Geometric Entropy: When Trajectory Diversity Helps and Hurts in Imitation Learning

Qian Luo[1,2], Ruizhe Liu[1], Pei Zhou[1,2], Xunzhe Zhou[1,2], and Yanchao Yang[1,2]

***Abstract*— We study how *trajectory-shape diversity* in demonstrations affects imitation learning (IL) performance across models, tasks, and data scales. We introduce Geometric Entropy (H_G), a task-agnostic metric that quantifies intrinsic diversity of transit trajectories after normalizing away extrinsic variation (e.g., goal pose and workspace scale) via target-frame alignment. Across multiple IL architectures and both simulation and real-robot contact-rich manipulation tasks, we observe a consistent inverted-U relationship between success and H_G: increasing geometric diversity improves robustness in low-diversity regimes but degrades performance once diversity induces strategy ambiguity. Moreover, the optimal entropy shifts toward *lower* values as task mastery increases (via more data, easier tasks, or stronger priors), and for a pretrained vision-language-action model the trend becomes effectively monotonic decreasing. Practically, H_G enables fast pre-training auditing of demonstration datasets and offers a simple guideline for calibrating demonstrations toward the learnable regime. Project page: `https://geometric-entropy.github.io/`.**

I. INTRODUCTION

Modern imitation learning (IL) is increasingly data-centric. As policy classes scale to high-capacity architectures such as diffusion policies and vision-language-action (VLA) models, the bottleneck often shifts from algorithm design to *what* demonstrations contain and *how* they are curated. Prior work studies data quality along axes such as observation fidelity, label accuracy, and task or scene coverage. Yet a more fundamental question remains underexplored: *how diverse should the **trajectory shapes** in a demonstration dataset be?*

Trajectory-shape diversity is a double-edged sword. On one hand, varied approach geometries can improve robustness by expanding effective coverage of relevant state–action regions, reducing overfitting to a narrow motion corridor. On the other hand, excessive diversity can introduce *mode inconsistency*: multiple distinct strategies accomplish the same goal, but the conditioning signal may be insufficient to reliably select among them. In contact-rich manipulation, even small confusion (or averaging across incompatible strategies) can yield infeasible motions and sharp performance drops. This motivates a non-monotonic relationship between geometric diversity and success: too little diversity leads to brittleness, while too much becomes interference.

To study this systematically, we need a pre-training metric that quantifies *between-trajectory* **geometric shape diversity** rather than marginal variability, and that is invariant to extrinsic episode factors (e.g., target pose, workspace placement, and scale) that can inflate naive spread measures without reflecting true strategy differences. We therefore propose **Geometric Entropy** (H_G), a lightweight, task-agnostic scalar designed to meet these requirements. H_G isolates the learnable geometric signal by focusing on free-space *transit* segments and expressing each segment in a canonical, target-anchored representation: trajectories are arc-length resampled to a fixed length, translated to the segment endpoint, rotated into the target frame, and normalized by the start–end displacement to remove scale. This alignment ensures that trajectories that differ only by where the task happens map to the same descriptor, while genuine strategy differences (approach angle, curvature, altitude profile, detours) remain. H_G then summarizes the intrinsic complexity of the resulting shape manifold with a stable spectral statistic of the aligned trajectory descriptors, capturing both the number of independent shape modes and their overall spread in shape space. Because it is computationally cheap and sample-efficient, H_G provides a practical audit signal for dataset curation *before* training.

Using controlled demonstrations with tunable geometric diversity and evaluating across multiple tasks, models, and data scales in both simulation and on real robots, we observe a consistent **Inverted-U** pattern: performance often improves as H_G increases from low values, but degrades when H_G becomes too large. Moreover, the peak shifts toward **lower** H_G as the learner becomes more proficient. This peak-shift supports a unifying **task mastery** hypothesis: when mastery is low, additional geometric diversity is beneficial by expanding effective coverage; when mastery is high, the tolerable diversity shrinks and excessive variation acts as interference that reduces success.

Our contributions are:

1) We provide a **formal treatment** of trajectory **geometric shape diversity** in imitation learning, clarifying why it is a distinct dataset attribute that can be beneficial yet harmful beyond a learnable regime.
2) We propose **Geometric Entropy** (H_G), a practical metric that quantifies intrinsic trajectory-shape diversity and can be computed prior to training for dataset auditing and curation.
3) Through systematic evaluation across simulation and real-world contact-rich manipulation tasks, we report a consistent **Inverted-U** trend and a **peak-shift** with increasing task mastery, providing actionable guidance for calibrating demonstration diversity.

The paper is organized as follows. Section II surveys related work. Section III presents the metric. Section IV

[1]InfoBodied AI Lab, The University of Hong Kong. [2]Transcengram. e-mail: {luoqian1,zrllrz360,pezhou,xunzhe_zhou}@connect.hku.hk; yanchaoy@hku.hk.

describes the experimental setup, and Section V reports results and analysis. Section VI discusses implications and limitations.

II. RELATED WORK

A. Data Scaling in Imitation Learning

Recent robotics imitation learning has shifted towards scaling, with policies and agents improving via large, heterogeneous robot datasets and multi-embodiment trajectories, as seen in RT-1 [1], RT-2 [2], Open X-Embodiment/RT-X [3], DROID [4], Octo [5], and OpenVLA [6]. Efforts like BC-Z [7] and RoboCat [8] further advocate for broad generalization through scaling task diversity, while embodied multimodal models bridge internet-scale pretraining with robot control [9]. In parallel, simulation-scale imitation benchmarks and prompted generalist manipulation agents operationalize scaling along tasks and prompts [10]. Hybrid offline pipelines [11], show the importance of large-scale experience combined with strong policy architectures. However, these approaches typically treat "diversity" as a coarse dataset attribute (tasks/objects/embodiments/modalities) without focusing on trajectory geometry, a gap our work addresses through geometric entropy and an IL-specific inverted-U trend that indicates when geometric diversity aids or hinders learning.

