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 Robust Deconfounder (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 Benchmarked 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
Recommender, Unmeasured confounder, Deconfounder

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM Reference Format:

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
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 $U = \{u\}$ and $I = \{i\}$ 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 $D = U \times I$.
- **Feature**: the feature $x_{u,i}$ describes user-item pair $(u, i)$.
- **Treatment**: $a_{u,i} \in \{1, 0\}$. It is the exposure status of $(u, i)$, where $a_{u,i} = 1$ 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 $a_{u,i}$ had been set to $o$.

Let $P$ and $E$ be the distribution and expectation on the target population, and $O = \{(u, i) \in D, a_{u,i} = 1\}$ be the set of exposed units. In recommendation, the target estimand [12, 16, 41] is $E(r_{u,i}(1) \mid x_{u,i})^{3}$, 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 $u, i$, we have

$$o_{u,i} \perp r_{u,i}(1) \mid (x_{u,i}, h_{u,i}), \quad o_{u,i} \perp r_{u,i}(1) \mid x_{u,i},$$

(1)

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

---

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

2In 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].
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. In this section, we first show that the existing propensity-based methods are biased in the presence of unmeasured confounders. To remove all the confounding bias, one can never accurately estimate the nominal propensity score for the exposed events.

**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 L2 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 \( p_{u,i} \), denoted as \( \hat{p}_{u,i} \), the IPS estimator of the prediction inaccuracy is presented as

\[
L_{\text{IPS}}(\phi) = \frac{1}{|D|} \sum_{(u,i) \in D} e_{u,i} \hat{p}_{u,i}.
\]

**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 \( g_{u,i} = \mathbb{E}[e_{u,i} \mid x_{u,i}] \). Given the learned \( \hat{p}_{u,i} \) and \( \hat{e}_{u,i} \), the DR estimator is

\[
L_{\text{DR}}(\phi, \theta) = \frac{1}{|D|} \sum_{(u,i) \in D} \left[ \hat{e}_{u,i} + \frac{o_{u,i} \hat{e}_{u,i} - \hat{e}_{u,i}}{\hat{p}_{u,i}} \right].
\]

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

\[
L_{\text{DR}^+}(\phi, \theta) = \frac{1}{|D|} \sum_{(u,i) \in D} \left( \hat{e}_{u,i} - e_{u,i} \right)^2 / \hat{p}_{u,i}.
\]

In this paper estimator and loss function are interchangeable.

**3.2 Robust Deconfounder Framework**

Sensitivity analysis [9, 29] is a basic tool to assess the robustness of 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 \( p_{u,i} \) 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

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

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

\[
1 \leq \frac{1 - p_{u,i}}{\bar{p}_{u,i}(1 - \bar{p}_{u,i})} \leq \Gamma.
\]

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

\[
a_{u,i} \leq \bar{w}_{u,i} \leq b_{u,i},
\]

\[
a_{u,i} = 1 + (1/p_{u,i} - 1)/\Gamma, \quad b_{u,i} = 1 + (1/p_{u,i} - 1)\Gamma.
\]

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

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

where \( W = \{ w_{u,i} : (u,i) \in D \} \), \( \hat{a}_{u,i} \) and \( \hat{b}_{u,i} \) are the estimates of \( a_{u,i} \) and \( b_{u,i} \). and the parameters \( \phi \) and \( \theta \) are estimated by solving Eq. (5) and Eq. (6) alternately. Without unmeasured confounders, it is well known that \( L_{\text{IPS}}(\phi) \) 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, (a) both the IPS and DR estimators are biased, even \( \hat{p}_{u,i} \) and \( \hat{e}_{u,i} \) estimate \( p_{u,i} \) and \( e_{u,i} \) accurately.

(b) if we define the true propensity score as

\[
\hat{p}_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}, h_{u,i}),
\]

and assume that \( \hat{p}_{u,i} \) is an accurate estimate of \( p_{u,i} \), 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 \( p_{u,i} \) due to the full missingness of unmeasured confounders without imposing strong assumptions.
Algorithm 1: Robust Deconfounder IPS (RD-IPS)

**Input:** Data \( \mathcal{D} \), nominal propensity score \( p_{u,i} \).
**Output:** An optimized recommender model.
1. Initialize a recommender model with parameters \( \phi \).
2. Calculate the bound of true propensity score based on Eq. (10) with hyper-parameter \( \Gamma \).
3. Generate the uncertainty set \( \mathcal{W} = [W] \) based on Eq. (11).
4. while Stop condition is not reached do
   5. Fetch \( (u, i) \) from \( \mathcal{D} \).
   6. Calculate the prediction error \( e_{u,i} \) based on \( \phi \) and Eq. (3).
   7. Maximize the loss Eq. (12) to update \( W \).
   8. Minimize the loss Eq. (12) to optimize \( \phi \).
9. end
10. Return an optimized recommender model with \( \phi \).

