CONDITIONAL MEAN DEPENDENCE GOVERNS ONE-STEP ENDPOINT REGRESSION

Anonymous authors
Paper under double-blind review

ABSTRACT

One-step generative policies may match iterative samplers or collapse to a conditional mean. We characterize this outcome for L^2 endpoint regression. Let ρ denote the ratio of target within-condition variance reproduced by a one-step map. For the optimal map, $\rho = \rho^*$, the squared correlation ratio between source noise and target given context. Collapse occurs exactly when the target's conditional mean is independent of the source noise; full statistical independence is sufficient but unnecessary. Moreover, ρ^* depends on the training pairing, not on the target marginal alone. A fixed-marginal coupling family makes ρ^* span $[0, 1]$, giving every target-marginal-only estimator worst-case absolute error at least $1/2$; paired training samples permit estimation under regularity assumptions. Synthetic experiments recover the predicted curve. A cost-matched construction further shows that quadratic transport cost does not identify endpoint retention. An area-downsampled full-image MNIST check recovers the predicted ordering but exposes the failure of a generic high-dimensional estimator. An exploratory OGBench test recovers the map-spread ordering yet refutes the behavioral prediction: the collapsed policy succeeds more often. Thus the criterion diagnoses endpoint-map spread, not task utility.

1 INTRODUCTION

Generative policies—diffusion and flow-matching action heads—are a common way to represent multimodal behavior in continuous-control policies (Chi et al., 2023; Wang et al., 2022). Their cost is iterative sampling, so a large literature compresses sampling to one or a few network evaluations: consistency models (Song et al., 2023), shortcut models (Frans et al., 2024), mean flows (Geng et al., 2025), rectified-flow reflow (Liu et al., 2022), distillation (Salimans & Ho, 2022; Yin et al., 2023), and in RL one-step flow policies (Nguyen et al., 2026; Li et al., 2026a; Zhan et al., 2026).

The reported behavior varies across tasks: one-step policies can outperform longer diffusion samplers (Li et al., 2026a; Nguyen et al., 2026), but their advantage can shrink with weaker conditioning or longer action horizons (Chen et al., 2026).

We ask when a single endpoint-regression step retains conditional spread. Prior work proves collapse under independent pairing. We show that the exact answer depends on the source–target pairing and is therefore not identified by $p(c, x_1)$ alone (Fig. 1).

What we mean by conditional-spread retention. For a one-step map f we define the realised conditional-spread ratio

$$\rho(f) \;=\; \frac{\mathbb{E}_c\big[\,\mathrm{tr}\,\mathrm{Var}_{x_0}\left(f(x_0, c) \mid c\right)\big]}{\mathbb{E}_c\big[\,\mathrm{tr}\,\mathrm{Var}(x_1 \mid c)\big]} \;\in\; [0, \infty), \tag{1}$$

i.e. the ratio of the target's conditional spread that the map reproduces, where Var_{x_0} is over the source noise at fixed c and the denominator is assumed positive and finite. $\rho \approx 0$ signals collapse onto a per-condition point mass; $\rho \approx 1$ only matches aggregate spread and does *not* imply the correct conditional law, mode weights, or sample quality.

Contributions.

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

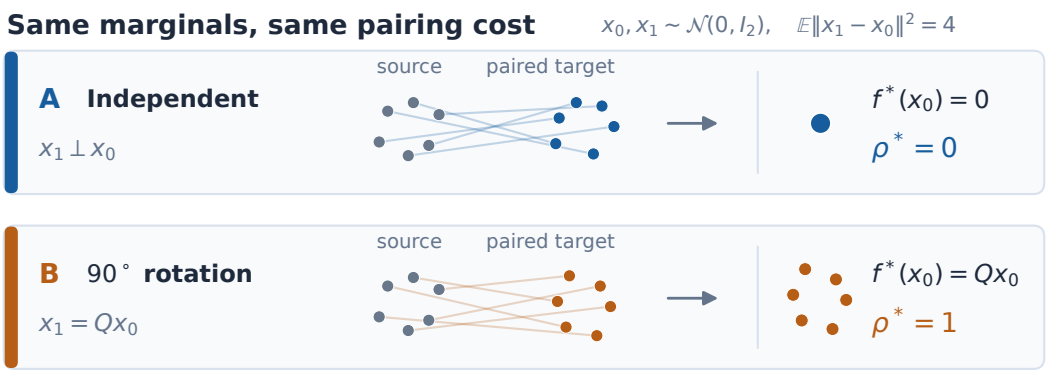


Figure 1: **One target law, two endpoint outcomes.** Under the two couplings, both marginals and the expected quadratic pairing cost agree, yet the optimal pointwise L^2 maps retain opposite amounts of source variation (Prop. 1). Dots schematize the pairings; they are not experimental measurements.

- **C1 Exact criterion.** The L^2-optimal endpoint map is $f^*(x_0, c) = \mathbb{E}_\pi[x_1 \mid x_0, c]$, and $\rho(f^*)$ is the classical squared multivariate correlation ratio (Rényi, 1959; Sampson, 1984). Collapse is equivalent to conditional mean independence, not full independence (Thm. 1).
- **C2 Non-identifiability and its escape.** A fixed-marginal coupling family has $\rho^*(\alpha) = \alpha^2$, so any target-marginal-only estimator has minimax absolute-error risk $1/2$ (Thms. 2–3). The paired law identifies ρ^*, while equal quadratic transport costs can have opposite retention (Prop. 1; Sec. 5.5).
- **C3 Empirical scope.** Controlled ring and full-image MNIST studies support the criterion, isolate matching scope and pairing, and separate transport cost from retention (Secs. 5.4 and 5.6; Appendices D.2–D.3). An exploratory OGBench test reproduces the map-level ordering but refutes its behavioral implication (Appendix E).

Scope. Claims are either exact theorems or measurements under a specified training target. Pilot-calibrated checks and post-hoc choices are marked explicitly; Appendices G and B give the full scope and corrections.

2 Setup and the Classical Dependence Quantity

Let $c \in \mathcal{C}$ be a conditioning variable (observation, state, goal, embedding) and $x_1 \in \mathbb{R}^d$ a target (action, action chunk, image), with joint law $p(c, x_1)$ observed only through i.i.d. samples. A one-step generative policy is a map $f : \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}^d$ applied to source noise $x_0 \sim \mathcal{N}(0, I_d)$. Flow and diffusion policies typically sample iteratively; flow matching learns a velocity field from a *coupling* π between x_0 and x_1. Throughout, π is the joint law of (x_0, x_1) given c, with both marginals fixed: $x_0 \sim \mathcal{N}(0, I_d)$ and $x_1 \sim p(\cdot \mid c)$. The two couplings that matter in practice are the *independent* coupling $x_0 \perp\!\!\!\perp x_1 \mid c$ used by vanilla conditional flow matching (Lipman et al., 2022; Tong et al., 2023), and a *deterministic* coupling $x_1 = T(x_0, c)$, which a deterministic teacher or reflow (Liu et al., 2022) can induce.

Theorem 1 (Optimal endpoint map and mean-dependence criterion). *Assume $\mathbb{E}\|x_1\|^2 < \infty$ and $0 < \mathbb{E}_c \operatorname{tr} \operatorname{Var}(x_1 \mid c) < \infty$, and fix a coupling π. The minimizer of the pointwise L^2 risk $\mathbb{E}_\pi\|f(x_0, c) - x_1\|^2$ over all f with $\mathbb{E}\|f(x_0, c)\|^2 < \infty$ is $f^*(x_0, c) = \mathbb{E}_\pi[x_1 \mid x_0, c]$, and*

$$\rho(f^*) = \frac{\mathbb{E}_c\big[\operatorname{tr} \operatorname{Var}\big(\mathbb{E}[x_1 \mid x_0, c] \mid c\big)\big]}{\mathbb{E}_c\big[\operatorname{tr} \operatorname{Var}(x_1 \mid c)\big]} \in [0, 1],$$

with the bound following from the law of total variance, $\operatorname{Var}(x_1 \mid c) = \mathbb{E}[\operatorname{Var}(x_1 \mid x_0, c) \mid c] + \operatorname{Var}(\mathbb{E}[x_1 \mid x_0, c] \mid c)$. Moreover, $\rho(f^*) = 0$ iff $\mathbb{E}[x_1 \mid x_0, c] = \mathbb{E}[x_1 \mid c]$ almost surely (conditional mean independence), and $\rho(f^*) = 1$ iff $\operatorname{Var}(x_1 \mid x_0, c) = 0$ almost surely.

Proof sketch. The conditional mean is the L^2 projection; the law of total variance gives the decomposition and hence $\rho \le 1$; $\rho \ge 0$ is trivial. Appendix A.

The ratio in Thm. 1 is the squared multivariate correlation ratio, taken within each context and therefore normalised by within-context variance (Rényi, 1959; Sampson, 1984). Concurrent statistical work calls the same trace-normalised quantity the normalised conditional mean discrepancy and gives a consistent nearest-neighbour-graph estimator and test (Chatterjee et al., 2026). Neither the projection identity, the measure, nor general-purpose estimation is new. What is new is only the use: we replace full independence (too strong) by mean independence, the exact condition governing squared-loss endpoint collapse. Changing the loss can also avoid the pointwise conditional mean; no universal converse is claimed.