B. Data Curation and Collection for Imitation Learning

Beyond scale, recent work formalizes IL data quality and develops curation/collection mechanisms. Data quality in IL decomposes performance under distribution shift into action divergence and transition diversity [12]; interactive IL such as DAgger collects labels on the learner-induced distribution to mitigate compounding error [13], with related robustness via noise-injected supervision [14]. Offline learning is also sensitive to demonstration quality [15], motivating reweighting/filtering from suboptimal data [16], transition-level self-supervised curation [17], influence-based attribution to closed-loop performance [18], and online experience-guided demo selection [19], as well as learning with imperfect demonstrations via self-supervision [20]. Complementary directions synthesize data or inject variability to expand coverage/occupancy, including MimicGen [21], DemoGen [22], ManiBox [23], adversarial perturbation [24], FieldGen [25], and MOVE [26]. Unlike methods that primarily optimize selection, robustness, or coverage, we directly diagnose trajectory geometry as an intrinsic dataset property and provide a predictive metric plus an inverted-U trend to guide whether to increase or reduce geometric diversity.

C. Trajectory Geometry, Diversity, and Mode Coverage

Trajectory diversity is commonly studied as multimodality/mode coverage in conditional action distributions: multimodal IL learns stochastic multi-skill policies [27], and modern sequence/generative policies model multimodality via discretization or generation [28], [29], with related formulations including implicit energy-based action modeling [30], hierarchical tokenization [31], and efficient multi-task transformers [32]. Diversity is also evaluated via benchmarks (D3IL) [33] and distributional views such as occupancy-measure matching [34] , while analyses show its effect depends on which factors vary and can even hinder learning [35] ; orthogonally, classical sequence-alignment tools quantify pairwise trajectory similarity/structure [36]. However, these lines do not provide an IL-specific, dataset-level criterion for *how much* geometric diversity is learnable under different priors; we address this with extrinsic-invariant geometric entropy and a predictive inverted-U trend (with peak shifts across priors) for curation.

III. METHODOLOGY

We seek a metric that answers a single question: given a set of demonstrations for a manipulation task, how geometrically diverse are the trajectory *shapes*, independent of where in the workspace the task happens to take place? This requires three components: a formal decomposition of the factors that shape raw trajectories (Section III-A), a phase-aware alignment pipeline that isolates the geometric signal (Sections III-B and III-C), and a spectral diversity measure robust to the high-dimensional, moderate-sample regime of robotics (Section III-D).

A. Problem Formulation

Let a demonstration dataset $\mathcal{D} = \{\tau_1, \ldots, \tau_M\}$ consist of M trajectories, where each trajectory $\tau_i = \{\mathbf{p}_1^{(i)}, \ldots, \mathbf{p}_{L_i}^{(i)}\}$ is a variable-length sequence of end-effector positions $\mathbf{p}_t \in \mathbb{R}^3$. In a typical manipulation episode, the raw positions are shaped by three entangled factors:

- **Extrinsic context** $\mathcal{E}$: the object/goal *pose* (position and orientation) and workspace scale, which vary across episodes due to domain randomization or different environment configurations. These factors determine *where* and *how* the task is situated in the workspace but carry no information about the demonstrator's motion strategy. Our alignment normalizes this context by expressing trajectories in a target-anchored local frame and removing scale (Section III-C).
- **Geometric strategy** $\mathcal{G}$: the intrinsic shape of the motion path—its approach angle, curvature profile, altitude trajectory, and detour geometry—which reflects the demonstrator's choice of *how* to reach the target. This is the quantity whose diversity we wish to measure.
- **Execution noise** ϵ: controller jitter, motion planner nondeterminism, and other sources of irreproducibility that add stochastic variation unrelated to either context or strategy.

The central challenge is to quantify the diversity of $\mathcal{G}$ across the dataset while marginalizing away $\mathcal{E}$ and ϵ. Failure to separate these factors leads to inflated diversity estimates: a dataset of identical straight-line approaches executed at randomized goal locations will appear highly diverse under any metric that operates on raw positions, yet it contains zero genuine strategy variation. Our pipeline addresses this by first decomposing episodes into semantically meaningful

segments (Section III-B), then applying a normalization that strips extrinsic context while preserving the geometric signal (Section III-C).

B. Phase Decomposition: Transit vs. Fine Manipulation

A manipulation episode typically interleaves two qualitatively different motion regimes:

- **Transit motion**: free-space movement of the end-effector toward (or away from) a task-relevant target. Its geometry—the approach angle, curvature, and altitude profile—is the primary carrier of demonstrator strategy diversity and the locus of learnable geometric variation.
- **Fine manipulation** [25]: contact-rich actions such as grasping, insertion, pivoting, and release, where the end-effector trajectory is tightly constrained by object geometry and contact mechanics, leaving little room for meaningful geometric variation.

We focus exclusively on transit segments. Phase boundaries are identified by task-specific semantic events: for example, gripper closure marks the transition from approach to grasp, and gripper opening marks the transition from transport to release. This decomposition generalizes to any multi-phase task—a pick-and-place episode yields pick-transit and place-transit segments, while a more complex assembly task might yield three or more transit phases separated by distinct contact events.

Each transit segment is aligned and scored independently (Sections III-C–III-D). This per-phase treatment is essential: mixing transit and fine-manipulation segments would conflate genuine strategy diversity with contact-constrained motion, diluting the diagnostic power of the metric.

C. Endpoint-Anchored Frame Alignment

We transform each transit trajectory into a canonical shape descriptor $\boldsymbol{\xi} \in \mathbb{R}^{3T}$ via four steps designed to remove extrinsic variation while preserving intrinsic geometric strategy.

Step 1: Arc-length resampling. Raw trajectories vary in length due to differences in execution speed and controller frequency. We resample each trajectory to exactly T waypoints (We set T=50 in our experiments) at uniform arc-length intervals via linear interpolation, ensuring that the descriptor captures *where* the end-effector goes but not *when* it arrives, and that all trajectories live in the same fixed-dimensional space $\mathbb{R}^{3T}$.

