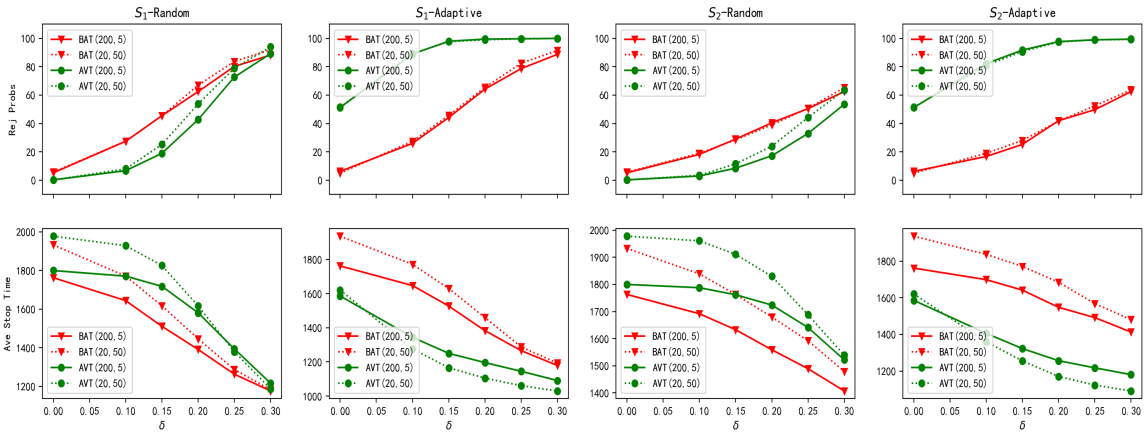

Figure 3: Rejection probabilities and average stopping times of the proposed test when $\alpha_1(\cdot)$ is chosen as the spending function. From left to right: Scenario 1 with random design, Scenario 1 with ϵ-greedy design, Scenario 2 with random design and Scenario 2 with ϵ-greedy design.

under the null hypothesis. This is because under such a design, the time-varying variables will confound the treatment and the outcome. As commented before, the always valid test is established under settings where no confounders exist. It is expected to fail under the adaptive design.

		method	BAT				LIL			
			Random		Adaptive		Random		Adaptive	
	(n, K)	δ	rej probs	E[stop]	rej probs	E[stop]	rej probs	E[stop]	rej probs	E[stop]
S1	(200, 5)	0.00	5.5(1.1)	3522(16)	5.8(1.2)	3522(16)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
		0.05	20.2(2.0)	3355(27)	16.8(1.9)	3382(26)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
		0.10	0.49.8(2.5)	2985(36)	46.5(2.5)	3013(36)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
		0.15	79.2(2.0)	2554(35)	76.2(2.1)	2572(35)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
	(20, 50)	0.00	6.2(1.2)	3856(20)	6.2(1.2)	3864(19)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
		0.05	27.0(2.2)	3571(35)	28.7(2.3)	3545(36)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
		0.10	68.8(2.3)	2945(42)	69.0(2.3)	2929(42)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
		0.15	95.2(1.1)	2334(29)	94.0(1.2)	2320(28)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
S2	(200, 5)	0.00	5.5(1.1)	3522(16)	5.8(1.2)	3522(16)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
		0.05	12.0(1.6)	3469(20)	11.8(1.6)	3468(20)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
		0.10	23.2(2.1)	3317(28)	23.2(2.1)	3321(28)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
		0.15	43.8(2.5)	3111(33)	40.5(2.5)	3134(33)	0.0(0.0)	3600(0)	0.0(0.0)	3600(0)
	(20, 50)	0.00	6.2(1.2)	3856(20)	6.2(1.2)	3864(19)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
		0.05	30.5(2.3)	3584(32)	34.2(2.4)	3533(34)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
		0.10	84.2(1.8)	2789(35)	84.0(1.8)	2778(36)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)
		0.15	99.8(0.2)	2268(18)	99.2(0.4)	2252(17)	0.0(0.0)	3960(0)	0.0(0.0)	3960(0)

Table 1: QTE: rejection probabilities (multiplied by 100) and average stopping times under Scenarios 1 and 2 when $\alpha_1(\cdot)$ is chosen as the spending function. Standard errors are reported in the parentheses.

4.2 Real data analysis

In this section, we apply the proposed method to a Yahoo! Today Module user click log dataset[1], which contains 45,811,883 user visits to the Today Module, during the first

1. `https://webscope.sandbox.yahoo.com/catalog.php?datatype=r&did=49`

		method	BAT				AVT			
			Random		Adaptive		Random		Adaptive	
	(n, K)	δ	rej probs	E[stop]	rej probs	E[stop]	rej probs	E[stop]	rej probs	E[stop]
S1	(200, 5)	0.00	5.2(1.1)	1763(8)	6.2(1.2)	1762(8)	0.2(0.2)	1800(0)	51.2(2.5)	1586(11)
		0.10	27.5(2.2)	1644(15)	26.0(2.2)	1647(14)	6.8(1.3)	1771(6)	89.0(1.6)	1344(9)
		0.15	45.5(2.5)	1511(17)	44.2(2.5)	1527(17)	19.0(2.0)	1718(10)	98.0(0.7)	1250(6)
		0.20	62.5(2.4)	1391(18)	64.2(2.4)	1383(18)	42.8(2.5)	1581(15)	99.5(0.4)	1196(6)
		0.25	80.2(2.0)	1263(17)	78.8(2.0)	1266(17)	72.8(2.2)	1394(17)	99.8(0.2)	1145(5)
		0.30	88.2(1.6)	1176(14)	88.8(1.6)	1179(14)	89.2(1.5)	1216(14)	100.0(0.0)	1091(5)
	(20, 50)	0.00	5.8(1.2)	1933(9)	5.0(1.1)	1936(9)	0.2(0.2)	1978(1)	51.5(2.5)	1621(17)
		0.10	27.5(2.2)	1771(18)	27.5(2.2)	1771(18)	8.0(1.4)	1929(9)	89.2(1.5)	1276(13)
		0.15	45.5(2.5)	1617(22)	45.8(2.5)	1630(21)	25.2(2.2)	1826(15)	97.8(0.7)	1166(7)
		0.20	67.0(2.4)	1446(22)	65.5(2.4)	1459(22)	53.8(2.5)	1617(20)	99.0(0.5)	1105(6)
		0.25	83.8(1.8)	1287(19)	82.5(1.9)	1288(19)	79.0(2.0)	1379(19)	99.8(0.2)	1061(4)
		0.30	92.0(1.4)	1182(16)	91.5(1.4)	1193(16)	94.0(1.2)	1187(15)	100.0(0.0)	1030(2)
S2	(200, 5)	0.00	5.2(1.1)	1763(8)	6.2(1.2)	1762(8)	0.2(0.2)	1800(0)	51.2(2.5)	1586(11)
		0.10	18.2(1.9)	1692(12)	16.8(1.9)	1699(12)	3.0(0.9)	1788(3)	82.2(1.9)	1406(10)
		0.15	29.0(2.3)	1633(15)	25.2(2.2)	1642(15)	8.5(1.4)	1762(7)	91.8(1.4)	1323(9)
		0.20	40.5(2.5)	1559(17)	42.0(2.5)	1548(17)	17.2(1.9)	1724(10)	97.8(0.7)	1257(7)
		0.25	50.5(2.5)	1489(18)	49.8(2.5)	1492(18)	33.0(2.4)	1641(14)	99.0(0.5)	1218(6)
		0.30	62.5(2.4)	1407(18)	62.5(2.4)	1413(18)	53.5(2.5)	1522(16)	99.5(0.4)	1181(6)
	(20, 50)	0.00	5.8(1.2)	1933(9)	5.0(1.1)	1936(9)	0.2(0.2)	1978(1)	51.5(2.5)	1621(17)
		0.10	19.0(2.0)	1839(16)	19.0(2.0)	1837(16)	3.5(0.9)	1961(5)	81.0(2.0)	1360(15)
		0.15	28.5(2.3)	1763(19)	28.0(2.2)	1771(18)	11.5(1.6)	1911(10)	90.8(1.4)	1256(12)
		0.20	39.0(2.4)	1680(21)	41.8(2.5)	1685(20)	24.0(2.1)	1830(15)	97.5(0.8)	1171(8)
		0.25	50.7(2.5)	1592(22)	52.5(2.5)	1568(22)	44.2(2.5)	1688(19)	99.0(0.5)	1124(6)
		0.30	65.2(2.4)	1479(22)	63.7(2.4)	1481(22)	63.5(2.4)	1539(20)	99.2(0.4)	1092(5)

Table 2: ATE: rejection probabilities (multiplied by 100) and average stopping times under Scenarios 1 and 2 when $\alpha_1(\cdot)$ is chosen as the spending function. Standard errors are reported in the parentheses.

ten days in May 2009. For the ith visit, the dataset contains an ID of the new article recommended to the user, a binary response variable Y_i indicating whether the user clicked the article or not, and a five dimensional feature vector summarizing information of the user. Due to privacy concerns, feature definitions and article names were not included in the data. Each feature vector sums up to 1. Therefore, we took the first three and the fifth elements to form the covariates X_i. For illustration, we only consider a subset of data that contains visits on May 1st where the recommended article ID is either 109510 or 109520. These two articles were being recommended most on that day. This gives us a total of 405888 visits. On the reduced dataset, define $A_i = 1$ if the recommended article is 109510 and $A_i = 0$ otherwise.

We first conduct A/A experiments (which compare these two articles against themselves) to examine the validity of our test. The A/A experiments are done when every 2000 more users are available, we randomly assign 1000 users to arm A, and the other 1000 users in arm B. We expect our test will not reject H_0 in A/A experiments, since the articles being recommended are the same. Then, we conduct A/B experiment to test the QTE of these two articles. The test statistics and their corresponding critical values are plotted in Figure 4. On average it takes several seconds to implement our test. It can be seen that our test is able to be reject H_0 after obtaining the first one-third of the observations, in the A/B experiment. In the A/A experiments, we fail to reject H_0, as expected.

5. Proof of Theorem 1

We present the proof of Theorem 1 in this section. Proofs of other theorems are given in the appendix. We begin with some notations. For any matrix Mat, we use $\|\text{Mat}\|_p$ to denote the matrix norm induced by the corresponding ℓ_p norm of vectors, for $1 \le p \le +\infty$. For two nonnegative sequences $\{s_{1,n}\}_n$ and $\{s_{2,n}\}_n$, we will use the notation $s_{1,n} \preceq s_{2,n}$ to represent that $s_{1,n} \le \bar{c} s_{2,n}$ for some universal constant $\bar{c} > 0$ whose value is allowed to change from place to place. When a matrix Mat is degenerate, Mat^{-1} denotes the Moore-Penrose inverse of Mat. For any vector ψ, we use $\psi^{(i)}$ to denote its i-th element.

Let $n(\cdot)$ be the realization of the counting process $N(\cdot)$. We will show the assertion in Theorem 1 holds for any such realizations that satisfy $n(t_1) < n(t_2) < \cdots < n(t_K)$. The case where some of the $n(t_k)$'s are the same can be similarly discussed.

For any $j \ge 1$, define $\sigma(\mathcal{F}_j)$ to be the σ-algebra generated by $\mathcal{F}_j$. For $a \in \{0,1\}$, define

$$\widehat{\Sigma}_{a,j} = \frac{1}{j}\sum_{i=1}^{j} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i) \quad \text{and} \quad \widehat{\beta}_{a,j} = \widehat{\Sigma}_{a,j}^{-1}\left(\frac{1}{j}\sum_{i=1}^{j} \mathbb{I}(A_i = a)\varphi(X_i)Y_i\right).$$

It is immediate to see that $\widehat{\Sigma}_a(t) = \widehat{\Sigma}_{a,n(t)}$ and $\widehat{\beta}_a(t) = \widehat{\beta}_{a,n(t)}$. Define $\delta_n = qn^{-\alpha_0}\log^{\alpha_0} n$. We state the following lemmas before proving Theorem 1.

Lemma 3 *There exists some constant $0 < \epsilon_0 < 1$ such that $\lambda_{\min}[E\varphi(X)\varphi^\top(X)] \ge \epsilon_0$, $\lambda_{\max}[E\varphi(X)\varphi^\top(X)] \le \epsilon_0^{-1}$, $\sup_x \|\varphi(x)\|_2 \le \sup_x \|\varphi(x)\|_1 \le \epsilon_0^{-1}\sqrt{q}$, $\min_{a\in\{0,1\}} \lambda_{\min}[\Sigma_a] \ge \epsilon_0$, $\max_{a\in\{0,1\}} \|\beta_a\|_2 \le \epsilon_0^{-1}$, and $\sup_x \max_{a\in\{0,1\}} |Q_0(x,a)| \le \epsilon_0^{-1}$.*

Lemma 4 *Assume the conditions in Theorem 1 hold. Then for any sequence $\{j_n\}_n$ that satisfies $j_n^{\alpha_0}/\log^{\alpha_0}(j_n) \gg q^2$, we have with probability at least $1 - O(j_n^{-\alpha_0})$ that for any $a \in \{0,1\}$ and any $k \ge j_n$,*

$$\|(\widehat{\Sigma}_{a,k} - \Sigma_a)\|_2 \preceq q\delta_k + \sqrt{qk^{-1}\log k}, \tag{12}$$

$$\|(\widehat{\Sigma}_{a,k}^{-1} - \Sigma_a^{-1})\|_2 \preceq q\delta_k + \sqrt{qk^{-1}\log k}. \tag{13}$$

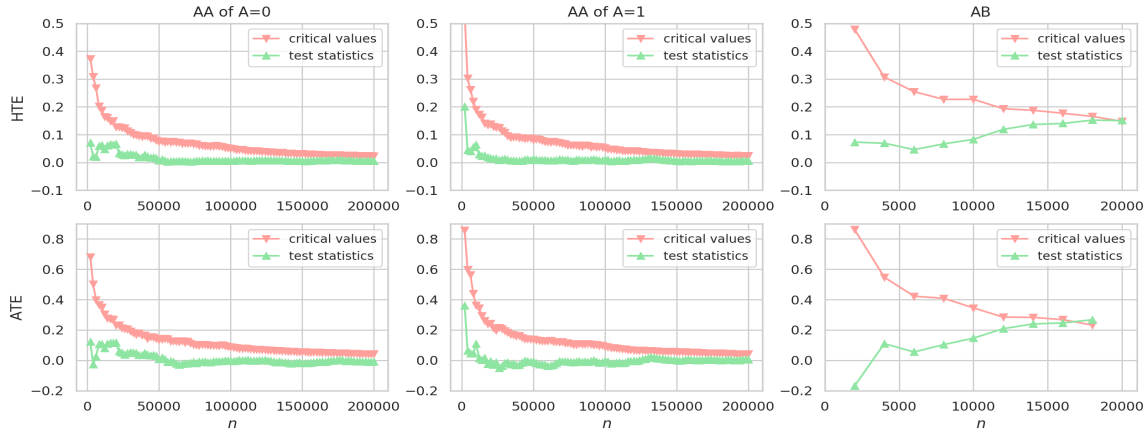

Figure 4: Critical values and test statistics.

