Fast and Furious Symmetric Learning in Zero-Sum Games: Gradient Descent as Fictitious Play

John Lazarsfeld JOHN_LAZARSFELD@SUTD.EDU.SG
Singapore University of Technology and Design

Georgios Piliouras GPIL@GOOGLE.COM
Google DeepMind

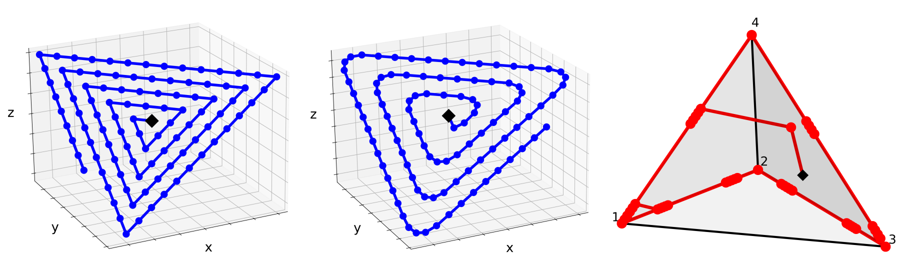
Ryann Sim RYANN_SIM@SUTD.EDU.SG
Singapore University of Technology and Design

Andre Wibisono ANDRE.WIBISONO@YALE.EDU
Yale University

Editors: Nika Haghtalab and Ankur Moitra

Abstract

This paper investigates the sublinear regret guarantees of two *non*-no-regret algorithms in zero-sum games: Fictitious Play, and Online Gradient Descent with *constant* stepsizes. In general adversarial online learning settings, both algorithms may exhibit instability and linear regret due to no regularization (Fictitious Play) or small amounts of regularization (Gradient Descent). However, their ability to obtain tighter regret bounds in two-player zero-sum games is less understood. In this work, we obtain strong new regret guarantees for both algorithms on a class of symmetric zero-sum games that generalize the classic three-strategy Rock-Paper-Scissors to a weighted, n-dimensional regime. Under *symmetric initializations* of the players' strategies, we prove that Fictitious Play with *any tiebreaking rule* has $O(\sqrt{T})$ regret, establishing a new class of games for which Karlin's Fictitious Play conjecture holds. Moreover, by leveraging a connection between the geometry of the iterates of Fictitious Play and Gradient Descent in the dual space of payoff vectors, we prove that Gradient Descent, for *almost all* symmetric initializations, obtains a similar $O(\sqrt{T})$ regret bound when its stepsize is a *sufficiently large* constant. For Gradient Descent, this establishes the first "fast and furious" behavior (i.e., sublinear regret *without* time-vanishing stepsizes) for zero-sum games larger than 2×2.



(a) (b) (c)

Figure 1: In $n = 3$ Rock-Paper-Scissors for 200 iterations and initialized at $x^0 = [1, 0, 0]$: (a) the dual iterates of Fictitious Play, and (b) the dual iterates of Gradient Descent with stepsize $\eta = 0.5$. In $n = 4$ Rock-Paper-Scissors initialized at $x^0 = [0.05, 0.35, 0.39, 0.21]$: ($c$) the primal iterates of Gradient Descent using stepsize $\eta = 1$. (◆) denotes the initial iterate of the dynamics.

Keywords: Fictitious Play, Gradient Descent, Online Learning in Zero-Sum Games

1. Introduction

In adversarial or adaptive online learning settings, regularization is the key ingredient for obtaining optimal $O(\sqrt{T})$ regret bounds over T iterations (Cesa-Bianchi and Lugosi, 2006; Hazan et al., 2016). Indeed, *un*-regularized algorithms like Follow-The-Leader (FTL) are known to be highly-sensitive to even slight oscillations in reward sequences, leading to regret scaling linearly in T in the worst case. Using online Gradient Descent (GD), an instantiation of the Follow-The-Regularized-Leader (FTRL) family, this sensitivity to oscillating rewards is circumvented via regularization, the magnitude of which is controlled via a *stepsize* parameter $\eta > 0$: smaller settings of η correspond to more regularization, and the iterates of the algorithm become more stable. Using time-horizon-dependent settings of $\eta \approx 1/\sqrt{T}$, it is well known that Gradient Descent and other instantiations of FTRL like Multiplicative Weights Update obtain optimal $O(\sqrt{T})$ regret bounds for adversarial online learning (see e.g. Shalev-Shwartz et al. (2012)).

In the setting of two-player zero-sum games, a well-known connection with online learning establishes that the time-averaged iterates of two players using *no-regret algorithms* (e.g., algorithms that guarantee *sublinear* regret in any adaptive setting) converge to a Nash Equilibrium of the underlying game (Hannan, 1957; Freund and Schapire, 1999). While using Gradient Descent or any FTRL algorithm with a time-horizon-dependent or time-decaying stepsize guarantees such convergence, in the context of modeling strategic interactions between agents, selecting stepsizes depending inversely on T is somewhat unrealistic, and it is natural to question whether algorithms for learning in games require the same amount of regularization as in worst-case, adversarial settings. For example, the Fictitious Play (FP) algorithm of Brown (1951) is the result of both players using the *un*regularized FTL algorithm. While FTL is *not* a no-regret algorithm in general, Robinson (1951) proved that the sum of the players' regrets *is* in fact sublinear in the zero-sum game setting, although scaling like $O(T^{1-1/(2n-2)})$ for $n \times n$ games. Later, Karlin (1959) conjectured that in all zero-sum games, Fictitious Play achieves an even stronger regret guarantee of only $O(\sqrt{T})$.

Karlin's conjecture was seemingly put to rest by Daskalakis and Pan (2014), who showed that even for the n-dimensional identity payoff matrix, when using a specific *adversarial* tiebreaking rule, Fictitious Play has regret scaling like $\Omega(T^{1-1/n})$. Moreover, for FTRL algorithms like Gradient Descent in the *under*-regularized regime of *constant stepsizes* (e.g., with no dependence on T), Bailey and Piliouras (2018) showed that the day-to-day iterates of these algorithms are *repelled* from interior Nash equilibria and converge towards the boundary of the simplex. Using FTRL algorithms with constant stepsizes was also shown to exhibit formally chaotic behavior, both in congestion/coordination games (Palaiopanos et al., 2017; Chotibut et al., 2020, 2021; Bielawski et al., 2021) as well as (symmetric) zero-sum games (Cheung and Piliouras, 2019, 2020; Cheung et al., 2022). In fact, in congestion/coordination games, chaotic as well as more generally unstable behavior (i.e., limit cycles) can result in provably non-vanishing regret (Blum et al., 2006; Chotibut et al., 2020). Such results seemed to indicate a much narrower path for success for Fictitious Play and other under-regularized, *non*-no-regret learning algorithms to achieve strong regret guarantees in zero-sum games.

On the other hand, several newer works have painted a more optimistic picture: for example, Bailey and Piliouras (2019a) showed that the day-to-day *non*-equilibration behavior of Gradient Descent can actually be leveraged as a *tool* for controlling the algorithm's total regret. For 2×2 zero-sum games with a unique interior equilibrium, they proved that running Gradient Descent with constant stepsizes still leads to $O(\sqrt{T})$ regret, exhibiting so-called "fast and furious" behavior (i.e., regret-minimization with constant stepsizes instead of vanishing/decreasing ones). However, their result left open the question of establishing similar regret guarantees for classes of larger, high-dimensional zero-sum games. Abernethy et al. (2021) have also provided a more nuanced view of the negative result of Daskalakis and Pan (2014) for Fictitious Play: they proved for all $n \times n$ *diagonal* payoff matrices that Fictitious Play with a fixed *lexicographical* tiebreaking rule indeed obtains $O(\sqrt{T})$ regret, thus proving (the weak version of) Karlin's Conjecture for this class.

However, little progress has since been made on obtaining similar, tighter regret bounds for either Fictitious Play or Gradient Descent in the constant stepsize regime. Despite the increasing interest in the *optimistic variants* of FTRL algorithms (which can be used to obtain last-iterate convergence in zero-sum games (Rakhlin and Sridharan, 2013; Syrgkanis et al., 2015; Cai et al., 2024)), the behavior and guarantees of their standard, non-optimistic counterparts still lacks a more fine-grained understanding, even for simplified settings such as *symmetric* zero-sum games under *symmetric learning* (i.e., when the players use the same initialization). Can Fictitious Play and FTRL with constant stepsizes minimize regret in such settings, despite their un(der)-regularization?

Our Contributions. In this paper, we make new progress towards obtaining tighter regret bounds for both Fictitious Play and Gradient Descent in zero-sum games. We study these dynamics under *symmetric learning* on a large class of *high-dimensional rock-paper-scissors (RPS)* payoff matrices (Definition 3). These games generalize the canonical three-strategy Rock-Paper-Scissors game – perhaps the most well-studied symmetric zero-sum game – to a weighted, n-dimensional regime. For *all n-dimensional RPS matrices*, we prove that under *symmetric learning*:

- (Theorem 11): Fictitious Play from any initialization obtains worst-case $O(\sqrt{T})$ regret using *any* arbitrary tie-breaking rule.
- (Theorem 17): Gradient Descent, from *almost all* initializations and with *sufficiently large constant stepsizes*, obtains $O(\sqrt{T})$ regret.

Our results provide new evidence on the robustness of these *non*-no-regret algorithms in obtaining sublinear regret in zero-sum games (and as a consequence, fast time-averaged convergence to Nash equilibria). For Fictitious Play, Theorem 11 establishes a new class of high-dimensional zero-sum games for which Karlin's Conjecture holds, and in contrast to Abernethy et al. (2021) it does not rely on using a specific tiebreaking rule. For Gradient Descent with constant stepsizes, Theorem 17 establishes the first class of *large* zero-sum games (beyond the 2×2 case studied by Bailey and Piliouras (2019a)) with provable $O(\sqrt{T})$ regret.

Our proof techniques build off of recent work identifying the Hamiltonian structure of *continuous-time* FTRL dynamics (Mertikopoulos et al., 2018; Bailey and Piliouras, 2019b; Wibisono et al., 2022). We show in the dual space of payoff vectors that the iterates of both *discrete-time* Fictitious Play and Gradient Descent can be viewed as a *skew-gradient descent* with respect to a corresponding *energy function*, which is closely related to regret. Using this geometric perspective, we establish a *cycling* property in the primal space that is shared by the iterates of both algorithms on RPS matrices. In turn, this leads to a regularity in energy growth that allows us to obtain strong sublinear regret guarantees.

Somewhat surprisingly, we show for Gradient Descent that a *sufficiently large* constant stepsize is the key driver for establishing this regularity. This comes in stark contrast to much of the conventional wisdom in online learning (where small or even vanishing stepsizes are believed necessary for obtaining low regret), as well as in games settings (where numerous results suggest that using large, constant stepsizes can lead to unpredictable, chaotic behavior). To the contrary, we show that large stepsizes are precisely the catalyst for proving the desirable "fast and furious" behavior, and they allow for viewing Gradient Descent and Fictitious Play from a shared perspective.

2. Preliminaries

Let $[n] := \{1, \ldots, n\}$. Let $\Delta_n = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i\in[n]} x_i = 1,\ x_i \geq 0 \ \forall\ i \in [n]\}$ denote the probability simplex in $\mathbb{R}^n$. For $x \in \Delta_n$, let $\text{supp}(x) = \{i \in [n] : x_i > 0\}$ denote the support of x. If $\text{supp}(x) = [n]$, then we say x is *interior*. For $x, y \in \mathbb{R}^n$, we denote the ℓ_2-inner product by $\langle x, y\rangle = x^\top y = \sum_{i=1}^n x_i y_i$, and the ℓ_2-norm by $\|x\|_2 = \sqrt{\langle x, x\rangle}$.

2.1. Online Learning in Two-Player Zero-Sum Games

We consider a setting where two players (Player 1 and Player 2) repeatedly play a zero-sum game with payoff matrix $A \in \mathbb{R}^{m\times n}$ over a sequence of T rounds. At each round $t = 0, \ldots, T$, the players choose mixed strategies $x_1^t \in \Delta_m$ and $x_2^t \in \Delta_n$ belonging to the m and n-dimensional simplices. Players 1 and 2 then obtain expected payoffs $\langle x_1^t, Ax_2^t\rangle$ and $-\langle x_2^t, A^\top x_1^t\rangle$, and they observe the vector feedback Ax_2^t and $-A^\top x_1^t$, respectively.

The goal of each player is to maximize their cumulative expected payoffs. In this online learning setting, we quantify Player i's performance by measuring its *regret* $\text{Reg}_i(T)$, which is the difference between its cumulative expected payoff and the cumulative payoff of the best fixed strategy in hindsight after T rounds. More precisely, we define the regrets of Player 1 and Player 2 to be:

$$\text{Reg}_1(T) := \max_{x_1\in\Delta_m} \sum_{t=0}^{T} \langle x_1, Ax_2^t\rangle - \sum_{t=0}^{T} \langle x_1^t, Ax_2^t\rangle$$
$$\text{Reg}_2(T) := \sum_{t=0}^{T} \langle x_1^t, Ax_2^t\rangle - \min_{x_2\in\Delta_n} \sum_{t=0}^{T} \langle x_2, A^\top x_1^t\rangle\ .$$

We define the *total regret* of the two players to be $\text{Reg}(T) := \text{Reg}_1(T) + \text{Reg}_2(T)$.

A well-known connection between online learning and game theory establishes that bounds on the growth rate of $\text{Reg}(T)$ implies the convergence of the players' time-averaged mixed strategies to a *Nash Equilibrium* (NE), the canonical solution concept when two players play A for a single round. Recall that an NE for A is a pair $(x_1^*, x_2^*) \in \Delta_m \times \Delta_n$ satisfying the inequalities

$$\max_{x_1\in\Delta_m} \langle x_1, Ax_2^*\rangle \leq \langle x_1^*, Ax_2^*\rangle \leq \min_{x_2\in\Delta_n} \langle x_1^*, Ax_2\rangle\ .$$

Due to von Neumann's minimax theorem (von Neumann, 1928), every zero-sum game has at least one NE. Then the following relationship holds (see Appendix 8.1 for a proof):

Proposition 1 *Let $\tilde{x}_1^T := (\sum_{t=0}^T x_1^t)/T$ and $\tilde{x}_2^T := (\sum_{t=0}^T x_2^t)/T$ denote the time-averaged strategies of Players 1 and 2, respectively. If $\text{Reg}(T) = o(T)$, then $(\tilde{x}_1^T, \tilde{x}_2^T)$ converges to an NE (x_1^*, x_2^*) in duality gap at a rate of $\text{Reg}(T)/T = o(1)$.*

2.2. Symmetric Zero-Sum Games and Symmetric Learning

A zero-sum game with payoff matrix $A \in \mathbb{R}^{n \times n}$ is called *symmetric* if $A = -A^\top$ (i.e., A is a skew-symmetric matrix). In this paper, we consider such symmetric zero-sum games under a *symmetric initialization* (i.e., when $x_1^0 = x_2^0$), and when the players use the *same algorithm* to update their strategies. It follows that $x_1^t = x_2^t$ for all $t = 0, \dots, T$, and so we call this setting *symmetric learning*. Letting x^t denote the common strategy at time t, the total regret $\mathrm{Reg}(T)$ can be written as

$$\mathrm{Reg}(T) \;:=\; 2 \cdot \max_{x \in \Delta_n} \sum_{t=0}^{T} \langle x, Ax^t \rangle \;. \tag{1}$$

Letting $\tilde{x}^T = (\sum_{t=0}^T x^t)/T$, it follows by Proposition 1 that the strategy profile $(\tilde{x}^T, \tilde{x}^T)$ converges to an NE of A at a rate of $\mathrm{Reg}(T)/T$. In such cases where an NE $(x_1^*, x_2^*) = (x^*, x^*)$ is symmetric, we simply say that x^* is an NE of A. Nash equilibria for symmetric zero-sum games have the following property (see Appendix 8.2 for a proof):

Proposition 2 *Let $A = -A^\top$ be a symmetric zero-sum game, and let x^* be an NE for A. Then $Ax^* = 0$ (where $0 \in \mathbb{R}^n$ is the all-zeros vector).*

Symmetric learning has its origins in both classical and evolutionary game theory (von Neumann, 1928; Brown and Von Neumann, 1950; Gale et al., 1950; Weibull, 1997) and has received increasing interest due to its connections with self-play in modern multi-agent learning settings (Lanctot et al., 2017; Balduzzi et al., 2019). Recent work has also shown that *swap*-regret minimization in symmetric learning settings results in stable, Nash-convergent dynamics (Leme et al., 2024).

