Homological Barriers to Stable Learning of First-Order Nash Equilibria

Ashkan Soleymani
Massachusetts Institute of Technology
ashkanso@mit.edu

Gabriele Farina
Massachusetts Institute of Technology
gfarina@mit.edu

Patrick Jaillet
Massachusetts Institute of Technology
jaillet@mit.edu

Georgios Piliouras
Google Deepmind
gpil@google.com

Abstract

We identify a homological barrier to stable learning of local Nash equilibria, also known as first-order Nash equilibria (FONE), in constrained non-concave differentiable games. We consider continuous deterministic learning dynamics that fix every FONE point and aim to attract all initial conditions to the full FONE set. We show that this convergence guarantee is impossible when the target set has nontrivial positive-degree cohomology, such as that arising from a loop, and the state space is compact and contractible. We exhibit this barrier in a simple non-concave quadratic zero-sum game on a box whose exact FONE set consists of a diagonal six-edge loop and an isolated point. Consequently, no such dynamics can make the full FONE set a stable global attractor. We further show that degree two is the minimal polynomial degree at which this barrier can arise and extend the construction to obtain nontrivial homology in arbitrarily high degrees. Finally, we establish that the barrier persists for sufficiently small approximate-FONE targets under small perturbations, yielding an open, positive-volume family of quadratic zero-sum games with the same homological barrier.

1 Introduction

First-order equilibrium conditions are the workhorse of learning in non-concave games. When global Nash equilibria are difficult to compute or may fail to exist, the standard fallback is to seek a first-order Nash equilibrium (FONE), where each player satisfies the first-order optimality conditions for feasible unilateral deviations. In differentiable games with constrained strategy sets, this means that no player can improve their payoff to first order along a feasible direction. Although this condition is weaker than requiring each player to be a global best response, it provides a natural solution concept for studying the behavior and convergence of learning dynamics. FONE and related first-order stationarity notions are widely studied as solution targets and algorithmic benchmarks in the min-max and non-concave game literature [Daskalakis and Panageas, 2018, Adolphs et al., 2019, Nouiehed et al., 2019, Mazumdar et al., 2020a, Daskalakis et al., 2021, Ostrovskii et al., 2021, Daskalakis, 2022, Rockafellar, 2024, Bernasconi et al., 2024, Bernasconi and Castiglioni, 2026].

We identify a topological barrier to learning FONE. The barrier is not algorithm-specific or computational; it arises from the topology of the target set. If the FONE set has nontrivial first homology, as in the presence of a nontrivial loop, then no continuous learning rule can both fix every FONE point and make the full FONE set a stable global attractor on a compact contractible state space. In our construction, projection is the source of this phenomenon. On the boundary of a box, normal-cone conditions permit outward-pointing gradients and can create cycles of projected stationary points, even for affine fields. We realize such a cycle in an explicit quadratic zero-sum game on a box. The

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

resulting FONE set consists of a diagonal boundary cycle and an isolated point; the cycle creates the topological barrier, while the isolated point contributes no first homology. Thus, even simple non-concave quadratic games can defeat the usual pointwise-stationary global-convergence template. Our findings are in line the "game dynamics as the meaning of a game" perspective [Papadimitriou and Piliouras, 2019], showing that understanding the long-run behavior of learning dynamics requires looking beyond the existence of equilibrium solutions.

Our contributions can be summarized as follows.

- We formulate a general homological barrier to stable learning of FONE. Specifically, a compact target with nontrivial positive-degree Čech cohomology cannot be fixed pointwise and uniformly attracting under continuous dynamics on a compact contractible state space. This barrier applies to autonomous maps, time-varying deterministic maps, continuous-time semiflows through their time maps, and finite-memory methods with compact state spaces.
- We construct a quadratic zero-sum game on $[0,1]^3 \times [0,1]^3$ whose exact FONE set consists of a diagonal six-edge cycle and an isolated point. The cycle gives the FONE set nontrivial first cohomology, and this triggers the barrier.
- We prove that affine zero-sum games on boxes have contractible FONE sets, making degree two the minimum polynomial degree at which the barrier can arise. We further construct quadratic games whose FONE sets have nontrivial homology in arbitrarily high degrees. Finally, we show that the barrier remains under sufficiently small approximate-FONE targets and small C^1 perturbations, and identify an open, positive-volume family of quadratic zero-sum games with this property.

2 Related Work

Classical work on attractor stability and shape theory relates the topology of attracting sets to that of the ambient state space [La Salle, 1976, Hastings, 1978, Garay, 1991, Kapitanski and Rodnianski, 2000, Bhatia and Szegö, 2002]. In particular, a compact set that is fixed pointwise and uniformly attracting under continuous dynamics on a compact contractible state space must have trivial positive-degree Čech cohomology. This restriction provides the topological foundation for our analysis of stable learning dynamics and FONE sets in constrained games.

The closest game-dynamics antecedent is the Conley-theoretic impossibility theorem of Milionis et al. [2023]. Their barrier arises from a *degenerate finite game* whose mixed-strategy Nash equilibrium set is homeomorphic to a circle. More recently, Biggar et al. [2026] investigate the computational difficulty of implementing Nash-convergent dynamics in nondegenerate finite games. In our setting, the nontrivial cohomology has a different origin. It arises from the *projected first-order stationarity* conditions in smooth constrained games, where the normal-cone geometry of a box can produce cycles of FONE points even for affine fields associated with quadratic zero-sum payoffs.

Beyond the topology of equilibrium sets, a complementary line of work studies the complex long-run behavior of learning dynamics. This includes Poincaré recurrence in regularized learning for zero-sum games [Piliouras and Shamma, 2014, Mertikopoulos et al., 2018], nonconvergence of multiplicative weights [Bailey and Piliouras, 2018], and recurrence in replicator dynamics [Boone and Piliouras, 2019]. Going further, P. Andrade et al. [2023] show that no-regret learning can simulate universal computation. These results illustrate that the existence of equilibria alone provides limited insight into the long-run behavior of learning dynamics.

Another line of work studies the informational limitations of equilibrium learning. In uncoupled dynamics, where each player observes only its own payoff function and the history of play, convergence to Nash equilibrium cannot be guaranteed across all games, even when the equilibrium is unique [Hart and Mas-Colell, 2003]. Subsequent work investigates equilibrium learning under decentralized or payoff-based information, including the role of stochastic dynamics and memory [Hart and Mas-Colell, 2006, Babichenko, 2012, Jaggard et al., 2014, Toonsi and Shamma, 2023].

Complementing these dynamical and informational results, a separate line of work investigates the computational complexity of equilibrium problems. The introduction of PPAD by Papadimitriou [1992] provided a framework for studying fixed-point computation, and Nash equilibrium computation was subsequently shown to be PPAD-complete even in two-player finite games [Daskalakis et al., 2009, Chen et al., 2009]. For nonconvex–nonconcave min-max optimization, Daskalakis et al. [2021]

proved PPAD-hardness under coupled constraints, while Bernasconi and Castiglioni [2026] extended these results to product constraints, including hypercubes. Recent work has further shown that quadratic objectives already suffice for such hardness. In particular, Anagnostides et al. [2025] proved PPAD-completeness for symmetric first-order equilibria in symmetric quadratic min-max problems, and later Bernasconi et al. [2026] proved PPAD-hardness for computing approximate first-order stationary points of quadratic min-max objectives over hypercubes.

These works focus on whether an equilibrium can be computed efficiently. Our barrier, by contrast, concerns the possibility of stable global convergence, independently of computational resources. Even for a non-concave zero-sum game with quadratic payoffs and product hypercube constraints, no continuous learning dynamics that fix every FONE point can make the full FONE set a stable global attractor. Further related work is discussed in Section A.

3 Local Nash (FONE) and Projected Stationarity

We consider a two-player differentiable game $G = (u_1, u_2)$ with compact convex feasible sets $\mathcal{X} \subset \mathbb{R}^m$ and $\mathcal{Y} \subset \mathbb{R}^n$, and differentiable payoff functions $u_1, u_2 : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$. Both players are maximizers. We are particularly interested in the *zero-sum* games where $u_2 = -u_1$.

Definition 3.1 (Local Nash equilibrium)**.** *Fix any step size $\eta > 0$. A profile $(x, y) \in \mathcal{X} \times \mathcal{Y}$ is a* local Nash equilibrium *(projected first-order Nash equilibrium) if*

$$x = \mathrm{Proj}_{\mathcal{X}}\big(x + \eta \nabla_x u_1(x, y)\big), \qquad y = \mathrm{Proj}_{\mathcal{Y}}\big(y + \eta \nabla_y u_2(x, y)\big). \tag{1}$$

The set of all such profiles is denoted by $\mathrm{FONE}(G)$*, or equivalently* $\mathrm{FONE}(u_1, u_2)$.

The choice of $\eta > 0$ does not affect the exact solution set. Indeed, for any closed convex set K, the projection identity $z = \mathrm{Proj}_K(z + \eta g)$ is equivalent to the variational inequality

$$\langle g, w - z \rangle \le 0 \qquad \text{for all } w \in K. \tag{2}$$

Equivalently, $g \in N_K(z)$, where $N_K(z) = \{v : \langle v, w - z \rangle \le 0 \text{ for all } w \in K\}$ is the outward normal cone of K at z. Since $N_K(z)$ is a cone, multiplying g by any positive scalar leaves the exact first-order condition unchanged. The step size η nevertheless matters when defining approximate projected residuals. For box constraints, the variational inequality admits the following coordinatewise characterization.

Lemma 3.2. *Let $s \in [a, b]$, $g \in \mathbb{R}$, and $\eta > 0$. Then*

$$s = \mathrm{Proj}_{[a,b]}(s + \eta g)$$

if and only if

$$s = a \Rightarrow g \le 0, \qquad a < s < b \Rightarrow g = 0, \qquad s = b \Rightarrow g \ge 0. \tag{3}$$

At an interior coordinate, the corresponding component of the player's payoff gradient must vanish, while at a boundary it may point outward from the feasible interval. We apply Lemma 3.2 coordinatewise to characterize the exact FONE sets throughout the paper.

