TEACHER IMITATION IS NOT ENOUGH: EXPOSURE-AWARE OPTIMAL-TRANSPORT DISTILLATION FOR RECOMMENDATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Knowledge distillation (KD) for recommendation tasks typically encourages a compact student to imitate a high-capacity teacher. Yet teachers are trained on logged interactions shaped by uneven historical exposure and popularity feedback, so direct distillation can propagate popularity bias together with useful ranking knowledge. This raises a fundamental question: should a student imitate everything its teacher has learned? We propose EOT-KD, an exposure-aware optimal-transport distillation framework that separates *teacher correction* from *teacher transfer*. Using interaction frequency as an exposure proxy, EOT-KD first constructs a minimum-distortion exposure correction of the teacher distribution; the closest distribution in Kullback–Leibler (KL) divergence while penalizing expected historical exposure. This correction admits a closed-form solution and provides monotonic control over the exposure score through a single parameter. The corrected target is then transferred through entropic optimal transport, with pairwise costs defined by cosine distance between teacher-derived item representations, assigning lower transport costs to more similar items. Across four recommendation benchmarks, EOT-KD improves Tail-NDCG@20 over the teacher by 25.9–48.9% while retaining 94.8–99.7% of the teacher's NDCG@20. Teacher and OT computations are used only during training. At inference, only the graph-free student is retained, requiring no teacher, graph propagation, or optimal transport.

1 INTRODUCTION

Knowledge distillation (KD) provides an effective mechanism for transferring predictive knowledge from high-capacity models to compact students Hinton et al. (2015). In recommender systems, KD is particularly attractive because deployment often requires ranking large item catalogues under strict latency and memory constraints (Tang & Wang, 2018; Kang et al., 2020). Existing recommender and graph distillation methods transfer teacher knowledge through soft-label matching, ranking-based supervision, latent-representation alignment, and relational or structural knowledge transfer (Tang & Wang, 2018; Kang et al., 2020; 2022; Chen et al., 2023a; Zhang et al., 2021; Wu et al., 2023; Feng et al., 2024; Forouzandeh et al., 2025). Despite these different mechanisms, most methods treat teacher supervision as a target to reproduce, which can propagate teacher bias.

In recommendation, however, the teacher is learned from observational user–item interactions. Such feedback can reflect not only user preference but also selection bias, historical exposure, and popularity effects Schnabel et al. (2016); Joachims et al. (2017); Chen et al. (2023b). In the absence of impression logs, we use training interaction frequency as an empirical proxy for historical exposure. As shown in Fig. 1, interactions are strongly skewed toward a small set of frequently interacted items, so the teacher may encode popularity-biased behavior alongside useful personalized ranking knowledge.

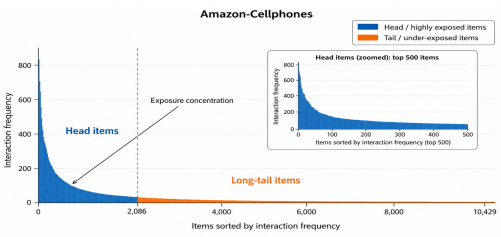


Figure 1: Long-tail item-popularity distribution on Amazon-Cellphones. Items are sorted by training interaction frequency; the dashed line marks the 80/20 head–tail partition.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Prior work shows that recommender distillation can amplify popularity bias, with popular items recommended more heavily and long-tail items further under-represented after distillation (Chen et al., 2023a). UnKD addresses this by partitioning items into a fixed set of popularity groups and distilling ranking knowledge within each group. The partition itself is fixed and coarse, giving no explicit, continuously controllable correction of a user's predictive distribution and no indication of how strongly that distribution should be reshaped toward the tail. Broader long-tail exposure also does not by itself guarantee stronger retrieval of *relevant* long-tail items, since low interaction frequency need not imply relevance to a particular user. A correction is therefore needed that acts directly on each user's distribution, reduces popularity concentration in a controlled and continuous manner, and preserves the teacher's personalized ranking structure. Such a correction must answer two questions: where should probability mass be redistributed within a user's candidate set, and how should this redistribution be realized by a capacity-limited student while respecting relationships among candidate items?

To address these questions, we introduce **EOT-KD**, an **E**xposure-Aware **O**ptimal-**T**ransport **K**nowledge **D**istillation framework that allows *teacher correction* before *teacher transfer*. Rather than directly suppressing popular items or treating inverse popularity as an independent supervision signal, EOT-KD formulates teacher correction as a minimum-distortion projection. Over a finite candidate set, EOT-KD constructs a corrected teacher distribution by minimizing its KL divergence from the original teacher distribution while penalizing the expected log-exposure score. The resulting projection admits a closed-form solution, remains conditioned on the teacher's personalized distribution, and provides a single parameter ρ that controls the strength of exposure correction, as detailed in Section 2.2. Importantly, ρ has an interpretable monotonic effect. Increasing ρ monotonically decreases the expected log-exposure score under the projected teacher distribution. The method therefore provides a continuous trade-off between fidelity to the teacher and correction of exposure concentration, rather than arbitrarily promoting low-frequency items.

The corrected teacher distribution determines *what* should be transferred, but not *how* the corrected probability mass should be realized by a capacity-limited student. Common distribution-matching objectives in KD, such as KL divergence, penalize item-wise discrepancies independently and therefore ignore relationships among candidate items. To account for such relationships, EOT-KD therefore transfers the projected distribution using entropic optimal transport (OT) under a teacher-induced item geometry. Specifically, the ground cost is defined by cosine distance between teacher-derived item representations, so assigning mass to a more similar item incurs a smaller transport penalty than assigning it to a dissimilar item. This yields the decomposition

$$\underbrace{p_u^T \longrightarrow p_u^\star}_{\text{minimum-distortion exposure correction}} \qquad \underbrace{p_u^\star \xrightarrow{\text{OT}} p_u^S}_{\text{geometry-aware knowledge transfer}} \quad .$$

The broader significance of EOT-KD lies in reconsidering what should be transferred from a recommender teacher. Teacher predictions can combine useful personalized ranking information with popularity concentration inherited from training-interaction frequency. EOT-KD addresses this tension by correcting the teacher target before transfer, improving relevant long-tail retrieval while largely preserving ranking utility. Our main contributions are as follows:

- **Minimum-distortion exposure correction.** We formulate exposure-aware distillation as a KL-regularized projection of the teacher distribution, $p^\star_{u,i}(\rho) \propto p^T_{u,i} \exp(-\rho\phi_i)$, rather than direct inverse-popularity reweighting. The projection remains conditioned on the teacher's personalized distribution, admits a closed-form solution, and provides provably monotonic control over the expected log-exposure score through the parameter ρ.
- **Geometry-aware transfer of the corrected teacher.** We transfer the exposure-corrected teacher distribution through entropic optimal transport under a teacher-induced item geometry. The transport cost encourages probability mass shifted toward lower-exposure items to be assigned preferentially to semantically or collaboratively related candidates.
- **Empirical utility–long-tail trade-off with lightweight deployment.** Across four real-world benchmarks, EOT-KD improves Tail-NDCG@20 over the teacher by 25.9–48.9% while retaining 94.8–99.7% of the teacher's NDCG@20. All teacher and OT computations are confined to training. At inference, only the lightweight graph-free student is retained, requiring neither the teacher, graph propagation, nor OT computation.

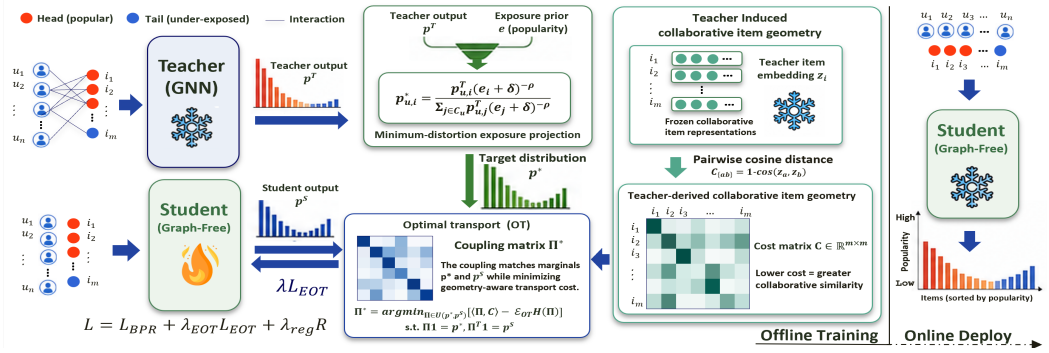


Figure 2: Overview of EOT-KD. The frozen teacher distribution is first transformed through a minimum-distortion exposure projection. Teacher-derived item representations define the transport geometry, and entropic optimal transport aligns the corrected teacher distribution with the student.

2 Methodology

We propose EOT-KD (Exposure-Aware Optimal Transport Knowledge Distillation), which decomposes recommender distillation into two sequential stages: *teacher correction* and *teacher transfer*. First, rather than treating the teacher distribution as immutable supervision, EOT-KD computes a minimum-distortion exposure-corrected projection of that distribution. Second, the corrected distribution is transferred to the student through optimal transport under an item-level geometry. This geometry preserves the teacher's relational structure during redistribution. These components have distinct roles: the projection determines what teacher distribution should be transferred, whereas the transport geometry determines how discrepancies between that corrected distribution and the student should be resolved. Figure 2 provides an overview of the framework.

