



Debiased Recommendation with Noisy Feedback

Haoxuan Li
Peking University
Beijing, China
hxli@stu.pku.edu.cn

Chunyuan Zheng
Peking University
Beijing, China
zhengchunyuan99@gmail.com

Wenjie Wang
National University of Singapore
Singapore, Singapore
wenjiawang96@gmail.com

Hao Wang
Zhejiang University
Hangzhou, China
haohaow@zju.edu.cn

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

Xiao-Hua Zhou*
Peking University
Beijing, China
azhou@math.pku.edu.cn

Abstract

Ratings of a user to most items in recommender systems are usually missing not at random (MNAR), largely because users are free to choose which items to rate. To achieve unbiased learning of the prediction model under MNAR data, three typical solutions have been proposed, including error-imputation-based (EIB), inverse-propensity-scoring (IPS), and doubly robust (DR) methods. However, these methods ignore an alternative form of bias caused by the inconsistency between the observed ratings and the users' true preferences, also known as noisy feedback or outcome measurement errors (OME), *e.g.*, due to public opinion or low-quality data collection process. In this work, we study intersectional threats to the unbiased learning of the prediction model from data MNAR and OME in the collected data. First, we design OME-EIB, OME-IPS, and OME-DR estimators, which largely extend the existing estimators to combat OME in real-world recommendation scenarios. Next, we theoretically prove the unbiasedness and generalization bound of the proposed estimators. We further propose an alternate denoising training approach to achieve unbiased learning of the prediction model under MNAR data with OME. Extensive experiments are conducted on three real-world datasets and one semi-synthetic dataset to show the effectiveness of our proposed approaches. The code is available at <https://github.com/haoxuanli-pku/KDD24-OME-DR>.

CCS Concepts

• Information systems → Recommender systems.

Keywords

Bias, Debias, Noisy Feedback, Recommender Systems

ACM Reference Format:

Haoxuan Li, Chunyuan Zheng, Wenjie Wang, Hao Wang, Fuli Feng, and Xiao-Hua Zhou. 2024. Debiased Recommendation with Noisy Feedback. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and*

*Corresponding author

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 the author(s) 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.

KDD '24, August 25–29, 2024, Barcelona, Spain.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0490-1/24/08

<https://doi.org/10.1145/3637528.3671915>

Data Mining (KDD '24), August 25–29, 2024, Barcelona, Spain. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3637528.3671915>

1 Introduction

Recommender systems (RS) are designed to generate meaningful recommendations to a collection of users for items or products that might interest them, which have made a number of significant advancements in recent years [53, 72, 85, 88]. Nevertheless, direct use of these advanced models in real-world scenarios while ignoring the presence of numerous biases in the collected data can lead to sub-optimal performance of the rating prediction model [6, 80, 81]. Among these, the problem of missing data is particularly prevalent, as users are free to choose items to rate and the ratings with a lower value are more likely to be missing [12], leading to the collected data in RS is always missing not at random (MNAR) [41]. MNAR ratings pose a serious challenge to the unbiased evaluation and learning of the prediction model because the observed data might not faithfully represent the entirety of user-item pairs [59, 73].

To tackle this problem, previous studies evaluate the performance of a prediction model by computing the *prediction inaccuracy*: the average of the prediction errors (*e.g.*, the squared difference between a predicted rating and the *potentially observed* rating) for all ratings [54, 56]. To unbiasedly estimate the prediction inaccuracy when the ratings are partially observable, three typical approaches have been proposed, including: (1) The error-imputation-based (EIB) approaches [18, 65], which compute an imputed error for each missing rating. (2) The inverse-propensity-scoring (IPS) approaches [58, 59], which inversely weight the prediction error for each observed rating with the probability of observing that rating. (3) The doubly robust (DR) approaches [28–30, 64], which use both the error imputation model and the propensity model to estimate the prediction inaccuracy, and the estimation is unbiased when either the imputed errors or the learned propensities are accurate.

Despite the widespread use of these methods, they ignore an alternative form of bias caused by the inconsistency between the observed ratings and the users' true preferences, also known as noisy feedback or outcome measurement errors (OME). Similar to selection bias, OME also arises from systematic bias during the data collection. For example, the collected user feedback may differ from the true user preferences due to the influence of public opinions [89]. Meanwhile, low-quality data collection such as recommender attacks [16] or carelessly filling out the after-sales assessment can also result in noisy user feedback [71]. Therefore, to make current debiased recommendation techniques more applicable to real-world

scenarios that may contain OME, it is worthwhile to develop new model evaluation criteria and corresponding estimators to achieve unbiased learning under OME.

In this work, we study intersectional threats to the unbiased learning of the prediction model introduced by data MNAR and OME from historical interactions, where the observed ratings are not only MNAR but may differ from the true ratings due to the presence of OME. In such a context, a natural evaluation criterion for the prediction models is proposed, called *true prediction inaccuracy*: the average of the *true* prediction errors (e.g., the squared difference between a predicted rating and the *true* rating) for all ratings. Given the collected data with OME, the true prediction errors are difficult to obtain even for the observed ratings, making the previous debiasing estimators severely biased in estimating the true prediction inaccuracy and training the rating prediction model.

To combat the influence of OME on the performance of the prediction model, many data-driven error parameter estimation methods are developed in recent machine learning literature [45, 60, 61, 79]. Meanwhile, given knowledge of measurement error parameters, recent studies propose unbiased risk minimization approaches for learning under noisy labels [9, 44, 49, 67]. Despite the prevalence of OME in real-world recommendation scenarios, there is still only limited work focusing on denoising in RS [4, 38, 50, 82]. For example, [87] proposes to overcome noisy confounders by leveraging the sensitivity analysis in the statistics literature. By noticing that noisy feedback typically has large loss values in the early stages, [71] proposes adaptive denoising training and [15] proposes self-guided denoising learning for implicit feedback. However, most of these methods are heuristic and lack theoretical guarantees of statistical unbiasedness for estimating the true prediction inaccuracy.

To address the above problem, we extend the widely adopted EIB, IPS, and DR estimators to achieve unbiased estimations with the true prediction inaccuracy under OME, named OME-EIB, OME-IPS, and OME-DR estimators, respectively. Our methods are built upon the existing weak separability assumption [42], which states that there exist "perfectly positive" and "perfectly negative" samples among the entire user-item pairs. Specifically, we discuss the rationality of the weak separability assumption in real-world recommendation scenarios and build up the linkage between the observed and true outcome probabilities at specific instances, from which we are able to obtain unbiased estimations of the measurement error parameters. We also derive explicit forms of the biases and the generalization bounds of the proposed OME-EIB, OME-IPS, and OME-DR estimators under inaccurately estimated measurement error parameters, from which we prove the double robustness of the OME-DR estimator. We further propose an alternating denoise training approach to achieve unbiased learning of the prediction model, which corrects for data MNAR and OME in parallel. In particular, the imputation model learns to accurately estimate the prediction errors made by the prediction model, while the prediction model learns from the imputation model to reduce the prediction errors in itself. In this way, the prediction and imputation models mutually regularize each other to reduce both prediction and imputation inaccuracies. The effectiveness of our proposed approaches is validated on three real-world datasets and one semi-synthetic dataset with varying MNAR levels and OME rates. To the best of our knowledge, our holistic evaluation is the first to examine how OME with

its measurement error parameters' estimation in the context of selection bias interact to affect the debiased recommendations.

The contributions of this paper are summarized as follows.

- We formulate OME caused by the inconsistency between the observed ratings and the users' true preferences, and establish an evaluation criterion for the unbiased learning of the prediction model under OME, named true prediction inaccuracy.
- We develop OME-EIB, OME-IPS, and OME-DR estimators for unbiasedly estimating the true prediction inaccuracy of the prediction model under MNAR data with OME, and theoretically analyze the biases and generalization bounds of the estimators.
- We further propose an alternating denoise training approach to estimate the measurement error parameters and achieve unbiased learning of the prediction model under MNAR data with OME, which corrects for data MNAR and OME in parallel.
- We conduct extensive experiments on three real-world datasets and one semi-synthetic dataset, and the results demonstrate the superiority of our methods with varying MNAR and OME rates.

2 Preliminaries

2.1 Task Formulation without OME

Let $\mathcal{U} = \{u_1, \dots, u_N\}$ be a set of users, $\mathcal{I} = \{i_1, \dots, i_M\}$ a set of items, and $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ the collection of all user-item pairs. The potentially observed rating matrix $\mathbf{R} \in \mathbb{R}^{N \times M}$ comprises potentially observed ratings $r_{u,i}$, which represents the observed rating if user u had rated the item i . In RS, given the user-item features $x_{u,i}$, the prediction model $f_\theta(x_{u,i})$ parameterized by θ aims to accurately predict the ratings of all users for all items, and then recommend to the user the items with the highest predicted ratings, as described in [51]. The prediction matrix $\hat{\mathbf{R}} \in \mathbb{R}^{N \times M}$ comprises the predicted ratings $f_\theta(x_{u,i})$ obtained from this prediction model. If the rating matrix \mathbf{R} had been fully observed, then the *prediction inaccuracy* \mathcal{P} of the prediction model can be measured using

$$\mathcal{P} = \mathcal{P}(\hat{\mathbf{R}}, \mathbf{R}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} e_{u,i},$$

where $e_{u,i} = \ell(f_\theta(x_{u,i}), r_{u,i})$ is the prediction error, and $\ell(\cdot, \cdot)$ is a pre-defined loss, such as the mean square error (MSE), i.e., $e_{u,i} = (f_\theta(x_{u,i}) - r_{u,i})^2$. Let $\mathbf{O} \in \{0, 1\}^{N \times M}$ be an indicator matrix, where each entry $o_{u,i}$ is an observation indicator: $o_{u,i} = 1$ if the rating $r_{u,i}$ is observed, and $o_{u,i} = 0$ if the rating $r_{u,i}$ is missing. Given \mathbf{R}^o as the set of the observed entries in the rating matrix \mathbf{R} , the rating prediction model aims to train the prediction model that minimizes the prediction inaccuracy \mathcal{P} . Nonetheless, as users are free to choose items to rate, leading to the collection of observational data that is always missing not at random (MNAR) [12, 41, 83], e.g., the ratings with a lower value are more likely to be missing. Let $\mathcal{O} = \{(u, i) \mid (u, i) \in \mathcal{D}, o_{u,i} = 1\}$ be the set of user-item pairs for the observed ratings, the direct use of the naive estimator $\mathcal{E}_N = \mathcal{E}_N(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{O}} e_{u,i}$ that computes on the observed data would yield severely biased estimation [59, 73].

2.2 Existing Estimators

To unbiasedly estimate the prediction inaccuracy \mathcal{P} given the partially observed ratings \mathbf{R}^o , the error-imputation-based (EIB) approaches [18, 65] compute an imputed error $\hat{e}_{u,i}$ for each missing

rating, and estimate the prediction inaccuracy with

$$\mathcal{E}_{\text{EIB}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} (o_{u,i} e_{u,i} + (1 - o_{u,i}) \hat{e}_{u,i}),$$

which is an unbiased estimator of the prediction inaccuracy when the imputed errors are accurate, *i.e.*, $\hat{e}_{u,i} = e_{u,i}$.

