# LEARNING WHEN TO TRUST: RISK-ADAPTIVE SUPERVISION FOR FLOW MAP MODELS

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

## ABSTRACT

Flow map models (FMMs) achieve efficient one-step generation by directly learning finite-time transport, yet existing training strategies largely treat supervision reliability as globally uniform. We challenge this assumption and show that the relative reliability of different supervision signals can vary substantially across transport states. To formalize this observation, we develop a flow-map-aware risk framework that characterizes state-dependent supervision reliability and derives the optimal trade-off between complementary supervision sources. Building on this framework, we propose **Risk-Adaptive Flow Map Learning (RAFL)**, which estimates state-dependent reliability from target-space disagreement and uses it to construct adaptive finite-time supervision. RAFL provides a lightweight post-training refinement for pretrained FMMs without modifying their inference procedure or introducing additional inference-time cost. Extensive experiments across diverse FMMs reveal substantial supervision heterogeneity and demonstrate consistent improvements over existing supervision strategies.

## 1 INTRODUCTION

Diffusion and flow-based generative models have achieved remarkable progress in high-fidelity generation, yet their practical efficiency remains constrained by iterative sampling (Ma et al., 2024; Liu et al., 2023; Yu et al., 2025). Recent *flow map models* (FMMs) address this limitation by directly learning finite-time transport (Frans et al., 2025; Wang et al., 2026; Hu et al., 2026a), compressing long generative trajectories into only one or a few neural function evaluations. MeanFlow (MF) (Geng et al., 2026a;b) provides a representative realization of this paradigm, while subsequent studies have explored improved optimization, representation reuse, distillation, and alternative formulations of flow-map learning (Lee et al., 2026; Fan et al., 2026; Sabour et al., 2026; Zheng et al., 2026; Luo et al., 2026; Tong et al., 2026). Despite this rapid progress, accurately learning long-range transport remains challenging because the available supervision can be imperfect and unevenly reliable across transport states (Zhang et al., 2026; Li et al., 2026).

When a pretrained FMM is available, flow-map learning can exploit two complementary supervision sources (Kim et al., 2026; Huang et al., 2026). The conditional flow-matching target is unbiased in expectation but can be highly stochastic for an individual transport state, whereas pretrained guidance is typically more stable but may contain state-dependent approximation error (Gao et al., 2026). Existing approaches usually balance these signals using globally shared rules, such as fixed mixtures, transport-dependent schedules, or staged supervision (Zhao et al., 2026; Zhang et al., 2026; Lu et al., 2026; Hu et al., 2026b). However, we find that states sharing the same transport condition can still prefer very different supervision sources. This exposes a limitation of globally shared supervision and raises a natural question: *how much should each supervision source be trusted at each transport state?*

We study this question from the perspective of the finite-time target optimized by an FMM. For MF, instantaneous supervision errors are transformed in a state-dependent manner before they affect the finite-time training target. This motivates a *flow-map-aware supervision risk* that captures both stochastic uncertainty in conditional supervision and approximation error in pretrained guidance. Our analysis derives the optimal state-dependent balance between the two sources and shows that globally shared policies become suboptimal whenever states sharing the same transport condition require different levels of trust. Because the ideal state-wise policy depends on local risk statistics



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.

that are not directly observable, we further study adaptive supervision under restricted information and characterize how much of the available adaptation opportunity can be captured by learned representations. This gives representation quality a task-specific interpretation: useful representations should distinguish states that require different supervision strategies.

Building on this analysis, we introduce **Risk-Adaptive Flow Map Learning (RAFL)**, a state-adaptive supervision framework for FMMs. RAFL measures disagreement between complementary supervision signals in the finite-time target space and uses learned representations to estimate their local reliability. These estimates determine a state-dependent supervision coefficient that controls how much each source contributes at each transport state. All RAFL-specific components are used only during training, making RAFL a lightweight post-training refinement that preserves the original sampling procedure and introduces no additional inference-time cost.

Our experiments support both the theoretical motivation and the practical effectiveness of RAFL. Across multiple FMM families, we observe substantial state-dependent supervision heterogeneity and find that learned representations capture a meaningful portion of this variation. Flow-map-aware reliability estimation consistently outperforms instantaneous-space alternatives, while risk-based adaptation improves over fixed mixtures, global schedules, staged supervision, and directly learned adaptive gates. Overall, RAFL consistently improves one-step generation with lightweight post-training refinement while preserving the original inference cost.

Our contributions are threefold: 1) We formulate supervision reliability in FMMs as a state-dependent risk problem and derive the optimal balance between stochastic conditional supervision and potentially biased pretrained guidance. 2) We characterize when state-adaptive supervision can outperform globally shared strategies and quantify how much of this adaptation opportunity can be captured by learned representations. 3) We introduce RAFL, a lightweight post-training framework that estimates state-dependent supervision reliability and consistently improves diverse pretrained FMMs without additional inference-time cost.

## 2 PRELIMINARIES

We briefly review Flow Matching (FM) and the finite-time supervision construction of FMMs, using MF as a representative formulation whose affine target structure forms the basis of our analysis.

**Flow Matching.** Let $p_0$ and $p_1$ denote the data and prior distributions, respectively. FM constructs an interpolated state $z_t = \alpha_t x + \sigma_t e$, with $x \sim p_0$, $e \sim p_1$, and conditional velocity $\hat{v}_t = \dot{\alpha}_t x + \dot{\sigma}_t e$. Its marginal velocity is $v^\star(z_t, t) = \mathbb{E}[\hat{v}_t \mid z_t, t]$. For a transport state $\xi = (z_t, r, t)$ with $0 \le r < t \le 1$, we write $v^\star_\xi := v^\star(z_t, t)$ and $\hat{v}_\xi = v^\star_\xi + \zeta_\xi$, where $\mathbb{E}[\zeta_\xi \mid \xi] = 0$. Thus, $\hat{v}_\xi$ provides unbiased but stochastic instantaneous supervision.

**Flow-Map Models.** FMMs directly learn finite-time transport between arbitrary time points. MF defines the average velocity over $[r, t]$ as

$$u(z_t, r, t) = \frac{1}{t-r} \int_r^t v(z_\tau, \tau)\, d\tau. \tag{1}$$

Differentiating with respect to $t$ gives $u = v - (t-r)\frac{du}{dt}$ and $\frac{du}{dt} = \nabla_{z_t} u\, v + \partial_t u$. MF parameterizes $u$ with $u_\theta$ and constructs the target using a frozen branch $u_{\bar{\theta}}$, where $\bar{\theta} = \mathrm{sg}(\theta)$. Replacing the unavailable marginal velocity with $\hat{v}_\xi$ yields the standard MF regression target, which we rewrite below in a form convenient for our analysis.

For our analysis, define $J_\xi := \nabla_{z_t} u_{\bar{\theta}}(z_t, r, t)$ and $c_\xi := \partial_t u_{\bar{\theta}}(z_t, r, t)$. The finite-time target induced by an arbitrary instantaneous supervision signal $w$ is

$$\mathcal{T}_\xi(w) = w - (t-r)(J_\xi w + c_\xi).$$

Hence, with $A_\xi := I - (t-r)J_\xi$,

$$\boxed{\mathcal{T}_\xi(w_1) - \mathcal{T}_\xi(w_2) = A_\xi(w_1 - w_2)}. \tag{2}$$

This affine relation characterizes how instantaneous supervision errors propagate into finite-time target space and provides the local geometry used in Sec. 3.

**Complementary Supervision.** Besides the stochastic conditional signal $\hat{v}_\xi$, we consider pretrained guidance $v_{\phi,\xi} = v^\star_\xi + b_\xi$, where $b_\xi$ is the state-dependent teacher bias. Thus, $\hat{v}_\xi$ is unbiased but stochastic, whereas $v_{\phi,\xi}$ is deterministic conditioned on $\xi$ but may be biased. Because their relative reliability can vary across transport states, a globally shared combination rule can be suboptimal, motivating the flow-map-aware, state-adaptive supervision analysis in Sec. 3.

## 3 Theoretical Analysis

We develop a risk-based theory for state-adaptive supervision in flow-map learning. Building on the affine finite-time target structure from Sec. 2, we characterize when state adaptation improves over global mixing and how much of this opportunity can be recovered from a given representation. Our analysis connects local supervision risk, information-constrained adaptation, and representation-dependent risk reduction under a fixed target-network snapshot.

### 3.1 Flow-Map-Aware Supervision Risk

Supervision reliability should be evaluated by the error it induces in the finite-time target optimized by an FMM, rather than solely in instantaneous velocity space. From Eq. (2), for an instantaneous supervision error $\eta$, $\mathcal{T}_\xi(v^\star_\xi + \eta) - \mathcal{T}_\xi(v^\star_\xi) = A_\xi \eta$. This induces the state-dependent metric $G_\xi := A_\xi^\top A_\xi$ and $\left\|\mathcal{T}_\xi(v^\star_\xi + \eta) - \mathcal{T}_\xi(v^\star_\xi)\right\|_2^2 = \eta^\top G_\xi \eta$. Recall from Sec. 2 that $\hat{v}_\xi = v^\star_\xi + \zeta_\xi$, $\mathbb{E}[\zeta_\xi \mid \xi] = 0$, and $v_{\phi,\xi} = v^\star_\xi + b_\xi$. We define the corresponding flow-map-aware risks as $V_\xi := \mathbb{E}\left[\zeta_\xi^\top G_\xi \zeta_\xi \mid \xi\right]$ and $B_\xi := b_\xi^\top G_\xi b_\xi$. Here, $V_\xi$ measures transformed stochastic uncertainty, whereas $B_\xi$ measures the squared teacher bias after propagation into the finite-time target space.

For a teacher coefficient $\lambda \in [0,1]$, consider $w_{\lambda,\xi} = (1-\lambda)\hat{v}_\xi + \lambda v_{\phi,\xi}$, with conditional flow-map risk

$$R_\xi(\lambda) := \mathbb{E}\left[\left\|\mathcal{T}_\xi(w_{\lambda,\xi}) - \mathcal{T}_\xi(v^\star_\xi)\right\|_2^2 \mid \xi\right].$$

**Proposition 3.1.** *For every transport state $\xi$, $R_\xi(\lambda) = (1-\lambda)^2 V_\xi + \lambda^2 B_\xi$. Let $W_\xi := V_\xi + B_\xi$. Whenever $W_\xi > 0$, the unique risk-minimizing coefficient and the corresponding minimum risk are*

$$\lambda^\star_\xi = \frac{V_\xi}{V_\xi + B_\xi}, \qquad R_\xi(\lambda^\star_\xi) = \frac{V_\xi B_\xi}{V_\xi + B_\xi}.$$

*When $W_\xi = 0$, $R_\xi(\lambda) = 0$ for all $\lambda$; we set $\lambda^\star_\xi = 0$ by convention. Moreover,*

$$\boxed{R_\xi(\lambda) = R_\xi(\lambda^\star_\xi) + W_\xi(\lambda - \lambda^\star_\xi)^2}. \tag{3}$$

Proposition 3.1 reveals a state-wise uncertainty–bias trade-off: large $V_\xi$ favors pretrained guidance, whereas large $B_\xi$ favors the unbiased conditional signal. Eq. (3) further shows that coefficient mismatch is more costly at states with larger total supervision risk $W_\xi$.