2.1 Problem Setup

Let $\mathcal{U}$ and $\mathcal{I}$ denote the user and item sets, and $\mathcal{D}^+ \subseteq \mathcal{U} \times \mathcal{I}$ the observed training interactions. We use a pretrained teacher f_T and a lightweight student f_S, with the teacher frozen during student training. For each triplet (u, i_u^+, i_u^-), we construct $\mathcal{C}_u = \{i_u^+, i_u^-\} \cup \mathcal{R}_u \cup \mathcal{T}_u = \{c_{u,1}, \ldots, c_{u,K}\}$, where $\mathcal{R}_u$ and $\mathcal{T}_u$ contain randomly sampled and long-tail candidate items, respectively. Let $s_{u,k}^T = f_T(u, c_{u,k})$ and $s_{u,k}^S = f_S(u, c_{u,k})$ denote the teacher and student scores for user u and candidate $c_{u,k}$. Their candidate-level distributions are

$$p_{u,k}^T = \frac{\exp(s_{u,k}^T/\tau)}{\sum_{r=1}^K \exp(s_{u,r}^T/\tau)}, \qquad p_{u,k}^S = \frac{\exp(s_{u,k}^S/\tau)}{\sum_{r=1}^K \exp(s_{u,r}^S/\tau)}, \tag{1}$$

where $\tau > 0$ is the distillation temperature.

2.2 Minimum-Distortion Exposure Projection

Our objective is to correct exposure-related concentration in the teacher distribution without discarding its personalized ranking structure. We therefore seek a corrected distribution that reduces expected historical exposure while remaining as close as possible to the original teacher supervision. This leads naturally to a minimum-distortion projection in distribution space.

Let $e_i \geq 0$ denote an item-level exposure signal for item i. EOT-KD is agnostic to how this signal is obtained: explicit impression or propensity information may be used when available. In standard implicit-feedback benchmarks, where such logs are unavailable, we use training interaction frequency as an empirical proxy for historical exposure.

Specifically, let $n_i = \sum_{u \in \mathcal{U}} \mathbb{I}[(u, i) \in \mathcal{D}^+]$ denote the number of observed training interactions involving item i. We normalize this quantity as

$$e_i = \frac{n_i}{\bar{n}}, \qquad \bar{n} = \frac{1}{|\mathcal{I}^+|} \sum_{j \in \mathcal{I}^+} n_j, \qquad \mathcal{I}^+ = \{j : n_j > 0\}. \tag{2}$$

For candidate item $i \in \mathcal{C}_u$, define the log-exposure score $\phi_i = \log(e_i + \delta)$, where $\delta > 0$ is a numerical stabilizer. We define the corrected teacher distribution as

$$p_u^\star(\rho) = \arg\min_{q \in \Delta(\mathcal{C}_u)} D_{\mathrm{KL}}\big(q \| p_u^T\big) + \rho \mathbb{E}_{i\sim q}[\phi_i], \tag{3}$$

where $\Delta(\mathcal{C}_u)$ denotes the probability simplex over the candidate set and $\rho \geq 0$ controls the strength of exposure correction. The logarithm compresses the heavy-tailed range of interaction frequencies, yielding a smoother exposure score. The expectation $\mathbb{E}_{i\sim q}[\phi_i] = \sum_{i\in\mathcal{C}_u} q_i \phi_i$ measures the exposure level of the probability mass assigned by q; placing more mass on highly exposed items increases this term. The KL term preserves the personalized structure of the teacher distribution, while the second term penalizes expected log-exposure. As established formally in Proposition 1, the minimum-distortion projection admits the closed-form solution

$$p_{u,i}^\star(\rho) = \frac{p_{u,i}^T \exp(-\rho\phi_i)}{\sum_{j\in\mathcal{C}_u} p_{u,j}^T \exp(-\rho\phi_j)} = \frac{p_{u,i}^T (e_i+\delta)^{-\rho}}{\sum_{j\in\mathcal{C}_u} p_{u,j}^T (e_j+\delta)^{-\rho}}. \tag{4}$$

When $\rho = 0$, the projection reduces to the original teacher distribution, i.e., $p_u^\star = p_u^T$. The projection also admits a constrained interpretation:

$$\begin{aligned} \min_{q\in\Delta(\mathcal{C}_u)} \quad & D_{\mathrm{KL}}\big(q\|p_u^T\big) \\ \text{s.t.} \quad & \mathbb{E}_{i\sim q}[\phi_i] \leq b_u. \end{aligned} \tag{5}$$

Its Lagrangian recovers Eq. 3, with $\rho \geq 0$ as the dual variable associated with the exposure constraint. Hence, ρ parameterizes a family of minimum-distortion solutions corresponding to different exposure budgets.

2.3 Relevance-Aware Transport Geometry

The minimum-distortion projection determines the corrected teacher distribution, but it does not specify how discrepancies between this distribution and a capacity-limited student should be resolved. EOT-KD therefore equips the candidate space with a relevance-aware transport geometry.

Let $\mathbf{z}_i$ denote an item representation used to define this geometry. We consider two instantiations:

$$\mathbf{z}_i = \begin{cases} \mathbf{e}_i^T, & \text{collaborative geometry,} \\ \mathbf{r}_i, & \text{semantic geometry,} \end{cases} \tag{6}$$

where $\mathbf{e}_i^T$ denotes the frozen teacher-derived collaborative representation of item i, while $\mathbf{r}_i$ denotes its semantic representation constructed from training-set review text. Specifically, reviews associated with item i are encoded independently using a pretrained sentence encoder, and the resulting review-level embeddings are aggregated by mean pooling to obtain a fixed item-level representation.

For candidates $c_{u,a}$ and $c_{u,b}$, the ground cost is

$$C_{u,ab} = 1 - \cos\left(\mathbf{z}_{c_{u,a}}, \mathbf{z}_{c_{u,b}}\right). \tag{7}$$

Let $\mathbf{C}_u = [C_{u,ab}]_{a,b=1}^K$ denote the resulting ground-cost matrix. Hence, assigning probability to a related alternative incurs a smaller transport cost than assigning it to an unrelated item.

2.4 Exposure-Aware Optimal Transport Distillation

The projection in Eq. 3 determines *what* teacher distribution should be transferred. We next address *how* this corrected distribution should be realized by the student. Given $p_u^\star$, p_u^S, and the ground-cost matrix $\mathbf{C}_u$, define the set of feasible couplings

$$\mathcal{U}(p_u^*, p_u^S) = \left\{\Pi \in \mathbb{R}_+^{K\times K} : \Pi\mathbf{1} = p_u^*,\ \Pi^\top\mathbf{1} = p_u^S\right\}. \tag{8}$$

We obtain a smooth transport plan using entropy-regularized OT:

$$\Pi_{u,\varepsilon}^* = \arg\min_{\Pi\in\mathcal{U}(p_u^*, p_u^S)} \left[\langle \Pi, \mathbf{C}_u\rangle - \varepsilon_{\mathrm{ot}} H(\Pi)\right], \tag{9}$$

where $H(\Pi) = -\sum_{a,b} \Pi_{ab} \log \Pi_{ab}$ denotes the transport-plan entropy and $\varepsilon_{\text{ot}} > 0$ controls entropic regularization. We compute $\Pi^*_{u,\varepsilon}$ using differentiable Sinkhorn iterations.

We use entropic regularization to obtain a smooth, differentiable transport coupling via Sinkhorn iterations Cuturi (2013). The resulting coupling depends on the entropy regularization, while the distillation loss evaluates the ground transport cost under that coupling:

$$\mathcal{L}_{\text{EOT}} = \frac{1}{|\mathcal{B}|} \sum_{u \in \mathcal{B}} \left\langle \Pi^*_{u,\varepsilon}, \mathbf{C}_u \right\rangle. \tag{10}$$

Unlike coordinate-wise divergence, the transport objective accounts for relationships among candidate items. The two stages therefore play complementary roles. The minimum-distortion projection modifies the *source supervision* by determining which teacher probability mass should be retained or reduced under the exposure penalty. Optimal transport then determines how discrepancies between this corrected source distribution and the student can be resolved under collaborative or semantic item geometry.

2.5 Learning Objective and Deployment

The student is jointly optimized for recommendation and exposure-aware knowledge transfer:

$$\mathcal{L} = \mathcal{L}_{\text{BPR}} + \lambda_{\text{EOT}} \mathcal{L}_{\text{EOT}} + \lambda_{\text{reg}} \mathcal{R}, \tag{11}$$

where

$$\mathcal{L}_{\text{BPR}} = -\frac{1}{|\mathcal{B}|} \sum_{(u,i^+,i^-) \in \mathcal{B}} \log \sigma\left(s^S_{u,i^+_u} - s^S_{u,i^-_u} \right). \tag{12}$$

and $\mathcal{R}$ denotes embedding regularization. All teacher, exposure-reshaping, semantic-representation, and OT computations are used only during training. At inference, EOT-KD requires only the graph-free student and therefore introduces no additional deployment-time cost.

3 Theoretical Properties of Exposure Projection

We next characterize the exposure projection introduced in Eq. 3. The analysis establishes two properties central to EOT-KD: first, the exposure-adjusted target used by the algorithm is the exact minimum-distortion solution; second, the correction parameter ρ provides monotonic control over the expected exposure score. Proofs are provided in Appendix 12.

Proposition 1 (Closed-form minimum-distortion solution)**.** *For a teacher distribution p^T_u with positive support over $\mathcal{C}_u$, the objective*

$$\min_{q \in \Delta(\mathcal{C}_u)} D_{\text{KL}}(q \| p^T_u) + \rho \mathbb{E}_{i \sim q}[\phi_i] \tag{13}$$

has the unique solution

$$p^\star_{u,i}(\rho) = \frac{p^T_{u,i} \exp(-\rho \phi_i)}{\sum_j p^T_{u,j} \exp(-\rho \phi_j)}. \tag{14}$$

Consequently, the exposure-weighted target implemented by EOT-KD is the exact optimizer of the minimum-distortion objective.