2.3. High-Dimensional RPS Matrices

Our primary focus in this work is on a family of symmetric, n-dimensional zero-sum games that generalize the classic three-strategy Rock-Paper-Scissors (RPS). These games are specified by skew-symmetric payoff matrices with the following structure:

Definition 3 *For any $n \geq 3$, we say $A \in \mathbb{R}^{n \times n}$ is an n-dimensional RPS matrix with positive constants $a_1, \dots, a_n > 0$ if its entries $A_{i,j}$ are given by*

$$A_{i,j} \;:=\; \begin{cases} -a_i & \text{if } j = i+1 \pmod{n} \\ a_{i-1} & \text{if } j = i-1 \pmod{n} \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } i, j \in [n] \;.$$

Concretely, an n-dimensional RPS matrix has the following structure:

$$\text{For } n = 3: \quad A = \begin{pmatrix} 0 & -a_1 & a_3 \\ a_1 & 0 & -a_2 \\ -a_3 & a_2 & 0 \end{pmatrix}. \quad \text{For } n > 3\text{:} \;\; A = \begin{pmatrix} 0 & -a_1 & 0 & \dots & a_n \\ a_1 & 0 & -a_2 & 0 & \dots \\ 0 & a_2 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & -a_{n-1} \\ -a_n & 0 & \dots & a_{n-1} & 0 \end{pmatrix}.$$

When all $a_1 = \cdots = a_n = 1$, we say that A is the *unweighted* n-dimensional RPS matrix. We write $a_{\max}$ and $a_{\min}$ to denote the maximum and minimum entries among $\{a_i\}$, which we assume are absolute constants, and we write $A_i \in \mathbb{R}^n$ to denote the i-th column vector of A. By slight abuse of notation, the use of $\pmod n$ assumes the indices stay within the set $\{1, \ldots, n\}$ in the natural way. For readability, we will usually omit the $\pmod n$ operator when referencing the indices of $\{a_i\}$: it will be clear from context that if $i = n$, then $i + 1 = 1$, and if $i = 1$, then $i - 1 = n$.

In Appendix 8, we also prove that every n-dimensional RPS matrix A has a (not-necessarily unique) *interior* NE x^*.

Proposition 4 *Let A be an n-dimensional RPS matrix from Definition 3 with positive constants $a_1, \ldots, a_n > 0$. Then A has an interior Nash equilibrium x^*.*

RPS variants have been studied extensively in classical game theory (Von Neumann and Morgenstern, 1944) and in evolutionary game theory (Hofbauer and Sigmund, 1998). Other lines of work focusing on this fundamental class are discussed in Appendix 7.

3. A Unifying View of FP and GD in Symmetric Games

3.1. Leader-Based Algorithms for Symmetric Learning

Leader-based algorithms are the most ubiquitious methods for online learning in games, and Fictitious Play and Gradient Descent can both be viewed in this perspective. We introduce these algorithms in the context of symmetric learning with skew-symmetric payoff matrices $A = -A^\top$.

Fictitious Play. For symmetric learning in symmetric games, Fictitious Play (FP) is initialized at a strategy $x^0 \in \Delta_n$. At each step $t + 1$, the algorithm selects the (pure) strategy x^{t+1} given by

$$x^{t+1} \;:=\; \operatorname*{argmax}_{x \in \{e_i :\, i \in [n]\}} \; \left\langle x, \sum_{k=0}^{t} A x^k \right\rangle . \tag{FP}$$

Here, $\{e_i : \; i \in [n]\}$ is the set of standard basis vectors in $\mathbb{R}^n$, which corresponds to the vertices of Δ_n. For convenience, we use the shorthand $\{e_i\}$ to denote this set. Moreover, we make the following remark on the behavior of the argmax function in the update rule:

Remark 5 (Tiebreaking Rules) *At time t, the set $M^t = \{i \in [n] : \langle e_i, y^t \rangle = \max_{j \in [n]} \langle e_j, y^t \rangle\}$ may contain multiple vertices. For this, we assume the function $\operatorname{argmax}_{x \in \{e_i\}} \langle x, y^t \rangle$ encodes a tiebreaking rule that always returns a single element from M^t. Unless otherwise specified, we make no assumptions on the tiebreaking rule (e.g., it may be adaptive/adversarial with respect to the history of previous iterates). This is in contrast to Daskalakis and Pan (2014) (who assumed a specific adversarial tiebreaking rule) and Abernethy et al. (2021) (who assumed a fixed lexicographical tiebreaking rule) for diagonal payoff matrices.*

FTRL and Gradient Descent. Gradient Descent (GD) is an instantiation of the more general Follow-the-Regularized-Leader (FTRL) algorithm. Using a strictly convex regularizer $\phi : \Delta_n \to \mathbb{R}$ and a fixed stepsize $\eta > 0$, the iterates of FTRL update as:

$$x^{t+1} \;:=\; \operatorname*{argmax}_{x \in \Delta_n} \; \left\langle x, \sum_{k=0}^{t} A x^k \right\rangle - \frac{\phi(x)}{\eta} \; . \tag{FTRL}$$

Due to the strict convexity of ϕ, the maximization problem (FTRL) is strictly concave and has a unique solution (and thus no tiebreaking is needed). In this work, we focus on FTRL instantiated with the ℓ^2 regularizer $\phi(x) = \frac{1}{2}\|x\|_2^2$, which is 1-strongly convex. This results in the online Gradient Descent (GD) algorithm.

Interpolating Between FTRL and FP. Our particular regime of interest is when $\eta = \Theta(1)$ is a fixed absolute constant with respect to T. Note that using the time-horizon-dependent setting of $\eta \approx 1/\sqrt{T}$ (which is used to obtain $O(\sqrt{T})$ regret bounds in general online learning setings) implies that the maximization problem at each time $t+1$ places a *weight* on the regularizer of order $\sqrt{T}$. In contrast, when $\eta = \Theta(1)$, this weight is an absolute constant for all t, so as $\eta \to \infty$, the (FTRL) update approaches that of (FP).

3.2. Geometry of the Dual Dynamics

Primal and Dual Updates. The primal iterates $x^t \in \Delta_n$ of both Fictitious Play and FTRL can be expressed in terms of a sequence of *dual payoff vectors* $y^t \in \mathbb{R}^n$. Let $y^0 = 0 \in \mathbb{R}^n$ be the all-zeros vector. For a fixed $\eta > 0$, we define for each $t \geq 1$:

$$y^{t+1} \; = \; y^t + \eta \cdot Ax^t \; . \tag{Dual Vector}$$

Then, the primal iterates of Fictitious Play and FTRL are given by:

$$\begin{aligned} \text{For FP:} \quad & \eta = 1 \;\text{ and }\; & x^{t+1} = \underset{x \in \{e_i\}}{\operatorname{argmax}} \; \langle x, y^{t+1} \rangle \\ \text{For FTRL:} \quad & \eta > 0 \;\text{ and }\; & x^{t+1} = \underset{x \in \Delta_n}{\operatorname{argmax}} \; \langle x, y^{t+1} \rangle - \phi(x) \; . \end{aligned} \tag{2}$$

For each algorithm, we call the corresponding sequence $\{y^t\}$ the set of *dual iterates*.

Regret and Energy of Dual Iterates. Using the definition of (Dual Vector) and the primal iterates of either algorithm from expression (2), the regret definition from (1) can be rewritten as:

$$\text{Reg}(T) \; = \; \frac{2}{\eta} \cdot \max_{x \in \Delta_n} \; \langle x, y^{T+1} \rangle \; .$$

For both algorithms, regret is closely related to a corresponding *energy function* defined over the dual space $\mathbb{R}^n$. For Fictitious Play, the energy is the *support function* Ψ of Δ_n (which is the convex conjugate of the indicator function on Δ_n). For FTRL, the energy is the *convex conjugate* ϕ^* of the regularizer ϕ over the domain Δ_n. Specifically, for $y \in \mathbb{R}^n$, we define:

$$\begin{aligned} \text{Energy function for FP:} \quad & \Psi(y) \; = \; \max_{x \in \Delta_n} \; \langle x, y \rangle \\ \text{Energy function for FTRL:} \quad & \phi^*(y) \; = \; \max_{x \in \Delta_n} \; \langle x, y \rangle - \phi(x) \; . \end{aligned} \tag{3}$$

For Fictitious Play, the following relationship between Ψ and regret is then immediate:

Proposition 6 *Let $\{y^t\}$ be the dual iterates of FP (with $\eta = 1$). Then* $\text{Reg}(T) = 2 \cdot \Psi(y^{T+1})$.

For FTRL, we also establish the following similar relationship (see Appendix 9.3 for a proof):

Proposition 7 *Let $\{y^t\}$ be the dual iterates of FTRL with $\eta > 0$. Let $M = \max_{x \in \Delta_n} \phi(x)$. Then:*

$$\text{Reg}(T) \; \leq \; \frac{2 \cdot \phi^*(y^{T+1})}{\eta} + \frac{2M}{\eta} \; .$$

Energy Conservation in Continuous Time. Several recent works have identified that the continuous time variants (i.e., the limit of vanishing step size $\eta \to 0$) of both FTRL (Mertikopoulos et al., 2018; Bailey and Piliouras, 2019b; Wibisono et al., 2022) and Fictitious Play (Ostrovski and van Strien, 2011; Van Strien, 2011; Abernethy et al., 2021) have a Hamiltonian structure: the dual iterates follow a *skew-gradient flow* that exactly conserves the corresponding energy function over time. In both cases, continuous-time energy conservation corresponds to constant regret bounds.

Skew-Gradient Descent in Discrete Time. In discrete time, the dual iterates of each algorithm follow a first-order forward discretization of the skew-gradient flow with respect to its corresponding energy function. By the convexity of the energy function, the energy along this *skew-(sub)-gradient descent* is always non-decreasing. Formally, following the presentation of Abernethy et al. (2021) and Wibisono et al. (2022), let $\{y^t\}$ be the dual iterates of either Fictitious Play (with $\eta = 1$) or FTRL (with $\eta > 0$) from (2), and let $H : \mathbb{R}^n \to \mathbb{R}$ denote its corresponding energy function from (3). By a slight abuse of notation, let $\partial H(y^t)$ denote any vector in the subgradient set[1] of H at y^t. It is then straightforward to show the following (see Appendix 9.4 for a derivation):

Proposition 8 *Let $\{y^t\}$ be the dual iterates of either Fictitious Play ($\eta = 1$) or FTRL ($\eta > 0$), and let H be its corresponding energy function from* (3)*. Then for every $t \geq 1$, it holds that*

$$y^{t+1} \;=\; y^t + \eta A \partial H(y^t)\;. \tag{4}$$

In particular: for FP, each $x^t \in \partial\Psi(y^t)$, and for FTRL, each $x^t = \nabla\phi^*(y^t)$. Moreover, for all $t \geq 1$: $\Delta H(y^t) = H(y^{t+1}) - H(y^t) \geq 0$.

3.3. Bounds on Regret via Controlling the Energy Growth

In light of the geometric perspective given in Section 3.2 and of the relationships between energy and regret from Propositions 6 and 7, our approach to obtain regret bounds for both algorithms is to control the energy growth of their respective dual iterates over time. While smoothness properties of the energy function can be used to derive one-step, worst-case growth bounds (e.g., as in Wibisono et al. (2022)), this approach may be overly pessimistic.

In this work, we instead aim for a more fine-grained analysis (leveraging structural properties of the underlying payoff matrices) that captures the non-uniform energy growth implied by the geometric perspective developed above. Below, we provide intuition as to why such an approach may be possible:

Intuition for Fictitious Play. For example, for Fictitious Play, we show in Section 4 (see Proposition 12) that when $x^t = x^{t+1} = e_i$ for some $i \in [n]$, then $\Psi(y^{t+1}) - \Psi(y^t) = 0$ (and only when $x^t \neq x^{t+1}$ can Ψ increase). Here, the intuition is the following: by definition of Ψ, the energy function is linear in the coordinate y_i within the region $L_i \subseteq \mathbb{R}^n$, where

$$L_i \;=\; \{y \in \mathbb{R}^n : y_i > \max_{j \neq i} y_j\}.$$

Moreover, for any time t such that $y^t \in L_i$, the definition of (15) shows that $\Delta y^t = y^{t+1} - y^t = A_i$, and Proposition 8 implies that Δy^t follows a linear skew-(sub)gradient step with respect to Ψ. Thus the coordinate $y_i^{t+1} = y_i^t$ is unchanged, and if y^{t+1} also falls in L_i, then y^{t+1} and y^t must be on

1. For convex H, its subgradient set at $y \in \mathbb{R}^n$ is given by $\partial H(y) = \{g \in \mathbb{R}^n : \forall z \in \mathbb{R}^n, H(z) \geq H(y) + \langle g, z - y\rangle\}$.

the same level set of Ψ and energy is conserved (in other words, the skew-gradient discretization behaves the same as the skew-gradient flow).

On the other hand, if (due to the discretization) y^{t+1} lands in a new region L_j, the energy increases, and Ψ becomes linear in y_j. This explains the expanding (but linear) trajectory of the FP dual iterates in Figure 1(*a*), and it roughly implies that the total energy growth (and regret) can be controlled by understanding how frequently the dual iterates switch between the L_i regions.

Intuition for FTRL and Gradient Descent. For Fictitious Play, energy conservation between consecutive dual iterates is guaranteed when the two corresponding primal iterates lie on the same vertex of Δ_n. While the primal iterate of Fictitious Play is *always* at a vertex (for $t \geq 1$), for FTRL instantiated with a *Legendre* regularizer ϕ (for example, the negative entropy function corresponding to the Multiplicative Weights algorithm), the primal iterates will always remain on the *interior* of the simplex (see e.g., Wibisono et al. (2022)). Thus in general, the geometry of the dual iterates of FTRL will not exactly coincide with those of Fictitious Play.

On the other hand, in this work we focus on Gradient Descent, the FTRL instantiation with $\phi(x) = \frac{1}{2}\|x\|_2^2$. For Gradient Descent, the primal update rule of (2) may require a projection onto Δ_n, and thus the primal iterates may in general lie on the boundary of the simplex. In particular, when $x^t = e_i$ for some $i \in [n]$, the dual iterate y^t *must lie in a region of* $\mathbb{R}^n$ *where the geometry of* ϕ^* *and* Ψ *exactly aligns* (specifically, ϕ^* is linear in y_i in this region). Within these regions (defined formally in Section 5) the dual trajectory of Gradient Descent is identical to that of Fictitious Play (both aligning with a linear skew-gradient flow), and energy is conserved each step.

However, when the primal iterate x^t is on a non-vertex boundary (or in the interior of Δ_n), the energy ϕ^* is *quadratic* in y^t, and thus the first-order discretization of the skew-gradient flow will strictly increase the energy. This explains the expanding trajectory of the dual iterates of Gradient Descent in Figure 1(*b*), and analogously to Fictitious Play, it implies that the energy growth can be controlled by analyzing how frequently the dual iterates switch between the linear and quadratic regions. We give more details on the Gradient Descent primal update and energy function in Section 5 and Appendix 12.

Challenges in High Dimension. The preceding intuition establishes similarities between the dual iterates and energy functions for Fictitious Play and Gradient Descent when consecutive sequences of the primal iterates lie on the same vertex of Δ_n. While this intuition holds for symmetric learning on any skew-symmetric payoff matrix A, without tight control over how the primal (and dual) iterates evolve over longer sequences of steps, controlling the energy growth (and thus regret) can be challenging, especially in high dimensions. However, for the class of n-dimensional RPS matrices, we prove that such long-term control for both algorithms is possible, subsequently leading to strong regret guarantees. We present these analyses in the following sections.

4. Analysis of Fictitious Play on High-Dimensional RPS

In this section, we introduce our analysis of Fictitious Play on high-dimensional RPS matrices, for which we prove a worst-case $O(\sqrt{T})$ regret bound. Using the dual perspective from Section 3, recall that the iterates of Fictitious Play are given by

$$\begin{cases} y^{t+1} &= y^t + Ax^t \\ x^{t+1} &= \operatorname{argmax}_{x \in \{e_i\}} \langle x, y^{t+1} \rangle \ . \end{cases} \tag{FP Primal-Dual}$$

Throughout this section, we let $\{x^t\}$ and $\{y^t\}$ denote these primal and dual iterates, respectively.

