Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis

Sihao Ding
dsihao@mail.ustc.edu.cn
University of Science and Technology of China

Peng Wu*
wupeng@bicmr.pku.edu.cn
Beijing Technology and Business University

Fuli Feng
fulifeng93@gmail.com
University of Science and Technology of China

Yitong Wang
wangyitong@mail.ustc.edu.cn
University of Science and Technology of China

Xiangnan He
xiangnanhe@gmail.com
University of Science and Technology of China

Yong Liao
yliao@ustc.edu.cn
University of Science and Technology of China

Yongdong Zhang
zhyd73@ustc.edu.cn
University of Science and Technology of China

ABSTRACT

Recommender systems should answer the intervention question "if recommending an item to a user, what would the feedback be", calling for estimating the causal effect of a recommendation on user feedback. Generally, this requires blocking the effect of confounders that simultaneously affect the recommendation and feedback. To mitigate the confounding bias, a strategy is incorporating *propensity* into model learning. However, existing methods forgo possible unmeasured confounders (*e.g.,* user financial status), which can result in biased propensities and hurt recommendation performance. This work combats the risk of unmeasured confounders in recommender systems.

Towards this end, we propose ***R**obust **D**econfounder* (RD) that accounts for the effect of unmeasured confounders on propensities, under the mild assumption that the effect is bounded. It estimates the bound with sensitivity analysis, learning a recommender model robust to unmeasured confounders within the bound by adversarial learning. However, pursuing robustness within a bound may restrict model accuracy. To avoid the trade-off between robustness and accuracy, we further propose ***B**enchmarked RD* (BRD) that incorporates a pre-trained model into the learning as the benchmark. Theoretical analyses prove the stronger robustness of our methods compared to existing propensity-based deconfounders, and also prove the no-harm property of BRD. Our methods are applicable to any propensity-based estimators, where we select three representative ones: IPS, Doubly Robust, and AutoDebias. We conduct experiments on three real-world datasets to demonstrate the effectiveness of our methods.

CCS CONCEPTS

• **Information systems → Recommender systems**.

KEYWORDS

Recommendation, Unmeasured confounder, Deconfounder

ACM Reference Format:
Sihao Ding, Peng Wu, Fuli Feng, Yitong Wang, Xiangnan He, Yong Liao, and Yongdong Zhang. 2022. Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), August 14–18, 2022, Washington, DC, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3534678.3539240

1 INTRODUCTION

Recommender systems play an important role in a wide range of applications such as content-sharing [8, 36], social media [7], and e-commerce [20, 40]. They should answer the intervention question "if recommending an item to a user, what would the feedback be". From the perspective of causal inference [27], this is a *potential outcome* [25, 30] question calling for estimating the causal effect of system exposure (*treatment*) on user feedback (*outcome*). Estimating the causal effect directly from historical data may suffer from confounding biases, which are caused by the *confounders* between the treatment and the output. For example, item popularity affects user feedback due to the herd mentality of users [6] and affects system exposure due to the uneven distribution of historical data used for model training [45]. Neglecting such confounders will result in popularity bias [40], *e.g.,* over-recommending popular items regardless of item quality. It is critical to block confounding biases in recommendation.

Existing methods mitigate such bias with confounder adjustment, which are in two main categories.

- *Structural Causal Models* (SCM) [27] based methods, which perform adjustment (*e.g.,* backdoor adjustment [45]) according to

*Corresponding author: Peng Wu.

KDD '22, August 14–18, 2022, Washington, DC, USA.

ACM ISBN 978-1-4503-9385-0/22/08…$15.00
https://doi.org/10.1145/3534678.3539240

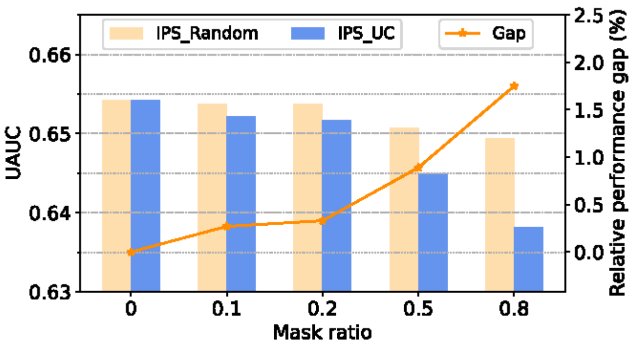


Figure 1: Performances of IPS on the Yahoo!R3 dataset in the presence of unmeasured confounders (IPS_UC) or not (IPS_Random). Gap denotes the relative performance difference from IPS_Random to IPS_UC.

the causal relations in a predefined causal graph. As relying on causal graph, these methods can only handle a small portion of confounders with known causal relations.

- *Propensity*-based methods [5, 32, 37] handle confounders without knowing causal relations to pursue unbiased model estimators. They estimate a *propensity* for each training interaction and incorporate the propensity to the loss function. Nevertheless, they are based on the *strong ignorability assumption* [13, 39], *i.e.,* no unmeasured confounders.

Recommender systems face unmeasured confounders for various reasons like technical difficulties or privacy restrictions [23, 35]. For example, user financial status directly affects the feedback and exposure probability[1], but be immeasurable to many recommender systems. Ignoring such confounders may mislead the system to overemphasize item price, *e.g.,* over-recommending cheap items. Our theoretical analyses show that propensity-based deconfounders are biased in the presence of unmeasured confounders (*cf.* Section 3.1). Figure 1 further shows an empirical evidence where we compare a representative method IPS [32] on the Yahoo!R3 [37] dataset under the scenarios with and without simulated unmeasured confounders. IPS_UC denotes selectively masking positive feedback to simulate unmeasured confounders. IPS_Random is a control group that randomly masks feedback. The performance of IPS_UC (*i.e.,* under unmeasured confounders) is worse than IPS_Random in all cases and deteriorates faster as the mask ratio increases. It is thus critical to consider the unmeasured confounders in recommender systems.

To combat the risk of unmeasured confounders, a possible solution is employing relevant causal inference methods such as *instrumental variable estimation* [12] and *front door adjustment* [27]. Nevertheless, it is difficult to find a convincing instrument variable or a mediator that satisfies the front door criterion[2] [12, 15]. We instead explore propensity-based methods to combat unmeasured confounders, in which knowledge of causal relations are unnecessary. Existing methods fail since their **nominal** propensity is not the **true** propensity to achieve the unbiased estimator [17]. To address this issue, we consider approaching the true propensity from the nominal one. We assume the true propensity is near to the nominal one within a bound, which is decided by the strength of the unmeasured confounder. Through this way, we can leverage the propensities within the bound to enhance the robustness of methods against unmeasured confounders.

Specifically, we propose *Robust Deconfounder* (RD) for recommender learning. We first quantify the potential effect of unmeasured confounders on propensities with the sensitivity analysis [9, 29], obtaining the bound of true propensity around the nominal one. To account for all propensities within the bound, we then optimize the model with adversarial learning, which alternatively searches for the worst-case of the propensity within the bound and updates model parameters accordingly. As that adversarial learning may cause the trade-off between robustness and accuracy [3], we further propose *Benchmarked Robust Deconfounder* (BRD) that incorporates a pre-trained model into the training. We prove in theory that BRD is no worse than the benchmark model, and both RD and BRD are more robust to unmeasured confounders. We instantiate RD and BRD on three representative propensity-based methods: IPS [32], DR [37], and AutoDebias [5], and conduct experiments on three real-world datasets to demonstrate our methods.

To summarize, our main contributions are as follows:

- We reveal the risk of unmeasured confounders in recommender systems with theoretical and empirical analyses.
- We propose a robust deconfounding framework that combats unmeasured confounders with theoretical accuracy guarantee.
- We conduct extensive experiments, that justified the effectiveness and robustness of our methods to unmeasured confounders.

2 PROBLEM FORMULATION

In this section, we formulate the problem of confounders adjustment in the presence of unmeasured confounders using *potential outcome framework* [25, 30]. Let $\mathcal{U} = \{u\}$ and $\mathcal{I} = \{i\}$ denote the set of users and items, respectively, *potential outcome framework* consists of the key components as follows:

- *Unit*: a user-item pair (u, i).
- *Target population*: the set of all user-item pairs $\mathcal{D} = \mathcal{U} \times \mathcal{I}$.
- *Feature*: the feature $x_{u,i}$ describes user-item pair (u, i).
- *Treatment*: $o_{u,i} \in \{1, 0\}$. It is the exposure status of (u, i), where $o_{u,i} = 1$ or 0 denotes item i is exposed to user u or not.
- *Outcome*: the feedback $r_{u,i}$ of user-item pair (u, i).
- *Potential outcome*: $r_{u,i}(o)$ for $o \in \{0, 1\}$. It is the outcome that would be observed if $o_{u,i}$ had been set to o.

Let $\mathbb{P}$ and $\mathbb{E}$ be the distribution and expectation on the target population, and $O = \{(u, i) \mid (u, i) \in \mathcal{D}, o_{u,i} = 1\}$ be the set of exposed units. In recommendation, the target estimand [12, 16, 41] is $\mathbb{E}(r_{u,i}(1) \mid x_{u,i})$[3], it requires to predict the potential outcome $r_{u,i}(1)$ using feature $x_{u,i}$. In this article, we focus on the scenario of unmeasured confounders. Considering both of the measured confounders $x_{u,i}$ and the unmeasured confounders $h_{u,i}$, we have

$$o_{u,i} \perp r_{u,i}(1) \mid (x_{u,i}, h_{u,i}), \quad o_{u,i} \not\perp r_{u,i}(1) \mid x_{u,i}, \tag{1}$$

where the symbols $\perp$ and $\not\perp$ represent independent and not independent. A typical causal graph is displayed in Figure 2.

[1]Historical interactions of users reflect their financial status. Learning recommender models from historical interactions will bring the influence of financial status.