Corollary 1 (Independent coupling collapses)**.** *If $x_0 \perp\!\!\!\perp x_1 \mid c$ then $\mathbb{E}[x_1 \mid x_0, c] = \mathbb{E}[x_1 \mid c] = m(c)$ is constant in x_0, so $\rho(f^*) = 0$: the optimal one-step map is a per-condition point mass.*

The converse of Cor. 1 is false. For example, with no context, let $x_0 \sim \mathcal{N}(0, 1)$, let S be an independent Rademacher sign, and set $x_1 = S|x_0|$. Then $x_1 \sim \mathcal{N}(0, 1)$ and the coupling is dependent because $|x_1| = |x_0|$, yet $\mathbb{E}[x_1 \mid x_0] = 0$ and $\rho(f^*) = 0$. Thus a non-independent coupling need not prevent mean collapse.

Corollary 2 (Deterministic coupling does not)**.** *If $x_1 = T(x_0, c)$ with $T(\cdot, c)_{\#}\mathcal{N}(0, I_d) = p(\cdot \mid c)$, then $f^* = T$ and $\rho(f^*) = 1$.*

Proposition 1 (Transport cost does not identify retention)**.** *Let d be even and Q be block-diagonal with 90° rotation blocks. For $x_0 \sim \mathcal{N}(0, I_d)$, both the independent coupling $x_1 \perp\!\!\!\perp x_0$ and the deterministic coupling $x_1 = Qx_0$ have target marginal $\mathcal{N}(0, I_d)$ and $\mathbb{E}\|x_1 - x_0\|^2 = 2d$, but their optimal endpoint ratios are respectively 0 and 1. Hence even the population quadratic coupling cost cannot determine endpoint retention.*

Independence and determinism give the two ratios; since $\operatorname{tr} Q = 0$, $\mathbb{E}\|(Q - I)x_0\|^2 = 2d$. Appendix D.2 verifies the separation with finite networks and adds mean-independent and higher-cost deterministic controls.

Corollary 1 in closed form. On a Gaussian-mixture conditional whose exact velocity field integrates in closed form, the unit Euler endpoint matches $m(c)$ to 8.882×10^{-16} (multi-step ratios in Appendix A). Under a multi-condition version, changing *only* the coupling moves one-step ρ from 0.02722 to 0.8776 at $k = 30$—a factor of 32.24—while the data statistic $\hat{D}$ takes the value 0.9116 in *both* arms, to a measured difference of 6.661×10^{-16}. That invariance is exact, not a near-miss: $\hat{D}$ is a functional of the (c, x_1) pairs alone, and switching the coupling only reorders those pairs within each condition.

3 The target marginal cannot decide

Mean independence and deterministic transport are properties of the training coupling. We now ask whether the observed target law $p(c, x_1)$ alone determines which regime applies. A fixed-marginal construction gives a negative answer and an exact estimation lower bound.

3.1 A family with fixed data and sweeping ρ

Fix $p(\cdot \mid c)$ with finite second moment. Assume a measurable conditional map $T(\cdot, c)$ pushes $\mathcal{N}(0, I_d)$ to $p(\cdot \mid c)$; a deterministic teacher is one possible source of such a map. For $\alpha \in [0, 1]$ define the α-*interpolated coupling*

$$\begin{aligned} &x_1 = B\,T(x_0, c) + (1 - B)X, \\ &B \sim \text{Bernoulli}(\alpha), \quad X \sim p(\cdot \mid c), \\ &B \perp\!\!\!\perp (x_0, X) \mid c, \quad X \perp\!\!\!\perp x_0 \mid c. \end{aligned} \tag{2}$$

Theorem 2 (Fixed data, sweeping ρ)**.** *For the coupling* (2)*, with $\mathbb{E}\|x_1\|^2 < \infty$:* (i) *both branches have marginal $p(\cdot \mid c)$, so the law of (c, x_1) does not depend on α;* (ii) $f^*_\alpha(x_0, c) = \alpha T(x_0, c) + (1 - \alpha)m(c)$*; and* (iii) $\rho^*(\alpha) = \alpha^2$.

Proof. (i) $T(x_0, c) \sim p(\cdot \mid c)$ because T is a transport map, and $X \sim p(\cdot \mid c)$ by construction; a mixture of two identical laws is that law. (ii) $\mathbb{E}[x_1 \mid x_0, c] = \alpha\mathbb{E}[T(x_0, c) \mid x_0, c] + (1 - \alpha)\mathbb{E}[X \mid c]$.

(iii) Given c, the conditional variance of $\alpha T(x_0, c) + (1-\alpha)m(c)$ over x_0 is $\alpha^2 \operatorname{Var}(x_1 \mid c)$, since the second term is constant and $\operatorname{Var}(T(x_0, c) \mid c) = \operatorname{Var}(x_1 \mid c)$.

The family is not exotic: $\alpha = 0$ is the standard independent coupling of vanilla conditional flow matching, $\alpha = 1$ is a deterministic coupling attainable by a teacher map, and intermediate α is a mixture of the two. What varies is a design choice; what does not vary is the law of a single observed pair $(c_i, x_{1,i})$.

3.2 Minimax form: the target marginal is insufficient

Theorem 3 (Impossibility, minimax form)**.** *Let $\mathcal{A}$ be any estimator of ρ^* measurable with respect to n i.i.d. draws from $p(c, x_1)$—randomised, and allowed to know the population law $p(c, x_1)$ exactly. Then for the family of Theorem 2,*

$$\inf_{\mathcal{A}} \sup_{\alpha \in [0,1]} \mathbb{E}_\alpha \big|\mathcal{A} - \rho^*(\alpha)\big| \;=\; \tfrac{1}{2} \;=\; \inf_{a \in \mathbb{R}} \sup_{\alpha \in [0,1]} \big|a - \alpha^2\big|,$$

and the infimum on the left is attained by the constant $\mathcal{A} \equiv 1/2$.

Proof. By Theorem 2(i) the law of the sample is the same for every α, so the law of $\mathcal{A}$—and in particular $\bar{a} := \mathbb{E}[\mathcal{A}]$—does not depend on α. By Jensen, $\mathbb{E}_\alpha|\mathcal{A} - \alpha^2| \geq |\bar{a} - \alpha^2|$ for each α, hence $\sup_\alpha \mathbb{E}_\alpha|\mathcal{A} - \alpha^2| \geq \sup_\alpha |\bar{a} - \alpha^2| \geq \max\{|\bar{a}|, |\bar{a} - 1|\} \geq 1/2$, the last step because α^2 attains both 0 and 1. Conversely $\mathcal{A} \equiv 1/2$ achieves $\sup_\alpha |1/2 - \alpha^2| = 1/2$. The second equality is the same argument applied to constant predictors.

The bound follows from the fixed-marginal construction: no function of the observed (c, x_1) law can distinguish its couplings.

Four remarks. (a) The bound is free of n: it is an information deficiency, not an estimation problem, and $n = \infty$ does not help. (b) It uses only that the *distribution* of $\mathcal{A}$ is α-free, so it also covers statistics given the population law or the conditioning c. (c) The value $1/2$ is attained, so within this constructed family the target marginal has zero decision value for estimating ρ^*; this statement does not apply once the training pairing is observed. (d) Both endpoints are couplings in routine use ($\alpha = 0$ and 1), so the bound forces every data-side predictor to miss by at least $1/2$ on one of them; a practitioner holding an intermediate mixture cannot even tell which regime they are in from the target marginal alone.

What the theorem does not say. It does not say ρ is unknowable. ρ uses the pairing, and the pairing is available: every training step of a flow model draws a paired batch $(x_0^{(i)}, x_1^{(i)})$. Estimating ρ from those pairs requires computation but no additional data or trained sampler, and can be done before the one-step model exists. The theorem says the pairing carries information that the target marginal omits.

4 Two design levers in the controlled study

4.1 Lever 1: change the coupling

Corollaries 1–2 give two endpoints, while the counterexample after Cor. 1 shows that "dependent" alone is insufficient. Mini-batch OT (Tong et al., 2023), semidiscrete OT (Mousavi-Hosseini et al., 2026), reflow (Liu et al., 2022), and teacher-map distillation often strengthen source–target mean dependence. ODE-free Neural Flow Matching already proved mean collapse for independent endpoint regression and proposed OT pairings (Shou, 2026); our distinction is that the exact condition is mean dependence, not merely non-independence. Table 1 is a controlled implementation study, not an exhaustive classification of published methods.

4.2 LEVER 2: CHANGE THE LOSS FORM

The less explored channel is to keep the independent coupling and change the loss. Let $q_c = f(\cdot, c)_{\#}\mathcal{N}(0, I_d)$ denote the law the one-step model produces for condition c, and let the training batch contain M paired samples $\{x_0^{(i)}, x_1^{(i)}, c^{(i)}\}_{i=1}^M$ with $G_i = f(x_0^{(i)}, c^{(i)})$.