4.1. Cycling of Primal Iterates Under Arbitrary Tiebreaking

For any n-dimensional RPS matrix A, and using any (possibly adversarial) tiebreaking rule, our analysis begins by proving the following *cycling* behavior of the Ficitious Play iterates: if the energy of the dual iterates ever increases from its initial value, then the subsequent primal iterates *cycle through a fixed order of the vertices of* Δ_n *for the remainder of the dynamics*. Specifically, starting from some vertex $e_i \in \Delta_n$, the iterates $\{x^t\}$ cycle in the order

$$e_i \to e_{i+1} \to \cdots \to e_n \to e_1 \to \cdots \to e_i \to \ldots$$

We call a sequence of consecutive iterates at the same vertex a *phase*, defined formally as follows:

Definition 9 *Fix a time $t_0 > t$. For each $k \geq 1$, let $t_k := \min\,\{t > t_{k-1} : x^t \neq x^{t_{k-1}}\}$. Then Phase k is the sequence of iterates at times $t = t_k, t_k + 1, \ldots, t_{k+1} - 1$. Let $\tau_k = t_{k+1} - t_k$ denote the length of Phase k. Let $K \geq 0$ be the total number of phases in T rounds, where $T = \sum_{k=0}^{K} \tau_k$.*

Then for every RPS matrix A, and using any tiebreaking method (cf., Remark 5), the following behavior occurs: if $x^t = e_j$ for $j \in [n]$, then $x^{t'} = e_{j+1 \ (\mathrm{mod}\ n)}$, where $t' > t$ is the next time the primal iterate changes. Formally:

Lemma 10 *Let t_0 be the first $t > 0$ where $\Psi(y^{t_0}) > \Psi(y^1)$, and suppose $x^{t_0} = e_i$ for some $i \in [n]$. Then $x^{t_k} = e_{i+k \ (\mathrm{mod}\ n)}$ for all $k \geq 1$.*

The proof of the lemma (see Appendix 10) relies on exactly characterizing the *linear* trajectory of the dual iterates within a phase. Specifically, suppose in the current phase that the primal iterate is at vertex e_i. Then by definition of the update rule of (FP Primal-Dual), $y_i^t \geq \max_{j \neq i} y_j^t$ for all iterates t within the phase, and thus the velocity $\Delta y^t = y^{t+1} - y^t = A_i$ is a fixed constant vector.

Using the structure of RPS matrices, we can then track the evolution of the coordinates of y^t under this fixed velocity, and we prove that, if ever $y_i^t = y_j^t$ for some $j \neq i$ at time t within the current phase, then $j = i + 1$. In other words, tiebreaking scenarios can only occur between adjacent coordinates of the dual variable. By leveraging the structure of the velocity $\Delta y^t = A_i$ for RPS matrices, it is then straightforward to establish that under any tiebreaking rule, the primal iterate must eventually switch to vertex e_{i+1} in the next phase.

4.2. Cycling Implies Worst-Case $O(\sqrt{T})$ Regret

The cycling behavior of Lemma 10 establishes a *regularity* in the trajectory and energy growth of the dual iterates. Given the relationship between energy and regret from Proposition 6, this regularity ultimately allows us to establish the following worst-case $O(\sqrt{T})$ regret bound:

Theorem 11 *Let A be an n-dimensional RPS payoff matrix, and let $\{x^t\}$ and $\{y^t\}$ be the iterates of (FP Primal-Dual) on A from any $x^0 \in \Delta_n$. Then using any tiebreaking rule,* $\mathrm{Reg}(T) \leq O(\sqrt{T})$.

As mentioned in Section 1, Theorem 11 establishes the first class of zero-sum games beyond the case of diagonal payoff matrices (from Abernethy et al. (2021)) for which Fictitious Play has provable $O(\sqrt{T})$ regret, and by Proposition 1, this also guarantees a $O(1/\sqrt{T})$ convergence rate (in duality gap) of the time-averaged iterates to an NE of the game. The full proof of the theorem is developed in Appendix 10, but we sketch the main ideas below.

Proof Sketch of Theorem 11. In addition to the cycling behavior of Lemma 10, the proof relies on (i) bounds on the energy growth between phases, and (ii) bounds on the length of each phase.

(i) Bounds on energy growth. First, we establish the following two cases of energy growth:

Proposition 12 *For any t, define $\Delta\Psi(y^t) := \Psi(y^{t+1}) - \Psi(y^t)$. Then:*

(a) If $x^t = x^{t+1}$, then $\Delta\Psi(y^t) = 0$.

(b) If $x^t \neq x^{t+1}$, then $0 \leq \Delta\Psi(y^t) \leq a_{\max}$.

In particular, the proposition implies that Ψ can only increase when entering a new phase, and the total energy $\Psi(y^{T+1})$ is proportional to the number of phases in which Ψ has strictly increased. The proof of the proposition follows along the lines of the intuition introduced in Section 3.3 for the energy growth of skew-gradient descent.

(ii) Bounds on phase length. Then, using the cycling behavior of Lemma 10, we prove that the *length* τ_k of each Phase k is roughly proportional to the energy at the start of the phase:

Lemma 13 *For $k = 1, 2, \ldots, K$, let $\gamma_k = \Psi(y^{t_k})$ be the energy at the start of Phase k. Then $\tau_k \geq \alpha_k \cdot \gamma_k - \beta_k$, where $\alpha_k > 0$ and $\beta_k > 0$ are absolute constants.*

The proof of the lemma uses the following intuition: first, assume for simplicity that $x^{t_k} = e_i$ at the start of Phase k. Then due to the fixed cycling order, we show that the phase length τ_k is bounded below by

$$\tau_k \geq \Omega\big(y_i^{t_k} - y_{i+1}^{t_k}\big) \;=\; \Omega\big(\Psi(y^{t_k}) - y_{i+1}^{t_k}\big)\,. \tag{5}$$

the cycling order and structure of RPS matrices further implies that each coordinate of the dual vector can increase and decrease in exactly one phase during every consecutive sequence of n phases. In turn, this allows for controlling the growth of coordinate $(i+1)$ over time, which ensures in expression (5) that roughly $y_{i+1}^{t_k} \leq O(\Psi(y^{t_k}))$.

Together, Proposition 12 and Lemma 13 establish a quadratic relationship between the total number of phases and the total energy of the dual iterates, which is a similar property to those leveraged by both Bailey and Piliouras (2019a) and Abernethy et al. (2021). To see how such a relationship leads to $O(\sqrt{T})$ regret, assume for simplicity that Ψ strictly increases between phases, and thus $\Psi(y^{t_k}) - \Psi(y^{t_{k-1}}) = \Theta(1)$ for all k. By Proposition 12, this implies that $\Psi(y^{T+1}) \leq O(K)$, where K is the total number of phases in T rounds. By Lemma 13, this also implies that $\tau_k \geq \Omega(\Psi(y^{t_k})) \geq \Omega(k)$. Together, we find:

$$T = \sum_{k=0}^{K} \tau_k \geq \sum_{k=0}^{k} \Omega(k) \geq \Omega(K^2) \implies K = O(\sqrt{T}) \implies \Psi(y^{T+1}) \leq O(\sqrt{T})\,. \tag{6}$$

By Proposition 6, this proves the claimed regret bound. Note that Ψ might not be strictly increasing between each phase, and in the full proof we account for this behavior.

5. Analysis of Gradient Descent on High-Dimensional RPS

In this section, we turn toward analyzing Gradient Descent on high-dimensional RPS matrices. Using the dual perspective from Section 3, recall that the iterates of Gradient Descent with stepsize $\eta > 0$ are given by:

$$\begin{cases} y^{t+1} & = y^t + \eta A x^t \\ x^{t+1} & = \text{argmax}_{x \in \Delta_n} \langle x, y^{t+1} \rangle - \frac{\|x\|_2^2}{2} \, . \end{cases} \tag{GD Primal-Dual}$$

Throughout this section, we let $\{x^t\}$ and $\{y^t\}$ denote these primal and dual iterates, respectively.

Closed-Form Expressions and Primal-Dual Map. As mentioned in Section 3, the primal update rule of (GD Primal-Dual) may require a projection onto the boundary of Δ_n. For this, we state in Proposition 25 (derived originally in Bailey and Piliouras (2019a)) a closed-form characterization of the primal iterates x^t, which also leads to a closed-form characterization of the energy function $\phi^*(y^t)$. To streamline the presentation, we defer these details to Appendix 12. However, central to our analysis of Gradient Descent is to identify regions of the dual space $\mathbb{R}^n$ that, under the update, map to the vertices and (a subset of) edges of Δ_n. Specifically, we define the regions P_i and $P_{i\sim(i+1)}$ as follows:

Definition 14 *For each $i \in [n]$, let $P_i \subset \mathbb{R}^n$ and $P_{i\sim(i+1)} \subset \mathbb{R}^n$ be the following sets:*

$$P_i := \Big\{ y \in \mathbb{R}^n : \quad y_i - y_j > 1 \quad \textit{for } j \in [n] \setminus \{i\} \Big\}$$
$$P_{i\sim(i+1)} := \left\{ y \in \mathbb{R}^n : \quad \begin{array}{l} |y_i - y_{i+1}| \leq 1 \textit{ and} \\ \frac{1}{2}(y_i + y_{i+1}) - y_j > \frac{1}{2} \quad \textit{for } j \in [n] \setminus \{i, i+1\} \end{array} \right\} .$$

Then, in Appendix 12.2, we prove the following relationship:

Proposition 15 *For any $i \in [n]$: $y^t \in P_i$ if and only if $x^t = e_i$, and $y^t \in P_{i\sim(i+1)}$ if and only if* $\text{supp}(x^t) = \{i, i+1\}$.

In other words, if $y^t \in P_i$, then the primal iterate x^t must be at the vertex e_i, and if $y^t \in P_{i\sim(i+1)}$, then the primal iterate is on the edge of Δ_n between the vertices e_i and e_{i+1} (and vice versa).

5.1. $O(\sqrt{T})$ Regret with Large Stepsizes

The behavior of Fictitious Play on RPS matrices established in Section 4, together with the shared geometric characterization of Fictitious Play and Gradient Descent introduced in Section 3, suggests the following intuition: if the primal iterates of Gradient Descent eventually reach a *vertex* of Δ_n, then we could expect all subsequent primal iterates to demonstrate a similar cycling behavior as Fictitious Play, and to show a similar regularity in energy growth and regret.

In this section, we prove in Theorem 17 that this intuition is indeed correct in the regime of *large* constant stepsizes: for almost all initializations $x^0 \in \Delta_n$, we prove that when the stepsize $\eta > 0$ is a *sufficiently large* constant, then Gradient Descent obtains $O(\sqrt{T})$ regret on every n-dimensional RPS matrix. The proof of this result relies on establishing that the primal iterates (i) converge to a vertex of Δ_n, and (ii) exhibit a cycling behavior that leads to patterns of energy growth and phase lengths similar to those of Fictitious Play. We introduce these components below:

Fast Convergence to a Vertex. First, we prove that using a large enough stepsize ensures that the primal iterate reaches a vertex after only a single iteration. For this, fixing an RPS matrix A, for any $x \in \Delta_n$, let $\gamma(x)$ be the constant

$$\gamma(x) := \min_{k \neq \ell \in [n]} \Big| \big(a_{k-1} \cdot x_{k-1} - a_k \cdot x_{k+1}\big) - \big(a_{\ell-1} \cdot x_{\ell-1} - a_\ell \cdot x_{\ell+1}\big) \Big| . \tag{7}$$

Then the following holds:

Lemma 16 *If $x^0 \in \Delta_n$ is such that $\gamma(x^0) > 0$, then along one step of* (GD Primal-Dual) *with stepsize $\eta > 1/\gamma(x^0)$, the iterate x^1 is a vertex of Δ_n.*

To prove the lemma (see Appendix 13.1), we show that when $\eta > 1/\gamma(x^0)$, the dual iterate $y^1 = \eta A x^0$ must fall in some region P_i (from Definition 14), and thus by Proposition 15, $x^1 = e_i$. By definition of γ, the set of points $x \in \Delta_n$ where $\gamma(x) = 0$ are the solutions to the linear constraint in (7) and has Lebesgue measure zero (note also by the definition of RPS matrices and Proposition 2 that every Nash equilibrium x^* of A has $\gamma(x^*) = 0$).

Cycling, Energy Growth, and Phase Lengths. In this large stepsize regime, we further establish that the primal iterates eventually cycle between vertices of Δ_n in the same order as in Lemma 10 for Fictitious Play. However, between the vertices e_i and e_{i+1}, the iterates may spend a constant number of steps on the edge of Δ_n where $\mathrm{supp}(x) = \{i, i+1\}$ (which by Proposition 15 means that the corresponding dual iterates lie in the region $P_{i\sim(i+1)}$). This explains the behavior of the primal iterates on $n = 4$ unweighted RPS in Figure 1(*c*) and is formally captured in Lemma 28, which we state and prove in Appendix 13.

Lemma 28 additionally gives bounds on the energy growth between phases (at most an absolute constant) and the length of each phase (growing proportionally to energy) similar to the analysis of Fictitious Play from Section 4.2. In turn, this leads to a similar worst-case $O(\sqrt{T})$ regret bound for Gradient Descent on high-dimensional RPS matrices:

Theorem 17 *Let A be an n-dimensional RPS matrix. Then for nearly all initial distributions $x^0 \in \Delta_n$, the following holds: letting $\{x^t\}$ be the iterates of running* (GD Primal-Dual) *on A with $\eta > \min\{2/a_{\min}, 1/\gamma(x^0)\}$, then* $\mathrm{Reg}(T) \leq O(\sqrt{T})$.

The full proof of Theorem 17 is developed in Appendix 13 and follows similarly to that of Theorem 11 for FP. We remark that while the constraint on η may depend on the initialization x^0, the theorem and its proof yield the following moral conclusion: with large enough constant stepsizes, the dual trajectory of Gradient Descent becomes increasingly similar to that of Fictitious Play, and GD thus inherits the same cycling and regularity in energy growth that are sufficient for establishing $O(\sqrt{T})$ regret.

5.2. Behavior of Gradient Descent with Smaller Stepsizes

In light of the regret guarantees for Gradient Descent on RPS matrices in the *large* stepsize regime, a natural question to ask is how the algorithm (and its regret) behaves using smaller stepsizes. Fully answering this question is more challenging, and our work leaves open a complete characterization of the behavior of GD on RPS matrices. However, in Appendix 14, we prove several auxiliary results in this direction. To summarize:

- In Theorem 32, we prove a *boundary invariance* property that holds for any $\eta > 0$: when the energy ever exceeds a game-dependent constant value, then every subsequent primal iterate of Gradient Descent must lie on the boundary of Δ_n. In particular, this extends the boundary invariance result of Bailey and Piliouras (2019a) for 2×2 games to the present n-dimensional symmetric setting. However, beyond boundary invariance, obtaining $O(\sqrt{T})$ regret bounds for Gradient Descent using *any* constant stepsize remains open.

- At the other extreme, we prove in Lemma 33 that with a small time-horizon-dependent stepsize (e.g., $\eta = \Theta(1/\sqrt{T})$), even if all primal iterates remain *interior*, we can still only guarantee a worst-case regret bound for Gradient Descent scaling like $O(\sqrt{T})$. Interestingly, in Appendix 15, we present simulations showing empirically that Gradient Descent with large constant stepsizes obtains tighter regret than its time-dependent $\eta = \Theta(1/\sqrt{T})$ counterpart.

We give the precise statements of these results and more discussion in Appendix 14.

6. Discussion and Future Work

This paper establishes new $O(\sqrt{T})$ regret guarantees (and thus $O(1/\sqrt{T})$ time-averaged convergence to Nash equilibria) for two *non*-no-regret algorithms for symmetric learning in zero-sum games. For Fictitious Play, our regret bound for high-dimensional RPS establishes a new class of matrices for which Karlin's Conjecture is true, and importantly, this result holds using *any* tiebreaking rule. Interestingly, we show in Theorem 24 of Appendix 11 that on unweighted RPS, and using a specific *fixed* tiebreaking rule, FP obtains only *constant* regret. Better understanding the interplay between tiebreaking and regret for other classes of matrices is left as open.