Step 2: Endpoint anchoring (translation normalization). We translate each resampled trajectory so that its terminal point—the task target (e.g., the grasp pose)—sits at the origin: $\mathbf{p}'_k = \bar{\mathbf{p}}_k - \bar{\mathbf{p}}_T$. Anchoring at the terminal point cancels extrinsic positional offsets due to randomized target locations.

Step 3: Target-frame alignment (rotation normalization). To remove extrinsic rotational variation, we express the anchored trajectory in a canonical local frame attached to the task target. Let $\mathbf{R}_T \in SO(3)$ denote the target (or object) orientation used to define the local task frame for the corresponding phase.[1] We rotate the anchored points into this frame: $\mathbf{p}''_k = \mathbf{R}_T^\top \mathbf{p}'_k$. This step ensures that trajectories that are identical up to a global rotation (e.g., due to different object yaw) map to the same shape descriptor.

Step 4: Displacement scaling. We normalize by the Euclidean distance between the start and end of the segment: $\tilde{\mathbf{p}}_k = \mathbf{p}''_k \,/\, (\|\bar{\mathbf{p}}_1 - \bar{\mathbf{p}}_T\| + \epsilon)$, with $\epsilon = 10^{-9}$. This removes workspace-scale effects so that a short reach and a long reach with the same curvature profile yield identical descriptors—we measure *shape*, not physical extent.

The canonical descriptor is the flattened vector $\boldsymbol{\xi} = [\tilde{\mathbf{p}}_1^\top, \ldots, \tilde{\mathbf{p}}_T^\top]^\top \in \mathbb{R}^{3T}$. After this pipeline, two trajectories produce the same $\boldsymbol{\xi}$ if and only if they share the same intrinsic shape relative to the target frame (up to residual execution noise).

D. Geometric Entropy (H_G)

Given M aligned descriptors $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_M\} \subset \mathbb{R}^{3T}$, we seek a single scalar that summarizes geometric shape diversity. In our setting, the descriptor dimension is moderately high ($3T$=150) while demonstration counts are typically $M \sim 10^2$–10^3, where fully non-parametric density estimation can be unreliable. We therefore summarize diversity through the covariance eigenspectrum, and later empirically compare this choice against common alternatives in Section V-D.

Spectral decomposition. To avoid estimating a probability density in a high-dimensional descriptor space ($3T$=150) with moderate sample sizes ($M \sim 10^2$–10^3), we summarize diversity through the covariance eigenspectrum. Let $\mathbf{X} \in \mathbb{R}^{M \times 3T}$ be the mean-centered descriptor matrix. Its SVD yields eigenvalues $\{\lambda_j\}$ of the sample covariance, where each λ_j captures variance along one orthogonal shape mode.

A useful diversity metric should capture two factors: *how many* independent shape modes exist (structural complexity) and *how spread out* they are (variation magnitude). We quantify the former via the **effective dimensionality** d_{eff}, the exponentiated Shannon entropy of the normalized spectrum $p_j = \lambda_j / \sum_k \lambda_k$:

$$d_{\text{eff}} = \exp\Big(-\sum_{j:\, p_j > 0} p_j \ln p_j\Big). \tag{1}$$

When variance concentrates on a single mode, d_{eff}=1; when it is uniform across D modes, d_{eff}=D.

We then define the **Geometric Entropy**:

$$H_G = \ln d_{\text{eff}} + \ln\Big(\sum_j \lambda_j\Big). \tag{2}$$

The first term increases as the dataset populates more independent approach strategies; the second increases as these modes spread further apart in aligned shape space. Both are necessary: many modes separated by tiny angles yield high d_{eff} but negligible magnitude, whereas a single dominant mode spanning a large arc yields low d_{eff} but

[1] In simulation, $\mathbf{R}_T$ is available from environment state; on real robots, it is obtained from the same perception/state estimation used to define the task target.

high magnitude. Unlike $\log\det(\Sigma)$, which is ill-conditioned under $M \ll 3T$ and requires ad-hoc regularization, the total-variance term $\log\big(\sum_j \lambda_j\big)$ remains stable in our regime.

Rather than claiming a strict approximation of differential entropy, we use H_G as a **lightweight spectral surrogate**: it is non-parametric (no bandwidth to tune), computationally cheap (a single SVD), and well-behaved at practical M. We later contrast H_G with representative baselines in Section V-D.

Multi-phase aggregation. For tasks with P transit phases, we compute H_G per phase and define the overall dataset entropy as the size-weighted average:

$$H_G^{\text{avg}} = \frac{\sum_{p=1}^{P} n_p \cdot H_G^{(p)}}{\sum_{p=1}^{P} n_p}, \tag{3}$$

where n_p is the number of successful trajectories for phase p.

IV. EXPERIMENTAL SETUP

Our experiments are designed to answer three questions: (1) Is H_G a stable and discriminative descriptor of trajectory-shape diversity? (2) How does imitation learning performance evolve as a function of H_G across tasks, models, and data scales? (3) Can we identify an organizing principle that governs the optimal diversity level?

To isolate the effect of *geometric strategy diversity* from other factors, we require datasets whose trajectory shapes can be varied *continuously and independently* while holding task definition and observation/action interfaces fixed. We therefore build a parameterized demonstration generator with controllable diversity and then validate the observed trends on real-robot hardware.

A. Simulation Experiments

1) Parameterized Demonstration Generation: Our simulation pipeline is built on the **PandaArmMotionPlanningSolver** from ManiSkill3 [37]. We inject controlled geometric diversity by introducing an *approach waypoint* that the end-effector must pass through before reaching the task target. Two parameters control diversity:

Approach radius r. For each demonstration, we sample a waypoint on a circle of radius r centered above the target:

$$\mathbf{w} = \mathbf{g} + [r\cos\theta,\ r\sin\theta,\ \Delta z]^\top, \tag{4}$$

where $\mathbf{g}$ is the target position, θ is the approach angle, and Δz is a fixed height offset. When r=0, trajectories are near-identical straight descents; larger r yields increasingly curved, laterally displaced approaches (Fig. 1).

