UNIFIED FRAMEWORK FOR GENERATIVE SEQUENTIAL RECOMMENDATION

Anonymous authors
Paper under double-blind review

ABSTRACT

How much variation a generative recommender's sampler introduces on the way to a sample, its stochasticity level, is a property of the user rather than of the task. A long history already determines the next item, so exploration can only cost; a short history, or interests spanning several catalog categories, leaves the target under-determined. No existing generative recommender can act on this, because none can vary the level at all: it is determined jointly by the interpolant and the training objective. We remove that constraint. Denoising diffusion (DDPM), its deterministic counterpart (DDIM), Flow Matching and Schrödinger Bridges are one SDE at different values of that scalar; the theory unifying them holds it fixed across the population, and extending that theory frees it to be any measurable function of the user, every conditional marginal preserved. The level appears in neither training objective, so a trained checkpoint reproduces these processes at inference rather than committing to one of them before training. Serving it per user pays in diversity at no cost in accuracy, on corpora whose users genuinely differ, and when the next several items are predicted at once.

1 INTRODUCTION

Generative models have displaced point predictors across much of machine learning. Among generative models, we study the *iterative-refinement* family, i.e., those that transport a source distribution to the data distribution over a sequence of denoising steps. They return a *distribution* rather than one output, they accept *conditioning at inference*, which makes the generative process an object in its own right: separable from the weights, swappable without retraining. Four generative processes dominate practice: Flow Matching (FM) (Lipman et al., 2023; Liu et al., 2023b), Schrödinger Bridges (SB) (De Bortoli et al., 2021; Shi et al., 2023), and the two of Diffusion (DM) that differ in nothing but the sampler, stochastic Denoising Diffusion Probabilistic Models (DDPM) and deterministic Denoising Diffusion Implicit Models (DDIM) (Ho et al., 2020; Song et al., 2021a).

Which generative models to use is a settled question in domains like image and audio synthesis (Karras et al., 2022; Esser et al., 2024; Kong et al., 2021), because samples are interchangeable. Sequential recommendation has no such answer (Kang & McAuley, 2018; Sun et al., 2019). Two users with comparable histories, and still more two users whose histories differ, need different amounts of spread in the search over the catalog: a single narrow interest is best served by a process that lands in the same place every time, and several interests by one that can reach several. Generative recommenders nonetheless deploy these models as they are, with one amount of spread for every user, across sequential prediction (Yang et al., 2023; Li et al., 2023; Liu et al., 2025), collaborative filtering (Wang et al., 2023; Walker et al., 2022; Zhu et al., 2024) and generative retrieval (Rajput et al., 2023; Zhai et al., 2024), c.f., App. A for a full discussion.

What differentiates generative models is the spread their sampler injects between source and sample, and therefore how much their output varies from one draw to another. In sequential recommendation, a long history already determines the next item, so injected noise can only cost; a short history, or interests spanning several catalog categories, leaves the target under-determined and is better served by a sampler that can reach more than one answer. Personalizing the level is therefore worth doing in *sequential* recommendation in a way it is not elsewhere. It is warranted, because users vary in exactly the two quantities that set the optimum, from a handful of interactions to several hundred and from one interest to many; and it is feasible. The question is not which generative model, but

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.

which one *for whom*. No existing generative recommender can ask that question, because none can vary the spread injected by the sampler for each user: the amount is pinned by the interpolant and the training objective, and the sampler follows from those two choices, so it is fixed before a single user is seen. Varying the spread from user to user is therefore not an idea that has been tried and rejected; it is structurally unavailable. Our contributions could be summarized as follows.

- **A stochasticity level that may depend on the user.** Stochastic interpolants index flows and diffusions by one level shared across the population (Albergo & Vanden-Eijnden, 2023). We extend that family to a level measurable in the user's history, every conditional marginal preserved (Thm. 3), which is what makes which level for whom a question that can be asked.
- **STIR, which makes that level available.** The level appears in neither training objective, so it is chosen at inference rather than before training: a trained checkpoint reproduces the established processes to within seed noise, and two users can be served two of them out of the same weights.
- **The recommendation target fixes the remaining choices.** We model the sequential recommendation task itself to derive from it theoretical links between the generative process and user behaviour.
- **Where per-user structure pays.** On seven public benchmarks it is real and collectible: in diversity at no accuracy cost, where users genuinely differ, and when several next items are predicted together.

2 PROBLEM FORMULATION

Sequential recommendation ranks a catalog of N items, indexed by $i \in \{1, \dots, N\}$, for a user given the items already consumed. Each item i carries a fixed ℓ_2-normalized embedding $e_i \in \mathbb{R}^d$, the i-th row of the item table $E \in \mathbb{R}^{N \times d}$. Each user is represented by a history, i.e., a sequence of T items consumed in order, written through their embeddings as $h = (x_1, \dots, x_T)$, and encoded into $h_u = \mathrm{Enc}(h)$. While a conventional recommender scores every candidate directly, a *generative* one learns the distribution a history induces over next items, one history being consistent with several. It draws a point $\hat{x} \sim \hat{p}(\cdot \mid h)$ in embedding space, approximating the law $p(x_{T+1} \mid h)$ of the next item's embedding, and returns the K items whose embeddings score highest against it,

$$\text{Top-K}(h) \;=\; \arg\operatorname{topK}_{i \in [N]} \langle \hat{x}, e_i \rangle. \tag{1}$$

We denote by ρ_1 the target law, i.e., $\rho_1 := p(\cdot \mid h)$. Every generative recommender realizes $\hat{p}$ as the terminal law of a transport process from a source ρ_0 to ρ_1. That process fixes four components:

- a *source* ρ_0 generation starts from, usually a standard Gaussian, $x_0 \sim \rho_0$;
- a *path* linking ρ_0 to ρ_1, along which a training point x_t runs from $x_0 \sim \rho_0$ at $t{=}0$ to $x_1 \sim \rho_1$ at $t{=}1$, and how much noise is mixed into it;
- a *denoising objective* $\mathcal{L}$ where a network $\hat{f}_\theta$ learns an exact field f, e.g., $\mathcal{L} = \mathbb{E}\|\hat{f}_\theta(t, x_t, h) - f\|^2$;
- a *sampler* $\mathcal{A}$, an ordinary or stochastic differential equation (ODE, SDE) integrating $\hat{f}_\theta$ from $x_0 \sim \rho_0$ to the endpoint $\hat{x}$, and with it the *level* of noise it injects on the way.

The first three components are fixed during training, and $\mathcal{A}$ is the only one that runs at inference, and recommenders simply keep the sampler their backbone shipped with (Yang et al., 2023; Li et al., 2023; Wang et al., 2023; Liu et al., 2025). What the sampler $\mathcal{A}$ controls is how much the output varies from one draw to the next, i.e., if we run the sampler twice on the same user and it returns two different points, and that spread is what lets it reach a mode a conditional mean would miss (Song et al., 2021b; Karras et al., 2022). The literature already classifies methods by how much of it they carry, from the deterministic ODEs of FM and DDIM through a bridge's fixed diffusion coefficient to a full reverse SDE, and exposes it as a dial in DDIM's noise parameter and in the churn of Karras et al. (2022). Two mechanisms produce it, worth separating: the *source-propagated* part, carried from x_0 drawn afresh at every request, and the *injected* part, the noise the sampler adds along the path. Def. 1 measures that variation, the *dispersion* of the output, and separates the two.

Definition 1 (Dispersion of a generative recommender)**.** *With the generative model trained and the history h held fixed, let $\mathcal{A}$ be the sampler: an ODE or an SDE run from a source draw $x_0 \sim \rho_0$ to an endpoint $\hat{X}_{\mathcal{A}}(h) \in \mathbb{R}^d$, whose only randomness is x_0 and, for an SDE, the Brownian path it follows. Its* dispersion *at h is the total variance of that endpoint, which the law of total variance splits into*

$$\mathrm{R}_{\mathcal{A}}(h) = \operatorname{tr}\operatorname{Cov}\big[\hat{X}_{\mathcal{A}}(h)\big] = \underbrace{\mathbb{E}_{x_0 \sim \rho_0}\big[\operatorname{tr}\operatorname{Cov}(\hat{X}_{\mathcal{A}} \mid x_0)\big]}_{\textit{injected}} + \underbrace{\operatorname{tr}\operatorname{Cov}_{x_0 \sim \rho_0}\big(\mathbb{E}[\hat{X}_{\mathcal{A}} \mid x_0]\big)}_{\textit{source-propagated}}. \tag{2}$$

We ran the four native methods on sequential recommendation and measured both terms. All four were trained on the same task, generating the embedding of the item a user consumes next from that

user's history, and each follows its own framework: its own noising path, its own objective and its own sampler. We estimate the two terms of Eq. (2) empirically from repeated sampler draws per user, and report the median across users. As shown in Fig. 1 on the dataset ML-1M, the four methods deliver measurably different dispersion and differ in where it comes from: the two ODEs (FM, DDIM) inject none and inherit all of theirs from the source, while the two SDEs (SB, DDPM) draw between a quarter and three fifths of theirs from injection. Full details of the experiment are in App. K.

If the processes deliver different dispersion, then how much of it a given user needs. Exploring the catalog is not equally useful to everyone: a user who returns to one narrow interest is best served by a process that lands in the same place every time, whereas a user spread over several categories needs one that can reach more than one of them. How much exploration helps is therefore a property of the user, not of the corpus. That is the hypothesis this paper tests. It concerns the user's dispersion $\mathrm{R}^\star(h)$ on it conditional marginal rather than any model's output, i.e., $\mathrm{R}^\star(h) = \operatorname{tr}\operatorname{Cov}\big[x_{T+1} \mid h\big]$ where $x_{T+1} \sim p(\cdot \mid h)$. We test whether it is constant across the population or user-dependent, i.e.,

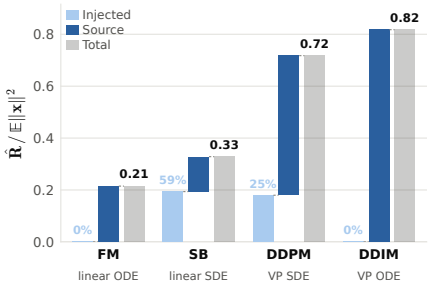


Figure 1: **The four native processes deliver different dispersion.** Median over 4k ML-1M users, S=32 sampler runs at frozen weights, split into the two terms of Eq. (2).

$$\mathcal{H}_{\text{corpus}}: \quad \exists \mathrm{R}^\star \geq 0 \ \ \forall h, \ \ \mathrm{R}^\star(h) \equiv \mathrm{R}^\star \qquad \text{versus} \qquad \mathcal{H}_{\text{user}}: \ \ \operatorname{Var}_h\big[\mathrm{R}^\star(h)\big] > 0. \tag{3}$$

Serving users with different levels of dispersion then means one training run per level, or one per subgroup at best, which is infeasible due to high computational cost, and the data scarcity on each subgroup. Therefore, we need instead a single model in which all four methods could be recovered during inference using one unified framework that requires only a single training, which we build in §3.

3 The Stochastic Interpolant Recommender

In no existing model can the level be changed at inference, still less chosen differently for one user than for another. Raising it means injecting noise while sampling, and injected noise has to be cancelled by a second field, or the sampler stops drawing from the distribution it was trained to produce. We present more details about this limitation in App. E. We build one object in which the level is free, the *Stochastic Interpolant Recommender* (STIR). A generative process is named by three quantities. The *path* $\mathcal{I}(t, x_0, x_1)$ is the route a training point takes from the source to the target. The *envelope* $\gamma(t)$ is how much noise is mixed into that route. The *level* ϵ is the hyperparameter whose values give us the sampler of FM, SB, DDPM or DDIM in turn. The first two are chosen before training and the third at inference, which is what lets one checkpoint recover different generative recommenders:

$$\underbrace{(\mathcal{I}, \gamma)}_{\text{fixed at training}} + \underbrace{\epsilon}_{\text{set at inference}} \longmapsto \text{one of the four, or a process none of them can reach.}$$

3.1 The Interpolant and the Two Fields It Induces

The path $(\mathcal{I}, \gamma)$ must come from a family wide enough to hold every process we mean to recover. In Def.2, we present a unified framework for interpolants, blurred by an envelope that vanishes at both ends so each endpoint is reached exactly.

Definition 2 (Stochastic interpolant (Albergo et al., 2025)). *For independent $x_0 \sim \rho_0$, $x_1 \sim \rho_1$, and $z \sim \mathcal{N}(0, I_d)$,*

$$x_t = \mathcal{I}(t, x_0, x_1) + \gamma(t)\, z, \qquad t \in [0, 1], \tag{4}$$

with $\mathcal{I}(0, \cdot, \cdot) = x_0$, $\mathcal{I}(1, \cdot, \cdot) = x_1$, and $\gamma(0) = \gamma(1) = 0$, $\gamma \geq 0$; ρ_t denotes the law of x_t.

The family includes FM ($\gamma \equiv 0$), SB, and the Variance-Preserving (VP) path of DDPM and DDIM. All four live in the affine subfamily $\mathcal{I}(t, x_0, x_1) = \alpha(t)x_0 + \beta(t)x_1$, which is where we work from here, c.f., Table 1. Once the interpolant is fixed, writing $\rho(t, \cdot \mid h)$ for the law of x_t when the endpoint is drawn from $p(\cdot \mid h)$, everything a sampler can use is carried by two fields that law induces, the *velocity* and the *score*,

$$b(t, x, h) := \mathbb{E}\big[\dot{x}_t \mid x_t = x, \, h\big], \qquad s(t, x, h) := \nabla_x \log \rho(t, x \mid h). \tag{5}$$

Table 1: Each existing recommender is Eq. (6) at one fixed level (App. D). The two *Sampler* columns are the same process written twice, as the method publishes it and as STIR read at the level beside it.

	Interpolant (fixed at training)		Sampler				
Method	**Path** $\mathcal{I}(t, x_0, x_1)$	**Envelope** $\gamma(t)$	**The process**	**Equivalent process in STIR's parameterization**		**Free at inference**	**Reason**
Linear interpolant $\mathcal{I}_{\text{lin}}$							
FM	$(1-t)\,x_0 + t\,x_1$	$\gamma_\sigma, \sigma{=}0$	ODE $\mathrm{d}X = b\,\mathrm{d}t$	$\epsilon{=}0$	$\mathrm{d}X = b\,\mathrm{d}t$	✗	No score head
SB	$(1-t)\,x_0 + t\,x_1$	$\gamma_\sigma, \sigma{>}0$	SDE $\mathrm{d}X = \big(b + \frac{\sigma^2}{2} s\big)\mathrm{d}t + \sigma\,\mathrm{d}W$	$\epsilon{=}\frac{\sigma^2}{2}$	$\mathrm{d}X = (b + \epsilon s)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W$	✗	Emits the two fused
VP interpolant $\mathcal{I}_{\text{vp}}$							
DDPM	$\alpha_{\text{vp}}(t)\,x_0 + \beta_{\text{vp}}(t)\,x_1$	0	SDE $\mathrm{d}X = \omega_{\text{vp}}\big(X + 2s\big)\mathrm{d}t + \sqrt{2\omega_{\text{vp}}}\,\mathrm{d}W$	$\epsilon{=}\omega_{\text{vp}}$	$\mathrm{d}X = (b + \omega_{\text{vp}} s)\mathrm{d}t + \sqrt{2\omega_{\text{vp}}}\,\mathrm{d}W$	✗	Tied to the schedule
DDIM	$\alpha_{\text{vp}}(t)\,x_0 + \beta_{\text{vp}}(t)\,x_1$	0	ODE $\mathrm{d}X = \omega_{\text{vp}}\big(X + s\big)\mathrm{d}t$	$\epsilon{=}0$	$\mathrm{d}X = b\,\mathrm{d}t$	✗	Level forced to 0
STIR (ours)	any $\alpha(t)\,x_0 + \beta(t)\,x_1$	any $\gamma \geq 0$	either $\mathrm{d}X = (b + \epsilon s)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W$, any $\epsilon(t, h) \geq 0$			✓	Two heads, any $\epsilon(t, h)$

3.2 ONE SAMPLER FOR THE FOUR

Having fixed the interpolant and its envelope in Eq. 4, we seek a parametrized SDE that generalizes the samplers of the four established processes, i.e. FM, SB, DDPM and DDIM. Its hyperparameter should be independent of the training and set at inference, and its values should recover the exact SDE or ODE of each of them in turn, and with it their results. It should also keep the guarantee each native sampler comes with, that we recover the target law ρ_1, at every value of that hyperparameter. We construct that SDE in Eq. 6 and prove the second property in Thm. 3.

Theorem 3 (Marginal Preservation; Proof in App. H.1)**.** *Let b, s be the exact fields of Def. 2 and let $\epsilon : [0, 1] \times \mathcal{H} \to \mathbb{R}_{\geq 0}$ be* any *measurable function of time and conditioning. Then*

$$\mathrm{d}X_t = \big(b(t, X_t, h) + \epsilon(t, h)\, s(t, X_t, h)\big)\mathrm{d}t + \sqrt{2\epsilon(t, h)}\,\mathrm{d}W_t, \qquad X_0 \sim \rho_0, \tag{6}$$

has marginals $\rho(t, \cdot \mid h)$ for every t and every h; in particular $X_1 \sim p(\cdot \mid h)$.

The SDE in Eq. (6) is parametrized by the *level* function ϵ, the rate at which our sampler adds fresh Brownian noise as it runs. As Thm. 3 proves, the noise $\sqrt{2\epsilon}\,\mathrm{d}W_t$ and the score term $\epsilon\, s$ cancel in the marginals at every t, so we reach ρ_1 however we set ϵ. What is new here is the dependence on h: existing stochastic-interpolant theory fixes one level, or at most a schedule $\epsilon(t)$, for every sample (Albergo et al., 2025), whereas Eq. (6) fixes a map $h \mapsto \epsilon(\cdot, h)$ and preserves every conditional marginal at once. The statement is therefore about assigning users to samplers, not about choosing one.[1] As derived in App. D and Table 1, we can recover the four methods by setting specific values for the level ϵ:

$$\epsilon = 0 : \quad \mathrm{d}X = b\,\mathrm{d}t \qquad \text{(FM, DDIM)}$$

$$\epsilon \in \{\tfrac{\sigma^2}{2},\, \omega_{\text{vp}}\} : \quad \mathrm{d}X = \big(b + \epsilon\, s\big)\,\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W \qquad \text{(SB, DDPM)}$$

with σ the Bridge's diffusion coefficient and ω_{vp} fixed by DDPM's noise schedule. Eq. (6) reads b and s separately, since ϵ multiplies only s.

