# PERSONALIZED CONTENT GENERATION AS PREFERENCE DISTRIBUTION COMPLETION

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Personalized content generation is becoming increasingly important in modern generative systems. Existing approaches typically personalize provided prompts, instructions, or references; directly generating content from user behavior requires inferring preferences from limited observations. This makes it difficult to determine both what to generate and how generation probabilities should be distributed across interests. In this paper, we formulate the personalized content generation task as preference distribution completion. To this end, we develop a collaborative framework that draws on cross-user behavioral patterns to supplement sparse individual preferences. The framework combines source-distribution modeling with history-conditioned Riemannian flow matching, which together determine how probability mass is allocated in a semantic latent space. A decoder maps samples from the resulting distributions directly to content, without explicit user prompts. Experiments on short-video user interaction benchmarks, with textual descriptions as the generation targets, demonstrate the effectiveness of our approach.[1]

## 1 INTRODUCTION

Recent advances in generative modeling have opened up the possibility of creating content tailored to individual users, extending personalization from content selection to content creation (Rombach et al., 2022; Shen et al., 2024; Xu et al., 2025a). In contrast to generation guided by explicit predefined instructions or reference targets (Li et al., 2024; Xu et al., 2025b), a more natural and scalable paradigm is behavior-driven personalized content generation: given a user's interaction history, the system directly synthesizes content that aligns with the user's implicit tastes. A user's preferences can encompass multiple interests (Li et al., 2019; Cen et al., 2020), and reflecting their relative importance matters beyond merely covering a diverse set of topics (Steck, 2018). Accordingly, the fundamental objective shifts toward modeling a user's overall content preference distribution and generating content proportionally to that distribution, rather than producing only the single highest-scoring output or maximizing diversity alone, as illustrated in Figure 1.

Achieving this distributional alignment, however, poses a significant challenge, as it requires inferring the underlying preferences from a finite interaction history that provides only partial observations (Zhang et al., 2026; Ni et al., 2026). Some interests may be weakly represented or entirely absent from the empirical interactions, while even the relative importance of the observed interests remains highly uncertain. Prior approaches address catalog recommendation (Rajput et al., 2023; Wang et al., 2024; Yang et al., 2023), request-conditioned personalization (Salemi et al., 2024; Liu et al., 2025; Shi et al., 2025; Zhang et al., 2025), or preference-aligned creation (Dang et al., 2025; Dunlop et al., 2025; Patron et al., 2025), but leave insufficiently examined the joint problem of inferring a user's overall content preference distribution from sparse behavioral interactions and generating content proportionally across diverse interests.

To address this problem, we formulate personalized content generation as preference distribution completion and theoretically analyze the conditions under which cross-user interactions can support this completion from finite observations. Building on this foundation, we develop a collaborative generative framework that integrates dual-view conditioned Riemannian flow matching (Lipman et al., 2023; Chen & Lipman, 2024) with source-distribution calibration, as illustrated in Figure 2. By decoupling preference modeling in a semantic latent space from content decoding, our approach

[1]The code is available for review at `https://anonymous.4open.science/r/PCGaPDC-485E/`.



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.

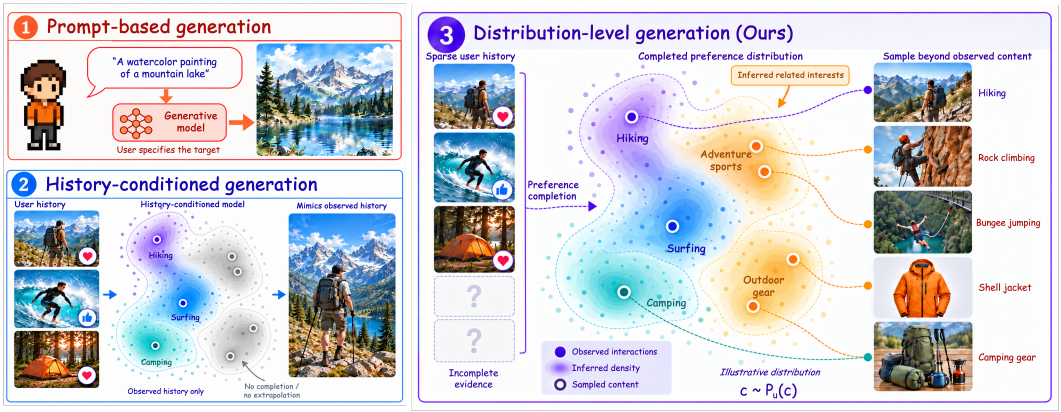


Figure 1: **Motivation and challenges of personalized content generation.** **(1)** Prompt-based generation produces content guided by explicit instructions or prompts. **(2)** History-conditioned generation personalizes outputs via observed interactions, but may remain concentrated around the observed preference evidence and fail to extrapolate to other potential interests. **(3)** Our distribution-level generation completes the user's preference distribution over sparse observations and samples content aligned with the inferred probability mass.

leverages collaborative cross-user patterns to supplement sparse individual histories, jointly conditioning the transport dynamics. A subsequent source-calibration phase refines the probability mass allocation to ensure the learned distribution accurately captures the user's multi-faceted interests. Consequently, personalization is naturally realized by sampling from this continuous distribution and decoding directly into content, bypassing explicit text prompts and post-hoc catalog reranking. Extensive evaluations on two short-video benchmarks, combining quantitative distribution-level metrics and LLM-based assessments (Zheng et al., 2023), demonstrate that our framework consistently outperforms baselines. Our contributions are summarized as follows:

- We formulate personalized content generation as preference distribution completion and provide a conditional analysis of how cross-user interactions can support completion from finite observations.
- We develop a collaborative generative framework that combines source-distribution modeling, dual-view conditioned Riemannian flow matching, and semantic decoding to generate personalized content directly.
- We evaluate the framework on short-video user interaction benchmarks, demonstrating consistent improvements over baselines across distribution-level metrics and LLM-based assessments, while ablation studies validate the contributions of core components.

## 2 Related Work

**Personalized Content Generation.** Existing research has explored several ways to incorporate personal preferences into generation. Request-conditioned approaches customize outputs for a predefined query or target: LaMP retrieves relevant profile entries, PPlug encodes historical contexts, and StyleVector steers generation through personalized activation directions (Salemi et al., 2024; Liu et al., 2025; Zhang et al., 2025). PMG, Pigeon, and DRC combine behavioral preferences with target content or multimodal instructions for visual synthesis (Shen et al., 2024; Xu et al., 2025a;b), while CFRAG, ClusterRAG, and CoPersona leverage cross-user evidence to improve query-conditioned generation (Shi et al., 2025; Nkhata et al., 2026; Zhang et al., 2026). Building on diffusion preference optimization (Wallace et al., 2024), PPD and C-DPO align image generation and editing with individual preferences using pairwise feedback (Dang et al., 2025; Dunlop et al., 2025). Beyond explicit prompts, REBECA learns ID- and rating-conditioned embedding distributions for image synthesis (Patron et al., 2025). These request-conditioned methods assume a supplied query or reference; existing behavior-driven approaches enable content creation but fail to jointly complete under-observed interests and match their generation probabilities. In contrast, our approach leverages interaction histories and collaborative information for request-free generation, targeting the full multi-interest preference distribution rather than merely the relevance of individual outputs.

**Multi-Interest and Distributional Preference Modeling.** Capturing multifaceted preferences has been studied in both recommendation and pluralistic preference modeling. Multi-interest frameworks represent heterogeneous user intents via multiple vectors (Li et al., 2019; Cen et al., 2020; Xie et al., 2023). Calibrated recommendation matches output proportions to historical interest distributions (Steck, 2018; Abdollahpouri et al., 2020; Feijer et al., 2025), distinct from balancing relevance and diversity (Wilhelm et al., 2018). Concurrently, probabilistic approaches model latent preferences and sequential uncertainty from incomplete interactions (Liang et al., 2016; 2018; Fan et al., 2022). Pluralistic frameworks study coverage, steerability, and population-level calibration (Sorensen et al., 2024), while distributional approaches model utility or match reward distributions via optimal transport (Siththaranjan et al., 2024; Melnyk et al., 2024). However, multi-vector representations and utility distributions do not specify generative content distributions, and matching empirical history alone cannot recover interests absent from finite observations. In contrast, our approach infers a semantic content distribution from individual and collaborative information, recovering under-observed interests and allocating generation probabilities proportionally across them.

# 3 Problem Formulation

We formulate personalized content generation as the completion of a user's preference distribution from finite positive interactions. The goal is to directly generate content aligned with the user's preferences by sampling from the inferred distribution.

## 3.1 Problem Setting

Let $(\mathcal{X}, \mathcal{A})$ be a measurable content space. For each user $u$, let $P_u^* \in \mathcal{P}(\mathcal{X})$ denote the user's unknown preference distribution over the content space. For simplicity, we omit the unobserved exposure process and assume that the observed finite interaction history $H_u = (X_{u,1}, \ldots, X_{u,n_u})$ satisfies $X_{u,1}, \ldots, X_{u,n_u} \mid P_u^* \overset{\text{iid}}{\sim} P_u^*$.

Given $H_u$ and group-level interaction data $\mathcal{D}$ from other users, we aim to learn a completion operator $Q_{\mathcal{D},H_u} = \widehat{\mathcal{C}}(\mathcal{D}, H_u)$. Personalized content is then generated directly by sampling $\widehat{X} \sim Q_{\mathcal{D},H_u}$.

Crucially, matching a preference distribution ensures that generated content proportionally reflects the user's varied interests. This distributional approach fundamentally differs from optimizing a top-$k$ ordering (the classic ranking problem). As we prove in Appendix A, an ordering alone cannot uniquely determine a generative distribution. It also goes beyond merely maximizing diversity, which does not guarantee alignment with underlying preferences.

## 3.2 Collaborative Preference Completion

**Cross-user information for completion.** The finite history of a target user provides only partial information about their true preferences and does not uniquely determine the underlying preference distribution. Specifically, multiple distributions can match the same observed interactions while assigning arbitrary probability mass to unobserved content regions (Appendix B). Cross-user interactions provide a potential source of complementary preference information. Agreement on observed items can help identify users whose additional feedback is informative about a target user's unobserved preferences, which is the central intuition behind collaborative filtering (Resnick et al., 1994). More broadly, interaction patterns can reveal shared preference structures linking users and content, allowing observations from different users to inform individual preference estimation (Koren et al., 2009). This motivates using information shared across users to restrict plausible completions.

**Theoretical justification.** In Appendix C, we provide a theoretical analysis detailing when cross-user interactions can support preference distribution estimation. Assuming that users share an underlying family of preference distributions satisfying a Hölder smoothness condition over a latent collaborative state, we prove that transferring preferences from localized neighbors bounds the estimation error. **Proposition 1** establishes a finite-sample error bound for a local collaborative estimator, proving that the distributional approximation error is tightly controlled by the interplay between cross-user transfer bias and the effective reference sample size. This confirms that collaborative completion is theoretically viable when related users can be identified with sufficient precision.

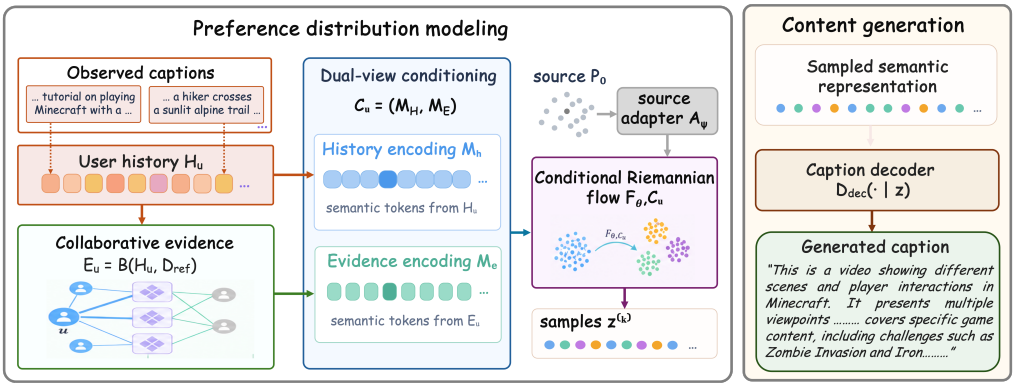


Figure 2: **Overview of our collaborative generative framework.** The framework combines source-distribution modeling with dual-view conditioned Riemannian flow matching to complete personalized preference distributions and generate user-preferred content.