[2]In addition, to ensure the identifiability, they rely on additional assumptions unrealistic in recommender system. For example, front door criterion assumes there is no unmeasured confounder between the mediator and outcome [15].

[3]It is equivalent to $\mathbb{P}(r_{u,i} \mid x_{u,i}, do(o_{u,i} = 1))$ using *do-calculus* in SCM framework.

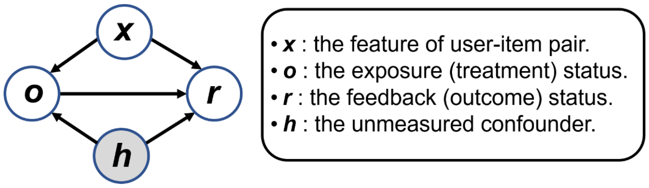


Figure 2: A typical causal graph of unmeasured confounders.

Let $f_\phi(\cdot)$ be a recommender model with parameters ϕ and $\hat{r}_{u,i}(1) = f_\phi(x_{u,i})$ be the predicted $\mathbb{E}(r_{u,i}(1) \mid x_{u,i})$. If all potential outcomes $\{r_{u,i}(1) : (u,i) \in \mathcal{D}\}$ were observed, the ideal loss function for training ϕ is

$$\mathcal{L}_{ideal}(\phi) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} e_{u,i}, \tag{2}$$

where $e_{u,i}$ is the prediction error, such as the least square loss:

$$e_{u,i} = (\hat{r}_{u,i}(1) - r_{u,i}(1))^2. \tag{3}$$

Noticing that $e_{u,i}$ is computable only when $o_{u,i} = 1$, $\mathcal{L}_{ideal}(\phi)$ is infeasible due to the inaccessibility of $\{r_{u,i}(1) : o_{u,i} = 0\}$. As such, our target is constructing robust estimators[4] that approximate to $\mathcal{L}_{ideal}(\phi)$ to combat unmeasured confounders.

3 PROPOSED METHODS

In this section, we first show that the existing propensity-based methods are biased in the presence of unmeasured confounders. Then, we propose RD and BRD frameworks that can effectively alleviate this problem, and further extend the proposed frameworks to a general propensity-based method. Throughout, we omit the L_2 normalization term in all loss functions for brevity.

3.1 Motivation

The IPS [14, 31, 32] and doubly robust (DR) learning [11, 37, 38, 44] are main strategies to cope with confounding bias. Yet, most existing IPS/DR methods are designed only for addressing measured confounders x and ignore the unmeasured confounders h. In this case, the **nominal** propensity score is defined as $p_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i})$.

IPS estimator. Given the estimate of $p_{u,i}$[5], denoted as $\hat{p}_{u,i}$, the IPS estimator of the prediction inaccuracy is presented as

$$\mathcal{L}_{IPS}(\phi) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}}. \tag{4}$$

DR estimator. It can be constructed in the augmented IPS form [2, 37]. Specifically, let $\hat{e}_{u,i} = g_\theta(x_{u,i})$ be an error imputation model to fit the prediction error $e_{u,i}$ using $x_{u,i}$, i.e., it estimates $g_{u,i} = \mathbb{E}[e_{u,i}|x_{u,i}]$. Given the learned $\hat{p}_{u,i}$ and $\hat{e}_{u,i}$, the DR estimator is

$$\mathcal{L}_{DR}(\phi,\theta) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \left[\hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})}{\hat{p}_{u,i}}\right]. \tag{5}$$

Joint learning [37] is a common technique for DR estimator, which obtains $\hat{e}_{u,i}$ by minimizing

$$\mathcal{L}^{DR}_{Imp}(\phi,\theta) = \frac{1}{|\mathcal{O}|} \sum_{(u,i)\in\mathcal{O}} \frac{(\hat{e}_{u,i} - e_{u,i})^2}{\hat{p}_{u,i}}, \tag{6}$$

and the parameters ϕ and θ are estimated by solving Eq. (5) and Eq. (6) alternately. Without unmeasured confounders, it is well known that $\mathcal{L}_{IPS}(\phi)$ is an unbiased estimate of the ideal loss function, provided that $\hat{p}_{u,i}$ equals $p_{u,i}$ accurately for the exposed events. And the DR estimator is unbiased if either $\hat{p}_{u,i}$ or $\hat{e}_{u,i}$ is estimated precisely. Nevertheless, the existence of unmeasured confounders will invalidate the existing IPS and DR methods.

Theorem 3.1. *In the presence of unmeasured confounders h,*
(a) both the IPS and DR estimators are biased, even $\hat{p}_{u,i}$ and $\hat{e}_{u,i}$ estimate $p_{u,i}$ and $g_{u,i}$ accurately.
(b) if we define the **true** *propensity score as*

$$\tilde{p}_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}, h_{u,i}), \tag{7}$$

and assume that $\hat{p}_{u,i}$ is an accurate estimate of $\tilde{p}_{u,i}$, then both the IPS and DR estimators are unbiased.

Theorem 3.1 indicates that the definition of propensity score plays a key role in the unbiasedness of IPS and DR estimators. Actually, using the measured confounders x to estimate the propensity score only controls the confounding bias created by x and cannot rule out the bias induced by h. To remove all the confounding bias, both the measured confounders x and the unmeasured confounders h should be used to estimate the propensity score. Unfortunately, one can never accurately estimate $\tilde{p}_{u,i}$ due to the full missingness of unmeasured confounders without imposing strong assumptions.

3.2 Robust Deconfounder Framework

Sensitivity analysis [9, 29] is a basic tool to assess the robustness of inferences to the assumption of no unmeasured confounders. Based on the idea of sensitivity analysis on propensity score model [17, 29], we obtain the uncertainty set of $\tilde{p}_{u,i}$ for each (u,i) by restricting the strength of unmeasured confounding on the treatment. Concretely, assume the nominal propensity score can be expressed as

$$p_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}) = \frac{\exp(m(x_{u,i}))}{1 + \exp(m(x_{u,i}))},$$

where m is an arbitrary function. Given a bound $\Gamma \geq 1$, consider an additive model of true propensity score that

$$\tilde{p}_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}, h_{u,i}) = \frac{\exp(m(x_{u,i}) + \varphi(h_{u,i}))}{1 + \exp(m(x_{u,i}) + \varphi(h_{u,i}))},$$

φ is an arbitrary function. By assuming the strength of unmeasured confounders is bounded as $|\varphi(h)| \leq \log(\Gamma)$, we have

$$\frac{1}{\Gamma} \leq \frac{(1 - p_{u,i})\tilde{p}_{u,i}}{p_{u,i}(1 - \tilde{p}_{u,i})} \leq \Gamma. \tag{8}$$

Eq. (8) restricts the value range of $\tilde{w}_{u,i} = 1/\tilde{p}_{u,i}$ as

$$a_{u,i} \leq \tilde{w}_{u,i} \leq b_{u,i}, \tag{9}$$

$$a_{u,i} = 1 + (1/p_{u,i} - 1)/\Gamma,\ b_{u,i} = 1 + (1/p_{u,i} - 1)\Gamma. \tag{10}$$

The hyper-parameter Γ corresponds to the strength of unmeasured confounding, and $\Gamma = 1$ means no unmeasured confounding. Let

$$\mathcal{W} = \{W \in \mathbb{R}_+^{|\mathcal{D}|} : \hat{a}_{u,i} \leq w_{u,i} \leq \hat{b}_{u,i}\}, \tag{11}$$

where $W = \{w_{u,i} : (u,i) \in \mathcal{D}\}$, $\hat{a}_{u,i}$ and $\hat{b}_{u,i}$ are the estimates of $a_{u,i}$ and $b_{u,i}$.

[4] In this paper *estimator* and *loss function* are interchangeable.
[5] We refer [32] for the methods to estimate the nominal propensity score $p_{u,i}$.

Algorithm 1: Robust Deconfounder IPS (RD-IPS)

Input: Data $\mathcal{D}$, nominal propensity score $p_{u,i}$.
Output: An optimized recommender model.

```
Initialize a recommender model with parameters φ;
Calculate the bound of true propensity score based on
  Eq. (10) with hyper-parameter Γ;
Generate the uncertainty set 𝒲 = {W} based on Eq. (11)
while Stop condition is not reached do
    Fetch (u, i) from 𝒟;
    Calculate the prediction error e_{u,i} based on φ and Eq. (3);
    Maximize the loss Eq. (12) to update W;
    Minimize the loss Eq. (12) to optimize φ;
end
Return an optimized recommender model with φ;
```

The uncertainty set $\mathcal{W}$ is the key for proposing the RD framework, which introduces the adversarial learning technique by fluctuating the inverse of estimated nominal propensity scores within $\mathcal{W}$. The RD framework provides an excellent opportunity to design robust propensity-based methods. Specifically, the RD-IPS estimator can be constructed as

$$\mathcal{L}_{RD-IPS}(\phi) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} o_{u,i} e_{u,i} w_{u,i}, \tag{12}$$

In addition, a new RD-DR estimator is formulated as

$$\mathcal{L}_{RD-DR}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} [\hat{e}_{u,i} + o_{u,i}(e_{u,i} - \hat{e}_{u,i}) w_{u,i}], \tag{13}$$

and the corresponding loss function for the imputation model is

$$\mathcal{L}_{Imp}^{RD-DR}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{O}} (\hat{e}_{u,i} - e_{u,i})^2 w_{u,i}. \tag{14}$$

The parameters ϕ and θ are solved alternately: given $\hat{\theta}$, ϕ is updated by minimizing (13); given $\hat{\phi}$, θ is updated by minimizing (14). Algorithm 1 summarizes the procedure of RD-IPS, and Algorithm 2 in Appendix A.2 summarizes the procedure of RD-DR.