Lemma 5 *Assume the conditions in Theorem 1 hold. The for any sequence $\{j_n\}_n$ that satisfies $j_n/\log(j_n) \gg q$, we have with probability at least $1-O(j_n^{-1})$ that for any $a \in \{0,1\}$ and any $k \geq j_n$,*

$$\left\| \sum_{i=1}^{k} \varphi(X_i)\mathbb{I}(A_i = a)\{Y_i - Q_0(X_i, a)\} \right\|_2 \preceq \sqrt{qk \log k}.$$

Lemma 6 *Assume the conditions in Theorem 1 hold. Then for any sequence $\{j_n\}_n$ that satisfies $j_n^{\alpha_0}/\log^{\alpha_0} j_n \gg q^2$, we have with probability at least $1 - O(j_n^{-\alpha_0})$ that*

$$\|\widehat{\beta}_{a,k} - \beta_a\|_2 \preceq q^{1/2}k^{-1/2}\sqrt{\log k}, \quad \forall a \in \{0,1\}, \forall k \geq j_n.$$

Lemma 7 *Assume the conditions in Theorem 1 hold. Then for any sequence $\{j_n\}_n$ that satisfies $j_n^{\alpha_0}/\log^{\alpha_0} j_n \gg q^2$, we have with probability at least $1 - O(j_n^{-\alpha_0})$ that*

$$\left\| \frac{1}{k}\sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\beta_a\}^2 - \Phi_a \right\|_2 \preceq q\delta_k + q^{1/2}k^{-1/2}\sqrt{\log k}, \quad \forall a \in \{0,1\}, k \geq j_n.$$

We now start our proof. We first approximate $\widehat{\beta}_{a,k} - \beta_a$ by a sum of independent mean zero random variables. For $a \in \{0,1\}$,

$$\begin{aligned}\widehat{\beta}_{a,k} - \beta_a = \widehat{\Sigma}_{a,k}^{-1}\left[\frac{1}{k}\sum_{i=1}^{k}\mathbb{I}(A_i = a)\varphi(X_i)\{Y_i - Q_0(X_i,a)\}\right] \\ +\widehat{\Sigma}_{a,k}^{-1}\left[\frac{1}{k}\sum_{i=1}^{k}\mathbb{I}(A_i = a)\varphi(X_i)\{Q_0(X_i,a) - \varphi^\top(X_i)\beta_a\}\right].\end{aligned}$$

Under the given conditions, using similar arguments in proving (E.40) and (E.41) in Shi et al. (2021), we can show that the second term is $O(\text{err})$. It follows that

$$\begin{aligned} & \left\| \widehat{\beta}_{a,k} - \beta_a - \Sigma_a^{-1}\left[\frac{1}{k}\sum_{i=1}^{k}\mathbb{I}(A_i = a)\varphi(X_i)\{Y_i - Q_0(X_i,a)\}\right] \right\|_2 \\ \leq & \ \|\widehat{\Sigma}_{a,k}^{-1} - \Sigma_a^{-1}\|_2 \left\| \frac{1}{k}\sum_{i=1}^{k}\mathbb{I}(A_i = a)\varphi(X_i)\{Y_i - Q_0(X_i,a)\} \right\|_2 + O(\text{err}) \\ \preceq & \ (q\delta_{n(t_k)} + \sqrt{qk^{-1}\log k})q^{1/2}k^{-1/2}\log^{1/2} k + \text{err}, \quad \forall k \geq j_n, \end{aligned} \tag{14}$$

with probability at least $1 - O(j_n^{-\alpha_0})$, by Lemma 4 and Lemma 5. Define

$$B^*(t) = \frac{1}{\sqrt{n(t)}}\sum_{i=1}^{n(t)}[\Sigma_1^{-1}\varphi(X_i)A_i\{Y_i - Q_0(X_i,1)\} - \Sigma_0^{-1}\varphi(X_i)(1-A_i)\{Y_i - Q_0(X_i,0)\}].$$

It follows that

$$\begin{aligned}\|B^*(t_k) - B(t_k)\|_2 \preceq \{q^{3/2}\delta_{n(t_k)} + q\sqrt{n^{-1}(t_k)\log n(t_k)}\}\log^{1/2} n(t_k) \\ +\sqrt{n(t_K)}\text{err}, \quad \forall k \geq 1,\end{aligned} \tag{15}$$

with probability at least $1 - O(n^{-\alpha_0}(t_1))$. By Lemmas 3, 4 and 7, the denominator in $S(t)$ is of the same order of magnitude as $O(\|\varphi(x)\|_2)$. It follows that

$$\left\| \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\widehat{\Sigma}_a^{-1}(t_k)\widehat{\Phi}_a(t_k)\widehat{\Sigma}_a^{-1}(t_k)}} - \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)B(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\widehat{\Sigma}_a^{-1}(t_k)\widehat{\Phi}_a(t_k)\widehat{\Sigma}_a^{-1}(t_k)}} \right\|_2$$
$$\leq \bar{c}\{q^{3/2}\delta_{n(t_k)} + q\sqrt{n^{-1}(t_k)\log n(t_k)}\}\sqrt{\log n(t_k)} + \bar{c}\sqrt{n(t_K)}\text{err}, \quad \forall k \geq 1,$$

with probability at least $1 - O(n^{-\alpha_0}(t_1))$, for some constant $\bar{c} > 0$, by (43).

Define $S^*(t)$ to be a version of our test with $B(t)$ replaced by $B^*(t)$. For any given $z_1, z_2, \ldots, z_K$, we obtain

$$\begin{aligned} & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(S^*(t_k) - z^0_{k,-}\right) \leq 0 \right\} - O(n^{-\alpha_0}(t_1)) \\ \leq\ & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(S(t_k) - z_k\right) \leq 0 \right\} \\ \leq\ & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(S^*(t_k) - z^0_{k,+}\right) \leq 0 \right\} + O(n^{-\alpha_0}(t_1)), \end{aligned} \tag{16}$$

where

$$\begin{aligned} z^0_{k,-} &= z_k - \bar{c}\{q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k) + \sqrt{n(t_K)}\text{err}\}/2, \\ z^0_{k,+} &= z_k + \bar{c}\{q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k) + \sqrt{n(t_K)}\text{err}\}/2. \end{aligned}$$

This completes the first step of the proof.

In the next step, we focus on bounding the difference between $S^*(t)$ and $S^{**}(t)$, the latter being a version of $S^*(t)$ with the denominator replaced with the oracle value

$$\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}.$$

Similarly, under Lemmas 4, 5 and 7, we can show that difference $|S^*(t_k) - S^{**}(t_k)|$ can be upper bounded by $O(q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k))$, uniformly for any k. Combine this together with (16) yields

$$\begin{aligned} & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(S^{**}(t_k) - z_{k,-}\right) \leq 0 \right\} - O(n^{-\alpha_0}(t_1)) \\ \leq\ & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(S(t_k) - z_k\right) \leq 0 \right\} \\ \leq\ & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(S^{**}(t_k) - z_{k,+}\right) \leq 0 \right\} + O(n^{-\alpha_0}(t_1)), \end{aligned} \tag{17}$$

where

$$\begin{aligned} z_{k,-} &= z_k - \bar{c}\{q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k) + \sqrt{n(t_K)}\text{err}\}, \\ z_{k,+} &= z_k + \bar{c}\{q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k) + \sqrt{n(t_K)}\text{err}\}. \end{aligned}$$

This completes the second step.

In the last step, we aim to apply the high-dimensional Gaussian approximation technique developed by Belloni and Oliveira (2018) to approximate $S^{**}(t)$ by its Gaussian analogue. For any $i \geq 1$, $1 \leq k \leq K$, define a q-dimensional vector

$$\xi_{i,k} = \frac{1}{\sqrt{n(t_k)}}[\Sigma_1^{-1}\varphi(X_i)A_i\{Y_i - Q_0(X_i,1)\} - \Sigma_0^{-1}\varphi(X_i)(1-A_i)\{Y_i - Q_0(X_i,0)\}]\mathbb{I}(i \leq n(t_k)),$$

or equivalently,

$$\xi_{i,k} = \frac{1}{\sqrt{n(t_k)}}[\Sigma_1^{-1}\varphi(X_i)A_i\{Y_i^*(1) - Q_0(X_i,1)\} - \Sigma_0^{-1}\varphi(X_i)(1-A_i)\{Y_i^*(0) - Q_0(X_i,0)\}]\mathbb{I}(i \leq n(t_k)),$$

by Condition (A1). Let $\boldsymbol{\xi}_i = (\xi_{i,1}^\top, \xi_{i,2}^\top, \cdots, \xi_{i,K}^\top)^\top$ and $\mathcal{M}_j = \sum_{i=1}^j \boldsymbol{\xi}_i$. The sequence $\{\mathcal{M}_i\}_{i\geq 1}$ forms a multivariate martingale with respect to the filtration $\{\sigma(\mathcal{F}_i) : i \geq 1\}$, since

$$\mathrm{E}(\xi_{i,k}|\mathcal{F}_i) = [\{\mathrm{E}(\xi_{i,k}|A_i, X_i, \mathcal{F}_i)\}|\mathcal{F}_i] = 0,$$

by (A2). Let $n(t_0) = 0$. For any i such that $n(t_{k-1}) < i \leq n(t_k)$ for some $1 \leq k \leq K$, we have

$$\begin{aligned}\|\boldsymbol{\xi}_i\|_\infty &\leq \frac{1}{\sqrt{n(t_k)}}\{\|\Sigma_1^{-1}\varphi(X_i)\{Y_i^*(1) - Q_0(X_i,1)\beta_1\}\|_2 + \|\Sigma_0^{-1}\varphi(X_i)\{Y_i^*(0) - Q_0(X_i,0)\}\|_2\} \\ &\leq \sqrt{q}n^{-1/2}(t_k)\epsilon_0^{-2}(2\epsilon_0^{-1} + |Y_i^*(0)| + |Y_i^*(1)|),\end{aligned}$$

where the second inequality is due to Lemma 3. Under the sub-Gaussianity assumption, $Y^*(0)$ and $Y^*(1)$ have moments of all orders. Therefore,

$$\mathrm{E}\|\boldsymbol{\xi}_i\|_\infty^3 \preceq \frac{q^{3/2}}{n^{3/2}(t_k)}.$$

It follows that

$$\begin{aligned}&\sum_{i=1}^{n(t_K)} \mathrm{E}\|\boldsymbol{\xi}_i\|_\infty^3 = \sum_{k=1}^K \sum_{i=n(t_{k-1})+1}^{n(t_k)} \mathrm{E}\|\boldsymbol{\xi}_i\|_\infty^3 \preceq q^{3/2}\sum_{k=1}^K \frac{n(t_k) - n(t_{k-1})}{n^{3/2}(t_k)} \qquad (18) \\ \leq\ & \frac{q^{3/2}}{\sqrt{n(t_1)}} + q^{3/2}\sum_{k=2}^K \frac{n(t_k) - n(t_{k-1})}{n^{3/2}(t_k)} \leq q^{3/2}n^{-1/2}(t_1) + q^{3/2}\int_{n(t_1)}^{+\infty} x^{-3/2}dx = 3q^{3/2}n^{-1/2}(t_1).\end{aligned}$$

Define a sequence of independent Gaussian vectors $\{\boldsymbol{\eta}_i\}_{i\geq 1}$ that satisfy $\boldsymbol{\eta}_i \sim N(0, \mathrm{E}(\boldsymbol{\xi}_i\boldsymbol{\xi}_i^\top|\mathcal{F}_{i-1}))$ for any $i \geq 1$. Then the distribution of $\boldsymbol{\eta}_i$ is the same as

$$\left(\frac{\mathbb{I}(i \leq n(t_1))}{\sqrt{n(t_1)}}Z^\top, \frac{\mathbb{I}(i \leq n(t_2))}{\sqrt{n(t_2)}}Z^\top, \cdots, \frac{\mathbb{I}(i \leq n(t_K))}{\sqrt{n(t_K)}}Z^\top\right),$$

where Z is a p-dimensional mean-zero Gaussian vector with covariance matrix

$$\begin{aligned}
&\mathrm{cov}[\sum_{a\in\{0,1\}} \Sigma_a^{-1}\varphi(X_i)\mathbb{I}(A_i=a)\{Y_i^*(a)-Q_0(X_i,a)\}|\mathcal{F}_{i-1}] \\
&= \sum_{a\in\{0,1\}} \Sigma_a^{-1}\mathrm{E}[\varphi(X_i)\varphi^\top(X_i)\mathbb{I}(A_i=a)\{Y_i^*(a)-Q_0(X_i,a)\}^2|\mathcal{F}_{i-1}]\Sigma_a^{-1} \\
&= \sum_{a\in\{0,1\}} \Sigma_a^{-1}\mathrm{E}\{\varphi(X_i)\varphi^\top(X_i)\mathbb{I}(A_i=a)\sigma^2(a,X_i)|\mathcal{F}_{i-1}\}\Sigma_a^{-1} \\
&= \sum_{a\in\{0,1\}} \Sigma_a^{-1}\mathrm{E}\{\varphi(X_i)\varphi^\top(X_i)\pi_{i-1}(a,X_i)\sigma^2(a,X_i)|\mathcal{F}_{i-1}\}\Sigma_a^{-1} \\
&\equiv \sum_{a\in\{0,1\}} \Sigma_a^{-1}\mathrm{E}^{\mathcal{F}_{i-1}}\pi_{i-1}(a,X)\sigma^2(a,X)\varphi(X)\varphi^\top(X)\Sigma_a^{-1},
\end{aligned} \tag{19}$$

where the second equality follows from (A2) and Lemma 8, the third equality is due to the definition of π_{i-1} and the last equality follows from Lemma 8. See the proof of Lemma 1 for details.

Similar to (18), we can show that

$$\sum_{i=1}^{n(t_K)} \mathrm{E}\|\boldsymbol{\eta}_i\|_\infty^3 \preceq q^{3/2}n^{-1/2}(t_1). \tag{20}$$

Using similar arguments in (19), we can show that for any $1\le k_1\le k_2\le K$,

$$\sum_{i=1}^{n(t_K)} \mathrm{E}\{\xi_{i,k_1}\xi_{i,k_2}^\top|\mathcal{F}_{i-1}\} = \frac{1}{\sqrt{n(t_{k_1})n(t_{k_2})}}\sum_{i=1}^{n(t_{k_1})}\sum_{a\in\{0,1\}} \Sigma_a^{-1}\mathrm{E}^{\mathcal{F}_{i-1}}\pi_{i-1}(a,X)\sigma^2(a,X)\varphi(X)\varphi^\top(X)\Sigma_a^{-1}.$$

Let

$$\begin{aligned}
V(k_1,k_2) &= \frac{1}{\sqrt{n(t_{k_1})n(t_{k_2})}}\sum_{i=1}^{n(t_{k_1})}\sum_{a\in\{0,1\}} \Sigma_a^{-1}\mathrm{E}^{\mathcal{F}_{i-1}}\pi^*(a,X)\sigma^2(a,X)\varphi(X)\varphi^\top(X)\Sigma_a^{-1} \\
&= \frac{1}{\sqrt{n(t_{k_1})n(t_{k_2})}}\sum_{i=1}^{n(t_{k_1})}\sum_{a\in\{0,1\}} \Sigma_a^{-1}\Phi_a\Sigma_a^{-1} = \frac{\sqrt{n(t_{k_1})}}{\sqrt{n(t_{k_2})}}\sum_{a\in\{0,1\}} \Sigma_a^{-1}\Phi_a\Sigma_a^{-1}.
\end{aligned}$$

Consider an arbitrary sequence of $\mathbb{R}^{p+1}$ vectors $\{b_k\}_{1\le k\le K}$. Under the given conditions, we have

$$\begin{aligned}
&\left|b_{k_1}^\top\left(\sum_{i=1}^{n(t_K)}\mathrm{E}(\xi_{i,k_1}\xi_{i,k_2}^\top|\mathcal{F}_{i-1}) - V(k_1,k_2)\right)b_{k_2}\right| \\
\preceq\ & \frac{1}{n(t_{k_1})}\sum_{a\in\{0,1\}}\left\|\sum_{i=1}^{n(t_{k_1})}\mathrm{E}^{\mathcal{F}_{i-1}}\{\pi_{i-1}(a,X)-\pi^*(a,X)\}\sigma^2(a,X)\varphi(X)\varphi^\top(X)\right\|_2\|b_{k_1}\|_2\|b_{k_2}\|_2.
\end{aligned}$$

Define a matrix $\boldsymbol{V}$ as

$$\boldsymbol{V} = \begin{pmatrix} V(1,1) & V(1,2) & \dots & V(1,K) \\ V(2,1) & V(2,2) & \dots & V(2,K) \\ \vdots & \vdots & & \vdots \\ V(K,1) & V(K,2) & \dots & V(K,K) \end{pmatrix}. \tag{21}$$

It follows that

$$\left\| \sum_{i=1}^{n(t_K)} \mathrm{E}(\boldsymbol{\xi}_i \boldsymbol{\xi}_i^\top | \mathcal{F}_{i-1}) - \boldsymbol{V} \right\|_2 \preceq \sup_{\substack{a\in\{0,1\} \\ j \geq n(t_1)}} \left\| \frac{1}{j} \sum_{i=1}^{j} \mathrm{E}^{\mathcal{F}_{i-1}} \{\pi_{i-1}(a,X) - \pi^*(a,X)\} \sigma^2(a,X) \varphi(X) \varphi^\top(X) \right\|_2.$$

Using similar arguments in proving (12), we can show the RHS of the above equation is upper bounded by

$$\epsilon_0^{-2} q \sup_{\substack{a\in\{0,1\} \\ x\in\mathbb{X}, j \geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right|,$$

and hence by $\epsilon_0^{-2} q \delta_{n(t_1)}$, with probability at least $1 - O(n^{-\alpha_0}(t_1))$. Therefore, we have

$$\lambda_{\min} \left[\boldsymbol{V} + \delta_{n(t_1)} \boldsymbol{I}_{Kp \times Kp} - \sum_{i=1}^{n(t_K)} \mathrm{E}(\boldsymbol{\xi}_i \boldsymbol{\xi}_i^\top | \mathcal{F}_{i-1}) \right] \geq 0, \tag{22}$$

with probability at least $1 - O(n^{-\alpha_0}(t_1))$, where $\boldsymbol{I}_{Kp\times Kp}$ denotes a $Kp \times Kp$ identity matrix.