**Learning from finite interactions.** We now return to the generative task. While the local estimator establishes the statistical viability of collaborative preference completion, it does not directly yield a generalizable generative model. To realize this principle, we formulate a shared, distribution-valued mapping:

$$\mathcal{C}_\Theta : (\mathcal{D}, H_u) \longmapsto Q_\Theta(\cdot \mid H_u, \mathcal{D}) \in \mathcal{P}(\mathcal{X}), \tag{1}$$

where $\Theta$ denotes parameters shared across all users. After training, $\widehat{\mathcal{C}} = \mathcal{C}_\Theta$ instantiates the operator, yielding $Q_{\mathcal{D},H_u} = Q_\Theta(\cdot \mid H_u, \mathcal{D})$. Personalized content is then generated directly by sampling $\widehat{X} \sim Q_\Theta(\cdot \mid H_u, \mathcal{D})$. Independent draws reflect the learned distribution, incurring a finite Monte Carlo sampling error distinct from the distribution discrepancy with $P_u^*$ (see Appendix D).

Learning this parametric map requires supervision for the conditional output distribution. Although the true preference distribution $P_u^*$ is unavailable as an explicit label, the assumption that observed interactions are drawn independently from $P_u^*$ provides a natural mechanism for sample-based evaluation. Specifically, any unseen interaction $X \sim P_u^*$ independent of the conditioning history $H_u$ yields an unbiased estimate $\mathbb{E}[f(X) \mid H_u, \mathcal{D}] = \int f \, \mathrm{d}P_u^*$ for any preference functional $f \in \mathcal{F}$.

Consequently, finite interactions provide sample-level supervision of the target preference distribution without requiring complete distributional labels. We leverage this principle by splitting each user's history into conditioning sequences and target items, driving a tractable sample-based learning objective. Section 4 details the realization of this framework.

## 4 The Method

We propose a collaborative framework that combines source calibration with dual-view Riemannian flow matching, separating preference distribution modeling from content decoding (Figure 2). We model user preferences in a fixed semantic latent space and use Riemannian flow matching to learn a conditional transport under a dual-view condition, where the user's observed history and supplementary evidence derived from cross-user interactions are encoded through separate pathways. We then calibrate the source distribution to rectify residual mismatch and refine probability mass allocation across preference modes. At inference, given the user's observed history, a frozen caption decoder directly maps semantic samples drawn from the resulting personalized distribution into video captions aligned with the user's preferences.

### 4.1 Decoupled Preference Modeling and Content Generation

We separate preference distribution modeling from content decoding: the former infers a semantic distribution from interactions, while the latter decodes samples drawn from this distribution into captions. To implement the completion rule in Equation (1), let $R_{\theta,\psi}(\cdot \mid H_u, \mathcal{D})$ denote a personalized distribution over a semantic space $\mathcal{Z}$, and let $K_{\mathrm{dec}}(\cdot \mid z) = \delta_{D_{\mathrm{dec}}(z)}(\cdot)$ denote the distribution induced by a fixed deterministic caption decoder $D_{\mathrm{dec}} : \mathcal{Z} \to \mathcal{X}$. The resulting content distribution

is

$$Q_\Theta(A \mid H_u, \mathcal{D}) = \int_{\mathcal{Z}} K_{\mathrm{dec}}(A \mid z)\, R_{\theta,\psi}(\mathrm{d}z \mid H_u, \mathcal{D}), \qquad A \in \mathcal{A}. \tag{2}$$

Here $\Theta = (\theta, \psi)$, where $\theta$ parameterizes the dual-view transport and its conditioning networks, and $\psi$ parameterizes the source adapter. Together, they define the semantic sampling distribution.

We instantiate the semantic space as the unit hypersphere $\mathcal{Z} = \mathbb{S}^{d-1}$ via a frozen, $\ell_2$-normalized CLIP text encoder $e : \mathcal{X} \to \mathbb{S}^{d-1}$ (Radford et al., 2021; Yang et al., 2022). This normalization defines the spherical geometry for conditional transport and induces the target semantic distribution $P_u^{\mathcal{Z}} = e_{\#}P_u^*$, for which encoded held-out interactions $e(Y_{u,j}) \sim P_u^{\mathcal{Z}}$ directly provide sample-level supervision. Conversely, we implement $K_{\mathrm{dec}}$ as a frozen, user-shared caption decoder (Qwen Team, 2026) conditioned solely on sampled spherical coordinates rather than interaction histories, isolating personalization entirely within the learned semantic distribution (see Appendix E for backbone details and counterfactual cycle-consistency checks).

### 4.2 Dual-view Riemannian Flow Matching

We learn a conditional transport in the fixed semantic space $\mathcal{Z}$, using separate pathways for the user's observed history and interaction-derived evidence. Following Section 3.2, we construct training episodes by partitioning a user's available interactions into a visible conditioning history $H_u$ and held-out target interactions $Y_{u,1}, \ldots, Y_{u,\ell_u}$. The history $H_u$ supplies the observed preference evidence, while the encoded targets $z_{u,j} = e(Y_{u,j})$ supervise the transport.

**Interaction-based evidence.** We obtain collaborative evidence via a frozen retrieval module $E_u = B(H_u; \mathcal{D}_{\mathrm{ref}})$, where $\mathcal{D}_{\mathrm{ref}}$ contains reference interactions from $\mathcal{D}$. We first learn item representations from the reference interaction graph using XSimGCL (Yu et al., 2024) and partition items into behavioral groups. By matching the groups covered in $H_u$ against reference users, we identify behaviorally aligned neighbors whose remaining interactions supply candidate captions. In this way, collaborative evidence is discovered through behavioral overlaps across interaction sets, operationalizing the local reference neighborhood condition in Eq. (18). The graph representations are used only for discovering such collaborative evidence. The generative model itself receives the fixed semantic encodings of the retrieved captions, which serve as additional conditioning information rather than positive feedback. Crucially, because retrieval is driven solely by visible history $H_u$, unseen users require neither presence in the interaction graph nor prelearned graph embeddings.

**Dual-view conditioning.** We encode observed history and retrieved collaborative evidence via separate set encoders, yielding $M_H = \mathrm{Enc}_\theta^H(e(H_u))$ and $M_E = \mathrm{Enc}_\theta^E(e(E_u))$, where $e$ independently encodes each caption and the evidence encoder ignores padded positions. The two memories form the condition $C_u = (M_H, M_E)$. Each flow block attends to the history and evidence through separate cross-attention modules (Vaswani et al., 2017). Given the hidden state $h^{(b)}$ at block $b$, their outputs are combined as

$$c^{(b)} = \sum_{V \in \{H,E\}} \omega_V^{(b)}\, \mathrm{Attn}_V^{(b)}(h^{(b)}, M_V), \tag{3}$$

where $(\omega_H^{(b)}, \omega_E^{(b)})$ are produced by a softmax gate conditioned on $h^{(b)}$. The memories remain fixed along each trajectory, while the gate balances both views as the flow state evolves.

**Conditional flow matching.** We train the transport using Riemannian flow matching (Lipman et al., 2023; Chen & Lipman, 2024). We parameterize the velocity field by projecting the ambient network output onto the tangent space of the sphere at the current representation:

$$v_\theta(t, z; C_u) = \Pi_z \widetilde{v}_\theta(t, z; C_u), \qquad \Pi_z = I - zz^\top, \tag{4}$$

thereby ensuring that the predicted velocity remains in $T_z\mathbb{S}^{d-1}$ throughout the transport.

We construct training episodes via a masked-set scheme, partitioning a user's interactions into a visible support set $H_u$ (retrieving evidence $E_u$ to form condition $C_u$) and masked targets $\{z_{u,j} = e(Y_{u,j})\}_{j=1}^{\ell_u}$. We define $P_0$ as a fixed population source distribution on $\mathcal{Z}$ constructed from training semantic representations. For each user, source samples $r \sim P_0$ and masked targets $z_1 \in \{z_{u,j}\}$

are paired within the same condition using a minimum-cost spherical assignment (Pooladian et al., 2023), ensuring every masked target is used once. Along each spherical geodesic $\gamma_t(r, z_1)$, we regress the predicted velocity field directly toward the analytical target velocity $v_t^\star = \partial_t \gamma_t(r, z_1)$:

$$\mathcal{L}_{\mathrm{RFM}}(\theta) = \mathbb{E}_{t,(r,z_1),C_u}\left[\|v_\theta(t, \gamma_t(r, z_1); C_u) - \partial_t \gamma_t(r, z_1)\|_2^2\right]. \tag{5}$$

To retain predictive historical cues during transport learning, we jointly optimize the condition representations via an auxiliary multi-positive context InfoNCE loss $\mathcal{L}_{\mathrm{ctx}}$ (Van den Oord et al., 2018):

$$\mathcal{L}_{\mathrm{transport}} = \mathcal{L}_{\mathrm{RFM}}(\theta) + \lambda_{\mathrm{ctx}}\mathcal{L}_{\mathrm{ctx}}. \tag{6}$$

Integrating $\dot{z}_t = v_\theta(t, z_t; C_u)$ from $t = 0$ to 1 yields the endpoint map $F_{\theta,C_u}$, defining the uncalibrated semantic pushforward $R^{(0)}_{\theta,u} = (F_{\theta,C_u})_\# P_0$.

### 4.3 Conditional Source Calibration

We calibrate the source distribution of the frozen transport to reduce residual mismatch between the generated endpoints and the user's preference samples without disturbing the learned flow dynamics. We parameterize the source adapter $A_\psi$ via a constrained tangent displacement followed by the spherical exponential map:

$$A_\psi(r; C_u) = \mathrm{Exp}_r\big(\delta_\psi(r; C_u)\big), \qquad \|\delta_\psi(r; C_u)\|_2 \le \rho, \tag{7}$$

where an MLP predicts $\delta_\psi$ conditioned on $C_u$ and initialized to zero, preserving an exact identity mapping at initialization. This adapter induces the conditional source $q_\psi(\cdot \mid C_u) = (A_\psi(\cdot; C_u))_\# P_0$, yielding the calibrated semantic distribution $R_{\theta,\psi} = (F_{\theta,C_u})_\# q_\psi$.

We map calibrated source samples through the frozen transport to produce endpoints $g_{u,k} = F_{\theta,C_u}(A_\psi(r_{u,k}; C_u))$ with $r_{u,k} \sim P_0$. We supervise $A_\psi$ against masked targets $\{z_{u,j}\}_{j=1}^{\ell_u}$ using a composite objective:

$$\mathcal{L}_{\mathrm{source}}(\psi) = \mathbb{E}\left[\widehat{\mathcal{L}}_{\mathrm{E},u} + \lambda_{\mathrm{cov}}\widehat{\mathcal{L}}_{\mathrm{cov},u} + \lambda_{\mathrm{move}}\widehat{\mathcal{L}}_{\mathrm{move},u}\right]. \tag{8}$$

Specifically, we use a lower-is-better energy-score objective (Gneiting & Raftery, 2007), $\widehat{\mathcal{L}}_{\mathrm{E},u} = \frac{1}{K\ell_u}\sum_{k=1}^{K}\sum_{j=1}^{\ell_u} d_{\mathrm{E}}(g_{u,k}, z_{u,j}) - \frac{1}{2K(K-1)}\sum_{k\neq k'} d_{\mathrm{E}}(g_{u,k}, g_{u,k'})$, where $d_{\mathrm{E}}(a, b) = \|a - b\|_2$ is the Euclidean chordal distance. This objective evaluates distributional agreement in the fixed semantic space by combining distances to held-out preference samples with a dispersion term over generated endpoints. We introduce the coverage term $\widehat{\mathcal{L}}_{\mathrm{cov},u} = \frac{1}{\ell_u}\sum_{j=1}^{\ell_u}\min_{1\le k\le K}\arccos(g_{u,k}^\top z_{u,j})$ to penalize poorly represented targets, and add the penalty $\widehat{\mathcal{L}}_{\mathrm{move},u} = \frac{1}{K}\sum_{k=1}^{K}\|\delta_\psi(r_{u,k}; C_u)\|_2^2$ to regularize source displacement. Finally, we backpropagate gradients directly through the numerical ODE solver to update $\psi$.