Training Without the Level. We fit the two fields of Eq. (5) with a head apiece, and *neither objective contains* ϵ:

$$\mathcal{L}_b[\hat{b}_\theta] = \int_0^1 \mathbb{E}\big|\hat{b}_\theta(t, x_t, h) - \dot{x}_t\big|^2 \mathrm{d}t\,, \qquad \mathcal{L}_s[\hat{s}_\phi] = \int_0^1 \mathbb{E}\big|\hat{s}_\phi(t, x_t, h) - s(t, x_t, h)\big|^2 \mathrm{d}t\,. \tag{7}$$

We therefore train once and decide the level afterwards. At inference we read the two heads into Eq. (6), integrate its learned counterpart, i.e., we sample from the following SDE,

$$\mathrm{d}X_t = \big(\hat{b}_\theta(t, X_t, h) + \epsilon(t, h), \hat{s}_\phi(t, X_t, h)\big)\mathrm{d}t\sqrt{2\epsilon(t, h)}, \mathrm{d}W_t. \tag{8}$$

Reading the Four Processes off One Checkpoint. In §3.2 we recovered each native sampler as Eq. (6) at one value of ϵ, written in the exact fields b and s: at that level we integrate the same equation they do, so we draw from the same law as each of the four native methods. But during deployment at inference, neither they nor we have those exact fields. Each native method integrates its own equation on the field its own loss produced, and we integrate ours on the two fields Eq. (7) produces. Recovering a method therefore needs its loss and ours to deliver the same field, and this is where the four part company.

[1] Stated for $\epsilon(t, h)$, not $\epsilon(t, x, h)$: if the level varied with x, the cancellation in App. H.1 would leave a residual $\nabla \cdot (\rho \nabla_x \epsilon)$, and the drift would need an extra $+\nabla_x \epsilon$ to remove it. Our levels do not depend on x, so Eq. (6) carries no such term. More details could be found in App. F.

The four split into two regimes. For FM and DDPM the native loss *is* one of ours, up to a weight depending on time alone, and such a reweighting moves no minimizer: the two objectives are the same objective, so the two checkpoints agree at *any* weights and the recovery is exact. For SB and DDIM , neither method regresses one of our two targets on its own, we can claim recovery at the population optimum. We derive both regimes in App. D and measure them in Table 2.

What is left to decide before training is not which process to adopt but which interpolant suits sequential recommendation, and §4 settles that. Once $(\mathcal{I}, \gamma)$ is fixed, a process is nothing but its level. Write $\mathcal{A}_\epsilon$ for our checkpoint decoded at a level ϵ, and $\mathcal{M}_\epsilon$ for a native method trained at that same level. Recovery is then the statement $\mathcal{A}_\epsilon \approx \mathcal{M}_\epsilon$. Suppose two users want $\epsilon_1 \neq \epsilon_2$. Instead of training two different models $\mathcal{M}_{\epsilon_1}, \mathcal{M}_{\epsilon_2}$, we can use instead STIR, and serve both users with the same model, i.e., $\{\mathcal{M}_{\epsilon_1}, \mathcal{M}_{\epsilon_2}\} \equiv \{\mathcal{A}_{\epsilon_1}, \mathcal{A}_{\epsilon_2}\}$: two training runs on the left, one checkpoint on the right. If instead a single level serves the whole corpus, the checkpoint still pays: the level is chosen after training, so the four processes become four settings of one model rather than four models to train.

4 DETERMINING THE PROCESS FOR RECOMMENDATION

Three choices are still to make, the *path* $\mathcal{I}$, the *envelope* γ and the *level* ϵ. Reasoning about them needs a tractable picture of what a recommender generates, so we take the next-item law to be a mixture over catalog modes. We use it to derive theoretical intuitions, not to describe any dataset, and we verify the resulting predictions experimentally in §5.

4.1 MODELING THE TARGET

Unlike other domains, recommendation draws its target from a catalog we hold in advance. Items fall into modes, a user is a weighting over them, and the distribution we generate therefore has a form we can write down. Assumption 4 does exactly that.

Assumption 4 (Mode Mixture)**.** *The next-item distribution in embedding space is a finite Gaussian mixture over M catalog* modes*, whose centroids μ_m and common isotropic covariance $\Sigma = \varsigma_1^2 I_d$ belong to the catalog and are shared across the population, so that* all *user-specific information sits in the weights $\pi(h) = (\pi_1(h), \dots, \pi_M(h))$ with $\sum_m \pi_m(h) = 1$, i.e.,*

$$p(x_1 \mid h) \;=\; \textstyle\sum_{m=1}^{M} \pi_m(h)\, \mathcal{N}\big(x_1;\, \mu_m,\, \Sigma\big). \tag{9}$$

In Eq. (9) the modes and their spread belong to the catalog and every user shares them; only the weights $\pi(h)$ are the user's own. The modes are countable, *catalog categories* being one instance and a habit or a shopping occasion another. The assumption constrains the prediction-time conditional only, leaving histories free to be non-stationary and non-Markov. And Gaussian mixtures are universal approximators, dense in the space of densities, so the parametric form costs little and the shared centroids carry the weight.

4.2 PATH AND ENVELOPE

Both $\mathcal{I}$ and γ are committed before any user is seen, and we show in Thm 5 that Assumption 4 determines both, so neither is a hyperparameter here.

Theorem 5 (The recommendation interpolant; Proof in App. H.1)**.** *Let ρ_0 be Gaussian and independent of a target obeying Assumption 4.*

(a) Path. *Over affine paths $\mathcal{I} = \alpha(t)x_0 + \beta(t)x_1$, the choice $\alpha^\star(t) = 1 - t$, $\beta^\star(t) = t$ uniquely minimizes the transport cost $\int_0^1 \mathbb{E}\|\dot{x}_t\|^2 \mathrm{d}t$ and is the only one with finite irreducible velocity loss; on a variance-preserving path both terms diverge.*

(b) Envelope. *Fixing that path, training on n samples in d dimensions and writing $H_\rho = \sup_t \mathbb{E}\|\nabla^2 \log \rho_t\|_F^2$, the total error is $O(\sigma^2 H_\rho) + \Theta\big(d/(n\sigma^2)\big)$, uniquely minimized by $\gamma^\star(t) = \sigma^\star\sqrt{t(1-t)}$ with $\sigma^\star = \Theta((d/nH_\rho)^{1/4})$.*

Both parts point away from what the field imported. Part (a) singles out the linear path and rules out the variance-preserving one, on which the transport cost and the irreducible velocity loss both diverge: the path DDPM and DDIM carry over from image diffusion is the one our target excludes, and Table 2 shows the cost of keeping it. Part (b) makes the envelope a function of d/n alone, so its

amplitude is derived rather than swept, and its shape $\sqrt{t(1-t)}$ is the Bridge's, which leaves σ as the only quantity to fix. With d=64 and n in the hundreds of thousands to millions, $\sigma^\star$ is small: the theory prescribes near-deterministic transport, where image diffusion, with d in the thousands, sits at the other end of the same expression.

4.3 The Two Field Errors

Every method here comes with the same guarantee as ours in Thm. 3, i.e. the sampler reaches ρ_1. But as explained in §3.2, each method integrates on the learned field, so the distance from ρ_1 is set by that fit, which is why the four in Fig. 1 agree in composition, as Thm. 3 requires, but differ in total dispersion. The two fields are not equally easy to fit. A history is consistent with several next items, and the velocity is a conditional mean, Eq. (5), so on an ambiguous history it can only average the continuations that history admits and keeps an irreducible loss whatever the model; the score is a local gradient of the same law and keeps none. A method that regresses the velocity therefore starts from a worse fit than one that regresses the score, exactly where users are ambiguous. In our sampler both fits appear at once, as a hat on each field in Eq. (8), which leaves the cancellation between the injected noise and the score term only approximate. We name the two errors that residue is built from,

$$\delta_b(h) = \int_0^1 \mathbb{E}_{x_t\sim\rho(t,\cdot|h)} \big\|\hat{b}_\theta - b\big\|^2 \mathrm{d}t, \qquad \delta_s(h) = \int_0^1 \mathbb{E}_{x_t\sim\rho(t,\cdot|h)} \big\|\hat{s}_\phi - s\big\|^2 \mathrm{d}t, \tag{10}$$

both fields evaluated at (t, x_t, h) and each expectation running over that user's own interpolant marginal. Since the ambiguity that costs the velocity is the user's own, the ratio $\delta_b(h)/\delta_s(h)$ moves with the user, and it is the quantity ϵ acts on: raising the level buys accuracy in the weaker field by admitting noise the stronger one can still correct.

4.4 The Stochasticity Level

As explained in §4.3, the ratio $r(h) = \delta_b(h)/\delta_s(h)$ moves with the user, and it is what should ϵ acts on. We confirm this intuition in Thm. 6. The level that minimizes a user's delivered error explicitly depends on $r(h)$, up to one scale shared by the whole population.

Theorem 6 (Threshold Form of the Optimum; Proof in App. I)**.** *Under mild regularity, the Wasserstein-2 distance between the target $p(\cdot \mid h)$ and the delivered $\hat{p}^\epsilon(\cdot \mid h)$ obeys $\mathcal{W}_2(p, \hat{p}^\epsilon) \le \kappa(\epsilon)\big(\sqrt{\delta_b(h)} + \epsilon\sqrt{\delta_s(h)}\big)$, with κ* decreasing *in ϵ and $\kappa(0)$ finite, and its convex surrogate is minimized by,*

$$\epsilon^\star_\lambda(h) = \max\Big(0,\ \tfrac{1}{\lambda}\big(\sqrt{\lambda\, r(h)} - 1\big)\Big). \tag{11}$$

Therefore what sets a user's level is which of the two fields their own behaviour makes easier to learn. A user who returns to one narrow interest leaves the velocity an easy target: the continuations agree, the conditional mean is close to all of them, $r(h)$ is small and the rule sends that user to $\epsilon = 0$, a deterministic process. A user whose history spans several catalog categories leaves the velocity averaging continuations that point in different directions, while the score still reads the same local gradient wherever the state happens to be; $r(h)$ is large, and the rule buys accuracy in the weaker field by admitting noise the stronger one can still correct. Two histories with the same ratio are owed the same level however else they differ, and the users either side of the threshold are owed an ODE and an SDE out of the same checkpoint, so no single ϵ is right for everyone unless the population shares one ratio. §4.5 identifies the two behaviours that move it.

4.5 Which Users Sit Where

We show in Prop. 7 which users carry a high or low level ϵ, through two behaviours a recommender can observe. We define $\mathcal{L}_b^*(h) = \int_0^1 \mathbb{E}_{x_t}[\mathrm{Var}(\dot{x}_t \mid x_t, h)]\mathrm{d}t$, the part of the velocity loss Eq. (7) that no model can remove, c.f., Eq (28) in App. H.

Proposition 7 (Recommendation semantics; Proof in App. J.1)**.** *Under Assumption 4 with the linear interpolant and a Gaussian source:*

(a) Preference spread. *$\mathcal{L}_b^*(h) \ge c\,\mathrm{Var}_{\pi(h)}(\mu_m)$ for an explicit $c > 0$, where $\mathrm{Var}_{\pi(h)}(\mu_m)$ is the spread of the mode centroids under the user's own weights $\pi(h)$.*

Table 2: **One checkpoint, decoded at four levels.** NDCG@20 against the full catalog. Each upper block is a *native*, then STIR decoded at the level derived for it. Subscripts are standard deviations over 5 seeds in units of the last digit ($0.0240_{\pm 8}$ means 0.0240 ± 0.0008).

	BEAUTY	TOYS	SPORTS	ML-1M	YELP	GOWALLA	GOODREADS
Flow Matching							
native	$0.0230_{\pm 12}$	$0.0199_{\pm 12}$	$0.0112_{\pm 4}$	$0.0840_{\pm 22}$	$0.0191_{\pm 7}$	$0.0305_{\pm 6}$	$0.0991_{\pm 26}$
STIR $(\mathcal{I}_{\mathrm{lin}}, \gamma_\sigma)$, ϵ=0	$0.0232_{\pm 9}$	$0.0200_{\pm 4}$	$0.0110_{\pm 6}$	$0.0812_{\pm 35}$	$0.0188_{\pm 16}$	$0.0300_{\pm 12}$	$0.0919_{\pm 132}$
STIR $(\mathcal{I}_{\mathrm{lin}}, 0)$, ϵ=0	$0.0233_{\pm 6}$	$0.0203_{\pm 9}$	$0.0098_{\pm 7}$	$0.0812_{\pm 16}$	$0.0179_{\pm 13}$	$0.0283_{\pm 12}$	$0.1005_{\pm 25}$
Schrödinger Bridge							
native	$0.0202_{\pm 17}$	$0.0159_{\pm 11}$	$0.0095_{\pm 7}$	$0.0807_{\pm 7}$	$0.0156_{\pm 10}$	$0.0276_{\pm 23}$	$0.0783_{\pm 62}$
STIR $(\mathcal{I}_{\mathrm{lin}}, \gamma_\sigma)$, $\epsilon=\sigma^2/2$	$0.0190_{\pm 14}$	$0.0164_{\pm 11}$	$0.0097_{\pm 10}$	$0.0766_{\pm 44}$	$0.0157_{\pm 8}$	$0.0288_{\pm 7}$	$0.0768_{\pm 95}$
STIR $(\mathcal{I}_{\mathrm{lin}}, 0)$, same level	$0.0150_{\pm 25}$	$0.0133_{\pm 20}$	$0.0070_{\pm 5}$	$0.0594_{\pm 52}$	$0.0049_{\pm 6}$	$0.0183_{\pm 34}$	$0.0102_{\pm 48}$
DDPM (variance-preserving)							
native	$0.0184_{\pm 13}$	$0.0126_{\pm 5}$	$0.0062_{\pm 9}$	$0.0442_{\pm 27}$	$0.0121_{\pm 11}$	$0.0235_{\pm 9}$	$0.0823_{\pm 24}$
STIR $(\mathcal{I}_{\mathrm{vp}}, 0)$, $\epsilon=\omega_{\mathrm{vp}}$	$0.0184_{\pm 16}$	$0.0141_{\pm 13}$	$0.0064_{\pm 10}$	$0.0637_{\pm 43}$	$0.0128_{\pm 10}$	$0.0270_{\pm 7}$	$0.0997_{\pm 40}$
STIR $(\mathcal{I}_{\mathrm{lin}}, \gamma_\sigma)$, same level	$0.0222_{\pm 13}$	$0.0178_{\pm 12}$	$0.0093_{\pm 5}$	$0.0811_{\pm 70}$	$0.0158_{\pm 8}$	$0.0271_{\pm 14}$	$0.1005_{\pm 51}$
DDIM (variance-preserving)							
native	$0.0143_{\pm 6}$	$0.0101_{\pm 7}$	$0.0042_{\pm 2}$	$0.0330_{\pm 12}$	$0.0095_{\pm 4}$	$0.0202_{\pm 13}$	$0.0705_{\pm 49}$
STIR $(\mathcal{I}_{\mathrm{vp}}, 0)$, ϵ=0	$0.0135_{\pm 12}$	$0.0102_{\pm 7}$	$0.0039_{\pm 3}$	$0.0327_{\pm 27}$	$0.0097_{\pm 5}$	$0.0206_{\pm 9}$	$0.0679_{\pm 40}$
STIR $(\mathcal{I}_{\mathrm{lin}}, \gamma_\sigma)$, same level	$0.0232_{\pm 9}$	$0.0200_{\pm 4}$	$0.0110_{\pm 6}$	$0.0812_{\pm 35}$	$0.0188_{\pm 16}$	$0.0300_{\pm 12}$	$0.0919_{\pm 132}$
One level chosen at inference for the whole corpus							
STIR $(\mathcal{I}_{\mathrm{lin}}, \gamma_\sigma)$, best global ϵ	$\mathbf{0.0240}_{\pm 8}$	$\mathbf{0.0211}_{\pm 2}$	$\mathbf{0.0130}_{\pm 4}$	$\mathbf{0.0883}_{\pm 38}$	$\mathbf{0.0223}_{\pm 14}$	$\mathbf{0.0332}_{\pm 8}$	$\mathbf{0.1023}_{\pm 125}$
Inside published backbones (Yang et al., 2023; Liu et al., 2025)							
DreamRec, native	$0.0136_{\pm 14}$	$0.0134_{\pm 9}$	$0.0102_{\pm 3}$	$0.1091_{\pm 32}$	$0.0106_{\pm 3}$	$0.0120_{\pm 3}$	$0.0579_{\pm 8}$
+ STIR, best global ϵ	$\mathbf{0.0229}_{\pm 9}$	$\mathbf{0.0224}_{\pm 9}$	$\mathbf{0.0111}_{\pm 4}$	$\mathbf{0.1103}_{\pm 12}$	$\mathbf{0.0210}_{\pm 5}$	$\mathbf{0.0315}_{\pm 3}$	$\mathbf{0.1235}_{\pm 13}$
PreferDiff, native	$0.0213_{\pm 8}$	$0.0201_{\pm 11}$	$0.0105_{\pm 14}$	$0.1045_{\pm 154}$	$0.0054_{\pm 7}$	$0.0141_{\pm 8}$	$0.0679_{\pm 13}$
+ STIR, best global ϵ	$\mathbf{0.0217}_{\pm 4}$	$\mathbf{0.0282}_{\pm 7}$	$\mathbf{0.0115}_{\pm 3}$	$\mathbf{0.1071}_{\pm 23}$	$\mathbf{0.0210}_{\pm 4}$	$\mathbf{0.0311}_{\pm 3}$	$\mathbf{0.1224}_{\pm 10}$

(b) History length. *Under a Bayesian model in which interactions reveal the user's mode,* $\epsilon^\star(h) = O(e^{-c'T/2})$ *for a constant* $c' > 0$*, and* $\epsilon^\star(h) = 0$ exactly *past a finite* T.

(c) Shape of the trade. *The dispersion* $\mathrm{R}_{\hat{\mathcal{A}}_\epsilon}(h) = \mathrm{trCov}[X_1^\epsilon \mid h]$ *satisfies* $\mathrm{R}_{\hat{\mathcal{A}}_\epsilon}(h) = \mathrm{R}_{\hat{\mathcal{A}}_0}(h) + c(h)\,\epsilon\,\delta_s(h) + o(\delta_s)$ *with* $c(h) > 0$*, and the accuracy–diversity front is* concave.

From Prop. 7, a multi-modal user's velocity target points in different directions depending on the mode and a conditional mean can only average them, whereas the score accommodates several basins; exploration switches off past a finite history rather than annealing; and because the ϵ term in (c) carries δ_s, the level moves dispersion only through the score error.

5 EXPERIMENTS

We evaluate STIR on three questions: whether one trained model reproduces the four published processes, what the level changes once it is free and who needs it changed, and how to set a level per user and where that helps. Datasets, metrics and implementation: App. B.1.

5.1 RECONSTRUCTION OF THE NATIVE PROCESSES

Reconstruction Holds Within an Interpolant. A process is named by three things, $(\mathcal{I}, \gamma)$ fixed at training and ϵ set at inference, so a method is identified once all three are. We train one checkpoint per native interpolant, three in all: $(\mathcal{I}_{\mathrm{lin}}, 0)$, which is FM's own path; $(\mathcal{I}_{\mathrm{lin}}, \gamma_\sigma)$ at σ=0.5, which is SB's; and $(\mathcal{I}_{\mathrm{vp}}, 0)$, which is DDPM's and DDIM's. Each is then decoded at the *exact* level the framework derives for the method in question, with nothing fitted and no sweep, and we see if the native's numbers are recovered. In Table 2, the two linear checkpoints reproduce FM and SB inside five seed standard deviations on all seven corpora, and the VP one recovers native DDPM and DDIM.

