TRAJECTORY-REGULARIZED DISTRIBUTION DISTILLATION FOR ONE-STEP GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Distilling flow matching models for one-step generation generally follows two paradigms. Trajectory alignment supervises the student with the teacher's probability flow ODE, but its performance degrades under very small sampling budgets. Distribution matching directly optimizes the generated distribution, but estimating its reverse Kullback-Leibler gradient typically requires a separate network to track the evolving student score. We propose trajectory-regularized distribution distillation (TRD), which combines distribution-level and trajectory-level supervision through a momentum-mixture policy predicted in a single student forward pass. This policy serves simultaneously as a one-step generator, an estimator of the student's marginal velocity that removes the need for a second trainable network in distribution matching, and a closed-form time-dependent velocity field supervised along the teacher ODE trajectory. Experiments on SANA-1.6B and Stable Diffusion 3.5 show that TRD outperforms representative consistency and distribution matching baselines in generation quality and training efficiency for both one-step and few-step generation, with a trade-off in sample diversity.

1 INTRODUCTION

Diffusion and flow matching models support high-resolution image synthesis (Rombach et al., 2022; Peebles & Xie, 2023; Esser et al., 2024; Xie et al., 2025). Deterministic samplers for score-based diffusion and flow matching models numerically integrate a generative ODE, often requiring tens of network evaluations (Song et al., 2021; Karras et al., 2022; Lipman et al., 2023). Distillation seeks to reduce this cost to one or two evaluations (Salimans & Ho, 2022; Song et al., 2023; Yin et al., 2024b); we organize prior approaches around two primary supervision objects.

Trajectory-oriented methods reduce long generative paths through recursive teacher-step compression (Salimans & Ho, 2022), consistency along solution curves (Song et al., 2023; Luo et al., 2023a; Kim et al., 2024; Lu & Song, 2025), path straightening (Liu et al., 2023; 2024), or interval-level policies (Chen et al., 2026; Yang et al., 2026). Related direct-training objectives learn long transitions through improved consistency training, shortcuts or average velocities without requiring a fixed teacher (Song & Dhariwal, 2024; Frans et al., 2025; Geng et al., 2025). For linear-shortcut distillation, Yang et al. (2026) attribute quality loss at two to four evaluations to approximating a teacher path whose tangent direction changes over time.

Distribution matching instead aligns the student and teacher output laws through score differences (Luo et al., 2023b; Yin et al., 2024b;a) or a Fisher-divergence objective (Zhou et al., 2024). It avoids prescribing a one-to-one teacher sampling path. The cited score-difference and Fisher methods typically train a second network to track the evolving student score; inaccurate tracking can destabilize training (Yin et al., 2024a), while the tracker adds memory and update overhead that becomes substantial at larger scales (Cheng et al., 2026; Kim et al., 2026).

TRD uses one policy-valued student output as a shared representation for generation, distribution matching, and trajectory alignment. It reuses the momentum-mixture parameterization of ArcFlow for closed-form curved generation (Yang et al., 2026), and additionally uses the same output for student marginal velocity estimation and trajectory-grid prediction. Self flow matching, distribution matching, trajectory alignment, and distribution preservation jointly train this field. The controlled estimator ablation in Table 7 shows that sharing improves overall generation quality and teacher-

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.

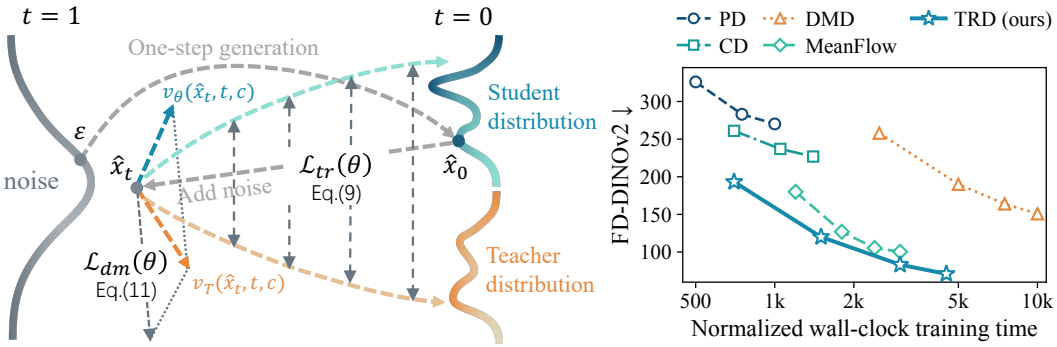


Figure 1: Overview of TRD. Left: one student forward pass at a noised student state returns a momentum-mixture policy that supplies both the distribution matching direction of equation 11 and the trajectory-grid velocities of equation 9. Right: one-step FD-DINOv2 against normalized wall-clock training time when distilling SANA-1.6B. Appendix B describes both panels.

distribution agreement relative to a DMD-style separate fake network, while avoiding its auxiliary optimizer state and gradient memory.

Our contributions are summarized as follows:

- We use the momentum-mixture policy as a shared student representation for closed-form one-step generation, student marginal velocity estimation, and trajectory-grid velocity prediction.
- We formulate TRD to combine distribution-level and trajectory-level supervision through a shared self-estimator, removing the need for a separately trained fake model and its auxiliary optimizer state and gradient memory. Distribution preservation further stabilizes updates to the evolving student distribution.
- We establish ideal-state correctness for the combined objective, show that trajectory supervision of the shared start-time output rescales the reverse Kullback-Leibler update while preserving its direction, and provide finite-error bounds with complete proofs in section 4.5 and Appendix C.
- We evaluate TRD on SANA-1.6B (Xie et al., 2025) and Stable Diffusion 3.5 Medium, whose architecture follows the multimodal rectified-flow transformer family of Esser et al. (2024), through matched-NFE comparisons and controlled ablations that isolate the shared estimator, the objective terms and the policy capacity.

2 Related Work

Trajectory-oriented methods. Trajectory-oriented methods learn long transitions tied to sampling paths. Endpoint, progressive and bootstrapped distillation enlarge deterministic teacher transitions (Luhman & Luhman, 2021; Salimans & Ho, 2022; Gu et al., 2024), while rectified flow and reflow straighten their couplings (Liu et al., 2023; 2024). Consistency and flow-map objectives enforce agreement across ODE points or intervals (Song et al., 2023; Song & Dhariwal, 2024; Luo et al., 2023a; Lu & Song, 2025; Kim et al., 2024; Boffi et al., 2025; Sabour et al., 2025). Moment matching distils conditional moments (Salimans et al., 2024); MeanFlow and shortcut models directly learn average velocities or long transitions (Geng et al., 2025; Frans et al., 2025); and policy-based methods distil interval dynamics (Chen et al., 2026; Yang et al., 2026). TRD adds distribution matching to pointwise teacher-trajectory supervision through one policy.

Distribution distillation with a learned student-score model. Distribution methods compare noised generated and target laws. VSD uses target and learned variational scores (Wang et al., 2023; Nguyen & Tran, 2024), and Diff-Instruct integrates divergences over time (Luo et al., 2023b). DMD combines a score-difference update with teacher-output regression (Yin et al., 2024b); DMD2 removes the regression loss, stabilizes fake-score training with a two-time-scale update rule and adds

adversarial training (Yin et al., 2024a). SiD and Guided SiD use Fisher-type objectives (Zhou et al., 2024; 2025a), whereas SIM generalizes to a broader family of score-based divergences (Luo et al., 2024). Each of these methods trains an auxiliary model to follow the evolving student score. TRD estimates the corresponding marginal velocity with its generation policy.

Hybrid and fake-score-free alternatives. Hybrid methods combine teacher-following with distributional or adversarial alignment. ADD, LADD, Diffusion2GAN, Flash Diffusion and SANA-Sprint mix score, paired-output or consistency supervision with discriminators or distribution losses (Sauer et al., 2024b;a; Kang et al., 2024; Chadebec et al., 2025; Chen et al., 2025). SiDA adds adversarial supervision to score identity matching (Zhou et al., 2025b), ADM uses adversarial distribution matching (Lu et al., 2025), and TDM matches trajectories as distributions (Luo et al., 2025). Other methods avoid a separately trained fake-score network through energy distance, contrastive energy estimation, endpoint pseudo-velocities, self-adversarial twin trajectories or network self-evaluation (Shao et al., 2025; Zhu et al., 2025; Kim et al., 2026; Cheng et al., 2026; Yu et al., 2026). TRD instead couples pointwise teacher velocities with reverse Kullback-Leibler matching through one policy, without a separate score model or discriminator.

3 PRELIMINARIES

Flow matching. We use the convention that noise sits at $t = 1$ and data at $t = 0$. For a data sample $\boldsymbol{x}_0$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$, the linear path is

$$\boldsymbol{x}_t = (1-t)\boldsymbol{x}_0 + t\boldsymbol{\epsilon}, \qquad t \in [0,1]. \tag{1}$$

Its conditional velocity is $\boldsymbol{\epsilon} - \boldsymbol{x}_0$. A teacher trained by conditional flow matching (Lipman et al., 2023; Liu et al., 2023) predicts the marginal velocity $\boldsymbol{v}_T(\boldsymbol{x}, t, c)$ under a condition c, and sampling integrates $\dot{\boldsymbol{x}} = \boldsymbol{v}_T(\boldsymbol{x}, t, c)$ from $t = 1$ to $t = 0$. Conditioning on c is implicit in distribution notation. The core objective below uses the conditional teacher field directly; Appendix D gives the classifier-free-guidance extension.

Distribution matching distillation. Let G_θ be a student generator with trainable parameters θ. Given a condition c and initial noise $\boldsymbol{x}_1 \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$, it produces the endpoint $\hat{\boldsymbol{x}}_0 = G_\theta(\boldsymbol{x}_1, c)$. Draw $t \sim \rho$ on $[t_{\min}, t_{\max}] \subset (0,1)$ and independent noise $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ to form

$$\hat{\boldsymbol{x}}_t = (1-t)\hat{\boldsymbol{x}}_0 + t\boldsymbol{\epsilon}. \tag{2}$$

We use $q_{\theta,t}$ and $p_{T,t}$ for the noised student and teacher laws, including their endpoint laws at $t = 0$. Following integral divergence and distribution matching distillation updates (Luo et al., 2023b; Yin et al., 2024b), we consider

$$\mathcal{J}_{\mathrm{D}}(\theta) = \mathbb{E}_{c,t\sim\rho}\big[w_{\mathrm{D}}(t) D_{\mathrm{KL}}(q_{\theta,t} \| p_{T,t})\big], \qquad w_{\mathrm{D}}(t) \geq 0. \tag{3}$$

Writing $\boldsymbol{s}_{q_\theta,t}(\boldsymbol{x}) := \nabla_{\boldsymbol{x}} \log q_{\theta,t}(\boldsymbol{x})$ and $\boldsymbol{s}_{p_T,t}(\boldsymbol{x}) := \nabla_{\boldsymbol{x}} \log p_{T,t}(\boldsymbol{x})$, its gradient under the regularity conditions in Appendix C is

$$\nabla_\theta \mathcal{J}_{\mathrm{D}} = \mathbb{E}\big[w_{\mathrm{D}}(t)(J_\theta \hat{\boldsymbol{x}}_t)^\top \big(\boldsymbol{s}_{q_\theta,t}(\hat{\boldsymbol{x}}_t) - \boldsymbol{s}_{p_T,t}(\hat{\boldsymbol{x}}_t)\big)\big]. \tag{4}$$

The teacher supplies the target score, but the generator does not automatically supply the score of its own evolving output law. DMD therefore trains an auxiliary velocity model $\boldsymbol{v}_\phi$ on generated samples with the flow-matching target $\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0$, while the frozen teacher provides $\boldsymbol{v}_T$. On the linear path equation 1, DMD replaces the score difference in equation 4 by the velocity-form plug-in

$$\boldsymbol{s}_{q_\theta,t}(\boldsymbol{x}) - \boldsymbol{s}_{p_T,t}(\boldsymbol{x}) \;\longrightarrow\; -\frac{1-t}{t}\big(\boldsymbol{v}_\phi(\boldsymbol{x}, t, c) - \boldsymbol{v}_T(\boldsymbol{x}, t, c)\big). \tag{5}$$

The generator is held fixed when $\boldsymbol{v}_\phi$ is trained. We set $\lambda(t) = w_{\mathrm{D}}(t)(1-t)/t$ below; the additional regression and adversarial terms of DMD and DMD2 are discussed in section 2.

ArcFlow. ArcFlow replaces a constant-velocity shortcut with a momentum-mixture velocity parameterization that represents velocity evolution over an entire sampling interval (Yang et al., 2026). Its backbone feeds three output projections that produce K base velocities, softmax-normalized

gate weights and positive momentum factors, which we call momentum rates below. One network evaluation at a fixed input $(\boldsymbol{x}_t, t, c)$ returns the momentum-mixture parameters

$$\begin{aligned} \mathcal{P}_\theta(\boldsymbol{x}_t, t, c) &= \{\boldsymbol{v}_k, \pi_k, \gamma_k\}_{k=1}^K, \qquad \pi_k > 0, \quad \sum_k \pi_k = 1, \quad \gamma_k > 0, \\ \boldsymbol{v}_{t,s}^\theta &:= \tilde{\boldsymbol{v}}_\theta\big(s; \mathcal{P}_\theta(\boldsymbol{x}_t, t, c)\big) = \sum_{k=1}^K \pi_k \boldsymbol{v}_k \gamma_k^{1-s}, \qquad 0 \le s \le t. \end{aligned} \tag{6}$$

In $\boldsymbol{v}_{t,s}^\theta$, t denotes the time of the input state $\boldsymbol{x}_t$, while $s \in [0, t]$ denotes the time at which the resulting velocity is evaluated. The starting state and condition are implicit in $\boldsymbol{v}_{t,s}^\theta$. With the predicted coefficients fixed, $s \mapsto \boldsymbol{v}_{t,s}^\theta$ defines the student's velocity parameterization over the interval $[0, t]$. Its exact integral defines the endpoint generator

$$G_\theta(\boldsymbol{x}_t, t, c) = \boldsymbol{x}_t - \int_0^t \boldsymbol{v}_{t,s}^\theta \, \mathrm{d}s = \boldsymbol{x}_t - \sum_{k=1}^K \pi_k \boldsymbol{v}_k \gamma_k^{1-t} \int_0^t \gamma_k^u \, \mathrm{d}u.$$

We abbreviate $G_\theta(\boldsymbol{x}_1, 1, c)$ as $G_\theta(\boldsymbol{x}_1, c)$, and equation 16 in Appendix C.1 states the rollout to an arbitrary earlier time.

4 Trajectory-regularized distribution distillation

Trajectory-regularized distribution distillation (TRD) combines distribution matching and teacher-trajectory alignment through a shared ArcFlow parameterization. Distribution matching typically trains an auxiliary model to track the marginal velocity of the student's generated distribution; because an ArcFlow student already outputs a time-dependent velocity, TRD trains the student itself to provide this estimate. We first combine the two signals while retaining the auxiliary model (section 4.1), and then replace it with self flow matching (section 4.2).

4.1 Distribution matching with trajectory regularization

Distribution matching with an auxiliary estimator. We first apply the conventional DMD construction of section 3 to the ArcFlow generator G_θ, retaining $\boldsymbol{v}_\phi$ as the auxiliary velocity model. Using the velocity-form substitution in equation 5, the gradient in equation 4 is implemented by

$$\mathcal{L}_{\mathrm{dm}}(\theta; \phi) = \mathbb{E}\Big[\Big\langle \mathrm{sg}\big[-\lambda(t)(\boldsymbol{v}_\phi(\hat{\boldsymbol{x}}_t, t, c) - \boldsymbol{v}_T(\hat{\boldsymbol{x}}_t, t, c))\big], \hat{\boldsymbol{x}}_t \Big\rangle\Big]. \tag{7}$$

We write $\mathrm{sg}[\cdot]$ for stop-gradient. Because the bracketed velocity difference is detached, automatic differentiation acts only through the path $\hat{\boldsymbol{x}}_t \to \hat{\boldsymbol{x}}_0 \to \theta$. The resulting gradient agrees with equation 4 when $\boldsymbol{v}_\phi$ and $\boldsymbol{v}_T$ recover their corresponding marginal velocities; the scalar value of the surrogate is not the KL divergence itself.

Trajectory alignment from the re-noised state. The state $\hat{\boldsymbol{x}}_t$ is obtained by adding independent noise to the generated endpoint $\hat{\boldsymbol{x}}_0$; it need not lie on the original generation trajectory. Given a decreasing timestep grid $\mathcal{S}_t = \{t, s_1, \ldots, s_M\}$ with $t > s_1 > \cdots > s_M \ge 0$, let $\hat{\boldsymbol{x}}_{t,s}^T$ be the teacher state reached at time s, starting from $\hat{\boldsymbol{x}}_t$ at t. For consecutive grid times $s > s'$, explicit Euler gives

$$\hat{\boldsymbol{x}}_{t,t}^T = \hat{\boldsymbol{x}}_t, \qquad \boldsymbol{v}_{t,s}^T = \boldsymbol{v}_T(\hat{\boldsymbol{x}}_{t,s}^T, s, c), \ s \in \mathcal{S}_t, \qquad \hat{\boldsymbol{x}}_{t,s'}^T = \hat{\boldsymbol{x}}_{t,s}^T - (s - s')\boldsymbol{v}_{t,s}^T. \tag{8}$$

The entire teacher rollout is computed without gradients. The student is evaluated once at $(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)$, giving the coefficients of equation 6. Thus, evaluating $\boldsymbol{v}_{t,s}^\theta$ at every grid time requires no additional student evaluations. The trajectory alignment loss is defined as

$$\mathcal{L}_{\mathrm{tr}}(\theta) = \mathbb{E}\Bigg[\frac{1}{M+1} \sum_{s \in \mathcal{S}_t} \big\|\boldsymbol{v}_{t,s}^\theta - \mathrm{sg}[\boldsymbol{v}_{t,s}^T]\big\|_2^2\Bigg]. \tag{9}$$

This loss teaches the student's velocity parameterization to represent the teacher's velocity field over the whole grid. It compares velocities, but not student endpoints or student states along the grid. Proposition 2 and Theorem 3 in Appendix C give the corresponding expressivity and finite-grid error results.

