Jie Peng

Data driven individualized decision making problems have received a lot of attentions in recent years. In particular, decision makers aim to determine the optimal Individualized Treatment Rule (ITR) so that the expected speci ed outcome averaging over heterogeneous patient-speci c characteristics is maximized. Many existing methods deal with binary or a moderate number of treatment arms and may not take potential treatment e ect structure into account. However, the e ectiveness of these methods may deteriorate when the number of treatment arms becomes large. In this article, we propose GRoup Outcome Weighted Learning (GROWL) to estimate the latent structure in the treatment space and the optimal group-structured ITRs through a single optimization. In particular, for estimating group-structured ITRs, we utilize the Reinforced Angle based Multicategory Support Vector Machines (RAMSVM) to learn group-based decision rules under the weighted angle based multi-class classi cation framework. Fisher consistency, the excess risk bound, and the convergence rate of the value function are established to provide a theoretical guarantee for GROWL. Extensive empirical results in simulation studies and real data analysis demonstrate that GROWL enjoys better performance than several other existing methods.

Angle-based multicategory classi cation, Group structure, Individualized treatment rules, Precision medicine, Support Vector Machine.

https://creativecommons.org/licenses/by/4.0/
http://jmlr.org/papers/v24/22-0520.html

A common data-driven individualized decision making problem seeks to optimize the expected value of a speci ed outcome, by carefully determining the Individualized Treatment Rule (ITR) based on individual characteristics and contextual information. Since the treatment e ect may contain signi cant heterogeneity, it is necessary to tailor treatment decision rules to di erent subgroups of individuals. For example, using a large-scale Electronic Health Records (EHR) database, a physician may assign an optimal individualized therapy based on a patient s speci c characteristics to maximize the quality of health care (Wu et al., 2020).

Machine learning based approaches for estimating an optimal ITR have been studied intensively in the literature. These methods can be usually classi ed into two categories. The rst category consists of model-based indirect learning methods such as modeling the conditional treatment e ects given the individual characteristics (Q-learning) (Watkins, 1989; Qian and Murphy, 2011), modeling the contrast between two candidate treatment e ects (A-learning) (Murphy, 2003), sub-group identi cation methods based on a weighted loss minimization problem (Tian et al., 2014; Chen et al., 2017), and direct learning methods (D-learning) (Qi and Liu, 2018). The second category circumvents the need for modeling conditional mean functions by directly estimating the ITR that maximizes the value function based on Inverse Probability Weighting (IPW) (Zhao et al., 2012, 2015). To combine the advantages of methods in the two categories discussed above, Zhang et al. (2012), Liu et al. (2018) and Athey and Wager (2021) proposed doubly robust augmented IPW estimation to overcome model misspeci cation issues. In addition, extensions to more than two treatments were studied in Zhang et al. (2021) and Qi et al. (2020).

Despite great development for estimating the optimal ITRs with a moderate number of treatments in the literature as discussed above, in some clinical problems, there can be many treatment options available. For instance, Rashid et al. (2021) analyzed the Patient-Derived Xenograft (PDX) dataset which permits the evaluation of more than 20 treatments in the allowable treatment space. Another potential challenge for learning the optimal ITR is the situation with unbalanced structure of treatment propensity scores. For example, in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study (Rush et al., 2004), the ratio of the number of patients who were provided with the cognitive therapy and the number of patients who received venlafaxine is only around 1:3. Another example is that, when studying Type 2 Diabetes (T2D) treatment patterns, Montvida et al. (2018) concluded that the baseline treatments such as Metformin and Insulin would dominate other treatment options in the EHR database.

With many treatments but limited data size, model-based indirect methods are di - cult to model conditional treatment e ect due to the large number of interaction terms between treatments and features. In addition, it can be impractical to t a useful regression model without enough observations for certain treatments. Therefore, the estimated optimal ITR induced by the indirect methods can be inaccurate with large variability due to the poor performance of the regression model. On the other hand, IPW-based direct learning methods utilize a plug-in approach for possible unbalanced propensities that appear in the denominator of IPW-based value function. Su ering from unbalanced structure

of propensities in the presence of many treatments, small values in propensity scores can lead to large variability of the estimated value function.

It is interesting to point out that many treatments may work similarly for patients, due to the fact that the development of drugs is often based on intervening the same disease symptoms and mechanisms. For example, for treating depression in the STAR*D study, the 7 treatment options at Level 2 are often combined with one class of treatments involving selective serotonin reuptake inhibitors (SSRI) and the other class of treatments without SSRI because the treatments within the same class have similar treatment e ects (Liu et al., 2018; Pan and Zhao, 2021). Hence, it can be helpful to identify such latent structure in the treatment space. Moreover, utilizing this latent cluster treatment structure allows us to group homogeneous treatments together and helps reducing the dimension of the treatment space. This motivates us to explore speci c latent structure for treatments to identify optimal treatment groups.

To the best of our knowledge, not much has been done in the literature for estimating the optimal ITR with latent structure for treatments. Rashid et al. (2021) imposed a hierarchy binary group structure based on the conditional treatment e ects for the treatments in the PDX study. This estimated group structure helps producing high-quality ITRs and identifying the important genes that are known to be associated with response to treatment. In addition, several existing methods explored combining treatment decision rules for different patients when the conditional treatment e ects cannot be distinguished. Speci cally, Laber et al. (2014), Ertefaie et al. (2016) and Meng et al. (2020) proposed recommending a set of near-optimal individualized treatment recommendations that are alternative to each other to a patient. However, these methods are not tailored to deal with many treatment options.

In this article, we propose estimating the latent multiple group structure of treatments and associated optimal group-structured ITRs within a single optimization. Considering grouping structure, our proposed method reduces the dimension of the treatment space and automatically clusters the treatments with similar treatment e ects into the same group. In particular, we de ne our value function associated with both treatment partition and group-based decision rules in the IPW-based direct learning framework. The optimal treatment partition and group-based decision rules are obtained by maximizing the value function. When the treatment e ects employ exact homogeneous group structure, our de ned optimal partition can induce the same expected homogeneous group structure. Under the estimated optimal partition for the treatment space, the estimated group-structured ITR uses a random treatment assignment strategy, determined by randomly sampling treatment based on speci c strategies within the estimated optimal treatment group. Speci cally, the Reinforced Angle based Multicategory Support Vector Machines (RAMSVM) based surrogate loss function (Zhang et al., 2016) is tailored for estimating the optimal treatment group decision rules robustly in the interpretable angle-based weighted multiclass classi cation framework (Zhang and Liu, 2014). The group-based decision functions can give linear or non-linear decision rules to deal with complicate decision boundaries. Moreover, we prove that the surrogate loss function enjoys Fisher consistency for both group structure and group-structured ITRs. Furthermore, we present comprehensive theoretical justi cation on the excess risk bound, nite sample regret bound and convergence rate for our method, and allow the number of the treatment groups diverge to in nity as the sample size increases.

Finally, we implement e cient algorithms to solve the non-convex integer programming problem to search for the optimal partition, and the coordinate descent algorithm to solve the dual problem of RAMSVM based weighted classi cation problem.

The main contributions of this article are summarized as follows. Our proposed method learns the optimal ITR by identifying the latent treatment group structure in a possible large treatment space. We cluster the treatments with similar treatment e ects into the same group to reduce the dimension of the possible large treatment space. In contrast to existing methods (Zhao et al., 2012; Liu et al., 2018), our method avoids using weights involving the inverse of individual treatment propensity scores, which can be close to 0 when there are many treatments. Using the treatment group propensity scores, our method can obtain more stable estimate of the value function. In addition, our method simultaneously learns the optimal group-structured ITR and clusters the treatments. Di erent from the two-step method (Rashid et al., 2021), we combine both supervised learning (learn the optimal ITR) and unsupervised learning (cluster the treatments) through one single optimization. Moreover, we propose an e ective procedure to determine the number of unknown treatment groups. This procedure is motivated by the trade-o between the bene t and the variability of the value function. It is worth noting that our theoretical contributions are di erent from that in the Outcome Weighted Learning (OWL) literature (Zhao et al., 2012). In particular, we establish the generalized Fisher consistency, excess risk bound, and nite sample regret bound with respect to both and under the angle-based multi-class classi cation framework.

The remainder of this article is organized as follows. In Section 2, we introduce the methodology and implementation details of our proposed GRoup Outcome Weighted Learning (GROWL) method. In Section 3, we provide theoretical guarantees of GROWL. In Section 4, we conduct simulation studies to evaluate the performance of GROWL. Our method is then illustrated using the data from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study in Section 5. We conclude this article and discuss some future extensions in Section 6.

In this section, we rst introduce the framework of estimating optimal ITRs. Then we propose our GROWL method to estimate group-structured ITRs from the IPW-based value function.

Consider the i.i.d. training data ( ) for 1 , where $\in$ $\sim$ denotes the patient s prognostic variables, $\in$ $\{1\ 2$ $\}$: [ ] is the treatment assignment, and $\in$ is the observed outcome for each patient . Suppose that the number of treatments may diverge to in nity with a certain rate as the sample size increases since we consider the large treatment space. We assume that the larger outcome is better and is bounded. Let ( ) $_{\in}$ $\in$ be the potential outcome. In addition, de ne the propensity function (|) : (|) and the unknown mean-outcome function (|) : [( )|] for $\in$ [ ]. An Individualized Treatment Rule (ITR) $\in$ is a map from the covariate space to the treatment space and $\sim$ is a prespeci ed

ITR class. Our goal is to nd the optimal ITR $\in$, that maximizes the expected outcome, known as the value function (Zhao et al., 2012). Speci cally, the value of an ITR is de ned as

$$( ) \quad [( )] ( ) \quad [( )] (|)$$

Next we state the following identi ability assumptions (Rubin, 1974): (1) Consistency: $\in [ ] ( )$; (2) Unconfoundedness: for each $\in [ ] ( ) \perp\!\!\!\perp |$; (3) Positivity: $(|) \quad 0$ for any $\in$. If the above assumptions are satis ed, $( )$ can be written as the following two equivalent forms:

$$( ) \quad [( )] [|] \tag{1}$$

$$\frac{[( )]}{(|)} \tag{2}$$

Based on (1), model-based Q-learning methods (Qian and Murphy, 2011) rst give an estimate for $\widehat{[|]}$ (Q-function), then the optimal ITR $( )$ is estimated from solving $\arg\max_{\in} \widehat{[|]}$. However, due to the large number of treatment options and possible unbalanced structure of the propensity score $(|)$, we may not have enough observations for some speci c treatments to t the regression model. Consequently, $\widehat{[|]}$ can be inaccurate due to potential poor performance of the regression model and the estimated optimal ITR may have large variability. Another common approach is to estimate the value function based on (2) using empirical data, and then directly search for the optimal ITR that maximizes the empirical value function $\frac{[( )]}{(|)}$ (Zhao et al., 2012). Note that the propensity score $(|)$ appears in the denominator of $\frac{[( )]}{(|)}$. For the case with many treatments where insu cient data are observed for some speci c treatments, it is likely to have the propensity score $(|)$ close to 0 for some treatments. Hence, this can cause large variability of the empirical estimate for the value function.

Next we introduce our proposed GROWL method using the idea of latent group structure for the treatment space. We consider treatments can be partitioned into disjoint latent groups where 2 . We allow go to in nity with a certain rate as the sample size increases. Denote as the partition of , which is a map from to $\{1 \; 2 \; \}$: $[ ]$. Under , denote $\{ | ( ) \in \}$ as the -th treatment set for $\in [ ]$. Intuitively, for a reasonable partition, the treatments that belong to the same treatment group should have similar treatment e ects. In contrast, the treatment e ects from di erent treatment groups should have relatively large di erences. Hence, we need to rst de ne the optimal partition that maximizes the expected outcome.

To start with, we de ne the following group-structured ITR class, denoted as $[$. Speci cally, associated with a partition , a group-structured ITR in is obtained from a

random treatment assignment strategy given as

$$(\mid) \quad [( ) \quad ( )] \frac{(\mid)}{(( ) \mid)} \tag{3}$$

where is a group-based decision rule mapping from to treatment group space [ ], and (( ) |) : [$\in$ ( ) |] is the propensity score for the ( )-th treatment group under . Then, for a given partition and a group-based decision rule , the value function of group-structured ITR equals to the expectation of weighted conditional treatment e ects. With these notations in place, we can express the value of group-structured ITR ( ) as follows:

$$\begin{aligned} (\quad) \quad & [\quad (\quad) \quad] \quad \frac{(\mid)}{(( ) \mid)} \; [\mid \quad] \\ & [\quad (\quad) \quad (\quad)] \frac{}{((\quad) \mid \quad)} \end{aligned} \tag{4}$$

( )

( )

(|)

(( ) |)

( ) ( )

For any given , the optimal group-based decision rule is given by

$$\in \arg\max_{\rightarrow [\quad]} \quad (\quad) \tag{5}$$

and the corresponding optimal value for is

$$( ) \quad (\quad) \tag{6}$$

The optimal partition is de ned as

$$\in \arg\max \quad ( ) : \tag{7}$$

where is the optimal equivalent partition class and each element in this set achieves the maximum value. Observing that ( ) $\frac{(\mid)}{(( ) \mid)}$ [\|] [\| $\in$] for $\in$ [ ], the optimal value for in (6) can be written as

$$( ) \quad \max_{\rightarrow [\quad]} \quad [\quad (\quad) \quad] \; [\mid \in \quad]$$

$$\max_{\in [\quad]} \; [\mid \in \quad]$$

Hence, the optimal partition has the following interpretation. Averaging over the marginal distribution of , the maximum of conditional treatment e ects under the group domain, which is a mixture mean of conditional treatment e ects under the individual treatment domain, is optimized under the optimal partition .

It is worth noting that, when treatment e ects have homogeneous structure, our de ned optimal partition in (7) would lead to this expected natural group structure. In particular, treatment e ects have homogeneous structure if treatments can be partitioned into homogeneous groups { } and the treatment e ects are identical within each treatment group set for ∈ [ ]. For each ∈ [ ] and each pair of treatments within the same treatment group , we have [|] [|] [| ∈] a.e. in Meanwhile, for each pair of treatments that belong to two di erent treatment groups respectively, [|] [|] holds with a positive probability in . In this case, GROWL aims to combine treatments with identical treatment e ects based on homogeneous structure to reduce the dimension of treatment space and learn the optimal group-structured ITR. Denote to be the partition that induces the group structure , then the following Lemma 2 holds, which demonstrates that our is properly de ned for the expected homogeneous structure.