For Gradient Descent, our result establishes the first sublinear regret guarantees using constant stepsizes in high-dimensional zero-sum games, beyond the 2×2 case. Our analysis leveraged the shared geometry of the two algorithms that emerges under *large constant stepsizes*, and we believe this insight may be useful for deriving regret bounds for both algorithms in other classes of symmetric zero-sum games. Finally, obtaining regret guarantees for other instantiations of FTRL (beyond Gradient Descent) with constant stepsizes remains an important open challenge.

Acknowledgements This research was supported by the MOE Tier 2 Grant (MOE-T2EP20223-0018), the National Research Foundation, Singapore, under its QEP2.0 programme (NRF2021-QEP2-02-P05), the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (Award Number: AISG2-RP-2020-016). AW was supported by NSF Award CCF #2403391. The authors thank Yang Cai, Anas Barakat, and Joseph Sakos for helpful discussions.

References

Jacob Abernethy, Kevin A Lai, and Andre Wibisono. Fast convergence of fictitious play for diagonal payoff matrices. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1387–1404. SIAM, 2021.

Ioannis Anagnostides, Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Noah Golowich, and Tuomas Sandholm. Near-optimal no-regret learning for correlated equilibria in multi-player general-sum games. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 736–749, 2022.

James Bailey and Georgios Piliouras. Fast and furious learning in zero-sum games: Vanishing regret with non-vanishing step sizes. *Advances in Neural Information Processing Systems*, 32, 2019a.

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.

James P Bailey and Georgios Piliouras. Multi-agent learning in network zero-sum games is a Hamiltonian system. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 233–241, 2019b.

James P Bailey, Gauthier Gidel, and Georgios Piliouras. Finite regret and cycles with fixed step-size via alternating gradient descent-ascent. In *Conference on Learning Theory*, pages 391–407. PMLR, 2020.

David Balduzzi, Marta Garnelo, Yoram Bachrach, Wojciech Czarnecki, Julien Perolat, Max Jaderberg, and Thore Graepel. Open-ended learning in symmetric zero-sum games. In *International Conference on Machine Learning*, pages 434–443. PMLR, 2019.

Lucas Baudin and Rida Laraki. Fictitious play and best-response dynamics in identical interest and zero-sum stochastic games. In *International Conference on Machine Learning*, pages 1664–1690. PMLR, 2022.

Ulrich Berger. Fictitious play in $2 \times n$ games. *Journal of Economic Theory*, 120(2):139–154, 2005.

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.

Avrim Blum, Eyal Even-Dar, and Katrina Ligett. Routing without regret: On convergence to nash equilibria of regret-minimizing algorithms in routing games. In *Proceedings of the Twenty-Fifth Annual ACM Symposium on Principles of Distributed Computing*, pages 45–52, 2006.

Stephen Boyd and Lieven Vandenberghe. Convex optimization, 2004.

Felix Brandt, Felix Fischer, and Paul Harrenstein. On the rate of convergence of fictitious play. In *Algorithmic Game Theory: Third International Symposium, SAGT 2010, Athens, Greece, October 18-20, 2010. Proceedings 3*, pages 102–113. Springer, 2010.

George W Brown. *Some notes on computation of games solutions*. Rand Corporation, 1949.

George W Brown. Iterative solution of games by fictitious play. *Act. Anal. Prod Allocation*, 13(1): 374, 1951.

George W Brown and John Von Neumann. *Solutions of games by differential equations*. Rand Corporation, 1950.

Yang Cai, Gabriele Farina, Julien Grand-Clément, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Weiqiang Zheng. Fast last-iterate convergence of learning in games requires forgetful algorithms. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

Timothy N Cason, Daniel Friedman, and Ed Hopkins. Cycles and instability in a rock–paper–scissors population game: A continuous time experiment. *Review of Economic Studies*, 81(1): 112–136, 2014.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

Xi Chen and Binghui Peng. Hedging in games: Faster convergence of external and swap regrets. *Advances in Neural Information Processing Systems*, 33:18990–18999, 2020.

Yun Kuen Cheung and Georgios Piliouras. Vortices instead of equilibria in minmax optimization: Chaos and butterfly effects of online learning in zero-sum games. In *Conference on Learning Theory*, pages 807–834. PMLR, 2019.

Yun Kuen Cheung and Georgios Piliouras. Chaos, extremism and optimism: Volume analysis of learning in games. *Advances in Neural Information Processing Systems*, 33:9039–9049, 2020.

Yun Kuen Cheung, Georgios Piliouras, and Yixin Tao. The evolution of uncertainty of learning in games. In *International Conference on Learning Representations*, 2022.

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.

Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. Family of chaotic maps from game theory. *Dynamical Systems*, 36(1):48–63, 2021.

Constantinos Daskalakis and Qinxuan Pan. A counter-example to Karlin's strong conjecture for fictitious play. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 11–20. IEEE, 2014.

Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. *arXiv preprint arXiv:1807.04252*, 2018.

Constantinos Daskalakis, Alan Deckelbaum, and Anthony Kim. Near-optimal no-regret algorithms for zero-sum games. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 235–254. SIAM, 2011.

Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. Near-optimal no-regret learning in general games. *Advances in Neural Information Processing Systems*, 34:27604–27616, 2021.

Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

David Gale, Harold W Kuhn, and Albert W Tucker. On symmetric games. *Contributions to the Theory of Games*, 1:81–87, 1950.

Christopher Griffin, Li Feng, and Rongling Wu. Spatial dynamics of higher order rock-paper-scissors and generalisations. *Journal of Physics A: Mathematical and Theoretical*, 57(18): 185701, 2024.

James Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3(2):97–139, 1957.

Christopher Harris. On the rate of convergence of continuous-time fictitious play. *Games and Economic Behavior*, 22(2):238–259, 1998.

Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

Johannes Heinrich, Marc Lanctot, and David Silver. Fictitious self-play in extensive-form games. In *International Conference on Machine Learning*, pages 805–813. PMLR, 2015.

Josef Hofbauer and William H Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.

Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.

Yu-Guan Hsieh, Kimon Antonakopoulos, and Panayotis Mertikopoulos. Adaptive learning in continuous games: Optimal regret bounds and convergence to Nash equilibrium. In *Conference on Learning Theory*, pages 2388–2422. PMLR, 2021.

Samuel Karlin. *Mathematical methods and theory in games, programming, and economics*. Addison-Wesley, 1959.

Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Pérolat, David Silver, and Thore Graepel. A unified game-theoretic approach to multiagent reinforcement learning. *Advances in Neural Information Processing Systems*, 30, 2017.

Renato Paes Leme, Georgios Piliouras, and Jon Schneider. Convergence of no-swap-regret dynamics in self-play. In *The Thirty-Eighth Annual Conference on Neural Information Processing Systems*, 2024.

Tung Mai, Milena Mihail, Ioannis Panageas, Will Ratcliff, Vijay Vazirani, and Peter Yunker. Cycles in zero-sum differential games and biological diversity. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 339–350, 2018.

Robert M May and Warren J Leonard. Nonlinear aspects of competition between three species. *SIAM Journal on Applied Mathematics*, 29(2):243–253, 1975.

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.

K Miyasawa. On the convergence of the learning process in a 2×2 non-zero-sum game, Princeton University Econometric Research Program. *Research Memorandum*, 1963.

Dov Monderer and Aner Sela. A 2×2 game without the fictitious play property. *Games and Economic Behavior*, 14(1):144–148, 1996.

Dov Monderer and Lloyd S Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.

Georg Ostrovski and Sebastian van Strien. Piecewise linear Hamiltonian flows associated to zero-sum games: transition combinatorics and questions on ergodicity. *Regular and Chaotic Dynamics*, 16:128–153, 2011.

Joshua Paik and Christopher Griffin. Completely integrable replicator dynamics associated to competitive networks. *Physical Review E*, 107(5):L052202, 2023.

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.

Ioannis Panageas, Nikolas Patris, Stratis Skoulakis, and Volkan Cevher. Exponential lower bounds for fictitious play in potential games. *Advances in Neural Information Processing Systems*, 36, 2024.

Sarah Perrin, Julien Pérolat, Mathieu Laurière, Matthieu Geist, Romuald Elie, and Olivier Pietquin. Fictitious play for mean field games: Continuous time analysis and applications. *Advances in Neural Information Processing Systems*, 33:13199–13213, 2020.

Georgios Piliouras, Ryann Sim, and Stratis Skoulakis. Beyond time-average convergence: Near-optimal uncoupled online learning via clairvoyant multiplicative weights update. *Advances in Neural Information Processing Systems*, 35:22258–22269, 2022.

Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. *Advances in Neural Information Processing Systems*, 26, 2013.

Julia Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54(2):296–301, 1951.

William H Sandholm. *Population games and evolutionary dynamics*. MIT press, 2010.

Muhammed O Sayin, Francesca Parise, and Asuman Ozdaglar. Fictitious play in zero-sum stochastic games. *SIAM Journal on Control and Optimization*, 60(4):2095–2114, 2022a.

Muhammed O Sayin, Kaiqing Zhang, and Asuman Ozdaglar. Fictitious play in Markov games with single controller. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 919–936, 2022b.

Aner Sela. Fictitious play in 'one-against-all' multi-player games. *Economic Theory*, 14(3):635–651, 1999.

Dirk Semmann, Hans-Jürgen Krambeck, and Manfred Milinski. Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature*, 425(6956):390–393, 2003.

Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

Lloyd S Shapley et al. *Some topics in two-person games*. Rand Corporation, 1963.

J Maynard Smith. The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology*, 47(1):209–221, 1974.

Brian Swenson and Soummya Kar. On the exponential rate of convergence of fictitious play in potential games. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 275–279. IEEE, 2017.

Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E Schapire. Fast convergence of regularized learning in games. *Advances in Neural Information Processing Systems*, 28, 2015.

Attila Szolnoki, Mauro Mobilia, Luo-Luo Jiang, Bartosz Szczesny, Alastair M Rucklidge, and Matjaž Perc. Cyclic dominance in evolutionary games: a review. *Journal of the Royal Society Interface*, 11(100):20140735, 2014.

Sebastian Van Strien. Hamiltonian flows with random-walk behaviour originating from zero-sum games and fictitious play. *Nonlinearity*, 24(6):1715, 2011.

Lieven Vandenberghe. Lecture notes on conjugate functions, 2012. URL https://www.seas.ucla.edu/~vandenbe/236C/lectures/conj.pdf.

Matthew Visomirski and Christopher Griffin. Integrability of generalised skew-symmetric replicator equations via graph embeddings. *Journal of Physics A: Mathematical and Theoretical*, 58(1):015701, 2024.

John von Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

John Von Neumann and Oskar Morgenstern. Theory of games and economic behavior. In *Theory of Games and Economic Behavior*. Princeton university press, 1944.

Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.

Andre Wibisono, Molei Tao, and Georgios Piliouras. Alternating mirror descent for constrained min-max games. *Advances in Neural Information Processing Systems*, 35:35201–35212, 2022.

Bin Xu, Hai-Jun Zhou, and Zhijian Wang. Cycle frequency in standard rock–paper–scissors games: evidence from experimental economics. *Physica A: Statistical Mechanics and its Applications*, 392(20):4997–5005, 2013.

Table of Contents

7. Additional Related Work

Rock-Paper-Scissors Variants. Rock-Paper-Scissor is the canonical example of a symmetric zero-sum game (Von Neumann and Morgenstern, 1944; Weibull, 1997; Sandholm, 2010), and variants of RPS matrices have been studied extensively in the context of evolutionary game theory (Smith, 1974; May and Leonard, 1975; Szolnoki et al., 2014; Mai et al., 2018) and population dynamics and economics (Semmann et al., 2003; Xu et al., 2013; Cason et al., 2014). Recent works have also framed RPS as the prototypical example of a 'cyclically dominant' zero-sum game defined on tournament graphs (Paik and Griffin, 2023; Visomirski and Griffin, 2024; Griffin et al., 2024).

Fictitious Play. Fictitious Play was originally introduced by Brown (1949, 1951). Though it may fail to converge in general-sum games, (Shapley et al., 1963; Monderer and Sela, 1996), FP has been shown to converge to equilibria in zero-sum games (Robinson, 1951; Harris, 1998), identical-interest/potential games (Monderer and Shapley, 1996) and even in restricted classes of general-sum games (Berger, 2005; Miyasawa, 1963; Sela, 1999). Beyond the works studying the convergence *rates* of FP in zero-sum games described in Section 1, the convergence rate of FP has also been studied in potential games (Panageas et al., 2024; Swenson and Kar, 2017) and general (non zero-sum) games (Brandt et al., 2010). Moreover, the simplicity of the FP algorithm has led to numerous applications in multi-agent learning (Heinrich et al., 2015; Sayin et al., 2022a,b; Baudin and Laraki, 2022; Perrin et al., 2020; Hofbauer and Sandholm, 2002).

Fast Regret Minimization in Games. As described in Section 1, the black-box $O(\sqrt{T})$ regret bound obtained by FTRL with a time-dependent $\theta(1/\sqrt{T})$ stepsize holds for general (and possibly adversarial) online learning settings. Many recent works have focused on improving over this worst-case regret bound when using FTRL in games settings, which in general implies faster convergence to various classes of equilibria. The most widely studied modification to the vanilla FTRL setup is that of 'optimism' (Syrgkanis et al., 2015; Rakhlin and Sridharan, 2013), where learners update their strategies taking into account their previously encountered payoffs in addition to the payoffs observed in the current timestep. This modification has led to algorithms that achieve faster time-average convergence (Daskalakis et al., 2021; Chen and Peng, 2020; Anagnostides et al., 2022) and even last-iterate convergence (Daskalakis and Panageas, 2018; Cai et al., 2024) in various game settings. Indeed, optimism can be applied to other algorithms in the literature (Daskalakis et al., 2011; Hsieh et al., 2021), often resulting in improved regret bounds.

Optimism is far from the only approach in the literature that can obtain sharper regret bounds in games. In line with the constant stepsize regime studied in this paper, several works have been able to obtain good regret bound with *absolute constant* stepsizes by modifying the standard FTRL algorithm. In particular, Bailey et al. (2020); Wibisono et al. (2022) study *alternating* variants of FTRL, while Piliouras et al. (2022) introduced a 'clairvoyant' version of MWU, both approaches resulting in good regret guarantees in their respective settings.

8. Symmetric Zero-Sum Games

8.1. Convergence of No-Regret Learning to Nash in Two-Player Zero-Sum Games

Recall the *duality gap* $\mathrm{DG}\colon \Delta_m \times \Delta_n \to \mathbb{R}$ for the game A is defined by, for all $(x_1, x_2) \in \Delta_m \times \Delta_n$:

$$\mathrm{DG}(x_1, x_2) \;:=\; \max_{x_1' \in \Delta_m} \langle x_1', Ax_2 \rangle - \min_{x_2' \in \Delta_n} \langle x_1, Ax_2' \rangle \;.$$

Note by construction, $\mathrm{DG}(x_1, x_2) \geq 0$, and $\mathrm{DG}(x_1, x_2) = 0$ if and only if (x_1, x_2) is an NE for the game A. Therefore, we can use the duality gap as a measure of convergence to NE.