Discrete mode count k. We sample $\theta \sim \text{Uniform}(0, 2\pi)$ for k=0, or from k equally spaced discrete values for k>0. This decouples *magnitude* from *modality*: large r with k=1 produces long but consistent (single-strategy) approaches, while small r with k=0 produces short but angularly scattered trajectories. This controlled separation lets us distinguish magnitude-driven variation from modality-driven ambiguity in Section V.

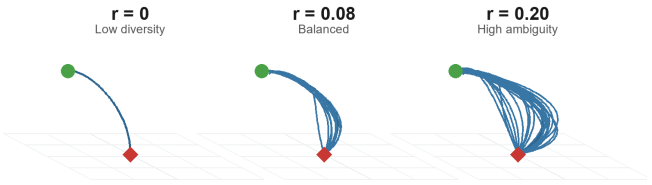


Fig. 1: Trajectory manifolds under varying approach radius r. At $r \approx 0$ (left), demonstrations follow near-identical descents. At moderate r (center), approach paths form a cone of curved arcs. At large r (right), the manifold expands further and can enter a regime where excessive geometric diversity induces mode ambiguity that degrades learning.

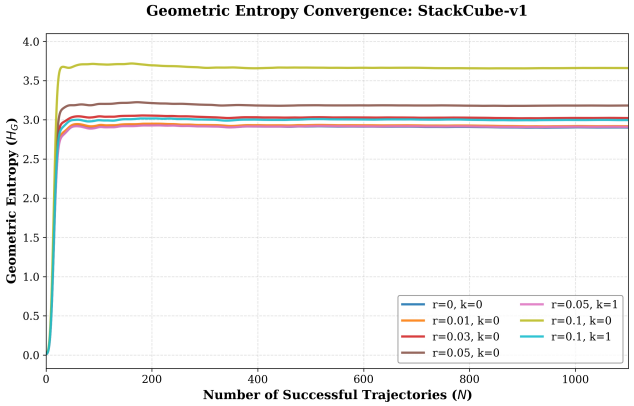


Fig. 2: H_G convergence on `StackCube-v1`. Across all (r, k) settings, H_G stabilizes within the first 50–100 successful trajectories, enabling reliable pre-training audits at practical collection scales.

For each (r, k) configuration, we generate datasets of size $N \in \{100, 200, 500, 1000\}$, segment each episode into transit phases based on gripper events, and compute H_G per phase, averaged using Eq. 3. All simulation results are reported over three independent seeds.

2) Convergence and Stability of H_G: A practical dataset-audit metric should be sample-efficient and stable. Fig. 2 plots H_G on `StackCube-v1` as a function of the number of successful trajectories N across all (r, k) configurations. In all cases, H_G stabilizes within the first 50–100 demonstrations; for $N \gtrsim 100$, residual fluctuations are small compared to gaps between configurations.

To make ordering explicit, Table I reports converged H_G^{avg} (Eq. 3). For k=0, H_G increases monotonically with r, confirming that enlarging the waypoint radius expands the aligned shape manifold. Enforcing a single approach mode (k=1) substantially lowers H_G even at large r, reflecting reduced mode ambiguity. The non-zero baseline at (r, k)=$(0, 0)$ captures residual variation from environment randomization and execution noise.

3) Task Environments: We evaluate two ManiSkill3 tasks spanning distinct manipulation regimes:

StackCube-v1: pick up a cube and stack it on a second cube. Success requires robust grasping under geometric

TABLE I: Converged H_G^{avg} on `StackCube-v1`.

r/k	0/0	0.01/0	0.03/0	0.05/0	0.05/1	0.10/0	0.10/1
H_G^{avg}	2.901	2.917	3.016	3.183	2.909	3.667	2.997

variation and precise placement alignment.

PegInsertionSide-v1: laterally insert a peg into a hole with sub-centimeter clearance. This high-precision task reaches near-saturation performance at moderate N, providing a "high mastery" contrast to StackCube.

4) Imitation Learning Models: We evaluate two architectures spanning the prior-knowledge spectrum:

Diffusion Policy (DP) [29]: a specialist model trained from scratch on each dataset.

$\pi_{0.5}$ **[38]**: a vision-language-action (VLA) model developed by **Physical Intelligence**, fine-tuned on StackCube-v1 in our VLA study. Its large-scale pre-training provides a strong prior that we hypothesize shifts the optimal diversity toward lower H_G.

B. Real-Robot Experiments

To verify that simulation trends are not artifacts of simulated dynamics, we run real-robot experiments on an ARX robot arm [39] across three tasks: **StackCube**, **PlacePanda** (place a toy panda into a basket), and **OpenDrawer**. Demonstrations are collected at five diversity levels by three collectors, with N=100 demonstrations per configuration. We compute H_G using the same alignment pipeline and train **ACT** [40] policies. Success is evaluated over 40 test episodes per configuration. All real-robot results are reported over three independent seeds.

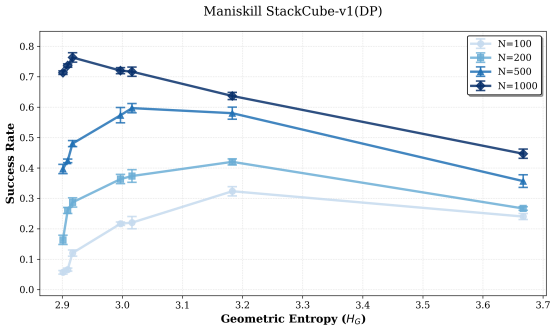
V. RESULTS AND ANALYSIS

We analyze how *trajectory geometric shape diversity* impacts imitation learning performance. Across models, tasks, and data scales, the effect of diversity is not fixed: it depends on the learner's proficiency, which we summarize as *task mastery*.