Moreover, notice that

$$\sup_{\substack{a\in\{0,1\} \\ x\in\mathbb{X}, j \geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right|$$

is bounded between 0 and 1. For any $a \in \{0,1\}$ and any $z > 0$, we have

$$\begin{aligned} & \mathrm{E} \sup_{\substack{a\in\{0,1\} \\ x\in\mathbb{X}, j \geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right| \\ \leq\ & \mathrm{E} \sup_{\substack{a\in\{0,1\} \\ x\in\mathbb{X}, j \geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right| \mathbb{I} \left(\sup_{\substack{a\in\{0,1\} \\ x\in\mathbb{X}, j \geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right| \leq z \right) \\ +\ & \Pr \left(\sup_{\substack{a\in\{0,1\} \\ x\in\mathbb{X}, j \geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right| > z \right). \end{aligned}$$

Under the given conditions, we have

$$\mathrm{E} \sup_{\substack{a\in\{0,1\}\\ x\in\mathbb{X}, j\geq n(t_1)}} \left| \frac{1}{j} \sum_{i=1}^{j} \{\pi_{i-1}(a,x) - \pi^*(a,x)\} \right| \preceq \delta_{n(t_1)} + O(n^{-\alpha_0}(t_1)).$$

Therefore, we obtain

$$\mathrm{E} \left\| \sum_{i=1}^{n(t_K)} \mathrm{E}(\boldsymbol{\xi}_i \boldsymbol{\xi}_i^\top | \mathcal{F}_{i-1}) - \boldsymbol{V} \right\|_2 \preceq q n^{-\alpha_0}(t_1) + q\delta_{n(t_1)},$$

or

$$\mathrm{E} \left\| \sum_{i=1}^{n(t_K)} \mathrm{E}(\boldsymbol{\xi}_i \boldsymbol{\xi}_i^\top | \mathcal{F}_{i-1}) - \boldsymbol{V} \right\|_2 \preceq q\delta_{n(t_1)}, \tag{23}$$

since $n^{-\alpha_0}(t_1) \ll \delta_{n(t_1)}$. Combining (18) with (20), (22) and (23), an application of Theorem 2.1 in Belloni and Oliveira (2018) yields that

$$\begin{aligned} & |\mathrm{E}\psi(\mathcal{M}_{n(t_K)}) - \mathrm{E}\psi(N(0,\boldsymbol{V}))| \\ \preceq \;& c_0(\psi) n^{-\alpha_0}(t_1) + c_2(\psi) q \delta_{n(t_1)} + c_3(\psi) q^{3/2} n^{-1/2}(t_1), \end{aligned} \tag{24}$$

for any thrice differential function $\psi(\cdot)$, and

$$c_0(\psi) = \sup_{z,z'\in\mathbb{R}^{pK}} |\psi(z) - \psi(z')| \text{ and } c_i = \sup_{z\in\mathbb{R}^{pK}} \sum_{j_1,\cdots,j_i} |\partial_{j_1}\partial_{j_2}\cdots\partial_{j_i}\psi(z)|, i = 2, 3,$$

where $\partial_j g(z)$ denotes the partial derivative $\partial g(z)/\partial z^{(j)}$ for any function g$(\cdot)$ and $z^{(j)}$ stands for the j-th element of z.

Let $\mathbb{X}_{k,0}$ be an ε-net of $\mathbb{X}$ that satisfies the following: for any $x \in \mathbb{X}$, there exists some $x_0 \in \mathbb{X}_0$ such that $\|x - x_0\|_2 \leq \varepsilon$. Note that we require $\mathbb{X}$ to be a compact set. To simplify the proof, we assume $\mathbb{X} = [0,1]^d$. In cases where $\mathbb{X} \neq [0,1]^d$, we could conduct the min-max normalization to rescale the range of features to $[0,1]$. Set $\varepsilon = \sqrt{d}/n^4(t_1)$. There exists some $\mathbb{X}_0$ with

$$|\mathbb{X}_0| \leq n^{4d}(t_1), \tag{25}$$

where $|\mathbb{X}_0|$ denotes the number of elements in $\mathbb{X}_0$. Under Condition (A3), we have

$$\sup_{x\in\mathbb{X}} \inf_{x_0\in\mathbb{X}_0} \|\varphi(x) - \varphi(x_0)\|_2 \preceq \frac{\sqrt{q}}{n^4(t_1)}.$$

It follows that

$$\sup_{\|\nu\|_2=1} |\sup_{x\in\mathbb{X}} \varphi^\top(x)\nu - \sup_{x\in\mathbb{X}_0} \varphi^\top(x)\nu| \preceq \frac{\sqrt{q}}{n^4(t_1)}. \tag{26}$$

Using similar arguments in showing (15), we can show the following event occurs with probability at least $1 - O(n^{-1}(t_1))$,

$$\|B^*(t_k)\|_2 \preceq q^{1/2} \log^{1/2} n(t_k), \quad \forall k \geq 1.$$

This together with (26) and the fact that the denominator $\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}/\|\varphi(x)\|_2$ is uniformly bounded away from zero yields

$$\begin{aligned}
&\left| \max_{k\in\{1,\ldots,K\}} \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - \max_{k\in\{1,\ldots,K\}} \sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} \right| \\
&\quad \leq \max_{k\in\{1,\ldots,K\}} \left| \sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - \sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} \right| \\
&\quad \preceq \frac{\sqrt{q\log n(t_K)}}{n^4(t_1)},
\end{aligned}$$

with probability at least $1 - O(n^{-1}(t_1))$. Under the given conditions, we have $n(t_1) \gg \max(q, \log n(t_K))$. It follows that there exists some constant $\bar{c}^* > 0$ such that

$$\left| \max_{k\in\{1,\ldots,K\}} \sup_{x\in\mathbb{X}} \varphi^\top(x)B^*(t_k) - \max_{k\in\{1,\ldots,K\}} \sup_{x\in\mathbb{X}_0} \varphi^\top(x)B^*(t_k) \right| \leq \bar{c}^* n^{-2}(t_1), \tag{27}$$

with probability at least $1 - O(n^{-1}(t_1))$.

Define

$$\begin{aligned}
z^*_{k,-} &= z_k - \bar{c}\{q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k) + \sqrt{n(t_K)}\mathrm{err}\} - \bar{c}^* n^{-2}(t_1), \\
z^*_{k,+} &= z_k + \bar{c}\{q^{3/2}\delta_{n(t_k)}\sqrt{\log n(t_k)} + q\sqrt{n^{-1}(t_k)}\log n(t_k) + \sqrt{n(t_K)}\mathrm{err}\} + \bar{c}^* n^{-2}(t_1).
\end{aligned}$$

Combining (27) with (17) yields

$$\begin{aligned}
&\Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z^*_{k,-} \right) \leq 0 \right\} - O(n^{-\alpha_0}(t_1)) \\
&\quad \leq \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}} \frac{\varphi^\top(x)B(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\widehat{\Sigma}_a^{-1}(t_k)\widehat{\Phi}_a(t_k)\widehat{\Sigma}_a^{-1}(t_k)\varphi(x)}} - z_k \right) \leq 0 \right\} \qquad (28) \\
&\leq \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z^*_{k,+} \right) \leq 0 \right\} + O(n^{-\alpha_0}(t_1)).
\end{aligned}$$

Notice that $\mathcal{M}_{n(t_K)} = \{B^*(t_1)^\top, B^*(t_2)^\top, \cdots, B^*(t_K)^\top\}^\top$. By (25) and the fact that the denominator $\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)} \geq \bar{c}\|\varphi(x)\|_2$ for some constant $\bar{c} > 0$, there

exist a set of vectors $d_1, d_2, \ldots, d_L \in \mathbb{R}^{qK}$ with $L \leq n^{4d}(t_1)K$, $\max_j \|d_j\|_1 \leq \epsilon^{-1}$ for some $0 < \epsilon < 1$ and a function $k(\cdot)$ that maps $\{1, \ldots, L\}$ into $\{1, \ldots, K\}$ such that

$$\max_{k\in\{1,\ldots,K\}} \left\{ \sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - \nu_k \right\} = \max_{1\leq j\leq L} \{d_j^\top \mathcal{M}_{n(t_K)} - \nu_{k(j)}\}, \quad (29)$$

for any $\{\nu_k\}_{k=1}^K$. For any $\eta > 0$, $m \in \mathbb{R}^{qK}$, consider the function $\phi_{\eta,\{\nu_k\}_k} : \mathbb{R}^{qK} \to \mathbb{R}$, defined as

$$\phi_{\eta,\{\nu_k\}_k}(m) = \frac{1}{\eta} \log \left\{ \sum_{j=1}^{L} \exp[\eta\{d_j^\top m - \eta\nu_{k(j)}\}] \right\}.$$

It has the following property:

$$\begin{aligned} & \max_{1\leq j\leq L} \{d_j^\top m - \nu_{k(j)}\} \leq \phi_{\eta,\{\nu_k\}_k}(m) \leq \max_{1\leq j\leq L} \{d_j^\top m - \nu_{k(j)}\} + \eta^{-1}\log L \\ \leq & \max_{1\leq j\leq L} \{d_j^\top m - \nu_{k(j)}\} + \eta^{-1}\{\log K + 4d\log n(t_1)\} \\ = & \max_{1\leq j\leq L} [d_j^\top m - \{\nu_{k(j)} - \eta^{-1}\log K - \eta^{-1}4d\log n(t_1)\}]. \end{aligned}$$

It follows that

$$\begin{aligned} & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z_{k,+}^* \right) \leq 0 \right\} \quad (30) \\ \leq & \Pr\left\{ \phi_{\eta,\{z_{k,+}^{**}\}_k}(\mathcal{M}_{n(t_K)}) \leq 0 \right\}, \end{aligned}$$

and that

$$\begin{aligned} & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - z_{k,-}^* \right) \leq 0 \right\} \quad (31) \\ = & \Pr\left\{ \max_{k\in\{1,\ldots,K\}} \left(\sup_{x\in\mathbb{X}_0} \frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}} - (z_{k,-}^* - 3\delta) \right) \leq 3\delta \right\} \\ \geq & \Pr\left\{ \phi_{\eta,\{z_{k,-}^{**}\}_k}(\mathcal{M}_{n(t_K)}) \leq 3\delta \right\}, \end{aligned}$$

where

$$z_{k,+}^{**} = z_{k,+}^* + \eta^{-1}\{\log K + 4d\log n(t_1)\} \text{ and } z_{k,-}^{**} = z_{k,-}^* - 3\delta.$$

The value of δ will be specified later. In addition, with some calculations, we have

$$\begin{aligned}
\partial_j \phi_{\eta,\{\nu_k\}_k}(m) &= \frac{\sum_{i=1}^L d_i^{(j)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)}{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)}, \\
\partial_{j_1}\partial_{j_2} \phi_{\eta,\{\nu_k\}_k}(m) &= \eta \frac{\sum_{i=1}^L d_i^{(j_1)} d_i^{(j_2)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)}{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)} \\
&- \eta \frac{\prod_{l=1,2}\left\{\sum_{i=1}^L d_i^{(j_l)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}}{\left\{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}^2},
\end{aligned}$$

$$\begin{aligned}
\partial_{j_1}\partial_{j_2}\partial_{j_3} \phi_{\eta,\{\nu_k\}_k}(m) &= \eta^2 \frac{\sum_{i=1}^L d_i^{(j_1)} d_i^{(j_2)} d_i^{(j_3)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)}{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)} \\
&- 3\eta^2 \frac{\left\{\sum_{i=1}^L d_i^{(j_1)} d_i^{(j_2)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}}{\left\{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}} \\
&\times \frac{\left\{\sum_{i=1}^L d_i^{(j_3)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}}{\left\{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}} \\
&+ 2\eta^2 \frac{\prod_{l=1,2,3}\left(\sum_{i=1}^L d_i^{(j_l)} \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right)}{\left\{\sum_{i=1}^L \exp\left(\eta[d_i^\top m - \nu_{k(i)}]\right)\right\}^3}.
\end{aligned}$$

Since $\max_i \|d_i\|_1 \le \epsilon^{-1}$, we obtain that

$$\begin{aligned}
&\sum_j |\partial_j \phi_{\eta,\{\nu_k\}_k}(m)| \le \epsilon^{-1}, \quad \sum_{j_1,j_2} |\partial_{j_1}\partial_{j_2} \phi_{\eta,\{\nu_k\}_k}(m)| \le 2\eta\epsilon^{-2}, \\
&\qquad \sum_{j_1,j_2,j_3} |\partial_{j_1}\partial_{j_2}\partial_{j_3} \phi_{\eta,\{\nu_k\}_k}(m)| \le 6\eta^2\epsilon^{-3}.
\end{aligned} \tag{32}$$

By Lemma 5.1 of Chernozhukov et al. (2016), for any $\delta > 0$, there exists some function $\mathrm{g}_\delta(\cdot) : \mathbb{R} \to \mathbb{R}$ with $\|\mathrm{g}_\delta'\|_\infty \le \delta^{-1}$, $\|\mathrm{g}_\delta''\|_\infty \le K_0\delta^{-2}$, $\|\mathrm{g}_\delta'''\|_\infty \le K_0\delta^{-3}$ for some constant $K_0 > 0$ such that

$$\mathbb{I}(z_0 \le 0) \le \mathrm{g}_\delta(z_0) \le \mathbb{I}(z_0 \le 3\delta), \ \forall \delta \in \mathbb{R}.$$

It follows that

$$\mathbb{I}(\phi_{\eta,\{\nu_k\}_k}(m) \le 0) \le \mathrm{g} \circ \phi_{\eta,\{\nu_k\}_k}(m) \le \mathbb{I}(\phi_{\eta,\{\nu_k\}_k}(m) \le 3\delta),$$

for any $m \in \mathbb{R}^{qK}$. Combining this together with (29), (30) and (31), we obtain that

$$\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}}\varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}}-z^*_{k,+}\right)\le 0\right\}\le \mathrm{Eg}_\delta\circ\phi_{\eta,\{z^{**}_{k,+}\}_k}(\mathcal{M}_{n(t_K)}), \quad (33)$$

$$\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}}\varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}}-z^*_{k,-}\right)\le 0\right\}\ge \mathrm{Eg}_\delta\circ\phi_{\eta,\{z^{**}_{k,-}\}_k}(\mathcal{M}_{n(t_K)}). \quad (34)$$

Consider the function $\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k}$. Apparently, we have

$$\sup_{\delta,\eta,\{\nu_k\}_k} c_0(\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k})\le 1. \quad (35)$$

By (32), we can show that

$$\begin{aligned}\sup_{\delta,\eta,\{\nu_k\}_k} c_2(\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k})&\preceq \delta^{-2}+\delta^{-1}\eta,\\ \sup_{\delta,\eta,\{\nu_k\}_k} c_3(\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k})&\preceq \delta^{-3}+\delta^{-2}\eta+\delta^{-1}\eta^2.\end{aligned} \quad (36)$$

Set $\delta=\eta^{-1}\{\log K+4d\log n(t_1)\}$, we obtain

$$\sup_{\eta,\{\nu_k\}_k} c_i(\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k})\preceq \eta^i\{\log^i K+\log^i n(t_1)\},\quad i=2,3.$$

Combining (36) together with (24) and (35) yields

$$\begin{aligned}&\sup_{\delta,\eta,\{\nu_k\}_k}\left|\mathrm{Eg}_\delta\circ\phi_{\eta,\{\nu_k\}_k}(\mathcal{M}_{n(t_K)})-\mathrm{Eg}_\delta\circ\phi_{\eta,\{\nu_k\}_k}(N(0,\boldsymbol{V}))\right|\\ \preceq\ & n^{-1/2}(t_1)\eta^3\{\log^3 K+\log^3 n(t_1)\}+\eta^2\{\log^2 K+\log^2 n(t_1)\}\delta_{n(t_1)}+n^{-\alpha_0}(t_1).\end{aligned}$$

This together with (33) and (34) yields

$$\begin{aligned}&\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}}\varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}}-z^*_{k,+}\right)\le 0\right\}-\mathrm{Eg}_\delta\circ\phi_{\eta,\{z^{**}_{k,+}\}_k}(N(0,\boldsymbol{V})) \quad (37)\\ &\preceq n^{-1/2}(t_1)\eta^3\{\log^3 K+\log^3 n(t_1)\}+\eta^2\{\log^2 K+\log^2 n(t_1)\}\delta_{n(t_1)}+n^{-\alpha_0}(t_1),\\ &\mathrm{Eg}_\delta\circ\phi_{\eta,\{z^{**}_{k,-}\}_k}(N(0,\boldsymbol{V}))-\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)B^*(t_k)}{\sqrt{\sum_{a\in\{0,1\}}\varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)}}-z^*_{k,-}\right)\le 0\right\} \quad (38)\\ &\preceq n^{-1/2}(t_1)\eta^3\{\log^3 K+\log^3 n(t_1)\}+\eta^2\{\log^2 K+\log^2 n(t_1)\}\delta_{n(t_1)}+n^{-\alpha_0}(t_1).\end{aligned}$$