The inverse-propensity-scoring (IPS) approaches [58, 59] first learn $\hat{p}_{u,i}$ as the estimate of the propensity $p_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i})$, *i.e.*, the probability of observing the rating, then inversely weight the prediction error for each observed rating with the learned propensity, and estimate the prediction inaccuracy with

$$\mathcal{E}_{\text{IPS}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}},$$

which is an unbiased estimator of the prediction inaccuracy when the learned propensities are accurate, *i.e.*, $\hat{p}_{u,i} = p_{u,i}$.

The doubly robust (DR) approaches [24–26, 73] use both the error imputation model and the propensity model to estimate the prediction inaccuracy with

$$\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})}{\hat{p}_{u,i}} \right),$$

which is an unbiased estimator of the prediction inaccuracy when either the imputed errors or the learned propensities are accurate, *i.e.*, $\hat{e}_{u,i} = e_{u,i}$ or $\hat{p}_{u,i} = p_{u,i}$, this is also known as double robustness.

3 Methodology

3.1 Task Formulation under OME

Despite many methods have been proposed for achieving unbiased learning to tackle the data MNAR problem [27, 56, 59, 70, 73], they ignore an alternative form of bias caused by the inconsistency between the observed ratings and the users' true preferences, also known as noisy feedback or outcome measurement errors (OME). Both data MNAR and OME arise from systematic bias during the data collection. In RS, two common scenarios that cause incorrect user feedback signals include the influence of public opinions [89], and the low-quality data collection such as recommender attacks [16] or carelessly filling out the after-sales assessments [71]. Formally, we denote \mathbf{R}^* as the user's true preference matrix, with $r_{u,i}^*$ as its entries, which may deviate from the potentially observed ratings $r_{u,i}$. Let $\mathbb{P}(r_{u,i} = 0 \mid r_{u,i}^* = 1) = \rho_{01}$ and $\mathbb{P}(r_{u,i} = 1 \mid r_{u,i}^* = 0) = \rho_{10}$ be the false negative rate and false positive rate, respectively, where $\rho_{01} + \rho_{10} < 1$, then we have

$$\mathbb{P}(r_{u,i} = 1 \mid x_{u,i}) = (1 - \rho_{01}) \cdot \mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}) + \rho_{10} \cdot \mathbb{P}(r_{u,i}^* = 0 \mid x_{u,i}).$$

To make current debiased recommendation techniques more applicable to real-world scenarios that may contain OME, we then establish a new model evaluation criterion for the unbiased learning under OME, named *true prediction inaccuracy*, formally defined as

$$\mathcal{P}^* = \mathcal{P}(\hat{\mathbf{R}}, \mathbf{R}^*) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} e_{u,i}^*,$$

where $e_{u,i}^* = \ell(f_{\theta}(x_{u,i}), r_{u,i}^*)$ is called the *true prediction error*.

3.2 Proposed OME-EIB, OME-IPS, and OME-DR Estimators

To achieve unbiased learning of $r_{u,i}^*$ using MNAR data with OME, a typical class of methods [36, 44] propose the use of a surrogate loss $\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})$ based on the observed labels $r_{u,i}$, which satisfies

$$\mathbb{E}_{r \mid r^*} [\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})] = \ell(f_{\theta}(x_{u,i}), r_{u,i}^*).$$

Considering the cases $r_{u,i}^* = 1$ and $r_{u,i}^* = 0$ separately, we have

$$\begin{aligned} (1 - \rho_{01}) \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 1) + \rho_{01} \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 0) &= \ell(f_{\theta}(x_{u,i}), 1), \\ (1 - \rho_{10}) \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 0) + \rho_{10} \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 1) &= \ell(f_{\theta}(x_{u,i}), 0). \end{aligned}$$

Solving these equations for $\tilde{\ell}(f_{\theta}(x_{u,i}), 1)$ and $\tilde{\ell}(f_{\theta}(x_{u,i}), 0)$ gives

$$\begin{aligned} \tilde{\ell}(f_{\theta}(x_{u,i}), 1) &= \frac{(1 - \rho_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \rho_{01} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \rho_{01} - \rho_{10}}, \\ \tilde{\ell}(f_{\theta}(x_{u,i}), 0) &= \frac{(1 - \rho_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \rho_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \rho_{01} - \rho_{10}}. \end{aligned}$$

However, the existing surrogate loss-based methods require fully observed noisy labels $r_{u,i}$, which prevents the direct use of such methods for RS in the presence of missing data. To fill this gap, motivated by the broad usage of EIB, IPS, and DR estimators in the missing data literature [2, 32], we extend the above surrogate loss-based methods to address the data MNAR and OME in parallel.

Specifically, given knowledge of error parameters ρ_{01} and ρ_{10} , let $\tilde{e}_{u,i} = r_{u,i} \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 1) + (1 - r_{u,i}) \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 0)$ be the surrogate loss with the rating $r_{u,i}$, the OME-EIB estimator estimates the true prediction inaccuracy $\mathcal{P}^* = \mathcal{P}(\hat{\mathbf{R}}, \mathbf{R}^*)$ with

$$\begin{aligned} \mathcal{E}_{\text{OME-EIB}}(\hat{\mathbf{R}}, \mathbf{R}^o; \rho_{01}, \rho_{10}) &= \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} (1 - o_{u,i}) \bar{e}_{u,i} + \\ &\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} r_{u,i} \{ (1 - \rho_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \rho_{01} \ell(f_{\theta}(x_{u,i}), 0) \}}{1 - \rho_{01} - \rho_{10}} + \\ &\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} (1 - r_{u,i}) \{ (1 - \rho_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \rho_{10} \ell(f_{\theta}(x_{u,i}), 1) \}}{1 - \rho_{01} - \rho_{10}}, \end{aligned}$$

where $\bar{e}_{u,i}$ is the imputed error for estimating $\tilde{e}_{u,i}$, and OME-EIB is an unbiased estimator of the true prediction inaccuracy when the imputed errors are accurate, *i.e.*, $\bar{e}_{u,i} = \tilde{e}_{u,i}$. Since the proof is not trivial, we postpone to show unbiasedness of OME-EIB estimator and the following OME-IPS and OME-DR estimators in Sec. 3.4.

Similarly, the OME-IPS estimator estimates $\mathcal{P}^* = \mathcal{P}(\hat{\mathbf{R}}, \mathbf{R}^*)$ with

$$\begin{aligned} \mathcal{E}_{\text{OME-IPS}}(\hat{\mathbf{R}}, \mathbf{R}^o; \rho_{01}, \rho_{10}) &= \\ &\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} r_{u,i}}{\hat{p}_{u,i}} \cdot \frac{(1 - \rho_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \rho_{01} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \rho_{01} - \rho_{10}} + \\ &\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} (1 - r_{u,i})}{\hat{p}_{u,i}} \cdot \frac{(1 - \rho_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \rho_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \rho_{01} - \rho_{10}}, \end{aligned}$$

where $\hat{p}_{u,i}$ is the learned propensity for estimating $p_{u,i}$, and OME-IPS is an unbiased estimator of the true prediction inaccuracy when the learned propensities are accurate, *i.e.*, $\hat{p}_{u,i} = p_{u,i}$.

The OME-DR estimator estimates $\mathcal{P}^* = \mathcal{P}(\hat{\mathbf{R}}, \mathbf{R}^*)$ with

$$\begin{aligned} \mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \rho_{01}, \rho_{10}) &= \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left(1 - \frac{o_{u,i}}{\hat{p}_{u,i}}\right) \bar{e}_{u,i} + \\ &\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} r_{u,i}}{\hat{p}_{u,i}} \cdot \frac{(1 - \rho_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \rho_{10} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \rho_{01} - \rho_{10}} + \\ &\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} (1 - r_{u,i})}{\hat{p}_{u,i}} \cdot \frac{(1 - \rho_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \rho_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \rho_{01} - \rho_{10}}, \end{aligned}$$

which is an unbiased estimate of the true prediction inaccuracy when either the imputed errors or the learned propensities are accurate, *i.e.*, $\bar{e}_{u,i} = \tilde{e}_{u,i}$ or $\hat{p}_{u,i} = p_{u,i}$.

3.3 Identification and Estimation of ρ_{01} and ρ_{10}

The proposed OME-EIB, OME-IPS, and OME-DR estimators require knowledge of the known error parameters ρ_{01} and ρ_{10} , which are usually not directly available from the collected data. By building upon the existing weak separability assumption [42], we present a data-driven identification and estimation method of ρ_{01} and ρ_{10} .

We impose the following weak separability assumption that

$$\inf_{(u,i) \in \mathcal{D}} \mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}) = 0 \quad \text{and} \quad \sup_{(u,i) \in \mathcal{D}} \mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}) = 1,$$

which also known as mutual irreducibility [60, 61] in the observational label noise literature. This does not require the true ratings to be separable, *i.e.*, $\mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}) \in \{0, 1\}$ for all user-item pairs, but instead stipulates that there exist "perfectly positive" and "perfectly negative" samples among the entire user-item pairs. In real-world recommendation scenarios, the weak separability assumption is easily satisfied, providing there exists at least one "perfectly positive" feedback and one "perfectly negative" feedback among the thousands of collected ratings. For example, in the movie rating scenario, "perfectly positive" feedback refers to at least one of the users who made a positive review of the movie is fully reliable, and we do not need to know who that user exactly is.

By noting the linkage between the observed and the true ratings

$$\begin{aligned} \mathbb{P}(r_{u,i} = 1 \mid x_{u,i}) &= (1 - \rho_{01})\mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}) + \rho_{10}\mathbb{P}(r_{u,i}^* = 0 \mid x_{u,i}) \\ &= (1 - \rho_{01} - \rho_{10})\mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}) + \rho_{10}, \end{aligned}$$

which demonstrates the monotonicity between $\mathbb{P}(r_{u,i} = 1 \mid x_{u,i})$ and $\mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i})$ under $\rho_{01} + \rho_{10} < 1$, we have

$$(u^<, i^<) = \arg \min_{(u,i) \in \mathcal{D}} \mathbb{P}(r_{u,i} = 1 \mid x_{u,i}) = \arg \min_{(u,i) \in \mathcal{D}} \mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}),$$

$$(u^>, i^>) = \arg \max_{(u,i) \in \mathcal{D}} \mathbb{P}(r_{u,i} = 1 \mid x_{u,i}) = \arg \max_{(u,i) \in \mathcal{D}} \mathbb{P}(r_{u,i}^* = 1 \mid x_{u,i}).$$

Then, we can identify the error parameters ρ_{01} and ρ_{10} via

$$\rho_{01} = 1 - \mathbb{P}(r_{u^>, i^>} = 1 \mid x_{u^>, i^>}) \quad \text{and} \quad \rho_{10} = \mathbb{P}(r_{u^<, i^<} = 1 \mid x_{u^<, i^<}),$$

where $\mathbb{P}(r_{u^>, i^>} = 1 \mid x_{u^>, i^>})$ and $\mathbb{P}(r_{u^<, i^<} = 1 \mid x_{u^<, i^<})$ can be unbiasedly estimated from the existing EIB, IPS, or DR estimators for estimating $\mathbb{P}(r_{u,i} = 1 \mid x_{u,i})$ without considering the OME.

