CURVATURE SIGN DETERMINES GEOMETRIC ERROR-PROPAGATION REGIMES IN BEHAVIORAL CLONING

Anonymous authors
Paper under double-blind review

ABSTRACT

Behavioral cloning (BC) optimizes one-step prediction, but deployment feeds predictions back into future states, so local errors compound. We ask how rollout geometry determines the regime in which these errors accumulate without online correction. On homogeneous constant-curvature manifolds, we derive a parameter-free local recurrence separating inherited-error propagation from fresh-error injection. Under direction averaging and zero state-dependent policy amplification ($L = 0$), the reduced recurrence yields three regimes: a bounded squared-error floor under positive curvature, diffusive $\sqrt{T}$ root-mean-square growth in flat space, and exponential amplification under negative curvature; a conditional orientation band extends the result beyond exact averaging. Simulations match the theory within 3% in second moment on $\mathbb{S}^2$ and $\mathbb{R}^2$, and $\mathrm{SO}(3)$ exhibits the predicted stabilization. Learned-policy evaluations on $\mathrm{SO}(3)$ and in an articulated execution chain show that matched one-step error can coexist with different rollout errors, consistent with measured propagation response. At fixed decision/noise-injection counts, longer action chunks strengthen positive-curvature contraction and negative-curvature amplification, while the negative-curvature regime offers a geometric perspective on fragility in hyperbolic reinforcement learning. Curvature sign sets the accumulation regime, so sequential evaluation should measure propagation, not just error magnitude.

1 INTRODUCTION

Behavioral cloning (BC) reduces imitation learning to supervised prediction on expert demonstrations, but deployment is sequential: each prediction changes the state on which the next prediction is made. Small local mistakes can therefore alter future inputs and accumulate throughout a rollout, the compounding-error problem (Ross & Bagnell, 2010; Ross et al., 2011). This creates a fundamental mismatch between the one-step objective optimized during BC training and the long-horizon behavior that matters at deployment: low one-step imitation error alone need not predict closed-loop behavior, and policies with similar local accuracy can differ substantially in rollout error. Understanding how local errors evolve after they occur is therefore essential for connecting supervised imitation accuracy to sequential performance.

Classical BC and Dataset Aggregation (DAgger) analyses characterize compounding through distribution shift and its horizon dependence (Ross & Bagnell, 2010; Ross et al., 2011), while recent work shows that state-dependent policy response can amplify deviations exponentially even under stable Euclidean dynamics (Simchowitz et al., 2025). Beyond these compounding-error analyses, several widely used settings make rollout geometry consequential: robot orientations lie on the positively curved rotation manifold $\mathrm{SO}(3)$ (Zhou et al., 2019), hyperbolic reinforcement learning (RL) uses negatively curved representations (Cetin et al., 2023), and action chunking changes the effective step length over which geometry acts (Zhao et al., 2023; Chi et al., 2023). Yet these lines of work do not isolate how rollout geometry propagates state separation already created by past prediction errors. This gap motivates our question: *In behavioral cloning without online correction, how does rollout geometry determine the regime in which local prediction errors accumulate?*

Answering this question requires a geometric decomposition that existing compounding-error analyses do not make. At each step, the learner carries an inherited state separation from past predictions and introduces a fresh action error through the current prediction. Geometrically, the former perturbs initial position under parallel initial velocities, whereas the latter perturbs initial velocity from

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.

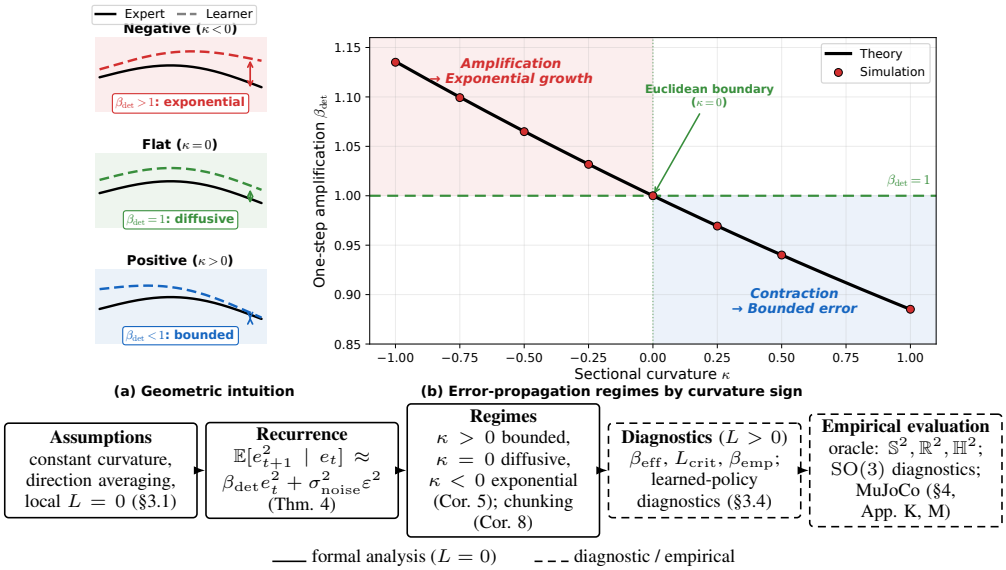


Figure 1: **Curvature sign determines geometric error-propagation regimes.** *(a)* Local geometric intuition under parallel initial velocities. *(b)* Direction-averaged inherited-error propagation factor $\beta_{\det}$ at $v = 0.5$ (Theorem 4): theory and simulation without injected noise. Bottom: overview; solid boxes mark formal $L = 0$ results, dashed boxes diagnostics and empirical evaluation.

a common state, so curvature acts on them through different one-step maps. Moreover, because curvature acts only on motion-transverse components, inherited-error propagation depends on the error–action orientation. Thus, a curvature-based account must separate error injection from error propagation and account for the orientation dependence of inherited-error propagation, rather than reduce compounding to a single scalar error magnitude.

We answer this question in a direction-averaged local model on homogeneous constant-curvature manifolds with zero state-dependent policy amplification ($L = 0$): curvature sign determines the regime of leading-order geometric error propagation (Figure 1). This result follows from a Jacobi-field analysis that yields separate closed-form factors for inherited-error propagation and fresh-error injection and combines them into a local second-moment recurrence. Specifically, the reduced recurrence yields a time-uniform squared-error floor proportional to the squared local error scale under positive curvature, squared error that grows linearly in the horizon T—equivalently, diffusive $\sqrt{T}$ root-mean-square (RMS) growth—in flat space, and exponential amplification under negative curvature. Beyond exact direction averaging, a conditional orientation band gives explicit upper and lower recurrences. Empirically, we test the recurrence in controlled constant-curvature simulations and evaluate its propagation diagnostics on learned $\mathrm{SO}(3)$ policies and an articulated execution chain: the recurrence is quantitatively accurate in the tested positive- and zero-curvature settings and captures the local amplification trend under negative curvature, while matched-error comparisons show that measured propagation response carries information beyond one-step error.

Our contributions are:

- **A curvature-explicit theory of error-propagation regimes.** We derive a local recurrence separating inherited-error propagation from fresh-error injection and show that curvature sign determines the error-propagation regime (bounded, diffusive, or exponential); a conditional orientation band extends the theory beyond exact direction averaging, indicating whether inherited error is attenuated, preserved, or amplified.
- **Propagation-aware evaluation of learned policies.** On $\mathrm{SO}(3)$ and an articulated execution chain, closely matched one-step error can coexist with different rollout errors, showing that measured propagation response can distinguish policies that one-step error cannot.
- **A curvature-dependent effect of action chunking.** At fixed decision/noise-injection counts, longer chunks strengthen positive-curvature contraction and negative-curvature amplification over the studied range, while flat-space propagation is unchanged; curvature sign thus indicates whether longer chunks damp or amplify inherited error.

- **A geometric implication for hyperbolic RL.** A curvature surrogate links negative-curvature amplification to fragility reported in hyperbolic reinforcement learning (Cetin et al., 2023), providing a complementary geometric interpretation.

2 Related Work

Compounding error in imitation learning. Classical BC and Dataset Aggregation (DAgger) analyses (Ross & Bagnell, 2010; Ross et al., 2011) yield cumulative-cost guarantees with quadratic and linear horizon dependence, respectively. Subsequent work mitigates BC compounding through noise injection and stability-oriented training (Laskey et al., 2017; Mehta et al., 2025). Recent work sharpens this picture: exponential compounding error can be unavoidable for smooth deterministic policies in continuous Euclidean state-action spaces, even under stable dynamics (Simchowitz et al., 2025). In that line of work, the exponential effect arises in a Euclidean setting with state-dependent amplification, whereas the exponential regime studied here arises from negative manifold curvature in the oracle ($L = 0$) setting. Action chunking and exploratory data collection can mitigate compounding in Euclidean settings, where recent work analyzes both through control-theoretic stability of the open-loop dynamics (Zhang et al., 2026), a mechanism distinct from the effective-step geometric channel studied here. Curvature-induced contraction is classical for Markov kernels and diffusion semigroups (Ollivier, 2009; von Renesse & Sturm, 2005), while contraction on Riemannian state spaces is well studied for dynamical systems (Simpson-Porco & Bullo, 2014). These frameworks characterize contraction of a prescribed kernel or vector field; our BC analysis instead separates inherited state separation from fresh action error, which induce distinct Jacobi-field initial conditions and enter the rollout recurrence differently.

Geometry in learning and robotics. Riemannian and hyperbolic methods support representation learning (Nickel & Kiela, 2017; Ganea et al., 2018; Chami et al., 2019), reinforcement learning (Cetin et al., 2023), and manifold optimization (Bonnabel, 2013; Bécigneul & Ganea, 2019), with numerical challenges studied in Peng et al. (2022); Mishne et al. (2023). Manifold-valued imitation and motion learning address representation, skill encoding, and motion synthesis (Calinon, 2020; Zeestraten et al., 2017; Beik-Mohammadi et al., 2021; Braun et al., 2024). In robotics (Ravichandar et al., 2020), learned rotation representations (Zhou et al., 2019; Levinson et al., 2020; Peretroukhin et al., 2020) and action chunking (Zhao et al., 2023; Chi et al., 2023) motivate the rollout geometries studied here. These methods use geometry for representation, optimization, skill encoding, or policy design, but do not characterize how already-present error propagates through closed-loop rollouts on the state manifold; we complement them with a curvature-explicit local recurrence separating inherited-error propagation from fresh-error injection.

3 Theory

3.1 Setup

To isolate the effect of geometry on error propagation, we analyze a minimal local model on a complete d-dimensional Riemannian manifold (M, g) with constant sectional curvature κ. States s_t lie on M, and actions are tangent increments in $T_{s_t}M$ that generate the next state through the exponential map. An expert policy π^* maps states to tangent vectors: $\pi^*(s) \in T_sM$ with $\|\pi^*(s)\| = v$ for a fixed step size $v > 0$. The expert trajectory is $s^*_{t+1} = \exp_{s^*_t}(\pi^*(s^*_t))$; the action direction may change at every step, so the trajectory need not be geodesic. Discrete $\mathrm{SO}(3)$ orientation updates provide a direct kinematic instance (Appendices B and K). In the $L = 0$ oracle comparison, the learner executes $\hat{a}_t = \Gamma_{s^*_t \to \hat{s}_t}\, \pi^*(s^*_t) + \xi_t$: the parallel transport Γ of the expert action to the learner's state, plus a fresh perturbation ξ_t in the learner's tangent space. Conditioned on the current action and the pre-noise history, $\xi_t = \varepsilon z_t$, where z_t follows a fixed zero-mean isotropic law with bounded support and $\mathbb{E}[\|z_t\|^2] = 1$, representing fresh local action error; throughout, $d \geq 2$. The learner's trajectory is $\hat{s}_{t+1} = \exp_{\hat{s}_t}(\hat{a}_t)$, and we study the geodesic tracking error $e_t = d(s^*_t, \hat{s}_t)$.

Action steps are taken within the injectivity radius when it is finite, and local comparisons use unique minimizing geodesics with sufficiently small state separations and perturbations. For learned policies, we define the *policy sensitivity* $L \geq 0$ of the deterministic action map $\hat{\pi}$ before injected noise by $\|\hat{\pi}(s_1) - \Gamma_{s_2 \to s_1}\hat{\pi}(s_2)\| \leq L \cdot d(s_1, s_2)$, where Γ denotes parallel transport along the minimizing geodesic. The $L = 0$ oracle isolates geometric propagation; $L > 0$ policies are evaluated through the diagnostic in Section 3.4. Geodesic stepping specifies the update map, while direction averaging specifies the error–action orientation distribution; Proposition 6 relaxes this assumption. Complete $L = 0$ proofs are in Appendix A.

3.2 Main Result: Two Geometric Mechanisms

The key insight is that error propagation on a curved manifold involves two geometric mechanisms, each governed by a different Jacobi field initial condition (do Carmo, 1992, Ch. 5).

Lemma 1 (Inherited-Error Propagation). *Let $p, q \in M$ with $d(p,q) = \delta$ (small). Let $a \in T_pM$ with $\|a\| = v$ be a random direction uniform on the sphere in T_pM, and let $\tilde{a} = \Gamma_{p\to q}\, a \in T_qM$ be its parallel transport. Then:*

$$\mathbb{E}_a\big[d(\exp_p(a),\, \exp_q(\tilde{a}))^2\big] = \beta_{\det}(v) \cdot \delta^2 + O(\delta^3) \tag{1}$$

*where the **inherited-error propagation factor** is:*

$$\beta_{\det}(v) = \frac{1}{d} + \frac{d-1}{d}\, \mathrm{c}_\kappa^2(v), \qquad \mathrm{c}_\kappa(v) = \begin{cases} \cos(\sqrt{\kappa}\, v) & \kappa > 0 \\ 1 & \kappa = 0 \\ \cosh(\sqrt{|\kappa|}\, v) & \kappa < 0 \end{cases} \tag{2}$$

Proof. Consider geodesics $\gamma_1(t) = \exp_p(ta/v)$ and $\gamma_2(t) = \exp_q(t\tilde{a}/v)$ with parallel initial velocities. Their separation defines a Jacobi field $J(t)$ along γ_1 with $J(0) = \log_p(q)$, where $\log_p$ denotes the local Riemannian logarithm, $\|J(0)\| = \delta$, and $\nabla_t J(0) = 0$ (velocities are parallel). Decompose $J(0) = J_\parallel(0) + J_\perp(0)$ into components parallel and perpendicular to $\dot{\gamma}_1(0)$. On a constant-curvature space: the tangential component satisfies $J_\parallel(v) = J_\parallel(0)$ (preserved), while the normal component satisfies $J_\perp(v) = \mathrm{c}_\kappa(v) \cdot J_\perp(0)$. Thus $d(\gamma_1(v), \gamma_2(v))^2 = \|J_\parallel(0)\|^2 + \mathrm{c}_\kappa^2(v)\|J_\perp(0)\|^2 + O(\delta^3)$. Since a is uniformly random, $\mathbb{E}[\|J_\parallel(0)\|^2] = \delta^2/d$ and $\mathbb{E}[\|J_\perp(0)\|^2] = (d-1)\delta^2/d$, yielding $\beta_{\det} \cdot \delta^2$. □

Fresh action error follows a different Jacobi-field initial condition, yielding a distinct one-step factor.

Lemma 2 (Noise Amplification through Exponential Map). *Let $a \in T_pM$ with $\|a\| = v$, and let $\xi = \varepsilon z \in T_pM$ follow the noise model of Section 3.1. Then:*

$$\mathbb{E}_\xi\big[d(\exp_p(a),\, \exp_p(a+\xi))^2\big] = \sigma^2_{\mathrm{noise}}(v) \cdot \varepsilon^2 + O(\varepsilon^3) \tag{3}$$

*where the **noise amplification factor** is:*

$$\sigma^2_{\mathrm{noise}}(v) = \frac{1}{d} + \frac{d-1}{d}\, \mathrm{s}_\kappa^2(v), \qquad \mathrm{s}_\kappa(v) = \begin{cases} \sin(\sqrt{\kappa}\, v)/(\sqrt{\kappa}\, v) & \kappa > 0 \\ 1 & \kappa = 0 \\ \sinh(\sqrt{|\kappa|}\, v)/(\sqrt{|\kappa|}\, v) & \kappa < 0 \end{cases} \tag{4}$$

Proof. The differential $(d\exp_p)_a$ maps $\xi = \xi_\parallel + \xi_\perp$ as follows: the radial component satisfies $\|(d\exp_p)_a\, \xi_\parallel\| = \|\xi_\parallel\|$ (isometry along the geodesic), while the transverse component satisfies $\|(d\exp_p)_a\, \xi_\perp\| = \mathrm{s}_\kappa(v)\, \|\xi_\perp\|$ (Jacobi field with $J(0) = 0$, $\nabla_t J(0) = \xi_\perp/v$, giving $J(v) = \mathrm{s}_\kappa(v)\, \xi_\perp$). For isotropic ξ: $\mathbb{E}[\|\xi_\parallel\|^2] = \varepsilon^2/d$ and $\mathbb{E}[\|\xi_\perp\|^2] = (d-1)\varepsilon^2/d$, yielding $\sigma^2_{\mathrm{noise}} \cdot \varepsilon^2$. □

Remark 3 (Why $\cos^2$ vs. sinc^2). *The two factors arise from* different *Jacobi field initial conditions. Inherited-error propagation (Lemma 1) uses $\|J(0)\| = \delta$, $\nabla_t J(0) = 0$, yielding the* endpoint *value $\mathrm{c}_\kappa(v)$. Noise amplification (Lemma 2) uses $J(0) = 0$, $\nabla_t J(0) = \xi/v$, yielding the* average-like *value $\mathrm{s}_\kappa(v)$. Since $\mathrm{sinc}(x) = \frac{1}{x}\int_0^x \cos(u)\, du > \cos(x) > 0$ for $x \in (0, \pi/2)$, we have $\mathrm{c}_\kappa^2(v) < \mathrm{s}_\kappa^2(v)$ for $\sqrt{\kappa}\, v \in (0, \pi/2)$: deterministic contraction is stronger than noise contraction over this step-size range when $\kappa > 0$. The inherited-error coefficient $\beta_{\det}$, rather than the fresh-error coefficient $\sigma^2_{\mathrm{noise}}$, governs how past error propagates through the recurrence. For the unit action direction $u = a/v$, the small-step expansion is $\beta_{\det}(v) = 1 - \frac{d-1}{d}\kappa v^2 + O(v^4) = 1 - \frac{1}{d}\,\mathrm{Ric}(u,u)\, v^2 + O(v^4)$, consistent with the classical curvature–contraction connection.*

3.3 Error Recurrence

Combining the inherited-error and fresh-error contributions gives the following local recurrence.