Comparing Eq. (14) with Eq. (6), the RD-DR method also adds an adversarial procedure for the imputation model. The reason is that using $\hat{p}_{u,i}$ directly will lead to a biased imputation model and raise the risk of performance degradation. Thus, applying the adversarial procedure may reduce the unstable factors resulting from unmeasured confounders. Intuitively, the proposed RD-IPS/DR approaches can avoid the worst-case caused by unmeasured confounders and are thus more robust than non-adversarial estimators.

In practice, addressing the unmeasured confounding is a very challenging problem, because the functional relations among h, o, and r are elusive and can be in any form. Instead of aiming to eliminate the unmeasured confounding thoroughly, the proposed RD framework calibrates the loss function with uncertainty sets by leveraging the sensitivity analysis techniques in causal inference, which provides a flexible way to mitigate the unmeasured confounding and avoid the dangers of relying on unrealistic assumptions.

3.3 Benchmarked RD Framework

The proposed RD framework will improve the robustness of the existing propensity-based methods. However, there is no theoretical evidence that the robustness improvement can boost prediction accuracy. To tackle the problem, we propose another Benchmarked RD framework, which guarantees the prediction accuracy advancement by setting a benchmark estimator.

For clarity, the prediction error $e_{u,i}$ is written as $e_{u,i}(\phi)$, a function of ϕ. Suppose we have access to an estimator of ϕ, denoted as $\hat{\phi}^{(0)}$, via using the existing approaches. Then the corresponding BRD-IPS estimator is

$$\mathcal{L}_{BRD-IPS}(\phi) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} o_{u,i}\{e_{u,i}(\phi) - e_{u,i}(\hat{\phi}^{(0)})\} w_{u,i}. \tag{15}$$

The BRD-IPS estimator is different from the RD-IPS estimator. Compared with Eq. (12), setting a benchmark would change the solution of $w_{u,i}$ in Eq. (15), and thereby calibrate the estimation of ϕ. Analogously, the BRD-DR estimator and the associated imputation model are given as

$$\begin{aligned}\mathcal{L}_{BRD-DR}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \Big\{ [\hat{e}_{u,i}(\theta) - \hat{e}_{u,i}(\hat{\theta}^{(0)})] + \\ [o_{u,i}(e_{u,i}(\phi) - \hat{e}_{u,i}(\theta)) - o_{u,i}(e_{u,i}(\hat{\phi}^{(0)}) - \hat{e}_{u,i}(\hat{\theta}^{(0)}))] w_{u,i} \Big\}, \\ \mathcal{L}_{Imp}^{BRD-DR}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{D}} \Big\{ [\hat{e}_{u,i}(\theta) - e_{u,i}(\phi)]^2 \\ - [(\hat{e}_{u,i}(\hat{\theta}^{(0)}) - e_{u,i}(\hat{\phi}^{(0)})]^2 \Big\} w_{u,i},\end{aligned}$$

where $\hat{\phi}^{(0)}$ and $\hat{\theta}^{(0)}$ are the benchmark estimates of ϕ and θ.

The proposed BRD-IPS/DR approaches are very flexible due to the free choice of the benchmark estimators. In addition, note that $\mathcal{L}_{BRD-IPS}(\phi) = 0$ if $\phi = \hat{\phi}^{(0)}$, which implies that the minimum value of $\mathcal{L}_{BRD-IPS}(\phi)$ is non-positive. Similarly, it is not hard to derive that $\min_\phi \mathcal{L}_{BRD-DR}(\phi) \le 0$. Thus, the BRD-IPS/DR estimators achieve smaller losses than the corresponding benchmark estimators. Furthermore, the following Theorem 3.2 shows that the BRD-IPS/DR estimators outperform the corresponding benchmark estimators in terms of generalization error. Define

$$\phi^\dagger = \arg\min_\phi \mathcal{L}_{BRD-IPS}(\phi), \quad \phi^\ddagger = \arg\min_\phi \mathcal{L}_{BRD-DR}(\phi). \tag{16}$$

Theorem 3.2 ("no-harm" property). *Suppose that $\phi^{(0)}$ is a benchmark estimator of ϕ and $|\mathcal{D}|$ is large enough, then*

(a) under the conditions of Theorem 4.1, if $\mathcal{L}_{BRD-IPS}(\phi^\dagger) < 0$, we have with probability at least $1 - \eta$,

$$\mathcal{L}_{ideal}(\phi^\dagger) < \mathcal{L}_{ideal}(\hat{\phi}^{(0)}). \tag{17}$$

(b) under the conditions of Theorem 4.2, if $\mathcal{L}_{BRD-DR}(\phi^\ddagger) < 0$, we have with probability at least $1 - \eta$,

$$\mathcal{L}_{ideal}(\phi^\ddagger) < \mathcal{L}_{ideal}(\hat{\phi}^{(0)}). \tag{18}$$

Theorem 3.2 indicates that the BRD-IPS/DR estimators enjoy the "no-harm" property: they are no worse than the corresponding benchmark estimators, and setting a pre-trained model as the benchmark would guarantee no accuracy sacrifice issue. This is an appealing and exciting result that means we can comfortably implement BRD in practical recommender systems by setting the current recommender model as the benchmark. In Appendix A.2, Algorithm 3 summarizes the procedures of BRD-IPS. RD and BRD are general frameworks that can be applied in normal propensity-based methods. Recent debias work in recommendation [5, 28, 47] generalizes these propensity-based methods. We then consider extending RD and BRD to a recent advance AutoDebias [5].

3.4 RD and BRD AutoDebias

AutoDebias [5] is a general propensity-based method that learns the propensity score from uniform data and imputes ratings of unobserved data in training phase to boost the prediction accuracy. Based on the RD framework, the new RD AutoDebias (RD-Auto) estimator is given as

$$\mathcal{L}_{RD-Auto}(\phi) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[o_{u,i} e_{u,i}(\phi) w_{u,i} + \mathcal{L}_{Imp}^{Auto}(\phi) \right],$$

where $\mathcal{L}_{Imp}^{Auto}(\phi)$ is the loss that measures the distance between model predictions and the imputation ratings. Moreover, the BRD AutoDebias (BRD-Auto) estimator is presented as

$$\begin{aligned} \mathcal{L}_{BRD-Auto}(\phi) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \Big\{ \mathcal{L}_{Imp}^{Auto}(\phi) \\ - \mathcal{L}_{Imp}^{Auto}(\hat{\phi}^{(0)}) + o_{u,i}(e_{u,i}(\phi) - e_{u,i}(\hat{\phi}^{(0)})) w_{u,i} \Big\}. \end{aligned} \tag{19}$$

4 THEORETICAL ANALYSIS

We present the generalization bounds of the proposed methods. The bounds depend on the complexity of the prediction model class. Letting $\mathcal{F}$ be a class of functions and assume $f_\phi \in \mathcal{F}$, we define the Rademacher complexity

$$\mathcal{R}(\mathcal{F}) = \mathbb{E}_{\sigma \sim \{-1,+1\}^{|\mathcal{D}|}} \sup_{f_\phi \in \mathcal{F}} \left[\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \sigma_{u,i} e_{u,i} \right],$$

where $\sigma = \{\sigma_{u,i} : (u,i) \in \mathcal{D}\}$ is a Rademacher sequence [24]. Assume that $\mathcal{F}$ has a vanishing complexities, i.e., $\mathcal{R}(\mathcal{F}) \to 0$ as $|\mathcal{D}| \to \infty$. This is a very weak assumption, which is satisfied by common models including matrix factorization [21, 34] considered in this paper.

Theorem 4.1 (generalization bound of RD-IPS and BRD-IPS). *Suppose that $\tilde{w}_{u,i} \in [\hat{a}_{u,i}, \hat{b}_{u,i}]$, $e_{u,i} \le C_1$, and $\tilde{w}_{u,i} \le C_2$ for all (u,i) pairs. Then for any $f_\phi \in \mathcal{F}$ and $\eta > 0$, we have that with probability at least $1-\eta$,*

$$\mathcal{L}_{ideal}(\phi) \le \mathcal{L}_{RD-IPS}(\phi) + \mathcal{B}(\eta, \mathcal{D}, \mathcal{F}), \tag{20}$$

$$\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{ideal}(\hat{\phi}^{(0)}) \le \mathcal{L}_{BRD-IPS}(\phi) + 2\mathcal{B}(\eta, \mathcal{D}, \mathcal{F}). \tag{21}$$

where $\phi^{(0)}$ is a pre-trained IPS estimator of ϕ,

$$\mathcal{B}(\eta, \mathcal{D}, \mathcal{F}) = 2(C_2+1)\mathcal{R}(\mathcal{F}) + C_1(C_2+1)\sqrt{\frac{18\log(2/\eta)}{|\mathcal{D}|}}.$$

Eq. (20) and Eq. (21) hold for any ϕ. In particular, let $\phi^* = \arg\min_\phi \mathcal{L}_{RD-IPS}(\phi)$, the right side of (20) reaches the minimum. Thus, the RD-IPS estimator is asymptotically an upper bound of the ideal estimator by noting that $\mathcal{B}(\eta, \mathcal{D}, \mathcal{F}) \to 0$ as $|\mathcal{D}| \to \infty$. In addition, Theorem 3.2(a) can be deduced by Eq. (21). For $\phi^\dagger$ defined in Eq. (16), if $\mathcal{L}_{BRD-IPS}(\phi^\dagger) < 0$, then we have $\mathcal{L}_{BRD-IPS}(\phi^\dagger) \le -\epsilon$ for a positive constant ϵ. When $|\mathcal{D}|$ is large enough, such that $\mathcal{B}(\eta, \mathcal{D}, \mathcal{F}) < \epsilon/2$. Then Theorem 3.2(a) follows immediately from the truth that the right side of (21) is negative by setting $\phi = \phi^\dagger$. For RD-DR and BRD-DR estimators, we have the similar results as shown in the following Theorem 4.2. And Theorem 3.2(b) also can be deduced by Theorem 4.2.