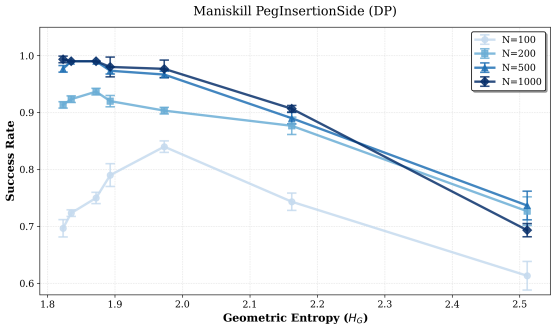
A. Simulation Results

Diffusion Policy: Inverted-U Pattern and Peak-Shift. Fig. 3 reports Diffusion Policy performance on both tasks as a function of H_G. Two consistent effects emerge:

1) *Inverted-U empirical pattern.* Performance often improves as H_G increases from low values, but degrades once H_G becomes too large. At low entropy, demonstrations occupy a narrow motion corridor; when test-time initial conditions deviate, the policy lacks coverage to recover. Increasing H_G expands geometric coverage and improves robustness. Beyond a task- and model-dependent peak H_G^*, however, demonstrations contain multiple qualitatively different strategies for the same conditioning signal, introducing *mode ambiguity*: the learner may execute an incorrect mode or interpolate between incompatible strategies, producing infeasible motions—especially near contact.



(a) StackCube-v1



(b) PegInsertionSide-v1

Fig. 3: Diffusion Policy success rate vs. Geometric Entropy (H_G). Both tasks exhibit an **Inverted-U** trend and a **peak-shift** toward lower H_G as dataset size N increases. The easier `PegInsertion` task saturates faster than `StackCube`, consistent with the mastery hypothesis. Error bars denote variation over three seeds.

2) *Peak-shift with data volume.* The optimal H_G^* shifts leftward as dataset size N increases. With more demonstrations, the training distribution covers a wider range of initial conditions even without injected geometric diversity, reducing the marginal benefit of higher H_G while the cost of ambiguity persists. Larger datasets thus favor more consistent demonstrations.

Comparison across tasks. The contrast between tasks highlights how difficulty interacts with diversity. `StackCube-v1` remains challenging even at large N; here, moderate increases in H_G are beneficial because added coverage directly improves robustness. `PegInsertionSide-v1` reaches near-saturation at moderate N; once the motion corridor is learned, additional geometric diversity provides little benefit and primarily acts as interference, leading to a sharper post-peak drop. This supports the view that higher mastery reduces the tolerable diversity range.

Generalist foundation model ($\pi_{0.5}$): monotonic decay. We fine-tune $\pi_{0.5}$ on `StackCube-v1` (Fig. 4). When plotted against H_G, performance decreases approximately monotonically. This is consistent with a high-mastery regime induced by strong pretraining: the model already brings substantial geometric coverage, so additional diversity in the fine-tuning set is redundant and can be interpreted

as conflicting behavior, degrading execution of the motion corridors preferred by the prior. In effect, the Inverted-U peak has shifted so far left that no beneficial diversity regime is visible within the tested range.

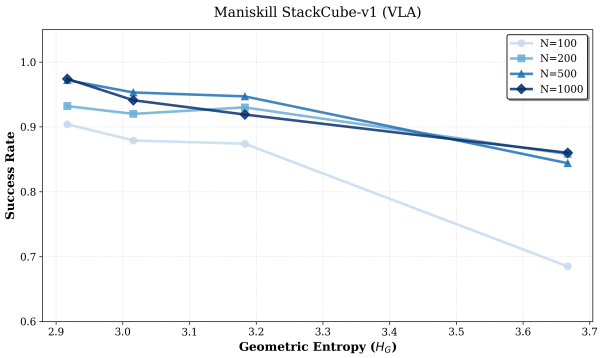


Fig. 4: $\pi_{0.5}$ success rate vs. H_G on StackCube-v1. The approximately monotonically decreasing curve is consistent with a high-mastery regime where additional geometric diversity acts as interference.

B. The Mastery–Entropy Principle

Synthesizing results across tasks, data scales, and model architectures suggests an organizing principle: the optimal geometric entropy H_G^* is governed by the learner's **task mastery**. We use *task mastery* to denote how proficient the learner is relative to a task's demands, shaped by (i) training data volume/coverage (N), (ii) task difficulty, and (iii) model priors (pretraining and inductive biases). As mastery increases, the learner requires less geometric diversity to achieve robustness, and tolerance for conflicting modes shrinks.

a) Operationalizing mastery.: Task mastery is a latent notion rather than a single measured variable, but it can be approximated with simple proxies. A direct proxy is the success rate achieved when training on a *low-diversity* (or "no-injected-diversity") demonstration set under the same evaluation protocol. High baseline success indicates that the learner already captures the relevant motion corridor (high mastery), so additional geometric diversity is more likely to act as interference; low baseline success suggests insufficient coverage (low mastery), where increasing diversity can improve robustness. Increasing N, reducing task difficulty, or strengthening priors typically increases this baseline and shifts the optimal H_G^* toward lower values.

- **Low mastery** (e.g., `StackCube` at small N, or a model trained from scratch): moderate increases in H_G help by expanding effective coverage; the peak lies at higher H_G.
- **High mastery** (e.g., `PegInsertion` at large N, or a pre-trained VLA): consistent demonstrations already suffice; additional diversity is primarily interference, and the relationship becomes effectively monotonic decreasing.

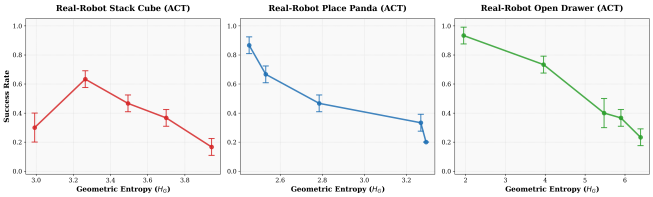


Fig. 5: Real-robot success rate vs. H_G for three tasks. StackCube (left) exhibits a clear Inverted-U, peaking at $H_G \approx 3.3$. PlacePanda (center) and OpenDrawer (right) show monotonic decline, consistent with high-mastery tasks where additional diversity is interference. Error bars denote variation over three seeds.

