LATENT STIFFNESS IN MEANFLOW TRAINING

Anonymous authors
Paper under double-blind review

ABSTRACT

MeanFlow trains one-step generative models using targets that depend on the network's own time derivative. We interpret this objective as a semi-gradient bootstrap, analogous to temporal-difference (TD) learning: each update fits a target that subsequent updates can change. We introduce a matrix-free Jacobian probe of *stiffness*, the derivative channel's parameter sensitivity relative to the output channel. In the TD interpretation, stiffness scaled by the time gap measures how strongly an output-visible parameter change can alter the bootstrapped target relative to the prediction being fitted. The probe therefore exposes target sensitivity that a small prediction residual need not constrain. Across CIFAR-10 and ImageNet-256 checkpoints with up to 677M parameters, stiffness grows by one to two orders of magnitude, especially at large gaps, while the weighted loss remains bounded and FID improves. Controlled experiments distinguish gradual stiffening without bootstrapping from additional amplification under bootstrapped large-gap training whose onset tracks learning-rate warmup. Released α-Flow and Decoupled MeanFlow checkpoints achieve lower large-gap self-consistency error and better one-step FID than MeanFlow despite retaining high stiffness. Increasing large-gap supervision improves the best CIFAR-10 finetuning FID from 6.03 to 3.95 but degrades an ImageNet-256 transformer. These results connect sampling and learning-rate choices to hidden changes in bootstrap sensitivity and give a diagnostic for studying the training dynamics and transfer limits of one-step generators.

1 INTRODUCTION

Diffusion and flow models (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2020; Lipman et al., 2022; Liu et al., 2022; Albergo et al., 2025) generate by integrating a velocity field over many steps; MeanFlow (Geng et al., 2026a) is a strong one-step alternative. It trains a network $u_\theta(z, r, t)$ to predict the average velocity of a flow-matching trajectory between two time points $r < t$, so that a sample is produced in a single evaluation, $\hat{x} = z_1 - u_\theta(z_1, 0, 1)$, rather than by many function evaluations. Its success, however, rests on constants that are tuned but not explained: the fraction of instantaneous samples ($r = t$), set between 25% and 75% depending on the dataset; a time-sampling distribution whose logit-normal parameters also change per dataset; and an adaptive loss weighting. In the original ablation, removing average-velocity samples altogether (so that training reduces to flow matching) gives a one-step FID (Heusel et al., 2017) of 329, while within MeanFlow's working range the $r{\neq}t$ fraction, the time sampler and the weighting power move it by up to 6, 5 and 19 points (Geng et al., 2026a). The constants matter most when converting a pretrained flow-matching model into MeanFlow, the natural route to cheap one-step models, which is fragile (Zhang et al., 2025); several remedies have been proposed: objective annealing (α-Flow; Zhang et al., 2026), architectural conversion (Decoupled MeanFlow, DMF; Lee et al., 2026), reformulated targets (iMF; Geng et al., 2026b), and reflowed trajectories (Re-MeanFlow; Zhang et al., 2025).

The cause is visible in the objective itself. Along a flow-matching trajectory (Lipman et al., 2022; Liu et al., 2022) $z_t = (1-t)x + t\epsilon$ with instantaneous velocity $v = \epsilon - x$, the average velocity over $[r, t]$ is $u(z_t, r, t) = \frac{1}{t-r}\int_r^t v(z_\tau, \tau)\, d\tau$. Its true value is unavailable, so MeanFlow samples pairs (r, t) from a schedule and regresses onto a target obtained by differentiating the definition in t:

$$u_\theta(z_t, r, t) \;\leftarrow\; \mathrm{sg}\big[\, v - (t-r)\, \tfrac{d}{dt} u_\theta(z_t, r, t)\big]\,, \tag{1}$$

where sg stops gradients through the target and the total derivative is evaluated by a Jacobian-vector product (Baydin et al., 2018). The target substitutes the current network's derivative into an identity

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

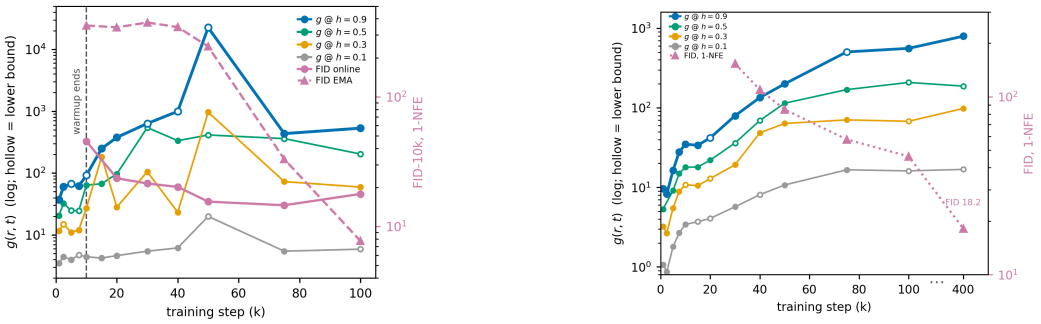

Figure 1: Latent stiffness during from-scratch MeanFlow training on CIFAR-10 (Krizhevsky et al., 2009) (left) and ImageNet-256 with a 12-block DiT-B/4 Transformer, 130M parameters (right): $g(r, t)$ at four gaps (log scale; hollow markers are lower bounds) and one-step FID. On CIFAR-10, g accelerates after the warmup boundary and peaks near 50k while the FID of the same weights keeps improving and the EMA, seeded at initialization, lags through 75k; on the Transformer (no warmup) g rises from the first checkpoint to 792 at 400k while FID falls from 153 to 18.

for the true average velocity, producing a semi-gradient bootstrap. Table 1 summarizes its structural correspondence with TD learning. This connection motivates examining how bootstrapping interacts with uneven supervision, without transferring TD convergence guarantees to MeanFlow.

Our probe compares the parameter sensitivity of du_θ/dt with that of u_θ at frozen checkpoints, producing a time-resolved diagnostic $g(r, t)$ (§3). In from-scratch training, large-gap estimates grow by nearly two orders of magnitude on an ImageNet-256 (Deng et al., 2009) Transformer and by more than one on CIFAR-10, while small-gap values stay small, the logged loss stays flat, and FID improves (Fig. 1). The large gaps include the one-step query $(r, t) = (0, 1)$ at which the model is used.

Controlled CIFAR-10 runs show stiffening even in a control that never trains an $r \neq t$ pair, and further amplification under bootstrapped large-gap training whose onset tracks the learning-rate warmup boundary (§4). Finetuning a flow-matching model with more large-gap and no instantaneous samples improves one-step FID at every matched step (best 3.95 against 6.03; §5.1).

On ImageNet-256 (Deng et al., 2009) with a 677M transformer (Peebles & Xie, 2023; Ma et al., 2024), the same intervention degrades and destabilizes finetuning, with large-gap targets dominated by the derivative bootstrap term (§5.2). This is consistent with the finding of Kim et al. (2026) that large-gap training can degrade even pretrained instantaneous velocities. Released α-Flow and Decoupled MeanFlow (DMF) checkpoints achieve lower large-gap self-consistency error and better one-step FID than MeanFlow, showing that their gains come with more consistent large-gap predictions despite persistently high stiffness.

Our contributions are summarized as follows:

- **A TD interpretation and a scalable stiffness probe.** We interpret MeanFlow as a semi-gradient bootstrap and introduce $g(r, t)$ to measure derivative sensitivity relative to output sensitivity. The gap-scaled quantity $(t-r)g$ bounds local bootstrap amplification over output-visible perturbations. A matrix-free estimator enables measurements on frozen checkpoints.
- **Latent stiffness and its training dynamics.** Across CIFAR-10 and ImageNet-256, measured stiffness grows by one to two orders of magnitude, especially at large gaps, despite bounded logged loss and improving FID. Controlled CIFAR-10 experiments distinguish drift without bootstrapping from additional amplification under bootstrapped large-gap training whose onset tracks learning-rate warmup, and locate the dominant measured response in the spatial derivative.
- **Large-gap supervision: benefits and transfer limits.** A sampling intervention improves the best CIFAR-10 finetuning FID from 6.03 to 3.95 but degrades finetuning of a 677M-parameter ImageNet-256 transformer. Released α-Flow and DMF checkpoints achieve better one-step FID and lower large-gap self-consistency error than the MeanFlow baseline while retaining high stiff-

ness, identifying improved large-gap prediction consistency as a shared feature of successful recipes.

2 Related Work

2.1 One-step generative models and training stability

One-step and few-step models reduce the sampling cost of diffusion and flow matching (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2020; Lipman et al., 2022; Liu et al., 2022; Albergo et al., 2025). Examples include consistency models (Song et al., 2023; Song & Dhariwal, 2024; Geng et al., 2025; Lu & Song, 2025), shortcut and flow-map models (Frans et al., 2025; Kim et al., 2024; Boffi et al., 2024; Sabour et al., 2026; Zhou et al., 2025), and distillation from pretrained teachers (Salimans & Ho, 2022; Yin et al., 2024; Zhou et al., 2024). MeanFlow (Geng et al., 2026a) predicts the average velocity between two time points, allowing a single network evaluation to map noise to data over the full interval. Many of these methods use model-generated targets, coupling target construction with optimization. Their recipes employ robust loss weighting (Song & Dhariwal, 2024; Geng et al., 2025; Karras et al., 2024), tangent normalization and warmup (Lu & Song, 2025; Sabour et al., 2026), and nonuniform time sampling (Esser et al., 2024); guidance can also modify the training target (Ho & Salimans, 2022; Tang et al., 2025; Kynkäänniemi et al., 2024).