Next we illustrate how to solve ( ) ∈ arg max →[ ] ( ) in order to get an estimate for the optimal partition and the associated optimal group-based decision rule under the IPW-based direct learning framework. Since

$$( ) \quad ( ) \quad \frac{}{(( )|)}$$

where the risk function ( ) : $\frac{}{(( )|)}$ [( ) ( )] . Hence, maximizing ( ) is equivalent to minimizing the generalized risk function:

$$\tilde{}( ) \quad ( ) \quad \frac{}{(( )|)}$$

In practice, we use empirical risk minimization to approximate the generalized risk function $\tilde{}( )$ by $\frac{}{(( )|)}$ [( ) ( )] $\frac{}{(( )|)}$, where is the empirical average based on the training data. To alleviate the di culty of the discontinuity and nonconvexity of the 0-1 loss in the treatment group-based weighted misclassi cation error, for each ∈ and ∈ , we propose replacing the 0-1 loss function [( ) ( )] by a angle-based loss ( ) ( ) , as proposed in Zhang and Liu (2014) and Zhang et al. (2016). The group-based decision rule is determined by the decision function mapping from to . Speci cally, We encode the -th treatment group as a vector ∈ with

$$\begin{cases} (\quad 1) \ / \quad 1 \\ (1 \quad \sqrt{\ }) / (\quad 1) \ / \quad (——) \ / \quad 2 \ 3 \end{cases}$$

where is a vector of ones of length 1, and ∈ is a vector with the -th element equal to one, and zero elsewhere. Speci cally, when 2, we have 1 and 1, which corresponds to the standard coding procedure in the binary classi cation problem. In addition, based on this coding procedure, one can check that, this treatment group simplex is symmetric with all vertices share an equal distance from each other in . We refer Zhang and Liu (2014) for more details about the angle-based classi cation method. The RAMSVM-based loss consists of a convex combination of two loss functions

$$( ) \quad ( ) : \quad (1 \quad) \qquad (1 \quad \langle \qquad ( )\rangle) \qquad (\qquad 1 \quad \langle \quad _{( )} \quad ( )\rangle) \tag{8}$$

where ∈ [0 1]. The nal group-based decision rule is obtained from

$$( ) \quad \arg\max_{\in[\quad]} \langle \qquad ( )\rangle$$

The corresponding optimization problem is

$$\min_{\in} \left\{ \quad \frac{}{(\ (\ )| \ )} \quad ( ) \quad ( ) \qquad \frac{}{(\ (\ )| \ )} \quad \| \ \| \right\} \tag{9}$$

where is a pre-speci ed function class of { : → }, is a tuning parameter, and ‖ ‖ is the functional penalty associated with to overcome over tting.

We introduce e cient algorithms to solve the optimization problem (9). To this end, we follow the procedure proposed by Liu et al. (2018) to replace with the residual ( ). The rational is that removing the main e ect that is independent of treatment should not a ect the treatment decision while using residuals can signi cantly reduce the variability of weights to improve algorithm performance. However, note that the residual ( ) can take negative values, which would break the convexity of the minimization problem. In this case, we can switch the treatment group to other di erent treatment groups under the uniform sampling procedure. Speci cally, it can be checked that, for any xed , the following two optimization problems are equivalent:

$$\begin{aligned} \min_{\boldsymbol{f}} \quad & \frac{(\quad)}{(\ (\ )| \ )} \quad (\ (\ ) \quad (\ )) \iff \min_{\boldsymbol{f}} \quad \frac{(\qquad (\ ))}{(\ (\ )| \ )} \quad (\ (\ ) \quad (\ )) \\ & (\qquad 1) \quad \frac{(\qquad (\ ))}{(\ (\ )| \ )} \quad (\tilde{}(\ ) \quad (\ )) \end{aligned} \tag{10}$$

where max(0) max(0), and the conditional distribution of the random variable ˜( ) is determined by (˜( ) | ( )) —— for ( ) and 0 for ( ). In this way, the weight term can be easily computed by $\frac{(\qquad (\ ))}{(\ (\ )| \ )}$ (1)$\frac{(\qquad (\ ))}{(\ (\ )| \ )}$. For simplicity of notations, in the following of this section, we use to denote (( )) (1)(( )) . The derivations of why the two optimization problems in (10) are equivalent can be seen in Appendix C.

Next we specify the decision function : → in a product Reproducing Kernel Hilbert Space (RKHS) ⊗ . We develop e cient algorithms to solve (9) after

replacing with and switching treatments for observations with negative residuals. Our implementation consists of two steps. Step 1: under any xed partition candidate , we convert the RAMSVM-based weighted classi cation problem (9) to a dual quadratic programming problem with box constraints. Then we solve the dual problem using coordinate descent algorithm to obtain the estimated optimal decision function under , denoted as $\hat{}$; Step 2: Treatment partition estimation step: after plugging ($\hat{}$) back into (9) to get the value (smaller is preferred) for the candidate , we propose to use the genetics algorithm (Goldberg and Holland, 1988), which is a stochastic search and evolutionary algorithm to obtain the optimal $\hat{}$. Alternatively, we can also use the coordinate descent type of greedy algorithm to adjust the partition.

For step 1, we propose the following algorithm to solve the weighted classi cation problem when specifying in the product linear space or product RKHS. Speci cally, let $\frac{}{(\ (\ \ )|\ \ )}$ be the weight for subject $\in [\ ]$. For $\in$, denote ( ) (( ) ( )) and represents the -th element of , where $\in [\ ]$ and $\in [\ \ 1]$.

For linear decision functions, we assume ( ) with $\in [\ \ 1]$, where s are our parameters of interest. The penalty term $\|\ \|$. Note that we include the intercepts in to simplify notation. After introducing slack variables for (9) and taking partial derivative of the Lagrangian function with respect to each and slack variables, we can derive the following dual problem with respect to the Lagrangian multiplier and obtain $(\hat{}\ )_{\in[\ ]\ \in[\ \ ]}$ by solving

$$\min_{(\ \ )_{\in[\ ]\ \in[\ \ ]}} \frac{1}{2} \quad (\ ) \quad (\ ) \quad (\ )$$

$$(\ ) \quad (\ ) \quad (\ )$$

$$(\ )(\ \ \ 1) \quad (\ )$$

$$\text{s.t. } 0 \quad (\ [\ \ (\ )]\ \ (1\ \ )\ [\ \ (\ )])\ (\ \in [\ ];\ \in [\ \ ]) \tag{11}$$

Moreover, we can calculate

$$\hat{} \quad \frac{1}{} \quad \hat{} \quad (\ ) \quad (\ ) \quad \hat{} \quad (\ )$$

Note that one can verify that the quadratic optimization function in (11) is strictly convex with respect to each . The constraints in (11) are box constraints. Therefore, (11) can be solved e ciently by the well-known coordinate descent algorithm. Compared with standard Quadratic Programming (QP) algorithms for solving the dual problem, the coordinate descent algorithm can enjoy a faster computational speed and obtain more accurate solutions (Zhang et al., 2016). The nal estimated group-based ITR is obtained by

$\hat{}(\ ) \quad \arg\max_{\in[\quad]}\langle \quad \hat{}(\ )\rangle$, where $\hat{}(\ ) \quad (\hat{}(\ ) \quad \hat{} \quad (\ ))$ and $\hat{}(\ ) \quad \hat{}$ for $\in [\quad 1]$.

To deal with more complicated functions, we generalize the linear approach to obtain a nonlinear decision function in RKHS. To begin with, denote to be the corresponding kernel function and $(\quad ,) \ _{,\in[\ ]}$ to be the gram matrix. We assume is invertible. Denote to be the -th column of . By using the norm in $\otimes$ for the penalty term, i.e., $\| \ \|$, we can represent the decision function as $(\ )$ $(\quad)$ for $\in [\quad 1]$. Here, $(\quad)$ is our kernel product coe cient vector for $\in [\quad 1]$. Similar to the steps in linear case, $(\hat{}\ )_{\in[\ ]\ \in[\ ]}$ can be obtained by solving the following dual problem

$$
\begin{aligned}
(\quad)\min_{\in[\ ]\ \in[\quad]} \ & \frac{1}{2} \qquad (\ ) \quad (\ ) \\
& \qquad\qquad\qquad\qquad\qquad\qquad (\ ) \\
& \qquad (\ ) \quad (\ ) \\
& \qquad\qquad\qquad\qquad (\ ) \\
& \frac{1}{2} \qquad\qquad (\ ) \quad (\ ) \\
& \qquad\qquad\qquad\qquad\qquad (\ ) \\
& \qquad (\ )(\quad 1) \\
& \qquad\qquad\qquad\qquad (\ ) \\
\text{s.t. } 0 & \qquad (\ [\quad (\ )] \quad (1 \quad) \ [\quad (\ )]) \ (\ \in[\ ]; \ \in[\quad]) \qquad (12)
\end{aligned}
$$

Furthermore, we can obtain

$$
\begin{aligned}
\hat{} \quad & \frac{1}{} \qquad\qquad \hat{} \ (\ ) \quad (\ ) \qquad\qquad\qquad \hat{} \\
& \qquad\qquad\qquad\qquad\qquad\qquad (\ ) \\
\hat{} \quad & \frac{1}{} \qquad \hat{} \ (\ ) \quad (\ ) \qquad\qquad\qquad \hat{} \\
& \qquad\qquad\qquad\qquad\qquad (\ )
\end{aligned}
$$

One can check that (12) can be solved in an analogous manner as (11). The nal decision function is obtained from $\hat{}(\ ) \quad \hat{} \qquad \hat{} \quad (\quad)$ for $\in [\quad 1]$. More details about how the original problem (9) is transformed to the dual problems (11) and (12) in step 1 are provided in Appendix C.

For step 2, after we plug $(\quad \hat{})$ back to (9), we get the value for the candidate partition . We formulate the partition space as the discrete problem of partitioning numbers into groups. To solve this non-convex integer programming problem, when and are relatively small, we can implement the genetics algorithm using the R package called `GA` introduced in Scrucca (2013). Furthermore, if $\ll$ and both and are large, then the total number of partitions can be very large. Consequently, the genetics

algorithm can be time consuming. Hence, to deal with this case, we propose a coordinate descent type of greedy algorithm to search for the optimal partition iteratively. Speci cally, at each iteration, we minimize (9) by successively adjusting the group assignment for one speci c treatment while holding the assignment of other treatments xed. We go through each treatment in a cyclic fashion until convergence. The initial partition can be obtained via clustering the tted conditional expected outcome for each treatment. The conditional expected outcome can be roughly estimated by penalized regression (Qian and Murphy, 2011), random forest or latent supervised clustering using the pairwise fusion penalty (Chen et al., 2021).

Our analysis so far treats the group number as given. However, is typically unknown in practice. We propose the following e ective procedure to determine . We rst randomly split the observed data $\{(\quad)\}$ into two folds. For each group number $1 \quad$, denote the $\hat{}$ and $\hat{}$ as the estimated optimal partition and associated group-based decision rule learned from one fold of the training data based on the implementations discussed in Section 2.3. Then, we calculate the estimated value function $\hat{}(\hat{} \quad \hat{} \quad)$ for each using

$$\hat{}(\hat{} \quad \hat{} \quad) \quad \frac{[\hat{} \quad (\quad) \quad \hat{} \quad (\quad)]/ \ (\hat{} \quad (\quad)| \quad)}{[\hat{} \quad (\quad) \quad \hat{} \quad (\quad)]/ \ (\hat{} \quad (\quad)| \quad)} \tag{13}$$

where denotes the empirical mean of the other fold of observed data. Note that when 1, all the treatments are grouped together and the associated $\hat{}(\hat{} \quad \hat{} \quad)$ corresponds to the value when we randomly recommend treatments. Thus, we can obtain the estimated bene t function of the estimated group-structured ITR for each with

$$\hat{} \quad (\quad) \quad \hat{}(\hat{} \quad \hat{} \quad) \quad \hat{}(\hat{} \quad \hat{} \quad)$$
$$\hat{}(\hat{} \quad \hat{} \quad) \quad [\quad]$$

We replicate the above process times. For each replication $1 \ 2 \quad$, denote the bene t function as $\hat{}^{(\ )}(\quad)$. We propose the following procedure that can be interpreted as the trade-o between the bene t and the variability of the estimated group-structured ITR to determine the optimal :

$$\hat{} \quad \arg\max \left\{ \texttt{mean} \ \{ \hat{}^{(\ )}(\quad)\} \quad \right\} \tag{14}$$

One can also replace with $\hat{}(\quad)$ in (13) to remove variability coming from estimating the main e ect.

The group number estimator (14) can be interpreted as follows. Denote as the optimal partition when the group number is . For $1 \quad$, let $(\quad)$: $\max_{\in[\ ]} [\ | \ \in \quad]$ be the maximum bene t when the group number is speci ed as and : $\max_{\in} [\ | \quad] \quad (\quad)$ be the optimal bene t. The optimal bene t corresponds to the case that we do not consider any

group structure in the treatment space. We rst consider the case that the treatments have homogeneous group structure discussed in Section 2.2 and the true value of group number equals to . Then, similar to the proof of Lemma 2 shown in Appendix C, one can check that if setting , then the optimal partition de ned in (7) would result in over identi ed group structures. These over identi ed optimal group structures can be any re nement of . In particular, these re ned optimal partitions s of all lead to the same optimal bene t ( ) when . In this case, the bias of the value function is 0. However, the stochastic error bound and the convergence rate of the estimated value function shown in Theorem 6 in Section 3.3 increases with a polynomial rate ( ) as becomes larger. This demonstrates that as increases, the variability of the group-structured ITR becomes larger. Based on Xia et al. (2009), this variability is involved in (14) by using times of sample splitting. Therefore, the group number selection procedure (14) incorporates the penalization of variability to avoid the over identi ed group structures when for the homogeneous case.

When the homogeneous case does not hold, then our optimal bene t ( ) may be strictly less than the optimal bene t for all 1 . One can check that ( ) is a non-decreasing function as increases by induction, and nally equals to when . For the non-homogeneous case, together with the same analysis of the increasing variability as increases, the selection of group number can be interpreted as a trade-o between the bene t and the variability of the group-structured ITR.

$\hat{}$

$\hat{}$

$\hat{}$

In this section, we establish Fisher consistency of the optimal partition and associated group-based ITR for GROWL. We further obtain an excess risk bound and derive the convergence rate for the value function with the diverging group number .

Denote ( ) as the optimal partition and associated optimal decision function under the generalized risk function $\tilde{}$ ( ):

$$(\quad) \in \underset{\rightarrow}{\arg\min} \left\{ \tilde{} (\quad) : \quad (\quad) \quad \frac{}{(\ (\ )| \ )} \right\} \tag{15}$$

where the risk function ( ) is de ned as

$$( ) : \quad \overline{(( )|)} \quad ( ) \quad ( ) \tag{16}$$

Let be the optimal decision function for any xed : arg min $\to$ $\tilde{}$ ( ) Under the angle-based weighted classi cation framework, a classi er is said to be Fisher consistent if for each partition and $\in$, the predicted treatment group has the maximum conditional group treatment e ect under :

$$\arg\max_{\in[ ]}\langle \quad ( )\rangle \quad \arg\max_{\in[ ]} [\ | \ \in \quad]$$

For our problem, we establish the generalized Fisher consistency results for both partition and the decision rule under the group domain if we choose to be the surrogate loss function, i.e., the derived optimal decision rule is the same as the one using the 0-1 loss. In particular, the following generalized Fisher consistency holds:

$$( ) \qquad \qquad ( ) \qquad \in$$
$$\arg\max_{\in[ ]} [\ | \ \in \quad] \quad \in [0 \ -] \qquad \in$$
$$\arg\max_{\in[ ]}\langle \quad ( )\rangle \quad ( )$$

Next we establish the excess risk of 0-1 loss can be upper bounded by that of RAMSVM loss. To start with, we introduce the following notations. For any group-based ITR , there exists a decision function : $\to$ such that ( ) $\arg\max_{\in[ ]}\langle \quad ( )\rangle$. Similar to the de nition of ( ), we de ne

$$( ) \quad [\arg\max_{\in[ ]}\langle \quad ( )\rangle \quad ( )]\overline{(( )|)}$$

and denote $\tilde{}$( ) ( ) [/ (( )|)]. Then the generalized Bayesian risk is denoted as $\tilde{}$ inf $\left\{ \tilde{}( )| \ : \ \to \right\}$. In terms of the value function ( ) de ned in (4), we observe that ( ) ( ) $\tilde{}$( ) $\tilde{}$. Note that in GROWL, we replace the 0-1 loss with the loss. Recall we have de ned the generalized risk function $\tilde{}$ ( ) in (15) and (16). Similarly, the in mum of generalized risk function is de ned as $\tilde{}$ inf $\left\{ \tilde{} ( )| \ : \ \to \right\}$. In addition, under any xed , let arg min $\to$ ( ) be the optimal decision function under the group domain. Denote arg min $\to$ ( ).