Table 1: The statistics of Yahoo!R3, Coat, and Product.

Dataset	#User	#Item	#Biased Data	#Unbiased Data
Yahoo!R3	15,400	1,000	311,704	54,000
Coat	290	300	6,960	4,640
Product	7,176	10,729	1,325,197	270,180

Theorem 4.2 (generalization bound of RD-DR and BRD-DR). *Suppose that $\tilde{w}_{u,i} \in [\hat{a}_{u,i}, \hat{b}_{u,i}]$, $|e_{u,i} - \hat{e}_{u,i}| \le C_3$, and $\tilde{w}_{u,i} \le C_2$, given imputed errors $\{\hat{e}_{u,i} : (u,i) \in \mathcal{D}\}$, then for any $f_\phi \in \mathcal{F}$ and $\eta > 0$, we have that with probability at least $1-\eta$,*

$$\mathcal{L}_{ideal}(\phi) \le \mathcal{L}_{RD-DR}(\phi) + \mathcal{B}_2(\eta, \mathcal{D}, \mathcal{F}),$$

$$\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{ideal}(\hat{\phi}^{(0)}) \le \mathcal{L}_{BRD-DR}(\phi) + 2\mathcal{B}_2(\eta, \mathcal{D}, \mathcal{F}),$$

where $\phi^{(0)}$ is a pre-trained DR estimator of ϕ,

$$\mathcal{B}_2(\eta, \mathcal{D}, \mathcal{F}) = 2(C_2+1)\mathcal{R}(\mathcal{F}) + (2C_1 + C_3)(C_2+1)\sqrt{\frac{2\log(2/\eta)}{|\mathcal{D}|}}.$$

5 EXPERIMENTS

We conduct experiments to answer the following questions:

- **RQ1:** Do the proposed RD and BRD boost the performance of propensity-based methods?
- **RQ2:** Do our methods stably perform well under the scenarios with different extents of unmeasured confounder?
- **RQ3:** What factors influence the effectiveness of our methods?

5.1 Experimental Settings

Datasets. To validate the effectiveness of RD and BRD, we utilize three datasets with unbiased data [32] in different application domains: 1) Yahoo!R3[6], 2) Coat[7], and 3) Product[8], which are obtained from the music, coat, and micro-video recommendations, respectively. All datasets contain a set of biased data collecting the normal interactions of users in the platform, and a set of unbiased data collecting from a *randomized controlled trial* where items are assigned randomly (See Table. 1). Following [5, 19], the biased data is used to train, and we extract a small part (5% on Yahoo!R3 and Coat, 1% on Product) of unbiased data as unbiased training data, and set 5% unbiased data as the validation set for choosing hyper-parameters, and the rest of unbiased is used to test. As to Yahoo!R3 and Coat, explicit feedback with ratings larger than 3 is treated as 1, otherwise the feedback is labeled as −1. For the Product dataset, the feedback is defined by the platform based on user's playing time.

Compared methods. We implement RD and BRD in three basic propensity-based methods Inverse Propensity Score (IPS) [32],

[6]https://webscope.sandbox.yahoo.com/.
[7]https://www.cs.cornell.edu/~schnabts/mnar/.
[8]It is a popular micro-video sharing platform.

Table 2: Recommendation performances on Yahoo!R3, Coat, and Product. The best results relevant to each basic propensity-based method are highlighted with bold. RI refers to the relative improvement of RD or BRD over the corresponding baseline.

Datasets	Yahoo!R3				Coat				Product			
	UAUC	RI	NDCG@5	RI	UAUC	RI	NDCG@5	RI	UAUC	RI	NDCG@50	RI
Base model	0.6507	-	0.5449	-	0.6575	-	0.4761	-	0.6269	-	0.0914	-
DCF	0.6542	-	0.5489	-	0.6490	-	0.5016	-	0.6680	-	0.1204	-
IPS	0.6542	-	0.5525	-	0.6612	-	0.4858	-	0.6587	-	0.1131	-
RD-IPS	0.6791	3.8%	0.5808	5.1%	0.6712	1.5%	**0.5145**	5.9%	0.6680	1.4%	0.1266	12%
BRD-IPS	**0.6810**	4.1%	**0.5825**	5.4%	**0.6819**	3.1%	0.5028	3.5%	**0.6753**	2.5%	**0.1300**	15.0%
DR	0.6633	-	0.5622	-	0.6689	-	0.4949	-	0.6612	-	0.1144	-
RD-DR	0.6785	2.3%	0.5799	3.1%	**0.6803**	1.7%	**0.5092**	2.9%	0.6787	2.6%	0.1277	11.6%
BRD-DR	**0.6801**	2.5%	**0.5842**	3.9%	0.6770	1.2%	0.5080	2.8%	**0.6832**	3.3%	**0.1428**	24.8%
AutoDebias	0.7279	-	0.6421	-	0.6857	-	0.5264	-	0.6879	-	0.1365	-
RD-AutoDebias	0.7328	0.7%	0.6453	0.6%	0.6891	0.5%	0.5337	1.4%	0.6962	1.2%	**0.2183**	59.9%
BRD-AutoDebias	**0.7400**	1.7%	**0.6580**	2.6%	**0.6950**	1.4%	**0.5647**	7.3%	**0.6989**	1.6%	0.1493	9.4%

Doubly Robust (DR) [37], and AutoDebias [5] to validate their effectiveness. we also compare the methods above with DCF [39] that can alleviate the effect of unmeasured confounders.

Base model [18]. We employ widely used Matrix Factorization (MF) as the base model for propensity-based methods, and adopt the public implement [5] to train it with biased data.
IPS [32]. IPS is a classical propensity-based method. It estimates the propensity, and uses the inverse propensity score to weight the loss function. We adopt the public implementation [5] and employ MF as the base model of IPS with the biased data for training.
DR [37]. DR is another classical propensity-based method. It trains a base model and a imputation model with loss functions adjusted by the propensity score. In this work we use the unbiased training data to estimate the propensity score, and implement DR base on MF with public implementation [5].
AutoDebias [5]. AutoDebias is a general propensity-based method that learns the propensity score from unbiased training data with a meta-learning module. We adopt the source code and hyper-parameter ranges for grid search released in the original paper.
DCF [39]. DCF is a deconfounded recommender that tries to make MF more robust to unobserved confounders. It trains an exposure model with exposure data to predict which item will be exposed to a user. By leveraging the training data and the predictions of the exposure model, DCF can learn a model that is more robust to unobserved confounders than MF. In our work, we use the combination of biased data and unbiased training data to train DCF. To obtain the exposure data, we set all -1 and 1 feedback in the training data as 1. We adopt the public implementation and tune the hyper-parameters following the official code.
RD. RD is our proposed method, and we apply RD in IPS, DR, and AutoDebias as RD-IPS, RD-DR, and RD-AutoDebias respectively.[9]
BRD. BRD is our proposed method, and we apply BRD in IPS, DR, and AutoDebias as BRD-IPS, BRD-DR, and BRD-AutoDebias respectively.

Evaluation Metrics. We adopt *UAUC* and *NDCG@K* to evaluate performances. For each user, we calculate the AUC [5] and NDCG@K [5] over the exposed items in the unbiased testing data, and then take the average scores of all users[10] to obtain UAUC and NDCG@K, respectively. Here K is set to 5 for Yahoo!R3 and Coat, and is set to 50 for Product due to its low positive ratio (2.12%).

5.2 Performance Comparison (RQ1)

Table 2 reports the performance comparison on three datasets. From the table, we have the following observations:

- Both RD and BRD boost the performances of corresponding IPS, DR, and AutoDebias regarding UAUC and NDCG across all datasets, which demonstrates the effectiveness of RD and BRD. We attribute the performance gain to calibrating the loss function with uncertainty sets obtained through sensitivity analysis, which accounts for the impact of unmeasured confounders.
- BRD-IPS, BRD-DR, and BRD-AutoDebias outperform the corresponding IPS, DR, and AutoDebias in all cases. These results validate the "no-harm" property of BRD, which are consistent with our theoretical analysis (Theorem 3.2). That is, incorporating the pre-trained propensity-based model into the loss function can guarantee no worse performance. Therefore, it will be safe to directly apply BRD to existing models.
- BRD achieves larger performance gains than RD in most cases. We postulate the reason is that BRD further incorporates a pre-trained propensity-based model as the benchmark. It is also reasonable that RD achieves the best performance in a few cases if the impact of unmeasured confounders is strong and the benchmark model is largely affected. Moreover, in some cases BRD achieves higher UAUC but lower NDCG@K than RD. In these cases, BRD ranks the tail items in the ranking list more accurately than RD. We postulate the reason is the benchmark model in BRD ensures the mismatched items don't receive higher scores during adversarial training.
- As to the previous propensity-based methods, IPS, AutoDebias, and DR all achieve better performances than the base model regarding both UAUC and NDCG across all datasets. It validates that using the nominal propensity $p_{u,i}$ to construct the estimator

[9]https://github.com/Dingseewhole/Robust_Deconfounder_master/

[10]In order to compare each method more accurately, we remove the testing users that only have positive (*i.e.,* 1) or only have negative (*i.e.,* -1) testing feedback, since their NDCG and AUC of all methods are always -1 or 1.