We also define the max projected residual

$$\mathcal{R}_\eta^G(x, y) = \max\left\{ \left\| x - \mathrm{Proj}_{\mathcal{X}}\big(x + \eta \nabla_x u_1(x, y)\big) \right\|_2, \left\| y - \mathrm{Proj}_{\mathcal{Y}}\big(y + \eta \nabla_y u_2(x, y)\big) \right\|_2 \right\}. \tag{4}$$

Then $\mathrm{FONE}(G) = \{(x, y) : \mathcal{R}_\eta^G(x, y) = 0\}$, and the ε-approximate first-order Nash set is $\mathrm{AFONE}_\varepsilon(G) = \{(x, y) \in \mathcal{X} \times \mathcal{Y} : \mathcal{R}_\eta^G(x, y) \le \varepsilon\}$. We write $\mathrm{PStat}(G) = \{z \in K : z = \mathrm{Proj}_K(z + \eta G(z))\}$ for the projected stationary set of G on K, equivalently $G(z) \in N_K(z)$.

4 General Homological Barrier to Learning Dynamics

We begin with a general topological barrier for continuous deterministic learning dynamics, with applications beyond games, including optimization and fixed-point methods. Throughout, $\mathbb{F}$ denotes a fixed field, and $\check{H}^q(\cdot; \mathbb{F})$ denotes Čech cohomology. We use Čech cohomology to accommodate potentially irregular attracting sets. For compact polyhedra, including the FONE sets constructed below, it agrees with singular cohomology, while its continuity property over neighborhoods plays a central role in our proof [Spanier, 1966, Mardešic and Segal, 1982].

Table 1: Algorithmic interpretations of the master barrier in Theorem 4.2.

Algorithmic family	State space	Covered?	Reference
Continuous-time gradient play	K	yes	Semiflow (Corollary 4.5).
Projected GDA	K	yes	(Corollary 4.5).
Extragradient / optimistic methods	K or $K \times M$	yes	Finite-memory (Corollary 4.6).
Heavy-ball / momentum methods	$K \times M$	yes	Contractible memory (Corollary 4.6).
Time-varying step sizes	K	yes	Nonautonomous (Corollary 4.5).
Discontinuous or randomized methods	varies	no	(Remark 4.7)
Algorithms moving along $\mathrm{FONE}(G)$	varies	no	(Remark 4.7)

Definition 4.1. *Let K be a compact metric space, let $A \subseteq K$ be compact, and let $(T_t)_{t\geq 0}$ be a sequence of continuous maps $T_t : K \to K$. Define*

$$\Psi_N = T_{N-1} \circ \cdots \circ T_0.$$

We say that $(T_t)_{t\geq 0}$ is pointwise-fixed and uniformly attracting to A *if:*

1. *$T_t(a) = a$ for every $a \in A$ and every $t \geq 0$; and*
2. *for every open neighborhood $U \supseteq A$, there exists N such that $\Psi_N(K) \subseteq U$.*

Both conditions in Definition 4.1 have natural interpretations for learning dynamics. The pointwise-fixed condition captures the usual stationarity requirement for an equilibrium solver, ensuring that once an iterate reaches a solution $a \in A$, all subsequent updates leave it unchanged. The second condition requires the dynamics to approach the target set A uniformly over all initial conditions. The formulation includes autonomous iterations, corresponding to $T_t = T$ for all t, and continuous-time semiflows $(\Phi_t)_{t\geq 0}$ through a fixed time map Φ_T. The topological barrier in the following theorem therefore applies to a broad class of learning algorithms, both in games and beyond, as discussed in Corollaries 4.4 to 4.6 and summarized in Table 1.

Theorem 4.2 (Master Barrier Theorem)**.** *Let K be compact and contractible, and let $A \subseteq K$ be compact. If*

$$\check{H}^q(A; \mathbb{F}) \neq 0$$

for some $q \geq 1$, then there is no sequence of continuous maps $T_t : K \to K$ that is pointwise-fixed and uniformly attracting to A.

Proof. Suppose, for contradiction, that such a sequence exists. Let $U \supseteq A$ be an open neighborhood. By uniform attraction, there is an integer N such that $\Psi_N(K) \subseteq U$. Since each T_t fixes A pointwise, the composition Ψ_N also fixes A pointwise.

Let $i : A \hookrightarrow K$ and $j : A \hookrightarrow U$ denote the inclusion maps. Viewing Ψ_N as a map from K to U, we have $\Psi_N \circ i = j$. By contravariant functoriality of Čech cohomology, $j^* = i^* \circ \Psi_N^*$. Since K is contractible, $\check{H}^q(K; \mathbb{F}) = 0$ for every $q \geq 1$. Therefore $j^* = 0$. Thus, for every neighborhood U of A, the canonical map $\check{H}^q(U; \mathbb{F}) \longrightarrow \check{H}^q(A; \mathbb{F})$ is zero. Since Čech cohomology of A is the direct limit of the cohomologies of its neighborhoods, $\check{H}^q(A; \mathbb{F}) \cong \varinjlim_{U \supseteq A} \check{H}^q(U; \mathbb{F})$, this forces $\check{H}^q(A; \mathbb{F}) = 0$, contradicting the hypothesis. □

The structure and proof of our barrier are closely related to classical shape-theoretic results connecting the shape of global attractors to that of the ambient phase space [Hastings, 1978, Garay, 1991, Kapitanski and Rodnianski, 2000, Milionis et al., 2023]. The following results show that, for dynamics that fix the target pointwise, Lyapunov stability together with pointwise attraction implies the uniform attraction required by Definition 4.1. Consequently, the barrier in Theorem 4.2 applies to this broader class of dynamics.

Lemma 4.3. *Let $T : K \to K$ be a continuous map on a compact metric space, and let $A \subseteq K$ be compact and invariant. Suppose that A is Lyapunov stable and pointwise attracting; that is,*

$$\mathrm{dist}(T^n z, A) \to 0 \qquad \text{for every } z \in K.$$

Then A is uniformly attracting: for every open neighborhood $U \supseteq A$, there exists N such that $T^n(K) \subseteq U$ for all $n \geq N$.

Proof. Fix an open neighborhood $U \supseteq A$. By Lyapunov stability, there exists a neighborhood $V \supseteq A$ such that every trajectory starting in V remains in U. For each $z \in K$, pointwise attraction gives an integer n_z such that $T^{n_z} z \in V$. By continuity of T^{n_z}, there is a neighborhood W_z of z such that $T^{n_z}(W_z) \subseteq V$. Compactness of K gives a finite subcover $W_{z_1}, \dots, W_{z_m}$ of K. Let $N = \max_i n_{z_i}$. If $w \in W_{z_i}$ and $n \geq N$, then $T^{n_{z_i}} w \in V$, and Lyapunov stability keeps all subsequent iterates in U. Hence $T^n(K) \subseteq U$ for all $n \geq N$. □

Corollary 4.4 (No strict Lyapunov-style certificate)**.** *Let K be compact and contractible, and let $A \subseteq K$ be compact with $\check{H}^q(A; \mathbb{F}) \neq 0$ for some $q \geq 1$. There is no continuous map $T : K \to K$ and continuous function $V : K \to \mathbb{R}_{\geq 0}$ satisfying:*

1. *$T(a) = a$ for every $a \in A$;*
2. *$V^{-1}(0) = A$; and*
3. *$V(Tz) < V(z)$ for every $z \notin A$.*

The same conclusion holds for a continuous-time semiflow whenever its time-one map satisfies the strict decrease condition. Consequently, no global strict Lyapunov argument can certify convergence to such a pointwise-fixed target.

Proof. We show that such a Lyapunov certificate would imply uniform attraction, thereby contradicting Theorem 4.2. Let $U \supseteq A$ be open. Since $V^{-1}(0) = A$ and $K \setminus U$ is compact and disjoint from A, $a_U := \min_{z \in K \setminus U} V(z) > 0$. On the compact set $\{z \in K : V(z) \geq a_U\}$, the continuous function $z \mapsto V(z) - V(Tz)$ is strictly positive. Hence it has a positive minimum, say $\delta_U > 0$. Therefore, as long as $V(T^n z) \geq a_U$, the Lyapunov value decreases by at least δ_U at the next step. Since V is bounded on the compact set K, every orbit enters $\{V < a_U\}$ after a uniformly bounded number of steps. Because $\{V < a_U\} \subseteq U$, the target A is pointwise-fixed and uniformly attracting, contradicting Theorem 4.2. □

Corollary 4.5 (Obstructed dynamics)**.** *Let K and A be as in Theorem 4.2. Then A cannot be:*

1. *the pointwise-fixed uniformly attracting target of a continuous semiflow on K;*
2. *the pointwise-fixed uniformly attracting target of an autonomous continuous map $T : K \to K$;*
3. *the pointwise-fixed uniformly attracting target of a time-varying deterministic algorithm on K;*
4. *the pointwise-fixed uniformly attracting target $B \subseteq K \times M$ of a finite-memory algorithm, whenever M is compact and contractible and $\check{H}^q(B; \mathbb{F}) \neq 0$ for some $q \geq 1$.*

In particular, if the intended memory target is $B = A \times \{m_\star\}$ with $m_\star$ as in Corollary 4.6, then $\check{H}^q(B; \mathbb{F}) \cong \check{H}^q(A; \mathbb{F})$, so adding compact contractible memory does not remove the barrier.

Corollary 4.6 (Finite-memory gradient methods)**.** *Let G be a game with compact contractible strategy space K and FONE target $A = \mathrm{FONE}(G)$. Suppose that $\check{H}^q(A; \mathbb{F}) \neq 0$ for some $q \geq 1$. Consider any deterministic finite-memory method with compact contractible memory M and continuous update rule*

$$(z_{t+1}, m_{t+1}) = T_G(z_t, m_t), \qquad (z_t, m_t) \in K \times M.$$

Let $m_\star \in M$ be a distinguished consistent-memory state. If

$$T_G(a, m_\star) = (a, m_\star) \qquad \text{for every } a \in A,$$

then the target $A \times \{m_\star\}$ cannot be uniformly globally attracting in $K \times M$.

In particular, projected GDA with finitely many auxiliary variables, extragradient and optimistic variants represented as finite-memory methods, and momentum methods cannot admit a universal Lyapunov-style convergence proof that both fixes every FONE and uniformly attracts to $\mathrm{FONE}(G) \times \{m_\star\}$ on such games.