### 3.2 Information-Constrained Supervision

The oracle $\lambda^\star_\xi$ depends on state-conditional risk statistics that are not directly observable from a single training sample. A practical supervision policy must therefore operate under restricted information.

Let $\kappa := (r,t)$ denote the globally shared transport condition, and let $\mathcal{I}$ denote the information accessible to the policy, satisfying $\sigma(\kappa) \subseteq \mathcal{I} \subseteq \sigma(\xi)$. For an $\mathcal{I}$-measurable coefficient $\lambda_\mathcal{I} \in [0,1]$, define $R^\star_\mathcal{I}(\kappa) := \inf_{\lambda_\mathcal{I}} \mathbb{E}\left[R_\xi(\lambda_\mathcal{I}) \mid \kappa\right]$, and $R^\star_{\text{local}}(\kappa) := \mathbb{E}\left[R_\xi(\lambda^\star_\xi) \mid \kappa\right]$.

**Theorem 3.2.** *Assume $\mathbb{E}[W_\xi \mid \mathcal{I}] > 0$ almost surely. Then the optimal $\mathcal{I}$-measurable coefficient is*

$$\boxed{\lambda^\star_\mathcal{I} = \frac{\mathbb{E}[V_\xi \mid \mathcal{I}]}{\mathbb{E}[V_\xi + B_\xi \mid \mathcal{I}]} = \frac{\mathbb{E}[W_\xi \lambda^\star_\xi \mid \mathcal{I}]}{\mathbb{E}[W_\xi \mid \mathcal{I}]}}. \tag{4}$$

*Its excess risk relative to fully state-dependent supervision is*

$$R^\star_{\mathcal{I}}(\kappa) - R^\star_{\text{local}}(\kappa) = \mathbb{E}\left[W_\xi \left(\lambda^\star_\xi - \lambda^\star_{\mathcal{I}}\right)^2 \mid \kappa\right]. \tag{5}$$

Theorem 3.2 gives adaptive supervision a weighted-projection interpretation: the optimal policy under restricted information is the $W_\xi$-weighted conditional projection of the state-wise oracle. Hence, restricted information incurs excess risk precisely when it groups together states requiring different risk-optimal trust.

### 3.3 Supervision Heterogeneity and Representation Resolution

We now specialize Theorem 3.2 to the information hierarchy relevant to RAFL: $\mathcal{I}_{\text{global}} = \sigma(\kappa)$, $\mathcal{I}_h = \sigma(\kappa, h_\xi)$, and $\mathcal{I}_{\text{local}} = \sigma(\xi)$, where $h_\xi = \mathcal{H}(\xi)$ denotes a transport representation. The corresponding optimal global and representation-conditioned coefficients are:

$$\lambda^\star_\kappa = \frac{\mathbb{E}[V_\xi \mid \kappa]}{\mathbb{E}[V_\xi + B_\xi \mid \kappa]}, \qquad \lambda^\star_h = \frac{\mathbb{E}[V_\xi \mid h_\xi, \kappa]}{\mathbb{E}[V_\xi + B_\xi \mid h_\xi, \kappa]}.$$

**Corollary 3.3.** *The information hierarchy implies* $R^\star_{\text{local}}(\kappa) \le R^\star_h(\kappa) \le R^\star_{\text{global}}(\kappa)$. *Define the* supervision heterogeneity*:*

$$\boxed{\mathcal{H}_{\text{sup}}(\kappa) := R^\star_{\text{global}}(\kappa) - R^\star_{\text{local}}(\kappa) = \mathbb{E}\left[W_\xi(\lambda^\star_\xi - \lambda^\star_\kappa)^2 \mid \kappa\right].} \tag{6}$$

*Then*

$$\begin{aligned}\mathcal{H}_{\text{sup}}(\kappa) &= \underbrace{R^\star_{\text{global}}(\kappa) - R^\star_h(\kappa)}_{\textit{resolved by } h_\xi} + \underbrace{R^\star_h(\kappa) - R^\star_{\text{local}}(\kappa)}_{\textit{unresolved by } h_\xi} \\ &= \underbrace{\mathbb{E}\left[W_\xi(\lambda^\star_h - \lambda^\star_\kappa)^2 \mid \kappa\right]}_{\textit{resolved by } h_\xi} + \underbrace{\mathbb{E}\left[W_\xi(\lambda^\star_\xi - \lambda^\star_h)^2 \mid \kappa\right]}_{\textit{unresolved by } h_\xi}.\end{aligned} \tag{7}$$

Eq. (6) identifies the exact opportunity for state adaptation within the considered supervision mixture family under a fixed target-network snapshot. In particular, $\mathcal{H}_{\text{sup}}(\kappa) = 0 \iff W_\xi(\lambda^\star_\xi - \lambda^\star_\kappa)^2 = 0$, a.s. given $\kappa$. Equivalently, all states with $W_\xi > 0$ share the same risk-optimal trust almost surely under a common transport condition. When $\mathcal{H}_{\text{sup}}(\kappa) > 0$, state-dependent information can further reduce supervision risk. Eq. (7) separates this opportunity into the portion resolved by $h_\xi$ and the residual gap to fully local adaptation. Whenever $\mathcal{H}_{\text{sup}}(\kappa) > 0$, define:

$$\boxed{\rho_h(\kappa) := \frac{R^\star_{\text{global}}(\kappa) - R^\star_h(\kappa)}{R^\star_{\text{global}}(\kappa) - R^\star_{\text{local}}(\kappa)} = \frac{\mathbb{E}\left[W_\xi(\lambda^\star_h - \lambda^\star_\kappa)^2 \mid \kappa\right]}{\mathcal{H}_{\text{sup}}(\kappa)} \in [0, 1].}$$

Thus, $\rho_h(\kappa) \approx 0$ indicates that the representation resolves little supervision-relevant heterogeneity beyond the transport condition, whereas $\rho_h(\kappa) \approx 1$ indicates that it captures most of the available adaptation opportunity. A useful representation for adaptive supervision is therefore one that distinguishes states requiring different risk-optimal trust.

**Design Principle.** The analysis identifies the representation-conditioned quantities required for adaptive supervision:

$$\boxed{\mathbb{E}[V_\xi \mid h_\xi, \kappa], \qquad \mathbb{E}[B_\xi \mid h_\xi, \kappa] \quad \Longrightarrow \quad \lambda^\star_h.}$$

In the next section, we develop RAFL to construct estimable surrogates for these quantities from target-space disagreement and connect them to the optimal representation-conditioned policy.

Proofs of Proposition 3.1, Theorem 3.2, and Corollary 3.3, together with the general nested-information decomposition, are deferred to the appendix.

# 4 RISK-ADAPTIVE FLOW MAP LEARNING

Sec. 3 identifies $\lambda_h^\star$ as the optimal supervision policy under representation-level information, but the required conditional risks are not directly observable. RAFL approximates this policy through

$$\lambda_\xi^\star \;\to\; \lambda_h^\star \;\to\; \lambda_{\mathrm{cal}}^\star \;\to\; \hat{\lambda}_\xi,$$

progressing from the state-wise oracle to the representation-conditioned optimum, a disagreement-calibrated surrogate, and finally an amortized training coefficient. RAFL operates in finite-time target space. We use MF for explicit derivations because its affine target transformation enables exact computation of target-space disagreement with a Jacobian–vector product. The exact derivation is given for MF, while RAFL applies to other FMMs by constructing disagreement in their corresponding finite-time supervision spaces.

## 4.1 TARGET-SPACE RELIABILITY ESTIMATION

Let $v_{\phi,\xi}$ denote the frozen pretrained instantaneous guidance introduced in Sec. 2, and let $H_\eta$ be a frozen representation encoder. For each transport state $\xi = (z_t, r, t)$, we extract $h_\xi = \mathrm{sg}\left[\mathrm{Pool}\left(H_\eta^\ell(z_t, t)\right)\right]$. The encoder $H_\eta$ may use features from the pretrained generative backbone or an external frozen representation model. We study several choices in Table 5. Together with $\kappa = (r, t)$, $h_\xi$ specifies the information available to the representation-conditioned policy.

During training, both the stochastic conditional signal $\hat{v}_\xi$ and pretrained guidance $v_{\phi,\xi}$ are available. We measure their disagreement after the finite-time target transformation: $\delta_\xi := \mathcal{T}_\xi(\hat{v}_\xi) - \mathcal{T}_\xi(v_{\phi,\xi})$. For MF, Eq. (2) gives $\delta_\xi = A_\xi(\hat{v}_\xi - v_{\phi,\xi}) = A_\xi(\zeta_\xi - b_\xi)$. Conditioned on the exact transport state, $\mathbb{E}[\delta_\xi \mid \xi] = -A_\xi b_\xi$, and therefore $\|\mathbb{E}[\delta_\xi \mid \xi]\|_2^2 = B_\xi$, $\mathrm{Tr}\left(\mathrm{Cov}[\delta_\xi \mid \xi]\right) = V_\xi$. Thus, the conditional mean and variation of target-space disagreement recover the transformed teacher bias and stochastic uncertainty that determine the oracle coefficient in Proposition 3.1.

For efficient computation, let $d_\xi := \hat{v}_\xi - v_{\phi,\xi}$. Since $A_\xi = I - (t-r)J_\xi$, we have $\delta_\xi = d_\xi - (t-r)J_\xi d_\xi$, which requires only a Jacobian–vector product.

**Representation-conditioned calibration.** Exact-state moments cannot be estimated from a single continuous training sample. We therefore condition disagreement statistics on the information $(h_\xi, \kappa)$ available to the representation-level policy: $B_{\mathrm{cal}}(h,\kappa) := \|\mathbb{E}[\delta_\xi \mid h_\xi = h, \kappa]\|_2^2$, $V_{\mathrm{cal}}(h,\kappa) := \mathrm{Tr}\left(\mathrm{Cov}[\delta_\xi \mid h_\xi = h, \kappa]\right)$.