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

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

(12)

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

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

(13)

and the corresponding loss function for the imputation model

\[
\mathcal{L}_{\text{Imp}}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{O}} (\hat{e}_{u,i} - e_{u,i})^2 w_{u,i}.
\]

(14)

The parameters \( \phi \) and \( \theta \) are solved alternately: given \( \hat{\theta} \), \( \phi \) is updated by minimizing (13); given \( \hat{\phi} \), \( \theta \) 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 \( p_{u,i} \) 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 \( h, \sigma, \) 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 \( \phi \). Suppose we have access to an estimator of \( \phi \), denoted as \( \hat{\phi}(0) \), via using the existing approaches. Then the corresponding BRD-IPS estimator is

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

(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 \( \phi \). Analogously, the BRD-DR estimator and the associated imputation model are given as

\[
\mathcal{L}_{\text{BRD-DR}}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[ [\hat{e}_{u,i}(\theta) - e_{u,i}(\hat{\theta}(0))] + \right. \\
\left. \left[ o_{u,i}(e_{u,i}(\phi) - \hat{e}_{u,i}(\theta)) - o_{u,i}(e_{u,i}(\hat{\phi}(0)) - \hat{e}_{u,i}(\hat{\theta}(0))) \right] \right] w_{u,i},
\]

\[
\mathcal{L}_{\text{Imp}}(\phi, \theta) = \max_{W \in \mathcal{W}} \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{O}} \left[ \hat{e}_{u,i}(\theta) - e_{u,i}(\hat{\phi}(\theta)) \right]^2 w_{u,i}.
\]

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

The proposed BRD-IPS/DR approaches are very flexible due to the free choice of the benchmark estimators. In addition, note that \( \mathcal{L}_{\text{BRD-IPS}}(\phi) = 0 \) if \( \phi = \hat{\phi}(0) \), which implies that the minimum value of \( \mathcal{L}_{\text{BRD-IPS}}(\phi) \) is non-positive. Similarly, it is not hard to derive that \( \min_{\phi} \mathcal{L}_{\text{BRD-DR}}(\phi) \leq 0 \). 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

\[
\hat{\phi}^\dagger = \arg \min_{\phi} \mathcal{L}_{\text{BRD-IPS}}(\phi), \quad \hat{\phi}^\ddagger = \arg \min_{\phi} \mathcal{L}_{\text{BRD-DR}}(\phi).
\]

(16)

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

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

\[
\mathcal{L}_{\text{Ideal}}(\hat{\phi}^\dagger) < \mathcal{L}_{\text{Ideal}}(\hat{\phi}(0)).
\]

(17)

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

\[
\mathcal{L}_{\text{Ideal}}(\hat{\phi}^\ddagger) < \mathcal{L}_{\text{Ideal}}(\hat{\phi}(0)).
\]

(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

\[
L_{RD-Auto} (\phi) = \max_{W \in W} \left\{ \frac{1}{|D|} \sum_{(u,i) \in D} [ \epsilon_{u,i} \phi_{u,i} + L_{Auto} (\phi)] \right\},
\]

where \( L_{Auto} (\phi) \) is the loss that measures the distance between model predictions and the imputation ratings. Moreover, the RD AutoDebias (BRD-Auto) estimator is presented as

\[
L_{BRD-Auto} (\phi) = \max_{W \in W} \left\{ \frac{1}{|D|} \sum_{(u,i) \in D} \left[ L_{Auto} (\phi) - L_{Imp} (\phi) \right] + \epsilon_{u,i} (\phi_{u,i} - \phi(0)) \right\}.
\]

