ACQUIRE ONCE, DEPLOY MANY: ZERO-TARGET-FIT TEMPORAL MEASURES FOR FEW-STEP FLOW MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Few-step flow samplers choose both the number and timing of field evaluations. Existing scheduling methods often repeat this decision for each target model. We instead acquire a temporal measure once and reuse it. At a budget of eight function evaluations (NFE = 8), explicit-midpoint second-order Runge–Kutta (RK2) has four executable intervals, represented by four positive widths. We fit these widths on an inexpensive unconditional CIFAR-10 Diffusion Transformer (DiT) using projected terminal Fréchet risk, freeze them, and copy them unchanged to ImageNet receivers. Deployment uses no target trajectory, feature, gradient, extra parameter, auxiliary network, or additional field evaluation. Canonical Fréchet Inception Distance (FID), evaluated with the Ablated Diffusion Model (ADM) suite over 50,000 samples (FID50k), falls from 8.3674 to 7.1696 on Scalable Interpolant Transformer (SiT-XL/2) and from 17.9714 to 14.6264 on Self-Flow; paired 95% bootstrap intervals for the FID reductions (uniform minus the acquired measure) are $[1.1088, 1.2975]$ and $[3.1486, 3.4903]$. A second, independently seeded acquisition also improves fresh paired FID10k on SiT, Self-Flow, and REPresentation Alignment (REPA). A one-parameter power measure retains 83.7–91.8% of the flexible four-interval gain, with a receiver-dependent residual. The frozen measure transfers to an unseen Ralston RK2 tableau, and the allocation ordering agrees between ADM FID and independent DINOv2 metrics. Applied separately to forecasting, the same acquire-then-freeze protocol improves ensemble continuous ranked probability score in all six fresh receiver cells across electricity and short-horizon Lorenz forecasting at NFE = 8. The evidence supports amortized post-training temporal allocation within stated objective, source family, receiver family, solver, and budget contracts.

1 INTRODUCTION

Flow matching learns a velocity field that transports a simple source to a data distribution (Lipman et al., 2023; Tong et al., 2024). A numerical solver samples that field at selected times. When only a few evaluations are available, their placement can matter as much as their number. Under a few-step budget, sampling time is a finite deployment resource.

Recent methods improve few-step generation by changing the model or fitting inference-time controls. Shortcut models and adaptive coefficient modules retrain or augment the receiver (Frans et al., 2025; Lee et al., 2025). Schedulers, velocity or sharpness profiles, and instance-aware grids instead change how a fixed receiver is sampled (Min et al., 2026; Bondar, 2026; Gupta et al., 2026; Yuan et al., 2026). When a control is fitted or profiled on a receiver, its acquisition cost recurs as receivers change. We ask whether a cheap source can absorb that cost once, before any target receiver is queried.

Our deployed object is a positive temporal measure. At midpoint NFE = 8, four interval masses determine all four executable steps with three effective degrees of freedom. We optimize those masses against a terminal distribution objective on one inexpensive source model, freeze the result, and copy it to class-conditional ImageNet receivers. No target trajectory, feature, gradient, quality metric, or selection query enters acquisition.

The acquired measure improves canonical FID50k by 14.3% on SiT and 18.6% on Self-Flow. Repeating acquisition with a new seed yields different masses but again improves SiT, Self-Flow, and an independently trained REPA receiver on fresh paired streams. A directly acquired power measure recovers most of the gain. Transfer follows the temporal allocation, not one optimizer trajectory.

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.

The paper makes three contributions:

- We formulate temporal allocation as a positive measure that fully specifies a finite-step solver action and is acquired independently of target receivers.
- We demonstrate zero-target-fit transfer to official SiT, Self-Flow, and REPA receivers through independent reacquisition, paired uncertainty, DINOv2 witnesses, and matched-cost controls.
- Frozen-clock interventions, solver and budget refinements, and history-conditioned forecasting experiments identify the contract under which the measure is reusable.

2 Related Work

Flow matching and stochastic interpolants provide the continuous transport models used here (Lipman et al., 2023; Albergo et al., 2025; Tong et al., 2024). Path straightening and shortcut models modify training or the receiver (Liu et al., 2023; Shankar & Geffner, 2025; Frans et al., 2025); we keep receiver weights fixed. Classical step allocation, mesh redistribution, and adjoint-weighted goal adaptivity motivate temporal measures (Lindberg, 1977; Estep, 1995; Carey & Dinh, 1985; Becker & Rannacher, 2001; Giles & Süli, 2002). Work on generative schedules uses geometry-inspired time scheduling (GITS), field sharpness, trajectory regularity, or hierarchical midpoint-error objectives (Gupta et al., 2026; Chen et al., 2024; Zhu et al., 2026). Model-specific solver learning instead adapts integrator coefficients or non-stationary solver structure to a target flow (Shaul et al., 2024a;b). Learning to Discretize Denoising Diffusion ODEs (LD3) learns target time discretizations, and Adaptive Reparameterized Time (ART) learns a reusable clock control (Tong et al., 2025; Huang et al., 2026). DPM-Solver-v3, instance-aware discretizations (INDIS), and Stabilized Taylor Orthogonal Runge–Kutta (STORK) alter solver statistics, instance grids, or integration structure (Zheng et al., 2023; Yuan et al., 2026; Tan et al., 2026). Align Your Steps (AYS), constant total rotation schedules (TORS), and Optimize Your Sampling (OYS) optimize deployment-specific steps, schedules, or metrics (Xue et al., 2024; Sabour et al., 2024; Zhou et al., 2026; Zhang et al., 2026). Delay Flow Matching adds delayed state dependence to the learned dynamics (Zhao et al., 2026). Our setting changes neither the receiver nor the integrator. It acquires a terminal-objective temporal measure on a cheap source and tests that frozen object on receivers unseen during acquisition.

Training-coupled methods alter the sampled field, and target-adaptive schedules query the deployed receiver. Frozen-source transfer instead reuses a source decision without observing the target, so we report target access and acquisition cost separately from sampling NFE.

3 Finite-Step Temporal Control

3.1 Clock intervention and induced grids

Let $u_\theta(t, x)$ be a frozen velocity field, $I(t) = t$ the identity clock, and $C_\phi(t)$ a learned monotone clock with fixed endpoints. We define the function-space homotopy

$$C_\lambda(t) = (1 - \lambda)t + \lambda C_\phi(t), \qquad \lambda \in [0, 1]. \tag{1}$$

The corresponding time-only dynamics is

$$\dot{x} = C'_\lambda(t) u_\theta(C_\lambda(t), x). \tag{2}$$

For any strictly monotone endpoint-preserving C_λ, the exact solution satisfies $x_\lambda(t) = x_I(C_\lambda(t))$, so $x_\lambda(1) = x_I(1)$. A finite Runge–Kutta rule need not preserve this equivalence because its stages and quadrature weights remain fixed in physical time.

For a uniform physical grid t_n, denote the numerical endpoint of Eq. 2 by $\mathcal{D}_h^{\mathrm{clock}}$. An induced-grid baseline instead applies $u_\theta(\tau, x)$ on $\tau_n = C_\lambda(t_n)$ and produces $\mathcal{D}_h^{\mathrm{grid}}$. For a nonlinear clock and finite h, internal stages query different times and states, so generally

$$\mathcal{D}_h^{\mathrm{clock}} \neq \mathcal{D}_h^{\mathrm{grid}}, \qquad \|\mathcal{D}_h^{\mathrm{clock}} - \mathcal{D}_h^{\mathrm{grid}}\| \to 0 \quad \text{as } h \to 0. \tag{3}$$

Consistency and stability give the limiting relation, but say nothing about monotone contraction at the measured budgets or the direction of a quality change. We test those two questions separately and keep every receiver parameter fixed.

3.2 GOAL-ORIENTED TEMPORAL MEASURE

For explicit midpoint RK2 with K intervals and exact NFE $2K$, let $m = (m_1, \ldots, m_K)$ satisfy $m_j > 0$ and $\sum_j m_j = 1$. Ordered nodes are $t_j = \sum_{\ell \le j} m_\ell$, and interval j applies

$$\begin{aligned} k_{j,1} &= u_\theta(t_{j-1}, x_{j-1}), \\ k_{j,2} &= u_\theta(t_{j-1} + m_j/2, x_{j-1} + m_j k_{j,1}/2), \\ x_j &= x_{j-1} + m_j k_{j,2}. \end{aligned} \tag{4}$$

We parameterize $m = \mathrm{softmax}(\eta)$, which preserves positivity, ordered nodes, fixed endpoints, the midpoint tableau, and second-order consistency.

The vector m completely specifies the deployed midpoint operator at exact NFE $2K$. By comparison, a clocked finite-step operator depends on both C_ϕ and its derivative at solver stages; its timestep nodes alone do not determine the action. The temporal measure stores only K positive widths and requires no learned-function evaluation on the receiver.

The four-bin representation follows from the midpoint NFE = 8 budget: two field evaluations per step expose four physical intervals and three free masses. Additional bins would not describe executable NFE = 8 steps. At NFE = 16, subdividing each mass instead preserves the normalized density over eight valid intervals.

The source objective compares generated and real distributions after a frozen feature map, normalization, and random orthogonal projection. With empirical means μ_m, μ_r and covariances Σ_m, Σ_r, the projected terminal Fréchet risk is

$$R(m) = \|\mu_m - \mu_r\|_2^2 + \mathrm{tr}\Big(\Sigma_m + \Sigma_r - 2(\Sigma_m^{1/2}\Sigma_r\Sigma_m^{1/2})^{1/2}\Big). \tag{5}$$

We compare differentiation through the solver with matched-cost Sobol initialization followed by Gaussian process (GP) expected improvement. Both procedures search the same temporal measure under the same source budget. We freeze the selected measure before target evaluation, separating the search procedure from the object that is deployed.

Acquire-once/deploy-many contract. Figure 1 summarizes the contract. Before search, we freeze the source receiver, calibration noises, real-feature bank, projection, solver, and exact NFE. Differentiable or Sobol/GP search spends a fixed source budget, and an untouched source stream chooses the measure before any target run. We then copy its masses to every NFE = 8 target without interpolation, reversal, or rescaling. Target samples are used only for the reported metrics; they never update the measure. Independent acquisition and NFE refinement use the same freeze-before-evaluation rule.

The deployment record stores the masses, terminal objective, feature witness, solver tableau, NFE, and source provenance. Changing the objective, source family, receiver family, solver, or budget defines a new contract.

3.3 CONDITIONAL ASYMPTOTIC MOTIVATION AND BUDGET REFINEMENT

The discrete measure has a continuous fixed-budget interpretation. For a stable order-p one-step method with N integration intervals, suppose the leading terminal-goal error has a nonnegative density $a(t)$. With normalized step density $\rho(t)$, the leading risk is

$$\mathcal{E}_N(\rho) = N^{-p} \int_0^1 a(t)\rho(t)^{-p}\, dt + o(N^{-p}), \qquad \int_0^1 \rho(t)\, dt = 1. \tag{6}$$

A Lagrange multiplier yields

$$\rho^\star(t) = \frac{a(t)^{1/(p+1)}}{\int_0^1 a(s)^{1/(p+1)}\, ds}. \tag{7}$$

The function $a(t)$ combines terminal adjoint sensitivity with local numerical defect, unlike a curvature score. This dependence on the goal motivates our terminal objective and the controls in Section 5.6. At finite budget, we optimize four-interval masses directly instead of estimating $a(t)$.

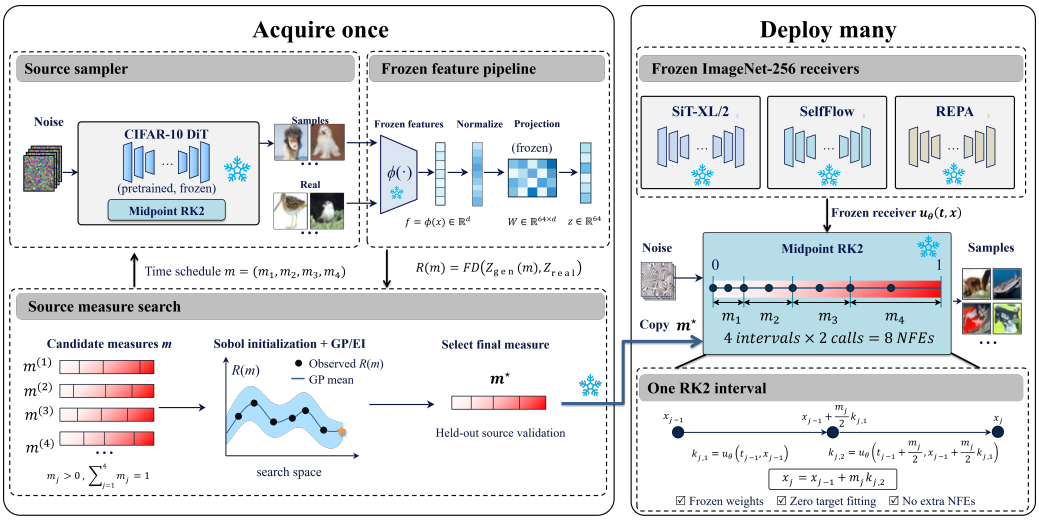

Figure 1: **Acquire once, deploy many.** A four-interval temporal measure is selected on a frozen CIFAR-10 source using a fixed terminal witness and held-out source validation. The measure is then copied unchanged to frozen SiT, Self-Flow, and REPA receivers at exact NFE = 8.

Conditional asymptotic certificate. Let $J(\rho) = \int_0^1 a(t)\rho(t)^{-p}\,dt$ and $Z = \int_0^1 a(t)^{1/(p+1)}\,dt$. For any positive normalized density, Hölder's inequality gives

$$Z = \int_0^1 \left(a\rho^{-p}\right)^{1/(p+1)} \rho^{p/(p+1)} dt \le J(\rho)^{1/(p+1)}, \tag{8}$$

so $J(\rho) \ge Z^{p+1}$, with equality exactly at Eq. 7 (up to zero-density sets). The leading excess terminal error is $N^{-p}[J(\rho) - Z^{p+1}]$. Because equality depends on the terminal sensitivity $a(t)$, the relevant allocation is tied to a particular objective, receiver, solver, and metric.

The interval masses induce the piecewise-constant density $\rho_m(t) = (Km_j)^{-1}$ on interval j. Replacing each mass by two children $(m_j/2, m_j/2)$ doubles K while leaving ρ_m unchanged. This deterministic subdivision carries a measure to the next midpoint budget without fitting new parameters.

Each m_j is a physical interval width, not an evaluation probability. At fixed K, the solver assigns one interval and two midpoint field calls to that width, so increasing m_j lowers the local evaluation density. This distinction also makes the refinement rule unambiguous: it preserves density rather than reusing the numerical widths as sampling weights.

4 Experimental Protocol

We diagnose clocks on independently implemented corrected-DiT (Peebles & Xie, 2023) and TorchCFM CIFAR-10 families. With receiver weights fixed, the tests cross Euler, explicit-midpoint second-order Runge–Kutta (RK2), Heun RK2, and fourth-order Runge–Kutta (RK4) at exact $\mathrm{NFE} \in \{8, 16, 24, 32, 64\}$. Midpoint, Heun, and Ralston use distinct RK2 tableaux; interval masses always denote physical widths. Independently trained artifacts are the statistical units, and paired comparisons share initial noises. The core DiT cohorts contain five artifacts. The TorchCFM confirmation contains ten field-only and joint pairs.

Table 1 defines the trainable parameters, losses, and test-time clocks. In DiT, θ_L contains only the sixth output block and final layer. A field-only run never updates clock parameters; at evaluation, it uses either I or the frozen clock learned by its matched joint run.

For measure acquisition, we freeze one unconditional CIFAR-10 DiT at midpoint NFE = 8. Gradient, pure Sobol, Sobol/GP, and direct power searches each receive 128 calibration noises, one frozen 64-dimensional projected witness, and 131,072 model-sample NFEs, the product of source-sample count and per-sample field evaluations. The direct power search uses $C_\gamma(r) = r^\gamma$ over $\gamma \in [0.25, 4]$ and induces $t_j = C_\gamma(j/K)$ and $m_j = t_j - t_{j-1}$ rather than fitting the four-bin solution. A disjoint

Table 1: **Clock-diagnosis arms.** Adaptation minimizes terminal rollout risk and flow-matching mean squared error (MSE). The field-only arm fixes $C = I$ during adaptation; its paired evaluation uses the frozen C_{ϕ_i} learned by the matched joint run.

Arm	Trainables	Adaptation loss	Evaluation clock
Field-only	θ_L; $C = I$ fixed	terminal rollout Fréchet risk + flow-matching MSE	I or paired C_{ϕ_i}
Joint	θ_L, ϕ	same losses + linear-path anchor + clock identity regularization	I or own C_{ϕ_i}

CleanFID10k stream evaluates source quality, and a second independently seeded Sobol/GP run tests acquisition repeatability.

We report source search, receiver-specific adaptation, and generation NFE separately. Matched source methods receive equal budgets; target-profiled baselines report receiver calls, while frozen measures incur none.

We freeze every temporal measure before deploying it to official SiT-XL/2 (Ma et al., 2024), 675M-parameter Self-Flow (Chefer et al., 2026), and REPA (Yu et al., 2025) ImageNet-256 receivers. SiT and REPA use classifier-free guidance (CFG) scale 1.5; Self-Flow uses its native conditional sampler without CFG. Guidance strength does not change across schedules. Paired ADM FID2k/FID10k screens precede canonical FID50k on SiT and Self-Flow. A fresh paired FID10k experiment crosses all three receivers with uniform, original Sobol/GP, independent Sobol/GP, and direct power. The 1,000 common-index feature bootstraps condition on the fixed checkpoints, generation streams, and ADM reference. We use CleanFID10k for the CIFAR-10 source and ADM FID2k/FID10k/FID50k for ImageNet receivers (Parmar et al., 2022; Heusel et al., 2017). Additional frozen tests use DINOv2 features (Oquab et al., 2024), Ralston RK2, Kutta's third-order Runge–Kutta (RK3), midpoint NFE = 16, GITS-style target profiling, SharpEuler, and an LD3 adaptation (Tong et al., 2025) based on the released official code. It is trained with learned perceptual image patch similarity (LPIPS) (Zhang et al., 2018) on 100 SiT NFE = 64 teacher images. Here, NFE denotes the number of function evaluations, with one NFE equal to one neural-field evaluation.