States sharing the same representation and transport condition may still exhibit different transformed teacher biases. Define $C_b(h,\kappa) := \mathrm{Tr}\left(\mathrm{Cov}[A_\xi b_\xi \mid h_\xi = h, \kappa]\right)$. By the law of total covariance, $V_{\mathrm{cal}}(h,\kappa) = \mathbb{E}[V_\xi \mid h_\xi = h, \kappa] + C_b(h,\kappa)$, and $B_{\mathrm{cal}}(h,\kappa) = \mathbb{E}[B_\xi \mid h_\xi = h, \kappa] - C_b(h,\kappa)$. Their sum is preserved: $V_{\mathrm{cal}}(h,\kappa) + B_{\mathrm{cal}}(h,\kappa) = \mathbb{E}[W_\xi \mid h_\xi = h, \kappa]$.

The resulting disagreement-calibrated coefficient is

$$\lambda_{\mathrm{cal}}^\star(h,\kappa) := \frac{V_{\mathrm{cal}}(h,\kappa)}{V_{\mathrm{cal}}(h,\kappa) + B_{\mathrm{cal}}(h,\kappa)}.$$

From Theorem 3.2,

$$\lambda_h^\star(h,\kappa) = \frac{\mathbb{E}[V_\xi \mid h_\xi = h, \kappa]}{\mathbb{E}[W_\xi \mid h_\xi = h, \kappa]}.$$

Hence,

$$\boxed{\lambda_{\mathrm{cal}}^\star(h,\kappa) - \lambda_h^\star(h,\kappa) = \frac{C_b(h,\kappa)}{\mathbb{E}[W_\xi \mid h_\xi = h, \kappa]}.} \tag{8}$$

Thus, disagreement calibration exactly recovers the optimal representation-conditioned policy when $C_b(h,\kappa) = 0$. More generally, its excess risk relative to $\lambda_h^\star$ is $\frac{C_b(h,\kappa)^2}{\mathbb{E}[W_\xi \mid h_\xi = h, \kappa]}$, so the induced risk penalty is second-order in unresolved transformed teacher-bias variation; see appendix.

## 4.2 PRACTICAL RISK ESTIMATION

**Local calibration.** The conditional moments above are unavailable directly. We maintain a bounded FIFO buffer of recent tuples $(h_j, r_j, t_j, \delta_j)$ and, for each query state $i$, retrieve $K$ nearby samples

$\mathcal{N}_i$ under compatible transport conditions using cosine similarity in the frozen representation space. The normalized neighborhood weights are

$$a_{ij} = \frac{\exp(\mathrm{sim}(h_i, h_j)/\tau)}{\sum_{k\in\mathcal{N}_i} \exp(\mathrm{sim}(h_i, h_k)/\tau)}, \qquad \mathrm{sim}(h_i, h_j) = \frac{h_i^\top h_j}{\|h_i\|_2 \|h_j\|_2}.$$

Define $s_i = \sum_{j\in\mathcal{N}_i} a_{ij}^2$ and $\hat{\mu}_i = \sum_{j\in\mathcal{N}_i} a_{ij}\delta_j$. We estimate the calibrated risks by

$$\widehat{V}_i^{\mathrm{cal}} = \frac{\sum_{j\in\mathcal{N}_i} a_{ij}\|\delta_j - \hat{\mu}_i\|_2^2}{1 - s_i},$$

and

$$\widehat{B}_i^{\mathrm{cal}} = \left[\|\hat{\mu}_i\|_2^2 - s_i \widehat{V}_i^{\mathrm{cal}}\right]_+,$$

where $[x]_+ := \max(x, 0)$. The correction removes the finite-sample bias of the weighted moment estimators under local stationarity; finite neighborhoods may still introduce representation-space smoothing bias. Derivations and retrieval details are provided in the appendix.

**Amortized risk prediction.** Neighborhood calibration provides local reliability targets but need not be queried for every flow-map update. We therefore amortize these estimates with a lightweight predictor $(\widetilde{V}_\xi, \widetilde{B}_\xi) = \mathrm{Softplus}\,(Q_\psi(h_\xi, r, t))$, where Softplus ensures non-negative predictions. The predictor is trained against the calibrated targets in log space:

$$\mathcal{L}_{\mathrm{rel}} = \ell\left(\log(\widetilde{V}_\xi + \varepsilon), \mathrm{sg}\left[\log(\widehat{V}_\xi^{\mathrm{cal}} + \varepsilon)\right]\right) + \ell\left(\log(\widetilde{B}_\xi + \varepsilon), \mathrm{sg}\left[\log(\widehat{B}_\xi^{\mathrm{cal}} + \varepsilon)\right]\right),$$

where $\ell$ is a robust regression loss and $\varepsilon > 0$ is a numerical constant.

Following the risk-optimal form in Theorem 3.2, the predicted risks determine

$$\boxed{\hat{\lambda}_\xi = \frac{\widetilde{V}_\xi}{\widetilde{V}_\xi + \widetilde{B}_\xi + \varepsilon}.} \tag{9}$$

Larger predicted stochastic uncertainty increases reliance on pretrained guidance, whereas larger predicted teacher bias shifts supervision toward the conditional signal. Unlike an unconstrained adaptive gate, $\hat{\lambda}_\xi$ is explicitly parameterized through estimated supervision risks.

### 4.3 Risk-Adaptive Flow-Map Training

RAFL converts the predicted coefficient into the adaptive instantaneous supervision signal $w_\xi^{\mathrm{RAFL}} = (1 - \bar{\lambda}_\xi)\hat{v}_\xi + \bar{\lambda}_\xi v_{\phi,\xi}$, $\bar{\lambda}_\xi = \mathrm{sg}[\hat{\lambda}_\xi]$. Stopping gradients through $\hat{\lambda}_\xi$ decouples reliability estimation from flow-map regression and prevents the generative objective from directly turning the reliability predictor into an unconstrained gating function.

The corresponding finite-time supervision target is $y_\xi^{\mathrm{RAFL}} = \mathcal{T}_\xi\left(w_\xi^{\mathrm{RAFL}}\right)$. For the affine MF target transformation, $y_\xi^{\mathrm{RAFL}} = (1 - \bar{\lambda}_\xi)\mathcal{T}_\xi(\hat{v}_\xi) + \bar{\lambda}_\xi\mathcal{T}_\xi(v_{\phi,\xi})$, so the same state-dependent coefficient directly controls the balance between complementary supervision in finite-time target space.

The FMM predictor is optimized with

$$\mathcal{L}_{\mathrm{RAFL}} = \mathbb{E}_\xi\left[\omega(r,t)\left\|u_\theta(z_t, r, t) - \mathrm{sg}\left[y_\xi^{\mathrm{RAFL}}\right]\right\|_2^2\right], \tag{10}$$

where $\omega(r, t)$ follows the weighting of the underlying flow-map objective.

**Joint optimization.** The complete training objective is

$$\boxed{\mathcal{L} = \mathcal{L}_{\mathrm{RAFL}} + \beta\mathcal{L}_{\mathrm{rel}}.} \tag{11}$$

The pretrained guidance model and representation encoder remain frozen throughout training. Thus, $Q_\psi$ is trained only through $\mathcal{L}_{\mathrm{rel}}$ to predict representation-conditioned supervision reliability, while $u_\theta$ is optimized through $\mathcal{L}_{\mathrm{RAFL}}$ using the resulting risk-adaptive targets.

**Inference.** RAFL introduces no additional inference-time components. After training, the pretrained guidance model, calibration buffer, and reliability predictor are discarded, leaving only the optimized FMM. RAFL therefore preserves the original sampling procedure and inference cost. For MF, the optimized model uses $z_r = z_t - (t - r)u_{\theta^\star}(z_t, r, t)$, with one-step generation obtained by setting $r = 0$, $t = 1$, and sampling the initial state from $p_1$.

Table 1: **System-level evaluation on class-conditional ImageNet-256.** We evaluate RAFL across representative FMMs, including MF (Geng et al., 2026a), MF-RAE (Hu et al., 2026b), iMF (Geng et al., 2026b), Shortcut Models (Frans et al., 2025), TiM (Wang et al., 2026), and $\alpha$-Flow (Zhang et al., 2026). For MF, we use the pretrained checkpoint released by $\alpha$-Flow. When sampling budgets are unavailable in the original reports, only RAFL results are shown to characterize NFE scaling.

| Models | FID↓ | NFE↓ | Iter. | Batch | #Params |
|---|---|---|---|---|---|
| MF-B/2 | 6.04 | 1 | 300k | 256 | 131M |
| MF-B/2 | 5.17 | 2 | 300k | 256 | 131M |
| +RAFL | 5.81 | 1 | +20k | 256 | 131M |
| +RAFL | 4.97 | 2 | +20k | 256 | 131M |
| +RAFL | 4.91 | 4 | +20k | 256 | 131M |
| +RAFL | **4.88** | 8 | +20k | 256 | 131M |
| MF-XL/2 | 3.47 | 1 | 300k | 256 | 675M |
| MF-XL/2 | 2.46 | 2 | 300k | 256 | 675M |
| +RAFL | 3.01 | 1 | +20k | 256 | 675M |
| +RAFL | 2.17 | 2 | +20k | 256 | 675M |
| +RAFL | 2.06 | 4 | +20k | 256 | 675M |
| +RAFL | **2.02** | 8 | +20k | 256 | 675M |
| MF-RAE | 2.03 | 1 | 1207k | 256 | 841M |
| MF-RAE | 1.89 | 2 | 1207k | 256 | 841M |
| +RAFL | 1.91 | 1 | +30k | 256 | 841M |
| +RAFL | 1.80 | 2 | +30k | 256 | 841M |
| +RAFL | 1.76 | 4 | +30k | 256 | 841M |
| +RAFL | **1.73** | 8 | +30k | 256 | 841M |
| iMF-XL/2 | 1.72 | 1 | 1000k | 1024 | 610M |
| iMF-XL/2 | 1.54 | 2 | 1000k | 1024 | 610M |
| +RAFL | 1.63 | 1 | +40k | 1024 | 610M |
| +RAFL | 1.49 | 2 | +40k | 1024 | 610M |
| +RAFL | 1.45 | 4 | +40k | 1024 | 610M |
| +RAFL | **1.42** | 8 | +40k | 1024 | 610M |

| Models | FID↓ | NFE↓ | Iter. | Batch | #Params |
|---|---|---|---|---|---|
| Shortcut-XL/2 | 10.60 | 1 | 800k | 256 | 675M |
| Shortcut-XL/2 | 7.80 | 4 | 800k | 256 | 675M |
| +RAFL | 9.78 | 1 | +30k | 256 | 675M |
| +RAFL | 8.39 | 2 | +30k | 256 | 675M |
| +RAFL | 7.26 | 4 | +30k | 256 | 675M |
| +RAFL | **6.57** | 8 | +30k | 256 | 675M |
| TiM-XL/2 | 7.11 | 1 | 750k | 512 | 664M |
| TiM-XL/2 | 6.14 | 2 | 750k | 512 | 664M |
| TiM-XL/2 | 3.61 | 4 | 750k | 512 | 664M |
| TiM-XL/2 | 2.62 | 8 | 750k | 512 | 664M |
| +RAFL | 6.74 | 1 | +35k | 512 | 664M |
| +RAFL | 5.96 | 2 | +35k | 512 | 664M |
| +RAFL | 3.52 | 4 | +35k | 512 | 664M |
| +RAFL | **2.48** | 8 | +35k | 512 | 664M |
| $\alpha$-Flow-XL/2 | 2.95 | 1 | 1200k | 256 | 676M |
| $\alpha$-Flow-XL/2 | 2.16 | 2 | 1200k | 256 | 676M |
| +RAFL | 2.81 | 1 | +40k | 256 | 676M |
| +RAFL | 2.02 | 2 | +40k | 256 | 676M |
| +RAFL | 1.98 | 4 | +40k | 256 | 676M |
| +RAFL | **1.95** | 8 | +40k | 256 | 676M |
| $\alpha$-Flow-XL/2+ | 2.58 | 1 | 1275k | 1024 | 676M |
| $\alpha$-Flow-XL/2+ | 1.95 | 2 | 1275k | 1024 | 676M |
| +RAFL | 2.46 | 1 | +35k | 1024 | 676M |
| +RAFL | 1.87 | 2 | +35k | 1024 | 676M |
| +RAFL | 1.83 | 4 | +35k | 1024 | 676M |
| +RAFL | **1.81** | 8 | +35k | 1024 | 676M |