Similar to (30)-(34), we can show

$$\begin{aligned}&\mathrm{Eg}_\delta\circ\phi_{\eta,\{z^{**}_{k,+}\}_k}(N(0,\boldsymbol{V}))\le\Pr\left(\phi_{\eta,\{z^{**}_{k,+}\}_k}(N(0,\boldsymbol{V}))\le 3\delta\right)\\ \le\ &\Pr\left(\max_{1\le j\le L}\{d_j^\top N(0,\boldsymbol{V})-z^{**}_{k(j),+}\}\le 3\delta\right)=\Pr\left(\max_{1\le j\le L}\{d_j^\top N(0,\boldsymbol{V})-z^{***}_{k(j),+}\}\le 0\right),\\ &\mathrm{Eg}_\delta\circ\phi_{\eta,\{z^{**}_{k,-}\}_k}(N(0,\boldsymbol{V}))\ge\Pr\left(\phi_{\eta,\{z^{**}_{k,-}\}_k}(N(0,\boldsymbol{V}))\le 0\right)\\ \ge\ &\Pr\left(\max_{1\le j\le L}\{d_j^\top N(0,\boldsymbol{V})-z^{***}_{k(j),-}\}\le 0\right),\end{aligned}$$

where

$$z^{***}_{k,+} = z^{*}_{k,+} + \eta^{-1}\{\log K + 4d\log n(t_1)\} + 3\delta \text{ and } z^{***}_{k,-} = z^{*}_{k,-} - \eta^{-1}\{\log K + 4d\log n(t_1)\} - 3\delta,$$

for each k. Let $\sigma(x) = \sum_{a\in\{0,1\}} \varphi^\top(x)\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\varphi(x)$ and $\widehat{\sigma}(x,t) = \sum_{a\in\{0,1\}} \varphi^\top(x)\widehat{\Sigma}_a^{-1}(t)\widehat{\Phi}_a(t)\widehat{\Sigma}_a^{-1}(t)\varphi(x)$. Notice that for any $\{\nu_k\}_k$, we have

$$\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\sigma^{-1}(x)\varphi^\top(x)G(t_k) - \nu_k\right) \le 0\right\} = \Pr\left(\max_{1\le j\le L}\{d_j^\top N(0,\boldsymbol{V}) - \nu_{k(j)}\} \le 0\right).$$

This together with (37) and (38) yields

$$\begin{aligned}
&\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)B^*(t_k)}{\sigma(x)} - z^{*}_{k,+}\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x)} - z^{***}_{k,+}\right) \le 0\right\}\\
&\qquad \preceq n^{-1/2}(t_1)\eta^3\{\log^3 K + \log^3 n(t_1)\} + \eta^2\{\log^2 K + \log^2 n(t_1)\}\delta_{n(t_1)} + n^{-\alpha_0}(t_1),\\
&\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x)} - z^{***}_{k,-}\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)B^*(t_k)}{\sigma(x)} - z^{*}_{k,-}\right) \le 0\right\}\\
&\qquad \preceq n^{-1/2}(t_1)\eta^3\{\log^3 K + \log^3 n(t_1)\} + \eta^2\{\log^2 K + \log^2 n(t_1)\}\delta_{n(t_1)} + n^{-\alpha_0}(t_1).
\end{aligned}$$

In view of (28), we have shown that

$$\begin{aligned}
&\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)B(t_k)}{\sigma(x)} - z_k\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x)} - z^{***}_{k,+}\right) \le 0\right\}\\
&\qquad \preceq n^{-1/2}(t_1)\eta^3\{\log^3 K + \log^3 n(t_1)\} + \eta^2\{\log^2 K + \log^2 n(t_1)\}\delta_{n(t_1)} + n^{-\alpha_0}(t_1),\\
&\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x)} - z^{***}_{k,-}\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)B(t_k)}{\sigma(x)} - z_k\right) \le 0\right\}\\
&\qquad \preceq n^{-1/2}(t_1)\eta^3\{\log^3 K + \log^3 n(t_1)\} + \eta^2\{\log^2 K + \log^2 n(t_1)\}\delta_{n(t_1)} + n^{-\alpha_0}(t_1).
\end{aligned}$$

By Theorem 1 of Chernozhukov et al. (2017), we obtain that

$$\begin{aligned}
&\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x)} - z^{***}_{k,+}\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x)} - z^{***}_{k,-}\right) \le 0\right\}\\
\preceq\ & \eta^{-1}\{\log K + \log n(t_1)\}^{3/2} + q^{3/2}\delta_{n(t_1)}\{\log K + \log n(t_1)\} + q\sqrt{n^{-1}(t_1)}\{\log K + \log n(t_1)\}^{3/2}\\
+\ & \sqrt{n(t_K)}\mathrm{err}\{\log K + \log n(t_1)\}^{1/2}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
&\left|\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\varphi^\top(x)B(t_k) - z_k\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\varphi^\top(x)G(t_k) - z_k\right) \le 0\right\}\right|\\
\preceq\ & n^{-1/2}(t_1)\eta^3\{\log^3 K + \log^3 n(t_1)\} + \eta^2\{\log^2 K + \log^2 n(t_1)\}\delta_{n(t_1)} + n^{-\alpha_0}(t_1)\\
+\ & \eta^{-1}\{\log K + \log n(t_1)\}^{3/2} + q^{3/2}\delta_{n(t_1)}\{\log K + \log n(t_1)\} + q\sqrt{n^{-1}(t_1)}\log n(t_1)\{\log K + \log n(t_1)\}^{3/2}\\
+\ & \sqrt{n(t_K)}\mathrm{err}\{\log K + \log n(t_1)\}^{1/2}.
\end{aligned}$$

Setting $\eta = \min(n^{1/8}(t_1)\log^{-3/8}\{Kn(t_1)\}, n^{-\alpha_0/3}(t_1)\log^{-\alpha_0/3-1/6}\{Kn(t_1)\})$ yields the desired results. The proof is hence completed.

6. Discussion

6.1 Growing number of basis functions

In the current proposal, we use the same set of basis functions at each interim stage. It would be more desirable to allow the number of basis functions q to increase with k to deal with the model approximation error.

In this section, we extend our proposal to allow q to vary with k. We focus on the case where the maximum number of interim stages K is known and the basis functions to be used at each interim stage are predetermined. Let $\{\varphi_k\}_k$ denote the sets of basis functions and q_k the dimension of φ_k for any k. Let $\widehat{\Sigma}_a^{(k)}(t) = N^{-1}(t)\sum_{i=1}^{N(t)} \mathbb{I}(A_i = a)\varphi_k(X_i)\varphi_k^\top(X_i)$. To implement the resulting test, we need to compute the regression coefficients

$$\widehat{\beta}_a^{(k)}(t_k) = \{\widehat{\Sigma}_a^{(k)}(t_k)\}^{-1}\left\{\frac{1}{N(t_k)}\sum_{i=1}^{N(t_k)} \mathbb{I}(A_i = a)\varphi_k(X_i)Y_i\right\},$$

at the kth interim stage. To allow for online updating, at the kth interim stage, we not only compute $\widehat{\Sigma}_a^{(k)}(t_k)$ and $N^{-1}(t_k)\sum_{i=1}^{N(t_k)} \mathbb{I}(A_i = a)\varphi_k(X_i)Y_i$, but $\{\widehat{\Sigma}_a^{(\kappa)}(t_k)\}_{k<\kappa\le K}$ and $\{\gamma_{a,\kappa}(t_k)\}_{k<\kappa\le K}$ as well where $\gamma_{a,\kappa}(t_k) = N(t_k)^{-1}\sum_{i=1}^{N(t_k)} \mathbb{I}(A_i = a)\varphi_\kappa(X_i)Y_i$. These quantities allow us to compute $\widehat{\beta}_a^{(\kappa)}(t_\kappa)$ for $\kappa > k$, without storing historical data. As such, the proposed test statistic at the kth interim stage is given by

$$\sup_x \frac{\varphi_k^\top(x)\{\widehat{\beta}_1^{(k)}(t_k) - \widehat{\beta}_0^{(k)}(t_k)\}}{\widehat{s.e}[\varphi_k^\top(x)\{\widehat{\beta}_1^{(k)}(t_k) - \widehat{\beta}_0^{(k)}(t_k)\}]},$$

where $\widehat{s.e}[\cdot]$ denotes some consistent standard error estimator.

In addition, we extend the proposed bootstrap algorithm to determine the stopping boundary. We aim to construct bootstrap statistics $\{\widehat{\beta}_a^{(k),\mathrm{MB}^*}(t_k)\}_k$ to approximate the distribution of $\{\widehat{\beta}_a^{(k)}(t_k)\}_k$. To allow for online updating, at the kth interim stage, we compute $\{\widehat{\beta}_a^{(k),\mathrm{MB}^*}(t_k), \widehat{\beta}_a^{(k+1),\mathrm{MB}^*}(t_k), \cdots, \widehat{\beta}_a^{(K),\mathrm{MB}^*}(t_k)\}^\top$ as

$$\frac{1}{N(t_k)}\sum_{j=1}^{k}\left(\sum_{i=N(t_{j-1})+1}^{N(t_j)} \mathbb{I}(A_i = a)\phi_{[k:K]}(X_i, Y_i)\phi_{[k:K]}^\top(X_i, Y_i)\right)^{1/2} e_{j,a}, \tag{39}$$

where $\phi_{[k:K]}(X_i, Y_i)$ denotes the vector

$$\left\{[\{\widehat{\Sigma}_a^{(k)}(t_j)\}^{-1}\varphi_k(X_i)\{Y_i - \varphi_k(X_i)\}]^\top, \cdots, [\{\widehat{\Sigma}_a^{(K)}(t_j)\}^{-1}\varphi_K(X_i)\{Y_i - \varphi_K(X_i)\}]^\top\right\}^\top,$$

and $e_{j,a}$ denotes a multivariate normal random vector with zero mean and identity covariance matrix. This yields the bootstrap test statistic

$$\sup_x \frac{\varphi_k^\top(x)\{\widehat{\beta}_1^{(k),\mathrm{MB}^*}(t_k) - \widehat{\beta}_0^{(k),\mathrm{MB}^*}(t_k)\}}{\widehat{s.e}[\varphi_k^\top(x)\{\widehat{\beta}_1^{(k)}(t_k) - \widehat{\beta}_0^{(k)}(t_k)\}]},$$

based on which the α-spending approach is applicable (see Equation (9) for details). The resulting test is valid, under certain regularity conditions on $\{\varphi_k\}_k$. See Appendix A for details.

Finally, we discuss some drawbacks of the resulting test. Compared to our proposed test in the main text, it would be much more computationally intensive to implement such a test. For instance, in order to compute the test statistic, at the kth interim stage, we need to compute not only $\gamma_{a,k}(t_k)$ and $\widehat{\Sigma}_a^{(k)}(t_k)$, but $\{\gamma_{a,\kappa}(t_k)\}_{\kappa>k}$ and $\{\widehat{\Sigma}_a^{(\kappa)}(t_k)\}_{\kappa>k}$ as well. More important, in order to compute the bootstrap statistic, we need to do a Cholesky decomposition on a $(\sum_{j=k}^K q_k) \times (\sum_{j=k}^K q_k)$ matrix at the kth interim stage (see Equation (39)). In cases where $K-k$ is large, this is much more computationally intensive than the proposed procedure that only requires to do a Cholesky decomposition on a $q \times q$ matrix.

6.2 Extensions to the two-sided test

In this paper, we focus on the null hypothesis H_0^a that the heterogeneous treatment effect (HTE) is nonpositive for any realization of the baseline covariates. It is also interesting to consider the two-sided null hypothesis that HTE is either always nonpositive (H_0^a), or always nonnegative (denoted by H_0^b). Notice that the latter corresponds to a union of H_0^a and H_0^b. To test such a null, one can separately test H_0^a and H_0^b using the proposed test, obtain the corresponding p-values p_a and p_b, and derive the p-value using the union-intersection principle, i.e., $p = \max(p_a, p_b)$.

Acknowledgement

The authors thank the AE, and two anonymous reviewers for their constructive comments, which have led to a significant improvement of the proposed methodology.

Appendix A. More on the basis function

(A3)(i) Assume $\lambda_{\min}[\mathrm{E}\varphi(X)\varphi^\top(X)] \asymp 1$, $\lambda_{\max}[\mathrm{E}\varphi(X)\varphi^\top(X)] \asymp 1$, $\sup_x \|\varphi(x)\|_1 = O(q^{1/2})$, $\liminf_q \inf_{x\in\mathbb{X}} \|\varphi(x)\|_2 > 0$. In addition, assume

$$\sup_{\substack{x,y\in\mathbb{X}\\ x\neq y}} \frac{\|\varphi(x)-\varphi(y)\|_2}{\|x-y\|_2} \preceq q^{1/2}. \tag{40}$$

(ii) Suppose

$$\mathrm{err} \equiv \inf_{\beta_0,\beta_1\in\mathbb{R}^q} \sup_{x\in\mathbb{X},a\in\{0,1\}} |Q_0(x,a) - Q(x,a;\beta_0,\beta_1)| = o(\{N(T)\}^{-1/2}). \tag{41}$$

When a tensor-product B-spline is used (see Section 6 of Chen and Christensen, 2015, for a brief overview of tenor-product B-splines), (A3)(i) is automatically satisfied. Specifically, $\lambda_{\min}[\mathrm{E}\varphi(X)\varphi^\top(X)] \asymp 1$, $\lambda_{\max}[\mathrm{E}\varphi(X)\varphi^\top(X)] \asymp 1$ follow from Theorem 3.3 of (Burman and Chen, 1989). $\sup_x \|\varphi(x)\|_1 = O(q^{1/2})$ follows by noting that the absolute value of each element in $\varphi(x)$ is bounded by $O(q^{1/2})$ and that the number of nonzero elements in $\varphi(x)$ is finite. $\liminf_q \inf_{x\in\mathbb{X}} \|\varphi(x)\|_2 > 0$ follows from the arguments used in the proof of Lemma E.4 of Shi et al. (2021). The last condition in (40) holds by noting that each function in the vector $\varphi(\cdot)$ is Lipschitz continuous when a tensor-product B-spline is used.

Suppose the Q-function $Q_0(\cdot,a)$ is p-smooth (see the definition of p-smoothness in Stone, 1982), for $a \in [0,1]$. When a tensor-product B-spline or Wavelet basis is used for $\varphi(\cdot)$, then there exist some β_0^* and β_1^* that satisfy

$$\inf_{\beta_0,\beta_1\in\mathbb{R}^q} \sup_{x\in\mathbb{X},a\in\{0,1\}} |Q_0(x,a) - Q(x,a;\beta_0,\beta_1)| = O(q^{-p/d}).$$

See Section 2.2 of Huang (1998) for detailed discussions on the approximation power of these basis functions. Condition (41) is thus automatically satisfied when

$$q \gg \{N(T)\}^{d/(2p)}.$$

As we have commented, the normalized test requires a weaker condition than the unnormalized test without standardization. Specifically, if we use the unnormalized test, we would require the approximation error to decay at a rate at $o\{q^{-1/2}N^{-1/2}(t)\}$, strictly faster than the RHS of (41) when q grows to infinity. To elaborate, the denominator in the normalized test is of the same order of magnitude as $N^{-1/2}(t)\|\phi(x)\|_2$, uniformly in x. This ensures the bias of the ratio $\varphi(x)^T(\widehat{\beta}_1-\widehat{\beta}_0)/\widehat{s.e.}[\varphi(x)^T(\widehat{\beta}_1-\widehat{\beta}_0)]$ is of the same order of magnitude as the bias of $\widehat{\beta}_1-\widehat{\beta}_0$, uniformly in x, eliminate the effect of the approximation error. Without standardization, the bias of the test would be of the same order of magnitude as the bias of $q^{1/2}(\widehat{\beta}_1-\widehat{\beta}_0)$.

Finally, we present the regularity conditions on $\{\varphi_k\}_k$ to ensure the validity of the proposed test in Section 6.1. They are very similar to those imposed in (A3).

(A3*)(i) For any k, assume $\lambda_{\min}[\mathrm{E}\varphi_k(X)\varphi_k^\top(X)] \asymp 1$, $\lambda_{\max}[\mathrm{E}\varphi_k(X)\varphi_k^\top(X)] \asymp 1$, $\sup_x \|\varphi_k(x)\|_1 = O(q_k^{1/2})$, $\liminf_k \inf_{x\in\mathbb{X}} \|\varphi_k(x)\|_2 > 0$. In addition, assume

$$\sup_{\substack{x,y\in\mathbb{X}\\ x\neq y}} \frac{\|\varphi_k(x)-\varphi_k(y)\|_2}{\|x-y\|_2} \preceq q_k^{1/2}.$$

(ii) Suppose

$$\inf_{\beta\in\mathbb{R}^{q_k}} \sup_{x\in\mathbb{X},a\in\{0,1\}} |Q_0(x,a) - \varphi_k^\top(x)\beta| = o(\{N(t_k)\}^{-1/2}).$$

Appendix B. Proofs

B.1 Proof of Lemma 1

Set $\mathcal{F}_0 = \emptyset$. We state the following lemma before proving Lemma 1.