Paired intervals quantify cached-stream and FID-estimator uncertainty; independent acquisition and fresh receiver seeds test the remaining sources of variation.

The data split fixes zero-target fitting: calibration noises score source candidates and a disjoint source stream selects the measure; target data only evaluate it. Competing schedules share latents, classes, and reference statistics, and the bootstrap preserves this pairing.

The history-conditioned tests map 96 Electricity Transformer Temperature minute-level 1 (ETTm1) (Zhou et al., 2021) observations to 24 future steps and 32 Lorenz (Lorenz, 1963) states to the next 32. Two source receivers screen 64 candidates, after which the selected measure is frozen for three fresh receiver seeds. Ensemble continuous ranked probability score (CRPS) at midpoint NFE = 8 is the primary metric. NFE = 4 and NFE = 16, point metrics, and dense-reference errors test the budget and objective contracts. Canonical records contain the complete splits, seeds, hashes, metrics, and secondary cells.

5 Results

5.1 Zero-target-fit transfer across ImageNet receivers

A four-interval measure acquired on an unconditional CIFAR-10 source transfers unchanged to three class-conditional ImageNet receivers. Canonical ADM FID50k improves by 14.3% on SiT and 18.6% on Self-Flow. On fresh FID10k streams, a second, independently acquired measure improves SiT $11.4937 \to 10.3454$, Self-Flow $20.8918 \to 16.9490$, and REPA $10.1904 \to 9.4505$. In every case, the receiver weights, parameter count, and NFE remain unchanged.

On SiT and Self-Flow, ADM FID, DINOv2 feature Fréchet distance (FD) (Oquab et al., 2024), Kernel Inception Distance (KID) (Bińkowski et al., 2018), and precision/recall (Kynkäänniemi et al., 2019) agree on the all-arm order: Sobol/GP, pure Sobol, gradient, uniform, then GITS-style. Source-acquired measures improve every DINOv2 metric over uniform in all 24 predefined directions. For

Sobol/GP, FD decreases by 13.7% and 12.6%, and KID by 22.7% and 19.0%, on SiT and Self-Flow, respectively. Recall increases by 0.0340 and 0.0374, respectively, while precision also improves on both receivers. This audit reuses paired 10,000-sample streams and adds no receiver calls, separating feature-witness agreement from additional sampling.

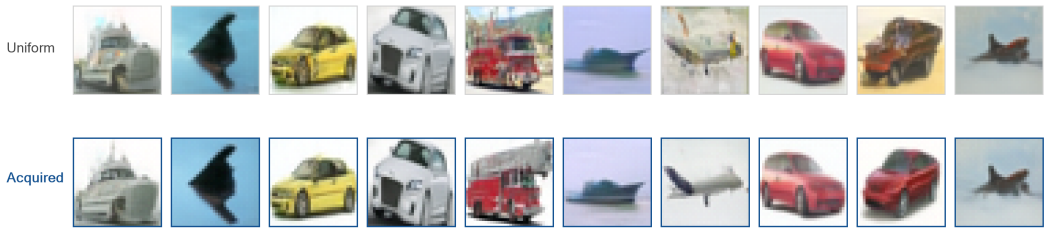

Figure 2: **Matched CIFAR-10 source examples under uniform and acquired temporal allocation.** Ten corrected-DiT sample pairs under uniform allocation (top) and the source-acquired four-interval measure (bottom), both with explicit midpoint at NFE = 8. Each column uses the same initial noise for both schedules. These examples are illustrative rather than an aggregate sharpness evaluation.

5.2 Frozen clocks locate the finite-step mechanism

At Heun NFE = 8, changing only the corrected-DiT receiver field shifts the clock effect from -2.040 to $+31.531$ FID, where positive denotes improvement. The corresponding TorchCFM effects are $+5.457$ and $+5.946$. The clock's finite-step effect is receiver dependent.

At Heun NFE = 8, the selected clock beats its induced grid by 7.671 and 33.876 FID for corrected-DiT field-only and joint receivers, respectively. Across four solvers and five budgets, 39 of 40 receiver–solver discrepancy slopes contract with NFE. Mean gain contracts from 31.86 to -0.13 for Heun and from 7.68 to 0.35 for RK4 between NFE = 8 and NFE = 64. As Eq. 3 predicts, clocked and grid operators remain distinct at low budget and approach one another as the grid is refined.

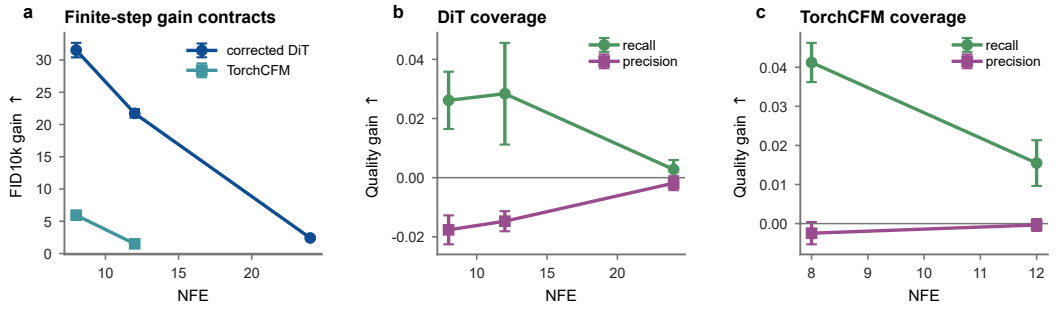

Figure 3: **Finite-step and coverage mechanism.** **a**, Clock FID gains contract with NFE. **b–c**, Low-NFE recall rises while precision changes little; intervals and points summarize five independent artifacts.

The maximum clock–identity endpoint discrepancy reaches 8.79×10^{-4} at RK4-512. At NFE = 8, recall rises by 0.0261 on DiT and 0.0412 on TorchCFM (Figure 3), while precision changes little. A frozen dose selector improves all 20 fresh TorchCFM receivers at 6.4% of one FID10k evaluation's NFE = 8 cost.

5.3 A source-acquired temporal measure enables zero-target-fit transfer

At matched cost, pure Sobol and Sobol/GP search the same three-dimensional mass space. Sobol/GP returns $[0.0903, 0.2063, 0.2761, 0.4272]$ and has the lowest source FID (Table 2).

One acquisition corresponds to 16,384 source trajectories at NFE = 8 and adds no target computation. An independent Sobol/GP run produces $[0.1206, 0.1787, 0.1949, 0.5058]$; both vectors place their largest mass in the final interval and beat uniform on all eight cross-bank streams.

Table 2: **Matched-cost source acquisition at midpoint NFE = 8.** Wall time excludes shared setup; FID is an untouched paired CleanFID10k read. Boldface marks the lowest FID.

	Acquisition cost		Source quality
Allocation	Model-sample NFEs	Time (s)	FID10k ↓
Uniform	–	–	26.4430
Gradient	131,072	63.65	22.4064
Pure Sobol	131,072	48.12	20.3816
Sobol/GP	131,072	54.64	**20.0774**

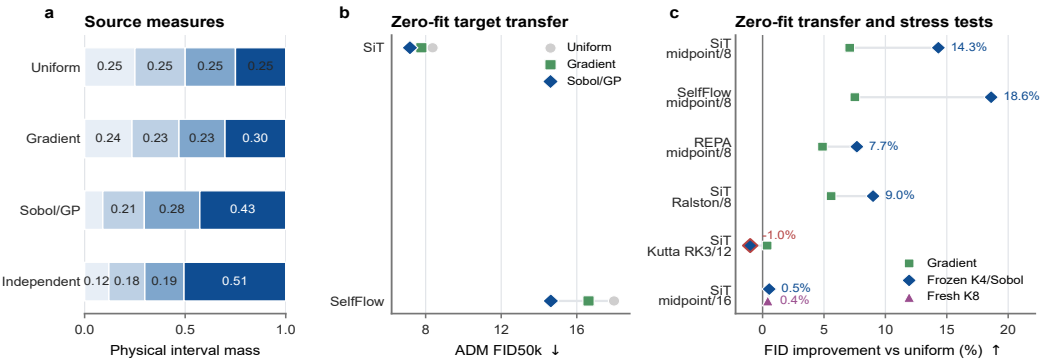

Figure 4: **Source-acquired temporal measures transfer with zero-target fitting. a**, Four source interval masses (width, not evaluation density). **b**, Canonical ADM FID50k. **c**, Relative effects across REPA and solver/budget changes; acquisition uses no target information.

Their mean cross-bank source risks are 6.8542 and 6.7824, versus 8.4468 for uniform, while the two acquisitions split their head-to-head comparison 4/4. Thus repeatability lies in the late-loaded family and its held-out advantage, not in recovery of one four-decimal mass vector.

At equal capacity, a Bernstein clock reaches source FID10k 22.7623 versus 20.0774 for Sobol/GP, but zero-target-fit SiT/Self-Flow favor the measure ($25.4282/34.0226$ versus $26.9500/36.4354$). The source clock benefit does not transfer, unlike the interval masses.

Table 3 gives the exact reads and Figure 4 the transfer and refinement effects. The frozen measure improves REPA by 7.68%; paired SiT/Self-Flow FID-reduction intervals are $[1.1088, 1.2975]$ and $[3.1486, 3.4903]$.

The target-profiled GITS-style control spends 30,720 target model-sample NFEs per receiver, whereas our acquisition spends 131,072 source model-sample NFEs once and none on targets. On matched FID10k cells at NFE = 8, the source measure lowers FID by 2.9626 on SiT and 7.4826 on Self-Flow. At NFE = 16, it lowers SiT FID by 1.1209. Profiling accesses receiver trajectories but no real images or FID, so this comparison isolates per-receiver trajectory fitting from reuse of a frozen source artifact.

Direct power ($\gamma = 1.7067$) retains 91.5%, 83.7%, and 91.8% of the flexible four-interval gain across the three targets. Both acquisitions retain additional four-interval gain on Self-Flow. Receiver-trained LD3 reaches 10.8276 on SiT after 30,400 target model-sample NFEs, versus 10.3227 for zero-target-fit Sobol/GP; neither adds a network at inference.

LD3 improves uniform by 5.80%, confirming that receiver-trained learned scheduling is effective in this cell. The original and independently acquired source measures remain lower by 0.5049 and 0.4822 FID, respectively, without the SiT teacher generation and LPIPS training used for LD3 construction.

For Sobol/GP minus direct power, paired 95% intervals exclude zero on Self-Flow for both acquisitions ($[-0.6627, -0.4626]$ and $[-1.2809, -0.8566]$) but cross zero on independent SiT and REPA. Residual four-interval value is therefore receiver-dependent rather than universal.

Table 3: **Zero-target-fit transfer and target-profiled control.** Lower is better. Bottom rows use paired generation streams; MP/Ral. denote midpoint/Ralston. Boldface marks the lowest FID in each comparable row.

Receiver	Evaluation	Uniform	Gradient	Sobol/GP	GITS-style
SiT	MP/8, FID50k	8.3674	7.7736	**7.1696**	–
Self-Flow	MP/8, FID50k	17.9714	16.6210	**14.6264**	–
SiT	Ral./8, FID10k	12.2741	11.5892	**11.1698**	–
SiT	MP/8, FID10k	11.1602	10.5158	**9.8663**	12.8289
Self-Flow	MP/8, FID10k	20.7335	19.2975	**17.1338**	24.6164
REPA	MP/8, FID10k	10.2939	9.7918	**9.5035**	–
SiT	MP/16, FID10k	9.0224	8.9056	**8.7705**	9.8914

Table 4: **Source-acquired controls versus receiver-trained LD3.** Fresh paired ADM FID10k at midpoint NFE = 8; lower is better. LD3 is adapted and trained on SiT; the other nonuniform columns are acquired without target fitting. Boldface marks the lowest FID in each receiver row.

Receiver	Uniform	Original Sobol/GP	Independent Sobol/GP	Direct power	LD3
SiT	11.4937	**10.3227**	10.3454	10.4218	10.8276
Self-Flow	20.8918	17.4640	**16.9490**	18.0225	–
REPA	10.1904	**9.3245**	9.4505	9.3955	–

5.4 Domain-local acquisition transfers across receiver seeds

One measure per domain is acquired from two source receivers and deployed on three fresh receivers, spanning ETTm1, Free Spoken Digit Dataset (FSDD) (Jackson, 2016), AG News (Zhang et al., 2015), Lorenz, Adult (Becker & Kohavi, 1996), and Covertype (Blackard, 1998).

Across ETTm1, FSDD, AG News, and Lorenz, selected measures improve primary risk in 11/12 receiver cells (mean log gain 0.3577, receiver-cell bootstrap 95% interval $[0.1042, 0.6353]$). ETTm1 and AG News improve both witnesses, FSDD is mixed, and Lorenz gains only acquisition risk. In the tabular extension, Adult and Covertype improve conditional sliced 2-Wasserstein (SW2) and projected Fréchet risk in all 6/6 cells (mean log gain 0.1225, receiver-cell bootstrap 95% interval $[0.0356, 0.2177]$).

Each domain supplies its own source receivers, objective, and frozen measure. Thus, receiver variation tests deployment reuse, whereas a domain or objective change triggers a new source-side acquisition.

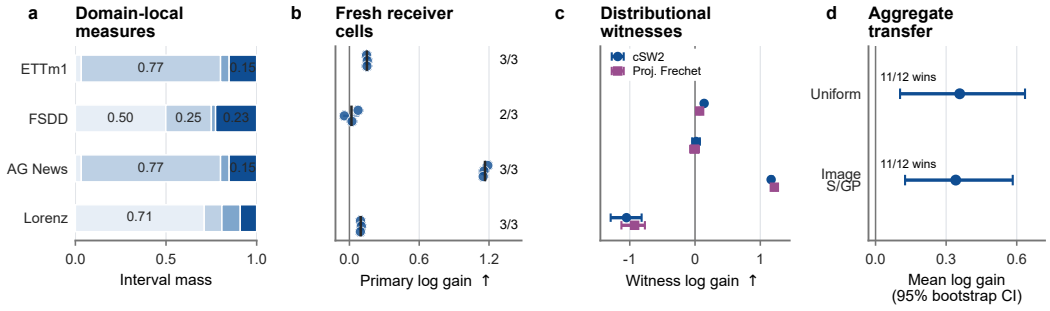

Figure 5: **Domain-local measures concentrate gains on acquisition objectives.** **a**, Masses; **b**, fresh-receiver gains; **c**, distribution witnesses; **d**, aggregate transfer.

5.5 Task-local acquisition transfers to history-conditioned forecasts

Two source receivers select one task-local measure before deployment to three fresh seeds. At midpoint NFE = 8, ensemble CRPS improves in all six cells (mean log gain 0.1195) and beats a source-Heun-defect allocation in 6/6 cells (mean log gain 0.0942). The respective receiver-cell bootstrap 95% intervals are $[0.0370, 0.2081]$ and $[0.0289, 0.1655]$; ETTm1 improves by 0.16–0.29% and short-horizon Lorenz by 19.1–22.1%.

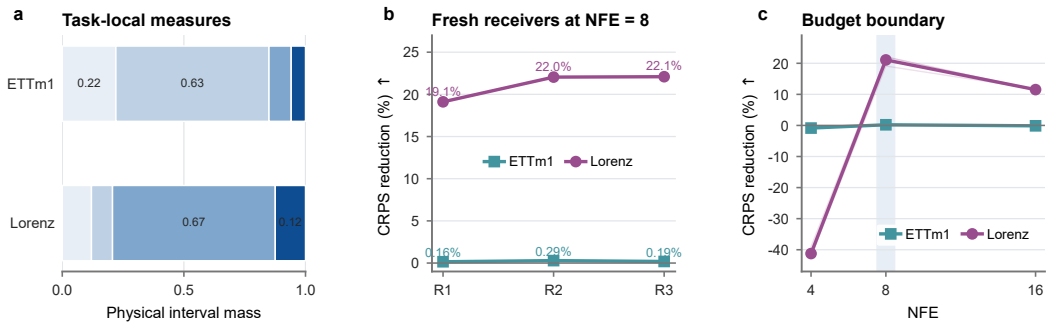

Figure 6: **Task-local measures improve history-conditioned forecasts at the acquired budget. a**, Masses; **b**, CRPS at NFE = 8; **c**, budget boundary.

At NFE = 8, all six CRPS cells improve; all three Lorenz gains remain at NFE = 16, whereas NFE = 4 changes the ordering. Neither mean-path root mean squared error (RMSE) nor distance to the uniform NFE = 128 endpoint tracks CRPS, bounding the contract to the acquired budget and probabilistic objective.

5.6 Goal-specific controls and modern-receiver checks

Goal-specific controls separate terminal quality from numerical adaptation: a dense teacher lowers endpoint MSE by 63.5% yet gives FID2k 29.3383, versus 26.7461 for uniform and 26.0756 for our measure; TORS-style rotation and SharpEuler agree.

Kutta RK3 at NFE = 12 gives SiT FID2k values of 25.7679, 25.6727, and 26.0278 for uniform, gradient, and Sobol/GP, respectively. Ralston remains positive while Kutta changes the ordering, favoring RK2 reuse.

At midpoint NFE = 16, SiT FID2k is 24.8199 for uniform, 24.6858 for the $K = 4$ subdivision, and 24.7154 for a fresh $K = 8$ acquisition. Relative to uniform, the two refinements improve FID by 0.1341 and 0.1045, respectively, and differ by 0.0296. The former adds no fitted parameters, whereas the latter reacquires seven free masses, supporting no-refit subdivision without claiming equivalence to reacquisition.

Across four fresh streams and two witnesses, all eight cells preserve ordering. Under the original and independently seeded projection witnesses, the directional derivatives are -15.98 and -18.61, respectively. The corresponding tangent-gradient norm ratios are 0.395 and 0.376. This supports discrete marginal-risk balance, not empirical validation of the continuous density law.

Together, these controls span frozen source acquisition, target trajectory profiling, and target-teacher training. The transferred measure changes neither receiver parameters nor inference NFE, so its gains do not arise from hidden target fitting or additional test-time computation.

6 Discussion and Conclusion

A four-interval temporal measure is a deployable post-training object, rather than merely a numerical grid. It is acquired once from a frozen source terminal objective and then copied without receiver trajectories, features, gradients, or additional field evaluations. The ImageNet results show that this object can survive independently trained receivers, an independent acquisition, a second RK2 tableau, and a feature-space audit.