## 5 EXPERIMENTS

We evaluate RAFL by addressing the following questions: **(1)** Does risk-adaptive supervision consistently improve diverse FMMs? **(2)** Is supervision heterogeneity practically significant and resolvable from learned representations? **(3)** Does state-adaptive risk modeling outperform globally shared and unconstrained supervision policies? **(4)** Is flow-map-aware risk estimation necessary? **(5)** How much does each RAFL component contribute to performance? Our experiments are organized around system-level evaluation, supervision heterogeneity analysis, policy comparison, flow-map-aware risk ablation, and component analysis.

**Experimental Setup.** We evaluate RAFL on CIFAR-10 and class-conditional ImageNet-256 using FID and IS, with emphasis on one-step generation and the quality–efficiency trade-off. To assess generality, we apply RAFL to representative FMMs via lightweight fine-tuning of pretrained checkpoints while preserving their original sampling procedures. All RAFL-specific components are discarded at inference, introducing no additional inference-time overhead. Unless otherwise stated, controlled analyses in Tables 3–6 and Figures 1–2 use the same pretrained MF-XL/2 checkpoint, a batch size of 256, and 20k fine-tuning iterations under identical optimization budgets.

**System-Level Evaluation.** Tables 1 and 2 show that RAFL improves diverse FMMs across datasets and sampling budgets without altering their inference procedures. On ImageNet-256, RAFL improves one-step FID across all evaluated backbones, e.g., from 3.47 to 3.01 for MF-XL/2 and from 1.72 to 1.63 for iMF-XL/2, with similar gains at higher NFEs. On CIFAR-10, RAFL improves pretrained MF from 2.92 to 2.57 FID using 15k additional iterations. These results show that RAFL provides an effective and lightweight post-training strategy across diverse FMM formulations.

**Table 2: Performance on unconditional CIFAR-10.**

| Model | Precond. | NFE↓ | FID↓ |
|---|---|---|---|
| iCT | EDM | 1 | 2.83 |
| ECT | EDM | 1 | 3.60 |
| sCT | EDM | 1 | 2.97 |
| IMM | EDM | 1 | 3.20 |
| MF | none | 1 | 2.92 |
| **MF+RAFL** | MF | 1 | **2.57** |

**Table 3: Ablation of flow-map-aware reliability estimation on ImageNet-256.** All variants share the same representation, calibration, and risk predictor.

| **Disagreement** | $\hat{R}\downarrow$ | FID ↓ | IS ↑ |
|---|---|---|---|
| Instantaneous velocity $\|\hat{v}_\xi - v_{\phi,\xi}\|$ | 0.1086 | 3.17 | 258.9 |
| Target-space, Jacobian-free | 0.1001 | 3.09 | 262.4 |
| **Flow-map-aware** $\|A_\xi(\hat{v}_\xi - v_{\phi,\xi})\|$ | **0.0942** | **3.01** | **265.2** |

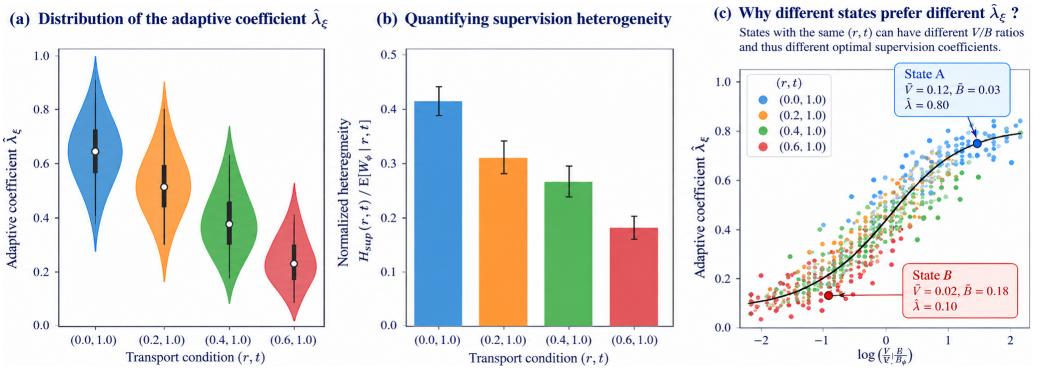



Figure 1: **Supervision heterogeneity across transport states.** (a) Adaptive coefficients vary substantially even among states sharing the same transport condition $(r, t)$. (b) Supervision heterogeneity remains significant across transport conditions, revealing an opportunity for state-dependent adaptation. (c) This variation is explained by the local uncertainty-bias trade-off: states with the same $(r, t)$ can have different risk ratios and therefore favor different supervision coefficients.

Table 4: **Comparison of supervision policies on ImageNet-256.**

| Supervision policy | Policy information | FID ↓ | IS ↑ | $\hat{R}_h\downarrow$ |
|---|---|---|---|---|
| Conditional only ($\lambda = 0$) | Global | 3.36 | 251.7 | 0.1218 |
| Teacher only ($\lambda = 1$) | Global | 3.31 | 253.4 | 0.1185 |
| Best fixed mixture | Global | 3.18 | 258.6 | 0.1097 |
| Global schedule | $(r, t)$ | 3.13 | 260.8 | 0.1051 |
| Staged supervision | Training progress | 3.11 | 261.6 | 0.1029 |
| Direct adaptive gate | $(h_\xi, r, t)$ | 3.07 | 263.4 | 0.0987 |
| **RAFL** | Risk from $(h_\xi, r, t)$ | **3.01** | **265.2** | **0.0942** |

**Supervision Heterogeneity.** Figure 1 shows state-wise variation in adaptive coefficients even under the same $(r, t)$, driven by uncertainty–bias profiles. This indicates that transport conditions alone are insufficient to characterize supervision reliability, motivating state-dependent adaptation.

**Flow-Map-Aware Risk.** Table 3 isolates the effect of reliability geometry while keeping the representation, calibration, and predictor fixed. Compared with instantaneous-space disagreement, flow-map-aware disagreement reduces FID from 3.17 to 3.01 and estimated risk from 0.1086 to 0.0942, while improving IS from 258.9 to 265.2. This supports evaluating supervision reliability in the finite-time target space rather than solely in instantaneous velocity space.

**Supervision Policies.** Table 4 compares policies under the same initialization and budget. Performance generally improves as supervision incorporates increasingly state-specific information. A direct adaptive gate reaches FID 3.07, while RAFL further improves to 3.01 with lower estimated risk (0.0942 vs. 0.0987). Under this controlled setting, explicitly structuring adaptation through estimated risk is more effective than learning an unconstrained gate.

**Representation Choice.** Table 5 shows that MF backbone features already outperform the raw state, indicating that RAFL does not require an external discriminative encoder. DINOv2 generally resolves more supervision-relevant heterogeneity and yields better generation quality, with DINOv2-

Table 5: **Component-wise analysis on ImageNet-256.** All results are reported under 1-NFE sampling. We study the effects of representation family, model scale, feature depth, and retrieval metric on supervision heterogeneity resolution and generation performance.

| Iter. | Target Repr. | Depth | Retrieval Metric | $\widehat{\rho}_h \uparrow$ | $\widehat{R}_h \downarrow$ | $\mathrm{Gflops}_{\mathrm{Infer.}} \downarrow$ | FID ↓ | IS ↑ |
|---|---|---|---|---|---|---|---|---|
| 20k | Raw state $x_t$ | – | Cos. sim. | 0.196 | 0.1151 | 119.0 | 3.39 | 250.6 |
| 20k | MF-XL/2 | 6 | Cos. sim. | 0.536 | 0.1012 | 119.0 | 3.17 | 258.4 |
| 20k | MF-XL/2 | 8 | Cos. sim. | 0.602 | 0.0979 | 119.0 | 3.09 | 261.2 |
| 20k | MF-XL/2 | 10 | Cos. sim. | 0.571 | 0.0994 | 119.0 | 3.13 | 259.8 |
| 20k | MoCov3-L | 8 | Cos. sim. | 0.487 | 0.1027 | 119.0 | 3.20 | 256.1 |
| 20k | CLIP-L | 8 | Cos. sim. | 0.594 | 0.0981 | 119.0 | 3.10 | 260.9 |
| 20k | DINOv2-B | 8 | Cos. sim. | 0.552 | 0.0999 | 119.0 | 3.15 | 258.7 |
| 20k | DINOv2-g | 8 | Cos. sim. | 0.651 | 0.0956 | 119.0 | 3.05 | 263.7 |
| 20k | DINOv2-L | 6 | Cos. sim. | 0.576 | 0.0989 | 119.0 | 3.13 | 260.3 |
| 20k | DINOv2-L | 8 | Cos. sim. | **0.684** | **0.0942** | 119.0 | **3.01** | **265.2** |
| 20k | DINOv2-L | 10 | Cos. sim. | 0.642 | 0.0960 | 119.0 | 3.06 | 263.6 |
| 20k | DINOv2-L | 12 | Cos. sim. | 0.538 | 0.1005 | 119.0 | 3.14 | 259.7 |
| 20k | DINOv2-L | 14 | Cos. sim. | 0.447 | 0.1044 | 119.0 | 3.23 | 256.4 |
| 20k | DINOv2-L | 16 | Cos. sim. | 0.361 | 0.1081 | 119.0 | 3.31 | 253.8 |
| 20k | DINOv2-L | 6 | Euclidean | 0.548 | 0.1006 | 119.0 | 3.17 | 258.8 |
| 20k | DINOv2-L | 8 | Euclidean | 0.623 | 0.0968 | 119.0 | 3.08 | 262.1 |
| 20k | DINOv2-L | 10 | Euclidean | 0.591 | 0.0982 | 119.0 | 3.12 | 260.7 |