Proposition 2 (Monotonic exposure control)**.** *Let*

$$\mu_u(\rho) = \mathbb{E}_{i \sim p^\star_u(\rho)}[\phi_i] \tag{15}$$

denote the expected log-exposure score under the projected teacher distribution. Then

$$\frac{\partial \mu_u(\rho)}{\partial \rho} = -\operatorname{Var}_{p^\star_u(\rho)}[\phi_i] \leq 0. \tag{16}$$

The inequality is strict whenever the candidate items do not all have identical exposure scores.

Proposition 2 gives ρ a direct interpretation: increasing ρ monotonically decreases the expected log-exposure score under the projected teacher distribution. This provides a distribution-level characterization of the exposure correction rather than describing only pairwise reweighting effects.

Corollary 1 (Teacher recovery and pairwise correction)**.** *At $\rho = 0$, $p^\star_u(0) = p^T_u$. Moreover, for any candidates a and b,*

$$\log \frac{p^\star_{u,a}(\rho)}{p^\star_{u,b}(\rho)} = \log \frac{p^T_{u,a}}{p^T_{u,b}} - \rho(\phi_a - \phi_b). \tag{17}$$

Hence, if $\phi_a < \phi_b$, increasing ρ monotonically increases the relative probability of the lower-exposure item a with respect to b.

Role of transport geometry. The preceding results characterize the correction stage, but they do not determine which student mismatches should be considered similar. If the ground cost were replaced by a discrete metric $C_{u,ab} = \mathbb{I}[a \neq b]$, transport would distinguish only item identity. In contrast, the collaborative or semantic cost used by EOT-KD assigns smaller transport penalties between related items. The empirical contribution of this geometry is evaluated through the cost-function ablations in Section 4.

4 Experiments

We evaluate EOT-KD through five questions: (i) whether it preserves overall recommendation utility; (ii) whether increased long-tail exposure translates into retrieval of *relevant* tail items; (iii) how the exposure projection and optimal transport contribute individually; (iv) how the correction strength and transport geometry affect the utility–long-tail trade-off; and (v) what computational overhead is introduced during training and deployment.

4.1 Experimental Setup

We evaluate EOT-KD on four implicit-feedback benchmarks: LastFM Cantador et al. (2011), Amazon-Cellphones, Amazon-Beauty, and Amazon-Sports McAuley et al. (2015). These datasets span different domains and popularity distributions. Item popularity is computed solely from training interactions. Items are ranked by training interaction frequency; the top $\lceil 0.2|\mathcal{I}| \rceil$ define the head and the remainder the long tail. No validation or test interactions are used to construct the popularity partition or exposure proxy. We use a 3-layer LightGCN teacher with 64-dimensional embeddings and a 64-dimensional Bayesian Personalized Ranking Matrix Factorization (BPR-MF) student. The teacher is trained separately for each dataset and seed and frozen before distillation. For each seed, all methods use the same data split and teacher checkpoint. The graph-free BPR-MF student trained without distillation is reported as No-KD. We compare EOT-KD with the pretrained LightGCN teacher, the non-distilled BPR-MF student, conventional distillation baselines KD-MSE and KD-KL, and recommendation-oriented methods including RD, CD, DE-RRD, HTD, KRD, and UnKD. These baselines cover direct teacher imitation, ranking-based and relational distillation, and popularity-aware knowledge transfer. Additional details are provided in Appendix 13.

4.2 Evaluation Metrics

We evaluate EOT-KD along three complementary dimensions: *ranking utility*, *relevant long-tail retrieval*, and *exposure balance*. For ranking utility, we report Recall@K and NDCG@K for $K \in \{10, 20\}$, with $K = 20$ used for the primary comparison. To distinguish relevant long-tail retrieval from indiscriminate tail-item promotion, we additionally report Tail-Recall@K and Tail-NDCG@K, computed using held-out relevant tail items. Exposure balance is evaluated using Catalogue Coverage@K, APLT@K, and Gini@K. Precision@K and AvgPop@K are reported in additional analyses. Higher Coverage and APLT and lower Gini indicate broader and less popularity-concentrated recommendation. Full metric definitions are provided in Appendix 9.

4.3 Overall Recommendation Utility

We first examine whether EOT-KD preserves conventional recommendation utility while modifying the transferred teacher signal. Table 1 compares EOT-KD with the LightGCN teacher, the non-distilled MF student, and the popularity-aware UnKD baseline at $K \in \{10, 20\}$. EOT-KD remains close to the teacher at both cutoffs while substantially improving over the non-distilled student. At $K = 20$, for example, EOT-KD reaches NDCG values of 0.2039, 0.0730, 0.0698, and 0.0420 on

Table 1: Recommendation utility at $K \in \{10, 20\}$. The teacher and MF are shown as a reference.

		LastFM		Amazon-Cellphones		Amazon-Beauty		Amazon-Sports	
K	Method	R@K	N@K	R@K	N@K	R@K	N@K	R@K	N@K
20	Teacher	0.2662±0.0002	0.2046±0.0007	0.1537±0.0005	0.0755±0.0002	0.1350±0.0007	0.0720±0.0004	0.0904±0.0023	0.0443±0.0009
	MF (No KD)	0.1943±0.0002	0.1463±0.0026	0.0612±0.0009	0.0303±0.0001	0.0732±0.0005	0.0397±0.0001	0.0324±0.0008	0.0161±0.0005
	UnKD	0.2501±0.0018	0.1884±0.0006	0.1395±0.0002	0.0685±0.0000	0.1068±0.0013	0.0555±0.0005	0.0728±0.0004	0.0366±0.0004
	EOT-KD	**0.2627±0.0035**	**0.2039±0.0047**	**0.1469±0.0002**	**0.0730±0.0003**	**0.1289±0.0003**	**0.0698±0.0005**	**0.0840±0.0005**	**0.0420±0.0008**
10	Teacher	0.1776±0.00044	0.1669±0.00076	0.1076±0.00094	0.0632±0.00023	0.0957±0.00118	0.0606±0.00053	0.0608±0.00028	0.0359±0.00025
	MF (No KD)	0.1253±0.00175	0.1168±0.00300	0.0436±0.00002	0.0255±0.00013	0.0510±0.00151	0.0331±0.00035	0.0212±0.00106	0.0129±0.00061
	UnKD	0.1628±0.00084	0.1509±0.00090	0.0985±0.00168	0.0576±0.00049	0.0745±0.00057	0.0461±0.00096	0.0500±0.00068	0.0302±0.00073
	EOT-KD	**0.1745±0.00330**	**0.1662±0.00460**	**0.1031±0.00019**	**0.0612±0.00025**	**0.0925±0.00058**	**0.0593±0.00048**	**0.0571±0.00063**	**0.0344±0.00079**

Table 2: Relevant long-tail retrieval and exposure balance at $K = 20$. TR and TN denote Tail-Recall and Tail-NDCG, respectively. The teacher and MF student without KD are shown as references. Bold and underlining indicate the best and second-best non-reference results, respectively.