### 4.4 Training and Direct Generation

We train the framework in two decoupled stages: we first optimize the condition encoders and velocity field using $\mathcal{L}_{\mathrm{transport}}$, and subsequently freeze the transport to train the source adapter using $\mathcal{L}_{\mathrm{source}}$. The base content encoder and caption decoder remain frozen across both stages.

At inference, we compute the condition $C_u$ once from the full history $H_u$ and retrieved evidence $E_u$, generating personalized content via sequential forward mapping:

$$r \sim P_0 \xrightarrow{A_\psi} \tilde{r} \xrightarrow{F_{\theta,C_u}} z \xrightarrow{D_{\mathrm{dec}}} \widehat{X}, \tag{9}$$

where $D_{\mathrm{dec}}$ denotes deterministic greedy decoding. This operation samples directly from the induced pushforward $Q_\Theta(\cdot \mid H_u, \mathcal{D}) = (D_{\mathrm{dec}} \circ F_{\theta,C_u} \circ A_\psi)_\# P_0$, producing personalized content directly without user-written prompts or post-generation catalog reranking.

## 5 Experimental Results

### 5.1 Experimental Settings

**Datasets.** We conduct experiments on the short-video domain of RecIF (OneRec Team, 2025) and MicroLens-100K (Ni et al., 2025), using video captions as the generation targets. RecIF contains

Table 1: **Quantitative evaluation on RecIF and MicroLens-100K.** We assess the generated content across three dimensions: (1) *Historical replication*, measured by History-nearest (**HN**); (2) *Target and mode-level alignment*, measured by Target-best (**TB**), Coverage (**Cov.**), and Future-mode coverage (**FMC**), which collectively evaluate semantic alignment with held-out targets and coverage of their distinct modes; (3) *Set-level distributional agreement*, measured by Energy score (**ES**) and $\mathbf{MMD^2}$**-U**, quantifying the overall discrepancy with the target preference distribution. ↑ indicates higher is better, ↓ indicates lower is better.

| Method | HN | TB ↑ | Cov. ↓ | FMC ↑ | ES ↓ | $\mathrm{MMD}^2$-U ↓ |
|---|---|---|---|---|---|---|
| *Dataset: RecIF* | | | | | | |
| Task-only (Qwen Team, 2026) | 0.710 | 0.567 | 0.963 | 0.105 | 0.674 | 0.073 |
| Full-history (Qwen Team, 2026) | 0.774 | 0.661 | 0.841 | 0.362 | 0.602 | 0.054 |
| LaMP-IPA (Salemi et al., 2024) | 0.776 | 0.655 | 0.848 | 0.317 | 0.624 | 0.071 |
| PPlug (Liu et al., 2025) | 0.757 | 0.722 | 0.755 | 0.601 | 0.491 | 0.003 |
| CFRAG (Shi et al., 2025) | 0.776 | 0.668 | 0.832 | 0.354 | 0.583 | 0.034 |
| ClusterRAG (Nkhata et al., 2026) | 0.768 | 0.639 | 0.870 | 0.279 | 0.657 | 0.103 |
| **Ours** | 0.758 | **0.755** | **0.706** | **0.691** | **0.478** | **0.001** |
| *Dataset: MicroLens-100K* | | | | | | |
| Task-only (Qwen Team, 2026) | 0.478 | 0.465 | 1.085 | 0.012 | 0.757 | 0.291 |
| Full-history (Qwen Team, 2026) | 0.538 | 0.531 | 1.006 | 0.075 | 0.699 | 0.231 |
| LaMP-IPA (Salemi et al., 2024) | 0.543 | 0.535 | 1.001 | 0.088 | 0.698 | 0.230 |
| PPlug (Liu et al., 2025) | 0.635 | 0.704 | 0.777 | 0.545 | 0.517 | 0.029 |
| CFRAG (Shi et al., 2025) | 0.664 | 0.543 | 0.990 | 0.109 | 0.714 | 0.240 |
| ClusterRAG (Nkhata et al., 2026) | 0.584 | 0.539 | 0.995 | 0.128 | 0.733 | 0.263 |
| **Ours** | 0.618 | **0.729** | **0.746** | **0.632** | **0.500** | **0.011** |

short-video interactions with multiple behavioral signals and dense captions. We treat long views, likes, follows, and forwards as positive interactions. MicroLens-100K contains comment-derived implicit feedback, with video captions obtained from LLMPopcorn (Fu et al., 2026). Detailed dataset statistics are summarized in Appendix F (Table 5). We split users into train, validation, and test sets (8:1:1). Model training and collaborative reference construction use training users only.

**Implementation details.** We conduct experiments on eight NVIDIA A100 GPUs with 80 GB memory each. We use a frozen Chinese-CLIP encoder (Yang et al., 2022) with $\ell_2$-normalized embeddings and a frozen Qwen3.5-9B decoder (Qwen Team, 2026) with continuous-prefix conditioning (Mokady et al., 2021). Frozen XSimGCL (Yu et al., 2024) retrieves at most eight evidence items per user. We use AdamW (Loshchilov & Hutter, 2019) with weight decay 0.01 and gradient clipping at 1.0. We generate 64 captions per user using 32-step spherical Heun integration and greedy decoding, without reranking. Further details are provided in Appendix G.

**Baselines.** We compare our approach with Qwen3.5-based Task-only and Full-history baselines (Qwen Team, 2026), together with task-adapted LaMP-IPA (Salemi et al., 2024), PPlug (Liu et al., 2025), CFRAG (Shi et al., 2025), and ClusterRAG (Nkhata et al., 2026).

**Evaluation.** For each test user $u$, we partition interactions into observed history $H_u$ and held-out positive targets $T_u$, and evaluate the generated caption collection $G_u = (\widehat{X}_{u,k})_{k=1}^{K}$ with $K = 64$ from two complementary perspectives:

- **Quantitative semantic metrics.** In the normalized representation space, we assess historical replication using History-nearest (HN), target and mode-level alignment using Target-best (TB, ↑), Coverage (Cov., ↓), and Future-mode coverage (FMC, ↑), and set-level distributional agreement with held-out targets using Energy score (ES, ↓) and $\mathrm{MMD}^2$-U (↓). Detailed formulations are provided in Appendix H.
- **LLM-based assessment.** Following the LLM-as-a-judge paradigm (Zheng et al., 2023), we conduct pairwise evaluations using a diverse panel of models (GPT-5.6 Sol (OpenAI, 2026), Gemini 3.8 Flash (Google, 2026), and GLM-5.3-Flash (GLM-5-Team et al., 2026)), averaging their judgments to rigorously mitigate single-model evaluation bias. We assess generations along three corresponding dimensions: overall preference alignment (**Q1**), target-level support and genera-

Table 2: **Set-level LLM evaluation on RecIF and MicroLens-100K.** We conduct pairwise evaluations comparing our approach against baselines across three dimensions: (1) *Overall preference* (**Q1**), measured by Win rate (**WR**); (2) *Target-level support* (**Q2**), measured by improvements ($\Delta$) in Best-match (**BM**, graded on a 0-3 scale) and Valid coverage (**VC**), which evaluates accurate target coverage by high-quality generations; (3) *Incremental completion* (**Q3**), measured by improvements ($\Delta$) in Incremental valid coverage (**IVC**), which evaluates the out-of-history preference modes to verify true distribution completion. WR $> 50\%$ and positive $\Delta$ values favor our approach. $\uparrow$ indicates higher is better.

| **Ours** *vs.* Baseline | **Q1: Overall Preference** | **Q2: Target-level Support** | | **Q3: Incremental Completion** |
|---|---|---|---|---|
| | WR $\uparrow$ (%) | $\Delta$ BM $\uparrow$ | $\Delta$ VC $\uparrow$ (pp) | $\Delta$ IVC $\uparrow$ (pp) |
| *Dataset: RecIF* | | | | |
| Task-only | 99.7 | +1.37 | +47.7 | +24.3 |
| Full-history | 71.0 | +0.35 | +15.5 | +12.2 |
| LaMP-IPA | 78.4 | +0.52 | +20.8 | +13.4 |
| PPlug | 63.4 | +0.14 | + 4.1 | + 6.0 |
| CFRAG | 68.3 | +0.47 | +20.6 | +14.5 |
| ClusterRAG | 85.2 | +0.67 | +24.4 | +17.5 |
| *Dataset: MicroLens-100K* | | | | |
| Task-only | 97.9 | +1.26 | +42.5 | +39.2 |
| Full-history | 86.8 | +0.99 | +38.0 | +36.1 |
| LaMP-IPA | 84.3 | +0.92 | +35.2 | +33.1 |
| PPlug | 58.8 | +0.28 | +16.5 | +17.0 |
| CFRAG | 84.6 | +1.04 | +40.1 | +37.3 |
| ClusterRAG | 78.9 | +1.03 | +37.5 | +35.4 |

tion quality (**Q2**), and incremental completion beyond observed history (**Q3**). Further protocol details and metric formulations are provided in Appendices H.1 and H.2.

### 5.2 Main Results

As shown in Table 1, our approach consistently outperforms all baselines across both the RecIF and MicroLens-100K datasets. Compared to the strongest baseline, PPlug, our model demonstrates a significant improvement in Future-mode coverage (FMC). This substantial gain indicates that our approach is highly capable of capturing diverse, exploratory user intents rather than merely collapsing to historical or nearest-neighbor patterns. Moreover, higher Target-best scores and lower Coverage values indicate that held-out targets are more closely matched by at least one generated caption, complementing the broader target-mode coverage reflected by FMC. Furthermore, the significant reductions in Energy and $\mathrm{MMD}^2$-U confirm that this expanded target coverage is not achieved by indiscriminately dispersing noisy samples, but rather by aligning the overall generated distribution much more faithfully with the target preference distribution. Taken together, these results support the effectiveness of our framework for preference distribution completion.

To further validate the distributional alignment and target coverage of the generated captions, Table 2 reports the set-level LLM evaluation results. Our approach consistently emerges as the preferred method across all baseline comparisons. Specifically, for overall preference (Q1), our method achieves a Win rate (WR) well above 50% against all baselines. This confirms that our approach maintains generation quality while providing a more accurate approximation of the target preference distribution. Beyond overall preference, our method demonstrates superior capability in Target-level support (Q2) and Incremental completion (Q3). Across all baselines, we observe positive improvements in Best-match (BM) and Valid coverage (VC). This demonstrates that our method generates higher-quality content and achieves broader coverage of the target preference distribution. Notably, our approach yields substantial gains in Incremental valid coverage (IVC). This robust qualitative superiority corroborates our quantitative findings: the proposed model goes beyond merely retrieving past behaviors, effectively leveraging collaborative patterns to accurately infer unobserved interests and complete the user's target preference distribution.