Table 6: **Component ablation of RAFL on ImageNet-256.** $\rho_\lambda$ denotes the Spearman correlation between the predicted coefficient and an *offline disagreement-calibrated reference* constructed from a disjoint held-out calibration bank that is never used to optimize $Q_\psi$. The reference is used only as a diagnostic of policy alignment, while FID and IS provide independent downstream evaluation.

| Variant | Local | Finite-sample correction | $Q_\psi$ | SG | $\rho_\lambda \uparrow$ | FID↓ | IS↑ |
|---|---|---|---|---|---|---|---|
| Global calibration | – | – | ✓ | ✓ | 0.61 | 3.19 | 257.8 |
| w/o finite-sample correction | ✓ | – | ✓ | ✓ | 0.76 | 3.10 | 261.7 |
| w/o amortization | ✓ | ✓ | – | ✓ | 0.81 | 3.06 | 263.1 |
| w/o stop-gradient | ✓ | ✓ | ✓ | – | 0.72 | 3.13 | 260.4 |
| **RAFL** | ✓ | ✓ | ✓ | ✓ | **0.85** | **3.01** | **265.2** |

L at depth 8 achieving the best $\widehat{\rho}_h = 0.684$, risk $0.0942$, and FID $3.01$. This suggests that DINOv2 is preferred empirically for its more effective feature geometry, rather than because RAFL depends on a particular encoder family.

**Component Analysis.** Table 6 evaluates each RAFL component. Replacing local calibration with global statistics causes the largest degradation, reducing $\rho_\lambda$ from $0.85$ to $0.61$ and worsening FID from $3.01$ to $3.19$. Removing finite-sample correction, amortization, or stop-gradient also degrades alignment and generation quality, supporting their complementary roles.

## 6 Conclusion

This work reframes supervision in flow-map learning as a state-dependent reliability problem rather than a globally uniform one. Our analysis shows that complementary supervision signals exhibit heterogeneous uncertainty–bias trade-offs across transport states, creating a clear opportunity for adaptive supervision. Building on this insight, we introduce RAFL, which estimates representation-conditioned supervision reliability and converts it into risk-adaptive finite-time supervision. Across diverse FMMs, RAFL consistently improves generation quality while preserving the inference procedure and cost. More broadly, our results suggest that explicitly modeling supervision heterogeneity is a promising principle for improving the reliability and effectiveness of flow-map learning.

## AI USE STATEMENT

In this work, generative AI tools were used to polish the grammar and improve the overall readability of the manuscript, as well as to check language consistency and logical flow throughout the paper. The tools were used only for language editing and presentation refinement and were not involved in developing the scientific ideas, methodology, experimental design, analyses, or conclusions. All scientific content was developed by the authors, and all AI-assisted revisions were carefully reviewed and verified by the authors, who take full responsibility for the final manuscript.

## REFERENCES

Nicholas Boffi, Michael Albergo, and Eric Vanden-Eijnden. How to build a consistency model: Learning flow maps via self-distillation. *Advances in Neural Information Processing Systems*, 38: 33346–33382, 2026.

Hansheng Chen, Kai Zhang, Hao Tan, Leonidas Guibas, Gordon Wetzstein, and Sai Bi. pi-flow: Policy-based few-step generation via imitation distillation. In *International Conference on Learning Representations*, volume 2026, pp. 151521–151547, 2026.

Xiangyu Fan, Zesong Qiu, Zhuguanyu Wu, Fanzhou Wang, Zhiqian Lin, Tianxiang Ren, Dahua Lin, Ruihao Gong, and Lei Yang. Phased DMD: Few-step distribution matching distillation via score matching within subintervals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 41667–41676, 2026.

Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models. In *International Conference on Learning Representations*, 2025.

Huanlin Gao, Ping Chen, Fuyuan Shi, Ruijia Wu, Li YanTao, Qiang Hui, Ting Lu, Chao Tan, Shaoan Zhao, Zhaoxiang Liu, et al. MeanCache: From instantaneous to average velocity for accelerating flow matching inference. In *International Conference on Learning Representations*, volume 2026, pp. 60466–60485, 2026.

Zhengyang Geng, Mingyang Deng, Xingjian Bai, Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. *Advances in Neural Information Processing Systems*, 38:75460–75482, 2026a.

Zhengyang Geng, Yiyang Lu, Zongze Wu, Eli Shechtman, J Zico Kolter, and Kaiming He. Improved Mean Flows: On the challenges of fastforward generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 30467–30476, 2026b.

Yi Guo, Wei Wang, Zhihang Yuan, Rong Cao, Kuan Chen, Zhengyang Chen, Yuanyuan Huo, Yang Zhang, Yuping Wang, Shouda Liu, et al. Splitmeanflow: Interval splitting consistency in few-step generative modeling. *arXiv preprint arXiv:2507.16884*, 2025.

Zheyuan Hu, Chieh-Hsin Lai, Yuki Mitsufuji, and Stefano Ermon. CMT: Mid-training for efficient learning of consistency, mean flow, and flow-map models. In *International Conference on Learning Representations*, 2026a.

Zheyuan Hu, Chieh-Hsin Lai, Ge Wu, Yuki Mitsufuji, and Stefano Ermon. Meanflow transformers with representation autoencoders. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 25709–25718, 2026b.

Yuhao Huang, Shih-Hsin Wang, Andrea Bertozzi, and Bao Wang. RMFlow: Refined mean flow by a noise-injection step for multimodal generation. In *International Conference on Learning Representations*, volume 2026, pp. 76267–76286, 2026.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 35:26565–26577, 2022.

Jin-Young Kim, Hyojun Go, Lea Bogensperger, Julius Erbach, Nikolai Kalischek, Federico Tombari, Konrad Schindler, and Dominik Narnhofer. Understanding, accelerating, and improving meanflow training. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 37992–38003, 2026.

Kyungmin Lee, Sihyun Yu, and Jinwoo Shin. Decoupled MeanFlow: Turning flow models into flow maps for accelerated sampling. In *International Conference on Learning Representations*, 2026.

Xingjian Leng, Jaskirat Singh, Yunzhong Hou, Zhenchang Xing, Saining Xie, and Liang Zheng. REPA-e: Unlocking vae for end-to-end tuning of latent diffusion transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 18262–18272, 2025.

Zhiqi Li, Yuchen Sun, Greg Turk, and Bo Zhu. Functional mean flow in hilbert space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1928–1938, 2026.

Xingchao Liu, Xiwen Zhang, Jianzhu Ma, Jian Peng, et al. Instaflow: One step is enough for high-quality diffusion-based text-to-image generation. In *International Conference on Learning Representations*, 2023.

Cheng Lu and Yang Song. Simplifying, stabilizing and scaling continuous-time consistency models. In *International Conference on Learning Representations*, volume 2025, pp. 50611–50649, 2025.

Yiyang Lu, Susie Lu, Qiao Sun, Hanhong Zhao, Zhicheng Jiang, Xianbang Wang, Tianhong Li, Zhengyang Geng, and Kaiming He. One-step latent-free image generation with pixel mean flows. *arXiv preprint arXiv:2601.22158*, 2026.

Simian Luo, Yiqin Tan, Longbo Huang, Jian Li, and Hang Zhao. Latent consistency models: Synthesizing high-resolution images with few-step inference, 2023.

Tianze Luo, Haotian Yuan, and Zhuang Liu. SoFlow: Solution flow models for one-step generative modeling. In *International Conference on Learning Representations*, volume 2026, pp. 101623–101641, 2026.

Nanye Ma, Mark Goldstein, Michael S Albergo, Nicholas M Boffi, Eric Vanden-Eijnden, and Saining Xie. SiT: Exploring flow and diffusion-based generative models with scalable interpolant transformers. In *European Conference on Computer Vision*, pp. 23–40. Springer, 2024.

Amirmojtaba Sabour, Sanja Fidler, and Karsten Kreis. Align your flow: Scaling continuous-time flow map distillation. *Advances in Neural Information Processing Systems*, 38:146459–146512, 2026.

Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. In *International Conference on Learning Representations*, 2022.

Jaskirat Singh, Xingjian Leng, Zongze Wu, Liang Zheng, Richard Zhang, Eli Shechtman, and Saining Xie. What matters for representation alignment: Global information or spatial structure? In *The Fourteenth International Conference on Learning Representations*, 2026.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv:2010.02502*, October 2020.

Yang Song and Prafulla Dhariwal. Improved techniques for training consistency models. In *The Twelfth International Conference on Learning Representations*, 2024.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *International Conference on Machine Learning*, pp. 32211–32252. PMLR, 2023.

Shangyuan Tong, Nanye Ma, Saining Xie, and Tommi Jaakkola. Flow map distillation without data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 33973–33984, 2026.

Zidong Wang, Yiyuan Zhang, Xiaoyu Yue, Xiangyu Yue, Yangguang Li, Wanli Ouyang, and Lei Bai. Transition models: Rethinking the generative learning objective. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 29178–29189, 2026.

Tianwei Yin, Michaël Gharbi, Taesung Park, Richard Zhang, Eli Shechtman, Fredo Durand, and William T Freeman. Improved distribution matching distillation for fast image synthesis. In *NeurIPS*, 2024a.

Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Frédo Durand, William T Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *CVPR*, 2024b.

Sihyun Yu, Sangkyung Kwak, Huiwon Jang, Jongheon Jeong, Jonathan Huang, Jinwoo Shin, and Saining Xie. Representation alignment for generation: Training diffusion transformers is easier than you think. In *International Conference on Learning Representations*, 2025.

Huijie Zhang, Aliaksandr Siarohin, Willi Menapace, Michael Vasilkovsky, Sergey Tulyakov, Qing Qu, and Ivan Skorokhodov. AlphaFlow: Understanding and improving meanflow models. In *International Conference on Learning Representations*, 2026.

An Zhao, Shengyuan Zhang, Zhongjian Sun, Yixiang Zhou, Zejian Li, Ling Yang, Tianrun Chen, and Lingyun Sun. Mean flow distillation: Robust and stable distillation for flow matching models. *arXiv preprint arXiv:2606.11155*, 2026.

Kaiwen Zheng, Yuji Wang, Qianli Ma, Huayu Chen, Jintao Zhang, Yogesh Balaji, Jianfei Chen, Ming-Yu Liu, Jun Zhu, and Qinsheng Zhang. Large scale diffusion distillation via score-regularized continuous-time consistency. In *International Conference on Learning Representations*, volume 2026, pp. 2582–2603, 2026.