Dataset	Method	TR@20 ↑	TN@20 ↑	Coverage (%) ↑	APLT ↑	Gini ↓
LastFM	Teacher	0.0741±0.0011	0.0379±0.0001	46.40±0.16	0.1388±0.0003	0.8399±0.0004
	MF (No KD)	0.0540±0.0015	0.0277±0.0007	50.76±0.46	0.1539±0.0094	0.8171±0.0069
	KD-MSE	0.0713±0.0005	0.0362±0.0007	45.62±0.28	0.1316±0.0021	0.8441±0.0006
	KD-KL	0.0741±0.0063	0.0380±0.0013	47.27±0.47	0.1422±0.0004	0.8364±0.0020
	RD	0.0493±0.0117	0.0252±0.0058	41.47±8.21	0.1203±0.0435	0.8583±0.0358
	CD	0.0608±0.0058	0.0297±0.0014	45.11±1.45	0.1369±0.0061	0.8426±0.0046
	DE-RRD	0.0575±0.0045	0.0271±0.0019	49.37±0.88	0.1548±0.0011	0.8234±0.0026
	HTD	0.0757±0.0013	0.0373±0.0020	48.86±1.09	0.1580±0.0112	0.8208±0.0089
	UnKD	0.0969±0.0020	0.0473±0.0005	54.03±0.99	0.2200±0.0005	0.7813±0.0022
	KRD	0.0765±0.0056	0.0383±0.0022	47.34±0.32	0.1430±0.0002	0.8346±0.0012
	EOT-KD	**0.1076±0.0077**	**0.0537±0.0024**	**60.33±0.09**	**0.2329±0.0012**	**0.7634±0.0002**
Cellphones	Teacher	0.0305±0.0005	0.0133±0.0001	54.21±0.26	0.0634±0.0012	0.9234±0.0001
	MF (No KD)	0.0203±0.0012	0.0100±0.0008	87.01±0.47	0.2417±0.0042	0.7351±0.0041
	KD-MSE	0.0307±0.0011	0.0134±0.0002	54.90±0.41	0.0653±0.0009	0.9202±0.0003
	KD-KL	0.0303±0.0011	0.0131±0.0003	56.78±0.20	0.0672±0.0012	0.9184±0.0001
	RD	0.0154±0.0020	0.0075±0.0006	62.25±5.19	0.0664±0.0160	0.9058±0.0174
	CD	0.0160±0.0025	0.0076±0.0007	53.54±4.73	0.0530±0.0101	0.9226±0.0121
	DE-RRD	0.0150±0.0006	0.0064±0.0000	27.45±0.42	0.0218±0.0001	0.9569±0.0005
	HTD	0.0271±0.0017	0.0117±0.0005	52.83±2.41	0.0565±0.0052	0.9207±0.0072
	UnKD	0.0282±0.0019	0.0112±0.0003	66.46±0.56	0.1112±0.0001	0.8842±0.0012
	KRD	0.0311±0.0005	0.0135±0.0000	57.90±0.41	0.0688±0.0012	0.9145±0.0002
	EOT-KD	**0.0394±0.0002**	**0.0171±0.0001**	**82.25±0.14**	**0.1478±0.0015**	**0.8384±0.0002**
Beauty	Teacher	0.0447±0.0005	0.0197±0.0002	75.49±0.24	0.1662±0.0028	0.8172±0.0009
	MF (No KD)	0.0264±0.0020	0.0119±0.0006	86.77±0.06	0.2915±0.0006	0.6962±0.0007
	KD-MSE	0.0447±0.0004	0.0198±0.0002	75.63±0.23	0.1672±0.0027	0.8161±0.0008
	KD-KL	0.0440±0.0001	0.0195±0.0004	77.50±0.32	0.1764±0.0045	0.8070±0.0022
	RD	0.0169±0.0003	0.0072±0.0004	61.81±1.53	0.1019±0.0057	0.8839±0.0017
	CD	0.0165±0.0031	0.0072±0.0012	57.69±5.24	0.0953±0.0165	0.8890±0.0177
	DE-RRD	0.0161±0.0003	0.0071±0.0002	18.94±0.15	0.0516±0.0018	0.9744±0.0003
	HTD	0.0447±0.0001	0.0200±0.0004	82.46±0.19	0.2209±0.0042	0.7566±0.0024
	UnKD	0.0253±0.0023	0.0107±0.0014	70.20±7.75	0.1822±0.0516	0.8224±0.0425
	KRD	0.0450±0.0008	0.0199±0.0001	76.08±0.21	0.1699±0.0023	0.8129±0.0006
	EOT-KD	**0.0544±0.0025**	**0.0248±0.0005**	**92.79±2.09**	**0.3043±0.0139**	**0.6867±0.0161**
Sports	Teacher	0.0111±0.0006	0.0045±0.0003	50.66±2.18	0.0609±0.0041	0.9304±0.0035
	MF (No KD)	0.0072±0.0013	0.0029±0.0005	80.65±0.27	0.2238±0.0025	0.7540±0.0013
	KD-MSE	0.0118±0.0005	0.0048±0.0003	51.26±2.06	0.0622±0.0040	0.9284±0.0033
	KD-KL	0.0115±0.0005	0.0045±0.0003	53.54±1.80	0.0658±0.0038	0.9245±0.0031
	RD	0.0028±0.0004	0.0011±0.0002	44.67±0.64	0.0390±0.0007	0.9518±0.0000
	CD	0.0053±0.0009	0.0021±0.0003	52.17±7.73	0.0661±0.0164	0.9272±0.0181
	DE-RRD	0.0039±0.0010	0.0016±0.0003	20.44±3.04	0.0127±0.0034	0.9725±0.0050
	HTD	0.0102±0.0006	0.0039±0.0001	51.09±2.77	0.0617±0.0076	0.9210±0.0074
	UnKD	0.0102±0.0013	0.0038±0.0004	71.92±5.74	**0.1690±0.0419**	0.8498±0.0274
	KRD	0.0114±0.0005	0.0047±0.0003	53.30±1.86	0.0656±0.0038	0.9219±0.0032
	EOT-KD	**0.0163±0.0009**	**0.0067±0.0004**	**79.31±1.16**	0.1585±0.0051	**0.8365±0.0051**

LastFM, Cellphones, Beauty, and Sports, respectively, compared with 0.2046, 0.0755, 0.0720, and 0.0443 for the teacher. EOT-KD also achieves higher Recall and NDCG than UnKD on all four datasets at both $K = 10$ and $K = 20$. These results indicate that EOT-KD preserves strong overall ranking quality while modifying the transferred supervision to reduce exposure-driven popularity concentration. Bold indicates the best student result among the reported methods.

4.4 Relevant Long-Tail Recommendation

We next test whether broader tail exposure corresponds to *relevant* long-tail retrieval. Table 2 reports relevant long-tail retrieval and exposure-balance results across all datasets. EOT-KD improves both Tail-Recall and Tail-NDCG over the teacher on all datasets. Relative to the teacher, Tail-NDCG increases by approximately 41.7% on LastFM, 28.6% on Cellphones, 25.9% on Beauty, and 48.9% on Sports. The comparison also illustrates why exposure-oriented metrics alone are insufficient. The non-distilled MF model often produces high Coverage or APLT while having substantially weaker

Table 3: Component ablation on Amazon-Cellphones at $K = 20$.

Method	Projection	OT	N@20 ↑	TN@20 ↑	Coverage (%) ↑
KD-KL	✗	✗	0.0752±0.0006	0.0131±0.0003	56.78±0.20
Projection-KL	✓	✗	0.0754±0.0003	0.0147±0.0000	63.31±0.70
OT-$\rho = 0$	✗	✓	**0.0755±0.0001**	0.0141±0.0002	60.60±0.61
EOT-KD	✓	✓	0.0730±0.0003	**0.0171±0.0001**	**82.25±0.14**

Table 4: Effect of transport geometry on Amazon-Cellphones at $K = 20$.

Cost	R@20 ↑	N@20 ↑	TN@20 ↑	Coverage (%) ↑
Identity	0.1458±0.0006	0.0723±0.0009	0.0167±0.0004	77.19±1.13
Collaborative	0.1469±0.0002	**0.0730±0.0003**	0.0171±0.0001	82.25±0.14
Text	0.1170±0.0009	0.0588±0.0005	0.0167±0.0002	**82.90±0.26**
Hybrid	**0.1472±0.0007**	**0.0730±0.0005**	**0.0173±0.0002**	81.52±0.20

relevant-tail retrieval. Similarly, UnKD generally increases tail exposure, but this does not consistently translate into higher Tail-Recall or Tail-NDCG on Cellphones, Beauty, and Sports. EOT-KD instead increases tail exposure while also improving retrieval of held-out relevant tail interactions.

4.5 Ablation study

We isolate the contributions of the two central components of EOT-KD through a 2×2 ablation. KD-KL uses neither exposure-aware projection nor optimal transport. Projection-KL applies the minimum-distortion exposure projection but matches the projected target using KL divergence. OT-ρ=0 uses the OT formulation without exposure correction, and EOT-KD combines exposure-aware projection with OT-based distillation. All variants use the same candidate support, student architecture, data splits, and pretrained teacher checkpoints. The ablation reveals distinct roles for the two components. Relative to KD-KL, Projection-KL improves Tail-NDCG from 0.0131 to 0.0147 and Coverage from 56.78% to 63.31%, while preserving overall NDCG. OT-ρ=0 exhibits a complementary effect: using transport geometry alone yields the highest NDCG@20 (0.0755) and also improves Tail-NDCG and Coverage over KD-KL. Combining the two components changes the trade-off more substantially. EOT-KD increases Tail-NDCG to 0.0171 and Coverage to 82.25%, the highest values among all variants. Compared with OT-ρ=0, this corresponds to gains of approximately 21.3% in Tail-NDCG and 35.7% in Coverage, while NDCG@20 decreases by only about 3.3%. These results indicate that OT alone is effective at transferring ranking information, whereas exposure-aware projection redirects the transferred signal toward relevant long-tail items. Their combination therefore provides the intended utility–exposure trade-off of EOT-KD.

4.6 Effect of Transport Geometry

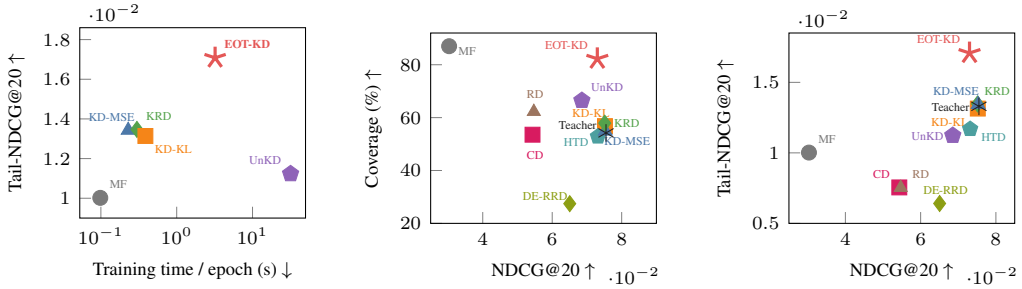
We compare four transport geometries on Amazon-Cellphones: identity, teacher-derived collaborative, review-text semantic, and hybrid collaborative–text costs. The choice of transport geometry affects the balance between ranking utility and exposure-oriented performance. Compared with identity matching, the collaborative cost improves all reported metrics. Text-based transport achieves the highest Coverage (82.90%) but substantially reduces ranking utility, suggesting that semantic similarity is not fully aligned with collaborative preference structure. The hybrid cost recovers utility and attains the highest Recall and Tail-NDCG. Overall, the collaborative geometry provides a strong balance between utility, long-tail retrieval, and coverage without requiring external side information, and is thus used as the default ground cost.

4.7 Sensitivity to Exposure Strength and Tail Support

We next examine the controllability of EOT-KD. Table 5 varies the exposure parameter ρ and the number of tail-support candidates on LastFM while fixing $K_{\mathrm{OT}} = 50$. Increasing ρ consistently shifts the recommendation distribution toward the long tail. For example, increasing ρ from 0.15 to 0.20 with $K_{\mathrm{tail}} = 20$ improves Tail-NDCG from 0.0538 to 0.0599 and Coverage from 60.43% to 67.02%, while reducing Gini from 0.7629 to 0.7233. This stronger exposure correction, however,

Table 5: Sensitivity to exposure strength ρ and tail-support size on LastFM.