Lemma 8 *For any $j \geq 1$, $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp \mathcal{F}_{j-1}$.*

For any $a \in \{0,1\}$, $i \geq 1$, notice that

$$\begin{aligned}\mathrm{E}\mathbb{I}(A_i = a)\{Y_i - Q_0(X_i,a)\} &= \mathrm{E}\mathbb{I}(A_i = a)\{Y_i^*(a) - Q_0(X_i,a)\} \\ &= \mathrm{E}\mathrm{E}^{X_i,\mathcal{F}_{i-1}}[\mathbb{I}(A_i = a)\{Y_i^*(a) - Q_0(X_i,a)\}],\end{aligned}$$

where the first equation is due to Assumption (A1) and $\mathrm{E}^{X_i,\mathcal{F}_{i-1}}$ denotes the conditional expectation given $\mathcal{F}_{i-1}$ and X_i. By Assumption (A2), we have

$$\mathrm{E}^{X_i,\mathcal{F}_{i-1}}[\mathbb{I}(A_i = a)\{Y_i^*(a) - Q_0(X_i,a)\}] = \{\mathrm{E}^{X_i,\mathcal{F}_{i-1}}\mathbb{I}(A_i = a)\}[\mathrm{E}^{X_i,\mathcal{F}_{i-1}}\{Y_i^*(a) - Q_0(X_i,a)\}].$$

The second term on the RHS equals zero due to Lemma 8 and our model assumption $\mathrm{E}\{Y_i^*(a)|X_i\} = Q_0(X_i,a)$. The proof is hence completed.

B.2 Proof of Lemma 8

The assertion trivially holds for $j = 1$. We prove it holds for any $j \geq 2$, by induction. By (A2), we have $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp A_1|X_1$. Since $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp (X_1, Y_1^*(0), Y_1^*(1))$, this further implies $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp A_1$ and hence $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp (X_1, A_1, Y_1^*(0), Y_1^*(1))$. By (A1), Y_1 is completely determined by A_1, $Y_1^*(0)$ and $Y_1^*(1)$. Therefore, we obtain $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp \mathcal{F}_1$.

Suppose we have shown that $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp \mathcal{F}_k$ for some $k < j-1$. To prove $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp \mathcal{F}_{k+1}$, it suffices to show $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp (X_{k+1}, A_{k+1}, Y_{k+1})$. By (A1), Y_{k+1} is determined by A_{k+1}, $Y_{k+1}^*(0)$ and $Y_{k+1}^*(1)$. Since $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp (X_{k+1}, Y_{k+1}^*(0), Y_{k+1}^*(1))$, it suffices to show $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp A_{k+1}$. This is implied by $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp A_{k+1}|X_{k+1}, \mathcal{F}_k$ and that $(X_j, Y_j^*(0), Y_j^*(1)) \perp\!\!\!\perp X_{k+1}, \mathcal{F}_k$. The proof is hence completed.

B.3 Proof of Lemma 3

The assertions

$$\epsilon_0 \leq \lambda_{\min}[\mathrm{E}\varphi(X)\varphi^\top(X)] \leq \lambda_{\max}[\mathrm{E}\varphi(X)\varphi^\top(X)] \leq \epsilon_0^{-1}, \tag{42}$$

and

$$\sup_x \|\varphi(x)\|_1 \leq \epsilon_0^{-1}\sqrt{q}, \tag{43}$$

for some $0 < \epsilon_0 < 1$ are directly implied by the conditions that $\lambda_{\min}[\mathrm{E}\varphi(X)\varphi^\top(X)] \asymp 1$, $\lambda_{\max}[\mathrm{E}\varphi(X)\varphi^\top(X)] \asymp 1$, $\sup_x \|\varphi(x)\|_1 \le \epsilon_0^{-1}\sqrt{q}$. Since $\|\varphi(x)\|_2 \le \|\varphi(x)\|_1$, we obtain $\sup_x \|\varphi(x)\|_2 \le \sup_x \|\varphi(x)\|_1 \le \epsilon_0^{-1}\sqrt{q}$.

Under the condition $\inf_{a,x} \pi^*(a,x) > 0$, we can similarly show that $\lambda_{\min}[\Sigma_a] \ge \epsilon_0$ for some $\epsilon_0 > 0$.

Notice that $Q_0(x,a) = \mathrm{E}\{Y^*(a)|X = x\}$. Under the condition that $\mathrm{E}[\{Y^*(a)\}^2|X]$ is bounded, we obtain $\sup_{x\in\mathbb{X}} \max_{a\in\{0,1\}} |Q_0(x,a)| \le \epsilon_0^{-1}$ for some $0 < \epsilon < 1$.

Notice that $\beta_a = \Sigma_a^{-1}\mathrm{E}\varphi^\top(X)Y^*(a)$. Since $\lambda_{\min}[\Sigma_a]$ is bounded away from 0, it suffices to show $\|\mathrm{E}\varphi^\top(X)Y^*(a)\|_2 = O(1)$, or equivalently,

$$\sup_{\nu\in\mathbb{R}^p, \|\nu\|_2=1} |\mathrm{E}\nu^\top\varphi(X)Y^*(a)| = O(1).$$

By Cauchy-Schwarz inequality, it suffices to show

$$\sup_{\nu\in\mathbb{R}^p, \|\nu\|_2=1} \mathrm{E}|Y^*(a)|^2\mathrm{E}|\nu^\top\varphi(X)|^2 = O(1).$$

We have by the condition $\lambda_{\max}[\mathrm{E}\varphi(X)\varphi^\top(X)] = O(1)$ that

$$\sup_{\nu\in\mathbb{R}^p, \|\nu\|_2=1} \mathrm{E}|\nu^\top\varphi(X)|^2 = \sup_{\nu\in\mathbb{R}^p, \|\nu\|_2=1} \nu^\top\mathrm{E}\varphi(X)\varphi^\top(X)\nu \le \lambda_{\max}[\mathrm{E}\varphi(X)\varphi^\top(X)] = O(1).$$

The sub-Gaussianity of $Y^*(a)$ implies that it has bounded second moment. The proof is hence completed.

B.4 Proof of Lemma 4

B.4.1 Proof of (12)

Notice that

$$\begin{aligned}\|j(\widehat{\Sigma}_{1,j} - \Sigma_1)\|_2 = &\left\|\sum_{i=1}^{j}\{A_i\varphi(X_i)\varphi^\top(X_i) - \mathrm{E}^{\mathcal{F}_{i-1}}\pi_{i-1}(1,X)\varphi(X)\varphi^\top(X)\}\right\|_2 \\ &+j\left\|\mathrm{E}^{\mathcal{F}_{i-1}}\varphi(X)\varphi^\top(X)\left(\frac{1}{j}\sum_{i=1}^{j}\pi_{i-1}(1,X) - \pi^*(1,X)\right)\right\|_2.\end{aligned} \tag{44}$$

By Lemma 3, we have

$$\begin{aligned}\left\|\mathrm{E}^{\mathcal{F}_{i-1}}\varphi(X)\varphi^\top(X)\left(\frac{1}{j}\sum_{i=1}^{j}\pi_{i-1}(1,X) - \pi^*(1,X)\right)\right\|_2 &\le \varepsilon_0^{-2}q\mathrm{E}^{\mathcal{F}_{i-1}}\left|\frac{1}{j}\sum_{i=1}^{j}\pi_{i-1}(1,X) - \pi^*(1,X)\right| \\ &\le \varepsilon_0^{-2}q^2 j^{-\alpha_0}\log^{\alpha_0} j, \quad \forall j \ge j_n,\end{aligned}$$

with probability at least $1 - O(j_n^{-\alpha_0})$.

Consider the first term on the RHS of (44). For any $i \ge 1$, define $M_i = \varphi(X_i)\varphi^\top(X_i)\{A_i - \pi_{i-1}(1,X_i)\}$. Notice that $\{M_i\}_{i\ge1}$ forms a martingale difference sequence with respect to the filtration $\{\sigma(\mathcal{F}_{i-1}) : i \ge 2\}$, since

$$\begin{aligned}&\mathrm{E}[\varphi(X_i)\varphi^\top(X_i)\{A_i - \pi_{i-1}(X_i)\}|\mathcal{F}_{i-1}] \\ =\ &\mathrm{E}^{\mathcal{F}_{i-1}}[\mathrm{E}(\varphi(X_i)\varphi(X_i)^\top\{A_i - \pi_{i-1}(X_i)\}|\mathcal{F}_{i-1}, X_i)] = 0,\end{aligned} \tag{45}$$

where $\mathrm{E}^{\mathcal{F}_i, X_i}$ denotes the conditional expectation given X_i and $\mathcal{F}_i$. Here, the first equality is due to that $X_i \perp\!\!\!\perp \mathcal{F}_{i-1}$, implied by Lemma 8. Under the given conditions on the basis function $\varphi(\cdot)$, using similar arguments in proving Equation (C.15) of Shi et al. (2021), we can show that the following event occurs with probability at least $1 - O(j^{-2})$,

$$\left\| \sum_{i=1}^{j} M_i \right\|_2 \preceq \sqrt{qj \log(j)}.$$

Notice that $\sum_{k \geq j} k^{-2} \leq j^{-2} + \sum_{k>j} \{k(k-1)\}^{-1} = j^{-2} + j^{-1}$. Thus, the following occurs with probability at least $1 - O(j_n^{-1})$,

$$\left\| \sum_{i=1}^{j} \{A_i \varphi(X_i) \varphi^\top(X_i) - \mathrm{E}^{\mathcal{F}_{i-1}} \pi_{i-1}(1, X) \varphi(X) \varphi^\top(X)\} \right\|_2 \preceq \sqrt{qj \log j}, \quad \forall j \geq j_n. \tag{46}$$

It follows that

$$\|(\widehat{\Sigma}_{1,k} - \Sigma_1)\|_2 \preceq q\delta_k + \sqrt{qk^{-1} \log k}, \quad \forall k \geq j_n,$$

with probability at least $1 - O(j^{-\alpha_0})$. Similarly, we can show

$$\|(\widehat{\Sigma}_{0,k} - \Sigma_0)\|_2 \preceq q\delta_k + \sqrt{qk^{-1} \log k}, \quad \forall k \geq j_n,$$

with probability at least $1 - O(j_n^{-\alpha_0})$. The proof is hence completed.

B.4.2 Proof of (13)

When j_n satisfies $j_n^{\alpha_0} / \log^{\alpha_0}(j_n) \gg q^2$, it follows from (12) and (42) that

$$\lambda_{\min}[\widehat{\Sigma}_{a,k}] \geq \lambda_{\min}[\Sigma_a] - \|\widehat{\Sigma}_{a,k} - \Sigma_a\|_2 \geq 2^{-1}\varepsilon_0, \quad \forall k \geq j_n, \tag{47}$$

with probability at least $1 - O(j_n^{-\alpha_0})$. Combining (42) with (47) and (12), we obtain

$$\begin{aligned} \|\widehat{\Sigma}_{a,k}^{-1} - \Sigma_a^{-1}\|_2 = \|\widehat{\Sigma}_{a,k}^{-1}(\widehat{\Sigma}_{a,k} - \Sigma_a)\Sigma_a^{-1}\|_2 &\leq \lambda_{\min}[\Sigma_a]\lambda_{\min}[\widehat{\Sigma}_{a,k})\|\widehat{\Sigma}_{a,k} - \Sigma_a\|_2 \\ &\preceq q\delta_k + \sqrt{qk^{-1} \log k}, \quad \forall k \geq j_n, \end{aligned}$$

with probability at least $1 - O(j_n^{-\alpha_0})$. The proof is hence completed.

B.5 Proof of Lemma 5

For any $l \in \{1, \ldots, q\}$ and $i \geq 1$, define $M_i(l) = \varphi^{(l)}(X_i) A_i \{Y_i - Q_0(X_i, a)\}$. Here, $\varphi^{(l)}(X_i)$ corresponds to the l-th element of $\varphi(X_i)$. Similar to (45), we can show $\{M_i(l)\}_{i \geq 1}$ forms a martingale difference sequence with respect to the filtration $\{\sigma(\mathcal{F}_{i-1}) : i \geq 1\}$. By (45), we have for any l,

$$\mathrm{E}\{\varphi^{(l)}(X_i)\}^2 \leq \lambda_{\max}[\varphi(X_i)\varphi^\top(X_i)] \leq \epsilon_0^{-1}. \tag{48}$$

Similar to Lemma 3, we can show that $\sigma^2(a,x)$ is uniformly bounded by $4\epsilon^{-1}$ for some $\epsilon > 0$ as well. It follows that

$$\begin{aligned} & \mathrm{E}\{M_i^2(l)|\mathcal{F}_{i-1}\} = \mathrm{E}[\{\varphi^{(l)}(X_i)\}^2 A_i \{Y_i^*(1) - Q_0(X_i,1)\}^2|\mathcal{F}_{i-1}] \\ \leq & \mathrm{E}[\{\varphi^{(l)}(X_i)\}^2 \{Y_i^*(1) - Q_0(X_i,1)\}^2|\mathcal{F}_{i-1}] = \mathrm{E}\sigma^2(1,X_i)\{\varphi^{(l)}(X_i)\}^2 \\ \leq & 4\epsilon_0^{-2}\mathrm{E}\{\varphi^{(l)}(X_i)\}^2 \leq 4\epsilon_0^{-3}, \end{aligned}$$

where the first equality is due to (A1), the first inequality is due to that A is bounded between 0 and 1, the second equality follows from Lemma 8, the second inequality follows from Lemma 3, and the last inequality is due to (48). It follows that

$$\sum_{i=1}^{k} \mathrm{E}\{M_i^2(l)|\mathcal{F}_{i-1}\} \leq 4k\epsilon_0^{-3}. \tag{49}$$

Similarly, by (A1) and Lemma 3, we have

$$\sum_{i=1}^{k} M_i^2(l) \leq 2\sum_{i=1}^{k} \{\psi^{(l)}(X_i)\}^2(\epsilon_0^{-2} + Y_i^2). \tag{50}$$

Under sub-Gaussianity, Y_i^2 has bounded sub-exponential tail. Similar to (46), it follows from Bernstein's inequality (Lemma 2.2.11 van der Vaart and Wellner, 1996) and Bonferroni's inequality that, with probability at least $1 - O(j^{-1})$ that

$$\sum_{i=1}^{k} [M_i^2(l) - \mathrm{E}\{M_i^2(l)|\mathcal{F}_{i-1}\}] \preceq \sqrt{qk\log k}, \quad \forall k \geq j.$$

Thus, for any sequence j_n that satisfies $j_n/\log(j_n) \gg q$, we have by (49) that

$$\sum_{i=1}^{k} M_i^2(l) + \sum_{i=1}^{k} \mathrm{E}\{M_i^2(l)|\mathcal{F}_{i-1}\} \leq \bar{c}k, \quad \forall k \geq j_n,$$

for some constant $\bar{c} > 0$, with probability at least $1 - O(j_n^{-1})$. It follows that

$$\begin{aligned}
& \Pr\left(\bigcap_{k\geq j_n}\{|\sum_{i=1}^{k} M_i(l)| \leq 2\sqrt{\bar{c}k\log k}\}\right) \\
\geq\ & \Pr\left(\left\{\bigcap_{k\geq j_n}\{|\sum_{i=1}^{k} M_i(l)| \leq 2\sqrt{\bar{c}k\log k}\}\right\}\bigcap\left\{\bigcap_{k\geq j_n}\{\sum_{i=1}^{k}[M_i^2(l) + \{M_i^2(l)|\mathcal{F}_{i-1}\}] \leq \bar{c}k\}\right\}\right) \\
-\ & O(j_n^{-1}) \geq \Pr\left(\left\{\bigcap_{k\geq j_n}\{\sum_{i=1}^{k}[M_i^2(l) + \{M_i^2(l)|\mathcal{F}_{i-1}\}] \leq \bar{c}k\}\right\}\right) - O(j_n^{-1}) \\
-\ & \Pr\left(\left\{\bigcup_{k\geq j_n}\{|\sum_{i=1}^{k} M_i(l)| > 2\sqrt{\bar{c}k\log k}\}\right\}\bigcap\left\{\bigcap_{k\geq j_n}\{\sum_{i=1}^{k}[M_i^2(l) + \{M_i^2(l)|\mathcal{F}_{i-1}\}] \leq \bar{c}k\}\right\}\right) \\
\geq\ & 1 - \Pr\left(\left\{\bigcup_{k\geq j_n}\{|\sum_{i=1}^{k} M_i(l)| > 2\sqrt{\bar{c}k\log k}\}\right\}\bigcap\left\{\bigcap_{k\geq j_n}\{\sum_{i=1}^{k}[M_i^2(l) + \{M_i^2(l)|\mathcal{F}_{i-1}\}] \leq \bar{c}k\}\right\}\right) \\
-\ & O(j_n^{-1}).
\end{aligned}$$

By Bonferroni's inequality and Theorem 2.1 of Bercu and Touati (2008), we have

$$\begin{aligned}
& \Pr\left(\bigcap_{k\geq j_n}\{|\sum_{i=1}^{k} M_i(l)| \leq 2\sqrt{\bar{c}k\log k}\}\right) \geq 1 - O(j_n^{-1}) \\
-\ & \sum_{k\geq j_n}\Pr\left(\{|\sum_{i=1}^{k} M_i(l)| > 2\sqrt{\bar{c}k\log k}\}\bigcap\left\{\bigcap_{k'\geq j_n}\{\sum_{i=1}^{k'}[M_i^2(l) + \{M_i^2(l)|\mathcal{F}_{i-1}\}] \leq \bar{c}k'\}\right\}\right) \\
\geq\ & 1 - O(j_n^{-1}) - \sum_{k\geq j_n}\Pr\left(\{|\sum_{i=1}^{k} M_i(l)| > 2\sqrt{\bar{c}k\log k}\}\bigcap\left\{\sum_{i=1}^{k}[M_i^2(l) + \{M_i^2(l)|\mathcal{F}_{i-1}\}] \leq \bar{c}k\right\}\right) \\
\geq\ & 1 - O(j_n^{-1}) - 2\sum_{k\geq j_n}\exp\left(-\frac{4\bar{c}k\log k}{2\bar{c}k}\right) = 1 - O(j_n^{-1}) - \sum_{k\geq j_n} 2k^{-2}. \qquad (51)
\end{aligned}$$

The last term on the RHS of (51) is $1 - O(j_n^{-1})$. To summarize, we have shown that the following event occurs with probability at least $1 - O(j_n^{-1})$,

$$\bigcap_{k\geq j_n}\left\{|\sum_{i=1}^{k} M_i(l)| \leq 2\sqrt{\bar{c}k\log k}\right\}.$$

By Bonferroni's inequality, we have

$$\bigcap_{k\geq j_n}\left\{\left\|\sum_{i=1}^{k}\varphi(X_i)A_i\{Y_i - Q_0(X_i,1)\}\right\|_2 \leq 2\sqrt{\bar{c}qk\log k}\right\},$$

with probability at least $1 - O(j_n^{-1/2})$. Similarly, we can show

$$\bigcap_{k \geq j_n} \left\{ \left\| \sum_{i=1}^{k} \varphi(X_i)(1 - A_i)\{Y_i - Q_0(X_i, 0)\} \right\|_2 \leq c\sqrt{qk \log k} \right\},$$

for some constant $c > 0$, with probability at least $1 - O(j_n^{-1})$. The proof is hence completed.