Which Interpolant to Train On. DDPM and DDIM are defined on the variance-preserving path and FM and SB on the linear one, each welded to the interpolant its derivation assumes. STIR trains on whichever we supply, so the three checkpoints above are also three answers to *which path recommendation should use*, and Thm. 5 calls both halves of the choice. On the path $(\mathcal{I}_{\mathrm{lin}}, 0)$, the linear checkpoints at ϵ=0 beat the VP one at its best level on all seven corpora, by $1.0\times$ to $1.5\times$ in NDCG@20 (Table 2). The two families also disagree about noise as predicted: *within* VP accuracy rises with the level, whereas on the linear path ϵ=0 is best. Stochasticity there compensates for a path whose irreducible velocity loss diverges, and it does not recover what the path costs. On the envelope γ_σ, switching it off still matches at ϵ=0 on four corpora but cannot be read anywhere else,

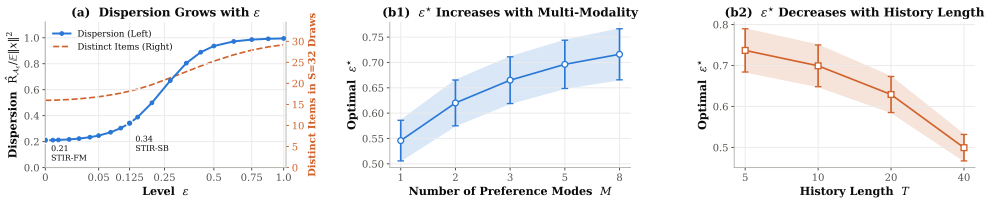


Figure 2: **What the level does, and who needs it. (a)** ML-1M, swept over ϵ: dispersion relative to $\mathbb{E}\|x\|^2$ and distinct items over $S{=}32$ draws, native levels marked. **(b1, b2)** Synthetic corpus (App. M): the optimum rises with modes M and falls with history T, as Prop. 7(a,b) predicts.

keeping 10–73% of its own $\epsilon{=}0$ accuracy at levels where the envelope-trained checkpoint keeps 82–96%: the envelope is what makes the level free, not what optimizes it. We therefore train on the linear path with the envelope on and leave ϵ to inference, the one commitment STIR cannot postpone.

The Choice Moves After Training. Choosing among the four is currently a training-time decision, since each is defined by its own objective and its own sampler, so comparing them means one training run per process. With STIR the commitment is only to an interpolant: the process is named at inference, so one set of weights answers what four separate runs used to answer. Picking which one to serve is itself cheap, since we sweep ϵ on a grid and score each value on a single batch of 512 validation users, keeping the best on NDCG@20 (the *best global* ϵ row of Table 2), so the whole selection is one batch of inference and costs no gradient step. Those weights also answer outside the four, since ϵ ranges over the whole half-line while the published methods occupy four points of it, so a level none of them offers is reachable without a further run.

Agnostic to the Backbone. Our goal is not to beat the published recommenders with a different architecture and training schemes. They all run one of the four processes, fixed once for the whole population, and they already differ from one another in training scheme, in architecture and in how they use the catalog, so a win obtained by changing any of those would say nothing about which process a given user should be served with. STIR changes the process alone, so it inherits whatever training scheme, architecture and objective a paper already uses, and the comparison becomes one between levels rather than between systems. As a sanity check, we therefore evaluated it on the DreamRec and PreferDiff backbones (Liu et al., 2025; Yang et al., 2023), keeping each one's architecture, training scheme and selecting one global ϵ on a batch of validation users exactly as above (last block of Table 2). In both backbones the level we select serves better than the one the backbone ships, on almost every corpus.

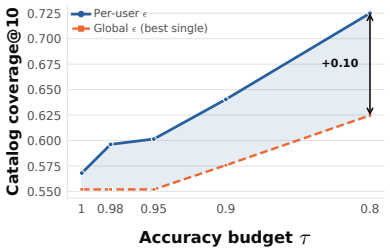


Figure 3: **More coverage at every budget.** On ML-1M a per-user ϵ beats every global ϵ at accuracy budget τ.

5.2 What the Level Buys, and Who Needs It

Dispersion Follows the Level. Sampling integrates Eq. (8) by Euler-Maruyama, c.f., Alg. 1. At $\epsilon{=}0$ the trajectory is a deterministic function of x_0 and all dispersion is source-propagated; raising ϵ opens the injected term of Def. 1 while the score contracts the trajectories. Prop. 7(c) predicts the total rises like $c(h)\,\epsilon\,\delta_s(h)$, first order in the score error and in no other quantity, along a concave frontier. Fig. 2(a) is the measurement of Fig. 1 run again with STIR in place of the four natives: same corpus, same $S{=}32$ sampler runs per user at frozen weights, one checkpoint instead of four. Sweeping ϵ traces a continuous curve through every value that figure reports. It starts at 0.21 of $\mathbb{E}\|x\|^2$ at $\epsilon{=}0$, which is FM's number exactly, passes 0.34 at SB's level against SB's 0.33, and carries on past DDPM's 0.72 and DDIM's 0.82 to 0.94 at the diffusion end. We measure the same effect a second way, in item space. Running the sampler $S{=}32$ times for one user therefore yields up to 32 such items, and we count how many of them are distinct. *A deployment can therefore ask for a dispersion no published method offers.* The four natives land on four values, whereas the curve is continuous in ϵ: any level we want between the two ends is reached by setting that one number.

Who Needs It: Modes and History. We verify empirically Prop. 7 (a,b), i.e., the level a user wants should rise with the spread of the modes they draw from and fall with the length of the history that resolves them. We therefore generate a synthetic corpus that assigns both, 20,000 users drawing histories from M of 20 item categories. For each user, we choose the number of categories M and the

history length T randomly and independently c.f., App. M. We then train STIR technique on it. We partition users by M and by T, and sweep ϵ per group to look for the optimal level ϵ. The group optimum $\epsilon^\star$ rises concavely from $\epsilon^\star \approx 0.55$ at $M{=}1$ to 0.72 at $M{=}8$, most of the climb falling between one and three modes, and it drops from 0.74 at $T{=}5$ to 0.50 at $T{=}40$ (Fig. 2b1, b2). This confirms the theory of Prop. 7, i.e., modes raise the irreducible velocity error δ_b because a conditional mean has to average targets pointing into different categories, and history lowers it by resolving which category the user is in. We also test the length effect outside the synthetic design: on four real dataset we hold one trained model fixed and truncate each test user's history at inference, so that length is the only quantity that changes, then group the test users by the resulting length and read the mean per-user $\epsilon^\star$ of each group, which leaves the fitted slope negative on all four corpora, c.f., App. N.

5.3 CHOOSING THE LEVEL, AND WHERE IT HELPS

Level Per User. We choose the level on validation data, and the only question is how many we choose: one for the corpus, one per user, or one per group. We form the groups from the model's own history embeddings, clustered by K-means into C of them, without using any label, and we use $C{=}5$ throughout. We then decode every user at each level on the grid, score the candidates on a batch of validation users, keep in each group the level that scores best on the ones it contains, and read each test user at its group's level. C sets the trade-off: raising C lets the level vary more finely and moves what is reachable toward the per-user oracle, but leaves fewer validation users behind each choice, until at one cluster per user a single held-out item decides and the choice no longer transfers.

Where It Helps. We distinguish two cases. The first is *coverage at no cost in accuracy*. We fix a budget τ, the fraction of the deterministic model's NDCG@20 that must be retained, and ask how much of the catalog a method reaches inside it. At $\tau{=}1.00$, where no accuracy may be given up, a level chosen per group already reaches further than any single global level, and it keeps that lead at every budget we test (Fig. 3), the cohorts that move being the cold ones as Prop. 7(b) requires. The budget is enforced on validation: for each method we sweep its own knob, discard every setting whose validation NDCG@20 falls below τ times the deterministic model's, and report the test coverage of the best setting that survives; the settings selected this way go on to satisfy the same constraint on test. Re-ranking cannot buy coverage this way. Our baselines are the standard post-hoc diversifiers, temperature sampling of the readout scores, maximal marginal relevance (Carbonell & Goldstein, 1998) and a determinantal re-rank (Chen et al., 2018), and all three stay within 0.005 of the deterministic sampler at $\tau{=}1.00$, since an operation that preserves the metric exactly can only permute inside the dead zone of Lemma 12 (App. L.1). We compare against them in App. L.

Table 3: **Predicting the next q items.** NDCG@q over 20 seeds.

STIR	$q{=}5$	$q{=}10$	$q{=}15$
ML-1M			
best global	$0.0720_{\pm 14}$	$0.1012_{\pm 15}$	$0.1155_{\pm 21}$
per user	$\mathbf{0.0816}_{\pm 29}$	$\mathbf{0.1129}_{\pm 43}$	$\mathbf{0.1248}_{\pm 33}$
oracle	$0.1194_{\pm 25}$	$0.1506_{\pm 13}$	$0.1595_{\pm 22}$
GOODREADS			
best global	$0.0741_{\pm 42}$	$0.0828_{\pm 14}$	$0.0770_{\pm 26}$
per user	$\mathbf{0.0806}_{\pm 42}$	$\mathbf{0.0892}_{\pm 13}$	$\mathbf{0.0813}_{\pm 39}$
oracle	$0.1130_{\pm 76}$	$0.1173_{\pm 6}$	$0.1081_{\pm 26}$

The second is *predicting the next q items rather than one*. Leave-one-out scores a whole list against a single held-out item, so two samplers that disagree about most of that list are recorded as identical whenever they agree on the one item scored. We therefore hold out the last q interactions, run the sampler q times, take at each draw the nearest item not already returned, and score the resulting set against the held-out set with NDCG@q. Only ML-1M and GOODREADS have histories long enough to surrender q interactions without emptying the conditioning. In Table 3 the level it chooses per group beats the global level in all six cells.

6 CONCLUSION

Generative recommenders imported their models from other domains like image generation, and with them a dispersion level each process fixes in advance. Recommendation cannot leave it fixed: how much spread a user needs depends on how many catalog categories their interests span and how much history resolves them. We showed that FM, SB, DDPM and DDIM are one SDE at four values of that level, that it enters neither training objective, and that it may vary with the user with every conditional marginal preserved, so one STIR checkpoint delivers all four and the continuum between them. The path and envelope remain training-time commitments; the level does not, and setting it per user pays, the headroom it reveals still open to whoever can estimate it without labels.

REPRODUCIBILITY STATEMENT

All modelling and training details needed to reproduce our results are in the Appendix : dataset construction and splits, the exact architecture and parameter counts, the interpolant, sampler and inference algorithms, the losses and optimizer. All theoretical claims state their assumptions explicitly in the main text and are proved in full in Appendix.

ETHICS STATEMENT

This work studies the generative process inside recommender systems and uses only public benchmark datasets (Amazon reviews, MovieLens-1M, Yelp) containing no personally identifying information beyond anonymized user identifiers, under their published terms of use. We introduce no new data collection and no human subjects.

AI USE STATEMENT

In this work, we used generative AI tools principally for paraphrasing and copy-editing of author-written prose, and to a lesser extent for debugging code. We have not used generative AI tools for research ideation, for the formulation or proof of any theoretical result, for experimental design, or for the analysis of results; dataset creation is not applicable, as we use only public benchmarks. We have reviewed all AI-assisted work: every paraphrased passage was read and revised by the authors, and AI-assisted code was verified by rerunning the affected stage and confirming that all reported metrics were unchanged.

REFERENCES

Michael S. Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *International Conference on Learning Representations (ICLR)*, 2023.

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 2025. arXiv:2303.08797.

Jaime Carbonell and Jade Goldstein. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 1998.

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*, pp. 2942–2951, 2020.

Laming Chen, Guoxin Zhang, and Eric Zhou. Fast greedy MAP inference for determinantal point process to improve recommendation diversity. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

Eunjoon Cho, Seth A. Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1082–1090, 2011.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, Dustin Podell, Tim Dockhorn, Zion English, Kyle Lacey, Alex Goodwin, Yannik Marek, and Robin Rombach. Scaling rectified flow transformers for high-resolution image synthesis. In *International Conference on Machine Learning (ICML)*, 2024.

Alessio Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *Journal of Functional Analysis*, 254(1):109–153, 2008.

F. Maxwell Harper and Joseph A. Konstan. The MovieLens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(4), 2015.

Ruining He and Julian McAuley. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th International Conference on World Wide Web (WWW)*, pp. 507–517, 2016.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Wang-Cheng Kang and Julian McAuley. Self-attentive sequential recommendation. In *IEEE International Conference on Data Mining (ICDM)*, 2018.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Zhifeng Kong, Wei Ping, Jiaji Huang, Kexin Zhao, and Bryan Catanzaro. DiffWave: A versatile diffusion model for audio synthesis. In *International Conference on Learning Representations (ICLR)*, 2021.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 2014.

Zihao Li, Aixin Sun, and Chenliang Li. DiffuRec: A diffusion model for sequential recommendation. *ACM Transactions on Information Systems (TOIS)*, 2023.

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 (ICLR)*, 2023.

Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos A Theodorou, Weili Nie, and Anima Anandkumar. I^2SB: Image-to-image Schrödinger bridge. In *ICML*, 2023a.

Shuo Liu, Zihan Liu, Yuanhao Wang, Jieming Zhang, and Jian Xu. Preference diffusion for recommendation. In *International Conference on Learning Representations (ICLR)*, 2025.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations (ICLR)*, 2023b.

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.

Shashank Rajput, Nikhil Mehta, Anima Singh, et al. Recommender systems with generative retrieval (TIGER). In *NeurIPS*, 2023.

Hannes Risken. *The Fokker–Planck Equation: Methods of Solution and Applications*. Springer, 2nd edition, 1996.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion schrödinger bridge matching. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations (ICLR)*, 2021a. arXiv:2010.02502.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations (ICLR)*, 2021b.

Harald Steck. Calibrated recommendations. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys)*, pp. 154–162, 2018.

Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *ACM International Conference on Information and Knowledge Management (CIKM)*, 2019.

Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Computation*, 23(7):1661–1674, 2011.

Joojo Walker, Ting Zhong, Fengli Zhang, Qiang Gao, and Fan Zhou. Recommendation via collaborative diffusion generative model (CoDiGEM). In *KSEM*, 2022.

Mengting Wan and Julian McAuley. Item recommendation on monotonic behavior chains. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 86–94, 2018.

Wenjie Wang, Yiyan Xu, Fuli Feng, Xinyu Lin, Xiangnan He, and Tat-Seng Chua. Diffusion recommender model. In *ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2023.

Yu Wang, Zhiwei Zhang, Hongru Liu, Kun Zhou, and Philip S. Yu. Conditional denoising diffusion for sequential recommendation. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2024.

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

Yelp. Yelp open dataset. `https://www.yelp.com/dataset`, 2023.

Jiaqi Zhai, Lucy Liao, Xing Liu, et al. Actions speak louder than words: Trillion-parameter sequential transducers for generative recommendations (HSTU). In *ICML*, 2024.

Yunqin Zhu, Chao Wang, Qi Zhang, and Hui Xiong. Graph signal diffusion model for collaborative filtering. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2024.

Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web (WWW)*, pp. 22–32, 2005.

Part A - STIR: Construction and Experimental Setup

A Related Work

Generative Recommendation. Sequential recommenders classically *score* a catalog, encoding the interaction history and taking an inner product against a learned item table (Kang & McAuley, 2018; Sun et al., 2019), an idea since scaled with Semantic Identifiers (SIDs) and long-context architectures (Rajput et al., 2023; Zhai et al., 2024). A more recent line instead *synthesizes* a target and retrieves its nearest catalog neighbours. These divide by what they generate: one group diffuses the next-item embedding (Yang et al., 2023; Li et al., 2023; Liu et al., 2025; Wang et al., 2024), another the user's whole interaction vector for collaborative filtering (Wang et al., 2023; Walker et al., 2022; Zhu et al., 2024), and Flow Matching (Lipman et al., 2023; Liu et al., 2023b) and Schrödinger Bridges (De Bortoli et al., 2021; Shi et al., 2023) supply the deterministic and intermediate alternatives to both. What these share matters more here than what separates them. They differ in objective and training scheme, in architecture, in how the catalog is used and in how the sampler is run, so their reported numbers are not directly comparable; yet every one of them builds on one of a handful of generative frameworks, Flow Matching, denoising diffusion or a Bridge, and inherits from that framework how much stochasticity its process injects. Across this literature that amount is fixed once, at design time, and the mechanism that fixes it differs by family: diffusion recommenders tie it to the noise schedule, so it cannot be varied without changing the corruption process; Flow Matching trains no score head, so positive stochasticity is unreachable from a released model; and Bridge networks emit the velocity and score as an inseparable sum, from which recovering either requires already knowing the other. Table 1 makes this per-family lock precise, and removing it is what STIR contributes.

Transport-Based Generative Modelling. A model trained to output a single point can only return a conditional mean, which is the right answer when the target is determined and the wrong one whenever several outputs are equally valid: the average of two plausible images is a blurred one, and the average of two plausible next items is an embedding that names neither. Wherever the target is genuinely under-determined, what is wanted is the conditional distribution itself, so the object to be learned is a law rather than a value. The family this paper works in learns that law from a simple observation: a distribution is easier to learn through the noise that destroys it than directly. Denoising score matching made that precise, fitting the gradient of the log-density from corrupted samples (Vincent, 2011), and denoising diffusion turned it into a generative recipe by corrupting data to Gaussian noise along a fixed schedule and learning to reverse it step by step (Ho et al., 2020). Song et al. (2021b) then showed that the corruption and its reversal are a pair of stochastic differential equations, which placed the construction in the language of transport and exposed a deterministic counterpart, the probability-flow ODE, sampled in practice as DDIM (Song et al., 2021a). The appeal is structural rather than aesthetic: training is a regression onto a target available in closed form, with no adversary and no likelihood, while sampling is the numerical integration of a field, so quality, cost and stochasticity are all settled after training rather than before it.

Flow Matching and rectified flow drop the corruption schedule and regress directly onto the velocity of a straight path between the two distributions (Lipman et al., 2023; Liu et al., 2023b), which shortens trajectories and has since carried to large-scale systems (Esser et al., 2024). Schrödinger Bridges (Léonard, 2014; De Bortoli et al., 2021; Shi et al., 2023; Liu et al., 2023a) pose the same problem as entropic optimal transport between two prescribed marginals and occupy the ground between the deterministic and the fully stochastic ends, their sampler carrying a diffusion coefficient that sets how much noise is injected on the way. Read side by side, the three differ in the path they take between the two distributions and in how much stochasticity that path admits, while sharing the same two objects underneath, a velocity and a score.

Diversity, exploration, and cold start. Accuracy-diversity control has been studied for two decades (Carbonell & Goldstein, 1998; Ziegler et al., 2005; Chen et al., 2018; Steck, 2018; Cen et al., 2020), almost always as post-hoc re-ranking applied to an already-generated candidate list, with no feedback into the generative process itself. The stochasticity level acts *inside* that process instead (Prop. 7(c)), so the two mechanisms are orthogonal and composable; we claim first-principles control rather than superiority over a tuned re-ranker. Separately, a large bandit literature treats exploration as a decision-theoretic problem whose intensity should decay as posterior uncertainty

resolves. Proposition 7(b) is the transport-model analogue of that principle, with the optimal generative stochasticity for a user decaying as their history concentrates the posterior, and switching off exactly past a finite history length. Exploration in the *generative process* rather than in the *policy* seems to us worth developing further, and we sketch it as future work rather than claiming it here.