The clock–measure comparison clarifies what transfers. A continuous clock can change finite-step source trajectories, whereas interval masses specify physical boundaries realized as solver stages. Direct-power and target-profiled controls show that the central value is transferable source-side allocation, rather than universal schedule optimality or receiver-specific fitting.

In forecasting, the same acquisition–deployment logic improves all six NFE = 8 cells across ETTm1 and Lorenz. Reuse is task-local: changing the objective, source, receiver, solver, or budget defines a new contract and calls for reacquisition.

Reproducibility Statement

The experimental protocol specifies the receiver families, solver semantics, exact NFE budgets, acquisition and evaluation splits, paired generation streams, sample counts, metrics, and bootstrap procedure. The supplementary material records the frozen interval masses, random seeds, implementation details, additional controls, and canonical artifact hashes. The accompanying anonymous code archive provides reference implementations of the temporal measure, fixed-grid solvers, projected Fréchet risk, and matched-budget Sobol/GP acquisition, together with frozen schedules, dependency requirements, a CPU smoke test, and file checksums. Third-party datasets and pretrained checkpoints are not redistributed; their sources are identified in the paper and supplementary material. All temporal measures were frozen before their corresponding target evaluations.

Ethics Statement

This work studies numerical time allocation for generative and forecasting models. It involved no new collection of human-participant data and no human-subject intervention. Experiments use established research datasets under their existing access terms, and the study does not attempt to identify or re-identify individuals. The method changes the numerical sampling schedule but does not alter the data or remove biases and misuse risks inherited from the underlying datasets and pretrained models. The authors have read and adhered to the ICLR Code of Ethics.

AI Use Statement

Generative AI tools were used to polish the manuscript and assist with code development and debugging. The authors reviewed and verified all AI-assisted text and code and retain full responsibility for the manuscript, implementation, and claims.

References

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025. URL https://www.jmlr.org/papers/v26/23-1605.html.

Barry Becker and Ronny Kohavi. Adult. UCI Machine Learning Repository, 1996. URL https://archive.ics.uci.edu/dataset/2/adult.

Roland Becker and Rolf Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001. doi: 10.1017/S0962492901000010.

Mikołaj Bińkowski, Danica J. Sutherland, Michael Arbel, and Arthur Gretton. Demystifying MMD GANs. In *International Conference on Learning Representations*, 2018.

Jock Blackard. Covertype. UCI Machine Learning Repository, 1998. URL https://archive.ics.uci.edu/dataset/31/covertype.

Vitalii Bondar. Velocity scheduled flow matching. *arXiv preprint arXiv:2607.11442*, 2026.

Graham F. Carey and Hung T. Dinh. Grading functions and mesh redistribution. *SIAM Journal on Numerical Analysis*, 22(5):1028–1040, 1985. doi: 10.1137/0722061.

Hila Chefer, Patrick Esser, Dominik Lorenz, Dustin Podell, Vikash Raja, Vinh Tong, Antonio Torralba, and Robin Rombach. Self-supervised flow matching for scalable multi-modal synthesis. In *International Conference on Machine Learning*, 2026. URL https://icml.cc/virtual/2026/poster/65011.

Defang Chen, Zhenyu Zhou, Can Wang, Chunhua Shen, and Siwei Lyu. On the trajectory regularity of ODE-based diffusion sampling. In *International Conference on Machine Learning*, 2024.

Donald Estep. A posteriori error bounds and global error control for approximation of ordinary differential equations. *SIAM Journal on Numerical Analysis*, 32(1):1–48, 1995. doi: 10.1137/0732001.

Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/559a0998fab1d19b80e7e43a5852401c-Abstract-Conference.html.

Michael B. Giles and Endre Süli. Adjoint methods for PDEs: A posteriori error analysis and postprocessing by duality. *Acta Numerica*, 11:145–236, 2002. doi: 10.1017/S096249290200003X.

Aditi Gupta, Soon Hoe Lim, Annan Yu, and N. Benjamin Erichson. Sharpen your flow: Sharpness-aware sampling for flow matching. *arXiv preprint arXiv:2605.11547*, 2026.

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*, 2017.

Yilie Huang, Wenpin Tang, and Xunyu Zhou. ART for diffusion sampling: A reinforcement learning approach to timestep schedule. *arXiv preprint arXiv:2601.18681*, 2026.

Zohar Jackson. Free spoken digit dataset. https://github.com/Jakobovski/free-spoken-digit-dataset, 2016.

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*, 2019.

Dohoon Lee, Jaehyun Park, Hyunwoo J. Kim, and Kyogu Lee. Multidimensional adaptive coefficient for inference trajectory optimization in flow and diffusion. In *International Conference on Machine Learning*, 2025.

Bengt Lindberg. Characterization of optimal stepsize sequences for methods for stiff differential equations. *SIAM Journal on Numerical Analysis*, 14(5):859–887, 1977. doi: 10.1137/0714058.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=XVjTT1nw5z.

Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2): 130–141, 1963.

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

Yunhong Min, Juil Koo, Seungwoo Yoo, and Minhyuk Sung. BézierFlow: Learning Bézier stochastic interpolant schedulers for few-step generation. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/0bd32794b26cfc99214b89313764da8e-Paper-Conference.pdf.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, et al. DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research*, 2024.

Gaurav Parmar, Richard Zhang, and Jun-Yan Zhu. On aliased resizing and surprising subtleties in GAN evaluation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11410–11420, 2022.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *IEEE/CVF International Conference on Computer Vision*, pp. 4172–4182, 2023. doi: 10.1109/ICCV51070.2023.00387.

Amirmojtaba Sabour, Sanja Fidler, and Karsten Kreis. Align your steps: Optimizing sampling schedules in diffusion models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 42947–42975, 2024. URL https://proceedings.mlr.press/v235/sabour24a.html.

Shiv Shankar and Tomas Geffner. Learning straight flows by learning curved interpolants. *arXiv preprint arXiv:2503.20719*, 2025.

Neta Shaul, Juan Perez, Ricky T. Q. Chen, Ali Thabet, Albert Pumarola, and Yaron Lipman. Bespoke solvers for generative flow models. In *The Twelfth International Conference on Learning Representations*, 2024a. URL https://openreview.net/forum?id=1PXEY7ofFX.

Neta Shaul, Uriel Singer, Ricky T. Q. Chen, Matthew Le, Ali Thabet, Albert Pumarola, and Yaron Lipman. Bespoke non-stationary solvers for fast sampling of diffusion and flow models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 44603–44627, 2024b. URL https://proceedings.mlr.press/v235/shaul24a.html.

Zheng Tan, Weizhen Wang, Andrea L. Bertozzi, and Ernest K. Ryu. STORK: Faster diffusion and flow matching sampling by resolving both stiffness and structure-dependence. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/7a641b8ec86162fc875fb9f6456a542f-Paper-Conference.pdf.

Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024.

Vinh Tong, Trung-Dung Hoang, Anji Liu, Guy Van den Broeck, and Mathias Niepert. Learning to discretize denoising diffusion ODEs. In *International Conference on Learning Representations*, 2025.

Shuchen Xue, Zhaoqiang Liu, Fei Chen, Shifeng Zhang, Tianyang Hu, Enze Xie, and Zhenguo Li. Accelerating diffusion sampling with optimized time steps. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.

Sihyun Yu, Sangkyung Kwak, Huiwon Jang, Jongheon Jeong, Jonathan Huang, Jinwoo Shin, and Saining Xie. Representation alignment for generation: Training diffusion transformers is easier than you think. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/d9e42b4d7163931f3689d6d6fbaa11d0-Abstract-Conference.html.

Liangyu Yuan, Ruoyu Wang, Tong Zhao, Dingwen Fu, Mingkun Lei, Beier Zhu, and Chi Zhang. Few-step diffusion sampling through instance-aware discretizations. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 35882–35892, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Yuan_Few-Step_Diffusion_Sampling_Through_Instance-Aware_Discretizations_CVPR_2026_paper.html.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 586–595, 2018. doi: 10.1109/CVPR.2018.00068.

Travis Zhang, Christian Belardi, Justin Lovelace, Jin Peng Zhou, Saebyeol Shin, Carla P. Gomes, and Kilian Q. Weinberger. Optimize your sampling: Tuned diffusion sampling with bayesian optimization. *arXiv preprint arXiv:2608.18040*, 2026.

Xiang Zhang, Junbo Zhao, and Yann LeCun. Character-level convolutional networks for text classification. In *Advances in Neural Information Processing Systems*, volume 28, 2015.

Bolin Zhao, Xiaoyu Zhang, Yuting Dong, Xin Lu, Wei Lin, and Qunxi Zhu. Delay flow matching. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=6lH1XblLpo.

Kaiwen Zheng, Cheng Lu, Jianfei Chen, and Jun Zhu. DPM-Solver-v3: Improved diffusion ODE solver with empirical model statistics. In *Advances in Neural Information Processing Systems*, 2023.

Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *AAAI Conference on Artificial Intelligence*, volume 35, pp. 11106–11115, 2021.

Zhenyu Zhou, Defang Chen, Siwei Lyu, Chun Chen, and Can Wang. Analyzing and improving training-free fast sampling of text-to-image diffusion models. In *Computer Vision – ECCV 2026*, volume 17012 of *Lecture Notes in Computer Science*, pp. 355–372. Springer, 2026. doi: 10.1007/978-3-032-37044-0_20. URL https://doi.org/10.1007/978-3-032-37044-0_20.

Aihua Zhu, Rui Su, Qinglin Zhao, Li Feng, Meng Shen, and Shibo He. Hierarchical schedule optimization for fast and robust diffusion model sampling. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(16):13907–13915, 2026. doi: 10.1609/aaai.v40i16.38400. URL https://ojs.aaai.org/index.php/AAAI/article/view/38400.

ACQUIRE ONCE, DEPLOY MANY: ZERO-TARGET-FIT TEMPORAL MEASURES FOR FEW-STEP FLOW MODELS SUPPLEMENTARY MATERIAL

Anonymous authors
Paper under double-blind review

1 SUPPLEMENT OVERVIEW

This supplement records the frozen protocols, artifact-level statistical design, negative-result chronology, baseline controls, controller audit, and scope limitations. It additionally reports receiver-local numerical spline selection on the official FlowTurbo and REPA ImageNet-256 checkpoints, one-parameter REPA compression, independent DINOv2 metrics, and matched-NFE latency. The final sections give the complete goal-oriented temporal-measure chain: matched-cost source acquisition, independent source replication, zero-target SiT/SelfFlow transfer, canonical target FID50k, transfer to an unseen Ralston RK2 tableau, analytic compression, exact density refinement, numerical-teacher and rotation controls, target-access costs, and canonical hashes. A subsequent preregistered extension records the failure of direct image-schedule transfer to offline-action receivers and the positive, domain-local acquisition audit on time-series, speech, text, and chaotic-system representations. We additionally record the frozen direct cube-root profile failure, finite-NFE8 mass-geometry diagnostic, and the separately acquired history-conditioned ETTm1/Lorenz forecast transfer. The order matters: the binary compatibility rule was tested and closed before the dose-aware rule was frozen; fresh DiT artifacts were trained only after the new rule was fixed; and final FID streams were not available to the selector. The later robust-consensus sequence is recorded separately: individual-schedule and gated failures precede the aggregation rule, while all fresh SelfFlow and transferred DiT evaluation streams follow the frozen median/mean construction. The temporal-measure schedule is likewise frozen on the CIFAR source before any target receiver call.

2 NOTATION AND INTERVENTION DETAILS

For a receiver field F, identity clock $I(t) = t$, learned monotone clock C, and dose λ, the deployed clock is

$$C_\lambda(t) = (1 - \lambda)t + \lambda C(t). \tag{1}$$

The dose-controller development experiments use the grid $\{0, .25, .5, .75, 1\}$; the frozen large-scale budget-policy confirmation uses only $\{0, .75, 1\}$ and performs no new dose search. A convex combination of monotone endpoint-preserving clocks is itself monotone and endpoint preserving. The intervention changes only the solver time map; receiver weights, initial noises, solver, image count, and exact NFE are held fixed.

2.1 EXECUTABLE CLOCK SEMANTICS

For the time-only clock, the implementation evaluates $v_\lambda(t, x) = C'_\lambda(t)u_\theta(C_\lambda(t), x)$. A single Heun interval $[t_n, t_{n+1}]$ is therefore

$$k_1 = C'_\lambda(t_n)u_\theta(C_\lambda(t_n), x_n), \tag{2}$$
$$\widetilde{x} = x_n + hk_1, \tag{3}$$
$$k_2 = C'_\lambda(t_{n+1})u_\theta(C_\lambda(t_{n+1}), \widetilde{x}), \tag{4}$$
$$x_{n+1} = x_n + \tfrac{h}{2}(k_1 + k_2). \tag{5}$$

Euler, midpoint, and RK4 apply the same transformation independently at every stage. Stage count, rather than interval count, defines exact NFE. This is implemented by `converted_velocity`

and `fixed_step_solve`; no solver node is silently replaced by a learned node. The sign in the converter is $(\sigma\dot{\alpha} - \alpha\dot{\sigma})/s$, consistent with TorchCFM's forward $u = x_{\text{data}} - x_{\text{noise}}$ convention.

For clarity, the finite-step claim does not identify a clocked RK operator with RK on the induced node grid. For an internal abscissa c_j, clocked RK queries $C(t_n + c_j h)$, whereas RK on the induced interval queries $C(t_n) + c_j[C(t_{n+1}) - C(t_n)]$; these are generally unequal for nonlinear C. Even when endpoint stages align, as in the Euler/Heun audit below, the clock predictor uses $hC'(t_n)$ and derivative-weighted endpoint quadrature, whereas the grid predictor uses $\Delta\tau_n$. Under the usual smoothness, consistency, and stability conditions, both operators converge to the same continuous endpoint, so their difference vanishes asymptotically. This argument predicts neither monotonicity at finite NFE nor the sign of a sample metric difference.

More explicitly, an explicit RK rule with coefficients (a_{ij}, b_i, c_i) forms clocked stages

$$Y_i^{\mathrm{c}} = x_n + h \sum_{j<i} a_{ij} C'(t_n + c_j h) u(C(t_n + c_j h), Y_j^{\mathrm{c}}), \tag{6}$$

$$x_{n+1}^{\mathrm{c}} = x_n + h \sum_{i} b_i C'(t_n + c_i h) u(C(t_n + c_i h), Y_i^{\mathrm{c}}), \tag{7}$$

whereas the induced-grid interval forms

$$Y_i^{\mathrm{g}} = x_n + \Delta\tau_n \sum_{j<i} a_{ij} u(\tau_n + c_j \Delta\tau_n, Y_j^{\mathrm{g}}), \tag{8}$$

$$x_{n+1}^{\mathrm{g}} = x_n + \Delta\tau_n \sum_{i} b_i u(\tau_n + c_i \Delta\tau_n, Y_i^{\mathrm{g}}). \tag{9}$$

The two updates coincide intervalwise when C is affine on that interval, but not in general. If C' is bounded, then $\max_n \Delta\tau_n = O(h)$.

Proof of the convergence statement. Fix x_0 and let $\Phi(x_0)$ be the common exact endpoint of the clocked and canonical-time ODEs. Smoothness, consistency, and stability give $\|\mathcal{D}_h^{\mathrm{clock}}(x_0) - \Phi(x_0)\| = O(h^p)$ and $\|\mathcal{D}_h^{\mathrm{grid}}(x_0) - \Phi(x_0)\| = O(h^p)$. Therefore

$$\|\mathcal{D}_h^{\mathrm{clock}}(x_0) - \mathcal{D}_h^{\mathrm{grid}}(x_0)\| \le \|\mathcal{D}_h^{\mathrm{clock}}(x_0) - \Phi(x_0)\| + \|\mathcal{D}_h^{\mathrm{grid}}(x_0) - \Phi(x_0)\| = O(h^p),$$

which tends to zero as $h \to 0$. This asymptotic order is not fitted to the measured NFE8/16/32 points.

2.2 Minimal selection protocol

The complete deterministic selection path is:

1. draw 128 latent noises from the recorded analysis seed;
2. generate five paired batches at $\lambda \in \{0, .25, .5, .75, 1\}$ using the frozen receiver, solver, and exact NFE;
3. extract `legacy_pytorch` Inception 2048-D features;
4. standardize each coordinate using 1,024 fixed real fit examples starting at CIFAR-10 training index 35,840;
5. project with the 64 orthonormal columns produced by reduced QR of a 2048×64 standard-Gaussian matrix with seed 11,991;
6. compute empirical Gaussian Fréchet distance to the following 2,048 disjoint real examples and select the smallest minimizing dose.

The final CleanFID10k stream, its generated images, and its feature moments are never read by the selector. The projected risk is deliberately Inception directed, so this is a no-final-FID-feedback protocol, not a metric-agnostic controller. Dimension 64 is inherited from the frozen post-training witness; it was not selected using confirmation FID. The multi-feature and resampling audits below quantify its sensitivity rather than asserting universal feature agreement.

We use "quality-oriented gain" so that positive values always denote an improvement. For FID and KID this is baseline minus intervention; for precision and recall it is intervention minus baseline.

Unless a section states otherwise, intervals are two-sided artifact-level t intervals with $n = 5$. The fresh budget-policy confirmation uses $n = 10$ learned schedules after averaging its two streams within each schedule; grid-transfer audits use eight paired generation streams. Images are never treated as independent replicates.

3 Artifact map and frozen streams

Table 1: Artifact families used by the controller experiments.