Proposition 1 *Let $\tilde{x}_1^T := (\sum_{t=0}^T x_1^t)/T$ and $\tilde{x}_2^T := (\sum_{t=0}^T x_2^t)/T$ denote the time-averaged strategies of Players 1 and 2, respectively. If $\mathrm{Reg}(T) = o(T)$, then $(\tilde{x}_1^T, \tilde{x}_2^T)$ converges to an NE (x_1^*, x_2^*) in duality gap at a rate of $\mathrm{Reg}(T)/T = o(1)$.*

Proof By definition of $\mathrm{Reg}(T) = \mathrm{Reg}_1(T) + \mathrm{Reg}_2(T)$ from Section 2.1, we have

$$\mathrm{Reg}(T) \;=\; \max_{x_1 \in \Delta_m} \sum_{t=0}^{T} \langle x_1, Ax_2^t \rangle - \min_{x_2 \in \Delta_n} \sum_{t=0}^{T} \langle x_1^t, Ax_2 \rangle \;.$$

Then recalling that $\tilde{x}_1^T := (\sum_{t=0}^T x_1^t)/T$ and $\tilde{x}_2^T := (\sum_{t=0}^T x_2^t)/T$, observe that

$$\mathrm{DG}(\tilde{x}_1^T, \tilde{x}_2^T) \;=\; \max_{x_1 \in \Delta_m} \langle x_1, A\tilde{x}_2^T \rangle - \min_{x_2 \in \Delta_m} \langle \tilde{x}_1^T, Ax_2 \rangle \;=\; \frac{\mathrm{Reg}(T)}{T} \;.$$

Since $\mathrm{Reg}(T) = o(T)$ by assumption, the average iterate $(\tilde{x}_1^T, \tilde{x}_2^T)$ converges to an NE of A in duality gap at a rate of $\mathrm{Reg}(T)/T = o(1)$. ∎

8.2. Property of Nash Equilibria in Symmetric Zero-Sum Games

Proposition 2 *Let $A = -A^\top$ be a symmetric zero-sum game, and let x^* be an NE for A. Then $Ax^* = 0$ (where $0 \in \mathbb{R}^n$ is the all-zeros vector).*

Proof Let x^* be a Nash equilibrium for A. By definition of Nash, $Ax^* = c \cdot \mathbf{1}$ is a constant vector for some $c \in \mathbb{R}$. Since A is skew-symmetric, this implies that $0 = \langle x^*, Ax^* \rangle = \langle cx^*, \mathbf{1} \rangle = c$, so $Ax^* = 0$. ∎

8.3. Existence of Interior Equilibrium for High-Dimensional RPS

Proposition 4 *Let A be an n-dimensional RPS matrix from Definition 3 with positive constants $a_1, \ldots, a_n > 0$. Then A has an interior Nash equilibrium x^*.*

Proof Let A be an n-dimensional RPS matrix. At a Nash equilibrium, we have from Proposition 2 that $(Ax)_i = 0$ for all $i \in [n]$. This results in a system of n linear equations of the form:

$$a_{i-1}x_{i-1(\text{mod } n)} - a_i x_{i+1(\text{mod } n)} = 0$$

Moreover, x lies in the simplex, so $\sum_{i\in[n]} x_i = 1$. Thus, we have a system of $n+1$ linear equations and n variables. In such a system, there must always be a solution where all $x_i > 0$. Indeed, due to the fact that all $a_{i-1}, a_i > 0$ and $x_i \geq 0$ by definition, if we set some $x_i = 0$ then it follows that $x_i = 0$ for all $i \in [n]$, which violates the simplex constraint. Since a solution exists by consistency of the linear system, the solution x must be such that $x_i > 0$, implying it is interior. ■

Note that when n is odd, the interior Nash equilibrium is unique, and when n is even, there exists a continuum of interior Nash Equilibria (see (Sandholm, 2010, Chapter 9)).

9. Dual Dynamics of Fictitious Play and Gradient Descent

9.1. Properties of Conjugate Functions

Proposition 18 *Let $\phi : \Delta_n \to \mathbb{R}$ be a strictly convex regularizer. Let $\phi^* : \mathbb{R}^n \to \mathbb{R}$ and $Q : \mathbb{R}^n \to \Delta_n$ be the functions given by*

$$\begin{aligned} \phi^*(y) &= \max_{x\in\Delta_n} \langle x, y\rangle - \phi(x) \\ Q(y) &= \operatorname*{argmax}_{x\in\Delta_n} \langle x, y\rangle - \phi(x) \ . \end{aligned}$$

Then the following properties hold:

(i) ϕ^* is convex and continuously differentiable, and $\nabla\phi^*(y) = Q(y)$ for all $y \in \mathbb{R}^n$.

(ii) The map Q is surjective: for every $x \in \Delta_n$, there exists some $y \in \mathbb{R}^n$ such that $Q(y) = x$.

Proof Claim (i) follows from standard arguments of conjugate functions and using the strict convexity of ϕ^* over Δ_n (see e.g., Boyd and Vandenberghe (2004), Shalev-Shwartz et al. (2012), and Vandenberghe (2012)).

For Claim (ii), fix $x \in \Delta_n$, and let $y = \nabla\phi(x) \in \mathbb{R}^n$. We will show that $Q(y) = x$, which by the maximizing principle is equivalent to showing that $\phi^*(y) = \langle x, y\rangle - \phi(x)$. For this, let $f_y : \Delta_n \to \mathbb{R}$ be the function $f_y(x') = \langle x', y\rangle - \phi(x')$ for $x' \in \Delta_n$. Observe by the strict convexity of ϕ that f_y is strictly concave. Moreover, since we defined $y = \nabla\phi(x)$, then $\nabla f_y(x) = y - \nabla\phi(x) = 0$. Thus by concavity, $x \in \Delta_n$ is the unique maximizer of f_y, and $\max_{x'\in\Delta_n} f_y(x') = \langle x, y\rangle - \phi(x)$. Toegether, we have

$$\phi^*(y) \ = \ \max_{x'\in\Delta_n} \langle x', y\rangle - \phi(x') \ = \ \max_{x'\in\Delta_n} f_y(x') \ = \ \langle x, y\rangle - \phi(x) \ ,$$

which implies that $Q(y) = x$, as desired. ■

Note that Claim (ii) of Proposition 18 is equivalent to saying that $\nabla\phi^*(\nabla\phi(x)) = x$ for every $x \in \Delta_n$. However, $\nabla\phi^*$ is in general not the inverse of $\nabla\phi$, as $Q(y) = \nabla\phi^*(y)$ is not necessarily injective. For example, when $\phi = (\|\cdot\|_2^2)/2$ as in Gradient Descent, the preimage under $\nabla\phi^*$ of any x on the boundary of Δ_n may consist of multiple $y \in \mathbb{R}^n$. On the other hand, this inverse property does hold for the case of Legendre regularizers (such as negative entropy) as studied by Wibisono et al. (2022).

9.2. Properties of Bregman Divergences

Definition 19 (Bregman Divergence) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then D_f is the Bregman Divergence of f, where for all $x, x' \in \mathbb{R}^n$:*

$$D_f(x', x) \;=\; f(x') - f(x) - \langle \nabla f(x), x' - x \rangle \;.$$

Note that when f is convex, we have $D_f(x, x') \geq 0$ for all $x, x' \in \mathbb{R}^n$. We recall that D_f satisfies the following *three-point identity* (see e.g., Wibisono et al. (2022)).

Proposition 20 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then for any $w, x, y, z \in \mathbb{R}^n$: Then for any $x, y, z \in \mathbb{R}^n$:*

$$\langle \nabla f(z) - \nabla f(y), x - z \rangle \;=\; D_f(x, y) - D_f(x, z) - D_f(z, y) \;.$$

9.3. Energy-Based Regret Bounds for Fictitious Play and FTRL

Proposition 6 *Let $\{y^t\}$ be the dual iterates of FP (with $\eta = 1$). Then* $\mathrm{Reg}(T) = 2 \cdot \Psi(y^{T+1})$.

Proof Recall from expression (1) that $\mathrm{Reg}(t) = 2 \cdot \max_{x \in \Delta_n} \sum_{t=0}^{T} \langle x, Ax^t \rangle$. Moreover, we have by definition of (FP Primal-Dual) that $y^{t+1} = \sum_{k=0}^{t} Ax^k$ for each t, which means $y^{T+1} = \sum_{k=0}^{T} Ax^k$. Thus we can write

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

where the second equality follows from the definition of Ψ from expression (3). ■

Proposition 7 *Let $\{y^t\}$ be the dual iterates of FTRL with $\eta > 0$. Let $M = \max_{x \in \Delta_n} \phi(x)$. Then:*

$$\mathrm{Reg}(T) \;\leq\; \frac{2 \cdot \phi^*(y^{T+1})}{\eta} + \frac{2M}{\eta} \;.$$

Proof The proof of the proposition follows similarly to Wibisono et al. (2022, Theorem C.2), but specialized to the present symmetric learning setting. For this, recall by definition of (FTRL) and (Dual Vector) that the primal and dual iterates evolve as

$$\begin{aligned} y^{t+1} \;&=\; y^t + \eta A x^t \\ x^{t+1} \;&=\; \operatorname*{argmax}_{x \in \Delta_n} \langle x, y^{t+1} \rangle - \phi(x) \;, \end{aligned}$$

where $y^0 = 0$ is the zero vector. By claim (i) of Proposition 18, we then have for each t that $x^t = \nabla\phi^*(y^t)$. Now observe by definition of $\text{Reg}(T)$ from expression (1) and by the skew-symmetry of A, we can write

$$\text{Reg}(T) \;=\; 2 \cdot \max_{x' \in \Delta_n} \sum_{t=0}^{T} \langle x', Ax^t \rangle \;=\; 2 \cdot \max_{x' \in \Delta_n} \sum_{t=0}^{T} \langle x' - x^t, Ax^t \rangle \; . \tag{8}$$

Our goal will then be to derive a uniform upper bound on $\sum_{t=0}^{T} \langle x' - x^t, Ax^t \rangle$ over all $x' \in \Delta^n$. For this, fix some $x \in \Delta_n$, and let $y \in \mathbb{R}^n$ be a vector satisfying $\nabla\phi^*(y) = x$ and $\phi^*(y) = \langle x, y\rangle - \phi(x)$, which we know must exist from claim (ii) of Proposition 18. Then at each time t, we use the three-point identity of Bregman divergences (Proposition 20) and the fact that $y^{t+1} - y^t = \eta A x^t$ to write

$$\begin{aligned} \langle x - x^t, Ax^t \rangle \;&=\; \frac{1}{\eta}\big\langle \nabla\phi^*(y) - \nabla\phi^*(y^t), y^{t+1} - y^t \big\rangle \\ &=\; \frac{1}{\eta}\big(D_{\phi^*}(y^{t+1}, y^t) + D_{\phi^*}(y^t, y) - D_{\phi^*}(y^{t+1}, y)\big) \; . \end{aligned}$$

Then summing over all t and telescoping, we find

$$\begin{aligned} \sum_{t=0}^{T} \langle x - x^t, Ax^t \rangle \;&=\; \frac{1}{\eta}\sum_{t=0}^{T} D_{\phi^*}(y^{t+1}, y^t) + \frac{1}{\eta}\sum_{t=0}^{T} \big(D_{\phi^*}(y^t, y) - D_{\phi^*}(y^{t+1}, y)\big) \\ &=\; \frac{1}{\eta}\sum_{t=0}^{T} D_{\phi^*}(y^{t+1}, y^t) + \frac{1}{\eta}\big(D_{\phi^*}(y^0, y) - D_{\phi^*}(y^{T+1}, y)\big) \\ &\le\; \frac{1}{\eta}\sum_{t=0}^{T} D_{\phi^*}(y^{t+1}, y^t) + \frac{1}{\eta} D_{\phi^*}(y^0, y) \; , \end{aligned} \tag{9}$$

where the final inequality comes from the non-negative of $D_{\phi^*}(y^{T+1}, y)$ given that ϕ^* is convex. Observe further that we can write

$$\begin{aligned} D_{\phi^*}(y^0, y) \;&=\; \phi^*(y^0) - \phi^*(y) - \langle \nabla\phi^*(y), y^0 - y \rangle \\ &=\; \phi^*(y^0) - (\langle x, y\rangle - \phi(x)) - \langle x, y^0 - y\rangle \\ &=\; \phi^*(y^0) + \phi(x) - \langle x, y\rangle + \langle x, y\rangle - \langle x, y^0\rangle \\ &=\; \phi^*(y^0) + \phi(x) \; . \end{aligned} \tag{10}$$

Here, we use in the second equality that $\phi^*(y) = \langle x, y\rangle - \phi(x)$ and $\nabla\phi^*(y) = x$ by definition of x and y, and in the last equality, we use the fact that $y^0 = 0$.

Now observe also for each t that

$$\begin{aligned} D_{\phi^*}(y^{t+1}, y^t) \;&=\; \phi^*(y^{t+1}) - \phi^*(y^t) - \langle \nabla\phi^*(y^t), y^{t+1} - y^t \rangle \\ &=\; \phi^*(y^{t+1}) - \phi^*(y^t) - \eta\langle x^t, Ax^t \rangle \\ &=\; \phi^*(y^{t+1}) - \phi^*(y^t) \; , \end{aligned} \tag{11}$$

where the final two equalities follow by definition of x^t and the dual update rule, and by the skew-symmetry of A. Then substituting expressions (10) and (11) back into (9) and summing, we find

$$\begin{aligned} \sum_{t=0}^{T} \langle x - x^t, Ax^t \rangle &\leq \frac{1}{\eta} \sum_{t=0}^{T} \left(\phi^*(y^{t+1}) - \phi^*(y^t) \right) + \frac{\phi^*(y^0) + \phi(x)}{\eta} && (12) \\ &= \frac{\phi^*(y^{T+1}) - \phi^*(y^0)}{\eta} + \frac{\phi^*(y^0) + \phi(x)}{\eta} && (13) \\ &= \frac{\phi^*(y^{T+1})}{\eta} + \frac{\phi(x)}{\eta} && (14) \end{aligned}$$

Maximizing over all $x \in \Delta^n$, substituting the inequality into (8), and recalling the definition of M from the statement of the proposition then yields the desired claim. ■

9.4. Dual Dynamics as Skew-Gradient Descent

Proposition 8 *Let $\{y^t\}$ be the dual iterates of either Fictitious Play ($\eta = 1$) or FTRL ($\eta > 0$), and let H be its corresponding energy function from (3). Then for every $t \geq 1$, it holds that*

$$y^{t+1} = y^t + \eta A \partial H(y^t) . \tag{4}$$

In particular: for FP, each $x^t \in \partial \Psi(y^t)$, and for FTRL, each $x^t = \nabla \phi^*(y^t)$. Moreover, for all $t \geq 1$: $\Delta H(y^t) = H(y^{t+1}) - H(y^t) \geq 0$.

Proof Recall for a convex function $H : \mathbb{R}^n \to \mathbb{R}$, that its subgradient set at $y \in \mathbb{R}^n$ is given by

$$\partial H(y) = \{ g \in \mathbb{R}^n : \forall z \in \mathbb{R}^n, H(z) \geq H(y) + \langle g, z - y \rangle \} .$$

For Fictitious Play, the subgradient set of the energy function $\Psi(y) = \max_{x \in \Delta_n} \langle x, y \rangle$ is the set of maximizers $\partial \Psi(y) = \operatorname{argmax}_{x \in \Delta_n} \langle x, y \rangle$. Thus by the definition of Fictitious Play from (2), $x^t \in \partial \Psi(y^t)$ for all $t \geq 1$. For FTRL with strictly convex regularizer ϕ, Proposition 18 implies for all $t \geq 1$ that $x^t = \nabla \phi^*(y^t)$, and thus $x^t \in \partial \phi^*(y^t)$ by definition.

Letting H be the energy function for either Fictitious Play or FTRL and $\{x^t\}$ and $\{y^t\}$ the corresponding iterates, then it follows from (Dual Vector) that

$$y^{t+1} = y^t + \eta A x^t = y^t + \eta A \partial H(y^t) ,$$

as claimed. For the second statement, observe by definition of the subgradient set of H and the fact that $x^t \in \partial H(y^t)$,

$$H(y^{t+1}) - H(y^t) \geq \langle x^t, y^{t+1} - y^t \rangle = \eta \langle x^t, Ax^t \rangle = 0 ,$$

where the final equality follows by the skew-symmetry of A. ■

Skew-gradient flow. The continuous-time limit (i.e., the limit as $\eta \to 0$) of the skew-gradient descent algorithm (4) is the following skew-gradient flow dynamics:

$$\dot{Y}_t = A\nabla H(Y_t).$$

Here, we assume H is differentiable for simplicity. Since A is skew-symmetric, the vector field $A\nabla H(Y_t)$ is orthogonal to the level set of H, and thus the dynamics conserves the energy function H. Concretely, we can compute:

$$\frac{d}{dt}H(Y_t) \;=\; \langle \nabla H(Y_t), \dot{Y}_t\rangle \;=\; \langle \nabla H(Y_t), A\nabla H(Y_t)\rangle \;=\; 0$$

so $H(Y_t) = H(Y_0)$ for all $t \geq 0$.

We note that for a zero-sum game with a general payoff matrix, it is the *joint strategy* of the two players in the dual space that becomes a skew-gradient flow in continuous time (see e.g., Wibisono et al. (2022, Section 3.1)). In the present setting of symmetric learning with a skew-symmetric payoff matrix, the strategy of each player (which is the same for both players) itself follows the skew-gradient flow in the dual space in continuous time, as discussed above.