3.4 Theoretical Analyses

Since the OME-DR estimator degenerates to the OME-IPS estimator when $\bar{e}_{u,i} = 0$, and degenerates to the OME-EIB estimator when $\hat{p}_{u,i} = 1$, without loss of generality, we only analyze the explicit bias

form of the OME-DR estimator. Following existing literature [59, 73], we assume that the indicator matrix \mathbf{O} contains independent random variables and each $o_{u,i}$ follows a Bernoulli distribution with probability $p_{u,i}$. In addition, due to the presence of OME, we also consider the randomness of the potentially observed ratings $r_{u,i}$ given the true ratings $r_{u,i}^*$, *e.g.*, $r_{u,i} = 1$ with probability $1 - \rho_{01}$ given $r_{u,i}^* = 1$, and $r_{u,i} = 0$ with probability ρ_{01} given $r_{u,i}^* = 1$.

THEOREM 3.1 (BIAS OF OME-DR ESTIMATOR). *Given $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ with $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$, imputed errors $\bar{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$ with $\hat{p}_{u,i} > 0$ for all user-item pairs, the bias of the OME-DR estimator is*

$$\begin{aligned} \text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \left(1 - \frac{o_{u,i}}{\hat{p}_{u,i}}\right) \bar{e}_{u,i} + \right. \\ &\sum_{(u,i): r_{u,i}^* = 1} \left(\frac{p_{u,i} \omega_{11} - \hat{p}_{u,i}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{p_{u,i} \omega_{01}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 0) \right) + \\ &\left. \sum_{(u,i): r_{u,i}^* = 0} \left(\frac{p_{u,i} \omega_{10}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{p_{u,i} \omega_{00} - \hat{p}_{u,i}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 0) \right) \right|, \end{aligned}$$

where ω_{11} , ω_{01} , ω_{10} , and ω_{00} are given by

$$\begin{aligned} \omega_{11} &= \frac{1 - \rho_{01} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, & \omega_{01} &= \frac{\rho_{01} - \hat{\rho}_{01}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \\ \omega_{10} &= \frac{\rho_{10} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, & \omega_{00} &= \frac{1 - \hat{\rho}_{01} - \rho_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}. \end{aligned}$$

The bias of the OME-DR estimator includes the three terms: (1) The first term shares a similar form to the bias of the previous IPS estimator and leads to smaller bias when $\hat{p}_{u,i} \approx p_{u,i}$. (2) The second term is novel for OME, specifically focusing on the estimated false negative rate $\hat{\rho}_{01}$ that corresponds to the positive samples $r_{u,i}^* = 1$. Moreover, we find that $\omega_{11} = 1$ and $\omega_{01} = 0$ when the the estimated false negative rate $\hat{\rho}_{01}$ is accurate, *i.e.*, $\hat{\rho}_{01} = \rho_{01}$, which results in smaller bias when $\bar{e}_{u,i} \approx \tilde{e}_{u,i}$. (3) The third term is similar to the second term, but instead focuses on the estimated false positive rate $\hat{\rho}_{10}$ that corresponds to the negative samples $r_{u,i}^* = 0$, and also results in smaller bias when $\bar{e}_{u,i} \approx \tilde{e}_{u,i}$. Given the importance of bias derivation for constructing estimators under OME, we provide a proof sketch as below (see Appendix A for more details).

PROOF SKETCH. By definition, bias of the OME-DR estimator is

$$\text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] = \left| \mathbb{E}_{\mathbf{R}, \mathbf{O}} [\mathcal{E}_{\text{OME-DR}}] - \mathcal{P}^* \right|.$$

Then, we can derive the bias of the OME-DR estimator as follows

$$\begin{aligned} \text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] &= \left| \mathbb{E}_{\mathbf{R}, \mathbf{O}} [\mathbb{E}_{\mathbf{O}} [\mathcal{E}_{\text{OME-DR}}]] - \mathcal{P}^* \right| \\ &= \left| \frac{1}{|\mathcal{D}|} \sum_{(u,i): r_{u,i}^* = 1} \mathbb{E}_{\mathbf{R}, \mathbf{O}} \left[\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \bar{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right] - \ell(f_{\theta}(x_{u,i}), 1) \right. \\ &\left. + \frac{1}{|\mathcal{D}|} \sum_{(u,i): r_{u,i}^* = 0} \mathbb{E}_{\mathbf{R}, \mathbf{O}} \left[\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \bar{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right] - \ell(f_{\theta}(x_{u,i}), 0) \right|. \end{aligned}$$

For the user-item pairs with $r_{u,i}^* = 1$, we have

$$\begin{aligned} & \mathbb{E}_{\mathbf{R}|\mathbf{O}} [\tilde{e}_{u,i}] \\ &= (1 - \rho_{01}) \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 1; \hat{\rho}_{01}, \hat{\rho}_{10}) + \rho_{01} \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 0; \hat{\rho}_{01}, \hat{\rho}_{10}) \\ &= (1 - \rho_{01}) \cdot \frac{(1 - \hat{\rho}_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \hat{\rho}_{01} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \\ &\quad + \rho_{01} \cdot \frac{(1 - \hat{\rho}_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \hat{\rho}_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \\ &= \frac{1 - \rho_{01} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{\rho_{01} - \hat{\rho}_{01}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 0) \\ &:= \omega_{11} \ell(f_{\theta}(x_{u,i}), 1) + \omega_{01} \ell(f_{\theta}(x_{u,i}), 0), \end{aligned}$$

where ω_{11} and ω_{01} are given by

$$\omega_{11} = \frac{1 - \rho_{01} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \quad \omega_{01} = \frac{\rho_{01} - \hat{\rho}_{01}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}.$$

Similar results hold for the user-item pairs with $r_{u,i}^* = 0$ that

$$\begin{aligned} & \mathbb{E}_{\mathbf{R}|\mathbf{O}} [\tilde{e}_{u,i}] \\ &= \frac{\rho_{10} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{1 - \hat{\rho}_{01} - \rho_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 0) \\ &:= \omega_{10} \ell(f_{\theta}(x_{u,i}), 1) + \omega_{00} \ell(f_{\theta}(x_{u,i}), 0), \end{aligned}$$

where ω_{10} and ω_{00} are given by

$$\omega_{10} = \frac{\rho_{10} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \quad \omega_{00} = \frac{1 - \hat{\rho}_{01} - \rho_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}.$$

This completes the proof. \square

We formally describe double robustness under OME as follows.

COROLLARY 3.2 (DOUBLE ROBUSTNESS). *Given $\hat{\rho}_{01} = \rho_{01}$ and $\hat{\rho}_{10} = \rho_{10}$, the OME-DR estimator is unbiased when either imputed errors $\tilde{\mathbf{E}}$ or learned propensities $\hat{\mathbf{P}}$ are accurate for all user-item pairs.*

The above result is obtained via substituting either $\tilde{e}_{u,i} = \tilde{e}_{u,i}$ or $\hat{p}_{u,i} = p_{u,i}$ into the bias of the OME-DR estimator in Theorem 3.1. Given the estimated error parameters $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$, we obtain the optimal prediction model under OME by minimizing the OME-DR estimator over a hypothesis space \mathcal{F} of the prediction models f_{θ}

$$\hat{\mathbf{R}}^{\ddagger} = \arg \min_{f_{\theta} \in \mathcal{F}} \{\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^{\circ}; \hat{\rho}_{01}, \hat{\rho}_{10})\}.$$

We next derive the generalization bound of the optimal prediction model in terms of the empirical Rademacher complexity [62]. The basic idea of proving the performance guarantee under OME is to exploit the inheritance of the Lipschitz continuity from the true prediction loss $\ell(f_{\theta}(x_{u,i}), r_{u,i}^*)$ to the surrogate loss $\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})$.

LEMMA 3.3 (LIPSCHITZ CONTINUITY). *Given $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ with $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$, if $\ell(f_{\theta}(x_{u,i}), r_{u,i}^*)$ is L -Lipschitz in $f_{\theta}(x_{u,i})$ for all $r_{u,i}^*$, then $\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})$ is $\frac{2L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}$ -Lipschitz in $f_{\theta}(x_{u,i})$ for all $r_{u,i}$.*

The above lemma immediately leads to a generalization bound with respect to the true ratings by using the contraction principle for Rademacher complexity (see Appendix A for proofs).

THEOREM 3.4 (GENERALIZATION BOUND). *Given $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ with $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$, suppose $\ell(f_{\theta}(x_{u,i}), r_{u,i}^*)$ is L -Lipschitz in $f_{\theta}(x_{u,i})$ for all $r_{u,i}^*$, and $\hat{p}_{u,i} \geq C_p$, $|\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})| \leq C_l$ for all $r_{u,i}$, then with probability $1 - \eta$, the true prediction inaccuracy $\mathcal{P}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^*)$ of the*

Algorithm 1: Alternating Denoise Training with OME-DR.

Input: observed ratings \mathbf{R}° , learned propensities $\hat{p}_{u,i}$, initial estimates $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$, a pre-trained model $h_{\theta^*}(x_{u,i})$ for estimating $\mathbb{P}(r_{u,i} = 1 | x_{u,i})$ via existing method.

- 1 **while** stopping criteria is not satisfied **do**
- 2 **for** number of steps for training the denoising prediction model **do**
- 3 Sample a batch $\{(u_j, i_j)\}_{j=1}^J$ from \mathcal{D} ;
- 4 Update θ by descending along the gradient
 $\nabla_{\theta} \mathcal{E}_{\text{OME-DR}}(\theta, \phi; \hat{\rho}_{01}, \hat{\rho}_{10})$;
- 5 $(u^<, i^<) \leftarrow \arg \min_{(u,i) \in \mathcal{D}} f_{\theta}(x_{u,i})$;
- 6 $(u^>, i^>) \leftarrow \arg \max_{(u,i) \in \mathcal{D}} f_{\theta}(x_{u,i})$;
- 7 **end**
- 8 **for** number of steps for training the denoising imputation model **do**
- 9 $\hat{\rho}_{01} \leftarrow 1 - h_{\theta^*}(x_{u^>, i^>})$ and $\hat{\rho}_{10} \leftarrow h_{\theta^*}(x_{u^<, i^<})$;
- 10 Sample a batch $\{(u_k, i_k)\}_{k=1}^K$ from \mathcal{O} ;
- 11 Update ϕ by descending along the gradient
 $\nabla_{\phi} \mathcal{L}_{\tilde{\mathbf{E}}}(\theta, \phi; \hat{\rho}_{01}, \hat{\rho}_{10})$;
- 12 **end**
- 13 **end**

optimal prediction matrix using the OME-DR estimator with imputed errors $\tilde{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$ has the upper bound

$$\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^{\circ}; \hat{\rho}_{01}, \hat{\rho}_{10}) + \text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^{\circ}; \hat{\rho}_{01}, \hat{\rho}_{10})] + \left(1 + \frac{2}{C_p}\right) \left[\frac{4L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \mathcal{R}(\mathcal{F}) + \left(C_l + \frac{4L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}\right) \sqrt{\frac{2 \log(4/\eta)}{|\mathcal{D}|}} \right],$$

where $\mathcal{R}(\mathcal{F})$ is the empirical Rademacher complexity defined as

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

in which $\sigma = \{\sigma_{u,i} : (u, i) \in \mathcal{D}\}$, and $\sigma_{u,i}$ are independent uniform random variables taking values in $\{-1, +1\}$. The random variables $\sigma_{u,i}$ are called Rademacher variables.