### 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

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

where \( \sigma = \{ \sigma_{u,i} : (u,i) \in D \} \) is a Rademacher sequence [24]. Assume that \( \mathcal{F} \) has a vanishing complexities, i.e., \( \mathcal{R}(\mathcal{F}) \to 0 \) as \( |D| \to \infty \). 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 \( \tilde{\omega}_{u,i} \in [\tilde{a}_{u,i}, \tilde{b}_{u,i}] \), \( \epsilon_{u,i} \leq C_1 \), and \( \tilde{\omega}_{u,i} \leq C_2 \) for all \( (u, i) \) pairs. Then for any \( f_{\phi} \in \mathcal{F} \) and \( \eta > 0 \), we have that with probability at least \( 1 - \eta \),

\[
L_{ideal}(\phi) \leq L_{RD-IPS}(\phi) + B(\eta, D, \mathcal{F}), \tag{20}
\]

\[
L_{ideal}(\phi) - L_{ideal}(\phi^{(0)}) \leq L_{BRD-IPS}(\phi) + 2B(\eta, D, \mathcal{F}). \tag{21}
\]

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

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

Eq. (20) and Eq. (21) hold for any \( \phi \). In particular, let \( \phi^* = \arg \min_{\phi} L_{RD-IPS}(\phi) \), 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 \( B(\eta, D, \mathcal{F}) \to 0 \) as \( |D| \to \infty \). In addition, Theorem 3.2(a) can be deduced by Eq. (21). For \( \phi^\dagger \) defined in Eq. (16), if \( L_{BRD-IPS}(\phi^\dagger) < 0 \), then we have \( L_{BRD-IPS}(\phi^\dagger) \leq -\epsilon \eta \) for a positive constant \( \epsilon \). When \( |D| \) is large enough, such that \( B(\eta, D, \mathcal{F}) < \epsilon / 2 \). Then Theorem 3.2(a) follows immediately from the fact 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.

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

\[
L_{ideal}(\phi) \leq L_{RD-DR}(\phi) + 2B_2(\eta, D, \mathcal{F}),
\]

\[
L_{ideal}(\phi) - L_{ideal}(\phi^{(0)}) \leq L_{BRD-DR}(\phi) + 2B_2(\eta, D, \mathcal{F}),
\]

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

\[
B_2(\eta, D, \mathcal{F}) = 2(C_2 + 1)\mathcal{R}(\mathcal{F}) + (2C_1 + C_3)\mathcal{R}(\mathcal{F}) \sqrt{\frac{18 \log(2/\eta)}{|D|}}.
\]

### 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\(^\dagger\), 2) Coat\(^\dagger\), and 3) Product\(^\dagger\), 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 collected from a random 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% randomly (See Table. 1).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#User</th>
<th>#Item</th>
<th>#Biased Data</th>
<th>#Unbiased Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yahoo!R3</td>
<td>15,400</td>
<td>1,000</td>
<td>311,704</td>
<td>54,000</td>
</tr>
<tr>
<td>Coat</td>
<td>290</td>
<td>300</td>
<td>6,960</td>
<td>4,640</td>
</tr>
<tr>
<td>Product</td>
<td>7,176</td>
<td>10,729</td>
<td>1,325,197</td>
<td>270,180</td>
</tr>
</tbody>
</table>

\(^\dagger\)https://webscope.sandbox.yahoo.com/.

\(^\dagger\)https://www.cs.cornell.edu/~schnabts/mnar/.

\(^\dagger\)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.

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

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 propensity 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 hyperparameter 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. RD-IPS, RD-DR, and RD-AutoDebias as RD-IPS, RD-DR, and RD-AutoDebias respectively.

**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 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 ensures 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 than RD. In these cases, RD ranks the 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 \( p_{ui} \) to construct the estimator

9https://github.com/Dingsewhole/Robust_Deconfounder_master/
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 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.
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 \( W \). 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 \( W \) 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 \((\tilde{b}_{u,i} - \tilde{a}_{u,i})/2\). 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 \( (i.e., \text{Eq. (13)} \text{ and Eq. (14))}. \) 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., \text{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 hyperparameter $\Gamma$ 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 (61972572, U21B0206, 62121002), and the CCCD Key Lab of Ministry of Culture and Tourism.

REFERENCES

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

A.1 Theoretical Proof

Lemma A.1 (McDiarmid’s inequality). Let $X_1, \ldots, X_m \in \mathbb{X}^m$ be a set of $m \geq 1$ independent random variables and assume that there exist $c_1, \ldots, c_m > 0$ such that $f : \mathbb{X}^m \to \mathbb{R}$ satisfies the following conditions:

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

for all $i \in \{1, 2, \ldots, m\}$ and any points $x_1, \ldots, x_m, x_i' \in \mathbb{X}$. Let $f(S)$ denote $f(X_1, \ldots, X_m)$, 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 $X \in \mathbb{X}$ be a random variable with distribution $\mathbb{P}$, $X_1, \ldots, X_m$ be a set of independent copies of $X$, $\mathcal{G}$ be a class of real-valued functions on $X$. Then we have