Prior analyses identify derivative terms as sources of training difficulty. Boffi et al. (2026) connect instability in Eulerian self-distillation to the spatial Jacobian, while Lu & Song (2025) and Sabour et al. (2026) discuss difficulties involving time derivatives and time embeddings. We complement these accounts by measuring the parameter sensitivity of MeanFlow's derivative-based bootstrap throughout training. Like studies of the edge of stability (Cohen et al., 2021) and Hessian spectra (Ghorbani et al., 2019), our work examines optimization through numerical measurements, but targets bootstrap sensitivity relative to prediction sensitivity rather than loss curvature.

2.2 Understanding and improving MeanFlow

Recent work studies MeanFlow's optimization and conversion from pretrained flow models. α-Flow (Zhang et al., 2026) attributes slow convergence to conflicting gradients between flow-matching and consistency objectives and introduces a curriculum that shifts their relative influence; its EMA-target ablation is consistent with our observations. iMF (Geng et al., 2026b) reformulates the target to reduce noise introduced by the Jacobian-vector-product tangent. While iMF modifies the derivative signal used in training, we measure its sensitivity to parameter changes. Kim et al. (2026) show that accurate instantaneous velocities support large-gap learning, whereas emphasizing large gaps prematurely can degrade the instantaneous field. Our experiments complement this account by examining sensitivity during training and the consequences of withdrawing large-gap coverage (§5.2). Decoupled MeanFlow (Lee et al., 2026) and Re-MeanFlow (Zhang et al., 2025) convert pretrained flow models through an encoder–decoder separation with flow-matching warmup and reflowed trajectories, respectively. Further work develops stabilization methods and curricula for large-scale training (He et al., 2026; Hsu & Wood, 2026). We complement these methods with measurements across training and released checkpoints, relating derivative sensitivity to optimization and large-gap prediction consistency.

2.3 Temporal-difference learning and bootstrapped targets

Our analysis imports the stability theory of Temporal-difference (TD) learning (Sutton, 1988; Sutton et al., 1998): convergence of semi-gradient TD under on-policy sampling (Tsitsiklis & Van Roy, 1996), divergence when the update distribution departs from the bootstrap (Baird, 1995), and the deadly triad in deep RL (Van Hasselt et al., 2018). Three of its remedies have direct counterparts in our study, which we label as inspirations: emphatic TD's interest function (Sutton et al., 2016) for coverage, target networks (Mnih et al., 2015) for the EMA teacher we tested, and the conservatism of offline RL (Fujimoto et al., 2019; Kumar et al., 2020; Kostrikov et al., 2021) for the conservative-bootstrap probe of §5.2. The MeanFlow residual is a continuous-time TD error (Doya, 2000); NTK analyses characterize when deep TD diverges (Achiam et al., 2019), and TD-Flow bootstraps flow matching (Farebrother et al., 2025).

Table 1: The term-by-term correspondence between MeanFlow training and semi-gradient TD learning. Each MeanFlow quantity plays the same role in the fixed-point iteration as its TD counterpart; the correspondence organizes the analysis, it does not import TD's convergence guarantees.

MeanFlow	TD learning	Role in the iteration
instantaneous velocity v	reward $r(s)$	data-grounded target
$(t-r)\,\frac{d}{dt}u_\theta$	bootstrapped value $\gamma V_\theta(s')$	network-supplied target
gap $t-r$	bootstrap weight (the role of γ)	bootstrap weight
$\mathrm{sg}[\cdot]$	semi-gradient	no gradient through target
$r=t$ samples	reward-only targets (as at terminal states)	bootstrap-free targets
time sampler over (r,t)	behavior distribution d_μ	which states are updated

3 MeanFlow as a Bootstrapped Iteration

3.1 The objective as a fixed-point iteration

Write the MeanFlow target of Eq. 1 as an operator acting on the current network,

$$T(u)(z_t,r,t) \;=\; \underbrace{v(z_t,t)}_{\text{anchor}} \;-\; \underbrace{(t-r)\,\tfrac{d}{dt}u(z_t,r,t)}_{\text{bootstrap}}, \qquad \mathcal{L}(\theta)=\mathbb{E}\left\|u_\theta-\mathrm{sg}[T(u_\theta)]\right\|^2. \tag{2}$$

The data provide the anchor $v=\epsilon-x$; the network supplies the derivative through a JVP along $(v,0,1)$. Stop-gradient makes the update a semi-gradient: optimization differentiates the output but not the target. The true average velocity satisfies $u^\star=T(u^\star)$ (in conditional expectation over data–noise pairs, since the anchor and the JVP tangent use the conditional velocity), but this identity does not establish convergence from an arbitrary initialization.

Table 1 summarizes the TD correspondence. Precisely, with displacement $D=(t-r)\,u$ the residual $u-T(u)$ equals $\frac{d}{dt}D-v$, an undiscounted continuous-time TD error (Doya, 2000) with reward rate v. In the u-parameterization the bootstrap's share of the target grows with the gap, which is the sense in which the gap plays the role of γ; it is largest at the one-step query $(0,1)$, while the time sampler determines which pairs receive updates.

3.2 What to measure, and why the loss cannot show it

The network exposes two channels from one parameter vector: the output channel u_θ, which the loss supervises directly, and the derivative channel $\frac{d}{dt}u_\theta$, which the loss never supervises directly but which the bootstrap consumes as a target. Let

$$J_A=\frac{\partial u_\theta}{\partial\theta}, \qquad J_B=\frac{\partial}{\partial\theta}\Big(\frac{d}{dt}u_\theta\Big), \tag{3}$$

both evaluated at a fixed (r,t) on a fixed batch of z_t. A perturbation $\delta\theta$ moves the output by $J_A\delta\theta$ and the bootstrap term of the target by $-(t-r)J_B\delta\theta$; at the largest gaps, where the one-step query lies, the factor $t-r$ is near one and g is the target-to-output sensitivity itself. The iteration is well-conditioned when the derivative channel is no more sensitive than the output channel, and ill-conditioned when there exist perturbations that leave the output nearly unchanged while moving the target substantially: such perturbations are invisible to the loss and yet rewrite what the network is trained toward.

The natural quantity is therefore the worst-case ratio of derivative response to output response. The naive form $\max_{\delta\theta}\|J_B\delta\theta\|/\|J_A\delta\theta\|$ is ill-posed, because J_A has a null space (parameter directions that do not move the output at all) and the ratio is infinite there. The well-posed version restricts attention to output-visible perturbations:

$$g(r,t) \;=\; \left\|J_B\,J_A^{+}\right\|, \tag{4}$$

the operator norm of the derivative response induced by a unit output response, with J_A^+ the pseudoinverse. A value near one means the two channels move together; a value of 10^4 means the bootstrap consumes a channel four orders of magnitude more sensitive than the one the loss controls.

The loss in Eq. 2 measures prediction–target mismatch at sampled time pairs. Because the target is stop-gradient, the training gradient depends on the output Jacobian and the current residual, without directly differentiating the target. Small loss and gradient norms can therefore coexist with large g. We call this *latent stiffness*: high relative derivative sensitivity that standard training diagnostics may not reveal. We report g pointwise, as a scalar at fixed (r, t) and batch, on a grid of gaps that reaches the one-step query $(0, 1)$.

A local stability interpretation of g. Each update changes both the prediction and the target it is trying to fit. The gap-scaled stiffness $(t-r)g$ bounds the target's response relative to the prediction's over output-visible perturbations. To make this precise, freeze J_A and J_B at one checkpoint, time pair, and batch, and assume J_A has full row rank. For the semi-gradient flow $\dot{\theta} = -J_A^\top e$, where $e = u_\theta - T(u_\theta)$, the residual evolves as

$$\dot{e} = -\big(I + (t-r)J_B J_A^+\big) J_A J_A^\top e.$$

If $(t-r)g < 1$, the weighted residual norm $\|(J_A J_A^\top)^{1/2} e\|$ decreases: target feedback is too weak to overturn the update's corrective effect. This sufficient condition gives g a local stability interpretation, not a convergence guarantee for full training. Violating it does not imply instability, and a lower-bound estimate below the threshold cannot certify it. We therefore examine separately how measured stiffness relates to training outcomes (§5.4).

3.3 Estimator and stability

Neither Jacobian is ever formed: for a deep neural network they are prohibitively large. What automatic differentiation provides is the action of each Jacobian on a vector, $J\,d$ by a forward-mode Jacobian-vector product and $J^\top w$ by a reverse-mode vector-Jacobian product (Baydin et al., 2018). The estimator is built from these two operations alone (Algorithm 1).

Estimator. Eq. 4 is the largest generalized singular value of (J_B, J_A) over output-visible perturbations. We estimate it by a power iteration carried in output space: for a unit output-space direction w, the least-norm pre-image $x = (J_A^\top J_A + \mu I)^{-1} J_A^\top w$ lies in the range of $J_A^\top$ by construction, and the iteration applies J_B, $J_B^\top$, the same ridge solve, and J_A in turn, reporting $\|J_B x\|/\|J_A x\|$ at convergence (Algorithm 1). The ridge solve uses conjugate gradients (Hestenes et al., 1952); μ is fixed per protocol relative to the mean eigenvalue of $J_A^\top J_A$ ($\mu_{\mathrm{rel}} = 10^{-2}$ on CIFAR-10, 10^{-3} on ImageNet-256), a choice validated by a scan from 10^{-2} to 10^{-5} on a calibration checkpoint.