C. Real-Robot Validation

We validate that the mastery–entropy behavior is not an artifact of simulation by training ACT policies on the ARX arm across three tasks (Fig. 5); all results are averaged over three seeds.

StackCube: clear Inverted-U. On StackCube (Fig. 5, left), success peaks at $H_G \approx 3.3$. Below this peak, demonstrations are too geometrically narrow to handle physical perturbations and pose variation; above it, conflicting approach geometries increase ambiguity near contact, reducing success—mirroring the low-to-mid mastery regime observed in simulation.

PlacePanda and OpenDrawer: monotonic decline. For PlacePanda and OpenDrawer (Fig. 5, center and right), success decreases with H_G, consistent with higher-mastery settings where low-entropy demonstrations already suffice and additional geometric variation acts as noise rather than useful coverage.

Sim-to-real consistency. The qualitative match—Inverted-U for the harder task and monotonic decay for easier tasks—supports both the mastery–entropy principle and the use of H_G as a predictive pre-training diagnostic for real robot data.

D. Robustness of H_G: Overcoming Metric Aliasing

We test whether H_G captures learning-relevant geometric structure missed by common diversity proxies. We compare against four baselines, all computed on the *same aligned descriptors* as H_G (Section III-C): (1) **Mean Variance** σ^2_{avg}; (2) **LogDet Covariance** $\log\det(\Sigma + \alpha I)$; (3) **Participation Ratio** $(\sum_j \lambda_j)^2 / \sum_j \lambda_j^2$; and (4) **kNN Differential Entropy** (Kozachenko–Leonenko). We include αI only to stabilize LogDet under $M \ll 3T$. Fig. 6 shows their behavior on `StackCube-v1` across the same (r, k) configurations summarized by Table I.

Baseline failure mode 1: marginal spread ignores mode organization (mean variance). σ^2_{avg} aggregates waypoint-wise spread and therefore discards cross-time correlations and mode structure. As a result, low-radius multi-mode settings $(r, k){=}(0, 0)$, $(0.01, 0)$, and $(0.03, 0)$ collapse to nearly the same value despite increasing H_G^{avg} (Table I). Moreover, large-excursion single-mode settings $(0.05, 1)$ and $(0.10, 1)$ can be mis-ordered relative to these cases, since

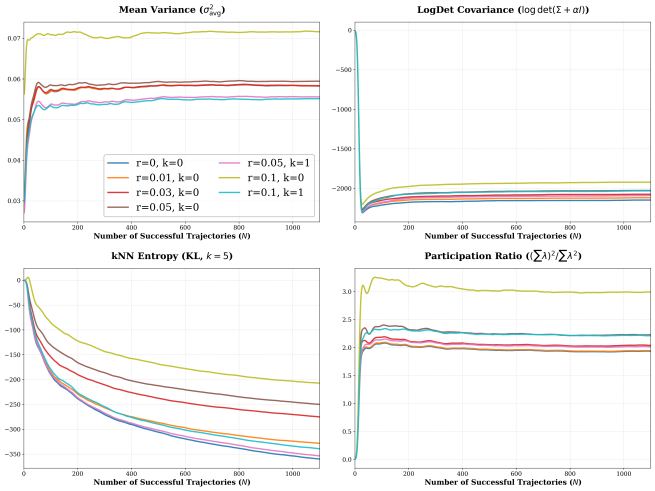


Fig. 6: Baseline metric convergence on `StackCube-v1`. Mean variance, LogDet, and participation ratio converge quickly but collapse distinct (r, k) settings (aliasing/misordering), e.g., $(0, 0)$, $(0.01, 0)$, and $(0.03, 0)$ are nearly indistinguishable under variance despite increasing H_G. kNN entropy shows pronounced sample-size drift in the D=150 descriptor space.

marginal spread alone cannot distinguish "many tight modes" from "one consistent corridor."

Baseline failure mode 2: second-order summaries alias multimodality (LogDet and PR). LogDet and participation ratio incorporate spectral information but still compress the dataset to a single covariance summary, so distinct multimodal organizations can share similar values, leading to aliasing across (r, k). For example, participation ratio can place $(0.10, 1)$ close to $(0.05, 0)$ even though H_G^{avg} ranks $(0.10, 1)$ much lower, consistent with reduced strategy ambiguity under k=1. In addition, LogDet requires explicit regularization under rank deficiency, making it sensitive to α.

Baseline failure mode 3: instability at practical M (kNN entropy). kNN differential entropy yields a broadly reasonable ordering but shows pronounced sample-size drift over $N \leq 1000$ in the D=150 descriptor space, reducing its usefulness as a stable pre-training audit signal from $O(10^2)$ demonstrations.

Why H_G is better behaved. In contrast, H_G combines (i) the number of independent shape modes (effective dimensionality; Eq. 1) and (ii) their overall spread (Eq. 2), directly reflecting both multimodality and geometric magnitude in aligned shape space. This makes it less prone to collapsing distinct strategy organizations while converging rapidly (Fig. 2).

VI. DISCUSSION

A. From Heuristic to Principle: Mastery–Entropy Matching

Practitioners have long observed that heavily randomized demonstrations can hurt imitation learning, but these effects are often treated as idiosyncratic. Our results suggest a more general organizing principle: a learner benefits from geometric diversity only up to a limit set by its current *task mastery*, beyond which additional modes introduce ambiguity and degrade performance. This reframes "data quality" along the trajectory-shape axis: instead of maximizing diversity by default, one should match diversity to what the learner can reliably absorb.

Practically, H_G enables a simple pre-training diagnostic: estimate H_G on a pilot dataset and compare it to expected trend shapes for the target model and data scale. If H_G is very low, the dataset likely covers a narrow corridor and may be brittle; if H_G is very high, the dataset likely mixes incompatible strategies and may induce mode interference.

B. Limitations

Our study focuses on a set of representative manipulation tasks in simulation and on hardware. Extending the evaluation to a broader range of task families (e.g., deformables, tool use, less contact-constrained skills, and longer-horizon assembly) would help further assess how widely the mastery–entropy behavior carries over and refine practical guidance across domains.