ρ	K_{tail}	N@20 ↑	TN@20 ↑	Coverage (%) ↑	APLT ↑	Gini ↓
0.10	10	0.2025±0.0048	0.0473±0.0013	55.63±0.47	0.1990±0.0031	0.7907±0.0006
0.10	20	0.2024±0.0050	0.0495±0.0017	55.97±0.43	0.2050±0.0017	0.7881±0.0010
0.15	10	**0.2039±0.0047**	0.0537±0.0024	60.33±0.10	0.2329±0.0012	0.7634±0.0002
0.15	20	0.2012±0.0023	0.0538±0.0004	60.43±0.11	0.2371±0.0033	0.7629±0.0027
0.20	10	0.1987±0.0024	0.0580±0.0004	65.71±1.09	0.2744±0.0083	0.7300±0.0067
0.20	20	0.2003±0.0069	**0.0599±0.0022**	**67.02±0.14**	**0.2836±0.0026**	**0.7233±0.0006**



(a) (b) (c)

Figure 3: Computational and utility–exposure trade-offs on Amazon-Cellphones.

is accompanied by a modest reduction in overall NDCG. The results also show that ρ has a stronger effect than the tail-support size. At $\rho = 0.15$, increasing K_{tail} from 10 to 20 changes Tail-NDCG only from 0.0537 to 0.0538 and Coverage from 60.33% to 60.43%, while NDCG decreases from 0.2039 to 0.2012. In contrast, changing ρ produces a substantially larger shift in the utility–exposure trade-off. Notably, the $\rho = 0.15$, $K_{\text{tail}} = 10$ configuration attains the highest NDCG@20 among the tested exposure-aware settings, whereas stronger exposure correction yields progressively better long-tail and coverage-oriented metrics.

4.8 Computational Trade-off

EOT-KD introduces additional cost only during offline distillation. As shown in Fig. 3(a), Sinkhorn transport makes training more expensive than KD-KL, but EOT-KD achieves the highest Tail-NDCG among the compared methods while remaining approximately $10\times$ faster per epoch than UnKD. Figures 3(b)–(c) illustrate the resulting utility–exposure trade-off. Compared with KD-KL, EOT-KD increases Coverage from 56.78% to 82.25% and Tail-NDCG from 0.0131 to 0.0171, while retaining similar overall NDCG (0.0730 vs. 0.0752). Notably, although UnKD broadens catalogue exposure, this does not translate into stronger relevant-tail ranking on Cellphones. The additional cost does not carry over to deployment, the teacher and OT module are discarded after distillation, leaving the same MF student architecture. EOT-KD has essentially the same inference latency as MF, while remaining faster than the LightGCN teacher and requiring no graph propagation at deployment.

5 Conclusion

We introduced EOT-KD, an exposure-aware knowledge-distillation framework for recommendation that explicitly separates *what* should be transferred from *how* it should be transferred. Rather than directly imitating the teacher, it first constructs a minimum-distortion, exposure-corrected teacher distribution and then transfers this corrected signal through geometry-aware optimal transport. This formulation turns distillation from direct teacher reproduction into a controllable redistribution-and-transfer process. Across four benchmarks, EOT-KD improves relevant long-tail retrieval and catalogue exposure while preserving most of the teacher's ranking utility. Ablation studies confirm complementary roles for exposure projection and transport geometry: the projection redirects probability mass away from exposure-dominated items, while optimal transport preserves useful ranking structure during transfer. The exposure parameter ρ provides an explicit control over the utility–exposure trade-off. Since the teacher and OT module are used only during training, EOT-KD retains the same graph-free student architecture and essentially unchanged inference cost at deployment.

AI USE STATEMENT

Generative AI tools were used to assist with proofreading, language clarity, and LaTeX formatting. All suggested changes were reviewed and verified by the authors.

REPRODUCIBILITY STATEMENT

Implementation details, dataset preprocessing, evaluation protocols, hyperparameter settings, and additional experimental results are provided in the appendix. The anonymous source code is included in the supplementary material. For paired comparisons, methods under the same random seed use the same data split and pretrained teacher checkpoint.

REFERENCES

Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Controlling popularity bias in learning-to-rank recommendation. In *Proceedings of the eleventh ACM conference on recommender systems*, pp. 42–46, 2017.

Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Popularity-aware item weighting for long-tail recommendation. *arXiv preprint arXiv:1802.05382*, 2018.

Iván Cantador, Peter Brusilovsky, and Tsvi Kuflik. 2nd workshop on information heterogeneity and fusion in recommender systems (hetrec 2011). In *Proceedings of the 5th ACM Conference on Recommender Systems*, 2011.

Gang Chen, Jiawei Chen, Fuli Feng, Sheng Zhou, and Xiangnan He. Unbiased knowledge distillation for recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, WSDM '23, pp. 976–984. ACM, February 2023a. doi: 10.1145/3539597.3570477. URL http://dx.doi.org/10.1145/3539597.3570477.

Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3):1–39, 2023b.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26, 2013.

Yifan Feng, Yihe Luo, Shihui Ying, and Yue Gao. Lighthgnn: Distilling hypergraph neural networks into mlps for 100x faster inference. In *The Twelfth International Conference on Learning Representations*, 2024.

Saman Forouzandeh, Parham Moradi DW, and Mahdi Jalili. Distillhgnn: A knowledge distillation approach for high-speed hypergraph neural networks. In *The Thirteenth International Conference on Learning Representations*, 2025.

Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, pp. 639–648, 2020.

Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015.

Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. Unbiased learning-to-rank with biased feedback. In *Proceedings of the tenth ACM international conference on web search and data mining*, pp. 781–789, 2017.

SeongKu Kang, Junyoung Hwang, Wonbin Kweon, and Hwanjo Yu. De-rrd: A knowledge distillation framework for recommender system. In *Proceedings of the 29th ACM International Conference on Information amp; Knowledge Management*, CIKM '20, pp. 605–614. ACM, October 2020. doi: 10.1145/3340531.3412005. URL http://dx.doi.org/10.1145/3340531.3412005.

SeongKu Kang, Junyoung Hwang, Wonbin Kweon, and Hwanjo Yu. Topology distillation for recommender system. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 829–839, 2021.

SeongKu Kang, Dongha Lee, Wonbin Kweon, and Hwanjo Yu. Personalized knowledge distillation for recommender system. *Knowledge-Based Systems*, 239:107958, 2022.

Adit Krishnan, Ashish Sharma, Aravind Sankar, and Hari Sundaram. An adversarial approach to improve long-tail performance in neural collaborative filtering. In *Proceedings of the 27th ACM International Conference on information and knowledge management*, pp. 1491–1494, 2018.

Jae-woong Lee, Minjin Choi, Jongwuk Lee, and Hyunjung Shim. Collaborative distillation for top-n recommendation. In *2019 IEEE international conference on data mining (ICDM)*, pp. 369–378. IEEE, 2019.

Julian McAuley, Christopher Targett, Qinfeng Shi, and Anton van den Hengel. Image-based recommendations on styles and substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 43–52, 2015.

Gabriel Peyré and Marco Cuturi. *Computational optimal transport: With applications to data science*. Now Foundations and Trends, 2019.

Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618*, 2012.

Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. Recommendations as treatments: Debiasing learning and evaluation. In *international conference on machine learning*, pp. 1670–1679. PMLR, 2016.

Jiaxi Tang and Ke Wang. Ranking distillation: Learning compact ranking models with high performance for recommender system. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 2289–2298, 2018.

Tianxin Wei, Fuli Feng, Jiawei Chen, Ziwei Wu, Jinfeng Yi, and Xiangnan He. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, pp. 1791–1800, 2021.

Lirong Wu, Haitao Lin, Yufei Huang, and Stan Z Li. Quantifying the knowledge in gnns for reliable distillation into mlps. In *International conference on machine learning*, pp. 37571–37581. PMLR, 2023.

Shichang Zhang, Yozen Liu, Yizhou Sun, and Neil Shah. Graph-less neural networks: Teaching old mlps new tricks via distillation. *arXiv preprint arXiv:2110.08727*, 2021.

6 Related Work

Knowledge Distillation for Recommendation and Graph Models. Knowledge distillation (KD) transfers knowledge from a high-capacity teacher to a more compact student model (Hinton et al., 2015). In recommender systems, prior work has explored several forms of teacher supervision, including predicted scores, ranking relations, latent representations, and relational structure. Ranking Distillation (RD) transfers ranking information from the teacher by emphasizing highly ranked items (Tang & Wang, 2018). Collaborative Distillation (CD) (Lee et al., 2019) uses teacher predictions to construct informative distillation instances, while DE-RRD (Kang et al., 2020) combines list-wise ranking distillation with representation-level knowledge transfer. HTD further preserves relational structure by distilling the topology of the teacher embedding space (Kang et al., 2021). Knowledge distillation has also been explored in graph learning, where graph-based teachers transfer their structural knowledge to lightweight or graph-free students. GLNN (Zhang et al., 2021) distills the predictions of a GNN into an MLP to enable graph-free inference, while KRD improves GNN-to-MLP distillation by identifying reliable teacher knowledge (Wu et al., 2023). Beyond improving transfer effectiveness and efficiency, recent work has also examined popularity-related bias in teacher

supervision. More recently, Unbiased Knowledge Distillation (UnKD) (Chen et al., 2023a) studies how popularity bias in the teacher's soft labels can propagate to the student. Rather than distilling the teacher ranking globally, UnKD partitions items according to popularity and transfers ranking knowledge within popularity-aware groups.

Exposure, Popularity, and Long-Tail Bias in Recommendation. Recommendation models are typically learned from observational user–item interactions. An interaction can occur only after an item has been exposed to the user; consequently, observed interaction frequency reflects not only user preference but also the historical exposure process. Items receiving greater exposure can accumulate more interactions and become increasingly dominant in the training data, reinforcing the recommendation of popular items and reducing the visibility of the long tail.