Theorem 4 (Population limits of three objectives)**.** *Assume finite second moments, that each $p(\cdot \mid c)$ is a pushforward of $\mathcal{N}(0, I_d)$ by a measurable map, and the standard empirical-measure convergence conditions; for **(L2)** restrict both $p(\cdot \mid c)$ and candidate q_c to compact support. Under the independent coupling, as $M \to \infty$: **(L1)** the pointwise loss $\frac{1}{M}\sum_i \|G_i - x_1^{(i)}\|^2$ has unique population minimizer $q_c = \delta_{m(c)}$, giving $\rho = 0$; **(L2)** the* unbalanced *collection loss, with both* $\arg\min$*s taken* within the condition*,*

$$\frac{1}{M}\sum_i \min_j \|x_1^{(i)} - G_j\|^2 + \frac{1}{M}\sum_j \min_i \|G_j - x_1^{(i)}\|^2 \;\longrightarrow\; C(p_c, q_c),$$

$$C(p_c, q_c) := \mathbb{E}_{y\sim p_c}\big[\min_{x\in \operatorname{supp} q_c} \|y - x\|^2\big] + \mathbb{E}_{x\sim q_c}\big[\min_{y\in \operatorname{supp} p_c} \|x - y\|^2\big],$$

*the sum of two nearest-neighbour distances: it vanishes when the two closed supports agree and is otherwise silent about how mass is split across them (supports identified only up to measure-zero/closure qualifications); **(L3)** the* balanced *collection loss, $\min_\sigma \frac{1}{M}\sum_i \|G_i - x_1^{(\sigma(i))}\|^2$ over bijections σ within each condition, converges to $W_2^2(p_c, q_c)$, whose unique minimizer is $q_c = p_c$ and $\rho = 1$.*

Proof sketch. L1 is Cor. 1. For L2, the two terms are non-negative and vanish exactly when each support is contained in the other's closure; no term depends on the mass assigned to a region once it is non-empty. For L3, minimizing over bijections is exactly the optimal-transport problem between the empirical measures, whose $M \to \infty$ limit is W_2^2, minimized uniquely at $q_c = p_c$. Appendix A.

The L2/L3 distinction concerns population minimizer sets, not finite-network training dynamics. Whether an optimizer uses the freedom in L2 is empirical. In our task it does not: the observed flattening appears only when matching is pooled across conditions.

5 EXPERIMENTS

Common design. Within each stated comparison we hold the synthetic data, MLP width/depth, optimizer, batch size and per-stage update count fixed, so capacity or optimization differences cannot explain the measured contrasts. Methods with a teacher or reflow stage use more total compute, so this is not a compute-matched ranking. Metrics are the spread ratio ρ of (1) (with a kNN estimator, Sec. 5.5), mode coverage (a mode counts as covered if any sample falls within 0.6 mode-widths of its center), and conditional sliced W_2 (cSW) normalized by $\sqrt{\operatorname{tr}\operatorname{Var}(x_1 \mid c)}$. The synthetic tables use five seeds and the MNIST study uses 10; the analytic α construction uses three independent Monte Carlo repetitions. Runs reported in the main tables and Sections 5.1–5.7 are CPU-only; the separate dimension-stress check is the only GPU run (Appendix D.3).

5.1 $\rho^*(\alpha) = \alpha^2$ WITH THE DATA HELD FIXED

We instantiate (2) with 4 conditions, each an 8-mode ring of width σ (a scale, not the permutation σ of Thm. 4), and T the 512-step Euler flow map of the exact conditional velocity field. We sweep $\alpha \in \{0,\ 0.25,\ 0.5,\ 0.75,\ 1\}$ with a *fresh* independent sample at every level—reusing one realization would make the data-side statistics constant by construction, restating the theorem rather than testing it. The kNN estimate of ρ moves 0.05006→0.1054→0.2854→0.5747→0.9766, against the analytic α^2 and its ideal local-neighbour finite-k correction $\alpha^2 + (1-\alpha^2)/k$ at $k = 20$; the mean absolute deviation from the corrected prediction is 0.007836 (pre-registered criterion < 0.08). Over the same sweep, the data-side statistics move by at most 0.002993 ($\hat{D}$) and 0.00028 ($\hat{\Lambda}$)—i.e. by Monte-Carlo noise—while ρ^* moves by 1.0 (Fig. 2a). All five pre-registered criteria pass (verdict PASS).

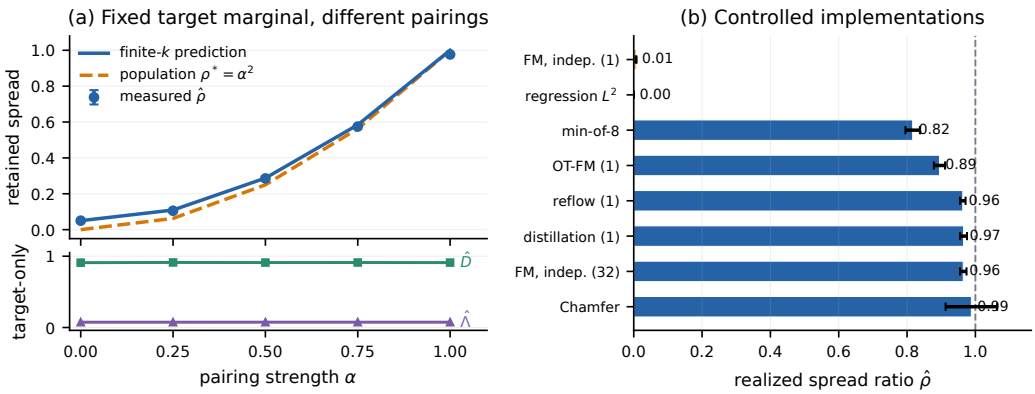


Figure 2: **Target marginal versus training pairing.** (a) ρ against α on the family of Thm. 2: measured (circles) versus the analytic α^2 (dashed) and its finite-k correction (solid); two target-only statistics remain flat while ρ traverses $[0, 1]$. (b) Controlled method implementations with cross-seed SD: a finite comparison, not a universal dichotomy.

5.2 A BATTERY OF DATA-SIDE STATISTICS IS FLAT IN α

Theorem 3 is exact and needs no experiment; what needs checking is its empirical premise—that statistics one would naturally use are, in fact, flat. We test seven: $\hat{D}$ (unexplained conditional-variance fraction), $\hat{\Lambda}$ (Gaussian deficit), the number of clusters selected by Calinski–Harabasz, mean nearest-neighbour distance, the k-means between/within variance ratio at $K = 8$, excess kurtosis, and the marginal trace variance.

Because these statistics differ by orders of magnitude in their own sampling noise, a fixed relative threshold tests the noise rather than the α effect; we therefore test with a one-way ANOVA over α with seeds as replicates. At level 0.05 with five seeds per level, no statistic shows a detectable α effect; the largest relative deviation from $\alpha = 0$ is 0.06374. By contrast the paired estimator's mean absolute deviation from α^2 is 0.01405—35.58$\times$ smaller than the 0.5 that Theorem 3 imposes on *any* data-only predictor.

Sample size does not rescue the data side. Increasing n from 1000 to 16000 leaves the $\alpha = 0$-versus-$\alpha = 1$ difference of every $[0, 1]$-scaled statistic below 0.003—at most 0.3% of the 1.0 a useful predictor must resolve—while the paired estimator, at the same budget, tracks α^2: a real α-signal would converge to a nonzero constant as n grows, and these differences do not. (The two ratio-scaled statistics are covered by the scale-free ANOVA above.)

5.3 WHAT DOES DECIDE: THE METHOD MAP

Table 1 runs six method families plus two one-step-only variants on the same conditional ring mixture. In these implementations, both collapsing rows use independent targets with pointwise L^2: flow matching at $N = 1$ gives $\rho = 0.00669 \pm 0.00159$ with coverage 0.05/8, and a direct one-step regression gives 9.815×10^{-5}. Removing *either* condition removes the collapse in this task—by the coupling (mini-batch OT 0.8949, reflow 0.9631, distillation 0.9654) or by the loss form with the coupling left independent (pooled bidirectional Chamfer 0.9882, re-measured as L2a in Sec. 5.4). The same field that collapses at $N = 1$ recovers at $N = 32$ (0.9646), showing that the same learned velocity field can recover the distribution when integrated. This finite method set does not establish the if-and-only-if claim, which is stronger than what a finite comparison can support.

5.4 LOSS COMPARISON: MATCHING SCOPE MATTERS IN THIS TASK

We hold the independent coupling fixed and vary only the loss: pointwise L^2 (L1), pooled Chamfer (L2a), within-condition Chamfer (L2b), hard balanced assignment (L3), and its Sinkhorn relaxation (S). L2a–L2b isolates matching scope; L2b–L3 isolates balancedness; L3–S isolates hard versus soft balance. Four ring radii make incorrect condition-specific spread directly visible.

