Probabilistic Geometry of No-Regret Learning in Repeated Games

Tianyang Han*
Department of Computer Science
University of Pennsylvania
Philadelphia, PA, 19104
norahty@seas.upenn.edu

Abstract

We analyze learning dynamics in repeated games through the geometry of their strategy trajectories. Standard guarantees for no-regret learning describe the behavior of time-averaged play, but they do not characterize the *path* traced by the last iterate, which often exhibits persistent cycling even when averages approach equilibrium sets. Because mixed strategies are probability distributions, their natural geometry is information-theoretic and curved; consequently, Euclidean projections such as PCA can obscure or distort trajectory structure. We therefore embed joint play distributions using Intensive Principal Component Analysis (InPCA), which applies a Hellinger (square-root) lift before dimensionality reduction. In 3×3 games, this embedding reveals clean geometric signatures associated with learning rules: smooth ring-like orbits for multiplicative weights dynamics, polygonal (hexagon-like) cycles under online gradient descent due to projection effects, and triangular cycles for hybrid interactions. These observations support the view that common learning algorithms leave distinct, measurable "geometric fingerprints" in the space of joint play.

1 Introduction

Repeated-play games are a canonical model of strategic adaptation: at each round, agents update mixed strategies based on observed feedback. A standard lens is regret minimization, which yields strong guarantees about *time-averaged* play (e.g., convergence to equilibrium-like sets), and provides concrete update rules such as Multiplicative Weights (MWU) and Online Gradient Descent (OGD) [Roth, 2025]. However, these guarantees often leave the *last iterate* underdetermined: in many games and learning rules, the sequence (p_t, q_t) does not converge but instead cycles or recurrently wanders, even when averages behave predictably. Such nonconvergence is not a nuisance artifact—it is a qualitative dynamical feature that can govern stability, transient regimes, and potential predictability of future play [Mertikopoulos et al., 2018].

A core obstacle to understanding last-iterate behavior is geometric. Mixed strategies and joint play distributions are probability vectors, so their natural notion of proximity is information-theoretic rather than Euclidean. Consequently, standard visualization tools that treat trajectories as points in a flat space (e.g., applying PCA directly to probabilities) can smear structured motion into diffuse clouds and make distinct dynamical regimes appear indistinguishable. To address this, we adopt an information-geometric representation and visualization method. We log the joint distribution of play at each round, $P_t = p_t \otimes q_t$, and embed the resulting sequence using Intensive Principal Component Analysis (InPCA), which performs PCA after a Hellinger (square-root) lift that is well-adapted to probability simplices [Quinn et al., 2019].

*Supervised by Professor Aaron Roth

Our goal is to answer a concrete trajectory-level question: *Do common no-regret update rules induce reproducible, algorithm-specific geometric structure in the space of joint play, and if so, what mechanism explains it?* Empirically, in 3×3 games, InPCA reveals distinct geometric signatures: MWU produces smooth rotational orbits, whereas projected-gradient dynamics (OGD) produce polygonal attractors with sharp corners; in hybrid pairings, the geometry interpolates in a structured way (e.g., triangular cycles when OGD interacts with MWU). Moreover, the corners of these polygonal attractors align with phases where the OGD player is driven close to simplex vertices (near-pure actions), suggesting a mechanistic link between projection-induced boundary interactions and the observed turning points.

Contributions. Our analysis makes three contributions. First, we provide a reproducible pipeline for trajectory-level analysis of repeated-game learning: logging joint play P_t, embedding trajectories via InPCA, and pairing them with simple geometric diagnostics (e.g., curvature and angle summaries). Second, we empirically identify geometric fingerprints that distinguish learning rules beyond regret guarantees: MWU-like multiplicative dynamics produce smooth cyclic orbits, whereas OGD-like projected dynamics produce polygonal attractors with corners associated with near-pure phases. Third, we articulate a plausible mechanistic hypothesis for the polygonal geometry, namely, that sharp turns arise when projection steps drive OGD iterates toward the simplex boundary, while leaving open whether, and in what settings, these geometric structures have predictive or exploitability implications.

Together, these contributions document a consistent geometric phenomenon not visible from time-averaged analyses and establish a foundation for future work assessing its practical or strategic significance.

2 Background and Prior Work

We study learning dynamics in repeated two-player games, where at each round t the row and column players select mixed strategies $p_t \in \Delta_m$ and $q_t \in \Delta_n$, and the row payoff is

$$u(p_t, q_t) \;=\; p_t^\top A q_t.$$

Classical equilibrium theory characterizes stationary solution concepts (e.g., Nash equilibria and minimax solutions). In particular, in zero-sum games the minimax theorem guarantees a value v^* such that

$$\max_{p\in\Delta_m} \min_{q\in\Delta_n} p^\top A q \;=\; \min_{q\in\Delta_n} \max_{p\in\Delta_m} p^\top A q \;=:\; v^*.$$

When both players use no-regret algorithms, their empirical play distribution converges to the set of coarse correlated equilibria; in zero-sum games, $(\bar{p}_T, \bar{q}_T)$ approaches the minimax set and achieves near-optimal value v^* [Roth, 2025].