10. Worst-Case Regret Bound for FP on High-Dimensional RPS

In this section, we develop the proof of Theorem 11, which shows Fictitious Play obtains worst-case $O(\sqrt{T})$ regret on every n-dimensional RPS matrix.

10.1. Cycling Behavior

We begin by establishing the primal cycling behavior of Fictitious Play:

Lemma 10 *Let t_0 be the first $t > 0$ where $\Psi(y^{t_0}) > \Psi(y^1)$, and suppose $x^{t_0} = e_i$ for some $i \in [n]$. Then $x^{t_k} = e_{i+k \pmod n}$ for all $k \geq 1$.*

For this, we first state prove the following proposition, which characterizes the trajectory of the dual iterates $\{y^t\}$ for a sequence of consecutive primal iterates all at the same vertex e_i:

Proposition 21 *Suppose at time t that $x^t = e_i$ and $y^t_i \geq y^t_{i-1} > y^t_j$ for all other $j \in [n] \setminus \{i, i+1\}$. Let $\tau := \left\lceil \frac{y_i^t - y_{i+1}^t}{a_i} \right\rceil$. Then $y^{t+s} = y^t + s \cdot A_i$ for all $1 \leq s \leq \tau$.*

Proof Without loss of generality (and for readability), assume $1 < i < n$, and thus $i+1 \pmod n = i+1$ and $i-1 \pmod n = i-1$. Observe from (FP Primal-Dual) that for any time ℓ such that $x^\ell = e_i$, we have $\Delta y^\ell = y^{\ell+1} - y^\ell = A_i$, whose coordinates are given by

$$\Delta y^\ell_k = A_{k,i} := \begin{cases} -a_{i-1} & \text{if } k = i-1 \\ +a_i & \text{if } k = i+1 \\ 0 & \text{otherwise .} \end{cases} \tag{15}$$

This implies that at time $t+1$, we have $y_i^{t+1} = y_i^t$, $y_{i+1}^{t+1} = y_{i+1}^t + a_i$, $y_{i-1}^{t+1} = y_{i-1}^t - a_{i-1}$, and $y_j^{t+1} = y_j^t$ for all other $j \in [n] \setminus \{i-1, i, i+1\}$.

If $y_i^t - y_{i+1}^t \leq a_i$, then the statement of the proposition trivially holds for $\tau = 1$. Otherwise, $y_i^t - y_{i+1}^t > a_i$, and we have $y_{i+1}^{t+1} < y_i^{t+1}$ and also $y_{i-1}^{t+1} < y_{i-1}^t \leq y_i^t = y_i^{t+1}$. Thus for $\tau \geq 1$, we have $y_i^{t+1} > \max_{j \neq i} y_j^{t+1}$ which means $x^{t+1} = e_i$ and $\Delta y^{t+1} = A_i$. It follows by induction that for $s \geq 1$, we have $x^{t+s} = e_i$ (and $\Delta y^{t+s} = A_i$), so long as

$$y_{i+1}^t + s \cdot a_i < y_i^{t+s} = y_i^t \qquad \Longleftrightarrow \qquad s < \frac{y_i^t - y_{i+1}^t}{a_i} \ .$$

Recall that $\tau := \left\lceil \frac{y_i^t - y_{i+1}^t}{a_i} \right\rceil$, and thus the latter condition holds for all $s \leq \tau - 1$. It then follows for all $1 \leq s \leq \tau$ that $y^{t+s} = y^t + s \cdot A_i$, which yields the statement of the proposition. ■

We now use Proposition 21 to prove the cycling property of Lemma 10:

Proof (of Lemma 10) We will prove the claim by induction on k. First, we show that $x^{t_1} = e_{i+1}$. For this, observe by definition of t_0 that $\Psi(y^{t_0}) > \Psi(y^1)$. Since $x^{t_0} = e_i$ by assumption, this implies that $y_i^{t_0} > \max_{j \neq i} y_j^{t_0}$. Then setting $\tau = \left\lceil \frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i} \right\rceil$, Proposition 21 implies that $x^{t_0+s} = e_i$ for all $1 \leq s \leq \tau - 1$ and that $y^{t_0+\tau} = y^{t_0} + \tau A_i$. This means that $t_1 > t_0 + \tau - 1$, and also

$$\begin{aligned}
y_i^{t_0+\tau} &= y_i^{t_0} \\
y_{i+1}^{t_0+\tau} &= y_{i+1}^{t_0} + \tau \cdot a_i \\
y_j^{t_0+\tau} &\leq y_j^{t_0} \ < \ y_i^{t_0} \ = \ y_i^{t_0+\tau} \quad \text{for all other } j \neq i \neq i+1 \ .
\end{aligned}$$

From here, we separate the argument into two cases: first, when $\frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i} \notin \mathbb{N}$, and second when $\frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i} \in \mathbb{N}$. In the first case, the condition $\frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i} \notin \mathbb{N}$ implies that $\tau > \frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i}$ and thus

$$\begin{aligned}
y_{i+1}^{t_0+\tau} &= y_{i+1}^{t_0} + \tau \cdot a_i \\
&> y_{i+1}^{t_0} + \left(\frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i} \right) \cdot a_i \ = \ y_i^{t_0} \ = \ y_i^{t_0+\tau} \ .
\end{aligned}$$

It follows that $y_{i+1}^{t_0+\tau} > \max_{j \neq i+1} y_j^{t_0+\tau}$, which means $x^{t_0+\tau} = e_{i+1}$, and thus the desired claim holds for $t_1 = t_0 + \tau$.

In the second case, when $\frac{y_i^{t_0} - y_{i+1}^{t_0}}{a_i} \in \mathbb{N}$, we instead have $y_{i+1}^{t_0+\tau} = y_{i+1}^{t_0} + \tau \cdot a_i = y_i^{t_0} = y_i^{t_0+\tau}$. Thus in this case, a tiebreaking rule determines whether $x^{t_0+\tau} = e_i$ or $x^{t_0+\tau} = e_{i+1}$. If $x^{t_0+\tau} = e_{i+1}$, then the desired claim again holds for $t_1 = t_0 + \tau$. Moreover, we have $y_{i+1}^{t_0+\tau} \geq y_i^{t_0+\tau} > y_j^{t_0+\tau}$ for all other $j \in [n] \setminus \{i, i+1\}$.

On the other hand, if $x^{t_0+\tau} = e_i$, then $y^{t_0+\tau+1} = y^{t_0+\tau} + A_i$, meaning:

$$\begin{aligned}
y_i^{t_0+\tau+1} &= y_i^{t_0+\tau} \\
y_{i+1}^{t_0+\tau+1} &= y_{i+1}^{t_0+\tau} + a_i \ > \ y_i^{t_0+\tau} \ = \ y_i^{t_0+\tau+1} \\
y_j^{t_0+\tau+1} &\leq y_j^{t_0+\tau} \ < \ y_i^{t_0+\tau} \ = \ y_i^{t_0+\tau+1} \quad \text{for all other } j \in [n] \setminus \{i, i+1\} \ .
\end{aligned}$$

It follows in this case that $y_{i+1}^{t_0+\tau+1} > \max_{j \neq i+1} y_j^{t_0+\tau+1}$, which means $x^{t_0+\tau+1} = e_{i+1}$, and thus the desired claim is satisfied with $t_1 = t_0 + \tau + 1$.

In all cases, we find that $x^{t_1} = e_{i+1}$, and that $y^{t_1}_{i+1} \geq y^{t_1}_i > y^{t_1}_j$ for all other $j \neq i \neq i+1$. Now assume the claim holds up to phase k, meaning $y^{t_k}_{i+k \text{ (mod } n)} > y^{t_k}_j$ for all other j. Then setting $\tau = \left\lceil \frac{y^{t_k}_{i+k \text{ (mod } n)} - y^{t_k}_{i+k+1 \text{ (mod } n)}}{a_{i+k \text{ (mod } n)}} \right\rceil$ and using identical calculations via an application of Proposition 21, we find that $x^{t_{k+1}} = e_{i+k+1 \text{ (mod } n)}$, and thus the claim also holds at phase $k+1$. By induction, this proves the statement of the lemma. ■

10.2. Energy Growth Bound

Proposition 12 *For any t, define $\Delta\Psi(y^t) := \Psi(y^{t+1}) - \Psi(y^t)$. Then:*

(a) If $x^t = x^{t+1}$, then $\Delta\Psi(y^t) = 0$.

(b) If $x^t \neq x^{t+1}$, then $0 \leq \Delta\Psi(y^t) \leq a_{\max}$.

Proof For the first case, assume $x^t = x^{t+1}$. Recall that $y^{t+1} = y^t + Ax^t$, and using the definition of Ψ and the maximizing principle: $\Psi(y^{t+1}) = \langle x^{t+1}, y^{t+1}\rangle$ and $\Psi(y^t) = \langle x^t, y^t\rangle$. Together, this means:

$$\begin{aligned}\Psi(y^{t+1}) - \Psi(y^t) &= \langle x^{t+1}, y^{t+1}\rangle - \langle x^t, y^t\rangle \\ &= \langle x^t, y^t + Ax^t\rangle - \langle x^t, y^t\rangle \\ &= \langle x^t, y^t\rangle + \langle x^t, Ax^t\rangle - \langle x^t, y^t\rangle \;=\; 0\;.\end{aligned}$$

Here, the final equality is due to the skew-symmetry of A (and thus $\langle x, Ax\rangle = 0$ for all $x \in \mathbb{R}^n$).

For the second case, where $x^t \neq x^{t+1}$, suppose $x^t = e_i$ and $x^{t+1} = e_j$ for some $i \neq j \in [n]$. Then observe that

$$\begin{aligned}\Psi(y^{t+1}) - \Psi(y^t) &= \langle x^{t+1}, y^{t+1}\rangle - \langle x^t, y^t\rangle \\ &= \langle e_j, y^{t+1}\rangle - \langle e_i, y^t\rangle \\ &= y^{t+1}_j - y^t_i\;.\end{aligned}$$

Let $\alpha^t = y^{t+1}_j - y^t_i$. By definition of (FP Primal-Dual) and the assumption that $x^t = e_i$, we have

$$\begin{aligned}y^{t+1}_j &= y^t_j + (Ax^t)_j &&= y^t_j + A_{ji}\;. \\ y^{t+1}_i &= y^t_i + (Ax^t)_i &&= y^t_i + A_{ii} = y^t_i\;.\end{aligned}$$

It follows that $\alpha^t \;=\; y^{t+1}_j - y^t_i \;=\; y^t_j + A_{ji} - y^t_i \;\leq\; y^t_i + A_{ji} - y^t_i = A_{ji} \leq a_{\max}$. ■

10.3. Phase Length Bound

Lemma 13 *For $k = 1, 2, \ldots, K$, let $\gamma_k = \Psi(y^{t_k})$ be the energy at the start of Phase k. Then $\tau_k \geq \alpha_k \cdot \gamma_k - \beta_k$, where $\alpha_k > 0$ and $\beta_k > 0$ are absolute constants.*

The cycling property of the primal iterates helps to establish this phase length bound in the following way: supposing $x^{t_k} = e_i$, then we show roughly $\tau_k \approx \gamma_k - y_{i+1}^{t_k}$. Using the fixed cycling order, and using the fact that energy increases by at most a constant between phases, we can deduce that the magnitude of $y_{i+1}^{t_k}$ is at most a constant fraction of γ_k, and the claim follows.

Proof We prove the claim by induction on k. First, recall from Lemma 10 that for any $\ell \geq 0$, $x^{t_\ell} = e_{i+\ell \text{ (mod } n)}$ and $x^{t_{\ell+1}} = e_{i+\ell+1 \text{ (mod } n)}$. By Proposition 21 and the fact that $y^{t_\ell+1} = y^{t_\ell} + Ax^{t_\ell}$, this implies that τ_ℓ must satisfy:

$$y_{i+\ell+1 \text{ (mod } n)}^{t_{\ell+1}} \;=\; y_{i+\ell+1 \text{ (mod } n)}^{t_\ell} + \tau_\ell \cdot a_{i+\ell \text{ (mod } n)} \;\geq\; \gamma_\ell \quad\implies\quad \tau_\ell \;\geq\; \frac{\gamma_\ell - y_{i+\ell+1 \text{ (mod } n)}^{t_\ell}}{a_{i+\ell \text{ (mod } n)}} \;.$$

For phases $\ell = 0, \ldots, n$, note first that by definition of (FP Primal-Dual), $y^1 = Ax^0$. Thus the coordinates of y^1 are absolute constants. Moreover, as $\Psi(y^{t_0}) - \Psi(y^0) < O(1)$ due to Proposition 12, then each coordinate of y^{t_0} is also an absolute constant. By the cycling property of Lemma 10, and using expression (15), observe that for each coordinate $i + \ell + 1 \pmod n$ for $\ell = 0, \ldots, n$, we must have $\Delta y_{i+\ell+1 \text{ (mod } n)}^t \leq 0$ until $t \geq t_\ell$. It follows that $y_{i+\ell+1 \text{ (mod } n)}^{t_\ell} \leq \kappa$ for some absolute constant $\kappa > 0$, and thus $\tau_\ell \geq (\gamma_\ell - \kappa)/a_{i+\ell \text{ (mod } n)}$. Thus for each phase $0 \leq \ell \leq n$, the statement of the lemma holds for $\alpha_\ell = 1/a_{i+\ell \text{ (mod } n)} > 0$ and $\beta_\ell = \kappa/a_{i+\ell \text{ (mod } n)}$.

Now assume that the claim holds for all phases up to $k-1$. We will prove it holds also for phase $k > n$. For this, recall that we have

$$\tau_k \;\geq\; \frac{\gamma_k - y_{i+k+1 \text{ (mod } n)}^{t_k}}{a_{i+k \text{ (mod } n)}} \;. \tag{16}$$

Now by the cycling property of Lemma 10 and using expression (15), it follows that $\Delta y_{i+k+1 \text{ (mod } n)}^t = 0$ for all $t_{k+3-n} \leq t < t_k$, and thus $y_{i+k+1 \text{ (mod } n)}^{t_k} = y_{i+k+1 \text{ (mod } n)}^{t_{k+3-n}}$. Moreover, we also have from Proposition 21 that

$$\begin{aligned} y_{i+k+1 \text{ (mod } n)}^{t_k} \;=\; y_{i+k+1 \text{ (mod } n)}^{t_{k+3-n}} \;&=\; y_{i+k+1 \text{ (mod } n)}^{t_{k+2-n}} - a_{i+k+1 \text{ (mod } n)} \cdot \tau_{k+2-n} \\ &\leq\; \gamma_{k+2-n} + a_{\max} - a_{i+k+1 \text{ (mod } n)} \cdot \tau_{k+2-n} \;, \end{aligned} \tag{17}$$

where $a_{\max} > 0$ is an absolute constant, and where the inequality follows from the energy growth bound of Proposition 12. By the inductive hypothesis, we have $\tau_{k+2-n} \geq \alpha_{k+2-n} \cdot \gamma_{k+2-n} - \beta_{k+2-n}$ for absolute constants $\alpha_{k+2-n}, \beta_{k+2-n} > 0$. Substituting this into (17), we then find:

$$y_{i+k+1 \text{ (mod } n)}^{t_k} \;\leq\; \gamma_{k+2-n} - a_{i+k+1 \text{ (mod } n)} \cdot (\alpha_{k+2-n} \cdot \gamma_{k+2-n} - \beta_{k+2-n}) + a_{\max} \tag{18}$$

$$\;=\; \gamma_{k+2-n} \cdot (1 - a_{i+k+1 \text{ (mod } n)} \cdot \alpha_{k+2-n}) + \beta_{k+2-n} + a_{\max} \;. \tag{19}$$