3.5 Alternating Denoise Training Approach

We further propose an alternating denoise training approach to achieve unbiased learning of the prediction model, which corrects for data MNAR and OME in parallel. Specifically, we first train a propensity model using the observed MNAR data. Based on the estimated propensities, the denoising prediction and imputation models are alternately updated, which also facilitates the accurate estimations $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ of the error parameters ρ_{01} and ρ_{10} .

Propensity Estimation via Logistic Regression. Since unbiased ratings are difficult to be collected in the real-world scenarios [57], following previous studies [23, 59], we adopt logistic regression to train a propensity model $\hat{p}_{u,i} = \sigma(w^{\top} x_{u,i} + \beta_u + \gamma_i)$ parameterized by $\psi = (w, \beta_1, \dots, \beta_N, \gamma_1, \dots, \gamma_M)$ using the observed MNAR data, where $\sigma(\cdot)$ is the sigmoid function, $w \in \mathbb{R}^{N+M}$ is the weight

Table 1: RE on ML-100K dataset with $\rho_{01} = 0.2$ and $\rho_{10} = 0.1$. The best two results are bolded and the best baseline is underlined.

	ROTATE	SKEW	CRS	ONE	THREE	FIVE
Naive [20]	0.125 ± 0.002	0.179 ± 0.001	0.175 ± 0.002	0.241 ± 0.002	0.264 ± 0.003	0.299 ± 0.003
OME ($\hat{\rho}_{01}, \hat{\rho}_{10}$)	0.087 ± 0.006	0.163 ± 0.002	0.104 ± 0.014	0.161 ± 0.013	0.179 ± 0.014	0.213 ± 0.014
OME (ρ_{01}, ρ_{10})	<u>0.024 ± 0.003</u>	0.136 ± 0.002	0.105 ± 0.004	<u>0.067 ± 0.004</u>	<u>0.076 ± 0.004</u>	<u>0.100 ± 0.004</u>
MRDR [17]	0.098 ± 0.003	0.089 ± 0.001	<u>0.099 ± 0.002</u>	0.172 ± 0.004	0.190 ± 0.004	0.194 ± 0.004
SDR [29]	0.097 ± 0.002	0.089 ± 0.001	0.102 ± 0.002	0.172 ± 0.003	0.191 ± 0.003	0.194 ± 0.003
TDR [23]	0.092 ± 0.002	<u>0.080 ± 0.001</u>	0.103 ± 0.002	0.171 ± 0.003	0.189 ± 0.003	0.195 ± 0.003
EIB [65]	0.366 ± 0.001	0.256 ± 0.001	0.150 ± 0.001	0.562 ± 0.001	0.605 ± 0.001	0.635 ± 0.001
OME-EIB ($\hat{\rho}_{01}, \hat{\rho}_{10}$)	0.362 ± 0.001	0.255 ± 0.001	0.144 ± 0.001	0.554 ± 0.001	0.598 ± 0.001	0.627 ± 0.001
OME-EIB (ρ_{01}, ρ_{10})	0.357 ± 0.001	0.253 ± 0.001	0.144 ± 0.001	0.546 ± 0.001	0.589 ± 0.001	0.618 ± 0.001
IPS [59]	0.110 ± 0.002	0.116 ± 0.002	0.134 ± 0.003	0.212 ± 0.004	0.231 ± 0.004	0.254 ± 0.004
OME-IPS ($\hat{\rho}_{01}, \hat{\rho}_{10}$)	0.060 ± 0.003	0.096 ± 0.002	0.075 ± 0.017	0.111 ± 0.006	0.125 ± 0.007	0.144 ± 0.007
OME-IPS (ρ_{01}, ρ_{10})	0.013 ± 0.003*	0.068 ± 0.003*	0.052 ± 0.004*	0.034 ± 0.005*	0.038 ± 0.006*	0.050 ± 0.005*
DR [56]	0.106 ± 0.002	0.087 ± 0.001	0.104 ± 0.002	0.190 ± 0.003	0.209 ± 0.003	0.216 ± 0.003
OME-DR ($\hat{\rho}_{01}, \hat{\rho}_{10}$)	0.056 ± 0.004	0.067 ± 0.001	0.045 ± 0.018	0.090 ± 0.006	0.102 ± 0.006	0.106 ± 0.007
OME-DR (ρ_{01}, ρ_{10})	0.009 ± 0.003*	0.039 ± 0.002*	0.022 ± 0.003*	0.013 ± 0.004*	0.016 ± 0.005*	0.012 ± 0.004*

Note: * means statistically significant results (p-value ≤ 0.05) using the paired-t-test compared with the best baseline.

vector, β_u and γ_i are the user-specific and item-specific constants, respectively. We train the propensity model by minimizing the loss

$$\mathcal{L}_p(\psi) = -\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} o_{u,i} \log(\hat{p}_{u,i}) + (1 - o_{u,i}) \log(1 - \hat{p}_{u,i}).$$

Denosing Prediction Model Training. Based on the learned propensities $\hat{p}_{u,i}$, imputed errors $\bar{e}_{u,i}(\phi)$, and initial estimates $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$, we train the denosing prediction model by minimizing the estimated true prediction inaccuracy from the proposed OME-DR estimator $\mathcal{E}_{\text{OME-DR}}(\theta, \phi; \hat{\rho}_{01}, \hat{\rho}_{10})$. To obtain more accurate estimates $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ from the training loop, we then compute the minimum and maximum predicted ratings in the training batch to update ($u^<, i^<$) and ($u^>, i^>$), respectively. The choice not to employ a separate noisy rating prediction model for computing ($u^<, i^<$) and ($u^>, i^>$) is guaranteed by the monotonicity between $\mathbb{P}(r_{u,i} = 1 | x_{u,i})$ and $\mathbb{P}(r_{u,i}^* = 1 | x_{u,i})$ (see Sec. 3.3 for proofs), where the latter has a larger gap between the minimum and the maximum, making it easier to distinguish ($u^<, i^<$) and ($u^>, i^>$).

Denosing Imputation Model Training. Given a pre-trained model $h_{\theta^*}(x_{u,i})$ for estimating $\mathbb{P}(r_{u,i} = 1 | x_{u,i})$ via existing methods [59, 73], we update the estimates $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ by computing $1 - h_{\theta^*}(x_{u^>, i^>})$ and $h_{\theta^*}(x_{u^<, i^<})$, respectively. We finally train the denosing imputation model $\bar{e}_{u,i}(\phi)$ by minimizing the loss

$$\mathcal{L}_{\bar{e}}(\theta, \phi; \hat{\rho}_{01}, \hat{\rho}_{10}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i}(\bar{e}_{u,i} - \bar{e}_{u,i})^2}{\hat{p}_{u,i}}.$$

The overall training process with OME-DR is summarized in Alg. 1.

4 Semi-Synthetic Experiments

To investigate if the proposed estimators are able to estimate the true prediction inaccuracy accurately in the presence of OME and MNAR effect, several semi-synthetic experiments are conducted on a widely-used dataset MovieLens-100K (ML-100K). We first adopt MF to complete the five-scaled rating matrix \mathbf{R} . However, the completed matrix will have an unrealistic rating distribution, thus we

sort the matrix entries in ascending order and assign a positive feedback probability of 0.1 for the lowest p_1 proportion, a positive feedback probability of 0.3 for the next p_2 proportion, and so on to obtain a true positive feedback probability $\gamma_{u,i}$ for each user-item pair. The adjusted probability matrix contains $\gamma_{u,i} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ with proportion $[p_1, p_2, p_3, p_4, p_5]$, respectively.

Second, following the previous studies [17, 59, 73], we use six matrices below as the prediction matrix $\hat{\mathbf{R}}$:

- **ROTATE:** Set predicted $\hat{r}_{u,i} = \gamma_{u,i} - 0.2$ when $\gamma_{u,i} \geq 0.3$, and $\hat{r}_{u,i} = 0.9$ when $\gamma_{u,i} = 0.1$.
- **SKEW:** Predicted $\hat{r}_{u,i}$ are sampled from the Gaussian distribution $\mathcal{N}(\mu = \gamma_{u,i}, \sigma = (1 - \gamma_{u,i})/2)$, and clipped to the interval $[0.1, 0.9]$.
- **CRS:** If the true $\gamma_{u,i} \leq 0.6$, then $\hat{r}_{u,i} = 0.2$. Otherwise, $\hat{r}_{u,i} = 0.6$.
- **ONE:** The predicted matrix $\hat{\mathbf{R}}$ is identical to the true positive feedback probability matrix, except that randomly select $\gamma_{u,i} = 0.1$ with total amount $|\{(u,i) | \gamma_{u,i} = 0.9\}|$ are flipped to 0.9.
- **THREE:** Same as **ONE**, but flipping $\gamma_{u,i} = 0.3$ instead.
- **FIVE:** Same as **ONE**, but flipping $\gamma_{u,i} = 0.5$ instead.

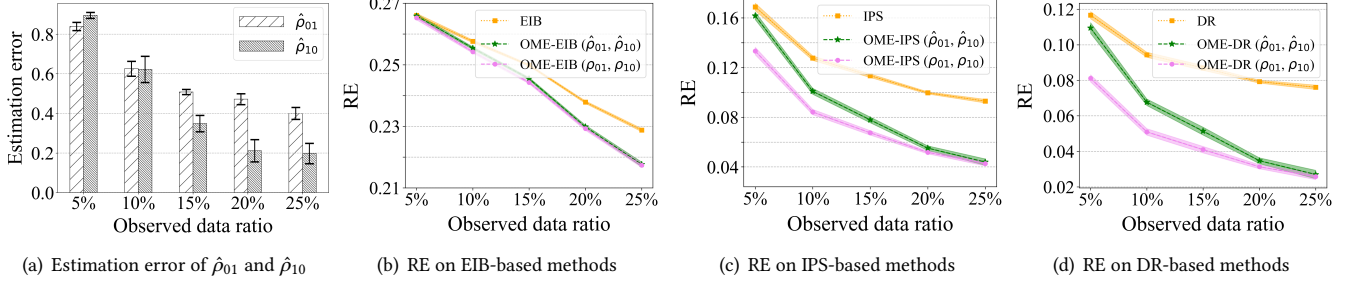
Next, we assign the propensity $p_{u,i} = p\alpha^{\min(4, 6 - r_{u,i})}$ for each user-item pair and obtain the estimate propensities

$$\frac{1}{\hat{p}_{u,i}} = \frac{1 - \beta}{p_{u,i}} + \frac{\beta}{p_e},$$

where $p_e = |\mathcal{D}|^{-1} \sum_{(u,i) \in \mathcal{D}} o_{u,i}$, p is set to 1 in our experiment and β is randomly sampled from a uniform distribution $U(0, 1)$ to introduce noises. Then, we sample the binary true feedback matrix \mathbf{R}^* and binary observation matrix \mathbf{O} as follows:

$$o_{u,i} \sim \text{Bern}(p_{u,i}), \forall (u,i) \in \mathcal{D}, \quad r_{u,i}^* \sim \text{Bern}(\gamma_{u,i}), \forall (u,i) \in \mathcal{D},$$

where $\text{Bern}(\cdot)$ denotes the Bernoulli distribution. Then we flip the feedback matrix according to the ρ_{01} and ρ_{10} to generate a binary noise feedback matrix \mathbf{R} . The absolute relative error (RE) is used for evaluation, which is defined as $\text{RE}(\mathcal{E}_{est}) = |\mathcal{P}^* - \mathcal{E}_{est}(\hat{\mathbf{R}}, \mathbf{R})|/\mathcal{P}^*$, where \mathcal{E}_{est} denotes the estimate prediction inaccuracy. The smaller the RE, the more accurate the estimation. In addition, we provide both results for using estimated $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ to estimate \mathcal{P}^* and directly using true ρ_{01} and ρ_{10} to estimate \mathcal{P}^* .