A key distinction in this literature is between *time-averaged play* and the *last iterate*. Define

$$\bar{p}_T := \frac{1}{T}\sum_{t=1}^{T} p_t, \qquad \bar{q}_T := \frac{1}{T}\sum_{t=1}^{T} q_t,$$

and compare $(\bar{p}_T, \bar{q}_T)$ to (p_T, q_T). No-regret theory primarily constrains the former: when both players use no-regret algorithms, empirical play approaches equilibrium-like sets (e.g., coarse correlated equilibria in general, and the minimax set in zero-sum games). However, the last iterate can fail to converge and instead exhibit persistent cycling or recurrent motion even when regret is small. Recent work on Follow-the-Regularized-Leader (FTRL) dynamics makes this separation explicit: time-averages converge to equilibrium-like objects, while the actual trajectory (p_t, q_t) can be recurrent and nonconvergent [Mertikopoulos et al., 2018]. This motivates a trajectory-level analysis that focuses on the *shape* of the path traced in strategy space rather than only on averaged outcomes.

We focus on two representative no-regret families used throughout this analysis.

Multiplicative Weights Update (MWU) updates probabilities multiplicatively, scaling action weights exponentially according to observed feedback; its iterates typically move smoothly in the interior of the simplex because the update is continuous in the current distribution and does not require hard feasibility corrections. *Online Gradient Descent (OGD)* takes a linear gradient step followed by a projection back to the feasible set (the simplex). Because this projection is piecewise-defined, OGD

can undergo abrupt directional changes when iterates interact with simplex faces or vertices. In 3×3 games such as Rock-Paper-Scissors, these structural differences yield markedly different trajectory geometry despite similar regret guarantees: as we show below, MWU produces smooth rotational orbits, whereas OGD produces polygonal cycles with sharp turns.

Existing analysis and visualization tools only partially capture this behavior. Regret curves and payoff traces summarize dynamics using one-dimensional statistics and primarily reflect the behavior of averages. Phase-portrait visualizations are informative in low-dimensional continuous-time systems but become difficult to interpret for discrete, higher-dimensional trajectories. Finally, applying PCA directly to probability vectors treats the simplex as a flat Euclidean space, which can collapse structured cycles into diffuse clouds. These limitations motivate an approach that respects the intrinsic geometry of probability distributions.

3 Information-Geometric Representation and Embedding

InPCA was originally introduced as a general manifold-learning method for probabilistic models [Quinn et al., 2019]. Quinn et al. demonstrated that the same Hellinger-based embedding can (i) recover the low-dimensional structure of statistical physics models such as the Ising model, (ii) track the training trajectory of a convolutional neural network on MNIST as digit clusters separate over epochs, and (iii) visualize the parameter manifold of the ΛCDM cosmological model, capturing both local and global structure more faithfully than t-SNE or diffusion maps. These applications support viewing InPCA as a natural tool for studying probability-valued dynamical systems; in our setting, the "model manifold" is the set of joint play distributions generated by repeated-game learning algorithms.

Our analysis operates on *joint play* rather than marginals. In our 3×3 games, each player uses a mixed strategy

$$p_t \in \Delta^2, \qquad q_t \in \Delta^2,$$

where $\Delta^2 := \{p \in \mathbb{R}^3 : p_i \geq 0,\ \sum_i p_i = 1\}$ is the three-action simplex. We form the joint distribution over *action pairs* via the outer product

$$P_t \;=\; p_t \otimes q_t \;\in\; \Delta^8,$$

which can be viewed as a length-9 probability vector (one entry for each pair (i, j)). Concretely,

$$P_t(i,j) \;=\; p_t(i)\, q_t(j),$$

the probability that the row player plays action i and the column player plays action j at round t. This representation preserves the interaction structure that determines payoffs and allows algorithm-dependent differences to appear in the coupled behavior of both players.

Because P_t lies on a probability simplex, Euclidean geometry is not well-aligned with natural notions of statistical proximity. We therefore adopt the Hellinger geometry used by Intensive Principal Component Analysis (InPCA). We apply the entrywise square-root lift

$$S_t \;=\; \sqrt{P_t},$$

which embeds the simplex into the positive orthant of a sphere and makes Hellinger distances linear in the ambient Euclidean space. We then center the lifted data and perform SVD/PCA:

$$\tilde{S}_t \;=\; S_t - \frac{1}{T}\sum_{s=1}^{T} S_s, \qquad \tilde{S} \;=\; U\Sigma V^\top.$$

Using the leading components, we define the three-dimensional embedding

$$X_t \;=\; U_{t,1:3}\,\frac{\Sigma_{1:3}}{\sqrt{2}} \;\in\; \mathbb{R}^3,$$

yielding a trajectory that captures dominant modes of variation in the Hellinger-lifted space.

To compare multiple runs, we normalize time to $[0, 1]$ by setting

$$\tau \;=\; \frac{t - t_{\min}}{t_{\max} - t_{\min}} \in [0,1],$$

since early iterations often move rapidly while later iterations evolve more slowly. This normalization enables coherent overlays of replicas and averaged trajectories without conflating "early" and "late" regimes.