Family	Purpose	Training seeds	Analysis/generation seeds
Corrected DiT	mechanism/development	20268201–20268205	frozen paired streams
Fresh corrected DiT	prospective confirmation	20268401–20268405	20280001–05 / 20281001–05
TorchCFM original-FM	initial controls	20267511–20267515	20282001–05
Fresh TorchCFM	prospective confirmation	20268701–20268710	20287001–10 / 20288001–10
SelfFlow fresh fits	consensus development	20269501–08 / 20269901–08	20269701–08 / 20270101–08
SelfFlow consensus	prospective evaluation	accepted 20269901–08	20270201–08
Fresh SelfFlow dose	modern prospective selection	20270801–08	20271801–08 / 20270901–08
Fresh SelfFlow policy	frozen budget-policy confirmation	20271401–10	20271501–10 / 20271601–10
Analytic source grid	generic-grid control	frozen SelfFlow median	existing paired SelfFlow/DiT streams
Corrected DiT grid transfer	zero-fit transfer	frozen 1M EMA	20270301–08 / 20270401–08 / 20270501–08
FlowTurbo receiver spline	numerical-only local selection	frozen official predictor	20272801–02 / 20272841–48 / 20272981–82
REPA receiver spline	independent local selection	frozen official 4M checkpoint	20273001–02 / 20273041–48 / 20273101–04
REPA analytic grid	L_2 cubic compression and replication	$a = -.2655287$	20273041–48 / 20273101–04 / 20273301–08 / 20273401–04
Goal measure targets	canonical zero-fit ADM FID50k	frozen source measures	20315201–02
SiT Ralston transfer	unseen RK2 zero-fit ADM FID10k	frozen source measures	20315210–11

The main corrected-DiT factorial uses 10,000 generated samples per arm and six metrics. The full 5×5 transplant adds 20 off-diagonal cells and reuses five matched cells. The TorchCFM parent study contains 80 frozen cells across NFE8/12, field states, clock doses, and metrics. The DiT controller confirmation contains five proxy curves and 40 logical FID arms; exact aliases are used when selected and full dose coincide. Dense-dose development and confirmation contribute 410 cells each. Solver/NFE development and confirmation contribute 600 cells each. The fresh TorchCFM experiment uses ten independent proxy streams followed by 80 fixed 10k-FID cells. Robust-consensus evidence adds 16 fresh SelfFlow schedules, five failed-seed replications, 24 fresh SelfFlow consensus FID cells, and 120 transferred DiT FID cells across five budgets. The final modern-dose cohort adds eight completely new SelfFlow fits, sixteen disjoint five-dose proxy curves, and 80 untouched NFE8/32 ADM FID10k cells. The subsequent budget-policy confirmation adds ten new fits and 120 final cells: two independent streams, two budgets, and three frozen arms per fit. The generic-grid control adds 48 source-only analytic-grid cells on the same paired SelfFlow and DiT streams used by the uniform/median comparison. The modern-receiver audit adds two numerical-only nine-candidate selection matrices, eight FlowTurbo confirmation streams, eight plus four REPA cubic streams, 24 first-cohort analytic cells, 24 prospective matched-analytic cells, and independent DINOv2 and matched-NFE latency evaluations. The FlowTurbo and REPA sign counts use paired generation streams; neither count is presented as independent receiver replication. The canonical goal-measure extension adds six FID50k cells across SiT and SelfFlow, followed by one exact-parity check and three FID10k cells under the unseen Ralston RK2 tableau.

4 Causal factorial results

Table 2: NFE8 quality-oriented CleanFID clock effects.

Architecture	Receiver	Mean gain	95% interval	Positive
Corrected DiT	field-only	-2.040	$[-3.767, -.313]$	0/5
Corrected DiT	joint	31.531	$[30.407, 32.655]$	5/5
TorchCFM	field-only	5.457	$[5.113, 5.801]$	5/5
TorchCFM	joint	5.946	$[5.371, 6.521]$	5/5

For corrected DiT, the factorial interaction is 33.57 $[32.13, 35.01]$, 5/5 positive. For TorchCFM it is 0.489 FID at NFE8 and 0.057 $[0.012, 0.102]$ at NFE12. This difference is why the paper claims field-state modulation rather than universal co-training dependence.

5 DOSE RESPONSE AND FINITE-STEP EVIDENCE

Table 3: Joint-field clock gain contracts with exact NFE.

Architecture	NFE8	NFE12	NFE24
Corrected DiT	31.531	21.731	2.413
TorchCFM	5.946	1.507	–

At the primary budget, the five-point joint-field FID homotopies are monotone for every artifact. Corrected-DiT field-only is nonmonotone: the mean quality-oriented FID gain is positive at partial dose but negative at full dose. This empirical shape motivates a controller over dose rather than a binary endpoint decision. The NFE contraction supports a finite-solver interpretation; it does not constitute a theorem about sample metrics.

Table 4: Quality-oriented precision/recall effects of the joint clock.

Architecture	NFE	Precision	Recall
Corrected DiT	8	-0.01764	$+0.02614$
Corrected DiT	12	-0.01472	$+0.02838$
Corrected DiT	24	-0.00186	$+0.00278$
TorchCFM	8	-0.00246	$+0.04122$
TorchCFM	12	-0.00036	$+0.01548$

Recall is positive for all five artifacts at NFE8/12 in both architectures; precision is small or negative. We therefore describe the clock as mainly repairing finite-step coverage, not as improving FID through mode contraction.

6 DENSE DOSE LANDSCAPES AND ORACLE REGRET

We evaluate all 41 doses $\lambda = 0, .025, \ldots, 1$ for both receiver fields and all five artifacts in a development cohort, freeze global baselines, and repeat the same 410-cell matrix on an untouched cohort. Development chooses the field-state doses .475 for field-only and .95 for joint. The fresh artifact optima are .625, .475, .600, .475, .575 and .95, .95, .90, .95, .95, respectively.

Table 5: Mean dense-FID oracle regret on the untouched DiT cohort. Lower is better. Every frozen rule is evaluated against each artifact's 41-dose oracle.

Receiver	Identity	Full	Architecture-global	Five-dose proxy	Field-state
field-only	6.695	8.060	3.285	.855	**.115**
joint	32.684	.301	2.013	.301	**.098**

The coarse five-dose oracle itself has only $.196/.301$ mean regret. The dense experiment therefore sharpens the conclusion from "dose can vary" to "most useful variation is predictable from receiver field state." It also explains why the sole .75 choice in the later TorchCFM cohort is not evidence for a special artifact: the selection noise is smaller than the stable regime-level full-dose effect.

The 17-point set is the nearest distinct approximately uniform subset of the 41 actually measured doses; no interpolation or FID fitting is used. This analysis concerns the frozen field-state rule, not a post-hoc replacement of the prospective per-artifact selector.

7 SOLVER, NFE, AND CONTINUOUS-LIMIT CONFIRMATION

The mechanism matrix crosses two receiver fields, three clock actions (identity, selected, full), four solvers (Euler, midpoint, Heun, RK4), five NFE budgets, and five artifacts. Development and confirmation each contain 600 canonical FID cells.

Table 6: Frozen field-state grid resolution. Doses are selected from development mean curves and evaluated by mean oracle regret on the untouched cohort.

Measured doses	Cost vs. 5	Field-only dose/regret	Joint dose/regret
5	$1.0\times$	$.50/.206$	$1.00/.301$
9	$1.8\times$	$.50/.206$	$1.00/.301$
17	$3.4\times$	$.50/.206$	$.95/.098$
41	$8.2\times$	$.475/.115$	$.95/.098$

Table 7: Joint full-clock CleanFID gain across exact NFE. Development and confirmation were fixed, independent five-artifact cohorts.

Cohort	Solver	8	16	24	32	64
development	Heun	30.840	11.109	2.443	.422	$-.139$
	RK4	7.196	7.118	3.099	1.237	.338
confirmation	Heun	31.859	11.489	2.497	.384	$-.125$
	RK4	7.675	7.279	3.115	1.272	.354

Euler is harmful in this regime; midpoint, Heun, and RK4 expose different finite-step utilities. The effect is therefore solver-conditioned rather than a universal preference for one nonuniform grid. Independent endpoint matrices compare identity and full-clock maps against high-accuracy RK4 references. The fresh cohort passes the continuous-equivalence gate directly at RK4-512 with maximum normalized RMS 8.7859×10^{-4}; all mean slopes and 39/40 receiver-level slopes contract. In development, one cell is narrowly above 10^{-3} at RK4-512 (1.0486×10^{-3}); the frozen adjudication at RK4-1024 reduces the maximum to 1.1815×10^{-4}, with 4.37–11.59 fold contraction for every full-clock discrepancy.

The cached CleanFID real moments are a computational optimization only. Five preselected cells reconstruct direct CleanFID within maximum absolute error 5.5034×10^{-5} under the recorded 10^{-4} cache-equivalence tolerance; this tolerance is not used to select a scientific endpoint.

7.1 Controller transfer across deployment regimes

We froze independent five-dose calibration curves for Heun-NFE8, Heun-NFE16, RK4-NFE8, and RK4-NFE16 before generating a complete 200-cell FID oracle matrix. Table 8 contrasts matched calibration with mechanically transferring the original Heun-NFE8 selector. Each number is the mean excess FID over the evaluation regime's measured five-dose oracle across five receivers.

Table 8: Five-dose oracle regret for deployment-regime calibration.

Evaluation	Field	Matched	Heun-8 transfer
Heun-8	field-only	.386	.386
	joint	.340	.340
Heun-16	field-only	.000	2.589
	joint	.000	.678
RK4-8	field-only	.000	1.039
	joint	.000	.362
RK4-16	field-only	.413	.257
	joint	.000	.285

The selected dose changes materially with the receiver–solver–budget block: field-only chooses mean dose $.45$ at Heun-8, 1 at Heun-16, 0 at RK4-8, and $.60$ at RK4-16; joint chooses $.95, 1, 1, 1$. This prospective matrix turns the solver-conditioned observation into an actionable rule: calibrate under the intended deployment solver and budget. The single RK4-16 field-only reversal is retained and quantifies the finite-sample limitation of the 128-noise witness.

8 CLOCK DYNAMICS VERSUS THE MATCHED INDUCED GRID

For each frozen selected clock C_λ, the matched grid uses $\tau_k = C_\lambda(k/K)$ but integrates the original canonical field with step $\Delta\tau_k$. The clock arm uses uniform physical steps and evaluates $C'_\lambda(t)u(C_\lambda(t), x)$ at every RK stage. This completed comparison is restricted to Euler and Heun: canonical left/end-point query times align, but predicted stage states and quadrature weights generally do not. It does not assert matched midpoint/RK4 internal stages. Receiver, noise, dose, solver, NFE, feature witness, and sample count are identical. Table 9 reports quality-oriented projected-risk gains; positive clock-versus-grid values favor clock dynamics.

Table 9: Five-receiver matched clock-induced-grid audit. "Clock–id." and "grid–id." are gains over identity; C–G is clock gain over its matched grid.

Field	Solver–NFE	Clock–id.	Grid–id.	C–G	C wins
field-only	Euler–8	$-.696$	$-.012$	$-.684$	0/5
field-only	Euler–16	$-.029$	$+.026$	$-.055$	1/5
field-only	Euler–32	$-.053$	$+.040$	$-.093$	0/5
field-only	Heun–8	$+2.232$	$-.597$	$+2.829$	5/5
field-only	Heun–16	$+.740$	$-.093$	$+.834$	5/5
field-only	Heun–32	$+.027$	$-.011$	$+.038$	4/5
joint	Euler–8	-1.549	$+.032$	-1.581	0/5
joint	Euler–16	$+.161$	$+.035$	$+.126$	4/5
joint	Euler–32	$+.033$	$+.001$	$+.032$	2/5
joint	Heun–8	$+11.129$	$-.915$	$+12.044$	5/5
joint	Heun–16	$+3.463$	$-.280$	$+3.743$	5/5
joint	Heun–32	$+.094$	$-.362$	$+.456$	5/5

The endpoint RMS separation also contracts with refinement: for selected field-only/joint Heun clocks it is $.02893/.06007$ at NFE8, $.00765/.01551$ at NFE16, and $.00195/.00301$ at NFE32. The Heun result rejects the hypothesis that the effect is exhausted by placing stages at $C_\lambda(t)$; the Euler reversal simultaneously rejects a solver-agnostic clock advantage. The ten-cell CleanFID confirmation was frozen after this proxy separation became visible and is labeled post-proxy rather than preregistered discovery evidence. It nevertheless provides a direct final-metric check on the mechanism: the matched grid is worse than identity by 1.540 FID for field-only and 2.017 for joint receivers, with 0/5 wins in both fields. Selected clock dynamics beats the same grid by 7.671 and 33.876 FID, both 5/5. Table 10 reports every paired artifact; the identity, clock, and independently optimized-schedule columns are immutable cells from the strong-baseline matrix.

Table 10: Post-proxy CleanFID10k confirmation at Heun-NFE8. Lower is better; all methods use the same 10,000-noise stream within each row.

Field	Seed	Dose	Identity	Induced grid	Clock	Opt. schedule
field-only	20268401	.50	64.532	66.626	58.236	60.404
field-only	20268402	.50	64.747	66.941	58.866	64.626
field-only	20268403	.50	66.720	67.672	59.867	62.746
field-only	20268404	.50	65.755	68.007	59.242	59.896
field-only	20268405	.75	66.122	66.328	61.009	59.057
joint	20268401	1.00	70.828	72.375	39.591	44.441
joint	20268402	1.00	69.963	72.601	37.750	38.762
joint	20268403	1.00	71.244	72.721	40.086	41.704
joint	20268404	1.00	70.761	73.279	38.337	43.285
joint	20268405	1.00	71.437	73.340	39.175	43.887
field-only mean	–	–	65.575	67.115	59.444	61.346
joint mean	–	–	70.847	72.863	38.988	42.416

9 STRONG TRAINING-FREE RUNTIME BASELINES

The schedule-search comparison uses a deterministic 81-candidate bank and the same 128-noise projected-risk calibration budget. SharpEuler uses its released schedule builder and a 1,024-path, 100-step field-variation profile. Table 11 reports paired means; Heun and Euler are kept in separate blocks because identical NFE does not make different solvers a causal schedule comparison.

Table 11: Corrected-DiT runtime baseline means at exact NFE8.

Solver	Comparison (left vs. right)	Left	Right	Left wins
Heun	schedule vs. uniform, field-only	61.346	65.575	5/5
Heun	clock vs. schedule, field-only	59.444	61.346	3/5
Heun	schedule+clock vs. schedule, field-only	55.735	61.346	5/5
Heun	schedule vs. uniform, joint	42.416	70.847	5/5
Heun	clock vs. schedule, joint	38.988	42.416	4/5
Heun	schedule+clock vs. schedule, joint	40.077	42.416	4/5
Euler	SharpEuler vs. uniform, field-only	40.259	31.264	0/5
Euler	SharpEuler vs. uniform, joint	27.930	25.652	0/5

Schedule+clock beats clock-only in 3/5 receivers for each field, but its mean is 3.709 better for field-only and 1.089 worse for joint. We therefore do not describe schedule and clock as orthogonal. The narrower supported claim is that clock dose remains competitive with an independently optimized schedule and can add value on top of it, while their interaction is receiver-dependent.

10 CLOCK TRANSPLANT AND FAMILY GEOMETRY

The corrected-DiT 5×5 matrix fixes each receiver row and replaces only the donor clock. Table 12 reports matched minus the mean of unmatched donors.

Table 12: Corrected-DiT clock transplant.

Metric	Matched advantage	95% interval	Assignment p
Inception FID	-0.0412	$[-.1351, .0527]$	.900
CLIP-FID	$+0.0083$	$[-.0364, .0529]$	.300
DINOv2 FD	-0.4194	$[-2.0274, 1.1886]$	.875
DINOv2 KID	-0.00053	$[-.00321, .00214]$	.675
Precision	-0.00094	$[-.00383, .00195]$	.900
Recall	$+0.00073$	$[-.00697, .00843]$	.392

No metric resolves a seed-specific diagonal advantage. TorchCFM geometry gives the same conclusion: between-seed clock RMS is 0.000399, compared with mean warp RMS 0.008542. A receiver-excluded mean clock is unresolved from the individual clock. The learned object is a shared regime-level clock family.

11 LEARNED VERSUS ANALYTIC CLOCK

Without reading quality, we project the mean learned TorchCFM clock onto $C_a(t) = t + at(1-t)$. The fitted coefficient is $a = -0.0462302$; the minimum slope is 0.95377 and RMSE to the mean learned clock is 0.001327.

Table 13: Quality-blind analytic clock-family control at NFE8 (CleanFID10k).

Receiver	Identity	Analytic	LOO mean	Individual
field-only	28.362	24.370	22.969	22.904
joint	29.287	24.833	23.464	23.341

The analytic clock improves identity by 3.992 and 4.454 FID, both 5/5. Individual learned clocks further improve analytic by 1.465 $[1.084, 1.846]$ and 1.492 $[1.076, 1.909]$, both 5/5. Independent VGG-KID and recall reproduce the learned-over-analytic hierarchy. This baseline shows that generic allocation is important while preserving a learned-shape effect.

12 Negative binary rule and matched-dose diagnosis

The first compatibility hypothesis used

$$s(F, C) = \frac{R(F, I) - R(F, C_{.25})}{.25}, \qquad s > 0 \Rightarrow \text{enable } C_1. \tag{10}$$

It was frozen before evaluation and required all four architecture–field signs to agree with known endpoint FID.

Table 14: Preregistered binary compatibility result.

Architecture	Receiver	Score mean	Full-clock FID gain	Correct
Corrected DiT	field-only	4.889	-2.040	no
Corrected DiT	joint	12.732	31.531	yes
TorchCFM	field-only	1.314	5.457	yes
TorchCFM	joint	1.336	5.946	yes

The gate fails at 3/4 and is permanently closed; no threshold, projection, or classifier was tuned. A separately frozen matched-dose diagnostic then tests whether the proxy is wrong or the endpoint extrapolation is wrong.

Table 15: Corrected-DiT field-only matched-dose diagnosis.

Quantity	Mean gain	95% interval	Positive
Projected risk at $\lambda = .25$	1.283	$[.845, 1.721]$	5/5
CleanFID at $\lambda = .25$	3.522	$[2.897, 4.147]$	5/5
CleanFID at $\lambda = 1$	-2.040	$[-3.767, -.313]$	0/5

The proxy is locally correct. The failure is nonlinear dose extrapolation. Importantly, the development dose curve exposed by this diagnostic is not used for the final confirmation; a new protocol and fresh artifacts are required.

13 Controller protocol and prospective result

For each receiver, the controller uses 128 fixed noises, five doses, exact Heun NFE8, a frozen Inception feature extractor, and the inherited 64-dimensional projection bank. It selects the smallest projected-risk minimizer. The selection JSON is written and hashed before any final FID job starts.

Table 16: Fresh corrected-DiT confirmation. Positive values favor selected.