Numerical reliability. We check each estimate using independent restarts and monitor whether it stabilizes within the computation budget. Reported values are directional lower bounds on g: repeatability does not guarantee that the most sensitive direction has been found. We mark estimates that continue to increase and omit points whose restarts disagree. On CIFAR-10 and the B/4 Transformer, incomplete inner solves make the estimates sensitive to the solver budget, so comparisons between checkpoints require care. Appendix A reports the grading criteria, budget checks, and variation across probe batches. On a synthetic problem with known g, the estimator agrees within 2%, whereas random-direction sampling underestimates it by more than an order of magnitude. The probe uses fp32 and takes minutes per time pair on a frozen checkpoint.

4 Latent Stiffness, Measured at Two Scales

Setup. We apply the probe of §3.3 to checkpoints from MeanFlow training on two testbeds, measured on the same checkpoint schedule. On CIFAR-10 (Krizhevsky et al., 2009) we train the recipe's convolutional backbone with its default settings (effective batch 1024; 100k steps instead of the original 800k; final one-step FID 5.84). On ImageNet-256 (Deng et al., 2009) we train MeanFlow-B/4 from scratch with the default recipe of a public PyTorch reproduction of MeanFlow (effective batch

Algorithm 1 Estimate of $g(r,t) = \|J_B J_A^+\|$ at a fixed (r,t) and batch (output-space iteration)

Require: frozen θ; batch $\{z_t\}$; ridge μ; restarts R; iterations K
Ensure: $g(r,t)$ with its grade
for $\rho = 1, \dots, R$ **do**
  draw a random unit direction w_0 in output space
  for $k = 1, \dots, K$ **do**
    solve $(J_A^\top J_A + \mu I)\, x = J_A^\top w_{k-1}$ by conjugate gradients
    $y \leftarrow J_B x$; $\hat{g}_k \leftarrow \|y\| / \|J_A x\|$
    solve $(J_A^\top J_A + \mu I)\, x' = J_B^\top y$ by conjugate gradients
    $w_k \leftarrow J_A x' / \|J_A x'\|$
  end for
  $\hat{g}^{(\rho)} \leftarrow \hat{g}_K$; $p^{(\rho)} \leftarrow$ [$\hat{g}_k$ plateaued]
end for
grade $\leftarrow$ *converged* if all $p^{(\rho)}$ and the $\hat{g}^{(\rho)}$ agree within tolerance
    lower bound if the $\hat{g}^{(\rho)}$ agree but did not plateau
    uncertifiable if the $\hat{g}^{(\rho)}$ disagree
report $\max_\rho \hat{g}^{(\rho)}$ with its grade

256, 400k steps, final one-step FID 18.2) and additionally probe 677M-parameter XL/2 Transformers (Peebles & Xie, 2023; Ma et al., 2024) in the SD-VAE latent space (Rombach et al., 2022): the finetuning checkpoints of §5.2 and the MeanFlow-XL/2 released with α-Flow (Zhang et al., 2026) (Appendix B). Every g carries its numerical reliability (§3.3) and protocol; absolute values are not comparable across backbones. Configurations, protocols, and full measurement tables are in Appendix B.

Finding 1: latent stiffness grows as generation improves. Large-gap probe estimates increase while one-step FID improves on both testbeds (Fig. 1). On CIFAR-10, sensitivity accelerates after learning-rate warmup and peaks near 50k steps, while the FID of the same online weights continues to improve through 75k. ImageNet-256 B/4 checkpoints exhibit a similar increase in sensitivity alongside improving FID under a schedule without warmup.

The flat logged loss conceals both the sensitivity growth and changes in large-gap residuals. On CIFAR-10, the sampled, weighted loss remains near 5 while the unweighted residual at $h = 0.9$ approximately triples (Table 7). Adaptive weighting compresses these changes: for $R \gg c$, $R/(R+c)^p \approx R^{1-p}$, giving fourth-root growth at $p = 0.75$ and saturation near one at the ImageNet setting $p = 1$. Moreover, residual magnitude does not directly constrain the target's parameter sensitivity. We therefore report large-gap sensitivity and unweighted residuals separately: the probe exposes changes in the derivative channel, while the residual measures prediction–target mismatch that the logged loss can obscure.

Finding 2: large gaps concentrate derivative sensitivity. Large-gap probe estimates exceed small-gap estimates across the CIFAR-10 U-Net, ImageNet-256 B/4, and released XL/2 models (Tables 4 to 6). This pattern does not simply follow local sampling density: on CIFAR-10, the large-gap probe points receive more nearby samples than the small-gap points (Appendix F). Two other factors provide a plausible explanation. The marginal sampling mass decreases toward large gaps, leaving those gap values less frequently supervised, while the multiplier $(t - r)$ increases the derivative's contribution to the target. Separating limited supervision from derivative-based targets is therefore necessary. Moreover, at every reported point with $h \geq 0.5$, from the first checkpoint onward, the lower-bound estimate already exceeds the small-gain threshold of §3.2, by up to four orders of magnitude. Large-gap training therefore operates outside this sufficient condition, so feedback directions and optimization dynamics must be examined directly.

Finding 3: drift without bootstrapping, amplification with it. The CIFAR-10 control trained exclusively on instantaneous samples also develops increasing large-gap sensitivity (Fig. 2, left). Stiffening therefore does not require derivative-based training targets: shared-parameter updates can increase sensitivity at time pairs that receive no direct supervision. The canonical MeanFlow run exhibits substantially stronger growth, with median estimates approximately 19–28$\times$ those of the

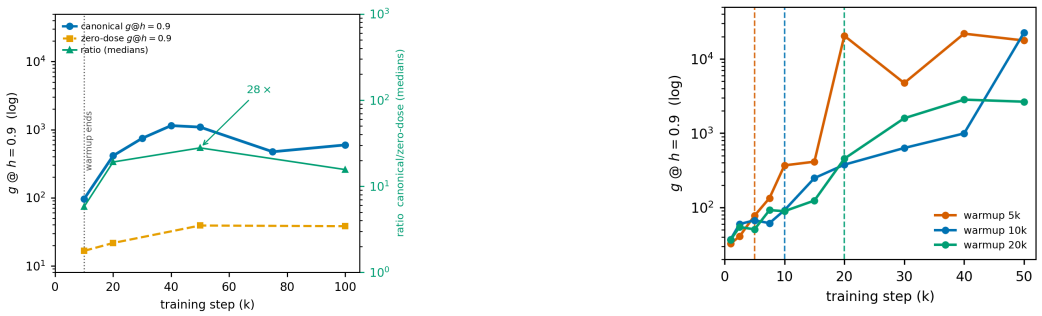

Figure 2: Two components (CIFAR-10). Left: g at h=0.9 for the canonical run and a zero-dose control (medians of three probe batches) and the ratio of the medians (right axis). Right: runs differing only in warmup length (5k, 10k, 20k; dashed: end of warmup), one probe batch per point.

control at 20k and 50k steps. This comparison measures the combined effect of introducing large-gap supervision and bootstrapping. Changing the warmup length shifts the timing of the largest excursions (Fig. 2, right), linking the amplification to the learning-rate schedule.

The dominant measured response is spatial. Along the probe's estimated maximizing direction, the spatial term $\nabla_z u_\theta \, v$ contributes 83–96% of the total derivative response, compared with 4–18% for the explicit time-derivative term (Table 8). This connects the measured stiffness to the spatial-Jacobian difficulties reported for Eulerian self-distillation (Boffi et al., 2026).

Finding 4: a stability interpretation of MeanFlow's design choices. The preceding results suggest complementary roles for sampling and weighting. The instantaneous-sample fraction balances data-supervised velocity fitting at $r = t$ against derivative-based targets at nonzero gaps; the time sampler determines coverage across the time plane; and adaptive loss weighting (Song & Dhariwal, 2024; Geng et al., 2025) limits the influence of large residuals. The roughly $20\times$ larger per-sample gradient norms on large-gap draws during finetuning, even after weighting (§5.1), illustrate the update imbalance these choices must manage. These roles are interpretations of the diagnostics; we did not isolate them by individual from-scratch ablations.

5 Finetuning Flow Matching into MeanFlow

5.1 Finetuning on CIFAR-10

We finetune the 100k-step flow-matching checkpoint of §4. Abruptly switching to MeanFlow produces a gradient shock: large-gap samples have approximately $20\times$ larger median gradient norms than small-gap samples, even after adaptive weighting, and clipping remains active for a sustained period. Training nevertheless remains bounded and generation improves, while large-gap sensitivity stays high (Table 2). Ramping the sample fraction or learning rate reduces the initial shock but yields worse FID at 30k steps and leaves sensitivity at a comparable level (Appendix C). Thus, smoothing the objective switch alone does not improve generation within this budget.

We next change the allocation of supervision, increasing large-gap sampling three- or fivefold while removing instantaneous samples (Appendix B). Both variants improve early finetuning FID and produce lower probe estimates than the abrupt switch (Table 2). The fivefold variant outperforms both the abrupt switch and our α-Flow-style curriculum at every matched evaluation step (Fig. 3; Table 9), matching the abrupt baseline's 100k-step FID by 60k steps. Its best one-step FID is 3.95 at 150k steps, compared with 6.03 at 110k for the abrupt baseline (Appendix D).

These results support reallocating supervision toward large gaps during finetuning from a pretrained flow-matching field. The intervention jointly changes gap coverage and the instantaneous sample fraction, so their individual contributions remain unresolved. Its success also does not establish that instantaneous samples can be removed when training from scratch.