Table 1: Controlled method implementations. Data, network width/depth, optimizer, batch size and per-stage step count are fixed; total compute is not equal because reflow and distillation include a teacher stage. The rows support the stated implementations only, not a converse over all one-step methods.

method (all at N= 1)	coupling	loss	ρ	modes / 8
flow matching	independent	pointwise	0.00669±0.00159	0.05
same field, N= 32	independent	pointwise	0.9646±0.009531	8
+ mini-batch OT	≈det.	pointwise	0.8949±0.0164	8
+ reflow	deterministic	pointwise	0.9631±0.008115	8
distillation of the 32-step map	deterministic	pointwise	0.9654±0.01008	8
one-step regression	independent	pointwise	9.815×10^{-5}±8.811×10^{-5}	0
one-step, min-of-M	independent	collection	0.8165±0.02117	8
one-step, bidirectional Chamfer (pooled)	independent	collection	0.9882±0.07542	7.6

Table 2: One-step loss comparison under the same independent coupling. L1 collapses; L2a flattens condition-specific spread; L2b restores it. L3 is best among hard assignments by conditional sliced W_2; Sinkhorn improves as ε decreases. $\mathrm{disp} = \mathrm{std}_j(\rho_j)$ and $|\rho_j - 1|$ is mean spread error.

loss	ρ	modes / 8	disp	$\lvert\rho_j - 1\rvert$	cSW
L1 pointwise L^2	$9.716 \times 10^{-5} \pm 8.884 \times 10^{-5}$	0	0.0001205	0.9999	0.7024
L2a unbalanced, pooled over cond.	0.9474 ± 0.04917	7.8	0.5082	0.4331	0.3973
L2b unbalanced, within cond.	0.972 ± 0.01528	8	0.01885	0.0271	0.1294
L3 balanced, within cond.	0.8991 ± 0.02395	8	0.01805	0.09696	0.07917
S Sinkhorn, within cond. ($\varepsilon = 1$)	0.8017 ± 0.03997	8	0.07585	0.2262	0.1176
S Sinkhorn, within cond. ($\varepsilon = 0.1$)	0.9238 ± 0.01795	8	0.0204	0.08063	0.08111
S Sinkhorn, within cond. ($\varepsilon = 0.01$)	0.9914 ± 0.01054	8	0.01516	0.01854	0.06141

Aggregate ρ misses the main failure: L2a and L2b score 0.9474 and 0.972, yet Fig. 3a shows L2a nearly horizontal and L2b near the diagonal. Thus pooling, rather than missing bijection, flattens condition-specific spread in this task. L2b and L3 both track the diagonal; L3 has lower cSW (0.07917 versus 0.1294), whereas L2b's spread error is lower by more than a factor of three (0.0271 versus 0.09696).

Dropping balancedness at fixed scope (L2b versus L3) did not flatten spread, refuting our planned prediction. Soft balance is sufficient: all Sinkhorn settings escape, and $\varepsilon = 0.01$ achieves the lowest cSW and spread error in Table 2. These results concern this controlled task; they do not establish a universal loss ranking.

A corrected mechanism: min-of-M. min-of-M (IMLE) also escapes ($\rho = 0.8165$, coverage $8/8$, against 9.815×10^{-5} for pointwise L^2) once its M candidates come from *independent* source noise; a negative finding from our own pilot run was an artifact of repeating one source draw, which makes the objective literally pointwise L^2 (Appendix B(iii)).

5.5 Estimating ρ from paired samples

ρ^* is the variance of a conditional mean, so the natural estimator is kNN regression of x_1 on x_0 within each condition. Under the ideal zero-neighbourhood approximation and an independent coupling, the k-neighbour average of i.i.d. targets has variance $\mathrm{Var}(x_1 \mid c)/k$, so $\hat{\rho}_{\mathrm{kNN}}(k) \approx \rho^* + (1 - \rho^*)/k$; at $\rho^* = 0$ this is the $1/k$ law, which we use as a check. Measured $k \cdot \hat{\rho}$ under the independent coupling is $1.002, 0.9775, 0.9369, 0.8885, 0.8257$ at $k = 1, 2, 5, 10, 20$: the law holds within about 11% up to $k = 10$ and degrades to 0.8257 by $k = 20$, where the zero-neighbourhood approximation weakens.

5.6 MNIST-derived endpoint study

We construct four endpoint tasks from MNIST images (LeCun et al., 1998), using area downsampling of the complete 28×28 frames to 16×16 and $x_0 \sim \mathcal{N}(0, I_{256})$. This follows an initial crop-based pilot; the four couplings, model, optimizer, seeds, and nine pilot-calibrated checks were held fixed for the full-frame run (Appendix C). Independent, mini-batch OT, and sorted-quantile pairings preserve the empirical image marginal. A separate heteroscedastic target, $x_1 = \mu + (1 + 0.8 \tanh x_{0,1})(\tilde{x} - \mu)$, is dependent on x_0 but satisfies $\mathbb{E}[x_1 \mid x_0] = \mu$. It does not

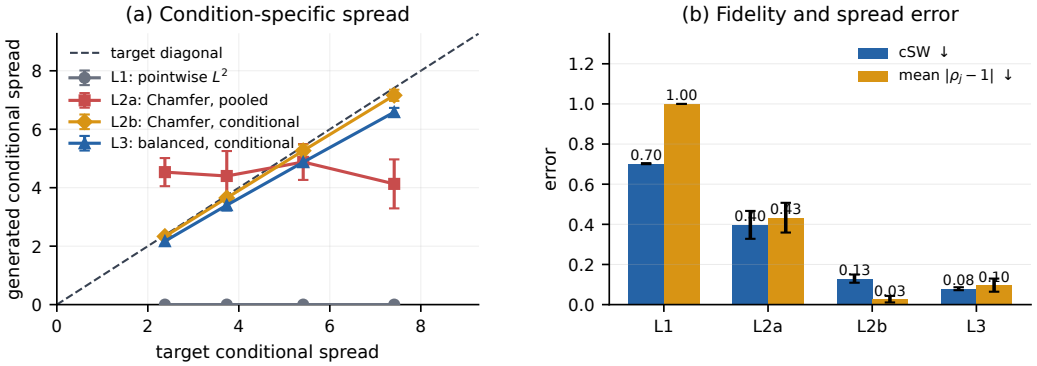


Figure 3: **Matching scope controls spread in this task.** Error bars are cross-seed SD. (a) L1 collapses; pooled L2a flattens condition-specific spread; within-condition L2b and L3 follow the target diagonal. (b) Conditional sliced W_2 and mean spread error (lower is better).

preserve the MNIST target marginal and may leave the image range, so it tests mean independence rather than fixed-marginal non-identifiability.

Across 10 seeds, independent and heteroscedastic targets collapse ($\rho = 0.0180 \pm 0.0006$ and 0.0158 ± 0.0006), despite measurable dependence in the latter. Sorted-quantile pairing retains intermediate spread (0.1578 ± 0.0086), while mini-batch OT is highest (0.2763 ± 0.0078). All nine previously fixed checks pass, and cSW fidelity follows the same ordering. Only the independent, sorted, and OT arms form a fixed-target-marginal comparison.

Seed 0, used to calibrate the crop protocol, appears in both ten-seed aggregates; the full-frame test checks preprocessing robustness rather than an independent benchmark. Generic kNN estimation remains uninformative at $d = 256$ and $n \approx 10^3$; the constructed couplings nevertheless predict the trained-map ordering (Appendix C).

5.7 Exploratory downstream stress test

The controlled experiments test the endpoint map directly. To check whether that map-level ordering transfers to decision making, we ran the same two couplings in the OGBench `pointmaze-medium-navigate-v0` suite. Each arm uses the same pointwise L^2 action head and training schedule; only the source–target pairing changes. The independent arm has nearly zero paired $\hat{\rho}$, whereas the sorted arm has a positive value, and the trained-map spread follows the same ordering on all ten seeds. Each arm and seed is evaluated on 50 episodes and five goals, giving 250 rollouts.

The behavioral result is the opposite of the simplest transfer hypothesis. The collapsed conditional-mean arm reaches success 0.2564 ± 0.1253, versus 0.1624 ± 0.05896 for the spread-preserving arm; the independent arm wins on $8/10$ seeds, with paired gap -0.094 and 95% t interval $[-0.1633, -0.02466]$. We therefore treat this experiment as a falsification of behavioral transfer, not as evidence for a policy improvement.

The negative result has a simple population explanation. Let $Y \sim p(\cdot \mid c)$ be an evaluation action drawn independently of x_0 given c, and write $m(c) = \mathbb{E}[Y \mid c]$. For any deployed map f,

$$\mathbb{E}\|f(x_0, c) - Y\|^2 = \mathbb{E}\,\mathrm{tr}\,\mathrm{Var}(Y \mid c) + \mathbb{E}\|\mathbb{E}[f \mid c] - m(c)\|^2 + \mathbb{E}\,\mathrm{tr}\,\mathrm{Var}(f \mid c). \tag{3}$$

Thus, if f has the correct conditional mean, its excess independent-draw imitation risk over the collapsed map is exactly its output variance, or $\rho(f)$ times the denominator of Eq. (1). This identity is classical and makes no claim about reward, but it shows why retaining behavior-policy noise can hurt imitation risk. In this suite the measured conditional spread does not increase with the conditioning horizon, so the sorted coupling retains action noise that the conditional mean removes. Task utility therefore needs a separate test that the retained modes are reward-relevant. Full protocol deviations and per-seed records are given in Appendix E.

6 Related work

Dependence and endpoint regression. ρ^* is the classical multivariate correlation ratio (Rényi, 1959; Sampson, 1984); concurrent work studies the same quantity with a nearest-neighbour graph estimator (Chatterjee et al., 2026). ODE-free Neural Flow Matching establishes collapse for independent endpoint pairs (Shou, 2026). We give the fixed-pairing criterion, including dependent but mean-independent pairs.

One-step and few-step generative models. Consistency and shortcut models, mean flows, progressive and distribution-matching distillation, and InstaFlow reduce the number of sampling steps (Song et al., 2023; Yang et al., 2024; Frans et al., 2024; Geng et al., 2025; Salimans & Ho, 2022; Yin et al., 2023; Liu et al., 2023b). Each targets a prescribed step budget rather than the fixed-coupling population optimum of direct endpoint regression, so an empirically successful step count does not settle whether that optimum is degenerate. Our analysis is complementary and does not rank their finite-model performance.