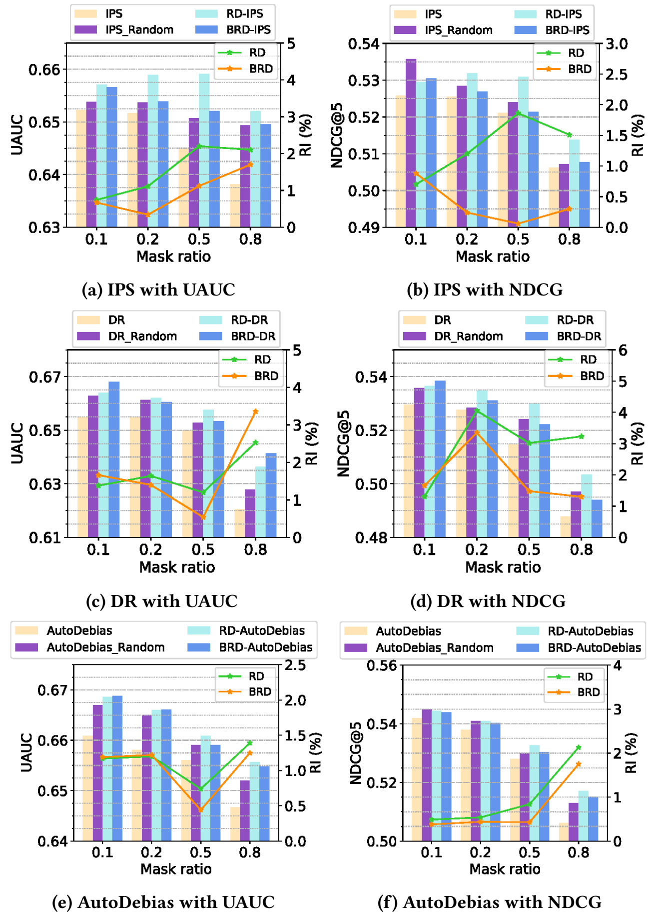
can still help deconfounding. It is because adjusting the loss function with $p_{u,i}$ can block the effect of measured confounders.
- DCF also considers unmeasured confounders, which outperforms MF and IPS in most cases, while performs worse than RD and BRD. We postulate the reason is that DCF assumes all confounders are relevant to the probability of exposure, which is not exactly satisfied. Moreover, the two-stage training strategy of DCF may raise the error amplification issue: the error of the exposure model may cause the performance degradation of DCF.

5.3 In-depth Analysis (RQ2, RQ3)

Study on Confounding Strength. We further investigate how RD and BRD perform as the impact of unmeasured confounder getting stronger. We test RD, BRD and the three basic propensity-based methods on semi-simulated datasets from Yahoo!R3. The semi-simulated data is generated by **selectively** masking a ratio of **positive** feedback in training data. The masking can be viewed as having unmeasured confounders that impact treatment (item exposure) and outcome (positive feedback). The mask ratios are set to {0.1, 0.2, 0.5, 0.8}, and the larger ratio means the stronger effect of the unmeasured confounder. To block the effect of training data reduction, we take random masking as a reference, which **randomly** masks training data regardless of the label at the ratios of {0.1, 0.2, 0.5, 0.8}. We denote the reference performance of IPS, DR, and Autodebias as IPS_Random, DR_Random, and AutoDebias_Random.

The performance of all methods are shown in Figure 3. As to each mask ratio, we also show the relative improvement (*i.e.,* RI) of RD and BRD over the basic propensity-based methods. From the figures, we have the following observations: 1) the performance of IPS, DR, and AutoDebias deteriorates as the mask ratio increases, *i.e.,* the effect of unmeasured confounder increases. 2) IPS, DR, and AutoDebias perform worse than the corresponding IPS_Random, DR_Random, AutoDebias_Random across all mask ratios. These results confirm the performance degradation of the basic propensity-based methods are mainly caused by the unmeasured confounder that we simulated. 3) RD and BRD stably outperform the corresponding IPS, DR, and AutoDebias *w.r.t.* both UAUC and NDCG across all mask ratios. It indicates that our methods make the propensity-based methods robust to unmeasured confounders with different strengths. 4) In most cases, the largest RI is achieved under the highest mask ratio when the confounding effect is the strongest. In other words, our methods achieve the most outstanding performance gain when the effect of unmeasured confounders is significant. We believe the reason is our methods leverage the adversarial learning to consider the worst-case caused by unmeasured confounders. 5) In some instances, basic propensity-based methods perform better under the randomly masking setting (*e.g.,* IPS_Random). Since the original training data of Yahoo!R3 is biased, not all training samples play a positive role of unbiased test. We believe that, in these cases, the masked samples are the ones that are difficult to debias, and masking these samples improves the unbiasedness of prediction.

Effects on Propensities. The adversarial learning process of RD makes perturbations on every nominal propensity given by IPS or DR. To investigate the effects of RD, we calculate the absolute gap between the inverse of the nominal propensity and final propensity



(a) IPS with UAUC

(b) IPS with NDCG

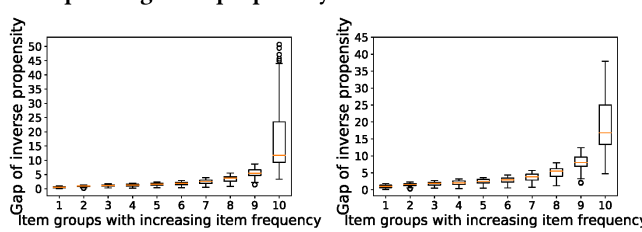
(c) DR with UAUC

(d) DR with NDCG

(e) AutoDebias with UAUC

(f) AutoDebias with NDCG

Figure 3: Performances of basic propensity-based methods under settings with and without the simulated unmeasured confounders, and the performance of RD and BRD applied in corresponding basic propensity-based methods on Yahoo!R3.



(a) IPS

(b) DR

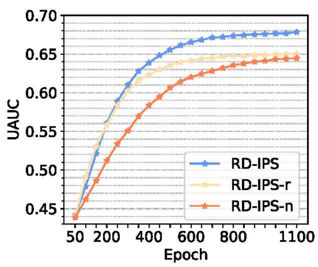
Figure 4: Distribution of inverse nominal propensity scores and those given by RD after the optimization.

given by RD. Figure 4 shows that the average gap on each item is positively correlated to item frequency. In Theorem 3.1 we proved the objective function of IPS/DR is biased since they use the nominal propensity. Also, the frequent items have more critical bias issue, because their embedding are trained more frequently to fit the biased objective function. Therefore, RD-IPS/DR makes larger perturbations for the frequent items' propensities to debias.

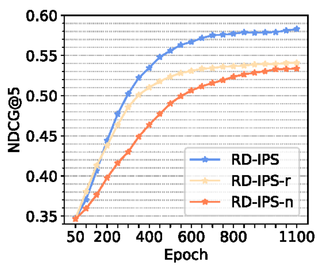
To further evaluate whether the larger perturbations are beneficial, we report the group-wise performance of IPS, DR, RD-IPS/DR

Table 3: Group-wise performance of IPS, DR, and RD-IPS/DR.

Item set	High frequency				Low frequency			
	UAUC	RI	NDCG@5	RI	UAUC	RI	NDCG@5	RI
IPS	0.6536	-	0.7370	-	0.6254	-	0.7151	-
RD-IPS	0.6867	5.06%	0.7576	2.80%	0.6310	0.90%	0.7164	0.18%
DR	0.6713	-	0.7464	-	0.6206	-	0.7118	-
RD-DR	0.6845	1.97%	0.7540	1.02%	0.6250	0.71%	0.7138	0.28%



(a) UAUC



(b) NDCG

Figure 5: Performances of different optimizing methods.

Table 4: Performances of DR, RD-DR and RD-NI-DR.

Dataset	Yahoo!R3		Coat		Product	
	UAUC	NDCG@5	UAUC	NDCG@5	UAUC	NDCG@50
DR	0.6633	0.5622	0.6689	0.4949	0.6612	0.1144
RD-NI-DR	0.6664	0.5642	0.6734	0.5027	0.6782	0.1265
RD-DR	**0.6785**	**0.5799**	**0.6803**	**0.5092**	**0.6787**	**0.1277**

on the high/low frequency item sets of Yahoo!R3. According to the item frequency in the training set, the high frequency set includes the testing samples related to the 50% highest frequency items, and the low frequency set includes the rest testing samples. As shown in Table 3, RD-IPS and RD-DR achieve higher relative performance gains on the high frequency set as compared to the corresponding IPS and DR. These results validate the benefit of adding larger perturbations on frequent items.

Ablation Study. Recall that we adopt adversarial learning to optimize the worst-case in the uncertainty set $\mathcal{W}$. To further validate the effectiveness of this design, we compare it with two variants: RD-IPS-r and RD-IPS-n. In each training epoch, RD-IPS-r randomly selects inverse propensities from $\mathcal{W}$ to construct the loss function in Eq. (12). RD-IPS-n adds a Gaussian white noise to the nominal inverse propensity score $w_{u,i}$ in each training epoch to construct the loss function. We set the standard deviation of the Gaussian white noise as $(\hat{b}_{u,i} - \hat{a}_{u,i})/2$. Figure 5 shows that RD-IPS achieves better performances than its variants, validating the effectiveness of the adversarial learning.

Recall that DR estimator consists of two losses Eq. (5) and Eq. (6), which are both adjusted by propensities. When implementing RD-DR, we calibrate both losses with the uncertainty set (*i.e.,* Eq. (13) and Eq. (14)). To validate the necessity of using adversarial learning to train the imputation model with Eq. (14), we make an ablation study of RD-DR by comparing it with RD-NI-DR. RD-NI-DR only employs the adversarial procedure for base model, *i.e.,* RD-NI-DR jointly optimizes Eq. (13) and Eq. (6). Table. 4 shows the performance of DR, RN-NI-DR, and RD-DR. Notably, RD-DR outperforms RD-NI-DR in all cases, which indicates that adding an adversarial procedure for the imputation model reduces the unstable factors resulting from unmeasured confounders. RD-NI-DR still outperforms DR, which further validates the rationality of our methods.

6 RELATED WORK

In this section, we review existing work on Causal Inference in Recommendation and Debias in Recommendation, which are most relevant with this work.