4 Experimental Setup and Diagnostics

4.1 Games and learning rules

We focus on 3×3 games to keep the state space small enough for dense trajectory logging while still exhibiting nontrivial cycling. Our primary environment is Rock-Paper-Scissors (RPS), a zero-sum game with payoff matrix $A \in \mathbb{R}^{3\times 3}$ for the row player and $-A$ for the column player, and a unique mixed Nash equilibrium at the uniform distribution $(1/3, 1/3, 1/3)$.

On top of this fixed game, we vary only the learning rules. We study three pairings:

1. **MWU vs. Constant:** the row player runs Multiplicative Weights Update, while the column player plays a fixed pure strategy ("always Scissors" in the experiments).
2. **OGD vs. OGD:** both players run Online Gradient Descent on their own payoffs, with projection onto the simplex each round.
3. **OGD vs. MWU:** the row player runs OGD, the column player runs MWU. This hybrid pairing lets us probe how a "sharp" projected dynamic interacts with a smoother multiplicative one.

Unless otherwise stated, we use step sizes $\eta_r = \eta_c = 0.25$ and simulate $R = 20$ independent runs per condition, each for $T = 300$ rounds. Initial strategies are drawn by adding small normal noise to the uniform distribution and renormalizing, so trajectories start near but not exactly at the mixed equilibrium.

4.2 Logging protocol and joint-play representation

For each algorithm pair and each run, we log both the *last iterate* and the *time-average*. At round t we record

$$(p_t, q_t) \quad \text{and} \quad (\bar{p}_t, \bar{q}_t) := \Big(\frac{1}{t}\sum_{s=1}^{t} p_s,\ \frac{1}{t}\sum_{s=1}^{t} q_s\Big),$$

and form the corresponding joint distributions of play

$$P_t^{\text{last}} = p_t \otimes q_t, \qquad P_t^{\text{avg}} = \bar{p}_t \otimes \bar{q}_t \ \in\ \Delta^8.$$

We store these as length-9 vectors (one coordinate per action pair) together with metadata (run index, time step, and whether the point corresponds to "last" or "avg"). This gives us dense trajectories in the space of joint play, rather than in the marginal simplices.

To compare runs with different time horizons or transient lengths, we normalize time to $[0, 1]$ separately for each $(\text{kind}, \text{run})$ pair:

$$\tau_t \ =\ \frac{t - t_{\min}}{t_{\max} - t_{\min} + 10^{-12}} \in [0, 1].$$

This normalization is used whenever we aggregate or bin trajectories across runs, for example to compute mean "highways" or radius-over-time summaries.

4.3 InPCA embedding: implementation details

The Information-Geometric embedding described in Section 3 is implemented as follows. We stack all recorded joint distributions into a matrix $P \in \mathbb{R}^{M\times 9}$, where M is the total number of logged checkpoints across runs and kinds. We then apply the entrywise square-root lift $S = \sqrt{P}$, center across checkpoints, and compute a rank-3 Singular Value Decomposition:

$$S - \bar{S} \ =\ U\Sigma V^\top, \qquad \bar{S} = \frac{1}{M}\mathbf{1}\mathbf{1}^\top S.$$

Following Quinn et al. [2019], we define the 3-dimensional embedding

$$X_t = U_{t,1:3}\frac{\Sigma_{1:3}}{\sqrt{2}} \ \in\ \mathbb{R}^3.$$

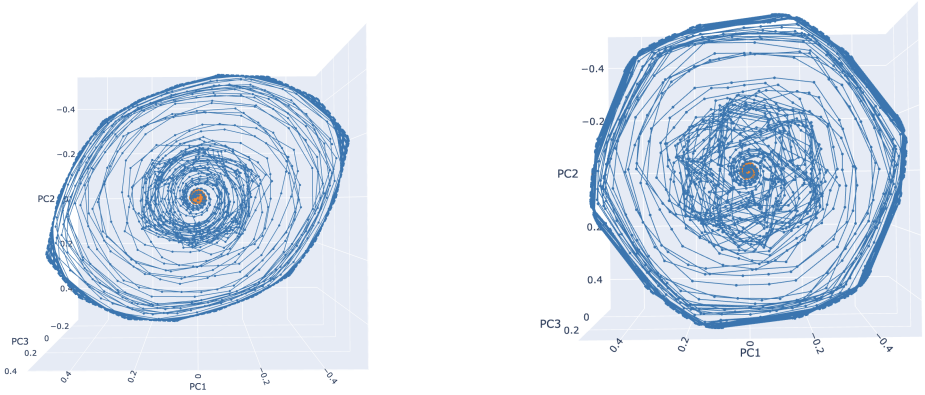
This representation preserves Hellinger distances between distributions up to second order and makes the dominant directions of variation in joint play explicit in the Euclidean coordinates $(\mathrm{PC1}, \mathrm{PC2}, \mathrm{PC3})$.