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

where $\sigma = (\sigma_1, \ldots, \sigma_m)$ 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 $\hat{p}_{u,i} = \hat{p}_{u,i}$, we have

$$\mathbb{E}[L_{IPS}(\phi)] = \mathbb{E}\left[\frac{\sum_{u,i} e_{u,i}}{p_{u,i}}\right]$$

$$= \mathbb{E}\left[\frac{\sum_{u,i} e_{u,i} | x_{u,i}, h_{u,i}}{p_{u,i}}\right]$$

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

Similarly, if $\hat{p}_{u,i} = \hat{p}_{u,i}, \hat{e}_{u,i} = g_{u,i}, \mathbb{E}[L_{DR}(\phi)]$ equals

$$\mathbb{E}[L_{ideal}(\phi)] + \mathbb{E}\left[\frac{\sum_{u,i} e_{u,i} | x_{u,i}, h_{u,i}}{p_{u,i}}\right]$$

$$= \mathbb{E}[\hat{p}_{u,i} \mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$= \mathbb{E}[\hat{p}_{u,i} [e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$= 0.$$

The last equation holds due to $\mathbb{E}[\hat{p}_{u,i} - o_{u,i} | x_{u,i}, h_{u,i}] = 0$.

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

$$\mathbb{E}\left[L_{IPS}(\phi)\right] = \mathbb{E}\left[\frac{\sum_{u,i} e_{u,i} | x_{u,i}}{p_{u,i}}\right]$$

$$\leq \mathbb{E}\left[\frac{\sum_{u,i} e_{u,i} | x_{u,i}}{p_{u,i}}\right]$$

$$= \mathbb{E}[\mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$= \mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]$$

$$= \mathbb{E}[\hat{p}_{u,i} \mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$= \mathbb{E}[\hat{p}_{u,i} [e_{u,i} | x_{u,i}, h_{u,i}]]$$

The last equation holds due to $\mathbb{E}[\hat{p}_{u,i} - o_{u,i} | x_{u,i}, h_{u,i}] = 0$.

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

$$\mathbb{E}\left[L_{IPS}(\phi)\right] = \mathbb{E}\left[\frac{\sum_{u,i} e_{u,i} | x_{u,i}}{p_{u,i}}\right]$$

$$\leq \mathbb{E}[\mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$= \mathbb{E}[\hat{p}_{u,i} \mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$= \mathbb{E}[\hat{p}_{u,i} [e_{u,i} | x_{u,i}, h_{u,i}]]$$

The last equation holds due to $\mathbb{E}[\hat{p}_{u,i} - o_{u,i} | x_{u,i}, h_{u,i}] = 0$. By Lemma A.2,

$$\mathbb{E}[\hat{p}_{u,i} \mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

$$\leq \mathbb{E}[\mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

By McDiarmid’s inequality again, we have

$$\mathbb{E}[\mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$$

Letting $\epsilon = \sqrt{2} C_1$ gives that $\mathbb{E}[\mathbb{E}[e_{u,i} | x_{u,i}, h_{u,i}]]$
For BRD-IPS estimator with a benchmark \( f_{\hat{\phi}(0)} \), the conclusion follows immediately from the truth that
\[
L_{\text{ideal}}(\phi) - L_{\text{ideal}}(\hat{\phi}(0)) - L_{\text{BRD-IPS}}(\phi)
\leq \frac{1}{|D|} \sum_{(u,i) \in D} e_{u,i}(\phi)(1 - o_{u,i} \hat{w}_{u,i}) - \frac{1}{|D|} \sum_{(u,i) \in D} e_{u,i}(\hat{\phi}(0))(1 - o_{u,i} \hat{w}_{u,i})
\leq 2 \sup_{f_{\hat{\phi}} \in F} \left( \frac{1}{|D|} \sum_{(u,i) \in D} e_{u,i}(\phi)(1 - o_{u,i} \hat{w}_{u,i}) \right).
\]