Without loss of generality, we will assume that expression (19) is positive. Otherwise, we would trivially have in expression (16) that $\tau_k \geq \gamma_k/a_{i+k \text{ (mod } n)}$, which would yield the desired claim. Similarly, we assume additionally that $1 - a_{i+k+1 \text{ (mod } n)} \cdot \alpha_{k+2-n} > 0$, as otherwise the bound in expression (19) is at most an absolute constant $\epsilon > 0$, which would imply $\tau_k \geq (\gamma_k - \epsilon)/a_{i+k \text{ (mod } n)}$, again trivially yielding the desired claim. Then under these assumptions, we can substitute expression (19) back into (16) to find

$$\tau_k \;\geq\; \frac{\gamma_k - \gamma_{k+2-n} \cdot (1 - a_{i+k+1 \pmod n} \cdot \alpha_{k+2-n})}{a_{i+k \pmod n}} - \frac{\beta_{k+2-n} + a_{\max}}{a_{i+k \pmod n}} \tag{20}$$

$$\geq\; \frac{\gamma_k \cdot a_{i+k+1 \pmod n} \cdot \alpha_{k+2-n}}{a_{i+k \pmod n}} - \frac{\beta_{k+2-n} + a_{\max}}{a_{i+k \pmod n}} \;, \tag{21}$$

where the final inequality comes from the fact that $\gamma_{k+2-n} \leq \gamma_k$ and the assumption that $1 - a_{i+k+1 \pmod n} \cdot \alpha_{k+2-n} > 0$. It follows that $\tau_k \geq \alpha_k \gamma_k - \beta_k$ for

$$\alpha_k \;:=\; \frac{a_{i+k+1 \pmod n} \cdot \alpha_{k+2-n}}{a_{i+k \pmod n}} > 0 \quad \text{and} \quad \beta_k \;:=\; \frac{\beta_{k+2-n} + a_{\max}}{a_{i+k \pmod n}} > 0 \;,$$

which completes the proof. ■

10.4. Proof of Theorem 11

Recall from Definition 9 that $K \geq 1$ is the total number of phases undergone by the dynamics in T rounds. For each phase $k = 0, 1 \ldots, K$ define the indicator variable $c_k \in \{0, 1\}$ as follows:

$$c_k \;:=\; \begin{cases} 0 & \text{if } \Psi(y^{t_k}) = \Psi(y^{t_{k-1}}) \\ 1 & \text{if } \Psi(y^{t_k}) > \Psi(y^{t_{k-1}}) \end{cases} .$$

In other words, $c_k = 1$ if and only if the energy of the dynamics strictly increases from phase $k-1$ to phase k. Observe that the energy growth bound of Proposition 12 further implies that if $c_k = 1$, then $\Psi(y^{t_k}) - \Psi(y^{t_{k-1}}) = \Theta(1) = \Theta(c_k)$. This yields the following corollary to Proposition 12:

Corollary 22 *Assume the setting of Theorem 11. Then* $\mathrm{Reg}(T) = 2 \cdot \Psi(y^{T+1}) = \Theta\Big(\sum_{k=1}^{K} c_k \Big)$.

We can now prove the main result of Theorem 11. In particular, the use of the indicators $\{c_k\}$ allows for handling the case of non-increasing energy between phases, as mentioned in the sketch presented in Section 4.

Proof For readability, we will use the notation $f \lesssim g$ and $f \gtrsim g$ to indicate $f = O(g)$ and $f = \Omega(g)$ respectively. Now without loss of generality, assume $t_0 < T$, as otherwise $\Psi(y^T) = \Psi(y^1)$ is a constant, and the statement of the theorem trivially holds. Then by Corollary 22, our goal is to derive an upper bound on $\sum_{k=0}^{K} c_k$. For this, recall from Definition 9 that $T = \sum_{k=0}^{K} \tau_k$. By Lemma 13, we have for all k that:

$$\tau_k \;\geq\; \alpha_k \cdot \Psi(y^{t_k}) - \beta_k \;\gtrsim\; \sum_{i=1}^{k} c_i - \beta_k \;,$$

for positive absolute constants $\alpha_k, \beta_k > 0$. It follows that we can write

$$T \;=\; \sum_{k=0}^{K} \tau_k \;\gtrsim\; \sum_{k=0}^{K} \sum_{i=1}^{k} c_i - \sum_{k=0}^{K} \beta_k \;\gtrsim\; \sum_{k=0}^{K} \sum_{i=1}^{k} c_i - T \;, \tag{22}$$

where the final inequality comes from the fact that $K \leq T$. Observe that we can also write

$$\sum_{k=0}^{K}\sum_{i=1}^{k} c_i \;=\; Kc_1 + (K-1)c_2 + \cdots + c_K \;=\; \sum_{k=0}^{K}(K-k+1)c_k \tag{23}$$

Now let L be the set of indices in $\{0, 1, \ldots, K\}$ where $c_i = 1$, and let $\mathbb{1}_L$ be the indicator function of membership to the set L. Since there are $|L|$ non-zero elements of $\{c_i\}$, we can write:

$$\sum_{k=0}^{K}(K-k+1)c_k \;=\; |L| \cdot K - \sum_{k=0}^{K}(\mathbb{1}_L \cdot k) + |L| \tag{24}$$

The second term is the sum of indices in L, which can be upper-bounded by the sum of the top-$|L|$ indices. This means:

$$|L| \cdot K - \sum_{k=0}^{K}(\mathbb{1}_L \cdot k) + |L| \;\geq\; |L| \cdot K - \frac{|L|}{2}(2K - |L|) + |L| \tag{25}$$

$$= \frac{|L|^2}{2} + |L| \;\geq\; \frac{1}{2}|L|^2 \,. \tag{26}$$

By definition, $|L| = \sum_{k=0}^{K} c_k$, and thus we conclude $\sum_{k=0}^{K}\sum_{i=1}^{k} c_i \gtrsim (|L|)^2 = \Big(\sum_{k=0}^{K} c_k\Big)^2$. Substituting this into expression (22) and rearranging, we then obtain via Corollary 22:

$$\Big(\sum_{k=0}^{K} c_k\Big)^2 \;\lesssim\; T \quad\implies\quad \mathrm{Reg}(T) \;\lesssim\; \sum_{k=0}^{K} c_k \;\lesssim\; \sqrt{T}\,,$$

which completes the proof. ■

11. Constant Regret for FP via Tournament Tiebreaking

The cycling behavior of Lemma 10 also leads to an improved regret bound under the following "tournament" tiebreaking rule, which mirrors the cyclical tournament graph structure latent in the definition of generalized RPS:

Definition 23 (Tournament Tiebreaking Rule) *Let A be an n-dimensional RPS matrix from Definition 3. Using the tournament tiebreaking rule, ties between coordinates $i \in [n]$ and $j \in [n]$ are broken lexicographically, except for ties between coordinates 1 and n, which are broken in favor of coordinate 1.*

For *unweighted* RPS matrices, we show in Theorem 24 that when x^0 is a vertex $e_i \in \Delta_n$, the energy of the dual iterates is exactly conserved over time, and the regret is thus constant.

Theorem 24 *For any $n \geq 3$, let A be an n-dimensional RPS matrix from Definition 3 with all $a_i = 1$. Let $\{x^t\}$ and $\{y^t\}$ be the iterates of (FP Primal-Dual) initialized at $x^0 = e_i$ for some $i \in [n]$. Then using the tournament tiebreaking rule, $\Psi(y^T) = \Psi(y^1)$ and $\mathrm{Reg}(T) = O(1)$.*

Proof We will show that under the tournament tiebreaking rule of Definition 3, for any t such that $x^t \neq x^{t+1}$, the energy $\Psi(y^t) = \Psi(y^{t+1})$. First, since all $a_i = 1$ in the unweighted case, we have that $y_i^1 = Ax_i^0 \in \mathbb{Z}$ for $i \in [n]$. Recall that by definition, $y^t = y^1 + \sum_{k=0}^{t-1} Ax^k$. Since, under (FP Primal-Dual), each $x^k = e_j$ for some $j \in [n]$, it follows for each t that we can write $y^t = y^1 + \sum_{i=1}^{n} c_i^t \cdot A_i$ for non-negative integer constants $c_1^t, \ldots, c_n^t$. Altogether, we have $y_i^t = y_i^1 + c_{i-1}^t - c_{i+1}^t = y_i^1 + d_i^t$ for any $i \in [n]$, where $d_i^t \in \mathbb{Z}$. Thus for any $i, j \in [n]$, we have: $y_i^t - y_j^t = y_i^1 - y_j^1 + d_i^t - d_j^t$. As argued above, $y_i^1 - y_j^1 \in \mathbb{Z}$ and thus $y_i^t - y_j^t \in \mathbb{Z}$.

Let $i := \operatorname{argmax}_{j\in[n]} y_j^t$ using the tournament tiebreaking rule of Definition 23, and set $\tau = \lceil y_i^t - y_{i+1}^t \rceil$. Without loss of generality, assume $i < n$, so that $i + 1 \pmod n = i + 1$. Then using identical calculations as in Proposition 21, we have

$$
\begin{aligned}
& y_i^{t+\tau} &=&\; y_i^t \\
& y_{i+1}^{t+\tau} &=&\; y_{i+1}^t + \lceil y_i^t - y_{i+1}^t \rceil \\
\text{and} \quad & y_j^{t+\tau} &\leq&\; y_j^t && \text{for all other } j \neq i \neq i+1 \,.
\end{aligned}
$$

Moreover, we also have $x^{t+s} = e_i$ and $\Psi(y^{t+s}) = y_i^{t+s} = \Psi(y^t)$ for all $1 \leq s \leq \tau - 1$. As $y_i^t - y_{i+1}^t \in \mathbb{Z}$ by the arguments above, we have $\lceil y_i^t - y_{i+1}^t \rceil \in \mathbb{Z}$ and thus $y_{i+1}^{t+\tau} = y_{i+1}^t + y_i^t - y_{i+1}^t = y_i^t = y_i^{t+\tau}$. Using the tournament tiebreaking rule, it follows that $x^{t+\tau} = e_{i+1} \neq e_i = x^{t+\tau-1}$. Moreover, $\Psi(y^{t+\tau}) = \Psi(y^{t+\tau-1}) = \cdots = \Psi(y^t) = \cdots = \Psi(y^1)$. Thus by Proposition 6 we have $\mathrm{Reg}(T) = 2 \cdot \Psi(y^T) = O(1)$, which proves the claim. ■

12. Details on Gradient Descent Under Symmetric Learning

12.1. Closed Form Primal Update and Energy Function

Under (GD Primal-Dual), the primal iterates can be characterized in closed-form using the KKT conditions of the constrained optimization problem over Δ_n. In particular, under Gradient Descent, Bailey and Piliouras (2019a) give the following characterization of the primal iterate x^t, which in turn leads to a closed-form characterization of the energy function $\phi^*(y^t)$. At time t, both expressions are defined with respect to a set S^t, which characterizes the support of x^t. The set S^t can be found using the iterative method of FINDSUPPORT given in Algorithm 1.

Proposition 25 (Bailey and Piliouras (2019a)) *Let $\{x^t\}$ and $\{y^t\}$ be iterates of (GD Primal-Dual). At each time $t \geq 1$, let $S^t = \text{FINDSUPPORT}(y^t) \subseteq [n]$, and let $|S^t| = m$. Then*

$$
x_i^t = \begin{cases} 0 & \text{if } i \notin S^t \\ y_i^t - \frac{1}{m} \cdot \sum_{j\in S^t} y_j^t + \frac{1}{m} & \text{if } i \in S^t \,. \end{cases} \tag{27}
$$

Moreover, the energy $\phi^*(y^t)$ is given by

$$
\phi^*(y^t) = \frac{1}{2} \sum_{j\in S^t} (y_j^t)^2 + \frac{1}{m} \sum_{j\in S^t} y_j^t - \frac{1}{2m} \Bigg(\sum_{j\in S^t} y_j^t \Bigg)^2 - \frac{1}{2m} \,. \tag{28}
$$

While Proposition 25 is stated with respect to the iterates of (GD Primal-Dual), given any dual vector $y \in \mathbb{R}^n$, under the maps of $\phi^* : \mathbb{R}^n \to \mathbb{R}$ and $Q : \mathbb{R}^n \to \Delta_n$ from Proposition 18, we can

Algorithm 1: FINDSUPPORT

Input: $y \in \mathbb{R}^n$

$S \leftarrow [n]$

Search():

  Select $i \in \text{argmin}_{j \in S}\, y_j$

  if $y_i - \frac{1}{|S|}\big(\sum_{j \in S} y_j^t\big) + \frac{1}{|S|} < 0$ **then:**

    $S \leftarrow S \setminus \{i\}$

    goto *Search()*

  else return S

more generally describe the corresponding primal iterate and energy value in closed form similar to expressions (27) and (28). Specifically, in the case of Gradient Descent where $\phi(x) = \frac{\|x\|_2^2}{2}$, recall for $y \in \mathbb{R}^n$ that ϕ^* and Q are given by:

$$\begin{aligned} \phi^*(y) &= \max_{x \in \Delta_n} \langle x, y \rangle - \frac{\|x\|_2^2}{2} \\ Q(y) &= \operatorname*{argmax}_{x \in \Delta_n} \langle x, y \rangle - \frac{\|x\|_2^2}{2} \ . \end{aligned} \tag{29}$$

Then as a corollary to Proposition 25 (and Proposition 18), we state the following (see Bailey and Piliouras (2019a, Appendix B)):

Corollary 26 *Let $\phi^* : \mathbb{R}_n \to \mathbb{R}$ and $Q : \mathbb{R}^n \to \Delta_n$ be the functions in expression (29). For any $y \in \mathbb{R}^n$, let $S = \text{FINDSUPPORT}(y) \subseteq [n]$, and let $|S| = m$. Let $x := Q(y) \in \Delta_n$. Then for $i \in S$, the coordinate x_i is given by*

$$x_i \;=\; y_i - \frac{1}{m} \cdot \sum_{j \in S} y_j + \frac{1}{m} \ ,$$

and $x_i = 0$ otherwise. Additionally, the energy $\phi^*(y)$ is given by

$$\phi^*(y) \;=\; \frac{1}{2} \sum_{j \in S} (y_j)^2 + \frac{1}{m} \sum_{j \in S} y_j - \frac{1}{2m} \Big(\sum_{j \in S} y_j \Big)^2 - \frac{1}{2m} \ . \tag{30}$$

Moreover, $Q(y) = \nabla \phi^*(y)$.

In the case where $\phi = (\|\cdot\|_2^2)/2$ as in Gradient Descent, the key property of Proposition 25 and Corollary 26 is that the map $Q(y) = \nabla \phi^*(y)$ is not injective: in particular, multiple dual vectors of $\mathbb{R}^n$ map to the same distribution on the boundary of Δ_n under Q.

12.2. Primal-Dual Mapping for Gradient Descent

In light of Proposition 25 and Corollary 26, under expression (30) (or equivalently, under the map Q from expression (29)), we define a convenient mapping between the primal and dual spaces that holds for certain regions of the boundary of the simplex. For this, we first recall the definition of the sets P_i and $P_{i \sim (i+1)}$ from Section 14.

Definition 14 *For each $i \in [n]$, let $P_i \subset \mathbb{R}^n$ and $P_{i\sim(i+1)} \subset \mathbb{R}^n$ be the following sets:*

$$P_i := \left\{ y \in \mathbb{R}^n : \quad y_i - y_j > 1 \quad \textit{for } j \in [n] \setminus \{i\} \right\}$$
$$P_{i\sim(i+1)} := \left\{ y \in \mathbb{R}^n : \quad \begin{array}{l} |y_i - y_{i+1}| \leq 1 \textit{ and} \\ \frac{1}{2}(y_i + y_{i+1}) - y_j > \frac{1}{2} \quad \textit{for } j \in [n] \setminus \{i, i+1\} \end{array} \right\} .$$

Then we prove the following relationship (for simplicity, we state the result for the iterates of (GD Primal-Dual), but the same statement holds for $x \in \Delta_n$ and $y \in \mathbb{R}^n$ such that $Q(y) = x$):

Proposition 15 *For any $i \in [n]$: $y^t \in P_i$ if and only if $x^t = e_i$, and $y^t \in P_{i\sim(i+1)}$ if and only if* $\mathrm{supp}(x^t) = \{i, i+1\}$.

Proof We first prove the equivalence $y^t \in P_i \iff x^t = e_i$. For the forward direction, assume $y^t \in P_i$. By definition of P_i, we have $y^t_j - y^t_i < -1$ for any $j \neq i$. Then by definition of FINDSUPPORT (Algorithm 1) it follows that $j \notin S^t$. Thus $S^t = \mathrm{supp}(x^t) = \{i\}$, meaning $x^t = e_i$.