The following theorem shows the relationship between the generalized excess 0-1 risk $\tilde{}$( ) $\tilde{}$ and generalized excess risk $\tilde{}$ ( ) $\tilde{}$ under some bounded restrictions for .

:

$$\rightarrow \qquad \langle \quad (\ )\rangle \in [\ 1 \quad 1] \qquad \forall \ \in \qquad \forall \ \in [\quad]$$
$$(\quad)$$
$$\tilde{}(\quad) \quad \tilde{} \quad \tilde{} (\quad) \quad \tilde{}$$

Note that Theorem 5 is di erent from Theorem 3.2 in Zhao et al. (2012) in the sense that we consider multiple treatments, and dealing with both partition and the decision function.

De ne the estimated optimal partition $\hat{}$ and group-based decision function $\hat{}$ as

$$(\hat{} \ \hat{}) \in \underset{\boldsymbol{f}\in\otimes \quad \mathcal{H}}{\arg\min} \left\{ \frac{}{(\ (\ )| \ )} \ (\ (\ ) \ (\ )) \quad \frac{}{(\ (\ )| \ )} \quad \| \ \|_{\mathcal{F}} \right\} \tag{17}$$

For a xed partition , denote the optimal estimated group-based decision function as

$$\hat{} \quad \underset{\boldsymbol{f}\in\otimes \quad \mathcal{H}}{\arg\min} \left\{ \frac{}{(\ (\ )| \ )} \ (\ (\ ) \ (\ )) \quad \frac{}{(\ (\ )| \ )} \quad \| \ \|_{\mathcal{F}} \right\} \tag{18}$$

Speci cally, for the decision function class, we restrict our consideration to the product RKHS associated with Radial Basis Function (RBF) kernels:

$$(\quad ') \quad \exp(\quad \| \quad \|) \quad ' \in$$

where 0 is a bandwidth parameter varying with . For theoretical convenience, we assume $\hat{}$ satis es the extra bounded constraint $\langle \quad \hat{} (\ )\rangle \in [\ 1 \quad 1]$ for $\forall \ \in$ and $\forall \ \in [\quad]$. This constraint does not show up in the algorithm discussed in Section 2.3 because it makes the computation algorithm in Section 2.3 become more complicate and ine cient. Our numerical experience suggests that removing the constraint for $\hat{}$ can yield better classi cation performance than including it.

Next we show that $\tilde{}(\hat{} \ \hat{})$ converges to $\tilde{}$ and equivalently, the value function $(\hat{} \ \hat{})$ converges to $(\quad)$ where the estimated group-based ITR $\hat{} \ (\ ) : \arg\max_{\in[\ ]}\langle \quad \hat{} (\ )\rangle$. We start with introducing the following quantity:

$$(\ ) \quad \inf_{\in\otimes} \quad \left\{ \ \| \ \| \quad \tilde{} (\quad) \quad \tilde{} (\quad) \right\}$$

For a xed , the term $(\ )$ describes how well the regularized RAMSVM-risk approximates the optimal RAMSVM-risk in the RKHS. This quantity is often referred as the approximation error term (Steinwart and Scovel, 2007). Speci cally, when 2, Steinwart and Scovel (2007) proposed a geometric noise assumption to upper bound $(\ )$ in the context of hinge loss based SVM classi cation problem under any xed . In this paper, we generalize the geometric noise assumption so that we can upper bound $(\ )$ for the multicategory group-based ITR problem under the RAMSVM-based loss. Under each , denote the di erence of two group treatment e ects as

$$(\ ) \quad [\ | \ \in \quad] \quad [\ | \ \in \quad] \quad \text{and} \quad \in [\quad]$$

De ne the decision regions for each pair of treatment groups ∈ [ ] to be { ∈ | ( ) 0} and { ∈ | ( ) 0} Then let [for ∈ [ ] be the subset of where the treatment e ect of group dominates any other treatment groups under partition . Denote the function ( ) [∈] sup | ( )| as the maximum di erence of the group treatment e ects for each region . Furthermore, denote the following distance function to the decision boundary as ( ) [∈] inf dist( ), where dist( ) is the distance between a point and a set . Then we de ne the following generalized geometric noise assumption:

∈ [ ]

0 0

(( )) | ( )| /

One can check that when 2, Assumption 1 is consistent with De nition 2.3 in Steinwart and Scovel (2007) and De nition 3.8 in Zhou et al. (2017). In some sense, this geometric noise exponent describes the concentration of the measure | ( )| near the decision boundary. In the case of complete separation, i.e., ( ) 0 for some constant , can be as large as possible.

Let be the total number of partitions. Recall the de nition of ( ) and in (6) and (7). Consider can diverge to in nity as the sample size increases. Then, for any , there exist a positive gap : – inf ∉ { ( ) ( )} 0, such that ˜( ) ˜ ( ) ( ) holds for any non-optimal partition ∉ . Here, can be interpreted as the signal to characterize the minimum distance of the value function between the optimal partitions and any other non-optimal partitions. Intuitively, we need the signal to be large enough so that we can distinguish the optimal partitions from non-optimal partitions. Now we are ready to present the main theorem for the convergence rate of GROWL.

ˆ

∈

0 ∞ | (( )|)|

( ) { } ( )

/ / / / / ( ) — ( )( ) — —

( )( )

/log → 0 → ∞ / → ∞ 0 0 2

/( )

0 0 2

(ˆ ∈) 1 ( );

˜(ˆ ˆ) ˜ ( )

For Theorem 6, we can choose ——, and let ( ) be su ciently small. When data are well separated under one of the optimal partitions, can be su ciently large. Note that if the group-based propensity score (( )|) has balanced structure under , then as → ∞, (( )|) would decay uniformly. Thus, we have |———————| ( ). In this case, the convergence rate for the value function can achieve ( ) (̂ ̂) ̃(̂ ̂) ̃ (¯)

Note that in Theorem 6 describes the rate of the stochastic error bound for the regret. Thus, the assumption / → ∞ implies that, the signal should be large enough and dominates the noise so that ̂ can nally belong to . To better under illustrate this assumption, we consider the following simple example. Suppose the number of treatment groups is xed. Then, under the homogeneous case, the best partition among the partition set without the optimal partitions corresponds to the case that only one treatment is misclustered. In this case, one can check that the de ned signal term (1/) based on the proof of Lemma 2. Hence, due to (log), the assumption / → ∞ is satis ed since decays at a polynomial rate of .

The pipeline for proving Theorem 6 is stated as the follow two steps: First, for any optimal partition ∈ , we establish a nite sample bound for the di erence between the expected outcome using the estimated group-based decision function ̂ based on the training data and that of the optimal group-based decision function under ; Second, due to / → ∞, as goes to in nity, the stochastic error of ̃( ) arising from using to estimate in (17) would be dominated by the gap for any ∉ . Hence, ̂ in (17) would nally belong to when is su cient large since ∈ maximizes the value function . Then, the convergence rate is determined by the rate of (̂ ∈) → 1 and the convergence rate of the rst step when treating the partition is xed as ∈ . Speci cally, the novelty of our technical proof arises from bounding the approximation bias term ( ) with order (/( )) and deriving the nite value reduction bound in a multicategory setting. The intermediate results deriving from the rst step generalize Lemma 3.9 in Zhou et al. (2017) and Theorem 3.4 in Zhao et al. (2012) from binary treatments to multiple treatment groups that may diverge to in nity as the sample size increases. More details are provided in Appendix B.

For the case that the number of treatments and treatment groups are xed, it is straightforward to derive the following Corollary 7 from Theorem 6. Note that for the xed group number case, / → ∞ is trivially satis ed since is a constant.

̂

0 ∞ ∈ |———————| ( )

→ ∞ /( )

0 0 2

(̂ ∈) 1 ( ) / ( ) — ( )( ) ( )( ) — — ;

$\tilde{}(\hat{}\ \hat{}) \quad \tilde{} \qquad (\quad) \ / \quad (\quad) \ — \ \frac{(\quad)(\quad)}{(\quad)(\quad)} \quad — \qquad —$

We evaluate the nite-sample performance of our proposed method using several simulation studies.

In this simulation study, we consider the setting where the treatment responses for the treatments in the same group are equivalent, but di er for the treatments across di erent groups. We generate 10-dimensional independent prognostic variables , following [1 1]. The outcome is normally distributed with [|] 1 2 0 5 ( ) and standard deviation 1, where ( ) re ects the interaction between the treatment and the prognostic variables. In addition, we assume that the treatment e ects have the homogeneous grouping structure induced by discussed in Section 2.2. Speci cally, we consider the following three scenarios:

10 2 $\{1\ 2\ 3\ 4\ 5\}\ \{6\ 7\ 8\ 9\ 10\}\}$ and ( ) 1 8 0 2 [$\in \{1\ 2\ 3\ 4\ 5\}$] (1) [$\in \{6\ 7\ 8\ 9\ 10\}$] 1 ;

10 2 $\{1\ 2\ 3\ 4\ 5\}\ \{6\ 7\ 8\ 9\ 10\}\}$ and ( ) 3 5 0 8 [$\in \{1\ 2\ 3\ 4\ 5\}$] (1) [$\in \{6\ 7\ 8\ 9\ 10\}$] 1 ;

15 3 $\{1\ 2\ 3\ 4\ 5\}\ \{6\ 7\ 8\ 9\ 10\}\ \{11\ 12\ 13\ 14\ 15\}\}$ and ( ) 5 (0 2 2) [$\in \{1\ 2\ 3\ 4\ 5\}$] (0 3 2) [$\in \{6\ 7\ 8\ 9\ 10\}$] (0 2 3) [$\in \{11\ 12\ 13\ 14\ 15\}$] .

Scenario 1 corresponds to 10 treatment arms belonging to two treatment groups with underlying linear decision boundaries whereas Scenario 2 considers the circle decision boundary. Scenario 3 includes 15 treatments compared with the rst two and deals with three treatment groups with linear decision boundary. Since our studies are especially interested in the case that the propensity score of some speci c treatments may be very small, we perform the following two designs varying from balanced to unbalanced designs within each scenario:

(a) Balanced Design: (|) — for each $\in$ [ ] and each $\in$;

(b) Unbalanced Design: The value of (|) for some speci c treatments can be very small compared with other treatments.

Under the unbalanced design, for each $\in$, the propensity scores for the rst two scenarios are set to be (— — — — -) (— — — — -) while the propensity scores equal to (— — — — —) (— — — — —) (— — — — —) for Scenario 3. In addition, we conduct more simulation settings when the propensity scores are more unbalanced and may depend on the covariates. These additional results are shown in Appendix D. For GROWL, we use the linear kernel for Scenarios 1 and 3 and utilize the Gaussian kernel for Scenario 2 corresponding to di erent shapes of the decision boundary. The tuning parameter is chosen to maximize the empirical value function [$\hat{}$( )]/ (|) / [$\hat{}$( )]/ (|) by 10-fold cross-validation among {— - - - 1 2 4 8 16}. For the Gaussian kernel, we x the inverse bandwidth of the kernel with $1/(2\hat{})$, where $\hat{}$ is the median of

the pairwise Euclidean distance of the covariates (Wu and Liu, 2007). The treatment group number is determined by the trade-o procedure (14) with 50 shown in Section 2.4.

The following four methods are compared under each scenario:

(a) SL: Super Learner based Q-learning method to estimate [|] (Polley and Van Der Laan, 2010);

(b) AD: Multi-armed Angle-based Direct learning using linear terms for Scenarios 1 and 3 and polynomial terms for Scenario 2 (Qi et al., 2020);

(c) PLS: -Penalized Least Squares method (Qian and Murphy, 2011), which estimates [|] using the basis sets (1) for Scenarios 1 and 3, and (1) for Scenario 2;

(d) GROWL: Our proposed method.

SL aims to nd the optimal combination of multiple estimated Q-functions by minimizing the cross-validated risk. For the collection of learning algorithms, we include ridge regression (Hoerl and Kennard, 1970), elastic net (Zou and Hastie, 2005), random forest (Breiman, 2001), XGBoosting (Chen and Guestrin, 2016), and neural network (Venables and Ripley, 2013). We refer Polley and Van Der Laan (2010) for more implementation details about SL. In addition, the tuning parameters in AD and PLS are selected to maximize [̂( )]/ (|) / [̂( )]/ (|) by 10-fold cross-validation.

We evaluate the above methods using the empirical value function and the group-based misclassi cation rate between the estimated group decision rules and the true group decision rules on an independently generated testing data of size 10 000. The empirical value function is calculated by the mean of treatment e ects under the empirical distribution of based on the estimated decision rule. Note that under the homogeneous setting, the maximum group-based treatment e ect equals to the maximum individual treatment based e ect. Hence, the misclassi cation rate under the group domain is equivalent to the misclassi cation rate under the individual treatment domain. For each scenario, the training sample sizes vary from 200, 400 to 600 and we replicate the simulations for 200 times.

We present the empirical value function of each scenario under di erent designs using boxplots in Figure 1. Results of group-based misclassi cation rates are included in Figure 5 of Appendix D. We also report the square root of Mean Square Error (MSE) of the empirical value function and the misclassi cation rate in Table 1. Based on Lemma 2 and Theorem 4, we have shown that should equal to the optimal partition de ned in (7), and should equal to the optimal partition derived from (15). Accordingly, the Ratio column in Table 1 reports the ratio of our estimated ̂ exactly being among the 200 replications. It can be seen that ̂ , the estimation of , converges to with a high ratio as becomes larger, which con rms part (I) of Theorem 6. In general, as the trial design becomes more unbalanced, all these methods perform worse, in the sense that MSE becomes larger for each scenario. Without considering the group structure for the treatments, SL, PLS and AD su er from the inaccurate estimation of functions related to individual-treatment e ects because of the small amount of observations for some speci c treatments. However, GROWL estimates the group-structured ITR, which reduces the dimension of the treatment space and clusters the treatments that employ similar treatment e ects into the same group. In addition, since GROWL estimates the ITR in the treatment group domain, the variance of the value function shrinks quicker than other methods as the training sample size increases. As is demonstrated in Figure 1 and Table 1, our method outperforms other methods in most

cases with higher empirical value functions, smaller misclass cation rates, and especially lower variabilities for both evaluation criteria.