4.2 REPLACING THE AUXILIARY MODEL WITH SELF FLOW MATCHING

We observe that the velocity parameterization already produces a start-time instantaneous velocity $\boldsymbol{v}_{t,t}^{\theta}$, which can estimate the marginal velocity of the student's generated law. We train this existing output to estimate the marginal velocity through a self flow matching loss, removing the separate auxiliary model:

$$\mathcal{L}_{\mathrm{fm}}(\theta) = \mathbb{E}\Big[\big\|\boldsymbol{v}_{t,t}^{\theta} - \big(\boldsymbol{\epsilon} - \mathrm{sg}[\hat{\boldsymbol{x}}_0]\big)\big\|_2^2\Big]. \tag{10}$$

Here the student is evaluated at the detached state $\mathrm{sg}[\hat{\boldsymbol{x}}_t]$, as in trajectory alignment; gradients update the parameters producing $\boldsymbol{v}_{t,t}^{\theta}$, while the input $\hat{\boldsymbol{x}}_t$ and endpoint target $\hat{\boldsymbol{x}}_0$ remain detached. For a fixed student law, the minimizer of this regression is $\mathbb{E}[\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0 \mid \hat{\boldsymbol{x}}_t, t, c]$, the student's marginal velocity at $\hat{\boldsymbol{x}}_t$; Proposition 3 gives the formal statement.

Replacing $\boldsymbol{v}_{\phi}$ by this shared estimate gives the TRD distribution update:

$$\mathcal{L}_{\mathrm{dm}}(\theta) = \mathbb{E}\Big[\big\langle \mathrm{sg}\big[-\lambda(t)(\boldsymbol{v}_{t,t}^{\theta} - \boldsymbol{v}_{t,t}^{T})\big], \hat{\boldsymbol{x}}_t\big\rangle\Big]. \tag{11}$$

Here $\langle\cdot,\cdot\rangle$ sums over all non-batch axes. Compared with the surrogate in equation 7, equation 11 replaces only the auxiliary estimate $\boldsymbol{v}_{\phi}(\hat{\boldsymbol{x}}_t, t, c)$ with the shared prediction $\boldsymbol{v}_{t,t}^{\theta}$; the stop-gradient construction and endpoint gradient path remain unchanged. When $\boldsymbol{v}_{t,t}^{\theta}$ and $\boldsymbol{v}_{t,t}^{T}$ recover the corresponding marginal velocities, Theorem 1 shows that its automatic-differentiation gradient equals $\nabla_{\theta}\mathcal{J}_{\mathrm{D}}$. Distribution matching and self flow matching share the start-time prediction $\boldsymbol{v}_{t,t}^{\theta}$, while trajectory alignment evaluates the same velocity parameterization over $s \in \mathcal{S}_t$.

Interaction with trajectory alignment. Because $t \in \mathcal{S}_t$, equation 9 pulls $\boldsymbol{v}_{t,t}^{\theta}$ toward the teacher velocity while equation 10 pulls it toward the student marginal velocity, and equation 11 uses their difference. Proposition 4 in Appendix C.2 shows that, when the start-time output is otherwise unconstrained, the joint minimizer is a convex combination of the two targets. The resulting update is the exact reverse-KL update scaled by a factor $\alpha \geq \lambda_{\mathrm{fm}}/(\lambda_{\mathrm{fm}} + \lambda_{\mathrm{tr}})$, which exceeds 0.9 for the weights in Table 4; trajectory alignment therefore reduces the distribution step size and leaves its direction and zeros unchanged.

4.3 STABILIZATION AND OVERALL OBJECTIVE

Distribution preservation. We stabilize distribution matching with a moving-average copy. Let $\bar{\theta}$ be the exponential moving average of θ at rate $\mu \in [0, 1)$. Distribution preservation matches paired endpoints generated from the same initial noise:

$$\mathcal{L}_{\mathrm{dp}}(\theta) = \mathbb{E}_{\boldsymbol{x}_1, c}\Big[\big\|G_{\theta}(\boldsymbol{x}_1, c) - \mathrm{sg}[G_{\bar{\theta}}(\boldsymbol{x}_1, c)]\big\|_2^2\Big]. \tag{12}$$

This loss keeps the current generator map close to its moving-average reference and vanishes when $\theta = \bar{\theta}$. It constrains the original generated endpoint $\hat{\boldsymbol{x}}_0$, not an endpoint of the rollout from $\hat{\boldsymbol{x}}_t$.

Overall objective. Trajectory alignment follows the teacher path, while distribution alignment combines reverse Kullback-Leibler matching, self flow matching and distribution preservation.

$$\mathcal{L}_{\mathrm{TRD}}(\theta) = \underbrace{\lambda_{\mathrm{tr}}\mathcal{L}_{\mathrm{tr}}(\theta)}_{\text{trajectory alignment }(9)} + \underbrace{\lambda_{\mathrm{dm}}\mathcal{L}_{\mathrm{dm}}(\theta)}_{\text{reverse-KL matching }(11)} + \underbrace{\lambda_{\mathrm{fm}}\mathcal{L}_{\mathrm{fm}}(\theta)}_{\text{self flow matching }(10)} + \underbrace{\lambda_{\mathrm{dp}}\mathcal{L}_{\mathrm{dp}}(\theta)}_{\text{distribution preservation }(12)}. \tag{13}$$

One ArcFlow parameterization serves three roles: its analytic integral generates endpoints, its start-time velocity estimates the student's marginal velocity, and its velocities at earlier times provide predictions for trajectory alignment. The student is initialized from the teacher. Its final L transformer blocks are copied into K momentum branches, where L denotes the duplicated depth and each branch predicts $(\boldsymbol{v}_k, \pi_k, \gamma_k)$ in equation 6. Only these branches and the low-rank backbone adapters are trained (Hu et al., 2022).

4.4 TRAINING PROCEDURE

Algorithm 1 summarizes one TRD training iteration. We first generate and re-noise a student endpoint, evaluate the student at the detached noised state to obtain its velocity parameterization, and roll out the frozen teacher from the same state. This requires two student forward passes: one for generation and one for the shared velocity predictions. The teacher trajectory and moving-average endpoint define the remaining targets. All four losses are formed before one joint student update, after which the moving average is updated.

Algorithm 1 One TRD training iteration for one-step distillation

Require: student θ, moving average $\bar{\theta}$ with rate μ, frozen teacher $\boldsymbol{v}_T$, time distribution ρ on $[t_{\min}, t_{\max}]$, grid $\mathcal{S}_t = \{t, s_1, \ldots, s_M\}$ with $t > s_1 > \cdots > s_M \geq 0$, loss weights $\lambda_{\mathrm{tr}}, \lambda_{\mathrm{dm}}, \lambda_{\mathrm{fm}}, \lambda_{\mathrm{dp}}$ of equation 13 and time weight $\lambda(t)$

1: sample c, initial noise $\boldsymbol{x}_1 \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$, $t \sim \rho$ and independent $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
2: **generate:** evaluate the student at $(\boldsymbol{x}_1, 1, c)$ and integrate the output velocity to $\hat{\boldsymbol{x}}_0 = G_\theta(\boldsymbol{x}_1, c)$
3: **re-noise:** form $\hat{\boldsymbol{x}}_t = (1-t)\hat{\boldsymbol{x}}_0 + t\boldsymbol{\epsilon}$, retaining its gradient path to θ
4: **predict velocities:** evaluate the student at $(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)$; obtain $\{\boldsymbol{v}^\theta_{t,s} : s \in \mathcal{S}_t\}$ from equation 6, without further student evaluations
5: **teacher rollout, without gradients:** initialize $\hat{\boldsymbol{x}}^T_{t,t} = \hat{\boldsymbol{x}}_t$; for decreasing $s \in \mathcal{S}_t$, evaluate $\boldsymbol{v}^T_{t,s}$ and, if a next grid time s' exists, advance to $\hat{\boldsymbol{x}}^T_{t,s'}$ by equation 8
6: **reference endpoint, without gradients:** evaluate $G_{\bar{\theta}}(\boldsymbol{x}_1, c)$ on the same initial noise
7: **form the losses:** use the grid velocities for equation 9, $\boldsymbol{v}^\theta_{t,t}$ and $\boldsymbol{\epsilon} - \mathrm{sg}[\hat{\boldsymbol{x}}_0]$ for equation 10, the detached start-time velocity difference and $\hat{\boldsymbol{x}}_t$ for equation 11, and the paired endpoints for equation 12
8: **update:** backpropagate equation 13, take one optimizer step on θ, then set $\bar{\theta} \leftarrow \mu\bar{\theta} + (1-\mu)\theta$

Few-step extension. Inspired by DMD2's backward simulation (Yin et al., 2024a), few-step TRD uses random-length consistency sampling. Fix $1 = a_0 > \cdots > a_R = 0$ and sample $r \sim \mathrm{Unif}\{1, \ldots, R\}$. Starting from $\boldsymbol{x}_{a_0} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$, run $r-1$ detached stages that predict an endpoint and independently re-noise it to the next preset time; only $\hat{\boldsymbol{x}}_0 = G_\theta(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$ retains gradients. Then sample $t \sim \rho_r$ with $\mathrm{supp}(\rho_r) \subseteq [a_r, a_{r-1}]$, form equation 2 and reuse the four losses. Distribution preservation compares student and moving-average endpoints from the same $\boldsymbol{x}_{a_{r-1}}$. The stage times a_i differ from the trajectory-supervision times $s \in \mathcal{S}_t$; Algorithm 2 in Appendix E gives full details.

4.5 TRAINING TARGET ANALYSIS

The four terms of equation 13 supervise different objects, so we check that they admit a common ideal target. Assume exact expectations, an exact unguided teacher, $\lambda(t) > 0$ on a set of positive ρ-measure, and a student expressive enough to represent the teacher endpoint law, its own marginal velocity and the teacher trajectory.

Let $\boldsymbol{v}_{q_\theta,t}$ and $\boldsymbol{v}_{p_T,t}$ denote the true marginal velocity fields of $q_{\theta,t}$ and $p_{T,t}$, respectively.

Proposition 1 (Shared ideal target)**.** *Under the assumptions above, the following statements hold: (i) the population minimizer of equation 10 is the student marginal velocity $\boldsymbol{v}_{q_\theta,t}$; (ii) with that velocity and the exact teacher velocity in place, the automatic-differentiation gradient of equation 11 equals the gradient of*

$$\mathcal{J}_{\mathrm{D}}(\theta) = \mathbb{E}_{c,t\sim\rho}\big[w_{\mathrm{D}}(t) D_{\mathrm{KL}}(q_{\theta,t} \| p_{T,t})\big], \qquad w_{\mathrm{D}}(t) = \frac{t\lambda(t)}{1-t} \geq 0,$$

and every global minimizer of $\mathcal{J}_{\mathrm{D}}$ satisfies $q_{\theta,0} = p_{T,0}$; (iii) if the student attains $q_{\theta,0} = p_{T,0}$, predicts its own marginal velocity, and has zero continuous-time trajectory residual, and if $\bar{\theta} = \theta$, then the expected update induced by each of the four terms in equation 13 is zero.

Proposition 1 shows that replacing the auxiliary model with self flow matching preserves the ideal distribution target and that all four expected updates share a common zero. This common zero does not make every stationary point of the shared-parameter update correct, and it does not imply global

convergence. Appendix C gives the regularity conditions and finite-error bounds, and Appendix D treats guidance, gradient bias and local stability.

5 EXPERIMENTS

5.1 SETUP

Teachers, prompts and metrics. We distil SANA-1.6B (Xie et al., 2025) and Stable Diffusion 3.5 Medium, from the multimodal rectified-flow transformer family (Esser et al., 2024), at 1024×1024. Students initialize from their teachers and train only the policy heads and backbone adapters; TRD uses $K = 4$ and $L = 1$ unless stated otherwise. All runs use disjoint splits of a publicly available dataset (Schuhmann et al., 2022), including 2,500 test prompts. We evaluate teacher-relative fidelity and coverage, diversity, preference, visual appeal and text alignment. Appendix E gives the full configurations and protocols.

Baselines. Table 1 groups baselines by trajectory-oriented (Salimans & Ho, 2022; Song et al., 2023; Song & Dhariwal, 2024; Lu & Song, 2025; Yang et al., 2026; Geng et al., 2025) or distribution-oriented supervision (Yin et al., 2024b;a; Zhou et al., 2024; Kim et al., 2026; Cheng et al., 2026); TRD combines both. For comparability, we adapt the teacher-free MeanFlow and TwinFlow regimes to teacher initialization and the shared data, resolution and target-NFE protocol.

5.2 ONE-STEP AND FEW-STEP DISTILLATION

Table 1 reports one-step and few-step distillation results for SANA-1.6B, with the multi-step teacher as an upper reference; the matched comparison on Stable Diffusion 3.5 Medium is deferred to Table 5 in Appendix G. Figures 5 and 6 in Appendix G compare one-step and few-step samples from PD, CD, MeanFlow, DMD and TRD for the two teachers. Figures 7 and 8 provide larger sets of one-step TRD samples. At both NFEs, TRD leads the preference, aesthetic and text-alignment metrics and attains the lowest FD-DINOv2, and at two steps several of its preference and aesthetic scores exceed the multi-step teacher. It favors quality over diversity more strongly than DMD and related distribution-matching methods: its DINO Diversity falls behind the most diverse baselines at both NFEs, and at two steps its Recall-DINOv2 also falls behind DMD and DMD2, which cover more of the teacher manifold. Trajectory-oriented methods such as MeanFlow retain higher diversity, but lag on preference scores and Fréchet distance. Stable Diffusion 3.5 Medium shows the same pattern, except that DMD attains a slightly lower FD-DINOv2 at two steps.

5.3 TRAINING COST AND TIME-TO-QUALITY TRADE-OFFS

A practical motivation for shared self-estimation is to remove the persistent parameters, gradients and optimizer state of the auxiliary fake network used by DMD-style methods. We therefore compare the number of auxiliary networks, trainable parameters, measured peak memory, normalized per-iteration time and normalized time to a common FD-DINOv2 target. Table 2 reports these measurements for full-scale one-step training. Times are normalized to one DMD training step; Appendix F accounts for the underlying operations and persistent state. TRD uses $K = 4$ momentum modes and $L = 1$ duplicated final block. The right panel of figure 1 plots FD-DINOv2 against normalized training time.

Relative to DMD and DMD2, TRD removes one or two auxiliary networks, reducing trainable parameters and measured peak memory. Teacher-trajectory supervision makes each iteration more expensive, but TRD reaches the common target in substantially less normalized time. Within DMD-style distillation, TRD therefore trades per-iteration computation for less persistent training state and faster measured time to target. MeanFlow and CD remain cheaper in these measurements, so TRD is not the lowest-cost method across all distillation families.

5.4 DISSECTING THE TRD DESIGN

We examine three questions about the TRD design: whether shared self-estimation improves on a separate student-velocity estimator, how the objective terms complement one another and how the

Table 1: One-step and few-step results at 1024×1024 for students distilled from SANA-1.6B, with the multi-step teacher as a reference. Appendix E defines the metrics and Appendix B gives the table conventions.

Method	NFE	Extra models	FD-DINOv2 ↓	Pick ↑	HPS ↑	ImgRwd ↑	Aesth ↑	CLIP ↑	Recall ↑	Diversity ↑
SANA-1.6B teacher	20×2	–	–	0.1971	0.2360	0.5144	6.1879	28.7036	–	0.4100
Trajectory-oriented objectives										
PD	1	0	260.81	0.1764	0.1952	-1.3441	5.6703	24.9226	0.0338	0.3362
CD	1	0	226.80	0.1814	0.2058	-0.8208	5.7909	25.9214	0.0501	0.3181
sCD	1	0	202.58	0.1829	0.2103	-0.6012	5.7925	27.0117	0.0721	0.3290
ArcFlow	1	0	193.68	0.1816	0.2078	-0.4764	5.8057	26.8843	0.0755	0.3411
MeanFlow	1	0	97.21	0.1829	0.2109	-0.1739	5.9320	27.1596	0.3031	**0.4366**
Distribution-oriented objectives										
DMD	1	1	150.11	0.1840	0.2129	-0.3398	5.8483	28.1567	0.0854	0.3250
DMD2	1	2	112.47	0.1894	0.2192	0.1247	6.0468	28.1960	0.1740	0.3499
SiD	1	1	134.82	0.1877	0.2168	-0.0185	5.9812	27.4055	0.1478	0.3321
FSF-DMD	1	0	128.53	0.1873	0.2197	0.2259	6.0810	27.9147	0.0953	0.3142
TwinFlow	1	0	130.29	0.1861	0.2139	0.0526	5.9782	27.5482	0.0600	0.3233
Combined trajectory and distribution objective										
TRD (ours)	1	0	**70.34**	**0.1915**	**0.2223**	**0.3888**	**6.1169**	**28.3647**	**0.3162**	0.3204
Trajectory-oriented objectives										
PD	2	0	91.50	0.1833	0.2128	-0.3924	5.8369	26.7599	0.4250	0.4686
CD	2	0	177.65	0.1779	0.2133	-0.6014	5.6130	26.0449	0.1800	0.4027
sCD	2	0	137.31	0.1891	0.2124	0.1128	5.7723	26.6751	0.3691	0.4254
ArcFlow	2	0	138.74	0.1862	0.2091	0.0919	6.1177	26.6285	0.2825	**0.4715**
MeanFlow	2	0	83.89	0.1917	0.2151	0.1388	6.2768	27.0492	0.4934	0.4287
Distribution-oriented objectives										
DMD	2	1	29.05	0.1981	0.2339	0.4927	6.3019	28.4341	**0.8263**	0.3978
DMD2	2	2	35.92	0.1975	0.2346	0.4861	6.2987	28.0849	0.7251	0.3732
SiD	2	1	47.58	0.1974	0.2325	0.5037	6.2126	27.8064	0.4527	0.3929
FSF-DMD	2	0	39.41	0.1927	0.2308	0.4584	6.1869	28.0417	0.3308	0.3877
TwinFlow	2	0	36.57	0.1941	0.2285	0.4273	6.0078	28.0961	0.4068	0.3640
Combined trajectory and distribution objective										
TRD (ours)	2	0	**28.49**	**0.1993**	**0.2353**	**0.6746**	**6.3105**	**28.4355**	0.7151	0.3380

Table 2: Measured memory and time cost of one-step distillation from SANA-1.6B under full-scale training. Appendix B defines the reported columns.

Method	Aux. nets	Trainable params	Peak mem. (GB) ↓	Norm. time / iter ↓	Norm. time to target ↓
MeanFlow	0	1.6B	40	1.2	2400
CD	0	1.6B	40	0.7	1400
DMD	1	3.2B	75	1.0	10000
DMD2	2	3.3B	80	1.2	12000
TRD (ours)	0	1.9B	45	1.5	4500

capacity of the velocity parameterization affects performance. Unless stated otherwise, the comparisons use one-step SANA-1.6B distillation; Appendices E and G give the full settings and results.

Shared generation and velocity estimation. Table 7 replaces the shared velocity estimate with a separately trained DMD-style fake network under otherwise matched settings. The shared parameterization improves FD-DINOv2, preference metrics and teacher coverage, whereas the separate estimator retains slightly greater diversity. Shared self-estimation therefore provides a measured quality benefit in addition to removing the auxiliary trainable model and its associated state.