Our current pipeline leverages task semantics (e.g., gripper events) to segment episodes into transit phases. Generalizing the phase decomposition to settings without clear event markers—for example via automatic segmentation or state-based heuristics—is a natural next step.

Finally, the Inverted-U trend and peak-shift are empirical regularities supported by our experiments. Developing a more formal treatment that connects the optimal H_G^* to factors such as model capacity, conditioning information, and data coverage remains an interesting direction for future work.

VII. CONCLUSION

We introduced **Geometric Entropy** (H_G), a lightweight pre-training metric for trajectory-shape diversity in robot demonstration datasets. By aligning transit trajectories in a target-anchored frame to remove extrinsic variation and summarizing the aligned shape manifold with a stable spectral statistic, H_G provides a practical audit signal that can be computed before any policy training.

Across simulation and real-robot contact-rich manipulation tasks, we observe a consistent **Inverted-U** relationship between success and H_G: moderate diversity improves robustness via coverage, while excessive diversity harms performance via mode ambiguity. The optimal entropy H_G^* shifts toward lower values as task mastery increases through more data, easier tasks, or stronger priors, and can become effectively monotonically decreasing for pretrained VLAs. Together, these results challenge the "maximize diversity" heuristic and suggest a calibration principle: match demonstration diversity to the learner's mastery, using H_G as a fast diagnostic for data curation.

REFERENCES

[1] A. Brohan, N. Brown, J. Carbajal, Y. Chebotar, J. Dabis, C. Finn, K. Gopalakrishnan, K. Hausman, A. Herzog, J. Hsu *et al.*, "Rt-1: Robotics transformer for real-world control at scale," *arXiv preprint arXiv:2212.06817*, 2022.

[2] B. Zitkovich, T. Yu, S. Xu, P. Xu, T. Xiao, F. Xia, J. Wu, P. Wohlhart, S. Welker, A. Wahid *et al.*, "Rt-2: Vision-language-action models transfer web knowledge to robotic control," in *Conference on Robot Learning*. PMLR, 2023, pp. 2165–2183.

[3] A. O'Neill, A. Rehman, A. Maddukuri, A. Gupta, A. Padalkar, A. Lee, A. Pooley, A. Gupta, A. Mandlekar, A. Jain *et al.*, "Open x-embodiment: Robotic learning datasets and rt-x models: Open x-embodiment collaboration 0," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 6892–6903.

[4] A. Khazatsky, K. Pertsch, S. Nair, A. Balakrishna, S. Dasari, S. Karamcheti, S. Nasiriany, M. K. Srirama, L. Y. Chen, K. Ellis *et al.*, "Droid: A large-scale in-the-wild robot manipulation dataset," *arXiv preprint arXiv:2403.12945*, 2024.

[5] Octo Model Team, D. Ghosh, H. Walke, K. Pertsch, K. Black, O. Mees, S. Dasari, J. Hejna, C. Xu, J. Luo, T. Kreiman, Y. Tan, L. Y. Chen, P. Sanketi, Q. Vuong, T. Xiao, D. Sadigh, C. Finn, and S. Levine, "Octo: An open-source generalist robot policy," in *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, 2024.

[6] M. Kim, K. Pertsch, S. Karamcheti, T. Xiao, A. Balakrishna, S. Nair, R. Rafailov, E. Foster, G. Lam, P. Sanketi, Q. Vuong, T. Kollar, B. Burchfiel, R. Tedrake, D. Sadigh, S. Levine, P. Liang, and C. Finn, "Openvla: An open-source vision-language-action model," *arXiv preprint arXiv:2406.09246*, 2024.

[7] E. Jang, A. Irpan, M. Khansari, D. Kappler, F. Ebert, C. Lynch, S. Levine, and C. Finn, "Bc-z: Zero-shot task generalization with robotic imitation learning," in *conference on Robot Learning*. PMLR, 2022, pp. 991–1002.

[8] K. Bousmalis, G. Vezzani, D. Rao, C. Devin, A. X. Lee, M. Bauzá, T. Davchev, Y. Zhou, A. Gupta, A. Raju *et al.*, "Robocat: A self-improving generalist agent for robotic manipulation," *arXiv preprint arXiv:2306.11706*, 2023.

[9] D. Driess, F. Xia, M. S. Sajjadi, C. Lynch, A. Chowdhery, B. Ichter, A. Wahid, J. Tompson, Q. Vuong, T. Yu *et al.*, "Palm-e: an embodied multimodal language model," in *Proceedings of the 40th International Conference on Machine Learning*, 2023, pp. 8469–8488.

[10] Y. Jiang, A. Gupta, Z. Zhang, G. Wang, Y. Dou, Y. Chen, L. Fei-Fei, A. Anandkumar, Y. Zhu, and L. Fan, "Vima: General robot manipulation with multimodal prompts," in *Fortieth International Conference on Machine Learning*, 2023.

[11] Y. Lu, K. Hausman, Y. Chebotar, M. Yan, E. Jang, A. Herzog, T. Xiao, A. Irpan, M. Khansari, D. Kalashnikov *et al.*, "Aw-opt: Learning robotic skills with imitation andreinforcement at scale," in *Conference on Robot Learning*. PMLR, 2022, pp. 1078–1088.

[12] S. Belkhale, Y. Cui, and D. Sadigh, "Data quality in imitation learning," *Advances in neural information processing systems*, vol. 36, pp. 80 375–80 395, 2023.

[13] S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 627–635.

[14] M. Laskey, J. Lee, R. Fox, A. Dragan, and K. Goldberg, "Dart: Noise injection for robust imitation learning," in *Conference on robot learning*. PMLR, 2017, pp. 143–156.

[15] A. Mandlekar, D. Xu, J. Wong, S. Nasiriany, C. Wang, R. Kulkarni, L. Fei-Fei, S. Savarese, Y. Zhu, and R. Martín-Martín, "What matters in learning from offline human demonstrations for robot manipulation," in *Conference on Robot Learning*. PMLR, 2022, pp. 1678–1690.