Prior work has addressed biases arising from non-uniform observations through propensity weighting (Schnabel et al., 2016; Joachims et al., 2017), reweighting and regularization (Abdollahpouri et al., 2017), adversarial learning (Krishnan et al., 2018), causal adjustment (Wei et al., 2021), and other bias-aware recommendation objectives (Chen et al., 2023b). Related work on long-tail recommendation seeks to improve the visibility of less popular items while preserving personalized ranking quality (Abdollahpouri et al., 2018). Collectively, these studies highlight the importance of accounting for exposure and popularity effects in observational recommendation data.

7 PRELIMINARIES

This section introduces the basic notation and model definitions used throughout the paper, followed by background on knowledge distillation and optimal transport.

7.1 RECOMMENDATION SETTING

We consider implicit-feedback recommendation with a set of users $\mathcal{U}$ and a set of items $\mathcal{I}$. The observed training interactions are represented by $\mathcal{D}^+ \subseteq \mathcal{U} \times \mathcal{I}$, where $(u, i) \in \mathcal{D}^+$ indicates that user u has interacted with item i. For each user u, we denote the set of observed items by

$$\mathcal{I}_u^+ = \{i \in \mathcal{I} \mid (u, i) \in \mathcal{D}^+\}. \tag{18}$$

Items outside $\mathcal{I}_u^+$ are treated as unobserved and may be sampled as negatives during training.

A recommender model assigns a real-valued preference score s_{ui} to each user–item pair (u, i). At inference time, candidate items not observed during training are ranked according to these scores, and the highest-scoring items form the top-K recommendation list.

For pairwise implicit-feedback learning, training instances are constructed as triplets (u, i^+, i^-), where $i^+ \in \mathcal{I}_u^+$ is an observed item and $i^- \notin \mathcal{I}_u^+$ is an unobserved item. A standard Bayesian Personalized Ranking objective Rendle et al. (2012) is

$$\mathcal{L}_{\mathrm{BPR}} = -\frac{1}{|\mathcal{B}|} \sum_{(u,i^+,i^-)\in\mathcal{B}} \log \sigma(s_{ui^+} - s_{ui^-}), \tag{19}$$

where $\mathcal{B}$ denotes a training mini-batch and $\sigma(\cdot)$ is the sigmoid function.

For knowledge transfer, we additionally consider a candidate set $\mathcal{C}_u \subseteq \mathcal{I}$ for each user u. Given model scores $\{s_{ui}\}_{i\in\mathcal{C}_u}$, a probability distribution over the candidate items can be obtained using temperature-scaled softmax:

$$p_{ui} = \frac{\exp(s_{ui}/\tau)}{\sum_{j\in\mathcal{C}_u} \exp(s_{uj}/\tau)}, \qquad i \in \mathcal{C}_u, \tag{20}$$

where $\tau > 0$ controls the smoothness of the distribution.

7.2 LIGHTGCN TEACHER

LightGCN (He et al., 2020) represents user–item interactions as a bipartite graph and learns collaborative representations through neighborhood propagation without feature transformations or nonlinear activations. Let $\mathbf{e}_u^{(0)}$ and $\mathbf{e}_i^{(0)}$ denote the initial embeddings of user u and item i, respectively.

At layer ℓ, the embeddings are updated as

$$\mathbf{e}_u^{(\ell+1)} = \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u||\mathcal{N}_i|}} \mathbf{e}_i^{(\ell)}, \qquad \mathbf{e}_i^{(\ell+1)} = \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i||\mathcal{N}_u|}} \mathbf{e}_u^{(\ell)}. \tag{21}$$

The final user and item representations are obtained by combining the embeddings from all propagation layers:

$$\mathbf{e}_u = \sum_{\ell=0}^{L} \alpha_\ell \mathbf{e}_u^{(\ell)}, \qquad \mathbf{e}_i = \sum_{\ell=0}^{L} \alpha_\ell \mathbf{e}_i^{(\ell)}. \tag{22}$$

The preference score for a user–item pair is then computed by the inner product

$$s_{ui}^T = \mathbf{e}_u^\top \mathbf{e}_i. \tag{23}$$

7.3 BPR-MF Student

Our main student is a lightweight BPR-MF recommender with trainable user and item embeddings. Unlike the LightGCN teacher, the student does not perform neighborhood aggregation or graph propagation. Let $\mathbf{e}_u^S, \mathbf{e}_i^S \in \mathbb{R}^d$ denote the student embeddings of user u and item i, respectively. The preference score for a user–item pair is computed as

$$s_{ui}^S = (\mathbf{e}_u^S)^\top \mathbf{e}_i^S. \tag{24}$$

The student is trained for implicit-feedback recommendation using the BPR objective in Eq. 19, together with ℓ_2 regularization on the learned embeddings.

In the main experiments, both user and item embeddings have dimension $d = 64$. In the main BPR-MF experiments, the same student architecture is used for the No-KD student and all distilled methods, so differences in performance reflect the distillation strategy rather than changes in student capacity. At inference, candidate items are ranked directly by s_{ui}^S, requiring neither the interaction graph, the teacher, nor the optimal-transport module.

7.4 MLPCon Student

To examine whether EOT-KD generalizes beyond matrix-factorization students, we additionally consider a graph-free interaction MLP, denoted MLPCon. Unlike LightGCN, MLPCon does not perform neighborhood aggregation or graph propagation. Instead, it maintains trainable user and item embeddings and directly models their interaction through a multilayer perceptron.

Let $\mathbf{e}_u \in \mathbb{R}^d$ and $\mathbf{e}_i \in \mathbb{R}^d$ denote the trainable embeddings of user u and item i, respectively. For a user–item pair (u, i), the two embeddings are concatenated as

$$\mathbf{x}_{ui} = [\mathbf{e}_u \,\|\, \mathbf{e}_i] \in \mathbb{R}^{2d}, \tag{25}$$

where $\|$ denotes concatenation. The corresponding preference score is computed by an MLP f_ϕ:

$$s_{ui}^S = f_\phi(\mathbf{x}_{ui}), \tag{26}$$

where ϕ denotes the parameters of the interaction network.

The student is optimized for implicit-feedback recommendation using the BPR objective in Eq. 19, together with ℓ_2 regularization on the user and item embeddings.

At inference time, MLPCon scores candidate items independently using Eq. 26, without accessing the user–item graph. This makes the student structurally simpler than the graph-based teacher while retaining a nonlinear user–item interaction function.

7.5 Knowledge Distillation

Knowledge distillation (KD) transfers predictive knowledge from a high-capacity teacher model to a typically smaller student model Hinton et al. (2015). Let s_u^T and s_u^S denote the teacher and student

scores, respectively, over a candidate set $\mathcal{C}_u$ for user u. Temperature-scaled probability distributions are defined as

$$p_{ui}^T = \frac{\exp(s_{ui}^T/\tau)}{\sum_{j\in\mathcal{C}_u}\exp(s_{uj}^T/\tau)}, \qquad p_{ui}^S = \frac{\exp(s_{ui}^S/\tau)}{\sum_{j\in\mathcal{C}_u}\exp(s_{uj}^S/\tau)}. \tag{27}$$

where $\tau > 0$ is the temperature parameter.

A standard logit-based distillation objective encourages the student distribution to match the teacher distribution, commonly using the Kullback–Leibler divergence:

$$\mathcal{L}_{\mathrm{KD}} = \sum_u D_{\mathrm{KL}}\left(p_u^T \,\|\, p_u^S\right). \tag{28}$$

Thus, conventional KD directly treats the teacher distribution as the target to be reproduced by the student.

7.6 Optimal Transport

Optimal transport (OT) provides a framework for comparing probability distributions while taking into account the geometry of their support Peyré & Cuturi (2019). Let $\mu \in \mathbb{R}_+^n$ and $\nu \in \mathbb{R}_+^m$ be two probability distributions satisfying $\sum_i \mu_i = \sum_j \nu_j = 1$. Given a ground-cost matrix $C \in \mathbb{R}_+^{n\times m}$, where C_{ij} specifies the cost of transporting probability mass from element i to element j, OT seeks a coupling P with marginals μ and ν.

The set of admissible couplings is

$$\Pi(\mu,\nu) = \left\{P \in \mathbb{R}_+^{n\times m} \;\middle|\; P\mathbf{1}_m = \mu,\ P^\top \mathbf{1}_n = \nu\right\}, \tag{29}$$

where $\mathbf{1}_n$ denotes an n-dimensional vector of ones. The discrete OT problem is then

$$\mathcal{W}_C(\mu,\nu) = \min_{P\in\Pi(\mu,\nu)} \langle C, P\rangle = \min_{P\in\Pi(\mu,\nu)} \sum_{i=1}^n\sum_{j=1}^m C_{ij}P_{ij}, \tag{30}$$

where $\langle C, P\rangle$ denotes the Frobenius inner product.

For efficient optimization, we use the entropically regularized OT formulation Cuturi (2013):

$$\mathcal{W}_C^\varepsilon(\mu,\nu) = \min_{P\in\Pi(\mu,\nu)} \langle C, P\rangle - \varepsilon H(P), \tag{31}$$

where $\varepsilon > 0$ controls the regularization strength and

$$H(P) = -\sum_{i=1}^n\sum_{j=1}^m P_{ij}\big(\log P_{ij} - 1\big) \tag{32}$$

is the entropy of the transport plan. Entropic regularization yields a smooth transport problem that can be solved efficiently using Sinkhorn iterations Cuturi (2013). The resulting coupling P therefore depends jointly on the probability mass in the two distributions and on the pairwise geometry encoded by the ground-cost matrix C.