Table 2: Finetuning a flow-matching model into MeanFlow on CIFAR-10: one-step FID at matched steps (Table 9), each run's best one-step FID, and g at h=0.9 after 30k and 120k steps. The dosed arm both raises large-gap sampling and removes instantaneous samples. Protocols in Appendix B.

Method	FID@60k	FID@120k	best FID (step)	g @30k	g @120k
Default finetune (abrupt)	10.50	8.06	6.03 (110k)	1064	707
α-Flow-style curriculum	41.23	9.06	7.15 (120k)	71.6	1152
Ours (5× dose, no r=t)	**7.79**	**5.83**	**3.95** (150k)	567	760

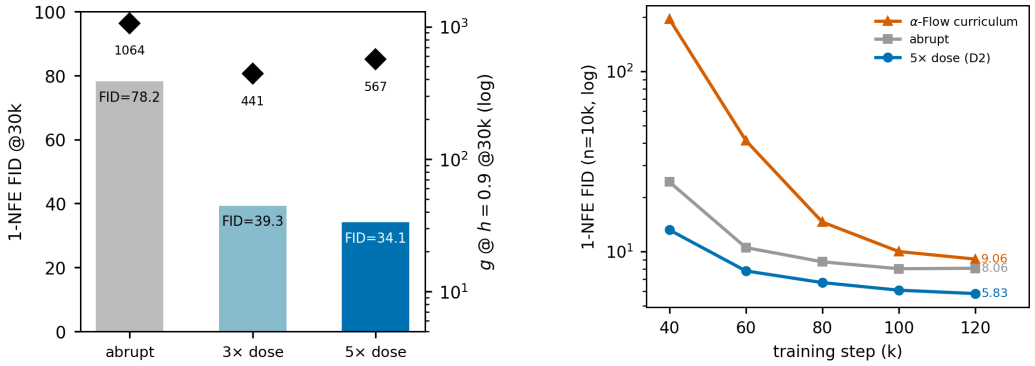

Figure 3: Large-gap supervision on the CIFAR-10 finetune. Left: one-step FID and g at h=0.9 at 30k steps for the abrupt switch and three- and fivefold doses. Right: one-step FID at matched steps for the fivefold dose, the abrupt switch, and the α-Flow-style curriculum baseline. The numbers are in Tables 2 and 9.

5.2 Transfer to ImageNet-256

The CIFAR-10 benefit does not transfer to the reproduction's 677M-parameter flow-matching XL/2 checkpoint. With zero-initialized second-time conditioning and matched evaluation settings, increasing large-gap sampling degrades finetuning, and none of the tested interventions recovers the improvement (Appendix E, Table 10).

Adjusting the timing of supervision does not resolve the failure. Delayed dosing eventually destabilizes an initially stable continuation, with damage concentrated in the one-step map. Ordering the same samples from small to large gaps also fails: loss on unvisited large-gap regions grows sharply while loss on visited regions changes little. These results expose two difficulties: introducing large-gap updates can destabilize adaptation, while postponing them leaves large-gap predictions without direct supervision.

5.3 Coverage and Target Feedback

The ImageNet diagnostics reveal a substantial contribution from model-generated targets. On large-gap draws, the median ratio of the bootstrap term's norm to the anchor's norm is 2.6, and gradient energy is distributed across parameter directions and network depth (Appendix E).

From the TD perspective, additional samples improve coverage but do not provide fixed labels: shared-parameter updates change both predictions and the derivative-based targets used by subsequent updates. Additional large-gap supervision may therefore reinforce errors in the evolving target field as well as constrain under-supervised predictions. Coverage differs substantially between the settings: the reproduction's sampler assigns 0.046% of its mass near the one-step query, compared with 1.8% on CIFAR-10. We hypothesize that the benefit of increasing this coverage depends on the reliability of the derivative-based targets during adaptation. Target magnitude alone does not establish target error, however, and differences in architecture, representation, and guidance prevent attributing the reversal to a single factor (Appendix G).

Table 3: Released ImageNet-256 XL/2 checkpoints; MeanFlow-XL/2 is the release of Zhang et al. (2026). Paper FID: the one-step FID reported for the probed checkpoint (Zhang et al., 2026; Lee et al., 2026). †: ours, under one protocol (Appendix B): one-/two-step FID, g at h=0.5 / 0.9 [1.0], self-consistency error $E(0.5)/E(1.0)$ (one jump from noise against two half jumps; Appendix B), and the training sampler's mass near the one-step query (t>0.9, h>0.7). The REPA base (Yu et al., 2024), DMF's starting point, is the flow-matching SiT-XL/2+REPA checkpoint evaluated as a flow map without finetuning.

Model	Paper FID	FID 1/2-NFE†	g@0.5 / 0.9 [1.0]†	$E(0.5)/E(1.0)$†	mass†
MeanFlow-XL/2	3.47	6.51 / 5.42	1825 / 6593 [3148]	0.145 / 0.323	0.094%
α-Flow-XL/2	2.95	5.92 / 5.45	2091 / uncert. [20797]	0.148 / 0.278	0.094%
DMF-XL/2	2.83	5.38 / 4.90	3289 / 6537 [7122]	0.113 / 0.266	0.416%
REPA base	n/a	341.4 / 271.6	177 / 458 [991]	0.306 / 0.515	0%

5.4 Comparison with α-Flow and Decoupled MeanFlow

Under a common evaluation protocol, released α-Flow and Decoupled MeanFlow checkpoints achieve lower large-gap self-consistency error and better one-step FID than MeanFlow (Fig. 4, Table 3). Across these three generators, the ordering of $E(1.0)$ matches the FID ordering, whereas the ordering of g changes with the probe gap. Decoupled MeanFlow also substantially improves generation and self-consistency relative to its flow-matching base while its measured derivative sensitivity increases. The observed gains therefore involve more consistent large-gap predictions without a corresponding reduction in stiffness.

Decoupled MeanFlow's sampler assigns approximately nine times as much mass near the one-step query as the reproduction's sampler. Its successful endpoint shows that increased coverage is compatible with good generation, despite the failures of our ImageNet intervention. These checkpoint comparisons do not isolate which training choices enable that outcome. What they identify is the behavior to be explained: successful recipes learn more consistent large-gap predictions while retaining a sensitive derivative field.

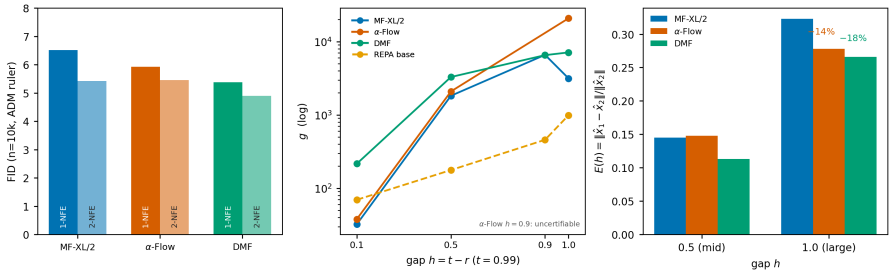

Figure 4: MeanFlow, α-Flow, and Decoupled MeanFlow on ImageNet-256, probed on released checkpoints under one protocol: one- and two-step FID (left), g against the gap with grade markers (middle), and flow-map self-consistency error (right); numbers in Table 3.

6 Conclusion

Our TD interpretation and Jacobian probe reveal latent stiffness in MeanFlow's derivative-based targets that the weighted training loss does not capture. Controlled experiments distinguish drift without bootstrapping from additional amplification under bootstrapped large-gap training. Successful α-Flow and DMF checkpoints achieve more consistent large-gap predictions while retaining high stiffness, whereas increasing large-gap supervision improves CIFAR-10 finetuning but degrades the tested ImageNet-256 configuration. These findings point to the interaction between supervision and the evolving target field during finetuning as the question to study next. Monitoring derivative sensitivity alongside unweighted large-gap residuals provides complementary information about target feedback and prediction error. Limitations are discussed in Appendix G.

AI USE STATEMENT

In this work, we used generative AI tools (large language models and an LLM-based coding assistant) to write and debug part of the code. We did not use generative AI tools to generate synthetic datasets, formulate or prove mathematical claims, clean or reformat datasets, assist with translation, design experiments, formulate hypotheses, or interpret experimental results; qualitative and thematic data analysis is not applicable to this work. Additionally, we used generative AI tools to edit the paper for readability and to identify and summarize relevant literature.

We have reviewed all AI-assisted work: all AI-assisted code was read, tested, and validated by the authors; all experimental designs and hypotheses were determined by the authors and evaluated by the authors against the recorded data; and every number, claim, and citation in the paper was checked by the authors against the experimental record and the cited sources. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work does not raise any ethical concerns. It does not involve human or animal subjects, personally identifiable information, or sensitive personal data. The datasets and experimental resources used in this work are publicly available and used in accordance with their respective terms of use.

REPRODUCIBILITY STATEMENT

We provide detailed descriptions of the proposed method, model architecture, training procedure, and experimental settings in the main paper and supplementary material. The datasets used in our experiments are publicly available, and their relevant preprocessing and evaluation protocols are described in the paper and supplementary material. We also provide implementation details, hyperparameters, and evaluation settings necessary to reproduce the reported results. A GitHub repository containing the source code and necessary scripts and configurations for reproducing our experiments will be made publicly available during the camera-ready stage.

REFERENCES

Joshua Achiam, Ethan Knight, and Pieter Abbeel. Towards characterizing divergence in deep q-learning. *arXiv preprint arXiv:1903.08894*, 2019.

Michael Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.