Pairing and coupling design. Flow matching can use independent or transport couplings (Tong et al., 2023; Liu et al., 2022); rectified flow and schedule design study path geometry and the training distribution (Liu et al., 2022; Esser et al., 2024), and recent designs add multisample (Pooladian et al., 2023), semidiscrete (Mousavi-Hosseini et al., 2026), and model-aligned or transport-based couplings (Lin et al., 2025; Malnick et al., 2026). Reference-guided flow matching uses a conditional endpoint mean at inference (Curvo et al., 2026), the object Thm. 1 identifies as the pointwise L^2 optimum. We show that pairing cost alone cannot determine retention and that matching within a condition differs from matching across conditions; further attributions appear in Appendix F.

Collection losses and matching scope. A collection loss scores a set of predictions against a set of targets, so it must specify which prediction is matched to which target: IMLE matches each target to its nearest prediction (Li & Malik, 2018), a one-step trajectory policy distills a flow with an IMLE-style objective (Dong et al., 2026), and W-Flow trains a one-step generator along a Wasserstein gradient flow (Han et al., 2026). At a fixed coupling, our ladder separates this choice of matching scope from balancedness, and only scope moved retention (Sec. 5.4).

Other failure modes of one-step generation. Self-distillation (Kim et al., 2026) and sub-mode conditioning (Lin et al., 2026) target distinct failure modes: training instability and loss of diversity. Our criterion is stated for whatever conditioning the model is given, so it also covers this lever: enriching the conditioning changes the conditional law that decides collapse here.

Control relevance. One-step policies for control report varied outcomes (Zhan et al., 2026; Im et al., 2026; Li et al., 2026b; Lee et al., 2026; Wang et al., 2026; Nguyen et al., 2026; Li et al., 2026a; Chen et al., 2026); the standard suites in this area span offline RL and multitask manipulation (Fu et al., 2020; Liu et al., 2023a; Park et al., 2024). Our OGBench test shows why map-spread retention should not be equated with improved control success.

7 Conclusion

For pointwise L^2 endpoint regression, conditional mean dependence exactly determines collapse. At fixed marginals, ρ^* spans $[0, 1]$ and target-only estimators incur worst-case absolute error at least $1/2$. The paired law resolves that ambiguity; matching scope, rather than balancedness alone, controlled retention in our loss experiments.

The criterion directs attention to the *actual training pairs*: near-zero conditional-mean variance predicts population mean collapse, while positive retention certifies neither distributional fidelity nor reward. Theorems concern population optima, and the generic kNN estimator was uninformative on our 256-dimensional MNIST task. The OGBench result further shows that retained action variation need not improve control. Testing this distinction across D4RL, LIBERO, or VLA settings (Fu et al., 2020; Liu et al., 2023a) requires an independent check that the extra variation is reward-relevant.

AI USE STATEMENT

In this work, we used generative AI tools for four required-disclosure tasks: research ideation (proposing and refining hypotheses, designing the experiments), research execution (implementing the methods and the audit scripts), retrieval and discovery (locating and summarizing literature), and writing (drafting parts of the manuscript and polishing it). We have not used generative AI tools to generate data sets or to produce the proofs: all data are produced by the released scripts from the distributions stated in this paper and from the public MNIST and OGBench datasets, and the proofs in the main text and in Appendix A are the author's own work, with AI tools used only to check algebraic steps. The remaining required-disclosure tasks are not applicable. Additionally, we used generative AI tools for recommended-disclosure tasks: writing the figure-generating code, formatting references, and suggesting structure and keywords. We have reviewed all AI-assisted work: AI-suggested claims were verified by manual derivation and a manual literature survey, and AI-assisted code is covered by the released regression tests. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Appendices C and D specify the data, compute, and execution protocols. The anonymous supplement contains source code, per-seed records, figure and table generators, and regression tests. Reported aggregates are generated from the released records; the audit checks them and verifies source-code fingerprints. Dataset-cache hashes were recorded after the original runs; the full-frame experiment records raw-data and transformed-array hashes before training.

ETHICS STATEMENT

This is an analysis of when one-step generative policies suffice. It uses synthetic data with exactly controlled ground truth and two public datasets (MNIST and OGBench); it involves no human subjects and releases no new dataset. The scope of every claim, and the checks that falsified earlier versions of it, are stated in Sec. 1 and Appendix G.

REFERENCES

Tadeusz Caliński and Jerzy Harabasz. A dendrite method for cluster analysis. *Communications in Statistics*, 3(1):1–27, 1974.

Anirban Chatterjee, Ziang Niu, and Bhaswar B. Bhattacharya. Conditional mean independence, global sensitivity analysis, and variable screening using nearest neighbor graphs. *arXiv preprint arXiv:2607.04692*, 2026.

Yitong Chen, Shiduo Zhang, Jingjing Gong, and Xipeng Qiu. Let it be simple: One-step action generation for vision-language-action models. *arXiv preprint arXiv:2606.05737*, 2026.

Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *arXiv preprint arXiv:2303.04137*, 2023.

Thomas M. Cover and Peter E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

Pedro M. P. Curvo, Maksim Zhdanov, Floor Eijkelboom, and Jan-Willem van de Meent. Follow the mean: Reference-guided flow matching. *arXiv preprint arXiv:2605.10302*, 2026.

Ju Dong, Liding Zhang, Lei Zhang, Yu Fu, Kaixin Bai, Zoltan-Csaba Marton, Zhenshan Bing, Zhaopeng Chen, Alois Christian Knoll, and Jianwei Zhang. From flow to one step: Real-time multi-modal trajectory policies via implicit maximum likelihood estimation-based distribution distillation. *arXiv preprint arXiv:2603.09415*, 2026.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, Dustin Podell, Tim Dockhorn, Zion English, Kyle Lacey, Alex Goodwin, Yannik Marek, and Robin Rombach. Scaling rectified flow transformers for high-resolution image synthesis. *arXiv preprint arXiv:2403.03206*, 2024.

Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models. *arXiv preprint arXiv:2410.12557*, 2024.

Justin Fu, Aviral Kumar, Ofir Nachum, George Tucker, and Sergey Levine. D4rl: Datasets for deep data-driven reinforcement learning. *arXiv preprint arXiv:2004.07219*, 2020.

Zhengyang Geng, Mingyang Deng, Xingjian Bai, J. Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. *arXiv preprint arXiv:2505.13447*, 2025. Tech report.

Jiaqi Han, Puheng Li, Qiushan Guo, Renyuan Xu, Stefano Ermon, and Emmanuel J. Candès. One-step generative modeling via wasserstein gradient flows. *arXiv preprint arXiv:2605.11755*, 2026.

Hokyun Im, Andrey Kolobov, Jianlong Fu, and Youngwoon Lee. Latent policy steering through one-step flow policies. *arXiv preprint arXiv:2603.05296*, 2026. Accepted to RSS 2026.

Youngjoong Kim, Duhoe Kim, Woosung Kim, and Jaesik Park. Stabilizing consistency training: A flow map analysis and self-distillation. *arXiv preprint arXiv:2601.22679*, 2026.

Soheil Kolouri, Kimia Nadjahi, Umut Simsekli, Roland Badeau, and Gustavo K. Rohde. Generalized sliced wasserstein distances. *arXiv preprint arXiv:1902.00434*, 2019.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Sungyoung Lee, Dohyeong Kim, Eshan Balachandar, Zelal Su Mustafaoglu, and Keshav Pingali. Towards efficient and expressive offline rl via flow-anchored noise-conditioned q-learning. *arXiv preprint arXiv:2605.01663*, 2026. ICML 2026.

Ke Li and Jitendra Malik. Implicit maximum likelihood estimation. *arXiv preprint arXiv:1809.09087*, 2018.

Shaolong Li, Lichao Sun, and Yongchao Chen. One-step flow policy: Self-distillation for fast visuomotor policies. *arXiv preprint arXiv:2603.12480*, 2026a.

Zeqiao Li, Yijing Wang, Haoyu Wang, Zheng Li, and Zhiqiang Zuo. Boosting maximum entropy reinforcement learning via one-step flow matching. *arXiv preprint arXiv:2602.01606*, 2026b.

Yexiong Lin, Yu Yao, and Tongliang Liu. Beyond optimal transport: Model-aligned coupling for flow matching. *arXiv preprint arXiv:2505.23346*, 2025.

Yexiong Lin, Jia Shi, Shanshan Ye, Wanyu Wang, Yu Yao, and Tongliang Liu. SubFlow: Sub-mode conditioned flow matching for diverse one-step generation. *arXiv preprint arXiv:2604.12273*, 2026.

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

Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. Libero: Benchmarking knowledge transfer for lifelong robot learning. *arXiv preprint arXiv:2306.03310*, 2023a.

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

Xingchao Liu, Xiwen Zhang, Jianzhu Ma, Jian Peng, and Qiang Liu. Instaflow: One step is enough for high-quality diffusion-based text-to-image generation. *arXiv preprint arXiv:2309.06380*, 2023b. ICLR 2024.

Stuart P. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28 (2):129–137, 1982.

James MacQueen. Some methods for classification and analysis of multivariate observations. In *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.

Shimon Malnick, Matan Rusanovsky, Ohad Fried, and Shai Avidan. Optimal transport flow matching by design. *arXiv preprint arXiv:2606.04092*, 2026.

Alireza Mousavi-Hosseini, Stephen Y. Zhang, Michal Klein, and Marco Cuturi. Flow matching with semidiscrete couplings. In *International Conference on Learning Representations*, 2026. arXiv:2509.25519.

Thanh Nguyen, Tri Ton, Hongbin Choe, Tung M. Luu, and Chang D. Yoo. Fast and highly expressive policy learning for offline reinforcement learning via bootstrapped flow q-learning. *arXiv preprint arXiv:2606.10613*, 2026. ICML 2026.