Complementary objective roles. Under matched settings, Table 6 shows that removing trajectory alignment degrades fidelity, preference and coverage while preserving similar diversity, demonstrating benefits beyond the distribution objectives. Jointly removing self flow matching and distribution matching broadly degrades performance, although this comparison does not isolate their individual effects. Removing distribution preservation collapses training, indicating that it is critical for stability in this setting.

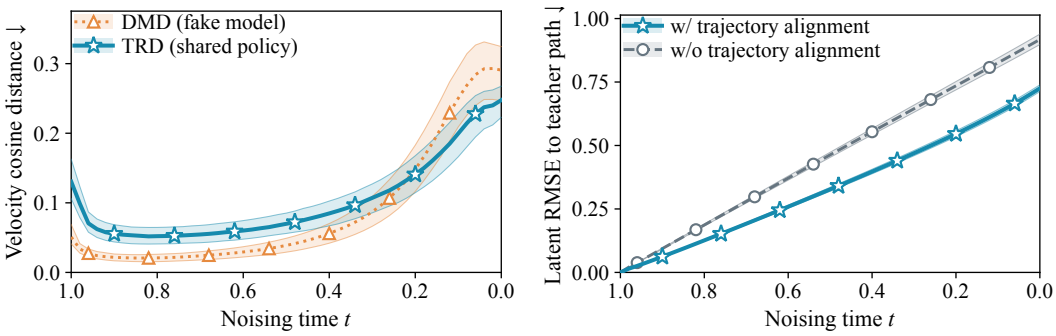


Figure 2: Mechanism analysis at the final checkpoint. Left: velocity cosine distance against noising time for DMD and TRD. Right: discrepancy from the teacher denoising path for TRD with and without trajectory alignment. Appendix B describes both panels.

Velocity-parameterization capacity. Table 8 varies the number of momentum modes K and the duplicated depth L. At $L = 1$, increasing K from one to two improves every reported metric, but four or eight modes reverse this trend despite the larger representational family of Proposition 2; figure 4 rules out simple collapse of the gates and rates at $t = 1$ but does not explain this reversal. At fixed K, increasing L consistently improves quality and coverage with nearly unchanged diversity. Duplicated depth therefore gives a predictable trade-off between quality and cost, whereas K should be tuned.

5.5 Understanding TRD's performance gains

figure 2 examines two possible sources of TRD's gains on final Stable Diffusion 3.5 Medium checkpoints; Appendix B gives the protocol.

Student velocity estimation. DMD uses a separately trained fake velocity estimator, while TRD reads the estimate from its generation policy. The left panel shows that TRD has weaker directional agreement at high noise but becomes more accurate near the data endpoint. Because this estimate supplies the distribution-matching direction in equation 11, the low-noise advantage is consistent with TRD's stronger one-step quality in Table 5. This diagnostic is correlational and does not establish causality.

Trajectory alignment. The right panel compares matched TRD variants with and without $\mathcal{L}_{\mathrm{tr}}$. Trajectory alignment lowers the discrepancy from the teacher path at every noninitial grid point, although each training grid covers only a short segment below its resampled start time (Table 4).

6 Conclusion

We introduced trajectory-regularized distribution distillation, whose single momentum-mixture policy supports closed-form generation, student marginal velocity estimation and teacher-trajectory prediction without a separately trained fake model. Our analysis identifies a common ideal target for the supervision signals, shows that trajectory supervision of the shared estimator preserves the direction of the distribution update, and bounds their finite-error complementarity. Experiments on two text-to-image models support the quality benefits of this design and its efficiency relative to fake-model distribution matching, while revealing a trade-off in teacher-relative coverage and diversity. The evaluation covers two public image models, measures fidelity against teacher samples, reports no run-to-run variability and uses feature-space diversity metrics that miss some forms of conditional diversity; why training collapses without distribution preservation also remains open. Future work should test broader model families, real-data benchmarks and guidance settings.

Reproducibility statement

The method is specified by equation 13, equation 6 through equation 12, and Algorithms 1 and 2. Appendix E records the teacher checkpoint identifiers, policy configuration, objective weights and schedules, trajectory grid and solver, optimizer settings, prompt split size and seed, shared image-seed schedule, metric definitions and model families. Appendix C formalizes the ideal argument of section 4.5 under Assumption 1 and gives complete proofs of Propositions 2, 3, 4 and 5, Theorems 1, 2 and 3, and Corollary 1. Appendix D gives the guidance extension, estimation-error analysis and convergence qualifications.

Ethics statement

TRD reduces the inference cost of pretrained text-to-image generators and requires no additional target images; the image experiments use prompts from a publicly available dataset during distillation. It therefore inherits the properties of its teachers, including any demographic or cultural bias in their training corpora and their capacity to produce misleading or harmful image content. Lowering inference cost can increase the rate at which such content is produced. The distillation objective in equation 11 is mode-seeking, so a distilled student can narrow the output distribution relative to its teacher in ways that aggregate quality metrics do not reveal. We recommend that any deployment evaluate diversity and safety on its own input distribution. The text-to-image teachers, evaluation datasets, benchmarks and metric models are publicly released and are used under their published licenses.

AI use statement

We used a large language model as a writing assistant for grammatical correction, structural revision and improving clarity in the introduction, method, experimental description, ideal-state correctness argument and appendix, and as a coding assistant for routine implementation of evaluation scripts and plotting utilities. Its mathematical role was limited to reorganizing and rephrasing existing author-provided definitions, claims and derivations. We did not use generative AI to originate research ideas, derive the method or theoretical claims, design experiments, or generate, alter or select any experimental result. All text produced with AI assistance was read and edited by at least two authors, every mathematical edit was checked against the original definitions and derivations, all code produced with AI assistance was reviewed line by line and validated against reference implementations of the corresponding metrics, and every reference was checked against the authoritative paper or technical report available for the cited work. The authors take full responsibility for the final content of this work, including its text, claims, figures, tables and released artifacts.

References

Nicholas Matthew Boffi, Michael Samuel Albergo, and Eric Vanden-Eijnden. Flow map matching with stochastic interpolants: A mathematical framework for consistency models. *Transactions on Machine Learning Research*, 2025.

Vivek S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

Clément Chadebec, Onur Taşar, Eyal Benaroche, and Benjamin Aubin. Flash diffusion: Accelerating any conditional diffusion model for few steps image generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 15686–15695, 2025. doi: 10.1609/aaai.v39i15.33722.

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*, 2026.

Junsong Chen, Shuchen Xue, Yuyang Zhao, Jincheng Yu, Sayak Paul, Junyu Chen, Han Cai, Song Han, and Enze Xie. SANA-Sprint: One-step diffusion with continuous-time consistency distillation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16185–16195, 2025. doi: 10.1109/ICCV51701.2025.01502.

Zhenglin Cheng, Peng Sun, Jianguo Li, and Tao Lin. TwinFlow: Realizing one-step generation on large models with self-adversarial flows. In *International Conference on Learning Representations*, 2026.

Timothée Darcet, Maxime Oquab, Julien Mairal, and Piotr Bojanowski. Vision transformers need registers. In *International Conference on Learning Representations*, 2024.

Tim Dettmers, Mike Lewis, Sam Shleifer, and Luke Zettlemoyer. 8-bit optimizers via block-wise quantization. In *International Conference on Learning Representations*, 2022.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, Dustin Podell, Tim Dockhorn, Zion English, and Robin Rombach. Scaling rectified flow transformers for high-resolution image synthesis. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12606–12633. PMLR, 2024.

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

Zhengyang Geng, Mingyang Deng, Xingjian Bai, J. Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. In *Advances in Neural Information Processing Systems*, volume 38, pp. 84032–84054, 2025. doi: 10.52202/085713-2534.

Jiatao Gu, Chen Wang, Shuangfei Zhai, Yizhe Zhang, Lingjie Liu, and Joshua M. Susskind. Data-free distillation of diffusion models with bootstrapping. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 16622–16646. PMLR, 2024.

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. In *Advances in Neural Information Processing Systems*, volume 30, pp. 6626–6637, 2017.

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

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

Minguk Kang, Richard Zhang, Connelly Barnes, Sylvain Paris, Suha Kwak, Jaesik Park, Eli Shechtman, Jun-Yan Zhu, and Taesung Park. Distilling diffusion models into conditional GANs. In *Computer Vision – ECCV 2024*, volume 15086 of *Lecture Notes in Computer Science*, pp. 428–447. Springer, 2024. doi: 10.1007/978-3-031-73390-1_25.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems*, volume 35, pp. 26565–26577, 2022. doi: 10.52202/068431-1926.

Dongjun Kim, Chieh-Hsin Lai, Wei-Hsiang 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*, 2024.

Youngjoong Kim, Deokyeong Lee, and Jaesik Park. Distribution matching distillation without fake score network. *arXiv preprint arXiv:2605.19256*, 2026.

Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-Pic: An open dataset of user preferences for text-to-image generation. In *Advances in Neural Information Processing Systems*, volume 36, pp. 36652–36663, 2023.

Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Improved precision and recall metric for assessing generative models. In *Advances in Neural Information Processing Systems*, volume 32, pp. 3927–3936, 2019.

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

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

Xingchao Liu, Xiwen Zhang, Jianzhu Ma, Jian Peng, and Qiang Liu. InstaFlow: One step is enough for high-quality diffusion-based text-to-image generation. In *International Conference on Learning Representations*, 2024.

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

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

Yanzuo Lu, Yuxi Ren, Xin Xia, Shanchuan Lin, Xing Wang, Xuefeng Xiao, Andy J. Ma, Xiaohua Xie, and Jian-Huang Lai. Adversarial distribution matching for diffusion distillation towards efficient image and video synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16818–16829, 2025. doi: 10.1109/ICCV51701.2025.01562.

Eric Luhman and Troy Luhman. Knowledge distillation in iterative generative models for improved sampling speed. *arXiv preprint arXiv:2101.02388*, 2021. doi: 10.48550/arXiv.2101.02388.

Simian Luo, Yiqin Tan, Longbo Huang, Jian Li, and Hang Zhao. Latent consistency models: Synthesizing high-resolution images with few-step inference. *arXiv preprint arXiv:2310.04378*, 2023a.

Weijian Luo, Tianyang Hu, Shifeng Zhang, Jiacheng Sun, Zhenguo Li, and Zhihua Zhang. Diff-instruct: A universal approach for transferring knowledge from pre-trained diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 76525–76546, 2023b. doi: 10.52202/075280-3344.

Weijian Luo, Zemin Huang, Zhengyang Geng, J. Zico Kolter, and Guo-Jun Qi. One-step diffusion distillation through score implicit matching. In *Advances in Neural Information Processing Systems*, volume 37, pp. 115377–115408, 2024. doi: 10.52202/079017-3664.

Yihong Luo, Tianyang Hu, Jiacheng Sun, Yujun Cai, and Jing Tang. Learning few-step diffusion models by trajectory distribution matching. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 17719–17728, 2025. doi: 10.1109/ICCV51701.2025.01646.

Thuan Hoang Nguyen and Anh Tran. SwiftBrush: One-step text-to-image diffusion model with variational score distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7807–7816, 2024. doi: 10.1109/CVPR52733.2024.00746.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy V. Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, Mahmoud Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Hervé Jégou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research*, 2024.

Felix Otto and Cédric Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000. doi: 10.1006/jfan.1999.3557.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4195–4205, 2023.

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

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695, 2022.

Amirmojtaba Sabour, Sanja Fidler, and Karsten Kreis. Align your flow: Scaling continuous-time flow map distillation. In *Advances in Neural Information Processing Systems*, volume 38, pp. 162568–162621, 2025. doi: 10.52202/085713-4900.

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

Tim Salimans, Thomas Mensink, Jonathan Heek, and Emiel Hoogeboom. Multistep distillation of diffusion models via moment matching. In *Advances in Neural Information Processing Systems*, volume 37, pp. 36046–36070, 2024. doi: 10.52202/079017-1136.

Axel Sauer, Frederic Boesel, Tim Dockhorn, Andreas Blattmann, Patrick Esser, and Robin Rombach. Fast high-resolution image synthesis with latent adversarial diffusion distillation. In *SIGGRAPH Asia 2024 Conference Papers*, pp. 1–11, 2024a. doi: 10.1145/3680528.3687625.

Axel Sauer, Dominik Lorenz, Andreas Blattmann, and Robin Rombach. Adversarial diffusion distillation. In *Computer Vision – ECCV 2024*, volume 15144 of *Lecture Notes in Computer Science*, pp. 87–103. Springer, 2024b. doi: 10.1007/978-3-031-73016-0_6.

Christoph Schuhmann, Romain Beaumont, Richard Vencu, Cade Gordon, Ross Wightman, Mehdi Cherti, Theo Coombes, Aarush Katta, Clayton Mullis, Mitchell Wortsman, Patrick Schramowski, Srivatsa Kundurthy, Katherine Crowson, Ludwig Schmidt, Robert Kaczmarczyk, and Jenia Jitsev. LAION-5B: An open large-scale dataset for training next generation image-text models. In *Advances in Neural Information Processing Systems, Datasets and Benchmarks Track*, volume 35, pp. 25278–25294, 2022.

Huiyang Shao, Xin Xia, Yuxi Ren, Xing Wang, and Xuefeng Xiao. VarFlow: Proper scoring-rule diffusion distillation via energy matching. In *Advances in Neural Information Processing Systems*, volume 38, pp. 61522–61538, 2025. doi: 10.52202/085713-1847.

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

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

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 32211–32252. PMLR, 2023.

Dmitrii Sorokin, Maksim Nakhodnov, Andrey Kuznetsov, and Aibek Alanov. ImageReFL: Balancing quality and diversity in human-aligned diffusion models. *arXiv preprint arXiv:2505.22569*, 2025. URL https://arxiv.org/abs/2505.22569.

George Stein, Jesse Cresswell, Rasa Hosseinzadeh, Yi Sui, Brendan Ross, Valentin Villecroze, Zhaoyan Liu, Anthony L. Caterini, Eric Taylor, and Gabriel Loaiza-Ganem. Exposing flaws of generative model evaluation metrics and their unfair treatment of diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 3732–3784, 2023.

Zhengyi Wang, Cheng Lu, Yikai Wang, Fan Bao, Chongxuan Li, Hang Su, and Jun Zhu. ProlificDreamer: High-fidelity and diverse text-to-3d generation with variational score distillation. In *Advances in Neural Information Processing Systems*, volume 36, pp. 8406–8441, 2023. doi: 10.52202/075280-0368.

Xiaoshi Wu, Yiming Hao, Keqiang Sun, Yixiong Chen, Feng Zhu, Rui Zhao, and Hongsheng Li. Human preference score v2: A solid benchmark for evaluating human preferences of text-to-image synthesis. *arXiv preprint arXiv:2306.09341*, 2023.

Enze Xie, Junsong Chen, Junyu Chen, Han Cai, Haotian Tang, Yujun Lin, Zhekai Zhang, Muyang Li, Ligeng Zhu, Yao Lu, and Song Han. SANA: Efficient high-resolution text-to-image synthesis with linear diffusion transformers. In *International Conference on Learning Representations*, 2025.

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. ImageReward: Learning and evaluating human preferences for text-to-image generation. In *Advances in Neural Information Processing Systems*, volume 36, pp. 15903–15935, 2023.

Zihan Yang, Shuyuan Tu, Licheng Zhang, Qi Dai, Yu-Gang Jiang, and Zuxuan Wu. ArcFlow: Unleashing 2-step text-to-image generation via high-precision non-linear flow distillation. *arXiv preprint arXiv:2602.09014*, 2026.

Tianwei Yin, Michaël Gharbi, Taesung Park, Richard Zhang, Eli Shechtman, Frédo Durand, and William T. Freeman. Improved distribution matching distillation for fast image synthesis. In *Advances in Neural Information Processing Systems*, volume 37, pp. 47455–47487, 2024a. doi: 10.52202/079017-1505.

Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Frédo Durand, William T. Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6613–6623, 2024b. doi: 10.1109/CVPR52733.2024.00632.

Xin Yu, Xiaojuan Qi, Zhengqi Li, Kai Zhang, Richard Zhang, Zhe Lin, Eli Shechtman, Tianyu Wang, and Yotam Nitzan. Self-evaluation unlocks any-step text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7816–7826, 2026.

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 *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 62307–62331. PMLR, 2024.

Mingyuan Zhou, Zhendong Wang, Huangjie Zheng, and Hai Huang. Guided score identity distillation for data-free one-step text-to-image generation. In *International Conference on Learning Representations*, 2025a.

Mingyuan Zhou, Huangjie Zheng, Yi Gu, Zhendong Wang, and Hai Huang. Adversarial score identity distillation: Rapidly surpassing the teacher in one step. In *International Conference on Learning Representations*, 2025b.

Huaisheng Zhu, Teng Xiao, Shijie Zhou, Zhimeng Guo, Hangfan Zhang, Siyuan Xu, and Vasant G. Honavar. Simple distillation for one-step diffusion models. In *Advances in Neural Information Processing Systems*, volume 38, pp. 102173–102196, 2025. doi: 10.52202/085713-3080.

A NOTATION

Table 3 collects only the recurring method symbols. Symbols used in one proof are defined where that proof begins.

B FIGURE AND TABLE CONVENTIONS

Overview figure. In the left panel of figure 1, the labeled arrows at the noised student state are the start-time velocities $\boldsymbol{v}^{\theta}_{t,t}$ and $\boldsymbol{v}^{T}_{t,t}$ whose difference forms the distribution matching direction of equation 11. The vertical double arrows mark the trajectory grid points at which equation 9 compares $\boldsymbol{v}^{\theta}_{t,s}$ with $\boldsymbol{v}^{T}_{t,s}$ for $s \in \mathcal{S}_t$, and the two dashed paths end at the student and teacher endpoint distributions. In the right panel, the time axis is logarithmic, all methods share the training settings of section 5.1, and line styles and markers identify the methods in grayscale.

Table 3: Core symbols used by the method.