Causal Inference in Recommendation. Causal inference [27] is the science that aims to estimate the causal effects between the variables, and has been widely used in natural language inference [10], computer vision [26], and recommendation [37]. The first line of causal inference in recommendation is on confounding effects [39, 42]. For example, [4] explores recommender systems are affected by algorithmic confounding. PD [45] reveals item popularity is a confounder that affects both item exposure and probability of interaction. By measuring the extent of item popularity, PD eliminates the confounding bias caused by popularity and boosts the recommendation performance. Another line is counterfactual learning methods [36, 40, 43]. For example, [36] imagines a counterfactual world to reduce the direct effect of the exposure features, and trains a recommender model that overcomes the clickbait issue. [43] proposes CPR that can generate counterfactual samples for promoting recommendation performance. MACR [40] proposes a counterfactual inference framework to remove the effect of item popularity for better recommendation. In this paper we propose a general solution that makes the propensity-based methods robust to unmeasured confounders, which is parallel with the deconfounding or counterfactual methods.

Debias in Recommendation. Bias is a common problem in recommender systems that getting more and more attention. Since recommender systems are trained with historical data that has intrinsic biases [6], the system always suffers from bias issues such as, popularity bias [45], position bias [1], and conformity bias [22]. To mitigate the bias issue, many solutions have been proposed. For example, propensity-based methods [6, 32, 37] estimate the propensity score and try to construct the unbiased estimator as the objective function for unbiased learning. Causal inference methods [40, 45] aim to block the effect of variables that cause the bias. Causal embedding-based methods [19, 46] aim to learn more rational causal embedding for mitigating the effect of bias. Different with existing works that aim to alleviate one or several bias, our work focuses on mitigating the confounding bias caused by unmeasured confounders.

7 CONCLUSION

In this work, we studied how to combat the unmeasured confounders in recommendation. By theoretically analyzing the existing propensity-based methods, we found that they are biased in the presence of unmeasured confounders due to adjusting the estimator with nominal propensities. However, the true propensity can never be accurately estimated, so we proposed *Robust Deconfounder* that leverages the sensitivity analysis to estimate the uncertainty set of the true propensity. To learn models that are robust to unmeasured confounders, RD leverages the adversarial learning technique to optimize the worst-case with the propensity in the uncertainty set.

Furthermore, we proposed *Benchmarked Robust Deconfounder* that incorporates a pre-trained propensity-based model as a benchmark to avoid the trade-off issue between robustness and accuracy, and we proved the no-harm property of BRD through theoretical analysis. We conducted experiments on three real-world datasets and did the theoretical analysis, proving the robustness and superiority of RD and BRD.

This work shows the limitation of propensity-based methods in recommendation, and proposes a general solution. In the future, we are interested in testing our methods on more recommender models such as graph convolutional network based models. Moreover, we would like to explore how to automatically estimate the hyper-parameter Γ from the data of different scenarios.

ACKNOWLEDGMENTS

This work is supported by the National Key Research and Development Program of China (2021ZD0111802), the National Natural Science Foundation of China (61972372, U21B2026, 62121002), and the CCCD Key Lab of Ministry of Culture and Tourism.

REFERENCES

[1] Qingyao Ai, Keping Bi, Cheng Luo, Jiafeng Guo, and W Bruce Croft. 2018. Unbiased learning to rank with unbiased propensity estimation. In *SIGIR*. 385–394.

[2] Heejung Bang and James M. Robins. 2005. Doubly robust estimation in missing data and causal inference models. *Biometrics* (2005), 962–972.

[3] Anirban Chakraborty, Manaar Alam, Vishal Dey, Anupam Chattopadhyay, and Debdeep Mukhopadhyay. 2018. Adversarial attacks and defences: A survey. *arXiv:1810.00069* (2018).

[4] Allison JB Chaney, Brandon M Stewart, and Barbara E Engelhardt. 2018. How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. In *RecSys*. 224–232.

[5] Jiawei Chen, Hande Dong, Yang Qiu, Xiangnan He, Xin Xin, Liang Chen, Guli Lin, and Keping Yang. 2021. AutoDebias: Learning to debias for recommendation. In *SIGIR*. 21–30.

[6] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. 2020. Bias and debias in recommender system: A survey and future directions. *arXiv:2010.03240* (2020).

[7] Jiawei Chen, Can Wang, Sheng Zhou, Qihao Shi, Yan Feng, and Chun Chen. 2019. Samwalker: Social recommendation with informative sampling strategy. In *WWW*. 228–239.

[8] Zhiyong Cheng, Fan Liu, Shenghan Mei, Yangyang Guo, Lei Zhu, and Liqiang Nie. 2022. Feature-Level Attentive ICF for Recommendation. *ACM Transactions on Information Systems* (2022), 1–24.

[9] H Christopher Frey and Sumeet R Patil. 2002. Identification and review of sensitivity analysis methods. *Risk analysis* (2002), 553–578.

[10] Fuli Feng, Jizhi Zhang, Xiangnan He, Hanwang Zhang, and Tat-Seng Chua. 2021. Empowering Language Understanding with Counterfactual Reasoning. In *ACL*. 2226–2236.

[11] Siyuan Guo, Lixin Zou, Yiding Liu, Wenwen Ye, Suqi Cheng, Shuaiqiang Wang, Hechang Chen, Dawei Yin, and Yi Chang. 2021. Enhanced Doubly Robust Learning for Debiasing Post-Click Conversion Rate Estimation. In *SIGIR*.

[12] M.A. Hernán and J. M. Robins. 2020. *Causal Inference: What If.* Boca Raton: Chapman and Hall/CRC.

[13] Paul W Holland. 1986. Statistics and causal inference. *Journal of the American statistical Association* (1986), 945–960.

[14] D. G. Horvitz and D. J. Thompson. 1952. A Generalization of Sampling Without Replacement From a Finite Universe. *Journal of the American Statistical Association* 47 (1952), 663–685.

[15] Guido W Imbens. 2020. Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature* (2020), 1129–79.

[16] G. W. Imbens and D. B. Rubin. 2015. *Causal Inference For Statistics Social and Biomedical Science.* Cambridge University Press.

[17] Nathan Kallus and Angela Zhou. 2018. Confounding-Robust Policy Improvement. In *NeurIPS*. 9289–9299.

[18] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* (2009), 30–37.

[19] Dugang Liu, Pengxiang Cheng, Zhenhua Dong, Xiuqiang He, Weike Pan, and Zhong Ming. 2020. A general knowledge distillation framework for counterfactual recommendation via uniform data. In *SIGIR*. 831–840.

[20] Fan Liu, Zhiyong Cheng, Lei Zhu, Zan Gao, and Liqiang Nie. 2021. Interest-aware message-passing gcn for recommendation. In *WWW*. 1296–1305.

[21] Tongliang Liu and Dacheng Tao. 2015. On the performance of manhattan nonnegative matrix factorization. *IEEE Transactions on Neural Networks and Learning Systems* (2015), 1851–1863.

[22] Yiming Liu, Xuezhi Cao, and Yong Yu. 2016. Are You Influenced by Others When Rating?: Improve Rating Prediction by Conformity Modeling. In *RecSys*. 269–272.

[23] Xuying Meng, Suhang Wang, Kai Shu, Jundong Li, Bo Chen, Huan Liu, and Yujun Zhang. 2018. Personalized privacy-preserving social recommendation. In *AAAI*.

[24] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. 2018. *Foundations of Machine Learning.* MIT Press.

[25] Jerzy Splawa Neyman. 1990. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statist. Sci.* 5 (1990), 465–472.

[26] Yulei Niu, Kaihua Tang, Hanwang Zhang, Zhiwu Lu, Xian-Sheng Hua, and Ji-Rong Wen. 2021. Counterfactual vqa: A cause-effect look at language bias. In *CVPR*. 12700–12710.

[27] Judea Pearl. 2009. *Causality.* Cambridge university press.

[28] Zhen Qin, Suming J. Chen, Donald Metzler, Yongwoo Noh, Jingzheng Qin, and Xuanhui Wang. 2020. Attribute-based Propensity for Unbiased Learning in Recommender Systems: Algorithm and Case Studies. In *KDD*. 2359–2367.

[29] Paul R. Rosenbaum. 2020. *Design of Observational Studies.* Springer Series in Statistics.

[30] D. B. Rubin. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational psychology* 66 (1974), 688–701.

[31] Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. 2020. Unbiased recommender learning from missing-not-at-random implicit feedback. In *WSDM*. 501–509.

[32] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *ICML*. 1670–1679.

[33] Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms.* Cambridge University Press.

[34] Nathan Srebro, Jason Rennie, and Tommi Jaakkola. 2004. Maximum-margin matrix factorization. *NeurIPS* (2004).

[35] Qinyong Wang, Hongzhi Yin, Tong Chen, Junliang Yu, Alexander Zhou, and Xiangliang Zhang. 2021. Fast-adapting and privacy-preserving federated recommender system. *The International Journal on Very Large Data Bases* (2021), 1–20.

[36] Wenjie Wang, Fuli Feng, Xiangnan He, Hanwang Zhang, and Tat-Seng Chua. 2021. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue. In *SIGIR*. 1288–1297.

[37] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2019. Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random. In *ICML*. 6638–6647.

[38] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2021. Combating Selection Biases in Recommender Systems with a Few Unbiased Ratings. In *WSDM*. 427–435.

[39] Yixin Wang, Dawen Liang, Laurent Charlin, and David M Blei. 2020. Causal inference for recommender systems. In *RecSys*. 426–431.

[40] Tianxin Wei, Fuli Feng, Jiawei Chen, Ziwei Wu, Jinfeng Yi, and Xiangnan He. 2021. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system. In *KDD*. 1791–1800.