Seohong Park, Kevin Frans, Benjamin Eysenbach, and Sergey Levine. Ogbench: Benchmarking offline goal-conditioned rl. *arXiv preprint arXiv:2410.20092*, 2024. ICLR 2025.

Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lipman, and Ricky T. Q. Chen. Multisample flow matching: Straightening flows with minibatch couplings. *arXiv preprint arXiv:2304.14772*, 2023. ICML.

Alfréd Rényi. On measures of dependence. *Acta Mathematica Academiae Scientiarum Hungaricae*, 10:441–451, 1959.

Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. *arXiv preprint arXiv:2202.00512*, 2022. Published as a conference paper at ICLR 2022.

Allan R. Sampson. A multivariate correlation ratio. *Statistics & Probability Letters*, 2(2):77–81, 1984. doi: 10.1016/0167-7152(84)90054-3.

Xiao Shou. ODE-free neural flow matching for one-step generative modeling. *arXiv preprint arXiv:2604.06413*, 2026.

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

Gábor J. Székely, Maria L. Rizzo, and Nail K. Bakirov. Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, 35(6):2769–2794, 2007. doi: 10.1214/009053607000000505.

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. *arXiv preprint arXiv:2302.00482*, 2023. TMLR.

Shaoxuan Wang, Guangting Zheng, Rui Huang, Zhipeng Tang, Sha Zhang, Jiajun Deng, and Yanyong Zhang. RoMAN-Flow: Taming autoregressive normalizing flows for offline reinforcement learning in robotic manipulation. *arXiv preprint arXiv:2608.20208*, 2026.

Zhendong Wang, Jonathan J Hunt, and Mingyuan Zhou. Diffusion policies as an expressive policy class for offline reinforcement learning. *arXiv preprint arXiv:2208.06193*, 2022. ICLR 2023.

Ling Yang, Zixiang Zhang, Zhilong Zhang, Xingchao Liu, Minkai Xu, Wentao Zhang, Chenlin Meng, Stefano Ermon, and Bin Cui. Consistency flow matching: Defining straight flows with velocity consistency. *arXiv preprint arXiv:2407.02398*, 2024.

Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Fredo Durand, William T. Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. *arXiv preprint arXiv:2311.18828*, 2023. CVPR 2024.

Guojian Zhan, Letian Tao, Pengcheng Wang, Yixiao Wang, Yiheng Li, Yuxin Chen, Hongyang Li, Masayoshi Tomizuka, and Shengbo Eben Li. Mean flow policy with instantaneous velocity constraint for one-step action generation. *arXiv preprint arXiv:2602.13810*, 2026. ICLR Oral Presentation.

A PROOFS

A.1 THEOREM 1

Since $\mathbb{E}\|x_1\|^2 < \infty$, the minimizer of $\mathbb{E}\|f(x_0, c) - x_1\|^2$ over measurable f with finite second moment is the conditional expectation, by the L^2 projection theorem applied conditionally on (x_0, c). For the range, condition on c and apply the law of total variance to the scalar components:

$$\mathrm{Var}(x_1 \mid c) = \mathbb{E}\big[\,\mathrm{Var}(x_1 \mid x_0, c) \;\big|\; c\big] + \mathrm{Var}\left(\mathbb{E}[x_1 \mid x_0, c] \;\big|\; c\right) \;\succeq\; \mathrm{Var}\left(\mathbb{E}[x_1 \mid x_0, c] \;\big|\; c\right)$$

in the positive-semidefinite order, so the trace of the conditional-mean term is at most the trace of $\mathrm{Var}(x_1 \mid c)$ and $\rho \le 1$; $\rho \ge 0$ is immediate. Equality holds iff $\mathrm{Var}(x_1 \mid x_0, c) = 0$ a.s., i.e. iff x_1 is a measurable function of (x_0, c). At the other endpoint, the expected trace of $\mathrm{Var}(\mathbb{E}[x_1 \mid x_0, c] \mid c)$ is zero iff the conditional mean is constant in x_0 almost surely; taking its conditional expectation identifies that constant as $\mathbb{E}[x_1 \mid c]$. This proves both equality statements in Thm. 1.

A.2 THEOREM 3, FULL STATEMENT

The proof in the main text is complete; we record two extensions. *(i) Randomisation does not help*: if $\mathcal{A}$ has access to auxiliary randomness $U \perp\!\!\!\perp$ data, the law of (data, U) is still α-free, so the same Jensen argument applies to $\bar{a} = \mathbb{E}[\mathcal{A}]$. *(ii) Knowledge of $p(c, x_1)$ does not help*: a statistic of the population law is a constant, and constants are covered by the second equality in the theorem. *(iii)* The same argument gives $\inf_{\mathcal{A}} \sup_{\alpha} \mathbb{E}(\mathcal{A} - \rho^*)^2 = 1/4$ for squared loss, attained at $1/2$ as well.

A.3 THEOREM 4, L3 LIMIT

For a batch of M samples within one condition, minimizing $\frac{1}{M}\sum_i \|G_i - x_1^{(\sigma(i))}\|^2$ over bijections σ is exactly the optimal-transport problem between the empirical measures $\hat{q}_c$ and $\hat{p}_c$ with quadratic cost; its value is $W_2^2(\hat{q}_c, \hat{p}_c)$. As $M \to \infty$ the empirical measures converge weakly and $W_2^2(\hat{q}_c, \hat{p}_c) \to W_2^2(q_c, p_c)$ under the usual moment conditions. The unique minimizer over probability measures is $q_c = p_c$; the minimizer over *pushforwards of $\mathcal{N}(0, I_d)$ by a measurable map* is any f with $f(\cdot, c)_{\#}\mathcal{N}(0, I_d) = p(\cdot \mid c)$, and every such f has $\rho = 1$.

A.4 THEOREM 4, L2 LIMIT

Write $d_{q_c}(y) = \min_{x \in \mathrm{supp}\, q_c} \|y - x\|$ and $d_{p_c}(x) = \min_{y \in \mathrm{supp}\, p_c} \|x - y\|$, and assume the supports are compact so that the empirical measures converge weakly, as in the L3 case. For fixed y we have $\min_j \|y - G_j\| \to d_{q_c}(y)$ almost surely: the lower bound $\min_j \|y - G_j\| \ge d_{q_c}(y)$ holds for every M, and for every $\varepsilon > 0$ an ε-net of $\mathrm{supp}\, q_c$ is almost surely eventually occupied, whence $\min_j \|y - G_j\| \le d_{q_c}(y) + \varepsilon$ for all large M. Both y and G_j range over compact sets, so the squared integrand is bounded; dominated convergence then gives the limit $\mathbb{E}_{y \sim p_c}[d_{q_c}(y)^2]$ of the first term, and the second is symmetric. The population functional is therefore a sum of two non-negative terms, vanishing exactly when each closed support is contained in the other's. It carries no term penalising the mass a realization assigns within a non-empty region, so it is minimized by every q_c with $\mathrm{supp}\, q_c = \mathrm{supp}\, p_c$ and is otherwise silent about weights—the L2/L3 separation.

A.5 MULTI-STEP RATIO VALUES OF SEC. 2

For the Gaussian-mixture construction of Sec. 2 the endpoint ratio ρ_N at step counts $N = 1, 2, 4, \ldots, 256$ is

N	1	2	4	8	16	32	64	128	256
ρ_N	3.936×10^{-33}	0.5907	0.8573	0.9233	0.9612	0.9815	0.992	0.9974	1

B CORRECTIONS MADE DURING THE STUDY

We record the errors we found and fixed, because each one changed a number reported in this paper; a reader re-implementing the experiments is likely to meet the same pitfalls.

(i) Do not pre-register a guessed effect size. The first $\hat{\Lambda}$ criterion required $\Lambda(\Delta = 6) \geq 0.50$ and failed. The failure was uninformative: the analytic ceiling of that quantity is 0.4042, so no estimator could have satisfied it. The first loss-ladder criterion required $\mathrm{disp}(\text{Chamfer}) > 2 \times \mathrm{disp}(\text{balanced})$ and failed at a 40-step smoke test, before the models had converged. Effect sizes depend on the state of optimization and were not knowable in advance; we now pre-register only *directional* and *qualitative* predictions and report magnitudes descriptively.

(ii) Test invariance against the statistic's own noise. Our first invariance check flagged the k-means between/within ratio as α-dependent. Under naïve random initialization the same statistic varies by more than a factor of two *across seeds at fixed* α, so a fixed relative threshold compares noise with noise. We switched to a one-way ANOVA with seeds as replicates—no statistic shows a detectable α effect at level 0.05 (Sec. 5.2)—and to restarted k-means++.

(iii) Two collection-level objectives could not test their own claim. *(a) min-of-M.* The first implementation drew the M candidates by repeating one source, $x_0^{(i)}, \ldots, x_0^{(i)}$. A one-step network is a deterministic map, so the M outputs coincide at every training state, not only at collapse: the $\arg\min$ is constant and the loss reduces to pointwise L^2 on the same batch. The two code paths agreed exactly—$\max|\Delta w| = 0$ over 200 steps of each, and the full runs produced the same spread ratio to every printed digit. That row therefore carried no information about IMLE. The corrected objective draws M *independent* sources per target, which is the original IMLE semantics; Table 1 uses the regenerated row. *(b) the unbalanced Chamfer row.* The first implementation took both $\arg\min$s over the *whole batch*, while the balanced row assigned *within each condition*; the comparison therefore varied balancedness and scope simultaneously, whereas Thm. 4(L2–L3) isolates the former. The corrected row matches within each condition. This is a stronger requirement rather than a weaker one—a minimum over a superset is smaller—so the support recovery we observe is not an artifact of the change. In both cases the lesson is the same: an objective whose purpose is to *choose* among candidates should assert in code that the candidate set and the choice are the ones the theorem is about.