Leemon Baird. Residual algorithms: Reinforcement learning with function approximation. In *Machine learning proceedings 1995*, pp. 30–37. Elsevier, 1995.

Atilim Gunes Baydin, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18(153):1–43, 2018.

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.

Nicholas M Boffi, Michael S Albergo, and Eric Vanden-Eijnden. Flow map matching with stochastic interpolants: A mathematical framework for consistency models. *arXiv preprint arXiv:2406.07507*, 2024.

Jeremy M Cohen, Simran Kaur, Yuanzhi Li, J Zico Kolter, and Ameet Talwalkar. Gradient descent on neural networks typically occurs at the edge of stability. *arXiv preprint arXiv:2103.00065*, 2021.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. Ieee, 2009.

Kenji Doya. Reinforcement learning in continuous time and space. *Neural computation*, 12(1): 219–245, 2000.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first international conference on machine learning*, 2024.

Jesse Farebrother, Matteo Pirotta, Andrea Tirinzoni, Rémi Munos, Alessandro Lazaric, and Ahmed Touati. Temporal difference flows. *arXiv preprint arXiv:2503.09817*, 2025.

Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models. In *International Conference on Learning Representations*, volume 2025, pp. 34668–34684, 2025.

Scott Fujimoto, David Meger, and Doina Precup. Off-policy deep reinforcement learning without exploration. In *International conference on machine learning*, pp. 2052–2062. PMLR, 2019.

Zhengyang Geng, Ashwini Pokle, Weijian Luo, Justin Lin, and Zico Kolter. Consistency models made easy. In *International Conference on Learning Representations*, volume 2025, pp. 75158–75185, 2025.

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.

Behrooz Ghorbani, Shankar Krishnan, and Ying Xiao. An investigation into neural net optimization via hessian eigenvalue density. In *International Conference on Machine Learning*, pp. 2232–2241. PMLR, 2019.

Xiao He, Yang Li, Peizhen Zhang, Songtao Liu, Zhao Zhong, and Nannan Wang. Stabilizing, scaling & enhancing meanflow for large-scale diffusion distillation. *arXiv preprint arXiv:2605.17834*, 2026.

Magnus R Hestenes, Eduard Stiefel, et al. Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49(6):409–436, 1952.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Chia-Hong Hsu and Frank Wood. Discrete meanflow training curriculum. *arXiv preprint arXiv:2604.08837*, 2026.

Tero Karras, Miika Aittala, Jaakko Lehtinen, Janne Hellsten, Timo Aila, and Samuli Laine. Analyzing and improving the training dynamics of diffusion models. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24174–24184. IEEE, 2024.

Dongjun Kim, Chieh-Hsin Lai, WeiHsiang Liao, Naoki Murata, Yuhta Takida, Toshimitsu Uesaka, Yutong He, Yuki Mitsufuji, and Stefano Ermon. Consistency trajectory models: Learning probability flow ode trajectory of diffusion. In *International Conference on Learning Representations*, volume 2024, pp. 44493–44525, 2024.

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.

Ilya Kostrikov, Ashvin Nair, and Sergey Levine. Offline reinforcement learning with implicit q-learning. *arXiv preprint arXiv:2110.06169*, 2021.

Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

Aviral Kumar, Aurick Zhou, George Tucker, and Sergey Levine. Conservative q-learning for offline reinforcement learning. *Advances in neural information processing systems*, 33:1179–1191, 2020.

Tuomas Kynkäänniemi, Miika Aittala, Tero Karras, Samuli Laine, Timo Aila, and Jaakko Lehtinen. Applying guidance in a limited interval improves sample and distribution quality in diffusion models. *Advances in Neural Information Processing Systems*, 37:122458–122483, 2024.

Kyungmin Lee, Sihyun Yu, and Jinwoo Shin. Decoupled meanflow: Turning flow models into flow maps for accelerated sampling. In *International Conference on Learning Representations*, volume 2026, pp. 64555–64583, 2026.

Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

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.

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.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4172–4182. IEEE, 2023.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 10674–10685. ieee, 2022.

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. *arXiv preprint arXiv:2202.00512*, 2022.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. pmlr, 2015.

Yang Song and Prafulla Dhariwal. Improved techniques for training consistency models. In *International Conference on Learning Representations*, volume 2024, pp. 15078–15097, 2024.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. *arXiv preprint arXiv:2303.01469*, 2023.

Richard S Sutton. Learning to predict by the methods of temporal differences. *Machine learning*, 3 (1):9–44, 1988.

Richard S Sutton, Andrew G Barto, and Andrew Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

Richard S Sutton, A Rupam Mahmood, and Martha White. An emphatic approach to the problem of off-policy temporal-difference learning. *Journal of Machine Learning Research*, 17(73):1–29, 2016.

Zhicong Tang, Jianmin Bao, Dong Chen, and Baining Guo. Diffusion models without classifier-free guidance. *arXiv preprint arXiv:2502.12154*, 2025.

John Tsitsiklis and Benjamin Van Roy. Analysis of temporal-diffference learning with function approximation. *Advances in neural information processing systems*, 9, 1996.

Hado Van Hasselt, Yotam Doron, Florian Strub, Matteo Hessel, Nicolas Sonnerat, and Joseph Modayil. Deep reinforcement learning and the deadly triad. *arXiv preprint arXiv:1812.02648*, 2018.

Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Fredo Durand, William T Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6613–6623. IEEE, 2024.

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. *arXiv preprint arXiv:2410.06940*, 2024.

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*, volume 2026, pp. 144056–144094, 2026.

Xinxi Zhang, Shiwei Tan, Quang Nguyen, Quan Dao, Ligong Han, Xiaoxiao He, Tunyu Zhang, Alen Mrdovic, and Dimitris Metaxas. Flow straighter and faster: Efficient one-step generative modeling via meanflow on rectified trajectories. *arXiv e-prints*, pp. arXiv–2511, 2025.

Linqi Zhou, Stefano Ermon, and Jiaming Song. Inductive moment matching. *arXiv preprint arXiv:2503.07565*, 2025.

Mingyuan Zhou, Huangjie Zheng, Zhendong Wang, Mingzhang Yin, and Hai Huang. Score identity distillation: Exponentially fast distillation of pretrained diffusion models for one-step generation. In *Forty-first International Conference on Machine Learning*, 2024.

A THE PROBE: ESTIMATOR AND VALIDATION

The object. The stiffness is $g(r,t) = \|J_B J_A^+\|$, the operator norm of J_B restricted to the output-visible perturbations $\mathrm{range}(J_A^\top)$, with $J_A = \partial u/\partial\theta$ and $J_B = \partial(\mathrm{d}u/\mathrm{d}t)/\partial\theta$. The unrestricted ratio $\sup_d \|J_B d\|/\|J_A d\|$ is unbounded, because J_A has a nontrivial null space of parameter directions that leave the output invariant while moving the derivative channel. The pseudoinverse restricts the ratio to directions the output can see; the null space is itself informative, consisting of output-invisible, derivative-active directions.

Recurrence. Algorithm 1 iterates in output space. For a unit direction w, the least-norm pre-image $x = (J_A^\top J_A + \mu I)^{-1} J_A^\top w$ lies in $\mathrm{range}(J_A^\top)$ by construction; each step computes $y = J_B x$, $x' = (J_A^\top J_A + \mu I)^{-1} J_B^\top y$ and $w \leftarrow J_A x'/\|J_A x'\|$, and reports $\|J_B x\|/\|J_A x\|$ at convergence. Both solves are conjugate gradients on $J_A^\top J_A + \mu I$. Iterating in parameter space instead is not equivalent: the iterates are not confined to $\mathrm{range}(J_A^\top)$ and drift into $\mathrm{null}(J_A)$, where only the ridge bounds the denominator. On MeanFlow-XL/2, where the inner solves converge, the parameter-space form overshoots Algorithm 1 by 15% at h=0.9; at small gaps, where J_A is well conditioned, the two

agree within 8% (at the canonical 20k checkpoint, 4.79 and 30.3 against 4.62 and 28.1 at h=0.1 and 0.3). The $h \leq 0.3$ columns of Table 4 use the parameter-space form; every other value in the paper uses Algorithm 1.

Ridge. $\mu = \mu_{\mathrm{rel}} \cdot \|J_A d_{\mathrm{probe}}\|^2$, with the scale a random-probe estimate of the mean eigenvalue of $J_A^\top J_A$; $\mu_{\mathrm{rel}} = 10^{-2}$ on CIFAR-10 and 10^{-3} on ImageNet-256. On a calibration checkpoint, a scan over $\mu_{\mathrm{rel}} \in \{10^{-2}, \ldots, 10^{-5}\}$ moved the estimate by less than 2% at the 40-iteration conjugate-gradient budget.

Inner-solver budget and probe batch. On the CIFAR-10 U-Net and the B/4 Transformer the ridge solves do not converge at 40 conjugate-gradient iterations (final relative residual 9–15 and 0.9–3.9), and the estimate grows with the budget: at the canonical CIFAR-10 20k checkpoint and h=0.9, 160 iterations give 994 against 411 at 40. Values on these two backbones are therefore loose lower bounds. Finetuned CIFAR-10 checkpoints converge better (residual 0.56 for the abrupt arm at 30k), and on the released XL/2 models the solves converge (residual ≤ 0.23) with repeated estimates within 2%. Values also depend on the probe batch: at h=0.9, three batches give 200, 414 and 857 at the canonical CIFAR-10 20k checkpoint and 461, 1090 and 2.4×10^4 at 50k, and span 153–532 and 218–828 at 50k and 100k of the B/4 run.