[41] Peng Wu, Haoxuan Li, Yuhao Deng, Wenjie Hu, Quanyu Dai, Zhenhua Dong, Jie Sun, Rui Zhang, and Xiao-Hua Zhou. 2022. Causal Analysis Framework for Recommendation. *arXiv:2201.06716* (2022).

[42] Shuyuan Xu, Juntao Tan, Shelby Heinecke, Jia Li, and Yongfeng Zhang. 2021. Deconfounded Causal Collaborative Filtering. *arXiv:2110.07122* (2021).

[43] Mengyue Yang, Quanyu Dai, Zhenhua Dong, Xu Chen, Xiuqiang He, and Jun Wang. 2021. Top-N Recommendation with Counterfactual User Preference Simulation. In *CIKM*. 2342–2351.

[44] Wenhao Zhang, Wentian Bao, Xiao-Yang Liu, Keping Yang, Quan Lin, Hong Wen, and Ramin Ramezani. 2020. Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning. In *WWW*. 2775–2781.

[45] Yang Zhang, Fuli Feng, Xiangnan He, Tianxin Wei, Chonggang Song, Guohui Ling, and Yongdong Zhang. 2021. Causal intervention for leveraging popularity bias in recommendation. In *SIGIR*. 11–20.

[46] Yu Zheng, Chen Gao, Xiang Li, Xiangnan He, Yong Li, and Depeng Jin. 2021. Disentangling User Interest and Conformity for Recommendation with Causal Embedding. In *WWW*. 2980–2991.

[47] Ziwei Zhu, Yun He, Yin Zhang, and James Caverlee. 2020. Unbiased Implicit Recommendation and Propensity Estimation via Combinational Joint Learning. In *RecSys*. 551–556.

A APPENDIX

A.1 Theoretical Proof

LEMMA A.1 (MCDIARMID'S INEQUALITY). *Let $X_1, ..., X_m \in \mathcal{X}^m$ be a set of $m \geq 1$ independent random variables and assume that there exist $c_1, ..., c_m > 0$ such that $f : \mathcal{X}^m \to \mathbb{R}$ satisfies the following conditions:*

$$|f(x_1, ..., x_i, ..., x_m) - f(x_1, ..., x_i', ..., x_m)| \leq c_i,$$

for all $i \in \{1, 2, ..., m\}$ and any points $x_1, ..., x_m, x_i' \in \mathcal{X}$. Let $f(S)$ denote $f(X_1, ..., X_m)$, then for all $\epsilon > 0$, the following inequalities hold:

$$\mathbb{P}[f(S) - \mathbb{E}\{f(S)\} \geq \epsilon] \leq \exp\left(-\frac{-2\epsilon^2}{\sum_{i=1}^m c_i^2}\right)$$

$$\mathbb{P}[f(S) - \mathbb{E}\{f(S)\} \leq -\epsilon] \leq \exp\left(-\frac{-2\epsilon^2}{\sum_{i=1}^m c_i^2}\right).$$

PROOF. The proof can be found in Appendix D.2 of [24]. □

LEMMA A.2 (RADEMACHER COMPARISON LEMMA). *Let $X \in \mathcal{X}$ be a random variable with distribution $\mathbb{P}, X_1, ..., X_m$ be a set of independent copies of X, $\mathcal{G}$ be a class of real-valued functions on $\mathcal{X}$. Then we have*

$$\mathbb{E} \sup_{g \in \mathcal{G}} \left| \frac{1}{m} \sum_{i=1}^m g(X_i) - \mathbb{E}(g(X_i)) \right| \leq 2\mathbb{E}\left[\mathbb{E}_{\sigma \sim \{-1,+1\}^m} \sup_{g \in \mathcal{G}} \left| \frac{1}{m} \sum_{i=1}^m g(X_i)\sigma_i \right| \right],$$

where $\sigma = (\sigma_1, ..., \sigma_m)$ is a Rademacher sequence.

PROOF. The proof can be found in Lemma 26.2 of [33]. □

PROOF OF THEOREM 3.1. We first prove conclusion (b). If $\hat{p}_{u,i} = \tilde{p}_{u,i}$, we have

$$\begin{aligned}
\mathbb{E}\Big(\mathcal{L}_{IPS}(\phi)\Big) &= \mathbb{E}\Big(\frac{o_{u,i}e_{u,i}}{\tilde{p}_{u,i}}\Big) \\
&= \mathbb{E}\Big\{\mathbb{E}\Big(\frac{o_{u,i}e_{u,i}}{\tilde{p}_{u,i}} \mid x_{u,i}, h_{u,i}\Big)\Big\} \\
&= \mathbb{E}\Big\{\frac{\mathbb{E}(o_{u,i}e_{u,i} \mid x_{u,i}, h_{u,i})}{\tilde{p}_{u,i}}\Big\} \\
&= \mathbb{E}\Big\{\frac{\mathbb{E}(o_{u,i} \mid x_{u,i}, h_{u,i})\mathbb{E}(e_{u,i} \mid x_{u,i}, h_{u,i})}{\tilde{p}_{u,i}}\Big\} \\
&= \mathbb{E}\Big\{\frac{\tilde{p}_{u,i}\mathbb{E}(e_{u,i} \mid x_{u,i}, h_{u,i})}{\tilde{p}_{u,i}}\Big\} \\
&= \mathbb{E}[\mathbb{E}(e_{u,i} \mid x_{u,i}, h_{u,i})] \\
&= \mathbb{E}[e_{u,i}] = \mathcal{L}_{ideal}(\phi).
\end{aligned}$$

where the second and the second last equations follow by the law of iterated expectations, the third equation follows from the truth that $\tilde{p}_{u,i}$ is a function of $(x_{u,i}, h_{u,i})$. Since $(x_{u,i}, h_{u,i})$ include all confounders, $o_{u,i} \perp r_{u,i}(1) \mid (x_{u,i}, h_{u,i})$, which leads to $o_{u,i} \perp e_{u,i} \mid (x_{u,i}, h_{u,i})$ and hence the fourth equation holds.

Similarly, if $\hat{p}_{u,i} = \tilde{p}_{u,i}$, $\hat{e}_{u,i} = g_{u,i}$, $\mathbb{E}(\mathcal{L}_{DR}(\phi))$ equals

$$\begin{aligned}
&\mathbb{E}\left[e_{u,i} + \frac{\{\tilde{p}_{u,i} - o_{u,i}\} \cdot \{g_{u,i} - e_{u,i}\}}{\tilde{p}_{u,i}}\right] \\
&= \mathcal{L}_{ideal}(\phi) + \mathbb{E}\left[\frac{\{\tilde{p}_{u,i} - o_{u,i}\} \cdot \{g_{u,i} - e_{u,i}\}}{\tilde{p}_{u,i}}\right] \\
&= \mathcal{L}_{ideal}(\phi) + \mathbb{E}\Big\{\frac{\mathbb{E}(\tilde{p}_{u,i} - o_{u,i} \mid x_{u,i}, h_{u,i})\mathbb{E}(g_{u,i} - e_{u,i} \mid x_{u,i}, h_{u,i})}{\tilde{p}_{u,i}}\Big\} \\
&= 0.
\end{aligned}$$

The last equation holds due to $\mathbb{E}(\tilde{p}_{u,i} - o_{u,i} \mid x_{u,i}, h_{u,i}) = 0$.

Next we prove conclusion (a). If $\hat{p}_{u,i} = p_{u,i}$, note that

$$\mathbb{E}\Big(\mathcal{L}_{IPS}(\phi)\Big) = \mathbb{E}\Big\{\mathbb{E}\Big(\frac{o_{u,i}e_{u,i}}{p_{u,i}} \mid x_{u,i}\Big)\Big\} = \mathbb{E}\Big\{\frac{\mathbb{E}(o_{u,i}e_{u,i} \mid x_{u,i})}{p_{u,i}}\Big\}$$

and

$$\mathcal{L}_{ideal}(\phi) = \mathbb{E}\{\mathbb{E}(e_{u,i} \mid x_{u,i})\} = \mathbb{E}\Big\{\frac{\mathbb{E}(o_{u,i} \mid x_{u,i}) \cdot \mathbb{E}(e_{u,i} \mid x_{u,i})}{p_{u,i}}\Big\}.$$

Since $e_{u,i}$ is associated with $o_{u,i}$ given $x_{u,i}$ due to the hidden confounder, we have $\mathbb{E}(o_{u,i}e_{u,i} \mid x_{u,i}) \neq \mathbb{E}(o_{u,i} \mid x_{u,i}) \cdot \mathbb{E}(e_{u,i} \mid x_{u,i})$, which implies $\mathbb{E}(\mathcal{L}_{IPS}(\phi)) \neq \mathcal{L}_{ideal}(\phi)$.

For DR estimator, if $\hat{p}_{u,i} = p_{u,i}$, $\hat{e}_{u,i} = g_{u,i}$, $\mathbb{E}(\mathcal{L}_{DR}(\phi))$ equals

$$\begin{aligned}
&\mathbb{E}\left[e_{u,i} + \frac{\{p_{u,i} - o_{u,i}\} \cdot \{g_{u,i} - e_{u,i}\}}{p_{u,i}}\right] \\
&= \mathcal{L}_{ideal}(\phi) + \mathbb{E}\left[\frac{\{p_{u,i} - o_{u,i}\} \cdot \{g_{u,i} - e_{u,i}\}}{p_{u,i}}\right] \\
&= \mathcal{L}_{ideal}(\phi) + \mathrm{Cov}((p_{u,i} - o_{u,i})/p_{u,i}, g_{u,i} - e_{u,i}).
\end{aligned}$$

The last equation follows by noting that $\mathbb{E}[(p_{u,i} - o_{u,i})/p_{u,i}] = 0$ and $\mathbb{E}(g_{u,i} - e_{u,i}) = 0$. In the presence of hidden confounder, $\mathrm{Cov}((p_{u,i} - o_{u,i})/p_{u,i}, g_{u,i} - e_{u,i}) \neq 0$. This completes the proof. □

PROOF OF THEOREM 4.1. For any $f_\phi \in \mathcal{F}$, since $\tilde{w}_{u,i} \in [\hat{a}_{u,i}, \hat{b}_{u,i}]$, we have

$$\begin{aligned}
&\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{RD-IPS}(\phi) \\
&\leq \mathcal{L}_{ideal}(\phi) - \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} o_{u,i}e_{u,i}\tilde{w}_{u,i} \\
&\leq \sup_{f_\phi \in \mathcal{F}} \Big(\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} e_{u,i}\{1 - o_{u,i}\tilde{w}_{u,i}\}\Big) \\
&\stackrel{def}{=} [B - \mathbb{E}(B)] + \mathbb{E}(B),
\end{aligned}$$

where $B = \sup_{f_\phi \in \mathcal{F}} \Big(\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} e_{u,i}\{1 - o_{u,i}\tilde{w}_{u,i}\}\Big)$.