Figure 1: Estimation errors of $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ and RE with varying observed data ratios.Table 2: Performance on AUC, NDCG@K, and Recall@K on the MAR test set of COAT, MUSIC and KUAIREC with $\rho_{01} = 0.2$ and $\rho_{10} = 0.1$. The best three results are bolded, and the best baseline is underlined.

Method	COAT			MUSIC			KUAIREC		
	AUC	NDCG@5	Recall@5	AUC	NDCG@5	Recall@5	AUC	NDCG@5	Recall@5
MF [20]	0.619±0.002	0.531±0.005	0.364±0.005	0.632±0.005	0.616±0.003	0.361±0.001	0.621±0.002	0.652±0.003	0.702±0.002
OME [36]	0.621±0.003	0.543±0.006	0.371±0.007	0.650±0.005	0.615±0.004	0.373±0.004	0.645±0.002	0.660±0.002	0.722±0.002
T-MF [71]	0.633±0.004	0.540±0.005	0.372±0.006	0.650±0.005	0.614±0.003	0.376±0.004	0.657±0.002	0.665±0.002	0.739±0.003
R-MF [71]	0.641±0.004	0.544±0.005	0.376±0.009	0.657±0.005	0.614±0.006	0.371±0.004	0.645±0.002	0.661±0.001	0.720±0.003
LCD-MF [11]	0.631±0.003	0.552±0.007	0.381±0.006	0.659±0.005	0.617±0.005	0.374±0.004	0.664±0.003	0.677±0.003	0.738±0.003
EIB [65]	0.621±0.003	0.547±0.006	0.368±0.008	0.648±0.003	0.620±0.002	0.371±0.002	0.641±0.001	0.653±0.001	0.713±0.002
IPS [59]	0.625±0.002	0.534±0.009	0.360±0.008	0.648±0.001	0.615±0.001	0.369±0.001	0.636±0.001	0.642±0.002	0.714±0.001
SNIPS [59]	0.631±0.004	0.544±0.008	0.369±0.007	0.659±0.005	0.617±0.005	0.375±0.004	0.657±0.002	0.670±0.001	0.735±0.001
CVIB [75]	0.626±0.001	0.552±0.009	0.379±0.004	0.661±0.001	0.612±0.002	0.381±0.001	0.653±0.001	0.644±0.001	0.736±0.001
DAMF [57]	0.631±0.004	0.548±0.006	0.367±0.006	0.640±0.003	0.627±0.002	0.377±0.004	0.657±0.003	0.669±0.003	0.735±0.004
DR [56]	0.633±0.009	0.548±0.015	0.360±0.011	0.654±0.009	0.613±0.007	0.372±0.005	0.643±0.002	0.650±0.001	0.721±0.001
DR-JL [73]	0.633±0.005	0.546±0.006	0.366±0.007	0.650±0.002	0.617±0.001	0.382±0.001	0.644±0.001	0.669±0.002	0.725±0.002
MRDR-JL [17]	0.634±0.005	0.538±0.010	0.373±0.008	0.659±0.002	0.624±0.007	0.376±0.006	0.658±0.003	0.670±0.002	0.741±0.003
DR-BIAS [10]	0.632±0.004	0.549±0.009	0.369±0.008	0.660±0.001	0.619±0.003	0.378±0.002	0.665±0.004	0.679±0.004	0.737±0.002
DR-MSE [10]	0.631±0.003	0.557±0.008	0.374±0.008	0.662±0.002	0.623±0.004	0.374±0.002	0.653±0.002	0.660±0.002	0.732±0.002
DIB [35]	0.626±0.003	0.550±0.009	0.376±0.006	0.647±0.005	0.626±0.005	0.382±0.006	0.641±0.002	0.642±0.003	0.738±0.003
MR [22]	0.630±0.005	0.546±0.009	0.376±0.007	0.656±0.005	0.623±0.004	0.380±0.004	0.637±0.002	0.644±0.004	0.723±0.002
SDR [29]	0.635±0.005	0.543±0.005	0.368±0.006	0.660±0.006	0.618±0.006	0.376±0.005	0.661±0.003	0.675±0.001	0.739±0.002
TDR [23]	0.638±0.007	0.556±0.014	0.370±0.013	0.657±0.004	0.616±0.002	0.371±0.003	0.650±0.002	0.670±0.001	0.723±0.003
IPS-V2 [26]	0.631±0.005	0.549±0.008	0.373±0.008	0.648±0.005	0.615±0.005	0.373±0.004	0.658±0.002	0.675±0.003	0.734±0.002
DR-V2 [26]	0.631±0.005	0.553±0.009	0.378±0.006	0.636±0.003	0.620±0.002	0.392±0.004	0.658±0.001	0.671±0.001	0.734±0.002
OME-EIB (ours)	0.627±0.002	0.563* ±0.002	0.387* ±0.002	0.664±0.003	0.622±0.004	0.384±0.005	0.667±0.003	0.682* ±0.002	0.742±0.002
OME-IPS (ours)	0.636±0.002	0.548±0.003	0.373±0.004	0.664±0.003	0.625±0.002	0.385±0.002	0.671* ±0.002	0.693* ±0.004	0.763* ±0.002
OME-DR (ours)	0.651* ±0.006	0.561±0.006	0.385* ±0.005	0.671* ±0.003	0.632* ±0.003	0.389±0.002	0.662±0.002	0.677±0.004	0.743±0.002

Note: * means statistically significant results (p-value ≤ 0.05) using the paired-t-test compared with the best baseline.

Performance Comparison. We compare the proposed methods with the Naive method [20], the EIB method [65], the IPS method [59], and the DR-based methods [17, 23, 29, 73]. The RE results are shown in Table 1 with $\rho_{01} = 0.2$ and $\rho_{10} = 0.1$. First, most debiasing and denoising methods have lower RE than the Naive method. In addition, our proposed methods stably and significantly outperform the corresponding debiasing baseline methods. Meanwhile, the OME methods are only able to denoising, thus they still have a large RE, even the true ρ is used for estimation. This shows the effectiveness of our method in the presence of both the OME and MNAR effects.

In Depth Analysis. We explore the effect of the proportion of observed data on the estimation accuracy of $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$, and the results when $\rho_{01} = 0.2$ and $\rho_{10} = 0.1$ are shown in Figure 1. We take ROTATE as an example, and the same phenomenon is observed for

the other five prediction matrices. First, our methods stably outperform the baseline methods in all scenarios. Second, the estimation error decreases significantly as the proportion of observed data increases. Moreover, even with the estimation error, our estimation results are only slightly worse than using the real ρ . This further validates the stability and practicality of our methods.

5 Real-World Experiments

Datasets and Experiment Details. We verify the effectiveness of our methods on three real-world datasets: COAT [59], MUSIC [59] and KUAIREC [14]. COAT includes 6,960 MNAR ratings and 4,640 missing-at-random (MAR) ratings from 290 users to 300 items. MUSIC has 311,704 MNAR ratings and 54,000 MAR ratings of 15,400 users to 1,000 items. For COAT and MUSIC, we binarize the ratings

less than three to 0 and otherwise to 1. KUAIREC is a fully exposed large-scale industrial dataset, which has 4,676,570 video watching ratio records from 1,411 users to 3,327 items. We binarize the records less than one to 0 and otherwise to 1. We adopt three common metrics, *i.e.*, AUC, NDCG@K, and Recall@K for performance evaluation. For COAT and MUSIC, K is set to 5. For KUAIREC, K is set to 50. All the experiments are implemented on PyTorch with Adam as the optimizer. For all experiments, we use GeForce RTX 3090 as the computing resource. Logistic regression is used as the propensity model for all the methods requiring propensity. The learning rate is tuned in $\{0.001, 0.005, 0.01, 0.05\}$ and the batch size is tuned in $\{64, 128, 256\}$ for COAT and $\{2048, 4096, 8192\}$ for MUSIC and KUAIREC. We tune the embedding dimension in $\{4, 8, 16, 32\}$ for COAT and $\{8, 16, 32, 64\}$ for MUSIC and KUAIREC. Moreover, we tune the weight decay rate in $[1e - 6, 5e - 3]$.

Baselines. We use matrix factorization (MF) [20] as the base model and compare our methods to the following debiasing methods: EIB [65], IPS [59], SNIPS [59], CVIB [75], DAMF [57], DR [56], DR-JL [73], MRDR-JL [17], DR-BIAS [10], DR-MSE [10], DIB [35], MR [22], SDR [29], TDR [23], IPS-V2 [26] and DR-V2 [26]. We also compare denoising methods including surrogate loss minimization (OME) [36], T-MF [71], R-MF [71], and LCD-MF [11].

Performance Comparison. Table 2 shows the prediction performance with varying baselines and our methods. First, most of the debiasing and denoising methods have better performance compared to the Naive method, which shows the necessity of debiasing and denoising. Meanwhile, our methods exhibit the most competitive performance in all three datasets, significantly outperforming the baselines including debiasing methods and denoising methods.

6 Related Work

6.1 Debaised Recommendation

The data collected in recommender systems are often systematically subject to varying types of bias, such as conformity bias [37], item popularity bias [77, 86], latent confounders [13, 25], and position bias [46, 47]. In order to achieve unbiased learning of the prediction model, three typical approaches have been proposed, including: (1) The error-imputation-based (EIB) approaches [18, 55, 65], which compute an imputed error for each missing rating. (2) The inverse-propensity-scoring (IPS) approaches [54, 58, 59, 66], which inversely weight the prediction error for each observed rating with the probability of observing that rating. (3) The doubly robust (DR) approaches [21, 56, 73], which use both the error imputation model and the propensity model, and the unbiased learning of the prediction model can be achieved when either the error imputation model or the propensity model is accurate. Based on the above EIB, IPS, and DR estimators, in terms of learning paradigms, recent studies have investigated flexible trade-offs between bias and variance [10, 17], parameter sharing in multi-task learning [40, 69, 84], and the use of a few unbiased ratings to improve the estimation of the prediction inaccuracy [3, 5, 31, 33, 34, 74]. In terms of statistical theory, recent studies have developed targeted DR (TDR) [23] and conservative DR (CDR) [64] to combat the inaccurate imputed errors, StableDR [29] to combat sparse data, and new propensity estimation methods based on balancing metrics [26, 39, 80]. In addition, [57] proposes to minimize the propensity-independent

generalization error bound via adversarial learning. [22] extends DR to multiple robust learning. These methods have also been applied to sequential recommendation [76] and social recommendation [7]. Furthermore, methods based on information bottlenecks [35, 75] and representation learning [48, 80] have also been proposed for debaised recommendation. In this work, we extend the previous debiasing methods to a more realistic RS scenario, in which the observed ratings and the users' true preferences may be different.