For each experimental condition (e.g., MWU vs. Constant), we construct a *single* embedding basis by concatenating all runs and both kinds (last and average) before performing SVD. This ensures that the resulting coordinates are directly comparable across runs and between last-iterate and averaged trajectories.

5 Additional Ablations and Controls

Outer product vs concatenation. To test whether joint-play coupling matters, we compare embedding $P_t = \mathrm{vec}(p_t \otimes q_t) \in \mathbb{R}^9$ against embedding the concatenated marginals $[p_t; q_t] \in \mathbb{R}^6$. The outer-product representation preserves action-pair structure that directly determines payoffs, and yields clearer algorithm-dependent geometric fingerprints.

Why PCA collapses structure. For OGD vs. OGD we directly compare PCA on P_t with InPCA applied to the Hellinger-lifted $\sqrt{P_t}$. PCA treats the joint-play probabilities as living in a flat Euclidean space and smears the polygonal orbit into a thick, entangled spiral. In contrast, the square-root lift aligns the embedding with Hellinger geometry, and the InPCA embedding recovers a clean, nearly regular polygonal cycle, preserving the algorithm-specific structure of the trajectory. Also see Figure 2 for a angled comparison.



(a) PCA on P_t

(b) InPCA on $\sqrt{P_t}$

Figure 1: OGD vs. OGD in Rock-Paper-Scissors under different embeddings. (**Left**) PCA on the joint-play probabilities P_t smears the polygonal orbit into a thick spiral. (**Right**) InPCA on the Hellinger-lifted $\sqrt{P_t}$ recovers a clean, nearly regular polygonal cycle, illustrating how respecting the probability geometry preserves algorithm-specific structure.

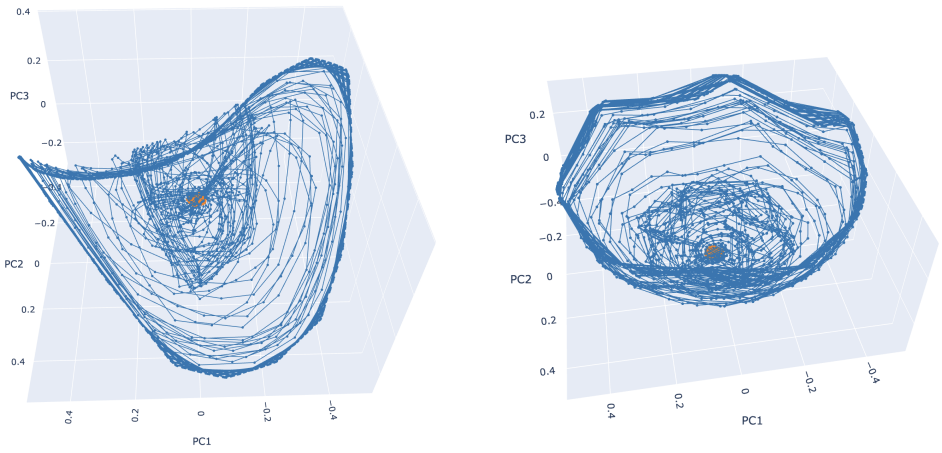
5.1 Geometric diagnostics

In addition to visual inspection of the 3D trajectories, we use several simple diagnostics that borrow ideas from dynamical systems and shape analysis.

The Euclidean radius of the embedded point,

$$r_t := \left\|X_t\right\|_2 = \sqrt{X_{t,1}^2 + X_{t,2}^2 + X_{t,3}^2},$$

measures how far the joint play is from the center of mass in Hellinger space. For each $(\mathrm{kind}, \mathrm{run})$ we treat r_t as a function of normalized time τ_t and then aggregate across runs by binning in τ and taking the median radius per bin. A last-iterate radius that stabilizes at a nonzero plateau while the time-averaged radius decays toward zero is evidence of persistent cycling coexisting with empirical convergence.



(a) PCA on P_t (b) InPCA on $\sqrt{P_t}$

Figure 2: OGD $\times$ OGD in Rock-Paper-Scissors: InPCA versus PCA.

Angular spectrum and symmetry. To quantify "how many corners" a cycle has, we project onto the $(\mathrm{PC1}, \mathrm{PC2})$ plane and compute the polar angle

$$\theta_t := \mathrm{atan2}(X_{t,2}, X_{t,1}) \ \in [0, 2\pi).$$

We form a histogram of θ_t over time and take the discrete Fourier transform of this histogram. The index k of the largest nonzero harmonic gives a crude estimate of the dominant rotational symmetry: $k \approx 1$ corresponds to a single ring-like orbit, $k \approx 3$ to a triangle-like pattern, and $k \approx 6$ to a hexagon-like pattern. We report these k values alongside the qualitative plots.