Remark 4.7. *The barrier applies to continuous deterministic dynamics on compact contractible state spaces that fix every point of the target and attract the full target uniformly, or satisfy the corresponding Lyapunov stability and pointwise-attraction conditions. Randomized or discontinuous algorithms, methods that converge to a selected FONE point, and dynamics that move along the*

FONE set fall outside these assumptions. For finite-memory methods with full state $(z, m) \in K \times M$, where M is compact and contractible, the barrier applies when the stationary target in the extended space has the same topology as $\mathrm{FONE}(G)$, as in $\mathrm{FONE}(G) \times \{0\}$. In contrast, a method may use memory to construct an extended attractor B whose projection onto strategy space is $\mathrm{FONE}(G)$ but whose total-space topology is contractible. Such a construction falls outside the scope of the barrier.

The proof of this result and further discussion appear in Sections C.1 and C.2. In view of Theorem 4.2 and Corollary 4.6, our focus and proof strategy in the following sections is to construct games whose FONE sets have nontrivial positive-degree Čech cohomology, namely $\check{H}^q(\mathrm{FONE}; \mathbb{F}) \neq 0$ for some degree $q \geq 1$.

5 Homological Barrier for Local (First-order) Nash

In this section, we instantiate the barrier from Section 4 by constructing games whose FONE sets have nontrivial topology. Such sets cannot be stable global attractors for continuous pointwise-stationary learning dynamics.

A nonconcave-concave quadratic zero-sum game on a box can have a FONE set with nontrivial first homology, as in the presence of a loop. Consequently, its exact FONE set cannot be a stable global attractor for any continuous pointwise-stationary learning rule.

We first present a degree-6 nonconcave-nonconcave scalar game as a warm-up to illustrate the underlying geometry. We then realize the same barrier in quadratic zero-sum games, showing that it can arise even at the lowest nontrivial polynomial degree.

5.1 Warm-up: A Minimal Example with Topologically Nontrivial FONE set

We begin with a simple example that illustrates the underlying geometry. To apply the topological barrier in Theorem 4.2, we seek to construct a game whose FONE set contains a circle. However, a closed curve of interior first-order stationary points necessarily forces additional stationary points in the region it encloses. The following elementary lemma explains why such a circle cannot form the entire FONE set.

Lemma 5.1. *Let $f \in C^1(D^2)$. If $\nabla f = 0$ on ∂D^2, then either f is constant on D^2 or f has a critical point in the interior.*

Proof. Since the tangential derivative vanishes on ∂D^2, the function is constant along the boundary. If f is not constant on D^2, then either its maximum or its minimum differs from the boundary value. By compactness, such an extremum is attained in the interior, and Fermat's rule implies the existence of an interior critical point. □

In particular, this result shows that any construction that enforces stationarity along a closed curve must also include at least one FONE point in the interior of that curve. We now design such a game.

Both players choose scalar strategies in $[-2, 2]$. Define $h(x, y) = x^2 + y^2 - 1, k(x, y) = 1 + y^2 - x^2$, and consider the zero-sum game

$$u_1(x, y) = h(x, y)^2 k(x, y), \qquad u_2(x, y) = -u_1(x, y). \tag{5}$$

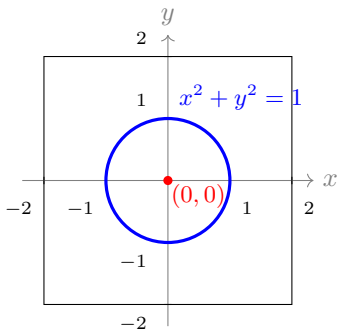
The factor h^2 enforces first-order stationarity on the unit circle $\{x^2 + y^2 = 1\}$, while the factor k eliminates all additional stationary points except for a single unavoidable interior point.

Theorem 5.2. *For the game (5),*

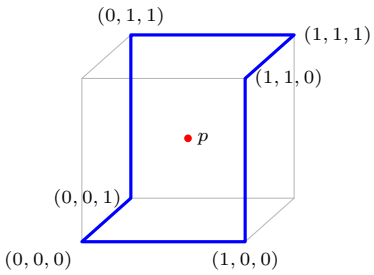
$$\mathrm{FONE}(u_1, u_2) = \{(x, y) : x^2 + y^2 = 1\} \cup \{(0, 0)\}.$$

In particular, $H_1(\mathrm{FONE}; \mathbb{Z}) \cong \mathbb{Z}$.

We defer the proof of this result to Section C.4. As a consequence of Theorem 5.2, even this scalar zero-sum example yields a FONE set with nontrivial topology as depicted in Figure 1a. Since $\check{H}^1(\mathrm{FONE}; \mathbb{F}) \neq 0$, Theorem 4.2 (with $q = 1$) rules out any continuous pointwise-stationary dynamics that globally attracts to FONE.



(a) The exact FONE set of the game (5).



(b) The quadratic projected-gradient gadget of (6).

Figure 1: Two FONE geometries, the circles carry the topological barrier. Left: the unit circle together with the isolated origin. Right: the six-edge cycle $\mathcal{C} \subset [0,1]^3$ with isolated interior zero p.

5.2 A Quadratic Zero-sum Game with Homological barrier

We now show that the topological barrier identified in Theorem 4.2 already arises in quadratic zero-sum games. The key idea is to first construct a projected-gradient gadget whose stationary set has nontrivial topology, and then embed this structure into a two-player game. This yields a quadratic zero-sum game whose FONE set inherits the same topological barrier, thereby ruling out continuous pointwise-stationary dynamics that uniformly attract to it.

5.2.1 A Quadratic Projected-Gradient Gadget

We now construct a quadratic gadget whose projected stationary set has nontrivial topology. Unlike the warm-up construction, where the stationary curve lies in the interior of the feasible domain and necessarily forces an additional critical point (Lemma 5.1), here the cycle is realized entirely on the boundary of the cube.

A natural attempt to exploit nonconcavity is to consider the gradient field induced by maximizing the convex function $\psi(x) = \|x\|^2$. However, as illustrated in Figure 3 in Section C, this does not produce the required nontrivial topology when restricted to the hypercube, since $H_1(\mathrm{PStat}(\nabla\psi); \mathbb{Z}) = 0$. Therefore, a different construction is required. Let

$$\phi(z_1, z_2, z_3) = z_1 z_2 - z_1 z_3 + z_2 z_3 - z_2, \quad \text{and} \quad G(z) = \nabla\phi(z) = \begin{pmatrix} z_2 - z_3 \\ z_1 + z_3 - 1 \\ z_2 - z_1 \end{pmatrix}. \tag{6}$$

Define the six-edge boundary cycle

$$\begin{aligned} \mathcal{C} = {} & \{(t,0,0) : 0 \le t \le 1\} \cup \{(1,t,0) : 0 \le t \le 1\} \cup \{(1,1,t) : 0 \le t \le 1\} \\ & \cup \{(t,1,1) : 0 \le t \le 1\} \cup \{(0,t,1) : 0 \le t \le 1\} \cup \{(0,0,t) : 0 \le t \le 1\}, \end{aligned} \tag{7}$$

and let $p = (1/2, 1/2, 1/2)$. Figure 1b illustrates the projection of the stationary set of the affine field $\nabla\phi(z)$ onto the hypercube.

Lemma 5.3. *For the affine field $G = \nabla\phi$ on $[0,1]^3$, the projected stationary set is exactly*

$$\mathrm{PStat}(G) = \mathcal{C} \cup \{p\}.$$

Consequently $H_1(\mathrm{PStat}(G); \mathbb{Z}) \cong \mathbb{Z}$.

Proof. Each edge of $\mathcal{C}$ satisfies the box sign conditions. For instance, on $E_1 = \{(t,0,0) : 0 \le t \le 1\}$, the first coordinate is interior and $G_1 = 0$, while the second and third coordinates are at lower endpoints with $G_2 = t - 1 \le 0$ and $G_3 = -t \le 0$. The remaining edges follow by identical checks. The point p lies in the interior and satisfies $G(p) = 0$, hence $\mathcal{C} \cup \{p\} \subseteq \mathrm{PStat}(G)$.

Let $z = (a,b,c) \in \mathrm{PStat}(G)$: i) If $0 < b < 1$, then coordinate 2 is interior, so $G_2 = 0$, i.e., $a + c = 1$. If $a, c \in (0,1)$, then $G_1 = G_3 = 0$, implying $a = b = c = 1/2$, i.e., $z = p$. If $a = 0$, then $c = 1$, giving $(0,b,1) \in E_5$; if $a = 1$, then $c = 0$, giving $(1,b,0) \in E_2$, ii) if $b = 0$, the lower-endpoint

condition gives $a + c - 1 \leq 0$. Interior conditions on a or c force $c = 0$ or $a = 0$, yielding $E_1 \cup E_6$. iii) if $b = 1$, the upper-endpoint condition gives $a + c - 1 \geq 0$. Interior conditions force $c = 1$ or $a = 1$, yielding $E_3 \cup E_4$. These cases exhaust $[0, 1]^3$, so $\mathrm{PStat}(G) = \mathcal{C} \cup \{p\}$.

Thus, the set $\mathcal{C}$ is a polygonal circle and p is isolated, so $H_1 \cong \mathbb{Z}$. □

Remark 5.4. *This gadget shows that topologically nontrivial projected stationary sets arise already in smooth quadratic optimization over a hypercube. In particular, for the quadratic function ϕ on $[0, 1]^3$, the projected stationary set is $\mathcal{C} \cup \{p\}$ and has nontrivial first homology. Therefore, by Theorem 4.2, no continuous pointwise-stationary projected dynamics can make the full projected stationary set a uniformly attracting global target. For example, projected gradient descent (PGD) is pointwise-stationary on $\mathrm{PStat}(\nabla\phi)$, but $\mathrm{PStat}(\nabla\phi)$ is not a stable global attractor. Concretely, it is easy to observe that for any neighborhood $U \supseteq \mathrm{PStat}(\nabla\phi)$, there exist initial points arbitrarily close to p whose iterates leave U under PGD. Thus, although PGD fixes every point in $\mathrm{PStat}(\nabla\phi)$, it fails to uniformly attract to the full stationary set.*

5.2.2 Diagonal Forcing and the Quadratic Zero-sum Game

We now lift the projected-gradient gadget to a two-player zero-sum game by diagonal forcing. Let $x, y \in [0, 1]^3$, and define

$$u_1(x, y) = \phi(x) + \tfrac{1}{2}\|x - y\|_2^2, \qquad u_2(x, y) = -u_1(x, y), \tag{8}$$

where ϕ is given in (6). The payoff is quadratic and zero-sum, and in contrast to the game developed in the warm-up (5), it is even nonconcave-concave. The diagonal forcing induced by $\frac{1}{2}\|x - y\|_2^2$ transfers the geometry of stationarity from the quadratic gadget in Section 5.2.1 to the FONE set of the quadratic zero-sum game. This is formalized in the following result.