6.2 Outcome Measurement Error

Outcome measurement error (OME) refers to the inconsistency between the observed outcomes and the true outcomes, also known as noisy outcomes in social science [52], biostatistics [19], and psychometrics [63]. The error models or transition matrices in OME establish the relation between the true outcomes and the observed outcomes, including uniform [1, 68], class-conditional [36, 42, 61], and instance-dependent [8, 78] structures of outcome misclassification. Many data-driven error parameter estimation methods are developed in recent machine learning literature [45, 60, 61, 79]. Meanwhile, given knowledge of measurement error parameters, recent studies propose unbiased risk minimization approaches for learning under noisy labels [9, 44, 49, 67].

Despite the prevalence of OME in real-world recommendation scenarios, there is still only limited work focusing on denoising in RS [4, 82, 87]. Recently, by noticing that noisy feedback typically has large loss values in the early stages, [71] proposes adaptive denoising training and [15] proposes self-guided denoising learning in implicit feedback. However, the existing methods are mostly heuristic and require the fully observed noisy labels, which prevents the direct use of such methods for RS in the presence of missing data. To fill this gap, we extend the surrogate loss-based methods to address the data MNAR and OME in parallel.

7 Conclusion

In this study, we explored the challenges characterized MNAR data with noisy feedback encountered in real-world recommendation scenarios. First, we introduced the concept of true prediction inaccuracy as a more comprehensive evaluation criterion that accounts for OME, and proposed OME-EIB, OME-IPS, and OME-DR estimators to provide unbiased estimations of the true prediction inaccuracy. Next, we derived the explicit forms of the biases and the generalization bounds of the proposed estimators, and proved the double robustness of our OME-DR estimator. We further proposed an alternating denoise training approach that corrects for MNAR data with OME. The effectiveness of our methods was validated on both semi-synthetic and real-world datasets with varying MNAR and OME rates. The potential limitations includes the usage of weak separability assumption, and accurate measurement error parameters' estimation for ensuring unbiasedness. In future research, we aim to develop unbiased learning methods applicable to more complex OME forms, study alternative practical assumptions for identifying and estimating the error parameters, and extend the applicability of the proposed methods to a wider range of RS settings.

Acknowledgements

This work was supported in part by National Natural Science Foundation of China (623B2002, 62272437).

References

- [1] Dana Angluin and Philip Laird. 1988. Learning from noisy examples. *Machine learning* (1988), 343–370.
- [2] Heejung Bang and James M Robins. 2005. Doubly robust estimation in missing data and causal inference models. *Biometrics* (2005), 962–973.
- [3] Stephen Bonner and Flavian Vasile. 2018. Causal embeddings for recommendation. In *RecSys*.
- [4] Huiyuan Chen, Yusan Lin, Menghai Pan, Lan Wang, Chin-Chia Michael Yeh, Xiaoting Li, Yan Zheng, Fei Wang, and Hao Yang. 2022. Denoising self-attentive sequential recommendation. In *RecSys*.
- [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*.
- [6] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. 2023. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems* (2023), 1–39.
- [7] Jiawei Chen, Can Wang, Martin Ester, Qihao Shi, Yan Feng, and Chun Chen. 2018. Social recommendation with missing not at random data. In *ICDM*.
- [8] Pengfei Chen, Junjie Ye, Guangyong Chen, Jingwei Zhao, and Pheng-Ann Heng. 2021. Beyond class-conditional assumption: A primary attempt to combat instance-dependent label noise. In *AAAI*.
- [9] Yu-Ting Chou, Gang Niu, Hsuan-Tien Lin, and Masashi Sugiyama. 2020. Unbiased risk estimators can mislead: A case study of learning with complementary labels. In *ICML*.
- [10] Quanyu Dai, Haoxuan Li, Peng Wu, Zhenhua Dong, Xiao-Hua Zhou, Rui Zhang, Xiuqiang He, Rui Zhang, and Jie Sun. 2022. A Generalized Doubly Robust Learning Framework for Debiasing Post-Click Conversion Rate Prediction. In *KDD*.
- [11] Quanyu Dai, Yalei Lv, Jieming Zhu, Junjie Ye, Zhenhua Dong, Rui Zhang, Shu-Tao Xia, and Ruiming Tang. 2022. LCD: Adaptive label correction for denoising music recommendation. In *CIKM*.
- [12] Arnaud De Myttenaere, Bénédicte Le Grand, Boris Golden, and Fabrice Rossi. 2014. Reducing offline evaluation bias in recommendation systems. *arXiv preprint arXiv:1407.0822* (2014).
- [13] Sihao Ding, Peng Wu, Fuli Feng, Xiangnan He, Yitong Wang, Yong Liao, and Yongdong Zhang. 2022. Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis. In *KDD*.
- [14] Chongming Gao, Shijun Li, Wenqiang Lei, Jiawei Chen, Biao Li, Peng Jiang, Xiangnan He, Jiaxin Mao, and Tat-Seng Chua. 2022. KuaiRec: A Fully-observed Dataset and Insights for Evaluating Recommender Systems. In *CIKM*.
- [15] Yunjun Gao, Yuntao Du, Yujia Hu, Lu Chen, Xinjun Zhu, Ziquan Fang, and Baihua Zheng. 2022. Self-guided learning to denoise for robust recommendation. In *SIGIR*.
- [16] Ihsan Gunes, Cihan Kaleli, Alper Bilge, and Huseyin Polat. 2014. Shilling attacks against recommender systems: a comprehensive survey. *Artificial Intelligence Review* (2014), 767–799.
- [17] 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*.
- [18] José Miguel Hernández-Lobato, Neil Houlsby, and Zoubin Ghahramani. 2014. Probabilistic matrix factorization with non-random missing data. In *ICML*.
- [19] Sui L Hui and Steven D Walter. 1980. Estimating the error rates of diagnostic tests. *Biometrics* (1980), 167–171.
- [20] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* (2009), 30–37.
- [21] Wonbin Kweon and Hwanjo Yu. 2024. Doubly Calibrated Estimator for Recommendation on Data Missing Not At Random. In *WWW*.
- [22] Haoxuan Li, Quanyu Dai, Yuru Li, Yan Lyu, Zhenhua Dong, Xiao-Hua Zhou, and Peng Wu. 2023. Multiple Robust Learning for Recommendation. In *AAAI*.
- [23] Haoxuan Li, Yan Lyu, Chunyuan Zheng, and Peng Wu. 2023. TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiasing Recommendations. In *ICLR*.
- [24] Haoxuan Li, Kunhan Wu, Chunyuan Zheng, Yanghao Xiao, Hao Wang, Zhi Geng, Fuli Feng, Xiangnan He, and Peng Wu. 2023. Removing Hidden Confounding in Recommendation: A Unified Multi-Task Learning Approach. In *NeurIPS*.
- [25] Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, and Peng Wu. 2023. Balancing unobserved confounding with a few unbiased ratings in debiased recommendations. In *WWW*.
- [26] Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, and Peng Cui. 2023. Propensity Matters: Measuring and Enhancing Balancing for Recommendation. In *ICML*.
- [27] Haoxuan Li, Chunyuan Zheng, Sihao Ding, Peng Wu, Zhi Geng, Fuli Feng, and Xiangnan He. 2024. Be Aware of the Neighborhood Effect: Modeling Selection Bias under Interference for Recommendation. In *ICLR*.
- [28] Haoxuan Li, Chunyuan Zheng, Shuyi Wang, Kunhan Wu, Hao Wang, peng Wu, Zhi Geng, Xu Chen, and Xiao-Hua Zhou. 2024. Relaxing the Accurate Imputation Assumption in Doubly Robust Learning for Debiasing Collaborative Filtering. In *ICML*.
- [29] Haoxuan Li, Chunyuan Zheng, and Peng Wu. 2023. StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random. In *ICLR*.
- [30] Haoxuan Li, Chunyuan Zheng, Yanghao Xiao, Peng Wu, Zhi Geng, Xu Chen, and Peng Cui. 2024. Debiased Collaborative Filtering with Kernel-Based Causal Balancing. In *ICLR*.
- [31] Zinan Lin, Dugang Liu, Weike Pan, Qiang Yang, and Zhong Ming. 2023. Transfer learning for collaborative recommendation with biased and unbiased data. *Artificial Intelligence* (2023), 103992.
- [32] Roderick JA Little and Donald B Rubin. 2019. *Statistical analysis with missing data*. Vol. 793. John Wiley & Sons.
- [33] 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*.
- [34] Dugang Liu, Pengxiang Cheng, Zinan Lin, Jinwei Luo, Zhenhua Dong, Xiuqiang He, Weike Pan, and Zhong Ming. 2022. KDCRec: Knowledge distillation for counterfactual recommendation via uniform data. *IEEE Transactions on Knowledge and Data Engineering* (2022).
- [35] Dugang Liu, Pengxiang Cheng, Hong Zhu, Zhenhua Dong, Xiuqiang He, Weike Pan, and Zhong Ming. 2021. Mitigating Confounding Bias in Recommendation via Information Bottleneck. In *RecSys*.
- [36] Tongliang Liu and Dacheng Tao. 2015. Classification with noisy labels by importance reweighting. *IEEE Transactions on pattern analysis and machine intelligence* (2015), 447–461.
- [37] Yiming Liu, Xuezhi Cao, and Yong Yu. 2016. Are You Influenced by Others When Rating? Improve Rating Prediction by Conformity Modeling. In *RecSys*.
- [38] Yiyu Liu, Qian Liu, Yu Tian, Changping Wang, Yanan Niu, Yang Song, and Chenliang Li. 2021. Concept-aware denoising graph neural network for micro-video recommendation. In *CIKM*.
- [39] Jinwei Luo, Dugang Liu, Weike Pan, and Zhong Ming. 2021. Unbiased recommendation model based on improved propensity score estimation. *Journal of Computer Applications* (2021), 3508.
- [40] Xiao Ma, Liqin Zhao, Guan Huang, Zhi Wang, Zelin Hu, Xiaoqiang Zhu, and Kun Gai. 2018. Entire Space Multi-Task Model: An Effective Approach for Estimating Post-Click Conversion Rate. In *SIGIR*.
- [41] Benjamin M Marlin and Richard S Zemel. 2009. Collaborative prediction and ranking with non-random missing data. In *RecSys*.
- [42] Aditya Menon, Brendan Van Rooyen, Cheng Soon Ong, and Bob Williamson. 2015. Learning from corrupted binary labels via class-probability estimation. In *ICML*.
- [43] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. 2018. *Foundations of machine learning*. MIT press.
- [44] Nagarajan Natarajan, Inderjit S Dhillon, Pradeep K Ravikumar, and Ambuj Tewari. 2013. Learning with noisy labels. In *NeurIPS*.
- [45] Curtis Northcutt, Lu Jiang, and Isaac Chuang. 2021. Confident learning: Estimating uncertainty in dataset labels. *Journal of Artificial Intelligence Research* (2021), 1373–1411.
- [46] Harrie Oosterhuis. 2022. Reaching the end of unbiasedness: Uncovering implicit limitations of click-based learning to rank. In *SIGIR*.
- [47] Harrie Oosterhuis. 2023. Doubly robust estimation for correcting position bias in click feedback for unbiased learning to rank. *ACM Transactions on Information Systems* (2023), 1–33.
- [48] Hang Pan, Jiawei Chen, Fuli Feng, Wentao Shi, Junkang Wu, and Xiangnan He. 2023. Discriminative-Invariant Representation Learning for Unbiased Recommendation. In *IJCAI*.
- [49] Giorgio Patrini, Alessandro Rozza, Aditya Krishna Menon, Richard Nock, and Lizhen Qu. 2017. Making deep neural networks robust to label noise: A loss correction approach. In *CVPR*.
- [50] Yuqi Qin, Pengfei Wang, and Chenliang Li. 2021. The world is binary: Contrastive learning for denoising next basket recommendation. In *SIGIR*.
- [51] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2010. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer, 1–35.
- [52] Fred S Roberts. 1985. Measurement theory. (1985).
- [53] Yong Rui, Vicente Ivan Sanchez Carmona, Mohsen Pourvali, Yun Xing, Wei-Wen Yi, Hui-Bin Ruan, and Yu Zhang. 2022. Knowledge mining: A cross-disciplinary survey. *Machine Intelligence Research* (2022), 89–114.
- [54] Yuta Saito. 2019. Unbiased Pairwise Learning from Implicit Feedback. In *NeurIPS Workshop*.
- [55] Yuta Saito. 2020. Asymmetric Tri-training for Debiasing Missing-Not-At-Random Explicit Feedback. In *SIGIR*.
- [56] Yuta Saito. 2020. Doubly robust estimator for ranking metrics with post-click conversions. In *RecSys*.
- [57] Yuta Saito and Masahiro Nomura. 2022. Towards Resolving Propensity Contradiction in Offline Recommender Learning. In *IJCAI*.
- [58] Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. 2020. Unbiased recommender learning from missing-not-at-random implicit feedback. In *WSDM*.