Discrete curvature and turning points. To localize sharp turns along a trajectory, we compute a discrete curvature proxy for each interior timestep:

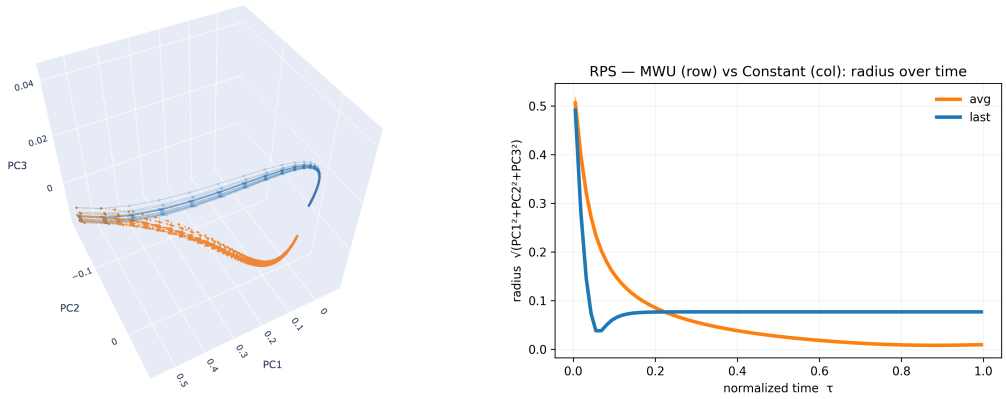
$$\kappa_t \ := \ \left\| X_{t+1} - 2X_t + X_{t-1} \right\|_2.$$

This is the Euclidean norm of the second finite difference and vanishes on a straight line. Within each run we identify local maxima of κ_t subject to a minimum separation constraint in t; these points correspond to "turning events" in the embedded space. By inspecting the underlying strategies (p_t, q_t) at these indices, we can relate geometric corners back to behavior on the probability simplex (e.g., near-pure phases under OGD).

6 Results

The RPS experiments illustrate that the geometry of learning trajectories is strongly algorithm-dependent and becomes visible only after embedding with an information-geometric method. Across all conditions we examined, the last iterate typically traces a stable orbit in the embedded space, while the time-average collapses inward toward the equilibrium region. This reproduces the conceptual separation emphasized earlier: empirical convergence can coexist with persistent last-iterate cycling.

MWU against a constant opponent. When one player runs MWU while the opponent plays a fixed pure strategy, the embedded joint distribution traces a smooth, curved path in InPCA space. The associated radius-over-time diagnostic in our experiments shows that the last iterate maintains a roughly stable radius, indicating sustained motion rather than convergence to a point, while the time-average radius decays toward the center, consistent with average-play stabilization. The salient observation is that InPCA preserves the smoothness and curvature of the motion, making the dynamical regime visually and quantitatively interpretable.

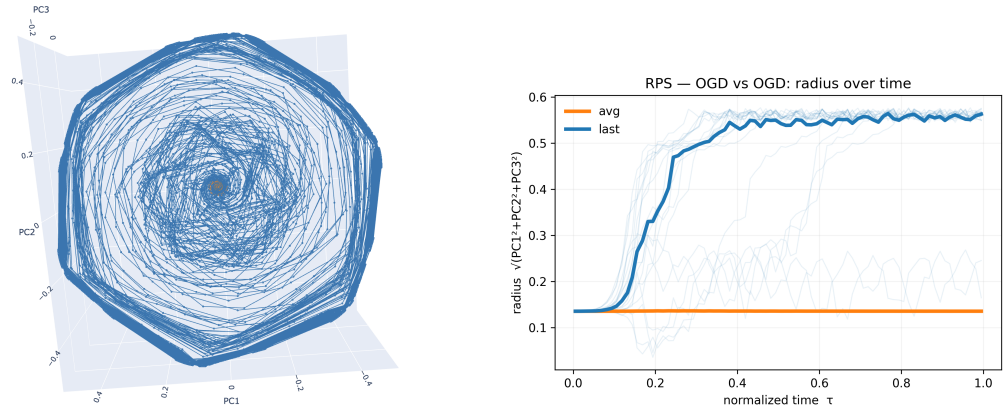


(a) InPCA trajectory

(b) Radius over normalized time

Figure 3: MWU against a constant opponent in Rock–Paper–Scissors. **Left**: InPCA embedding of joint play shows a smooth, curved orbit, indicating persistent last-iterate cycling. **Right**: Radius-over-time diagnostic: the last-iterate radius stabilizes at a nonzero value while the time-average contracts toward zero.

OGD versus OGD produces polygonal structure. When both players run Online Gradient Descent (OGD), the InPCA embedding reveals a distinctly cornered, hexagon-like cycle in the space of joint play. Unlike the smooth orbits produced by multiplicative dynamics, the trajectory exhibits sharp turns and piecewise-linear segments. This behavior is consistent with the structure of OGD updates on the simplex: linear gradient steps followed by projection induce regime changes when iterates interact with simplex faces and vertices.



(a) InPCA trajectory

(b) Radius over normalized time

Figure 4: OGD versus OGD in Rock-Paper-Scissors. **Left**: InPCA embedding of joint play reveals a polygonal, hexagon-like orbit with sharp corners. **Right**: Radius-over-time diagnostic shows a stable nonzero last-iterate radius alongside contracting time-averaged play.