B.6 Proof of Lemma 6

Combining Lemma 5 with Lemma 3 yields that

$$\left\| \Sigma_a^{-1} \left(\frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\{Y_i - Q_0(X_i, a)\} \right) \right\|_2 \preceq q^{1/2} k^{-1/2} \sqrt{\log k}, \quad \forall k \geq j_n, a \in \{0, 1\},$$

with probability at least $1 - O(j_n^{-1})$. Combining this together with (14) yields that

$$\|\widehat{\beta}_{a,k} - \beta_a\|_2 \preceq q^{1/2} k^{-1/2} \sqrt{\log k}, \quad \forall k \geq j_n, a \in \{0, 1\},$$

with probability at least $1 - O(j_n^{-1})$. The proof is hence completed.

B.7 Proof of Lemma 7

Notice that

$$\begin{aligned}
&\left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - Q_0(X_i, a)\}^2 - \Phi_a \right\|_2 \\
\leq &\left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)[\{Y_i - Q_0(X_i, a)\}^2 - \sigma^2(a, X_i)] \right\|_2 \\
+ &\left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\sigma^2(a, X_i) - \Phi_a \right\|_2 .
\end{aligned} \tag{52}$$

Similar to the proof of Lemma 4, we can show that the second term on the RHS of (52) is of the order $O(q\delta_k + \sqrt{qk^{-1} \log k})$, for any $a \in \{0, 1\}$ and any $k \geq j_n$, with probability at least $1 - O(j_n^{-\alpha_0})$. As for the first term, notice that each element in the matrix

$$\frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)[\{Y_i - Q_0(X_i, a)\}^2 - \sigma^2(a, X_i)] \tag{53}$$

corresponds to a martingale with respect to the filtration $\{\sigma(\mathcal{F}_{i-1}) : i \geq 1\}$, under (A1) and (A2). Using similar arguments in proving Lemma E.2 of Shi et al. (2021), we can show that

$$\left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)[\{Y_i - Q_0(X_i, a)\}^2 - \sigma^2(a, X_i)] \right\|_2 \preceq q^{1/2} k^{-1/2} \sqrt{\log k}, \quad \forall a \in \{0, 1\}, k \geq j_n,$$

with probability at least $1 - O(j_n^{-1})$.

Finally, we focus on providing an upper bound for the bias term

$$\begin{aligned}
&\left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a) \varphi(X_i) \varphi^\top(X_i) [\{Y_i - Q_0(X_i, a)\}^2 - \{Y_i - \varphi^\top(X_i)\beta_a\}^2] \right\| \\
\leq 2 &\left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a) \varphi(X_i) \varphi^\top(X_i) [\{Y_i - Q_0(X_i, a)\}\{Q_0(X_i, a) - \varphi^\top(X_i)\beta_a\}] \right\| \\
&+ \left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{I}(A_i = a) \varphi(X_i) \varphi^\top(X_i) \right\| o\{(NT)^{-1/2}\}.
\end{aligned}$$

Using similar arguments, the first term on the RHS can be upper bounded by $Cq^{1/2}k^{-1/2}\sqrt{\log k}$ for some constant $C > 0$. The second term is $o\{(NT)^{-1/2}\}$. The proof is hence completed.

B.8 Proof of Lemma 2

We begin by providing an upper bound for $\max_{a\in\{0,1\}} \|\widehat{\beta}_{a,k} - \beta_a\|_2$. With some calculations, we have

$$\begin{aligned}
\max_{a\in\{0,1\}} \|\widehat{\beta}_{a,k} - \beta_a\|_2 &= \max_{a\in\{0,1\}} \frac{1}{k} \left\| \widehat{\Sigma}_{a,k}^{-1} \left(\sum_{i=1}^{k} \varphi(X_i) \mathbb{I}(A_i = a) \{Y_i - \varphi^\top(X_i)\beta_a\} \right) \right\|_2 \\
&\leq \max_{a\in\{0,1\}} \left\| \widehat{\Sigma}_{a,k}^{-1} \right\|_2 \max_{a\in\{0,1\}} \frac{1}{k} \left\| \sum_{i=1}^{k} \varphi(X_i) \mathbb{I}(A_i = a) \{Y_i - \varphi^\top(X_i)\beta_a\} \right\|_2.
\end{aligned}$$

Using similar arguments in proving Theorem 1, we obtain with probability at least $1-O(j_n^{-1})$ that

$$\max_{a\in\{0,1\}} \frac{1}{k} \left\| \sum_{i=1}^{k} \varphi(X_i) \mathbb{I}(A_i = a) \{Y_i - \varphi^\top(X_i)\beta_a\} \right\|_2 \preceq q^{1/2}k^{-1/2}\sqrt{\log k}, \quad \forall k \geq j_n. \tag{54}$$

Similarly, we can show with probability at least $1 - O(j_n^{-1})$ that

$$\max_{a\in\{0,1\}} \frac{1}{k} \left\| \sum_{i=1}^{k} \varphi(X_i) \mathbb{I}(A_i = a) \{Y_i - \varphi^\top(X_i)\beta_a\} \right\|_2 \preceq q^{1/2}k^{-1/2}\sqrt{\log j_n}, \quad \forall 1 \leq k < j_n. \tag{55}$$

Similar to (44), we have

$$\begin{aligned}
\max_{a\in\{0,1\}} \lambda_{\min}[\widehat{\Sigma}_{a,k}] \geq &\min_{a\in\{0,1\}} \lambda_{\min} \left(\mathrm{E}^{\mathcal{F}_{i-1}} \varphi(X) \varphi^\top(X) \frac{1}{k} \sum_{i=1}^{k} \pi_{i-1}(a, X) \right) \\
- \max_{a\in\{0,1\}} \frac{1}{k} &\left\| \sum_{i=1}^{k} \{\mathbb{I}(A_i = a) \varphi(X_i) \varphi^\top(X_i) - \mathrm{E}^{\mathcal{F}_{i-1}} \pi_{i-1}(a, X) \varphi(X) \varphi^\top(X)\} \right\|_2.
\end{aligned}$$

Using similar arguments in proving (46), we can show that

$$\max_{a\in\{0,1\}}\left\|\sum_{i=1}^{k}\{\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)-\mathrm{E}^{\mathcal{F}_{i-1}}\pi_{i-1}(a,X)\varphi(X)\varphi^\top(X)\}\right\|_2 \preceq \sqrt{qk\log k}, \quad \forall k\geq j_n, \tag{56}$$

with probability at least $1-O(j_n^{-1})$. Similarly, we can show

$$\max_{a\in\{0,1\}}\left\|\sum_{i=1}^{k}\{\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)-\mathrm{E}^{\mathcal{F}_{i-1}}\pi_{i-1}(a,X)\varphi(X)\varphi^\top(X)\}\right\|_2 \preceq \sqrt{qk\log j_n}, \quad \forall 1\leq k<j_n, \tag{57}$$

with probability at least $1-O(j_n^{-1})$.

Without loss of generality, assume $\varepsilon_0\leq 1/2$. Notice that we have $\pi_{i-1}(a,x)\geq\varepsilon_0$, for any $a\in\{0,1\}$, $x\in\mathbb{X}$ and $i\geq N_0$. This together with Lemma (3) implies that

$$\inf_{a\in\{0,1\},n\geq j_n}\lambda_{\min}\left(\mathrm{E}^{\mathcal{F}_{i-1}}\varphi(X)\varphi^\top(X)\frac{1}{n}\sum_{i=1}^{n}\pi_{i-1}(a,X)\right)\geq\frac{n-N_0}{n}\varepsilon_0\geq\frac{j-N_0}{j}\varepsilon_0.$$

Combining this together with (56) and (57) yields

$$\max_{a\in\{0,1\}}\lambda_{\min}[\widehat{\Sigma}_{a,k}]\geq\frac{\varepsilon_0}{2}, \quad \forall k\geq L^*\sqrt{q\log j_n},$$

for some constant $L^*\geq 1$, with probability at least $1-O(j_n^{-1})$. This together with (54) and (55) yields that

$$\max_{a\in\{0,1\}}\|\widehat{\beta}_{a,k}-\beta_a\|_2\preceq q^{1/2}k^{-1/2}\sqrt{\log\max(k,j_n)}, \quad \forall k\geq L^*\sqrt{q\log j_n},$$

with probability at least $1-O(j_n^{-1})$.

By Condition (A3), we have

$$|\varphi^\top(X)(\widehat{\beta}_{1,k}-\widehat{\beta}_{0,k}-\beta_1+\beta_0)|\leq\bar{L}qk^{-1/2}\log^{1/2}\max(k,j_n), \quad \forall k\geq L^*\sqrt{q\log j_n}, \tag{58}$$

for some constant $\bar{L}>0$, with probability at least $1-O(j_n^{-1})$.

For any $z_1,z_2\in\mathbb{R}$, we have $\mathbb{I}(z_1>0)\neq\mathbb{I}(z_2>0)$ only when $|z_1-z_2|\geq|z_2|$. Hence, under the event defined in (58), the event $\mathbb{I}\{\varphi^\top(X)(\widehat{\beta}_{1,k}-\widehat{\beta}_{0,k})>0\}\neq\mathbb{I}\{\varphi^\top(X)(\beta_1-\beta_0)>0\}$ occurs only when

$$|\varphi^\top(X)(\beta_1-\beta_0)|\leq|\varphi^\top(X)(\widehat{\beta}_{1,k}-\widehat{\beta}_{0,k}-\beta_1+\beta_0)|\leq\bar{L}qk^{-1/2}\sqrt{\log\max(k,j_n)},$$

for any $k\geq j_n$. Since $Q_0(X,1)-Q_0(X,0)$ can be well-approximated by $\varphi^\top(X)(\beta_1-\beta_0)$, the above event occurs only when

$$|Q_0(X,1)-Q_0(X,0)|\leq 2\bar{L}qk^{-1/2}\sqrt{\log\max(k,j_n)}.$$

Under the given conditions, we have

$$\Pr\left(|Q_0(X,1)-Q_0(X,0)|\le 2\bar{L}qk^{-1/2}\log^{1/2}\max(k,j_n)\right)\le 2\bar{L}L_0qk^{-1/2}\log^{1/2}\max(k,j_n). \quad (59)$$

Notice that when $\mathbb{I}\{\varphi^\top(X)(\widehat{\beta}_{1,k}-\widehat{\beta}_{0,k})>0\}=\mathbb{I}\{\varphi^\top(X)(\beta_1-\beta_0)>0\}$, we have $\pi_k(a,X)=\pi^*(a,X)$. Thus, we obtain $\pi_k(a,X)=\pi^*(a,X)$ if $|\varphi^\top(X)(\beta_1-\beta_0)|>2\bar{L}qk^{-1/2}\sqrt{\log\max(k,j_n)}$, for any $k\ge L^*\sqrt{q\log j_n}$. Set $k_0=L^*\sqrt{q\log j_n}$. By (58) and (59), we have with probability at least $1-O(j_n^{-1})$ that

$$\begin{aligned}
&\sum_{a\in\{0,1\}}\mathrm{E}^{\mathcal{F}_{i-1}}\left|\sum_{i=1}^{k}\{\pi_{i-1}(a,X)-\pi^*(a,X)\}\right|\le\sum_{a\in\{0,1\}}\sum_{i=1}^{k_0}\mathrm{E}^{\mathcal{F}_{i-1}}|\pi_{i-1}(a,X)-\pi^*(a,X)|\\
+&\sum_{a\in\{0,1\}}\sum_{i=k_0+1}^{k}\mathrm{E}^{\mathcal{F}_{i-1}}|\pi_{i-1}(a,X)-\pi^*(a,X)|\le 2L^*\sqrt{q\log j_n}\\
+&\sum_{a\in\{0,1\}}\sum_{i=k_0+1}^{k}\mathrm{E}^{\mathcal{F}_{i-1}}|\pi_{i-1}(a,X)-\pi^*(a,X)|\mathbb{I}\{|\varphi^\top(X)(\beta_1-\beta_0)|>2\bar{L}qi^{-1/2}\log^{1/2}i\}\\
+&\sum_{a\in\{0,1\}}\sum_{i=k_0+1}^{k}\mathrm{E}^{\mathcal{F}_{i-1}}|\pi_{i-1}(a,X)-\pi^*(a,X)|\mathbb{I}\{|\varphi^\top(X)(\beta_1-\beta_0)|\le 2\bar{L}qi^{-1/2}\log^{1/2}i\}\\
\le&\ 2L^*\sqrt{q\log j_n}+\sum_{a\in\{0,1\}}\sum_{i=k_0+1}^{n}\Pr\left(|\varphi^\top(X)(\beta_1-\beta_0)|\le\bar{L}qi^{-1/2}\sqrt{\log i}\right)\preceq qk^{1/2}\log^{1/2}k,\quad\forall k\ge j_n.
\end{aligned}$$

The proof is hence completed.

B.9 Proof of Theorem 3

Similar to the proof of Theorem 1, we will show that the assertion in Theorem 3 holds for any $n(\cdot)$ that correspond to the realizations of $N(\cdot)$ that satisfy $n(t_1)<n(t_2)<\cdots<n(t_K)$. For any $1\le k_1\le k_2\le K$, define

$$\begin{aligned}
\widehat{V}(k_1,k_2)&=\sqrt{n(t_{k_1})n(t_{k_2})}\mathrm{cov}\left(\widehat{\beta}_1^{\mathrm{MB}*}(t_{k_1})-\widehat{\beta}_0^{\mathrm{MB}*}(t_{k_1}),\widehat{\beta}_1^{\mathrm{MB}*}(t_{k_2})-\widehat{\beta}_0^{\mathrm{MB}*}(t_{k_2})|\{(X_i,A_i,Y_i)\}_{i=1}^{+\infty}\right)\\
&=\frac{1}{\sqrt{n(t_{k_1})n(t_{k_2})}}\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\widehat{\Sigma}_a^{-1}(t_j)\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\widehat{\Sigma}_a^{-1}(t_j),
\end{aligned}$$

and

$$\widehat{\boldsymbol{V}}=\begin{pmatrix}\widehat{V}(1,1)&\widehat{V}(1,2)&\dots&\widehat{V}(1,K)\\\widehat{V}(2,1)&\widehat{V}(2,2)&\dots&\widehat{V}(2,K)\\\vdots&\vdots&&\vdots\\\widehat{V}(K,1)&\widehat{V}(K,2)&\dots&\widehat{V}(K,K)\end{pmatrix}.$$

We aim to bound the entrywise ℓ_∞ norm of $\widehat{\boldsymbol{V}}-\boldsymbol{V}$ where $\boldsymbol{V}$ is defined in (21). It suffices to bound $\max_{1\le k_1\le k_2\le K}\sup_{b_1,b_2\in\mathbb{R}^{p+1},\|b_1\|_2=\|b_2\|_2=1}|b_1^T\{\widehat{V}(k_1,k_2)-V(k_1,k_2)\}b_2|=$

$\max_{1\le k_1\le k_2\le K}\|\widehat{V}(k_1,k_2)-V(k_1,k_2)\|_2$. For any k_1,k_2, we decompose $\widehat{V}(k_1,k_2)-V(k_1,k_2)$ as

$$\widehat{V}(k_1,k_2)-V(k_1,k_2)=\widehat{V}(k_1,k_2)-\widehat{V}^*(k_1,k_2)+\widehat{V}^*(k_1,k_2)-\widehat{V}^{**}(k_1,k_2)+\widehat{V}^{**}(k_1,k_2)-V(k_1,k_2),$$

where

$$\widehat{V}^*(k_1,k_2)=\frac{1}{\sqrt{n(t_{k_1})n(t_{k_2})}}\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\Sigma_a^{-1},$$

$$\widehat{V}^{**}(k_1,k_2)=\frac{1}{\sqrt{n(t_{k_1})n(t_{k_2})}}\sum_{a=0}^{1}\sum_{j=1}^{n(t_{k_1})}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a^*\}^2\Sigma_a^{-1}.$$

By Lemma 3 and Lemma 7, we obtain that

$$\begin{aligned}
&\max_{1\le k_1\le k_2\le K}\|\widehat{V}^{**}(k_1,k_2)-V(k_1,k_2)\|_2\\
\le\ &\max_{1\le k_1\le K}\sum_{a=0}^{1}\left\|\frac{1}{n(t_{k_1})}\sum_{j=1}^{n(t_{k_1})}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a^*\}^2\Sigma_a^{-1}-\Sigma_a^{-1}\Phi_a\Sigma_a^{-1}\right\|_2\\
\le\ &\max_{1\le k_1\le K}\frac{1}{\epsilon_0^2}\sum_{a=0}^{1}\left\|\frac{1}{n(t_{k_1})}\sum_{j=1}^{n(t_{k_1})}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a^*\}^2-\Phi_a\right\|_2\\
\preceq\ &q\delta_{n(t_1)}+q^{1/2}n^{-1/2}(t_1)\sqrt{\log n(t_1)},
\end{aligned}\tag{60}$$

with probability at least $1-O(n^{-\alpha_0}(t_1))$.