Theorem 4 (Local Squared-Error Recurrence). *Under the setup of Section 3.1, with $L = 0$, direction averaging as in Lemma 1, and sufficiently small e_t and ε:*

$$\mathbb{E}[e_{t+1}^2 \mid e_t] \;=\; \beta_{\mathrm{det}}(v) \cdot e_t^2 \;+\; \sigma_{\mathrm{noise}}^2(v) \cdot \varepsilon^2 \;+\; O(e_t^3 + e_t\varepsilon^2 + \varepsilon^3)\,. \tag{5}$$

Proof sketch. Condition on the pre-noise configuration. The noise-free endpoint separation contributes $\beta_{\mathrm{det}}(v)\, e_t^2 + O(e_t^3)$ after direction averaging by Lemma 1, while Lemma 2 gives the fresh-noise contribution $\sigma_{\mathrm{noise}}^2(v)\, \varepsilon^2 + O(\varepsilon^3)$. In normal coordinates around the noise-free learner endpoint, the first-order noise displacement is linear in ξ_t; conditional zero mean therefore removes the leading cross term, leaving $O(e_t\varepsilon^2)$. Combining these terms and taking the conditional expectation given e_t yields equation 5; Appendix A.4 gives the full argument. □

Dropping the remainder in equation 5 and taking expectations gives the *reduced leading-order recurrence* $x_{t+1} = \beta_{\mathrm{det}}\, x_t + \sigma_{\mathrm{noise}}^2\, \varepsilon^2$ for the second moment x_t of the local model, with three regimes:

Corollary 5 (Three Error Regimes). *For this reduced recurrence with $x_0 = 0$ and $L = 0$: (i) **Positive curvature** ($\kappa > 0$): $\beta_{\mathrm{det}} < 1$, so $x_t \to \sigma_{\mathrm{noise}}^2\varepsilon^2/(1 - \beta_{\mathrm{det}})$—a bounded squared-error floor. (ii) **Zero curvature** ($\kappa = 0$): $\beta_{\mathrm{det}} = 1$, so $x_T = T \cdot \varepsilon^2$—linear growth of squared tracking error, i.e., diffusive $\sqrt{T}$ growth of root-mean-square tracking error. (iii) **Negative curvature** ($\kappa < 0$): $\beta_{\mathrm{det}} > 1$, so $x_T = \sigma_{\mathrm{noise}}^2\varepsilon^2(\beta_{\mathrm{det}}^T - 1)/(\beta_{\mathrm{det}} - 1)$—exponential.*

The following conditional band extends exact direction averaging; its proof is in Appendix A.7.

Proposition 6 (Conditional Orientation Band). *Under the setup of Section 3.1 with $L = 0$, let θ_t be the angle between the current separation direction and the action direction, and let $\mathcal{G}_t := \sigma(e_t, \mathcal{F}_{t-1})$ denote the sigma-algebra generated by the current error and the history $\mathcal{F}_{t-1}$ before the current action direction is realized. Suppose there exists $\eta \geq 0$ such that, for every t, almost surely,*

$$\left|\mathbb{E}[\cos^2\theta_t \mid \mathcal{G}_t] - \tfrac{1}{d}\right| \leq \eta\,. \tag{6}$$

Let $\Delta_\eta = |1 - \mathrm{c}_\kappa^2(v)|\,\eta$, $\beta_- = \max\{0,\, \beta_{\mathrm{det}} - \Delta_\eta\}$, and $\beta_+ = \beta_{\mathrm{det}} + \Delta_\eta$. Then the second moment $m_t := \mathbb{E}[e_t^2]$ of the reduced leading-order model satisfies

$$x_t^- \leq m_t \leq x_t^+, \qquad x_{t+1}^\pm = \beta_\pm\, x_t^\pm + \sigma_{\mathrm{noise}}^2\, \varepsilon^2, \qquad x_0^\pm = m_0\,. \tag{7}$$

The band permits temporal dependence and correlation with the current error. It certifies a bounded upper recurrence when $\beta_+ < 1$ and a supercritical lower recurrence when $\beta_- > 1$; when the interval straddles 1, $[\beta_-, \beta_+]$ makes explicit the remaining orientation-induced uncertainty in regime classification. Exact averaging recovers x_t; further consequences are in Appendix A.7.

Remark 7 (Angle-dependent recursion). *Before scalar averaging, the inherited-error propagation factor is $\beta(\theta_t) = \cos^2\theta_t + \mathrm{c}_\kappa^2(v)\sin^2\theta_t$: the tangential component is preserved, while curvature acts on the transverse component, so error exactly aligned with the action direction is unchanged at leading order. Modeling the angle process of a structured policy, or establishing its conditional band equation 6, extends this local description beyond exact direction averaging.*

3.4 Corollaries

Policy-sensitivity diagnostic. State dependence of the executed policy introduces action variation of order $L\, e_t$. We use $\beta_{\mathrm{eff}} = \beta_{\mathrm{det}}\,(1 + Lv)^2$ as a diagnostic proxy for the interaction between geometric propagation and policy sensitivity (motivated in Appendix A), not a certified bound. Its subcriticality condition $\beta_{\mathrm{eff}} < 1$ defines the nominal sensitivity margin

$$L_{\mathrm{crit}} = \frac{1/\sqrt{\beta_{\mathrm{det}}} - 1}{v}\,, \tag{8}$$

which is used together with a learned policy's measured one-step error and the empirical one-step state-amplification factor β_{emp}, measured as the ratio of squared next-state separation to squared input separation under a small state perturbation (Section 4.5, Appendix I), to assess learned policies.

Action chunking. Under a *fixed-event* protocol (decision/noise-injection count held constant), chunk size H rescales the effective step to $v_{\mathrm{eff}} = Hv$. Under positive curvature, $\beta_{\mathrm{det}}(Hv)$ decreases

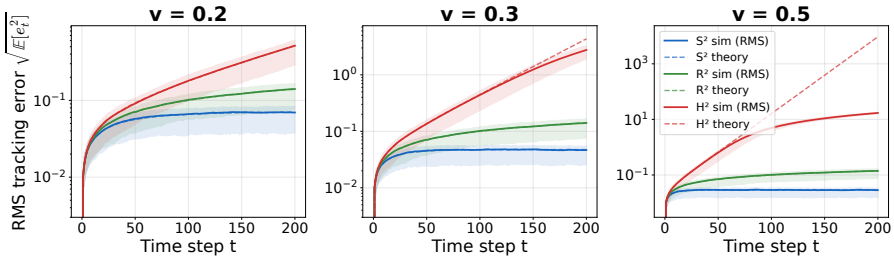


Figure 2: **Curvature sign determines three error regimes.** RMS tracking error $\sqrt{\mathbb{E}[e_t^2]}$ under oracle BC ($L = 0$, $\varepsilon = 0.01$) on $\mathbb{S}^2$ (blue), $\mathbb{R}^2$ (green), and $\mathbb{H}^2$ (red; expert-centred isometry frame, Appendix B). Solid: simulation RMS over 1500 trials; shaded band: interquartile range of e_t. Dashed: square root of the reduced recurrence (Theorem 4).

Table 1: Ratio of theoretical to simulated second moment $\mathbb{E}[e_t^2]$, both averaged over $t = 151$–200 of Experiment 1, on the manifolds with a steady state or diffusive growth.

	$v = 0.2$	$v = 0.3$	$v = 0.5$
$\mathbb{S}^2$ ($\kappa = 1$)	1.01×	1.03×	1.01×
$\mathbb{R}^2$ ($\kappa = 0$)	1.00×	1.00×	1.00×

while $0 < \sqrt{\kappa}\, Hv < \pi/2$, and over the step-size range studied here the nominal margin equation 8 evaluated at v_{eff} increases; the flat-space factor is unchanged, and under negative curvature $\beta_{\text{det}}(Hv)$ increases with H. This effective-step channel complements the Euclidean stability and data-collection mechanisms of Zhang et al. (2026).

Corollary 8 (Curvature-Dependent Chunking)**.** *At fixed decision/noise-injection counts and over the step-size range studied here, increasing chunk size enlarges the nominal sensitivity margin under positive curvature and strengthens geometric amplification under negative curvature.*

4 Experiments

Experiments 1–4 test the oracle recurrence on the sphere $\mathbb{S}^2$ ($\kappa = 1$), the Euclidean plane $\mathbb{R}^2$ ($\kappa = 0$), the hyperbolic plane $\mathbb{H}^2$ ($\kappa = -1$), and $\mathrm{SO}(3)$ ($\kappa = 1/4$); these spaces isolate the geometric mechanism and provide native rotation-kinematic grounding, with ε set directly as the injected one-step perturbation. Experiment 5 evaluates learned-policy diagnostics, with one-step error and policy response measured empirically. Implementations and full protocols are in Appendix B.

4.1 Curvature Sign Determines Three Error Regimes

Setup. We simulate oracle BC ($L = 0$) on $\mathbb{S}^2$, $\mathbb{R}^2$, and $\mathbb{H}^2$ with per-step noise $\varepsilon = 0.01$, horizon $T = 200$, and step sizes $v \in \{0.2, 0.3, 0.5\}$, averaging over 1500 Monte Carlo trials.

Results. Figure 2 shows the three predicted regimes clearly separated across all step sizes: $\mathbb{S}^2$ error saturates at a low bounded value, $\mathbb{R}^2$ error grows as $\sqrt{T}$, and $\mathbb{H}^2$ error exhibits local exponential amplification. The recurrence is quantitatively accurate in the canonical positive- and zero-curvature settings, with $\mathbb{E}[e^2]$ theory–simulation ratios of 1.00–1.03× on $\mathbb{S}^2$ and $\mathbb{R}^2$ (Table 1). At $v = 0.5$, the predicted $\mathbb{S}^2$ RMS floor is 0.029 rad, far below its geodesic diameter π, distinguishing the curvature-controlled floor from compactness alone. On $\mathbb{H}^2$, the RMS prediction agrees with simulation to approximately 2% over the full horizon at $v = 0.2$ and until $t \approx 100$ at $v = 0.3$, and to approximately 5% until $t \approx 50$ at $v = 0.5$ (Figure 2). At larger separations, the nonlinear rollout falls below the reduced exponential prediction while preserving the observed finite-horizon ordering. At $v = 0.5$ and $t = 150$, the $\mathbb{H}^2/\mathbb{S}^2$ RMS error ratio reaches 380×, illustrating the finite-horizon impact of geometric propagation. The same finite-horizon curvature ordering holds across random-walk, spiral, and straight-geodesic expert trajectories (Appendix C); deviations under the non-averaged straight-geodesic schedule are consistent with the angle-dependent mechanism in Remark 7.

4.2 Curvature-Dependent Geometric Effects of Chunking

Setup. We apply oracle BC ($L = 0$, $\varepsilon = 0.01$, $T = 150$) with base step $v = 0.1$ and chunk sizes $H \in \{1, 2, 3, 5, 8\}$, giving effective step sizes $v_{\text{eff}} \in \{0.1, 0.2, 0.3, 0.5, 0.8\}$. We use the fixed-event protocol with $T = 150$ decision/noise-injection events held constant across chunk sizes.

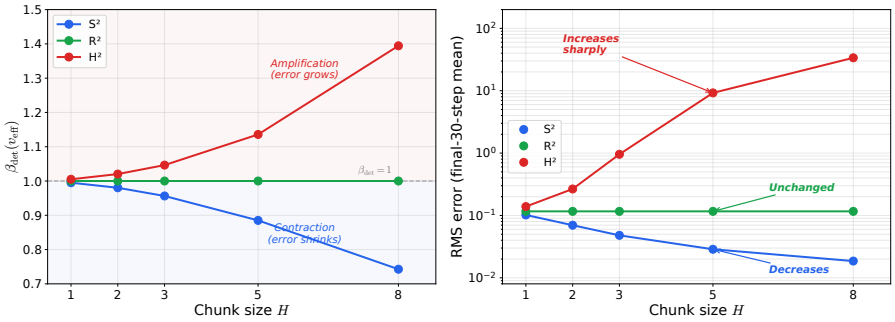


Figure 3: **Geometric effects of chunking at fixed event count.** *(a)* $\beta_{\det}(v_{\text{eff}})$ versus chunk size H with $v_{\text{eff}} = Hv$; the dashed line marks $\beta_{\det} = 1$. *(b)* Simulated RMS tracking error, averaged over the final 30 of $T = 150$ decision/noise-injection events. Large-step deviations on $\mathbb{H}^2$ are quantified in Appendix F.

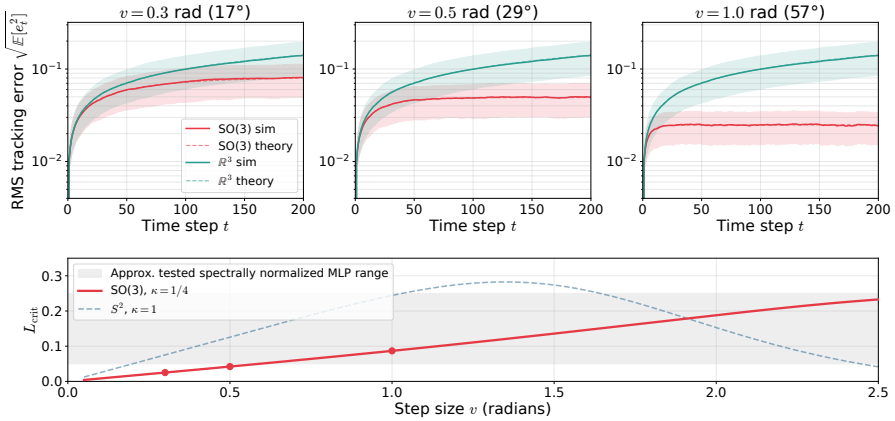


Figure 4: **SO(3) stabilization and nominal margin.** *Top (a):* Oracle BC ($L = 0$, expert direction resampled each step) on SO(3) ($\kappa = 1/4$, red) vs. $\mathbb{R}^3$ ($\kappa = 0$, green); curves show RMS error and bands show RMS $\pm\,\text{std}(e_t)$. *Bottom (b):* Nominal sensitivity margin L_{crit} vs. step size for SO(3) and $\mathbb{S}^2$; the shaded band (0.05–0.25) approximately spans the estimated sensitivities of the spectrally normalized MLPs evaluated in Section 4.5 and Appendix J.

Results. Figure 3 supports the fixed-event prediction of Corollary 8: larger effective steps strengthen contraction on $\mathbb{S}^2$ and amplification on $\mathbb{H}^2$, while leaving the geometric factor on $\mathbb{R}^2$ unchanged. Panel (b) shows the simulation consequence: on $\mathbb{S}^2$, the finite-horizon tracking error *decreases* monotonically with H; on $\mathbb{R}^2$, it remains constant; on $\mathbb{H}^2$, it *increases* by more than two orders of magnitude from $H = 1$ to $H = 8$ (0.14 to 34). By equation 8, chunking on $\mathbb{S}^2$ also enlarges the nominal sensitivity margin: L_{crit} grows from 0.025 at $H = 1$ to 0.200 at $H = 8$.

4.3 SO(3) Exhibits Predicted Positive-Curvature Stabilization

Setup. We compare oracle BC on SO(3) ($\kappa = 1/4$, quaternion representation) against a flat Euclidean baseline $\mathbb{R}^3$ ($\kappa = 0$) at step sizes $v \in \{0.3, 0.5, 1.0\}$ rad, with the expert direction resampled uniformly at every step as in Experiment 1 (the straight-geodesic variant is reported in Appendix C).

Results. In the oracle regime ($L = 0$), SO(3) ($\kappa = 1/4$) error saturates below $\mathbb{R}^3$ for all step sizes (Figure 4(a)), with terminal RMS within 2% of the recurrence (1.02×, 1.01×, 0.98× at $v = 0.3, 0.5, 1.0$); the gap grows with step size, consistent with stronger geometric contraction at larger v.

Positive curvature enlarges the nominal sensitivity margin, but the gain depends strongly on curvature magnitude (Figure 4(b)): $\mathbb{S}^2$ ($\kappa = 1$) enters the approximate sensitivity range observed for the tested spectrally normalized multilayer perceptrons (MLPs) already around $v \approx 0.3$, whereas SO(3) ($\kappa = 1/4$) does so only at substantially larger step sizes. For $v = 0.3$–0.5, the nominal SO(3) sensitivity margin is $L_{\text{crit}} \approx 0.025$–$0.042$; Section 4.5 compares the tested MLPs with this nominal margin.

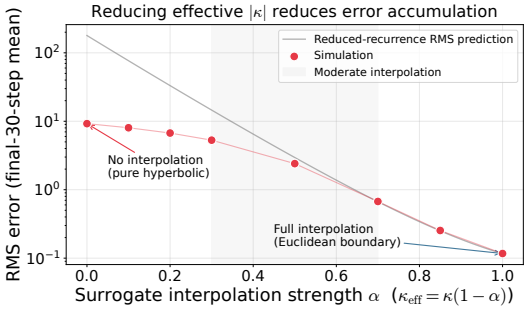


Figure 5: **Curvature surrogate sweep.** The surrogate interpolation strength α gives $\kappa_{\text{eff}} = \kappa(1-\alpha)$ with base $\kappa = -1$. Red: simulated RMS error averaged over the final 30 of $T = 150$ steps (800 trials). Gray: reduced-recurrence prediction with the same averaging window. Trajectories and phase diagrams are in Appendix E.

4.4 Geometric Lens on Hyperbolic RL Fragility

Setup. We sweep κ from -2 to $+1$ at fixed $v = 0.5$ and $\varepsilon = 0.01$ and measure finite-horizon RMS tracking error (Appendix E). We then vary the modeled curvature through $\kappa_{\text{eff}} = \kappa(1-\alpha)$, where $\alpha \in [0, 1]$ is the surrogate interpolation strength (Figure 5), moving the geometry toward the Euclidean boundary to probe its effect on error accumulation; κ_{eff} is a modeling surrogate rather than a specific regularizer.

Results. The phase diagram (Figure 1(b)) places the boundary $\beta_{\text{det}} = 1$ at $\kappa = 0$, and the error sweep (Appendix E) rises by orders of magnitude as curvature turns negative. Figure 5 shows that increasing α monotonically reduces error in this model, as predicted by the recurrence. At larger separations (small α, strong negative curvature), the nonlinear rollout falls below the reduced prediction (Appendix F). Thus, $\beta_{\text{det}} > 1$ suggests a complementary geometric mechanism relevant to fragility reported in hyperbolic RL and representation learning (Cetin et al., 2023; Mishne et al., 2023; Peng et al., 2022).

4.5 SO(3) Policies Exhibit a Sensitivity–Approximation Trade-off

Setup. We train four neural BC designs on $\mathrm{SO}(3)$, crossing action representation $\times$ chunk size: {six-dimensional (6D) continuous rotation representation (Zhou et al., 2019), geodesic tangent-vector representation} $\times$ $\{H{=}1, H{=}5\}$, across spectral-normalization output scales 0.02–1.0, plus unconstrained MLPs. We denote the designs by A (6D, $H{=}1$), B (geodesic, $H{=}1$), C (geodesic, $H{=}5$), and D (6D, $H{=}5$). For each model, we measure a finite-difference representation-space sensitivity estimate L, the representation error ε_{rep} (mean residual norm in the output representation), and the decoded next-state geodesic RMS error. For the rollout comparison in Figure 6(b), which uses the unconstrained model of each design, state evolution is closed loop while the previous-action input is teacher-forced from the expert sequence.

Results. Figure 6(a) compares the learned-policy L–ε_{rep} frontier with the nominal curvature-dependent margin: stronger sensitivity control lowers L while increasing representation error in the tested sweeps, and the configurations below L_{crit} are those with the largest ε_{rep}. For the $H{=}5$ designs, $L_{\text{crit}} = 0.087$, compared with estimated sensitivities $L = 3.2$–23.8 for the unconstrained models. In the spectrally normalized configurations of Appendix I, direct amplification is subcritical ($\beta_{\text{emp}} \approx 0.85$–$0.98$); values close to 1 retain most of the inherited squared separation from one step to the next, so measured amplification and one-step error carry complementary information.

Figure 6(b) provides a controlled comparison showing that matched one-step error need not imply matched rollout behavior. Output noise is calibrated through each decoder to match the noisy next-state geodesic RMS error within each chunk-size pair to 0.2%: approximately 0.180 for A/B and 0.215 for C/D (Appendix H). At $T = 150$, the terminal median tracking error is higher for B than for A (1.73 vs. 1.46) and for C than for D (1.31 vs. 1.14). In each pair, the higher-error policy also has the larger representation-space sensitivity estimate ($L \approx 24$ vs. 6.4 and 9.0 vs. 3.2). This supports the diagnostic view that one-step error and policy sensitivity should be evaluated jointly.