Symbol	Meaning
t, s, τ	input, policy-evaluation and generic proof times, with noise at 1 and data at 0
$\boldsymbol{x}_0, \boldsymbol{\epsilon}, \boldsymbol{x}_t$	data, Gaussian noise and their interpolation in equation 1
$\hat{\boldsymbol{x}}_0, \hat{\boldsymbol{x}}_t$	generated endpoint and its noised state
c, s_{cfg}	condition and classifier-free guidance scale
$\boldsymbol{v}_T, \boldsymbol{v}_{T,\mathrm{cfg}}$	teacher velocity and its guided version
$\theta, \bar{\theta}$	student parameters and their moving average
$\mathcal{P}_\theta, \{\boldsymbol{v}_k, \pi_k, \gamma_k\}_{k=1}^K$	policy and its K momentum modes
$G_\theta, \tilde{\boldsymbol{v}}_\theta$	analytic endpoint generator and momentum-mixture velocity parameterization
$q_{\theta,t}, p_{T,t}, \boldsymbol{v}_{q_\theta,t}, \boldsymbol{v}_{p_T,t}$	noised laws and their marginal velocities
$\lambda_{\mathrm{tr}}, \lambda_{\mathrm{dm}}, \lambda_{\mathrm{fm}}, \lambda_{\mathrm{dp}}$	fixed loss weights in equation 13
$\rho, \lambda(\cdot), w_{\mathrm{D}}(\cdot)$	time sampling and time-dependent distribution weights
a_i, r, ρ_r	few-step stage times, sampled rollout length and stage-conditioned time distribution
$\mathcal{S}_t, \hat{\boldsymbol{x}}_{t,s}^T, \boldsymbol{v}_{t,s}^\theta, \boldsymbol{v}_{t,s}^T$	trajectory grid, teacher states and student/teacher grid velocities
$\mathrm{sg}[\cdot]$	stop-gradient

Metric abbreviations and table styling. Recall abbreviates Recall-DINOv2, Diversity and DINO Div. abbreviate DINO Diversity, and Pick, HPS, ImgRwd and Aesth abbreviate PickScore, HPSv2, ImageReward and the aesthetic score. The extra-model and auxiliary-network columns count trainable parameter sets besides the student and exclude the frozen teacher. In the gray teacher reference rows, NFE is written as sampling steps times the two classifier-free-guidance branches, and FD-DINOv2 and Recall-DINOv2 are omitted because the teacher images define their reference. Within each NFE block of Tables 1 and 5, the best and second-best student values are bolded and underlined; the teacher row does not take part in this ranking. The last column of Table 2 reports the normalized time needed to reach the FD-DINOv2 target level of the right panel of figure 1. The upper block of Table 8 varies K at the default L, its lower block varies L at the default K, and its parameter column reports additional trainable parameters as a percentage of the teacher parameter count.

Mechanism figure. The left panel of figure 2 compares students distilled from Stable Diffusion 3.5 Medium at checkpoints of equal training length. Each method generates one sample for each of 512 held-out test prompts with a single unguided student step, and both methods share the prompts and the noise drawn for equation 2, so the two curves are paired. At each noising time, the start-time estimate $\boldsymbol{v}_{t,t}^\theta$ is compared by cosine distance with the sample-conditioned velocity $\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0$, where lower values indicate stronger directional agreement. The curves cross between $t = 0.24$ and $t = 0.22$, after which TRD has the lower distance toward $t = 0$. The time grid holds 51 points spaced by 0.02 from $t = 1$ to $t = 0$. Each curve is the median over samples, each band spans the 25th to the 75th percentile, and line styles and markers identify the methods in grayscale. The right panel compares matched one-step TRD variants with and without trajectory alignment at the same training checkpoint. For each of 512 held-out prompts, both students and the teacher start from the same noise; the analytic student paths and the teacher path are compared at the same 51 times. The teacher path uses 100 Euler evaluations with a guidance scale of 4.5, while each student uses one unguided evaluation. Each curve is the mean latent RMSE over prompts, and each band is a pointwise 95% percentile bootstrap interval over prompts from 2,000 resamples.

Training-loss figure. figure 3 reports every loss after applying its objective weight. All curves are solid, and their markers identify the four objectives in grayscale.

Policy-head distributions. figure 4 uses a $K = 4$ Stable Diffusion 3.5 Medium checkpoint without trajectory alignment. We evaluate one unguided conditional forward at $t = 1$ for each of 512 held-out prompts, using one initial-noise sample per prompt. Gate weights and momentum rates use separate pooled vertical ranges. Every head within one panel shares the same range, and no samples are clipped or rescaled by head. Violin width represents kernel density, the thin and thick vertical segments span the 5th to 95th and 25th to 75th percentiles, and the white marker denotes the median. All gate distributions remain well above zero and the rate distributions remain separated, so

neither plotted output exhibits collapse at $t = 1$. These statistics are descriptive and do not identify training-induced specialization or per-head agreement with the teacher trajectory.

Qualitative figures. Figures 5 and 6 use prompts from the held-out test split of the publicly available dataset (Schuhmann et al., 2022), and samples for the same prompt and sampling budget share their initial noise across methods. Figures 7 and 8 show additional one-step TRD samples for the two teachers.

C TECHNICAL ANALYSIS AND PROOFS

This appendix proves the statements collected in Proposition 1 and then describes what remains once the ideal assumptions are removed. The argument runs in three stages that depend on each other in this order. The first stage concerns the policy alone: one forward pass returns coefficients that define an explicit function of time, so the induced trajectory is available in closed form, and Proposition 2 records how many independent velocities such a field can carry on a finite grid. The second stage identifies what the two distribution terms optimize: Proposition 3 turns the shared student output into the marginal velocity of the current student law and converts velocities into scores, Theorem 1 uses both facts to show that the detached inner product of equation 11 differentiates to a time-averaged reverse Kullback-Leibler gradient, Proposition 4 shows that the trajectory target on the same start-time output only rescales that gradient, and Theorem 2 shows that the zeros of that divergence are exactly the students whose endpoint law is the teacher endpoint law. The third stage replaces exact quantities by errors: Theorem 3 bounds the endpoint error of the analytic rollout by the grid loss that is actually minimized, Corollary 1 adds the trajectory error and the distribution error in a single Wasserstein bound, and Proposition 5 states what distribution preservation controls. Appendix D then removes the remaining idealizations, namely the unguided teacher and exact optimization.

Conventions. All statements hold for almost every condition c, which is suppressed unless it is averaged explicitly. The generated endpoint $\hat{\boldsymbol{x}}_0 = G_\theta(\boldsymbol{x}_1, c)$ and its noised state $\hat{\boldsymbol{x}}_t$ from equation 2 are treated as random variables with laws $q_{\theta,0}$ and $q_{\theta,t}$, and the arguments of velocity and score fields are dropped whenever those fields are evaluated at the sampled state. For an endpoint law χ and $t \in (0,1)$, the Gaussian noising operator, the marginal velocity and the score are

$$\begin{aligned} &\boldsymbol{X} \sim \chi, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}) \text{ independently}, \quad \boldsymbol{X}_t = (1-t)\boldsymbol{X} + t\boldsymbol{\epsilon}, \quad \mathcal{N}_t\chi = \mathrm{Law}(\boldsymbol{X}_t), \\ &\boldsymbol{v}_{\chi,t}(\boldsymbol{x}) = \mathbb{E}\big[\boldsymbol{\epsilon} - \boldsymbol{X} \mid \boldsymbol{X}_t = \boldsymbol{x}\big], \qquad \boldsymbol{s}_{\chi,t}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}} \log\big(\mathcal{N}_t\chi\big)(\boldsymbol{x}). \end{aligned} \tag{14}$$

Applying equation 14 to $\chi = q_{\theta,0}$ and $\chi = p_{T,0}$ gives the noised laws $q_{\theta,t}$ and $p_{T,t}$ and the four fields $\boldsymbol{v}_{q_\theta,t}$, $\boldsymbol{v}_{p_T,t}$, $\boldsymbol{s}_{q_\theta,t}$ and $\boldsymbol{s}_{p_T,t}$ used below. We write $J_\theta G_\theta$ for the Jacobian of the generator in θ and $\|\cdot\|_{\mathrm{op}}$ for the operator norm.

Assumption 1 (Regularity)**.** *The following hold for almost every condition c. (i) ρ is supported in $[t_{\min}, t_{\max}] \subset (0,1)$. (ii) $\theta \mapsto G_\theta(\boldsymbol{x}_1, c)$ is differentiable with $\mathbb{E}\|J_\theta G_\theta\|_{\mathrm{op}}^2 < \infty$ and $\mathbb{E}\|G_\theta\|_2^2 < \infty$. (iii) For every $t \in (0,1)$ the laws $q_{\theta,t}$ and $p_{T,t}$ have positive densities that are continuously differentiable in $\boldsymbol{x}$, and $D_{\mathrm{KL}}(q_{\theta,t}\|p_{T,t}) < \infty$. (iv) Differentiation in θ may be exchanged with the integrals below, and the boundary terms of the integrations by parts vanish at infinity. (v) $\lambda(\cdot)$ and $t\lambda(t)/(1-t)$ are non-negative and ρ-integrable.*

Item (i) keeps the velocity-to-score conversion away from its singularity at $t = 0$ and keeps the noised law informative about the endpoint at $t = 1$; items (ii) to (iv) are what the gradient computation in Theorem 1 uses; item (v) makes the weighted divergence in equation 3 finite.

C.1 THE POLICY FIELD AND ITS REACH ON A TRAJECTORY GRID

The coefficients returned by one forward pass are held fixed over the interval, so the velocity in equation 6 is an explicit function of time and can be integrated in closed form. With

$$C(\gamma, \Delta) := \int_0^\Delta \gamma^u \,\mathrm{d}u = \begin{cases} (\gamma^\Delta - 1)/\log\gamma, & \gamma \neq 1, \\ \Delta, & \gamma = 1, \end{cases} \tag{15}$$

integrating equation 6 backward from $\boldsymbol{x}(t) = \boldsymbol{x}_t$ gives the state at any earlier time,

$$\boldsymbol{x}(\tau) = \boldsymbol{x}_t - \sum_{k=1}^{K} \pi_k \boldsymbol{v}_k \gamma_k^{1-t} C(\gamma_k, t-\tau), \qquad \tau \le t, \tag{16}$$

and the endpoint generator of section 3 is $G_\theta(\boldsymbol{x}_t, t, c) = \boldsymbol{x}(0)$. Yang et al. (2026) derive equation 16, its constant-velocity limit at $\gamma_k = 1$ and exact finite-grid interpolation with sufficiently many modes. TRD adopts the rollout and finite-grid supervision. We record the full rank statement to cover both undercomplete and overcomplete policies; TRD's methodological change is to couple this policy with distribution matching and self velocity estimation.

Fix a start $(\boldsymbol{x}_t, t, c)$ and write the distinct grid times as $\mathcal{S}_t = \{\tau_0, \dots, \tau_M\}$, where $t = \tau_0 > \cdots > \tau_M \ge 0$. Collect the mode coefficients at the start into $\boldsymbol{m}_k = \pi_k \boldsymbol{v}_k \gamma_k^{1-t}$ and define the evaluation matrix $\boldsymbol{V} \in \mathbb{R}^{(M+1)\times K}$ with entries $\boldsymbol{V}_{jk} = \gamma_k^{t-\tau_j}$, so that the grid velocities of equation 9 are the linear images

$$\boldsymbol{v}^\theta_{t,\tau_j} = \tilde{\boldsymbol{v}}_\theta(\tau_j; \mathcal{P}_\theta(\boldsymbol{x}_t, t, c)) = \sum_{k=1}^{K} \boldsymbol{V}_{jk} \boldsymbol{m}_k, \qquad j = 0, \dots, M. \tag{17}$$

Proposition 2 (Grid rank of the ArcFlow policy)**.** *If the grid times are distinct and $\gamma_1, \dots, \gamma_K$ are positive and pairwise distinct, then* $\operatorname{rank} \boldsymbol{V} = \min\{M+1, K\}$. *In particular, $K \ge M+1$ makes equation 17 solvable for any prescribed grid velocities, and for an equally spaced grid of spacing h_{grid} with $K = M+1$ the matrix $\boldsymbol{V}$ is a Vandermonde matrix in the values $\gamma_k^{h_{\mathrm{grid}}}$, with* $\det \boldsymbol{V} = \pm \prod_{p<q} (\gamma_q^{h_{\mathrm{grid}}} - \gamma_p^{h_{\mathrm{grid}}}) \neq 0$.

Proof. Write $\boldsymbol{V}_{jk} = e^{\kappa_k u_j}$ with $\kappa_k = \log \gamma_k$ and $u_j = t - \tau_j$, so that both families are pairwise distinct real numbers. The proof uses one elementary fact: a function $u \mapsto \sum_{k=1}^{K} c_k e^{\kappa_k u}$ with distinct real κ_k and some $c_k \neq 0$ has at most $K-1$ real zeros. This follows by induction on K, because dividing by $e^{\kappa_1 u}$ leaves the zeros unchanged and differentiating then removes the first term while, by Rolle's theorem, losing at most one zero.

If $K \le M+1$ and $\boldsymbol{V}\boldsymbol{c} = \boldsymbol{0}$ for some $\boldsymbol{c} \neq \boldsymbol{0}$, the function $u \mapsto \sum_k c_k e^{\kappa_k u}$ vanishes at the $M+1 \ge K$ distinct points u_j, contradicting that fact, so the columns are independent and $\operatorname{rank} \boldsymbol{V} = K$. If $K > M+1$ and $\boldsymbol{b}^\top \boldsymbol{V} = \boldsymbol{0}^\top$ for some $\boldsymbol{b} \neq \boldsymbol{0}$, the function $\kappa \mapsto \sum_j b_j e^{\kappa u_j}$ has $M+1$ terms with distinct rates u_j and vanishes at the $K > M+1$ distinct points κ_k, again a contradiction, so the rows are independent and $\operatorname{rank} \boldsymbol{V} = M+1$. Full row rank makes equation 17 solvable in $\{\boldsymbol{m}_k\}$ coordinate by coordinate, and any solution is realizable by the policy, because the softmax keeps $\pi_k > 0$ while the base velocities $\boldsymbol{v}_k$ are unconstrained. On an equally spaced grid $u_j = j h_{\mathrm{grid}}$, so $\boldsymbol{V}_{jk} = (\gamma_k^{h_{\mathrm{grid}}})^j$ and the determinant is the Vandermonde determinant. □

The interpolation consequence of Proposition 2 recovers the ArcFlow finite-grid result, while the full rank formula also states the undercomplete case. It is an existence statement at one fixed start and does not assert that training finds the interpolating coefficients. The rank formula also identifies conditioning limits: when $K \le M+1$, two momentum rates approaching each other drive the smallest singular value toward zero; when $M+1 \le K$, two grid times approaching each other have the analogous effect. The coefficients that realize prescribed grid velocities can therefore become sensitive to supervision error. The result supports the single-forward design only when the relevant rates and grid times remain separated.

C.2 What the distribution terms optimize

Self flow matching and the velocity-to-score dictionary. The first step fixes the meaning of the shared student output. Self flow matching regresses $\boldsymbol{v}^\theta_{t,t}$ on the conditional velocity $\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0$ of the generated sample, and the following statement identifies the resulting population target and records the identity that converts velocities into scores on the path equation 1.

Proposition 3 (Self flow matching identifies the student marginal velocity)**.** *For every square-integrable velocity predictor $\boldsymbol{v}$,*

$$\mathbb{E}\big\|\boldsymbol{v}(\hat{\boldsymbol{x}}_t, t, c) - (\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0)\big\|_2^2 = \mathbb{E}\big\|\boldsymbol{v}(\hat{\boldsymbol{x}}_t, t, c) - \boldsymbol{v}_{q_\theta, t}\big\|_2^2 + \mathbb{E}\big\|(\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0) - \boldsymbol{v}_{q_\theta, t}\big\|_2^2, \tag{18}$$

so $\boldsymbol{v}_{q_\theta,t}$ is the almost everywhere unique minimizer and the excess risk is exactly the squared velocity estimation error. Moreover, for endpoint laws χ and ζ noised by equation 14 and every $t \in (0,1)$,

$$\boldsymbol{s}_{\chi,t}(\boldsymbol{x}) = -\frac{\boldsymbol{x} + (1-t)\boldsymbol{v}_{\chi,t}(\boldsymbol{x})}{t}, \qquad \boldsymbol{s}_{\chi,t} - \boldsymbol{s}_{\zeta,t} = -\frac{1-t}{t}\big(\boldsymbol{v}_{\chi,t} - \boldsymbol{v}_{\zeta,t}\big). \tag{19}$$

Proof. Both claims follow from conditioning on the noised state. For equation 18, add and subtract $\boldsymbol{v}_{q_\theta,t}(\hat{\boldsymbol{x}}_t)$ inside the norm and expand the square. By equation 14 that field is the conditional mean of $\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0$ given $\hat{\boldsymbol{x}}_t$, so the residual $(\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0) - \boldsymbol{v}_{q_\theta,t}$ is orthogonal to every square-integrable function of $(\hat{\boldsymbol{x}}_t, t, c)$ and the cross term vanishes. Only the first term depends on $\boldsymbol{v}$, and it is zero exactly when $\boldsymbol{v} = \boldsymbol{v}_{q_\theta,t}$ up to null sets.

For equation 19, differentiate the Gaussian mixture density of $\boldsymbol{X}_t$ under the integral sign:

$$\boldsymbol{s}_{\chi,t}(\boldsymbol{x}) = \mathbb{E}\left[-\frac{\boldsymbol{x} - (1-t)\boldsymbol{X}}{t^2} \,\middle|\, \boldsymbol{X}_t = \boldsymbol{x}\right] = -\frac{1}{t}\,\mathbb{E}\big[\boldsymbol{\epsilon} \mid \boldsymbol{X}_t = \boldsymbol{x}\big], \tag{20}$$

where the second equality uses $\boldsymbol{x} - (1-t)\boldsymbol{X} = t\boldsymbol{\epsilon}$ on the event $\boldsymbol{X}_t = \boldsymbol{x}$. Since $\boldsymbol{\epsilon} = \boldsymbol{X}_t + (1-t)(\boldsymbol{\epsilon} - \boldsymbol{X})$, the same conditioning gives $\mathbb{E}[\boldsymbol{\epsilon} \mid \boldsymbol{X}_t = \boldsymbol{x}] = \boldsymbol{x} + (1-t)\boldsymbol{v}_{\chi,t}(\boldsymbol{x})$, which is the first identity in equation 19. Writing that identity for χ and for ζ and subtracting removes the common term $-\boldsymbol{x}/t$ and leaves the second. □

Equation 18 is what allows the student to act as its own velocity estimator: the loss that trains the estimator has a single population target, and the gap that remains when the target is not reached is exactly the quantity that Theorem 1 charges to the distribution gradient. Equation 19 is the dictionary between the two descriptions of a noised law, and the factor $(1-t)/t$ that it produces is the part of $\lambda(t)$ absorbed in the main text. Both are statements about functions, and whether the shared network realizes them remains an assumption.

Remark 1 (Why the self flow matching inputs are detached). At the optimal predictor, the value of equation 10 is the second term of equation 18, the conditional variance of the regression target given the noised state. Since $\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0 = (\hat{\boldsymbol{x}}_t - \hat{\boldsymbol{x}}_0)/t$, that value equals $t^{-2}\mathbb{E}\|\hat{\boldsymbol{x}}_0 - \mathbb{E}[\hat{\boldsymbol{x}}_0 \mid \hat{\boldsymbol{x}}_t]\|_2^2$, and for $t \in (0,1)$ the Gaussian likelihood is positive everywhere, so the conditional law of $\hat{\boldsymbol{x}}_0$ spreads over the whole support of $q_{\theta,0}$. The residual is therefore zero only when $q_{\theta,0}$ is a point mass. If gradients were allowed to flow through the generated sample, the generator would have a direct route to lower this loss by collapsing, so equation 10 detaches both the input state and the regression target.