Figure 3 illustrates this completion capability through a case study of a user with dense visible history. Sample counts reveal differences not only in which modes are covered, but in how samples

Table 3: Ablation study of different components.

| Variant | HN | TB ↑ | Cov. ↓ | FMC ↑ | ES ↓ | $MMD^2$-U ↓ |
|---|---|---|---|---|---|---|
| *Dataset: RecIF* | | | | | | |
| Semantic view RFM | 0.740 | 0.738 | 0.726 | 0.643 | 0.481 | 0.001 |
| Dual-view RFM | 0.745 | 0.745 | 0.722 | 0.670 | 0.479 | 0.001 |
| Dual-view RFM + Calibration | 0.758 | 0.755 | 0.706 | 0.691 | 0.478 | 0.001 |
| Reference PPlug | 0.757 | 0.722 | 0.755 | 0.601 | 0.491 | 0.003 |
| *Dataset: MicroLens-100K* | | | | | | |
| Semantic view RFM | 0.609 | 0.721 | 0.758 | 0.580 | 0.506 | 0.012 |
| Dual-view RFM | 0.617 | 0.725 | 0.748 | 0.620 | 0.505 | 0.012 |
| Dual-view RFM + Calibration | 0.618 | 0.729 | 0.746 | 0.632 | 0.500 | 0.011 |
| Reference PPlug | 0.635 | 0.704 | 0.777 | 0.545 | 0.517 | 0.029 |

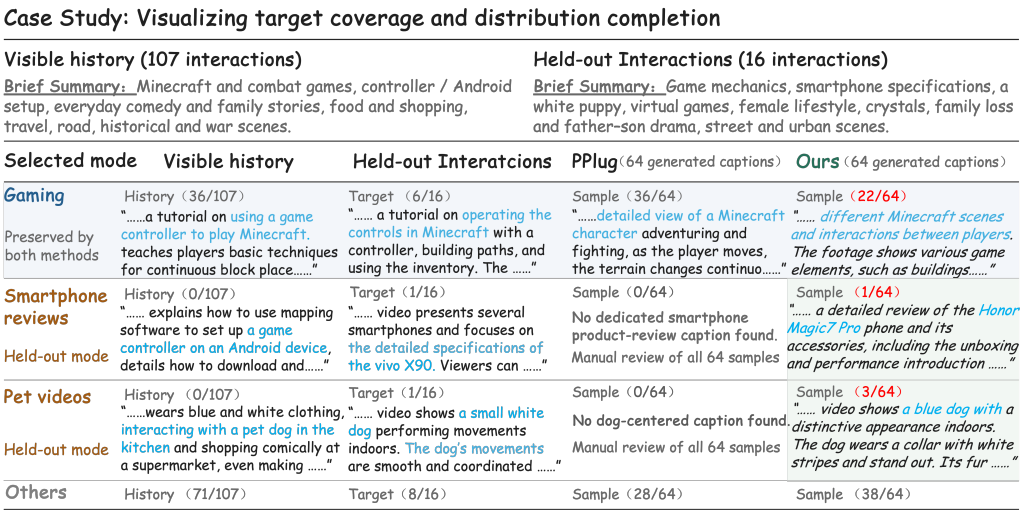


Figure 3: **Case study of distributional completion.**

are distributed across them. Both methods preserve the prominent historical mode, *Gaming*, but PPlug allocates 36/64 samples (56.3%) to it. In contrast, our approach generates 22 gaming captions (34.4%), closely matching proportions in both visible history (36/107, 33.6%) and held-out interactions (6/16, 37.5%). Beyond gaming, our approach produces captions in two target-supported modes, *Smartphone reviews* (1/64) and *Pet videos* (3/64), whereas neither appears among PPlug's 64 samples. Each accounts for 1/16 held-out interactions. These observations confirm two aspects of distributional completion: covering missing interests backed by held-out feedback and more accurately matching empirical distributions across modes, rather than merely increasing topical diversity.

### 5.3 Ablation Study

To evaluate each component's contribution, we conduct an ablation study by progressively adding dual-view conditioning and source calibration. We compare Semantic view RFM, Dual-view RFM, the full model with source calibration, and PPlug as a strong baseline. As shown in Table 3, the consistent gains across both datasets validate the effectiveness of each proposed component.

## 6 Conclusion

We formulated personalized content generation as preference distribution completion and provided a conditional analysis of how cross-user behavioral patterns can support this completion from finite observations. To operationalize this principle, we introduced a collaborative generative framework that integrates dual-view conditioned Riemannian flow matching with source-distribution calibration. This architecture allows the model to infer and directly sample from a continuous semantic preference distribution. Extensive evaluations via quantitative metrics and a multi-model LLM panel demonstrate that our approach consistently outperforms baselines.

### AI USAGE STATEMENT

In the preparation of this manuscript, we utilized Gemini (`https://gemini.google.com`) to help improve the clarity and readability of certain paragraphs in the text. OpenAI Codex (`https://openai.com/index/openai-codex`) was employed to assist with code drafting and implementation for the experimental pipeline. All scientific ideas, theoretical formulations, experimental design, and data analysis were solely conceived and conducted by the authors. All outputs from the large language models were carefully reviewed, manually verified, and edited to ensure technical accuracy and strict consistency with our research findings.

### ETHICS STATEMENT

This work involves learning from user interaction histories to generate personalized content. All experiments are conducted using publicly available, anonymized benchmark datasets (RecIF and MicroLens-100K). We ensure that no personally identifiable information is accessed, processed, or compromised during the training and evaluation phases. We recognize risks such as algorithmic bias, echo chambers, and inherited biases from base models (e.g., Chinese-CLIP, Qwen3.5). Real-world deployments should implement safety guardrails and fairness auditing.

### REPRODUCIBILITY STATEMENT

We provide documentation across the appendices and supplementary materials for the reproduction of our empirical and theoretical results. Complete mathematical proofs for the local collaborative estimation bounds and generation consistency are detailed in Appendices C and D. The dataset preprocessing pipelines, implementation details, and LLM-as-a-judge evaluation prompts are documented in Appendices F, G, and H. The code is anonymously available at `https://anonymous.4open.science/r/PCGaPDC-485E/`.

## REFERENCES

Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. The connection between popularity bias, calibration, and fairness in recommendation. In *Fourteenth ACM Conference on Recommender Systems*, 2020. doi: 10.1145/3383313.3418487.

Yukuo Cen, Jianwei Zhang, Xu Zou, Chang Zhou, Hongxia Yang, and Jie Tang. Controllable multi-interest framework for recommendation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, 2020.

Ricky T. Q. Chen and Yaron Lipman. Flow matching on general geometries. In *International Conference on Learning Representations*, 2024.

Meihua Dang, Anikait Singh, Linqi Zhou, Stefano Ermon, and Jiaming Song. Personalized preference fine-tuning of diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8020–8030, 2025.

Connor Dunlop, Matthew Zheng, Kavana Venkatesh, and Pinar Yanardag. Personalized image editing in text-to-image diffusion models via collaborative direct preference optimization. *arXiv preprint arXiv:2511.05616*, 2025.

Ziwei Fan, Zhiwei Liu, Alice Wang, Zahra Nazari, Lei Zheng, Hao Peng, and Philip S. Yu. Sequential recommendation via stochastic self-attention. In *Proceedings of the ACM Web Conference 2022*, 2022. doi: 10.1145/3485447.3512077.

Diego Feijer, Himan Abdollahpouri, Sanket Gupta, Alexander Clare, Yuxiao Wen, Todd Wasson, Maria Dimakopoulou, Zahra Nazari, Kyle Kretschman, and Mounia Lalmas. Calibrated recommendations with contextual bandits, 2025. Presented at the CONSEQUENCES Workshop at ACM RecSys 2025.

Junchen Fu, Xuri Ge, Kaiwen Zheng, Alexandros Karatzoglou, Ioannis Arapakis, Xin Xin, Yongxin Ni, and Joemon M. Jose. LLMPopcorn: Exploring LLMs as assistants for popular micro-video

generation. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 12007–12011. IEEE, 2026. doi: 10.1109/ICASSP55912.2026.11463708.

GLM-5-Team, :, Aohan Zeng, Xin Lv, Zhenyu Hou, Zhengxiao Du, Qinkai Zheng, Bin Chen, Da Yin, Chendi Ge, Chenghua Huang, Chengxing Xie, Chenzheng Zhu, Congfeng Yin, Cunxiang Wang, Gengzheng Pan, Hao Zeng, Haoke Zhang, Haoran Wang, Huilong Chen, Jiajie Zhang, Jian Jiao, Jiaqi Guo, Jingsen Wang, Jingzhao Du, Jinzhu Wu, Kedong Wang, Lei Li, Lin Fan, Lucen Zhong, Mingdao Liu, Mingming Zhao, Pengfan Du, Qian Dong, Rui Lu, Shuang-Li, Shulin Cao, Song Liu, Ting Jiang, Xiaodong Chen, Xiaohan Zhang, Xuancheng Huang, Xuezhen Dong, Yabo Xu, Yao Wei, Yifan An, Yilin Niu, Yitong Zhu, Yuanhao Wen, Yukuo Cen, Yushi Bai, Zhongpei Qiao, Zihan Wang, Zikang Wang, Zilin Zhu, Ziqiang Liu, Zixuan Li, Bojie Wang, Bosi Wen, Can Huang, Changpeng Cai, Chao Yu, Chen Li, Chengwei Hu, Chenhui Zhang, Dan Zhang, Daoyan Lin, Dayong Yang, Di Wang, Ding Ai, Erle Zhu, Fangzhou Yi, Feiyu Chen, Guohong Wen, Hailong Sun, Haisha Zhao, Haiyi Hu, Hanchen Zhang, Hanrui Liu, Hanyu Zhang, Hao Peng, Hao Tai, Haobo Zhang, He Liu, Hongwei Wang, Hongxi Yan, Hongyu Ge, Huan Liu, Huanpeng Chu, Jia'ni Zhao, Jiachen Wang, Jiajing Zhao, Jiamin Ren, Jiapeng Wang, Jiaxin Zhang, Jiayi Gui, Jiayue Zhao, Jijie Li, Jing An, Jing Li, Jingwei Yuan, Jinhua Du, Jinxin Liu, Junkai Zhi, Junwen Duan, Kaiyue Zhou, Kangjian Wei, Ke Wang, Keyun Luo, Laiqiang Zhang, Leigang Sha, Liang Xu, Lindong Wu, Lintao Ding, Lu Chen, Minghao Li, Nianyi Lin, Pan Ta, Qiang Zou, Rongjun Song, Ruiqi Yang, Shangqing Tu, Shangtong Yang, Shaoxiang Wu, Shengyan Zhang, Shijie Li, Shuang Li, Shuyi Fan, Wei Qin, Wei Tian, Weining Zhang, Wenbo Yu, Wenjie Liang, Xiang Kuang, Xiangmeng Cheng, Xiangyang Li, Xiaoquan Yan, Xiaowei Hu, Xiaoying Ling, Xing Fan, Xingye Xia, Xinyuan Zhang, Xinze Zhang, Xirui Pan, Xu Zou, Xunkai Zhang, Yadi Liu, Yandong Wu, Yanfu Li, Yidong Wang, Yifan Zhu, Yijun Tan, Yilin Zhou, Yiming Pan, Ying Zhang, Yinpei Su, Yipeng Geng, Yong Yan, Yonglin Tan, Yuean Bi, Yuhan Shen, Yuhao Yang, Yujiang Li, Yunan Liu, Yunqing Wang, Yuntao Li, Yurong Wu, Yutao Zhang, Yuxi Duan, Yuxuan Zhang, Zezhen Liu, Zhengtao Jiang, Zhenhe Yan, Zheyu Zhang, Zhixiang Wei, Zhuo Chen, Zhuoer Feng, Zijun Yao, Ziwei Chai, Ziyuan Wang, Zuzhou Zhang, Bin Xu, Minlie Huang, Hongning Wang, Juanzi Li, Yuxiao Dong, and Jie Tang. Glm-5: from vibe coding to agentic engineering, 2026.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437.

Google. Gemini 3.8 Flash. Google AI for Developers, 2026. Accessed: 2026-09-25.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830.

Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009. doi: 10.1109/MC.2009.263.