Theorem 5.5 (FONE set of our quadratic zero-sum game)**.** *For the quadratic zero-sum game (8),*

$$\mathrm{FONE}(u_1, u_2) = \{(z, z) : z \in \mathcal{C}\} \cup \{(p, p)\}.$$

Consequently $H_1(\mathrm{FONE}; \mathbb{Z}) \cong \mathbb{Z}$ and, by the universal coefficient theorem, $H^1(\mathrm{FONE}; \mathbb{F}) \cong \mathbb{F}$ for every field $\mathbb{F}$.

Therefore, we can invoke the master barrier theorem from Theorem 4.2 for this quadratic game.

Corollary 5.6 (Impossibility for quadratic games)**.** *For the quadratic zero-sum game (8), no pointwise-fixed, continuous, uniformly attracting dynamics of the type covered by Theorem 4.2 and Corollary 4.5 can globally learn the exact FONE set. In particular, no Lyapunov-style argument satisfying Corollary 4.4 can establish convergence to all FONE points.*

We conclude this section by showing that degree-two zero-sum games form the minimal nonconcave class that exhibits such homological barriers for learning dynamics to find FONE.

Proposition 5.7 (Degree two is minimal)**.** *For any two-player zero-sum game on a box whose payoff has total degree at most one, the FONE set is a product of faces of the box. In particular, it is contractible and has trivial first homology.*

Full proofs of Theorem 5.5, Corollary 5.6, and Proposition 5.7 are deferred to Section C.5 for brevity.

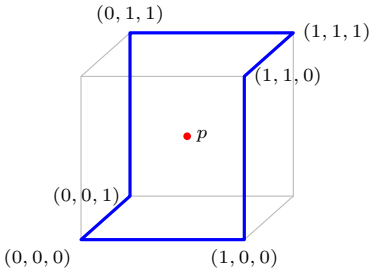
5.3 Arbitrarily High-dimensional Homological Barriers

The single-cycle construction extends to higher-dimensional topology by combining independent copies of the quadratic gadget. The resulting zero-sum games have quadratic payoffs, while their FONE sets contain higher-dimensional torus components and have nontrivial homology in arbitrarily high degrees.

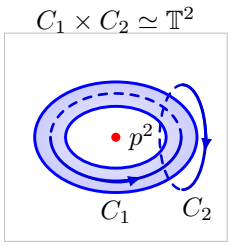
For $r \geq 1$, write $z = (z^{(1)}, \ldots, z^{(r)}) \in [0, 1]^{3r}$, with each block $z^{(j)} \in [0, 1]^3$. Define the block-separable quadratic potential

$$\Phi_r(z) = \sum_{j=1}^{r} \phi(z^{(j)}).$$

Its gradient acts independently on each block, giving the direct product of r copies of $\nabla\phi$.



(a) One block: the quadratic gadget gives the six-edge cycle $C \subset [0,1]^3$ together with an isolated point p.



(b) Two independent copies: the component $C_1 \times C_2$ is shown as a torus.

Figure 2: Product construction. One block has a cycle component C and an isolated point p. With two independent copies, the topology-carrying component becomes $C_1 \times C_2 \simeq \mathbb{T}^2$, shown schematically as a torus with a visible hole. The point $p^2 = (p, p)$ remains as a separate isolated component. The mixed circle components $C_1 \times \{p\}$ and $\{p\} \times C_2$ are omitted from the schematic.

Theorem 5.8 (Quadratic games with arbitrary homological degree)**.** *For every $r \geq 1$, the quadratic zero-sum game*

$$u_1^{(r)}(x, y) = \Phi_r(x) + \tfrac{1}{2}\|x - y\|_2^2, \qquad u_2^{(r)} = -u_1^{(r)},$$

has exact FONE set

$$\mathrm{FONE}(u_1^{(r)}, u_2^{(r)}) = \{(z, z) : z \in (\mathcal{C} \cup \{p\})^r\}. \tag{9}$$

Consequently,

$$H_r\big(\mathrm{FONE}(u_1^{(r)}, u_2^{(r)}); \mathbb{Z}\big) \neq 0, \qquad \check{H}^r\big(\mathrm{FONE}(u_1^{(r)}, u_2^{(r)}); \mathbb{F}\big) \neq 0.$$

In particular, quadratic zero-sum games realize homological barriers in arbitrarily high degree.

Proof. Since Φ_r is block-separable, projected stationarity for $\nabla\Phi_r$ decouples across blocks. By Lemma 5.3, each block has projected stationary set $\mathcal{C} \cup \{p\}$, so $\mathrm{PStat}(\nabla\Phi_r) = (\mathcal{C} \cup \{p\})^r$. Applying Lemma C.1 in dimension $3r$ yields (9). The subset $\mathcal{C}^r \subset (\mathcal{C} \cup \{p\})^r$ is homeomorphic to the r-torus $\mathbb{T}^r$, and $H_r(\mathbb{T}^r; \mathbb{Z}) \cong \mathbb{Z}$. Since $\mathcal{C}^r$ is a connected component of the FONE set, it follows that $H_r(\mathrm{FONE}; \mathbb{Z}) \neq 0$. For compact polyhedra, singular and Čech cohomology agree, and the universal coefficient theorem over a field yields $\check{H}^r \neq 0$. □

The theorem extends the barrier in Corollary 5.6 to arbitrarily high cohomological degrees, with the obstruction arising in degree r. A schematic illustration of this construction for the degree $r = 2$ is given in Figure 2.

5.4 Robust Approximate barrier

We now show that the topological barrier extends to approximate-FONE targets and small perturbations of the game. For sufficiently small residual tolerances, the approximate-FONE set contains the core cycle and remains within a regular neighborhood where the cycle is cohomologically essential. We then prove that the barrier also holds under small C^1 perturbations, covering an open, positive-volume family of quadratic games.

Consider the quadratic game $G_\star$ defined in (8), with projected residual $\mathcal{R}_\eta^\star$. Let

$$\Delta\mathcal{C} = \{(z, z) : z \in \mathcal{C}\}$$

denote the diagonal copy of the cycle. Choose a tubular neighborhood T of $\Delta\mathcal{C}$ and a small closed ball B around (p, p), disjoint from T, and set $O = \mathrm{int}(T \cup B)$. Define the residual gap

$$m_O := \min_{w \in K \setminus O} \mathcal{R}_\eta^\star(w) > 0, \qquad K = [0,1]^3 \times [0,1]^3.$$

The strict positivity follows from the continuity of $\mathcal{R}_\eta^\star$, the compactness of $K \setminus O$, and the fact that its zero set is precisely $\mathrm{FONE}(G_\star) \subset O$. We use this gap to show that sufficiently small residual sublevel sets remain confined to O, where the diagonal cycle continues to provide a topological barrier.

Theorem 5.9 (Approximate FONE and one-cycle)**.** *For every $0 < \varepsilon < m_O$,*

$$\Delta\mathcal{C} \subseteq \mathrm{AFONE}_\varepsilon(G_\star) \subseteq O.$$

Moreover, $\check{H}^1(\mathrm{AFONE}_\varepsilon(G_\star); \mathbb{F}) \neq 0$. Consequently, $\mathrm{AFONE}_\varepsilon(G_\star)$ cannot be a pointwise-fixed uniformly attracting target on the contractible strategy cube.

In turn, we show that our homological barrier extends to perturbed versions of the games we construct; consequently, it applies to a positive-volume set of quadratic games.

Theorem 5.10. *Let $\widetilde{G}$ be a game whose own-gradients are δ-close to those of $G_\star$, and let $\tau = \eta\delta$. If $0 < \tau < m_O/4$ and $\tau \leq \varepsilon < m_O - \tau$, then $\Delta\mathcal{C} \subseteq \mathrm{AFONE}_\varepsilon(\widetilde{G}) \subseteq O$, and $\check{H}^1(\mathrm{AFONE}_\varepsilon(\widetilde{G}); \mathbb{F}) \neq 0$. Hence these approximate targets cannot be pointwise-fixed uniformly attracting targets.*

Corollary 5.11. *Fix any finite-dimensional parameterization of quadratic zero-sum games containing $G_\star$. There exists an open neighborhood $\mathcal{U}$ of $G_\star$ such that for every $\widetilde{G} \in \mathcal{U}$, there is a nonempty interval $I_{\widetilde{G}}$ with*

$$\check{H}^1(\mathrm{AFONE}_\varepsilon(\widetilde{G}); \mathbb{F}) \neq 0 \quad \textit{for all } \varepsilon \in I_{\widetilde{G}}.$$

Consequently, sufficiently small approximate-FONE targets fail to be uniformly attracting on an open, positive-volume set of quadratic games.

These results establish that the barrier is robust to both approximation and small perturbations. In particular, it is not a pathological artifact of a single construction, but instead characterizes an open set of quadratic zero-sum games of positive measure.

6 Conclusion

We showed that stable global learning of first-order Nash equilibria faces a topological barrier even in quadratic zero-sum games. Starting from an affine projected-gradient field on a cube with a stationary six-edge cycle, we construct a quadratic zero-sum game whose FONE set consists of this cycle and an isolated point. The nontrivial topology of this set rules out continuous pointwise-stationary dynamics that make the entire FONE set a stable global attractor on a compact contractible state space. We further show that the barrier extends to sufficiently small approximate-FONE targets and nearby games. More broadly, our results demonstrate how projection and boundary normal cones can give rise to homologically nontrivial stationary sets even for affine fields. This geometric mechanism may also have implications for constrained nonconvex optimization, variational inequalities, complementarity problems, dynamical systems, and multi-agent control.

Acknowledgement

A.S. and P.J. were partially supported by the ONR grant N00014-24-1-2470. Part of this research was performed while A.S. was visiting the Institute for Mathematical and Statistical Innovation (IMSI), which is supported by the National Science Foundation (Grant No. DMS-2425650). G.F. was supported in part by the National Science Foundation awards CCF-2443068 and IIS-2552046, the Office of Naval Research grant N000142512296, and an AI2050 Early Career Fellowship.

References

Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. *Advances in neural information processing systems*, 31, 2018.