The trade-off persists across the tested width/depth variants (Appendix J). On $\mathbb{S}^2$, $\mathbb{R}^2$, and $\mathbb{H}^2$, action-clamp tests isolate the geometric step response, while matched-error self-fed rollouts, which use the

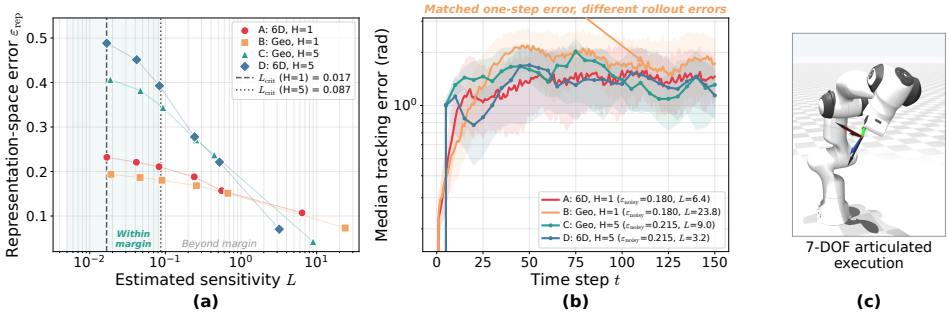


Figure 6: **Learned-policy diagnostics and articulated execution.** *(a)–(b)* Main $\mathrm{SO}(3)$ learned-policy results. *(a)* Estimated sensitivity L versus representation-space mean residual $\varepsilon_{\mathrm{rep}}$; vertical lines mark the nominal margins L_{crit} for $H{=}1$ and $H{=}5$. Unified geodesic errors are in Appendix I. *(b)* Median state-closed-loop tracking error (shaded: interquartile range) with teacher-forced previous-action inputs and noisy geodesic RMS error $\varepsilon_{\mathrm{noisy}}$ matched within each chunk-size pair; for $H{=}5$, only executed chunk endpoints are plotted. *(c)* Representative rendering of the separate complementary 7-DOF Panda execution-chain evaluation (arrows: end-effector frames); quantitative results are in Appendix M.

learner's own action history, preserve the finite-horizon curvature ordering (Appendix L). Single-body MuJoCo tests further relate measured amplification to closed-loop tracking (Appendix K).

A complementary seven-degree-of-freedom (7-DOF) Panda study, visualized in Figure 6(c), evaluates orientation commands executed through inverse kinematics, joint servoing, and MuJoCo robot dynamics (Appendix M). Within each of three training seeds, five cross-representation pairs are selected by matching one-step end-effector error after execution; all 15 retain the match on held-out calibration states (maximum relative gap 5.9%). In every pair, the 6D policy has lower measured β_{emp} and lower self-fed tracking error relative to the expert over the second half of the rollout; the corresponding 95% paired episode-bootstrap intervals for the 6D-minus-geodesic tracking-error difference lie below zero in all 15 comparisons. Together with the full 54-policy Panda sweep (Appendix M), these comparisons separate two roles: one-step accuracy captures the broad performance trend, while measured propagation response provides additional resolution once one-step error is closely matched.

5 Discussion and Conclusion

Scope and broader settings. Our formal analysis isolates geometric propagation in the direction-averaged local $L = 0$ setting on homogeneous constant-curvature spaces. Proposition 6 extends the scalar closure to conditional orientation bands, where the conditional error–action angle process governs structured policies. The piecewise-curvature and torus analyses in Appendices D.1–D.2 identify how curvature variation and initial-separation-rate terms ($J'(0) \equiv \nabla_t J(0) \neq 0$) enter beyond the homogeneous model; in the general smooth case, these terms enter through explicit Jacobi-field or comparison-geometry machinery (do Carmo, 1992, Ch. 10). For distinct observation and action spaces, Appendix D.3 decomposes observation, policy, and dynamics sensitivities through the closed-loop Jacobian. The learned-policy evaluations connect the geometric analysis to rotation kinematics and a simulated 7-DOF articulated execution chain. DAgger acts on a complementary axis by changing the training-state distribution through learner-visited expert labels, whereas our recurrence characterizes how residual local errors propagate geometrically.

Conclusion. Curvature is a first-class factor in the geometric propagation of accumulated tracking error in behavioral cloning, extending the view of geometry as a guiding principle from architecture design (Bronstein et al., 2021) to sequential prediction. Distinct inherited-error and fresh-error mechanisms yield the local trichotomy—a bounded squared-error floor, linear squared-error growth, or exponential amplification, with the Euclidean case at the boundary—while the conditional band quantifies departures from exact direction averaging. The fixed-event chunking analysis, the curvature surrogate for hyperbolic RL, and the $\mathrm{SO}(3)$ and articulated-arm evaluations motivate evaluating one-step accuracy jointly with geometric and policy response. More broadly, these results suggest a simple evaluation principle for sequential prediction on structured spaces: measure how errors propagate, not just how large they are.

AI USE STATEMENT

Generative AI tools were used to aid and polish the writing of this paper. The authors reviewed all AI-assisted text and take responsibility for the final content.

REPRODUCIBILITY STATEMENT

Complete proofs of all formal results are given in Appendix A. Appendix B specifies the manifold implementations, expert and learner protocols, hyperparameters (Table 2), randomization protocols, seed counts, and compute requirements. Appendices H–L provide supporting learned-policy measurements and evaluations, and Appendix M reports the articulated-arm execution-chain evaluation. Code, data files, and trained checkpoints needed to reproduce all experiments will be made publicly available upon acceptance.

REFERENCES

Gary Bécigneul and Octavian-Eugen Ganea. Riemannian adaptive optimization methods. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL https://openreview.net/forum?id=r1eiqi09K7.

Hadi Beik-Mohammadi, Søren Hauberg, Georgios Arvanitidis, Gerhard Neumann, and Leonel Rozo. Learning Riemannian manifolds for geodesic motion skills. In *Robotics: Science and Systems (RSS)*, 2021.

Silvère Bonnabel. Stochastic gradient descent on riemannian manifolds. *IEEE Trans. Autom. Control.*, 58(9):2217–2229, 2013. doi: 10.1109/TAC.2013.2254619. URL https://doi.org/10.1109/TAC.2013.2254619.

Max Braun, Noémie Jaquier, Leonel Rozo, and Tamim Asfour. Riemannian flow matching policy for robot motion learning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5144–5151, 2024.

Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Velickovic. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *CoRR*, abs/2104.13478, 2021. URL https://arxiv.org/abs/2104.13478.

Sylvain Calinon. Gaussians on riemannian manifolds: Applications for robot learning and adaptive control. *IEEE Robotics Autom. Mag.*, 27(2):33–45, 2020. doi: 10.1109/MRA.2020.2980548. URL https://doi.org/10.1109/MRA.2020.2980548.

Edoardo Cetin, Benjamin Paul Chamberlain, Michael M. Bronstein, and Jonathan J. Hunt. Hyperbolic deep reinforcement learning. In *The Eleventh International Conference on Learning Representations (ICLR)*, 2023.

Ines Chami, Zhitao Ying, Christopher Ré, and Jure Leskovec. Hyperbolic graph convolutional neural networks. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 4869–4880, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/0415740eaa4d9decbc8da001d3fd805f-Abstract.html.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *Robotics: Science and Systems (RSS)*, 2023.

Manfredo P. do Carmo. *Riemannian Geometry*. Birkhäuser, 1992.

Octavian-Eugen Ganea, Gary Bécigneul, and Thomas Hofmann. Hyperbolic neural networks. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pp. 5350–5360, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/dbab2adc8f9d078009ee3fa810bea142-Abstract.html.

Michael Laskey, Jonathan Lee, Roy Fox, Anca Dragan, and Ken Goldberg. DART: Noise injection for robust imitation learning. In *Proceedings of the 1st Annual Conference on Robot Learning (CoRL)*, pp. 143–156, 2017.

John M. Lee. *Introduction to Riemannian Manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2018. ISBN 978-3-319-91755-9. doi: 10.1007/978-3-319-91755-9.

Jake Levinson, Carlos Esteves, Kefan Chen, Noah Snavely, Angjoo Kanazawa, Afshin Rostamizadeh, and Ameesh Makadia. An analysis of SVD for deep rotation estimation. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/fec3392b0dc073244d38eba1feb8e6b7-Abstract.html.

Shaunak A. Mehta, Yusuf Umut Ciftci, Balamurugan Ramachandran, Somil Bansal, and Dylan P. Losey. Stable-BC: Controlling covariate shift with stable behavior cloning. *IEEE Robotics and Automation Letters*, 2025.

Gal Mishne, Zhengchao Wan, Yusu Wang, and Sheng Yang. The numerical stability of hyperbolic representation learning. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, pp. 24925–24949, 2023.

Maximillian Nickel and Douwe Kiela. Poincaré embeddings for learning hierarchical representations. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 6338–6347, 2017.

Maximillian Nickel and Douwe Kiela. Learning continuous hierarchies in the Lorentz model of hyperbolic geometry. In Jennifer Dy and Andreas Krause (eds.), *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 3779–3788. PMLR, 10–15 Jul 2018. URL https://proceedings.mlr.press/v80/nickel18a.html.

Yann Ollivier. Ricci curvature of Markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009. doi: 10.1016/j.jfa.2008.11.001.

Wei Peng, Tuomas Varanka, Abdelrahman Mostafa, Henglin Shi, and Guoying Zhao. Hyperbolic deep neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):10023–10044, 2022.

Valentin Peretroukhin, Matthew Giamou, W. Nicholas Greene, David M. Rosen, Jonathan Kelly, and Nicholas Roy. A smooth representation of belief over SO(3) for deep rotation learning with uncertainty. In Marc Toussaint, Antonio Bicchi, and Tucker Hermans (eds.), *Robotics: Science and Systems XVI, Virtual Event / Corvalis, Oregon, USA, July 12-16, 2020*, 2020. doi: 10.15607/RSS.2020.XVI.007. URL https://doi.org/10.15607/RSS.2020.XVI.007.

Harish Ravichandar, Athanasios S. Polydoros, Sonia Chernova, and Aude Billard. Recent advances in robot learning from demonstration. *Annu. Rev. Control. Robotics Auton. Syst.*, 3:297–330, 2020. doi: 10.1146/ANNUREV-CONTROL-100819-063206. URL https://doi.org/10.1146/annurev-control-100819-063206.

Stéphane Ross and J. Andrew Bagnell. Efficient reductions for imitation learning. In *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 661–668, 2010.

Stéphane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 627–635, 2011.

Max Simchowitz, Daniel Pfrommer, and Ali Jadbabaie. The pitfalls of imitation learning when actions are continuous. In *Proceedings of the 38th Conference on Learning Theory (COLT)*, pp. 5248–5351, 2025.

John W. Simpson-Porco and Francesco Bullo. Contraction theory on Riemannian manifolds. *Systems & Control Letters*, 65:74–80, 2014. doi: 10.1016/j.sysconle.2013.12.016.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5026–5033. IEEE, 2012. doi: 10.1109/IROS.2012.6386109.

Max-K. von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Communications on Pure and Applied Mathematics*, 58(7):923–940, 2005. doi: 10.1002/cpa.20060.

Martijn J.A. Zeestraten, Ioannis Havoutis, João Silvério, Sylvain Calinon, and Darwin G. Caldwell. An approach for imitation learning on Riemannian manifolds. *IEEE Robotics and Automation Letters*, 2(3):1240–1247, 2017.

Thomas T. Zhang, Daniel Pfrommer, Chaoyi Pan, Nikolai Matni, and Max Simchowitz. Action chunking and data augmentation yield exponential improvements in behavior cloning for continuous spaces. In *International Conference on Learning Representations (ICLR)*, 2026.

Tony Z. Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning fine-grained bimanual manipulation with low-cost hardware. In *Robotics: Science and Systems (RSS)*, 2023.

Yi Zhou, Connelly Barnes, Jingwan Lu, Jimei Yang, and Hao Li. On the continuity of rotation representations in neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5745–5753, 2019.

A Complete Proofs

This appendix provides detailed proofs for the formal $L = 0$ results and supporting derivations for the diagnostic quantities used in the main text; we follow the conventions of do Carmo (1992); Lee (2018) throughout. The main text contains compressed proofs; here we give the full Jacobi-field computations and supporting arguments. Jacobi fields are expressed in a parallel orthonormal frame along a unit-speed geodesic; primes denote covariant differentiation in that frame, so $J'(0) \equiv \nabla_t J(0)$.

A.1 Proof of Lemma 1 (Inherited-Error Propagation)

The main text provides the proof sketch; here we give additional detail on the Jacobi-field computation.

On a space of constant sectional curvature κ, the Jacobi equation $\nabla_t^2 J + R(\dot\gamma, J)\dot\gamma = 0$ reduces in the normal direction to:

$$J_\perp''(t) + \kappa\, J_\perp(t) = 0\,. \tag{9}$$

With initial conditions $J_\perp(0) = J_\perp^{(0)}$ (given) and $J_\perp'(0) = 0$ (parallel initial velocities), the solution is:

$$J_\perp(t) = \mathrm{c}_\kappa(t) \cdot J_\perp^{(0)} = \begin{cases} \cos(\sqrt{\kappa}\, t)\, J_\perp^{(0)} & \kappa > 0 \\ J_\perp^{(0)} & \kappa = 0 \\ \cosh(\sqrt{|\kappa|}\, t)\, J_\perp^{(0)} & \kappa < 0 \end{cases} \tag{10}$$

The tangential Jacobi field satisfies $J_\parallel''(t) = 0$ (no curvature in the direction of motion), so $J_\parallel(t) = J_\parallel(0) + t\, J_\parallel'(0) = J_\parallel(0)$ since $J_\parallel'(0) = 0$.

The expectation over the uniform direction a uses the fact that for a fixed unit vector $J(0)/\delta$ and a uniformly random unit vector a/v in T_pM, the expected squared projection is:

$$\mathbb{E}\left[\frac{|\langle J(0), a\rangle|^2}{\|J(0)\|^2\|a\|^2}\right] = \frac{1}{d}\,, \qquad \mathbb{E}\left[\frac{\|J(0) - \langle J(0), a/\|a\|\rangle\, a/\|a\|\|^2}{\|J(0)\|^2}\right] = \frac{d-1}{d}\,. \tag{11}$$

This is a standard result from the geometry of projections onto random subspaces.

For small δ, the endpoint distance is given by the Jacobi-field norm up to higher-order terms:

$$d(\exp_p(a),\, \exp_q(\tilde{a}))^2 = \|J_\parallel(v)\|^2 + \|J_\perp(v)\|^2 + O(\delta^3)\,. \tag{12}$$

Substituting $J_\parallel(v) = J_\parallel(0)$ and $J_\perp(v) = \mathrm{c}_\kappa(v)\, J_\perp(0)$ gives, conditioned on the angle between $J(0)$ and a,

$$d(\exp_p(a),\, \exp_q(\tilde{a}))^2 = \|J_\parallel(0)\|^2 + \mathrm{c}_\kappa^2(v)\, \|J_\perp(0)\|^2 + O(\delta^3)\,. \tag{13}$$

Taking expectation over the uniform direction a and applying the projection identities yields

$$\mathbb{E}_a[d(\exp_p(a),\, \exp_q(\tilde{a}))^2] = \left(\frac{1}{d} + \frac{d-1}{d}\, \mathrm{c}_\kappa^2(v)\right) \delta^2 + O(\delta^3)\,, \tag{14}$$

which is Lemma 1.

A.2 Proof of Lemma 2 (Noise Amplification)

The differential of the exponential map $(d\exp_p)_a : T_a(T_pM) \cong T_pM \to T_{\exp_p(a)}M$ is characterized by Jacobi fields with different initial conditions from Lemma 1. For a perturbation ξ at the base of the tangent vector a:

The Jacobi field has initial conditions $J(0) = 0$ and $J'(0) = \xi/v$. In the normal direction, the solution at $t = v$ gives:

$$\|J_\perp(v)\| = \mathrm{s}_\kappa(v)\, \|\xi_\perp\|\,, \tag{15}$$

where $\mathrm{s}_\kappa(v)$ is defined as in equation equation 4: $\sin(\sqrt{\kappa}\, v)/(\sqrt{\kappa}\, v)$ for $\kappa > 0$, 1 for $\kappa = 0$, and $\sinh(\sqrt{|\kappa|}\, v)/(\sqrt{|\kappa|}\, v)$ for $\kappa < 0$. This confirms that the transverse amplification factor is $\mathrm{s}_\kappa(v)$.

In the radial direction, $J_\parallel(t) = t\, \xi_\parallel/v$, so $\|J_\parallel(v)\| = \|\xi_\parallel\|$ (unit amplification).

Combining radial and transverse components, the squared one-step noise displacement is

$$d(\exp_p(a),\, \exp_p(a+\xi))^2 = \|\xi_\parallel\|^2 + \mathrm{s}_\kappa^2(v)\, \|\xi_\perp\|^2 + O(\|\xi\|^3)\,, \tag{16}$$

for every realization of ξ, with a remainder constant that is uniform on the bounded support of the noise model. Writing $\xi = \varepsilon z$ with z zero-mean, isotropic, $\mathbb{E}[\|z\|^2] = 1$, and bounded support, the radial–transverse decomposition gives

$$\mathbb{E}[\|\xi_\parallel\|^2] = \frac{\varepsilon^2}{d}\,, \qquad \mathbb{E}[\|\xi_\perp\|^2] = \frac{d-1}{d}\varepsilon^2\,. \tag{17}$$

and $\mathbb{E}[\|\xi\|^3] = \varepsilon^3\, \mathbb{E}[\|z\|^3] = O(\varepsilon^3)$. Therefore

$$\mathbb{E}_\xi[d(\exp_p(a),\, \exp_p(a+\xi))^2] = \left(\frac{1}{d} + \frac{d-1}{d}\, \mathrm{s}_\kappa^2(v)\right) \varepsilon^2 + O(\varepsilon^3)\,, \tag{18}$$

which is Lemma 2.

A.3 Cross-Term Estimate

We isolate the cross-term estimate used in the proof of Theorem 4.

Proposition 9 (Cross-term estimate)**.** *Let $p_0 \in M$ and $a \in T_{p_0}M$ with $\|a\| = v$ be fixed, and let $\xi = \varepsilon z \in T_{p_0}M$ follow the noise model of Section 3.1. Define $q_1 = \exp_{p_0}(a)$ and $q_2 = \exp_{p_0}(a+\xi)$, and let q_3 be a point fixed before ξ is drawn, with $d(q_1, q_3) = \delta$. Then, for δ and ε sufficiently small,*

$$\mathbb{E}_\xi[d(q_3, q_2)^2] = \delta^2 + \sigma_{\text{noise}}^2\, \varepsilon^2 + O(\delta\varepsilon^2 + \varepsilon^3)\,. \tag{19}$$

Proof. Work in normal coordinates centered at q_1. Let $u = \log_{q_1}(q_3)$, which is fixed with $\|u\| = \delta$, and $w = \log_{q_1}(q_2)$, which is random. For fixed u, the squared distance $F(w) := d(\exp_{q_1}(u), \exp_{q_1}(w))^2$ is smooth in w near 0 with $F(0) = \delta^2$, and by the standard normal-coordinate expansion $\nabla F(0) = -2u$ and $\nabla^2 F(0) = 2I + O(\delta^2)$, hence

$$F(w) = \delta^2 - 2\langle u, w\rangle + \|w\|^2 + O\big(\delta^2\|w\|^2 + \|w\|^3\big), \tag{20}$$

with a constant that is uniform on a normal neighborhood. Since $\delta \le 1$, the remainder is $O(\delta\|w\|^2 + \|w\|^3)$. Taking expectations term by term,

$$\mathbb{E}_\xi[F(w)] = \delta^2 - 2\langle u, \mathbb{E}_\xi[w]\rangle + \mathbb{E}_\xi\|w\|^2 + O\big(\delta\,\mathbb{E}_\xi\|w\|^2 + \mathbb{E}_\xi\|w\|^3\big). \tag{21}$$

Since $w = (d\exp_{p_0})_a\,\xi + O(\|\xi\|^2)$ with a uniform constant on the bounded support of z, and $\mathbb{E}[\xi] = 0$, we have $\mathbb{E}_\xi[w] = O(\varepsilon^2)$, so the linear cross term contributes $O(\delta\varepsilon^2)$. Lemma 2 gives $\mathbb{E}_\xi\|w\|^2 = \sigma^2_{\text{noise}}\varepsilon^2 + O(\varepsilon^3)$, and $\mathbb{E}_\xi\|w\|^3 = O(\varepsilon^3)$. Collecting terms yields the claim. □

A.4 Proof of Theorem 4 (Local Squared-Error Recurrence)

Proof. Let $\mathcal{H}_t$ denote the pre-noise history together with the realized expert action a_t. Conditioned on $\mathcal{H}_t$, the current states, the transported reference action $\Gamma_{s_t^*\to\hat{s}_t}a_t$, and the local differential of the exponential map are fixed. Define the noise-free learner position $\hat{s}^{(0)}_{t+1} = \exp_{\hat{s}_t}(\Gamma_{s_t^*\to\hat{s}_t}a_t)$ and the actual learner position $\hat{s}_{t+1} = \exp_{\hat{s}_t}(\Gamma_{s_t^*\to\hat{s}_t}a_t + \xi_t)$.