- [59] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *ICML*.
- [60] Clayton Scott. 2015. A rate of convergence for mixture proportion estimation, with application to learning from noisy labels. In *AISTATS*.
- [61] Clayton Scott, Gilles Blanchard, and Gregory Handy. 2013. Classification with asymmetric label noise: Consistency and maximal denoising. In *COLT*.
- [62] Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- [63] Patrick E Shrout and Sean P Lane. 2012. Psychometrics. (2012).
- [64] Zijie Song, Jiawei Chen, Sheng Zhou, Qihao Shi, Yan Feng, Chun Chen, and Can Wang. 2023. CDR: Conservative doubly robust learning for debiased recommendation. In *CIKM*.
- [65] Harald Steck. 2010. Training and testing of recommender systems on data missing not at random. In *KDD*.
- [66] Adith Swaminathan and Thorsten Joachims. 2015. The self-normalized estimator for counterfactual learning. In *NeurIPS*.
- [67] Brendan Van Rooyen et al. 2015. Machine learning via transitions. (2015).
- [68] Brendan Van Rooyen, Aditya Menon, and Robert C Williamson. 2015. Learning with symmetric label noise: The importance of being uninged. In *NeurIPS*.
- [69] Hao Wang, Tai-Wei Chang, Tianqiao Liu, Jianmin Huang, Zhichao Chen, Chao Yu, Ruopeng Li, and Wei Chu. 2022. ESCM²: Entire space counterfactual multi-task model for post-click conversion rate estimation. In *SIGIR*.
- [70] Jun Wang, Haoxuan Li, Chi Zhang, Dongxu Liang, Enyu Yu, Wenwu Ou, and Wenjia Wang. 2023. CounterCLR: Counterfactual contrastive learning with non-random missing data in recommendation. In *ICDM*.
- [71] Wenjie Wang, Fuli Feng, Xiangnan He, Liqiang Nie, and Tat-Seng Chua. 2021. Denoising implicit feedback for recommendation. In *WSDM*.
- [72] Wenjie Wang, Yang Zhang, Haoxuan Li, Peng Wu, Fuli Feng, and Xiangnan He. 2023. Causal Recommendation: Progresses and Future Directions. In *SIGIR*.
- [73] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2019. Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random. In *ICML*.
- [74] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2021. Combating Selection Biases in Recommender Systems with A Few Unbiased Ratings. In *WSDM*.
- [75] Zifeng Wang, Xi Chen, Shao-Lun Wen, Rui anFd Huang, Ercan E Kuruoglu, and Yefeng Zheng. 2020. Information Theoretic Counterfactual Learning from Missing-Not-At-Random Feedback. In *NeurIPS*.
- [76] Zhenlei Wang, Shiqi Shen, Zhipeng Wang, Bo Chen, Xu Chen, and Ji-Rong Wen. 2022. Unbiased sequential recommendation with latent confounders. In *WWW*.
- [77] 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*.
- [78] Xiaobo Xia, Tongliang Liu, Bo Han, Nannan Wang, Mingming Gong, Haifeng Liu, Gang Niu, Dacheng Tao, and Masashi Sugiyama. 2020. Part-dependent label noise: Towards instance-dependent label noise. In *NeurIPS*.
- [79] Xiaobo Xia, Tongliang Liu, Nannan Wang, Bo Han, Chen Gong, Gang Niu, and Masashi Sugiyama. 2019. Are anchor points really indispensable in label-noise learning?. In *NeurIPS*.
- [80] Mengyue Yang, Guohao Cai, Furui Liu, Jiarui Jin, Zhenhua Dong, Xiuqiang He, Jianye Hao, Weiqi Shao, Jun Wang, and Xu Chen. 2023. Debiased recommendation with user feature balancing. *ACM Transactions on Information Systems* (2023).
- [81] Mengyue Yang, Quanyu Dai, Zhenhua Dong, Xu Chen, Xiuqiang He, and Jun Wang. 2021. Top-N recommendation with counterfactual user preference simulation. In *CIKM*.
- [82] Chi Zhang, Rui Chen, Xiangyu Zhao, Qilong Han, and Li Li. 2023. Denoising and Prompt-Tuning for Multi-Behavior Recommendation. In *WWW*.
- [83] Honglei Zhang, Shuyi Wang, Haoxuan Li, Chunyuan Zheng, Xu Chen, Li Liu, Shanshan Luo, and Peng Wu. 2024. Uncovering the Propensity Identification Problem in Debiased Recommendations. In *ICDE*.
- [84] 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*.
- [85] Yongfeng Zhang, Xu Chen, et al. 2020. Explainable recommendation: A survey and new perspectives. *Foundations and Trends® in Information Retrieval* (2020), 1–101.
- [86] 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*.
- [87] Zhiheng Zhang, Quanyu Dai, Xu Chen, Zhenhua Dong, and Ruiming Tang. 2023. Robust causal inference for recommender system to overcome noisy confounders. In *SIGIR*.
- [88] Xiangyu Zhao, Long Xia, Jiliang Tang, and Dawei Yin. 2019. Deep reinforcement learning for search, recommendation, and online advertising: a survey. *ACM Sigweb Newsletter*, 1–15.
- [89] Yu Zheng, Chen Gao, Xiang Li, Xiangnan He, Depeng Jin, and Yong Li. 2021. Disentangling User Interest and Conformity for Recommendation with Causal Embedding. In *WWW*.

A Proofs

THEOREM 3.1 (BIAS OF OME-DR ESTIMATOR). *Given $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ with $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$, imputed errors \tilde{E} and learned propensities \hat{P} with $\hat{p}_{u,i} > 0$ for all user-item pairs, the bias of the OME-DR estimator is*

$$\begin{aligned} \text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \right. \\ &\quad \left. \sum_{(u,i): r_{u,i}^* = 1} \left(\frac{p_{u,i} \omega_{11} - \hat{p}_{u,i}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{p_{u,i} \omega_{01}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 0) \right) + \right. \\ &\quad \left. \sum_{(u,i): r_{u,i}^* = 0} \left(\frac{p_{u,i} \omega_{10}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{p_{u,i} \omega_{00} - \hat{p}_{u,i}}{\hat{p}_{u,i}} \ell(f_{\theta}(x_{u,i}), 0) \right) \right|, \end{aligned}$$

where ω_{11} , ω_{01} , ω_{10} , and ω_{00} are given by

$$\begin{aligned} \omega_{11} &= \frac{1 - \rho_{01} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, & \omega_{01} &= \frac{\rho_{01} - \hat{\rho}_{01}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \\ \omega_{10} &= \frac{\rho_{10} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, & \omega_{00} &= \frac{1 - \hat{\rho}_{01} - \rho_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}. \end{aligned}$$

PROOF. By definition, the bias of the OME-DR estimator is

$$\text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] = \left| \mathbb{E}_{\mathbf{R}, \mathbf{O}} [\mathcal{E}_{\text{OME-DR}}] - \mathcal{P}^* \right|.$$

By double expectation formula, the first term is

$$\begin{aligned} &\mathbb{E}_{\mathbf{R}, \mathbf{O}} [\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] \\ &= \mathbb{E}_{\mathbf{R} | \mathbf{O}} [\mathbb{E}_{\mathbf{O}} [\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})]] \\ &= \mathbb{E}_{\mathbf{R} | \mathbf{O}} \left[\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left(\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right) \right] \\ &= \frac{1}{|\mathcal{D}|} \sum_{(u,i): r_{u,i}^* = 1} \mathbb{E}_{\mathbf{R} | \mathbf{O}} \left[\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right] \\ &\quad + \frac{1}{|\mathcal{D}|} \sum_{(u,i): r_{u,i}^* = 0} \mathbb{E}_{\mathbf{R} | \mathbf{O}} \left[\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right]. \end{aligned}$$