(iv) A tolerance of half the mode spacing inverts the mode metric. Defining "on a mode" as within half the nearest-center distance shrinks the tolerance below the mode width as K grows, and is undefined at $K = 1$, where it conceals the collapse. We use a σ-scale ball measured from the mode *width*.

(v) Marginal sliced W_2 conflates the one-step floor with mode collapse. The ratio SW@1/SW@32 is not controlled by the floor argument, because it also penalizes an under-fit field whose conditional counterpart does not collapse. The floor bounds the conditional ratio cSW@1/cSW@32, which is what we use throughout.

(vi) The no-free-parameter prediction used the W_1 constant inside a W_2 argument. An earlier draft used $\sigma\sqrt{2/\pi}$ (the W_1 distance, equivalently the mean absolute deviation) where the W_2 distance is the scale gap σ, giving $\kappa_d = \sqrt{2/(\pi d)}$ instead of $1/\sqrt{d}$.

(vii) "Coverage exactly zero" is not a satisfiable criterion. The method map pre-registered "the independent coupling covers 0 of 8 modes at $N = 1$". Cor. 1 concerns the *exact* optimum, which sits on $m(c)$ and covers exactly 0; a trained network cannot attain the exact optimum, and over 2000 evaluation samples an occasional point falls inside a mode's tolerance ball, giving a mean coverage of $0.05/8$. The criterion as written asked for exact numerical equality. We relaxed it to at most 1 of 8—the observed value is two orders of magnitude from the passing side $(8/8)$—and report the change rather than the original verdict.

C Runtime and reproducibility

Every measurement reported in the tables and figures is produced by the scripts under `code/` from the raw per-run JSON files under `results/` and `logs/`, and is typeset through a generated macro rather than typed by hand. Protocol constants, pre-registered thresholds, and the

Table 3: MNIST preprocessing sensitivity over ten paired seeds. Entries are the trained-map spread ratio ρ (mean $\pm$ standard deviation). The predicted ordering survives when the complete frames replace upper-left crops.

Pairing	Upper-left crop	Whole-frame area downsample
Independent	0.0152 ± 0.0008	0.0180 ± 0.0006
Mean-independent	0.0136 ± 0.0011	0.0158 ± 0.0006
Sorted quantile	0.1820 ± 0.0072	0.1578 ± 0.0086
Mini-batch OT	0.3319 ± 0.0092	0.2763 ± 0.0078

historical debugging values quoted in Appendix B are labelled as such and are not measurements. Numbers in the theory sections come from `paper/numbers_theory.tex`, generated by `code/analysis/make_theory_macros.py`, whose every macro carries a comment naming the JSON field it came from.

Shared hyper-parameters. The method map, loss ladder, and Chamfer runs share one MLP (input $[x, t, \mathrm{onehot}(c)]$, four hidden layers of 256 with SiLU), Adam at learning rate 10^{-3}, batch size 384, and 8000 single-stage updates; the multi-step flow-matching baseline uses Adam at 3×10^{-4}, batch 512, 20000 steps. The MNIST study uses a smaller MLP (two hidden layers of 128, ReLU), Adam at 10^{-3} with weight decay 10^{-4}, batch size 256, 4000 updates, and 10 seeds; arms (i)/(ii) stop at the μ-baseline (the population optimum, Sec. 5.6), arms (iii)/(iv) run the full schedule, and no other early-stopping criterion is used. The Sinkhorn arm (Sec. 5.4) reuses the loss-ladder protocol with entropy regulariser $\varepsilon \in \{1, 0.1, 0.01\}$ and 100 log-domain iterations per batch, sharing the same 5 seeds and evaluation protocol.

Data-side ρ^* for the MNIST couplings (Sec. 5.6). Pairings (i)/(ii) have $\rho^* = 0$ analytically; the full-frame finite-bin diagnostic for (ii) is 0.0158 despite exact mean independence. Pairing (iii) tends to the deterministic-transport limit $\rho^* = 1$, while (iv) reduces to a one-dimensional conditional expectation over $x_{0,1}$, with a 256-bin estimate 0.1754. The generic kNN estimator returns ≈ 0 for *all four* couplings at this dimension—a curse-of-dimensionality failure consistent with its operating regime in Sec. 5.5. The original frozen image path took the upper-left 16×16 crop of each 28×28 source image: block size one in the preprocessing code makes this a crop, not a resize, retaining 256 of 784 pixel positions; the centred digits lose most of their ink. We kept that result and its source hashes, then reran the same ten seeds, couplings, network, optimizer, and checks on complete images area-downsampled to 16×16. Table 3 compares the two preprocessing paths. The full-frame script records raw IDX and transformed-array hashes before fitting; the crop cache was hashed only in a later audit. The early-stopping rule for (i)/(ii) and all nine checks were fixed after a crop seed-0 pilot; this seed remains in both aggregates, so the checks are descriptive rather than independent confirmatory tests.

D Estimators, and why they are deliberately classical

Every diagnostic in this paper is built from standard, assumption-light non-parametric tools, so that a reader can re-implement it without training any auxiliary model. The conditional mean behind ρ is estimated by kNN regression (Cover & Hart, 1967). A concurrent nearest-neighbour-graph estimator of the same population ratio is consistent under finite-moment and regularity conditions but has a dimension-dependent bias rate away from the null (Chatterjee et al., 2026); testing it in our finite-sample, high-dimensional regime is reported in Appendix D.1. The mode count and the between/within variance ratio use k-means (Lloyd, 1982; MacQueen, 1967) selected by the Calinski–Harabasz criterion (Caliński & Harabasz, 1974); our conditional sliced W_2 (cSW) averages standard sliced-Wasserstein distances within each condition (Kolouri et al., 2019). We use them on purpose: they need no hyper-parameter search and are *decoupled from the generative model being diagnosed*, which is what lets ρ be computed before any one-step model exists (Sec. 5.5). All statements are about the population ρ^*; we report the kNN bias of (1) rather than correcting it away. For the MNIST dependence check, the frozen experiment script double-centers the two pairwise-distance matrices and then removes their diagonals before forming the normalized inner product; permutation of target rows calibrates its null. The script comment calling this an "U-statistic version" is inaccurate: an unbiased U-centered distance covariance uses different row and grand-mean denominators. We

therefore treat the reported number as a permutation-calibrated, diagonal-removed dependence diagnostic, not as an unbiased estimate of population distance correlation (Székely et al., 2007). The central counterexample does not rely on this estimator: mean independence is analytic, and the scale depends on x_0 by construction.

D.1 Independent nearest-neighbour-graph estimator check

To avoid validating ρ^* only with our held-out regression estimator, we also implemented Eq. (3.6) of Chatterjee et al. (2026). On the analytically controlled family $x \sim \mathcal{N}(0, I_d)$ and $y_\alpha = \alpha x + \sqrt{1-\alpha^2}\,\epsilon$, with independent $\epsilon \sim \mathcal{N}(0, I_d)$, every target marginal is $\mathcal{N}(0, I_d)$ and the exact ratio is $\rho^* = \alpha^2$. We fixed $n = 4000$, directed $K = 5$ neighbours, five seeds, and the four low-dimensional acceptance criteria in the released script before execution. Table 4 reports means over seeds. The estimator passes all criteria in its low-dimensional operating regime (mean absolute error 0.007873), providing an independent reproduction of the fixed-marginal law. It remains monotone but is strongly attenuated as dimension grows: at $d = 256$, the deterministic endpoint is estimated as 0.1634 rather than one. This is evidence for the paper's estimator boundary, not a high-dimensional remedy, and agrees with the proved dimension-dependent bias rate.

Table 4: Nearest-neighbour-graph estimate on the Gaussian family. Exact values at $\alpha \in \{0, .5, 1\}$ are $0, .25, 1$; MAE averages all five α values and seeds.

d	$\hat{\rho}(0)$	$\hat{\rho}(.5)$	$\hat{\rho}(1)$	MAE
2	0.002639	0.2533	0.9872	0.007873
16	0.001851	0.1529	0.597	0.1495
64	4.22×10^{-6}	0.07956	0.3201	0.2553
256	0.0003308	0.04078	0.1634	0.3138

D.2 Cost-matched coupling zoo

Proposition 1 distinguishes the quadratic pairing cost from the information retained by endpoint regression. We test this distinction with eight couplings of $\mathcal{N}(0, I_2)$ to itself. Every arm uses the same two-layer 64-unit MLP, pointwise L^2, 1500 updates, and five seeds. The criteria were frozen in the released script before the first run. Table 5 reports cross-seed means. Independent and 90° rotation couplings have empirical costs differing by only 0.03221, yet their trained-map ratios differ by 0.9997 (descriptive paired-seed 95% t interval $[0.9991, 1.0002]$); the largest cross-seed standard deviation of any trained-map ratio is 0.02472. Exact cost equality is analytic, so empirical cost is a Monte Carlo sanity check rather than an inferential test. The antipodal map costs roughly twice as much as independence but retains all spread. Across all eight arms, the paired-sample nearest-neighbour-graph estimate and trained-map ratio have Spearman correlation 0.9762. All six pre-declared criteria pass. Thus low transport cost can be a useful design objective without being a sufficient statistic for endpoint retention.