Step 1 (Noise expectation at fixed configuration). Apply Proposition 9 with $p_0 = \hat{s}_t$, $a = \Gamma_{s_t^*\to\hat{s}_t}a_t$, $q_1 = \hat{s}^{(0)}_{t+1}$, $q_2 = \hat{s}_{t+1}$, and $q_3 = s^*_{t+1}$, with p_0, a, q_1, and q_3 fixed conditional on $\mathcal{H}_t$, while q_2 depends on the fresh noise ξ_t. Writing $d_{\text{det}} := d(s^*_{t+1}, \hat{s}^{(0)}_{t+1})$,

$$\mathbb{E}[e^2_{t+1} \mid \mathcal{H}_t] = d^2_{\text{det}} + \sigma^2_{\text{noise}}\,\varepsilon^2 + O(d_{\text{det}}\,\varepsilon^2 + \varepsilon^3)\,. \tag{22}$$

Step 2 (Direction averaging). In the direction-averaged local model, conditioned on e_t and the pre-action history, the direction of a_t is uniform on the sphere in $T_{s_t^*}M$. By Lemma 1, $\mathbb{E}[d^2_{\text{det}} \mid e_t] = \beta_{\text{det}}\,e_t^2 + O(e_t^3)$, and the pointwise bound $d_{\text{det}} = O(e_t)$ from the same Jacobi-field computation gives $\mathbb{E}[d_{\text{det}}\,\varepsilon^2 \mid e_t] = O(e_t\varepsilon^2)$.

Step 3 (Tower property). Taking the conditional expectation of Step 1 given e_t and substituting Step 2,

$$\mathbb{E}[e^2_{t+1} \mid e_t] = \beta_{\text{det}}\,e_t^2 + \sigma^2_{\text{noise}}\,\varepsilon^2 + O(e_t^3 + e_t\varepsilon^2 + \varepsilon^3)\,, \tag{23}$$

which is equation 5. □

A.5 Motivation for the $L > 0$ Diagnostic Proxy

For a learned policy $\hat{\pi}$ with sensitivity $L > 0$, decompose the executed action at the learner state relative to the transported expert action:

$$\hat{\pi}(\hat{s}_t) - \Gamma_{s_t^*\to\hat{s}_t}\,\pi^*(s_t^*) \;=\; \big[\hat{\pi}(\hat{s}_t) - \Gamma_{s_t^*\to\hat{s}_t}\,\hat{\pi}(s_t^*)\big] \;+\; \Gamma_{s_t^*\to\hat{s}_t}\big[\hat{\pi}(s_t^*) - \pi^*(s_t^*)\big]\,. \tag{24}$$

The first term is the policy's response to the state deviation and is bounded in norm by $L\,e_t$; the second is the learner's approximation residual at the expert state; in the oracle model, the corresponding fresh-error channel is represented by ξ_t. Inherited separation and policy-induced action variation pass through different local geometric maps, and their combined response depends on the radial–transverse decomposition and on the alignment of the action variation with the transported separation. A single trajectory-independent scalar therefore does not capture the exact $L > 0$ amplification without additional assumptions.

For the learned-policy diagnostics in the main text, we therefore adopt the closed-form proxy

$$\beta_{\text{eff}} \;=\; \beta_{\text{det}}(v)\,(1 + Lv)^2\,, \tag{25}$$

which preserves the curvature-controlled factor $\beta_{\text{det}}(v)$ while penalizing policy sensitivity through the dimensionless quantity Lv. The $(1+Lv)^2$ factor is a diagnostic proxy, not a certified bound; the empirical one-step amplification β_{emp} is measured directly in Appendix I.

The proxy condition $\beta_{\text{eff}} < 1$ defines the nominal sensitivity margin

$$L_{\text{crit}} \;=\; \frac{1/\sqrt{\beta_{\text{det}}} - 1}{v}\,, \tag{26}$$

used throughout the main text to interpret the sensitivity–approximation trade-offs of learned policies on $\mathrm{SO}(3)$.

A.6 Proof of Negative Curvature Exponential Growth

Proposition 10 (Exponential Growth in the Local Model for $\kappa < 0$)**.** *On a manifold with $\kappa < 0$, $\beta_{\mathrm{det}}(v) > 1$ for all $v > 0$, and, for the reduced leading-order recurrence $x_{t+1} = \beta_{\mathrm{det}}\, x_t + \sigma^2_{\mathrm{noise}}\varepsilon^2$ with $L = 0$ and $x_0 = 0$:*

$$x_T = \frac{\sigma^2_{\mathrm{noise}}\,\varepsilon^2\,(\beta^T_{\mathrm{det}} - 1)}{\beta_{\mathrm{det}} - 1}\,. \tag{27}$$

Proof. For $\kappa < 0$: $\mathrm{c}_\kappa(v) = \cosh(\sqrt{|\kappa|}\,v) > 1$ for all $v > 0$, so $\beta_{\mathrm{det}} = 1/d + (d-1)/d\,\cosh^2(\sqrt{|\kappa|}v) > 1/d + (d-1)/d = 1$. The reduced recurrence $x_{t+1} = \beta_{\mathrm{det}}\, x_t + \sigma^2_{\mathrm{noise}}\varepsilon^2$ with $x_0 = 0$ gives:

$$x_T = \sigma^2_{\mathrm{noise}}\varepsilon^2 \sum_{k=0}^{T-1} \beta^k_{\mathrm{det}} = \sigma^2_{\mathrm{noise}}\varepsilon^2\,\frac{\beta^T_{\mathrm{det}} - 1}{\beta_{\mathrm{det}} - 1}\,, \tag{28}$$

Thus the reduced recurrence grows exponentially for nonzero injected error, with growth factor $\beta_{\mathrm{det}} > 1$. □

A.7 Proof of Proposition 6 (Conditional Orientation Band)

We work under the constant-curvature local $L = 0$ setup and the fresh-noise assumptions of Theorem 4, and drop the same higher-order remainder, so that all statements concern the reduced leading-order model.

Step 1 (Componentwise factor). By the Jacobi-field computation in the proof of Lemma 1, before direction averaging the inherited squared error after one step is

$$\beta(\theta_t)\, e_t^2, \qquad \beta(\theta_t) = \cos^2\theta_t + \mathrm{c}^2_\kappa(v)\sin^2\theta_t = \mathrm{c}^2_\kappa(v) + \big[1 - \mathrm{c}^2_\kappa(v)\big]\cos^2\theta_t\,, \tag{29}$$

where θ_t is the angle between the separation direction $J(0)$ and the action direction. Averaging $\cos^2\theta_t$ uniformly gives $1/d$ and recovers $\beta_{\mathrm{det}} = \mathrm{c}^2_\kappa(v) + [1 - \mathrm{c}^2_\kappa(v)]/d$, which is equation 2.

Step 2 (Conditional bound). Since $\beta(\theta_t)$ is affine in $\cos^2\theta_t$ with slope $1 - \mathrm{c}^2_\kappa(v)$, the band $|\mathbb{E}[\cos^2\theta_t \mid \mathcal{G}_t] - 1/d| \le \eta$ implies

$$\big|\mathbb{E}[\beta(\theta_t) \mid \mathcal{G}_t] - \beta_{\mathrm{det}}\big| \le |1 - \mathrm{c}^2_\kappa(v)|\,\eta = \Delta_\eta\,, \tag{30}$$

so $\beta_- \le \mathbb{E}[\beta(\theta_t) \mid \mathcal{G}_t] \le \beta_+$, where the truncation $\beta_- = \max\{0, \beta_{\mathrm{det}} - \Delta_\eta\}$ uses $\beta(\theta_t) \ge 0$. No separate transverse condition is needed because $\sin^2\theta_t = 1 - \cos^2\theta_t$.

Step 3 (Fresh-error term and cross term). Conditioned on the realized action and the pre-noise history, the perturbation ξ_t is zero-mean and isotropic, so by Lemma 2 and Proposition 9 the fresh-error contribution is $\sigma^2_{\mathrm{noise}}\varepsilon^2$ and the cross term vanishes at leading order, exactly as in Theorem 4; the band affects only the inherited-error factor.

Step 4 (Tower property and induction). Since e_t is $\mathcal{G}_t$-measurable, the reduced leading-order model gives

$$\mathbb{E}[e^2_{t+1} \mid \mathcal{G}_t] = \mathbb{E}[\beta(\theta_t) \mid \mathcal{G}_t]\, e_t^2 + \sigma^2_{\mathrm{noise}}\varepsilon^2 \in \big[\beta_- e_t^2 + \sigma^2_{\mathrm{noise}}\varepsilon^2,\ \beta_+ e_t^2 + \sigma^2_{\mathrm{noise}}\varepsilon^2\big]\,. \tag{31}$$

Taking expectations, $\beta_- m_t + \sigma^2_{\mathrm{noise}}\varepsilon^2 \le m_{t+1} \le \beta_+ m_t + \sigma^2_{\mathrm{noise}}\varepsilon^2$ with $m_t = \mathbb{E}[e_t^2]$. The affine maps $x \mapsto \beta_\pm x + \sigma^2_{\mathrm{noise}}\varepsilon^2$ are nondecreasing because $\beta_\pm \ge 0$, so induction from $x_0^\pm = m_0$ yields $x_t^- \le m_t \le x_t^+$ for all t. No independence of θ_t across time is used: the band is imposed conditionally on $\mathcal{G}_t$, which may encode correlation with the error magnitude, task phase, or other history variables. □

Consequences. If $\beta_+ < 1$, the upper recurrence converges to $\sigma^2_{\mathrm{noise}}\varepsilon^2/(1 - \beta_+)$, so the inherited-error channel remains contractive and m_t bounded. If $\beta_- > 1$, the lower recurrence is supercritical. In the flat case $\mathrm{c}^2_\kappa(v) = 1$, so $\Delta_\eta = 0$ and the sandwich collapses to Theorem 4. If $\beta_- \le 1 \le \beta_+$, the interval $[\beta_-, \beta_+]$ makes explicit the remaining orientation-induced uncertainty in regime classification. Without scalar closure, the natural leading-order model is the nonautonomous recurrence with factor $\beta(\theta_t)$ (Remark 7): products of realized factors govern inherited error, while each fresh-error injection is weighted by the product of the factors after it. Empirical support for the componentwise factor in learned policies is reported in Appendix L.

A.8 Derivation of the Action-Chunking Prediction

At fixed decision/noise-injection counts, chunk size H corresponds to effective step size $v_{\text{eff}} = Hv$ in the reduced recurrence. Substituting into β_{det}:

For $\kappa > 0$: $\beta_{\text{det}}(Hv) = 1/d + (d-1)/d \cos^2(\sqrt{\kappa}\, Hv)$, which decreases with H while $0 < \sqrt{\kappa}\, Hv < \pi/2$. The nominal sensitivity margin $L_{\text{crit}} = (1/\sqrt{\beta_{\text{det}}} - 1)/v_{\text{eff}}$ combines this decrease with the growth of the denominator; over the step-size range used in Section 4.2 ($v_{\text{eff}} \leq 0.8$ on $\mathbb{S}^2$), the numerator dominates and L_{crit} increases with H.

For $\kappa = 0$: $\beta_{\text{det}} = 1$ for every v, so the reduced recurrence is unchanged under this fixed-event protocol.

For $\kappa < 0$: $\beta_{\text{det}}(Hv) = 1/d + (d-1)/d \cosh^2(\sqrt{|\kappa|}\, Hv)$. Since $\cosh^2$ is strictly increasing, larger H increases the reduced-recurrence amplification factor.

B Experimental Details

B.1 Manifold Implementations

Sphere $\mathbb{S}^2$ ($\kappa = 1$, $d = 2$). We use the unit sphere in $\mathbb{R}^3$. The exponential map is $\exp_p(v) = \cos(\|v\|)\, p + \sin(\|v\|)\, v/\|v\|$. The logarithmic map is $\log_p(q) = \theta\,(q - \cos\theta\, p)/\sin\theta$ where $\theta = \arccos(\langle p, q\rangle)$. Parallel transport of $u \in T_p\mathbb{S}^2$ to $T_q\mathbb{S}^2$ along the minimizing geodesic uses the closed-form formula

$$\Gamma_{p\to q} u \;=\; u \;-\; \frac{\langle u,\, q\rangle}{1 + \langle p,\, q\rangle}\,(p + q)\,. \tag{32}$$

Euclidean plane $\mathbb{R}^2$ ($\kappa = 0$, $d = 2$). $\exp_p(v) = p + v$, $\log_p(q) = q - p$, parallel transport is the identity.

Hyperbolic plane $\mathbb{H}^2$ ($\kappa = -1$, $d = 2$). We use the hyperboloid model $\{x \in \mathbb{R}^{2,1} : x_0^2 - x_1^2 - x_2^2 = 1,\ x_0 > 0\}$ with the Minkowski bilinear form $\langle u, v\rangle_{\text{M}} = u_0 v_0 - u_1 v_1 - u_2 v_2$. The exponential map is $\exp_p(v) = \cosh(\|v\|)\, p + \sinh(\|v\|)\, v/\|v\|$, where $\|v\| = \sqrt{-\langle v, v\rangle_{\text{M}}}$ for tangent vectors v, which satisfy $\langle v, v\rangle_{\text{M}} < 0$ under this convention. The logarithmic map uses $\theta = \operatorname{arccosh}(\langle p, q\rangle_{\text{M}})$, giving $\log_p(q) = \theta\,(q - \langle p, q\rangle_{\text{M}}\, p)/\sinh\theta$. Parallel transport of $u \in T_p\mathbb{H}^2$ to $T_q\mathbb{H}^2$ along the minimizing geodesic uses

$$\Gamma_{p\to q} u \;=\; u \;-\; \frac{\langle u,\, q\rangle_{\text{M}}}{1 + \langle p,\, q\rangle_{\text{M}}}\,(p + q)\,, \tag{33}$$

following Nickel & Kiela (2018); Ganea et al. (2018). Tangent projection is $P_p v = v - \langle v, p\rangle_{\text{M}}\, p$ (since $\langle p, p\rangle_{\text{M}} = 1$), and random tangent directions are drawn uniformly in the Minkowski-orthonormal polar frame $e_{\text{rad}} = (\sinh r, \cosh r\cos\phi, \cosh r \sin\phi)$, $e_{\text{ang}} = (0, -\sin\phi, \cos\phi)$ of $T_p\mathbb{H}^2$, where $p = (\cosh r, \sinh r\cos\phi, \sinh r \sin\phi)$. To maintain numerical precision at large radii, $\mathbb{H}^2$ rollouts use an expert-centred frame obtained by a Lorentz isometry after each step, so that the tracking error is $\operatorname{arccosh}(x_0)$ of the learner's relative position; tangent-vector operations are performed in the learner's local Minkowski-orthonormal polar frame. By isometry invariance, this representation preserves the dynamics, the isotropic noise model, and parallel transport. On $\mathbb{S}^2$, points are renormalized after each exponential update and chained parallel transports are re-projected onto the tangent space. For $\mathbb{H}^2$ (and $\mathbb{S}^2$) at general curvature κ (Experiment 4), the physical triple (κ, v, ε) is simulated on the unit-curvature space with $v' = \sqrt{|\kappa|}\, v$, $\varepsilon' = \sqrt{|\kappa|}\, \varepsilon$, and distances divided by $\sqrt{|\kappa|}$.

$\mathrm{SO}(3)$ ($\kappa = 1/4$, $d = 3$). We represent rotations by unit quaternions $q \in \mathbb{S}^3 \subset \mathbb{R}^4$ with the identification $q \sim -q$. We represent an $\mathrm{SO}(3)$ tangent increment by a q-orthogonal 4-vector v ($\langle q, v\rangle = 0$), scaled so that $\|v\|$ equals the corresponding rotation angle on $\mathrm{SO}(3)$. The exponential map is

$$\exp_q(v) \;=\; \cos\!\Big(\tfrac{\|v\|}{2}\Big)\, q \;+\; \sin\!\Big(\tfrac{\|v\|}{2}\Big)\, \frac{v}{\|v\|}\,, \tag{34}$$

where the factor of $1/2$ arises from the double cover $\mathbb{S}^3 \to \mathrm{SO}(3)$, so that $\|v\|$ is measured as a rotation angle in $\mathrm{SO}(3)$. Equivalently, our stored tangent coordinate is scaled so that its Euclidean

norm equals the induced $\mathrm{SO}(3)$ rotation angle; its corresponding $\mathbb{S}^3$ geodesic length is $\|v\|/2$. The geodesic distance is $d(q_1, q_2) = 2\arccos(|\langle q_1, q_2\rangle|)$. The logarithmic map first chooses the representative of q_2 with $\langle q_1, q_2\rangle \geq 0$, then computes

$$\log_{q_1}(q_2) \;=\; \frac{2\arccos(\langle q_1, q_2\rangle)}{\sin(\arccos(\langle q_1, q_2\rangle))}\left(q_2 - \langle q_1, q_2\rangle\, q_1\right). \tag{35}$$

The hemisphere identification $q \sim -q$ is handled *only* in the distance and logarithmic-map computations (not in the exponential-map update) to avoid trajectory discontinuities.

In all exponential and logarithmic maps above (sphere, hyperboloid, $\mathrm{SO}(3)$), the standard first-order limiting form is used when the norm or angle is near zero (i.e., $\exp_p(0) = p$ and $\log_p(p) = 0$, with the corresponding singular factors replaced by their leading Taylor coefficients).

B.2 Hyperparameters

For Experiments 1–4, the expert policy at each step draws a uniformly random tangent direction in T_sM with fixed norm v (or $v_{\text{eff}} = Hv$ for Experiment 2); the learner receives this expert action via parallel transport and adds isotropic noise of magnitude ε in its own tangent space. This non-goal-seeking protocol provides a controlled setting for evaluating the inherited-error propagation factor β_{det} without goal-directed feedback (Appendix C shows that the finite-horizon curvature ordering persists under spiral and straight-geodesic schedules).

Table 2: Hyperparameters for main-text Experiments 1–5. Experiment 5 reports learned one-step error metrics and calibrated rollout noise separately below.

	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5
Per-step noise ε	0.01	0.01	0.01	0.01	—
Step size v	0.2/0.3/0.5	0.1 (base)	0.3/0.5/1.0	0.5	0.2
Horizon T (rollout)	200	150	200	150	150
Training trajectory length	—	—	—	—	200
Monte Carlo trials	1500	1000	1500	800	—
Chunk size H	1	1/2/3/5/8	1	1	1/5
Neural network (Exp 5)					
Hidden size	—	—	—	—	128
Layers	—	—	—	—	3
Activation	—	—	—	—	ReLU
Epochs	—	—	—	—	200
Learning rate	—	—	—	—	10^{-3} (cosine annealed)
Batch size	—	—	—	—	256
Optimizer	—	—	—	—	Adam
Seeds per config	—	—	—	—	1

B.3 Exp 4: Curvature Sweep and Surrogate Interpolation Details

Experiment 4 has three components, all run at fixed $v = 0.5$, $\varepsilon = 0.01$, $T = 150$, and 800 Monte Carlo trials per simulated point. To realize a constant-curvature space at arbitrary κ, we use the standard rescaling: the physical triple (κ, v, ε) is simulated on $\mathbb{S}^2$ ($\kappa = +1$) for $\kappa > 0$ and on $\mathbb{H}^2$ ($\kappa = -1$) for $\kappa < 0$ with $v' = \sqrt{|\kappa|}\, v$ and $\varepsilon' = \sqrt{|\kappa|}\, \varepsilon$, and simulated distances are divided by $\sqrt{|\kappa|}$; $\kappa = 0$ is simulated on $\mathbb{R}^2$.