Table 1: Results for Ratio of nding the optimal partition and square root of of Empirical Value Function and Misclassi cation Rate evaluated on the independent test data under the settings. The best values are in bold.

	200			400			600		
	Ratio(%)	Value	Misclassi cation	Ratio(%)	Value	Misclassi cation	Ratio(%)	Value	Misclassi cation
					Scenario				
SL		0.109	0.112		0.046	0.069		0.028	0.052
AD		0.132	0.126		0.056	0.098		0.044	0.086
PLS		0.070	0.092		0.041	0.065		0.030	0.057
GROWL	97.0			99.5			100		
					Scenario				
SL		0.181	0.167		0.068	0.097		0.052	0.078
AD		0.106	0.133		0.073	0.111		0.069	0.107
PLS		0.086	0.112		0.047	0.079		0.041	0.069
GROWL	92.5			98.5			99.0		
					Scenario				
SL		0.567	0.260		0.169	0.154		0.119	0.108
AD		1.310	0.463		0.752	0.369		0.565	0.342
PLS					0.199	0.177		0.191	0.170
GROWL	67.5	0.511	0.269	92.0			97.5		
					Scenario				
SL		0.258	0.164		0.073	0.088		0.046	0.068
AD		0.416	0.216		0.270	0.145		0.075	0.101
PLS		0.146	0.129		0.045	0.067		0.032	0.058
GROWL	91.5			99.0			99.0		
					Scenario				
SL		0.255	0.199		0.127	0.139		0.082	0.102
AD		0.246	0.197		0.123	0.138		0.089	0.120
PLS		0.166	0.148		0.072	0.093		0.047	0.075
GROWL	83.5			97.5			98.0		
					Scenario				
SL		0.675	0.288		0.183	0.162		0.121	0.112
AD		1.484	0.481		0.814	0.386		0.664	0.355
PLS					0.231	0.184		0.198	0.174
GROWL	56.0	0.570	0.281	83.5			98.0		

In many cases, it is possible that the treatment e ects do not have exactly homogeneous grouping structure assumed in Section 4.1. In this section, we perform nearly homogeneous and nonhomogeneous scenarios to examine our method. Speci cally, we generalize Scenario 1 in Section 4.1 with the following Scenario 4 indexed by a parameter 0:
: 10, ( ) 1 8 0 2 [∈{1 2 3 4 5}] (1 −) [∈ {6 7 8 9 10}] (1 ——) .

The parameter determines the level of heterogeneity of treatment e ects. As becomes smaller, the treatment e ects become more diverse and thus the group structure

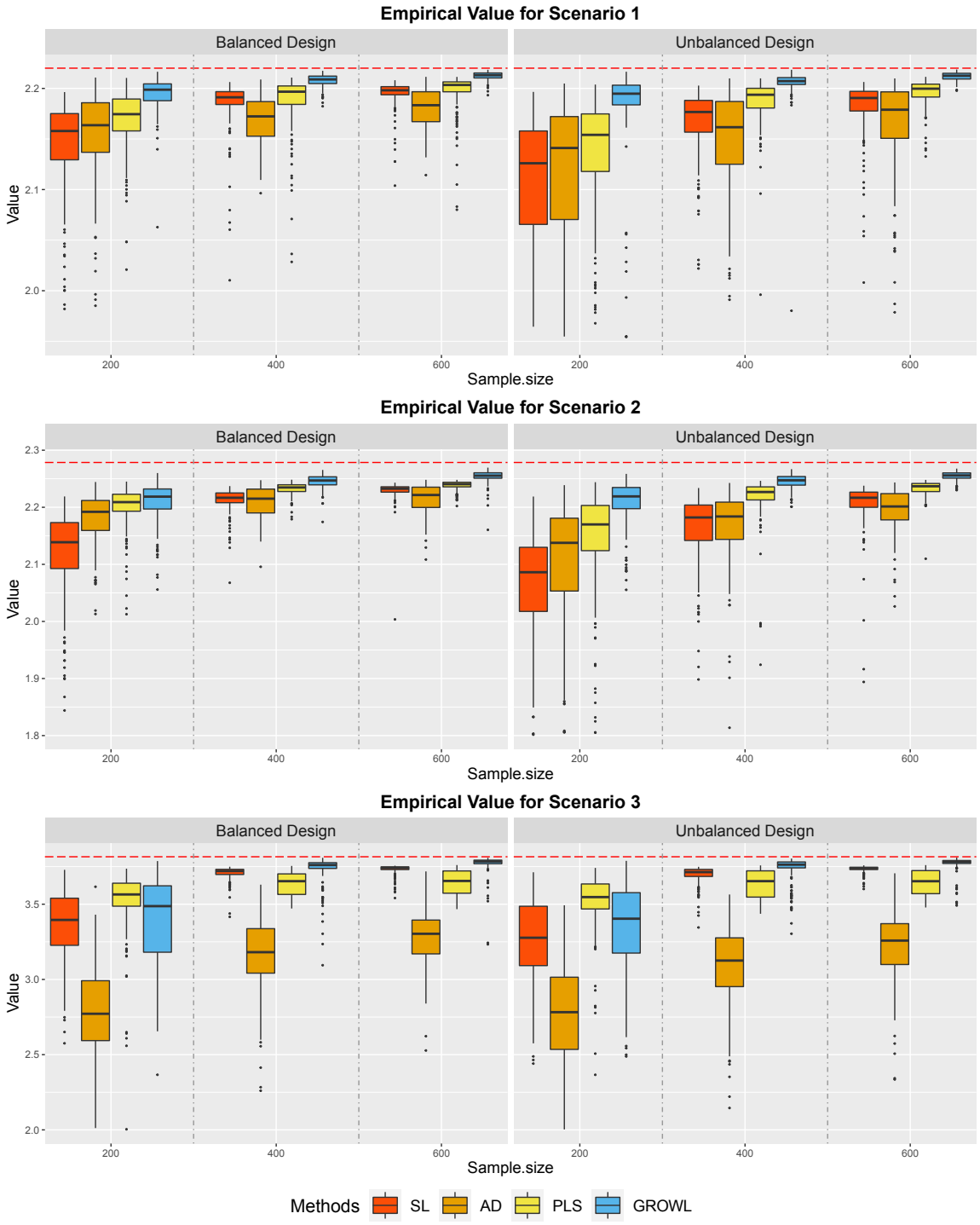

Figure 1: Boxplots of Empirical Value Function evaluated on the independent test data under the settings. Red horizontal dashed lines show oracle values for each scenario.

tends to disappear. When ∞, Scenario 4 has the exact homogeneous structure $\{1\ 2\ 3\ 4\ 5\}\ \{6\ 7\ 8\ 9\ 10\}\big\}$ shown in Scenario 1. We vary from 40 20 10 to 5. In particular, the simulation scenarios in this section only focus on the unbalanced design. For GROWL, the treatment group number is determined by the trade-o procedure in (14). Other simulation settings and comparison methods are the same as those in Section 4.1.

Similar to Section 4.1, we provide boxplots of the empirical value function in Figure 2, and MSE of the empirical value function in Table 2. When decreases from 40 to 5, the treatment e ects vary from nearly homogeneous structure to nonhomogeneous structure. For nearly homogeneous cases with 40 and 20, our method still outperforms other methods while for nonhomogeneous cases with 10 and 5, PLS performs the best. The results are consistent with Remark 3. For each 1 10, the maximum bene t of group-structured ITR ( ) is strictly less than the optimal bene t because the conditional treatment e ects within the same group are di erent under in nonhomogeneous cases. When ∞, these two values are equal. As decreases, the gap between these two values increases. Based on our simulation results, for each and training sample size, over 95% of the replications suggest $\widehat{\ }$ 2 based on the trade-o procedure (14), and over 90% of the estimated partition still equals to the same two-group structure $\{1\ 2\ 3\ 4\ 5\}\ \{6\ 7\ 8\ 9\ 10\}\big\}$ as the homogeneous setting. Hence, in Scenario 4, GROWL tends to sacri ce the bene t while retain small variability of the value function, and the gain of small variability continues to dominate the loss of bene t when decreases from 40 to 5. In Figure 2, one can observe that the empirical value of our method would not converge to the optimal value (shown by dashed lines) and a positive gap would exist. In addition, when sample size increases from 200 to 600, the relative improvement ratio in terms of the MSE for GROWL decreases when has smaller values. However, due to the usage of the group structure, the variability of our method is very small compared with others shown in Figure 2. Therefore, the trade-o between the bene t and variability of the value function for the group-structured ITR estimated by GROWL is clear. From Table 2, GROWL is still competitive in nonhomogeneous settings in terms of the MSE criterion.

In this section, we apply our proposed GROWL to analyze the data from the STAR*D study (Rush et al., 2004). The STAR*D study performed research on outpatients with nonpsychotic major depressive disorder. The goal of the study was to compare various treatment options for the patients who failed to obtain a satisfactory response with citalopram (CIT), an initial antidepressant treatment. The primary outcome was measured by the Quick Inventory of Depressive Symptomatology (QIDS) score ranging from 0 to 27, where higher scores indicate more severe depression.

The STAR*D data consist of four levels. In our analysis, we focus on the 1407 eligible patients who received treatments at Level 2. In particular, at Level 2, patients were asked to indicate their preference of either switching to one of the 4 di erent treatments, i.e., bupropion (BUP), cognitive therapy (CT), sertraline (SER), and venlafaxine (VEN), or augmenting their existing CIT with 3 options, i.e., CIT+BUP, CIT+buspirone (BUS), and CIT+CT. If a patient indicated no preference, then he/she was assigned to any of the above 7 treatments. To encourage the active collaboration and shared decision-making

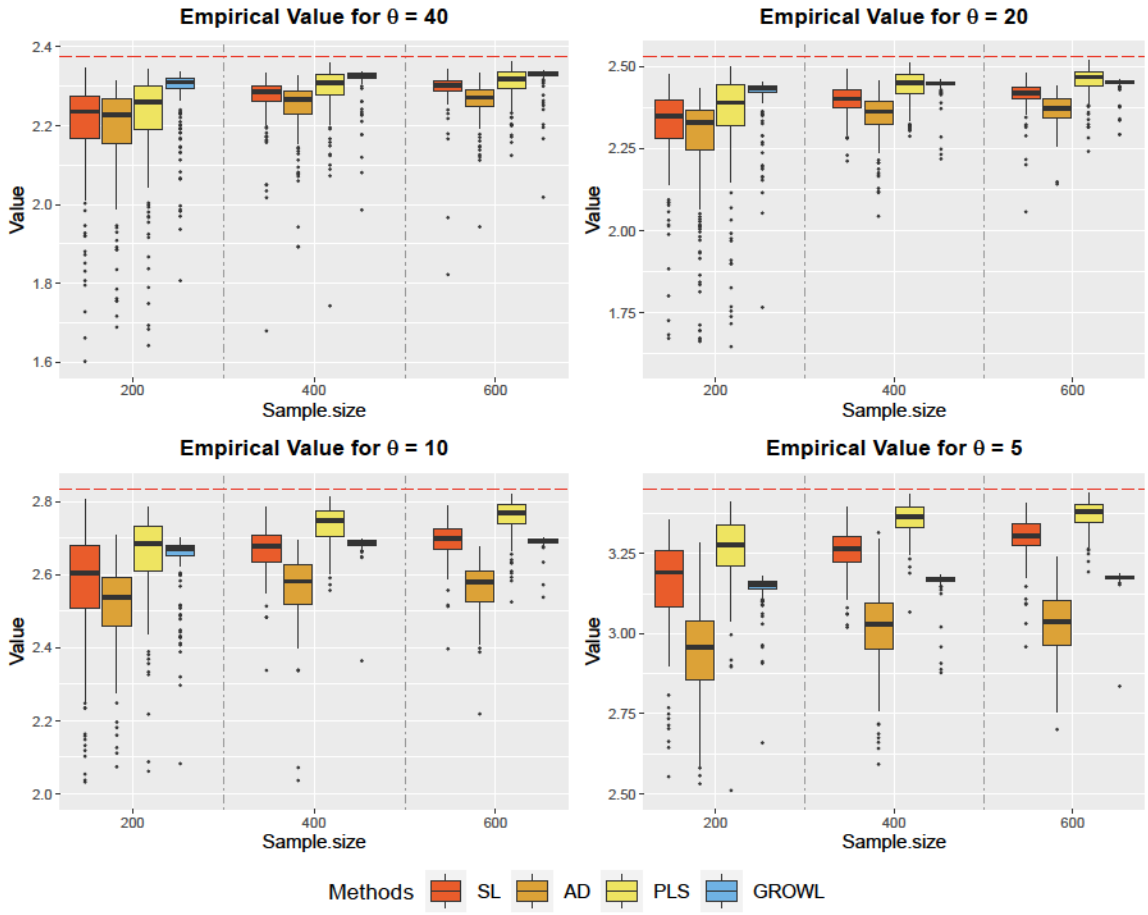

Figure 2: Boxplots of Empirical Value Function evaluated on the independent test data under the **nonhomogeneous** settings and unbalanced design. Red horizontal dashed lines show oracle values for each scenario.

Table 2: Results for square root of of Empirical Value Function evaluated on independent test data under the settings and unbalanced designs. The best values are in bold.

with patients, we consider the patients preference as part of the intervenable treatment options and assume the future patients preference can be intervened when recommending treatments. Hence, we have a total of 3 4 7 14 treatment options and these preference-related treatments are often called patient-centered medications in the literature (Robinson et al., 2008). Figure 3 demonstrates the distribution of the observation numbers for the 14 patient-centered treatments. Due to the relatively large treatment space and the unbalanced structure of propensities of treatments, it can be seen that only a few observations were obtained for many treatment options, especially for the treatments in the no-preference (Nop-) group.