Chao Li, Zhiyuan Liu, Mengmeng Wu, Yuchi Xu, Pipei Huang, Huan Zhao, Guoliang Kang, Qiwei Chen, Wei Li, and Dik Lun Lee. Multi-interest network with dynamic routing for recommendation at Tmall. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019. doi: 10.1145/3357384.3357814.

Cheng Li, Mingyang Zhang, Qiaozhu Mei, Weize Kong, and Michael Bendersky. Learning to rewrite prompts for personalized text generation. In *Proceedings of the ACM Web Conference 2024*. Association for Computing Machinery, 2024. doi: 10.1145/3589334.3645408.

Dawen Liang, Laurent Charlin, James McInerney, and David M. Blei. Modeling user exposure in recommendation. In *Proceedings of the 25th International Conference on World Wide Web*, 2016. doi: 10.1145/2872427.2883090.

Dawen Liang, Rahul G. Krishnan, Matthew D. Hoffman, and Tony Jebara. Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 World Wide Web Conference*, 2018. doi: 10.1145/3178876.3186150.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

Jiongnan Liu, Yutao Zhu, Shuting Wang, Xiaochi Wei, Erxue Min, Yu Lu, Shuaiqiang Wang, Dawei Yin, and Zhicheng Dou. LLMs + Persona-Plug = personalized LLMs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 9373–9385. Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.acl-long.461.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

Igor Melnyk, Youssef Mroueh, Brian Belgodere, Mattia Rigotti, Apoorva Nitsure, Mikhail Yurochkin, Kristjan Greenewald, Jiri Navratil, and Jarret Ross. Distributional preference alignment of LLMs via optimal transport. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-3317.

Ron Mokady, Amir Hertz, and Amit H. Bermano. ClipCap: CLIP prefix for image captioning. *arXiv preprint arXiv:2111.09734*, 2021.

Bo Ni, Branislav Kveton, Samyadeep Basu, Subhojyoti Mukherjee, Leyao Wang, Franck Dernoncourt, Sungchul Kim, Seunghyun Yoon, Zichao Wang, Ruiyi Zhang, Puneet Mathur, Jihyung Kil, Jiuxiang Gu, Nedim Lipka, Yu Wang, Ryan A. Rossi, and Tyler Derr. Reasoning-based personalized generation for users with sparse data. *arXiv preprint arXiv:2602.21219*, 2026. doi: 10.48550/arXiv.2602.21219.

Yongxin Ni, Yu Cheng, Xiangyan Liu, Junchen Fu, Youhua Li, Xiangnan He, Yongfeng Zhang, and Fajie Yuan. A content-driven micro-video recommendation dataset at scale. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, pp. 6486–6491. Association for Computing Machinery, 2025. doi: 10.1145/3746252.3761655.

Gibson Nkhata, Uttamasha Anjally Oyshi, Quan Mai, and Susan Gauch. ClusterRAG: Cluster-based collaborative filtering for personalized retrieval-augmented generation. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 20523–20539. Association for Computational Linguistics, 2026. doi: 10.18653/v1/2026.acl-long.940.

OneRec Team. OpenOneRec technical report, 2025.

OpenAI. GPT-5.6 System Card, July 2026. Accessed: 2026-09-25.

Gabriel A. Patron, Zhiwei Xu, Ishan Kapnadak, and Felipe Maia Polo. Personalized image generation for recommendations beyond catalogs. *arXiv preprint arXiv:2502.18477*, 2025. doi: 10.48550/arXiv.2502.18477.

Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lipman, and Ricky T. Q. Chen. Multisample flow matching: Straightening flows with minibatch couplings. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 28100–28127. PMLR, 2023.

Qwen Team. Qwen3.5: Towards native multimodal agents, February 2026.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8748–8763. PMLR, 2021.

Shashank Rajput, Nikhil Mehta, Anima Singh, Raghunandan Hulikal Keshavan, Trung Vu, Lukasz Heldt, Lichan Hong, Yi Tay, Vinh Tran, Jonah Samost, Maciej Kula, Ed Chi, and Maheswaran Sathiamoorthy. Recommender systems with generative retrieval. In *Advances in Neural Information Processing Systems*, volume 36, pp. 10299–10315, 2023.

Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. GroupLens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, pp. 175–186. ACM, 1994. doi: 10.1145/192844.192905.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

Alireza Salemi, Sheshera Mysore, Michael Bendersky, and Hamed Zamani. LaMP: When large language models meet personalization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 7370–7392. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.399.

Xiaoteng Shen, Rui Zhang, Xiaoyan Zhao, Jieming Zhu, and Xi Xiao. PMG: Personalized multimodal generation with large language models. In *Proceedings of the ACM Web Conference 2024*, pp. 3833–3843. Association for Computing Machinery, 2024. doi: 10.1145/3589334.3645633.

Teng Shi, Jun Xu, Xiao Zhang, Xiaoxue Zang, Kai Zheng, Yang Song, and Han Li. Retrieval augmented generation with collaborative filtering for personalized text generation. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2025. doi: 10.1145/3726302.3730075.

Anand Siththaranjan, Cassidy Laidlaw, and Dylan Hadfield-Menell. Distributional preference learning: Understanding and accounting for hidden context in RLHF. In *International Conference on Learning Representations*, 2024.

Taylor Sorensen, Jared Moore, Jillian Fisher, Mitchell L. Gordon, Niloofar Mireshghallah, Christopher Michael Rytting, Andre Ye, Liwei Jiang, Ximing Lu, Nouha Dziri, Tim Althoff, and Yejin Choi. Position: A roadmap to pluralistic alignment. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 46280–46302. PMLR, 2024.

Harald Steck. Calibrated recommendations. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 154–162, 2018. doi: 10.1145/3240323.3240372.

Aaron Van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Bram Wallace, Meihua Dang, Rafael Rafailov, Linqi Zhou, Aaron Lou, Senthil Purushwalkam, Stefano Ermon, Caiming Xiong, Shafiq Joty, and Nikhil Naik. Diffusion model alignment using direct preference optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.

Ye Wang, Jiahao Xun, Minjie Hong, Jieming Zhu, Tao Jin, Wang Lin, Haoyuan Li, Linjun Li, Yan Xia, Zhou Zhao, and Zhenhua Dong. EAGER: Two-stream generative recommender with behavior-semantic collaboration. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3245–3254, 2024. doi: 10.1145/3637528.3671775.

Mark Wilhelm, Ajith Ramanathan, Alexander Bonomo, Sagar Jain, Ed H. Chi, and Jennifer Gillenwater. Practical diversified recommendations on YouTube with determinantal point processes. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 2018. doi: 10.1145/3269206.3272018.

Yueqi Xie, Jingqi Gao, Peilin Zhou, Qichen Ye, Yining Hua, Jae Boum Kim, Fangzhao Wu, and Sunghun Kim. Rethinking multi-interest learning for candidate matching in recommender systems. In *Proceedings of the 17th ACM Conference on Recommender Systems*, 2023. doi: 10.1145/3604915.3608766.

Yiyan Xu, Wenjie Wang, Yang Zhang, Biao Tang, Peng Yan, Fuli Feng, and Xiangnan He. Personalized image generation with large multimodal models. In *Proceedings of the ACM Web Conference 2025*. Association for Computing Machinery, 2025a. doi: 10.1145/3696410.3714843.

Yiyan Xu, Wuqiang Zheng, Wenjie Wang, Fengbin Zhu, Xinting Hu, Yang Zhang, Fuli Feng, and Tat-Seng Chua. DRC: Enhancing personalized image generation via disentangled representation composition. In *Proceedings of the 33rd ACM International Conference on Multimedia*. Association for Computing Machinery, 2025b. doi: 10.1145/3746027.3755051.

An Yang, Junshu Pan, Junyang Lin, Rui Men, Yichang Zhang, Jingren Zhou, and Chang Zhou. Chinese CLIP: Contrastive vision-language pretraining in chinese. *arXiv preprint arXiv:2211.01335*, 2022.

Zhengyi Yang, Jiancan Wu, Zhicai Wang, Xiang Wang, Yancheng Yuan, and Xiangnan He. Generate what you prefer: Reshaping sequential recommendation via guided diffusion. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Junliang Yu, Xin Xia, Tong Chen, Lizhen Cui, Nguyen Quoc Viet Hung, and Hongzhi Yin. XSimGCL: Towards extremely simple graph contrastive learning for recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 36(2):913–926, 2024. doi: 10.1109/TKDE.2023.3288135.

Jinghao Zhang, Yuting Liu, Wenjie Wang, Qiang Liu, Shu Wu, Liang Wang, and Tat-Seng Chua. Personalized text generation with contrastive activation steering. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 7128–7141. Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.acl-long.353.

Yangtian Zhang, Leyao Wang, Hiren Madhu, Ngoc Bui, Walter Roznyatovskiy, and Rex Ying. CoPersona: Collaborative persona graphs for robust LLM personalization. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2026.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena, 2023.

## A Ranking Does Not Determine a Generative Distribution

To formalize why an item ranking alone cannot determine a personalized generative model, we formulate the relationship between continuous preference distributions and rankings as a mapping and show that this mapping is non-injective (many-to-one).

Let $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ be a finite content space with $n \geq 2$. Let $\Delta(\mathcal{X})$ denote the standard probability simplex of distributions over $\mathcal{X}$:

$$\Delta(\mathcal{X}) = \left\{ P \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} P(x_i) = 1,\ P(x_i) \geq 0,\ \forall i \right\}. \tag{10}$$

A distribution $P \in \Delta(\mathcal{X})$ induces a strict ranking $R$ (a permutation of $\mathcal{X}$) if its probability masses satisfy the chain of inequalities:

$$P(x_{i_1}) > P(x_{i_2}) > \cdots > P(x_{i_n}), \tag{11}$$

where $(i_1, i_2, \ldots, i_n)$ denotes an ordered permutation of $\{1, 2, \ldots, n\}$.

Let $\mathcal{R}$ be the set of all $n!$ strict rankings over $\mathcal{X}$, and let $f : \Delta(\mathcal{X}) \to \mathcal{R}$ denote the mapping that associates each non-degenerate distribution with its induced ranking. For any specific ranking $R$, its pre-image $f^{-1}(R)$ is given by the intersection of the simplex $\Delta(\mathcal{X})$ with the $n-1$ open half-spaces defined by $P(x_{i_k}) > P(x_{i_{k+1}})$ for $k = 1, \ldots, n-1$.

Because these $n-1$ strict linear inequalities do not reduce the affine dimension of the domain, the pre-image $f^{-1}(R)$ is a non-empty open convex polytope with non-zero $(n-1)$-dimensional

Lebesgue measure on the simplex. Consequently, for any valid ranking $R$, there exist uncountably infinitely many distributions $P \in \Delta(\mathcal{X})$ that satisfy $f(P) = R$. The mapping $f$ is therefore strictly non-injective, formally establishing that a ranking alone cannot specify a unique generative distribution.

To illustrate this geometrically, consider the minimal case of two items $\mathcal{X} = \{a, b\}$ ($n = 2$). The preference ordering $a \succ b$ defines the open half-line $P(a) > P(b)$. Under the simplex constraint $P(a) + P(b) = 1$, any distribution with $P(a) \in (0.5, 1]$ maps to $a \succ b$. For example, consider two candidate distributions:

$$\begin{aligned} P(a) &= 0.51, \quad P(b) = 0.49, \\ Q(a) &= 0.99, \quad Q(b) = 0.01. \end{aligned} \tag{12}$$

Both distributions induce the identical ranking $a \succ b$, yet they assign drastically different probability masses to the items. Thus, a ranking specifies only a broad continuous region of feasible distributions. Directly synthesizing content that respects a user's multi-interest profile fundamentally requires learning the underlying probability distribution, rather than simply optimizing a top-$k$ ordering.