Part A: Phase diagram simulation (Figure 1b in main text; empirical version in Figure 11, Appendix E). Figure 1(b) shows the conceptual phase diagram as the theoretical curve $\beta_{\text{det}}(\kappa)$ from Theorem 4, with $\kappa = 0$ as the diffusive Euclidean boundary between the positive-curvature bounded and negative-curvature amplified regimes. The empirical version (Figure 11) plots the long-horizon RMS error, averaged over the final 30 steps, versus κ, with simulation points overlaid on the theoretical prediction computed with the same averaging window. The continuous theory curve sweeps κ over 200 points in $[-2, +1]$. Simulation points are evaluated at $\kappa \in \{-2, -1.5, -1, -0.5, -0.25, 0, +0.25, +0.5, +1\}$ (nine values); the reported error at each point is the temporal mean of the RMS error $\sqrt{\mathbb{E}[e_t^2]}$ over the final 30 time steps of a $T = 150$ rollout.

Part B: Surrogate interpolation sweep (Figure 5). The base curvature is $\kappa_{\text{base}} = -1$, and the surrogate interpolation strength $\alpha \in \{0, 0.1, 0.2, 0.3, 0.5, 0.7, 0.85, 1.0\}$ is mapped to effective curvature via $\kappa_{\text{eff}} = \kappa_{\text{base}}(1-\alpha)$, so $\alpha = 0$ is pure $\mathbb{H}^2$ and $\alpha = 1$ is the Euclidean limit. The continuous theory curve uses 100 values of α uniformly in $[0, 1]$.

Part C: Trajectory subsets (Figure 10 in Appendix E). Error trajectories are reported for $\alpha \in \{0, 0.3, 0.7, 1.0\}$ to visualize that the late-horizon averages reported in Parts A–B are reached within the simulated horizon.

B.4 Exp 5: Sensitivity–Approximation Trade-off Details

Designs. Four designs in a 2×2 matrix: {six-dimensional (6D) continuous rotation representation, geodesic tangent-vector representation} $\times$ $\{H = 1, H = 5\}$. The 6D representation uses the first two columns of the rotation matrix (Zhou et al., 2019); the geodesic representation uses the logarithmic-map tangent vector in quaternion representation space (3 intrinsic degrees of freedom, embedded in $\mathbb{R}^4$ orthogonal to the current quaternion).

Expert trajectories. 300 training trajectories of length 200, generated by a non-goal-seeking smoothly varying expert policy on $\mathrm{SO}(3)$ with base step size $v = 0.2$ rad. Concretely, at each step the expert action is the projection onto $T_q\mathbb{S}^3$ of a slowly rotating ambient direction $(0, \cos\phi_t, \sin\phi_t, 0.3\sin(0.1t))$ with ϕ_t drifting at 0.03 rad/step, then renormalized to length v. The expert produces smoothly varying orientation trajectories without a fixed target, keeping the controlled comparison free of goal-directed feedback.

Training objective and decoding. All models are trained with mean-squared error in their output representation, using the Adam optimizer (learning rate 10^{-3}, cosine annealing) with batch size 256 over 200 epochs. For both $H = 1$ and $H = 5$, the supervised target corresponds to the aggregate transition from q_t to q_{t+H}: for $H = 1$ the next-step transition, and for $H = 5$ the actual endpoint displacement $\log_{q_t}(q_{t+H})$ aggregating five base-step transitions. The diagnostic of Section 3.4 uses Hv as the nominal effective step length for $H = 5$. The geodesic design represents this aggregate target as the tangent vector $\log_{q_t}(q_{t+H}) \in T_{q_t}\mathbb{S}^3$; rollout updates use $q_{t+H} \leftarrow \exp_{q_t}(\hat{v})$ with the predicted tangent $\hat{v}$. The 6D design represents states by the first two columns of the rotation matrix (Zhou et al., 2019) and predicts the 6D *difference* between the chunk-end and chunk-start representations; rollout updates apply this predicted difference in 6D space and then project back to a rotation matrix via Gram–Schmidt orthogonalization, which is then converted to a quaternion for geodesic evaluation. We report representation-space residuals as ε_{rep} and decoded next-state geodesic RMS errors as ε_{geo}; rollout tracking error is evaluated using the geodesic distance on $\mathrm{SO}(3)$ regardless of the training representation.

Sensitivity control. Spectral normalization is applied to all layers, with an output scaling factor $s \in \{0.02, 0.05, 0.1, 0.3, 1.0\}$. An additional unconstrained MLP (no spectral normalization) is included for comparison.

Sensitivity estimation. We estimate L as a representation-space finite-difference sensitivity proxy: the 95th percentile of $\|f(x+\delta) - f(x)\|/\|\delta\|$ over 3000 random perturbations with $\|\delta\| = 0.01$ in the network's input vector. This estimate is computed in each design's own representation space and serves as a representation-space sensitivity proxy for the manifold-level sensitivity defined in Section 3.1; comparisons across designs should be interpreted in this representation-space sense, and the geodesic state-step response is evaluated separately through β_{emp} (Appendix I).

Closed-loop rollout protocol (Figure 6b). Rollouts use teacher-forced inputs (each step the policy receives the previous *expert* action as part of the input, not the policy's own previous output) with additive isotropic noise injected on the predicted action in the design's output representation. The geodesic designs use output-noise norm $\varepsilon_{\text{add}} = 0.15$; the 6D noise magnitudes are calibrated within each chunk-size pair to match the noisy next-state geodesic RMS error (Appendix H). This fixes the auxiliary history input to the expert sequence while the learner state evolves in closed loop. For the 6D design, noise is added in the 6D output space; for the geodesic design, noise is added in the tangent space and projected onto $T_q\mathbb{S}^3$.

Data split. A random 90/10 train/validation split is used for all Experiment 5 models.

Compute. Experiments 1–4 are Monte Carlo simulations completed in minutes on a standard laptop CPU. Experiment 5 trains small MLPs ($\sim$50K parameters) for 200 epochs; total training and eval-

uation time for the main Experiment 5 configurations is approximately 2 hours on a single Apple M1 chip (Metal Performance Shaders (MPS) backend). The MuJoCo $\mathrm{SO}(3)$ diagnostic evaluation (Appendix K) uses the same Experiment 5 training procedure on the same hardware. The Panda execution-chain evaluation (Appendix M) uses PyTorch 2.5.1 and MuJoCo 3.8.0 on CPU; each complete seed of 18 trained policies and its evaluations takes approximately 1.6 hours, and the three seeds run independently.

C Robustness Across Expert Trajectories

We compare random-walk, spiral, and straight-geodesic schedules to examine how the finite-horizon curvature ordering and the agreement with the direction-averaged recurrence (Theorem 4) depend on the expert-direction sequence:

1. **Random walk:** At each step, the expert draws a uniformly random direction in T_sM with fixed norm v.
2. **Spiral:** At each step, the expert action direction rotates by 0.1 rad in a local orthonormal frame at the current state, producing a smoothly spiraling sequence of action directions with norm v.
3. **Straight geodesic:** The expert action at each step is the parallel transport of its previous-step action along the expert path, so the expert traces out a single geodesic on the manifold.

This comparison uses $v = 0.5$, $\varepsilon = 0.01$, $T = 150$, and 800 Monte Carlo trials per (manifold, expert-type) combination.

Across all three schedules the finite-horizon ordering $\mathbb{S}^2 < \mathbb{R}^2 < \mathbb{H}^2$ of the tracking error is preserved (Figure 7), and the direction-averaged recurrence is quantitatively accurate on $\mathbb{S}^2$ and $\mathbb{R}^2$ for the random-walk and spiral schedules (theory-to-simulation RMS ratios 1.00 and 0.99 for the random walk, 0.92 and 1.04 for the spiral); as in the main text, on $\mathbb{H}^2$ the nonlinear rollout falls below the averaged prediction once errors leave the local regime. The straight-geodesic schedule is not direction-averaged, and its behavior is consistent with the angle-dependent mechanism in Remark 7: the error component aligned with the fixed action direction is not contracted, so on $\mathbb{S}^2$ the RMS over the final 30 steps is 0.086, about three times the averaged prediction (0.029) although still below the $\mathbb{R}^2$ diffusive level (0.113); on $\mathbb{H}^2$ the RMS is 1.0, below the averaged exponential curve. The same effect appears on $\mathrm{SO}(3)$: with a straight-geodesic expert (previous action parallel transported) the $\mathrm{SO}(3)$ error does not saturate within $T = 200$ (terminal RMS $0.081/0.090/0.084$ at $v = 0.3/0.5/1.0$ versus the averaged prediction $0.080/0.049/0.025$), whereas resampling the expert direction at each step yields the saturation reported in Section 4.3. These schedules probe how departures from exact direction averaging affect finite-horizon prediction; Proposition 6 (Appendix A.7) gives the formal conditional orientation statistic that controls the scalar bounds.

D Curvature Variation Beyond a Single Homogeneous Manifold

Scope. The following two tests probe complementary aspects of the constant-curvature recurrence. Appendix D.1 tests long-horizon regime switching under a controlled schedule of homogeneous curvature segments, where the closed-form recurrence is valid within each segment. Appendix D.2 tests the local curvature mechanism on a single smooth variable-curvature torus, using a parallel-transport oracle to remove first-order vector-field sensitivity and match the $J'(0) = 0$ condition underlying the curvature factor. Together, these tests support the local Jacobi-field mechanism behind the curvature-dependent propagation factor.

D.1 Controlled Piecewise-Curvature Rollout

To test whether the local recurrence remains informative beyond a single homogeneous manifold, we construct a piecewise-constant curvature schedule that sequentially traverses three regimes: $\mathbb{S}^2$ ($\kappa = +1$, $t = 0$–29), $\mathbb{R}^2$ ($\kappa = 0$, $t = 30$–59), and $\mathbb{H}^2$ ($\kappa = -1$, $t = 60$–89). This controlled test connects homogeneous curvature segments rather than simulating a single smooth variable-curvature manifold; it asks whether a recurrence driven by the current curvature regime tracks accumulated

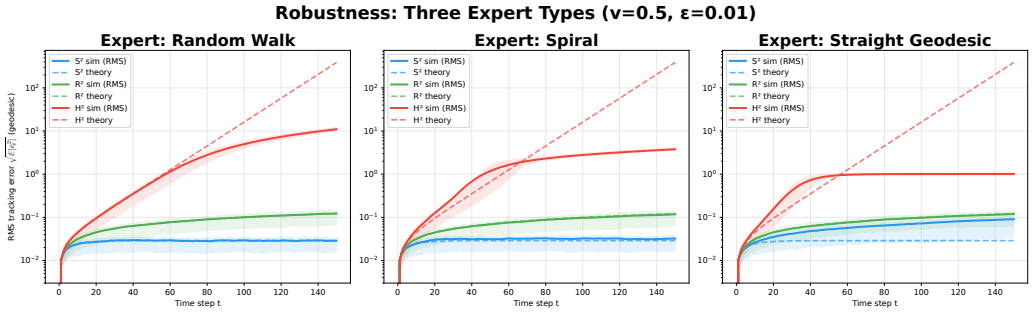


Figure 7: **Finite-horizon curvature ordering across expert schedules.** Tracking error under three expert types — random walk, spiral, and straight geodesic — on $\mathbb{S}^2$ (blue), $\mathbb{R}^2$ (green), $\mathbb{H}^2$ (red), with $v = 0.5$, $\varepsilon = 0.01$, $T = 150$, and 800 Monte Carlo trials. Solid: simulation RMS $\sqrt{\mathbb{E}[e_t^2]}$ (shaded band: interquartile range of e_t). Dashed: theory. The ordering $\mathbb{S}^2 < \mathbb{R}^2 < \mathbb{H}^2$ is preserved across all three schedules. The straight-geodesic schedule isolates the dependence on direction averaging: the component of error aligned with the fixed action direction is not contracted, so $\mathbb{S}^2$ lies above its averaged prediction while remaining below $\mathbb{R}^2$.

error better than a fixed-curvature surrogate. The candidate recurrence is

$$\mathbb{E}[e_{t+1}^2 \mid e_t] = \beta_{\det}(\kappa_{s(t)}, v)\, e_t^2 + \sigma_{\text{noise}}^2(\kappa_{s(t)}, v)\, \varepsilon^2 + \text{higher-order terms}, \tag{36}$$

where $\kappa_{s(t)}$ is the curvature of the active segment and the recurrence applies within segments only.

At each segment boundary, the next segment is initialized with a displacement whose geodesic magnitude matches the terminal error of the previous segment, with direction sampled isotropically in the new manifold's tangent space. This transfers the accumulated error distribution across regimes without requiring manifold-to-manifold state continuity. All three segments use the same $\mathbb{S}^2$, $\mathbb{R}^2$, and $\mathbb{H}^2$ implementations as Experiment 1, with $v = 0.3$, $\varepsilon = 0.01$, and 1000 Monte Carlo trials.

Figure 8 shows the result. The segment-wise homogeneous recurrence (blue) closely tracks the simulation curve (red) throughout all three regimes. In contrast, the constant $\bar{\kappa} = 0$ surrogate (gray) does not track the regime change: it is slightly too large in the $\mathbb{S}^2$ segment, where positive curvature suppresses error accumulation, and increasingly underestimates the trajectory in the $\mathbb{H}^2$ segment, where negative curvature amplifies error. This supports the controlled-regime interpretation that, when each segment is a homogeneous model space, accumulated error is better predicted by the active homogeneous segment's curvature than by a single trajectory-averaged curvature surrogate.

D.2 Curvature-Isolated Local Test on a Smooth Torus

The piecewise test (Appendix D.1) uses three homogeneous manifold segments. This experiment evaluates the local curvature factor $\beta_{\det}(v, \kappa)$ on a *single smooth manifold* whose curvature varies continuously, using parallel-transported actions that match the $J'(0) = 0$ initial condition of the constant-curvature model.

Manifold. We use a torus of revolution with major radius $R = 1.5$ and minor radius $r = 1$, parametrized by (θ, ϕ) with metric $\mathrm{d}s^2 = r^2\,\mathrm{d}\theta^2 + (R + r\cos\theta)^2\mathrm{d}\phi^2$, which for $r = 1$ reduces to $\mathrm{d}s^2 = \mathrm{d}\theta^2 + (R + \cos\theta)^2\mathrm{d}\phi^2$. The Gaussian curvature

$$K(\theta) = \frac{\cos\theta}{r(R + r\cos\theta)}, \tag{37}$$

which becomes $K(\theta) = \cos\theta/(R + \cos\theta)$ for $r = 1$, varies continuously from $K = +0.4$ at the outer equator ($\theta = 0$) through $K = 0$ ($\theta = \pi/2$) to $K = -2$ at the inner equator ($\theta = \pi$). This provides a single smooth two-dimensional manifold whose Gaussian curvature samples the positive, near-zero, and negative curvature signs appearing in the constant-curvature model.

Curvature-isolated oracle protocol. The inherited-error propagation factor $\beta_{\det}$ is derived under the Jacobi field initial condition $J'(0) = 0$ (parallel initial velocities). In the constant-curvature

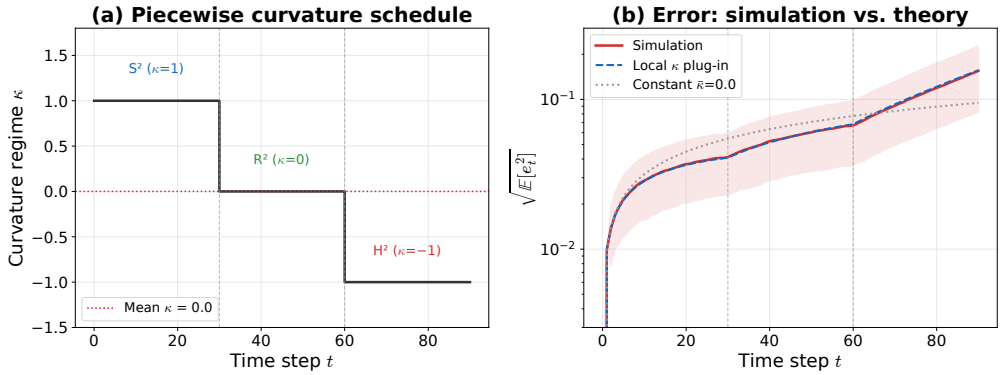


Figure 8: **Controlled piecewise-curvature rollout with matched error-magnitude transfer across segment boundaries.** *(a)* Curvature schedule: $\kappa = +1$ ($\mathbb{S}^2$), then $\kappa = 0$ ($\mathbb{R}^2$), then $\kappa = -1$ ($\mathbb{H}^2$), with segment boundaries at $t = 30$ and $t = 60$. *(b)* The segment-wise homogeneous recurrence tracks simulation (red, shaded band ± 1 standard deviation across 1000 trials) across all three regimes, while the constant $\bar{\kappa} = 0$ surrogate overestimates slightly in the positive-curvature segment and underestimates increasingly in the final negatively curved segment.

model analyzed in the main text, this condition is enforced by comparing geodesics with parallel initial velocities; equivalently, the learner's reference action at the perturbed state is the parallel transport of the expert action. On a non-homogeneous manifold such as the torus, using the same coordinate velocity instead of the transported action generally introduces an additional first-order term $J'(0) \neq 0$ proportional to the metric gradient. This metric-gradient effect can dominate the curvature signal and is outside the scope of the β_{det} formula.

To match the theorem's assumptions, we construct a *curvature-isolated oracle*: at each expert point $x_t^\star$, a small isotropic perturbation η (with $\|\eta\| = \rho = 10^{-3}$) is applied via the exponential map to produce the perturbed state $\hat{x}_t = \exp_{x_t^\star}(\eta)$. The expert's meridian action $(v, 0)$ is then parallel-transported from $x_t^\star$ to $\hat{x}_t$ by solving the parallel transport ordinary differential equation (ODE) along the connecting geodesic. Both expert and learner take one geodesic step, and the empirical amplification factor is

$$\beta_{\text{emp}}(\theta) = \mathbb{E}_\eta \left[\frac{d_{\text{loc}}^2(x_{t+1}^\star,\, \hat{x}_{t+1})}{d_{\text{loc}}^2(x_t^\star,\, \hat{x}_t)} \right], \tag{38}$$

where d_{loc} denotes the local quadratic metric approximation at the corresponding expert point, and the expectation is taken over 64 uniformly spaced perturbation directions. We also run a *natural meridian* ablation in which the learner uses the same coordinate velocity $(v, 0)$ without parallel transport, introducing the $J'(0) \neq 0$ metric-gradient term.

Results. Figure 9 reports the results for step sizes $v \in \{0.05, 0.10, 0.20\}$.

Panel (a) shows the *recovered curvature*: from the empirical β_{emp} at each θ, we invert the constant-curvature formula $\beta_{\text{det}} = \frac{1}{2} + \frac{1}{2}\text{c}_\kappa^2(v)$ to extract an effective curvature $\hat{K}$. The inversion uses the sign branch determined by β_{emp}: $\beta_{\text{emp}} < 1$ corresponds to the positive-curvature cosine branch, $\beta_{\text{emp}} = 1$ to the flat case, and $\beta_{\text{emp}} > 1$ to the negative-curvature hyperbolic-cosine branch. Under the curvature-isolated oracle, $\hat{K}(\theta)$ closely tracks the true Gaussian curvature $K(\theta)$ across the full range from $+0.4$ to -2.0, with agreement improving as $v \to 0$.