Notice that

$$\begin{aligned}
\sqrt{n(t_{k_1})n(t_{k_2})}\widehat{V}^*(k_1,k_2)&=\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\Sigma_a^{-1}\\
&=\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a^*+\varphi^\top(X_i)\beta_a^*-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\Sigma_a^{-1}\\
&=\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{\varphi^\top(X_i)\beta_a^*-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\Sigma_a^{-1}\\
&\quad+2\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a^*\}\varphi^\top(X_i)\{\beta_a^*-\widehat{\beta}_a(t_j)\}\Sigma_a^{-1}\\
&\quad+\sqrt{n(t_{k_1})n(t_{k_2})}\widehat{V}^{**}(k_1,k_2).
\end{aligned}$$

It follows that

$$\begin{aligned}
&\max_{1\le k_1\le k_2\le K}\left\|\widehat{V}^*(k_1,k_2)-\widehat{V}^{**}(k_1,k_2)\right\|_2\\
\le\ &\max_{1\le k_1\le K}\frac{1}{n(t_{k_1})}\left\|\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{\varphi^\top(X_i)\beta_a^*-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2\Sigma_a^{-1}\right\|_2\\
+\ &\max_{1\le k_1\le K}\frac{2}{n(t_{k_1})}\left\|\sum_{a=0}^{1}\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\Sigma_a^{-1}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a\}\varphi^\top(X_i)\{\beta_a^*-\widehat{\beta}_a(t_j)\}\Sigma_a^{-1}\right\|_2
\end{aligned}$$

By Lemma 3, we obtain that

$$\begin{aligned}
&\max_{1\le k_1\le k_2\le K}\left\|\widehat{V}^*(k_1,k_2)-\widehat{V}^{**}(k_1,k_2)\right\|_2 \qquad (61)\\
\preceq\ &\max_{\substack{1\le k_1\le K\\ a\in\{0,1\}}}\frac{1}{n(t_{k_1})}\Big\|\underbrace{\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{\varphi^\top(X_i)\beta_a^*-\varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2}_{\Psi_{1,a,k_1}}\Big\|_2\\
+\ &\max_{\substack{1\le k_1\le K\\ a\in\{0,1\}}}\frac{2}{n(t_{k_1})}\Big\|\underbrace{\sum_{j=1}^{k_1}\sum_{i=n(t_{j-1})+1}^{n(t_j)}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\{Y_i-\varphi^\top(X_i)\beta_a^*\}\varphi^\top(X_i)\{\beta_a^*-\widehat{\beta}_a(t_j)\}}_{\Psi_{2,a,k_1}}\Big\|_2.
\end{aligned}$$

By Lemmas 3 and 6, we have with probability at least $1-O(n^{-1}(t_1))$ that

$$\frac{1}{n(t_{k_1})}\|\Psi_{1,a,k_1}\|_2\preceq q^2n^{-1}(t_1)\log\{n(t_1)\}\left\|\frac{1}{n(t_{k_1})}\sum_{i=1}^{n(t_{k_1})}\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\right\|_2, \qquad (62)$$
$$\forall 1\le k_1\le K, a\in\{0,1\}.$$

Similar to Lemma 4, we can show there exists some constant $c_*>0$ that

$$\begin{aligned}
&\frac{1}{n(t_{k_1})}\left\|\sum_{i=1}^{n(t_{k_1})}[\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)-\mathrm{E}^{\mathcal{F}_{i-1}}\{\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\}]\right\|_2 \qquad (63)\\
&\quad\le c_*\{q\delta_{n(t_{k_1})}+q^{1/2}n^{-1/2}(t_{k_1})\sqrt{\log n(t_{k_1})}\},\quad \forall 1\le k_1\le K, a\in\{0,1\},
\end{aligned}$$

with probability at least $1-O(n^{-1}(t_1))$. By Lemma 3, we can show with probability at least $1-O(n^{-1}(t_1))$ that

$$\max_{1\le k_1\le K}\frac{1}{n(t_{k_1})}\left\|\sum_{i=1}^{n(t_{k_1})}\mathrm{E}^{\mathcal{F}_{i-1}}\{\mathbb{I}(A_i=a)\varphi(X_i)\varphi^\top(X_i)\}\right\|_2=O(1).$$

This together with (62) and (63) yields

$$n^{-1}(t_{k_1})\|\Psi_{1,a,k_1}\|_2\preceq q^2n^{-1}(t_1)\log\{n(t_1)\},\quad \forall 1\le k_1\le K, a\in\{0,1\}, \qquad (64)$$

with probability at least $1 - O(n^{-1}(t_1))$.

Moreover, using similar arguments in proving Lemma 7, we can show that for any $1 \le k_1 \le K$, the following event occurs with probability at least $1 - O(n^{-2}(t_{k_1}))$,

$$\frac{1}{n(t_{k_1})} \left\| \sum_{i=1}^{n(t_{k_1})} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\beta_a^*\}\varphi^{(l)}(X_i) \right\|_2 \preceq q^{1/2} n^{-1/2}(t_{k_1}) \sqrt{\log n(t_{k_1})}, \quad \forall 1 \le l \le q.$$

Since $\sum_{k_1=1}^K n^{-2}(t_{k_1}) \le n^{-1}(t_1)$, we obtain with probability at least $1 - O(n^{-1}(t_1))$ that

$$\frac{1}{n(t_{k_1})} \left\| \sum_{i=1}^{n(t_{k_1})} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\beta_a\}\varphi^{(l)}(X_i) \right\|_2 \preceq q^{1/2} n^{-1/2}(t_{k_1}) \sqrt{\log n(t_{k_1})}, \quad \forall 1 \le l \le q, 1 \le k_1 \le K.$$

In addition, it follows from Lemma 6 that

$$n^{-1}(t_{k_1}) \|\Psi_{2,a,k_1}\|_2 \preceq q^{3/2} n^{-1}(t_1) \log\{n(t_1)\}, \quad \forall 1 \le k_1 \le K, a \in \{0,1\}.$$

This together with (64) yields that

$$\max_{1 \le k_1 \le k_2 \le K} \left\| \widehat{V}^*(k_1,k_2) - \widehat{V}^{**}(k_1,k_2) \right\|_2 \preceq q^2 n^{-1}(t_1) \log n(t_1),$$

with probability at least $1 - O(n^{-1}(t_1))$. Under the given conditions, we have

$$\max_{1 \le k_1 \le k_2 \le K} \left\| \widehat{V}^*(k_1,k_2) - \widehat{V}^{**}(k_1,k_2) \right\|_2 \preceq q^{1/2} n^{-1/2}(t_1) \log^{1/2} n(t_1), \tag{65}$$

with probability at least $1 - O(n^{-1}(t_1))$.

Moreover, with some calculations, we can show that

$$\begin{aligned}
&\max_{1 \le k_1 \le k_2 \le K} \left\| \widehat{V}(k_1,k_2) - \widehat{V}^*(k_1,k_2) \right\|_2 \le \sum_{a=0}^{1} \max_{j \ge 1} \|\Sigma_a^{-1} - \widehat{\Sigma}_a^{-1}(t_j)\|_2 \\
\times \quad & \max_{1 \le k_1 \le K} \frac{2}{n(t_{k_1})} \left\| \sum_{j=1}^{k_1} \sum_{i=n(t_{j-1})+1}^{n(t_j)} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2 \Sigma_a^{-1} \right\|_2 \\
+ \quad & \sum_{a=0}^{1} \max_{1 \le k_1 \le K} \frac{1}{n(t_{k_1})} \left\| \sum_{j=1}^{k_1} \sum_{i=n(t_{j-1})+1}^{n(t_j)} \Sigma_a^{-1} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2 \Sigma_a^{-1} \right\|_2 \\
\times \quad & \max_{j \ge 1} \|\Sigma_a^{-1} - \widehat{\Sigma}_a^{-1}(t_j)\|_2^2.
\end{aligned}$$

In view of Lemma 3 and Lemma 4, we have with probability at least $1 - O(n^{-\alpha_0}(t_1))$ that

$$\begin{aligned}
&\max_{1 \le k_1 \le k_2 \le K} \left\| \widehat{V}(k_1,k_2) - \widehat{V}^*(k_1,k_2) \right\|_2 \le O(1)(q\delta_{n(t_1)} + \sqrt{qn^{-1}(t_1)\log n(t_1)}) \\
\times \quad & \max_{1 \le k_1 \le K} \frac{1}{n(t_{k_1})} \Big\| \underbrace{\sum_{j=1}^{k_1} \sum_{i=n(t_{j-1})+1}^{n(t_j)} \mathbb{I}(A_i = a)\varphi(X_i)\varphi^\top(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_a(t_j)\}^2}_{\Psi_{3,a,k_1}} \Big\|_2,
\end{aligned}$$

where $O(1)$ denotes some positive constant. Similar to (60) and (65), we can show with probability at least $1-O(n^{-\alpha_0}(t_1))$ that

$$\max_{a\in\{0,1\}}\max_{1\le k_1\le K}\left\|\frac{1}{n(t_{k_1})}\Psi_{3,a,k_1}-\Psi_a\right\|_2=o(1).$$

Similar to Lemma 3, we can show $\max_{a\in\{0,1\}}\|\Psi_a\|_2=O(1)$. It follows that

$$\max_{1\le k_1\le k_2\le K}\left\|\widehat{V}(k_1,k_2)-\widehat{V}^*(k_1,k_2)\right\|_2\preceq q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)},$$

with probability at least $1-O(n^{-\alpha_0}(t_1))$. Combining this together with (60) and (65), we obtain with probability at least $1-O(n^{-\alpha_0}(t_1))$ that

$$\max_{1\le k_1\le k_2\le K}\left\|\widehat{V}(k_1,k_2)-V(k_1,k_2)\right\|_2\preceq q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)}.$$

Consider the function $\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k}$ defined in the proof of Theorem 1. We fix $\delta=\eta^{-1}\{\log K+4d\log n(t_1)\}$. Based on Lemma A2 in Belloni and Oliveira (2018), we have with probability at least $1-O(n^{-\alpha_0}(t_1))$ that

$$\begin{aligned}&\sup_{\{\nu_k\}_k}\left|\mathrm{E}^*\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k}(N(0,\widehat{\boldsymbol{V}}))-\mathrm{E}\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_k\}_k}(N(0,\boldsymbol{V}))\right|\\&\preceq\eta^2\{\log^2K+\log^2n(t_1)\}\left(q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)}\right),\end{aligned}$$

where E^* denotes the expectation conditional on the observed data. For a given set of thresholds $\{\nu_k\}_k$, using similar arguments in proving (30), (31), (33) and (34), we can show with probability at least $1-O(n^{-\alpha_0}(t_1))$ that

$$\begin{aligned}&\mathrm{Pr}^*\left\{\max_{k\in\{1,\ldots,K\}}\left(\sqrt{n(t_k)}\widehat{S}^{\mathrm{MB}*}-\nu_k\right)\le0\right\}\le\mathrm{E}^*\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_{k,+}\}_k}(N(0,\widehat{\boldsymbol{V}}))\\\le\;&\mathrm{E}\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_{k,+}\}_k}(N(0,\boldsymbol{V}))+O(1)\eta^2\{\log^2K+\log^2n(t_1)\}\left(q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)}\right)\\\le\;&\mathrm{Pr}\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\varphi^\top(x)G(t_k)-\nu^*_{k,+}\right)\le0\right\}\\+\;&O(1)\eta^2\{\log^2K+\log^2n(t_1)\}\left(q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)}\right),\end{aligned}$$

and

$$\begin{aligned}&\mathrm{Pr}^*\left\{\max_{k\in\{1,\ldots,K\}}\left(\sqrt{n(t_k)}\widehat{S}^{\mathrm{MB}*}-\nu_k\right)\le0\right\}\ge\mathrm{E}^*\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_{k,-}\}_k}(N(0,\widehat{\boldsymbol{V}}))\\\ge\;&\mathrm{E}\mathrm{g}_\delta\circ\phi_{\eta,\{\nu_{k,-}\}_k}(N(0,\boldsymbol{V}))-O(1)\eta^2\{\log^2K+\log^2n(t_1)\}\left(q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)}\right)\\\ge\;&\mathrm{Pr}\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\varphi^\top(x)G(t_k)-\nu^*_{k,-}\right)\le0\right\}\\-\;&O(1)\eta^2\{\log^2K+\log^2n(t_1)\}\left(q\delta_{n(t_1)}+\sqrt{qn^{-1}(t_1)\log n(t_1)}\right),\end{aligned}$$

where $O(1)$ denotes some positive constant, and

$$\nu_{k,+} = \nu_k + \eta^{-1}\{4d\log n(t_1) + \log K\} + \bar{c}^* n^{-2}(t_1), \quad \nu^*_{k,+} = \nu_{k,+} + 3\eta^{-1}\{4d\log n(t_1) + \log K\},$$
$$\nu_{k,-} = \nu_k - 3\eta^{-1}\{4d\log n(t_1) + \log K\} - \bar{c}^* n^{-2}(t_1), \quad \nu^*_{k,-} = \nu_{k,-} - \eta^{-1}\{4d\log n(t_1) + \log K\}.$$

By Theorem 2 of Chernozhukov et al. (2017), we obtain that

$$\Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x,t_k)} - \nu^*_{k,+}\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}_0}\frac{\varphi^\top(x)G(t_k)}{\sigma(x,t_k)} - \nu^*_{k,-}\right) \le 0\right\}$$
$$\preceq \eta^{-1}\{\log^{3/2} n(t_1) + \log^{3/2} K\} + \bar{c}^* n^{-2}(t_1)\{\log^{1/2} n(t_1) + \log^{1/2} K\}.$$

It follows that

$$\sup_{\{\nu_k\}_k}\left|\Pr^*\left\{\max_{k\in\{1,\ldots,K\}}\left(\sqrt{n(t_k)}\widehat{S}^{\mathrm{MB}*} - \nu_k\right) \le 0\right\} - \Pr\left\{\max_{k\in\{1,\ldots,K\}}\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)G(t_k)}{\sigma(x,t_k)} - \nu_k\right) \le 0\right\}\right|$$
$$\preceq \eta^2\{\log^2 K + \log^2 n(t_1)\}\left(q\delta_{n(t_1)} + \sqrt{qn^{-1}(t_1)\log n(t_1)}\right)$$
$$+\eta^{-1}\{\log^{3/2} n(t_1) + \log^{3/2} K\} + \bar{c}^* n^{-2}(t_1)\{\log^{1/2} n(t_1) + \log^{1/2} K\},$$

with probability at least $1 - O(n^{-\alpha_0}(t_1))$. Set

$$\eta = \min[q^{-1/3}n^{\alpha_0/3}(t_1)\log^{-(1+2\alpha_0)/6}\{Kn(t_1)\}, q^{-1/6}n^{1/6}(t_1)\log^{-1/3}\{Kn(t_1)\}],$$

we obtain the desired result.

B.10 Proof of Theorem 5

We will show that under the current conditions, the stopping boundaries $\{\widehat{z}_k\}_k$ are upper bounded by $c\sqrt{\log N(t_1)}$ for some constant $c > 0$, with probability tending to 1. Using similar arguments in proving Lemma 5 and Equation (27), we can show that

$$\left|\sup_{x\in\mathbb{X}}\frac{\sqrt{N(t_K)}\varphi^\top(x)\{\widehat{\beta}_1(t_K) - \beta^*_1 - \widehat{\beta}_0(t_K) + \beta^*_0\}}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_K) - \widehat{\beta}_0(t_K)\}]}\right|$$

is upper bounded by $O\{\sqrt{\log N(t_1)}\}$ as well, with probability tending to 1.