Leonard Adolphs, Hadi Daneshmand, Aurelien Lucchi, and Thomas Hofmann. Local saddle point optimization: A curvature exploitation approach. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 486–495. PMLR, 2019.

Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D Lee, and Meisam Razaviyayn. Solving a class of non-convex min-max games using iterative first order methods. *Advances in Neural Information Processing Systems*, 32, 2019.

Eric Mazumdar, Lillian J Ratliff, and S Shankar Sastry. On gradient-based learning in continuous games. *SIAM Journal on Mathematics of Data Science*, 2(1):103–131, 2020a.

Constantinos Daskalakis, Stratis Skoulakis, and Manolis Zampetakis. The complexity of constrained min-max optimization. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1466–1478, 2021.

Dmitrii M Ostrovskii, Andrew Lowy, and Meisam Razaviyayn. Efficient search of first-order nash equilibria in nonconvex-concave smooth min-max problems. *SIAM Journal on Optimization*, 31 (4):2508–2538, 2021.

Constantinos Daskalakis. Non-concave games: A challenge for game theory's next 100 years. *Cowles Preprints*, page 4, 2022.

R Tyrrell Rockafellar. Generalized nash equilibrium from a robustness perspective in variational analysis. *Set-Valued and Variational Analysis*, 32(2):19, 2024.

Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Gabriele Farina. On the role of constraints in the complexity of min-max optimization. *arXiv preprint arXiv:2411.03248*, 2024.

Martino Bernasconi and Matteo Castiglioni. The complexity of min-max optimization with product constraints. *arXiv preprint arXiv:2602.04665*, 2026.

Christos Papadimitriou and Georgios Piliouras. Game dynamics as the meaning of a game. *ACM SIGecom Exchanges*, 16(2):53–63, 2019.

Joseph P La Salle. *The stability of dynamical systems*. SIAM, 1976.

Harold M Hastings. Shape theory and dynamical systems. *Nelson G. Markley Department of Mathematics University of Maryland College Park, MD 20742/USA*, page 150, 1978.

Barnabas Garay. Strong cellularity and global asymptotic stability. *Fundamenta Mathematicae*, 138 (2):147–154, 1991.

Lev Kapitanski and Igor Rodnianski. Shape and morse theory of attractors. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 53 (2):218–242, 2000.

Nam Parshad Bhatia and Giorgio P Szegö. *Stability theory of dynamical systems*. Springer Science & Business Media, 2002.

Jason Milionis, Christos Papadimitriou, Georgios Piliouras, and Kelly Spendlove. An impossibility theorem in game dynamics. *Proceedings of the National Academy of Sciences*, 120(41): e2305349120, 2023.

Oliver Biggar, Christos Papadimitriou, and Georgios Piliouras. Nash-convergence of game dynamics and complexity. *arXiv preprint arXiv:2602.16016*, 2026.

Georgios Piliouras and Jeff S Shamma. Optimization despite chaos: Convex relaxations to complex limit sets via poincaré recurrence. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 861–873. SIAM, 2014.

Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the twenty-ninth annual ACM-SIAM symposium on discrete algorithms*, pages 2703–2717. SIAM, 2018.

James P Bailey and Georgios Piliouras. Multiplicative weights update in zero-sum games. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 321–338, 2018.

Victor Boone and Georgios Piliouras. From darwin to poincaré and von neumann: Recurrence and cycles in evolutionary and algorithmic game theory. In *International Conference on Web and Internet Economics*, pages 85–99. Springer, 2019.

Gabriel P. Andrade, Rafael Frongillo, and Georgios Piliouras. No-regret learning in games is turing complete. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 111–111, 2023.

Sergiu Hart and Andreu Mas-Colell. Uncoupled dynamics do not lead to nash equilibrium. *American Economic Review*, 93(5):1830–1836, 2003.

Sergiu Hart and Andreu Mas-Colell. Stochastic uncoupled dynamics and nash equilibrium. *Games and economic behavior*, 57(2):286–303, 2006.

Yakov Babichenko. Completely uncoupled dynamics and nash equilibria. *Games and Economic Behavior*, 76(1):1–14, 2012.

Aaron D Jaggard, Neil Lutz, Michael Schapira, and Rebecca N Wright. Self-stabilizing uncoupled dynamics. In *International Symposium on Algorithmic Game Theory*, pages 74–85. Springer, 2014.

Sarah Toonsi and Jeff Shamma. Higher-order uncoupled dynamics do not lead to nash equilibrium-except when they do. *Advances in Neural Information Processing Systems*, 36:30267–30285, 2023.

Christos H Papadimitriou. On inefficient proofs of existence and complexity classes. In *Annals of Discrete Mathematics*, volume 51, pages 245–250. Elsevier, 1992.

Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a nash equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.

Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *Journal of the Association for Computing Machinery*, 56(3), 2009.

Ioannis Anagnostides, Ioannis Panageas, Tuomas Sandholm, and Jingming Yan. The complexity of symmetric equilibria in min-max optimization and team zero-sum games. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Alexandros Hollender. The complexity of min-max optimization for quadratic polynomials. *arXiv preprint arXiv:2606.17000*, 2026.

Edwin H Spanier. Algebraic topology. *Springer Books*, 1966.

Sibe Mardešic and Jack Segal. *Shape theory: the inverse system approach*, volume 26. Elsevier, 1982.

Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, Thanasis Lianeas, Panayotis Mertikopoulos, and Georgios Piliouras. No-regret learning and mixed nash equilibria: They do not mix. *Advances in Neural Information Processing Systems*, 33:1380–1391, 2020.

Ya-Ping Hsieh, Panayotis Mertikopoulos, and Volkan Cevher. The limits of min-max optimization algorithms: Convergence to spurious non-critical sets. In *International Conference on Machine Learning*, pages 4337–4348. PMLR, 2021.

Alistair Letcher. On the impossibility of global convergence in multi-loss optimization. In *International Conference on Learning Representations*, 2021.

Gerasimos Palaiopanos, Ioannis Panageas, and Georgios Piliouras. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. *Advances in Neural Information Processing Systems*, 30, 2017.

David Balduzzi, Sebastien Racaniere, James Martens, Jakob Foerster, Karl Tuyls, and Thore Graepel. The mechanics of n-player differentiable games. In *International Conference on Machine Learning*, pages 354–363. PMLR, 2018.

Sai Ganesh Nagarajan, David Balduzzi, and Georgios Piliouras. From chaos to order: Symmetry and conservation laws in game dynamics. In *International Conference on Machine Learning*, pages 7186–7196. PMLR, 2020.

Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. The route to chaos in routing games: When is price of anarchy too optimistic? *Advances in Neural Information Processing Systems*, 33:766–777, 2020.

Eric Mazumdar, Lillian J Ratliff, Michael I Jordan, and S Shankar Sastry. Policy-gradient algorithms have no guarantees of convergence in linear quadratic games. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, pages 860–868, 2020b.

Jakub Bielawski, Thiparat Chotibut, Fryderyk Falniowski, Grzegorz Kosiorowski, Michał Misiurewicz, and Georgios Piliouras. Follow-the-regularized-leader routes to chaos in routing games. In *International Conference on Machine Learning*, pages 925–935. PMLR, 2021.

Georgios Piliouras and Fang-Yi Yu. Multi-agent performative prediction: From global stability and optimality to chaos. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 1047–1074, 2023.

Renato Paes Leme, Georgios Piliouras, Jon Schneider, Kelly Spendlove, and Song Zuo. Complex dynamics in autobidding systems. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 75–100, 2024.

R Tyrrell Rockafellar and Roger JB Wets. *Variational analysis*. Springer, 1998.

Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer, 2003.

Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

Glen E Bredon. *Sheaf theory*, volume 170. Springer Science & Business Media, 1997.

Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.

A Nonconvergence and Instability for Learning in Games

A substantial body of work in game dynamics documents failures of equilibrium convergence, even when the equilibrium notion is well defined. For FTRL-type no-regret learning in finite games, Vlatakis-Gkaragkounis et al. [2020] show that stable attracting Nash equilibria must be strict, and hence pure, ruling out stable attraction to mixed Nash equilibria. Related limitations arise in min–max optimization. Hsieh et al. [2021] show that broad classes of stochastic-approximation schemes can converge to internally chain-transitive sets containing no critical points, while Letcher [2021] construct analytic zero-sum multi-loss examples in which reasonable learning algorithms must either diverge or cycle. These works point to a common theme that natural learning rules can converge to the wrong objects, or fail to converge at all.

A related line of work investigates the recurrent and chaotic behavior of familiar learning dynamics across different game-theoretic settings. In congestion games, multiplicative weights can produce cycles and chaotic trajectories [Palaiopanos et al., 2017]. For differentiable games, decompositions into potential-like and Hamiltonian-like components help explain conservation laws and recurrent motion [Balduzzi et al., 2018, Nagarajan et al., 2020]. Similarly, no-regret and regularized learning in adversarial games can lead to sustained cycling and Poincaré recurrence [Piliouras and Shamma, 2014, Mertikopoulos et al., 2018]. Such behavior also arises in applied multi-agent settings, including performative prediction, autobidding auctions, and linear-quadratic games, where simple learning dynamics can transition from convergence to periodicity or chaos [Chotibut et al., 2020, Mazumdar et al., 2020b, Bielawski et al., 2021, Piliouras and Yu, 2023, Paes Leme et al., 2024]. Together, these findings motivate studying the long-run behavior of learning dynamics, which can reveal fundamental properties of games that are not apparent from their equilibrium sets alone.

The preceding results concern the behavior of particular learning rules. Our barrier instead arises from the topology of the target set and applies even to fully coupled continuous dynamics with complete knowledge of the game. In our quadratic zero-sum construction, the FONE set

$$\mathrm{FONE}(G_\star) = \{(z,z) : z \in \mathcal{C}\} \cup \{(p,p)\}$$

has nontrivial first homology, while the ambient strategy space is contractible. Consequently, no continuous dynamics can both fix every FONE point and uniformly attract to the entire FONE set. This limitation is independent of the update rule and does not rely on any particular cycling or divergent behavior.

B Preliminaries

In this section, we briefly review the preliminaries used throughout the paper. All homology and cohomology groups are taken with coefficients in a fixed field $\mathbb{F}$, unless integer coefficients are explicitly specified.