B Experimental Setup, Datasets and Reproducibility

B.1 Datasets and Preprocessing

Seven public benchmarks, listed with their sources in Table 4. Each is core-filtered: users and items with fewer than a stated number of interactions are dropped. The threshold is 5 for the three Amazon corpora, 10 for Yelp and 20 for Gowalla and Goodreads; ML-1M is used as distributed.

Table 4: Datasets after filtering.

Dataset	#Users	#Items	#Actions	Train pairs	Avg. length	Source
Beauty	22,363	12,101	198,502	131,413	8.9	(McAuley et al., 2015)
Toys	19,412	11,924	167,597	109,361	8.6	(McAuley et al., 2015)
Sports	35,598	18,357	296,337	189,543	8.3	(He & McAuley, 2016)
ML-1M	6,040	3,706	1,000,209	982,089	165.6	(Harper & Konstan, 2015)
Yelp	51,678	24,879	1,324,207	1,169,173	25.6	(Yelp, 2023)
Gowalla	23,427	37,398	1,849,431	1,779,150	78.9	(Cho et al., 2011)
Goodreads	61,509	34,342	5,749,471	5,564,944	93.5	(Wan & McAuley, 2018)

Interactions are sorted by timestamp per user, and item identifiers are mapped to $1..N$ with 0 reserved for padding. Histories are padded or truncated to $L{=}50$. Writing $s = (s_1, \ldots, s_{|s|})$ for a user's sorted sequence and q for the number of targets held out at a time, the final $2q$ interactions are used for validation and test evaluations:

$$\underbrace{s_1,\ \ldots,\ s_{|s|-2q}}_{\text{training}} \qquad \underbrace{s_{|s|-2q+1},\ \ldots,\ s_{|s|-q}}_{\text{validation}} \qquad \underbrace{s_{|s|-q+1},\ \ldots,\ s_{|s|}}_{\text{test}}$$

Each target is predicted from the prefix preceding it: one training pair $(s_1, \ldots, s_{j-1}) \to s_j$ is emitted for every position j in the training block, so a user contributes many pairs rather than one. The first interaction has no prefix and is never a target, so a user with $|s|$ interactions yields $|s| - 2q - 1$ pairs, and the Train-pairs column of Table 4 is $\#\text{Actions} - 3\,\#\text{Users}$ at $q{=}1$. Validation and test targets never appear as training targets, and we keep users with at least $2q{+}1$ interactions.

Metrics. Every evaluation ranks the entire catalog for a user rather than a sampled negative set, and keeps the top K. Write $\mathcal{Q}$ for that user's set of q held-out items and $\mathrm{Top}_{\mathrm{K}}(h)$ for the returned list. The accuracy metrics is,

$$\mathrm{NDCG@K} = \frac{\sum_{i=1}^{\mathrm{K}} \mathbf{1}\{\text{rank } i \in \mathcal{Q}\} \,/\, \log_2(1+i)}{\sum_{i=1}^{\min(q,\mathrm{K})} 1 \,/\, \log_2(1+i)},$$

each averaged over users. NDCG@K counts how many of the q held-out items were retrieved at all, and additionally rewards where they landed, its denominator being the best placement q items could achieve inside the cut-off. At $q{=}1$ the pair reduces to $\mathbf{1}\{r \le \mathrm{K}\}$ and $1/\log_2(1+r)$ for a target at rank r, the usual leave-one-out form.

We also evaluate generative recommenders using Coverage@K. Given $\mathcal{U}$ the set of evaluated users and N for the catalog size, we define Coverage@K as follows,

$$\mathrm{Coverage@K} = \frac{1}{N}\,\Big|\bigcup_{h\in\mathcal{U}} \mathrm{Top}_{\mathrm{K}}(h)\Big|,$$

so an item counts once however many users receive it.

C Implementing Stir

C.1 Objectives, Targets and the Parameterization

Fix any affine interpolant of Definition 2 with a Gaussian source,

$$x_t = \alpha(t)\,x_0 + \beta(t)\,x_1 + \gamma(t)\,z, \qquad x_0 \sim \mathcal{N}(0, \varsigma_0^2 I), \quad z \sim \mathcal{N}(0, I),$$

the two independent of each other and of the target x_1, with $\varsigma_0{=}1$ throughout. The velocity head has a closed-form target defined as follows,

$$\dot{x}_t = \dot{\alpha}(t)\, x_0 + \dot{\beta}(t)\, x_1 + \dot{\gamma}(t)\, z,$$

The score, on the other hand, has no such closed form. We use Lemma 8 to empirically approximate the score function at a time t.

Lemma 8 (Proof in App. G.1)**.** *For the interpolant of Definition 2, we define the function ς by $\varsigma(t)^2 = \alpha(t)^2\varsigma_0^2 + \gamma(t)^2$, and the normalized residual $\xi_t = (x_t - \beta(t)x_1)/\varsigma(t)$, so that $\xi_t \sim \mathcal{N}(0, I_d)$ at every t. Then for every t with $\varsigma(t) > 0$, we have,*

$$s(t,x) \;=\; \nabla_x \log \rho(t,x) \;=\; -\frac{1}{\varsigma(t)}\,\mathbb{E}[\,\xi_t \mid x_t = x] \;=\; -\frac{1}{\varsigma(t)^2}\,\mathbb{E}[\,x_t - \beta(t)x_1 \mid x_t = x]\,. \quad (12)$$

If in addition $\gamma(t) > 0$, conditioning further on x_0 recovers the familiar bridge-noise form $s(t,x) = \mathbb{E}[-z/\gamma(t) \mid x_t = x]$.

For learning the score function, c.f., Eq. (12), during training, we approximate it empirically. Having drawn $x_0 \sim \mathcal{N}(0, \varsigma_0^2 I)$, $z \sim \mathcal{N}(0, I)$ and a target x_1 from a training pair, the trajectory point x_t and the residual ξ_t are both known in closed form,

$$\xi_t \;=\; \frac{x_t - \beta(t)x_1}{\varsigma(t)} \;=\; \frac{\alpha(t)\, x_0 + \gamma(t)\, z}{\varsigma(t)}, \qquad \varsigma(t)^2 = \alpha(t)^2\varsigma_0^2 + \gamma(t)^2, \quad (13)$$

so that $-\xi_t/\varsigma(t)$ is computable per sample. It is the score of the conditional law rather than of the marginal: since $x_t \mid x_1 \sim \mathcal{N}\big(\beta(t)x_1, \varsigma(t)^2 I_d\big)$,

$$\nabla_x \log \rho(t, x_t \mid x_1) \;=\; -\frac{x_t - \beta(t)x_1}{\varsigma(t)^2} \;=\; -\frac{\xi_t}{\varsigma(t)}. \quad (14)$$

The two objectives. We therefore minimize the two following loss functions,

$$\mathcal{L}_b = \mathbb{E}\big\|\hat{b}_\theta(t, x_t, h) - \dot{x}_t\big\|^2, \qquad \mathcal{L}_s = \mathbb{E}\big\|\hat{s}_\phi(t, x_t, h) - \frac{\xi_t}{\varsigma(t)}\big\|^2, \quad (15)$$

the expectations over $t \sim \mathcal{U}(0,1)$, the draws x_0, z. Neither contains ϵ.

The retrieval term. Regression fits the two fields but leaves the item table free: nothing in $\mathcal{L}_b$ forces the generated endpoint to be close to its target *in inner product*, which is what the readout ranks by. We therefore add to $\mathcal{L}_b$ a (sampled-)softmax cross-entropy that scores an endpoint estimate against the catalog,

$$\mathcal{L}_b \;\longleftarrow\; \mathcal{L}_b + \lambda\,\mathrm{CE}_b, \qquad \mathrm{CE}_b = -\log \frac{\exp\langle \hat{x}_1,\, e_{i^\star}\rangle}{\sum_{j\in\mathcal{S}} \exp\langle \hat{x}_1,\, e_j\rangle}, \qquad \hat{x}_1 = x_{t'} + (1 - t')\,\hat{b}_\theta,$$

where $\hat{x}_1$ is the endpoint the velocity field predicts from $x_{t'}$ at $t' \sim \mathcal{U}(0,1)$, e_j is the j-th row of the item table, $i^\star$ indexes the target x_1, and $\mathcal{S}$ is the candidate set: the target together with sampled negatives, or the whole catalog when it is small. Coupling a ranking term to a generative objective is standard in this literature (Liu et al., 2025; Wang et al., 2024).

Both networks b_θ and s_ϕ are trained with Adam (Kingma & Ba, 2015) at learning rate 10^{-3}, batch size 512, dropout 0.1 and 50 epochs per head, with retrieval weight $\lambda{=}1$; on large catalogs CE_b uses a sampled softmax over 1024 uniform negatives, and both the negatives and t are redrawn every minibatch.

C.2 Model Architecture

The velocity and score are *separate* networks of identical design, each mapping the current point x_t, the history h and the time t to a vector in $\mathbb{R}^d$ with $d{=}64$:

$$\hat{f}(t, x_t, h) \;=\; \mathrm{MLP}\big[\; x_t \;;\; \mathrm{Enc}(h) \;;\; \tau(t) \;\big] \;\in\; \mathbb{R}^d.$$

- Enc, the history encoder: item embeddings plus a learned positional embedding through two transformer layers in the SASRec manner (Kang & McAuley, 2018) (two attention heads, feed-forward width $4d$, GELU, dropout 0.1), read at the last non-padding position.
- τ, the time encoder: the scalar t through two linear layers of width d with a SiLU between them.
- MLP consists of three layers of width $4d$ with SiLU activations, over the concatenation $[\,x_t; \mathrm{Enc}(h); \tau(t)\,] \in \mathbb{R}^{3d}$.

Algorithm 1 STIR inference: generate $\hat{x}_1$ and rank items (one user)

Require: history encoding h; trained heads $\hat{b}_\theta, \hat{s}_\phi$; frozen item table $E \in \mathbb{R}^{N\times d}$; steps n_{step}; level schedule $\epsilon(\cdot) \geq 0$; cutoff K

$X \sim \mathcal{N}(0, I_d)$ ▷ source ρ_0 at t=0 (pure noise)
$\Delta t \leftarrow 1/n_{\text{step}}$
for $i = 0, \ldots, n_{\text{step}} - 1$ **do**
  $t \leftarrow (i + \frac{1}{2})\,\Delta t$ ▷ midpoint time, kept inside $(0, 1)$
  $\epsilon \leftarrow \epsilon(t)$ ▷ level at this step: schedule, constant or per user
  $\text{drift} \leftarrow \hat{b}_\theta(t, X, h)$ ▷ velocity field
  if $\epsilon > 0$ **then**
    $\text{drift} \leftarrow \text{drift} + \epsilon\, \hat{s}_\phi(t, X, h)$ ▷ score term
  end if
  $X \leftarrow X + \text{drift} \cdot \Delta t$ ▷ Euler step on the drift of Eq. (16)
  if $\epsilon > 0$ **then**
    $\xi \sim \mathcal{N}(0, I_d); \quad X \leftarrow X + \sqrt{2\epsilon\,\Delta t}\,\xi$ ▷ Euler-Maruyama noise
  end if
end for
$\hat{x}_1 \leftarrow X$ ▷ generated endpoint at the data end t=1
$\text{scores} \leftarrow \hat{x}_1 E^\top$
return top-K items of scores

Only the item table, i.e., the learned item embeddings, is shared between b_θ and s_ϕ .

Training Cost. We do not train two generative models. We train the two heads in sequence. We first fit $\hat{b}_\theta$ alone, i.e., we learn the history encoder, the time encoder and the item table in this stage, and we do nothing in it that is specific to STIR. We then initialize $\hat{s}_\phi$ from the converged $\hat{b}_\theta$, and we freeze the history encoder and the item table at the values the first stage reached, so we backpropagate only into the time encoder and the MLP. We therefore learn the catalog-sized item table, which dominates the parameter count at every scale we run, once rather than twice, which makes the training computationally very cheap.

C.3 THE UNIVERSAL SAMPLER

Sampling integrates the universal SDE on the interpolant of §C.1,

$$\mathrm{d}X = \big(\hat{b}_\theta(t, X, h) + \epsilon\, \hat{s}_\phi(t, X, h)\big)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W, \tag{16}$$

by Euler-Maruyama on a uniform grid of n_{step}=50 steps, $\Delta t = 1/n_{\text{step}}$, the fields read at interval midpoints $t_i = (i + \frac{1}{2})\Delta t$:

$$X_{i+1} = X_i + \big(\hat{b}_\theta(t_i, X_i, h) + \epsilon_i\, \hat{s}_\phi(t_i, X_i, h)\big)\Delta t + \sqrt{2\epsilon_i \Delta t}\,\xi_i, \qquad \xi_i \sim \mathcal{N}(0, I), \quad \epsilon_i = \epsilon(t_i, h),$$

from $X_0 \sim \mathcal{N}(0, I)$; the endpoint $\hat{x}_1 = X_{n_{\text{step}}}$ is scored against the frozen item table by inner product. We detail the sampling procedure in Algorithm 1.

D RECOVERING THE FOUR NATIVE PROCESSES

We previously claimed that four STIR can recover the different native generative four recommenders as long as we use the same interpolant $(\mathcal{I}, \gamma)$. To verify this, we need to ensure that the field STIR regresses must be the field the native method regresses, or a reparameterization of it. The *sampler* of Eq 6 should also recover the SDE or ODE of the native methods at one value of ϵ. Therefore, only $(\mathcal{I}, \gamma)$ and ϵ are free.

Recovery is *exact* when the native objective is a special case of Eq. (7), so the processes agree at any weights, and holds *at the population optimum* when they agree only once the heads attain $\hat{b}^*_\theta = b$ and $\hat{s}^*_\phi = s$.

D.1 Reconstruction of Flow Matching

The components of a FM-based generative recommender are:

- *Interpolant.* $x_t = (1-t)\,x_0 + t\,x_1$, with $\gamma \equiv 0$.
- *Objective.* $\mathcal{L}_{\mathrm{FM}} = \int_0^1 \mathbb{E}\big\|\hat{v}_\theta - (x_1 - x_0)\big\|^2 \mathrm{d}t$.
- *Sampler.* $\mathrm{d}X = \hat{v}_\theta(t, X, h)\,\mathrm{d}t$.

With $\gamma \equiv 0$ the velocity target is $\dot{x}_t = x_1 - x_0$, so $\mathcal{L}_{\mathrm{FM}}$ coincides with $\mathcal{L}_b$, and at $\epsilon{=}0$ the score term and the noise both drop out of unified SDE, leaving the ODE, which makes the recovery exact.

Therefore, STIR reproduces Flow Matching, but a native Flow-Matching checkpoint reproduces nothing else: no score head was ever fitted, so every $\epsilon > 0$ asks the sampler for a field that does not exist. The asymmetry is the whole point of the construction.

D.2 Reconstruction of the Schrödinger Bridge

The components of a SB-based generative recommender are:

- *Interpolant.* $x_t = (1-t)\,x_0 + t\,x_1 + \gamma_\sigma(t)\,z$, with $\gamma_\sigma(t) = \sigma\sqrt{t(1-t)}$.
- *Objective.* iterative fitting of a single field, converging to $v^{\mathrm{SB}} = b + \frac{\sigma^2}{2} s$.
- *Sampler.* $\mathrm{d}X = \hat{v}^{\mathrm{SB}}(t, X, h)\,\mathrm{d}t + \sigma\,\mathrm{d}W$.

The interpolant and envelope noise of SB are a special case of the stochastic interpolant by matching $\sqrt{2\epsilon} = \sigma$, which gives $\epsilon = \sigma^2/2$. However, SB never fits b and s separately; iterative proportional fitting converges to their weighted sum, so the identification is $\hat{b}^*_\theta + \frac{\sigma^2}{2}\hat{s}^*_\phi = v^{\mathrm{SB}}$ and can only be claimed at the population optimum, i.e., where both heads are optimally optimized.

D.3 Reconstruction of DDPM and DDIM

The Variance-Preserving path is normally written as a single Gaussian perturbation of the target,

$$x_t = \sqrt{\bar{\alpha}_t}x_1 + \sqrt{1-\bar{\alpha}_t}z, \qquad \bar{\alpha}_t = \exp\Big(-\int_t^1 g^2(u)\,\mathrm{d}u\Big),$$

with $\bar{\alpha}_1{=}1$ at the data end by construction. At the noise end $\bar{\alpha}_0 = \exp(-\int_0^1 g^2)$, so the source is close to pure noise only when the schedule accumulates enough corruption, $\int_0^1 g^2(u)\mathrm{d}u \gg 1$. In the affine family $x_t = \alpha(t)x_0 + \beta(t)x_1 + \gamma(t)z$ of §C.1 the source x_0 and the envelope draw z are both $\mathcal{N}(0, I)$, so only their combined scale $\varsigma(t)^2 = \alpha^2 + \gamma^2$ is identifiable. We may therefore put all of it in the source, taking $\gamma \equiv 0$ and $x_0 := z$, which makes the path follow the unified framework as follows:

$$\alpha_{\mathrm{vp}}(t) = \sqrt{1-\bar{\alpha}_t}, \qquad \beta_{\mathrm{vp}}(t) = \sqrt{\bar{\alpha}_t}, \qquad \gamma \equiv 0, \qquad x_0 := z, \qquad \text{so} \qquad \alpha^2_{\mathrm{vp}} + \beta^2_{\mathrm{vp}} = 1.$$

In the same three lines as above, the compoenents of a DDIM and DDPM based generative models are:

- *Interpolant.* $x_t = \alpha_{\mathrm{vp}}(t)\,x_0 + \beta_{\mathrm{vp}}(t)\,x_1$, with $\gamma \equiv 0$ and $x_0 := z$.
- *Objective.* $\mathcal{L}_{\mathrm{DDPM}} = \int_0^1 \mathbb{E}\big\|\hat{z}_\theta(t, x_t, h) - z\big\|^2 \mathrm{d}t$.
- *Samplers.* Both are built from that same $\hat{z}_\theta$ and no second field. Writing $\beta = 2\omega_{\mathrm{vp}}$ for the variance schedule, in the source-to-data direction used throughout,

$$\text{DDPM}\ \ \mathrm{d}X = \Big[\tfrac{\beta}{2}X - \tfrac{\beta}{\alpha_{\mathrm{vp}}}\hat{z}_\theta\Big]\mathrm{d}t + \sqrt{\beta}\,\mathrm{d}W, \qquad \text{DDIM}\ \ \mathrm{d}X = \Big[\tfrac{\beta}{2}X - \tfrac{\beta}{2\alpha_{\mathrm{vp}}}\hat{z}_\theta\Big]\mathrm{d}t.$$

With $\gamma \equiv 0$ and $\varsigma_0{=}1$, the total scale is $\varsigma(t) = \alpha_{\mathrm{vp}}(t)$, so the residual of Lemma 8 collapses to $\xi_t = (x_t - \beta_{\mathrm{vp}}x_1)/\alpha_{\mathrm{vp}} = z$: on this path the latent *is* the noise DDPM predicts. Our $\mathcal{L}_s$ regresses $\hat{s}_\phi$ against $\xi_t/\varsigma(t)$, and setting $\hat{s}_\phi = -\hat{z}_\theta/\alpha_{\mathrm{vp}}$ turns it into $\int_0^1 \alpha_{\mathrm{vp}}(t)^{-2}\,\mathbb{E}\|\hat{z}_\theta - z\|^2\mathrm{d}t$, which is $\mathcal{L}_{\mathrm{DDPM}}$ under a weight that depends on t alone. The two therefore have the same population minimizer, and recovery of the objective is exact.