8 Dataset Statistics

We evaluate EOT-KD on four processed implicit-feedback recommendation benchmarks: LastFM, Amazon-Cellphones, Amazon-Beauty, and Amazon-Sports. Table 6 summarizes the statistics of the exact processed splits used in our experiments. We follow the processed train/validation/test splits provided by the UnKD preprocessing pipeline, corresponding approximately to an 80/10/10 split over the observed interactions. The same fixed splits are used for UnKD, EOT-KD, and all other compared methods. The interaction count denotes the total number of user–item interactions across the training, validation, and test splits. Dataset density is computed as $|\mathcal{D}|/(|\mathcal{U}||\mathcal{I}|)$.

All exposure-related quantities are computed exclusively from the training split. In particular, item interaction frequency, the exposure proxy, and the head–tail partition are constructed without using validation or test interactions. Items are ranked by training interaction frequency, with the most frequently interacted $\lceil 0.2|\mathcal{I}|\rceil$ items defining the head and the remaining items defining the long tail.

Table 6: Statistics of the processed datasets used in our experiments.

Dataset	#Users	#Items	Train	Validation	Test	Total	Density (%)
LastFM	1,892	4,489	42,135	5,210	5,323	52,668	0.6201
Amazon-Cellphones	27,879	10,429	157,180	18,705	18,554	194,439	0.0669
Amazon-Beauty	22,363	12,101	159,912	19,392	19,198	198,502	0.0734
Amazon-Sports	35,598	18,357	238,863	28,768	28,706	296,337	0.0453

9 Evaluation Metric Definitions

Let $\mathcal{U}$ and $\mathcal{I}$ denote the sets of users and items, respectively. For each user u, let $\mathcal{R}_u^K$ denote the top-K recommendation list and $\mathcal{G}_u^{\text{test}}$ the set of held-out relevant test items. We report metrics at $K \in \{10, 20\}$. Evaluation uses full-catalog ranking, with items observed in the training set excluded from the candidate set.

Overall Ranking Utility. Precision@K and Recall@K are defined as

$$\text{Precision@K} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{R}_u^K \cap \mathcal{G}_u^{\text{test}}|}{K}, \tag{33}$$

and

$$\text{Recall@K} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{R}_u^K \cap \mathcal{G}_u^{\text{test}}|}{|\mathcal{G}_u^{\text{test}}|}. \tag{34}$$

NDCG@K additionally accounts for the positions of relevant items. For user u,

$$\text{NDCG@K}_u = \frac{\text{DCG@K}_u}{\text{IDCG@K}_u}, \tag{35}$$

where

$$\text{DCG@K}_u = \sum_{r=1}^{K} \frac{\mathbb{III}(i_r \in \mathcal{G}_u^{\text{test}})}{\log_2(r+1)}, \tag{36}$$

i_r is the item ranked at position r, and IDCG@K_u denotes the corresponding ideal DCG. NDCG@K is averaged across users.

Relevant Long-Tail Recommendation. Items are ranked according to their interaction frequency in the training set. Following the protocol used throughout this work, the top 20% most frequently interacted items form the head $\mathcal{I}_{\text{head}}$, and the remaining 80% form the long tail $\mathcal{I}_{\text{tail}}$.

For each user, let

$$\mathcal{G}_{u,\text{tail}}^{\text{test}} = \mathcal{G}_u^{\text{test}} \cap \mathcal{I}_{\text{tail}}. \tag{37}$$

denote the held-out relevant tail items, and define

$$\mathcal{U}_{\text{tail}} = \left\{ u \in \mathcal{U} : |\mathcal{G}_{u,\text{tail}}^{\text{test}}| > 0 \right\}. \tag{38}$$

Tail-Recall@K is then

$$\text{TailRecall@K} = \frac{1}{|\mathcal{U}_{\text{tail}}|} \sum_{u \in \mathcal{U}_{\text{tail}}} \frac{|\mathcal{R}_u^K \cap \mathcal{G}_{u,\text{tail}}^{\text{test}}|}{|\mathcal{G}_{u,\text{tail}}^{\text{test}}|}. \tag{39}$$

Tail-NDCG@K is computed analogously to NDCG@K, with relevance defined only over $\mathcal{G}_{u,\text{tail}}^{\text{test}}$ and the corresponding ideal ranking constructed from these relevant tail items. The metric is averaged over $\mathcal{U}_{\text{tail}}$. These metrics measure whether increased long-tail exposure corresponds to retrieval of relevant tail items rather than indiscriminate promotion of low-popularity items.

Catalogue Coverage. Catalogue Coverage@K measures the fraction of catalogue items that appear at least once across users' recommendation lists:

$$\text{Coverage@K} = 100 \times \frac{\left|\bigcup_{u \in \mathcal{U}} \mathcal{R}_u^K\right|}{|\mathcal{I}|}. \tag{40}$$

Higher values indicate that recommendations cover a broader portion of the catalogue.

Average Percentage of Long-Tail Items. APLT@K measures the average proportion of recommended items belonging to the long tail:

$$\text{APLT@K} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{R}_u^K \cap \mathcal{I}_{\text{tail}}|}{K}. \tag{41}$$

Higher APLT indicates greater representation of long-tail items in the recommendation lists.

Average Popularity. Let n_i denote the number of training interactions associated with item i. AvgPop@K is defined as

$$\text{AvgPop@K} = \frac{1}{|\mathcal{U}|K} \sum_{u \in \mathcal{U}} \sum_{i \in \mathcal{R}_u^K} n_i. \tag{42}$$

Lower AvgPop indicates reduced concentration on frequently interacted items.

Exposure Concentration. Let c_i^K denote the number of times item i appears across all users' top-K recommendation lists. Sorting these counts in ascending order gives $c_{(1)}^K \leq \cdots \leq c_{(|\mathcal{I}|)}^K$. We compute Gini@$K$ as

$$\text{Gini@K} = \frac{2 \sum_{j=1}^{|\mathcal{I}|} j c_{(j)}^K}{|\mathcal{I}| \sum_{j=1}^{|\mathcal{I}|} c_{(j)}^K} - \frac{|\mathcal{I}| + 1}{|\mathcal{I}|}. \tag{43}$$

Lower Gini indicates a more balanced distribution of recommendation exposure across the catalogue.

10 Algorithm

Algorithm 1 EOT-KD Training

Input: Training interactions $\mathcal{D}^+$; frozen teacher f_T; student f_S; exposure scores $\{e_i\}$; correction strength ρ; temperature τ; OT regularization ε_{ot}; OT weight λ_{EOT}
Output: Trained student f_S
for each mini-batch $\mathcal{B}$ **do**
  Construct candidate sets $\{\mathcal{C}_u\}_{u \in \mathcal{B}}$
  Compute teacher and student distributions $\{p_u^T, p_u^S\}_{u \in \mathcal{B}}$
  for each $u \in \mathcal{B}$ **do**
    Compute the exposure-corrected teacher distribution p_u^* using Eq. 4
    Construct the ground-cost matrix $\mathbf{C}_u$ using Eq. 7
    Compute the regularized transport plan Π_u^* using Sinkhorn iterations in Eq. 9
  end for
  Compute $\mathcal{L}_{\text{EOT}}$ using Eq. 10
  Update f_S using Eq. 11
end for
return f_S

11 Computational Complexity

For a candidate set of size K, constructing the pairwise ground-cost matrix requires $\mathcal{O}(K^2 d)$ operations for d-dimensional item representations. Each Sinkhorn iteration requires $\mathcal{O}(K^2)$ time, while storing the transport quantities requires $\mathcal{O}(K^2)$ memory per user. With T_{sk} Sinkhorn iterations and mini-batch size B, the OT component therefore requires

$$\mathcal{O}\left(B(K^2 d + T_{\text{sk}} K^2)\right) \tag{44}$$

training-time computation and up to $\mathcal{O}(BK^2)$ transport memory when processed jointly. In practice, candidate sampling and batch chunking bound this overhead. Since the teacher and OT module are discarded after training, inference complexity and memory are identical to those of the student model.

12 PROOFS OF THEORETICAL RESULTS

12.1 PROOF OF PROPOSITION 1

Proof. For clarity, omit the user index u and write the objective as

$$\mathcal{J}(q) = \sum_{i \in \mathcal{C}_u} q_i \log \frac{q_i}{p_i^T} + \rho \sum_{i \in \mathcal{C}_u} q_i \phi_i, \tag{45}$$

subject to

$$q_i \geq 0, \qquad \sum_i q_i = 1. \tag{46}$$

Since $p_i^T > 0$, the KL term is strictly convex in q over the probability simplex, while the exposure term is linear. Therefore, $\mathcal{J}(q)$ is strictly convex and has a unique minimizer. Moreover, the minimizer lies in the interior of the simplex: as $q_i \to 0^+$, the derivative of $q_i \log(q_i/p_i^T)$ tends to $-\infty$, while the exposure term has finite derivative. Hence $q_i^\star > 0$ for all $i \in \mathcal{C}_u$, and it suffices to enforce the normalization constraint.