B.1 Projection, Normal Cones, and Box Stationarity

We use standard facts from convex analysis and variational inequalities, and refer to Rockafellar and Wets [1998], Facchinei and Pang [2003] for standard references. Let $K \subseteq \mathbb{R}^d$ be nonempty, closed, and convex. The Euclidean projection

$$\mathrm{Proj}_K(z) = \arg\min_{w \in K} \|z - w\|_2$$

exists and is unique.

Lemma B.1. *For $p \in K$ and $q \in \mathbb{R}^d$,*

$$p = \mathrm{Proj}_K(q) \quad \Longleftrightarrow \quad \langle q - p, w - p\rangle \le 0 \qquad \forall w \in K.$$

Proof. The function $F(w) = \frac{1}{2}\|q - w\|_2^2$ is convex and differentiable. A point $p \in K$ minimizes F over K if and only if every feasible directional derivative from p is nonnegative:

$$\left.\frac{d}{dt}\right|_{t=0^+} F((1-t)p + tw) \ge 0 \qquad \forall w \in K.$$

Computing the derivative gives $-\langle q - p, w - p\rangle \ge 0$, proving the claim. □

Taking $q = z + \eta g$ and $p = z$ gives the projected stationarity condition

$$z = \mathrm{Proj}_K(z + \eta g) \quad \Longleftrightarrow \quad \langle g, w - z \rangle \le 0 \qquad \forall w \in K.$$

Equivalently, $g \in N_K(z)$, where

$$N_K(z) = \{v : \langle v, w - z \rangle \le 0 \text{ for all } w \in K\}$$

is the outward normal cone. Thus interior projected stationarity requires $g = 0$, while boundary stationarity allows outward directions.

For the interval $[a, b]$, the normal cone is

$$N_{[a,b]}(s) = \begin{cases} (-\infty, 0], & s = a, \\ \{0\}, & a < s < b, \\ [0, \infty), & s = b. \end{cases}$$

Hence the scalar projected condition is equivalent to the box sign rule:

$$s = a \Rightarrow g \le 0, \qquad a < s < b \Rightarrow g = 0, \qquad s = b \Rightarrow g \ge 0.$$

Lemma 3.2. *Let $s \in [a, b]$, $g \in \mathbb{R}$, and $\eta > 0$. Then*

$$s = \mathrm{Proj}_{[a,b]}(s + \eta g)$$

if and only if

$$s = a \Rightarrow g \le 0, \qquad a < s < b \Rightarrow g = 0, \qquad s = b \Rightarrow g \ge 0. \tag{10}$$

Proof. If $a < s < b$, projection is locally the identity, so $s = \mathrm{Proj}_{[a,b]}(s + \eta g)$ forces $g = 0$. If $s = a$, the projection returns a exactly when $a + \eta g \le a$, equivalently $g \le 0$. If $s = b$, the projection returns b exactly when $b + \eta g \ge b$, equivalently $g \ge 0$. □

This boundary sign rule is the mechanism that creates cycles in our examples. In the quadratic gadget, each edge of $\mathcal{C}$ has one free coordinate with zero gradient, while the remaining coordinates are pinned at endpoints by correct normal-cone signs.

We also use the following nonexpansiveness property.

Lemma B.2. *For any nonempty closed convex set $K \subseteq \mathbb{R}^d$,*

$$\| \mathrm{Proj}_K(u) - \mathrm{Proj}_K(v) \|_2 \le \|u - v\|_2.$$

Proof. Let $p = \mathrm{Proj}_K(u)$ and $q = \mathrm{Proj}_K(v)$. By Lemma B.1,

$$\langle u - p, q - p \rangle \le 0, \qquad \langle v - q, p - q \rangle \le 0.$$

Adding gives

$$\|p - q\|_2^2 \le \langle u - v, p - q \rangle \le \|u - v\|_2 \|p - q\|_2,$$

and the result follows. □

Consequently, a uniform perturbation of at most δ in the players' own-action gradients changes the projected residual by at most $\eta\delta$ throughout the strategy box. This bound is used in Lemma C.3 and Theorem 5.10.

B.2 Homology and Cohomology

We recall the basic homology and cohomology facts used throughout the paper, following Hatcher [2002]. Singular homology allows us to characterize the topology of the constructed FONE sets, while cohomology enters the barrier theorem through its contravariance. In particular, a continuous map $f : X \to Y$ induces

$$f^* : H^q(Y; \mathbb{F}) \to H^q(X; \mathbb{F}),$$

which is the direction required by the factorization argument in Theorem 4.2.

For a circle and an isolated point, the relevant homology groups are

$$H_1(S^1;\mathbb{Z}) \cong \mathbb{Z}, \qquad H_1(\{p\};\mathbb{Z}) = 0.$$

By additivity over disjoint unions, the set $A = S^1 \sqcup \{p\}$ satisfies

$$H_1(A;\mathbb{Z}) \cong H_1(S^1;\mathbb{Z}) \oplus H_1(\{p\};\mathbb{Z}) \cong \mathbb{Z}.$$

Thus, the isolated point contributes nothing to first homology, and the nontrivial topology comes entirely from the loop.

For coefficients in a field $\mathbb{F}$, the universal coefficient theorem gives

$$H^q(X;\mathbb{F}) \cong \mathrm{Hom}_{\mathbb{F}}(H_q(X;\mathbb{F}),\mathbb{F}).$$

Hence nontrivial homology over $\mathbb{F}$ is equivalent to nontrivial cohomology over $\mathbb{F}$. In particular,

$$H_1(A;\mathbb{Z}) \cong \mathbb{Z} \quad \Longrightarrow \quad H^1(A;\mathbb{F}) \cong \mathbb{F}.$$

Homotopy invariance further implies that a deformation retract has the same homology and cohomology as the ambient space. In particular, a tubular neighborhood of a polygonal circle deformation retracts onto the circle and therefore has the same first cohomology.

Finally, the block-product construction uses the Künneth formula over a field,

$$H_n(X \times Y;\mathbb{F}) \cong \bigoplus_{i+j=n} H_i(X;\mathbb{F}) \otimes_{\mathbb{F}} H_j(Y;\mathbb{F}).$$

Applying this to the r-dimensional torus gives

$$H_r((S^1)^r;\mathbb{F}) \cong H_1(S^1;\mathbb{F})^{\otimes r} \cong \mathbb{F}.$$

B.3 Čech Cohomology

The exact FONE sets constructed in this paper are compact polyhedra, whereas approximate-FONE sets and residual sublevel sets under perturbations may have more irregular geometry. We therefore formulate the barrier for these approximate sets using Čech cohomology, following Spanier [1966], Bredon [1997], Mardešic and Segal [1982].

For a compact set $A \subseteq \mathbb{R}^d$, Čech cohomology admits the neighborhood characterization

$$\check{H}^q(A;\mathbb{F}) = \varinjlim_{U \supseteq A} H^q(U;\mathbb{F}),$$

where U ranges over open neighborhoods of A, with restriction maps as the neighborhoods shrink. Thus, a class in $\check{H}^q(A;\mathbb{F})$ can be represented by a cohomology class on some neighborhood of A, and two such representatives define the same class if their restrictions agree on a sufficiently small neighborhood.

Lemma B.3. *Let $A \subseteq \mathbb{R}^d$ be compact. If every open neighborhood $U \supseteq A$ contains a smaller neighborhood V of A such that the restriction map*

$$H^q(U;\mathbb{F}) \to H^q(V;\mathbb{F})$$

is zero, then $\check{H}^q(A;\mathbb{F}) = 0$.

Proof. Every Čech class is represented by a cohomology class on some neighborhood U of A. By assumption, the representative restricts to zero on a smaller neighborhood V. Hence every representative dies in the direct limit. □

For compact polyhedra, singular and Čech cohomology agree, so our exact FONE sets, including polygonal cycles and isolated points, can be analyzed using either theory. Approximate-FONE sets may have more irregular geometry, for which Čech cohomology is particularly useful because of its continuity over shrinking neighborhoods.

B.4 Dynamical Systems

We use standard notions of Lyapunov stability, attraction, and global attractors, following Bhatia and Szegö [2002], La Salle [1976]. The central dynamical assumption in Theorem 4.2 is uniform attraction, which requires that for every open neighborhood $U \supseteq A$, there exists an iterate that maps the entire compact state space into U.

This uniformity is essential to our proof. For each neighborhood U, the argument requires a single iterate N satisfying

$$\Psi_N(K) \subseteq U$$

simultaneously for all initial conditions. Pointwise convergence allows the required number of iterations to depend on the initial condition and therefore does not guarantee the uniform inclusion needed for the cohomological factorization argument.

B.5 Semialgebraic Regularity

A subset of $\mathbb{R}^d$ is semialgebraic if it can be described by a finite Boolean combination of polynomial equalities and inequalities. Such sets are triangulable and locally contractible [Bochnak et al., 2013]. For polynomial games on boxes, projection acts by clipping each coordinate, so the projected update is polynomial on each of finitely many clipping regions. Consequently, the squared projected residual is piecewise polynomial, and its sublevel sets, including the unperturbed approximate-FONE sets, are semialgebraic.

C Proofs and Details

This section provides supplementary details and omitted proofs from the main text.

C.1 The Master barrier Theorem and its Variants

The proof of Theorem 4.2 rests on a simple factorization argument. Let K be compact and contractible, and let $A \subseteq K$ be compact. Suppose that continuous maps $T_t : K \to K$ fix A pointwise and uniformly attract to A. For any open neighborhood $U \supseteq A$, uniform attraction provides an N such that

$$\Psi_N(K) \subseteq U, \qquad \Psi_N = T_{N-1} \circ \cdots \circ T_0.$$

Since each T_t fixes A pointwise, the inclusion $A \hookrightarrow U$ factors as

$$A \hookrightarrow K \xrightarrow{\Psi_N} U.$$

By contravariance of Čech cohomology, the induced restriction map

$$\check{H}^q(U;\mathbb{F}) \longrightarrow \check{H}^q(A;\mathbb{F})$$

therefore factors through $\check{H}^q(K;\mathbb{F}) = 0$ for every $q \geq 1$, since K is contractible. Thus every neighborhood-level cohomology class restricts trivially to A. Passing to the neighborhood direct limit gives $\check{H}^q(A;\mathbb{F}) = 0$ for all $q \geq 1$, contradicting the hypothesis of Theorem 4.2.

The argument relies on uniform attraction to choose a single iterate Ψ_N for all initial conditions. We note that pointwise convergence alone does not guarantee this factorization.