# A Appendix

## A.1 Related Works

**Diffusion and Flow-Based Generative Models.** Diffusion and flow-based generative models have become dominant paradigms for high-fidelity generation by learning continuous transformations between simple reference distributions and complex data distributions, but typically require iterative numerical integration at inference time. This has motivated extensive research on reducing sampling cost through improved parameterizations (Song et al., 2020; Karras et al., 2022), distillation (Salimans & Ho, 2022; Yin et al., 2024b;a; Zheng et al., 2026; Chen et al., 2026), consistency-based objectives (Song et al., 2023; Luo et al., 2023; Song & Dhariwal, 2024; Lu & Song, 2025), and trajectory compression (Liu et al., 2023; Frans et al., 2025). While these approaches primarily aim to shorten the generative trajectory, our work focuses on a complementary question: how reliable are the supervision signals used to learn long-range transport, and how should their relative influence vary across transport states?

**Flow-map models and long-range transport.** Recent works increasingly formulate efficient generation as directly learning finite-time transport rather than repeatedly approximating instantaneous dynamics. MeanFlow (MF) (Geng et al., 2026a;b) learns average velocity over finite intervals and enables high-quality generation with one or few neural function evaluations. Building on this paradigm, subsequent studies have investigated improved optimization and training strategies (Hu et al., 2026a; Zhang et al., 2026; Luo et al., 2026), alternative target constructions (Frans et al., 2025; Guo et al., 2025; Boffi et al., 2026; Wang et al., 2026), representation-based architecture and feature reuse (Hu et al., 2026b; Lee et al., 2026), and flow-map distillation and scalable training (Sabour et al., 2026; Wang et al., 2026) for such models. These works primarily improve the formulation, architecture, or optimization of finite-time transport itself. In contrast, RAFL is largely orthogonal to the underlying flow-map parameterization: it addresses the reliability of supervision used to construct the finite-time target, allowing the contribution of complementary supervision sources to vary across states while leaving the inference procedure unchanged.

**Representation-guided generative learning** Pretrained representations have also been increasingly used to facilitate generative modeling. Representation alignment methods (Yu et al., 2025; Leng et al., 2025; Singh et al., 2026) encourage generative features to match pretrained semantic features, while other works (Hu et al., 2026b; Sabour et al., 2026; Zheng et al., 2026; Fan et al., 2026) reuse representations from pretrained models to improve optimization, distillation, or sampling efficiency. These results suggest that pretrained representations capture structure useful beyond their original discriminative or generative objectives. RAFL exploits such representations for a different purpose. Rather than using them as additional generative features, we use them to identify transport states with similar supervision reliability. Our analysis gives this role a task-specific interpretation:

a useful representation is one that resolves states requiring different risk-optimal levels of trust in the available supervision signals. The representation therefore determines how much of the state-wise supervision heterogeneity can be exploited by an adaptive policy.

### A.2 Theoretical Derivations

In this section, we provide derivations for Proposition 3.1, Theorem 3.2, and Corollary 3.3. Throughout, the target-network snapshot is fixed, so $A_\xi$ and $G_\xi = A_\xi^\top A_\xi$ are deterministic conditioned on the transport state $\xi$.

**Proof of Proposition 3.1.** Recall

$$\hat{v}_\xi = v_\xi^\star + \zeta_\xi, \qquad v_{\phi,\xi} = v_\xi^\star + b_\xi, \qquad \mathbb{E}[\zeta_\xi \mid \xi] = 0.$$

For

$$w_{\lambda,\xi} = (1-\lambda)\hat{v}_\xi + \lambda v_{\phi,\xi},$$

we have

$$w_{\lambda,\xi} - v_\xi^\star = (1-\lambda)\zeta_\xi + \lambda b_\xi.$$

Using Eq. (2),

$$\mathcal{T}_\xi(w_{\lambda,\xi}) - \mathcal{T}_\xi(v_\xi^\star) = A_\xi\left[(1-\lambda)\zeta_\xi + \lambda b_\xi\right].$$

Therefore,

$$\begin{aligned} R_\xi(\lambda) &= \mathbb{E}\left[\left((1-\lambda)\zeta_\xi + \lambda b_\xi\right)^\top G_\xi \left((1-\lambda)\zeta_\xi + \lambda b_\xi\right) \mid \xi\right] \\ &= (1-\lambda)^2 V_\xi + \lambda^2 B_\xi + 2\lambda(1-\lambda)\mathbb{E}\left[\zeta_\xi^\top G_\xi b_\xi \mid \xi\right]. \end{aligned} \tag{12}$$

Since $G_\xi$ and $b_\xi$ are fixed conditioned on $\xi$,

$$\mathbb{E}\left[\zeta_\xi^\top G_\xi b_\xi \mid \xi\right] = \mathbb{E}[\zeta_\xi \mid \xi]^\top G_\xi b_\xi = 0.$$

Hence

$$R_\xi(\lambda) = (1-\lambda)^2 V_\xi + \lambda^2 B_\xi.$$

Let $W_\xi = V_\xi + B_\xi$. When $W_\xi > 0$, differentiating gives

$$\frac{\partial R_\xi(\lambda)}{\partial \lambda} = -2V_\xi + 2W_\xi\lambda,$$

so the unique minimizer is

$$\lambda_\xi^\star = \frac{V_\xi}{W_\xi} = \frac{V_\xi}{V_\xi + B_\xi}.$$

Since $V_\xi, B_\xi \geq 0$, we have $\lambda_\xi^\star \in [0,1]$. Completing the square yields

$$\begin{aligned} R_\xi(\lambda) &= \frac{V_\xi B_\xi}{W_\xi} + W_\xi(\lambda - \lambda_\xi^\star)^2 \\ &= R_\xi(\lambda_\xi^\star) + W_\xi(\lambda - \lambda_\xi^\star)^2, \end{aligned} \tag{13}$$

which proves Proposition 3.1. □

**Evolution of the Optimal Supervision Coefficient.** Along a transport trajectory parameterized by $s$, the optimal coefficient $\lambda_\xi^\star = V_\xi/(V_\xi + B_\xi)$ satisfies

$$\frac{d\lambda_\xi^\star}{ds} = \frac{B_\xi\, dV_\xi/ds - V_\xi\, dB_\xi/ds}{(V_\xi + B_\xi)^2}.$$

Therefore, decreasing noise or increasing representation consistency does not by itself imply monotonic convergence of $\lambda_\xi^\star$. Its evolution depends on the relative changes of stochastic uncertainty and teacher bias, and may increase, decrease, or remain stable across different transport states.

**Proof of Theorem 3.2.** From Proposition 3.1, for any $\mathcal{I}$-measurable coefficient $\lambda_{\mathcal{I}}$,

$$R_\xi(\lambda_{\mathcal{I}}) = R_\xi(\lambda_\xi^\star) + W_\xi(\lambda_{\mathcal{I}} - \lambda_\xi^\star)^2.$$

Taking the conditional expectation given $c$ gives

$$\mathbb{E}[R_\xi(\lambda_{\mathcal{I}}) \mid c] = R^\star_{\text{local}}(c) + \mathbb{E}\left[W_\xi(\lambda_{\mathcal{I}} - \lambda_\xi^\star)^2 \mid c\right]. \tag{14}$$

Because $\sigma(c) \subseteq \mathcal{I}$, the tower property gives

$$\mathbb{E}\left[W_\xi(\lambda_{\mathcal{I}} - \lambda_\xi^\star)^2 \mid c\right] = \mathbb{E}\left[\mathbb{E}\left[W_\xi(\lambda_{\mathcal{I}} - \lambda_\xi^\star)^2 \mid \mathcal{I}\right] \mid c\right].$$

For fixed $\mathcal{I}$, the inner term is

$$\mathbb{E}[W_\xi \mid \mathcal{I}]\lambda_{\mathcal{I}}^2 - 2\mathbb{E}[W_\xi\lambda_\xi^\star \mid \mathcal{I}]\lambda_{\mathcal{I}} + \mathbb{E}[W_\xi(\lambda_\xi^\star)^2 \mid \mathcal{I}]. \tag{15}$$

Under $\mathbb{E}[W_\xi \mid \mathcal{I}] > 0$, this quadratic is uniquely minimized by

$$\lambda_{\mathcal{I}}^\star = \frac{\mathbb{E}[W_\xi\lambda_\xi^\star \mid \mathcal{I}]}{\mathbb{E}[W_\xi \mid \mathcal{I}]}.$$

Since

$$W_\xi\lambda_\xi^\star = (V_\xi + B_\xi)\frac{V_\xi}{V_\xi + B_\xi} = V_\xi,$$

we obtain

$$\lambda_{\mathcal{I}}^\star = \frac{\mathbb{E}[V_\xi \mid \mathcal{I}]}{\mathbb{E}[V_\xi + B_\xi \mid \mathcal{I}]}.$$

Moreover, $0 \le \lambda_{\mathcal{I}}^\star \le 1$ because $V_\xi, B_\xi \ge 0$.

Substituting $\lambda_{\mathcal{I}}^\star$ into Eq. (14) yields

$$R^\star_{\mathcal{I}}(c) - R^\star_{\text{local}}(c) = \mathbb{E}\left[W_\xi\left(\lambda_\xi^\star - \lambda_{\mathcal{I}}^\star\right)^2 \mid c\right],$$

which proves Theorem 3.2. □

**Weighted-projection interpretation.** The optimality condition above further gives

$$\mathbb{E}\left[W_\xi\left(\lambda_\xi^\star - \lambda_{\mathcal{I}}^\star\right) \mid \mathcal{I}\right] = 0. \tag{16}$$

Thus, $\lambda_{\mathcal{I}}^\star$ is the $W_\xi$-weighted conditional projection of the state-wise oracle $\lambda_\xi^\star$ onto the information contained in $\mathcal{I}$.

**Proof of Corollary 3.3.** Consider

$$\mathcal{I}_{\text{global}} = \sigma(c), \qquad \mathcal{I}_h = \sigma(c, h_\xi), \qquad \mathcal{I}_{\text{local}} = \sigma(c, \xi),$$

with

$$\mathcal{I}_{\text{global}} \subseteq \mathcal{I}_h \subseteq \mathcal{I}_{\text{local}}.$$

Enlarging the available information cannot increase the minimum achievable risk, hence

$$R^\star_{\text{local}}(c) \le R^\star_h(c) \le R^\star_{\text{global}}(c).$$

Applying Theorem 3.2 with $\mathcal{I} = \sigma(c)$ gives

$$R^\star_{\text{global}}(c) - R^\star_{\text{local}}(c) = \mathbb{E}\left[W_\xi(\lambda_\xi^\star - \lambda_c^\star)^2 \mid c\right],$$

which is exactly $\mathcal{H}_{\text{sup}}(c)$.