Receiver	Selected doses	vs. identity	vs. fixed .25	vs. full
field-only	$.5, .5, .5, .5, .75$	5.756	1.794	7.252
95% CI		$[4.108, 7.404]$	$[.154, 3.434]$	$[5.584, 8.920]$
joint	$1, 1, 1, 1, 1$	32.411	24.708	0
95% CI		$[31.641, 33.181]$	$[23.823, 25.593]$	$[0, 0]$

The field-only selected FIDs are $58.160, 59.568, 60.063, 59.752, 62.130$; the corresponding identity FIDs are $64.676, 65.328, 66.423, 66.432, 65.594$. Four of five selected arms beat fixed .25, while all five beat identity and full. Joint selected and full are exact aliases because all selected doses are one.

For the final TorchCFM test, ten new seeds produce 20 independently trained field-only/joint artifacts. Their manifest is sealed before ten independent proxy streams are evaluated, and the selection file is

sealed before any final FID generation. Field-only selects full dose in 10/10; joint selects full in 9/10 and .75 once.

Table 17: Prospective fresh TorchCFM confirmation (CleanFID10k, NFE8).

Receiver	Identity	Selected	Gain vs. identity	Gain vs. fixed .25
field-only	28.8969	23.4114	5.4855 [5.3063, 5.6647]	3.8873 [3.6853, 4.0892]
joint	29.0183	23.2068	5.8114 [5.4035, 6.2194]	3.7584 [3.4374, 4.0793]

Both identity comparisons are positive for 10/10 independently trained pairs. The sole .75 choice is a calibrated-proxy near-tie: its proxy advantage over full is .0467, whereas final FID favors full by .6573. Consequently the scientific conclusion is prospective matched-clock generalization and a stable TorchCFM full-dose regime, not artifact-specific dose diversity.

14 Feature-space validation and Pareto structure

Five independent calibration artifacts contain 20 resampled banks at each of six frozen sample counts from 16 to 512. Controller uncertainty decreases with sample count. A development-only feature selector freezes Inception before confirmation and obtains mean dense-Inception-FID regret .240 in development and .461 on untouched artifacts. Exact reconstructions from original CIFAR-10 examples verify the frozen Inception, DINOv2-L, and CLIP real references.

The selected clocks improve CLIP-FID but often worsen DINOv2 FD/KID, with mixed precision/recall. This is useful mechanistic evidence: finite-step time control moves the generated distribution along a feature-dependent Pareto surface rather than optimizing every representation simultaneously. A DINO-selected clock can improve the DINO objective, confirming that the discrepancy is metric geometry rather than a broken evaluator. The raw-artifact audit passes all 1,013 provenance, selection, reconstruction, and regret checks.

The projection-dimension audit freezes one new 128-column Gaussian matrix, applies reduced QR once, and uses its first 32, 64, or 128 columns. The three subspaces are therefore exactly nested. This audit does not claim elementwise identity with the earlier standalone 64D witness; it tests dimensional sensitivity within a separately frozen nested bank. Across five artifacts and 20 fresh calibration banks per artifact, the selections are:

Table 18: Nested projection-dimension sensitivity (100 banks per field).

Field	Dimension	Dose histogram	Agreement with 64D	Mean $\lvert\Delta\lambda\rvert$
field-only	32	.25:2, .5:68, .75:30	.68	.080
field-only	64	.25:4, .5:88, .75:8	1.00	0
field-only	128	.5:52, .75:48	.56	.110
joint	32	.75:2, 1:98	.98	.005
joint	64	1:100	1.00	0
joint	128	1:100	1.00	0

This is strong robustness for the joint full-dose regime and weaker robustness for the exact field-only interior action. All field-only disagreements remain within the interior .25–.75 band and the mean shifts are below half a grid step. We therefore retain 64D prospectively and do not claim dimension-invariant per-artifact doses. The original one-factor projection audit is complemented by the full feature-family $\times$ dimension reconstruction already present in the frozen multi-feature canonical artifact.

These cells close the earlier interaction gap. They also sharpen the claim: dimension changes are secondary within a fixed feature family, whereas the feature family can change the selected regime. Inception is the predeclared target geometry; agreement with itself is therefore a reference check rather than evidence of feature invariance. CLIP preserves the joint/full regime but not the exact field-only interior dose, while DINOv2 defines a different Pareto direction. We accordingly describe the controller as target-calibrated and do not advertise a feature-agnostic selector.

Table 19: Feature-family × projection-dimension interaction on the confirmation cohort. Each entry is agreement with the frozen Inception action at 32/64/128D; parentheses give mean $|\Delta\lambda|$ and final primary Inception-FID regret, averaged over five receivers and 20 banks per receiver.

Feature	Field	32D	64D	128D
Inception	field-only	1.00 (.185, 2.338)	1.00 (.095, .779)	1.00 (.089, .544)
	joint	1.00 (.075, .677)	1.00 (.067, .467)	1.00 (.065, .377)
CLIP	field-only	.42 (.223, 3.425)	.23 (.210, 3.154)	.11 (.200, 2.824)
	joint	.86 (.060, .301)	.94 (.060, .301)	.97 (.060, .301)
DINOv2	field-only	.05 (.433, 5.076)	.07 (.480, 5.637)	.03 (.520, 6.225)
	joint	.09 (.560, 18.623)	.04 (.606, 20.552)	.00 (.590, 20.366)
Ensemble	field-only	.65 (.168, 2.205)	.57 (.142, 1.866)	.53 (.118, 1.510)
	joint	.68 (.103, 1.219)	.55 (.125, 1.625)	.39 (.144, 1.829)

15 SelfFlow ImageNet-256 dose scale-up

The scale audit uses the shared official SelfFlow checkpoint and five frozen fitted schedules, while the 35-cell node-dose homotopy is specific to the present runtime study. Ten NFE8 endpoint cells are exact aliases with recorded source hashes; 15 interior-dose NFE8 cells and ten NFE32 endpoints are newly generated and scored on the same paired streams.

Table 20: SelfFlow SiT-XL/2 ImageNet-256 mean ADM FID10k.

NFE	$\lambda = 0$	.25	.5	.75	1
8	36.9573	36.9604	36.9687	36.9726	36.9229
32	14.0176	–	–	–	14.0421

At NFE8, full dose beats identity in 5/5 pairs by .0344 in the mean; four artifacts have full as the measured five-dose optimum and one chooses .25. At NFE32, full loses to identity in 5/5 by .0245. The dose curve is shallow and not monotone in its interior, so it does not support transferring the CIFAR controller or claiming a practically large ImageNet gain. It does independently support the finite-budget boundary: a clock effect detectable at NFE8 can contract and reverse after solver refinement on a 675M conditional model.

15.1 Fully prospective five-dose confirmation

Eight new schedules (training seeds 20270801–20270808) are fitted without access to any result in the earlier scale cohort. Independent calibration seeds 20271801–20271808 evaluate doses $0, .25, .5, .75, 1$ separately at NFE8 and NFE32, freezing 16 actions before generation seeds 20270901–20270908 produce all 80 paired ADM FID10k cells. Table 21 reports every selected action and its five-dose oracle regret.

Across NFE8, the mean curve for doses $0, .25, .5, .75, 1$ is 36.9983, 37.0756, 36.9161, 36.8307, and 36.8490. Dose .75 is the global mean optimum, improves identity by .1676 with 6/8 wins, and has paired 95% CI $[.0110, .3242]$. The frozen selector improves by .0845 with 5/8 wins, but is .0648 worse than full dose and therefore fails the predeclared gate. At NFE32 the respective means are 13.9350, 13.9821, 14.0109, 14.0005, and 14.0261; every nonzero dose has only 1/8 wins and a negative paired interval. The selector loses .0378 on average. This separates a genuine low-budget schedule effect from an insufficiently resolved artifact-level proxy and from high-budget overcorrection.

15.2 Frozen ten-artifact budget-policy confirmation

The eight-fit result above is treated as development evidence. We freeze $\pi(8) = .75$ and $\pi(32) = 0$ before fitting training seeds 20271401–20271410. Each fit is evaluated at doses $0, .75, 1$ and exact Heun NFE8/32 on two new streams; the two streams are averaged before inference, so the statistical unit is the learned schedule ($n = 10$), not a stream or image. All 120 cells are retained. Table 22 reports the artifact means.

Table 21: Full prospective SelfFlow dose result. Gain is identity minus the selected action; oracle regret is selected minus the best measured dose.

Fit	NFE	Selected	Identity	Selected FID	Full FID	Gain	Oracle dose/regret
20270801	8	1	37.110	37.170	37.170	−.060	.5/.087
20270802	8	0	37.277	37.277	37.535	.000	.75/.009
20270803	8	.5	37.461	37.330	37.016	+.130	1/.315
20270804	8	1	37.136	37.176	37.176	−.039	0/.039
20270805	8	.5	36.690	36.367	36.127	+.324	1/.240
20270806	8	.75	36.756	36.526	36.484	+.230	1/.043
20270807	8	.5	37.116	37.088	36.908	+.028	1/.180
20270808	8	1	36.440	36.377	36.377	+.062	.75/.197
20270801	32	.75	14.303	14.322	14.392	−.019	.25/.031
20270802	32	0	13.772	13.772	13.950	.000	0/.000
20270803	32	.75	13.932	14.035	14.045	−.103	0/.103
20270804	32	.75	13.940	13.988	13.954	−.048	0/.048
20270805	32	.25	13.744	13.782	13.734	−.038	1/.049
20270806	32	.25	13.960	14.012	14.090	−.052	0/.052
20270807	32	.5	13.882	13.924	13.972	−.041	0/.041
20270808	32	0	13.946	13.946	14.072	.000	0/.000

Table 22: Full frozen SelfFlow budget-policy confirmation. Gain is identity minus dose .75, after averaging two paired ADM FID10k streams per artifact.

Fit	NFE	Identity	Dose .75	Full	.75 gain
20271401	8	37.235	37.205	37.366	+.029
20271402	8	36.769	36.543	36.732	+.226
20271403	8	36.930	36.676	36.906	+.254
20271404	8	37.158	36.812	36.861	+.347
20271405	8	37.432	36.936	36.966	+.496
20271406	8	36.828	36.828	36.828	+.000
20271407	8	36.250	36.250	36.250	+.000
20271408	8	37.157	36.968	37.193	+.189
20271409	8	36.759	36.520	36.629	+.239
20271410	8	37.107	37.107	37.107	+.000
20271401	32	14.224	14.274	14.289	−.050
20271402	32	13.930	13.980	14.101	−.051
20271403	32	14.088	14.197	14.266	−.109
20271404	32	13.934	14.012	14.014	−.078
20271405	32	14.297	14.354	14.377	−.058
20271406	32	13.943	13.943	13.943	+.000
20271407	32	13.834	13.834	13.834	+.000
20271408	32	14.140	14.179	14.237	−.038
20271409	32	14.048	14.106	14.228	−.058
20271410	32	14.107	14.107	14.107	+.000

At NFE8, identity, .75, and full average $36.9626, 36.7845, 36.8838$. The primary .75 gain is $+.1781$ with artifact-level 95% CI $[+.0568, +.2994]$ and $7\mathrm{W}/0\mathrm{L}/3\mathrm{T}$. Dose .75 also beats full by .0993 $[.0311, .1675]$. At NFE32 the respective means are $14.0545, 14.0987, 14.1397$: the predeclared .75 safety contrast is $-.0442$ $[-.0700, -.0184]$ with $0\mathrm{W}/7\mathrm{L}/3\mathrm{T}$, and full loses .0852 $[-.1372, -.0332]$. Fits 20271406, 20271407, and 20271410 selected validation iteration zero, so all doses are exactly identity and appear as ties rather than missing or excluded results. The other seven fits are 7/7 positive at NFE8 and 7/7 negative at NFE32; each of those seven nontrivial fits also has positive gain on all three pre-FID audit streams, whereas the three iteration-zero fits have three exact audit ties. Both primary intervals therefore satisfy the strong frozen confirmation rule and justify identity deployment at NFE32.

16 Robust SelfFlow consensus and zero-fit DiT transfer

16.1 Failure chronology before aggregation

The corrected direct-grid trainer uses 384 SPSA updates, batch 128, four fixed 64-D ADM-Inception witnesses, one disjoint 1,024-sample validation stream, and three 1,024-sample audit streams. The field, VAE, extractor, preprocessing, and exact-NFE8 Heun solver are frozen. Table 23 reports the first fully fresh eight-fit cohort. Seven schedules improve paired ADM FID10k and mean gain is $+.273$, but seed 20269504 violates the preregistered safety condition.

Table 23: First fresh SelfFlow individual-fit cohort. Positive FID gain favors the selected schedule.

Fit seed	Iteration	Validation gain	Audit wins	FID gain
20269501	288	.259	2/3	$+.567$
20269502	288	.132	2/3	$+.433$
20269503	160	.050	3/3	$+.024$
20269504	96	.009	3/3	$-.138$
20269505	352	.193	3/3	$+.659$
20269506	256	.338	3/3	$+.236$
20269507	128	.036	2/3	$+.215$
20269508	128	.178	3/3	$+.191$

The failed schedule was frozen and evaluated on five new generation seeds. It loses 0/5, with mean/median gains $-.082/-.097$; the original failure is therefore a repeatable proxy–FID mismatch rather than an isolated generation stream.

Table 24: Replication of failed SelfFlow fit 20269504.

Generation seed	Identity FID	Selected FID	Gain
20269801	37.465	37.488	$-.023$
20269802	36.822	36.840	$-.018$
20269803	37.727	37.890	$-.163$
20269804	37.083	37.181	$-.097$
20269805	37.485	37.594	$-.109$

A second prospective cohort freezes a deployment gate before training: accept only when validation gain is at least .03 and at least two of three audit streams improve. The gate accepts five schedules and abstains exactly to identity on three. Four accepted schedules improve and active mean gain is $+.141$, but seed 20269907 loses .126 despite 3/3 audit wins. The overall eight-stream mean, including exact-zero abstentions, is $+.088$.

Table 25: Frozen-gate SelfFlow confirmation.

Fit seed	Validation gain	Audit wins	Decision	FID gain
20269901	.378	3/3	accept	$+.267$
20269902	.289	1/3	abstain	.000
20269903	.263	1/3	abstain	.000
20269904	.307	2/3	accept	$+.191$
20269905	.120	3/3	accept	$+.240$
20269906	.179	2/3	accept	$+.131$
20269907	.072	3/3	accept	$-.126$
20269908	.182	1/3	abstain	.000

16.2 Frozen construction and prospective SelfFlow result

The five proxy-accepted schedules above define the aggregation set before their FID values are used. Each three-logit vector is centered to remove the softmax gauge. Coordinatewise median

followed by re-centering is primary; arithmetic mean is a sealed ablation. Their logits are respectively $[.09614, -.01472, -.08142]$ and $[.08986, -.03325, -.05660]$, producing NFE8 nodes $[0, .26756, .52233, .76594, 1]$ and $[0, .26557, .51859, .76232, 1]$. Neither construction reads FID.

On eight new generation streams, both aggregations win 8/8. Median mean gain is $+.3902$ (minimum $+.2143$); mean-consensus gain is $+.3100$ (minimum $+.1560$). Median beats mean in all eight paired cells.

Table 26: Prospective robust-consensus ADM FID10k on SelfFlow.

Seed	Identity	Median	Median gain	Mean	Mean gain
20270201	37.071	36.856	+.214	36.915	+.156
20270202	37.367	37.124	+.244	37.168	+.199
20270203	37.531	37.138	+.393	37.268	+.263
20270204	37.252	36.751	+.501	36.792	+.460
20270205	37.391	36.995	+.397	37.079	+.312
20270206	37.590	36.761	+.829	36.927	+.663
20270207	36.740	36.475	+.265	36.564	+.176
20270208	37.576	37.298	+.278	37.325	+.251

16.3 Zero-fit transfer to corrected DiT

The same logits are loaded as a direct-grid Bernstein schedule on the audited corrected 1M-step DiT EMA. This leaves the canonical linear path and every field weight unchanged. NFE8, NFE12/24, and NFE48/64 use three separately frozen groups of eight new generation seeds. CleanFID uses the validated cached CIFAR-10 train moments; each arm generates 10,000 paired images. Median wins 40/40, mean wins 40/40, and median is better than mean in every cell.

Table 27: Mean corrected-DiT CleanFID10k under frozen SelfFlow grids.

NFE	Identity	Median	Median gain	Mean	Mean gain	Wins
8	129.538	128.588	+.949	129.123	+.415	8/8
12	102.308	97.771	+4.537	99.407	+2.901	8/8
24	64.837	61.326	+3.511	62.400	+2.437	8/8
48	37.288	34.780	+2.508	35.439	+1.849	8/8
64	29.652	27.996	+1.655	28.401	+1.251	8/8

Table 28: All corrected-DiT consensus-transfer cells. Gains are identity minus intervention CleanFID10k.

NFE	Generation seed	Identity	Median	Median gain	Mean gain
8	20270301	129.407	128.484	+.923	+.385
8	20270302	130.020	128.808	+1.211	+.693
8	20270303	129.969	129.205	+.764	+.288
8	20270304	129.170	128.241	+.929	+.353
8	20270305	129.474	128.657	+.818	+.275
8	20270306	129.322	128.521	+.801	+.256
8	20270307	129.696	128.638	+1.058	+.532
8	20270308	129.243	128.154	+1.088	+.536
12	20270401	102.072	97.518	+4.555	+2.951
12	20270402	102.524	97.894	+4.630	+2.993
12	20270403	102.295	97.870	+4.425	+2.857
12	20270404	102.295	97.627	+4.669	+2.964
12	20270405	102.378	97.792	+4.586	+2.968
12	20270406	101.504	97.032	+4.471	+2.835
12	20270407	102.537	98.130	+4.406	+2.775
12	20270408	102.858	98.303	+4.555	+2.866
24	20270401	64.834	61.287	+3.547	+2.466
24	20270402	64.795	61.251	+3.544	+2.466

NFE	Generation seed	Identity	Median	Median gain	Mean gain
24	20270403	64.911	61.443	+3.467	+2.446
24	20270404	64.746	61.206	+3.541	+2.459
24	20270405	65.106	61.562	+3.545	+2.445
24	20270406	64.174	60.661	+3.513	+2.426
24	20270407	64.597	61.199	+3.398	+2.317
24	20270408	65.532	61.997	+3.535	+2.469
48	20270501	37.283	34.729	+2.554	+1.883
48	20270502	37.564	35.073	+2.491	+1.845
48	20270503	37.446	34.923	+2.523	+1.863
48	20270504	36.782	34.337	+2.445	+1.817
48	20270505	37.563	35.013	+2.550	+1.874
48	20270506	37.287	34.832	+2.455	+1.805
48	20270507	36.940	34.348	+2.592	+1.887
48	20270508	37.438	34.983	+2.455	+1.814
64	20270501	29.571	27.925	+1.646	+1.244
64	20270502	29.895	28.237	+1.658	+1.251
64	20270503	29.776	28.129	+1.647	+1.272
64	20270504	29.221	27.647	+1.575	+1.191
64	20270505	29.863	28.182	+1.680	+1.255
64	20270506	29.756	28.055	+1.701	+1.282
64	20270507	29.274	27.572	+1.702	+1.306
64	20270508	29.855	28.220	+1.635	+1.206

Absolute median gain rises from .949 at NFE8 to 4.537 at NFE12, then contracts through $3.511, 2.508, 1.655$ at NFE24/48/64. There is no high-budget reversal. This curve is consistent with a finite-step intervention whose largest effect occurs after the solver has enough stages to exploit the shared nonuniform allocation; it is not evidence of a different continuous ODE.