Stable attraction. The same barrier applies to stable global attraction. Suppose $T : K \to K$ is continuous, K is compact, and A is Lyapunov stable and pointwise attracting. Fix an open neighborhood $U \supseteq A$. Stability gives a neighborhood $V \supseteq A$ such that every trajectory starting in V remains in U. For each $z \in K$, pointwise attraction gives n_z with $T^{n_z} z \in V$. By continuity, a neighborhood of z also enters V at time n_z. Compactness gives a finite subcover, hence a uniform time after which every trajectory lies in U. Therefore stable pointwise attraction implies uniform attraction, and Theorem 4.2 applies.

Strict Lyapunov certificates. A strict Lyapunov certificate also implies uniform attraction on compact K. Suppose $V : K \to \mathbb{R}_{\geq 0}$ is continuous, $V^{-1}(0) = A$, and

$$V(Tz) < V(z) \qquad \forall z \notin A.$$

For any neighborhood $U \supseteq A$,

$$a_U = \min_{z \in K \setminus U} V(z) > 0.$$

On the compact set $\{z : V(z) \geq a_U\}$, the continuous function $V(z) - V(Tz)$ has a positive minimum. Hence every orbit enters $\{V < a_U\} \subseteq U$ in a uniformly bounded number of steps. Thus such a Lyapunov certificate would imply a pointwise-fixed uniformly attracting target, which is ruled out whenever $\check{H}^q(A; \mathbb{F}) \neq 0$.

C.2 Finite-memory Algorithms

A deterministic finite-memory method can be written on an extended state space

$$E = K \times M,$$

where M stores auxiliary variables such as momentum, previous gradients, optimism buffers, or adaptive statistics. If K and M are compact and contractible, then E is compact and contractible.

The barrier applies to any compact target $B \subseteq E$ with $\check{H}^q(B; \mathbb{F}) \neq 0$ that is fixed pointwise and uniformly attracting. The usual consistent-memory target is

$$B = \mathrm{FONE}(G) \times \{m_\star\},$$

where $m_\star$ is a distinguished memory state, such as zero momentum. Then

$$\check{H}^q(B; \mathbb{F}) \cong \check{H}^q(\mathrm{FONE}(G); \mathbb{F}),$$

so compact contractible memory does not remove the barrier.

This target assumption is important. A method could evade the theorem by attracting to a larger contractible set in extended state space whose projection is $\mathrm{FONE}(G)$. Such a method is not a pointwise-stationary solver for the exact FONE set in the sense studied here, because its actual attracting object does not have the topology of $\mathrm{FONE}(G)$.

Corollary 4.6 (Finite-memory gradient methods)**.** *Let G be a game with compact contractible strategy space K and FONE target $A = \mathrm{FONE}(G)$. Suppose that $\check{H}^q(A; \mathbb{F}) \neq 0$ for some $q \geq 1$. Consider any deterministic finite-memory method with compact contractible memory M and continuous update rule*

$$(z_{t+1}, m_{t+1}) = T_G(z_t, m_t), \qquad (z_t, m_t) \in K \times M.$$

Let $m_\star \in M$ be a distinguished consistent-memory state. If

$$T_G(a, m_\star) = (a, m_\star) \qquad \textit{for every } a \in A,$$

then the target $A \times \{m_\star\}$ cannot be uniformly globally attracting in $K \times M$.

In particular, projected GDA with finitely many auxiliary variables, extragradient and optimistic variants represented as finite-memory methods, and momentum methods cannot admit a universal Lyapunov-style convergence proof that both fixes every FONE and uniformly attracts to $\mathrm{FONE}(G) \times \{m_\star\}$ on such games.

Proof. The extended state space $K \times M$ is compact and contractible. The target $A \times \{m_\star\}$ has the same Čech cohomology as A. Applying Theorem 4.2 to the autonomous map T_G on $K \times M$ rules out uniform global attraction to $A \times \{m_\star\}$. The same argument applies to time-varying finite-memory updates by using the nonautonomous version of Theorem 4.2. □

C.3 Homology of the Constructions and Block Products

The warm-up FONE set and the quadratic projected-gradient gadget in Sections 5.1 and 5.2.1 both contain a circle component and an isolated point. Up to homeomorphism, they have the form

$$S^1 \sqcup \{p\}.$$

Therefore

$$H_1(S^1 \sqcup \{p\}; \mathbb{Z}) \cong H_1(S^1; \mathbb{Z}) \oplus H_1(\{p\}; \mathbb{Z}) \cong \mathbb{Z}.$$

The isolated point is homologically irrelevant in degree one; the barrier is carried by the loop.

For the quadratic gadget, the six-edge cycle $\mathcal{C}$ is a polygonal circle, so $\mathcal{C} \cong S^1$. Since

$$\mathrm{PStat}(\nabla\phi) = \mathcal{C} \cup \{p\},$$

we obtain

$$H_1(\mathrm{PStat}(\nabla\phi); \mathbb{Z}) \cong \mathbb{Z}.$$

The diagonal map $z \mapsto (z, z)$ is a homeomorphism onto its image, so the lifted zero-sum game has the same topology:

$$\mathrm{FONE}(u_1, u_2) = \{(z, z) : z \in \mathcal{C}\} \cup \{(p, p)\}.$$

Thus $\check{H}^1(\mathrm{FONE}(u_1, u_2); \mathbb{F}) \neq 0$ for every field $\mathbb{F}$.

The same calculation amplifies under products. Let $A = \mathcal{C} \cup \{p\}$. Then

$$A^r = \bigsqcup_{S \subseteq \{1,\dots,r\}} \left(\prod_{i \in S} \mathcal{C} \right) \times \left(\prod_{i \notin S} \{p\} \right).$$

The component corresponding to $S = \{1, \dots, r\}$ is

$$\mathcal{C}^r \cong (S^1)^r = \mathbb{T}^r.$$

By the Künneth formula over a field,

$$H_r(\mathbb{T}^r; \mathbb{F}) \cong H_1(S^1; \mathbb{F})^{\otimes r} \cong \mathbb{F}.$$

Hence $H_r(A^r; \mathbb{F}) \neq 0$ and $\check{H}^r(A^r; \mathbb{F}) \neq 0$. This is the topological calculation behind Theorem 5.8.

C.4 Missing Details and Proofs for Warm-up in Section 5.1

The warm-up construction tries to create a circle of stationary points in the interior. The following elementary argument explains why an additional interior critical point is unavoidable.

Let $f \in C^1(D^2)$ and suppose $\nabla f = 0$ on ∂D^2. Then the tangential derivative of f along ∂D^2 is zero, so f is constant on ∂D^2. If f is not constant on D^2, then either its maximum or minimum over the disk differs from this boundary value. That extremum is attained in the interior, and Fermat's rule gives an interior critical point.

This is why the quadratic construction moves the nontrivial cycle to the boundary of a cube. Boundary normal-cone signs can create a stationary loop without forcing the loop to be an unconstrained interior critical set.

Theorem 5.2. *For the game (5),*

$$\mathrm{FONE}(u_1, u_2) = \{(x, y) : x^2 + y^2 = 1\} \cup \{(0, 0)\}.$$

In particular, $H_1(\mathrm{FONE}; \mathbb{Z}) \cong \mathbb{Z}$.

Proof. A direct calculation yields

$$\partial_x u_1(x, y) = 2x\, h(x, y)\big(3 + y^2 - 3x^2\big), \qquad \partial_y u_2(x, y) = -2y\, h(x, y)\big(1 + 3y^2 - x^2\big). \quad (11)$$

We show that no boundary point satisfies the box sign conditions. At $x = 2$, we have $h = 3 + y^2 > 0$ and $3 + y^2 - 3x^2 = y^2 - 9 < 0$, hence $\partial_x u_1 < 0$, violating the upper-endpoint condition. At $x = -2$,

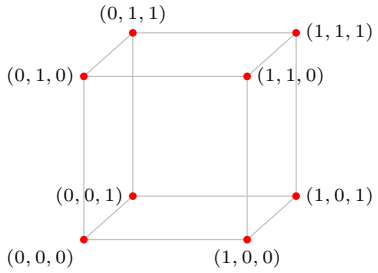


Figure 3: The projected stationary set of $\psi(x) = \frac{1}{2}\|x\|^2$ on $[0,1]^3$. In particular, we have $\mathrm{PStat}(\nabla\psi) = \{0,1\}^3$, and hence $H_1(\mathrm{PStat}(\nabla\psi);\mathbb{Z}) = 0$.

the sign reverses and violates the lower-endpoint condition. Similarly, at $y = 2$ we obtain $\partial_y u_2 < 0$, and at $y = -2$ we obtain $\partial_y u_2 > 0$, both contradictions. Hence every FONE point lies in the interior.

In turn, we know that interior FONE points satisfy $\nabla_x u_1 = 0$ and $\nabla_y u_2 = 0$, so (11) reduces to

$$x\,h(3 + y^2 - 3x^2) = 0, \qquad y\,h(1 + 3y^2 - x^2) = 0.$$

If $h = 0$, we recover the unit circle. Assume $h \neq 0$. If $x = 0$, then the second equation gives $y(1 + 3y^2) = 0$, hence $y = 0$. If $y = 0$, then the first equation gives $x(3 - 3x^2) = 0$; the solutions $x = \pm 1$ lie on $h = 0$, so off the circle only $x = 0$ remains. If $x, y \neq 0$, then

$$3 + y^2 - 3x^2 = 0, \qquad 1 + 3y^2 - x^2 = 0.$$

Substituting $x^2 = 1 + 3y^2$ into the first equation yields $-8y^2 = 0$, a contradiction. Hence the only off-circle solution is $(0, 0)$.

