# Fast and Furious Symmetric Learning in Zero-Sum Games

**John Lazarsfeld**
SUTD

*In collaboration with:*

**Georgios Piliouras**
SUTD/DeepMind

**Ryann Sim**
SUTD

**Andre Wibisono**
Yale

[COLT 2025]

arxiv:2506.13086

# What is this talk about?

- We study **Fictitious Play** in Zero-Sum Games.
- **Main result --** new sublinear regret guarantees:

On n-dimensional **Rock-Paper-Scissors** matrices, and **using any tiebreaking rule**:

Fictitious play has $O(T^{0.5})$ regret.

(New class of matrices for which Karlin's Conjecture holds.)

# Refresher: Online Learning in Zero-Sum Games

- **Payoff matrix A** (n by n) *(e.g., three-strategy Rock-Paper-Scissors)* $A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$

- Over **T rounds**, Players 1 and 2 **play mixed strategies:**

$$x_1^t \in \Delta_n \text{ and } x_2^t \in \Delta_n.$$

(e.g., distributions over strategy set {rock, paper, scissors}).

- Player 1 payoff: $\langle x_1^t, Ax_2^t \rangle$; Player 2 payoff: $-\langle x_1^t, Ax_2^t \rangle$;

- Players observe **payoff vector feedback**:
    - Player 1 observes $g_1^t = Ax_2^t \in \mathbb{R}^n.$
    - Player 2 observes $g_2^t = -A^\top x_1^t \in \mathbb{R}^n.$

# Refresher: Online Learning in Zero-Sum Games

- Players seek to **minimize their individual regrets**:

$$\mathrm{Reg}_1(T) := \max_{x_1 \in \Delta_n} \sum_{t=0}^{T} \langle x_1, Ax_2^t \rangle - \sum_{t=0}^{T} \langle x_1^t, Ax_2^t \rangle$$

$$\mathrm{Reg}_2(T) := \sum_{t=0}^{T} \langle x_1^t, Ax_2^t \rangle - \min_{x_2 \in \Delta_n} \sum_{t=0}^{T} \langle x_2, A^\top x_1^t \rangle \ .$$

- This work: interested in **total regret** (henceforth *regret*):

$$\mathrm{Reg}(T) = \mathrm{Reg}_1(T) + \mathrm{Reg}_2(T)$$

- **Equivalence:** regret minimization <--> convergence to Nash:

> If Reg(T) = $\alpha$ = o(T), then *time-averaged iterates* converge to Nash equilibrium of A at a rate of $\alpha$/T = o(1).

# Fictitious Play in Zero-Sum Games

- **Fictitious Play (FP)** [Brown, 1950]

$$\begin{cases} x_1^{t+1} := \text{argmax}_{x \in \Delta_n} \langle x, \sum_{k=0}^{t} A x_2^k \rangle \\ x_2^{t+1} := \text{argmax}_{x \in \Delta_n} \langle x, \sum_{k=0}^{t} -A^\top x_1^k \rangle \end{cases}$$

(*Note:* argmax **always returns vertex** of $\Delta_n$ and encodes **tiebreaking**)

*Equivalently*: FP is simultaneous **Follow-the-Leader (FTL):**

$$x^{t+1} := \text{argmax}_{x \in \Delta_n} \left\langle x, \sum_{k=1}^{t} g^k \right\rangle$$

For **general adversarial rewards**: Fictitious Play / FTL fails!

- Examples where **FTL has linear regret** Reg(T) = Ω(T).
- **Issue**: iterates of FTL can lack stability (no regularization!)

**Question:** can FP still obtain sublinear regret in zero-sum game setting?

# Fictitious Play in Zero-Sum Games

**Question:** can FP still obtain sublinear regret in zero-sum game setting?

[Robinson, 1951]: **Fictitious Play has sublinear regret** in all zero-sum games, but with $Reg_{FP}(T) \leqslant O(T^{1-1/n})$.

[Karlin, 1960]: **Karlin's Conjecture:** for all zero-sum games, $Reg_{FP}(T)=O(T^{0.5})$.