We apply four methods (SL, PLS, AD, GROWL) discussed in Section 4 to estimate the optimal treatment rules among the 14 treatment options for patients. Speci cally, the reward in our study is calculated by the reduction of QIDS score from the start to the end of Level 2. Hence, a higher value of is preferred. Feature variables include QIDS score at the start of Level 2, reduction of QIDS score during Level 1, and other demographic variables such as gender, race, age, education level, employment status, and marital status, etc. The estimated propensities $\hat{}(\ |\ )$ are obtained from tting the multinomial logistic regression model. For PLS, we use terms (1) to t the penalized linear regression to estimate conditional treatment e ects; For AD, polynomial terms are also included as the basis set of decision functions; For GROWL, we implement the Gaussian kernel for decision functions. Comparisons of all these methods were based on 200 repetitions of three-fold cross-validation, where two folds are used to train the model. For our proposed GROWL, we

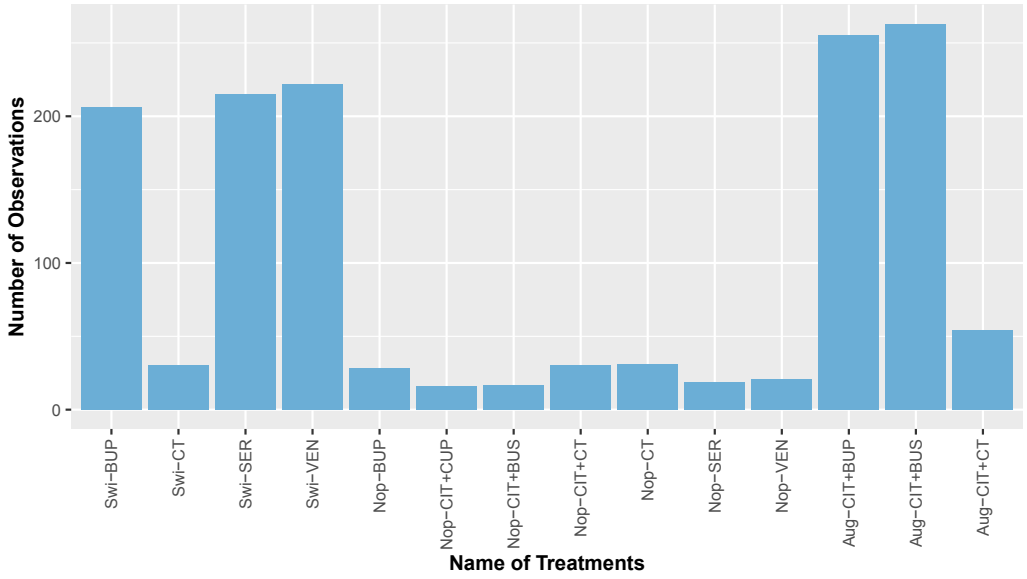

Figure 3: Distribution of observations for the 14 treatment options in the STAR*D dataset. Swi , Aug and Nop correspond to patients that switch to other treatments, augment existing treatments, and have no preference for two previous options.

follow equation (14) discussed in Section 2.4 to determine the number of groups with training data. We evaluate the four methods on the remaining one fold of testing data based on the empirical value function [̂( )]/ ̂(|) / [̂( )]/ ̂(|) , where denotes the empirical average of testing data.

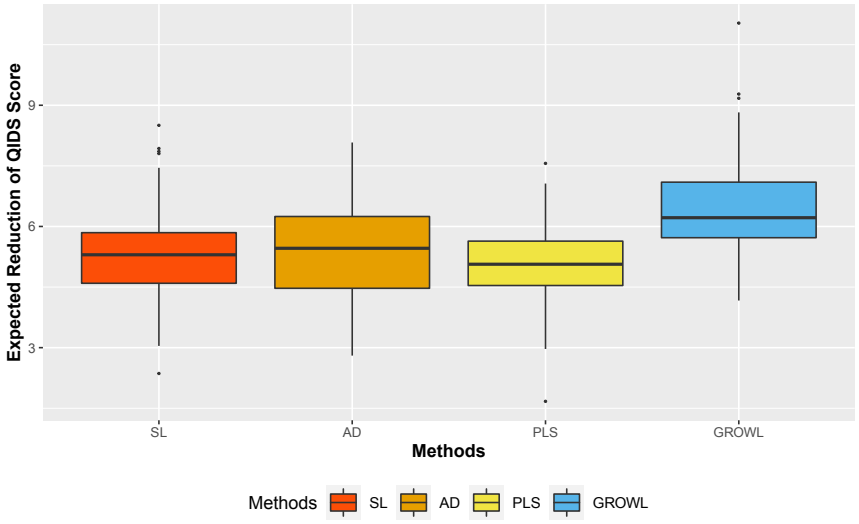

Figure 4: Boxplots of Expected Reduction of QIDS score during Level 2 for patients in testing data based on 200 repetitions for the STAR*D study (higher value is better).

The testing results are shown in Figure 4. The means of expected reduction of QIDS score during Level 2 by using GROWL is 6 44, which outperform the mean value of SL (5 28), AD (5 32) and PLS (5 04). Thus, compared with methods without considering the treatment partition, GROWL substantially improves the performance of the optimal ITR estimation. The estimated group numbers are 5 or 6 for most of the repetitions. Among the 200 estimated partitions of the treatment space, the patient-centered treatments containing SER, CIT+BUP, CIT+BUS, and CIT+CT are often combined within one group and the treatments containing BUP, CT and VEN are integrated with another group with high frequency. It is interesting to point out that, the treatments SER, CIT+BUP, CIT+BUS, and CIT+CT are often considered as one class of treatments including Selective Serotonin Reuptake Inhibitors (SSRI) while the treatments BUP, CT and VEN are non-SSRI treatments because the treatments within the same group have similar treatment e ects (Liu et al., 2018; Pan and Zhao, 2021). In addition, the patient-centered treatments with the same patients preference are often clustered into the same group. With the overall dataset, we implement the GROWL with the Gaussian kernel and obtain the nal estimated group structure with 5 treatment groups:

$$\hat{} \quad \left\{ \begin{array}{l} \{\text{Swi-BUP, Swi-VEN, Nop-VEN}\} \ \{\text{Swi-SER}\} \\ \{\text{Aug-CIT+BUP, Aug-CIT+BUS, Aug-CIT+CT}\} \\ \{\text{Swi-CT, Nop-CIT+CT, Nop-CT, Nop-VEN}\} \\ \{\text{Nop-CIT+BUP, Nop-CIT+BUS, Nop-SER}\} \end{array} \right\}$$

It can be seen that these patient-centered treatments with the same preference work similarly within the SSRI groups and the non-SSRI groups respectively.

To better interpret the decision rule and examine the e ects of the feature variables, we implement our GROWL with linear kernel based on $\hat{}$. For the STAR*D dataset, the mean of expected reduction of QIDS score for GROWL with linear kernel is 6 11, demonstrating that GROWL with the linear kernel still outperforms other methods. Note that we have 13 feature variables (including the intercept) and 5 treatment groups. Therefore, we obtain a 4 13 estimated coe cient matrix $\hat{}$ for the linear decision function. The -th column of Table 3 in Appendix D demonstrates $\hat{}$ for the -th treatment group where 1 2 5. We can see that nearly all feature variables play an important role in the estimated optimal ITR. In particular, for the important biomarker, the QIDS score, the patients with higher QIDS score reduction during Level 1 are suggested with augmenting the current CIT treatment implemented at Level 1, while patients with low QIDS reduction are recommended with switching to other treatments at Level 2.

In this article, we propose a new method called GROWL to cluster treatment options and at the same time, estimate the optimal group-structured ITR within one RAMSVM-based objective function. Other comparison methods estimate the ITR under the individual treatment domain while GROWL focuses on the group treatment domain. When homogeneous

or nearly-homogeneous treatment group structure is satis ed for the treatments, GROWL is able to nd the expected partition with high accuracy and the superior performance of GROWL is demonstrated in our numerical studies. Under heterogeneous settings when is small shown in Section 4.2, our method tends to sacri ce the bene t while reduce the variability signi cantly. In this case, our method is still superior to other methods when the sample size is small. Another advantage of GROWL is that it combines both supervised and unsupervised learning through one single optimization.

From a broad perspective, our method is not limited to ITR problems. It can be viewed as a multicategory classi cation technique. In particular, consider using observed data ( ) ( ) to classify the covariate $\in$ as a speci c class in a large class space . Due to the large number of labels, insu cient data are observed for some speci c labels. Consequently, standard classi cation methods can become ine ective. On the other hand, the conditional probability [|] may employ possible similar patterns for some classes $\in$. Then our method tends to estimate this pattern with group-based structure to reduce the dimension of the classi cation label space and classify the observations in the group domain. Although our paper mainly focuses on estimating the optimal ITR in the decision making framework, the essence is similar because the conditional treatment e ects [|] play a similar role as [|] in classi cation problems.

Several possible extensions can be explored for future studies. First, most scenarios considered in the paper is that the group structure of the treatment e ects is independent with the marginal distribution of feature variables . However, consider the case that homogeneous group structure is completely di erent with di erent values of $\in$, the optimal partition in (7) tends to sacri ce some subgroups of individuals. Consequently, more value would be lost because the optimal partition is obtained via averaging the marginal distribution of . For these more complex scenarios, it will be interesting to estimate di erent partitions targeting subgroups of individuals. Secondly, our method can be extended to learn group structures for multi-stage Dynamic Treatment Regimes (DTR) (Murphy, 2003; Zhao et al., 2015). This can be an interesting direction for future research.

The authors would like to thank the action editor and reviewers, whose helpful comments and suggestions led to a much improved presentation. This research was partially supported by NSF grants DMS 2100729, SES 2217440; and NIH grants GM126550, GM124104, and MH123487.

The appendix contains detailed proofs of the main results, additional gures and tables, and more implementation details. In Appendix A, we introduce some useful lemmas to prove the theorems in the main paper. In Appendix B, we give the detailed proofs of theorems in the main paper. In Appendix C, we prove the lemmas in the main paper and the lemmas in Appendix A. We also provide proofs of equation (10) and how the dual problems (11) and (12) are derived. Additional results of more simulations and real data analysis are given in Appendix D. Finally, implementation details are summarized in Appendix E.

: → ⟨ ( )⟩ ∈ [1 1] ∀ ∈
∀ ∈ [ ]
( ) ( ) ( ) ( )

Note that Lemma 8 generalizes the excess risk bound in Theorem 3.2 in Zhao et al. (2012) from the binary setting to the multi-class setting for any xed .

0 ∞
/( )
0
0
( ) /( )

∈ 0 ∞
|———————|
(( )|)
0 0 2 1
/( )

˜(ˆ) ˜ 1

( ) — ( )( ) — —— —
( )( )

Note that Lemma 10 is a generalization of Theorem 3.4 in Zhao et al. (2012) to deal with the multiclass ITR problem under the -based loss.

Based on the de nition of loss, ( ) equals to

$$\frac{[\ |\ ]}{(\ (\ )|\ )}\ [(\ )\ (\ \langle\ (\ )\rangle)\ (\ \langle\ _{(\ )}\ (\ )\rangle)\ ]$$

$$_{(\ )}\ \frac{[\ |\ ]\ (\ |\ )}{(\ (\ )|\ )}\ [(\ )\ (\ \langle\ (\ )\rangle)\ (\ \langle\ (\ )\rangle)\ ]$$

$$[\ |\ \in\ ]\ [(\ )\ (\ \langle\ (\ )\rangle)\ (\ \langle\ (\ )\rangle)\ ]$$

For each xed and xed $\in$, denote $(\ )$ $\arg\max_{\in[\ ]}[\ |\ \in\ ]$. Then based on the proof of Theorem 1 in Zhang et al. (2016), when $\in [0\ 1/2]$, we have $\langle\ _{(\ )}\ (\ )\rangle$ 1 and $\langle\ (\ )\rangle$ 1 for $(\ )$. Hence, classical Fisher Consistency holds for group-based decision rule for each xed . By plugging into above equation, we have

$$\inf_{\rightarrow}\ (\ )\ (\ )$$

$$_{(\ )}\ [\ |\ \in\ ]\ [(1\ )\ ]$$

$$[\ |\ \in\ ]\ \max_{\in[\ ]}[\ |\ \in\ ]$$

$$\frac{}{(\ (\ )|\ )}\ \max_{\in[\ ]}[\ |\ \in\ ]$$

Hence, we have

$$\in \arg\min\left\{\ ^{\sim}\ (\ )\ (\ )\ \frac{}{(\ (\ )|\ )}\right\}$$

$$\arg\max\ \max_{\in[\ ]}[\ |\ \in\ ]$$

Then $\arg\max_{\in[\ ]}\langle\ (\ )\rangle$ $\arg\max_{\in[\ ]}[\ |\ \in\ ]$ follows straightforward. ■

To prove the excess risk bound, we notice the following decomposition:

$$^{\sim}(\ )\ ^{\sim}\ ^{\sim}(\ )\ ^{\sim}(\ )\ ^{\sim}(\ )\ ^{\sim}(\ )$$

$$(\ )\ (\ )\ ^{\sim}(\ )\ ^{\sim}(\ )$$

$$(\ )\ (\ )\ ^{\sim}(\ )\ ^{\sim}(\ )$$

$$(\ )\ (\ )\ ^{\sim}(\ )\ ^{\sim}$$

$$^{\sim}(\ )\ ^{\sim}$$

The first inequality follows from Lemma 8. For the second inequality, notice that

$$\tilde{}(\quad)\quad\tilde{}(\quad)\qquad\tilde{}(\quad)\quad\tilde{}(\quad)\qquad\tilde{}(\quad)\quad\tilde{}$$

Then the proof is complete. ■

For each partition , take the following notations. De ne the random variable $\frac{}{(\ )|\ )}$. Denote $(\ )$ $(\quad)$ as the weighted loss. Similarly, under , let be the population measure of $(\ )\ /\ (\ (\ )|\ )$ and be the associated empirical measure. Recall that $-\min_{\notin}(\ )\quad(\ )\quad 0$. Then, for any $\notin$, we have $\inf\ \tilde{}(\quad)\quad\tilde{}$. To prove part (I) in this theorem, we need to control the probability of event $\left\{\hat{}\ \notin\ \right\}$. Take an optimal $\in$. Note that based on the de nition of $\hat{}$, the following two events are equivalent:

$$\left\{\hat{}\ \in\ \right\}\Leftrightarrow_{\notin}\left\{\quad(\quad(\hat{}))\qquad\hat{}\qquad(\quad)\qquad(\quad(\hat{}))\qquad\hat{}\qquad(\ )\right\}$$

Hence, $\hat{}\ \in$

$$_{\notin}\left\{\quad(\quad(\hat{}))\qquad\hat{}\qquad(\quad)\qquad(\quad(\hat{}))\qquad\hat{}\qquad(\ )\right\}$$

$$_{\notin}\quad(\quad(\hat{}))\qquad\hat{}\qquad(\quad)\qquad(\quad(\hat{}))\qquad\hat{}\qquad(\ )$$

$$_{\notin}\quad(\quad(\hat{}))\qquad\hat{}\qquad(\quad)\qquad(\quad(\hat{}))\qquad\hat{}\qquad(\ )$$

After rearranging both sides, we can only control the following probability for each xed $\notin$ in order to lower bound $\hat{}\ \in$:

$$\begin{aligned}&(\quad(\hat{}))\quad(\quad(\hat{}))\quad(\quad(\quad)\quad(\ ))\\&(\quad(\hat{}))\quad(\quad(\hat{}))\quad(\quad(\quad)\quad(\ ))\quad(\tilde{}\quad\tilde{}(\quad))\quad\tilde{}(\quad)\quad\tilde{}\\&(\quad(\hat{}))\quad(\quad(\hat{}))\quad(\quad(\quad)\quad(\ ))\quad(\tilde{}\quad\tilde{}(\quad))\\&[\qquad\qquad\qquad]\\&[\qquad\qquad\quad]\end{aligned}$$

where

$$\begin{aligned}&(\qquad)(\quad(\hat{}))\quad(\qquad)(\quad(\hat{}))\\&(\qquad)(\quad)\qquad(\qquad)(\quad)\\&(\quad(\hat{}\qquad))\\&(\quad(\hat{}\qquad))\end{aligned}$$

Note that the last inequality above is due to 0. Here, after removing , the advantage is that we only need the generalized geometric noise condition holds for to bound following from Lemma 10.