The distribution matching update is a weighted reverse Kullback-Leibler gradient. The second step connects the detached update of equation 11 to the divergence between the noised laws in equation 3. Its time weight and the effective velocity weight satisfy

$$w_{\mathrm{D}}(t) = \frac{t\lambda(t)}{1-t}. \tag{21}$$

Let $\boldsymbol{e}_\theta = \boldsymbol{v}^\theta_{t,t} - \boldsymbol{v}_{q_\theta,t}$ and $\boldsymbol{e}_T = \boldsymbol{v}^T_{t,t} - \boldsymbol{v}_{p_T,t}$ be the student and teacher velocity errors at the sampled state, with second moments bounded by δ_θ^2 and δ_T^2.

Theorem 1 (The detached update is a weighted reverse Kullback-Leibler gradient)**.** *When $\boldsymbol{e}_\theta = \boldsymbol{e}_T = \boldsymbol{0}$, the automatic-differentiation gradient of equation 11 equals the gradient of equation 3, and*

$$\nabla_\theta \mathcal{J}_{\mathrm{D}}(\theta) = -\mathbb{E}\Big[\lambda(t)(1-t)\big(J_\theta G_\theta\big)^\top\big(\boldsymbol{v}_{q_\theta,t} - \boldsymbol{v}_{p_T,t}\big)\Big]. \tag{22}$$

In general the expected update $\widehat{\boldsymbol{g}}$ obtained from estimated velocities obeys

$$\big\|\widehat{\boldsymbol{g}} - \nabla_\theta \mathcal{J}_{\mathrm{D}}(\theta)\big\|_2 \le C_v(\delta_\theta + \delta_T), \qquad C_v^2 = \mathbb{E}\big[\lambda(t)^2(1-t)^2\|J_\theta G_\theta\|_{\mathrm{op}}^2\big]. \tag{23}$$

Proof. Fix (t, c) and let $\boldsymbol{b}_t(\boldsymbol{x}) = \mathbb{E}[J_\theta \hat{\boldsymbol{x}}_t \mid \hat{\boldsymbol{x}}_t = \boldsymbol{x}]$ be the velocity that a parameter perturbation induces on the sample. Then

$$\nabla_\theta D_{\mathrm{KL}}(q_{\theta,t} \| p_{T,t}) = \int \big(\nabla_\theta q_{\theta,t}\big) \log \frac{q_{\theta,t}}{p_{T,t}} \, \mathrm{d}\boldsymbol{x} \tag{24}$$

$$= -\int \nabla_{\boldsymbol{x}} \cdot \big(q_{\theta,t} \boldsymbol{b}_t\big) \log \frac{q_{\theta,t}}{p_{T,t}} \, \mathrm{d}\boldsymbol{x} \tag{25}$$

$$= \int q_{\theta,t} \, \boldsymbol{b}_t^\top \big(\boldsymbol{s}_{q_\theta,t} - \boldsymbol{s}_{p_T,t}\big) \, \mathrm{d}\boldsymbol{x} \tag{26}$$

$$= \mathbb{E}\Big[\big(J_\theta \hat{\boldsymbol{x}}_t\big)^\top \big(\boldsymbol{s}_{q_\theta,t} - \boldsymbol{s}_{p_T,t}\big)\Big] \tag{27}$$

$$= -\frac{1-t}{t} \, \mathbb{E}\Big[\big(J_\theta \hat{\boldsymbol{x}}_t\big)^\top \big(\boldsymbol{v}_{q_\theta,t} - \boldsymbol{v}_{p_T,t}\big)\Big]. \tag{28}$$

Weighting equation 28 by $w_{\mathrm{D}}(t)$ from equation 21, averaging over (t, c) and substituting $J_\theta \hat{\boldsymbol{x}}_t = (1-t) J_\theta G_\theta$ gives equation 22. Automatic differentiation of equation 11 contracts the same Jacobian with the detached vector $-\lambda(t)(\boldsymbol{v}_{t,t}^\theta - \boldsymbol{v}_{t,t}^T)$, so

$$\widehat{\boldsymbol{g}} = -\mathbb{E}\big[\lambda(t)(1-t)(J_\theta G_\theta)^\top (\boldsymbol{v}_{t,t}^\theta - \boldsymbol{v}_{t,t}^T)\big] = \nabla_\theta \mathcal{J}_{\mathrm{D}}(\theta) - \mathbb{E}\big[\lambda(t)(1-t)(J_\theta G_\theta)^\top (\boldsymbol{e}_\theta - \boldsymbol{e}_T)\big]. \tag{29}$$

Step equation 24 differentiates under the integral sign and drops the term $\int \nabla_\theta q_{\theta,t} \, \mathrm{d}\boldsymbol{x} = \nabla_\theta 1 = 0$ produced by the derivative of the logarithm. Step equation 25 is the continuity equation, which states that the law of $\hat{\boldsymbol{x}}_t$ moves under the parameter perturbation by the transport of $q_{\theta,t}$ along $\boldsymbol{b}_t$. Step equation 26 integrates by parts, uses $\nabla_{\boldsymbol{x}} \log(q_{\theta,t}/p_{T,t}) = \boldsymbol{s}_{q_\theta,t} - \boldsymbol{s}_{p_T,t}$ and discards the boundary term by Assumption 1(iv). Step equation 27 rewrites the integral as an expectation and replaces $\boldsymbol{b}_t$ by $J_\theta \hat{\boldsymbol{x}}_t$ under the tower property, which is valid because the other factor depends on the sample only through $\hat{\boldsymbol{x}}_t$. Step equation 28 applies the score difference in equation 19, and the weight relation equation 21 is exactly the factor that turns $(1-t)^2/t$ into the $\lambda(t)(1-t)$ of equation 22. The first claim is the case $\boldsymbol{e}_\theta = \boldsymbol{e}_T = \boldsymbol{0}$ of equation 29. For equation 23, take norms in the error term of equation 29, bound $\|(J_\theta G_\theta)^\top \boldsymbol{u}\|_2 \le \|J_\theta G_\theta\|_{\mathrm{op}} \|\boldsymbol{u}\|_2$, apply Cauchy-Schwarz to the expectation, and use $(\mathbb{E}\|\boldsymbol{e}_\theta - \boldsymbol{e}_T\|_2^2)^{1/2} \le \delta_\theta + \delta_T$. □

Two properties of equation 23 matter in practice. The bias is linear in the velocity errors, so the shared estimator trained by equation 10 controls it directly through equation 18, and no separate network is needed to make the statement meaningful. The constant C_v carries $\lambda(t)$, which contains the score-to-velocity factor $(1-t)/t$ of equation 19 and grows without bound as $t \to 0$, which is why ρ is truncated away from the data endpoint in Assumption 1(i). Appendix D turns this bias into an optimization statement.

Trajectory alignment at the start time rescales the update. The grid $\mathcal{S}_t$ contains t, so the start-time output $\boldsymbol{v}_{t,t}^\theta$ receives the teacher target $\boldsymbol{v}_{t,t}^T$ from equation 9 in addition to the regression target of equation 10, and a nonzero trajectory weight makes $\boldsymbol{e}_\theta$ in Theorem 1 nonzero even at the population optimum. The next statement shows that this error has a special form. Fix (t, c) and write $\boldsymbol{u}(\boldsymbol{x})$ for the start-time output evaluated at the detached state $\boldsymbol{x} = \mathrm{sg}[\hat{\boldsymbol{x}}_t]$, so that $\boldsymbol{u}(\hat{\boldsymbol{x}}_t) = \boldsymbol{v}_{t,t}^\theta$. Let $\boldsymbol{v}_T(\cdot, t, c)$ be the teacher field that supplies $\boldsymbol{v}_{t,t}^T$ in both equation 9 and equation 11, and let $b = \lambda_{\mathrm{tr}}/(M+1)$ be the weight that equation 13 places on the start-time term of equation 9, where M may depend on t.

Proposition 4 (The start-time compromise preserves the distribution direction)**.** *Assume $\lambda_{\mathrm{fm}} > 0$ and that $\boldsymbol{u}$ ranges over square-integrable functions of $\hat{\boldsymbol{x}}_t$ independently of the other grid velocities and of the generator. The start-time terms of $\lambda_{\mathrm{fm}}\mathcal{L}_{\mathrm{fm}} + \lambda_{\mathrm{tr}}\mathcal{L}_{\mathrm{tr}}$ are minimized, uniquely up to null sets, by*

$$\boldsymbol{u}^\star = \frac{\lambda_{\mathrm{fm}} \boldsymbol{v}_{q_\theta,t} + b \, \boldsymbol{v}_T}{\lambda_{\mathrm{fm}} + b}, \qquad \boldsymbol{u}^\star - \boldsymbol{v}_T = \alpha\big(\boldsymbol{v}_{q_\theta,t} - \boldsymbol{v}_T\big), \qquad \alpha = \frac{\lambda_{\mathrm{fm}}}{\lambda_{\mathrm{fm}} + b} \in \Big[\frac{\lambda_{\mathrm{fm}}}{\lambda_{\mathrm{fm}} + \lambda_{\mathrm{tr}}}, 1\Big]. \tag{30}$$

If in addition $\boldsymbol{v}_T = \boldsymbol{v}_{p_T,t}$ and the student reports $\boldsymbol{v}_{t,t}^\theta = \boldsymbol{u}^\star(\hat{\boldsymbol{x}}_t)$, the automatic-differentiation gradient of equation 11 equals the gradient of equation 3 with $w_{\mathrm{D}}(t)$ replaced by $\alpha \, w_{\mathrm{D}}(t)$.

Proof. Write $a = \lambda_{\mathrm{fm}}$ and abbreviate the per-sample start-time objective and its population value as

$$\mathbb{E}\Big[a\big\|\boldsymbol{u} - (\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0)\big\|_2^2 + b\big\|\boldsymbol{u} - \boldsymbol{v}_T\big\|_2^2\Big] \tag{31}$$

$$= \mathbb{E}\Big[a\big\|\boldsymbol{u} - \boldsymbol{v}_{q_\theta,t}\big\|_2^2 + b\big\|\boldsymbol{u} - \boldsymbol{v}_T\big\|_2^2\Big] + a\,\mathbb{E}\big\|(\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0) - \boldsymbol{v}_{q_\theta,t}\big\|_2^2 \tag{32}$$

$$= (a+b)\,\mathbb{E}\big\|\boldsymbol{u} - \boldsymbol{u}^\star\big\|_2^2 + \frac{ab}{a+b}\,\mathbb{E}\big\|\boldsymbol{v}_{q_\theta,t} - \boldsymbol{v}_T\big\|_2^2 + a\,\mathbb{E}\big\|(\boldsymbol{\epsilon} - \hat{\boldsymbol{x}}_0) - \boldsymbol{v}_{q_\theta,t}\big\|_2^2. \tag{33}$$

Step equation 31 collects the start-time term of equation 10 and the $s = t$ summand of equation 9 with their weights from equation 13. Step equation 32 applies the decomposition equation 18 to the first term; the second term is unchanged because $\boldsymbol{v}_T$ is a function of $\hat{\boldsymbol{x}}_t$. Step equation 33 completes the square pointwise through $a\|\boldsymbol{u}-\boldsymbol{p}\|_2^2+b\|\boldsymbol{u}-\boldsymbol{q}\|_2^2 = (a+b)\|\boldsymbol{u}-(a\boldsymbol{p}+b\boldsymbol{q})/(a+b)\|_2^2+ab\|\boldsymbol{p}-\boldsymbol{q}\|_2^2/(a+b)$ with $\boldsymbol{p} = \boldsymbol{v}_{q_\theta,t}$ and $\boldsymbol{q} = \boldsymbol{v}_T$. Only the first term of equation 33 depends on $\boldsymbol{u}$, which gives the unique minimizer and, after subtracting $\boldsymbol{v}_T$, the second identity in equation 30. The range of α follows from $0 < b \le \lambda_{\mathrm{tr}}$. For the gradient claim, substituting $\boldsymbol{u}^\star - \boldsymbol{v}_{p_T,t} = \alpha(\boldsymbol{v}_{q_\theta,t} - \boldsymbol{v}_{p_T,t})$ into the detached vector of equation 11 turns it into the exact vector of Theorem 1 with $\lambda(t)$ replaced by $\alpha\lambda(t)$, and equation 21 carries the factor α to $w_{\mathrm{D}}(t)$. □

Proposition 4 controls the part of $\boldsymbol{e}_\theta$ that trajectory alignment introduces: at the joint optimum this part equals $-(1-\alpha)(\boldsymbol{v}_{q_\theta,t}-\boldsymbol{v}_{p_T,t})$, which is parallel to the exact distribution direction, so it rescales the time weight of equation 3 without changing the zeros used in Theorem 2 or the ideal equilibrium of Proposition 1. The middle term of equation 33 also shows that the two start-time residuals cannot vanish together before $\boldsymbol{v}_{q_\theta,t} = \boldsymbol{v}_T$. The statement relies on the start-time output being free. With shared parameters, and in particular when $K < M + 1$ couples the start-time value to the other grid velocities through equation 17, the realized error also contains a component that is not parallel to the distribution direction, and Theorem 1 bounds its effect.

A zero of the divergence forces endpoint agreement. Theorem 1 concerns noised laws at times inside $(0, 1)$, while the quantity of interest is the generated endpoint law. The next statement closes that gap at a global optimum.

Theorem 2 (Vanishing divergence forces endpoint agreement)**.** *Assume $p_{T,t} = \mathcal{N}_t p_{T,0}$, some $\theta^\dagger$ satisfies $q_{\theta^\dagger,0} = p_{T,0}$ for almost every c, and $\rho(\{t < 1 : \lambda(t) > 0\}) > 0$. Then $\inf_\theta \mathcal{J}_{\mathrm{D}}(\theta) = 0$, and every global minimizer θ^{opt} satisfies $q_{\theta^{\mathrm{opt}},0} = p_{T,0}$ for almost every c.*

Proof. The realizing parameter $\theta^\dagger$ makes every noisy divergence vanish, so the infimum is zero. Since the integrand of equation 3 is non-negative, a global minimizer satisfies $q_{\theta^{\mathrm{opt}},t} = p_{T,t}$ for ρ-almost every $t < 1$ with $\lambda(t) > 0$, and by Fubini's theorem such a t exists for almost every c. Fix one. Both laws are the same Gaussian convolution of their endpoint laws, so their characteristic functions factor as

$$\varphi_{\mathcal{N}_t\chi}(\boldsymbol{\xi}) = \varphi_\chi\big((1-t)\boldsymbol{\xi}\big)\exp\big(-\tfrac{1}{2}t^2\|\boldsymbol{\xi}\|_2^2\big)\,. \tag{34}$$

The Gaussian factor never vanishes and $1-t > 0$, so the two endpoint characteristic functions agree on all of $\mathbb{R}^d$, and the uniqueness theorem for characteristic functions gives the claim. □

Two hypotheses carry the conclusion. Both noisy families must come from noising one endpoint law, which is the hypothesis that classifier-free guidance breaks and that Appendix D examines, and the positively weighted times must include some $t < 1$, since the noised law at $t = 1$ is pure noise and retains no endpoint information. The conclusion also concerns the exact global minimum only. Noising contracts distances between endpoint laws, so a small positive value of $\mathcal{J}_{\mathrm{D}}$ does not by itself bound the endpoint discrepancy, and the statement that survives at finite error is Corollary 1 rather than a perturbed version of this theorem.

C.3 Finite-error behaviour and the complementarity of the two blocks

From grid velocity error to endpoint error. Theorem 1 already shows what inexact velocities cost the distribution gradient. The corresponding question for trajectory alignment is what the grid loss of equation 9, which compares velocities at $M + 1$ times only, implies about the distance between the student and teacher paths at all times.

Fix a realization of the training start $(\hat{\boldsymbol{x}}_t, t, c)$ and hold the predicted policy $\mathcal{P}_\theta(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)$ fixed over the interval. Let $\boldsymbol{x}_\theta(\tau)$ be the analytic student rollout equation 16 from that start and let $\boldsymbol{x}_T(\tau)$ be the exact teacher solution from the same start, so that both satisfy $\boldsymbol{x}_\theta(t) = \boldsymbol{x}_T(t) = \hat{\boldsymbol{x}}_t$. Write

$$\boldsymbol{v}_{T,\mathrm{path}}(\tau) = \boldsymbol{v}_T(\boldsymbol{x}_T(\tau), \tau, c), \qquad \boldsymbol{e}(\tau) = \tilde{\boldsymbol{v}}_\theta(\tau; \mathcal{P}_\theta(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)) - \boldsymbol{v}_{T,\mathrm{path}}(\tau)$$

for the teacher velocity along its own path and for the velocity discrepancy at time τ, and note that $\boldsymbol{e}(\tau_j) = \boldsymbol{v}^\theta_{t,\tau_j} - \boldsymbol{v}_{T,\mathrm{path}}(\tau_j)$ differs from the residual seen by training only through the accuracy of the teacher targets. For one sample, let $\ell_{\mathrm{tr}} = (M+1)^{-1} \sum_{j=0}^{M} \|\boldsymbol{v}^\theta_{t,\tau_j} - \boldsymbol{v}^T_{t,\tau_j}\|_2^2$ be the bracketed quantity of equation 9, let $\{A_j\}_{j=0}^M$ be a partition of $[0,t]$ with $\tau_j \in A_j$, and set $\Delta_{\max} = \max_j |A_j|$ and $\Delta_{\mathrm{mesh}} = \max_j \sup_{\tau \in A_j} |\tau - \tau_j|$.