17 POST-HOC SOURCE-ONLY ANALYTIC-GRID CONTROL

After observing the robust median-transfer result, we introduced this interpretive control to determine whether its exact coordinate geometry is necessary. We fit the one-parameter family $C_a(t) = t + at(1-t)$ by least squares to the three interior NFE8 nodes of the frozen SelfFlow median grid. The fit itself reads schedule coordinates only and returns $a = .0893154144$; the family and scalar were then frozen before any analytic-arm cell was evaluated. Thus the control is source-only and zero-fit at transfer time, but it is not a prospective primary confirmation. The curve defines a valid endpoint-preserving grid because $C_a'(t) = 1 + a(1-2t) > 0$. It is evaluated on exactly the eight paired SelfFlow streams and 40 paired DiT stream–budget cells used by the uniform/median audit.

Table 29: Canonical source-only analytic-grid summary. Parentheses give the paired 95% t interval across eight generation streams for analytic gain over uniform.

Receiver	NFE	Uniform	Analytic	Median	Analytic gain (95% CI)	Wins
SelfFlow	8	37.315	36.879	36.925	.436 [.282, .590]	8/8
DiT	8	129.538	129.059	128.588	.479 [.334, .623]	8/8
DiT	12	102.308	96.908	97.771	5.400 [5.324, 5.476]	8/8
DiT	24	64.837	61.366	61.326	3.471 [3.424, 3.518]	8/8
DiT	48	37.288	34.608	34.780	2.680 [2.645, 2.715]	8/8
DiT	64	29.652	27.935	27.996	1.716 [1.690, 1.743]	8/8

The analytic curve wins all 48 comparisons with uniform; for SelfFlow and at every DiT budget, its paired interval across the eight generation streams excludes zero. Relative to median, analytic wins 8/8 on SelfFlow and 8/8 on DiT NFE12, 48, and 64; median wins 8/8 at DiT NFE8 and 7/8 at NFE24. The paired median gains over analytic (analytic FID minus median FID) are $-.046, .470, -.862, .040, -.172, -.061$ for SelfFlow NFE8 and DiT NFE8/12/24/48/64, respectively. Thus exact median-node geometry is not the source of the broad transfer result. It exposes a smooth, compressible timing direction, while residual shape utility remains budget specific.

Table 30: All source-only analytic-grid cells. The first block is SelfFlow and the remaining rows are corrected DiT. $U - A$ is analytic gain over uniform; $A - M < 0$ means analytic has lower FID than median.

NFE	Seed	Uniform	Analytic	Median	$U - A$	$A - M$
8	20270201	37.071	36.800	36.856	+.271	−.057
8	20270202	37.367	37.036	37.124	+.331	−.087
8	20270203	37.531	37.098	37.138	+.433	−.040
8	20270204	37.252	36.732	36.751	+.520	−.019
8	20270205	37.391	36.937	36.995	+.454	−.058
8	20270206	37.590	36.749	36.761	+.841	−.011
8	20270207	36.740	36.431	36.475	+.309	−.044
8	20270208	37.576	37.247	37.298	+.329	−.051
8	20270301	129.407	128.928	128.484	+.479	+.444
8	20270302	130.020	129.197	128.808	+.823	+.389
8	20270303	129.969	129.656	129.205	+.313	+.451
8	20270304	129.170	128.778	128.241	+.393	+.537
8	20270305	129.474	129.129	128.657	+.346	+.472
8	20270306	129.322	128.985	128.521	+.337	+.464
8	20270307	129.696	129.110	128.638	+.586	+.472
8	20270308	129.243	128.689	128.154	+.554	+.535
12	20270401	102.072	96.652	97.518	+5.420	−.865
12	20270402	102.524	97.030	97.894	+5.493	−.863
12	20270403	102.295	97.013	97.870	+5.282	−.856
12	20270404	102.295	96.826	97.627	+5.469	−.801
12	20270405	102.378	96.868	97.792	+5.510	−.923
12	20270406	101.504	96.186	97.032	+5.318	−.847
12	20270407	102.537	97.242	98.130	+5.295	−.889
12	20270408	102.858	97.448	98.303	+5.409	−.855
24	20270401	64.834	61.355	61.287	+3.479	+.068
24	20270402	64.795	61.239	61.251	+3.556	−.012
24	20270403	64.911	61.455	61.443	+3.456	+.012
24	20270404	64.746	61.296	61.206	+3.450	+.090
24	20270405	65.106	61.609	61.562	+3.497	+.047
24	20270406	64.174	60.690	60.661	+3.485	+.028
24	20270407	64.597	61.241	61.199	+3.356	+.042
24	20270408	65.532	62.041	61.997	+3.491	+.044
48	20270501	37.283	34.579	34.729	+2.703	−.150
48	20270502	37.564	34.897	35.073	+2.668	−.177
48	20270503	37.446	34.737	34.923	+2.708	−.186
48	20270504	36.782	34.157	34.337	+2.625	−.180
48	20270505	37.563	34.850	35.013	+2.713	−.163
48	20270506	37.287	34.639	34.832	+2.647	−.193
48	20270507	36.940	34.202	34.348	+2.738	−.146
48	20270508	37.438	34.801	34.983	+2.637	−.183
64	20270501	29.571	27.877	27.925	+1.695	−.049
64	20270502	29.895	28.168	28.237	+1.727	−.069
64	20270503	29.776	28.057	28.129	+1.719	−.072
64	20270504	29.221	27.565	27.647	+1.656	−.082
64	20270505	29.863	28.121	28.182	+1.741	−.061
64	20270506	29.756	28.015	28.055	+1.742	−.040
64	20270507	29.274	27.522	27.572	+1.752	−.050
64	20270508	29.855	28.157	28.220	+1.698	−.064

18 Receiver-local numerical grids on modern ImageNet receivers

18.1 Frozen family, selection, and exact deployment semantics

Both receiver-local studies use the same nine-member monotone cubic Bézier family. Two logits range over $\{-.6, 0, .6\}^2$; a softmax converts them into three positive increments, so every candidate preserves endpoints and is strictly increasing. Two calibration streams per receiver select the eligible candidate with the smallest latent endpoint MSE between successive numerical budgets. Eligibility

requires at least 2% improvement in each stream and 3% on their mean, followed by a frozen refinement-safety check. No decoded image, FID, DINO, CLIP, VGG, or target reference statistic is available to selection.

Table 31: Numerically selected receiver-local cubic controls. Relative improvements compare the selected candidate with identity endpoint discrepancy.

Receiver	Candidate	Control points	Calibration seeds	coarse	refined
FlowTurbo	`m060_p060`	$[0, .162807, .703346, 1]$	20272801–02	39.698%	safety passed
REPA 4M	`m060_z000`	$[0, .215321, .607660, 1]$	20273001–02	12.564%	13.913%

FlowTurbo uses its released predictor-only SiT-XL/2 receiver, CFG 1.5, and eight Heun blocks, corresponding to 16 predictor evaluations. REPA uses its released SiT-XL/2 4M checkpoint, EMA VAE, CFG 1.5 over the full guidance interval, and explicit Euler at exact NFE8/16. Smoke audits verify strict checkpoint loading, reverse-time orientation, endpoint preservation, exact NFE, and identity parity before selection.

18.2 FlowTurbo FID, ranking, and independent features

Table 32: FlowTurbo H8 ADM FID10k. Positive gain is identity minus arm; intervals and signs use eight fresh paired generation streams.

Arm	Mean FID	Gain	95% CI	Wins
Identity	10.1386	0	–	–
Receiver cubic clock	9.9444	.1942	$[.1516, .2368]$	8/8
Receiver cubic grid	9.9180	.2206	$[.1786, .2626]$	8/8

A separate two-stream ranking audit evaluates all nine already frozen candidates. The numerical winner is the best-FID candidate for both clock and grid deployments. Numerical discrepancy versus mean FID has Spearman $\rho = .7333$ ($p = .02455$) and Kendall $\tau = .6111$ ($p = .02474$) for clock dynamics; for the induced grid the correlations strengthen to $\rho = .9167$ ($p = .000507$) and $\tau = .8333$ ($p = .000854$). This supports discrepancy as a directional receiver-local ranking signal rather than merely a lucky selected cell.

On the same eight confirmation streams, the induced grid improves DINOv2 FD by 2.7238 $[1.8622, 3.5855]$, KID by $.002292$ $[.001343, .003241]$, and recall by $.008013$ $[.004418, .011607]$, all 8/8. The clock arm independently improves DINOv2 FD by 1.4291 $[.4088, 2.4493]$ (6/8) and recall by $.014688$ $[.008944, .020431]$ (8/8).

18.3 REPA cubic selection and one-parameter compression

The independent REPA selection chooses cubic points $[0, .215321, .607660, 1]$. Its induced grid improves all NFE8 and NFE16 paired streams. We then compress this frozen cubic by the exact continuous unweighted L_2 projection onto $C_a(q) = q + aq(1-q)$. Symmetry gives $a = 1.5(p_1 + p_2 - 1) = -.265528662956$; this computation reads no image metric.

Table 33: REPA ADM FID10k confirmations. Cohort 1 evaluates the selected cubic and its analytic compression; cohort 2 prospectively repeats the already frozen analytic grid.

Cohort	Budget	Identity	Cubic grid	Analytic grid	Analytic gain [95% CI]	Wins
1	NFE8	14.2651	13.9538	13.7149	$.5502$ $[.4906, .6099]$	8/8
1	NFE16	9.5339	9.3682	9.3529	$.1810$ $[.1530, .2089]$	4/4
2	NFE8	14.5113	–	13.9442	$.5671$ $[.4728, .6615]$	8/8
2	NFE16	9.4903	–	9.2879	$.2024$ $[.1557, .2491]$	4/4

For completeness, the cubic-grid gains in cohort 1 are $.3113$ $[.2466, .3759]$ at NFE8 (8/8) and $.1657$ $[.1233, .2081]$ at NFE16 (4/4). The analytic grid beats the cubic by $.2390$ $[.2154, .2626]$ at NFE8

(8/8) and is statistically tied at NFE16. The result therefore compresses the useful receiver-local direction to one scalar; it does not require cubic flexibility.

18.4 Prospective DINOv2 agreement and matched-NFE latency

Table 34: Prospective REPA DINOv2 gains of the matched analytic grid over identity. All intervals exclude zero. FD/KID use baseline minus intervention; precision/recall use intervention minus baseline.

Budget	FD gain	KID gain	Precision gain	Recall gain
NFE8	13.3717 [12.3634, 14.3799]	.013634 [.012276, .014991]	.007975 [.003776, .012174]	.012137 [.006771, .017504]
NFE16	7.2927 [6.5954, 7.9900]	.007691 [.006721, .008660]	.005075 [.002083, .008067]	.008500 [.002664, .014336]

Every displayed metric improves on all eight NFE8 streams and all four NFE16 streams. Thus the ADM-FID result transfers to an independently cached DINOv2-L feature space and is accompanied by both precision and recall gains.

Table 35: Matched-NFE REPA latent-solver latency on one NVIDIA RTX 6000D, batch 64, two warmups, and ten CUDA-event repetitions. VAE decode is excluded.

Budget	Identity median (ms)	Analytic median (ms)	Difference
NFE8	4260.72	4259.62	$-.0258\%$
NFE16	8518.37	8519.71	$+.0158\%$

Both absolute differences are below $.04\%$; no additional network evaluation or measurable solver-time overhead is introduced by replacing uniform nodes with the analytic grid.

19 Goal-oriented temporal-measure acquisition and transfer

19.1 Development path into the order-preserving measure

Two controlled development results motivated the final representation. First, a positive C^1 piecewise-quadratic clock was constrained so that Heun's trapezoidal derivative mass equals every interval increment, eliminating the effective-dose drift of an earlier unconstrained clock. On the field-only receiver it improves projected risk over the learned clock by 8.41% in 16/16 precision/witness cells and beats its own induced grid 16/16. Frozen transfer to four further DiT receivers again beats the learned clock and the induced grid in 8/8 cells, by 7.60% and 17.64% on average. This isolates reproducible finite-step stage-shape headroom before direct grid optimization.

Second, a seven-parameter global generalized two-stage controller improves the learned-clock proxy by 23.47% in 16/16 cells. A state-conditioned extension is fully explained by its globally averaged actions under the shuffle audit, and the generalized update does not preserve the classical second-order conditions. We therefore retain the fixed global result only as a development upper bound; the paper's transferable method freezes classical midpoint and learns only positive interval masses.

19.2 Exact midpoint parameterization and acquisition contract

At exact NFE8, explicit midpoint uses four intervals and two field evaluations per interval. We optimize only their positive masses $m = \mathrm{softmax}(\eta)$; the stage location remains $1/2$, the first and second output weights remain $(0, 1)$, and the maximum second-order condition residual is zero. Thus this experiment learns neither a free RK tableau nor a receiver field. The source is corrected unconditional CIFAR-10 DiT artifact 20268401; both learned arms use calibration seed 20313301, 128 noises, the same 64-dimensional projected terminal Fréchet risk, and exactly 131,072 model-sample-NFEs.

The differentiable arm performs 64 solver-backpropagation iterations. The black-box arm uses 16 Sobol initial points followed by Gaussian-process expected improvement, 128 total evaluations, and a fixed 8,192-candidate pool. Table 36 records the resulting measures and untouched source FID. Wall

time excludes shared model and feature-bank setup; peak acquisition memory was not instrumented and is not reconstructed from later system snapshots.

Table 36: Source acquisition contract and result at explicit midpoint NFE8.

Arm	Four interval masses	Acquisition	Model-sample-NFEs	Time (s)	Calibration risk	CleanFID10k
Uniform	$[.2500, .2500, .2500, .2500]$	none	–	–	12.8886	26.4430
Gradient	$[.2354, .2338, .2289, .3019]$	solver gradient	131,072	63.65	12.4527	22.4064
Sobol/GP	$[.0903, .2063, .2761, .4272]$	black box	131,072	54.64	**10.9496**	**20.0774**

Before FID, the gradient measure beats uniform midpoint in 16/16 frozen cells crossing AMP/FP32 and two independent projection banks (mean risk change $-.6561$, 8.04%). Transferred unchanged to four other corrected-DiT artifacts, it wins all eight new source-proxy cells (mean change $-.8468$, 9.67%). The black-box measure then beats both gradient and uniform on four new streams:

Table 37: Held-out projected terminal risk for matched-cost source acquisition.

Stream	Uniform	Gradient	Sobol/GP
20314401	8.6310	8.2161	**7.3270**
20314402	9.5152	8.9402	**7.2310**
20314403	7.8096	7.3014	**6.3328**
20314404	9.1921	8.4411	**6.6682**

The gradient measure also receives an artifact-level paired CleanFID10k confirmation before the stronger black-box search is introduced:

Table 38: Frozen gradient-measure CleanFID10k across independently trained corrected-DiT receivers. Positive gain is uniform minus optimized.

Receiver	Seed	Uniform	Gradient measure	Gain
20268401	20313701	26.2296	22.2950	3.9346
20268402	20313702	27.0407	23.1993	3.8413
20268403	20313703	27.4068	23.4262	3.9807
20268404	20313704	27.3022	23.2535	4.0487
20268405	20313705	26.4895	22.5533	3.9362
Mean	–	26.8938	**22.9455**	3.9483 (14.68%)

The measure wins 5/5; the paired 95% Student-t interval for positive gain is $[3.8544, 4.0422]$. Both arms have identical parameter count and exact NFE. On the uncontended first pair, total generation-plus-FID time is 42.84 versus 42.78 s and peak CUDA allocation is identical.

19.3 Independent source acquisition

A second black-box acquisition uses new calibration seed 20314901 with every search setting and cost held fixed. It returns $[.120592, .178724, .194881, .505803]$, rather than reproducing the original vector exactly. On the union of the two four-stream banks, both original and replicated measures beat uniform 8/8. Their mean risks are 6.8542 and 6.7824, compared with 8.4468 for uniform, and they split their head-to-head comparison 4/4. The L_1 distance between mass vectors is .2176, but both increase the final-interval mass relative to uniform. This supports a stable direction/family and explicitly rules out interpreting the first four numbers as a unique universal optimum.

19.4 Prospective SiT and SelfFlow transfer

The original source measures are hashed and frozen before target generation. There is no SiT/SelfFlow trajectory collection, schedule objective, gradient, feature query, scalar dose, or candidate choice during fitting. Target samples are used only for the preregistered evaluation rows below.

All four reads favor the stronger source acquisition. Because the two targets are official class-conditional ImageNet-256 receivers while the source is an unconditional CIFAR-10 DiT, the transfer

Table 39: Complete zero-target transfer record at explicit midpoint NFE8. FID2k is the screening read; FID10k is a new confirmation seed.

Receiver	Read	Seed	Uniform	Gradient	Sobol/GP	Sobol/GP gain
SiT-XL/2	ADM FID2k	20314601	26.7992	26.2042	**25.4282**	.7760 vs. grad.
SiT-XL/2	ADM FID10k	20314611	–	10.5158	**9.8663**	.6495 (6.18%)
SelfFlow	ADM FID2k	20314602	36.1381	34.7386	**34.0226**	.7160 vs. grad.
SelfFlow	ADM FID10k	20314612	–	19.2975	**17.1338**	2.1637 (11.21%)

crosses architecture instance, dataset, conditioning, and absolute quality. It remains a two-checkpoint transfer result rather than a universal receiver theorem.