The sampler check needs one more step, because on this path the velocity is *not* an independent field. Write $\omega_{\rm vp} := \dot\beta_{\rm vp}/\beta_{\rm vp}$. Eliminating the target from $\dot x_t = \dot\alpha_{\rm vp} z + \dot\beta_{\rm vp} x_1$ using $x_1 = (x_t - \alpha_{\rm vp} z)/\beta_{\rm vp}$ leaves the state and the latent,

$$\dot x_t = \omega_{\rm vp}\, x_t + \Big(\dot\alpha_{\rm vp} - \omega_{\rm vp}\,\alpha_{\rm vp}\Big) z,$$

and taking conditional expectations, with $\mathbb{E}[z \mid x_t = x] = -\alpha_{\rm vp} s$ from Lemma 8 at $\xi_t = z$,

$$\begin{aligned} b(t,x) = \mathbb{E}\big[\dot x_t \mid x_t = x\big] &= \omega_{\rm vp}\, x + \big(\dot\alpha_{\rm vp} - \omega_{\rm vp}\,\alpha_{\rm vp}\big)\,\mathbb{E}\big[z \mid x_t = x\big] \\ &= \omega_{\rm vp}\, x - \alpha_{\rm vp}\big(\dot\alpha_{\rm vp} - \omega_{\rm vp}\,\alpha_{\rm vp}\big) s \\ &= \omega_{\rm vp}\, x + \omega_{\rm vp}\big(\alpha_{\rm vp}^2 + \beta_{\rm vp}^2\big) s = \omega_{\rm vp}\,(x + s), \end{aligned}$$

the middle step using $\alpha_{\rm vp}\dot\alpha_{\rm vp} = -\beta_{\rm vp}\dot\beta_{\rm vp}$ and the last $\alpha_{\rm vp}^2 + \beta_{\rm vp}^2 = 1$. The state is tied to the velocity: b is an affine readout of x and s, so there is no second field to learn, which is why DDPM trains one network and why the two-head form has nothing extra to fit on VP. Substituting $b = \omega_{\rm vp}(x + s)$ into the universal sampler leaves a single free parameter,

$$\mathrm{d}X = \big(b + \epsilon\, s\big)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W = \big(\omega_{\rm vp} X + (\omega_{\rm vp} + \epsilon)\, s\big)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W,$$

and the two VP methods are two of its values. Writing $\beta = 2\omega_{\rm vp}$ for the variance schedule and substituting $s = -\hat z_\theta/\alpha_{\rm vp}$,

$$\epsilon = 0 \;\;\text{(DDIM)}: \quad \mathrm{d}X = \omega_{\rm vp}\,(X + s)\,\mathrm{d}t = \Big[\tfrac{\beta}{2} X - \tfrac{\beta}{2\alpha_{\rm vp}}\hat z_\theta\Big]\mathrm{d}t,$$

$$\epsilon = \omega_{\rm vp} \;\;\text{(DDPM)}: \quad \mathrm{d}X = \omega_{\rm vp}\,(X + 2s)\,\mathrm{d}t + \sqrt{2\omega_{\rm vp}}\,\mathrm{d}W = \Big[\tfrac{\beta}{2} X - \tfrac{\beta}{\alpha_{\rm vp}}\hat z_\theta\Big]\mathrm{d}t + \sqrt{\beta}\,\mathrm{d}W,$$

which are the two initial samplers of DDIM and DDPM.

E Why the Native Methods Cannot Move Their Level

App. D showed that each native method *is* the universal sampler

$$\mathrm{d}X = \big(b + \epsilon\, s\big)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W$$

at one value of ϵ. This appendix asks the converse: given a released checkpoint of one of those methods, can ϵ be changed after training? For all four the answer is no, and the obstruction is different in each case, which is why no single patch would free them.

E.1 Flow Matching: the Second Field Was Never Fitted

Running the sampler at $\epsilon > 0$ requires both fields,

$$\mathrm{d}X = \big(\hat b_\theta + \epsilon\, \hat s_\phi\big)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W,$$

whereas Flow Matching trains one head against one target,

$$\mathcal{L}_{\rm FM}[\hat v] = \int_0^1 \mathbb{E}\big\|\hat v - (x_1 - x_0)\big\|^2 \mathrm{d}t, \qquad \hat v^* = \mathbb{E}\big[x_1 - x_0 \mid x_t, h\big] = b,$$

so the released weights supply b and nothing else. The score is not independent of it: §E.5 recalls that the two determine each other on any affine interpolant, and on this path, where $x_t = (1-t)x_0 + t x_1$ and $\varsigma(t) = (1-t)\varsigma_0$, the inversion is explicit,

$$s(t,x) = \frac{t\, b(t,x) - x}{1 - t}.$$

The obstruction is therefore conditioning rather than information. A velocity error δ enters the derived score as $\frac{t}{1-t}\,\delta$, which is 9δ at t=0.9 and 49δ at t=0.98: the amplification diverges exactly at the data end, where the sampler finishes and where the score term it needs is largest. A released Flow-Matching checkpoint thus has no usable score, and recovering one means training the second head. Nothing in the interpolant prevents that: $\varsigma(t) = (1-t)\varsigma_0 > 0$ for $t < 1$, so Lemma 8 applies on Flow Matching's own path and the head can be fitted without touching the velocity loss.

E.2 SCHRÖDINGER BRIDGE: THE TWO FIELDS ARRIVE FUSED

The Bridge trains a single drift whose population optimum is the weighted sum

$$\hat{v}^{\mathrm{SB},*} = b + \tfrac{\sigma^2}{2}\, s,$$

one vector per (t, x). The universal sampler at a different level needs the two terms apart,

$$b + \epsilon'\, s \;=\; \underbrace{\hat{v}^{\mathrm{SB}}}_{\text{known}} + \Big(\epsilon' - \tfrac{\sigma^2}{2}\Big)\, \underbrace{s}_{\text{unknown}}, \qquad \text{equivalently} \qquad s = \frac{2}{\sigma^2}\big(\hat{v}^{\mathrm{SB}} - b\big),$$

so recovering s presupposes b and recovering b presupposes s: one equation, two unknowns at every (t, x). Solving for either from the sum presupposes the other, and no post-processing of a sum recovers its terms. The fusion that makes the Bridge trainable is the same fact that welds its level.

E.3 DDPM AND DDIM: THE LEVEL IS THE SCHEDULE

On the variance-preserving path the level is not a free scalar at all. §D.3 derives $\epsilon = \omega_{\mathrm{vp}} = \dot{\beta}_{\mathrm{vp}}/\beta_{\mathrm{vp}}$, and that relation inverts:

$$\epsilon(t) = \frac{\dot{\beta}_{\mathrm{vp}}(t)}{\beta_{\mathrm{vp}}(t)} \qquad \Longleftrightarrow \qquad \beta_{\mathrm{vp}}(t) = \exp\Big(-\int_t^1 \epsilon(u)\, \mathrm{d}u\Big).$$

Choosing a level therefore *is* choosing a corruption schedule, hence an interpolant, hence a training distribution: a different ϵ is a different model, not a different decode. A second lock sits underneath. The noise predictor is calibrated pointwise in t: its target is

$$\xi_t = \frac{x_t - \beta_{\mathrm{vp}}(t)\, x_1}{\alpha_{\mathrm{vp}}(t)} = z, \qquad x_t \sim \mathcal{N}\big(\beta_{\mathrm{vp}}(t) x_1,\ \alpha_{\mathrm{vp}}(t)^2 I_d\big),$$

so $\hat{z}_\theta(t, \cdot)$ is fitted on states of scale $\alpha_{\mathrm{vp}}(t)$; running the sampler at another level changes the law of the states reaching time t, so $\hat{z}_\theta(t, \cdot)$ is evaluated away from the marginals it was trained on. The η knob is the one movement this family does allow, and it shows the limit: it interpolates between the probability-flow ODE and the ancestral SDE *at the same schedule*, spanning $\epsilon \in [0, \omega_{\mathrm{vp}}]$ with ω_{vp} fixed, and it cannot be made to depend on the user without giving each user a different β_{vp} and so a different trained model.

E.4 TWO SUBSTITUTES THAT DO NOT WORK

Faced with a locked checkpoint, there are two obvious things to try instead of a score-corrected level. Neither delivers what ϵ delivers, and the reasons are worth stating because both look reasonable.

Injecting noise without the correction. The first is to keep the trained velocity and simply add Brownian motion at sampling time, $\mathrm{d}X = \hat{b}_\theta\, \mathrm{d}t + \sqrt{2\epsilon}\, \mathrm{d}W$. This does produce dispersion, but not from the target. The density of the interpolant obeys the continuity equation $\partial_t \rho_t = -\nabla \cdot (b \rho_t)$, whereas that sampler evolves

$$\partial_t \hat{\rho}_t = -\nabla \cdot (b\, \hat{\rho}_t) + \epsilon\, \Delta \hat{\rho}_t,$$

an uncompensated heat term. Theorem 3 works precisely because the drift's $\epsilon\, s$ term contributes $-\epsilon \nabla \cdot (s \rho_t)$, which Step 2 of its proof shows equals $\epsilon \Delta \rho_t$ and cancels it; drop the score and nothing cancels. To leading order the terminal law is the target convolved with $\mathcal{N}\big(0, 2\int_0^1 \epsilon\, \mathrm{d}t\, I_d\big)$: an isotropic blur applied to every user alike, which is not the same object as a user visiting a different mode of their own conditional. The sampler is no longer drawing from $p(\cdot \mid h)$ at all, so the dispersion it reports is a property of the corruption, not of the recommendation.

Rescaling the spread instead. The second is to leave the sampler deterministic and rescale, either the source, $x_0 \sim \mathcal{N}(0, \tau^2 I_d)$, or the generated endpoint, $\hat{x}_1 \mapsto \tau\, \hat{x}_1$. Scaling the source breaks the interpolant's boundary condition, so $\hat{b}_\theta$ is evaluated on states it was never fitted on. Scaling the endpoint is worse: it does nothing at all. Item embeddings are L^2-normalized and scoring is by inner product (§C.2), so for every $\tau > 0$

$$\langle \tau\, \hat{x}_1,\, e_j \rangle = \tau\, \langle \hat{x}_1,\, e_j \rangle \quad \text{for all } j,$$

which is an order-preserving transformation of the scores: the Top-K list is *exactly* unchanged, for any τ, for every user. In the language of Def. 1 both substitutes act on the wrong term. Rescaling can

only inflate what the deterministic map already produces, leaving the injected component at zero, while uncorrected noise inflates the total without keeping the law it is supposed to be sampling.

E.5 What Changes Under Stir

The three obstructions are answered by one design. Both fields are emitted separately rather than fused, so the Bridge's decomposition problem does not arise; the score head is trained on any interpolant with $\varsigma(t) > 0$, so Flow Matching's missing field is not missing; and ϵ appears in neither loss of Eq. (15), so it is not a property of the trained weights at all. What remains is a conditioning question rather than an existence one, and Thm. 5(b) settles it: the envelope amplitude controls how well the score behaves at the data end, which is why the envelope-free checkpoint of Table 2 still matches at $\epsilon{=}0$ but degrades as soon as the level is raised.

F Relation to Stochastic Interpolants

Definition 2 is the stochastic interpolant of Albergo & Vanden-Eijnden (2023); Albergo et al. (2025), taken verbatim, and so is the fact that a one-parameter family of SDEs sharing a pair of fields transports the same marginals. This section we highlight and show the difference between both directions.

Fix an interpolant and write ρ_t for the law of x_t, $b(t,x) = \mathbb{E}[\dot{x}_t \mid x_t{=}x]$ and $s(t,x) = \nabla_x \log \rho_t(x)$. Albergo et al. (2025) show that for any schedule $\epsilon(t) \geq 0$ the SDE

$$\mathrm{d}X_t = \big(b(t,X_t) + \epsilon(t)\, s(t,X_t)\big)\mathrm{d}t + \sqrt{2\epsilon(t)}\,\mathrm{d}W_t, \qquad X_0 \sim \rho_0, \tag{17}$$

has marginals ρ_t for every t. The argument is the Fokker–Planck cancellation reproduced in App. H.1: the score term contributes $-\epsilon\,\nabla\cdot(s\rho_t)$, the identity $\nabla\cdot(s\rho_t) = \Delta\rho_t$ turns it into the diffusion term with the opposite sign, and uniqueness of weak solutions does the rest. The difference is which set the level is drawn from. Writing $\mathcal{H}$ for the space of histories, theirs and ours are

$$\mathcal{E}_{\mathrm{pop}} = \big\{\epsilon : [0,1] \to \mathbb{R}_{\geq 0}\big\}, \qquad \mathcal{E}_{\mathrm{usr}} = \big\{\epsilon : [0,1]\times\mathcal{H} \to \mathbb{R}_{\geq 0} \text{ measurable}\big\}, \qquad \mathcal{E}_{\mathrm{pop}} \subsetneq \mathcal{E}_{\mathrm{usr}}, \tag{18}$$

the inclusion sending a schedule to the map constant in h. Every statement below is about that strict inclusion.

Conditioning enters every object as a *parameter* and never as a variable of the PDE. For each h the conditional interpolant has marginals $\rho(t,\cdot \mid h)$ and fields $b(t,x,h),\ s(t,x,h)$, and the equation that must hold is

$$\partial_t \rho(t,x \mid h) = -\nabla_x\cdot\Big(\big(b(t,x,h) + \epsilon(t,h)\, s(t,x,h)\big)\rho(t,x \mid h)\Big) + \epsilon(t,h)\,\Delta_x \rho(t,x \mid h), \tag{19}$$

a PDE in x and t alone in which h appears only as a label. Fixing h therefore reduces Theorem 3 to Eq. (17) with the schedule $\epsilon(\cdot,h)$, and every step goes through unchanged: the continuity equation, the rewriting $\nabla\cdot(s\rho) = \Delta\rho$, the add-and-subtract, and the uniqueness appeal are all statements about one conditional law. Two users are two equations that never interact, so no coupling between them can spoil either, and measurability in h is needed only so that the induced population law is well defined.

Marginal preservation is an endpoint statement, not a statement about the process. Writing $\mathbb{P}^{\epsilon}_h$ for the law of the solution on path space $C([0,1],\mathbb{R}^d)$,

$$\big(\mathbb{P}^{\epsilon}_h\big)_{t=1} = p(\cdot \mid h) \ \text{ for every } \epsilon \in \mathcal{E}_{\mathrm{usr}}, \qquad \text{while} \qquad \mathbb{P}^{\epsilon}_h \neq \mathbb{P}^{\epsilon'}_h \ \text{ whenever } \epsilon \not\equiv \epsilon', \tag{20}$$

so the family is a set of genuinely different processes with one common endpoint. Under $\mathcal{E}_{\mathrm{usr}}$ the population terminal law $\int p(\cdot \mid h)\,\mathrm{d}\mu(h)$ is preserved as well, for any assignment of levels to users, which is what licenses serving two users two different processes from one checkpoint.

Allowing the level to depend on the state as well, $\epsilon = \epsilon(t,x)$, breaks it. The diffusion coefficient is then state dependent and the Fokker–Planck term appears in divergence form,

$$\partial_t \hat{\rho}_t = -\nabla\cdot\big((b+\epsilon s)\hat{\rho}_t\big) + \Delta\big(\epsilon\hat{\rho}_t\big), \qquad \Delta(\epsilon\rho) = \underbrace{\nabla\cdot(\epsilon\nabla\rho)}_{\text{cancels as before}} + \underbrace{\nabla\cdot(\rho\,\nabla_x\epsilon)}_{\text{residual}}, \tag{21}$$

and preservation is recovered only by putting the Itô correction back into the drift, $\mathrm{d}X_t = (b + \epsilon s + \nabla_x\epsilon)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W_t$, which requires differentiating the level with respect to the state at every step.

We use $\epsilon(t, h)$, constant in x for each user, so $\nabla_x \epsilon \equiv 0$ and the residual vanishes identically. The boundary names what the conditioning exploits: h is fixed along a trajectory while x is not.

Equations (17)-(19) hold for the *exact* fields. A deployed sampler runs on $\hat{b}_\theta, \hat{s}_\phi$, and subtracting the two Fokker–Planck equations isolates what the substitution costs. With $\hat{\rho}$ the law the learned sampler delivers,

$$\partial_t(\hat{\rho} - \rho) \;=\; -\nabla\cdot\Big(\underbrace{\big[(\hat{b}_\theta - b) + \epsilon\,(\hat{s}_\phi - s)\big]}_{\text{forcing}}\,\hat{\rho}\Big) \;-\; \nabla\cdot\big((b + \epsilon s)(\hat{\rho} - \rho)\big) \;+\; \epsilon\,\Delta(\hat{\rho} - \rho), \quad (22)$$

which is the display the rest of the paper turns on. The velocity error enters the forcing once, the score error enters it multiplied by ϵ, and the two are the quantities $\delta_b(h)$ and $\delta_s(h)$ of Eq. (10). Under exact fields the forcing vanishes and every ϵ delivers the same law, which is why the level is free in Theorem 3; under learned fields it does not vanish, its ϵ-dependence is first order in the score error, and the level becomes a quantity with a best value rather than an arbitrary one. That is the gap §4 fills.

Taken alone, the step from $\mathcal{E}_{\text{pop}}$ to $\mathcal{E}_{\text{usr}}$ is a small theorem and we do not present it as more. What it enables is not. Under Eq. (17) the level is a property of the sampler a paper ships, so which value to ship is decided once for a population; the extension turns it into a question that can be asked per user, Eq. (22) makes that question well posed, §4 answers it from a model of the recommendation target, and §5 measures that the answer differs across users and corpora.

Part B - Proofs and Extended Theory

G Proofs of the Lemmas

G.1 Proof of Lemma 8

Proof. We group all the Gaussian randomness of x_t into a single term,

$$x_t \;=\; \beta(t)\, x_1 + w_t, \qquad \text{with} \qquad w_t := \alpha(t)\, x_0 + \gamma(t)\, z.$$

Since x_0 and z are independent zero-mean Gaussians and α, γ are deterministic, $w_t \sim \mathcal{N}(0, \varsigma(t)^2 I_d)$ with $\varsigma(t)^2 = \alpha(t)^2 \varsigma_0^2 + \gamma(t)^2$, and since x_1 is independent of both, w_t is independent of x_1. Hence, conditionally on the target,

$$x_t \mid x_1 \;\sim\; \mathcal{N}\big(\beta(t)\, x_1,\; \varsigma(t)^2 I_d\big). \tag{23}$$

Therefore, x_t is a Gaussian-smoothed copy of the deterministic image $\beta(t)x_1$ of the target, at smoothing scale $\varsigma(t)$. Here $\beta(t)x_1$ is itself random, being a function of the random target, which is exactly why the marginal below is a mixture rather than a single Gaussian.