Thus, the FONE set is the disjoint union of a circle and a point, so $H_1(\mathrm{FONE};\mathbb{Z}) \cong \mathbb{Z}$. □

C.5 Missing Details and Proofs for Quadratic Zero-sum Game in Section 5.2.2

Lemma C.1. *For the game (8), every FONE point satisfies $x = y$. On the diagonal, player 1's FONE condition reduces to $x \in \mathrm{PStat}(\nabla\phi)$. Consequently,*

$$\mathrm{FONE}(u_1, u_2) = \{(z, z) : z \in \mathrm{PStat}(\nabla\phi)\}.$$

Proof. The own-gradients are

$$\nabla_x u_1(x, y) = \nabla\phi(x) + (x - y), \qquad \nabla_y u_2(x, y) = x - y.$$

Fix a coordinate i. Player 2's own-gradient in coordinate y_i is $x_i - y_i$. If $0 < y_i < 1$, the interior condition gives $x_i = y_i$. If $y_i = 0$, the lower-endpoint condition gives $x_i \leq 0$, and since $x_i \in [0, 1]$, this forces $x_i = 0 = y_i$. If $y_i = 1$, the upper-endpoint condition gives $x_i \geq 1$, and since $x_i \leq 1$, this forces $x_i = 1 = y_i$. Hence $x = y$. On the diagonal, $\nabla_x u_1(x, x) = \nabla\phi(x)$ and $\nabla_y u_2(x, x) = 0$, yielding the claim. □

Theorem 5.5 (FONE set of our quadratic zero-sum game)**.** *For the quadratic zero-sum game (8),*

$$\mathrm{FONE}(u_1, u_2) = \{(z, z) : z \in \mathcal{C}\} \cup \{(p, p)\}.$$

Consequently $H_1(\mathrm{FONE};\mathbb{Z}) \cong \mathbb{Z}$ and, by the universal coefficient theorem, $H^1(\mathrm{FONE};\mathbb{F}) \cong \mathbb{F}$ for every field $\mathbb{F}$.

Proof. Combine Lemma C.1 with Lemma 5.3. The diagonal map $z \mapsto (z, z)$ is a homeomorphism from $\mathcal{C} \cup \{p\}$ onto $\mathrm{FONE}(u_1, u_2)$, so the homology statement follows. Since the FONE set is a compact polyhedron, singular and Čech cohomology agree, and the cohomology statement follows from the universal coefficient theorem over a field. □

Corollary 5.6 (Impossibility for quadratic games)**.** *For the quadratic zero-sum game (8), no pointwise-fixed, continuous, uniformly attracting dynamics of the type covered by Theorem 4.2 and Corollary 4.5 can globally learn the exact FONE set. In particular, no Lyapunov-style argument satisfying Corollary 4.4 can establish convergence to all FONE points.*

Proof. The state space $[0,1]^3 \times [0,1]^3$ is compact and contractible, while Theorem 5.5 gives a FONE target with nonzero first Čech cohomology. The claim follows directly from Theorem 4.2 and Corollary 4.4, including finite-memory extensions. □

Proposition 5.7 (Degree two is minimal)**.** *For any two-player zero-sum game on a box whose payoff has total degree at most one, the FONE set is a product of faces of the box. In particular, it is contractible and has trivial first homology.*

Proof. Let $u_1(x,y) = a \cdot x + b \cdot y + c$ be affine and $u_2 = -u_1$. Player 1's own-gradient is the constant vector a, and player 2's own-gradient is the constant vector $-b$. On each coordinate, the box rule selects the upper endpoint if the gradient is positive, the lower endpoint if it is negative, and the entire interval if it is zero. Therefore, the FONE set is a Cartesian product of intervals and endpoints, i.e., a face product of the strategy box. Such a set is contractible. □

C.6 Missing Details and Proofs for Approximate FONE and Robustness in Theorem 5.9

Let

$$S = \Delta\mathcal{C} = \{(z,z) : z \in \mathcal{C}\}$$

be the diagonal cycle in the exact FONE set of the quadratic game. Let T be a tubular neighborhood of S, let B be a small ball around (p,p) disjoint from T, and set

$$O = \mathrm{int}(T \cup B).$$

The component of O containing S deformation retracts onto S.

Lemma C.2 (Cycle sandwich)**.** *Let A be compact and satisfy*

$$S \subseteq A \subseteq O.$$

Then $\check{H}^1(A;\mathbb{F}) \neq 0$.

Proof. Choose a cohomology class on O whose restriction to the tubular-neighborhood component of S generates $H^1(S;\mathbb{F})$, and whose restriction to the ball component is zero. Its restriction to S is nonzero because the tubular neighborhood deformation retracts onto S.

Suppose, for contradiction, that its image in $\check{H}^1(A;\mathbb{F})$ were zero. By the neighborhood definition of Čech cohomology, the class would restrict to zero on some open neighborhood W with

$$A \subseteq W \subseteq O.$$

But $S \subseteq A \subseteq W$, so the restriction to S would also be zero, a contradiction. Therefore $\check{H}^1(A;\mathbb{F}) \neq 0$. □

For the exact game $G_\star$, the residual vanishes exactly on $\mathrm{FONE}(G_\star)$. Since $K \setminus O$ is compact and contains no exact FONE points,

$$m_O = \min_{w \in K \setminus O} \mathcal{R}^\star_\eta(w) > 0.$$

Thus, for every $0 < \varepsilon < m_O$,

$$S \subseteq \mathrm{AFONE}_\varepsilon(G_\star) \subseteq O.$$

By Lemma C.2,

$$\check{H}^1(\mathrm{AFONE}_\varepsilon(G_\star);\mathbb{F}) \neq 0.$$

For a perturbed game $\widetilde{G}$, suppose the own-gradients are δ-close uniformly. By the residual stability estimate, the residuals are $\eta\delta$-close. Writing $\tau = \eta\delta$, the same sandwich holds whenever

$$\tau \le \varepsilon < m_O - \tau.$$

Indeed, S is contained in $\mathrm{AFONE}_\varepsilon(\widetilde{G})$ because the perturbed residual on S is at most τ, and no point outside O is ε-approximate because the perturbed residual there is at least $m_O - \tau$. Hence the approximate barrier survives small C^1 perturbations.

Consequently, the phenomenon is not a pathological artifact of one construction. In any finite-dimensional quadratic coefficient family containing $G_\star$, a sufficiently small open neighborhood of $G_\star$ consists of games whose approximate-FONE targets retain nontrivial first Čech cohomology for a nonempty interval of tolerances. We formalize this result in the following.

Theorem 5.9 (Approximate FONE and one-cycle)**.** *For every $0 < \varepsilon < m_O$,*

$$\Delta\mathcal{C} \subseteq \mathrm{AFONE}_\varepsilon(G_\star) \subseteq O.$$

Moreover, $\check{H}^1(\mathrm{AFONE}_\varepsilon(G_\star); \mathbb{F}) \neq 0$. Consequently, $\mathrm{AFONE}_\varepsilon(G_\star)$ cannot be a pointwise-fixed uniformly attracting target on the contractible strategy cube.

Proof. Since the residual vanishes on $\Delta\mathcal{C}$, we have $\Delta\mathcal{C} \subseteq \mathrm{AFONE}_\varepsilon(G_\star)$. By definition of m_O, any point outside O has residual at least $m_O > \varepsilon$, so $\mathrm{AFONE}_\varepsilon(G_\star) \subseteq O$.

Thus $\mathrm{AFONE}_\varepsilon(G_\star)$ contains the core cycle $\Delta\mathcal{C}$ and is contained in its neighborhood $T \cup B$. The inclusion of $\Delta\mathcal{C}$ into T induces a nontrivial map on first cohomology because T deformation retracts onto $\Delta\mathcal{C}$. Hence $\check{H}^1(\mathrm{AFONE}_\varepsilon(G_\star); \mathbb{F}) \neq 0$. The dynamical conclusion follows from Theorem 4.2. □

We now show that the residual remains stable under perturbations to the game.

Lemma C.3. *Let G and $\widetilde{G}$ be games on the same box. If*

$$\sup_K \|\nabla_x u_1 - \nabla_x \widetilde{u}_1\|_2 \le \delta, \qquad \sup_K \|\nabla_y u_2 - \nabla_y \widetilde{u}_2\|_2 \le \delta,$$

then $\sup_K |\mathcal{R}_\eta^G - \mathcal{R}_\eta^{\widetilde{G}}| \le \eta\delta$.

Proof. Projection onto a closed convex set is 1-Lipschitz, so each player's projected displacement changes by at most $\eta\delta$. Taking the maximum over the two players yields the claim. □

Theorem 5.10. *Let $\widetilde{G}$ be a game whose own-gradients are δ-close to those of $G_\star$, and let $\tau = \eta\delta$. If $0 < \tau < m_O/4$ and $\tau \le \varepsilon < m_O - \tau$, then $\Delta\mathcal{C} \subseteq \mathrm{AFONE}_\varepsilon(\widetilde{G}) \subseteq O$, and $\check{H}^1(\mathrm{AFONE}_\varepsilon(\widetilde{G}); \mathbb{F}) \neq 0$. Hence these approximate targets cannot be pointwise-fixed uniformly attracting targets.*

Proof. By Lemma C.3, $\|\mathcal{R}_\eta^{\widetilde{G}} - \mathcal{R}_\eta^\star\|_\infty \le \tau$. On $\Delta\mathcal{C}$, the residual is zero for $G_\star$, so $\mathcal{R}_\eta^{\widetilde{G}} \le \tau \le \varepsilon$. If $w \notin O$, then $\mathcal{R}_\eta^\star(w) \ge m_O$, hence $\mathcal{R}_\eta^{\widetilde{G}}(w) \ge m_O - \tau > \varepsilon$. The cohomology argument is identical to Theorem 5.9. □

Corollary 5.11. *Fix any finite-dimensional parameterization of quadratic zero-sum games containing $G_\star$. There exists an open neighborhood $\mathcal{U}$ of $G_\star$ such that for every $\widetilde{G} \in \mathcal{U}$, there is a nonempty interval $I_{\widetilde{G}}$ with*

$$\check{H}^1(\mathrm{AFONE}_\varepsilon(\widetilde{G}); \mathbb{F}) \neq 0 \quad \text{for all } \varepsilon \in I_{\widetilde{G}}.$$

Consequently, sufficiently small approximate-FONE targets fail to be uniformly attracting on an open, positive-volume set of quadratic games.

Proof. Coefficient closeness implies uniform closeness of gradients on a compact domain. Choosing $\mathcal{U}$ small enough ensures $\tau_{\widetilde{G}} < m_O/4$ for all $\widetilde{G} \in \mathcal{U}$. The interval

$$I_{\widetilde{G}} = [\tau_{\widetilde{G}},\ m_O - \tau_{\widetilde{G}})$$

is nonempty, and Theorem 5.10 applies. Openness implies positive Lebesgue measure. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Complete proofs are provided.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Concretely in remarks and tables.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: Complete proofs and clear assumptions are provided.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: [N/A]

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: This is primarily a theoretical study.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [No]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [No]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: [N/A]

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.