## B Distributional Underdetermination from Finite Histories

We formalize why a finite interaction history cannot uniquely identify a user's underlying preference distribution without structural priors. Consider a measurable content space containing two distinct points with measurable singletons, denoted by $\{0, 1\}$. Let $1 \leq n_u < \infty$ and consider two candidate preference distributions:

$$P^{(0)} = \text{Bernoulli}(1/3), \qquad P^{(1)} = \text{Bernoulli}(2/3). \tag{13}$$

Suppose the collective dataset $\mathcal{D}$ is statistically independent of the target history $H_u$ and follows a common marginal law $\mathsf{R}$ under both scenarios. The joint observation measures are therefore

$$\mathsf{P}_i = (P^{(i)})^{\otimes n_u} \otimes \mathsf{R}, \qquad i \in \{0, 1\}. \tag{14}$$

**Lemma 1.** *No measurable estimator $T(H_u, \mathcal{D})$ can achieve exact recovery almost surely under both worlds; that is, we cannot have $T(H_u, \mathcal{D}) = P^{(i)}$ $\mathsf{P}_i$-almost surely for both $i \in \{0, 1\}$.*

*Proof.* Every binary interaction sequence of length $n_u$ has strictly positive probability under both product measures $(P^{(0)})^{\otimes n_u}$ and $(P^{(1)})^{\otimes n_u}$. Consequently, these two product measures are mutually absolutely continuous, which immediately implies mutual absolute continuity for their joint product measures $\mathsf{P}_0$ and $\mathsf{P}_1$ with the common measure $\mathsf{R}$.

Now, suppose for contradiction that an estimator $T$ satisfies exact recovery in both worlds. Define the event sets $A_i = \{(H_u, \mathcal{D}) : T(H_u, \mathcal{D}) = P^{(i)}\}$ for $i \in \{0, 1\}$. By assumption, $\mathsf{P}_0(A_0) = 1$ and $\mathsf{P}_1(A_1) = 1$. By mutual absolute continuity, $\mathsf{P}_1(A_1) = 1$ requires that $\mathsf{P}_0(A_1) = 1$. Thus,

$$\mathsf{P}_0(A_0 \cap A_1) = \mathsf{P}_0(A_0) + \mathsf{P}_0(A_1) - \mathsf{P}_0(A_0 \cup A_1) = 1 + 1 - 1 = 1. \tag{15}$$

However, since $P^{(0)} \neq P^{(1)}$, the events are mutually disjoint ($A_0 \cap A_1 = \varnothing$), yielding $\mathsf{P}_0(\varnothing) = 1$, which is a contradiction. □

Concretely, an interaction history consisting solely of zeros has strictly positive probability under both distributions, even though the unobserved content point $\{1\}$ carries probability mass $1/3$ under $P^{(0)}$ and $2/3$ under $P^{(1)}$. More generally, an unobserved interest region with true probability mass $p > 0$ will be completely absent from an i.i.d. history of length $n_u$ with probability $(1 - p)^{n_u}$.

We emphasize that this counterexample reflects an intrinsic finite-sample limitation rather than an asymptotic non-identifiability claim, as the two sampling distributions are statistically distinct as $n_u \to \infty$. It also does not suggest that cross-user interactions cannot be informative. Rather, it highlights that finite individual interactions combined with uninformative group data are mathematically insufficient for exact distribution recovery, motivating the need for collaborative structural constraints.

# C Analysis of Local Collaborative Estimation

## C.1 Theoretical Setup and Error Bound

To specify the distributional properties considered in the analysis, we fix a nonempty finite collection $\mathcal{F} = \{f_1, \dots, f_J\}$ of bounded measurable test functions $f_j : \mathcal{X} \to [-1, 1]$, chosen prior to observing the data, and define the induced discrepancy:

$$d_{\mathcal{F}}(P, Q) = \max_{1 \le j \le J} \left| \int f_j \, \mathrm{d}P - \int f_j \, \mathrm{d}Q \right|. \tag{16}$$

Let $\mathcal{V}_M$ denote $M \ge 1$ reference users in $\mathcal{D}$. We assume users share an underlying family of preference distributions $\{P_s : s \in \mathcal{S}\}$ parameterized by a collaborative state space $\mathcal{S} \subseteq \mathbb{R}^p$. We assume that the preference distributions satisfy a Hölder smoothness condition with respect to the collaborative state: for fixed constants $0 \le L < \infty$ and $0 < \alpha \le 1$,

$$d_{\mathcal{F}}(P_s, P_{s'}) \le L \|s - s'\|^{\alpha}, \qquad \forall s, s' \in \mathcal{S}. \tag{17}$$

Let $\widehat{s}_u, \widehat{s}_v \in \mathcal{S}$ be state estimators, and let collaborative weights satisfy $w_{uv,M} \ge 0$ and $\sum_{v \in \mathcal{V}_M} w_{uv,M} = 1$. Given deterministic bounds $r_{u,n_u}$, $r_{\mathcal{D},M}$ and bandwidth $h_M$, we assume that with probability at least $1 - \eta_{u,M}$, the localization event $\mathcal{E}_{u,M}$ occurs such that:

$$\|\widehat{s}_u - s_u\| \le r_{u,n_u}, \quad \max_{v \in \mathcal{V}_M} \|\widehat{s}_v - s_v\| \le r_{\mathcal{D},M}, \quad \text{and} \quad w_{uv,M} > 0 \implies \|\widehat{s}_v - \widehat{s}_u\| \le h_M. \tag{18}$$

We construct the aggregated collaborative estimator $\widehat{P}^{\mathrm{coll}}_{u,M} = \sum_{v \in \mathcal{V}_M} w_{uv,M} \widehat{P}_v$, with effective reference sample size $m_{\mathrm{eff},u,M} = (\sum_{v \in \mathcal{V}_M} w^2_{uv,M}/m_v)^{-1}$.

**Proposition 1** (Local collaborative estimation)**.** *Under the preceding shared-structure, smoothness, localization, and sampling conditions, for any $\delta \in (0, 1)$, with probability at least $1 - \eta_{u,M} - \delta$,*

$$d_{\mathcal{F}}\Big(\widehat{P}^{\mathrm{coll}}_{u,M}, P^*_u\Big) \le L\big(h_M + r_{\mathcal{D},M} + r_{u,n_u}\big)^{\alpha} + \sqrt{\frac{2\log(2J/\delta)}{m_{\mathrm{eff},u,M}}}. \tag{19}$$

## C.2 Conditional Sampling Requirement

Let $\mathbf{s}_M = (s_u, (s_v)_{v \in \mathcal{V}_M})$ denote the collection of true collaborative states, and define the $\sigma$-algebra

$$\mathcal{G}_{u,M} = \sigma(H_u, \mathcal{D}^{\mathrm{match}}, \mathbf{s}_M).$$

As established in the main text, these states represent fixed but unknown quantities. By construction, all state estimates and collaborative weights are measurable with respect to $\sigma(H_u, \mathcal{D}^{\mathrm{match}})$. The required conditional sampling distribution for our theoretical analysis is

$$\mathcal{L}((Y_{v,1:m_v})_{v \in \mathcal{V}_M} \mid \mathcal{G}_{u,M}) = \bigotimes_{v \in \mathcal{V}_M} P_{s_v}^{\otimes m_v}, \tag{20}$$

where $Y_{v,1:m_v} = (Y_{v,1}, \dots, Y_{v,m_v})$ denotes the sequence of estimation samples for peer user $v$.

One sufficient procedure to satisfy this condition relies on rigorous sample splitting. Assuming independent interaction sequences across users, each drawn i.i.d. from its respective user's preference distribution, we partition the reference sequences using fixed sizes determined by a rule strictly independent of the empirical realizations. To prevent data leakage, any learned representation, interaction graph, or matching mechanism used to compute the collaborative weights must strictly exclude these designated estimation samples. Merely partitioning the dataset indices is insufficient to establish Equation (20) if the matching procedure implicitly absorbs information from both subsets. We emphasize that this strict conditional independence is a rigorous requirement for the theoretical estimator, rather than a descriptive property of arbitrary practical implementations.

We define the localization event $\mathcal{E}_{u,M}$ as the joint occurrence of the conditions in Equation (18), ensuring it is strictly $\mathcal{G}_{u,M}$-measurable. By definition, the non-negative collaborative weights are properly normalized for every realization. In cases where no admissible local neighborhood can be identified, the estimator reverts to a fixed, normalized fallback distribution; such realizations naturally fall outside the typical set $\mathcal{E}_{u,M}$ and are formally bounded in probability by the failure rate $\eta_{u,M}$.

### C.3 Proof of Proposition 1

*Proof.* For a probability distribution $P$ and a bounded measurable function $f$, we denote the expectation by $Pf = \int f \, \mathrm{d}P$. Define the weighted population distribution as

$$\overline{P}_{u,M} = \sum_{v \in \mathcal{V}_M} w_{uv,M} P_{s_v}.$$

By the triangle inequality, we decompose the total error as

$$\begin{aligned} d_{\mathcal{F}}(\widehat{P}^{\mathrm{coll}}_{u,M}, P^*_u) \le d_{\mathcal{F}}(\widehat{P}^{\mathrm{coll}}_{u,M}, \overline{P}_{u,M}) \\ + d_{\mathcal{F}}(\overline{P}_{u,M}, P^*_u). \end{aligned} \tag{21}$$

Conditioned on the localization event $\mathcal{E}_{u,M}$, any reference user $v$ with a strictly positive weight satisfies

$$\begin{aligned} \|s_v - s_u\| &\le \|s_v - \widehat{s}_v\| + \|\widehat{s}_v - \widehat{s}_u\| + \|\widehat{s}_u - s_u\| \\ &\le r_{\mathcal{D},M} + h_M + r_{u,n_u}. \end{aligned} \tag{22}$$

Leveraging the smoothness condition alongside the non-negativity and normalization of the weights, we bound the transfer bias as

$$\begin{aligned} d_{\mathcal{F}}(\overline{P}_{u,M}, P^*_u) &\le \sum_{v \in \mathcal{V}_M} w_{uv,M} d_{\mathcal{F}}(P_{s_v}, P_{s_u}) \\ &\le L\big(h_M + r_{\mathcal{D},M} + r_{u,n_u}\big)^{\alpha}. \end{aligned} \tag{23}$$

Next, we condition on the $\sigma$-algebra $\mathcal{G}_{u,M}$. Under this conditioning, the collaborative weights are deterministic, and Equation (20) guarantees that the estimation samples are independent, with each drawn from its respective true preference distribution. For any fixed test function $f \in \mathcal{F}$, the sampling error expands as

$$\begin{aligned} &\widehat{P}^{\mathrm{coll}}_{u,M} f - \overline{P}_{u,M} f \\ &\quad = \sum_{v \in \mathcal{V}_M} \sum_{i=1}^{m_v} \frac{w_{uv,M}}{m_v} \big(f(Y_{v,i}) - P_{s_v} f\big). \end{aligned} \tag{24}$$

Each summand is conditionally centered and almost surely bounded within an interval of length $2w_{uv,M}/m_v$. Consequently, the sum of the squared interval lengths is at most

$$4 \sum_{v \in \mathcal{V}_M} \frac{w^2_{uv,M}}{m_v} = \frac{4}{m_{\mathrm{eff},u,M}}.$$

Applying Hoeffding's inequality (Hoeffding, 1963) followed by a union bound over the $J$ fixed functions in $\mathcal{F}$, we obtain that for any $t > 0$,

$$\begin{aligned} &\Pr\Big(d_{\mathcal{F}}(\widehat{P}^{\mathrm{coll}}_{u,M}, \overline{P}_{u,M}) > t \,\Big|\, \mathcal{G}_{u,M}\Big) \\ &\quad \le 2J \exp\left(-\frac{t^2 m_{\mathrm{eff},u,M}}{2}\right). \end{aligned} \tag{25}$$