Panel (b) compares the prediction error $|\beta_{\text{emp}} - \beta_{\text{det}}|$ for the curvature-isolated oracle (blue) and the natural meridian field (red) at $v = 0.10$. In this test, the oracle error is 2–3 orders of magnitude smaller ($\sim 10^{-4}$ vs. $\sim 10^{-2}$). The reduction under parallel transport supports the role of the additional initial-separation-rate term $J'(0) \neq 0$ in the natural meridian field. Both errors become small near $\theta = 0$ and $\theta = \pi$, where $\sin\theta = 0$ by symmetry and the first-order metric-gradient effect disappears.

Panel (c) shows that the oracle error decreases as $v \to 0$, consistent with the local nature of the curvature approximation: variable-curvature and finite-step effects diminish in the small-v limit.

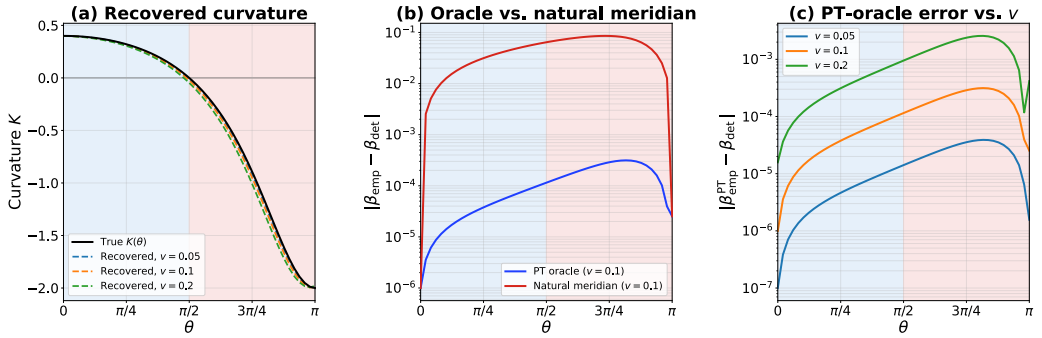


Figure 9: **Curvature-isolated local amplification test on a smooth torus** ($R = 1.5$, $r = 1$; K ranges from $+0.4$ to -2.0). *(a)* Under the curvature-isolated (parallel-transport) oracle, the curvature $\hat{K}$ recovered from the empirical one-step amplification factor closely tracks the true Gaussian curvature $K(\theta)$, with agreement improving as $v \to 0$. *(b)* Prediction error $|\beta_{\text{emp}} - \beta_{\text{det}}|$ at $v = 0.1$: the parallel-transport oracle (blue) yields error 2–3 orders of magnitude smaller than the natural meridian field (red), which carries an additional initial-separation-rate term. *(c)* Oracle error decreases as the step size is reduced, supporting the leading-order local character of β_{det}.

Discussion. This test is an assumption-matched local diagnostic: by parallel transporting the expert action to the perturbed state, we remove first-order vector-field sensitivity and enforce the $J'(0) = 0$ condition underlying the curvature factor in the constant-curvature model. Under this curvature-isolated oracle, the empirical one-step amplification recovers the local Gaussian curvature profile of the torus, with agreement improving as $v \to 0$. In contrast, the natural meridian field introduces a nonzero covariant variation term, causing the curvature-only prediction to deviate. Together with the piecewise-curvature rollout in Appendix D.1, this supports the local Jacobi-field mechanism while clarifying the scope of the theory: arbitrary smooth variable-curvature rollouts require an explicit Jacobi-field or comparison-geometry treatment, including curvature variation and vector-field sensitivity terms, rather than a pointwise substitution $\kappa \leftarrow K(x)$.

D.3 Observation–Policy–Dynamics Composition

The theorem takes actions as tangent increments on the state manifold itself. In the decoupled setting $o = h(s)$, $a = \pi(o)$, $s^+ = F(s, a)$, the closed-loop map $G(s) = F(s, \pi(h(s)))$ has differential

$$DG \;=\; D_1F \;+\; D_2F \circ D\pi \circ Dh\,, \tag{39}$$

which shows how observation, policy, and dynamics sensitivities enter a more general propagation analysis: D_1F captures direct state dependence at fixed action coordinates, while $D_2F \circ D\pi \circ Dh$ captures the policy response through the observation map.

E Experiment 4: Additional Figures

The surrogate interpolation sweep in Figure 5 (main text) shows the overall trend. Here we provide three supplementary views of the same experiment. Figure 10 shows RMS error trajectories at four surrogate interpolation strengths, illustrating how the theory–simulation gap grows as α decreases (stronger negative curvature). Figure 11 shows the full curvature phase diagram, the finite-horizon counterpart of the one-step factor shown in Figure 1, panel (b). Figure 12 visualizes the reduced recurrence's joint dependence on curvature and time, with the $\kappa = 0$ boundary separating the bounded and exponentially growing regimes.

F Local-Recurrence Accuracy on $\mathbb{H}^2$ Across Effective Step Sizes

On $\mathbb{H}^2$, the reduced recurrence describes amplification of small separations. With $\beta_{\text{det}} = \frac{1}{d} + \frac{d-1}{d}\cosh^2(\sqrt{|\kappa|}\,v)$ ($\beta_{\text{det}} \approx 1.136$ at $\kappa = -1$, $v = 0.5$), its RMS prediction $\sqrt{\mathbb{E}[e_T^2]}$ scales as

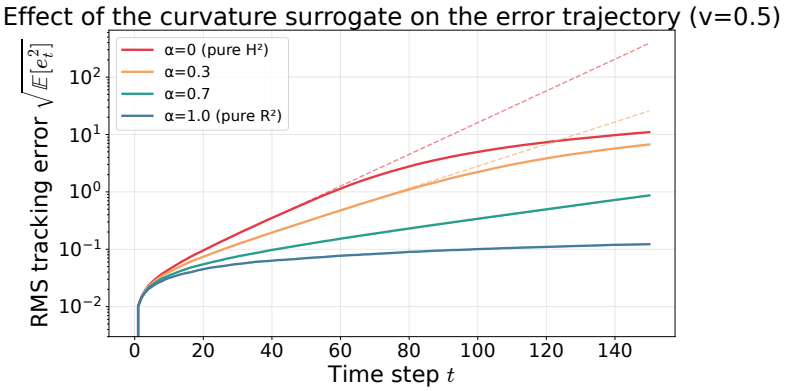


Figure 10: **Effect of the curvature surrogate on error trajectories.** RMS error vs. time at four surrogate interpolation strengths: $\alpha = 0$ (pure $\mathbb{H}^2$), $\alpha = 0.3$, $\alpha = 0.7$, and $\alpha = 1.0$ (pure $\mathbb{R}^2$). Solid: simulation RMS. Dashed: first-order theory. The recurrence is close to simulation near $\alpha = 1$; at small α, the nonlinear rollout falls below the local prediction at larger separations.

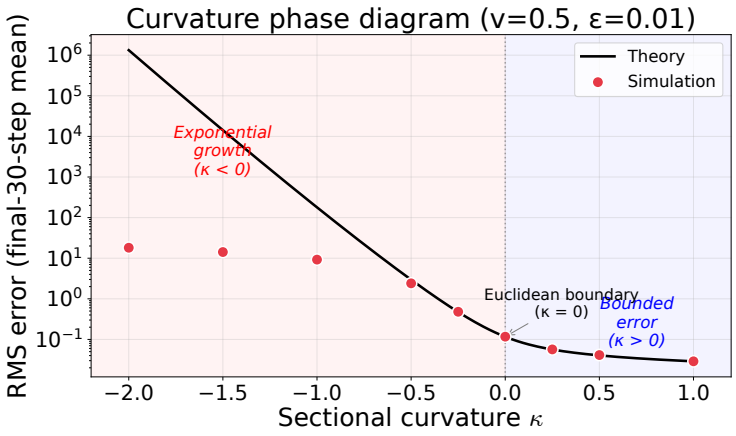


Figure 11: **Curvature phase diagram.** Long-horizon RMS error (averaged over the final 30 steps) vs. sectional curvature κ at $v = 0.5$, $\varepsilon = 0.01$, $T = 150$. Black line: theory with the same averaging window. Red points: simulation. $\kappa = 0$ marks the Euclidean boundary between the positive-curvature bounded and negative-curvature amplified regimes of the reduced recurrence. Theory matches simulation closely for $\kappa \geq 0$; under strong negative curvature ($\kappa \lesssim -0.5$), the leading-order recurrence overestimates the magnitude of divergence while preserving the regime ordering (Appendix F).

$\beta_{\text{det}}^{T/2}$ up to additive-noise constants. As accumulated separation leaves the local regime, higher-order finite-separation effects omitted by the leading-order recurrence become important, and the simulated RMS grows more slowly than the repeated local prediction. Because the recurrence is a small-separation expansion, repeated application need not remain quantitatively accurate at large accumulated separation. Table 3 quantifies the resulting theory-to-simulation ratios across effective step sizes in the action-chunking experiment (Section 4.2).

Across the tested chunk sizes, simulated finite-horizon RMS error increases with v_{eff}, preserving the fixed-event ordering predicted by the inherited-error propagation factor, while the reduced recurrence increasingly exceeds the simulated magnitude at larger effective steps. In the surrogate interpolation sweep (Figure 5), the corresponding RMS ratios are $0.99\times$ for $\alpha \geq 0.7$, $1.2\times$ at $\alpha = 0.5$, and $19\times$ at $\alpha = 0$.

For Experiment 1 on $\mathbb{H}^2$ at $v = 0.2$, the theory-to-simulation second-moment ratio over $t = 151$–200 is 1.005 ($1.01\times$ when rounded as in Table 1). Across the constant-curvature experiments,

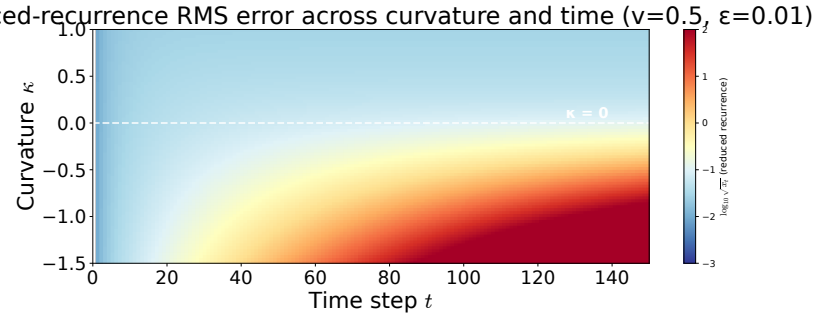


Figure 12: **Reduced-recurrence RMS error $\log_{10}\sqrt{x_t}$ as a function of curvature κ and time** t (computed from the recurrence, not from simulation; the $t = 0$ column, $x_0 = 0$, is left blank). The $\kappa = 0$ boundary (white dashed line) separates the positive-curvature bounded regime from the negative-curvature exponentially growing regime of the recurrence. The color gradient visualizes how curvature magnitude controls the predicted speed of divergence.

Table 3: Theory-to-simulation RMS ratio on $\mathbb{H}^2$ across chunk sizes. Ratios compare the reduced-recurrence RMS prediction $\sqrt{\mathbb{E}[e_t^2]}$ with the simulated RMS error, both averaged over the final 30 of the $T = 150$ decision events of Section 4.2.

H	v_{eff}	β_{det}	Theory/Sim ratio
1	0.1	1.005	1.00×
2	0.2	1.020	1.00×
3	0.3	1.046	1.07×
5	0.5	1.136	19×
8	0.8	1.394	$> 10^6\times$

the recurrence is quantitatively accurate on $\mathbb{S}^2$ and $\mathbb{R}^2$, and on $\mathbb{H}^2$ while separations remain local. Beyond that range, the observed finite-horizon sign and manifold ordering persist while the reduced recurrence overpredicts magnitude.

G SO(3) Quaternion Implementation Details

Hemisphere identification. Unit quaternions q and $-q$ represent the same rotation. To maintain a continuous quaternion lift, exponential-map updates retain the hemisphere reached by the geodesic; the identification $q \sim -q$ is applied only in distance and logarithmic-map computations.

Implementation. Concretely,

$$d(q_1, q_2) = 2\arccos\big(\min(|\langle q_1, q_2\rangle|, 1)\big), \qquad \log_{q_1}(q_2) : \text{if } \langle q_1, q_2\rangle < 0, \text{ replace } q_2 \leftarrow -q_2\,. \tag{40}$$

The exponential map returns whichever hemisphere the geodesic reaches, so quaternion trajectories remain continuous while distances are computed on $\mathrm{SO}(3)$.

Numerical consistency. Across random base quaternions and tangent directions at the step sizes used in the $\mathrm{SO}(3)$ experiments, the empirical one-step amplification agrees with β_{det} within 0.1%.

H Experiment 5 Supplementary: Unconstrained Teacher-Forced Designs

This appendix reports the per-design one-step errors and sensitivities underlying the four trajectories plotted in Figure 6(b). These configurations use unconstrained MLPs without spectral normalization, complementing the spectrally normalized measurements across output scales reported in Appendix I.

One-step measurements and decoder calibration. Two one-step error measures are reported on the same validation split: ε_{rep}, the mean residual norm in the design's own output representation (6D

difference or tangent vector), a representation-space error statistic that is the y-axis of Figure 6(a); and ε_{geo}, the RMS geodesic distance on $\mathrm{SO}(3)$ between the decoded predicted next state and the true next state, which is comparable across representations. During teacher-forced evaluation, a perturbation of norm $\varepsilon_{\text{add}} = 0.15$ is added in the design's output representation (tangent space for the geodesic design, 6D vector for the 6D design) before decoding; because the decoders are nonlinear, the same output-noise norm does not correspond to the same geodesic noise, so the resulting noisy one-step error $\varepsilon_{\text{noisy}}$ is measured directly through the decoder on the validation split (8 noise draws per sample) rather than computed as $\sqrt{\varepsilon_{\text{geo}}^2 + \varepsilon_{\text{add}}^2}$. Decoder response to a norm-0.15 output perturbation at zero residual gives geodesic RMS 0.150 for the tangent-space exponential decoder and 0.106 for the 6D Gram–Schmidt decoder, which is why the 6D designs have lower $\varepsilon_{\text{noisy}}$ than the plug-in formula would suggest. For $H{=}5$, ε_{geo} is the per-decision (chunk) error.

Matched-noise rollout protocol. Rollout errors are recorded only at executed states (chunk endpoints for $H{=}5$). For the rollouts of Figure 6(b), the output-noise norm of each 6D design is calibrated by bisection so that its $\varepsilon_{\text{noisy}}$ equals that of the geodesic design with the same chunk size ($\varepsilon_{\text{add}} = 0.194$ for A, 0.198 for D; the geodesic designs use 0.15); noise is matched separately within the $H{=}1$ and $H{=}5$ pairs (to within 0.2%), not across chunk sizes.

Table 4: Per-design measurements for the unconstrained teacher-forced configurations underlying Figure 6(b): representation-space mean error ε_{rep}, clean geodesic RMS next-state error ε_{geo}, noisy geodesic RMS error $\varepsilon_{\text{noisy}}$ measured through the decoder at output-noise norm $\varepsilon_{\text{add}} = 0.15$, the calibrated ε_{add} used for the matched-noise rollouts, representation-space sensitivity estimate L, and median rollout error at $T = 150$ at $\varepsilon_{\text{add}} = 0.15$ and at matched $\varepsilon_{\text{noisy}}$.

Method	ε_{rep}	ε_{geo}	$\varepsilon_{\text{noisy}}$ (ε_{add}=0.15)	matched ε_{add}	matched $\varepsilon_{\text{noisy}}$	L	e_{150} (ε_{add}=0.15)	e_{150} (matched)
A (6D, $H{=}1$)	0.107	0.115	0.157	0.194	0.180	6.41	1.33	1.46
B (Geo, $H{=}1$)	0.073	0.101	0.180	0.150	0.180	23.8	1.73	1.73
C (Geo, $H{=}5$)	0.042	0.153	0.215	0.150	0.215	8.97	1.31	1.31
D (6D, $H{=}5$)	0.070	0.162	0.195	0.198	0.215	3.20	1.13	1.14

The four-design comparison is directionally consistent with the L–ε_{rep} trade-off discussed in Section 4.5 and Figure 6(a): chunking ($H{=}5$) reduces ε_{rep} but raises the per-decision geodesic error ε_{geo} (a five-step target) while changing L in a representation-dependent way, and among the $H{=}1$ designs the geodesic representation (Design B) has the lower clean geodesic error but the largest estimated representation-space sensitivity and the highest rollout error. At a common output-noise norm the 6D decoder attenuates the injected noise, and the rollout ordering within each chunk size (A before B, D before C) coincides with the ordering of $\varepsilon_{\text{noisy}}$. The calibration equalizes decoded noisy one-step RMS error within each chunk-size pair to within 0.2%; at this matched level the design with the larger estimated representation-space sensitivity has the higher terminal median error (median 1.73 vs. 1.46 for B vs. A; 1.31 vs. 1.14 for C vs. D), with overlapping interquartile ranges across trajectories and curves that cross at intermediate times. Teacher forcing holds the auxiliary previous-action input on the expert sequence while the learner state evolves in closed loop. The matched comparison relates terminal tracking error to representation-space sensitivity at matched noisy one-step RMS error, complementary to the fully self-fed test in Appendix L.

I Direct One-Step Amplification Measurement on SO(3)

The analytic $L > 0$ proxy in Section 3.4 uses $\beta_{\text{eff}} = \beta_{\text{det}}(1 + Lv)^2$ as a nominal diagnostic amplification factor, with the proxy evaluated at the nominal effective step $v_{\text{eff}} = Hv$ for the chunked designs, to be compared with the empirically measured β_{emp} defined below. This experiment uses spectrally normalized policies across output scales, complementing Appendix H, which reports the unconstrained MLP points used in Figure 6(b), and compares directly measured state amplification with the sensitivity-based diagnostic proxy.

Protocol. Let $\Phi_h(s)$ denote the noiseless next state produced by the learned policy from state s with previous-action input h. For each trained model, we sample validation states s and construct a geodesic perturbation $\tilde{s}$ with $d(s, \tilde{s}) = \delta$, with direction sampled isotropically; both states are propagated with the same previous-action input h, so both next states are produced by the learned policy. The empirical amplification is $\beta_{\text{emp}} = \mathbb{E}[d(\Phi_h(s), \Phi_h(\tilde{s}))^2]/\delta^2$. We test $\delta \in \{0.01, 0.03, 0.05\}$ rad;

variation across δ is below 0.1%, indicating stability over the tested perturbation scales. No output noise is injected, so β_{emp} measures the learned next-state map's local response to state perturbations.

Results. Table 5 reports β_{emp} alongside the analytic proxy $\beta_{\text{eff}} = \beta_{\text{det}}(1+Lv)^2$ for all four designs at three sensitivity-control scales. The empirical amplification remains subcritical ($\beta_{\text{emp}} < 1$) in all 12 configurations, even when the analytic proxy exceeds 1. The deviation between β_{eff} and β_{emp} is configuration dependent: β_{eff} falls below β_{emp} for the two $H{=}5$ designs at scale 0.05 and exceeds it in the remaining ten configurations. At larger L, the analytic proxy increasingly exceeds the measured amplification, reaching about $1.5\times$ at the highest tested scale. Several configurations remain close to the critical boundary (e.g., $\beta_{\text{emp}} \approx 0.95$ at scale $= 0.1$), so empirical subcriticality can coexist with weak attenuation of inherited perturbations. Thus subcriticality ($\beta_{\text{emp}} < 1$) and strong attenuation of inherited perturbations are distinct properties.

Table 5: Empirical vs. analytic one-step amplification on SO(3). All β_{emp} values are means measured at $\delta = 0.01$ rad; the results at $\delta = 0.03$ and 0.05 are nearly identical. L is the representation-space sensitivity estimate of Appendix B. ε_{rep} is the mean residual norm in the design's output representation (a representation-space error statistic); ε_{geo} is the RMS geodesic next-state error on $\mathrm{SO}(3)$, computed from the same checkpoints and validation split, which is comparable across representations (per decision, i.e., per five-step chunk for $H{=}5$). The analytic proxy $\beta_{\text{eff}} = \beta_{\text{det}}(1+Lv)^2$ exceeds 1 in several configurations (bold), but the empirical amplification remains subcritical throughout.