Restarts, convergence, and grades. Each of R restarts starts from an independent unit Gaussian direction in output space. Restarts agree when the largest estimates lie within 10% of their median; the iteration has plateaued when the estimate varies by less than 2% over its last 30%. All reported points use 20 outer iterations, 40 conjugate-gradient iterations and $R = 2$, with a probe batch of 64 on CIFAR-10 and 24 on ImageNet-256. A point is *converged* if the iteration plateaued and the restarts agree, a *restart-consistent lower bound* if the restarts agree but the iteration was still climbing, and *uncertifiable* if the restarts disagree; uncertifiable points are omitted from figures. Every value is a lower bound on g: the grade certifies the outer iteration at the stated inner-solver budget, not the value of g. The separability projection of Appendix B uses a heavier budget (25 outer, 60 CG, $R = 3$) at three fixed large-gap probe points, $(r, t) \in \{(0.05, 0.95), (0.25, 0.90), (0.45, 0.95)\}$, and 256 per-sample gradients drawn over $h \in [0.5, 0.95]$; the energy fraction is $(g_i \cdot d)^2 / \|g_i\|^2$ for the unit top direction d.

Validation. On random linear pairs (A, B) the estimate agrees with the exact generalized singular value $\sqrt{\lambda_{\max}(B^\top B,\, A^\top A)}$ to within 2%, and on a rank-deficient pair with $\|J_B J_A^+\| = 10^6$ it recovers 10^6 at every μ_{rel}. On the network backbones the Jacobian-vector products agree with finite differences to a relative error of 1.6×10^{-6}. A maximum of $\|J_B d\| / \|J_A d\|$ over random directions underestimates the converged value by more than an order of magnitude, which is why the power iteration, not sampling, is the estimator.

B Experimental setup and protocols

CIFAR-10 canonical run. The MeanFlow recipe (Geng et al., 2026a) on its convolutional backbone: effective batch 1024, Adam with learning rate 6×10^{-4} and 10,000 warmup steps, gradient clipping at norm 1.0, logit-normal$(-2, 2)$ time sampler with a 25% r=t fraction, adaptive weighting ($p = 0.75$, $c = 10^{-3}$), EMA decay 0.99995 initialized from the step-0 weights, 100k steps. One-step FID 5.84; the reference implementation reports 6.35 after 800k steps.

CIFAR-10 finetuning arms. All arms start from the 100k-step flow-matching checkpoint of §4 with the optimizer of Appendix D. The abrupt switch uses the canonical sampler. The threefold dose (D1) draws both times from logit-normal$(0, 3.5)$, taking t as the larger and r as the smaller draw. The fivefold dose (D2) keeps the canonical logit-normal$(-2, 2)$ sampler and, for 60% of samples with $t > 0.5$, redraws the gap as $h \sim \mathcal{U}(0.5, t)$; samples with $t \leq 0.5$ keep their canonical (r, t). Neither arm uses r=t samples. The large-gap mass $P(h \geq 0.5)$ is 0.42 for D1 and 0.67 for D2, against 0.13 for the canonical sampler, hence "threefold" and "fivefold".

ImageNet-256 B/4 from-scratch recipe (public reproduction). SiT-B/4 (12 blocks, width 768, patch 4), class-conditional with guidance dropout 0.1, in the SD-VAE latent space ($4 \times 32 \times 32$);

Table 4: CIFAR-10 canonical run: $g(r,t)$ by gap and step (training-series protocol). Grade tags: c converged, ℓ restart-consistent lower bound. One-step FID of the EMA weights: 373.48@30k, 343.74@40k, 240.76@50k, 31.58@75k; of the probed online weights: 45.4@10k, 23.5@20k, 21.4@30k, 20.2@40k, 15.5@50k, 14.6@75k, 17.8@100k. The $h \leq 0.3$ columns use the parameter-space form of the probe (Appendix A).

step (k)	$g@h{=}0.9$	$g@h{=}0.5$	$g@h{=}0.3$	$g@h{=}0.1$
1	37.2^{c}	20.6^{c}	11.6^{c}	3.51^{c}
2.5	59.8^{c}	32.4^{c}	14.9^{ℓ}	4.42^{c}
5	67.2^{ℓ}	24.7^{ℓ}	11.0^{c}	3.97^{c}
7.5	61.6^{c}	24.6^{ℓ}	12.0^{c}	4.73^{ℓ}
10	91.9^{ℓ}	63.9^{c}	27.0^{c}	4.42^{c}
15	249^{c}	66.9^{c}	183^{c}	4.22^{c}
20	377^{c}	96.9^{c}	28.1^{c}	4.62^{c}
30	632^{ℓ}	544^{ℓ}	105^{c}	5.45^{c}
40	993^{ℓ}	333^{c}	23.2^{c}	6.15^{c}
50	$2.25{\times}10^{4\,\ell}$	412^{ℓ}	963^{c}	19.9^{ℓ}
75	434^{c}	360^{ℓ}	72.7^{c}	5.45^{c}
100	531^{c}	203^{ℓ}	59.3^{c}	5.87^{ℓ}

Adam with learning rate 10^{-4}, betas $(0.9, 0.95)$, no weight decay, gradient clipping at norm 1.0, constant learning rate with no warmup, batch 256; linear interpolant; adaptive weighting ($p = 1.0$); logit-normal$(-0.4, 1.0)$ time sampler with a 25% $r{\neq}t$ fraction; guidance folded into the target ($\omega = 3.0$, $\kappa = 0$); EMA decay 0.9999; 100k steps. The 400k checkpoint is the reproduction's released model of the same recipe.

Two XL/2 checkpoints. Two independently trained MeanFlow-XL/2 models appear in the paper, both 677M-parameter DiT-class Transformers (Peebles & Xie, 2023; Ma et al., 2024) in the SD-VAE latent space (Rombach et al., 2022). The *reproduction* (`https://github.com/zhuyu-cs/MeanFlow`) matches the original paper's one-step FIDs (B/4 15.43 against 15.53; XL/2 3.39 against 3.43) and trains with 25% $r{\neq}t$ samples; the B/4 recipe above, the flow-matching XL/2 that the finetunes of §5.2 convert, and the starting checkpoint of the post-hoc dosing arms (Appendix E) all come from it. The *α-Flow release* is the MeanFlow-XL/2 that Zhang et al. (2026) trained with 50% $r{\neq}t$ samples and released with α-Flow (reported one-step FID 3.47); the released-model panel (Tables 3 and 6, Fig. 4) probes it.

Protocols. *Training series (CIFAR-10).* g is measured along $t \in [0.9, 1]$ with $r = t - h$. *Released models and B/4 (ImageNet-256).* g is measured on the line $t = 0.99$, $r = \max(t - h, 0)$, which ends next to the one-step query. *Self-consistency error.* Starting from pure noise z_1, $\hat{x}_1 = z_1 - h\,u_\theta(z_1, 1{-}h, 1)$ is the point reached at $t = 1 - h$ in one jump and $\hat{x}_2$ the point reached through $t = 1 - h/2$ in two half-gap jumps; $E(h) = \|\hat{x}_1 - \hat{x}_2\|/\|\hat{x}_2\|$, averaged over 2000 draws. *Weights and tangent.* On CIFAR-10 the probe uses the online weights, which produce the training targets, while FID uses the EMA weights, so early EMA FIDs lag the probed weights (Table 4 lists both); on ImageNet-256 both use the released EMA weights. The probe's tangent is the conditional velocity $\epsilon - x$, as in training. *FID.* One- and two-step FID against the standard reference statistics (the CIFAR-10 training set; ADM for ImageNet-256), with 50k samples for CIFAR-10 endpoints and the ImageNet-256 finetuning arms, and 10k samples for the CIFAR-10 training series, the matched-step CIFAR-10 comparison and the released-model panel. Two steps are two deterministic half-gap jumps, $z_{1/2} = z_1 - \frac{1}{2}u_\theta(z_1, \frac{1}{2}, 1)$ and $\hat{x} = z_{1/2} - \frac{1}{2}u_\theta(z_{1/2}, 0, \frac{1}{2})$, without sampling-time guidance.

Diagnostics on the canonical CIFAR-10 run. *Residuals.* Table 7 reports the unweighted residual $R = \|u_\theta - \mathrm{sg}(v - (t-r)\,\mathrm{d}u_\theta/\mathrm{d}t)\|^2$ and its adaptively weighted value $R/(R + 10^{-3})^{0.75}$, which is the training loss, as medians over 2048 samples on the probe line and 8192 draws of the training sampler. *Tangent split.* At the probe's top direction $x^\star$ the tangent $(v, 0, 1)$ of the total derivative is split into a spatial part $(v, 0, 0)$ and a time part $(0, 0, 1)$; Table 8 gives each part's response relative to the full one (the two are nearly parallel, cosine 0.86–1.00). *Realized gain.* Let d be the negative gradient of the full training objective (all sampler terms, adaptive weighting included, before Adam's

Table 5: ImageNet-256 from-scratch B/4: $g(r,t)$ by gap and step (released-model protocol), through our 100k run and the released 400k checkpoint; at the one-step query (h=1.0) g is 201 at 50k and 708 at 100k. Grade tags: c converged, ℓ restart-consistent lower bound, u uncertifiable. One-step FID: 153.43@30k, 109.68@40k, 84.80@50k, 57.60@75k, 46.22@100k, 18.17@400k.