Now, we would bound the above three terms respectively using concentration techniques. First, to bound , we start with obtaining a bound for $\hat{}$. Since $\hat{}$ ($\hat{}$ ($\hat{}$) $\hat{}$) $\hat{}$ (( )) $\| \|$ holds for any and $\in \otimes$, we can select $\hat{}$ and to get

$$\hat{} \quad \frac{1}{} \frac{1}{} \quad (\hat{}) \quad \frac{2(\quad 1)}{}$$

so that only the constrained class $\left\{ \sqrt{} \in \otimes : \sqrt{} \sqrt{2(\quad 1)} \right\}$ would be considered. Note that RAMSVM loss is Lipschitz continuous with respect to with Lipschitz constant . Then, for each , McDiarmid s inequality (Bartlett and Mendelson, 2002) implies that with probability at least 1 ,

$$\sup_{\in} (\quad) \quad ( ) \quad \sup_{\in} (\quad) \quad ( ) \quad 2 \quad \sqrt{\frac{2}{}}$$

Next we de ne the Rademacher complexity over as R : sup $_{\in}$ (( )), where is the Rademacher variable independent of (( )) under . Then by standard symmetrization arguments and Lemma 22 in Bartlett and Mendelson (2002), we have

$$\sup_{\in} (\quad) \quad ( ) \quad 2\mathrm{R} \quad 2\sqrt{\frac{2(\quad 1)}{}}$$

Similar proof holds for . Therefore, we have [ ] 1 2 , where

$$4\sqrt{\frac{2(\quad 1)}{}} \quad 4 \quad \sqrt{\frac{2}{}}$$

Next we bound the second term . Note that by assumption, for each , (1)| | (1) . By Hoe ding inequality (Steinwart and Christmann, 2008), we have

$$[\quad] \quad 1$$

where

$$(\quad 1) \quad \sqrt{\frac{}{2}}$$

Finally, bounding with Lemma 10, we have for any 0 $\in [0 \ 2]$ such that for all 1 and $^{/(\quad)}$, we have

$$[\quad (\quad \hat{}) \quad] \quad [\quad] \quad 1$$

and

$$(\quad) \quad \frac{}{} \quad \frac{(\quad)(\quad)}{(\quad)(\quad)} \quad \frac{}{} \quad \frac{}{\quad} \quad \frac{}{}$$

where is a constant depending on and .

Therefore, combining the bound for , we have

[ ] 1 4

where and depend on and go to 0 as → ∞. Note that (/ / / / /) (/) and ( ) — — . Based on /( ) → ∞, we can nd su cient larget such that , then we have

[∈] 1 ( )

To prove (II), for each xed 1 0 0 2, consider

2

2

2

Next we bound the above two terms respectively. For the rst term, take su cient large such that , then

∈ ∈

∉ ∉

0 ∉

4

For the second term, Lemma 10 implies that Combining both parts, when , we have for all 1,

1 5

Therefore, we have

( )

Note that when is replaced with (( )) (1)(( )) discussed in Section 2.3, the nal convergence rate would also include the rate coming from estimating ( ). More details can be found in Liu et al. (2018). ■

For any xed : → [ ],

$$
\begin{aligned}
(\ ) \quad & \max_{\rightarrow[\ ]} \ [\ (\ )\ \ (\ )]\frac{}{(\ (\ )|\ )} \\
& \max_{\rightarrow[\ ]} \ \ [\ (\ )\ \ (\ )]\frac{(\ |\ )}{(\ (\ )|\ )}\ [\ |\ \ ] \\
& \max_{\rightarrow[\ ]} \ \ [\ (\ )\ \ ] \ \ \frac{(\ |\ )}{(\ (\ )|\ )}\ [\ |\ \ ] \\
& \max_{\in[\ ]} \ \ \frac{(\ |\ )}{(\ (\ )|\ )}\ [\ |\ \ ] \\
& \max_{\in} \ [\ |\ \ ] \\
& \max_{\in[\ ]} \ [\ |\ \ \in\ \ ] \ \ (\ )
\end{aligned}
$$

where the last equation is because of the homogeneous group structure. So, taking supremum over all of the partitions gives sup →[ ] ( ) ( ). Therefore, based on the de nition of , we have .

First, for the excess risk of 0-1 loss,

$$
\begin{aligned}
& (\ )\ \ (\ ) \\
& \frac{}{(\ (\ )|\ )}\ [\arg\max\langle\ \ (\ )\rangle\ \ (\ )]\ \ [\arg\max\langle\ \ (\ )\rangle\ \ (\ )] \\
& \quad [\ |\ \in\ \ ]\ [\arg\max\langle\ \ (\ )\rangle\ \ (\ )]\ \ [\arg\max\langle\ \ (\ )\rangle\ \ (\ )]
\end{aligned}
$$

Then, for the excess risk of RAMSVM loss,

$$
\begin{aligned}
& (\ )\ \ (\ ) \\
& \frac{}{(\ (\ )|\ )}\ (1\ \ )\ \ \ (1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ _{(\ )}\ (\ )\rangle) \\
& \frac{}{(\ (\ )|\ )}\ (1\ \ )\ \ \ (1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ _{(\ )}\ (\ )\rangle) \\
& \quad [\ |\ \in\ \ ]\ (1\ \ )\ \ \ (1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ _{(\ )}\ (\ )\rangle)
\end{aligned}
$$

$$[\ |\ \in\ ](1\ \ )\ \ (1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ \ _{(\ )}\ \ (\ )\rangle)$$
$$_{(\ )}$$

$$[\ |\ \in\ ](1\ \ )\ \ (1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ \ _{(\ )}\ \ (\ )\rangle)$$
$$_{(\ )}$$

$$[\ |\ \in\ ](1\ \ )\ \ (1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ \ _{(\ )}\ \ (\ )\rangle)$$
$$_{(\ )}$$

$$[\ |\ \in\ ](\ \ 1\ \langle\ \ (\ )\rangle)\ \ (\ \ 1\ \langle\ \ (\ )\rangle)$$

$$[\ |\ \in\ ]\langle\ \ (\ )\ \ (\ )\rangle$$

The inequality is because, taking out the positive operator would make the rst term smaller while leave the second term unchanged due to $\langle\ \ (\ )\rangle \in [\ 1\ \ 1]$ for every $\in$. Next, we can only prove, for each $\in$,

$$[\ |\ \in\ ]\langle\ \ (\ )\ \ (\ )\rangle$$

$$[\ |\ \in\ ]\ [\arg\max\langle\ \ (\ )\rangle\ \ (\ )]\ \ [\arg\max\langle\ \ (\ )\rangle\ \ (\ )]$$

(19)

Suppose $[\ |\ \in\ ]\ \ [\ |\ \in\ ]$ for every . Then based on Fisher Consistency for group-based decision rule, we have $\arg\max_{\in[\ \ ]}\langle\ \ (\ )\rangle$. Suppose $\arg\max\langle\ \ (\ )\rangle$ where . The right hand side of (19) becomes

$$[\ |\ \in\ ](1\ \ 1)\ \ [\ |\ \in\ ](0\ \ 1)\ \ [\ |\ \in\ ](1\ \ 0)$$

$$[\ |\ \in\ ]\ \ [\ |\ \in\ ]$$

Since , the left hand side of (19) equals to

$$[\ |\ \in\ ]\langle\ \ (\ )\ \ (\ )\rangle\ \ [\ |\ \in\ ]\langle\ \ (\ )\ \ (\ )\rangle$$

$$[\ |\ \in\ ]\ \ [\ |\ \in\ ]\langle\ \ (\ )\ \ (\ )\rangle$$

$$[\ |\ \in\ ]\ \ [\ |\ \in\ ](1\ \langle\ \ (\ )\rangle)$$

$$[\ |\ \in\ ]\ \ [\ |\ \in\ ](1\ \langle\ \ (\ )\rangle)$$

The last equation is based on $\langle\ \ (\ )\rangle\ \ 1$ for each (Theorem 1 in Zhang et al. (2016)). Then, by the de nition of and the constraint of bound, we have $\langle\ \ (\ )\rangle\ \ 1$

for and $\langle \quad (\ ) \rangle \quad 0$. Hence, (19) holds and taking expectation for in both sides of (19) gives the result. ■

Let ( ) be the RKHS of the RBF kernel with parameter . Then based on Steinwart and Scovel (2007), the following linear operator $: \otimes \quad (\ ) \rightarrow \otimes \quad (\ )$ de ned by

$$(\ ) \quad \frac{(2 \quad)^{/}}{^{/}} \Big[\quad \| \quad \| \ (\ ) \quad \in \otimes \quad (\ ) \quad \in$$

is an isometric isomorphism. Accordingly, for any xed , we obtain

$$(\ ) \quad \langle \quad \rangle \in [\quad \inf \quad] \forall \in [\quad] \ \| \ \| \quad (\quad) \quad (\quad)$$

Using the notation of Lemma 4.1 of Steinwart and Scovel (2007), de ne 3 . Let be similarly de ned as on the domain of . We x a speci c measurable function such that for each $\in [\ ]$ and $\in$, satis es $\langle \quad (\ ) \rangle \quad 1$ and $\langle \quad (\ ) \rangle \quad 1$ for . For $\in$, it can be checked that the above property is equivalent with ( ) (1) $[\in]$. Besides , de ne ( ) . By similar approach of Lemma 4.1 in Steinwart and Scovel (2007), we can make sure the ball (( )) for every $\in$ ($\in [\ ]$) on this enlarged support for . Choose $(\ / \ )^{/}$ and note that $\| \ \| \ 1$, we immediately obtain the following with Jensen inequality and assumption of bounded volume of ,

$$\| \ \| \quad (\quad 1)(\frac{81}{\quad})^{/} \ (\ ) \quad (\quad 1)(\frac{81}{\quad})^{/}$$

Similar to the proof of Lemma 2 in our paper,

$$(\quad) \quad (\quad)$$

$$[\ \in \ ] \ | \ [\ | \ \in \quad] \quad [\ | \ \in \quad] | \ | \langle \quad (\ ) \quad (\ ) \rangle |$$

$$[\ \in \ ] \ | \ (\ ) | \ | \langle \quad (\ ) \quad (\ ) \rangle |$$

Now for $\in$, we would bound $|\langle \quad (\ ) \quad (\ )\rangle|$ when . We observe

$$
\begin{aligned}
1 \quad & \langle \quad (\ )\rangle \\
& \langle \quad (\frac{2}{\text{—}})^{/} \Big[\quad \| \quad \| \quad (\ ) \quad \rangle \\
& (\frac{2}{\text{—}})^{/} \Big[\quad \| \quad \| \langle \quad (\ )\rangle \\
& (\frac{2}{\text{—}})^{/} \Big[\quad \| \quad \| \langle \quad (\ ) \quad (\quad 1) \quad \rangle \quad (\quad 1) \\
& (\frac{2}{\text{—}})^{/} \Big[\quad \| \quad \| \langle \quad (\ ) \quad (\quad 1) \quad \rangle \quad (\quad 1) \\
& \qquad\qquad\qquad\qquad\qquad\qquad (\ \ (\ )) \\
& (\frac{2}{\text{—}})^{/} \Big[\quad \| \quad \| \quad (\quad 1) \\
& \qquad\qquad\qquad\qquad\qquad\qquad (\ \ (\ )) \\
& 1 \qquad (|\ | \quad (\ ))
\end{aligned}
$$

where $(2 \ / \ )^{/}$ $^{|\ |}$ is a spherical Gaussian in . According to Ledoux and Talagrand (2013), we have $(|\ | \quad (\ )) \quad 4 \quad ^{(\ )/}$ and consequently,

$$1 \quad \langle \quad (\ )\rangle \quad 1 \quad 4 \quad ^{(\ )/}$$

Note that $\langle \quad (\ )\rangle \quad 1$ for $\in$ and . Observe that above derivation holds for all $\in [\quad]$. Hence, we conclude

$$|\langle \quad (\ ) \quad (\ )\rangle| \quad 4 \quad ^{(\ )/}$$

for all $\in \Big[$. Then generalized geometric noise assumption for — yields

$$(\quad) \quad (\quad) \quad 4 \quad (\quad 1) \quad (2\ )^{/}$$

Therefore, combining both part, we conclude that

$$(\quad) \quad ^{/(\quad)}$$

where depends on the dimension of covariates , geometric noise component and associated constant , and upper bound of volume of covariate space when we set $^{/(\quad)}$.

■

:

To begin with, according to 1, it su ces to prove the result for $\tilde{}\ (\quad \hat{}\ ) \quad \tilde{}$. Then, similar to the proof idea in Theorem 3.4 in Zhao et al. (2012), we observe the following decomposition:

$$
\begin{aligned}
& \tilde{}\ (\quad \hat{}\ ) \quad \tilde{} \qquad \hat{} \qquad \tilde{}\ (\quad \hat{}\ ) \quad \inf_{\in \otimes} \quad \| \ \| \quad \tilde{}\ (\quad) \\
& \inf_{\in \otimes} \quad \| \ \| \quad \tilde{}\ (\quad) \quad \tilde{}
\end{aligned}
$$

where the rst term is referred as stochastic error and the second term is called the approximation bias term. We will bound each term separately in the following. First, the approximation bias term has been bounded by ($^{/(}$ $^{)}$) in Lemma 9.

Next, to bound the stochastic error term, we follow the proof of Theorem 3.4 in Zhao et al. (2012). The proof of their theorem is basically derived by verifying the conditions of Theorem 5.6 in Steinwart and Scovel (2007). To achieve that, we rst point out that, with Cauchy-Schwarz inequality, RAMSVM loss (1) $_{(\ )}$(1 $\langle$ ( )$\rangle$) (1 $\langle$ $_{(\ )}$ ( )$\rangle$) is Lipschitz continuous with respect to with Lipschitz constant . In addition, let be the population measure of ( ) / (( )|) and be the associated empirical measure. Next, by the de nition of $\hat{}$, we have

$$\frac{}{(\ (\ )|\ )}\ (\ \hat{}\ )\qquad \hat{}\qquad \frac{}{(\ (\ )|\ )}\ (\ )\qquad \|\ \|$$

holds for any $\in \otimes$. Hence, we can select to get

$$\hat{}\qquad \frac{1}{}\frac{1}{}\ \frac{}{(\ (\ )|\ )}\ (\ )\ \frac{2}{}(\ 1)$$

Therefore, it su ces to consider a ball with radius $\sqrt{2(\ 1)\ /\ }$ in the product RKHS $\otimes$, denoted by $_{\otimes}$ ( ). We de ne the function class

$$\left\{\frac{}{(\ (\ )|\ )}\ (\ )\ \|\ \|\ \frac{}{(\ (\ )|\ )}\ (\ )\qquad \in\ _{\otimes}\ (\ )\right\}$$

where $\arg\min\left\{\ (\ )\ \|\ \|\ :\ \in\ _{\otimes}\ (\ )\right\}$. Then similar to the proof of Theorem 3.4 in Zhao et al. (2012), we verify the following three conditions: (i) By the Lipschitz continuity of , there exists constant such that $\sup_{\in}\ \|\ \|_{\infty}$ where ($^{/}$); (ii) Based on the convexity of , there exists constant such that [ ] [ ] for all $\in$ where ( ); (iii) Theorem 2.1 in Steinwart and Scovel (2007) gives that, for all 0 0 2 0 0, sup log (( )) $^{(\ /\ )(\ )}$. Then the result follows from Theorem 5.6 in Steinwart and Scovel (2007).