To obtain the representation decomposition, write

$$\lambda_\xi^\star - \lambda_c^\star = (\lambda_\xi^\star - \lambda_h^\star) + (\lambda_h^\star - \lambda_c^\star).$$

Expanding the weighted square yields

$$\mathcal{H}_{\text{sup}}(c) = \mathbb{E}\left[W_\xi(\lambda_\xi^\star - \lambda_h^\star)^2 \mid c\right] + \mathbb{E}\left[W_\xi(\lambda_h^\star - \lambda_c^\star)^2 \mid c\right] + 2\,\mathbb{E}\left[W_\xi(\lambda_\xi^\star - \lambda_h^\star)(\lambda_h^\star - \lambda_c^\star) \mid c\right].$$

The cross term vanishes. Indeed, since $\lambda_h^\star - \lambda_c^\star$ is $\mathcal{I}_h$-measurable,

$$\begin{aligned}&\mathbb{E}\left[W_\xi(\lambda_\xi^\star - \lambda_h^\star)(\lambda_h^\star - \lambda_c^\star) \mid c\right] \\ &\quad = \mathbb{E}\left[(\lambda_h^\star - \lambda_c^\star)\,\mathbb{E}\left[W_\xi(\lambda_\xi^\star - \lambda_h^\star) \mid \mathcal{I}_h\right] \mid c\right] = 0,\end{aligned} \tag{17}$$

where the last equality follows from Eq. (16). Therefore,

$$\mathcal{H}_{\sup}(c) = \mathbb{E}\left[W_\xi(\lambda_h^\star - \lambda_c^\star)^2 \mid c\right] + \mathbb{E}\left[W_\xi(\lambda_\xi^\star - \lambda_h^\star)^2 \mid c\right].$$

The first term is the heterogeneity resolved by $h_\xi$, while the second is the unresolved portion. Consequently,

$$R^\star_{\text{global}}(c) - R^\star_h(c) = \mathbb{E}\left[W_\xi(\lambda_h^\star - \lambda_c^\star)^2 \mid c\right],$$

and, whenever $\mathcal{H}_{\sup}(c) > 0$,

$$\rho_h(c) = \frac{R^\star_{\text{global}}(c) - R^\star_h(c)}{R^\star_{\text{global}}(c) - R^\star_{\text{local}}(c)} = \frac{\mathbb{E}[W_\xi(\lambda_h^\star - \lambda_c^\star)^2 \mid c]}{\mathcal{H}_{\sup}(c)}.$$

Since

$$R^\star_{\text{local}}(c) \le R^\star_h(c) \le R^\star_{\text{global}}(c),$$

we immediately obtain $\rho_h(c) \in [0, 1]$. □

**Excess risk of disagreement calibration.** We next characterize the risk induced by replacing the optimal representation-conditioned coefficient $\lambda_h^\star$ with the disagreement-calibrated coefficient $\lambda^\star_{\text{cal}}$.

**Corollary A.1** (Excess risk of disagreement calibration)**.** *Let*

$$\mathcal{R}_z(\lambda) := \mathbb{E}[R_\xi(\lambda) \,|\, z_\xi = z],$$

*and assume* $\mathbb{E}[W_\xi \mid z_\xi = z] > 0$*. Then the excess risk of the calibrated coefficient relative to the optimal representation-conditioned coefficient is*

$$\mathcal{R}_z(\lambda^\star_{\text{cal}}) - \mathcal{R}_z(\lambda_h^\star) = \frac{C_b(z)^2}{\mathbb{E}[W_\xi \mid z_\xi = z]}.$$

*Proof.* From Theorem 3.2, the optimal coefficient under information $z_\xi$ is

$$\lambda_h^\star(z) = \frac{\mathbb{E}[V_\xi \mid z_\xi = z]}{\mathbb{E}[W_\xi \mid z_\xi = z]},$$

and the corresponding conditional risk admits the quadratic decomposition

$$\mathcal{R}_z(\lambda) = \mathcal{R}_z(\lambda_h^\star) + \mathbb{E}[W_\xi \mid z_\xi = z]\,(\lambda - \lambda_h^\star)^2.$$

From Eq. (8), we have

$$\lambda^\star_{\text{cal}}(z) - \lambda_h^\star(z) = \frac{C_b(z)}{\mathbb{E}[W_\xi \mid z_\xi = z]}.$$

Substituting this relation into the quadratic decomposition gives

$$\begin{aligned}&\mathcal{R}_z(\lambda^\star_{\text{cal}}) - \mathcal{R}_z(\lambda_h^\star) \\ &\quad = \mathbb{E}[W_\xi \mid z_\xi = z]\left(\frac{C_b(z)}{\mathbb{E}[W_\xi \mid z_\xi = z]}\right)^2 \\ &\quad = \frac{C_b(z)^2}{\mathbb{E}[W_\xi \mid z_\xi = z]},\end{aligned} \tag{18}$$

which proves the result. □

Corollary A.1 shows that the discrepancy between disagreement calibration and the optimal representation-conditioned policy is controlled entirely by the transformed teacher-bias variation that remains unresolved within the representation class. In particular, when $C_b(z) = 0$, disagreement calibration exactly recovers $\lambda_h^\star$. More generally, while the coefficient deviation in Eq. (8) is linear in $C_b(z)$, the resulting excess supervision risk is quadratic in this unresolved component. Equivalently, writing $\eta(z) = C_b(z)/\mathbb{E}[W_\xi \mid z]$ for the coefficient deviation, the excess risk is $\mathbb{E}[W_\xi \mid z]\eta(z)^2$. Thus, the risk penalty is second-order in the calibration error

**General nested-information decomposition.** More generally, let $\mathcal{I}_1 \subseteq \mathcal{I}_2$ be two information sets containing $\sigma(c)$, with corresponding optimal coefficients $\lambda^\star_{\mathcal{I}_1}$ and $\lambda^\star_{\mathcal{I}_2}$. The same weighted-orthogonality argument gives

$$\mathbb{E}\left[W_\xi(\lambda^\star_\xi - \lambda^\star_{\mathcal{I}_1})^2 \mid c\right] = \mathbb{E}\left[W_\xi(\lambda^\star_{\mathcal{I}_2} - \lambda^\star_{\mathcal{I}_1})^2 \mid c\right] + \mathbb{E}\left[W_\xi(\lambda^\star_\xi - \lambda^\star_{\mathcal{I}_2})^2 \mid c\right].$$

Hence, additional information reduces supervision risk exactly by the portion of weighted oracle variation resolved by that information. Corollary 3.3 follows by taking $\mathcal{I}_1 = \sigma(c)$ and $\mathcal{I}_2 = \sigma(c, h_\xi)$.

### A.3 More Experiments

We provide additional analyses of RAFL focusing on optimization dynamics, matched post-training controls, sampling-budget robustness, hyperparameter sensitivity, and robustness to training configurations. Unless otherwise specified, we follow the ImageNet-256 evaluation protocol used in the main paper.

**Optimization Dynamics.** Figure 2 examines how different supervision strategies evolve during fine-tuning. RAFL consistently achieves lower one-step FID than continued training, fixed mixtures, global schedules, and direct adaptive gating throughout optimization. The estimated supervision risk exhibits a similar trend, with RAFL decreasing more rapidly and remaining lower over the course of training. Moreover, the predicted trust coefficients of RAFL align substantially better with the calibrated reference than those produced by the direct adaptive gate ($\rho = 0.86$ vs. $0.63$; MAE $0.06$ vs. $0.12$). These results indicate that the performance gain of RAFL is closely associated with more accurate estimation of state-dependent supervision reliability, rather than merely increased flexibility of the supervision policy.

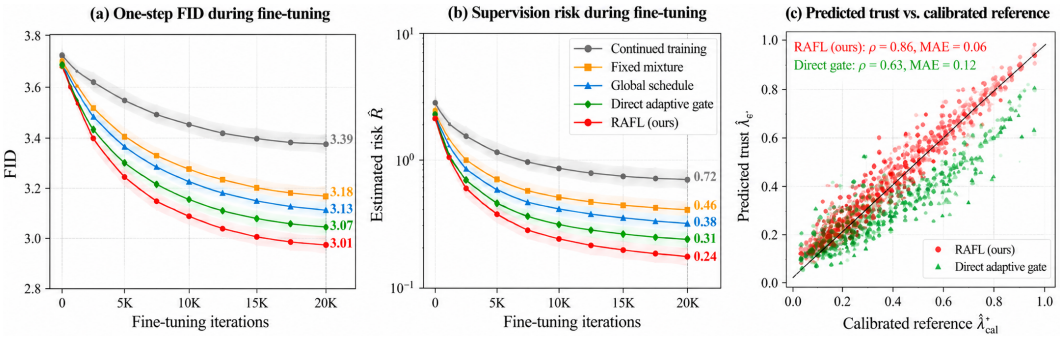



Figure 2: **Optimization dynamics under different supervision strategies.** (a) RAFL achieves consistently lower one-step FID throughout fine-tuning than continued training, fixed mixtures, global schedules, and direct adaptive gating. (b) RAFL also maintains the lowest estimated supervision risk during optimization, indicating more reliable target construction. (c) Its predicted trust coefficients align more closely with the calibrated reference than those of the direct adaptive gate ($\rho = 0.86$ vs. $0.63$; MAE $0.06$ vs. $0.12$). Together, these results connect improved generation quality with more accurate estimation of state-dependent supervision reliability.

**Hyperparameter Sensitivity.** Table 7 studies the sensitivity of RAFL to the neighborhood size $K$, similarity temperature $\tau$, calibration-buffer size $M$, and reliability-loss weight $\beta$. We vary one hyperparameter at a time while keeping the remaining settings fixed at their default values. RAFL remains stable across the evaluated ranges, with only moderate variation in policy alignment, estimated supervision risk, and generation quality. The default configuration consistently provides a favorable balance, indicating that RAFL does not rely on narrowly tuned hyperparameters.