We integrate Eq. (23) over $x_1 \sim \rho_1$ as follows,

$$\rho(t,x) \;=\; \int \rho_1(x_1)\, \mathcal{N}\big(x;\; \beta(t)x_1,\; \varsigma(t)^2 I_d\big)\, dx_1.$$

For $\varsigma(t) > 0$ the Gaussian kernel is smooth with derivative $\nabla_x \mathcal{N}(x; \mu, \varsigma^2 I_d) = -\mathcal{N}(x; \mu, \varsigma^2 I_d)\,(x - \mu)/\varsigma^2$, and the interchange is justified by dominated convergence under $\mathbb{E}\|x_1\|^2 < \infty$. Hence

$$\nabla_x \rho(t,x) \;=\; -\int \rho_1(x_1)\, \mathcal{N}\big(x; \beta(t)x_1, \varsigma^2 I_d\big)\, \frac{x - \beta(t)x_1}{\varsigma(t)^2}\, dx_1.$$

By Bayes' rule the ratio $\rho_1(x_1)\mathcal{N}(x; \beta(t)x_1, \varsigma^2 I_d)/\rho(t,x)$ is the posterior density $p(x_1 \mid x_t = x)$, so

$$\nabla \log \rho(t,x) \;=\; \frac{\nabla_x \rho(t,x)}{\rho(t,x)} \;=\; -\,\mathbb{E}\left[\frac{x_t - \beta(t)x_1}{\varsigma(t)^2} \;\middle|\; x_t = x\right],$$

which is Eq. (12) after substituting $x_t - \beta(t)x_1 = \varsigma(t)\, \xi_t$.

□

H Proofs of the Theorems

H.1 Proof of Theorem 3

Proof. We denote by ρ_t the marginal density of the stochastic interpolant defined by,

$$x_t = \mathcal{I}(t, x_0, x_1) + \gamma(t)\, z.$$

We prove that ρ_t satisfies the Fokker–Planck equation associated with the following SDE

$$\mathrm{d}X_t = \big(b(t, X_t, h) + \epsilon(t,h)\, s(t, X_t, h)\big)\mathrm{d}t + \sqrt{2\epsilon(t,h)}\, \mathrm{d}W_t, \qquad X_0 \sim \rho_0,$$

for *any* $\epsilon \geq 0$. Since the SDE solution X_t starts from the same initial condition ($X_0 \sim \rho_0$), uniqueness of Fokker–Planck solutions then implies that X_t has marginal ρ_t for all t.

The Continuity Equation for ρ_t.

The velocity field is defined as $b(t,x) := \mathbb{E}[\dot{x}_t \mid x_t = x]$. For any smooth, compactly supported test function $\varphi : \mathbb{R}^d \to \mathbb{R}$, we compute:

$$\frac{d}{\mathrm{d}t}\mathbb{E}[\varphi(x_t)] \;=\; \mathbb{E}[\nabla\varphi(x_t)\cdot \dot{x}_t] \;=\; \mathbb{E}\big[\nabla\varphi(x_t)\cdot b(t,x_t)\big] \;=\; \int \nabla\varphi(x)\cdot b(t,x)\, \rho_t(x)\, dx,$$

where the second equality uses the tower property of conditional expectation:

$$\mathbb{E}[\nabla\varphi(x_t)\cdot \dot{x}_t] = \mathbb{E}[\nabla\varphi(x_t)\cdot \mathbb{E}[\dot{x}_t \mid x_t]] = \mathbb{E}[\nabla\varphi(x_t)\cdot b(t,x_t)].$$

Integrating by parts on the right-hand side (moving ∇ from φ to $b\,\rho_t$), we have,

$$\tfrac{d}{\mathrm{d}t}\int \varphi(x)\,\rho_t(x)\,dx \;=\; -\int \varphi(x)\,\nabla\cdot(b\,\rho_t)(x)\,dx.$$

Since this holds for all test functions φ, we conclude that ρ_t satisfies the *continuity equation*:

$$\partial_t\rho_t + \nabla\cdot(b\,\rho_t) = 0. \tag{24}$$

Rewriting the Laplacian Using the Score.

By definition $s(t,x) = \nabla \log \rho_t(x)$ (Eq. (5); Lemma 8 gives its denoising form), which is equivalent to:

$$\nabla\rho_t(x) \;=\; \rho_t(x)\,s(t,x). \tag{25}$$

This is the standard identity $\nabla f = f\,\nabla \log f$ for any positive function f.

Using Eq. (25), we can rewrite the Laplacian of ρ_t as a divergence involving the score:

$$\begin{aligned}\Delta\rho_t &= \nabla\cdot(\nabla\rho_t)\\ &= \nabla\cdot\big(\rho_t\,\nabla\log\rho_t\big)\\ &= \nabla\cdot(\rho_t\,s).\end{aligned}$$

The Fokker–Planck Equation of the SDE.

The SDE Eq. (6) has drift $\mu(t,x) = b(t,x) + \epsilon\,s(t,x)$ and diffusion coefficient $\sigma = \sqrt{2\epsilon}$. The Fokker–Planck (forward Kolmogorov) equation governing the evolution of the marginal density $\hat{\rho}_t$ of the SDE solution X_t is (Risken, 1996):

$$\partial_t\hat{\rho}_t \;=\; -\nabla\cdot\big((b+\epsilon\,s)\,\hat{\rho}_t\big) + \epsilon\,\Delta\hat{\rho}_t. \tag{26}$$

Verify That ρ_t Satisfies the Fokker–Planck Equation.

We now show that the interpolant marginal ρ_t satisfies Eq. (26). Starting from the continuity Eq. (24):

$$\begin{aligned}\partial_t\rho_t &= -\nabla\cdot(b\,\rho_t)\\ &= -\nabla\cdot(b\,\rho_t) - \epsilon\,\nabla\cdot(s\,\rho_t) + \epsilon\,\nabla\cdot(s\,\rho_t) \qquad \text{(add and subtract)}\\ &= -\nabla\cdot\big((b+\epsilon\,s)\,\rho_t\big) + \epsilon\,\nabla\cdot(s\,\rho_t)\\ &= -\nabla\cdot\big((b+\epsilon\,s)\,\rho_t\big) + \epsilon\,\Delta\rho_t. \end{aligned}\tag{27}$$

The last equality uses the score rewriting above: $\nabla\cdot(s\,\rho_t) = \nabla\cdot(\rho_t\,\nabla\log\rho_t) = \nabla\cdot(\nabla\rho_t) = \Delta\rho_t$.

Eq. (27) is exactly the Fokker–Planck Eq. (26) with $\hat{\rho}_t$ replaced by ρ_t.

Conclude by Uniqueness.

We have established:

- ρ_t satisfies the Fokker–Planck Eq. (26) (verified above).
- $\hat{\rho}_t$ (the marginal of the SDE solution X_t) also satisfies Eq. (26) (standard SDE theory).
- Both share the same initial condition: $\hat{\rho}_0 = \rho_0$ (the SDE starts at $X_0 \sim \rho_0$).

Under standard regularity conditions (finite second moments and locally Lipschitz drift), the Fokker–Planck equation admits a *unique* weak solution (Figalli, 2008). Therefore: $\hat{\rho}_t = \rho_t$ for all $t \in [0,1]$. In particular, $X_1 \sim \rho_1 = p(\cdot \mid h)$, regardless of the value of ϵ. □

H.2 Proof of Theorem 5 (a)

Proof. We recall that $\mathcal{I}(t,x_0,x_1) = \alpha(t)x_0 + \beta(t)x_1$ is a C^1 affine path with $\gamma \equiv 0$ and $\alpha(0) = \beta(1) = 1$, $\alpha(1) = \beta(0) = 0$; the source is $x_0 \sim \mathcal{N}(0, \varsigma_0^2 I_d)$, independent of a target obeying Assumption 4 with common within-mode spread ς_1.

Using the Assumption 4, sampling x_1 is equivalent to sampling $m \in \{1, \dots, M\}$ with $\Pr[m = j \mid h] = \pi_j(h)$ and then $x_1 \mid m \sim \mathcal{N}(\mu_m, \varsigma_1^2 I_d)$.

So m is a random variable, the mode this particular target came from, μ_m is the random centroid it selects, and $\mathrm{Var}_{m|x_t}(\mu_m)$ below is the variance of μ_m under the posterior $p(m \mid x_t, h)$, which measures how much the state still leaves undecided about which mode is being generated.

We define the following notations which will be used in our proof,

$$V := \alpha^2\varsigma_0^2+\beta^2\varsigma_1^2, \qquad \Lambda := \dot\alpha\alpha\varsigma_0^2+\dot\beta\beta\varsigma_1^2, \qquad Q := \dot\alpha^2\varsigma_0^2+\dot\beta^2\varsigma_1^2-\frac{\Lambda^2}{V}, \qquad \Upsilon := \frac{\varsigma_0^2\,\alpha\,(\alpha\dot\beta-\beta\dot\alpha)}{V}.$$

With $\gamma \equiv 0$ we have $x_t = \alpha x_0 + \beta x_1$ and $\dot x_t = \dot\alpha x_0 + \dot\beta x_1$. For any measurable $\hat b(t, x_t, h)$, adding and subtracting the conditional mean $b = \mathbb{E}[\dot x_t \mid x_t, h]$ splits the velocity loss at each t,

$$\begin{aligned}\mathbb{E}\big\|\hat b - \dot x_t\big\|^2 &= \mathbb{E}\big\|(\hat b - b) + (b - \dot x_t)\big\|^2 \\ &= \mathbb{E}\big\|\hat b - b\big\|^2 \;+\; 2\,\mathbb{E}\big\langle \hat b - b,\; b - \dot x_t\big\rangle \;+\; \mathbb{E}\big\|b - \dot x_t\big\|^2.\end{aligned}$$

The middle term vanishes in three moves. Condition on (x_t, h); note that $\hat b$ and b are both functions of (t, x_t, h), hence constant under that conditioning and free to leave the inner expectation; then use $\mathbb{E}[\dot x_t \mid x_t, h] = b$, which is what b is defined to be:

$$\begin{aligned}\mathbb{E}\big\langle \hat b - b,\; b - \dot x_t\big\rangle &= \mathbb{E}\Big[\;\mathbb{E}\big[\langle \hat b - b,\; b - \dot x_t\rangle \;\big|\; x_t, h\big]\Big] && \text{(tower property)} \\ &= \mathbb{E}\Big[\big\langle \hat b - b,\; b - \mathbb{E}[\dot x_t \mid x_t, h]\big\rangle\Big] && (\hat b, b \text{ known given } x_t, h) \\ &= \mathbb{E}\big[\langle \hat b - b,\; b - b\rangle\big] \;=\; 0 && \text{(definition of } b).\end{aligned}$$

Therefore,

$$\mathcal{L}_b^* = \min_{\hat b} \int_0^1 \mathbb{E}\big\|\hat b - \dot x_t\big\|^2 \mathrm{d}t = \int_0^1 \mathbb{E}_{x_t}\big[\mathrm{Var}(\dot x_t \mid x_t, h)\big]\mathrm{d}t. \tag{28}$$

Now condition on m, then sampling $x_1 \sim \mathcal{N}(\mu_m, \varsigma_1^2 I_d)$ is independent of $x_0 \sim \mathcal{N}(0, \varsigma_0^2 I_d)$, so $(\dot x_t, x_t)$ is jointly Gaussian with

$$\begin{aligned}\mathbb{E}[x_t \mid m] &= \beta\mu_m, & \mathrm{Var}(x_t \mid m) &= (\alpha^2\varsigma_0^2 + \beta^2\varsigma_1^2) I_d = V I_d, \\ \mathbb{E}[\dot x_t \mid m] &= \dot\beta\mu_m, & \mathrm{Var}(\dot x_t \mid m) &= (\dot\alpha^2\varsigma_0^2 + \dot\beta^2\varsigma_1^2) I_d,\end{aligned}$$

and $\mathrm{Cov}(\dot x_t, x_t \mid m) = (\dot\alpha\alpha\varsigma_0^2 + \dot\beta\beta\varsigma_1^2) I_d = \Lambda I_d$, the cross terms in x_0, x_1 vanishing by independence.

Using the Gaussian conditioning formulas $\mathbb{E}[Y \mid X] = \mathbb{E}Y + \mathrm{Cov}(Y, X)\mathrm{Var}(X)^{-1}(X - \mathbb{E}X)$ and $\mathrm{Var}(Y \mid X) = \mathrm{Var}Y - \mathrm{Cov}(Y, X)\mathrm{Var}(X)^{-1}\mathrm{Cov}(X, Y)$, we obtain,

$$\mathbb{E}[\dot x_t \mid x_t, m] = \dot\beta\mu_m + \frac{\Lambda}{V}\big(x_t - \beta\mu_m\big) = \frac{\Lambda}{V}\,x_t + \Big(\dot\beta - \frac{\beta\Lambda}{V}\Big)\mu_m = \frac{\Lambda}{V}\,x_t + \Upsilon\,\mu_m,$$

$$\mathrm{Var}(\dot x_t \mid x_t, m) = \Big(\dot\alpha^2\varsigma_0^2 + \dot\beta^2\varsigma_1^2 - \frac{\Lambda^2}{V}\Big) I_d = Q\, I_d,$$

the variance free of m. The coefficient of μ_m is Υ because

$$\dot\beta - \frac{\beta\Lambda}{V} = \frac{\dot\beta\alpha^2\varsigma_0^2 + \dot\beta\beta^2\varsigma_1^2 - \beta\dot\alpha\alpha\varsigma_0^2 - \dot\beta\beta^2\varsigma_1^2}{V} = \frac{\varsigma_0^2\,\alpha\,(\alpha\dot\beta - \beta\dot\alpha)}{V} = \Upsilon,$$

the $\dot\beta\beta^2\varsigma_1^2$ terms cancelling. Applying the law of total variance over m conditionally on x_t, the $\frac{\Lambda}{V}x_t$ term being m-free, gives $\mathrm{Var}(\dot x_t \mid x_t, h) = Q\, I_d + \Upsilon^2\mathrm{Var}_{m|x_t}(\mu_m)$, so

$$\mathcal{L}_b^*(h) \;=\; \underbrace{\int_0^1 Q(t)\,\mathrm{d}t}_{\text{target energy}} \;+\; \underbrace{\int_0^1 \Upsilon(t)^2\;\mathbb{E}_{x_t}\big[\mathrm{Var}_{m|x_t}(\mu_m)\big]\mathrm{d}t}_{\text{mode confusion}}, \tag{29}$$

Only the second term sees the user, and it does so only through the posterior over modes.

On the linear path both terms are finite. Here $\dot\alpha = -1$, $\dot\beta = 1$, so $\dot\alpha^2\varsigma_0^2 + \dot\beta^2\varsigma_1^2 = \varsigma_0^2 + \varsigma_1^2$ is constant and $\Lambda^2 \le V(\varsigma_0^2 + \varsigma_1^2)$ by Cauchy–Schwarz, giving $0 \le Q \le \varsigma_0^2 + \varsigma_1^2$. The Wronskian is

$\alpha\dot\beta - \beta\dot\alpha = 1$, so $\Upsilon(t) = \varsigma_0^2(1-t)/\big((1-t)^2\varsigma_0^2 + t^2\varsigma_1^2\big)$ is continuous on $[0,1]$ with $g(0) = 1$ and $g(1) = 0$. Both integrands in Eq. (29) are bounded, so $\mathcal{L}_b^* < \infty$.

On the variance-preserving path both diverge. Write $a := \bar\alpha_t$, with $a \asymp t$ as $t \to 0$ and $1 - a \asymp 1 - t$ as $t \to 1$, so $\dot\alpha = -\dot a/(2\sqrt{1-a})$ and $\dot\beta = \dot a/(2\sqrt{a})$. At the data end $\dot\alpha^2\varsigma_0^2$ diverges while Λ stays bounded and $V \to \varsigma_1^2 > 0$, hence $Q = \Omega((1-t)^{-1})$ and $\int^1 Q\,\mathrm{d}t = \infty$: the regression target itself blows up, because the path must shed the source at unbounded speed. At the noise end the Wronskian is $\dot a/(2\sqrt{a(1-a)})$, so $\Upsilon = \varsigma_0^2\dot a/(2\sqrt{a}\,V)$ and, with $V \to \varsigma_0^2$, $\Upsilon^2 = \Omega(1/t)$; since $\mathbb{E}_{x_t}[\mathrm{Var}_{m|x_t}(\mu_m)] \to \mathrm{Var}_{\pi(h)}(\mu_m) > 0$ there, the mode-confusion integral diverges too. Both divergences are logarithmic under a symmetric truncation $t \in [\tau, 1-\tau]$. Among the paths considered here the linear one is therefore the only member with finite irreducible velocity loss.

The transport cost contains that loss. The second claim of the theorem concerns $\int_0^1 \mathbb{E}\|\dot x_t\|^2\mathrm{d}t$, which is not a new object: applying the decomposition above to the raw second moment of the same regression target gives

$$\mathbb{E}\|\dot x_t\|^2 = \mathbb{E}\|b\|^2 + \mathbb{E}_{x_t}\big[\mathrm{Var}(\dot x_t \mid x_t, h)\big], \qquad \text{hence} \qquad \int_0^1 \mathbb{E}\|\dot x_t\|^2\mathrm{d}t = \int_0^1 \mathbb{E}\|b\|^2\mathrm{d}t + \mathcal{L}_b^*(h).$$

The transport cost is therefore the *whole* energy of the velocity target, splitting into the part the state predicts and the part it cannot, the latter being Eq. (29). Minimizing it keeps the target small overall, and the bound below holds for every h because the source and target moments alone control it.

The linear path uniquely minimizes it. Since x_0 is zero-mean and independent of x_1, the cross term vanishes and $\mathbb{E}\|\dot x_t\|^2 = \dot\alpha^2\,\mathbb{E}\|x_0\|^2 + \dot\beta^2\,\mathbb{E}\|x_1\|^2$. The boundary conditions fix $\int_0^1 \dot\alpha\,\mathrm{d}t = -1$ and $\int_0^1 \dot\beta\,\mathrm{d}t = 1$, so Cauchy–Schwarz gives $\int_0^1 \dot\alpha^2\mathrm{d}t \geq 1$ and $\int_0^1 \dot\beta^2\mathrm{d}t \geq 1$, each with equality only when the integrand is constant. Adding them,

$$\int_0^1 \mathbb{E}\|\dot x_t\|^2\mathrm{d}t \;\geq\; \mathbb{E}\|x_0\|^2 + \mathbb{E}\|x_1\|^2,$$

with equality if and only if $\alpha(t) = 1 - t$ and $\beta(t) = t$ almost everywhere. No normalization of the path was needed: the boundary conditions alone pin both integrals. □

I Proof of Theorem 6

Assumption 9 (Path regularity)**.** *Along the interpolant path, for all $t \in [0,1]$:* (i) *the exact velocity $b(t,\cdot,h)$ is L-Lipschitz in x;* (ii) *the marginal $\rho_t(\cdot \mid h)$ is m-strongly log-concave, i.e. the exact score satisfies the one-sided condition $\langle s(t,x) - s(t,y),\, x - y\rangle \leq -m\|x-y\|^2$ for some $m > 0$;* (iii) *the learned fields have square-integrable errors, with time-local excess risks measured along the* sampler's own *trajectory, $\delta_b(t,h) := \mathbb{E}\|\hat b_\theta(t,\hat X_t) - b(t,\hat X_t)\|^2$ and likewise for $\delta_s(t,h)$.*

Proof of Theorem 6. Let X_t solve the exact-field SDE Eq. (6) and $\hat X_t$ the learned-field SDE Eq. (8), driven by the *same* Brownian motion and started from the same $X_0 \sim \rho_0$ (a synchronous coupling). Write $\Delta_t = \hat X_t - X_t$. Because the Brownian terms are identical and carry the same coefficient $\sqrt{2\epsilon}$, they cancel, and Δ_t solves a *random ODE*:

$$\frac{d\Delta_t}{\mathrm{d}t} \;=\; \underbrace{\big[b(t,\hat X_t) - b(t,X_t)\big]}_{\text{(I) velocity Lipschitz}} + \epsilon\,\underbrace{\big[s(t,\hat X_t) - s(t,X_t)\big]}_{\text{(II) score contraction}} + \underbrace{e_b(t,\hat X_t) + \epsilon\, e_s(t,\hat X_t)}_{\text{(III) approximation error}}, \tag{30}$$

where $e_b := \hat b_\theta - b$ and $e_s := \hat s_\phi - s$.