Design	scale	L	ε_{rep}	ε_{geo}	β_{det}	β_{emp}	β_{eff}
A (6D, $H{=}1$)	0.05	0.041	0.221	0.188	0.993	0.981	**1.010**
	0.10	0.083	0.211	0.182	0.993	0.969	**1.027**
	0.30	0.240	0.188	0.167	0.993	0.922	**1.091**
B (Geo, $H{=}1$)	0.05	0.047	0.187	0.189	0.993	0.971	**1.012**
	0.10	0.090	0.181	0.184	0.993	0.950	**1.030**
	0.30	0.262	0.168	0.174	0.993	0.895	**1.100**
C (Geo, $H{=}5$)	0.05	0.046	0.382	0.444	0.847	0.977	0.927
	0.10	0.093	0.344	0.414	0.847	0.952	**1.012**
	0.30	0.258	0.271	0.339	0.847	0.871	**1.339**
D (6D, $H{=}5$)	0.05	0.042	0.451	0.444	0.847	0.975	0.919
	0.10	0.083	0.393	0.410	0.847	0.948	0.992
	0.30	0.241	0.278	0.321	0.847	0.853	**1.304**

J Architecture Robustness (Experiment 5)

We evaluate the sensitivity–approximation trade-off of Section 4.5 across three MLP architectures—64×2, 128×3, 256×4 (width $\times$ depth)—at three spectral-normalization output scales ($\{0.05, 0.1, 0.3\}$), for all four designs. Table 6 reports the results.

Across all 36 configurations, the maximum variation in ε_{rep} across architectures at matched (method, scale) is approximately 3% (method D, scale $= 0.1$). Across the tested width/depth variants, matched design–scale configurations have similar representation errors and preserve the sensitivity–approximation trade-off observed in Section 4.5.

K MuJoCo Kinematic SO(3) Diagnostic Evaluation

Scope. This appendix evaluates whether a recurrence-inspired diagnostic score combining one-step error and measured amplification,

$$F_{\text{pred}} = \frac{\varepsilon_{\mathrm{SO}(3)}}{\sqrt{1-\beta_{\text{emp}}}}, \tag{41}$$

remains informative when expert trajectories, β_{emp} measurement, and closed-loop rollouts are all carried out through a standard MuJoCo (Todorov et al., 2012) kinematic integration pipeline. The setup uses a three-degree-of-freedom (3-DOF) ball joint without contact or joint-level control. All measured configurations satisfy $\beta_{\text{emp}} < 1$, so F_{pred} is well-defined without clipping.

Table 6: Architecture robustness: $(L, \varepsilon_{\text{rep}})$ across architectures at matched spectral-normalization output scale. The representation error ε_{rep} varies by less than 4% across architectures at each matched method/scale (approximately 3% for design D at scale 0.10); L shows somewhat larger variation but the overall L–ε_{rep} trade-off is preserved across architectures.

Method	Scale	64×2 $(L,\ \varepsilon_{\text{rep}})$	128×3 $(L,\ \varepsilon_{\text{rep}})$	256×4 $(L,\ \varepsilon_{\text{rep}})$
A (6D, H=1)	0.05	(0.042, 0.221)	(0.042, 0.221)	(0.039, 0.222)
	0.10	(0.083, 0.211)	(0.082, 0.211)	(0.081, 0.211)
	0.30	(0.240, 0.188)	(0.234, 0.188)	(0.234, 0.187)
B (Geo, H=1)	0.05	(0.046, 0.187)	(0.047, 0.187)	(0.046, 0.187)
	0.10	(0.091, 0.181)	(0.091, 0.181)	(0.090, 0.181)
	0.30	(0.258, 0.169)	(0.257, 0.168)	(0.256, 0.168)
C (Geo, H=5)	0.05	(0.046, 0.382)	(0.046, 0.382)	(0.044, 0.384)
	0.10	(0.093, 0.344)	(0.093, 0.344)	(0.088, 0.347)
	0.30	(0.257, 0.271)	(0.255, 0.271)	(0.256, 0.271)
D (6D, H=5)	0.05	(0.042, 0.451)	(0.042, 0.451)	(0.036, 0.456)
	0.10	(0.084, 0.391)	(0.084, 0.392)	(0.073, 0.403)
	0.30	(0.241, 0.276)	(0.243, 0.279)	(0.241, 0.280)

Setup. The environment is a single rigid body attached to the world by a 3-DOF ball joint, with identity inertia, no gravity, no actuator damping, and an Euler integrator at internal timestep 10^{-3} s. Each policy step consists of 10 MuJoCo substeps (effective policy $\Delta t = 0.01$ s). Body-frame angular velocity commands are sent to MuJoCo and integrated as standard quaternion updates; the controlled state is the body quaternion. The evaluation isolates kinematic orientation tracking in this single-body ball-joint setting. We use the same expert trajectory family, training procedure, and four designs (6D vs. geodesic; $H \in \{1, 5\}$) as Experiment 5, with three spectral-normalization output scales ($\{0.05, 0.1, 0.3\}$) and two seeds, for 24 trained policies in total. Rollouts are closed-loop in the state: the learner's predicted action updates the MuJoCo body orientation. As in Experiment 5, the auxiliary previous-action input is teacher-forced to the expert sequence while the learner state evolves in closed loop. Each closed-loop rollout runs for $T = 200$ policy steps, and the rollouts inject isotropic output noise of norm $\varepsilon_{\text{add}} = 0.03$ in each design's output representation. For chunked designs we record error only at chunk endpoints.

Integration consistency. Single-step MuJoCo updates agree with the analytic quaternion exponential used in Experiment 5 to within 10^{-7} rad across 20 random base orientations, with accumulated drift below 10^{-7} rad over 100-step trajectories; coordinate conversion errors remain at numerical precision across random orientations and round-trip conversions.

Quantities. For each policy we measure: (i) the SO(3) one-step geodesic error $\varepsilon_{\text{SO(3)}}$, the *mean* geodesic distance between the next-state quaternions decoded from the predicted and target actions (a mean rather than the RMS statistic ε_{geo} of Experiment 5, and clean rather than noisy; this geodesic metric is common across the two action representations and matches the metric of β_{emp} and of the observed rollout error); (ii) the representation-space sensitivity estimate L (95th-percentile finite-difference ratio); (iii) the empirical amplification β_{emp}, measured by the same protocol as Experiment 5 but with state evolution integrated by MuJoCo; and (iv) the late-half tracking error: the median across 100 state-closed-loop rollouts (with teacher-forced auxiliary previous-action inputs) at each executed state, averaged over the second half of the horizon ($t > 100$). The diagnostic score F_{pred} uses the clean one-step error $\varepsilon_{\text{SO(3)}}$ directly; the injected rollout noise is not folded into the score, because the same representation-space noise norm does not correspond to the same geodesic noise across decoders (Appendix H).

Results. Figure 13 reports four panels.

Panel (A) shows the L–$\varepsilon_{\text{SO(3)}}$ frontier with the two nominal margins as reference lines: all 12 H=1 policies exceed $L_{\text{crit}}(H{=}1) \approx 0.017$, and 6 of the 12 H=5 policies exceed $L_{\text{crit}}(H{=}5) \approx$

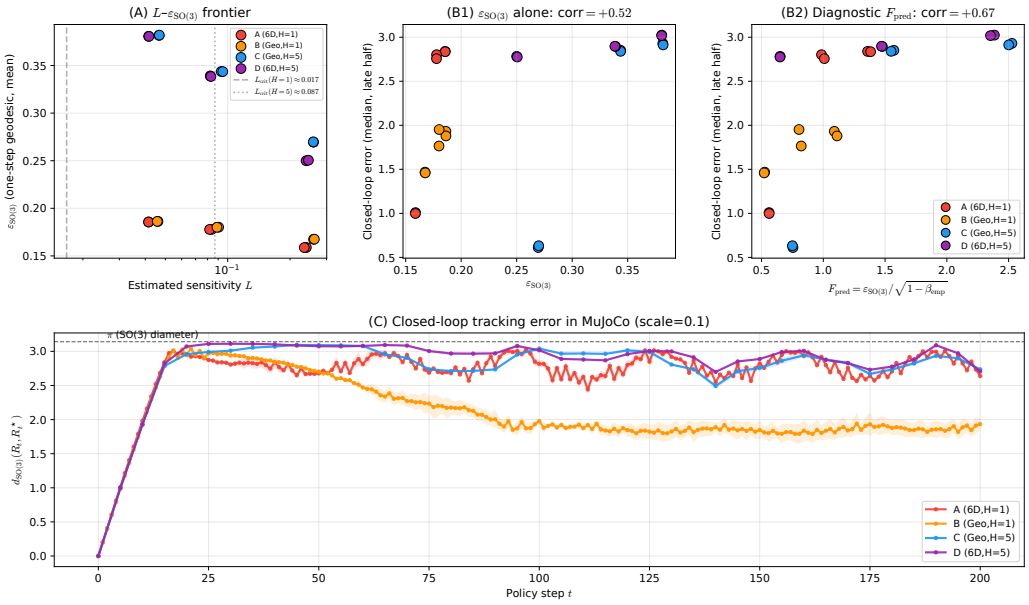


Figure 13: **MuJoCo kinematic SO(3) diagnostic evaluation** (24 policies: 4 designs × 3 scales × 2 seeds). *(A)* L–$\varepsilon_{\mathrm{SO}(3)}$ frontier; the two nominal margins are shown as reference lines. *(B1)* $\varepsilon_{\mathrm{SO}(3)}$ alone vs. observed late-half closed-loop error: Pearson $+0.52$. *(B2)* Diagnostic $F_{\mathrm{pred}} = \varepsilon_{\mathrm{SO}(3)}/\sqrt{1-\beta_{\mathrm{emp}}}$ vs. observed: Pearson $+0.67$ (Spearman $+0.86$); the higher pooled correlation appears after incorporating β_{emp}. *(C)* Closed-loop tracking error at scale 0.1: median over seeds, shaded bands show the seed-wise min–max range. Three designs rapidly approach the SO(3) diameter π; the geodesic $H{=}1$ design retains partial tracking and shows larger seed-to-seed variation, consistent with the fact that it does not saturate.

0.087, while the measured empirical amplification remains subcritical ($\beta_{\mathrm{emp}} < 1$) throughout. As in Section 4.5, $H{=}5$ designs trade higher $\varepsilon_{\mathrm{SO}(3)}$ for a more permissive nominal margin.

Panels (B1)–(B2) compare the two predictors of late-half closed-loop error across the 24 policies. Pearson correlations are:

$$\mathrm{corr}(\varepsilon_{\mathrm{SO}(3)}, \mathrm{obs}) = +0.52, \qquad \mathrm{corr}(F_{\mathrm{pred}}, \mathrm{obs}) = +0.67,$$

with the corresponding Spearman correlations $+0.72$ and $+0.86$. Seed-averaging the 24 points to 12 design×scale configurations yields the same Pearson values ($+0.52$ and $+0.67$) and a slightly stronger Spearman for F_{pred} ($+0.89$). A consolidated summary is in Table 7.

Table 7: Correlation between the two predictors and the late-half closed-loop tracking error in the MuJoCo SO(3) experiment.

Subset	Predictor	Pearson	Spearman
24 seed-level policies	$\varepsilon_{\mathrm{SO}(3)}$	$+0.52$	$+0.72$
	F_{pred}	$+0.67$	$+0.86$
12 seed-averaged configurations	$\varepsilon_{\mathrm{SO}(3)}$	$+0.52$	$+0.71$
	F_{pred}	$+0.67$	$+0.89$
Unsaturated subset (obs < 2.5 rad, $n = 10$)	$\varepsilon_{\mathrm{SO}(3)}$	-0.58	-0.03
	F_{pred}	$+0.51$	$+0.50$

Panel (C) plots median closed-loop tracking curves across the two training seeds at scale 0.1, with shaded bands indicating the seed-wise min–max range. Three of the four designs rapidly approach the SO(3) diameter π within roughly 15 policy steps, indicating near-complete loss of orientation tracking; the geodesic $H{=}1$ design retains partial tracking near 1.85 rad and shows visibly larger seed-to-seed variation, consistent with the fact that it does not saturate.

Diagnostic behavior across configurations. The diagnostic gain is mainly across the design×scale axis; within a fixed design (six points across three scales and two seeds), $\varepsilon_{\mathrm{SO}(3)}$ alone correlates strongly with observed error because changing the spectral scale moves $\varepsilon_{\mathrm{SO}(3)}$, L, and β_{emp} together; with six points per design, within-design correlations are descriptive. Second, because three of four designs saturate near π, we also report correlations on the unsaturated subset as a descriptive saturation diagnostic. Restricting to this subset (observed error < 2.5 rad, $n = 10$) gives Pearson (Spearman) correlations -0.58 (-0.03) for $\varepsilon_{\mathrm{SO}(3)}$ and $+0.51$ ($+0.50$) for F_{pred}: on this smaller subset, one-step error alone does not show a positive correlation, while F_{pred} does. Third, design D / scale 0.3 departs from the pooled score–error association: it has a low diagnostic score but high observed error, which the mean amplification factor does not capture.

Interpretation. First, a related sensitivity–one-step-error trade-off appears in this MuJoCo kinematic setting when expert generation, β_{emp} measurement, and rollout all pass through the integration pipeline. Second, the joint error/amplification score has higher pooled correlation with closed-loop error than $\varepsilon_{\mathrm{SO}(3)}$ alone in this setting.

L Learned-Policy Supporting Tests on $\mathbb{S}^2$, $\mathbb{R}^2$, and $\mathbb{H}^2$

This appendix reports two tests on learned policies that complement Section 4.5. The first isolates the rollout-geometric response conditional on the realized learned action; the second checks whether the resulting curvature ordering persists under self-fed execution.

Task and policies. On each of $\mathbb{S}^2$ ($\kappa = 1$), $\mathbb{R}^2$, and $\mathbb{H}^2$ ($\kappa = -1$) the expert takes geodesic steps of fixed length $v = 0.3$ whose heading, expressed in a frame parallel-transported along the trajectory, turns by $\omega + f \sin(\rho_t/2)$ per step, with $\omega = 0.05$, $f = 0.4$, and ρ_t the geodesic distance from a fixed origin; the turn rate therefore depends on the state. Training data consist of 240 expert trajectories of 40 steps with random initial heading and random start within geodesic radius 0.1–0.5 of the origin (on $\mathbb{R}^2$, a Gaussian start with standard deviation 0.5). The policy is an MLP with three hidden layers of width 128 (ReLU) mapping three state features and the previous action in frame coordinates (five inputs) to the next action in frame coordinates (two outputs); the state features are the ambient coordinates on $\mathbb{S}^2$, $(x_1, x_2, 0.1\rho)$ on $\mathbb{R}^2$, and $(0.1\rho\, x_1/\|x_{1:2}\|, 0.1\rho\, x_2/\|x_{1:2}\|, 0.1\rho)$ on $\mathbb{H}^2$, with ρ the geodesic distance from the origin, trained for 400 epochs with Adam (learning rate 10^{-3}, weight decay 10^{-4}, cosine schedule, batch size 256). We train three independent policy seeds per manifold. Replacing the state input by zeros, by a shuffled state, or by a previous-action-only policy raises the held-out one-step RMS error by 8–25× (previous-action-only 8.1–$10.5\times$, zero state 11.6–$24.7\times$, shuffled state 11.3–$14.2\times$ across seeds and manifolds), showing that the state features materially affect one-step prediction in this task. The finite-difference tangent-action sensitivity estimate L is the 95th percentile of $\|\Gamma_{q\to p}\hat{\pi}(q) - \hat{\pi}(p)\|/d(p,q)$ over perturbations at geodesic radius 0.01.

Action-clamp intervention. For each policy, 30 held-out expert trajectories provide 12 evaluation states each, and at each state 4 random tangent directions give perturbed states $\tilde{s} = \exp_s(\delta\, u)$, $\|u\| = 1$, at geodesic radius $\delta \in \{0.005, 0.01, 0.02\}$ (the same directions are reused across δ), yielding $1{,}440$ state–direction pairs per policy; the same directions are evaluated at each of the three perturbation radii. Both branches receive the same previous-action input. In the *full* branch, the policy is queried at both s and $\tilde{s}$. In the *clamped* branch, $\tilde{s}$ executes the action produced at s after parallel transport, i.e., the same physical tangent action expressed at the neighboring state. The clamp removes only the policy's differential response: the realized action and its approximation/representation error are held fixed rather than eliminated, and no output noise is added. Across three seeds and three perturbation radii, the clamped squared-separation ratio $d(s^+, \tilde{s}^+)^2/d(s,\tilde{s})^2$ follows the componentwise prediction of Remark 7,

$$\cos^2\theta + \mathrm{c}_\kappa^2(\|a\|)\sin^2\theta, \tag{42}$$

where θ is the initial separation–action angle and c_κ is the transverse factor from Lemma 1, with pairwise mean absolute error below 2×10^{-6} in every condition (maximum 1.7×10^{-6}). A control that clamps the learned direction at the fixed norm $\|a\| = v$ gives the ordering $\mathbb{S}^2 < \mathbb{R}^2 < \mathbb{H}^2$ in every seed. The full branch additionally includes the learned policy response; the full-minus-clamp contrast combines differential policy response with geometry–policy interaction, and the full-

branch ratio preserves the curvature ordering in all nine seed/radius conditions ($\mathbb{S}^2$: 0.936–0.940; $\mathbb{R}^2$: 0.981–0.983; $\mathbb{H}^2$: 1.025–1.029).

Matched-error self-fed rollouts. A separate set of three policy seeds per manifold was checkpointed every 20 epochs and their tangent-action RMS error (the RMS of the action residual in frame coordinates, not a decoded next-state error) evaluated on a calibration set of 100 expert trajectories covering the full 151-step rollout horizon. The common target error was the largest of the three per-manifold minima, and for each manifold the checkpoint whose calibration error was closest to the target was selected. The selected checkpoints were evaluated on an independent set of 100 expert trajectories, which also serve as the rollout references: the per-seed maximum-to-minimum one-step-error ratios across manifolds were 1.026, 1.047, and 1.097, all within the 1.10 matching criterion. Each rollout starts at the expert state s_1^* with the expert's first action as history and is executed self-fed for $T = 150$ steps with no injected noise, tracking error being the geodesic distance to the reference state; a trajectory is stopped at its first crossing of the threshold 1 (absorbing cap), after which its capped error is held at 1; we report absorbing-cap statistics (capped RMS, crossing fraction, and crossing time). Averaged over seeds (each statistic is computed per seed and then averaged), the tangent-action RMS errors were $0.00361/0.00364/0.00363$ on $\mathbb{S}^2/\mathbb{R}^2/\mathbb{H}^2$; the diagnostic multipliers $\beta_{\text{eff}} = \beta_{\det}(\kappa, v)(1 + Lv)^2$ were $0.991/1.029/1.074$; the RMS of the capped error at $t = 80$ was $0.373/0.676/0.862$; the fractions of trajectories crossing the threshold by $T = 150$ were $0.42/0.58/0.94$; and the median crossing steps among crossing trajectories were $119/86/57$. In every seed, β_{eff}, the capped RMS at $t = 80$, and the crossing fraction at $T = 150$ increase from $\mathbb{S}^2$ through $\mathbb{R}^2$ to $\mathbb{H}^2$, and the median crossing time decreases; these finite-horizon measurements preserve the curvature ordering under matched-error, self-fed execution. The tangent-action sensitivity estimates were $L = 0.059/0.048/0.044$; the finite-horizon ordering persists even though $\mathbb{H}^2$ has the smallest estimate.