Introducing a Lagrange multiplier λ for the normalization constraint gives

$$\mathcal{L}(q, \lambda) = \sum_i q_i \log \frac{q_i}{p_i^T} + \rho \sum_i q_i \phi_i + \lambda \left(\sum_i q_i - 1 \right). \tag{47}$$

At the optimum,

$$\frac{\partial \mathcal{L}}{\partial q_i} = \log \frac{q_i}{p_i^T} + 1 + \rho \phi_i + \lambda = 0. \tag{48}$$

Hence

$$q_i = p_i^T \exp(-\rho \phi_i) \exp(-1 - \lambda). \tag{49}$$

The factor $\exp(-1-\lambda)$ is shared by all candidates and is determined by the normalization constraint. Therefore,

$$q_i = \frac{p_i^T \exp(-\rho \phi_i)}{\sum_j p_j^T \exp(-\rho \phi_j)}. \tag{50}$$

Substituting $\phi_i = \log(e_i + \delta)$ yields

$$q_i = \frac{p_i^T (e_i + \delta)^{-\rho}}{\sum_j p_j^T (e_j + \delta)^{-\rho}}. \tag{51}$$

Thus $q = p_u^\star(\rho)$, proving the result. □

12.2 PROOF OF PROPOSITION 2

Proof. Define the partition function

$$Z_u(\rho) = \sum_{i \in \mathcal{C}_u} p_{u,i}^T \exp(-\rho \phi_i). \tag{52}$$

Then

$$p_{u,i}^\star(\rho) = \frac{p_{u,i}^T \exp(-\rho \phi_i)}{Z_u(\rho)}. \tag{53}$$

Differentiating the log-partition function gives

$$\frac{\partial}{\partial \rho} \log Z_u(\rho) = \frac{1}{Z_u(\rho)} \sum_i p_{u,i}^T (-\phi_i) \exp(-\rho \phi_i) \tag{54}$$

$$= -\sum_i p_{u,i}^\star(\rho) \phi_i \tag{55}$$

$$= -\mu_u(\rho). \tag{56}$$

Therefore,

$$\mu_u(\rho) = -\frac{\partial}{\partial \rho} \log Z_u(\rho). \tag{57}$$

Differentiating once more,

$$\frac{\partial \mu_u(\rho)}{\partial \rho} = -\frac{\partial^2}{\partial \rho^2} \log Z_u(\rho). \tag{58}$$

The second derivative of the log-partition function is

$$\frac{\partial^2}{\partial \rho^2} \log Z_u(\rho) = \mathbb{E}_{p_u^\star(\rho)}[\phi_i^2] - \left(\mathbb{E}_{p_u^\star(\rho)}[\phi_i]\right)^2 \tag{59}$$

$$= \mathrm{Var}_{p_u^\star(\rho)}[\phi_i]. \tag{60}$$

Consequently,

$$\frac{\partial \mu_u(\rho)}{\partial \rho} = -\mathrm{Var}_{p_u^\star(\rho)}[\phi_i] \leq 0. \tag{61}$$

The inequality is strict whenever the exposure scores are not constant over the support of $p_u^\star(\rho)$. □

12.3 Proof of Corollary 1

Proof. Setting $\rho = 0$ in Eq. 4 gives

$$p_{u,i}^\star(0) = \frac{p_{u,i}^T}{\sum_j p_{u,j}^T} = p_{u,i}^T, \tag{62}$$

since p_u^T is a probability distribution.

For two candidates a and b,

$$\frac{p_{u,a}^\star(\rho)}{p_{u,b}^\star(\rho)} = \frac{p_{u,a}^T}{p_{u,b}^T} \exp\left[-\rho(\phi_a - \phi_b)\right]. \tag{63}$$

Taking logarithms yields

$$\log \frac{p_{u,a}^\star(\rho)}{p_{u,b}^\star(\rho)} = \log \frac{p_{u,a}^T}{p_{u,b}^T} - \rho(\phi_a - \phi_b). \tag{64}$$

If $\phi_a < \phi_b$, then

$$\frac{\partial}{\partial \rho} \log \frac{p_{u,a}^\star}{p_{u,b}^\star} = \phi_b - \phi_a > 0, \tag{65}$$

which proves the result. □

13 Implementation Details

All models are implemented in PyTorch. The LightGCN teacher uses three propagation layers and 64-dimensional embeddings and is trained with Adam using a learning rate of 10^{-3}, weight decay 10^{-4}, BPR batch size 8192, and at most 1000 epochs. The BPR-MF student also uses 64-dimensional embeddings. For EOT-KD, we use $\rho = 0.15$, $K_{\mathrm{OT}} = 50$, $K_{\mathrm{tail}} = 10$, and $\delta = 0.01$ unless otherwise stated. The default transport cost is cosine distance between frozen teacher item representations, and the transport plan is computed using differentiable Sinkhorn iterations.

Item popularity and the head–tail partition are constructed from training interactions only. Hyperparameters are selected using validation data and kept fixed across random seeds. Baseline hyperparameters follow the settings reported in the corresponding original papers. All reported multi-seed comparisons use matched teacher checkpoints and data splits. Each experiment is repeated over five random seeds and we report mean ± standard deviation. KD-MSE and KD-KL directly match student predictions to teacher supervision using mean squared error and KL divergence, respectively. Experiments were conducted on a machine equipped with an NVIDIA A10G GPU (22 GiB memory), 8 vCPUs, and 32 GiB of system memory.

14 ADDITIONAL RESULTS

14.1 PRECISION–POPULARITY TRADE-OFF

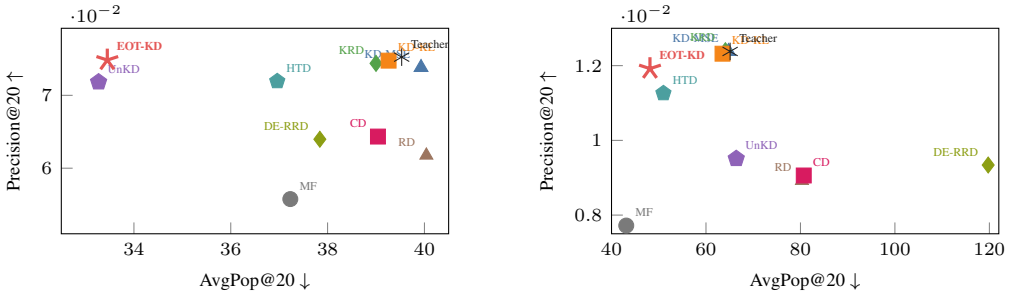
To complement the main results, we further examine the trade-off between ranking precision and recommendation popularity on two representative datasets. Figure 4 plots Precision@20 against AvgPop@20 for LastFM and Amazon-Beauty. Precision@20 measures the fraction of top-20 recommendations that correspond to held-out relevant items, while AvgPop@20 measures the average training interaction frequency of recommended items. Lower AvgPop therefore indicates less concentration on frequently interacted items, so the preferred region in each panel is the upper-left (higher Precision@20 with lower AvgPop@20).

The two datasets exhibit a consistent utility–popularity trade-off. Conventional teacher-matching methods such as KD-MSE and KD-KL remain close to the teacher in Precision@20, but also retain similar AvgPop levels. In contrast, the non-distilled MF student reduces AvgPop more strongly, but this is accompanied by a substantial loss in precision. This illustrates why lower recommendation popularity alone is not sufficient to characterize a useful long-tail recommendation model.

On LastFM, the teacher achieves Precision@20 of 0.07524 with AvgPop@20 of 39.53. EOT-KD reduces AvgPop@20 to 33.45 while retaining Precision@20 of 0.07484, remaining close to teacher-level precision. UnKD achieves a similarly low AvgPop@20 of 33.27, but with lower Precision@20 of 0.07182. Thus, EOT-KD combines a substantial reduction in recommendation popularity with limited loss in ranking precision.

A similar pattern is observed on Amazon-Beauty. The teacher obtains Precision@20 of 0.01238 with AvgPop@20 of 65.11, whereas EOT-KD reduces AvgPop@20 to 48.14 while maintaining Precision@20 of 0.01192. The non-distilled MF student reaches an even lower AvgPop@20 of 43.12, but its Precision@20 drops markedly to 0.00772. HTD also reduces AvgPop@20 to 50.99, although with lower precision than EOT-KD. Together, these results show that EOT-KD shifts recommendations toward less-frequently interacted items while preserving substantially more ranking utility than simply reducing popularity concentration.

Overall, the Precision–AvgPop analysis provides a complementary view of the utility–exposure trade-off studied in the main paper: the goal is not merely to recommend less-popular items, but to reduce popularity concentration while retaining useful recommendation quality.



(a) LastFM (b) Amazon-Beauty

Figure 4: Precision–popularity trade-offs on LastFM and Amazon-Beauty. The horizontal axis reports AvgPop@20 and the vertical axis reports Precision@20, so the preferred region is the upper-left, corresponding to higher precision with lower recommendation popularity.

14.2 TAIL-AWARE SUPPORT WITH AN ALTERNATIVE STUDENT

The main experiments use BPR-MF as the student. To examine whether the effect of tail-aware support is specific to this architecture, we additionally evaluate it with MLPCon, a nonlinear graph-free interaction model. Tail-aware support does not directly promote low-popularity items through the recommendation loss. Instead, it enlarges the candidate support available to the exposure projection and transport problem, allowing corrected teacher mass to reach low-exposure alternatives that may otherwise be absent from the sampled support.

Table 7 examines this effect with an MLPCon student. Introducing tail support ($K_{\text{tail}} = 10$) improves Recall, NDCG, and Tail-Recall over $K_{\text{tail}} = 0$. Increasing the support to 20 further improves Tail-NDCG, catalogue Coverage, and APLT, while reducing AvgPop and Gini, with only a small reduction in overall utility. These results indicate that tail-aware support controls the utility–exposure trade-off by expanding the alternatives available to the exposure correction rather than directly rewarding tail items.

Table 7: Effect of tail-support size with an MLPCon student. Higher is better except for AvgPop and Gini.

K_{tail}	R@20 ↑	N@20 ↑	TR@20 ↑	TN@20 ↑	Coverage (%) ↑	APLT ↑	AvgPop ↓	Gini ↓
0	0.23724	0.17973	0.06998	0.03684	54.24	0.18122	36.93	0.80843
10	**0.24229**	**0.18332**	**0.08194**	0.04128	54.00	0.19037	36.44	0.80453
20	0.23954	0.18235	0.08147	**0.04218**	**55.36**	**0.19715**	**36.37**	**0.79931**