Taking the inner product of Eq. (30) with Δ_t and using L-Lipschitzness for (I), the one-sided condition (ii) for (II), and Cauchy–Schwarz for (III),

$$\tfrac{1}{2}\tfrac{d}{\mathrm{d}t}\|\Delta_t\|^2 \;\leq\; (L - \epsilon m)\,\|\Delta_t\|^2 \;+\; \|\Delta_t\|\big(\|e_b\| + \epsilon\|e_s\|\big).$$

Wherever $\|\Delta_t\| > 0$ this gives $\frac{d}{\mathrm{d}t}\|\Delta_t\| \leq (L - \epsilon m)\|\Delta_t\| + \|e_b\| + \epsilon\|e_s\|$; the inequality extends to the whole interval by the standard regularization $\sqrt{\|\Delta_t\|^2 + \vartheta}$, $\vartheta \downarrow 0$.

Integrating from $\Delta_0 = 0$, $\|\Delta_1\| \leq \int_0^1 e^{(L-\epsilon m)(1-t)}(\|e_b(t,\hat X_t)\| + \epsilon\|e_s(t,\hat X_t)\|)\mathrm{d}t$ pathwise. Taking $L^2(\mathbb{P})$ norms and applying Minkowski's integral inequality (the L^2 norm of an integral is at most

the integral of the L^2 norms),

$$\big(\mathbb{E}\|\Delta_1\|^2\big)^{1/2} \;\le\; \int_0^1 e^{(L-\epsilon m)(1-t)}\Big(\sqrt{\delta_b(t,h)}+\epsilon\sqrt{\delta_s(t,h)}\Big)\mathrm{d}t.$$

Since $\hat{X}_1$ and X_1 form a coupling of $\hat{p}^\epsilon(\cdot \mid h)$ and $p(\cdot \mid h)$ (Thm. 3 gives $X_1 \sim p(\cdot \mid h)$ exactly, for every ϵ), the left-hand side upper-bounds $\mathcal{W}_2$. Bounding $\delta_\bullet(t,h) \le \bar{\delta}_\bullet(h) := \sup_t \delta_\bullet(t,h)$ and writing $\kappa(\epsilon) := \int_0^1 e^{(L-\epsilon m)(1-t)}\mathrm{d}t = \big(e^{L-\epsilon m}-1\big)/(L-\epsilon m)$,

$$\mathcal{W}_2\big(p,\hat{p}^\epsilon\big) \;\le\; \kappa(\epsilon)\Big(\sqrt{\bar{\delta}_b(h)}+\epsilon\sqrt{\bar{\delta}_s(h)}\Big). \tag{31}$$

(The time-averaged $\delta_\bullet(h)$ may be used instead at the cost of one further Cauchy–Schwarz and a constant; we keep the same symbols below.)

Substituting $r = 1-t$,

$$\kappa(\epsilon) = \int_0^1 e^{Lr}\, e^{-m\epsilon r}\, dr, \tag{32}$$

a Laplace transform of the finite positive measure $e^{Lr}\,dr$ on $[0,1]$, evaluated at $m\epsilon$. Three properties follow immediately and none require an expansion: κ is strictly decreasing in ϵ; κ is *finite* at $\epsilon = 0$, with $\kappa(0) = (e^L-1)/L$, which is precisely what $1/(2\epsilon)$ was not; and $\log\kappa$ is *convex* in ϵ, by Hölder's inequality applied to Eq. (32).

Fix the search range $[0,E]$. Convexity of $\log\kappa$ means its graph lies below the chord joining its endpoints, so with $\lambda_E := \frac{1}{E}\log\frac{\kappa(0)}{\kappa(E)} > 0$,

$$\kappa(\epsilon) \;\le\; \kappa(0)\, e^{-\lambda_E\,\epsilon} \qquad \text{for all } \epsilon \in [0,E]. \tag{33}$$

This is an exact inequality on the whole range, not a local approximation. (Because the chord slope of a convex function is at least its slope at the left endpoint, $\lambda_E \le -\kappa'(0)/\kappa(0) = m\int_0^1 re^{Lr}dr \,/\, \int_0^1 e^{Lr}dr$, so Eq. (33) is a genuine weakening of the tangent bound and $\lambda_E \uparrow -\kappa'(0)/\kappa(0)$ as $E \downarrow 0$.)

We set $\lambda := 2\lambda_E$, matching the definition in Thm. 6, and write $a = \sqrt{\bar{\delta}_b}$, $c = \sqrt{\bar{\delta}_s}$. Squaring Eq. (31) and applying, in order, Eq. (33); the elementary $e^{-u} \le 1/(1+u)$ for $u \ge 0$; the AM–GM bound $(a+\epsilon c)^2 \le 2a^2 + 2\epsilon^2 c^2$; and $\epsilon^2/(1+\lambda\epsilon) \le \epsilon/\lambda$:

$$\mathcal{W}_2^2\big(p,\hat{p}^\epsilon\big) \;\le\; \kappa(0)^2 e^{-\lambda\epsilon}\,(a+\epsilon c)^2 \;\le\; \frac{\kappa(0)^2\,(a+\epsilon c)^2}{1+\lambda\epsilon} \;\le\; \frac{2\kappa(0)^2\big(\bar{\delta}_b+\epsilon^2\bar{\delta}_s\big)}{1+\lambda\epsilon}$$

$$\le\; C\Big[\frac{\bar{\delta}_b(h)}{1+\lambda\epsilon}+\epsilon\,\bar{\delta}_s(h)\Big], \qquad C := 2\kappa(0)^2\max\{1,\lambda^{-1}\},$$

valid for every $\epsilon \in [0,E]$. Every step is an inequality between explicit quantities; no term is discarded as higher order. This is the bound claimed in Theorem 6.

Let $f(\epsilon) = \delta_b/(1+\lambda\epsilon) + \epsilon\,\delta_s$ on $\epsilon \ge 0$. Then $f'(\epsilon) = -\lambda\delta_b/(1+\lambda\epsilon)^2 + \delta_s$, which is increasing, so f is convex with a unique stationary point at $(1+\lambda\epsilon)^2 = \lambda\delta_b/\delta_s$, i.e. $\epsilon = \lambda^{-1}\big(\sqrt{\lambda\delta_b/\delta_s}-1\big)$. If that value is negative, equivalently $f'(0) = \delta_s - \lambda\delta_b \ge 0$, i.e. $\lambda\delta_b \le \delta_s$, then f is nondecreasing on $[0,\infty)$ and the constrained minimizer is $\epsilon^\star = 0$. Combining the two cases gives Eq. (11). □

J Proofs of the Propositions

J.1 Proof of Proposition 7

J.1.1 Proof of Proposition 7 (a)

Assumption 10 (Bayes risk drives excess risk)**.** *For the velocity head, the excess risk is monotone in the irreducible loss at fixed capacity and sample size:* $\delta_b(h) = \Theta\big(c_n\,\mathcal{L}_b^*(h)\big)$ *for a user-independent* $c_n > 0$ *decreasing in* n*.*

The velocity head solves a least-squares problem whose target $\dot{x}_t$ is noisy in a user-dependent way: $\mathcal{L}_b^*(h)$ is by definition the part of that target's variance no function of (t, x_t, h) can remove. For a well-specified class of effective dimension p fitted on n samples, the classical excess risk of least

squares is the noise level times p/n, so a user whose target is twice as noisy contributes twice the estimation error at the same capacity and the same data. Assumption 10 is that statement written per user, with $c_n = \Theta(p/n)$.

Every user is served by the same network, the same item table, the same optimizer and the same epoch budget, and the head is fitted once on pooled training pairs rather than separately per user. What differs across users is the conditioning h, which changes how noisy the target is but not the capacity spent on it or the number of parameters shared. The scaling factor is therefore a property of the estimator and the corpus, and the user enters only through $\mathcal{L}_b^*(h)$.

Proof. Take Assumption 4 with a common within-mode spread ς_1, the linear path, $\gamma \equiv 0$, and $x_0 \sim \mathcal{N}(0, \varsigma_0^2 I_d)$ (the lower bound extends to $\gamma > 0$ by data processing). Write $U := \dot{x}_t = x_1 - x_0$ and $W := x_t = (1-t)x_0 + tx_1$. The irreducible velocity loss is the integrated conditional variance, $\mathcal{L}_b^* = \int_0^1 \mathbb{E}_W\big[\mathrm{Var}(U \mid W, h)\big]\,\mathrm{d}t$.

Conditioned on the category k, the triple (x_0, x_1) is jointly Gaussian, so $\mathrm{Var}(U \mid W, k) = v(t)\, I_d$ with $v(t)$ *independent of* k: the components share ς_1, and changing k shifts the conditional mean but not the conditional covariance. Hence

$$\mathrm{Var}(U \mid W) = \underbrace{v(t)\, I_d}_{\text{within-category}} + \underbrace{(1 - t\,a(t))^2\, \mathrm{Var}_{k|W}(\mu_k)}_{\text{between-category}},$$

where $a(t) = t\varsigma_1^2/\big((1-t)^2\varsigma_0^2 + t^2\varsigma_1^2\big)$ is the Gaussian regression coefficient of x_1 on W within a component, and $\mathrm{Var}_{k|W}$ denotes the variance of μ_k under the posterior $p(k \mid W, h)$. Only the second term depends on the user's preference dispersion; the first is a user-independent floor. We bound the second from below.

This is the step at which a direct appeal to *conditioning reduces variance* would go the wrong way: the law of total variance gives $\mathbb{E}_W[\mathrm{Var}_{k|W}(\mu_k)] = \mathrm{Var}_{\pi(h)}(\mu_k) - \mathrm{Var}_W(\mathbb{E}[\mu_k \mid W]) \le \mathrm{Var}_{\pi(h)}(\mu_k)$, an *upper* bound. The correct route is through the endpoint. At $t = 0$ we have $W = x_0$, which is independent of k by construction of the interpolant (the source and the target are drawn independently, Definition 2), so the posterior equals the prior and $\mathbb{E}_W\big[\mathrm{Var}_{k|W}(\mu_k)\big]\Big|_{t=0} = \mathrm{Var}_{\pi(h)}(\mu_k)$. The map $t \mapsto \mathbb{E}_W[\mathrm{Var}_{k|W}(\mu_k)]$ is continuous on $[0, 1)$ (the posterior $p(k \mid W)$ is a ratio of Gaussian densities with parameters continuous in t, and μ_k ranges over a finite set), so there exists $t_0 = t_0(\varsigma_0, \varsigma_1, \{\mu_k\}) \in (0, 1)$ with

$$\mathbb{E}_W\big[\mathrm{Var}_{k|W}(\mu_k)\big] \;\ge\; \tfrac{1}{2}\,\mathrm{Var}_{\pi(h)}(\mu_k) \qquad \text{for all } t \in [0, t_0].$$

Discarding the (nonnegative) within-category term and the contribution of $t > t_0$,

$$\mathcal{L}_b^*(h) \;\ge\; \int_0^{t_0} (1 - t\,a(t))^2\, \mathbb{E}_W\big[\mathrm{Var}_{k|W}(\mu_k)\big]\,\mathrm{d}t \;\ge\; c\,\mathrm{Var}_{\pi(h)}(\mu_k),$$

$c := \frac{1}{2}\int_0^{t_0}(1 - t\,a(t))^2\,\mathrm{d}t \;>\; 0$, which is the claim, with c depending only on $(\varsigma_0, \varsigma_1, t_0)$ and not on the user. □

J.1.2 Proof of Proposition 7 (b)

We first prove the following Theorem:

Theorem 11 (Per-user, per-timestep generation bound)**.** *For any measurable schedule $\epsilon(t, h) > 0$,*

$$D_{\mathrm{KL}}(\, p(\cdot \mid h) \,\|\, \hat{p}^{\epsilon}(\cdot \mid h)\,) \;\le\; \int_0^1 \left[\, \frac{\delta_b(t,h)}{2\,\epsilon(t,h)} \;+\; \frac{\epsilon(t,h)\,\delta_s(t,h)}{2} \,\right] \mathrm{d}t,$$

and the right-hand side is minimized pointwise in t at

$$\epsilon^\star(t,h) = \sqrt{\frac{\delta_b(t,h)}{\delta_s(t,h)}}, \qquad \textit{attaining} \qquad \int_0^1 \sqrt{\delta_b(t,h)\,\delta_s(t,h)}\,\mathrm{d}t.$$

Proof. We first establish a KL dissipation identity for time-varying $\epsilon(t)$.

KL dissipation. Let ρ_t be the true interpolant marginal and $\hat{\rho}_t$ be the marginal of the universal sampler with approximate fields. Both satisfy Fokker–Planck equations with shared diffusion coefficient

$\epsilon(t)$ but different drifts:

$$\text{True:} \quad b_F(t) = b + \epsilon(t)\, s,$$
$$\text{Approximate:} \quad \hat{b}_F(t) = \hat{b}_\theta + \epsilon(t)\, \hat{s}_\phi.$$

Define $\ell(t,x) = \log(\rho_t(x)/\hat{\rho}_t(x))$. Differentiating $\mathrm{KL}(\rho_t\|\hat{\rho}_t) = \int \rho_t\, \ell\, dx$ and substituting both Fokker–Planck equations yields:

$$\frac{d}{\mathrm{d}t}\mathrm{KL}(\rho_t\|\hat{\rho}_t) = \int \rho_t\, (\nabla\ell)\cdot(\hat{b}_F - b_F)\, dx \;-\; \epsilon(t) \int \rho_t\, |\nabla\ell|^2\, dx.$$

Young's inequality. Integrating from $t=0$ to $t=1$ with $\mathrm{KL}(\rho_0\|\hat{\rho}_0) = 0$ (same initial condition):

$$\mathrm{KL}(\rho_1\|\hat{\rho}_1) = \int_0^1 \int \rho_t\, \nabla\ell\cdot(\hat{b}_F - b_F)\, dx\, \mathrm{d}t \;-\; \int_0^1 \epsilon(t) \int \rho_t\, |\nabla\ell|^2\, dx\, \mathrm{d}t.$$

Apply Young's inequality pointwise: $\nabla\ell\cdot(\hat{b}_F - b_F) \le \epsilon(t)\,|\nabla\ell|^2 + \frac{1}{4\epsilon(t)}|\hat{b}_F - b_F|^2$. The $\epsilon(t)\,|\nabla\ell|^2$ terms cancel:

$$\mathrm{KL}(\rho_1\|\hat{\rho}_1) \;\le\; \int_0^1 \frac{1}{4\epsilon(t)} \int \rho_t\, |\hat{b}_F - b_F|^2\, dx\, \mathrm{d}t. \tag{34}$$

Splitting the drift error into its velocity and score parts, substituting the two excess risks of Eq. (10), and minimizing the resulting AM–GM expression pointwise in t gives the stated bound at $\epsilon^\star(t,h) = \sqrt{\delta_b(t,h)/\delta_s(t,h)}$. □

We now provide the proof of Proposition 7 (b)

Proof. Under a Bayesian model in which interactions are i.i.d. draws from a latent category $m^\star$, the posterior concentrates: $\pi_{m^\star}(h) \;\ge\; 1 - (M-1)\, e^{-cT}$ after T interactions (standard Dirichlet–Multinomial concentration). Consequently

$$\mathrm{Var}_{\pi(h)}(\mu_m) \;\le\; (M-1)\, \Delta_{\max}^2\, e^{-cT}, \qquad \Delta_{\max} = \max_{j,k} \|\mu_j - \mu_k\|.$$

Proposition 7(a) together with Assumption 10 gives $\delta_b(h) = \Theta\big(c_n\, \mathrm{Var}_{\pi(h)}(\mu_m)\big) = O(e^{-cT})$, while the score excess risk satisfies $\delta_s(h) \ge \delta_s^{\min} > 0$ (the score problem does not become trivially easy as the history grows).

For the geometric-mean rule of Theorem 11 this already gives $\sqrt{\delta_b(h)/\delta_s(h)} = O(e^{-cT/2}) \to 0$. The threshold rule Eq. (11) says something strictly stronger. Since $\epsilon_\lambda^\star(h) \le \lambda^{-1}\sqrt{\lambda\delta_b/\delta_s} = \sqrt{\delta_b/(\lambda\delta_s)}$, the same $O(e^{-cT/2})$ decay holds; and because $\epsilon_\lambda^\star(h) = 0$ as soon as $\lambda\delta_b(h) \le \delta_s(h)$, the decay *terminates*: there is a finite $T_0 = \frac{1}{c}\log\frac{\lambda\,\Theta(c_n)\,(M-1)\Delta_{\max}^2}{\delta_s^{\min}}$ beyond which $\epsilon_\lambda^\star(h)$ is exactly zero, not merely small. This is the cold-start–stochasticity law in its sharp form: exploration should be switched off, not just annealed. □

J.1.3 Proof of Proposition 7 (c)

Proof. **Accuracy.** By Theorem 11, $A(\epsilon) = -D_{\mathrm{KL}}(p\,\|\,\hat{p}^\epsilon) \ge -B(\epsilon,h)$ with $B(\epsilon,h) = \delta_b/(2\epsilon) + \epsilon\,\delta_s/2$. The front derived below is therefore a front of the *bound*.