step (k)	g@h=0.9	g@h=0.5	g@h=0.3	g@h=0.1
1	9.61^{c}	5.34^{c}	3.21^{c}	1.07^{c}
2.5	8.24^{c}	u	2.66^{c}	0.87^{c}
5	16.3^{c}	9.11^{c}	5.50^{c}	1.79^{c}
7.5	27.9^{c}	15.0^{c}	8.85^{c}	2.69^{c}
10	34.8^{c}	17.9^{c}	10.8^{ℓ}	3.42^{c}
15	33.8^{c}	18.0^{c}	10.5^{c}	3.70^{ℓ}
20	41.9^{ℓ}	22.0^{c}	12.9^{ℓ}	4.07^{ℓ}
30	79.3^{c}	36.1^{ℓ}	19.2^{c}	5.70^{c}
40	134^{c}	69.2^{ℓ}	48.4^{c}	8.08^{ℓ}
50	200^{c}	114^{c}	63.5^{c}	10.7^{c}
75	503^{ℓ}	169^{c}	70.4^{c}	16.6^{c}
100	558^{c}	208^{ℓ}	67.6^{ℓ}	16.1^{ℓ}
400	792^{c}	187^{ℓ}	98.0^{c}	16.8^{ℓ}

Table 6: ImageNet-256 released XL/2 models (MeanFlow-XL/2 is the α-Flow release; Appendix B): g by gap (released-model protocol, $t = 0.99$; h=1.0 is the point $(r,t) = (0, 0.99)$). Grade tags: c converged, ℓ restart-consistent lower bound, u uncertifiable. The REPA base is the DMF starting checkpoint.

model	g@h=0.1	g@h=0.5	g@h=0.9	g@h=1.0
MeanFlow-XL/2	32.3^{ℓ}	1825^{c}	6593^{c}	3148^{c}
α-Flow-XL/2	37.9^{ℓ}	2091^{c}	u	20797^{c}
DMF-XL/2	216^{c}	3289^{c}	6537^{c}	7122^{c}
REPA base	69.6^{ℓ}	177^{c}	458^{c}	991^{c}

preconditioning) at the checkpoint's online weights on an independent batch of 1024 drawn from the training sampler, and let J_A, J_B be evaluated on the probe batch at h=0.9. Along this fixed direction the ratio $\rho = \|J_B d\|/\|J_A d\|$ needs no range restriction; it is 1.6, 21.5, 12.8, 154, 19.3 and 45.0 at 10k, 20k, 30k, 50k, 75k and 100k, against $g = 96$, 414, 746, 2.4×10^4, 473 and 596 on the same probe batch. *Separability.* At the 50k checkpoint the probe's top output-space direction at the large-gap probe points (itself converged) holds a median $\sim$0.001% of the per-sample gradient energy over 256 large-gap samples: training gradients carry almost none of their energy along the stiffest direction.

Table 7: Canonical CIFAR-10 run: unweighted residual R at probe gaps and at the one-step query, and the adaptively weighted training loss (medians).

step	R, h=0.1	R, h=0.5	R, h=0.9	R, query	loss, h=0.9	loss, sampler
1k	913	4268	13437	23996	10.8	5.13
20k	964	8602	29819	58399	13.1	5.03
50k	991	8864	32244	53084	13.4	4.99
100k	953	9965	39611	61710	14.1	4.98

C CURE RAMPS ON THE CIFAR-10 FINETUNE

Three ramp arms smooth the entry from flow matching into MeanFlow. **R** (ratio ramp): the $r \neq t$ fraction ramps from 0.05 to 0.75 linearly over the first 10k steps, with the abrupt learning-rate schedule. **L** (learning-rate ramp): full $r \neq t$ fraction from step 0, with warmup extended to 10k steps, the ramp suggested by Finding 3. **GL** (gap-dependent learning rate): full fraction and standard

Table 8: Canonical CIFAR-10 run at h=0.9: share of the derivative's response along the probe's top direction carried by the spatial and time parts of the tangent.

step	g	spatial / full	time / full
20k	413	0.83	0.18
50k	2.4×10^4	0.96	0.04
100k	599	0.94	0.07

Table 9: Step-matched one-step FID for the three CIFAR-10 finetunes.

step (k)	40	60	80	100	120
$5\times$ dose	13.16	7.79	6.71	6.09	**5.83**
abrupt	24.30	10.50	8.75	8.01	8.06
α-Flow-style curriculum	194.88	41.23	14.55	9.98	9.06

learning rate, with per-sample gradients scaled down at large gaps early and the scaling relaxed over 10k steps.

None of the three arms diverges, and none removes the latent stiffness. One-step FID at 30k is 88.9 (GL), 92.0 (R) and 111.2 (L), against 78.2 for the abrupt switch: the ramps that act on exposure to large gaps beat the ramp that acts on step size alone. Nor does any ramp lower the end-state stiffness: g at h=0.9 after 30k steps is 1011 (R), 1095 (L) and 1480 (GL) against 1064 for the abrupt switch. The ramps change how the stiff regime is entered, not whether it is reached.

D Dose versus curriculum versus abrupt finetuning (CIFAR-10)

The α-Flow-style curriculum baseline. Starting from the same 100k-step flow-matching checkpoint as the dose arms, the curriculum interpolates the target between a transport-flow-matching form and the MeanFlow bootstrap with weight α, annealed by a sigmoid: with $p = \min(1, \text{step}/T)$ and $T = 60{,}000$,

$$\alpha(\text{step}) \;=\; \frac{0.5}{1+\exp\big(-25\,(0.5-p)\big)} \;\in [0,\,0.5],$$

so α starts near 0.5 and anneals to ≈ 0 by T; for $\alpha < 5\times10^{-3}$ the loss is identical to the pure MeanFlow objective. In the interpolated regime the intermediate time is $s = \alpha r + (1-\alpha)t$, the target is $\alpha v_t + (1-\alpha)u_{\theta^-}(z_s, r, s)$ with θ^- the current stop-gradient weights (no EMA), and the weight is $\alpha/(\|\Delta\|^2 + 10^{-3})$. The optimizer is that of the dose and abrupt arms (learning rate 6×10^{-4}, effective batch 1024, no warmup). The run lasts 120k steps, so its tail (60–120k) is pure MeanFlow. Because the finetune starts from a flow-matching checkpoint, the schedule starts at $\alpha \approx 0.5$ instead of annealing from $\alpha = 1$; it is an α-Flow-style baseline, not a reproduction of the α-Flow recipe (Zhang et al., 2026).

Matched-step comparison. Table 9 evaluates the three arms at the same training steps under one protocol; the fivefold dose leads the abrupt switch, which leads the curriculum, at every step. At convergence the fivefold dose reaches 3.95 (150k), the abrupt switch 6.03 (110k), and the curriculum 7.15 (120k); the fivefold dose reaches the abrupt arm's best matched-step FID (8.01 at 100k) by 60k.

E The ImageNet-256 intervention map

None of the interventions below recovers the CIFAR-10 benefit on the 677M Transformer. For each we state the design and the result; FID is one-step unless stated.

E.1 Large-gap dose sweep

Design: increase the large-gap sampling mass of the finetune continued from the converged flow-matching field, to test whether the CIFAR-10 benefit transfers. *Result:* the zero-dose baseline

Table 10: Interventions on the ImageNet-256 finetune and their outcomes; none recovers the CIFAR-10 benefit. Designs and measurements are in the subsections below.

Intervention	Outcome
Large-gap dose (0.1–0.6)	monotonic degradation; g uncertifiable at large gaps
Gap-dependent gradient attenuation ($10^6\times$)	saturates; not carried by gradient magnitude
EMA target for the derivative	below baseline; matches α-Flow's ablation
Dose on the converged model	neutral under matched guidance configuration
Dose after partial training	instability at 87k; damage confined to the one-step map
Same samples, small gaps first	instability before any large gap is used
Off-support bootstrap penalty	instability at about 1k steps

reaches FID 68.2 at 30k; dose 0.2 reaches 398.5 with saturated samples (sample standard deviation 0.9–1.0 against 0.5 for the baseline), and dose 0.6 degrades further to $\sim$440 with the reference loss rising to 10^{10}. FID degrades monotonically with dose.

E.2 Gradient-weighted damping

Design: per-sample weighting that damps large-gap residuals, at doses 0.3 and 0.2. *Result:* g is uncertifiable at h=0.5 and 0.9 in both arms; sample standard deviation $\sim$90; FID 316–400, below the baseline. The weighting contains the loss (10^3–10^8 against 10^{14} without it) but does not restore a measurable channel.

E.3 EMA teacher

Design: bootstrap targets from an exponential moving average of the network (decay 0.95 or 0.99; dose 0.3). *Result:* FID 332–339 at 30k with sample standard deviation 84–91 at every checkpoint; g uncertifiable; the loss stays near 10^4 with spikes to 10^6.

E.4 Dose on the converged model

Design: finetune the reproduction's MeanFlow-XL/2 (Appendix B) under its own guidance configuration, with dose 0.3 and with dose 0. *Result:* the starting checkpoint scores 3.358 at one step and 4.187 at two; the two arms score 4.084 and 4.100 at one step. Dosing is neutral, and both arms sit about 0.7 above the starting checkpoint.

E.5 Dose after partial training

Design: dose 0.3 introduced into a continuation that had trained stably. *Result:* stable to about 85k steps, then unstable at 87k, with the damage confined to the one-step map: the gradients recover by 100k, but the self-consistency error $E(1.0)$, flat at 0.31–0.34 through 80k, rises to 0.41 by 90k and stays there. Dosing itself worsened large-gap self-consistency: $E(1.0)$ went from 0.28 at the branch point to 0.32 within 10k steps, while the undosed run improved to 0.26 at the same step.