19.5 Canonical FID50k and unseen-RK2 transfer

We next evaluated all three frozen arms at the standard 50,000-sample scale. Each receiver shared one generation seed across uniform, gradient, and Sobol/GP arms. No target trajectory, feature, gradient, FID, or schedule-fitting query was used before deployment. After this six-cell experiment completed, we deployed the same source masses unchanged with Ralston RK2 ($c_2 = a_{21} = 2/3$, $b_1 = 1/4$, $b_2 = 3/4$), an unseen member of the explicit two-stage second-order family.

Table 40: Canonical target-receiver FID50k and zero-fit solver transfer. All three arms within a row share the listed generation seed. Lower is better.

Receiver	Solver/read	Seed	Uniform	Gradient	Sobol/GP	Gain vs. uniform
SiT-XL/2	Midpoint FID50k	20315201	8.3674	7.7736	**7.1696**	14.32%
SelfFlow	Midpoint FID50k	20315202	17.9714	16.6210	**14.6264**	18.61%
SiT-XL/2	Ralston FID10k	20315211	12.2741	11.5892	**11.1698**	9.00%

The Sobol/GP measure beat both uniform and gradient acquisition on both FID50k receivers, improving over gradient by 7.77% on SiT and 12.00% on SelfFlow. Under Ralston, gradient improved over uniform by 5.58%, and Sobol/GP improved over uniform and gradient by 9.00% and 3.62%. Before generation, the Ralston implementation satisfied $b_1 + b_2 = 1$ and $b_2 c_2 = 1/2$ and matched the independently shared `fixed_grid_solve` implementation with zero latent maximum and RMS error at exact NFE8. All integrity and preregistered ordering gates passed. These results extend the earlier screens to canonical sample scale and one unseen RK2 tableau; they do not imply solver-independent optimality.

Before the matched-cost black-box comparison, the gradient measure had already passed target-vs.-uniform confirmation. On SiT it wins all three paired FID2k and all three paired FID10k streams; on SelfFlow it wins all three FID2k streams and the preregistered fresh FID10k pair:

Table 41: Independent zero-fit confirmation of the gradient source measure.

Receiver	Read	Pairs	Uniform	Gradient measure	Gain
SiT-XL/2	mean ADM FID2k	3	26.2678	25.7021	.5657 (2.16%)
SiT-XL/2	mean ADM FID10k	3	11.0910	**10.4856**	.6053 (5.46%)
SelfFlow	mean ADM FID2k	3	36.2082	35.0315	1.1767 (3.25%)
SelfFlow	fresh ADM FID10k	1	21.3108	**20.0190**	1.2918 (6.06%)

The SiT FID10k paired 95% interval for positive gain is $[.5279, .6828]$. These reads establish that the gradient schedule is already a transferable baseline; the later table shows that stronger source-only acquisition improves it on new target streams.

19.6 Analytic compression and the residual

Least-squares fitting of $C_\gamma(r) = r^\gamma$ to the frozen Sobol/GP nodes gives $\gamma = 1.8146211435$, masses $[.080814, .203464, .309034, .406687]$, and maximum node error $.020562$. This CPU-only fit reads

no target output. The analytic measure beats uniform and gradient on all four source proxy streams, while the full measure wins all four head-to-head comparisons. A separate zero-fit target pair gives:

Table 42: Analytic compression on new paired target FID2k streams. "Captured" is the analytic gain over uniform divided by the full-measure gain.

Receiver	Seed	Uniform	Power law	Full measure	Captured
SiT-XL/2	20314701	26.4176	25.6114	**25.3282**	74.0%
SelfFlow	20314702	35.6512	32.9634	**32.3948**	82.5%

Hence a one-parameter trend accounts for most of the target gain, but the source-acquired nonparametric residual remains useful on both targets and on every source stream.

19.7 Optimal-density derivation and exact budget refinement

Let an order-p stable one-step method deploy $N\rho(t)\,dt$ steps in an infinitesimal interval, with normalized $\rho \geq 0$ and local step size $h(t) \simeq [N\rho(t)]^{-1}$. If the leading local contribution to the terminal goal is $a(t)h(t)^{p+1}$, accumulation gives

$$\mathcal{E}_N(\rho) = N^{-p} \int_0^1 a(t)\rho(t)^{-p}\,dt + o(N^{-p}).$$

For the constrained functional $\int a\rho^{-p} + \lambda(\int \rho - 1)$, stationarity yields $-pa(t)\rho(t)^{-p-1} + \lambda = 0$. Normalization therefore gives

$$\rho^\star(t) = \frac{a(t)^{1/(p+1)}}{\int_0^1 a(s)^{1/(p+1)}\,ds}.$$

For midpoint $p = 2$, the optimum is proportional to the cube root of the leading goal-error density. In a terminal distribution problem, $a(t)$ contains both the local defect and its terminal adjoint/sensitivity weight. The derivation is explanatory and asymptotic; it is not a proof that the observed four-bin measure is a global finite-N optimum.

Frozen direct profile test. The preregistered theorem check estimates an absolute adjoint-weighted step-doubling profile with eight uniform midpoint intervals (NFE16) on three disjoint 128-noise streams. Replaying every stored full-step endpoint gives zero maximum error, and the three profiles are highly reproducible (pairwise cosine similarities .99483–.99711). The aggregate profile is $[6.15413, 3.83013, 1.60555, .57159, .25944, .39337, 1.66716, 9.21386]$. Its normalized cube-root density gives four equal-density-quantile masses $[.179337, .264532, .387251, .168880]$.

Table 43: Frozen held-out projected-risk test of the direct cube-root profile. Lower is better; the cube-root arm is not eligible for FID because its primary gate fails.

Schedule	Mean projected risk	Wins vs. uniform
Uniform midpoint	8.6319	–
Analytic power law	7.0961	4/4
Gradient terminal	8.0090	4/4
Sobol/GP terminal	**6.9040**	4/4
Cube-root profile	8.8359	1/4

The cube-root schedule worsens mean risk by .20395 relative to uniform. The failure is not explained by profile noise or an implementation mismatch. It instead exposes the gap between a leading nonnegative asymptotic bound and the nonlinear finite-sample distributional risk at four midpoint intervals. Signed cancellation, model error, higher-order terms, and cross-step interactions are all discarded by the absolute profile. The protocol forbids post-result changes to the exponent, floor, smoothing, or streams.

Frozen finite-NFE8 mass geometry. A separate preregistered diagnostic evaluates symmetric finite differences of the scalar terminal risk on a fixed three-dimensional simplex tangent basis. It

crosses four fresh 256-noise streams with the original and an independently seeded 64-dimensional projection witness. Every stencil mass remains positive and normalized. All 8/8 seed–witness cells preserve the complete ordering Sobol/GP $<$ analytic power $<$ gradient $<$ uniform.

Table 44: First- and second-order finite-NFE8 geometry. The directional derivative follows the unit direction from uniform toward Sobol/GP.

Witness	R_{unif}	R_{Sobol}	Direction	Grad. ratio	Linear RMSE	Quad. RMSE
Original	7.9365	**6.6618**	-15.9784	.3948	**.0620**	.2251
Independent	7.5085	**6.3793**	-18.6069	.3762	**.0762**	.1711

The negative directions and roughly 60% reduction in tangent-gradient norm support movement toward population-level first-order marginal-risk balance. The residual gradient is nonzero, so the result does not establish exact KKT stationarity. Quadratic prediction is worse than linear for both witnesses. Moreover, all two-radius Hessian comparisons exceed the frozen .5 resolution threshold. The corresponding eigenvalues are therefore not interpreted, and no stable local quadratic basin is claimed. No new schedule or FID read is created from either diagnostic.

For a discrete K-interval mass vector, define $\rho_m(t) = 1/(Km_j)$ on the physical interval of length m_j. Equal binary subdivision creates two children of length $m_j/2$ and changes the interval count to $2K$, so

$$\rho_{\mathrm{split}(m)}(t) = \frac{1}{(2K)(m_j/2)} = \frac{1}{Km_j} = \rho_m(t).$$

The continuous piecewise-linear time map is therefore identical; only its sampling resolution changes. Without NFE16 or target fitting, this refinement retains the source ordering on SiT:

Table 45: Density-preserving SiT refinement at exact midpoint NFE16.

Read	Seed	Uniform	Gradient	Sobol/GP	Gain vs. gradient
ADM FID2k	20314801	24.5108	24.3660	**24.2357**	.1303 (.53%)
ADM FID10k	20314811	–	8.9056	**8.7705**	.1351 (1.52%)

The refined gradient density had separately passed a target-vs.-uniform gate: three paired SiT FID2k streams improve from 24.6698 to 24.5695 (3/3; positive-gain interval $[.0567, .1437]$), and fresh FID10k seed 20314011 improves from 9.2130 to 9.0952 (gain $.1178$, 1.28%). The stronger source acquisition therefore preserves the ordering uniform $>$ gradient $>$ Sobol/GP under deterministic refinement.

19.8 Goal-specific controls and target-access ledger

The numerical-teacher control fits a SiT-specific mass vector against the latent endpoint of uniform explicit RK4-128. Across four held-out numerical streams it lowers midpoint-NFE8 endpoint MSE from $.125007$ (uniform) to $.045571$, a 63.5% reduction; the source goal measure gives $.118529$. The fitted teacher masses are $[.059319, .066495, .243566, .630620]$. Nevertheless, on three paired FID2k streams the teacher schedule worsens FID from 26.7461 to 29.3383, whereas the source goal measure improves it to 26.0756. Numerical endpoint fidelity is therefore not a sufficient selection objective for the distributional goal.

We also implement a TORS-style equal-total-rotation control using three PCA dimensions, a five-point local fit, and the recorded outlier rule. On the source it profiles 100 trajectories at 100 Euler steps (10,000 model-sample-NFEs, 7.45 s including geometry) and returns $[.024387, .400825, .205676, .369112]$. It loses to uniform, gradient, power-law, and full terminal measures on all four new source streams. For a native target access comparison, we re-profile the same construction on SiT using 100 guided target trajectories and 100 Euler steps: 10,000 target model-sample-NFEs, 20,000 conditional rows under batch-doubled CFG, 81.21 s, and no target quality metric during fitting. The resulting masses are $[.201528, .298410, .308669, .191393]$.

This control attributes the observed difference to the tested SiT/explicit-midpoint adaptation. It is not an official Flux/SD3 reproduction and does not support a universal claim against TORS. The

Table 46: Matched SiT midpoint-NFE8 FID2k mechanism control (seed 20315111).

Schedule	Target profile NFEs	Target metric queries in fit	ADM FID2k
Uniform	0	0	26.8912
Source terminal measure	0	0	**26.0330**
Target-profiled TORS-style	10,000	0	27.6034

defensible mechanism claim is narrower: in these cells, goal-aligned terminal distribution sensitivity transfers better than dense-teacher endpoint error or local trajectory rotation.

20 Cross-domain boundary and domain-local transfer

20.1 A fixed image schedule is not a modality-independent mesh

We first tested whether the frozen image Sobol/GP masses could be transplanted unchanged to four supplied offline-action datasets. Three independently trained conditional flow receivers per dataset used the same explicit midpoint/NFE8 deployment and no target-side schedule fitting. The image schedule beat uniform task MSE in 8/12 cells, but its aggregate mean was worse: $.115906$ versus $.115739$ for uniform, with mean paired gain -1.67×10^{-4} and receiver-cell bootstrap interval $[-.00276, .00209]$. Its endpoint MSE to uniform RK4/NFE128 was also worse, $.003139$ versus $.001070$, with the interval entirely on the harmful side. The preregistered RK3/NFE12 and RK4/NFE16 task-transfer gates also failed. These results reject a modality-independent interpretation of the four image masses and were not used to alter them.

20.2 Frozen domain-local design

The positive follow-up tested a different, narrower hypothesis: a schedule acquired on source receivers within a representation domain can transfer without refitting to new receiver seeds in that domain. Table 47 records all four representations and split sizes. Normalization uses training rows only. ETTm1 uses chronological regions with 5% gaps and a full-window purge; FSDD retains recording indices 0–4 as confirmation, 5–9 as calibration, and 10–49 as training; AG News uses the official test file only for confirmation; and Lorenz trajectories, rather than individual segments, are disjoint across splits.

Table 47: Frozen vector representations and splits for the domain-local audit. The split columns report train/calibration/confirmation rows.

Domain	Target representation	Condition	Split rows	Target dim.	Classes
ETTm1	flattened 32×7 window	quarter-day phase	10,445/2,606/2,606	224	4
FSDD	16×16 log-spectral vector	spoken digit	2,400/300/300	256	10
AG News	hashed unigram/bigram vector	topic	100,000/20,000/7,600	256	4
Lorenz	flattened 32×3 trajectory segment	$\rho \in \{24, 28, 32, 36\}$	4,608/1,536/1,536	96	4

Every domain uses the conditional optimal-transport path $x_t = (1-t)z + tx$ with target velocity $x - z$. The conditional field has four width-384 SiLU hidden layers and a fixed five-component time embedding. AdamW training uses learning rate 2×10^{-4}, weight decay 10^{-4}, gradient clipping at 5, batch size 512, and 6,000 updates. Source receiver seeds are 20317401 and 20317402; confirmation receiver seeds are 20317501, 20317502, and 20317503. The two groups share training data and architecture but not initialization, minibatch, or noise streams.

All domains receive the same ordered family of 64 positive four-mass candidates: uniform; the image gradient, analytic-power, and Sobol/GP schedules; the reversal of the image Sobol/GP schedule; and 59 logistic-normal candidates generated with seed 20317311, four fixed logit scales, and a 5% uniform mixture. The domain schedule minimizes mean calibration task risk over the two source receivers, with candidate order as the tie break. There is no continuous optimization, candidate extension, or confirmation-side selection. ETTm1 and Lorenz use normalized paired target MSE for the frozen primary risk; FSDD and AG News use mean class-conditional sliced-Wasserstein-2

over 64 directions. Acquisition uses projection seed 20317601. Confirmation averages independent projection seeds 20317611 and 20317612.

The acquisition record is committed before the confirmation array is opened. Table 48 reports the four selected candidates. ETTm1 and AG News independently select the same member of the shared candidate family; no schedule or result is passed between the domains.

Table 48: Domain-local source acquisition. Risk is averaged over two source receivers on calibration data; lower is better.

Domain	Selected masses	Uniform risk	Image Sobol/GP	Selected risk
ETTm1	$[.03249, .76833, .04759, .15160]$	1.96827	2.00724	**1.65428**
FSDD	$[.49874, .25131, .02337, .22657]$	.50617	.50737	**.50328**
AG News	$[.03249, .76833, .04759, .15160]$	.07060	.06491	**.02809**
Lorenz	$[.71067, .09775, .10112, .09046]$	1.38970	1.41779	**1.24520**

20.3 Fresh-receiver confirmation

The deployment comparison includes uniform, the domain-selected schedule, and all three frozen image schedules at explicit midpoint NFE8. Common confirmation rows and Gaussian sources are used within every receiver. Secondary metrics are normalized paired MSE, the two-witness conditional sliced-Wasserstein-2, projected Gaussian Fréchet, and endpoint MSE to uniform-grid RK4/NFE128. Table 49 reports every domain-level primary result rather than selecting favorable domains.

Table 49: Primary confirmation risk across three fresh receivers per domain. Log improvement is the mean of $\log(R_{\mathrm{unif}}/R_{\mathrm{sel}})$; positive is better.

Domain	Primary metric	Uniform	Selected	Log improvement	Wins
ETTm1	paired MSE	2.23298	**1.92104**	.15045	3/3
FSDD	conditional SW2	**.99990**	1.00674	.01745	2/3
AG News	conditional SW2	.062623	**.019513**	1.16561	3/3
Lorenz	paired MSE	1.40619	**1.27593**	.09710	3/3

Across all 12 domain–receiver cells, the selected schedule wins 11/12. The mean log improvement over uniform is $.357651$; a deterministic 20,000-replicate receiver-cell bootstrap with seed 20317899 gives the 95% interval $[.104153, .635253]$. The corresponding geometric mean selected-to- uniform risk ratio is $.6993$, and the one-sided sign probability is $.00317$. The selected schedules also beat the frozen image Sobol/GP schedule in 11/12 cells: mean log improvement $.340797$, interval $[.125424, .583055]$. Both comparisons pass the preregistered directional and strong gates. This cell-level bootstrap follows the frozen analysis plan; because cells share four domain selections, it does not imply population inference over all possible modalities.

The secondary distribution metrics materially qualify the aggregate result:

Table 50: Distributional confirmation audit. Entries are uniform $\rightarrow$ domain-selected means across three receivers.

Domain	Conditional SW2	Wins	Projected Fréchet	Wins
ETTm1	$.63391 \rightarrow \mathbf{.55123}$	3/3	$21.1128 \rightarrow \mathbf{19.6765}$	3/3
FSDD	$\mathbf{.99990} \rightarrow 1.00674$	2/3	$\mathbf{22.8370} \rightarrow 23.6658$	1/3
AG News	$.062623 \rightarrow \mathbf{.019513}$	3/3	$2.00151 \rightarrow \mathbf{.59359}$	3/3
Lorenz	$\mathbf{.064675} \rightarrow .179212$	0/3	$\mathbf{1.46553} \rightarrow 3.60976$	0/3

ETTm1 and AG News therefore supply the clearest domain-consistent evidence: both witnesses improve in every fresh receiver. FSDD is mixed; its arithmetic mean primary risk is slightly worse even though two receivers improve in log ratio. Lorenz improves the frozen paired MSE but loses both distribution witnesses in every receiver. Since independent Gaussian sources have no unique

pairing with held-out Lorenz targets, the paired-MSE result cannot support faithful chaotic-trajectory generation.

Finally, all 12 domain-selected schedules are farther from the dense endpoint. Mean dense-reference MSE increases from .012150 for uniform to .084940; mean log improvement is -2.78626 with bootstrap interval $[-3.64110, -1.94943]$. The domain-local gains are consequently properties of the task-aligned finite-step operator, not evidence for a more accurate generic ODE grid. The experiment concerns fixed vector representations, not raw waveform synthesis, language modeling, online forecasting, or conditional trajectory rollout.

21 History-conditioned forecasting and trajectory transfer

21.1 Frozen design and task-local acquisition

The causal extension tests future prediction rather than unconditional vector representations. ETTm1 maps a 96-step, seven-variable history to the next 24 steps. Its chronological split uses the first 60% for training, 65–80% for calibration, and the final 15% for confirmation, with two 5% gaps and full- window purging. Lorenz maps 32 observed three-state steps to the next 32 steps over four ρ regimes. For each regime, 48/16/16 trajectories are assigned to training/calibration/confirmation without sharing trajectory identities.