Under the alternative hypothesis and the assumption on the approximation error,

$$\sqrt{N(t_K)}\sup_{x\in\mathbb{X}}\varphi^\top(x)(\beta^*_1 - \beta^*_0) \gg \sqrt{q\log\{N(t_1)\}}.$$

Using similar arguments in the proof of Theorem 1, the standard error estimator is of the order of magnitude $O(\sqrt{q})$, uniformly for any x. It follows that

$$\sqrt{N(t_K)}\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)(\beta^*_1 - \beta^*_0)}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_K) - \widehat{\beta}_0(t_K)\}]} > \widehat{z}_K,$$

with probability tending to 1. The proof is hence completed.

It remains to show that the stopping boundaries are upper bounded by $c\sqrt{q\log N(t_1)}$. We will prove this assertion by induction. We first notice that $\widehat{z}_1$ satisfies

$$\Pr\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)\widehat{S}^{\mathrm{MB}*}(t_1)}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_K)-\widehat{\beta}_0(t_K)\}]}-\widehat{z}_1>0\middle|\,\mathrm{Data}\right)=\alpha(t_1).$$

The conditional variance of $\varphi^\top(x)\widehat{S}^{\mathrm{MB}*}(t_1)/\|\varphi(x)\|_2$ is upper bounded by $O(1)$. The denominator is lower bounded by $c\|\varphi(x)\|_2$ for some constant $c>0$. As such, there exists some constant $c_1>0$ such that

$$\Pr\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)\widehat{S}^{\mathrm{MB}*}(t_1)}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_1)-\widehat{\beta}_0(t_1)\}]}>c_1\sqrt{\log N(t_1)}\middle|\,\mathrm{Data}\right)\le\frac{1}{N^C(t_1)}.$$

As such, $\widehat{z}_1\le c_1\sqrt{\log N(t_1)}$.

Suppose we have shown that $\{\widehat{z}_j\}_{j=1}^k$ are upper bounded by $c_k\sqrt{\log N(t_1)}$. We aim to show $\widehat{z}_{k+1}$ is upper bounded by $c_{k+1}\sqrt{\log N(t_1)}$ for some constant $c_{k+1}>0$. A key observation is that

$$\begin{aligned}
&\Pr\left(\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)\widehat{S}^{\mathrm{MB}*}(t_{k+1})}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_1)-\widehat{\beta}_0(t_1)\}]}-\widehat{z}_{k+1}>0\middle|\,\mathrm{Data}\right)\\
\ge&\Pr\left(\max_{1\le j\le k+1}\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)\widehat{S}^{\mathrm{MB}*}(t_j)}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_1)-\widehat{\beta}_0(t_1)\}]}-\widehat{z}_j>0\middle|\,\mathrm{Data}\right)\\
-\Pr\left(\max_{1\le j\le k}\sup_{x\in\mathbb{X}}\frac{\varphi^\top(x)\widehat{S}^{\mathrm{MB}*}(t_j)}{\widehat{s.e.}[\varphi^\top(x)\{\widehat{\beta}_1(t_1)-\widehat{\beta}_0(t_1)\}]}-\widehat{z}_j>0\middle|\,\mathrm{Data}\right)&=\alpha(t_{k+1})-\alpha(t_k).
\end{aligned}$$

Since $\alpha(t_{k+1})-\alpha(t_k)\ge N^C(t_1)$, using similar arguments in proving $\widehat{z}_1\le c_1\sqrt{\log N(t_1)}$, we can show that $\widehat{z}_{k+1}\le c_{k+1}\sqrt{\log N(t_1)}$. This shows that the power of the proposed test approaches to one. Similarly, one can show that the stopping time is upper bounded by t_k, as long as

$$\sqrt{N(t_k)}\sup_{x\in\mathbb{X}}\varphi^\top(x)(\beta_1^*-\beta_0^*)\gg\sqrt{q\log\{N(t_1)\}}.$$

Appendix C. Comparison of the baseline method

We first introduce the test based on LIL. Consider our test statistic $S(t)$. Under H_0, it can be bounded from above by

$$\sup_{x\in\mathbb{X}}\varphi^\top(x)\{\widehat{\beta}_1(t)-\beta_1^*-\widehat{\beta}_0(t)+\beta_0^*\}. \tag{66}$$

It suffices to provide an upper bound for the above expression. By Cauchy-Schwarz inequality, (66) can be upper bounded by

$$\sup_{x\in\mathbb{X}}\|\varphi(x)\|_2\|\widehat{\beta}_1(t)-\beta_1^*-\widehat{\beta}_0(t)+\beta_0^*\|_2.$$

It suffices to provide anytime upper bound for $\|\widehat{\beta}_1(t) - \beta_1^* - \widehat{\beta}_0(t) - \beta_0^*\|_2$.

Recall that

$$\begin{aligned} & \widehat{\beta}_1(t) - \beta_1^* - \widehat{\beta}_0(t) + \beta_0^* \\ = \; & \frac{1}{N(t)} \sum_{i=1}^{N(t)} [\mathbb{I}(A_i = 1)\widehat{\Sigma}_1^{-1}(t)\varphi(X_i)\{Y_i - \varphi^\top(X_i)\beta_1^*\} - \mathbb{I}(A_i = 0)\widehat{\Sigma}_0^{-1}(t)\varphi(X_i)\{Y_i - \varphi^\top(X_i)\beta_0^*\}]. \end{aligned}$$

The above expression is asymptotically equivalent to

$$\frac{1}{N(t)} \sum_{i=1}^{N(t)} [\mathbb{I}(A_i = 1)\Sigma_1^{-1}\varphi(X_i)\{Y_i - \varphi^\top(X_i)\beta_1^*\} - \mathbb{I}(A_i = 0)\Sigma_0^{-1}\varphi(X_i)\{Y_i - \varphi^\top(X_i)\beta_0^*\}].$$

By the law of iterated logarithm, the ℓ-th dimension of the above expression can be upper bounded by

$$N^{-1/2}(t)\sqrt{2\sigma_\ell^2 \log\log\{N(t)\}},$$

where $\sum_\ell \widehat{\sigma}_\ell^2$ can be consistently estimated by

$$\frac{1}{N(t)} \sum_{i=1}^{N(t)} \|\mathbb{I}(A_i = 1)\widehat{\Sigma}_1^{-1}(t)\varphi(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_1(t)\} - \mathbb{I}(A_i = 0)\widehat{\Sigma}_0^{-1}(t)\varphi(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_0(t)\}\|_2^2.$$

As such, the finite error bound is given by

$$\sup_{x\in\mathbb{X}} \|\varphi(x)\|_2 \frac{\sqrt{2\log\log\{N(t)\}}}{\sqrt{N(t)}} \times$$
$$\sqrt{\frac{1}{N(t)} \sum_{i=1}^{N(t)} \|\mathbb{I}(A_i = 1)\widehat{\Sigma}_1^{-1}(t)\varphi(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_1(t)\} - \mathbb{I}(A_i = 0)\widehat{\Sigma}_0^{-1}(t)\varphi(X_i)\{Y_i - \varphi^\top(X_i)\widehat{\beta}_0(t)\}\|_2^2}.$$

We next discuss the test based on AVT. At time t, we compute the following test statistic

$$\sqrt{\frac{\widehat{\sigma}^2/(N_0^{-1}(t) + N_1^{-1}(t))}{\widehat{\sigma}^2/(N_0^{-1}(t) + N_1^{-1}(t)) + \tau^2}} \exp\left[\frac{\tau^2\{N_0^{-1}(t)\sum_{i=1}^{N(t)}(1 - A_i)Y_i - N_1^{-1}(t)\sum_{i=1}^{N(t)} A_i Y_i\}^2}{2\{\widehat{\sigma}^2/(N_0^{-1}(t) + N_1^{-1}(t))\}\{\widehat{\sigma}^2/(N_0^{-1}(t) + N_1^{-1}(t)) + \tau^2\}}\right],$$

where $N_a(t) = \sum_{i=1}^{N(t)} \mathbb{I}(A_i = a)$ and $\widehat{\sigma}^2$ is the pooled variance estimator $\{N(t)-2\}^{-1}[\{N_0(t)-1\}\widehat{\sigma}_0^2+\{N_1(t)-1\}\widehat{\sigma}_1^2]$ where $\widehat{\sigma}_a^2$ denotes the sampling variance estimator based on $\{Y_i\}_{\mathbb{I}(A_i=a)}$. The constant τ corresponds to a hyperparameter and we fix $\tau = 1$ in our implementation.

References

Donald W. K. Andrews and Xiaoxia Shi. Inference based on conditional moment inequalities. *Econometrica*, 81(2):609–666, 2013. ISSN 0012-9682. doi: 10.3982/ECTA9370. URL http://dx.doi.org/10.3982/ECTA9370.

Donald W. K. Andrews and Xiaoxia Shi. Nonparametric inference based on conditional moment inequalities. *J. Econometrics*, 179(1):31–45, 2014. ISSN 0304-4076. doi: 10.1016/j.jeconom.2013.10.005. URL http://dx.doi.org/10.1016/j.jeconom.2013.10.005.

Timothy B. Armstrong and Hock Peng Chan. Multiscale adaptive inference on conditional moment inequalities. *J. Econometrics*, 194(1):24–43, 2016. ISSN 0304-4076. doi: 10.1016/j.jeconom.2016.04.001. URL http://dx.doi.org/10.1016/j.jeconom.2016.04.001.

Susan Athey and Guido Imbens. Recursive partitioning for heterogeneous causal effects. *Proc. Natl. Acad. Sci. USA*, 113(27):7353–7360, 2016. ISSN 1091-6490. doi: 10.1073/pnas.1510489113. URL https://doi.org/10.1073/pnas.1510489113.

Alexandre Belloni and Roberto I Oliveira. A high dimensional central limit theorem for martingales, with applications to context tree models. *arXiv preprint arXiv:1809.02741*, 2018.

Alexandre Belloni, Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Some new asymptotic theory for least squares series: Pointwise and uniform results. *Journal of Econometrics*, 186(2):345–366, 2015.

Bernard Bercu and Abderrahmen Touati. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.*, 18(5):1848–1869, 2008. ISSN 1050-5164. doi: 10.1214/07-AAP506. URL https://doi.org/10.1214/07-AAP506.

Prabir Burman and Keh-Wei Chen. Nonparametric estimation of a regression function. *Ann. Statist.*, 17(4):1567–1596, 1989. ISSN 0090-5364. doi: 10.1214/aos/1176347382.

Minsu Chang, Sokbae Lee, and Yoon-Jae Whang. Nonparametric tests of conditional treatment effects with an application to single-sex schooling on academic achievements. *Econom. J.*, 18(3):307–346, 2015. ISSN 1368-4221.

Xiaohong Chen and Timothy M. Christensen. Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions. *J. Econometrics*, 188(2):447–465, 2015. ISSN 0304-4076. doi: 10.1016/j.jeconom.2015.03.010.

Xiaohong Chen and Timothy M Christensen. Optimal sup-norm rates and uniform inference on nonlinear functionals of nonparametric iv regression. *Quantitative Economics*, 9(1):39–84, 2018.

Victor Chernozhukov, Sokbae Lee, and Adam M. Rosen. Intersection bounds: estimation and inference. *Econometrica*, 81(2):667–737, 2013. ISSN 0012-9682. doi: 10.3982/ECTA8718. URL http://dx.doi.org/10.3982/ECTA8718.

Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related Gaussian couplings. *Stochastic Process. Appl.*, 126(12):3632–3651, 2016. ISSN 0304-4149. doi: 10.1016/j.spa.2016.04.009. URL `https://doi.org/10.1016/j.spa.2016.04.009`.

Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Detailed proof of nazarov's inequality. *arXiv preprint arXiv:1711.10696*, 2017.

M Gail and R Simon. Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*, 41(2):361–372, 1985.

L. Gunter, J. Zhu, and S. A. Murphy. Variable selection for qualitative interactions. *Stat. Methodol.*, 8(1):42–55, 2011. ISSN 1572-3127. doi: 10.1016/j.stamet.2009.05.003. URL `http://dx.doi.org/10.1016/j.stamet.2009.05.003`.

Lacey Gunter, Ji Zhu, and Susan Murphy. Variable selection for optimal decision making. In *Conference on Artificial Intelligence in Medicine in Europe*, pages 149–154. Springer, 2007.

Yu-Chin Hsu. Consistent tests for conditional treatment effects. *The Econometrics Journal*, 20(1):1–22, 2017.

Jianhua Z. Huang. Projection estimation in multiple regression with application to functional ANOVA models. *Ann. Statist.*, 26(1):242–272, 1998. ISSN 0090-5364. doi: 10.1214/aos/1030563984.

Christopher Jennison and Bruce W Turnbull. *Group sequential methods with applications to clinical trials*. Chapman and Hall/CRC, 1999.

Ramesh Johari, Leo Pekelis, and David J Walsh. Always valid inference: Bringing sequential analysis to a/b testing. *arXiv preprint arXiv:1512.04922*, 2015.

Ramesh Johari, Pete Koomen, Leonid Pekelis, and David Walsh. Peeking at a/b tests: Why it matters, and what to do about it. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1517–1525. ACM, 2017.

Nianqiao Ju, Diane Hu, Adam Henderson, and Liangjie Hong. A sequential test for selecting the better variant: Online a/b testing, adaptive allocation, and continuous monitoring. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pages 492–500. ACM, 2019.

Eugene Kharitonov, Aleksandr Vorobev, Craig Macdonald, Pavel Serdyukov, and Iadh Ounis. Sequential testing for early stopping of online experiments. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 473–482. ACM, 2015.

K. K. Gordon Lan and David L. DeMets. Discrete sequential boundaries for clinical trials. *Biometrika*, 70(3):659–663, 1983. ISSN 0006-3444. doi: 10.2307/2336502. URL `https://doi.org/10.2307/2336502`.

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

Gregory Morrow and Walter Philipp. An almost sure invariance principle for Hilbert space valued martingales. *Trans. Amer. Math. Soc.*, 273(1):231–251, 1982. ISSN 0002-9947. doi: 10.2307/1999203. URL https://doi.org/10.2307/1999203.

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

Jeremy Roth and Noah Simon. A framework for estimating and testing qualitative interactions with applications to predictive biomarkers. *Biostatistics*, 19(3):263–280, 2018. ISSN 1468-4357. doi: 10.1093/biostatistics/kxx038. URL https://doi.org/10.1093/biostatistics/kxx038.

Chengchun Shi, Wenbin Lu, and Rui Song. Breaking the curse of nonregularity with subagging: inference of the mean outcome under optimal treatment regimes. *Journal of Machine Learning Research*, 2020a.

Chengchun Shi, Wenbin Lu, and Rui Song. A sparse random projection-based test for overall qualitative treatment effects. *Journal of the American Statistical Association*, 115 (531):1201–1213, 2020b.

Chengchun Shi, Sheng Zhang, Rui Song, and Wenbin Lu. Statistical inference of the value function for reinforcement learning in infinite horizon settings. *Journal of the Royal Statistical Society, Series B*, accepted, 2021.

Charles J. Stone. Optimal global rates of convergence for nonparametric regression. *Ann. Statist.*, 10(4):1040–1053, 1982. ISSN 0090-5364.

Matt Taddy, Matt Gardner, Liyun Chen, and David Draper. A nonparametric bayesian analysis of heterogenous treatment effects in digital experimentation. *Journal of Business & Economic Statistics*, 34(4):661–672, 2016.

Aad W. van der Vaart and Jon A. Wellner. *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996. ISBN 0-387-94640-3. doi: 10.1007/978-1-4757-2545-2. URL http://dx.doi.org/10.1007/978-1-4757-2545-2. With applications to statistics.

Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *J. Amer. Statist. Assoc.*, 113(523):1228–1242, 2018. ISSN 0162-1459. doi: 10.1080/01621459.2017.1319839. URL https://doi.org/10.1080/01621459.2017.1319839.

C.-F. J. Wu. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.*, 14(4):1261–1350, 1986. ISSN 0090-5364. doi: 10.1214/aos/1176350142.

URL https://doi.org/10.1214/aos/1176350142. With discussion and a rejoinder by the author.

Fanny Yang, Aaditya Ramdas, Kevin G Jamieson, and Martin J Wainwright. A framework for multi-a (rmed)/b (andit) testing with online fdr control. In *Advances in Neural Information Processing Systems*, pages 5957–5966, 2017.

Miao Yu, Wenbin Lu, and Rui Song. A new framework for online testing of heterogeneous treatment effect. In *AAAI*, pages 10310–10317, 2020.

Baqun Zhang, Anastasios A Tsiatis, Eric B Laber, and Marie Davidian. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika*, 100(3):681–694, 2013.

Li-xin Zhang. Strong approximations of martingale vectors and their applications in Markov-chain adaptive designs. *Acta Math. Appl. Sin. Engl. Ser.*, 20(2):337–352, 2004. ISSN 0168-9673. doi: 10.1007/s10255-004-0173-z. URL https://doi.org/10.1007/s10255-004-0173-z.