Theorem 3 (Endpoint error of the single-forward rollout)**.** *If $\boldsymbol{e} \in L^2([0,t]; \mathbb{R}^d)$, then*

$$\boldsymbol{x}_\theta(\tau) - \boldsymbol{x}_T(\tau) = -\int_\tau^t \boldsymbol{e}(u)\,\mathrm{d}u, \qquad \big\|\boldsymbol{x}_\theta(0) - \boldsymbol{x}_T(0)\big\|_2 \le \sqrt{t}\,\big\|\boldsymbol{e}\big\|_{L^2([0,t])}. \tag{35}$$

If in addition $\boldsymbol{e}$ is L_e-Lipschitz in time and the teacher targets satisfy $\max_j \|\boldsymbol{v}^T_{t,\tau_j} - \boldsymbol{v}_{T,\mathrm{path}}(\tau_j)\|_2 \le \eta_{\max}$, then

$$\big\|\boldsymbol{x}_\theta(0) - \boldsymbol{x}_T(0)\big\|_2 \le \sqrt{t\,\Delta_{\max}(M+1)\ell_{\mathrm{tr}}} + t\big(\eta_{\max} + L_e\Delta_{\mathrm{mesh}}\big), \tag{36}$$

whose right-hand side becomes $t(\sqrt{\ell_{\mathrm{tr}}} + \eta_{\max} + L_e\Delta_{\mathrm{mesh}})$ when the cells have equal length. If the teacher targets come from an order p_{sol} solver with maximum step h_{sol} and global error constant C_{sol}, and the teacher field is L_x-Lipschitz in the state, then $\eta_{\max} \le L_x C_{\mathrm{sol}} h_{\mathrm{sol}}^{p_{\mathrm{sol}}}$.

Proof. The two rollouts share their state at time t, so subtracting their integral forms gives the identity in equation 35, and Cauchy-Schwarz applied to $\int_0^t \|\boldsymbol{e}(u)\|_2\,\mathrm{d}u$ gives the inequality. No Lipschitz condition on the teacher field and no Grönwall factor appear here, because $\boldsymbol{e}$ compares the student policy with the teacher velocity along the teacher path rather than at the student state.

For the finite grid, control $\boldsymbol{e}$ on each cell by its value at the grid point it contains:

$$\sup_{\tau \in A_j} \|\boldsymbol{e}(\tau)\|_2 \le \|\boldsymbol{v}^\theta_{t,\tau_j} - \boldsymbol{v}^T_{t,\tau_j}\|_2 + \eta_{\max} + L_e\Delta_{\mathrm{mesh}}, \tag{37}$$

$$\|\boldsymbol{e}\|_{L^2([0,t])} \le \Big(\sum_{j=0}^{M} |A_j|\,\|\boldsymbol{v}^\theta_{t,\tau_j} - \boldsymbol{v}^T_{t,\tau_j}\|_2^2\Big)^{1/2} + \sqrt{t}\,\big(\eta_{\max} + L_e\Delta_{\mathrm{mesh}}\big) \tag{38}$$

$$\le \sqrt{\Delta_{\max}(M+1)\ell_{\mathrm{tr}}} + \sqrt{t}\,\big(\eta_{\max} + L_e\Delta_{\mathrm{mesh}}\big). \tag{39}$$

Step equation 37 compares $\boldsymbol{e}(\tau)$ with $\boldsymbol{e}(\tau_j)$ through the Lipschitz property over a distance of at most Δ_{mesh}, then inserts $\boldsymbol{v}^T_{t,\tau_j}$ into $\boldsymbol{e}(\tau_j) = \boldsymbol{v}^\theta_{t,\tau_j} - \boldsymbol{v}_{T,\mathrm{path}}(\tau_j)$ and applies the target error bound $\eta_{\max}$. Step equation 38 integrates the square of equation 37 over each cell and applies Minkowski's inequality in the sequence space weighted by the cell lengths, where the constant contributes $(\sum_j |A_j|)^{1/2} = \sqrt{t}$. Step equation 39 replaces every $|A_j|$ by $\Delta_{\max}$ and uses $\sum_j \|\boldsymbol{v}^\theta_{t,\tau_j} - \boldsymbol{v}^T_{t,\tau_j}\|_2^2 = (M+1)\ell_{\mathrm{tr}}$. Combining equation 39 with equation 35 gives equation 36, and equal cells have $\Delta_{\max} = t/(M+1)$. For the last claim, the solver state error is at most $C_{\mathrm{sol}}h_{\mathrm{sol}}^{p_{\mathrm{sol}}}$ at every grid point, and evaluating the L_x-Lipschitz teacher field at those states turns it into a velocity error of at most $L_x C_{\mathrm{sol}} h_{\mathrm{sol}}^{p_{\mathrm{sol}}}$. □

Equation 36 separates three error sources that training treats very differently. The first term is the grid residual, which is what equation 9 minimizes. The second is the accuracy of the teacher targets, which is set by the solver and its step size rather than by the student. The third is the variation of the discrepancy between grid points, which no discrete loss can observe: without a regularity condition in time, a bump placed between two grid points leaves ℓ_{tr} unchanged and moves the endpoint arbitrarily far. The policy of equation 6 is a sum of K smooth exponentials, which is the structural reason a modest grid is informative about the whole interval, and Proposition 2 is the complementary statement that such a field can still match all grid points.

The configurations in Table 4 cap the grid span, so a single grid may end at $s_M > 0$. The identity in equation 35 holds at every $\tau \le t$, and repeating the argument on a partition of $[s_M, t]$ gives the

bounds of Theorem 3 for the state at s_M with t replaced by $t - s_M$. On $[0, s_M]$ the grid of that sample provides no supervision, and the full-interval bound for that sample relies on L_e over the long final cell. Lower times are supervised in other iterations, for policies started from states with smaller $t \sim \rho$.

Joint control of the law produced by the student policy. The two supervision blocks act on different objects: trajectory alignment compares maps from a common start, while distribution matching compares laws at a common time. The next statement puts both errors in one bound. Let Ψ_θ^t be the analytic policy map from time t to time 0, and let Φ_T^t be the exact teacher flow map over the same interval.

Corollary 1 (Joint bound on the law produced by the student policy)**.** *Assume* $(\Phi_T^t)_\# p_{T,t} = p_{T,0}$, $\mathrm{Lip}(\Phi_T^t) \le L_\Phi$*, and finite second moments. Define the mean squared trajectory map error from the current student law,*

$$\varepsilon_{\mathrm{tr}}^2(t) = \mathbb{E}_{\hat{\boldsymbol{x}}_t \sim q_{\theta,t}} \big\| \Psi_\theta^t(\hat{\boldsymbol{x}}_t) - \Phi_T^t(\hat{\boldsymbol{x}}_t) \big\|_2^2. \tag{40}$$

Then

$$W_2\big((\Psi_\theta^t)_\# q_{\theta,t}, p_{T,0}\big) \le \varepsilon_{\mathrm{tr}}(t) + L_\Phi W_2(q_{\theta,t}, p_{T,t}), \tag{41}$$

and if $p_{T,t}$ *satisfies a Talagrand* $T_2(C_T)$ *inequality (Otto & Villani, 2000), then*

$$W_2\big((\Psi_\theta^t)_\# q_{\theta,t}, p_{T,0}\big) \le \varepsilon_{\mathrm{tr}}(t) + L_\Phi \sqrt{2 C_T D_{\mathrm{KL}}(q_{\theta,t} \| p_{T,t})}. \tag{42}$$

Proof. Both terms come from the triangle inequality for W_2 applied at the intermediate law $(\Phi_T^t)_\# q_{\theta,t}$. For the first term, couple $(\Psi_\theta^t)_\# q_{\theta,t}$ and $(\Phi_T^t)_\# q_{\theta,t}$ by sending the same $\hat{\boldsymbol{x}}_t \sim q_{\theta,t}$ through both maps; the cost of that particular coupling is $\varepsilon_{\mathrm{tr}}^2(t)$, and W_2 is the infimum over all couplings. For the second term, push an optimal coupling of $q_{\theta,t}$ and $p_{T,t}$ through Φ_T^t; the map is L_Φ-Lipschitz, so the cost grows by at most L_Φ, and the assumption $(\Phi_T^t)_\# p_{T,t} = p_{T,0}$ identifies the second marginal. Equation 42 then substitutes the Talagrand inequality $W_2^2(q_{\theta,t}, p_{T,t}) \le 2 C_T D_{\mathrm{KL}}(q_{\theta,t} \| p_{T,t})$. □

The two summands of equation 41 are exactly the objects that the two training blocks control, and neither controls the other. Trajectory alignment bounds the first through equation 36, which applies to the integrand of equation 40 sample by sample, since Ψ_θ^t is the analytic rollout and Φ_T^t the exact teacher solution from the same start. Distribution matching bounds the second through the divergence of equation 3, in the form equation 42. A student whose noised law already matches the teacher law still needs an accurate map to turn it into samples, and an accurate map applied to a mismatched law reproduces the mismatch, which is the technical content of combining the two signals. The bound also states its own limit: the controlled object is the law obtained by pushing $q_{\theta,t}$ through the student policy, which need not equal the generator endpoint law $q_{\theta,0}$, so this is a complementarity statement rather than endpoint recovery at finite error.

What distribution preservation controls. The remaining term compares the current generator with its moving-average reference at a fixed $\bar{\theta}$. Since it is a squared distance between two maps evaluated at the same latent input, it is a proximal penalty in the space of generator maps, and it dominates the corresponding distances between the generated laws.

Proposition 5 (What distribution preservation controls)**.** *Let* $\mathcal{R}_{\mathrm{dp}}(\theta; \bar{\theta}) = \mathbb{E}_{\boldsymbol{x}_1, c} \| G_\theta(\boldsymbol{x}_1, c) - G_{\bar{\theta}}(\boldsymbol{x}_1, c) \|_2^2$*, which equals equation 12 when* $\bar{\theta}$ *is held fixed. Then*

$$\mathbb{E}_c\big[W_2^2(q_{\theta,0}(\cdot \mid c), q_{\bar{\theta},0}(\cdot \mid c))\big] \le \mathcal{R}_{\mathrm{dp}}, \qquad \mathbb{E}_c\big[W_2^2(q_{\theta,t}(\cdot \mid c), q_{\bar{\theta},t}(\cdot \mid c))\big] \le (1-t)^2 \mathcal{R}_{\mathrm{dp}}, \tag{43}$$

and for every $t \in (0, 1)$*,*

$$\mathbb{E}_c\Big[\max\big\{ D_{\mathrm{KL}}(q_{\theta,t}(\cdot \mid c) \| q_{\bar{\theta},t}(\cdot \mid c)),\ D_{\mathrm{KL}}(q_{\bar{\theta},t}(\cdot \mid c) \| q_{\theta,t}(\cdot \mid c)) \big\} \Big] \le \frac{(1-t)^2}{2t^2} \mathcal{R}_{\mathrm{dp}}. \tag{44}$$

Proof. Fix c. Sending the same latent $\boldsymbol{x}_1$ through both generators is one coupling of $q_{\theta,0}(\cdot \mid c)$ and $q_{\bar{\theta},0}(\cdot \mid c)$, and its cost is the summand of $\mathcal{R}_{\mathrm{dp}}$, which gives the first bound in equation 43. Sharing the Gaussian noise of equation 2 as well makes the two noised samples differ by $(1-t)(G_\theta - G_{\bar{\theta}})$, which gives the second bound.

For equation 44, keep the shared latent in the description and compare the joint laws of $(\boldsymbol{x}_1, \hat{\boldsymbol{x}}_t)$ under θ and under $\bar{\theta}$. The latent has the same law in both, so the KL between the joint laws is the average of the conditional KL given $\boldsymbol{x}_1$. Conditioned on the latent, the two noised outputs are Gaussians with the same covariance $t^2\boldsymbol{I}$ and means differing by $(1-t)(G_\theta - G_{\bar{\theta}})$, so their KL in either direction is the squared mean difference divided by $2t^2$. Discarding the latent cannot increase the KL, which transfers the bound to the marginal laws of $\hat{\boldsymbol{x}}_t$, and averaging over c finishes the proof. □

Proposition 5 is written for the one-step coupling in equation 12. For Algorithm 2, the same argument applies conditionally on the shared final-stage input $(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$ after replacing the student and moving-average generators by $G_\theta(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$ and $G_{\bar{\theta}}(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$, respectively.

Both bounds can be strict, because $\mathcal{R}_{\mathrm{dp}}$ fixes one latent coupling: two generators that induce the same law but permute which latent produces which sample have zero distance between their laws and a positive preservation loss. This is intended, since the term is meant to limit how far one update moves the generator map, and not to measure a distance between distributions. The divergence bound in equation 44 also degrades as $t \to 0$, so it is informative only at the noise levels that ρ actually visits.

D GUIDANCE AND OPTIMIZATION QUALIFICATIONS

Appendix C assumes an unguided teacher and, in Proposition 1, exact optimization of an expressive enough student. This appendix removes those assumptions one at a time. Classifier-free guidance changes the object that the distribution term matches, the residual velocity error of Theorem 1 sets a floor on the stationarity that term can reach, and the four updates together with the moving average form a coupled dynamical system whose equilibria have to be examined separately.

Classifier-free guidance. Classifier-free guidance (Ho & Salimans, 2022) replaces the teacher field in equation 8 and equation 11 by

$$\boldsymbol{v}_{T,\mathrm{cfg}}(\boldsymbol{x}, t, c) = \boldsymbol{v}_T(\boldsymbol{x}, t, \varnothing) + s_{\mathrm{cfg}}\big[\boldsymbol{v}_T(\boldsymbol{x}, t, c) - \boldsymbol{v}_T(\boldsymbol{x}, t, \varnothing)\big]. \tag{45}$$

Let $p_{T,\mathrm{c},t}(\boldsymbol{x} \mid c)$ and $p_{T,\mathrm{u},t}(\boldsymbol{x})$ be the conditional and unconditional teacher marginals, with scores $\boldsymbol{s}_{T,\mathrm{c},t}$ and $\boldsymbol{s}_{T,\mathrm{u},t}$. If both branches predict exact marginal velocities under the same noising kernel, equation 19 gives

$$\boldsymbol{s}_{T,\mathrm{cfg},t}(\boldsymbol{x}, c) = (1 - s_{\mathrm{cfg}})\boldsymbol{s}_{T,\mathrm{u},t}(\boldsymbol{x}) + s_{\mathrm{cfg}}\boldsymbol{s}_{T,\mathrm{c},t}(\boldsymbol{x} \mid c) = -\frac{\boldsymbol{x} + (1-t)\boldsymbol{v}_{T,\mathrm{cfg}}(\boldsymbol{x}, t, c)}{t}. \tag{46}$$

When normalizable, equation 46 is the score of the geometric combination $p_{T,\mathrm{c},t}^{\,s_{\mathrm{cfg}}} p_{T,\mathrm{u},t}^{\,1-s_{\mathrm{cfg}}}$, and Theorem 1 therefore applies separately at each time, with $p_{T,t}$ read as that combination. For guidance scales outside $[0, 1]$, normalizability additionally requires suitable tail control.

What guidance breaks is the hypothesis of Theorem 2. These time-indexed geometric combinations are in general not the noised versions of one endpoint law, so guided TRD matches a family of per-time score fields that need not describe a single target distribution. The endpoint conclusion applies directly to the conditional branch at $s_{\mathrm{cfg}} = 1$ and to the unconditional branch at $s_{\mathrm{cfg}} = 0$, and other scales require cross-time consistency as an additional assumption. The trajectory results of Appendix C.3 are unaffected, because they treat the guided field as the field being integrated. Guidance evaluates both teacher branches, which gives $2(M+1)$ teacher evaluations per iteration.

Stationarity floor from velocity estimation error. Theorem 1 bounds the bias of the distribution update by $\beta_{\mathrm{vel}} = C_v(\delta_\theta + \delta_T)$. The standard descent argument for biased stochastic gradients turns that bias into a floor on the stationarity that this term can reach.

Proposition 6 (Stationarity floor under a biased update)**.** *Assume $\mathcal{J}_{\mathrm{D}}$ is bounded below by $\mathcal{J}_{\mathrm{inf}}$ and has an L_{D}-Lipschitz gradient, the stochastic update $\widehat{\boldsymbol{g}}_k$ has conditional mean $\nabla\mathcal{J}_{\mathrm{D}}(\theta_k) + \boldsymbol{\delta}_k$ with $\|\boldsymbol{\delta}_k\|_2 \le \beta_{\mathrm{vel}}$ and conditional variance at most σ_g^2, and the step size satisfies $0 < \eta_{\mathrm{opt}} \le 1/(2L_{\mathrm{D}})$. Then the iterates $\theta_{k+1} = \theta_k - \eta_{\mathrm{opt}}\widehat{\boldsymbol{g}}_k$ obey*

$$\frac{1}{N}\sum_{k=0}^{N-1}\mathbb{E}\big\|\nabla\mathcal{J}_{\mathrm{D}}(\theta_k)\big\|_2^2 \le \frac{4\big(\mathcal{J}_{\mathrm{D}}(\theta_0) - \mathcal{J}_{\mathrm{inf}}\big)}{\eta_{\mathrm{opt}}N} + 6\beta_{\mathrm{vel}}^2 + 2L_{\mathrm{D}}\eta_{\mathrm{opt}}\sigma_g^2. \tag{47}$$

Proof. Write $\boldsymbol{g}_k = \nabla \mathcal{J}_{\mathrm{D}}(\theta_k)$. The descent lemma for an L_{D}-smooth function gives $\mathbb{E}_k[\mathcal{J}_{\mathrm{D}}(\theta_{k+1})] \le \mathcal{J}_{\mathrm{D}}(\theta_k) - \eta_{\mathrm{opt}} \boldsymbol{g}_k^\top (\boldsymbol{g}_k + \boldsymbol{\delta}_k) + \frac{1}{2} L_{\mathrm{D}} \eta_{\mathrm{opt}}^2 \mathbb{E}_k \|\widehat{\boldsymbol{g}}_k\|_2^2$. Bound the cross term by $-\boldsymbol{g}_k^\top \boldsymbol{\delta}_k \le \frac{1}{4}\|\boldsymbol{g}_k\|_2^2 + \|\boldsymbol{\delta}_k\|_2^2$ and the second moment by $\mathbb{E}_k\|\widehat{\boldsymbol{g}}_k\|_2^2 \le 2\|\boldsymbol{g}_k\|_2^2 + 2\|\boldsymbol{\delta}_k\|_2^2 + \sigma_g^2$. With $L_{\mathrm{D}}\eta_{\mathrm{opt}} \le 1/2$, the coefficient of $\|\boldsymbol{g}_k\|_2^2$ is at most $-\eta_{\mathrm{opt}}/4$ and the coefficient of $\|\boldsymbol{\delta}_k\|_2^2$ at most $3\eta_{\mathrm{opt}}/2$, so $\frac{\eta_{\mathrm{opt}}}{4}\mathbb{E}\|\boldsymbol{g}_k\|_2^2 \le \mathbb{E}[\mathcal{J}_{\mathrm{D}}(\theta_k) - \mathcal{J}_{\mathrm{D}}(\theta_{k+1})] + \frac{3}{2}\eta_{\mathrm{opt}}\beta_{\mathrm{vel}}^2 + \frac{1}{2}L_{\mathrm{D}}\eta_{\mathrm{opt}}^2\sigma_g^2$. Summing over k, telescoping and using the lower bound on $\mathcal{J}_{\mathrm{D}}$ gives equation 47. □

Only the first term of equation 47 decreases with N, so the bound retains the non-vanishing term $6\beta_{\mathrm{vel}}^2$ under persistent velocity error and does not establish convergence of the average gradient to zero. The constant C_v in equation 23 weights the velocity error by the time weight and the generator Jacobian, so the time range and weighting determine how strongly estimation error affects the update. The statement concerns the distribution term in isolation and does not imply convergence of the full TRD dynamics.