With an imperfect score, at each Euler–Maruyama step the injected noise $\sqrt{2\epsilon\,\Delta t}\,\xi$ is only partially corrected. The fraction that "leaks" per step is $1 - r_n$ with $r_n = 1 - O(\delta_s\,\Delta t)$ measuring score quality. Summing over $N = 1/\Delta t$ steps,

$$\mathrm{R}_{\hat{\mathcal{A}}_\epsilon}(h) = \mathrm{tr}\,\mathrm{Cov}(X_1^\epsilon \mid h) = \mathrm{R}_{\hat{\mathcal{A}}_0}(h) + c(h)\,\epsilon\,\delta_s + o(\delta_s), \qquad c(h) = 2\int_0^1 (1 - r(t,h))^2\, \mathrm{d}t > 0.$$

Write $u := \mathrm{R}_{\hat{\mathcal{A}}_\epsilon}(h) - \mathrm{R}_{\hat{\mathcal{A}}_0}(h) = c\,\epsilon\,\delta_s$, so $\epsilon = u/(c\,\delta_s)$. Substituting into the accuracy bound eliminates ϵ: $A \;\ge\; -\frac{c\,\delta_b\,\delta_s}{2u} - \frac{u}{2c} \;=:\; F(u), \qquad u > 0$. Then $F''(u) = -c\,\delta_b\,\delta_s/u^3 < 0$, so the attainable frontier is concave: the first units of diversity are cheap in accuracy and later ones progressively more expensive, with the curvature set by the product $\delta_b\delta_s$. □

Part B: Additional Results and Measurements

K Delivered Dispersion of the Native Samplers

All four native processes use the architecture of App. C.2, the same history encoder and item-table dimension, and the same optimizer, batch size and epoch budget (App. C.1). Each is trained independently with its own objective, and the processes differ only in the interpolant they are trained on and the sampler they are read with, so what the figure compares is processes rather than implementations.

Samplers. Each process is integrated on a uniform grid of n_{step} steps, $\Delta t = 1/n_{\text{step}}$, from $X_0 \sim \rho_0$, with the fields read at interval midpoints t_i and $\xi_i \sim \mathcal{N}(0, I_d)$ drawn independently at each step:

$$\begin{aligned}
\text{FM:} \quad & X_{i+1} = X_i + \hat{b}(t_i, X_i, h)\,\Delta t, \\
\text{SB:} \quad & X_{i+1} = X_i + \big(\hat{b} + \tfrac{\sigma^2}{2}\hat{s}\big)\Delta t + \sigma\sqrt{\Delta t}\,\xi_i, \\
\text{DDPM:} \quad & X_{i+1} = X_i + \omega_{\text{vp}}\big(X_i + 2\hat{s}\big)\Delta t + \sqrt{2\omega_{\text{vp}}\Delta t}\,\xi_i, \\
\text{DDIM:} \quad & X_{i+1} = X_i + \omega_{\text{vp}}\big(X_i + \hat{s}\big)\Delta t.
\end{aligned}$$

The two ODEs consume no ξ_i, which is why their injected term is exactly zero rather than merely small.

Estimator. We freeze the weights, the history and the item table, draw A source points $x_0^{(a)} \sim \rho_0$ and run B independent noise realizations from each, giving $S = AB = 32$ endpoints $\hat{x}^{(a,b)}$. Writing $\bar{x}^{(a)}$ for the mean over b and $\bar{x}$ for the mean over all S, the dispersion of Def. 1 is the trace of their empirical covariance,

$$\hat{\mathrm{R}}(h) \;=\; \frac{1}{S-1}\sum_{a,b}\big\|\hat{x}^{(a,b)} - \bar{x}\big\|^2,$$

and the law of total variance splits it into the two terms of Eq. (2),

$$\underbrace{\frac{1}{A}\sum_{a}\frac{1}{B-1}\sum_{b}\big\|\hat{x}^{(a,b)} - \bar{x}^{(a)}\big\|^2}_{\text{injected}} \;+\; \underbrace{\frac{1}{A-1}\sum_{a}\big\|\bar{x}^{(a)} - \bar{x}\big\|^2}_{\text{source-propagated}},$$

the first measuring spread at a fixed source draw, the second the spread of the per-source means. For the two deterministic samplers the first term vanishes identically, the endpoint being a function of x_0 alone. We report $\hat{\mathrm{R}}(h)$ relative to $\mathbb{E}\|\hat{x}\|^2$, and take medians over 4k users drawn uniformly from the test split.

Figure 1 reports ML-1M and Fig. 4 Beauty, where the composition and the ordering of totals both repeat, the Bridge drawing 59% and 55% of its dispersion from injection against DDPM's 25% and 20%. Read per user rather than at the median, the four families are disjoint over essentially the whole population.

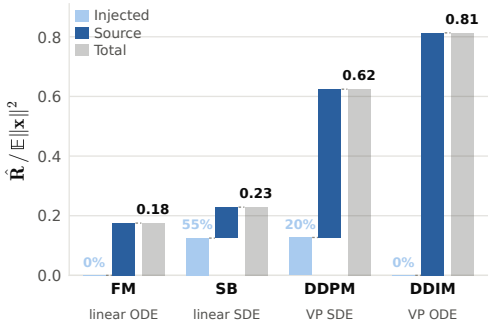


Figure 4: **Delivered Dispersion on Beauty**, as Fig. 1.

Three apply to every number quoted. Each method carries its own item table, so cross-method values are comparable only up to the learned geometry, which normalizing by $\mathbb{E}\|\hat{x}\|^2$ controls for scale but not for shape. The dispersion is taken over *sampler* seeds at one checkpoint, not over training seeds, so it measures the process and not the run-to-run variability of fitting one.

And absolute levels depend on that normalizer, so magnitudes across methods should be read as indicative only. The *composition* shares are immune to this, since the normalizer cancels between the two terms of one method, and so is the qualitative interpolant-versus-sampler contrast.

L Comparison with Post-hoc Diversification

Writing τ for the fraction of the deterministic sampler's NDCG@20 that a method must retain, we sweep τ from 1.00 down to 0.90 and report Coverage@20 for the best setting of each method meeting the budget. All methods read the same trained model. Per-user ϵ changes the process the three baselines leave the generated endpoint alone and operate on the list it produces.

L.1 Readout Invariance

What any of these methods can achieve at a tight budget is bounded by the readout itself. Retrieval against a fixed catalog is a rounding operation with a resolution: an endpoint can be displaced by any amount smaller than half the score gap at the cut-off without a single item changing place, and every metric computed from the resulting list is then unchanged *exactly*, not approximately. This is what makes coverage at τ=1.00 hard to buy by re-ranking. An operation constrained to leave NDCG@20 untouched can only permute items inside that dead zone, whereas moving ϵ changes which items are generated in the first place.

Lemma 12 (Readout invariance)**.** *Fix a user h and a generated endpoint $\hat{x}$, and let $\Delta_K(h) = \langle \hat{x}, e_{(K)} \rangle - \langle \hat{x}, e_{(K+1)} \rangle \geq 0$ be the score gap at the cut-off, where $e_{(j)}$ is the j-th ranked item. If $\|\hat{x}' - \hat{x}\| < \Delta_K(h)/2$ then* Top-K(h) *is unchanged, and every metric that is a function of that list, HR@K, NDCG@K, MRR truncated at K, takes* exactly *the same value.*

Proof. Item embeddings are ℓ_2-normalized, so $|\langle \hat{x}' - \hat{x}, e_j \rangle| \leq \|\hat{x}' - \hat{x}\| < \Delta_K/2$ for every j: each score moves by less than $\Delta_K/2$, so the gap Δ_K cannot be closed and no item crosses the cut-off. □

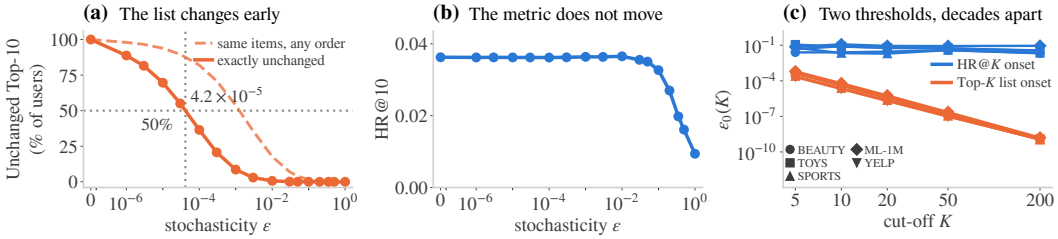


Figure 5: **The process moves; the measurement does not.** One checkpoint swept from ϵ=0 upward, everything else fixed. **(a)** Users still holding their ϵ=0 Top-10 fall through 50% early. **(b)** HR@10 over the same range is flat. **(c)** The two onsets are two to eight decades apart at every cut-off K, the gap being the quantizer of Lemma 12.

Fig. 5 measures the gap the lemma predicts: sweeping one checkpoint upward from ϵ=0, the users still holding their original Top-10 fall through half well before HR@10 has moved at all, the two onsets separated by two to eight decades at every cut-off.

L.2 The Post-hoc Baselines

Temperature. With p the mean softmax over the S draws, we sample K items without replacement from $\mathrm{softmax}(\log p/T)$, so $T \to 0$ recovers the deterministic top-K and larger T reaches further down the list. It is content-blind: two near-duplicates with similar scores remain equally likely to both appear.

MMR. We re-rank by maximal marginal relevance (Carbonell & Goldstein, 1998), trading each candidate's relevance against its greatest similarity to the items already placed.

DPP. On the top $M{=}100$ candidates we run the fast greedy MAP of Chen et al. (2018) with $L = \mathrm{diag}(q)\, S\, \mathrm{diag}(q)$, where S is the cosine Gram matrix of the ℓ_2-normalised item embeddings and $q_i = \exp(\theta(\mathrm{rel}_i - \mathrm{rel}_{\max}))$, so a set of high-scoring but mutually similar items loses to a slightly weaker one that is better spread.

L.3 Coverage at a Fixed Accuracy Budget

Table 5: **Coverage at a fixed accuracy budget.** Coverage@20 for the best setting of each method meeting the budget τ. Subscripts are standard deviations over 20 seeds in units of the last digit shown. Best per row in bold.

		STIR, $\epsilon{=}0$			STIR
Corpus	τ	+ Temperature	+ MMR	+ DPP	per-user ϵ
ML-1M	1.00	$0.5263_{\pm 15}$	$0.5260_{\pm 15}$	$0.5256_{\pm 15}$	$\mathbf{0.5560}_{\pm 30}$
	0.98	$0.5307_{\pm 21}$	$0.5260_{\pm 15}$	$0.5256_{\pm 13}$	$\mathbf{0.5583}_{\pm 34}$
	0.95	$0.5389_{\pm 12}$	$0.5260_{\pm 14}$	$0.5256_{\pm 15}$	$\mathbf{0.5642}_{\pm 21}$
	0.90	$0.5689_{\pm 24}$	$0.5260_{\pm 15}$	$0.5256_{\pm 15}$	$\mathbf{0.5693}_{\pm 10}$
BEAUTY	1.00	$0.6022_{\pm 22}$	$0.6019_{\pm 26}$	$0.5997_{\pm 43}$	$\mathbf{0.6186}_{\pm 29}$
	0.98	$0.6082_{\pm 28}$	$0.6077_{\pm 24}$	$0.6130_{\pm 25}$	$\mathbf{0.6262}_{\pm 21}$
	0.95	$0.6262_{\pm 40}$	$0.6183_{\pm 20}$	$0.6296_{\pm 26}$	$\mathbf{0.6327}_{\pm 21}$
	0.90	$\mathbf{0.6631}_{\pm 25}$	$0.6183_{\pm 20}$	$0.6529_{\pm 26}$	$0.6297_{\pm 14}$
TOYS	1.00	$0.6319_{\pm 15}$	$0.6318_{\pm 22}$	$0.6275_{\pm 19}$	$\mathbf{0.6560}_{\pm 12}$
	0.98	$\mathbf{0.6647}_{\pm 12}$	$0.6463_{\pm 20}$	$0.6519_{\pm 15}$	$0.6560_{\pm 12}$
	0.95	$0.6647_{\pm 11}$	$0.6521_{\pm 19}$	$0.6597_{\pm 18}$	$\mathbf{0.6741}_{\pm 14}$
	0.90	$0.6818_{\pm 15}$	$0.6599_{\pm 17}$	$\mathbf{0.6854}_{\pm 16}$	$0.6597_{\pm 14}$
SPORTS	1.00	$0.4951_{\pm 60}$	$0.4926_{\pm 58}$	$0.4926_{\pm 82}$	$\mathbf{0.5108}_{\pm 59}$
	0.98	$0.4964_{\pm 60}$	$0.4972_{\pm 54}$	$0.4937_{\pm 59}$	$\mathbf{0.5108}_{\pm 59}$
	0.95	$\mathbf{0.5138}_{\pm 68}$	$0.5009_{\pm 59}$	$0.5011_{\pm 56}$	$0.5132_{\pm 55}$
	0.90	$\mathbf{0.5358}_{\pm 44}$	$0.5055_{\pm 51}$	$0.5093_{\pm 58}$	$0.5290_{\pm 47}$
YELP	1.00	$0.5366_{\pm 94}$	$0.5371_{\pm 96}$	$0.5375_{\pm 92}$	$\mathbf{0.5882}_{\pm 88}$
	0.98	$0.5516_{\pm 96}$	$0.5378_{\pm 92}$	$0.5387_{\pm 95}$	$\mathbf{0.5882}_{\pm 85}$
	0.95	$0.5875_{\pm 13}$	$0.5381_{\pm 96}$	$0.5424_{\pm 94}$	$\mathbf{0.5940}_{\pm 91}$
	0.90	$0.5875_{\pm 13}$	$0.5381_{\pm 96}$	$0.5452_{\pm 92}$	$\mathbf{0.6099}_{\pm 93}$
GOWALLA	1.00	$0.8046_{\pm 20}$	$0.8037_{\pm 19}$	$0.8043_{\pm 16}$	$\mathbf{0.8077}_{\pm 14}$
	0.98	$0.8054_{\pm 23}$	$0.8037_{\pm 16}$	$0.8109_{\pm 15}$	$\mathbf{0.8143}_{\pm 19}$
	0.95	$0.8099_{\pm 20}$	$0.8056_{\pm 19}$	$0.8151_{\pm 18}$	$\mathbf{0.8170}_{\pm 13}$
	0.90	$\mathbf{0.8241}_{\pm 14}$	$0.8096_{\pm 20}$	$0.8166_{\pm 12}$	$0.8170_{\pm 13}$
GOODREADS	1.00	$0.5683_{\pm 39}$	$0.5682_{\pm 36}$	$0.5687_{\pm 39}$	$\mathbf{0.6724}_{\pm 65}$
	0.98	$0.5724_{\pm 43}$	$0.5682_{\pm 36}$	$0.5687_{\pm 39}$	$\mathbf{0.6724}_{\pm 65}$
	0.95	$0.5765_{\pm 38}$	$0.5682_{\pm 36}$	$0.5687_{\pm 39}$	$\mathbf{0.6724}_{\pm 65}$
	0.90	$0.5783_{\pm 39}$	$0.5682_{\pm 36}$	$0.5687_{\pm 39}$	$\mathbf{0.6724}_{\pm 65}$

At $\tau{=}1.00$ the three baselines buy almost nothing, staying within 0.005 of the deterministic sampler on every corpus, while per-user ϵ gains on all seven, from $+0.3$ points on GOWALLA to $+10.4$ on GOODREADS. Once accuracy is genuinely for sale the ordering shifts, and by $\tau{=}0.90$ temperature leads on BEAUTY, SPORTS and GOWALLA and DPP on TOYS. Flattening the softmax will always reach further if one is willing to pay for it; the level moves dispersion while holding the conditional marginals, so the regime where the two differ in kind is the tight one, $\tau \geq 0.98$.

M Controlled Synthetic Corpus

We use this corpus wherever a claim needs the number of preference modes or the history length to be *known* rather than inferred, which on a public corpus neither is. The two quantities Prop. 7(a,b) ranges over are assigned here by the generator, so we can partition users by either one exactly and read a group optimum off the partition.

Catalog. We use $N{=}1{,}000$ items in 20 categories of 50, item i belonging to category $\lfloor (i-1)/50 \rfloor$. The category is the only thing that distinguishes an item: there are no features, no popularity skew and no temporal drift. A category is therefore a preference mode in the sense of Assumption 4, by construction rather than by estimation.

Users. We generate $20{,}000$ users independently. For each user we

1. draw the number of modes M from the fixed distribution $\Pr(M{=}1,2,3,5,8) = (0.20,\ 0.25,\ 0.25,\ 0.20,\ 0.10)$;

2. draw the history length T uniformly from $\{5, 10, 20, 40\}$;
3. pick M distinct categories without replacement, with mixture weights $w \sim \mathrm{Dirichlet}(\mathbf{1}_M)$ over them;
4. draw $T+2$ categories i.i.d. from w and, inside each drawn category, one item uniformly;
5. write the sequence as T history items followed by one validation item and one test item.

N VARYING HISTORY LENGTH AT FIXED MODEL

The synthetic corpus of App. M assigns the history length; this experiment asks whether the same trend survives on real data, where it is not assigned. We train one model per corpus on the *full* sequences, with no fixed-length window and no truncation at training, so unlike the other experiments the model is never specialised to a particular history length. At evaluation we group the test users by the length of their own history and take the mean per-user $\epsilon^\star$ within each group, reading the same frozen checkpoint for every group so that nothing but length varies along a curve. The fitted slope is negative on all four corpora, from -0.005 per item on YELP to -0.038 on TOYS, though the curves are noisier than on the controlled corpus, BEAUTY visibly so, which is what entangling length with tenure and activity costs.

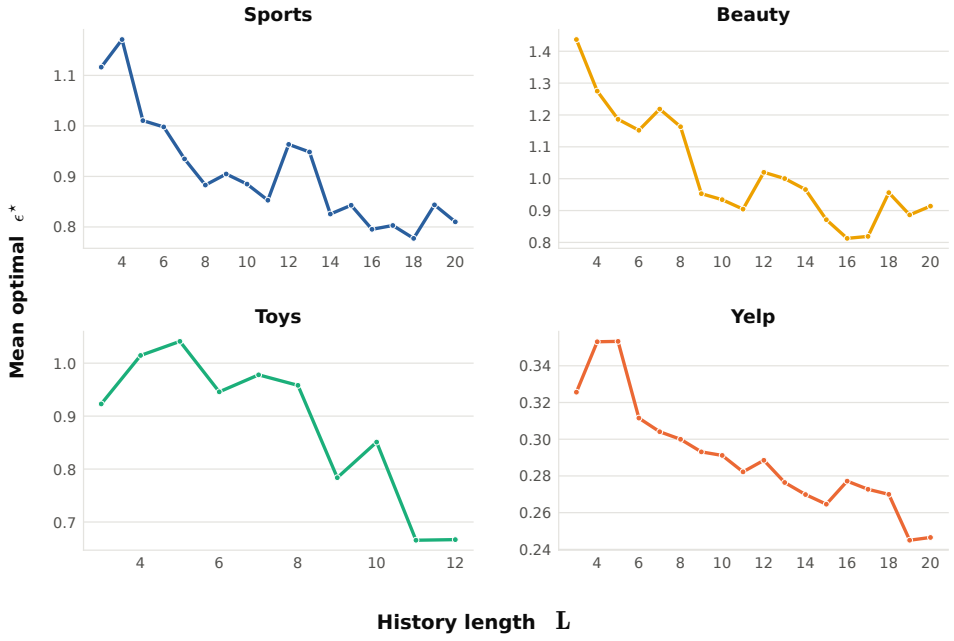


Figure 6: Mean per-user $\epsilon^\star$ against history length (written L on the axes, the T of Prop. 7(b)). One model per corpus trained on full sequences, with test users grouped by their own history length and read at the same frozen checkpoint. The fitted slope is negative on all four corpora.