[16] H. Xu, X. Zhan, H. Yin, and H. Qin, "Discriminator-weighted offline imitation learning from suboptimal demonstrations," in *International Conference on Machine Learning*. PMLR, 2022, pp. 24 725–24 742.

[17] Y. Zhang, Y. Xie, H. Liu, R. Shah, M. Wan, L. Fan, and Y. Zhu, "Scizor: Self-supervised data curation for large-scale imitation learning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2026.

[18] C. Agia, R. Sinha, J. Yang, R. Antonova, M. Pavone, H. Nishimura, M. Itkina, and J. Bohg, "Cupid: Curating data your robot loves with influence functions," *arXiv preprint arXiv:2506.19121*, 2025.

[19] A. S. Chen, A. M. Lessing, Y. Liu, and C. Finn, "Curating demonstrations using online experience," *arXiv preprint arXiv:2503.03707*, 2025.

[20] K. Wu, N. Liu, Z. Zhao, D. Qiu, J. Li, Z. Che, Z. Xu, and J. Tang, "Learning from imperfect demonstrations with self-supervision for robotic manipulation," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 16 899–16 906.

[21] A. Mandlekar, S. Nasiriany, B. Wen, I. Akinola, Y. Narang, L. Fan, Y. Zhu, and D. Fox, "Mimicgen: A data generation system for scalable robot learning using human demonstrations," in *Conference on Robot Learning*. PMLR, 2023, pp. 1820–1864.

[22] Z. Xue, S. Deng, Z. Chen, Y. Wang, Z. Yuan, and H. Xu, "Demogen: Synthetic demonstration generation for data-efficient visuomotor policy learning," in *7th Robot Learning Workshop: Towards Robots with Human-Level Abilities*, 2025.

[23] H. Tan, X. Xu, C. Ying, X. Mao, S. Liu, X. Zhang, H. Su, and J. Zhu, "Manibox: Enhancing spatial grasping generalization via scalable simulation data generation," *arXiv preprint arXiv:2411.01850*, 2024.

[24] S. Huang, Y. Liao, S. Feng, S. Jiang, S. Liu, H. Li, M. Yao, and G. Ren, "Adversarial data collection: Human-collaborative perturbations for efficient and robust robotic imitation learning," *arXiv preprint arXiv:2503.11646*, 2025.

[25] W. Wang, K. Ye, X. Zhou, T. Chen, C. Min, Q. Zhu, X. Yang, P. Luo, Y. Shen, Y. Yang *et al.*, "Fieldgen: From teleoperated pre-manipulation trajectories to field-guided data generation," *arXiv preprint arXiv:2510.20774*, 2025.

[26] H. Wang, C. B. Chen, Y. Yue, D. Tao, T. Guo, S. Xie, D. Huang, S. Song, G. Yao, and G. Huang, "Move: A simple motion-based data collection paradigm for spatial generalization in robotic manipulation," *arXiv preprint arXiv:2512.04813*, 2025.

[27] K. Hausman, Y. Chebotar, S. Schaal, G. Sukhatme, and J. J. Lim, "Multi-modal imitation learning from unstructured demonstrations using generative adversarial nets," *Advances in neural information processing systems*, vol. 30, 2017.

[28] N. M. Shafiullah, Z. Cui, A. A. Altanzaya, and L. Pinto, "Behavior transformers: Cloning k modes with one stone," *Advances in neural information processing systems*, vol. 35, pp. 22 955–22 968, 2022.

[29] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," *The International Journal of Robotics Research*, vol. 44, no. 10-11, pp. 1684–1704, 2025.

[30] P. Florence, C. Lynch, A. Zeng, O. A. Ramirez, A. Wahid, L. Downs, A. Wong, J. Lee, I. Mordatch, and J. Tompson, "Implicit behavioral cloning," in *Conference on robot learning*. PMLR, 2022, pp. 158–168.

[31] S. Lee, Y. Wang, H. Etukuru, H. J. Kim, N. M. M. Shafiullah, and L. Pinto, "Behavior generation with latent actions," in *International Conference on Machine Learning*. PMLR, 2024, pp. 26 991–27 008.

[32] S. Haldar, Z. Peng, and L. Pinto, "Baku: An efficient transformer for multi-task policy learning," *Advances in Neural Information Processing Systems*, vol. 37, pp. 141 208–141 239, 2024.

[33] X. Jia, D. Blessing, X. Jiang, M. Reuss, A. Donat, R. Lioutikov, and G. Neumann, "Towards diverse behaviors: A benchmark for imitation learning with human demonstrations," in *The Twelfth International Conference on Learning Representations*, 2024.

[34] J. Ho and S. Ermon, "Generative adversarial imitation learning," *Advances in neural information processing systems*, vol. 29, 2016.

[35] M. Shi, L. Chen, J. Chen, Y. Lu, C. Liu, G. Ren, P. Luo, D. Huang, M. Yao, and H. Li, "Is diversity all you need for scalable robotic manipulation?" *arXiv preprint arXiv:2507.06219*, 2025.

[36] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE transactions on acoustics, speech, and signal processing*, vol. 26, no. 1, pp. 43–49, 1978.

[37] S. Tao, F. Xiang, A. Shukla, Y. Qin, X. Hinrichsen, X. Yuan, C. Bao, X. Lin, Y. Liu, T.-k. Chan *et al.*, "Maniskill3: Gpu parallelized robotics simulation and rendering for generalizable embodied ai," *arXiv preprint arXiv:2410.00425*, 2024.

[38] P. Intelligence, K. Black, N. Brown, J. Darpinian, K. Dhabalia, D. Driess, A. Esmail, M. Equi, C. Finn, N. Fusai *et al.*, "$\pi_{0.5}$: a vision-language-action model with open-world generalization," *arXiv preprint arXiv:2504.16054*, 2025.

[39] "Arx robotics," https://arx-x.com/.

[40] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," *arXiv preprint arXiv:2304.13705*, 2023.