Equilibria and local stability of the full dynamics. Let $\boldsymbol{h}(\theta)$ denote the expected weighted update from self flow matching, distribution matching and trajectory alignment under the stop-gradient rules of Algorithm 1, and let $\boldsymbol{d}(\theta, \bar{\theta})$ denote the expected preservation update, with the weights of equation 13 absorbed into both. These are vector fields rather than gradients of one fixed scalar objective, because the states at which they are evaluated are produced by the current student and then detached. The associated mean dynamics are

$$\dot{\theta} = -\boldsymbol{h}(\theta) - \boldsymbol{d}(\theta, \bar{\theta}), \qquad \dot{\bar{\theta}} = r_{\mathrm{ema}}\big(\theta - \bar{\theta}\big), \qquad r_{\mathrm{ema}} > 0, \tag{48}$$

where r_{ema} is the positive continuous-time rate corresponding to the discrete decay μ; a unit-step forward-Euler correspondence gives $r_{\mathrm{ema}} = 1 - \mu$.

Proposition 7 (Equilibria and limits of the complete mean field)**.** *Assume $\boldsymbol{h}$ and $\boldsymbol{d}$ are continuous and $\boldsymbol{d}(\theta, \theta) = \boldsymbol{0}$. Every convergent solution of equation 48 has a limit with $\bar{\theta}_\infty = \theta_\infty$ and $\boldsymbol{h}(\theta_\infty) = \boldsymbol{0}$. In the ideal unguided setting, any $\theta^\star$ that simultaneously attains $q_{\theta^\star,0} = p_{T,0}$, its own marginal velocity and zero trajectory residual gives an equilibrium $(\theta^\star, \theta^\star)$. A general zero of $\boldsymbol{h}$ need not have these three properties, because the weighted updates can cancel each other.*

Proof. At the ideal target, Proposition 3 makes the expected self flow matching update zero, Theorem 1 makes the distribution update zero because the two velocity fields coincide, and the trajectory residual is zero by assumption; preservation vanishes because $\bar{\theta} = \theta$. This proves the equilibrium claim. For a convergent solution of a continuous autonomous system, the vector field vanishes at the limit. The second equation of equation 48 gives $\bar{\theta}_\infty = \theta_\infty$, and substituting this into the first equation together with $\boldsymbol{d}(\theta, \theta) = \boldsymbol{0}$ gives $\boldsymbol{h}(\theta_\infty) = \boldsymbol{0}$. □

Theorem 4 (Local stability of the full dynamics at the ideal target)**.** *Let $\theta^\star$ be an ideal equilibrium as in Proposition 7. Assume $\boldsymbol{h}$ and $\boldsymbol{d}$ are continuously differentiable nearby, that the symmetric part of $\boldsymbol{A} = \nabla \boldsymbol{h}(\theta^\star)$ is bounded below by $m\boldsymbol{I}$ for some $m > 0$, and that the preservation field linearizes as $\boldsymbol{B}(\theta - \bar{\theta})$ with $\boldsymbol{B} \succeq 0$. Then $(\theta^\star, \theta^\star)$ is locally asymptotically stable, so every mean trajectory started close enough to it converges to it, and along that limit the student endpoint law equals $p_{T,0}$, its velocity estimate is exact and its trajectory residual vanishes.*

Proof. Let $\delta\theta = \theta - \theta^\star$ and $\delta\bar{\theta} = \bar{\theta} - \theta^\star$. Linearizing equation 48 at the equilibrium gives the system matrix

$$\boldsymbol{L} = \begin{bmatrix} -(\boldsymbol{A} + \boldsymbol{B}) & \boldsymbol{B} \\ r_{\mathrm{ema}}\boldsymbol{I} & -r_{\mathrm{ema}}\boldsymbol{I} \end{bmatrix}, \tag{49}$$

so it suffices to show that every eigenvalue of $\boldsymbol{L}$ has negative real part. Suppose instead that $\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{y})^\top = z(\boldsymbol{x}, \boldsymbol{y})^\top$ for some eigenvector and some z with $\mathrm{Re}(z) \ge 0$. Then $z \ne -r_{\mathrm{ema}}$, so the second block row gives $\boldsymbol{y} = r_{\mathrm{ema}}\boldsymbol{x}/(z + r_{\mathrm{ema}})$, and $\boldsymbol{x} \ne \boldsymbol{0}$, since $\boldsymbol{x} = \boldsymbol{0}$ would force $\boldsymbol{y} = \boldsymbol{0}$. Substituting $\boldsymbol{y}$ into the first block row gives

$$\left[z\boldsymbol{I} + \boldsymbol{A} + \frac{z}{z + r_{\mathrm{ema}}}\boldsymbol{B} \right] \boldsymbol{x} = \boldsymbol{0}.$$

Take the inner product with $\boldsymbol{x}$ and keep the real part. The first term contributes $\mathrm{Re}(z)\|\boldsymbol{x}\|_2^2 \ge 0$. The second contributes at least $m\|\boldsymbol{x}\|_2^2 > 0$, because only the symmetric part of $\boldsymbol{A}$ survives in

the real part. The third contributes $\mathrm{Re}\left(z/(z+r_{\mathrm{ema}})\right)\boldsymbol{x}^*\boldsymbol{B}\boldsymbol{x} \geq 0$, because $\boldsymbol{B} \succeq 0$ and because $\mathrm{Re}(z) \geq 0$ with $r_{\mathrm{ema}} > 0$ makes that scalar non-negative. The real part is therefore strictly positive, which contradicts the displayed equation. Hence $\boldsymbol{L}$ is Hurwitz and the Lyapunov indirect method gives local asymptotic stability. For equation 12, the preservation field linearizes to $\boldsymbol{B} = 2\lambda_{\mathrm{dp}}\mathbb{E}[(J_\theta G_\theta)^\top J_\theta G_\theta]$ at $\theta = \bar{\theta}$, which is positive semidefinite as required. □

The conclusion is local, and the stability margin has to come from $\boldsymbol{h}$: the preservation term contributes a positive semidefinite $\boldsymbol{B}$, which can preserve an existing margin and cannot create one. The ordinary differential equation method relates decreasing-step stochastic updates to the mean dynamics under bounded iterates, Robbins and Monro step sizes, martingale difference noise and step-size-weighted summable bias (Borkar, 2008). Convergence to this particular equilibrium also requires eventual entry into its attraction basin. Persistent estimation error or a fixed step size prevents direct application of that result, and equation 47 only quantifies stationarity of the distribution term in isolation. Taken together with Corollary 1, the supported qualifications are that the ideal target is locally attracting under an explicit margin condition and that finite error yields a joint bound rather than endpoint recovery from a small noisy divergence.

E Implementation details

This appendix records the TRD training and evaluation settings in Table 4. SANA-1.6B is a linear diffusion transformer with a deeply compressed autoencoder and a decoder-only text encoder (Xie et al., 2025). Stable Diffusion 3.5 Medium follows the multimodal rectified-flow transformer family introduced by Esser et al. (2024); Table 4 gives the exact checkpoint identifier without assigning a parameter count whose published sources use different scopes. For both teachers, the last L student blocks are duplicated across the K policy modes of equation 6, and only these policy heads and the backbone adapters are trained. The baseline rows use the adaptations described in section 5.1; they are not reproductions of the original papers' training regimes. Unless an ablation states otherwise, the listed settings are fixed.

Prompt split and evaluation protocol. All image experiments use prompts from a publicly available dataset (Schuhmann et al., 2022). We partition the prompts into disjoint training and test sets, reserve 2,500 prompts for testing and use the remainder for training. Every reported metric uses this test set. Each method generates four images per prompt under a fixed four-seed schedule shared across methods, yielding 10,000 images, and the corresponding teacher generates a matched 10,000-image reference set.

Metrics. We extract DINOv2 ViT-L/14 features with registers (Oquab et al., 2024; Darcet et al., 2024) from the method and teacher images. FD-DINOv2 fits a Gaussian to each feature set and computes their Fréchet distance, following the DINOv2 evaluation of Stein et al. (2023) and the original Fréchet metric (Heusel et al., 2017). Recall-DINOv2 applies the improved precision and recall estimator to measure coverage of the teacher feature manifold (Kynkäänniemi et al., 2019). DINO Diversity adapts the mean pairwise DINOv2 cosine distance of Sorokin et al. (2025) to four images per prompt. We estimate human preference with PickScore (Kirstain et al., 2023), HPSv2 (Wu et al., 2023) and ImageReward (Xu et al., 2023), visual appeal with LAION-Aesthetics Predictor V2 on CLIP ViT-L/14 features (Schuhmann et al., 2022; Radford et al., 2021), and text alignment with 100 times the non-negative cosine similarity of OpenAI CLIP ViT-B/32 embeddings (Radford et al., 2021). Table 4 records the specified model variants and DINO preprocessing.

Few-step training. Algorithm 2 states the few-step iteration of section 4.4, whose outer noising interval follows the sampled step count. For a two-step schedule $1 = a_0 > a_1 > a_2 = 0$, one-step rollouts draw the outer noising time from $[a_1, a_0]$, while two-step rollouts draw it from $[a_2, a_1]$.

Algorithm 2 One TRD training iteration for few-step distillation with random-length consistency sampling

Require: student θ, moving average $\bar{\theta}$ with rate μ, frozen teacher $\boldsymbol{v}_T$, consistency-sampling times $1 = a_0 > \cdots > a_R = 0$, stage-conditioned time distributions $\{\rho_r\}_{r=1}^{R}$ with $\mathrm{supp}(\rho_r) \subseteq [a_r, a_{r-1}]$, grid $\mathcal{S}_t = \{t, s_1, \ldots, s_M\}$ with $t > s_1 > \cdots > s_M \geq 0$, loss weights $\lambda_{\mathrm{tr}}, \lambda_{\mathrm{dm}}, \lambda_{\mathrm{fm}}, \lambda_{\mathrm{dp}}$ of equation 13 and time weight $\lambda(t)$

sample c, initial noise $\boldsymbol{x}_{a_0} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ and step count $r \sim \mathrm{Unif}\{1, \ldots, R\}$
for $i = 0, \ldots, r-2$ **do**
  advance the sampler, without gradients: predict $\hat{\boldsymbol{x}}_0^{(i)} = G_\theta(\boldsymbol{x}_{a_i}, a_i, c)$, draw $\boldsymbol{\epsilon}_{i+1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ and re-noise it to $\boldsymbol{x}_{a_{i+1}} = (1 - a_{i+1})\hat{\boldsymbol{x}}_0^{(i)} + a_{i+1}\boldsymbol{\epsilon}_{i+1}$
end for
generate: evaluate the student at $(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$ and integrate the output velocity by equation 16 to $\hat{\boldsymbol{x}}_0 = G_\theta(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$, the only endpoint prediction that retains gradients
re-noise: draw $t \sim \rho_r$ and independent $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$; form $\hat{\boldsymbol{x}}_t = (1 - t)\hat{\boldsymbol{x}}_0 + t\boldsymbol{\epsilon}$, retaining its gradient path to θ
predict velocities: evaluate the student at $(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)$; obtain $\{\boldsymbol{v}_{t,s}^{\theta} : s \in \mathcal{S}_t\}$ from equation 6, without further student evaluations
teacher rollout, without gradients: initialize $\hat{\boldsymbol{x}}_{t,t}^{T} = \hat{\boldsymbol{x}}_t$; for decreasing $s \in \mathcal{S}_t$, evaluate $\boldsymbol{v}_{t,s}^{T}$ and, if a next grid time s' exists, advance to $\hat{\boldsymbol{x}}_{t,s'}^{T}$ by equation 8
reference endpoint, without gradients: evaluate $G_{\bar{\theta}}(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$ on the same final-stage input
form the losses: use the grid velocities for equation 9, $\boldsymbol{v}_{t,t}^{\theta}$ and $\boldsymbol{\epsilon} - \mathrm{sg}[\hat{\boldsymbol{x}}_0]$ for equation 10, the detached start-time velocity difference and $\hat{\boldsymbol{x}}_t$ for equation 11, and the paired final-stage endpoints for equation 12
update: backpropagate equation 13, take one optimizer step on θ, then set $\bar{\theta} \leftarrow \mu\bar{\theta} + (1 - \mu)\theta$

Table 4: Configuration of the TRD image experiments.

Item	SANA-1.6B	SD3.5 Medium
Teacher and reference generation		
Checkpoint	`Efficient-Large-Model/Sana_1600M_1024px_BF16_diffusers`	`stabilityai/stable-diffusion-3.5-medium`
Reference scheduler and steps	checkpoint scheduler; 20 steps	FlowMatch Euler scheduler; 40 steps
Reference guidance scale	4.5	4.5
Conditioning and image resolution	frozen Gemma2 text encoder; 1024×1024	three frozen checkpoint text encoders; 1024×1024
Student policy		
Number of modes K	4	4
Duplicated final blocks L and output branches	$L = 1$; velocity, gate and momentum projections per mode	$L = 1$; velocity, gate and momentum projections per mode
LoRA rank, α and dropout	64; 64; 0	32; 32; 0
LoRA targets and trainable components	all backbone linear layers except the output projection; full policy heads	attention projections in all transformer blocks; full policy heads
Teacher and student parameter sharing	one pretrained backbone; student adapters are disabled for teacher queries	one pretrained backbone; student adapters are disabled for teacher queries
Policy initialization	copied final block and teacher output projection; zero gate logits; γ_k log-spaced in $[0.4, 5.0]$ with one head at 1	copied final block and teacher output projection; zero gate logits; γ_k log-spaced in $[0.8, 1.2]$ with one head at 1
TRD objective		
Loss weights $(\lambda_{\mathrm{tr}}, \lambda_{\mathrm{dm}}, \lambda_{\mathrm{fm}}, \lambda_{\mathrm{dp}})$	$(0.1, 10.0, 1.0, 1.0)$	$(0.1, 30.0, 1.0, 1.0)$
Effective time weight $\lambda(t)$	constant 1 with mean reduction	constant 1 with mean reduction

Item	SANA-1.6B	SD3.5 Medium
Preservation reference and coupling	refreshed every step with EMA rate $\mu = 0.9$; shared final-stage input	refreshed every step with EMA rate $\mu = 0.9$; shared final-stage input
Teacher and student guidance scales	4.5; 1.0	4.5; 1.0
Noising-time distribution	$\mathrm{Unif}(0, 1)$	uniform within the selected rollout interval, with raw-time clip $[10, 980]$
Trajectory grid and solver	explicit Euler; step 0.05; maximum span 0.30; at most 7 grid points	explicit Euler; step 0.025; maximum span 0.20; at most 9 grid points
Few-step consistency schedule	scheduler-generated times; rollout length sampled uniformly from $\{1, \ldots, R\}$	$R = 2$ with raw times $(1000, 780)$; rollout depth induced by the clipped interval widths
Momentum scale and $\lvert \log \gamma_k \rvert$ clamp	1.0; 20.0	1.0; 20.0
Optimization		
Learning rates (backbone/basic; gate; momentum)	10^{-4}; 10^{-5}; 10^{-5}	10^{-4}; 10^{-5}; 10^{-5}
Optimizer and schedule	8-bit AdamW (Loshchilov & Hutter, 2019; Dettmers et al., 2022); $(\beta_1, \beta_2) = (0.9, 0.999)$; weight decay 10^{-4}; constant; no warmup	same
Batch, precision and clipping	global batch 8; bfloat16; gradient norm 1.0	global batch 8; bfloat16; gradient norm 1.0
Evaluation		
Held-out prompts and seeds	2,500 disjoint prompts selected with dataset seed 666; image seeds start at 42	same
Sample count	four images per prompt and 10,000 student images; matched teacher reference set	same
FD-DINOv2 features and preprocessing	`dinov2_vitl14_reg`; resize to 256, center crop to 224, ImageNet normalization	same
Recall-DINOv2	student 3-NN manifold balls; coverage of the matched 10,000-image teacher set	same
DINO Diversity	DINOv2 features; mean pairwise cosine distance within each four-image prompt group	same
Preference and aesthetic models	PickScore v1; HPSv2; ImageReward v1.0; LAION-Aesthetics Predictor V2 with CLIP ViT-L/14 (Schuhmann et al., 2022; Radford et al., 2021)	same
CLIP similarity backbone	OpenAI CLIP ViT-B/32; $100 \max(\cos, 0)$	same

F PER-ITERATION COST

This appendix accounts for the network evaluations and the persistent state of one TRD iteration, and compares them with a distribution distillation method that trains a separate fake score model. The counts follow Algorithm 1 under the settings of Table 4; section 5.3 reports what they cost in measured memory and time.

Network evaluations. One iteration evaluates the student twice, once at $(\boldsymbol{x}_1, 1, c)$ for generation and once at the detached noised state $(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)$ for the shared velocity predictions, and backpropagates through both. The frozen teacher is evaluated $M + 1$ times, once at each grid point of the

rollout in equation 8, and the moving-average copy once at $(\boldsymbol{x}_1, 1, c)$. Only the two student evaluations carry gradients, so the teacher rollout and the moving-average pass need no stored activations. The grid size therefore enters the cost through frozen forward passes alone. Because $\mathcal{P}_\theta(\mathrm{sg}[\hat{\boldsymbol{x}}_t], t, c)$ gives $\boldsymbol{v}^\theta_{t,s}$ for every $s \in \mathcal{S}_t$ in closed form, supervising $M+1$ grid points costs no additional student evaluation, whereas supervising the same points with a student that predicts one velocity per query would require one student pass per point. With classifier-free guidance each grid point evaluates the conditional and the unconditional branch, which raises the teacher count to $2(M+1)$ as recorded in Appendix D.

Parameters and optimizer state. TRD trains the policy heads and the low-rank backbone adapters and keeps the pretrained backbone frozen, so gradients and optimizer state exist only for that trainable subset. Teacher queries reuse the same backbone with the student adapters disabled, and no second copy of the backbone weights is held for the teacher. The moving-average copy tracks the trainable subset as well, so distribution preservation adds one parameter copy of that size and one forward pass, without gradients or optimizer state.

Comparison with a trained fake score model. A method that estimates the student score with a separate network holds that network at full size together with its gradients and optimizer state and updates it with its own forward and backward passes. The DMD2 measurement uses five fake-score updates per generator update, matching its ImageNet and SDXL configurations under the two-time-scale rule of Yin et al. (2024a); this ratio is a configuration choice rather than a universal property of DMD2. Estimating the same score with self flow matching on the shared policy removes those updates and that state, and adds in their place the $M+1$ frozen teacher passes of the trajectory rollout and the single moving-average pass. Table 2 and the right panel of figure 1 report where this exchange leaves the measured memory, per-iteration time and time to a target FD-DINOv2 level.