Angular diagnostic confirms six-fold structure. To quantify the polygonal symmetry, we project the InPCA trajectory onto the $(\mathrm{PC1}, \mathrm{PC2})$ plane, compute the polar angle $\theta_t = \mathrm{atan2}(\mathrm{PC2}_t, \mathrm{PC1}_t)$ for the last iterate, and form a histogram over θ_t. Figure 5 shows the resulting angle histogram for OGD$\times$OGD in RPS. The mass concentrates into six regularly spaced peaks, corresponding to six preferred directions of motion. This six-fold angular pattern matches the hexagon-like orbit in Figure 4, supporting the interpretation that OGD induces a cycle with *six effective vertices or turns* in the embedded space.

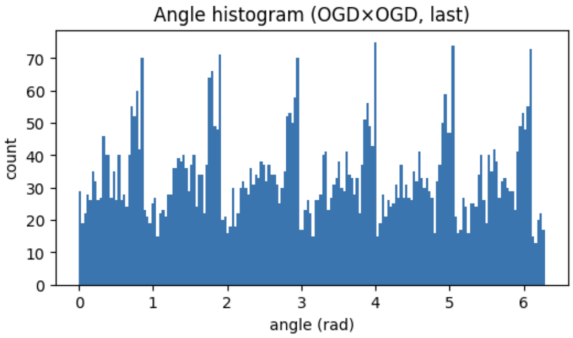
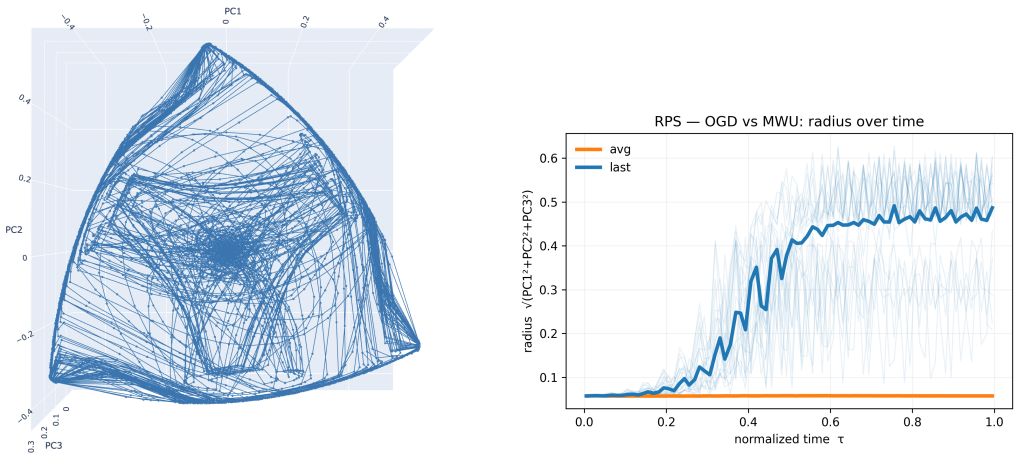


Figure 5: Angle histogram for the last-iterate InPCA trajectory in RPS under OGD×OGD. The six regularly spaced peaks in angle indicate six preferred directions around the cycle, consistent with a hexagon-like orbit with six effective vertices/turns in the embedded space.

OGD versus MWU yields triangular structure. In the hybrid pairing where one player uses OGD and the other uses MWU, the embedded trajectory concentrates into a triangle-like orbit rather than a ring or a hexagon. In our experiments, this manifests as three prominent corners, corresponding to phases where the OGD player is driven close to pure strategies. The associated angular analysis shows a dominant $k \approx 3$ mode, providing a quantitative counterpart to the triangular geometry. The hybrid case is informative because it indicates that polygonal structure is not merely a property of the game; it is tied to the update rule, and its interaction with a smoother multiplicative dynamic can reduce the effective symmetry from six-fold to three-fold.



(a) InPCA trajectory

(b) Radius over normalized time

Figure 6: OGD versus MWU in Rock-Paper-Scissors. **Left**: InPCA embedding of joint play reveals a triangle-like orbit with pronounced corners. **Right**: The last-iterate radius remains bounded away from zero while the time-average contracts inward, illustrating persistent cycling despite average-play stabilization.

Angular diagnostic confirms three-fold structure. We repeat the angular analysis for the hybrid OGD×MWU pairing by projecting the last-iterate trajectory onto the $(\mathrm{PC1}, \mathrm{PC2})$ plane and histogramming the polar angle θ_t. As shown in Figure 7, the mass now concentrates into *three* dominant peaks, roughly $2\pi/3$ radians apart. This three-fold angular pattern matches the triangle-like orbit in Figure 6 and supports the interpretation that the hybrid interaction produces a cycle with *three effective vertices or turns* in the embedded space.