Combining the bound for stochastic error term and approximation bias term, we complete the proof. ■

We rstly derive the equivalence of the two optimization problems based on the 0-1 loss. For any xed ,

$$\min \quad \frac{(\quad (\ ))}{(\ (\ )| \ )} [\quad (\ ) \quad (\ )] \quad (\quad 1) \quad \frac{(\quad (\ ))}{(\ (\ )| \ )} [\quad (\ ) \quad \tilde{} (\ )]$$

$$\min \quad _{(\ )} \quad \frac{[(\quad (\ )) \ | \ (\ ) \quad]}{(\ (\ )| \ )} [\quad (\ ) \quad (\ )]$$

$$(\quad 1) \quad _{(\ )} \quad \frac{[(\quad (\ )) \ | \ (\ ) \quad]}{(\ (\ )| \ )} \quad _{(\ )} \frac{1}{1} [\quad (\ ) \quad]$$

$$\min \quad _{(\ )} \quad \frac{[(\quad (\ )) \ | \ (\ ) \quad]}{(\ (\ )| \ )} [\quad (\ ) \quad (\ )]$$

$$_{(\ )} \quad \frac{[(\quad (\ )) \ | \ (\ ) \quad]}{(\ (\ )| \ )} \quad _{(\ )} [\quad (\ ) \quad]$$

$$\min \quad [(\quad (\ )) \ | \ (\ ) \quad] [\quad (\ ) \quad]$$

$$[(\quad (\ )) \ | \ (\ ) \quad] \quad [\quad (\ ) \quad]$$

$$\min \quad [(\quad (\ )) \ | \ (\ ) \quad] [\quad (\ ) \quad]$$

$$[(\quad (\ )) \ | \ (\ ) \quad] \quad [\quad (\ ) \quad] \quad [\quad (\ ) \quad]$$

$$\min \quad [(\quad (\ )) \ | \ (\ ) \quad] [\quad (\ ) \quad]$$

$$[(\quad (\ )) \ | \ (\ ) \quad] \quad 1 \quad [\quad (\ ) \quad]$$

$$\min \quad [\quad (\ )| \ (\ ) \quad] [\quad (\ ) \quad]$$

$$(\quad 1) \quad [(\quad (\ )) \ | \ (\ ) \quad]$$

$$\min \quad \frac{(\ )}{(\ (\ )| \ )} [\quad (\ ) \quad (\ )]$$

$$(\quad 1) \quad [(\quad (\ )) \ | \ (\ ) \quad]$$

Note that the second term in the last equation is not related to . Hence, the problem is equivalent with minimizing the rst term. This nishes the proof for the equivalence based on the 0-1 loss. The equivalence based on the RAMSVM loss can be directly guaranteed by Fisher consistency from Theorem 4. ■

For the linear kernel, we solve (9) with its dual form. After introducing new slack variables ( ) $\in[\ ]\ \in[\ ]$, the problem with linear learning can be written as

$$\min \frac{}{2} \qquad ( ) \quad (1 \quad) \qquad ( )$$

$$\text{s.t.} \qquad 0 \ (\ \in [\ ] \quad \in [\ ]);$$

$$( ) \ \langle \ (\ ) \quad (\ )\rangle \quad (\qquad 1) \quad 0 \ (\ \in [\ ]);$$

$$\langle \ (\ ) \quad \rangle \quad 1 \quad 0 \ (\ \in [\ ] \qquad (\ ))$$

The corresponding Lagrangian function is de ned as

$$\frac{}{2} \qquad ( ) \quad (1 \quad) \qquad ( )$$

$$( )[\quad ( ) \ \langle \ (\ ) \quad (\ )\rangle \quad (\qquad 1)]$$

$$[\qquad \langle \ (\ ) \quad \rangle \quad 1] \qquad ( )$$

where ( ) $\in[\ ]\ \in[\ ]$ and ( ) $\in[\ ]\ \in[\ ]$ are the Lagrangian multipliers. Furthermore, we can rewrite with

$$\frac{}{2} \qquad [\qquad (\ )] \quad (1 \quad) \ [\qquad (\ )]$$

$$( )(\qquad 1) \qquad ( )$$

$$( )\langle \ (\ ) \quad (\ )\rangle \qquad \langle \ (\ ) \quad \rangle \qquad ( )$$

Next we take partial derivative of with respect to ( ) $\in[\ ]\ \in[\ ]$ and ( ) $\in[\ ]$, and let them be 0. We have

$$\frac{\partial}{\partial} \qquad [\qquad (\ )] \quad (1 \quad) \ [\qquad (\ )] \qquad 0 \ (\ \in [\ ] \quad \in [\ ])$$

and

$$\frac{\partial}{\partial} \qquad ( ) \quad ( ) \qquad 0 \ (\ \in [\ 1]) \qquad ( )$$

When the partial derivatives for ( ) $\in[\ ]$ equal to 0, we have

$$\frac{1}{} \qquad ( ) \quad ( ) \qquad ( )$$

After plugging the above characterizations of into , we can simplify with

$$\frac{}{2} \qquad (\ )(\qquad 1)$$
$$(\ )$$

Note that maximizing with respect to is equivalent to minimizing , which solves the dual problem (11). The constraints of problem (11) come from $(\quad)_{\in[\ ]\ \in[\quad]} \quad 0$, $(\quad)_{\in[\ ]\ \in[\quad]} \quad 0$, and $(\partial \quad/\partial \quad)_{\in[\ ]\ \in[\quad]} \quad 0$.

For the general kernel, we similarly introduce the slack variables $(\quad)_{\in[\ ]\ \in[\quad]}$. If the intercepts $(\quad)_{\in[\quad]}$ are included in the penalty term, then (9) is equivalent to

$$\min \quad \frac{}{2} \qquad \frac{}{2} \qquad (\ ) \quad (1 \quad)$$
$$(\ )$$
$$\text{s.t.} \qquad 0\ (\ \in[\ ] \quad \in[\quad]);$$
$$(\ )\ \langle\ (\quad)\quad (\ )\rangle \quad (\qquad 1) \quad 0\ (\ \in[\ ]);$$
$$\langle\ (\quad)\quad\rangle \quad 1 \quad 0\ (\ \in[\ ] \qquad (\quad))$$

Similar to the linear case, we introduce the Lagrangian multipliers $(\quad)_{\in[\ ]\ \in[\quad]}$ and $(\quad)_{\in[\ ]\ \in[\quad]}$, calculate partial derivative with respect to $(\ )_{\in[\quad]}$ $(\quad)_{\in[\quad]}$, and $(\quad)_{\in[\ ]\ \in[\quad]}$, and set the derivatives to 0. Then we have that

$$\frac{1}{} \qquad (\ ) \quad (\ )$$
$$(\ )$$

and

$$\frac{1}{} \qquad (\ ) \quad (\ )$$
$$(\ )$$

where is the -th column of . After plugging the above characterizations of into the Lagrangian function and rewriting it, we get (12). ■

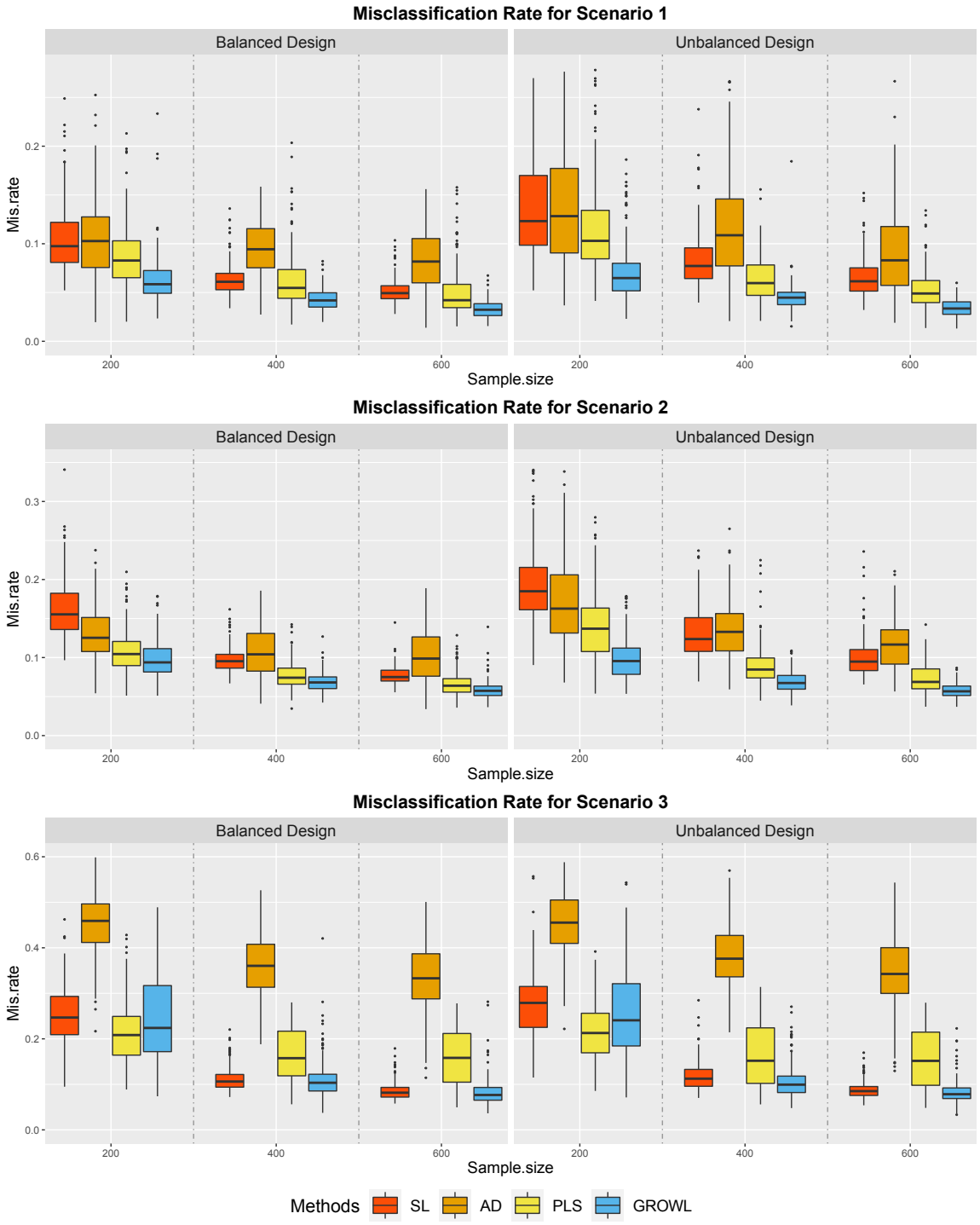

Figure 5: Boxplots of Misclassi caiton Rate evaluated on the independent test data under the settings.

To better demonstrate the performance of GROWL, we conduct several other simulations under more unbalanced designs for Scenario 3 in homogeneous setting. For the unbalanced design (II), the propensity scores of 15 treatments are set to be (0 035 0 035 0 035 0 114 0 114) (0 035 0 035 0 035 0 114 0 114) (0 035 0 035 0 088 0 088 0 088) , and for the unbalanced design (III), the propensity scores are set to be (0 020 0 020 0 020 0 137 0 137) (0 020 0 020 0 020 0 137 0 137) (0 020 0 020 0 098 0 098 0 098) . Varying from balanced design, unbalanced design (I) (same setting as shown in Section 4.1), unbalanced design (II), and unbalanced design (III), the treatment propensities become more and more unbalanced. In addition, we conduct another simulation setting where the propensity scores of one of the three treatment groups is extremely small. Speci cally, for this extreme case, the propensity scores are (0 070 0 070 0 087 0 087 0 087) (0 100 0 100 0 100 0 100 0 100) (0 020 0 020 0 020 0 020 0 020) . Overall, GROWL still performs the best in most cases shown in Figures 6 and 7. As treatment propensities become more unbalanced, GROWL may need more training data in order to learn the true partition correctly. In addition, a roughly correct $\hat{}$ is still helpful in terms of the performance of the value function. Compared with other methods that do not consider the treatment structure, GROWL is able to combine the similar treatments into the treatment groups. The decision rules learned from GROWL perform better because they are estimated under the treatment groups that have more observations than the individual treatments.

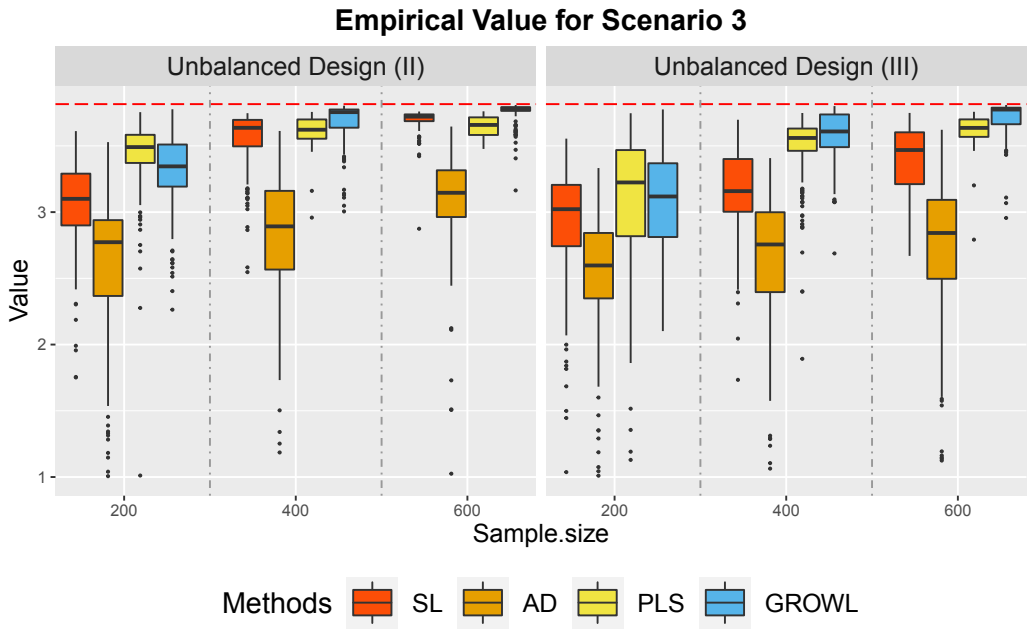

Figure 6: Boxplots of Empirical Value Function under more designs for the homogeneous case.