Few-step variant. Algorithm 2 prepends $r-1$ gradient-free student passes to the same iteration, where r is drawn uniformly from $\{1, \ldots, R\}$, so the expected additional cost is $(R-1)/2$ student forward passes per iteration. All remaining counts are unchanged after these detached passes. In particular, the moving average still requires one forward pass at $(\boldsymbol{x}_{a_{r-1}}, a_{r-1}, c)$; for $r = 1$, this reduces to the one-step input $(\boldsymbol{x}_1, 1, c)$.

G ADDITIONAL RESULTS

This appendix reports the Stable Diffusion 3.5 Medium comparison in Table 5, the objective, estimator and policy-capacity ablations in Tables 6, 7 and 8, the policy-head distribution analysis of figure 4, the Stable Diffusion 3.5 Medium training-loss curves in figure 3, and the qualitative results in Figures 5 to 8. All quantitative tables use the 10,000-image evaluation protocol of Appendix E.

Table 5: One-step and few-step results at 1024×1024 for students distilled from Stable Diffusion 3.5 Medium. Conventions follow Table 1.

Method	NFE	Extra models	FD-DINOv2 ↓	Pick ↑	HPS ↑	ImgRwd ↑	Aesth ↑	CLIP ↑	Recall ↑	Diversity ↑
SD3.5-Medium teacher	40×2	–	–	0.2081	0.2425	0.8684	6.2758	32.2167	–	0.2828
Trajectory-oriented objectives										
PD	1	0	453.30	0.1706	0.1823	-2.0211	5.0247	22.9582	0.0080	0.2476
CD	1	0	385.54	0.1737	0.1905	-1.7267	5.3324	24.3874	0.0120	0.2194
sCD	1	0	270.78	0.1713	0.1925	-1.2308	5.2646	24.8433	0.0361	0.2202
ArcFlow	1	0	293.60	0.1773	0.1964	-1.4753	5.1572	23.9076	0.0106	0.2317
MeanFlow	1	0	177.27	0.1808	0.2050	-0.9413	5.7254	27.1513	**0.1479**	**0.3934**
Distribution-oriented objectives										
DMD	1	1	133.24	0.1867	0.2116	-0.1388	5.4866	29.3045	0.0829	0.2322
DMD2	1	2	117.75	0.1827	0.2244	0.0090	5.5862	29.4182	0.1355	0.2161
SiD	1	1	125.33	0.1951	0.2188	0.1253	5.4694	29.0743	0.0960	0.2330
FSF-DMD	1	0	158.72	0.1748	0.1829	-1.3202	5.2847	25.0290	0.0138	0.2482
TwinFlow	1	0	137.35	0.1803	0.2128	-0.2031	5.3989	28.4420	0.0502	0.2247
Combined trajectory and distribution objective										
TRD (ours)	1	0	**86.66**	**0.2045**	**0.2433**	**0.8000**	**6.2299**	**31.3551**	0.1001	0.1904
Trajectory-oriented objectives										
PD	2	0	162.17	0.1831	0.2008	-1.1153	5.6754	27.7708	0.1630	0.3670
CD	2	0	161.58	0.1866	0.2171	-0.4251	5.9082	28.4025	0.1495	0.3383
sCD	2	0	161.38	0.1825	0.2198	-0.1937	5.9318	28.2862	0.1497	0.3293
ArcFlow	2	0	176.43	0.1840	0.2068	-0.3930	5.3749	27.3055	0.1038	0.3442
MeanFlow	2	0	159.13	0.1815	0.2048	-0.9471	5.8877	27.0103	0.2430	**0.4534**
Distribution-oriented objectives										
DMD	2	1	**49.12**	0.2021	0.2327	0.6139	6.0237	31.1482	**0.4597**	0.2375
DMD2	2	2	54.70	0.1988	0.2334	0.5822	6.0090	31.1444	0.3903	0.2452
SiD	2	1	53.66	0.1978	0.2344	0.5529	5.8950	29.9768	0.3887	0.2409
FSF-DMD	2	0	68.43	0.1958	0.2127	0.4285	5.7737	29.6091	0.2514	0.2212
TwinFlow	2	0	64.46	0.2001	0.2305	0.3867	5.9101	30.0574	0.3834	0.1962
Combined trajectory and distribution objective										
TRD (ours)	2	0	49.90	**0.2047**	**0.2435**	**0.9276**	**6.1135**	**31.3486**	0.2639	0.1932

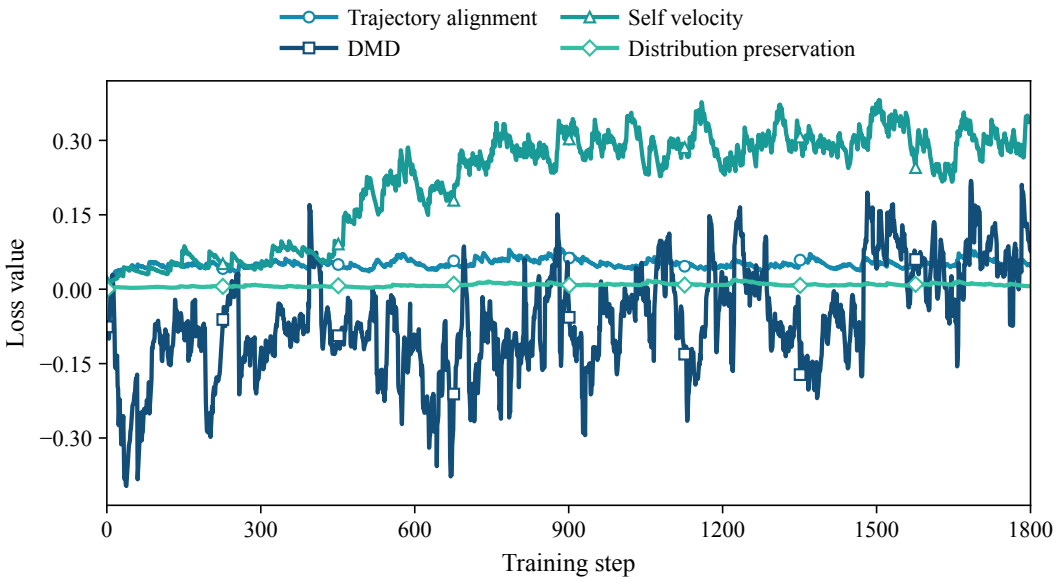


Figure 3: TRD training loss curves for Stable Diffusion 3.5 Medium. EMA smoothing is applied for clearer visualization.

Table 6: Contribution of the objective components in equation 13 for one-step distillation from SANA-1.6B.

Setting	FD-DINOv2 ↓	Pick ↑	HPS ↑	ImgRwd ↑	Aesth ↑	CLIP ↑	Recall ↑	Diversity ↑
w/o $\mathcal{L}_{\mathrm{tr}}$	107.01	0.1880	0.2179	0.3242	6.0409	28.0689	0.1703	0.3258
w/o $\mathcal{L}_{\mathrm{fm}}$ & $\mathcal{L}_{\mathrm{dm}}$	564.73	0.1657	0.1821	-2.0941	5.1234	20.4366	0.0052	0.1987
w/o $\mathcal{L}_{\mathrm{dp}}$	*training collapsed*							
full objective	70.34	0.1915	0.2223	0.3888	6.1169	28.3647	0.3162	0.3204

Table 7: Shared self-estimation against a separately trained fake velocity network for one-step distillation from SANA-1.6B.

Estimator	Extra models	FD-DINOv2 ↓	Pick ↑	HPS ↑	ImgRwd ↑	Aesth ↑	CLIP ↑	Recall ↑	Diversity ↑
Shared self-estimator (default)	0	70.34	0.1915	0.2223	0.3888	6.1169	28.3647	0.3162	0.3204
Separate fake velocity network	1	127.86	0.1861	0.2137	-0.0293	5.9505	27.7603	0.1772	0.3337

Table 8: Policy capacity for one-step distillation from SANA-1.6B, varying the number of momentum modes K and the number of duplicated final blocks L.

Setting	Added trainable params	FD-DINOv2 ↓	Pick ↑	HPS ↑	ImgRwd ↑	Aesth ↑	CLIP ↑	Recall ↑	Diversity ↑
$L=1, K=1$	5%	61.41	0.1913	0.2228	0.4081	6.1096	28.3633	0.3659	0.3260
$L=1, K=2$	10%	53.53	0.1936	0.2236	0.4545	6.1845	28.3687	0.4744	0.3406
$L=1, K=4$	20%	70.34	0.1915	0.2223	0.3888	6.1169	28.3647	0.3162	0.3204
$L=1, K=8$	40%	72.82	0.1917	0.2220	0.3404	6.1651	28.3131	0.3148	0.3357
$K=4, L=0$	$\sim 0\%$	79.65	0.1909	0.2206	0.3229	6.0988	28.3210	0.2718	0.3259
$K=4, L=1$	20%	70.34	0.1915	0.2223	0.3888	6.1169	28.3647	0.3162	0.3204
$K=4, L=2$	40%	65.24	0.1921	0.2240	0.3982	6.1258	28.5433	0.3467	0.3209

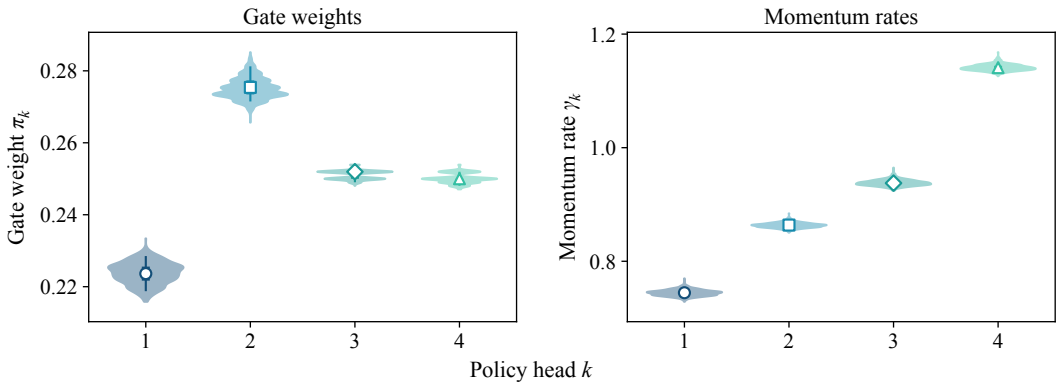


Figure 4: Distributions of the gate weights π_k and momentum rates γ_k produced at $t = 1$ by the four policy heads of a TRD checkpoint distilled from Stable Diffusion 3.5 Medium. Appendix B gives the plotting conventions.

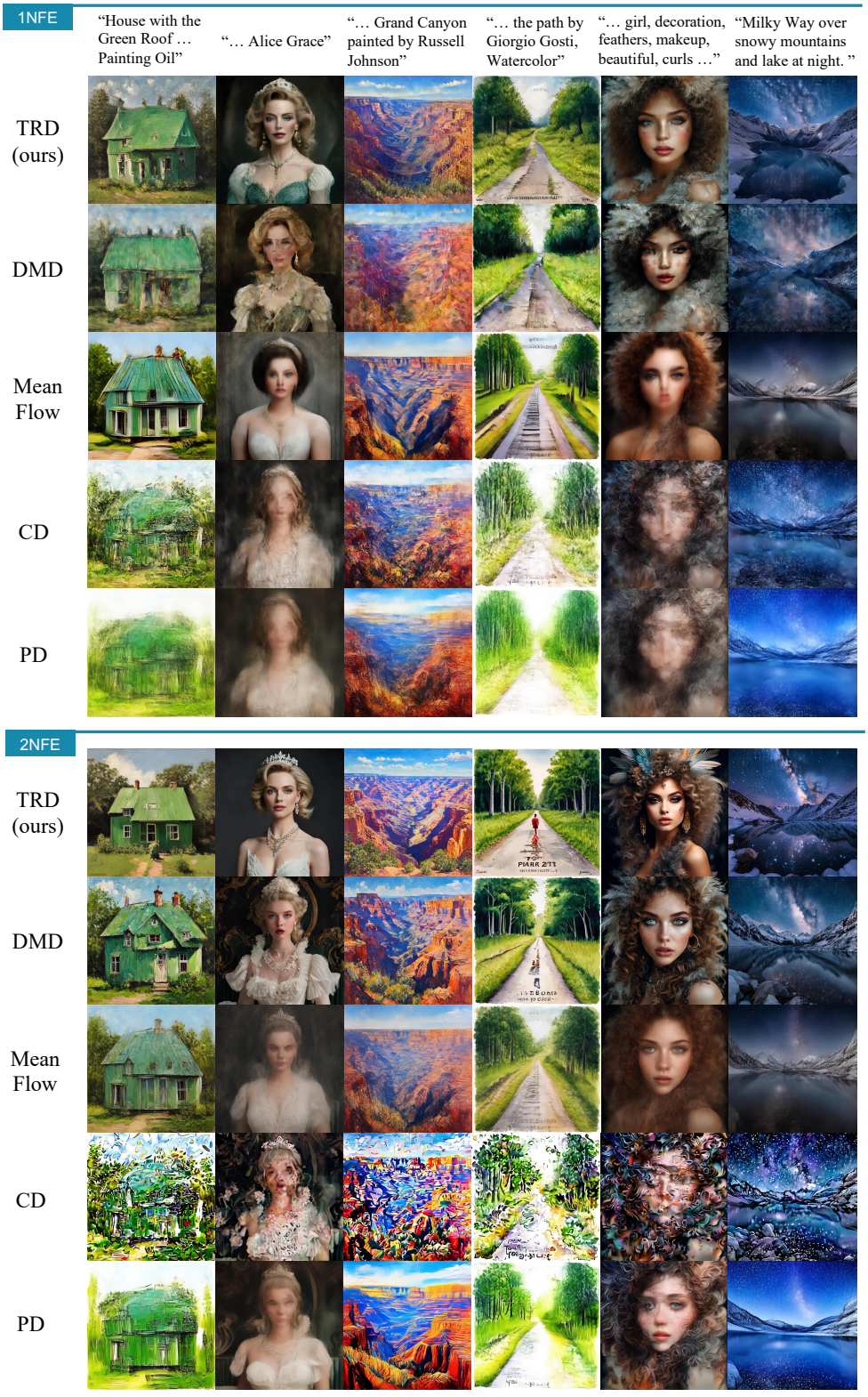


Figure 5: Qualitative comparison of PD, CD, MeanFlow, DMD and TRD at one and two steps, distilled from SANA-1.6B.

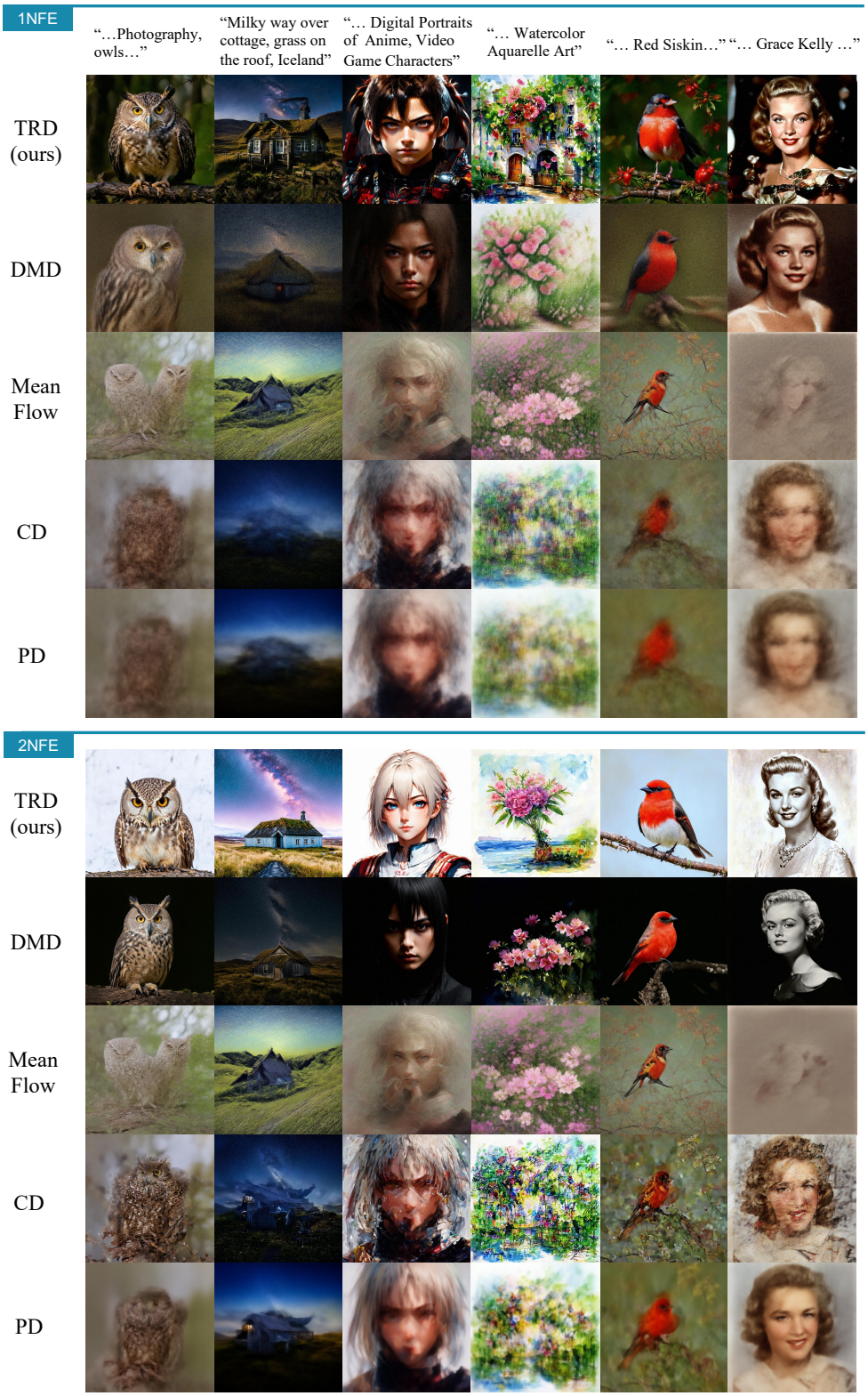


Figure 6: Qualitative comparison of PD, CD, MeanFlow, DMD and TRD at one and two steps, distilled from Stable Diffusion 3.5 Medium.



Figure 7: Additional one-step TRD samples distilled from SANA-1.6B.



Figure 8: Additional one-step TRD samples distilled from Stable Diffusion 3.5 Medium.