E.6 Small gaps first

Design: the same samples consumed in order of increasing gap. *Result:* unstable at 9–12k steps, before any large gap is used. Splitting the loss by region localizes the failure (Fig. 5): the term on visited gaps grows by $1.3\times$ ($3.6\times10^7 \rightarrow 4.8\times10^7$) while the term on the unvisited large gaps grows by $693\times$ ($4.06\times10^{10} \rightarrow 2.81\times10^{13}$).

E.7 Off-support bootstrap penalty

Design: a penalty on the bootstrap target where the sampler does not go, $\lambda = 3.75\times10^{-12}$ calibrated to the loss scale. *Result:* unstable at about step 1000; the penalty moderates the loss (1.2×10^{10} against 9.9×10^{15} at 1k without it) but does not prevent the instability.

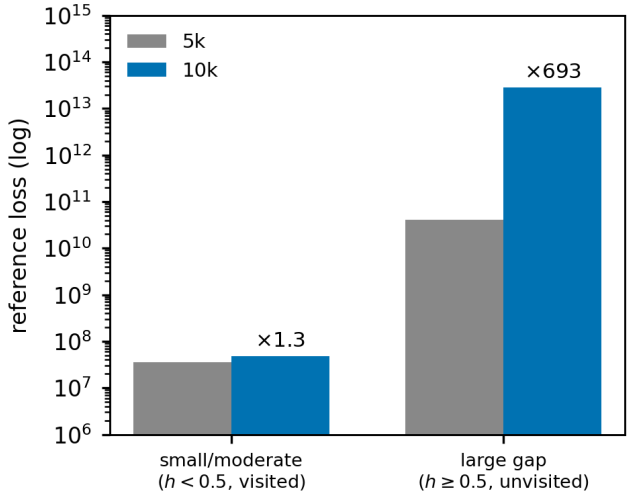

Figure 5: Loss by region under the small-gaps-first schedule, at 5k and 10k steps: the term on visited gaps grows by $1.3\times$; the term on the unvisited large gaps grows by $693\times$.

E.8 Diagnostics on the dosed arms

On the finetune targets: (a) the bootstrap term dominates large-gap targets, with median $\|(t-r)\,\mathrm{d}u/\mathrm{d}t\|/\|v\| = 2.62$ against ≈ 0 for the $r{=}t$ anchors; (b) large-gap gradients are diffuse, with a top-principal-component share of 0.28 against 0.76 for small-gap gradients; (c) large-gap gradient energy sits in the middle third of the network (first/middle/last thirds 0.266/0.585/0.149). There is no low-rank or depth-localized component that could be removed.

E.9 Two configuration requirements

A randomly initialized second-time embedder on a flow-matching checkpoint diverges at every dose; a zero-initialized one finetunes stably. Finetuning a guidance-distilled checkpoint without guidance degrades the one-step map most: one-step FID 8.2 against 4.3 at two steps, whereas the starting checkpoint, the reproduction's XL/2, scores 3.4 and 4.2. All finetunes in the paper use a zero-initialized embedder and the checkpoint's own guidance configuration.

E.10 Sampler coverage near the one-step query

Coverage is the fraction of the training sampler's (r, t) proposals with $t > 0.9$ and $h > 0.7$, near the one-step query $(0, 1)$: 0.046% for the ImageNet-256 reproduction (25% $r{\neq}t$), 0.094% for the MeanFlow and α-Flow releases (50% $r{\neq}t$), 0.416% for DMF, and 1.8% for CIFAR-10, raised to 20.6% and 8.3% by the threefold and fivefold doses (Monte Carlo over each recipe's sampler). Coverage near the query is higher for the one recipe, DMF, that succeeds at large-gap dosing; α-Flow succeeds without it, through the annealed target instead. From scratch, low coverage is not required: MeanFlow's uniform-sampler ablation, with $\sim 27\times$ more mass near the query than the reproduction's sampler, trains stably (Geng et al., 2026a).

F Sampler density at the probe points

Table 11 reports, for the training samplers of Appendix B, the share of all samples near each probe gap. Along the CIFAR-10 probe line the stiffest point ($h{=}0.9$) is also the most densely sampled, both locally and along the fixed-r chain that feeds it, whereas the marginal frequency of the gap value falls with h. On the ImageNet-256 probe line ($t = 0.99$) local density peaks near $h \approx 0.5$ and falls toward $h = 0.9$ while g rises monotonically.

Table 11: Share of all training samples within ± 0.05 of each probe gap (2×10^7 Monte Carlo draws of the samplers of Appendix B). *Local*: near the probe point on the protocol's probe line (for ImageNet-256, $|t-0.99| < 0.05$). *Chain*: along the fixed-r chain feeding the CIFAR-10 probe point at $t = 0.95$. *Marginal*: gap value near h, any t. ImageNet-256 rows use the reproduction's sampler (25% $r \neq t$); the MeanFlow and α-Flow releases train with 50%, which doubles each ImageNet-256 entry. D1 and D2 are the threefold and fivefold finetuning doses.

	h=0.1	h=0.3	h=0.5	h=0.9
CIFAR-10, local (%)	0.074	0.102	0.150	1.006
CIFAR-10, chain (%)	0.10	0.50	1.45	52.7
CIFAR-10, marginal (%)	17.5	7.9	4.9	1.4
CIFAR-10 D1, local (%)	4.92	2.70	2.44	8.44
CIFAR-10 D1, marginal (%)	14.3	8.5	7.2	9.7
CIFAR-10 D2, local (%)	0.040	0.055	1.56	3.11
CIFAR-10 D2, marginal (%)	9.3	4.2	18.0	3.7
ImageNet-256, local, t=0.99 (%)	0.0012	0.0044	0.0064	0.0017
ImageNet-256, marginal (%)	6.3	4.1	1.8	0.005

G LIMITATIONS AND EXTENDED DISCUSSION

No supervision-dose recipe of ours improves the Transformer finetune; the ImageNet-256 contribution is the map of where finetuning fails and the measurement of recipes that succeed. The controlled from-scratch ablations (the bootstrap-free control and the warmup arms) exist at CIFAR-10 scale only, and the from-scratch B/4 run reproduced the monotone stiffening but not the plateau-and-takeoff shape, which its warmup-free recipe cannot produce.

Every run uses a single seed. Every g is a lower bound, a loose one on the CIFAR-10 U-Net and the B/4 Transformer (Appendix A), and values vary across probe batches by more than an order of magnitude at the CIFAR-10 peak, so differences in g between runs are indicative rather than exact. g is a worst-case, same-point proxy for the operator-level contraction structure; it does not measure the directions the optimizer takes or the coupling between sampled and unsampled regions. The zero-dose control never trains large gaps, so it does not isolate the bootstrap from large-gap training; the clean test keeps coverage fixed and replaces the bootstrapped target by a teacher-integrated one.

The CIFAR-10 and ImageNet-256 finetunes differ in dataset, pixel versus latent space, backbone, guidance in the target and embedder initialization, so the reversal cannot be attributed to one factor, and we have not shown that g forecasts failure on held-out runs. The association between stiffness and one-step ability rests on the comparisons of §5.4 and is correlational; the ranking by $E(1.0)$ rests on three released models; the realized gain is measured along the raw gradient rather than Adam's update. The immature-derivative account of the ImageNet-256 failure is a hypothesis: we measured the bootstrap term's share of the target, not its error. ImageNet-256 measurements use re-trained releases, not the original weights (the reproduction's checkpoints, and the α-Flow release of MeanFlow-XL/2 for the released-model panel; Appendix B); the B/4 trajectory trained 400k checkpoint of the same recipe; and one method family is studied in depth.

The TD toolbox, read onto MeanFlow. MeanFlow's bootstrapped target is a semi-gradient temporal-difference fixed point (Sutton, 1988), and several field heuristics are its stabilizers under different names. Coverage, which gap values the sampler trains, plays the role of an emphatic-TD interest function over a designer-chosen visitation distribution (Sutton et al., 2016), with the g-map as a measured interest signal. The α-Flow curriculum, annealing from a transport target toward the bootstrap, is a TD(λ)-style interpolation between Monte-Carlo and bootstrapped returns (Zhang et al., 2026). The EMA teacher is a target network (Mnih et al., 2015). The conservative-bootstrap penalty is offline-RL conservatism on off-support targets (Kumar et al., 2020). DMF's success (Lee et al., 2026) rests on a mature anchor field (REPA pretraining (Yu et al., 2024)), a shared time embedding and a robust loss, not on lowering stiffness.

The largest gap is where gains and damages localize. Every effect concentrates at the large gap: stiffness is highest at the largest gaps; the finetuning instability is a large-gap per-sample gradient spike; dosing acts on exactly the large gaps and both helps (CIFAR-10) and, mis-timed, destabilizes

(ImageNet-256); DMF's value-level gain is a large-gap self-consistency improvement on a calm base. The same coordinate carries the benefit and the risk.

Transition versus destination. Coverage acts on the bootstrap-free drift and can move the *destination*: the fivefold dose is the only intervention that improved the CIFAR-10 end state, and it lowered g during the transition (567 against 1064 at 30k) while ending at the abrupt arm's level (760 against 707 at 120k). Ramps act on the *transition*: they change how the stiff regime is entered and which mechanism drives the gradient spike at entry, but not the end state (no ramp lowered the 30k stiffness below the abrupt arm's; Appendix C). Post-hoc combination for already-stiff checkpoints is left as future work.

Future work: a predictive test. The probe is used here diagnostically. The natural next test is predictive: on held-out training runs, whether the g trajectory forecasts finetune instability, and the coverage number which recipes succeed, before the loss or samples move.