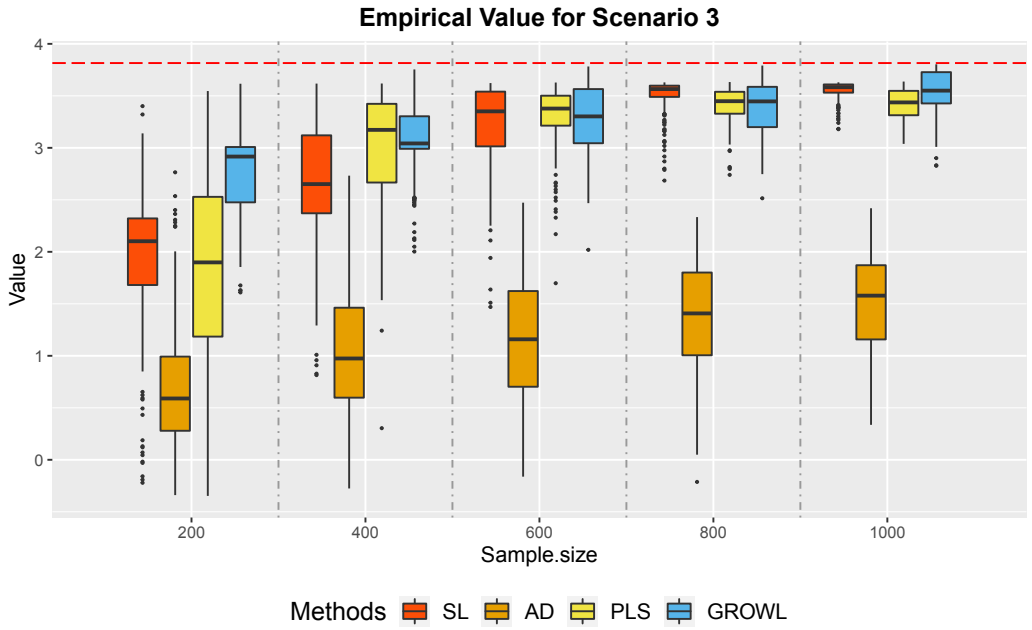

Figure 7: Boxplots of Empirical Value Function under designs for the homogeneous case.

We conduct more general simulation settings when the treatment propensity scores depend on the covariates. Speci cally, the treatment propensity score (|) is provided with the following multinomial model:

$$\log \frac{(\quad | \quad)}{(1| \quad)}$$

for 2 , where s are all generated independently from [0 1 0 1]. For the homogeneous case, we take Scenario 1 as an example. The value shown in Figure 8 demonstrates that GROWL still has superior performance over other methods. For the non-homogeneous case in Scenario 4, as shown in Figure 9, the overall trend of GROWL is similar to that of the unbalanced design in Figure 2. GROWL is still competitive, and especially has smaller variance than other methods.

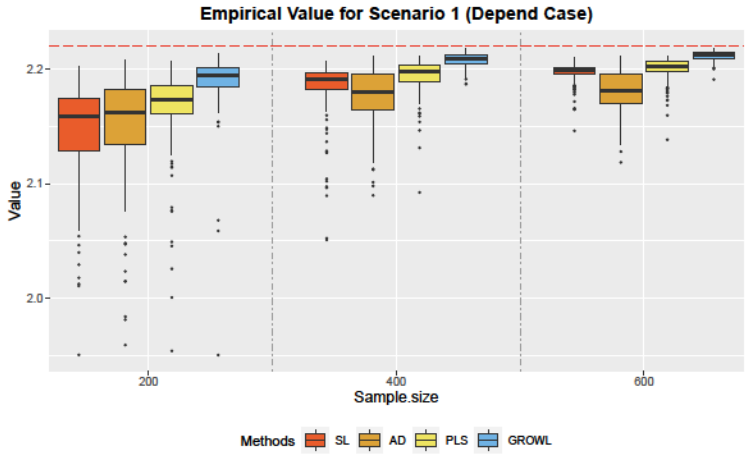

Figure 8: Boxplots of Empirical Value Function under the **dependent** design for the homogeneous case.

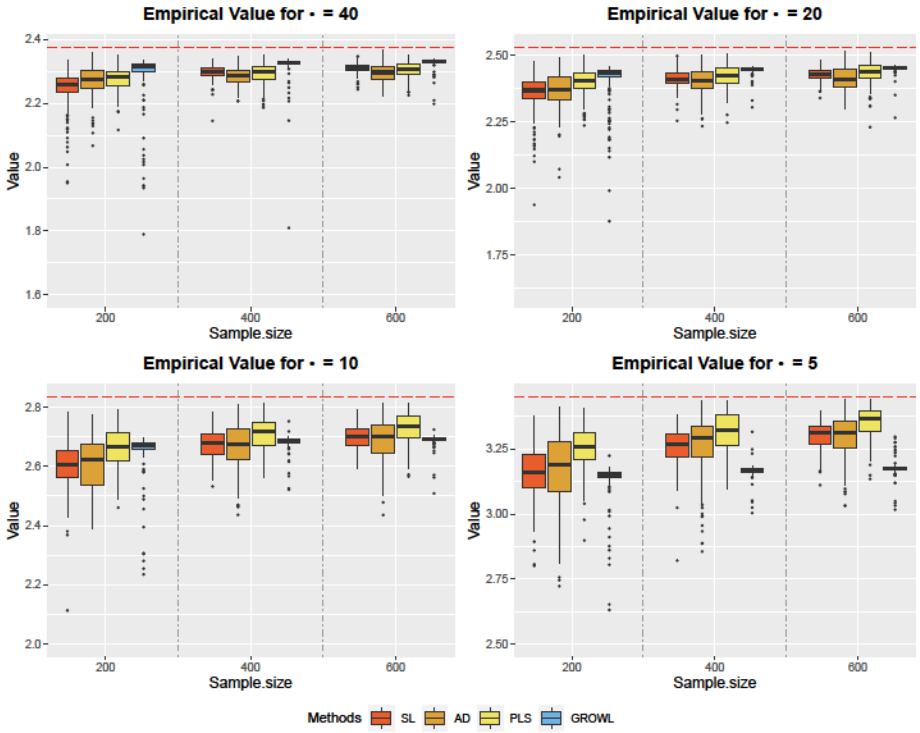

Figure 9: Boxplots of Empirical Value Function under the **dependent** design for the non-homogeneous case.

Table 3: Estimated coe cients of linear comparison function for GROWL on the STAR*D dataset. Larger coe cients encourage better reward.

Variable Name	Group 1	Group 2	Group 3	Group 4	Group 5
intercept	0.713	0.168	0.658	0.191	-1.730
gender (female)	-0.839	0.249	1.635	-0.444	-0.601
ethnic (white)	-0.878	-0.180	-0.764	0.221	1.601
age	-0.959	-0.319	0.579	-0.746	1.445
depression history (yes)	-1.427	0.021	-0.368	0.528	1.245
marital	1.049	0.371	-1.636	0.282	-0.066
school years	1.421	-0.151	-0.033	0.153	-1.390
education	0.526	-1.230	1.166	-0.841	0.379
student (yes)	-0.813	1.704	-0.541	-0.377	0.026
employment	0.178	-0.867	-1.085	1.369	0.405
volunteer work	0.066	-0.750	-0.056	-0.886	1.626
QIDS change during Level 1	-1.574	0.136	1.206	0.228	0.004
QIDS at the start of Level 2	0.743	0.576	-0.952	0.851	-1.218

.

a. Fit (penalized) linear regression with training data and get residuals ;
b. Input the estimated initial partition set .
For each partition in ,
a. Fit treatment into group ( ) based on
b. 0, Stay with the same assigned treatment ( ), and set ¯( ) ( );
Uniformly switch ( ) to arbitrary unassigned treatment ˜( ), and set ¯( ) ˜( );
c. Use (¯( )) to t RAMSVM with weights $\frac{|\ |}{(\ (\ )|\ )}$;
d. Get tted decision function ˆ based on RAMSVM;
e. Plug (ˆ) back into the empirical average of the risk function ˜ ;
f. Get the risk value for .
the set using the until convergence;
the optimal ˆ and corresponding ˆ under group domain;
one treatment from group ˆ ( ) based on , and nally get the ITR ( ).

.

a. Fit (penalized) linear regression with training data and get residuals ;
b. Input the estimated initial partition set .

,
1 2 1 2 ,

a. Adjust the assignment of the -th treatment and hold the assignment for others xed;
b. Get the risk value for the adjusted in the same way shown in Algorithm 1;
c. Obtain the locally best in cyclic fashion.

convergence.
the optimal ^ and corresponding ^ under group domain;
one treatment from group ^ ( ) based on , and nally get the ITR ( ).

Susan Athey and Stefan Wager. Policy learning with observational data. , 89 (1):133 161, 2021.

Peter L Bartlett and Shahar Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. , 3:463 482, 2002.

Leo Breiman. Random forests. , 45(1):5 32, 2001.

Jingxiang Chen, Quoc Tran-Dinh, Michael R Kosorok, and Yufeng Liu. Identifying heterogeneous e ect using latent supervised clustering with adaptive fusion. , 30(1):43 54, 2021.

Shuai Chen, Lu Tian, Tianxi Cai, and Menggang Yu. A general statistical framework for subgroup identi cation and comparative treatment scoring. , 73(4):1199 1209, 2017.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In , pages 785 794, 2016.

Ashkan Ertefaie, Tianshuang Wu, Kevin G Lynch, and Inbal Nahum-Shani. Identifying a set that contains the best dynamic treatment regimes. , 17(1):135 148, 2016.

David E Goldberg and John Henry Holland. . Kluwer Academic Publishers, 1988.

Arthur E Hoerl and Robert W Kennard. Ridge regression: applications to nonorthogonal problems. , 12(1):69 82, 1970.

Eric B Laber, Daniel J Lizotte, and Bradley Ferguson. Set-valued dynamic treatment regimes for competing outcomes. , 70(1):53 61, 2014.

Michel Ledoux and Michel Talagrand. . Springer Science & Business Media, 2013.

Ying Liu, Yuanjia Wang, Michael R Kosorok, Yingqi Zhao, and Donglin Zeng. Augmented outcome-weighted learning for estimating optimal dynamic treatment regimens. , 37(26):3776 3788, 2018.

Haomiao Meng, Yingqi Zhao, Haoda Fu, and Xingye Qiao. Near-optimal individualized treatment recommendations. , 21(183):1 28, 2020.

Olga Montvida, Jonathan Shaw, John J Atherton, Frances Stringer, and Sanjoy K Paul. Long-term trends in antidiabetes drug usage in the us: real-world evidence in patients newly diagnosed with type 2 diabetes. , 41(1):69 78, 2018.

Susan A Murphy. Optimal dynamic treatment regimes. , 65(2):331 355, 2003.

Yinghao Pan and Yingqi Zhao. Improved doubly robust estimation in learning optimal individualized treatment rules. , 116(533): 283 294, 2021.

Eric C Polley and Mark J Van Der Laan. Super learner in prediction. , 2010.

Zhengling Qi and Yufeng Liu. D-learning to estimate optimal individual treatment rules. , 12(2):3601 3638, 2018.

Zhengling Qi, Dacheng Liu, Haoda Fu, and Yufeng Liu. Multi-armed angle-based direct learning for estimating optimal individualized treatment rules with various outcomes. , 115(530):678 691, 2020.

Min Qian and Susan A Murphy. Performance guarantees for individualized treatment rules. , 39(2):1180 1210, 2011.

Naim U Rashid, Daniel J Luckett, Jingxiang Chen, Michael T Lawson, Longshaokan Wang, Yunshu Zhang, Eric B Laber, Yufeng Liu, Jen Jen Yeh, Donglin Zeng, and Michael R Kosorok. High-dimensional precision medicine from patient-derived xenografts. , 116(535):1140 1154, 2021.

Janice H Robinson, Lynn C Callister, Judith A Berry, and Karen A Dearing. Patient-centered care and adherence: De nitions and applications to improve outcomes. , 20(12):600 607, 2008.

Donald B Rubin. Estimating causal e ects of treatments in randomized and nonrandomized studies. , 66(5):688 701, 1974.

A John Rush, Maurizio Fava, Stephen R Wisniewski, Philip W Lavori, Madhukar H Trivedi, Harold A Sackeim, Michael E Thase, Andrew A Nierenberg, Frederic M Quitkin, T Michael Kashner, David J Kupfer, Jerrold F Rosenbaum, Jonathan Alpert, Jonathan W Stewart, Patrick J McGrath, Melanie M Biggs, Kathy Shores-Wilson, Barry D Lebowitz, Louise Ritz, and George Niederehe. Sequenced treatment alternatives to relieve depression (star* d): rationale and design. , 25(1):119 142, 2004.

Luca Scrucca. Ga: A package for genetic algorithms in r. ,
53(4):1 37, 2013.

Ingo Steinwart and Andreas Christmann. . Springer Science &
Business Media, 2008.

Ingo Steinwart and Clint Scovel. Fast rates for support vector machines using gaussian
kernels. , 35(2):575 607, 2007.

Lu Tian, Ash A Alizadeh, Andrew J Gentles, and Robert Tibshirani. A simple method for
estimating interactions between a treatment and a large number of covariates.
, 109(508):1517 1532, 2014.

William N Venables and Brian D Ripley. . Springer
Science & Business Media, 2013.

Christopher John Cornish Hellaby Watkins. . PhD thesis,
University of Cambridge, UK, 1989.

Peng Wu, Donglin Zeng, and Yuanjia Wang. Matched learning for optimizing individualized
treatment strategies using electronic health records.
, 115(529):380 392, 2020.

Yichao Wu and Yufeng Liu. Robust truncated hinge loss support vector machines.
, 102(479):974 983, 2007.

Yingcun Xia, Howell Tong, Wai Keung Li, and Li-Xing Zhu. An adaptive estimation of
dimension reduction space. In
, pages 299 346. World Scienti c, 2009.

Baqun Zhang, Anastasios A Tsiatis, Eric B Laber, and Marie Davidian. A robust method
for estimating optimal treatment regimes. , 68(4):1010 1018, 2012.

Chong Zhang and Yufeng Liu. Multicategory angle-based large-margin classi cation.
, 101(3):625 640, 2014.

Chong Zhang, Yufeng Liu, Junhui Wang, and Hongtu Zhu. Reinforced angle-based mul-
ticategory support vector machines. ,
25(3):806 825, 2016.

Chong Zhang, Jingxiang Chen, Haoda Fu, Xuanyao He, Yingqi Zhao, and Yufeng Liu. Mul-
ticategory outcome weighted margin-based learning for estimating individualized treat-
ment rules. , 30(4):1857 1879, 2021.

Yingqi Zhao, Donglin Zeng, A John Rush, and Michael R Kosorok. Estimating individual-
ized treatment rules using outcome weighted learning.
, 107(499):1106 1118, 2012.

Yingqi Zhao, Donglin Zeng, Eric B Laber, and Michael R Kosorok. New statistical learning
methods for estimating optimal dynamic treatment regimes.
, 110(510):583 598, 2015.

Xin Zhou, Nicole Mayer-Hamblett, Umer Khan, and Michael R Kosorok. Residual weighted learning for estimating individualized treatment rules.
, 112(517):169 187, 2017.

Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net.
, 67(2):301 320, 2005.