Some progress, ~50 years later:

[Daskalakis-Pan, 2014]: On n-dim. identity matrix, $Reg_{FP}(T)=\Omega(T^{1-1/n})$, but **using adversarial tiebreaking.**

[Abernethy-Lai-Wibisono, 2021]: On *all* diagonal payoff matrices, $Reg_{FP}(T)= O(T^{0.5})$ using **fixed lexicographical tiebreaking.**

# Main Result: New Regret Bounds for Fictitious Play

**Theorem [LPSW, 2025]**. On class of **n-dimensional Rock-Paper-Scissors matrices**, under **symmetric learning**, and using **any tiebreaking rule:**

Fictitious Play has regret $\mathrm{Reg}_{FP}(T)=O(T^{0.5})$.

- **n-dimensional RPS** – generalizes RPS to n-dim., weighted regime:

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \quad \text{-->} \quad A = \begin{pmatrix} 0 & -a_1 & a_3 \\ a_1 & 0 & -a_2 \\ -a_3 & a_2 & 0 \end{pmatrix} \quad \text{-->} \quad A = \begin{pmatrix} 0 & -a_1 & 0 & \dots & a_n \\ a_1 & 0 & -a_2 & 0 & \dots \\ 0 & a_2 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & -a_{n-1} \\ -a_n & 0 & \dots & a_{n-1} & 0 \end{pmatrix}.$$

  Canonical symmetric zero-sum game ($A=-A^T$) (See [Hobauer-Sigmund], [Sandholm])

- **Symmetric learning** – identical initializations $x_1^0 = x_2^0 \in \Delta_n \implies x^t := x_1^t = x_2^t$
- **No tiebreaking assumption!** New class where Karlin's conjecture holds.

**Secondary Result:** same bound for Online GD with *constant stepsizes.*
(i.e., "fast and furious" regime [Bailey-Piliouras, 2019])

**Core idea**: use geometric perspective of iterates in *dual space of payoffs.*

# One slide overview of analysis

- Study cumulative **dual payoff vectors**: $y^t = \sum_{k=0}^{t-1} Ax^k \in \mathbb{R}^n$
- Equivalence between **energy function** and regret:

$$\Psi(y) = \max_{x \in \Delta_n} \langle x, y \rangle \quad \Longrightarrow \quad \mathrm{Reg}(T) = 2 \cdot \Psi(y^{T+1})$$

**Underlying property**: dual iterates of FP follow skew-(sub)gradient descent wrt energy function. (cf., [Mertikopolous+'18], [Bailey-Pilouras, '19], [Abernethy+, '21])

Dual update under Fictitious Play (with $\gamma = 1$):

$$y^{t+1} = y^t + \gamma A \partial \Psi(y^t)$$

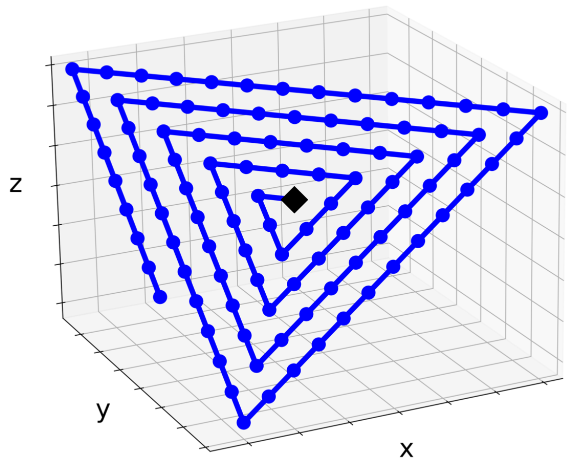
**Goal**: control (non-uniform) energy growth of dual iterates.

# One slide overview of analysis

**Goal**: control (non-uniform) energy growth of dual iterates.

- For RPS matrices, we prove a **cycling property** of dual iterates.

  *Holds regardless of tiebreaking.*




*(Dual iterates of n=3 RPS)*

- **Remaining intuition:**

Cycling --> regularity of energy growth --> $O(T^{0.5})$ total energy/regret.

# Conclusion

**Takeaway:** new evidence that *non*-no-regret algorithms can still learn (converge to Nash) in zero-sum games.

**Open:** establish same regret guarantee for FP on other classes of (symmetric) zero-sum games.

---

Thanks! Questions?

arxiv:2506.13086