Proof of Theorem 4.2. Given the imputed errors \( \{\hat{e}_{u,i} : (u,i) \in D\} \). For any \( f_{\hat{\phi}} \in F \), we have
\[
L_{\text{ideal}}(\phi) - L_{RD-DR}(\phi, \theta)
\leq \sup_{f_{\hat{\phi}} \in F} \left( \frac{1}{|D|} \sum_{(u,i) \in D} (e_{u,i} - \hat{e}_{u,i})(1 - o_{u,i} \hat{w}_{u,i}) \right)
\equiv 2B_2.
\]
Due to \( \{e_{u,i} - \hat{e}_{u,i} : (u,i) \in D\} \leq C_3(C_2 + 1) \), using McDiarmid’s inequality leads to that
\[
P(B_2 - \mathbb{E}(B_2) \geq \epsilon) \leq \exp \left\{ - \frac{\epsilon^2 |D|}{2C_3^2(C_2 + 1)^2} \right\}
\]
Next we focus on \( \mathbb{E}(B_2) \). By Lemma A.2 and the property of Rademacher complexity,
\[
\mathbb{E}(B_2) \leq 2\mathbb{E} \left( \sup_{f_{\hat{\phi}} \in F} \mathbb{E}_{\sigma \sim \{-1,1\}^{|D|}} \sigma_{u,i}(e_{u,i} - \hat{e}_{u,i})(1 - o_{u,i} \hat{w}_{u,i}) \right)
\leq 2(C_2 + 1)\mathbb{E} \left( \sup_{f_{\hat{\phi}} \in F} \mathbb{E}_{\sigma \sim \{-1,1\}^{|D|}} \sigma_{u,i}(e_{u,i} - \hat{e}_{u,i}) \right)
\leq 2(C_2 + 1)\mathbb{E}(R(F))
\]
By McDiarmid’s inequality again,
\[
P\left( R(F) - \mathbb{E}[R(F)] \geq \epsilon \right) \leq \exp \left\{ - \frac{\epsilon^2 |D|}{4C_3^2} \right\}
\]
Letting \( \epsilon \) be \( \frac{\eta}{2} \) and \( \epsilon \) be \( \frac{\eta}{2} \), we have with probability at least \( 1 - \eta \),
\[
L_{\text{ideal}}(\phi) - L_{RD-DR}(\phi)
\leq 2(C_2 + 1)R(F) + (2C_1 + C_3)(C_2 + 1)\sqrt{\frac{2 \log(2/\eta)}{|D|}}.
\]
The result of BRD-DR estimator with a benchmark \( f_{\hat{\phi}(0)} \) holds by noting that
\[
L_{\text{ideal}}(\phi) - L_{\text{ideal}}(\hat{\phi}(0)) - L_{BRD-IPS}(\phi)
\leq \frac{1}{|D|} \sum_{(u,i) \in D} \{e_{u,i}(\phi) - \hat{e}_{u,i}(\theta)\}(1 - o_{u,i} \hat{w}_{u,i}) +
\frac{1}{|D|} \sum_{(u,i) \in D} \{e_{u,i}(\hat{\phi}(0)) - \hat{e}_{u,i}(\theta)\}(1 - o_{u,i} \hat{w}_{u,i})
\leq 2 \sup_{f_{\hat{\phi}} \in F} \left( \frac{1}{|D|} \sum_{(u,i) \in D} \{e_{u,i}(\phi) - \hat{e}_{u,i}(\theta)\}(1 - o_{u,i} \hat{w}_{u,i}) \right).
\]

### A.2 Algorithm

**Algorithm 2: Robust Deconfounder DR (RD-DR)**

**Input:** Data \( D \), nominal propensity score \( p_{u,i} \).

**Output:** An optimized recommender model.

1. Initialize a recommender model with parameters \( \phi \);
2. Initialize an imputation model with parameters \( \theta \);
3. Calculate the bound of true propensity score based on Eq. (10) with hyper-parameter \( \Gamma \);
4. Generate the uncertainty set \( \mathcal{W} = \{\mathcal{W}\} \) based on Eq. (11);
5. **while** Stop condition is not reached **do**
   6. Fetch \((u,i)\) from \( D \);
   7. Calculate the prediction error \( e_{u,i}(\hat{\phi}) \) based on \( \phi \) and Eq. (3);
   8. Calculate \( \hat{e}_{u,i} = g(\hat{x}_{u,i}) \) based on \( \theta \);
   9. Maximize the loss Eq. (13) to update \( W \);
   10. Minimize the loss Eq. (13) to optimize \( \phi \);
   11. Maximize the loss Eq. (14) to update \( W \);
   12. Minimize the loss Eq. (14) to optimize \( \theta \);
5. **end**
6. Return an optimized recommender model with \( \phi \);

**Algorithm 3: Benchmarked Robust Deconfounder IPS**

**Input:** Data \( D \), nominal propensity score \( p_{u,i} \).

**Output:** An optimized recommender model.

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