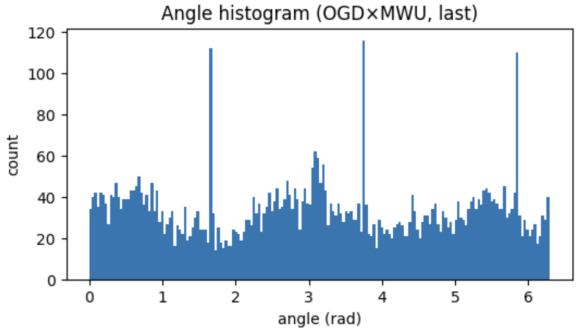


Figure 7: Angle histogram for the last-iterate InPCA trajectory in RPS under OGD×MWU. Three dominant peaks in angle indicate three preferred directions around the cycle, consistent with the triangle-like orbit and a three-vertex turning structure in the embedded space.

Turning points align with simplex vertices under OGD. We make the link between geometry and probability space explicit by overlaying the largest curvature values κ_t on the InPCA trajectories (Figure 9).

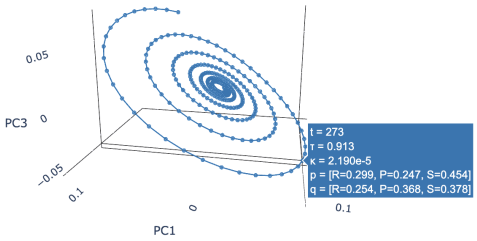


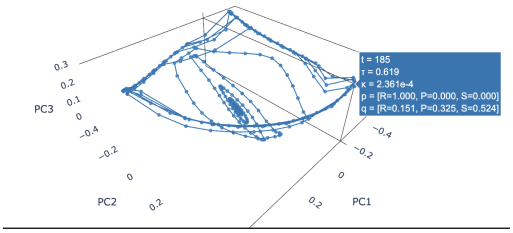
Figure 8: MWU×MWU: representative high-curvature timestep. Even at the largest observed κ_t, the trajectory remains in the interior of the simplex and the embedded path is smooth, with no polygonal corners, contrasting with the OGD cases.

For MWU×MWU (Figure 8), the highlighted point on the outer spiral has very small curvature ($\kappa_t \approx 2 \times 10^{-5}$) and both strategies remain strictly interior (e.g., $p_t \approx (0.30, 0.25, 0.45)$ and $q_t \approx (0.25, 0.37, 0.38)$). The trajectory bends smoothly without any abrupt change in direction, consistent with the continuously differentiable multiplicative update.

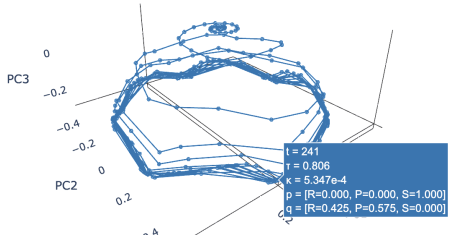
In contrast, the OGD-involving runs exhibit curvature peaks exactly when the OGD player is driven to the boundary. In the OGD×MWU case (left panel of Figure 9), a representative turning point with elevated curvature ($\kappa_t \approx 2.4 \times 10^{-4}$) occurs when the OGD player is essentially pure, $p_t \approx (1, 0, 0)$, while the MWU opponent remains mixed. For OGD×OGD (right panel of Figure 9), the largest curvature peak ($\kappa_t \approx 5.3 \times 10^{-4}$) occurs when one player is at a vertex, $p_t \approx (0, 0, 1)$, and the other lies on a simplex edge, $q_t \approx (0.43, 0.58, 0)$. These events correspond to the visually sharp corners of the polygonal orbit. Together, they support the projection-based mechanism: whenever OGD hits a face or vertex, the change in active constraints produces a discrete shift in update direction, which appears in the InPCA embedding as a high-curvature turning point and, at the level of the full cycle, as the polygonal edges and vertices characteristic of projection-driven dynamics.

7 Mechanistic Interpretation

Why MWU produces smooth orbits. MWU can be viewed as a discrete-time mirror-descent update with entropic regularization on the simplex. Each step rescales action probabilities multiplicatively and then renormalizes, so the iterate remains in the relative interior of the simplex and the update map $(p_t, q_t) \mapsto (p_{t+1}, q_{t+1})$ is smooth in all coordinates. The induced vector field in strategy space is therefore continuous, and for small step sizes its discrete trajectory is a smooth discretization



(a) OGD×MWU: high curvature near a simplex vertex.



(b) OGD×OGD: polygonal corner on the outer orbit.

Figure 9: Curvature-based turning points for InPCA trajectories involving OGD. In both pairings, large curvature κ_t occurs when the OGD player is near a simplex vertex (almost pure strategy), consistent with the projection-driven polygon hypothesis.

of the corresponding continuous-time flow. Since the Hellinger lift $P_t \mapsto \sqrt{P_t}$ and the subsequent InPCA projection are linear/smooth transformations of the joint-play distribution, they preserve this regularity, yielding the visually smooth, ring-like orbits seen in the embedded space.