Since the effective sample size $m_{\mathrm{eff},u,M}$ is $\mathcal{G}_{u,M}$-measurable, we are permitted to plug in the data-dependent conditional threshold

$$t = \sqrt{\frac{2\log(2J/\delta)}{m_{\mathrm{eff},u,M}}}.$$

This guarantees that the conditional failure probability is bounded by $\delta$. Taking the expectation over $\mathcal{G}_{u,M}$ unconditionally preserves this bound. By a union bound, intersecting this concentration event with the localization event $\mathcal{E}_{u,M}$ yields a joint success probability of at least $1 - \delta - \eta_{u,M}$. Combining the sampling error concentration with the transfer bias bound in Equation (23) proves Equation (19). □

### C.4 Consistency of the Local Estimator

Let the upper bound on the transfer bias be denoted by

$$b_{u,M} = L\big(h_M + r_{\mathcal{D},M} + r_{u,n_u}\big)^{\alpha}.$$

Under the asymptotic regime where $\eta_{u,M} \to 0$, $b_{u,M} \to 0$, and $m_{\mathrm{eff},u,M} \xrightarrow{\mathrm{p}} \infty$ as $M \to \infty$ (and $n_u \to \infty$), fix an arbitrary $\epsilon > 0$. For all sufficiently large $M$, it is guaranteed that $b_{u,M} \le \epsilon/2$. Restricted to the localization event $\mathcal{E}_{u,M}$, incurring a total estimation error greater than $\epsilon$ strictly implies that the sampling error must exceed $\epsilon/2$. For any fixed scalar threshold $R > 0$, decomposing the probability space based on the effective sample size yields the unconditional tail bound

$$\begin{aligned}&\Pr\Big(d_{\mathcal{F}}(\widehat{P}^{\mathrm{coll}}_{u,M}, P^*_u) > \epsilon\Big)\\ &\quad\le \eta_{u,M} + \Pr(m_{\mathrm{eff},u,M} < R) + 2J\exp(-\epsilon^2 R/8).\end{aligned} \tag{26}$$

We evaluate this limit sequentially. First, taking the limit as $M \to \infty$, the first two probability terms vanish by our regularity assumptions. Subsequently, taking the limit as $R \to \infty$ drives the remaining exponential term to zero. This establishes the asymptotic consistency $d_{\mathcal{F}}(\widehat{P}^{\mathrm{coll}}_{u,M}, P^*_u) \xrightarrow{\mathrm{p}} 0$ without necessitating the restrictive assumption that the effective sample size is deterministic.

## D Sampling from the Learned Distribution

The sampling operation in Equation (9) introduces a finite-sample error distinct from the estimation error of the learned preference distribution. Conditioned on the fitted parameters $\Theta$ and the available data $(H_u, \mathcal{D})$, we draw $K$ independent samples

$$\widehat{X}_1, \dots, \widehat{X}_K \overset{\mathrm{iid}}{\sim} Q_\Theta(\cdot \mid H_u, \mathcal{D}).$$

Let the empirical generation distribution be defined as

$$\widehat{Q}^{(K)}_{\Theta,u} = \frac{1}{K}\sum_{k=1}^{K} \delta_{\widehat{X}_k}. \tag{27}$$

Applying Hoeffding's inequality and a union bound over the fixed finite query class $\mathcal{F}$ yields, for any $t > 0$,

$$\begin{aligned}&\Pr\Big(d_{\mathcal{F}}\Big(\widehat{Q}^{(K)}_{\Theta,u}, Q_\Theta(\cdot \mid H_u, \mathcal{D})\Big) > t \,\Big|\, \Theta, H_u, \mathcal{D}\Big)\\ &\quad\le 2J\exp(-Kt^2/2).\end{aligned} \tag{28}$$

This tail bound holds because, conditioned on the fitted model and data, each evaluation $f(\widehat{X}_k)$ is strictly bounded in $[-1, 1]$ and has conditional expectation $\int f \,\mathrm{d}Q_\Theta(\cdot \mid H_u, \mathcal{D})$. Thus, the standard bounded-variable concentration argument from Appendix C.3 applies directly (Hoeffding, 1963).

Consequently, for any $\delta \in (0, 1)$, it holds with conditional probability at least $1 - \delta$ that

$$\begin{aligned}d_{\mathcal{F}}\Big(\widehat{Q}^{(K)}_{\Theta,u}, P^*_u\Big) \le\ & d_{\mathcal{F}}(Q_\Theta(\cdot \mid H_u, \mathcal{D}), P^*_u)\\ &+ \sqrt{\frac{2\log(2J/\delta)}{K}}.\end{aligned} \tag{29}$$

The first term captures the estimation error of the learned distribution, whereas the second term represents the empirical sampling error. Increasing the sample size $K$ controls only the latter.

For a sequence of learned models and datasets, achieving asymptotic consistency with respect to the true preference distribution additionally requires the estimation error to vanish in probability:

$$d_{\mathcal{F}}(Q_\Theta(\cdot \mid H_u, \mathcal{D}), P^*_u) \xrightarrow{\mathrm{p}} 0.$$

Provided this estimation consistency holds and $K \to \infty$, we have for any $\epsilon > 0$:

$$\begin{aligned}&\Pr\Big(d_{\mathcal{F}}(\widehat{Q}^{(K)}_{\Theta,u}, P^*_u) > \epsilon\Big)\\ &\quad\le \Pr(d_{\mathcal{F}}(Q_\Theta(\cdot \mid H_u, \mathcal{D}), P^*_u) > \epsilon/2)\\ &\qquad + 2J\exp(-K\epsilon^2/8) \longrightarrow 0.\end{aligned} \tag{30}$$

We note that Proposition 1 does not formally establish this required estimation consistency, as it analyzes a different theoretical estimator. Furthermore, all convergence statements in this section strictly apply to the finite query class $\mathcal{F}$ and do not imply weak convergence of the full content distribution.

# E Instantiation and Validation of the Semantic Encoder–Decoder

We instantiate the semantic encoder $e(\cdot)$ with the pretrained text encoder of Chinese-CLIP (Yang et al., 2022) and the decoding kernel $K_{\mathrm{dec}}$ with an embedding-conditioned caption decoder based on Qwen3.5-9B (Qwen Team, 2026). Specifically, given a caption $x \in \mathcal{X}$ with projected text feature $f_{\mathrm{text}}(x) \in \mathbb{R}^d$, $\ell_2$-normalization places the representation on the unit hypersphere $\mathcal{Z} = \mathbb{S}^{d-1}$ via $e(x) = f_{\mathrm{text}}(x)/\|f_{\mathrm{text}}(x)\|_2$. During preference learning, both $e(\cdot)$ and $K_{\mathrm{dec}}$ remain strictly frozen, and $K_{\mathrm{dec}}$ receives only the sampled coordinate $z \in \mathcal{Z}$ without user interaction histories. Although $K_{\mathrm{dec}}$ is not assumed to invert $e(\cdot)$ analytically, reliable distribution-level generation requires that $K_{\mathrm{dec}}$ achieve high semantic cycle-consistency and rely strictly on the input coordinate $z \in \mathcal{Z}$ rather than falling back on generic language priors. We validate this property on a held-out set of 2,000 caption targets.

For each ground-truth caption $x \in \mathcal{X}$ and its normalized embedding $z = e(x) \in \mathcal{Z}$, we evaluate the frozen Qwen3.5-9B decoder $K_{\mathrm{dec}}$ from two perspectives. First, we assess generation reconstruction by synthesizing $\widehat{x} = K_{\mathrm{dec}}(z)$ via deterministic greedy decoding and measuring the cosine similarity between $\widehat{x}$ and $x$ using both the in-domain encoder (Chinese-CLIP) and an independent external encoder (BGE). Second, we evaluate conditional likelihood via the token-level negative log-likelihood (NLL) and perplexity (PPL) of the ground-truth caption $x$. To test whether decoding is genuinely controlled by the conditioning representation, we compare conditioning on the matched embedding (Correct, $z = e(x)$) against two counterfactual baselines: omitting the embedding prefix (Null) and conditioning on a mismatched embedding randomly drawn from another sample (Shuffled, $z' = e(x')$).

Table 4: Counterfactual conditioning audit of the frozen caption decoder $K_{\mathrm{dec}}$. We report paired improvements ($\Delta$) of conditioning on the Correct embedding over counterfactual baselines across 2,000 validation targets. Here, $\Delta\mathrm{BGE}$ and $\Delta\mathrm{Chinese\text{-}CLIP}$ denote gains in reconstruction cosine similarity ($\mathrm{Cos}_{\mathrm{Correct}} - \mathrm{Cos}_{\mathrm{Baseline}}$), and $\Delta\mathrm{NLL}$ denotes the reduction in token-level negative log-likelihood ($\mathrm{NLL}_{\mathrm{Baseline}} - \mathrm{NLL}_{\mathrm{Correct}}$).

| Conditioning Comparison | $\Delta$ BGE $\uparrow$ | $\Delta$ Chinese-CLIP $\uparrow$ | $\Delta$ NLL $\uparrow$ |
|---|---|---|---|
| Correct vs. Null | $+0.422$ | $+0.408$ | $+0.791$ |
| Correct vs. Shuffled | $+0.344$ | $+0.418$ | $+16.356$ |

Under Correct conditioning, greedy decoding achieves a Chinese-CLIP cosine similarity of $0.911$ and an independent BGE cosine similarity of $0.844$, alongside a conditional NLL of $1.202$ ($\mathrm{PPL} = 3.33$), zero empty outputs, and $100.0\%$ unique generations. Furthermore, as shown in Table 4, removing or shuffling the input embedding causes substantial degradation in both semantic similarity and target likelihood, with all $95\%$ paired bootstrap confidence intervals strictly above zero. These results confirm that $K_{\mathrm{dec}}$ faithfully translates spherical representations into corresponding textual content.

# F Dataset Details

Table 5: Dataset statistics.

| Dataset | Users | Items | Interactions |
|---|---|---|---|
| RecIF | 156,245 | 11,589,098 | 75,671,142 |
| MicroLens-100K | 100,000 | 19,738 | 719,405 |

## G Additional Implementation Details

**Architecture specifications.** The semantic space $\mathcal{Z}$ is defined on $\mathbb{S}^{d-1}$ with a projected dimension of $d = 512$. The dual-view transport network consists of 10 transformer blocks with a hidden dimension of 512. The continuous prefix for the Qwen3.5-9B decoder is mapped via an MLP to continuous conditioning tokens. The frozen XSimGCL model uses two propagation layers and 128-dimensional embeddings.

**Context and transport training.** Supplementary context is retrieved using only the current visible history, with padding masks for shorter context sets. We train the dual-view transport using dynamically sampled episodes with disjoint history and target sets. The learning rate is set to $4 \times 10^{-5}$ for the transport and fusion modules, and $2 \times 10^{-5}$ for the condition encoders, accompanied by a linear warmup of 500 steps. The context InfoNCE loss weight is set to $\lambda_{\text{ctx}} \approx 0.012$. The content encoder, context retriever, and caption decoder remain frozen throughout preference learning.

**Source calibration and decoding.** After transport training, we freeze the transport and its conditioning encoders to optimize only the source adapter. The adapter is trained using a global batch size of 512 users. We use a learning rate of $10^{-4}$ with 500 warmup steps. The calibration objectives are weighted by $\lambda_{\text{cov}} = 0.10$ and $\lambda_{\text{move}} = 0.01$. Gradients propagate through the numerical ODE solver directly to the adapter. At inference, calibrated source samples are transported under history and context conditioning, then mapped to continuous prefixes for the frozen caption decoder. Generated captions are neither reranked nor filtered using held-out targets. Code and pretrained weights will be made publicly available upon publication.

## H Quantitative Evaluation Metrics

All quantitative semantic metrics are computed in the normalized representation space. For each test user $u$, let $G_u = \{g_i\}_{i=1}^{K}$ denote the set of generated samples (with generation size $K = 64$), $H_u = \{h_j\}_{j=1}^{L}$ denote the observed historical interactions, and $T_u = \{t_m\}_{m=1}^{M}$ denote the held-out target interactions. All feature vectors are extracted using a frozen Chinese-CLIP encoder and $\ell_2$-normalized to lie on the unit hypersphere $\mathbb{S}^{d-1}$. We denote the standard cosine similarity as $\text{sim}(x, y) = x^\top y$.