Note that $|e_{u,i}\{1-o_{u,i}\tilde{w}_{u,i}\}| \leq C_1(C_2+1)$, applying McDiarmid's inequality yields that

$$\mathbb{P}\Big(B - \mathbb{E}(B) \geq \epsilon\Big) \leq \exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_1^2(C_2+1)^2}\Big\}$$

Next we consider $\mathbb{E}(B)$. Since $\mathbb{E}[e_{u,i}(1-o_{u,i}\tilde{w}_{u,i})] = 0$. By Lemma A.2,

$$\begin{aligned}
\mathbb{E}(B) &\leq 2\mathbb{E}\Big\{\sup_{f_\phi \in \mathcal{F}} \mathbb{E}_{\sigma \sim \{-1,+1\}^{|\mathcal{D}|}} \sigma_{u,i}e_{u,i}(1 - o_{u,i}\tilde{w}_{u,i})\Big\} \\
&\leq 2(C_2 + 1)\mathbb{E}\{\mathcal{R}(\mathcal{F})\}.
\end{aligned}$$

By McDiarmid's inequality again, we have

$$\mathbb{P}\Big(\mathcal{R}(\mathcal{F}) - \mathbb{E}[\mathcal{R}(\mathcal{F})] \geq \epsilon\Big) \leq \exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_1^2}\Big\}.$$

Letting $\exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_1^2(C_2+1)^2}\Big\} = \eta/2$ leads to that $\epsilon = \sqrt{2}C_1(C_2 + 1)\sqrt{\frac{\log(2/\eta)}{|\mathcal{D}|}}$, and $\exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_1^2}\Big\} = \eta/2$ gives that $\epsilon = \sqrt{2}C_1\sqrt{\frac{\log(2/\eta)}{|\mathcal{D}|}}$.

Then, we have with probability at least $1 - \eta$,

$$\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{RD-IPS}(\phi) \leq 2(C_2+1)\mathcal{R}(\mathcal{F}) + C_1(C_2+1)\sqrt{\frac{18\log(2/\eta)}{|\mathcal{D}|}}.$$

For BRD-IPS estimator with a benchmark $f_{\hat{\phi}^{(0)}}$, the conclusion follows immediately from the truth that

$$
\begin{aligned}
&\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{ideal}(\hat{\phi}^{(0)}) - \mathcal{L}_{BRD-IPS}(\phi) \\
&\leq \mathcal{L}_{ideal}(\phi) - \mathcal{L}_{ideal}(\hat{\phi}^{(0)}) - \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} o_{u,i}\{e_{u,i}(\phi) - e_{u,i}(\hat{\phi}^{(0)})\}\tilde{w}_{u,i} \\
&= \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} e_{u,i}(\phi)\{1 - o_{u,i}\tilde{w}_{u,i}\} - \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} e_{u,i}(\hat{\phi}^{(0)})\{1 - o_{u,i}\tilde{w}_{u,i}\} \\
&\leq 2\sup_{f_\phi\in\mathcal{F}}\Big(\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} e_{u,i}(\phi)\{1 - o_{u,i}\tilde{w}_{u,i}\}\Big).
\end{aligned}
$$

□

Proof of Theorem 4.2. Given the imputed errors $\{\hat{e}_{u,i} : (u,i) \in \mathcal{D}\}$, For any $f_\phi \in \mathcal{F}$, we have

$$
\begin{aligned}
&\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{RD-DR}(\phi,\theta) \\
&\leq \sup_{f_\phi\in\mathcal{F}}\Big(\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}(e_{u,i} - \hat{e}_{u,i})\{1 - o_{u,i}\tilde{w}_{u,i}\}\Big) \\
&\overset{def}{=} B_2.
\end{aligned}
$$

Due to $|(e_{u,i} - \hat{e}_{u,i})\{1 - o_{u,i}\tilde{w}_{u,i}\}| \leq C_3(C_2+1)$, using McDiarmid's inequality leads to that

$$
\mathbb{P}\Big(B_2 - \mathbb{E}(B_2) \geq \epsilon\Big) \leq \exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_3^2(C_2+1)^2}\Big\}
$$

Next we focus on $\mathbb{E}(B_2)$. By Lemma A.2 and the property of Rademacher complexity,

$$
\begin{aligned}
\mathbb{E}(B_2) &\leq 2\mathbb{E}\Big\{\sup_{f_\phi\in\mathcal{F}}\mathbb{E}_{\sigma\sim\{-1,+1\}^{|\mathcal{D}|}}\sigma_{u,i}(e_{u,i} - \hat{e}_{u,i})(1 - o_{u,i}\tilde{w}_{u,i})\Big\} \\
&\leq 2(C_2+1)\mathbb{E}\Big\{\sup_{f_\phi\in\mathcal{F}}\mathbb{E}_{\sigma\sim\{-1,+1\}^{|\mathcal{D}|}}\sigma_{u,i}(e_{u,i} - \hat{e}_{u,i})\Big\} \\
&\leq 2(C_2+1)\mathbb{E}\{\mathcal{R}(\mathcal{F})\}.
\end{aligned}
$$

By McDiarmid's inequality again,

$$
\mathbb{P}\Big(\mathcal{R}(\mathcal{F}) - \mathbb{E}[\mathcal{R}(\mathcal{F})] \geq \epsilon\Big) \leq \exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_1^2}\Big\}.
$$

Letting $\exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_3^2(C_2+1)^2}\Big\} = \eta/2$ and $\exp\Big\{-\frac{\epsilon^2|\mathcal{D}|}{2C_1^2}\Big\} = \eta/2$, we have with probability at least $1-\eta$,

$$
\begin{aligned}
&\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{RD-DR}(\phi) \\
&\leq 2(C_2+1)\mathcal{R}(\mathcal{F}) + (2C_1 + C_3)(C_2+1)\sqrt{\frac{2\log(2/\eta)}{|\mathcal{D}|}}.
\end{aligned}
$$

The result of BRD-DR estimator with a benchmark $f_{\hat{\phi}^{(0)}}$ holds by noting that

$$
\begin{aligned}
&\mathcal{L}_{ideal}(\phi) - \mathcal{L}_{ideal}(\hat{\phi}^{(0)}) - \mathcal{L}_{BRD-DR}(\phi,\theta) \\
&\leq \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\{e_{u,i}(\phi) - \hat{e}_{u,i}(\theta)\}\{1 - o_{u,i}\tilde{w}_{u,i}\} + \\
&\quad \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\{e_{u,i}(\hat{\phi}^{(0)}) - \hat{e}_{u,i}(\theta^{(0)})\}\{1 - o_{u,i}\tilde{w}_{u,i}\} \\
&\leq 2\sup_{f_\phi\in\mathcal{F}}\Big(\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\{e_{u,i}(\phi) - \hat{e}_{u,i}(\theta)\}\{1 - o_{u,i}\tilde{w}_{u,i}\}\Big).
\end{aligned}
$$

□

A.2 Algorithm

Algorithm 2: Robust Deconfounder DR (RD-DR)

Input: Data $\mathcal{D}$, nominal propensity score $p_{u,i}$.
Output: An optimized recommender model.

```
Initialize a recommender model with parameters φ;
Initialize an imputation model with parameters θ;
Calculate the bound of true propensity score based on
  Eq. (10) with hyper-parameter Γ;
Generate the uncertainty set 𝒲 = {W} based on Eq. (11)
while Stop condition is not reached do
    Fetch (u, i) from 𝒟;
    Calculate the prediction error e_{u,i} based on φ and Eq. (3);
    Calculate ê_{u,i} = g_θ(x_{u,i}) based on θ;
    Maximize the loss Eq. (13) to update W;
    Minimize the loss Eq. (13) to optimize φ;
    Maximize the loss Eq. (14) to update W;
    Minimize the loss Eq. (14) to optimize θ;
end
Return an optimized recommender model with φ;
```

Algorithm 3: Benchmarked Robust Deconfounder IPS

Input: Data $\mathcal{D}$, nominal propensity score $p_{u,i}$.
Output: An optimized recommender model.

```
Use 𝒟 and p_{u,i} to train an IPS model by minimizing Eq. (4),
  and fix its model parameters φ̂^(0) as the benchmark;
Initialize a recommender model with model parameters φ;
Calculate the bound of true propensity score based on
  Eq. (10) with hyper-parameter Γ;
Generate the uncertainty set 𝒲 = {W} based on Eq. (11)
while Stop condition is not reached do
    Fetch (u, i) from 𝒟;
    Calculate the prediction error e_{u,i}(φ) based on φ and
      Eq. (3);
    Use the pre-trained IPS model to calculate e_{u,i}(φ̂^(0))
      based on φ̂^(0) and Eq. (3);
    Maximize the loss Eq. (15) to update W;
    Minimize the loss Eq. (15) to optimize φ;
end
Return an optimized recommender model with φ;
```