Why OGD produces polygonal cycles. In contrast, OGD performs an unconstrained Euclidean gradient step followed by projection back onto the simplex,

$$p_{t+1} \;=\; \Pi_\Delta\big(p_t - \eta \nabla_p \ell_t\big), \qquad q_{t+1} \;=\; \Pi_\Delta\big(q_t - \eta \nabla_q \ell_t\big).$$

The projection operator Π_Δ is piecewise affine: on each region with a fixed active set of constraints (which coordinates are zeroed out) it is linear, but its Jacobian changes discontinuously when the active set changes. As the iterates are driven toward faces and vertices of the simplex, these active-set switches induce abrupt changes in the update direction, so the trajectory in strategy space is a concatenation of nearly straight segments joined at sharp corners. After passing through the Hellinger lift and InPCA embedding, this piecewise-affine dynamics manifests as polygonal cycles with a small number of preferred directions. Empirically, curvature peaks along the embedded trajectory coincide with near-vertex phases in (p_t, q_t), supporting the interpretation that the observed polygonal geometry is generated by projection-induced boundary interactions under OGD.

8 Discussion

The experiments provide a coherent trajectory-level picture that complements standard regret-based guarantees. Across all RPS conditions, InPCA reveals stable last-iterate orbits in the space of joint play P_t, while the time-averaged trajectories contract toward the equilibrium region. This separation confirms that empirical convergence can coexist with highly structured recurrent motion, and that the geometry of the last iterate depends sensitively on the update rule rather than on the game alone.

The qualitative fingerprints are consistent across runs and diagnostics. Multiplicative dynamics (MWU) generate smooth ring-like orbits whose angular histograms are nearly uniform, while projected-gradient dynamics (OGD) produce polygonal cycles with a small, discrete set of preferred directions (six dominant peaks for OGD×OGD, three for OGD×MWU). Curvature-based turn detection shows that the corners of these polygons coincide with near-vertex phases in the probability simplex, supporting the projection-driven polygon hypothesis: linear gradient steps combined with piecewise Euclidean projection induce regime changes when active sets switch, yielding piecewise-linear segments stitched together at high-curvature events. MWU updates, by contrast, remain in the simplex interior and evolve smoothly in (p_t, q_t), so the Hellinger-aware embedding inherits this smoothness and produces approximately circular orbits without sharp turning points.

The ablation on PCA versus InPCA clarifies that this structure is not a visualization artifact. Applying ordinary PCA directly to P_t treats the simplex as flat Euclidean space and smears the polygonal cycles into thick, entangled spirals in which corners and symmetry are largely obscured. The Hellinger square-root lift restores the appropriate information-geometric metric, after which dimensionality reduction recovers clean, low-dimensional orbits that align with angular and curvature diagnostics.

Taken together, these observations support the view that common learning rules induce robust, algorithm-specific geometry in the space of joint play, and that information- geometric embeddings such as InPCA are an effective tool for exposing this structure in a way that can subsequently be analyzed, modeled, or exploited.

9 Future work

The experiments suggest that geometric structure is not just a visualization artifact but a potentially exploitable signal about the behavior of learning algorithms. We outline three directions that explicitly connect the observed polygonal geometry to questions of control and exploitation.

The hexagon- and triangle-like cycles induced by OGD (and OGD$\times$MWU) spend disproportionate time near a small set of "corner" phases, in a predictable order. One concrete next step is to model the last-iterate dynamics as a low-dimensional Markov chain on these corners and to study *geometric exploitability*: can a strategically chosen counter-update track the polygon and achieve average payoffs strictly better than minimax by systematically anticipating upcoming vertices? This would turn the observed edges into a predictive feature: if the opponent reveals that they are running an OGD-like rule, then the shape of their orbit may reveal when they are about to become nearly pure, and hence temporarily vulnerable to a simple best-response or timing-based attack.

A second direction is algorithm design: instead of treating polygonal cycles as an incidental pathology, we can ask how to *shape* trajectory geometry. For example, can we tune step sizes, regularizers, or projection operators to suppress corners (for better stability) or, conversely, to enforce cycles that encode a desired pattern of commitments? InPCA provides a low-dimensional feedback signal for such control problems: one could define regularizers or adaptive learning rates that penalize large curvature or high angular concentration, thereby steering the dynamics toward smoother or more spread-out orbits.

Third, it is natural to evaluate additional no-regret rules discussed in the learning-in-games literature (FTRL variants, optimistic and extragradient methods, optimistic MWU, etc.) and to test whether each induces a distinct geometric fingerprint in InPCA space across a broader class of games (e.g., Shapley, general-sum congestion or routing games). Beyond logging (p_t, q_t) and P_t, richer signals such as regret vectors, counterfactual payoffs, or response gradients could be tracked to relate trajectory geometry to quantities directly tied to stability and exploitability. Ultimately, the goal is to turn geometric diagnostics into actionable tools: given a short prefix of a trajectory, infer what update rule is being used, predict its future motion, and decide whether to exploit, stabilize, or reconfigure the interaction.

References

Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the 2018 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2703–2717, 2018.

K. N. Quinn, C. B. Clement, F. De Bernardis, M. D. Niemack, and J. P. Sethna. Visualizing probabilistic models and data with intensive principal component analysis. *Proceedings of the National Academy of Sciences*, 116(28):13762–13767, 2019. doi: 10.1073/pnas.1817218116.

Aaron Roth. Learning in games and games in learning, 2025. Draft textbook / course notes, University of Pennsylvania.