All metrics are first computed independently per user and then aggregated via a macro-average across the test population: $\frac{1}{U}\sum_{u=1}^{U} \text{Metric}_u$. To ensure rigorous statistical comparison, confidence intervals and significance tests between methods are calculated using user-level paired bootstrapping with 2,000 resamples.

**History-nearest.** This diagnostic metric measures how closely the generated samples replicate the user's observed history. It is defined as the average maximum cosine similarity between each generated sample and the historical interactions:

$$\text{HistoryNearest}_u = \frac{1}{K}\sum_{i=1}^{K} \max_{1\le j\le L} \text{sim}(g_i, h_j). \tag{31}$$

While a moderately high value indicates historical consistency, an excessively high score often implies generation collapse toward memorizing past behaviors rather than exploring unobserved interests.

**Target-best.** This target-centric metric measures how closely each held-out target is matched by its nearest generated caption:

$$\text{TargetBest}_u = \frac{1}{M}\sum_{m=1}^{M} \max_{1\le i\le K} \text{sim}(g_i, t_m). \tag{32}$$

Higher values indicate stronger best-match semantic alignment with held-out targets. This metric does not assess the quality or relevance of every generated caption.

**Coverage.** We measure target-side semantic coverage using the average nearest-neighbor angular distance from held-out targets to generated samples:

$$\text{Coverage}_u = \frac{1}{M}\sum_{m=1}^{M}\min_{1\le i\le K}\arccos(\text{sim}(g_i, t_m))\,. \tag{33}$$

Lower values indicate smaller average distances from targets to their closest generated counterparts. Like Target-best, this metric measures target-side coverage rather than generation-side precision or probability-mass calibration.

**Future-mode coverage.** To assess coverage across distinct interests represented in held-out interactions, we evaluate the proportion of held-out target modes supported by generated samples under a thresholded nearest-target assignment rule. We first cluster the target interactions $T_u$ into $J_u$ distinct modes using average-linkage hierarchical clustering under the cosine distance $d_{\cos}(a, b) = 1 - \text{sim}(a, b)$, cutting the dendrogram at a distance threshold of $1 - \tau$ (with $\tau \approx 0.69$ for both datasets). Let $c(t_m) \in \{1, \dots, J_u\}$ denote the mode label of target $t_m$. Each generated sample $g_i$ is then assigned to its nearest target $t_{m(i)}$, where $m(i) = \arg\max_{1\le m\le M}\text{sim}(g_i, t_m)$ (breaking ties deterministically), and supports mode $c(t_{m(i)})$ only if $\text{sim}(g_i, t_{m(i)}) \ge \tau$. The metric computes the fraction of unique target modes supported across all $K$ generations:

$$\text{FutureModeCoverage}_u = \frac{1}{J_u}\left|\left\{c(t_{m(i)}) : 1 \le i \le K,\ \text{sim}(g_i, t_{m(i)}) \ge \tau\right\}\right|. \tag{34}$$

Higher values indicate broader coverage of held-out target modes, where each generated sample supports at most one mode and each covered mode is counted once.

**Energy score.** We evaluate probabilistic predictions using a lower-is-better energy score (Gneiting & Raftery, 2007), based on the Euclidean chordal distance $d_{\text{chord}}(x, y) = \|x - y\|_2 = \sqrt{2 - 2x^\top y}$. We estimate the mean score over held-out targets as

$$\text{Energy}_u = \frac{1}{KM}\sum_{i=1}^{K}\sum_{m=1}^{M} d_{\text{chord}}(g_i, t_m) - \frac{1}{2K(K-1)}\sum_{i\ne j} d_{\text{chord}}(g_i, g_j). \tag{35}$$

Lower values indicate better distributional predictions under this score. For a fixed held-out target set with $M \ge 2$, the corresponding unbiased two-sample energy-distance estimate equals twice this score minus the mean pairwise distance between distinct targets. This target-only term does not affect comparisons between methods evaluated on the same users and targets. The reported score is not a zero-normalized distance.

**Unbiased Maximum Mean Discrepancy ($\text{MMD}^2$-U).** We evaluate distributional agreement in the fixed semantic representation space using the unbiased two-sample $\text{MMD}^2$ estimator (Gretton et al., 2012):

$$\text{MMD}^2_{U,u} = \frac{1}{K(K-1)}\sum_{i\ne j} k_\sigma(g_i, g_j) + \frac{1}{M(M-1)}\sum_{a\ne b} k_\sigma(t_a, t_b) - \frac{2}{KM}\sum_{i=1}^{K}\sum_{m=1}^{M} k_\sigma(g_i, t_m). \tag{36}$$

For both datasets, we use a Gaussian kernel based on the Euclidean chordal distance between normalized representations:

$$k_\sigma(x, y) = \exp\left(-\frac{\|x - y\|_2^2}{2\sigma^2}\right) = \exp\left(-\frac{1 - x^\top y}{\sigma^2}\right). \tag{37}$$

The bandwidth is fixed across all compared methods within each dataset. The estimator requires at least two generated samples and two held-out targets. Users with fewer than two held-out targets are excluded only from MMD aggregation, using the same eligible-user set for all methods.

### H.1 Detailed LLM-based Assessment Protocol

We design a highly structured, five-phase LLM-as-a-judge pipeline (P1–P5). This pipeline strictly isolates target information from history processing to prevent information leakage and employs bidirectional swapping to eliminate position bias. All evaluations are conducted using strict JSON schemas for structural validation.

**Phases 1 & 2: Target Aspect Extraction and History Support Annotation.** To objectively evaluate incremental completion (Q3), we must first identify which future targets are truly "unobserved" in the history. In **P1**, the evaluator extracts distinct primary semantic aspects strictly from the held-out targets without accessing any generated outputs. Synonymous interests are merged, and ambiguous items are discarded. In **P2**, the evaluator compares the observed history chunks against these frozen target aspects, categorizing the historical support into levels such as *observed*, *weakly observed*, *weakly partial*, or *not observed*. Targets falling into the weak or unobserved categories form the basis for evaluating incremental coverage.

**Phase 3: Intrinsic Caption Quality Assessment.** Before computing semantic coverage, **P3** filters out degenerate generations. The evaluator scores each generated caption on a 0-2 scale based solely on its intrinsic expressiveness, without seeing the target interactions:

- **0:** Empty, incomprehensible, or severely fragmented.
- **1:** Broadly understandable but generic, templated, or lacking specific detail.
- **2:** Clear, coherent, and expressing recognizable, specific content.

Only captions scoring a 2 are considered "valid" in subsequent coverage calculations, strictly penalizing models that generate diverse but nonsensical text.

**Phase 4: Set-Level Preference Comparison (Q1).** In **P4**, the evaluator compares two anonymous generated sets (Model A and Model B) in their entirety, conditioned on the user's history and held-out targets. The evaluation considers four dimensions: *Relevance* (alignment with observed behaviors), *Coverage* (representation of multiple distinct interests), *Frequency* (proportional allocation of generated modes), and *Validity* (absence of degenerate outputs).

To strictly mitigate position bias (the LLM's tendency to favor the first presented option), we employ a bidirectional evaluation protocol: every Model A vs. Model B comparison is evaluated twice by swapping their presentation order. A definitive win is recorded only if the evaluator consistently favors the same model across both permutations; otherwise, the result is recorded as a tie.

**Prompt Excerpt for Set-Level Preference Comparison (Phase 4)**

Given a user's historical interactions, true held-out future targets, and two anonymous generated candidate sets (Set A and Set B), you are an expert personalized preference analysis assistant. Please analyze the candidate sets and generate a Chain-of-Thought style evaluation.

**Scoring Dimensions:**

- **Relevance:** Does the set align with the user's observed historical evidence?
- **Coverage:** Are the multiple distinct interests from the targets covered?
- **Frequency:** Is the distribution of topics reasonably balanced?

**Your final output must strictly follow this JSON-parsable structure:**

```
<think>
1. Relevance Analysis:
Set A: {reasoning}; Set B: {reasoning}
2. Coverage Analysis:
Set A: {reasoning}; Set B: {reasoning}
...
</think>
<answer>
{
  "diagnostics": {"relevance": "...", "coverage": "..."},
  "overall_preference": "A" // Options: "A", "B", "tie", "insufficient"
}
</answer>
```

**Phase 5: Target-to-Candidate Semantic Matching (Q2 & Q3).** In **P5**, we evaluate fine-grained target-level support. The evaluator scores the semantic correspondence between a specific held-out target and each candidate in the fully anonymized and randomly shuffled generated set on a 0-3 scale:

- **0:** No meaningful semantic correspondence.
- **1:** Shares only broad domains or generic keywords.
- **2:** Matches the substantive interest/activity, but lacks sufficient detail.
- **3:** Highly matched core semantics and key details.

Candidates achieving a score of 2 or 3 are deemed to successfully cover the target. This phase drives the calculation for Target-level Support (**Q2**) and, when restricted to the weak-history targets identified in P2, computes Incremental Completion (**Q3**).

### H.2 Metrics for Set-level LLM Evaluation

We formalize the computation of the four metrics presented in Table 2 based on the LLM-as-a-judge pairwise evaluations in Section 5. Let $\mathcal{T}_u$ denote the set of held-out targets for user $u$, and $\mathcal{S}_u$ denote the generated set of captions.

**Win rate (WR).** For the overall preference (Q1), the LLM judge outputs a pairwise decision $d_u \in \{\text{Win}, \text{Lose}, \text{Tie}\}$ comparing our approach against a baseline. The Win rate is calculated as the proportion of users where our generated set is explicitly preferred:

$$\text{WR} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \mathbb{I}(d_u = \text{Win}) \times 100\%, \tag{38}$$

where $\mathcal{U}$ is the set of test users and $\mathbb{I}(\cdot)$ is the indicator function.

**Best-match (BM).** For target-level support quality (Q2), the LLM assigns a semantic alignment score $s(x,t) \in \{0,1,2,3\}$ for each generated caption $x \in \mathcal{S}_u$ against a target $t \in \mathcal{T}_u$. The Best-match score is the maximum alignment achieved for a target:

$$\text{BM} = \frac{1}{\sum_u |\mathcal{T}_u|} \sum_{u \in \mathcal{U}} \sum_{t \in \mathcal{T}_u} \max_{x \in \mathcal{S}_u} s(x,t). \tag{39}$$

**Valid coverage (VC).** Valid coverage strictly isolates successful hits driven by intrinsically high-quality generations. A generation $x$ is deemed valid only if it receives a maximum intrinsic quality score $q(x) = 2$ (on a 0-2 scale). A target $t$ is covered if there exists at least one valid generation $x \in \mathcal{S}_u$ that achieves a strong semantic match ($s(x,t) \geq 2$). The VC metric is defined as:

$$\text{VC} = \frac{1}{\sum_u |\mathcal{T}_u|} \sum_{u \in \mathcal{U}} \sum_{t \in \mathcal{T}_u} \mathbb{I}\Big(\exists x \in \mathcal{S}_u : q(x) = 2 \wedge s(x,t) \geq 2\Big). \tag{40}$$

**Incremental valid coverage (IVC).** For incremental completion (Q3), we focus on the subset of under-represented or unseen targets $\mathcal{T}_u^{\text{inc}} \subseteq \mathcal{T}_u$, which consists of held-out targets annotated in Phase 2 as either weakly supported (*weakly observed*, *weakly partial*) or completely absent (*not observed*) in the user's observed history. IVC applies the exact same valid coverage computation as Eq. (40), but restricted exclusively to $\mathcal{T}_u^{\text{inc}}$. This quantifies the model's ability to move beyond dominant historical behaviors and complete weakly evidenced or unobserved preference regions.