Then, we can derive the bias of the OME-DR estimator as follows

$$\begin{aligned} \text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}, \mathbf{R}^o; \hat{\rho}_{01}, \hat{\rho}_{10})] &= \left| \mathbb{E}_{\mathbf{R}, \mathbf{O}} [\mathbb{E}_{\mathbf{O}} [\mathcal{E}_{\text{OME-DR}}]] - \mathcal{P}^* \right| \\ &= \left| \frac{1}{|\mathcal{D}|} \sum_{(u,i): r_{u,i}^* = 1} \mathbb{E}_{\mathbf{R} | \mathbf{O}} \left[\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right] - \ell(f_{\theta}(x_{u,i}), 1) \right. \\ &\quad \left. + \frac{1}{|\mathcal{D}|} \sum_{(u,i): r_{u,i}^* = 0} \mathbb{E}_{\mathbf{R} | \mathbf{O}} \left[\left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \frac{p_{u,i} \tilde{e}_{u,i}}{\hat{p}_{u,i}} \right] - \ell(f_{\theta}(x_{u,i}), 0) \right| \\ &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i): r_{u,i}^* = 1} \left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} + \sum_{(u,i): r_{u,i}^* = 0} \left(1 - \frac{p_{u,i}}{\hat{p}_{u,i}}\right) \tilde{e}_{u,i} \right. \\ &\quad \left. + \sum_{(u,i): r_{u,i}^* = 1} \left(\frac{p_{u,i}}{\hat{p}_{u,i}} \cdot \mathbb{E}_{\mathbf{R} | \mathbf{O}} [\tilde{e}_{u,i}] - \ell(f_{\theta}(x_{u,i}), 1) \right) \right. \\ &\quad \left. + \sum_{(u,i): r_{u,i}^* = 0} \left(\frac{p_{u,i}}{\hat{p}_{u,i}} \cdot \mathbb{E}_{\mathbf{R} | \mathbf{O}} [\tilde{e}_{u,i}] - \ell(f_{\theta}(x_{u,i}), 0) \right) \right|. \end{aligned}$$

On one hand, for the user-item pairs with $r_{u,i}^* = 1$, we have

$$\begin{aligned} & \mathbb{E}_{\mathbf{R}|\mathbf{O}} [\tilde{e}_{u,i}] \\ &= (1 - \rho_{01}) \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 1; \hat{\rho}_{01}, \hat{\rho}_{10}) + \rho_{01} \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 0; \hat{\rho}_{01}, \hat{\rho}_{10}) \\ &= (1 - \rho_{01}) \cdot \frac{(1 - \hat{\rho}_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \hat{\rho}_{01} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \\ &\quad + \rho_{01} \cdot \frac{(1 - \hat{\rho}_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \hat{\rho}_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \\ &= \frac{1 - \rho_{01} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{\rho_{01} - \hat{\rho}_{01}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 0) \\ &:= \omega_{11} \ell(f_{\theta}(x_{u,i}), 1) + \omega_{01} \ell(f_{\theta}(x_{u,i}), 0), \end{aligned}$$

where ω_{11} and ω_{01} are given by

$$\omega_{11} = \frac{1 - \rho_{01} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \quad \omega_{01} = \frac{\rho_{01} - \hat{\rho}_{01}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}.$$

On the other hand, for the user-item pairs with $r_{u,i}^* = 0$, we have

$$\begin{aligned} & \mathbb{E}_{\mathbf{R}|\mathbf{O}} [\tilde{e}_{u,i}] \\ &= \rho_{10} \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 1; \hat{\rho}_{01}, \hat{\rho}_{10}) + (1 - \rho_{10}) \cdot \tilde{\ell}(f_{\theta}(x_{u,i}), 0; \hat{\rho}_{01}, \hat{\rho}_{10}) \\ &= \rho_{10} \cdot \frac{(1 - \hat{\rho}_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \hat{\rho}_{01} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \\ &\quad + (1 - \rho_{10}) \cdot \frac{(1 - \hat{\rho}_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \hat{\rho}_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \\ &= \frac{\rho_{10} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 1) + \frac{1 - \hat{\rho}_{01} - \rho_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \ell(f_{\theta}(x_{u,i}), 0) \\ &:= \omega_{10} \ell(f_{\theta}(x_{u,i}), 1) + \omega_{00} \ell(f_{\theta}(x_{u,i}), 0), \end{aligned}$$

where ω_{10} and ω_{00} are given by

$$\omega_{10} = \frac{\rho_{10} - \hat{\rho}_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \quad \omega_{00} = \frac{1 - \hat{\rho}_{01} - \rho_{10}}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}.$$

This completes the proof. \square

LEMMA 3.3 (LIPSCHITZ CONTINUITY). *Given $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ with $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$, if $\ell(f_{\theta}(x_{u,i}), r_{u,i}^*)$ is L -Lipschitz in $f_{\theta}(x_{u,i})$ for all $r_{u,i}^*$, then $\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})$ is $\frac{2L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}$ -Lipschitz in $f_{\theta}(x_{u,i})$ for all $r_{u,i}$.*

PROOF. The explicit form of $\tilde{\ell}(f_{\theta}(x_{u,i}), 1)$ and $\tilde{\ell}(f_{\theta}(x_{u,i}), 0)$ are

$$\begin{aligned} \tilde{\ell}(f_{\theta}(x_{u,i}), 1) &= \frac{(1 - \hat{\rho}_{10}) \ell(f_{\theta}(x_{u,i}), 1) - \hat{\rho}_{01} \ell(f_{\theta}(x_{u,i}), 0)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}, \\ \tilde{\ell}(f_{\theta}(x_{u,i}), 0) &= \frac{(1 - \hat{\rho}_{01}) \ell(f_{\theta}(x_{u,i}), 0) - \hat{\rho}_{10} \ell(f_{\theta}(x_{u,i}), 1)}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}. \end{aligned}$$

Then the conclusion can be derived directly from $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$. \square

LEMMA A.1 (McDIARMID'S INEQUALITY). *Let V be some set and let $f : V^m \rightarrow \mathbb{R}$ be a function of m variables such that for some $c > 0$, for all $i \in [m]$ and for all $x_1, \dots, x_m, x'_i \in V$ we have*

$$|f(x_1, \dots, x_m) - f(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_m)| \leq c$$

Let X_1, \dots, X_m be m independent random variables taking values in V . Then, with probability of at least $1 - \eta$ we have

$$|f(X_1, \dots, X_m) - \mathbb{E}[f(X_1, \dots, X_m)]| \leq c \sqrt{\log\left(\frac{2}{\eta}\right) m/2}.$$

PROOF. The proof can be found in Lemma 26.4 of [62]. \square

THEOREM 3.4 (GENERALIZATION BOUND). *Given $\hat{\rho}_{01}$ and $\hat{\rho}_{10}$ with $\hat{\rho}_{01} + \hat{\rho}_{10} < 1$, suppose $\ell(f_{\theta}(x_{u,i}), r_{u,i}^*)$ is L -Lipschitz in $f_{\theta}(x_{u,i})$ for all $r_{u,i}^*$, and $\hat{\rho}_{u,i} \geq C_p$, $|\tilde{\ell}(f_{\theta}(x_{u,i}), r_{u,i})| \leq C_l$ for all $r_{u,i}$, then with probability $1 - \eta$, the true prediction inaccuracy $\mathcal{P}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^*)$ of the optimal prediction matrix using the OME-DR estimator with imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$ has the upper bound*

$$\begin{aligned} & \mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^*; \hat{\rho}_{01}, \hat{\rho}_{10}) + \text{Bias}[\mathcal{E}_{\text{OME-DR}}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^*; \hat{\rho}_{01}, \hat{\rho}_{10})] + \\ & \left(1 + \frac{2}{C_p}\right) \left[\frac{4L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \mathcal{R}(\mathcal{F}) + \left(C_l + \frac{4L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}}\right) \sqrt{\frac{2 \log(4/\eta)}{|\mathcal{D}|}} \right]. \end{aligned}$$

PROOF. The true prediction inaccuracy can be decomposed as $\mathcal{P}^* = \mathcal{E}_{\text{OME-DR}} + (\mathcal{P}^* - \mathbb{E}[\mathcal{E}_{\text{OME-DR}}]) + (\mathbb{E}[\mathcal{E}_{\text{OME-DR}}] - \mathcal{E}_{\text{OME-DR}}) \leq \mathcal{E}_{\text{OME-DR}} + \text{Bias}(\mathcal{E}_{\text{OME-DR}}) + (\mathbb{E}[\mathcal{E}_{\text{OME-DR}}] - \mathcal{E}_{\text{OME-DR}}) \leq \mathcal{E}_{\text{OME-DR}} + \text{Bias}(\mathcal{E}_{\text{OME-DR}}) + \sup_{f_{\theta} \in \mathcal{F}} (\mathbb{E}[\mathcal{E}_{\text{OME-DR}}] - \mathcal{E}_{\text{OME-DR}})$.

To simplify notation, let $\mathcal{B}(\mathcal{F}) = \sup_{f_{\theta} \in \mathcal{F}} (\mathbb{E}[\mathcal{E}_{\text{OME-DR}}] - \mathcal{E}_{\text{OME-DR}})$, which can be decomposed as

$$\mathcal{B}(\mathcal{F}) = \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} [\mathcal{B}(\mathcal{F})] + \left\{ \mathcal{B}(\mathcal{F}) - \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} [\mathcal{B}(\mathcal{F})] \right\}.$$

For the first term $\mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} [\mathcal{B}(\mathcal{F})]$, according to the Rademacher Comparison Lemma [62] and Talagrand's Lemma [43], we have

$$\begin{aligned} & \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} [\mathcal{B}(\mathcal{F})] \\ & \leq 2 \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} \mathbb{E}_{\sigma} \sup_{f_{\theta} \in \mathcal{F}} \left[\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \sigma_{u,i} \left(\left(1 - \frac{\sigma_{u,i}}{\hat{\rho}_{u,i}}\right) \tilde{e}_{u,i} + \frac{\sigma_{u,i} \tilde{e}_{u,i}}{\hat{\rho}_{u,i}} \right) \right] \\ & \leq 2 \left(1 + \frac{2}{C_p}\right) \frac{2L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \cdot \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} \mathbb{E}_{\sigma} \sup_{f_{\theta} \in \mathcal{F}} \left[\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \sigma_{u,i} f_{\theta}(x_{u,i}) \right] \\ & = 2 \left(1 + \frac{2}{C_p}\right) \frac{2L}{1 - \hat{\rho}_{01} - \hat{\rho}_{10}} \cdot \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} \{\mathcal{R}(\mathcal{F})\}, \end{aligned}$$

where $\mathcal{R}(\mathcal{F})$ is the empirical Rademacher complexity of \mathcal{F}

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

By applying McDiarmid's inequality, and let $c = \frac{2}{|\mathcal{D}|}$, with probability at least $1 - \frac{\eta}{2}$,

$$\left| \mathcal{R}(\mathcal{F}) - \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} \{\mathcal{R}(\mathcal{F})\} \right| \leq 2 \sqrt{\frac{\log(4/\eta)}{2|\mathcal{D}|}} = \sqrt{\frac{2 \log(4/\eta)}{|\mathcal{D}|}}.$$

For the rest term $\mathcal{B}(\mathcal{F}) - \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} [\mathcal{B}(\mathcal{F})]$, by applying McDiarmid's inequality, and let $c = \frac{2}{|\mathcal{D}|} C_l \left(1 + \frac{2}{C_p}\right)$, with probability at least $1 - \frac{\eta}{2}$,

$$\left| \mathcal{B}(\mathcal{F}) - \mathbb{E}_{S \sim \mathbb{P}^{|\mathcal{D}|}} [\mathcal{B}(\mathcal{F})] \right| \leq C_l \left(1 + \frac{2}{C_p}\right) \sqrt{\frac{2 \log(4/\eta)}{|\mathcal{D}|}}.$$

After adding the inequalities above, we can rearrange the terms to obtain the stated results. \square