**Batch Size and Training Budget.** Table 8 evaluates RAFL across batch sizes of 128, 256, and 512 under both 20k and 40k fine-tuning iterations. RAFL consistently improves over continued training under every configuration, showing that its effectiveness is preserved across optimization regimes. Performance improves as the batch size increases from 128 to 256 and remains comparable between

Table 7: **Hyperparameter sensitivity on ImageNet-256.** We vary one hyperparameter at a time while fixing the remaining settings to $K = 32$, $\tau = 0.1$, $M = 16$k, and $\beta = 1.0$.

| Setting | $\rho_\lambda \uparrow$ | $\widehat{R}_h \downarrow$ | FID $\downarrow$ | IS $\uparrow$ |
|---|---|---|---|---|
| $K = 16$ | 0.82 | 0.0967 | 3.06 | 263.8 |
| $K = 32$ | **0.85** | **0.0942** | **3.01** | **265.2** |
| $K = 64$ | 0.84 | 0.0950 | 3.03 | 264.6 |
| $\tau = 0.05$ | 0.83 | 0.0956 | 3.04 | 264.3 |
| $\tau = 0.10$ | **0.85** | **0.0942** | **3.01** | **265.2** |
| $\tau = 0.20$ | 0.81 | 0.0972 | 3.07 | 263.4 |
| $M = 8$k | 0.82 | 0.0964 | 3.05 | 263.9 |
| $M = 16$k | **0.85** | **0.0942** | **3.01** | **265.2** |
| $M = 32$k | 0.84 | 0.0948 | 3.02 | 264.8 |
| $\beta = 0.5$ | 0.82 | 0.0961 | 3.05 | 264.0 |
| $\beta = 1.0$ | **0.85** | **0.0942** | **3.01** | **265.2** |
| $\beta = 2.0$ | 0.80 | 0.0984 | 3.09 | 262.8 |

256 and 512. Extending fine-tuning from 20k to 40k iterations yields only marginal changes, indicating that performance saturates under the default setting. These results support the use of batch size 256 and 20k iterations for the controlled analyses in the main paper.

Table 8: **Effect of batch size and fine-tuning budget on MF-XL/2 on ImageNet-256.** We evaluate RAFL with batch sizes of 128, 256, and 512 under 20k and 40k fine-tuning iterations.

| Method | Batch | Iter. | $\widehat{R}_h \downarrow$ | FID $\downarrow$ | IS $\uparrow$ |
|---|---|---|---|---|---|
| MF-XL/2 + continued training | 128 | 20k | 0.1261 | 3.44 | 248.7 |
| MF-XL/2 + RAFL | 128 | 20k | 0.0981 | 3.10 | 262.7 |
| MF-XL/2 + continued training | 128 | 40k | 0.1243 | 3.42 | 249.8 |
| MF-XL/2 + RAFL | 128 | 40k | 0.0970 | 3.08 | 263.3 |
| MF-XL/2 + continued training | 256 | 20k | 0.1238 | 3.40 | 250.0 |
| MF-XL/2 + RAFL | 256 | 20k | **0.0942** | **3.01** | **265.2** |
| MF-XL/2 + continued training | 256 | 40k | 0.1229 | 3.39 | 250.6 |
| MF-XL/2 + RAFL | 256 | 40k | 0.0944 | 3.02 | 265.0 |
| MF-XL/2 + continued training | 512 | 20k | 0.1236 | 3.39 | 250.1 |
| MF-XL/2 + RAFL | 512 | 20k | 0.0945 | 3.02 | 265.0 |
| MF-XL/2 + continued training | 512 | 40k | 0.1228 | 3.38 | 250.8 |
| MF-XL/2 + RAFL | 512 | 40k | 0.0943 | 3.01 | 265.1 |

**Matched Post-Training Controls across Sampling Budgets.** Since RAFL is applied through post-training, we additionally control for the possibility that its gains arise simply from extra optimization. For MF-XL/2 and iMF-XL/2, we construct matched baselines that start from exactly the same pretrained checkpoints and use the same additional training budgets as RAFL. These controls include continued training with the original objective, fixed-mixture supervision, and alternative adaptive supervision strategies. We evaluate all resulting models under multiple sampling NFEs to determine whether the effect of RAFL is specific to one-step generation.

Table 10 shows that additional optimization alone accounts for only a limited portion of the improvement. For MF-XL/2, continued training improves the 1-NFE FID from 3.47 to 3.39, whereas RAFL reaches 3.01 under the same 20k-step post-training budget. The advantage remains at larger sampling budgets: RAFL achieves FIDs of 2.17, 2.06, and 2.02 with 2, 4, and 8 NFEs, respectively. Likewise, on the stronger iMF-XL/2 backbone, RAFL improves the original 1-NFE FID from 1.72 to 1.63 and reaches 1.49, 1.45, and 1.42 as the sampling budget increases. These results show that the observed gains cannot be explained solely by additional post-training and that risk-adaptive supervision improves the learned flow map across different inference budgets.

Table 9: **Matched continued-training controls across FMMs on class-conditional ImageNet-256.** To isolate the effect of additional optimization, we continue training each pretrained FMM with its original objective for exactly the same number of iterations as RAFL. The continued-training and RAFL variants are initialized from the same checkpoint and use matched optimization settings. We additionally include a fixed-mixture baseline to control for the benefit of generic teacher-assisted supervision. All results are reported with 1-NFE sampling.

| Backbone | Extra Iter. | Original FID ↓ | Continued Training ↓ | Fixed Mixture ↓ | RAFL ↓ |
|---|---|---|---|---|---|
| MF-B/2 | 20k | 6.04 | 5.96 | 5.90 | **5.81** |
| MF-XL/2 | 20k | 3.47 | 3.39 | 3.18 | **3.01** |
| MF-RAE | 30k | 2.03 | 1.99 | 1.95 | **1.91** |
| iMF-XL/2 | 40k | 1.72 | 1.70 | 1.67 | **1.63** |
| Shortcut-XL/2 | 30k | 10.60 | 10.41 | 10.08 | **9.78** |
| TiM-XL/2 | 35k | 7.11 | 7.01 | 6.88 | **6.74** |
| $\alpha$-Flow-XL/2 | 40k | 2.95 | 2.91 | 2.86 | **2.81** |
| $\alpha$-Flow-XL/2+ | 35k | 2.58 | 2.55 | 2.51 | **2.46** |

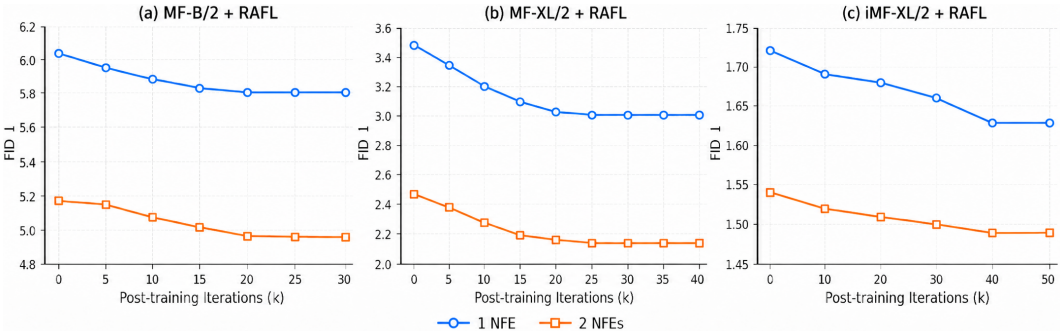



Figure 3: **Post-training behavior of RAFL across model scales.** FID steadily improves with additional RAFL fine-tuning for MF-B/2, MF-XL/2, and iMF-XL/2 under both 1-NFE and 2-NFE sampling. The gains become progressively smaller as training proceeds, indicating stable convergence and diminishing returns beyond the default fine-tuning budget.

Table 10: **Matched controls across different sampling NFEs on class-conditional ImageNet-256.** We compare RAFL with matched continued training and alternative supervision strategies under the same post-training budget. All variants are initialized from the same pretrained checkpoint and use matched optimization settings. Results are reported across different numbers of NFEs.

| Backbone | Method | Extra Iter. | 1 NFE ↓ | 2 NFE ↓ | 4 NFE ↓ | 8 NFE ↓ |
|---|---|---|---|---|---|---|
| MF-XL/2 | Original | – | 3.47 | 2.46 | – | – |
| | Continued training | 20k | 3.39 | 2.39 | 2.31 | 2.29 |
| | Best fixed mixture | 20k | 3.18 | 2.29 | 2.20 | 2.17 |
| | Global schedule | 20k | 3.13 | 2.25 | 2.15 | 2.12 |
| | Direct adaptive gate | 20k | 3.07 | 2.20 | 2.10 | 2.07 |
| | **RAFL** | 20k | **3.01** | **2.17** | **2.06** | **2.02** |
| iMF-XL/2 | Original | – | 1.72 | 1.54 | – | – |
| | Continued training | 40k | 1.70 | 1.52 | 1.50 | 1.48 |
| | Best fixed mixture | 40k | 1.67 | 1.51 | 1.48 | 1.46 |
| | **RAFL** | 40k | **1.63** | **1.49** | **1.45** | **1.42** |

**Post-Training Behavior Across Model Scales.** Figure 3 further examines RAFL fine-tuning on MF-B/2, MF-XL/2, and iMF-XL/2. Across model scales, FID improves smoothly with additional post-training under both 1-NFE and 2-NFE sampling, while the gains gradually diminish as optimization proceeds. This consistent behavior indicates stable convergence across different FMM capacities and suggests that RAFL does not depend on a narrowly selected post-training duration.

Table 11: **Run-to-run stability of RAFL on ImageNet-256 across model scales and sampling budgets.** FID is reported as mean ± standard deviation over ten independent fine-tuning runs initialized from the same pretrained checkpoint for each backbone. Only the training random seed is varied, while identical evaluation noise seeds are used across runs.

| Model | 1-NFE ↓ | 2-NFE ↓ | 4-NFE ↓ | 8-NFE ↓ |
|---|---|---|---|---|
| MF-B/2 + RAFL | $5.83 \pm 0.06$ | $4.99 \pm 0.05$ | $4.93 \pm 0.04$ | $4.89 \pm 0.04$ |
| MF-XL/2 + RAFL | $3.02 \pm 0.03$ | $2.18 \pm 0.03$ | $2.07 \pm 0.02$ | $2.03 \pm 0.02$ |
| iMF-XL/2 + RAFL | $1.64 \pm 0.03$ | $1.50 \pm 0.02$ | $1.46 \pm 0.02$ | $1.43 \pm 0.02$ |

**Run-to-Run Stability.** We assess the sensitivity of RAFL to training randomness on ImageNet-256 using ten independent fine-tuning runs for MF-B/2, MF-XL/2, and iMF-XL/2. All runs for a given backbone start from the same pretrained checkpoint and share identical training and evaluation settings, with only the random seed varied. As shown in Table 11, RAFL exhibits consistently low variance across model scales and sampling budgets. Together with the matched NFE comparisons in Table 10, these results suggest that RAFL is robust to training stochasticity and that its improvements persist from 1 to 8 NFEs.

**Summary.** Overall, these additional analyses demonstrate consistent RAFL behavior across optimization stages, hyperparameter choices, batch sizes, training budgets, model scales, random seeds, and sampling budgets. Together with the controlled comparisons in the main paper, they support RAFL as a stable training-time mechanism for adapting supervision to state-dependent reliability.

### A.4 Limitations

A remaining gap arises from teacher-bias variation unresolved by the representation, quantified by $C_b$ in Eq. (8). Consequently, disagreement calibration may systematically assign slightly more trust to pretrained guidance than the representation-conditioned oracle. Developing estimators that explicitly separate this component is an interesting direction for future work.