Measured alignment. We additionally measure the learner-side pre-escape alignment every 10 steps: at the learner state, the squared cosine of the angle between $\log_{\hat{s}}(s^*)$ and the learner action had mean $0.802/0.516/0.574$ on $\mathbb{S}^2/\mathbb{R}^2/\mathbb{H}^2$ (the isotropic value in two dimensions is 0.5). These values summarize marginal alignment on pre-escape states; Proposition 6 instead conditions the statistic on the current error and history. Using these marginal means only as an illustrative plug-in to the band width at $v = 0.3$ gives $\beta_+ \approx 0.983 < 1$ on $\mathbb{S}^2$ and $\beta_- \approx 1.040 > 1$ on $\mathbb{H}^2$, while the flat case remains exactly at 1. These values are not certificates of Proposition 6, which requires a conditional bound, but are consistent with the same regime classification under the observed non-isotropy.

Scope. The action-clamp intervention evaluates the angle-conditioned geometric response, while the self-fed rollouts measure finite-horizon error and threshold crossings; together, they connect the local mechanism to finite-horizon learned-policy behavior and support the curvature ordering for the tested MLP policies.

M 7-DOF Panda Execution-Chain Diagnostic Evaluation

This experiment evaluates the learned-policy propagation diagnostics after orientation commands pass through an articulated robot execution chain consisting of inverse kinematics, joint-position servoing, MuJoCo dynamics, and forward kinematics. The empirical amplification β_{emp} measures the local response of this complete executed map.

Task and execution chain. We use a simulated 7-DOF Franka Emika Panda in MuJoCo (Todorov et al., 2012) for goal-conditioned tool-orientation tracking about a fixed home position. Reference orientations follow piecewise-smooth trajectories within a 0.9-rad neighborhood of the home orientation, with occasional offset changes. At each step, the expert applies a proportional $\mathrm{SO}(3)$ orientation command with gain 0.7, capped at a 0.2-rad rotation increment.

Each policy output specifies a desired end-effector orientation increment. A damped least-squares inverse-kinematics solve maps the resulting orientation target to joint space while maintaining the end-effector position, with redundancy resolution toward the home posture and away from joint limits. The target is executed by the Panda joint-position servos for 200 MuJoCo substeps at a 2 ms

simulation timestep, giving a 0.4 s policy interval; joint commands are rate-limited and bias-force compensation is applied during execution. Figure 14 shows representative executions of the expert and of the near-exact Geo-0.2/6D-0.15 matched pair used in the quantitative comparison below.

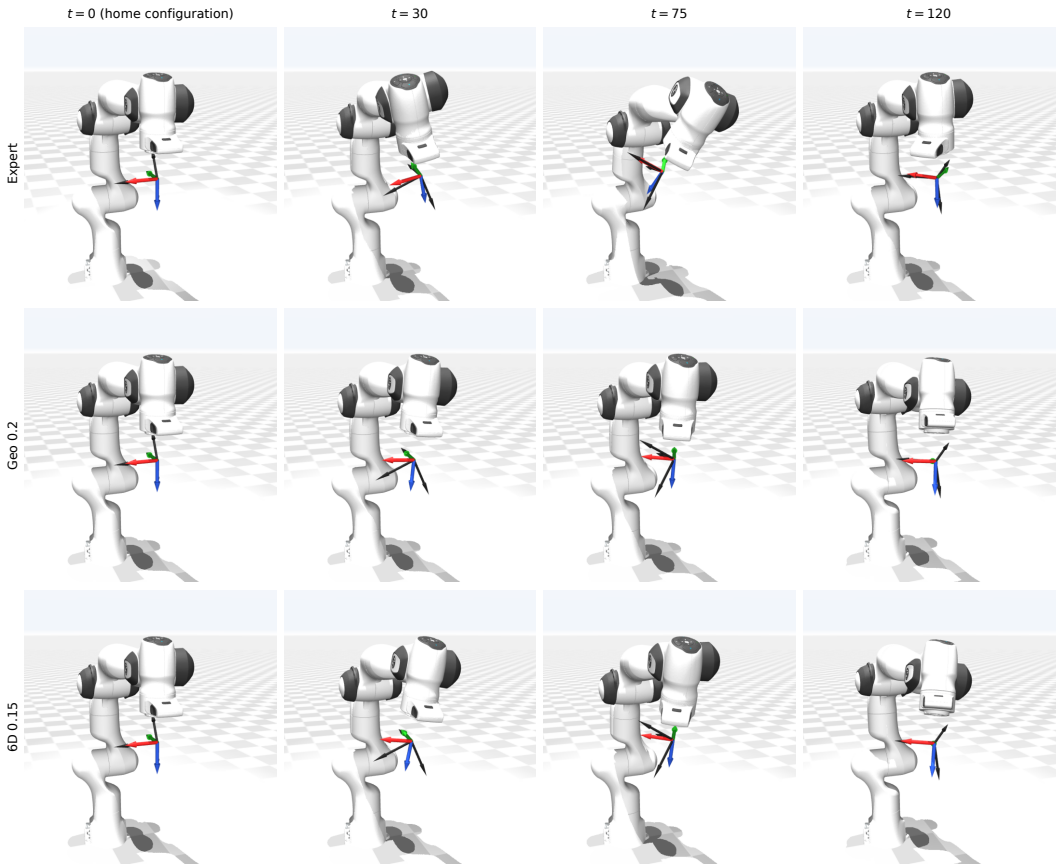


Figure 14: **Representative Panda execution-chain rollouts.** Expert, Geo-0.2, and 6D-0.15 executions (training seed 0) on one moving reference trajectory, shown at $t = 0$, 30, 75, and 120. The learned commands in the two policy rows pass through the same inverse-kinematics, joint-servo, MuJoCo-dynamics, and forward-kinematics pipeline as the expert. Colored arrows show the executed end-effector frame (x red, y green, z blue); dark arrows show the reference orientation at the same tool point. The Geo-0.2/6D-0.15 pair is the closely matched pair highlighted in Figure 15(B); quantitative comparisons across all matched pairs are reported in Figure 15 and Table 8.

The execution residual is the geodesic distance between the commanded end-effector orientation target and the orientation reached at the end of a policy step. Across seeds, the expert's mean execution residual is 0.0032–0.0034 rad and the 95th percentile is at most 0.0051 rad. Actuator torque limits remain active throughout execution; at least one actuator is saturated during approximately 39% of expert physics substeps.

All training demonstrations are retained for both representations. In one seed, 66 of the 45,000 training steps (0.15%, across two trajectories) meet the controller-event criterion of execution residual > 0.05 rad or nonpositive joint-limit margin; the calibration and test reference trajectories contain no such events.

Quaternion observations use the fixed hemisphere containing the home orientation; all expert, rollout, and probe states remain in its interior ($|\langle q, q_{\text{home}} \rangle| \geq 0.84$).

Policies and data. We compare the geodesic tangent-vector representation (Geo) and the 6D continuous rotation representation of Experiment 5 using three-hidden-layer, width-128 ReLU MLPs.

Spectral normalization is evaluated at output scales

$$\{0.05,\ 0.07,\ 0.1,\ 0.15,\ 0.2,\ 0.3,\ 0.5,\ 1.0\},$$

together with an unconstrained model. Models are trained for 200 epochs with Adam (learning rate 10^{-3}, batch size 256) and a cosine learning-rate schedule. Three independently trained seeds give 54 learned policies in total.

For each seed, the data comprise 300 training trajectories, 40 calibration trajectories, and 60 held-out test trajectories, all with horizon $T = 150$. The policy receives the current orientation, target orientation, and previous action in its corresponding representation. During self-fed rollouts, both the executed state and the previous-action input are carried forward from the learner's own preceding step; each rollout starts from the home configuration with a zero previous-action input, and no additional output noise is injected.

Executed one-step error and local amplification. The 40 calibration trajectories are split equally into a 20-trajectory *matching split* and a 20-trajectory *held-out calibration split*. From each split, 300 recorded simulator states are used for one-step evaluation. Starting from the same restored robot state, the expert command and learned command are separately passed through the complete execution chain. We define $\varepsilon_{\mathrm{exec}}$ as the mean geodesic distance on $\mathrm{SO}(3)$ between the resulting end-effector orientations.

The empirical amplification β_{emp} is measured on 300 probe states from the held-out calibration split. At each probe, an end-effector orientation perturbation of geodesic radius δ is realized by a local inverse-kinematics lift anchored at the recorded joint configuration and current end-effector position. The original and perturbed branches receive the same target and previous-action input and are each propagated through the learned policy, inverse kinematics, joint servoing, robot dynamics, and forward kinematics before their next-state separation is measured. For each probe, we divide the squared next-step end-effector separation by the squared initial separation actually realized by the local lift, then average these ratios over the retained probes; thus β_{emp} summarizes the local response of the complete executed map. Each seed supplies 300 state–direction probes; a probe is retained when the lift realizes the perturbation to within 10^{-5} rad, shifts the end-effector position by less than 10^{-6} m, keeps the realized radius within 10^{-4} rad of δ, and stays strictly inside the joint limits, and when both branch executions are finite without simulator warnings. These conditions retain 299, 300, and 300 probes for seeds 0, 1, and 2, respectively. We use $\delta = 0.01$ rad as the primary measurement and $\delta \in \{0.03, 0.05\}$ rad as finite-radius comparisons; the same probe states and perturbation directions are reused across the three radii and across all policies within a seed, and the radius comparison uses the mean over probes retained at all three radii (the *common-probe mean*).

Self-fed rollout metric. Each learned policy is evaluated on the same 60 held-out test trajectories. Tracking error is the geodesic distance between the learner's executed end-effector orientation and the corresponding executed expert reference. At each policy step we take the median across the 60 trajectories; the *late-half error* e_{late} is the average of these pointwise medians over $t = 76, \ldots, 150$.

Matched-error protocol. Pairs are selected independently within each training seed using only the matching split. The matching analysis includes Geo configurations with $\varepsilon_{\mathrm{exec}} \geq 0.01$ rad on the matching split. Eligible Geo models are processed in increasing order of their matching-split $\varepsilon_{\mathrm{exec}}$, and each is matched greedily, without replacement, to the unused 6D model minimizing the symmetric relative error gap

$$g(\varepsilon_1, \varepsilon_2) = \frac{2|\varepsilon_1 - \varepsilon_2|}{\varepsilon_1 + \varepsilon_2},$$

subject to $g \leq 0.10$. The resulting matching yields five pairs in each seed. The held-out calibration split is not used for pair selection and is used to evaluate whether the one-step-error match is retained.

Throughout the paired analysis, all differences are defined as 6D minus Geo. Finite evaluation-sample variability is quantified with 10,000 paired percentile-bootstrap resamples. For e_{late}, both members of a pair use the same resampled complete test trajectories and the late-half statistic is recomputed in each bootstrap replicate. For β_{emp}, the shared perturbation probes are resampled in pairs; for $\varepsilon_{\mathrm{exec}}$, the shared held-out calibration states are resampled in pairs. The 2.5th and 97.5th

percentiles define the reported 95% confidence intervals (CIs). These intervals quantify variability over the held-out evaluation samples conditional on each trained policy; each of the five design pairs is also repeated over three independently trained seeds.

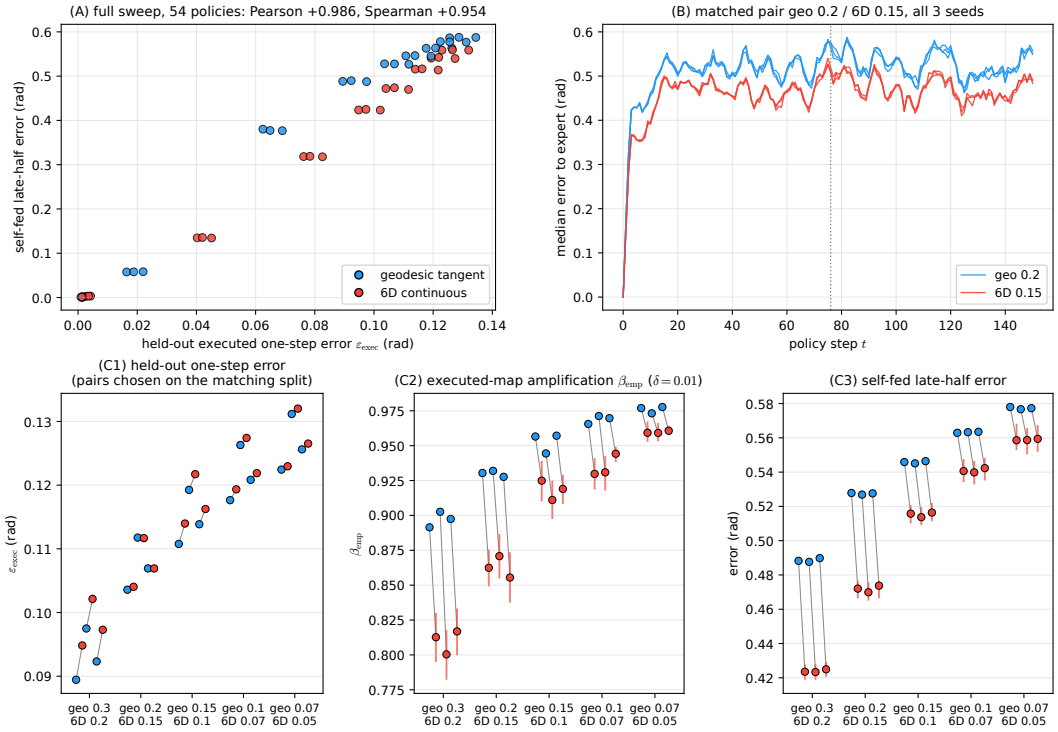


Figure 15: **7-DOF Panda execution-chain diagnostic evaluation.** *(A)* Across all 54 policies, held-out executed one-step error $\varepsilon_{\text{exec}}$ captures the broad sweep-wide variation in self-fed late-half error (Pearson 0.986, Spearman 0.954). *(B)* Self-fed pointwise-median tracking curves for the near-exact Geo-0.2/6D-0.15 pair, with all three training seeds shown separately; the dotted line marks the start of the late-half window. *(C1–C3)* Five Geo/6D design pairs repeated over three training seeds. Each marker is one training seed and each gray line joins the two members of a pair. Panel (C1) shows executed one-step error on the held-out calibration split after selection on the matching split; panel (C2) shows β_{emp} through the execution chain at $\delta = 0.01$; panel (C3) shows self-fed late-half error. Red vertical bars in panels (C1)–(C3) show the 95% paired-bootstrap intervals for the corresponding 6D–Geo differences, translated by the Geo value for display on the absolute-value axis.

Full-sweep behavior. Across the complete 54-policy sweep, executed one-step error captures the dominant accuracy trend: its Pearson and Spearman correlations with self-fed late-half error are 0.986 and 0.954, respectively. Measured amplification gives corresponding correlations of 0.974 and 0.994. The recurrence-inspired combination

$$\frac{\varepsilon_{\text{exec}}}{\sqrt{1-\beta_{\text{emp}}}}$$

has Pearson and Spearman correlations of 0.831 and 0.993. The matched-error comparison below examines the complementary setting in which the dominant variation in executed one-step accuracy has already been controlled.

Matched-error results. All 15 pairs selected on the matching split retain the 10% criterion on the held-out calibration split; the largest held-out relative gap is 5.9% and the mean gap is 1.9%. In every comparison, the 6D member has lower measured β_{emp} and lower self-fed late-half tracking error. The 95% paired-bootstrap interval of Δe_{late} lies entirely below zero in all 15 comparisons, with the largest upper endpoint at approximately -0.010 rad. The corresponding interval for $\Delta\beta_{\text{emp}}$

also lies below zero in all 15 comparisons, and the amplification ordering is unchanged when β_{emp} is measured at $\delta = 0.03$ or 0.05 (Table 8).

The one-step-error differences give a useful complementary view. In 13 of the 15 comparisons, the 6D member has a slightly larger point estimate of $\varepsilon_{\mathrm{exec}}$, and the paired interval lies above zero in all 13 of these cases. For the Geo-0.2/6D-0.15 pair, the held-out relative error gap is at most 0.45% in all three seeds, with the paired one-step-error interval containing zero in two seeds. After averaging over training seeds, the held-out executed one-step errors are 0.1074 and 0.1075 rad, while β_{emp} is 0.930 and 0.863 and self-fed late-half error is 0.527 and 0.472 rad for Geo and 6D, respectively. The paired rollout ordering therefore differs from the ordering of mean executed one-step error in most of the matched comparisons.

Table 8: Matched Panda comparisons. Differences are defined as $\Delta = \mathrm{6D} - \mathrm{Geo}$; $\Delta\varepsilon_{\mathrm{exec}}$ is in units of 10^{-3} rad, Δe_{late} in rad, and $\Delta\beta_{\mathrm{emp}}$ is dimensionless. Pairs are selected on the matching split; "held-out gap" is the symmetric relative gap g on the held-out calibration split, and all pairs remain within the 10% criterion. Brackets give 95% paired-bootstrap intervals.

Geo / 6D output scales	Seed	held-out gap	$\Delta\varepsilon_{\mathrm{exec}}$ (10^{-3} rad) [95% CI]	$\Delta\beta_{\mathrm{emp}}$ [95% CI]	Δe_{late} [95% CI]
0.30/0.20	0	5.85%	+5.39 [+4.79, +5.93]	−0.079 [-0.097, -0.061]	−0.065 [-0.069, -0.060]
	1	4.65%	+4.65 [+3.88, +5.32]	−0.102 [-0.120, -0.085]	−0.064 [-0.069, -0.060]
	2	5.25%	+4.97 [+4.41, +5.49]	−0.081 [-0.098, -0.064]	−0.065 [-0.069, -0.061]
0.20/0.15	0	0.45%	+0.46 [+0.03, +0.85]	−0.068 [-0.081, -0.055]	−0.056 [-0.061, -0.051]
	1	0.07%	−0.08 [-0.62, +0.41]	−0.061 [-0.077, -0.045]	−0.057 [-0.062, -0.051]
	2	0.004%	−0.004 [-0.40, +0.35]	−0.072 [-0.090, -0.054]	−0.054 [-0.061, -0.051]
0.15/0.10	0	2.85%	+3.20 [+2.77, +3.61]	−0.032 [-0.047, -0.017]	−0.030 [-0.036, -0.025]
	1	2.04%	+2.46 [+2.05, +2.84]	−0.033 [-0.047, -0.019]	−0.031 [-0.036, -0.025]
	2	2.08%	+2.39 [+2.07, +2.72]	−0.038 [-0.049, -0.028]	−0.030 [-0.035, -0.024]
0.10/0.07	0	1.43%	+1.70 [+1.39, +2.00]	−0.036 [-0.047, -0.024]	−0.022 [-0.029, -0.015]
	1	0.87%	+1.10 [+0.78, +1.42]	−0.040 [-0.053, -0.029]	−0.023 [-0.030, -0.017]
	2	0.86%	+1.05 [+0.77, +1.30]	−0.025 [-0.031, -0.021]	−0.021 [-0.028, -0.015]
0.07/0.05	0	0.43%	+0.52 [+0.33, +0.70]	−0.018 [-0.024, -0.010]	−0.019 [-0.025, -0.010]
	1	0.64%	+0.85 [+0.58, +1.10]	−0.014 [-0.020, -0.007]	−0.018 [-0.026, -0.011]
	2	0.72%	+0.91 [+0.68, +1.14]	−0.017 [-0.020, -0.013]	−0.018 [-0.025, -0.010]

Across the full sweep, the relative range of the common-probe mean β_{emp} across $\delta \in \{0.01, 0.03, 0.05\}$ reaches 21.1% (for an unconstrained model); for the configurations entering the matched-error comparison it is at most 2.1%, and the pairwise amplification ordering is unchanged at the two larger perturbation radii. The primary measurement remains at $\delta = 0.01$ rad throughout.

Together, the full sweep and matched comparisons separate two aspects of rollout behavior in this articulated execution chain: executed one-step error captures the broad accuracy trend, while measured propagation response resolves consistent rollout differences once one-step error is closely controlled.