Each task uses the fixed 64-candidate four-mass family. Candidate risk averages two source receivers (seeds 20321401 and 20321402), 256 calibration histories, and eight common Gaussian draws per history. The selected masses are committed before confirmation receivers 20321501, 20321502, and 20321503 are accessed. Confirmation uses 512 histories and 32 common Gaussian draws at exact midpoint NFE8. The primary metric is elementwise ensemble CRPS.

Table 51: Task-local causal acquisition. Mass is physical interval width, not evaluation density.

Task	Selected candidate	Four frozen masses
ETTm1 $96 \to 24$	logistic-normal 23	$[.22037, .62959, .09126, .05879]$
Lorenz $32 \to 32$	logistic-normal 19	$[.11998, .08660, .66907, .12435]$

21.2 Fresh-receiver CRPS confirmation

Table 52 reports every primary NFE8 cell and the frozen source-Heun-defect control. The selected measure beats uniform and the source-Heun allocation in all 6/6 receiver cells.

Table 52: Complete fresh-receiver CRPS record at exact midpoint NFE8. Lower is better.

Task	Receiver seed	Uniform	Source Heun	Selected	Reduction vs. uniform
ETTm1	20321501	.319793	.319974	**.319296**	.155%
ETTm1	20321502	.314213	.314390	**.313305**	.289%
ETTm1	20321503	.325181	.325372	**.324569**	.188%
Lorenz	20321501	.034605	.032875	**.027991**	19.114%
Lorenz	20321502	.034255	.032551	**.026703**	22.045%
Lorenz	20321503	.033791	.032098	**.026325**	22.093%

Across the six cells, mean $\log(\mathrm{CRPS}_{\mathrm{unif}}/\mathrm{CRPS}_{\mathrm{sel}})$ is .119527. A 20,000-replicate receiver-cell bootstrap with seed 20321899 gives the 95% interval $[.036984, .208053]$. Relative to source Heun, the corresponding mean is .094192 with interval $[.028913, .165487]$. Both task-level consistency rules pass, but the magnitudes differ: ETTm1 improves only .16–.29%, whereas short-horizon Lorenz improves 19.1–22.1%.

21.3 Budget, point-metric, and numerical boundaries

The NFE8-acquired measure is evaluated without modification at NFE4 and by equal mass subdivision at NFE16. These reads are frozen boundary tests rather than new schedule searches.

Table 53: Prespecified budget transfer and dense-endpoint boundary. Log improvement is against uniform CRPS; positive is favorable.

NFE	Mean log improvement	Wins	Selected MSE to dense	Uniform MSE to dense
4	−.176808	0/6	.027733	**.007129**
8	.119527	6/6	.004112	**.000256**
16	.060625	3/6	.000215	**.000010**

The schedule reverses at NFE4 in every cell. At NFE16 it retains all three Lorenz wins but loses all three ETTm1 cells. Ensemble-mean RMSE and MAE are slightly worse for ETTm1 in every receiver; Lorenz RMSE improves in one receiver and is slightly worse in two. At every budget, the selected endpoint is farther from uniform NFE128 than the matching uniform endpoint. The CRPS result is therefore not evidence of uniformly better mean trajectories or more accurate integration.

For qualitative inspection, four Lorenz confirmation rows are fixed before prediction by ground-truth future-volatility quantiles Q25/Q50/Q75/Q95. All four appear in the main-paper causal-forecasting figure. The selected NFE8 measure produces narrower 10–90% predictive bands while the truth remains inside almost all intervals. No example is selected by method difference. Empirical 80% coverage changes from 99.64–99.72% under uniform to 97.12–97.76% under the selected measure. The defensible claim is sharper probabilistic forecasts under the frozen CRPS objective at NFE8, not SOTA forecasting, long-horizon chaotic fidelity, or budget-universal transfer.

22 Calibration cost and sample-count robustness

The evaluator records curve-only time after model/feature-bank setup, total process time, peak CUDA allocation, trajectories, sample-NFE, and feature batches. A study job evaluates two receiver fields; deployment requires one.

Table 54: Frozen 128-noise calibration ledger.

Quantity	Corrected DiT	TorchCFM
Unique noises per receiver	128	128
Dose trajectories per receiver	640	640
Sample-NFE per receiver	5,120	5,120
Feature batches per receiver	10	10
Curve time per receiver (estimate)	3.06 s	2.56 s
Full two-field process time	11.29 s	8.96 s
Peak CUDA allocation	1.79 GiB	1.57 GiB
Fraction of 10k/NFE8 sample compute	6.4%	6.4%

The per-receiver wall time is half of measured two-field curve time; shared setup is excluded and the estimate is labelled as such. The chosen dose is a one-time property of the artifact and can be reused.

Table 55: Same-seed calibration sample-count stability.

Architecture	64 vs. 128	96 vs. 128	recomputed 128 vs. frozen
Corrected DiT	9/10	9/10	10/10
TorchCFM	10/10	10/10	10/10

The smaller-bank DiT disagreements are adjacent-dose changes; every joint DiT selection remains full. This analysis does not change the primary 128-noise rule.

23 Integrity and hashes

The principal machine-readable hashes are listed below. We insert explicit breaks only for typesetting; the concatenated strings are the file hashes.

TorchCFM parent `95f3baa2f6f384387005a06d46c68011 7914587294af14e79da77a4490cb7b3c`.

Fresh DiT selection `72ae6878c140f54487cd65375eaf09f6 4fb44f1869a318e66b62e7246f9d6019`.

TorchCFM selection `35e1b1ec9465b47ce738b28f528c4414 1faa84f348d3ad0adb144b2e1e6cc58e`.

Stability canonical `beb5ace3cbb2149468283114d944ea81 ab45ff014a8215eaeda9627b8925fc36`.

Dense confirmation `2478b844e552737f03a989559e73f86a 65809f79aec2dc13c2c086188c9f8f44`.

Dense global-dose rule `7d6a7994904b97c2c64eb3850c6bb24 ca9d068e7b866e5b69dbabd83e1ca6e8d`.

Solver/NFE development `5d28aabace0d439f602b1bbe6ea21c3 19881c4002e31d3fd3e62003c2f0fdfcd`.

Solver/NFE confirmation `d3e0d5ff0905f9855eea6ea6a2b6db 9a8f500b47b58e1db6d9db23b93736916b`.

Fresh TorchCFM manifest `f2c77ebf83d125c4246a77503d96f13 a982a3f70a7dc463653bfb7816106c18b`.

Fresh TorchCFM selection `e9637a4534ccbdca05d17131e632df81 5633c32f0cad82b49d825e9a7ac8bf0f`.

Fresh TorchCFM canonical `fc0b16ebc3c666d530e24a8f6e48612 d417e911596ca92639ed8c293653996ad`.

SelfFlow individual cohort `a08aad424d5239d6314a9cd1b02ca57 106bd80a8329bb1512f83906bb81095b2`.

SelfFlow failed-seed replication `7cf6ac2e9195a7bbdc78bb5afc88e4 b9a5cde5d47fb4b81ed034ff7ba2ad1be7`.

SelfFlow gated cohort `45c2e86fc1145f298a99ade15492e4cb e61fa9d74f81435b552879e23e981399`.

SelfFlow robust consensus `8ee4a3aebf76f262cb6e39deddc3de0 1bf88834806e08af52615727bca39c7b1`.

DiT consensus NFE8 `4e6ea02129c1589fe8145c215877310c bcc23041acbcd941a632d9291e9a2081`.

DiT consensus NFE12/24 `ab78c18f52d729b1ad2c2b557591b90 189a6907003f96c7100600dce34208670`.

DiT consensus NFE48/64 `a23788f0be5ffc906950943dffd6327c 4b097358cdf6c010832bd595aacba838`.

Clock–induced-grid FID `3b7630003f646a772e28427c686810f ff9095b15359aeb8b7964161e8551e08d`.

Prospective SelfFlow selection `1826964751117e4379bb7cae500ae8d 90991b5f8e41c7d68634a75f7fdcf6246`.

Prospective SelfFlow FID `e7ef0c11ab1bdae313002ec8fdfa2ea6 faa629735d7868d7e1f11daee38f115f`.

Budget-policy protocol `5801efb9d93aeab03fc68b396d29e08b 3276070dbe2dbf9c46cd2ef233fc2ae3`.

Frozen budget policy `02a41febf068188592502ed8945eebf21 77d527f731bd1c1fbd6109964924692`.

Budget-policy canonical `c7c1e92cabccc70482cb4d6a8d28999f 22a3f0d2635b398ebdf7bedd1956a381`.

Analytic-grid protocol `52f534a668fc30e388751f0971509b39 3a23d17854fa7cf24e28bfff9957ffa2`.

Analytic schedule `ac3dfd22223aebd22e48014739da0dc8147c2169e01c83df430b586b592c1e6e2`.

Analytic-grid canonical `3acba2c2dd13cd97a56311ba5c3a9b046f4b5d762da5fe57cf614b4913c33795`.

FlowTurbo receiver FID `c95ec7089a326edfddcf9971722e466fd878eccb507a04700050224ac17b8a89`.

FlowTurbo ranking `7b2b8e7dbfee248ee4acbcf8aa772d642acfca460550ca7745159fac8e74f7c6`.

FlowTurbo DINOv2 `96af91d140a863d25a069a65c744d9491c6e025482b6580f8520be3444cf48fb`.

REPA cubic canonical `1177ce7dc7e825a277ffb97483b9a3989c2c93e515a945d3a089e0168867d40b`.

REPA analytic control `1c14802c1bea90b9678a2ea6ed8cd2fe2f6dc6b42b360486719a7368b7ccf837`.

REPA prospective FID `873ffc2d35bd7b4860bddb0ea07c1f4abd32286b9c9f32460c7c6a1482055b50`.

REPA prospective DINOv2 `2a8886c235284e76fe5f31bcda67f48f1fbf62cf9c3d62cf077097047e3bb32c`.

REPA latency `fd23b1721b71cf6acfd87063877d4d9fc98ee37300690bc4696a2ac24076962f`.

Matched-cost terminal acquisition `ae8ead59bb2ef0c7ffdf60339a0d18e47f5dc39498894e7f87c4114220bc0acf`.

Gradient-measure DiT FID `4f3a1274d3a2da3beda946dfee2a6cc9dc72de4279abcfef5c6be02ad25929db`.

Gradient-measure SiT transfer `80771b4630227ce3994db6bbfbe419c0aed24e13f0e5bcfb4a48d54f16cbee5a`.

Gradient-measure SelfFlow transfer `20be396a4f35a1142372e6d6fe7ec6cf2d4408c2003f968bcc2a2f4d2af6cb5e`.

Gradient-measure NFE16 refinement `1eb790d56e13fb70bcd37a0bedc55a827cf435f07569eb3cb063a6127c55ad3d`.

Consistent-clock headroom audit `72451f4d69b08b249e60ab64884fcc076334dc91fc86cff818a06c1c510cf002`.

Dynamic-controller causality audit `30ee7b95e635126bb25a9a9090a8a22df044828bd7634f6b003d607850479859`.

Independent terminal acquisition `84202e6c5b9d916b76375c11d9124038cb7fa0a94bc291653cba7135fe672d7c`.

Source cross-bank replication `02fdb572735d648cf7d1b254ab2595699bc337d8941d61e5a4fc0dd6bd117fb6`.

SiT/SelfFlow zero-fit transfer `ba08ca08f9acdc9941ced7242f97e063e67c44d3256b8013fdb58e00e3873d38`.

Canonical target FID50k `20db7b727dc1db75932fde7f1ccb933981b8170a0eefa02258b46a53529aa343`.

Unseen Ralston transfer `ae13da1c09780159b8e6198a15fc9a5d87d0b9cc4f26a1bb62a84b2764db8c2e`.

Ralston parity `1ec9dbc1bf648c6e4bc146c07c7eda72e95c2e3142a17d0ae12d3b293f8264c7`.

Power-law schedule `4991e304d222e95a5d7d71fbabfbdbcd7315666e83542b385e3bf9e8410842b9`.

Power-law source audit `6191a4f94ca489119ef82a8ed0d371438d9465fd0b8b90ff3afbabcf82b538d4`.

Power-law target transfer `249b00911a67e58798eb1a29e8f6beac dfefc45f496b7408f5c4007cf07be785`.

NFE16 density refinement `5620c6b3478773cbe614b44ca87f2383 75245d65f45c7e218c3b08605be76942`.

Numerical-teacher FID `8adc3c3742d2e632e8736cf6be010abc f8d6a14b528d6cb1aa1a49e2cf26c225`.

Numerical-teacher endpoint audit `2d2debdb0a6869ca4b3c77bc3da0434 f353dc0007415d179c4848d0ef70189d4`.

Source TORS-style control `6546125493071a9b8c811e345ed1c6af 9bd9961bbe9b030ecd735568f5e517fd`.

Target TORS-style schedule `79a4021bddcd11287197bb5dae9a96da6 18ee0c733775c5ce6844b9c076c0fa6`.

Target TORS-style comparison `12f0ade5f84771eb4dbcd4344eb30b96 0c1fccdada2eebe6fcc43603fac424b5`.

Multidomain protocol `2e08e749fd4706eaa3834a4d5c450b66 197b770faa226b4e74d34a3c2b6fd46e`.

Multidomain canonical `8a18eba0569ecaf31825ebddf9f661222 f4b3b3d4a85c615af4780af3f5be332`.

Goal-density profile `536f8aadc577e6f012b13932fc3cb07c 8a0bd0c01d0ac55f06c214d0381ef3f4`.

Finite-NFE8 mass geometry `a0d431b809c6a3306efe30cd6538a33 af97bff74eb2c4f9261974adb440144c4`.

History-conditioned forecasting `83a00e4f44b6b8ca1f5d973d6ec331 e41cf72d1d4c75dfaccc41af7fcfd33025`.

Completion audit file `5d326c2dc3f91d5aff4368e3d3ba3786 ffe150cc15056f070e4c39d2c30c07d3`.

All long experiments use tmux workers, a five-minute watcher, terminal status files, exact artifact-count requirements, and success/failure notifications. The pre-consensus requirement-level audit independently verifies all frozen controller inputs, protocols, raw artifacts, proxies, final cells, selector argmins, intervals, gates, and canonical sidecars (144/144 requirements). The later consensus and analytic-grid chain is independently anchored by the canonical hashes above and by its frozen protocol files. The subsequent goal-measure chain is anchored separately by the source-acquisition, cross-bank, transfer, compression, refinement, teacher, and rotation-control hashes listed above. The multidomain protocol and canonical result are anchored separately; target-access fields are stored in the canonical records. The direct density-profile, finite-NFE8 geometry, and history-conditioned forecasting studies are independently anchored by their canonical hashes above.

An anonymous minimal reproducibility bundle accompanies the submission. It contains the frozen controller JSON, disjoint seed manifest, exact proxy and selection implementations, sealed expected selection, and an SHA256 manifest. It reproduces the path 128 noises $\to$ five doses $\to R_\lambda \to \widehat{\lambda}$ without access to final FID. An immutable external archive timestamp will be added only through a venue-compliant anonymous link.

24 Broader limitations

The evidence covers CIFAR-10, two independent model implementations, four solvers, five NFE budgets, fresh DiT cohorts, ten newly trained TorchCFM pairs, and four ImageNet-256 receiver configurations: SiT, SelfFlow, FlowTurbo, and REPA. Calibration uses a fixed feature space and projected Fréchet proxy; the repeated individual-fit failures show that this proxy is not a per-artifact safety certificate. Robust consensus succeeds on fresh SelfFlow streams and transfers to DiT, but it is built once from SelfFlow-aware fits and has not yet been rebuilt on a second large-scale training cohort. Its one-parameter analytic compression spans only one smooth family. The modern receiver-local study closes the previous strong-checkpoint gap, and the REPA analytic grid repeats across

two disjoint cohorts with DINOv2 agreement. FlowTurbo and REPA are nevertheless single frozen checkpoints, so their paired-stream sign counts do not substitute for independent training replication. The DiT transfer has weak absolute quality at NFE8 and should be read as numerical external validity, not an SOTA sampler comparison. We do not require every architecture to use an interior dose: the stable TorchCFM full-dose regime is itself a prospective outcome of the controller.

The goal-oriented measure transfers from CIFAR to SiT and SelfFlow and admits exact midpoint NFE8-to-NFE16 density refinement. Canonical FID50k compares all three measures on both targets; unseen-Ralston adds one zero-fit solver test. Both receivers remain single checkpoints with one paired FID50k stream, and Ralston remains one RK2 tableau. Independent acquisition supports a stable family, not a unique mass vector; the power law spans one analytic family, and the full residual is tested only for this four-bin source construction. The asymptotic density assumes a nonnegative leading goal-error density and does not certify a finite-bin global optimum. The failed direct cube-root profile shows that it does not identify the successful finite-NFE8 measure. The finite-dimensional audit supports a reproducible first-order direction but neither exact KKT stationarity nor stable quadratic/Hessian geometry. Per-sample timing remains untested. The TORS-style control is adapted to our SiT/midpoint setting rather than reproduced in its official Flux/SD3 setting.

The domain-local extension broadens receiver-seed evidence but not raw-modality coverage. ETTm1, FSDD, and AG News are deterministic vectorizations, and the Lorenz condition is a system parameter rather than a causal history. ETTm1 and AG News pass both distribution-witness checks; FSDD is inconsistent and Lorenz distribution fidelity fails. Moreover, every selected domain-local schedule worsens dense-reference endpoint MSE. These results support goal-specific finite-step transfer within the tested representations, not a common cross-modal mass vector, speech or text generation, forecasting, chaotic rollout, or generic numerical accuracy.

The separate history-conditioned study does establish genuine future prediction on ETTm1 and trajectory-disjoint Lorenz splits, but only at the frozen midpoint NFE8 contract. ETTm1 CRPS gains are small and do not extend to point RMSE; Lorenz gains cover a short horizon and largely sharpen an over-dispersed predictive distribution. The 6/6 NFE4 reversal and worse dense- endpoint agreement rule out budget-universal scheduling, uniformly better mean trajectories, or improved integration. The study is not an SOTA forecasting comparison and does not establish long-horizon chaotic fidelity.