Table 5: Same Gaussian marginals, different couplings. Cost is $\mathbb{E}\|x_1 - x_0\|^2$; ρ_{map} is measured after endpoint-network training; and $\hat{\rho}_{\text{graph}}$ is computed from pairs before training.

coupling	exact cost	empirical cost	exact ρ^*	ρ_{map} / $\hat{\rho}_{\text{graph}}$
independent	4	3.983	0	0.000188 / -0.007654
sign-symmetric	4	4.034	0	0.0001933 / 0.006006
mixture $\alpha = .25$	3	2.996	.0625	0.06247 / 0.07019
mixture $\alpha = .50$	2	2	.25	0.2451 / 0.2405
mixture $\alpha = .75$	1	0.9918	.5625	0.5601 / 0.5586
identity	0	0	1	0.9999 / 0.9865
90° rotation	4	4.015	1	0.9998 / 0.9868
antipodal	8	8.055	1	1 / 0.9872

D.3 Equal-cost dimension stress check

To test whether the analytic cost–retention contrast survives a larger finite regressor, we froze dimensions $d \in \{16, 64, 256\}$, five seeds, 1200 updates, batch size 512, and a two-hidden-layer SiLU network of width $\max(64, 2d)$ before running this check. Each target marginal is $\mathcal{N}(0, I_d)$; independent pairing, global random-sign pairing $x_1 = Sx_0$, and blockwise 90° rotation all have exact

quadratic cost $2d$, while their population ρ^* values are $0, 0, 1$. The random-sign pairing is dependent but conditionally mean-independent.

Table 6: Trained endpoint-map spread ratio across dimensions (five-seed means). Exact ρ^* is $0, 0, 1$ for the three columns.

d	exact cost	independent	random sign	90° rotation
16	32	0.0001564	0.0001501	0.9981
64	128	0.0001428	0.0001437	0.985
256	512	0.0002709	0.0002746	0.967

All three frozen checks pass against their pre-registered thresholds: both mean-independent arms remain below 0.05, the rotation stays above 0.90, and each empirical cost mean lies within 5% of $2d$. Per-seed results and the exact protocol are released. This is a synthetic finite-learner sanity check; it neither repairs the failed high-dimensional generic paired estimator nor establishes conditional-law fidelity or policy gains.

E Exploratory downstream stress test: design and execution

The criterion suggests a falsifiable prediction beyond the controlled tasks of this paper. We drafted the protocol below before execution, but changed the conditioning-strength operationalisation after inspecting a data-side diagnostic, so this section is exploratory rather than a strict pre-registration.

Task. One offline continuous-control suite with weak and strong conditioning variants (e.g. a single LIBERO task suite under two instruction-conditioning strengths, or the weakest-conditioning split of an offline dataset). No robot hardware is needed; the pairing is a training-time choice.

Arms. Same backbone, same loss, same schedule; the only difference is the coupling: (i) independent, (ii) paired with the sort-based or OT scheme of Sec. 5.6. Compute $\hat{\rho}$ from the paired batches before training (Sec. 5.5), at the training dimension.

Drafted predictions. (P1) the arm with lower $\hat{\rho}$ collapses to its conditional mean and fails on the weak-conditioning split while the arm with higher $\hat{\rho}$ does not; (P2) the failure gap grows with the conditioning gap, qualitatively tracking the ρ^* gap; (P3) the mean-dependence ordering of Sec. 5.6 reproduces on the policy's action distribution (ρ of the trained policy's one-step map).

Falsification. If both arms collapse despite a large $\hat{\rho}$ gap, or the high-$\hat{\rho}$ arm fails where the low-$\hat{\rho}$ arm succeeds, the criterion's transfer to policy pipelines is refuted and the paper's practical claim should be retreated to the controlled setting. Ten seeds per arm are used below; paired across-seed intervals quantify the behavioral contrast.

Executed protocol (suite and disclosed operationalisations). Suite: OGBench `pointmaze-medium-navigate-v0` (Park et al., 2024), CPU simulation; conditions $c = [\text{position}, \text{goal}]$ from HER-style relabelled transitions ($2{\times}10^5$ pairs). Arms: the same $3{\times}256$ MLP action head trained with pointwise L^2 for 8000 steps; the only difference is the coupling. Three operationalisations deviate from a literal reading of the drafted protocol and are disclosed here. (a) *Coupling within matched conditions*: batches are grouped into 24 anchor conditions $\times$ 16 near-duplicate pool states each, and the sorted arm applies the sorted-quantile scheme *within* each group. A global sort over mixed-condition batches leaks between-condition variance into the map's x_0-sensitivity (the trained map becomes "conditional mean + a condition-independent offset"), an implementation artifact rather than the coupling the criterion speaks about. (b) *Strength axis*: the HER-horizon split (short = strong, long = weak) shows *no* modulation of the conditional action spread in this dataset under the two comparable kNN bandwidths (weak/strong ratio 1.099 at $k = 8$ and 1.071 at $k = 16$; the grid estimator gives 5.109, but its cell count differs across splits—73 versus 22—so the two splits are not measured at a common resolution and we do not use it). The strength axis is the dataset's own local conditional spread $\sigma(c)$ (per-dim within-ball action variance, median split at 0.1624), which the drafted "weakest-conditioning split" option was intended to cover. (c) *Deployment*: the latent $x_0 \sim \mathcal{N}(0, I_8)$ is drawn once per episode and held fixed (identically for both arms); per-step re-drawing would inject i.i.d. action jitter unrelated to the criterion. $\hat{\rho}$ is the group-internal adjusted R^2 of $y \sim x_{0,1}$ within condition groups (centered at zero under independence); ρ of the trained map is measured within condition.

Execution results (10 seeds, 250 rollouts per arm and seed). The map-level predictions reproduce: the pairing-strength ordering ($\hat{\rho}$: -0.0002961 vs. 0.4058) and the within-condition retention ordering (ρ of the trained map: 0.004223 vs. 0.3988) are stable across all seeds, and the sorted arm's excess one-step risk over the independent arm grows with the conditioning stratum (low-σ 0.1543 vs. high-σ 0.1766). This risk gap is expected because the conditional mean minimizes L^2 risk and higher-variance strata magnify deviations from it; it is a sanity check, not evidence of downstream utility. The behavioral prediction P1 is *refuted* on this suite: the collapsed conditional-mean arm attains *higher* closed-loop success than the spread-preserving arm (0.2564 ± 0.1253 vs. 0.1624 ± 0.05896; 8/10 seeds), paired gap -0.094 with a 95% t interval [-0.1633, -0.02466]. Diagnosis: the measured spread in this suite does not grow with the conditioning horizon (operationalisation (b) above), which is consistent with its being dominated by per-step control noise rather than by route multimodality; the conditional mean is then a good action, while a spread-preserving map faithfully reproduces that noise into its actions. The criterion's behavioral transfer to this pipeline is not supported and the practical claim retreats to the controlled setting; only the map-level spread ordering stands. Aggregate JSON with per-seed records and protocol fingerprints is released.

Two further failure modes would challenge the central claims directly. First, if the paired estimator of Sec. 5.5 fails to track an independently computed population ρ^* wherever both are available, the measured ρ curves lose their interpretation as conditional-spread retention. Second, if the mean-independence criterion fails under the finite-moment and population-optimum assumptions of Thm. 1—for instance if a pointwise L^2 regressor retains conditional spread when the conditional mean does not depend on x_0—then the exact criterion is not exact. Data-side estimators that survive the high-dimensional regime where the current kNN estimator fails (Sec. 5.6) are the prerequisite for the first test.

F Additional related work

Sec. 6 compares against these works at the level of population objectives and matching scope; this appendix states the three qualifications that keep that comparison honest. First, the step-reduction methods (Song et al., 2023; Yang et al., 2024; Frans et al., 2024; Geng et al., 2025; Salimans & Ho, 2022; Yin et al., 2023; Liu et al., 2023b) optimize step-specific training objectives, and a step count that works in practice is compatible with either answer to our population question; our theorems therefore neither support nor contradict their reported step counts. Second, the pairing designs (Pooladian et al., 2023; Malnick et al., 2026; Curvo et al., 2026) choose a coupling in order to train a model, whereas we take the coupling as given and ask what its population optimum retains. Third, for the collection losses (Li & Malik, 2018; Dong et al., 2026; Han et al., 2026) our ladder isolates matching scope at a fixed coupling and fixed data; it is a controlled comparison of that one choice, not a finite-model ranking of the losses themselves.

G Scope and assumptions

- **Field:** machine learning — generative modeling for decision making, as an *analysis and diagnosis* contribution, not a new sampler.
- **Tasks:** synthetic conditional ring mixtures with closed-form ground truth; a whole-frame, area-downsampled MNIST 16×16 endpoint-regression study with constructed couplings; and an exploratory OGBench pointmaze execution (Appendix E) in which the map-level prediction reproduces and the behavioral prediction is refuted on that suite.
- **Datasets:** synthetic families, the public MNIST training images, and the public OGBench pointmaze-medium-navigate dataset; no proprietary data, no robot hardware, no pretrained backbone.
- **Compute:** the main-table and OGBench runs used CPU PyTorch on a 20-core workstation; the separate dimension-stress check (Appendix D.3) used one RTX 3060 Laptop GPU. No robot hardware.
- **Attribution:** each component is re-implemented from the published objective rather than reused from any author repository. Any error is ours.