For the backward direction, assume $x^t = e_i$, which means $S^t = \text{FINDSUPPORT}(y^t) = \{i\}$. By definition of FINDSUPPORT, this means for all $j \neq i$ that $y^t_j - y^t_i < -1$, and thus $y^t \in P_i$.

For the second part of the proposition, we prove the equivalence $y^t \in P_{i\sim(i+1)} \iff \mathrm{supp}(x^t) = \{i, i+1\}$. For the forward direction, assume $y^t \in P_{i\sim(i+1)}$. By definition of $P_{i\sim(i+1)}$, we have $\frac{y^t_i + y^t_{i+1}}{2} \leq y^t_i + \frac{1}{2}$ and $\frac{y^t_i + y^t_{i+1}}{2} \leq y^t_{i+1} + \frac{1}{2}$, which implies $\min\{y^t_i, y^t_{i+1}\} > y^t_j$ for all other $j \neq i \neq i+1$. Then by definition of FINDSUPPORT and $P_{i\sim(i+1)}$, it follows for all such j that the condition in line 4 is violated, and thus $j \notin S^t$. On the other hand, assuming without loss of generality that $y^t_{i+1} = \min\{y^t_i, y^t_{i+1}\}$, then by definition of $P_{i\sim(i+1)}$ we find $y^t_i - \frac{y^t_i + y^t_{i+1}}{2} = \frac{y^t_i - y^t_{i+1}}{2} \geq -1/2$. Thus the condition in line 4 of FINDSUPPORT fails, and we have $S^t = \mathrm{supp}(x^t) = \{i, i+1\}$.

For the backward direction, assume $\mathrm{supp}(x^t) = \{i, i+1\}$. Then by definition of FINDSUPPORT we must have $y^t_j - \frac{y^t_i + y^t_{i+1}}{2} < -\frac{1}{2}$ for all other j. Moreover, since $x^t_i, x^t_{i+1} > 0$, we find using expression (27) that $y^t_i - y^t_{i+1} \geq -1$ and $y^t_{i+1} - y^t_i \geq -1$. Thus $y^t \in P_{i\sim(i+1)}$. ■

12.3. Geometry of the Energy Function

Energy function on the simplex boundary. For (GD Primal-Dual), when the dual iterate y^t is in the region P_i or $P_{i\sim(i+1)}$, the energy function ϕ^* (from expression (28) in Proposition 25) has the following simplified form:

$$\phi^*(y^t) = y^t_i - \frac{1}{2} \qquad \text{for } y^t \in P_i. \tag{31}$$
$$\phi^*(y^t) = \frac{1}{4}(y^t_i - y^t_{i+1})^2 + \frac{1}{2}(y^t_i + y^t_{i+1}) - \frac{1}{4} \qquad \text{for } y^t \in P_{i\sim(i+1)}. \tag{32}$$

Then we have the following relationship between energy and the maximum coordinate of the dual iterate within these regions:

Lemma 27 *Let ϕ^* be the energy function from Proposition 25. Then for each $i \in [n]$:*

(i) If $y^t \in P_i$, then $\phi^*(y^t) \le y_i^t \le \phi^*(y^t) + \frac{1}{2}$.

(ii) If $y^t \in P_{i\sim(i+1)}$, then $\phi^*(y^t) - \frac{1}{2} \le y_i^t, y_{i+1}^t \le \phi^*(y^t) + \frac{3}{4}$

Proof Fix $y^t \in P_i$. Then by (31), we have $\phi^*(y^t) = y_i^t - \frac{1}{2}$. Thus $\phi^*(y^t) \le y_i^t \le \phi^*(y^t) + \frac{1}{2}$. For $y^t \in P_{i\sim(i+1)}$, we have by definition that $y_i^t - 1 \le y_{i+1}^t \le y_i^t + 1$. Then using the form of ϕ^* from (32), we can bound

$$y_i^t + \frac{1}{2} \;\ge\; \frac{y_i^t + y_{i+1}^t}{2} \;=\; \phi^*(y^t) - \frac{1}{4}(y_i^t - y_{i+1}^t)^2 + \frac{1}{4} \;\ge\; \phi^*(y^t)\;,$$

where the final inequality comes from the fact that $|y_i^t - y_{i+1}^t| \le 1$ when $y^t \in P_{i\sim(i+1)}$. Rearranging, we have $y_i^t \ge \phi^*(y^t) - \frac{1}{2}$. An identical calculation also finds that $y_{i+1}^t \ge \phi^*(y^t) - \frac{1}{2}$.

For the upper bound, we similarly have

$$y_i^t - \frac{1}{2} \;\le\; \frac{y_i^t + y_{i+1}^t}{2} \;=\; \phi^*(y^t) - \frac{1}{4}(y_i^t - y_{i+1}^t)^2 + \frac{1}{4} \;\le\; \phi^*(y^t) + \frac{1}{4}\;.$$

Rearranging, we have $y_i^t \le \phi^*(y^t) + \frac{3}{4}$. An identical calculation also yields $y_{i+1}^t \le \phi^*(y^t) + \frac{3}{4}$. ■

Energy of dual iterate when primal iterate has full support. For $y \in \mathbb{R}^n$ such that $S = \text{FINDSUPPORT}(y) = [n]$, then by Corollary 26, the primal iterate $x = Q(y)$ is *interior* and has full support (e.g., $\text{supp}(x) = [n]$). In this case, letting $\mathbf{1} \in \mathbb{R}^n$ denote the all-ones vector, the energy function can be simplified as

$$\phi^*(y^t) \;=\; \frac{\|y^t\|_2^2}{2} + \frac{\langle y^t, \mathbf{1}\rangle}{2n} - \frac{(\langle y^t, \mathbf{1}\rangle)^2}{2n} - \frac{1}{2n}\;. \tag{33}$$

13. Gradient Descent on High-Dimensional RPS in Large Stepsize Regime

In this section we develop the proof of Theorem 17, which establishes that Gradient Descent obtains $O(\sqrt{T})$ regret on high-dimensional RPS games in a regime of *large constant stepsizes*:

Theorem 17 *Let A be an n-dimensional RPS matrix. Then for nearly all initial distributions $x^0 \in \Delta_n$, the following holds: letting $\{x^t\}$ be the iterates of running* (GD Primal-Dual) *on A with $\eta > \min\{2/a_{\min}, 1/\gamma(x^0)\}$, then* $\text{Reg}(T) \le O(\sqrt{T})$.

The organization of this section is as follows:

- In Section 13.1, we give the proof of Lemma 16, which shows that, for nearly all initializations, the primal iterates of Gradient Descent quickly converge to a vertex of Δ_n when the stepsize is sufficiently large.
- In Section 13.2, we then develop the proof of Lemma 28. Here, we show in the large stepsize regime that the primal iterates of Gradient Descent cycle through the vertices of Δ_n in phases, in a manner analogous to Fictitious Play (c.f., Lemma 10). Lemma 28 also establishes similar bounds on the length of each phase and on the energy growth between phases (c.f., Lemma 13 and Proposition 12 for Fictitious Play).
- Using these ingredients, the full proof of Theorem 17 is given in Section 13.3 and again follows similarly to the proof of Theorem 11 for Fictitious Play.

13.1. Fast Convergence to a Vertex

Lemma 16 *If $x^0 \in \Delta_n$ is such that $\gamma(x^0) > 0$, then along one step of* (GD Primal-Dual) *with stepsize $\eta > 1/\gamma(x^0)$, the iterate x^1 is a vertex of Δ_n.*

Proof Fix $x^0 \in \Delta^n$, and recall by definition of (GD Primal-Dual) that $y^1 = \eta A x^0$. In particular, by the structure of the entries of an RPS matrix A, we have for all $i \in [n]$:

$$y_i^1 \;=\; \eta\big(a_{i-1} \cdot x_{i-1}^0 - a_i \cdot x_{i+1}^0\big) \;.$$

Without loss of generality, assume indices i and j are the largest and second largest coordinates of y^1 satisfying $y_i^1 = \max_{k\in[n]} y_k^1$ and $y_j^1 = \max_{k\in[n]\setminus\{i\}} y_k^1$. Observe from Proposition 15 that if $y_i^1 - y_j^1 > 1$, then $y^1 \in P_i$, and thus $x^1 = e_i$. In particular, this condition will be satisfied when

$$\eta\Big(\big(a_{i-1} \cdot x_{i-1}^0 - a_i \cdot x_{i+1}^0\big) - \big(a_{j-1} \cdot x_{j-1}^0 - a_j \cdot x_{j+1}^0\big)\Big) \;>\; 1 \;. \tag{34}$$

Recall from expression (7) that the constant $\gamma(x^0)$ is given by

$$\gamma(x^0) \;=\; \min_{k,\ell\in[n]} \Big|\big(a_{k-1} \cdot x_{k-1}^0 - a_k \cdot x_{k+1}^0\big) - \big(a_{\ell-1} \cdot x_{\ell-1}^0 - a_\ell \cdot x_{\ell+1}^0\big)\Big| \;. \tag{35}$$

Observe that if $\gamma(x^0) > 0$, and so long as $\eta > 1/\gamma(x^0)$, then the inequality in (34) is always satisfied, which ensures $y^1 \in P_i$. Fixing the matrix A, the set of points $x \in \Delta_n$ where $\gamma(x) = 0$ is restricted to the linear constraint in (35) and thus has (Lebesgue) measure zero. Thus for almost all initial conditions x^0, setting $\eta > 1/\gamma(x^0)$ guarantees $x^1 = e_i$. ■

13.2. Cycling, Energy Growth, and Phase Length Bounds

With sufficiently large stepsizes, once the primal iterate reaches a vertex e_i we prove in Lemma 28 below that the iterates subsequently cycle through the vertices in phases in the order

$$e_i \to e_{i+1} \to \cdots \to e_n \to e_1 \to \cdots \to e_i \to \ldots$$

similar to the behavior of Fictitious Play from Lemma 10. For a sequence of consecutive primal iterates all at the same vertex, the energy growth of ϕ^* is zero.

On the other hand, between vertices e_i and e_{i+1}, the primal iterates may spend a constant number of iterations on the simplex edge between vertices e_i and e_{i+1} (e.g., where $\mathrm{supp}(x) = \{i, i+1\}$ for $x \in \Delta_n$). Consecutive primal iterates on these edges correspond to a constant energy growth per step. However, using a sufficiently large constant stepsize, we can ensure that at most a *single round* is spent on any such edge before transitioning to the next vertex. This ensures a regularity in the cycling behavior closely mirroring that of Fictitious Play, ultimately leading to bounds on the energy growth and length of each phase analogous to those in Lemma 13.

To formally analyze this cycling behavior and its consequences, we reuse the notation of phases from the analysis of Fictitious Play, which we restate here:

Definition 9 *Fix a time $t_0 > t$. For each $k \geq 1$, let $t_k := \min\,\{t > t_{k-1} : x^t \neq x^{t_{k-1}}\}$. Then Phase k is the sequence of iterates at times $t = t_k, t_k+1, \ldots, t_{k+1}-1$. Let $\tau_k = t_{k+1} - t_k$ denote the length of Phase k. Let $K \geq 0$ be the total number of phases in T rounds, where $T = \sum_{k=0}^{K} \tau_k$.*

We then formally state the previously-described behavior in the following lemma:

Lemma 28 *Let $x^0 \in \Delta_n$ be such that $\gamma(x^0) > 0$ (for γ defined in (35)), and assume $\eta > \max\{2/a_{\min}, 1/\gamma(x_0)\}$. Then there exists $t_0 \le O(n)$ such that $x^{t_0} = e_i$ for some $i \in [n]$. Moreover, for each $k = 1, \ldots, K$:*

(i) $y^{t_k} \in P_{i+k \text{ (mod } n)}$, *and* $y^{t_{k+1}-1} \in \left\{P_{i+k-1 \text{ (mod } n)}, P_{(i+k+1 \text{ (mod } n))\sim(i+k(\text{mod } n))}\right\}$.

(ii) $\phi^*(y^{t_{k+1}}) - \phi^*(y^{t_k}) \le O(1)$

(iii) $\tau_k \ge \alpha_k \cdot \phi^*(y^{t_k}) - \beta_k$, *where $\alpha_k > 0$ and $\beta_k > 0$ are constants.*

Observe that, using the primal-dual mapping from Proposition 15, Part (i) of the lemma establishes the cycling of the primal iterates by characterizing the behavior of the corresponding dual iterates. In particular, Part (i) shows that $x^{t_k} = e_{i+k \text{ (mod } n)}$ and $x^{t_{k+1}} = e_{i+k+1 \text{ (mod } n)}$, and either $x^t = e_{i+k \text{ (mod } n)}$ for all $t_k \le t \le t_{k+1}$, or $\text{supp}(x^{t_{k+1}-1}) = \{i+k \text{ (mod } n), i+k+1 \text{ (mod } n)\}$. In other words, the primal iterates either transition directly from vertex $e_{i+k \text{ (mod } n)}$ to $e_{i+k+1 \text{ (mod } n)}$, or at most one iterate is spent on the simplex edge connecting these vertices before transitioning to $e_{i+k+1 \text{ (mod } n)}$. The constraint $\eta > \frac{2}{a_{\min}}$ ensures at most a single iterate is spent on this edge.

The proof of Lemma 28 relies on several helper propositions, which we develop in the following two subsections.

13.2.1. ESTABLISHING THE CYCLING ORDER

Recall that for any n-dimensional RPS matrix, for $x^t \in \Delta_n$, then $\Delta y^t = y^{t+1} - Y^t = \eta A x^t$ with coordinates given by

$$\Delta y_i^t \;=\; \eta\big(a_{i-1}x_{i-1}^t - a_i x_{i+1}^t\big)\ . \tag{36}$$

In the following two propositions, we establish the order in which the dual iterates cycle through the regions from Proposition 14.

Proposition 29 *Fix $i \in [n]$ and suppose $y^t \in P_i$. Assume $\eta > 2/a_{\min}$, and let $\tau := \left\lceil \frac{y_i^t - y_{i+1}^t - 1}{\eta \cdot a_i} \right\rceil$. Then $y^{t+s} \in P_i$ for all $1 \le s \le \tau - 1$. Moreover, at time $t+\tau$, we have either (i) $y^{t+\tau} \in P_{i+1}$, or (ii) $y^{t+\tau} \in P_{i\sim(i+1)}$, and additionally:*

- $y_i^{t+\tau} = y_i^t$
- $y_{i-1}^{t+\tau} < y_{i-1}^t - 2\tau$
- $y_{i+1}^{t+\tau} - y_j^{t+\tau} > 1$ *for all $j \in [n] \setminus \{i, i+1\}$.*

Proof By Proposition 15, if $y^\ell \in P_i$, then $x^\ell = e_i$. Thus $\Delta y^\ell = y^{\ell+1} - y^\ell = \eta A x^\ell = \eta A_i$, which is a constant vector. Inductively (and following the proof of Proposition 21 in the Fictitious Play case), in order to prove $y^{t+s} \in P_i$ for all $1 \le s \le \tau - 1$, it suffices to show that

$$y_j^{t+s} \;=\; y_j^t + s \cdot \eta A_{ji} \;<\; y_i^t + s \cdot \eta A_{ii} - 1 \;=\; y_i^t - 1 \quad \text{for all } j \ne i. \tag{37}$$

For this, we can directly check using the definition of the entries of A_i that

$$\begin{aligned} y_{i-1}^{t+s} &= y_{i-1}^t - s \cdot \eta a_{i-1} \\ y_{i+1}^{t+s} &= y_{i+1}^t + s \cdot \eta a_i \\ \text{and} \quad y_j^{t+s} &= y_j^t \qquad\qquad \text{for } j \in [n] \setminus \{i-1, i, i+1\} \ . \end{aligned} \tag{38}$$

Because $y^t \in P_i$, we have $y_i^t - y_{i-1}^t > 1$. Then using the positivity of a_{i-1}, we have at coordinate $i-1$ for all $s \geq 1$ that

$$y_{i-1}^{t+s} - y_i^{t+s} \ = \ y_{i-1}^t - y_i^t - s \cdot \eta a_{i-1} \ \leq \ -1 - 2s \ , \tag{39}$$

where the final inequality comes from the fact that $\eta > 2/a_{\min}$, and thus $\eta a_{i-1} \geq 2$. For coordinates $j \in [n] \setminus \{i, i+1\}$, as $y_i^t - y_j^t > 1$ by assumption of $y^t \in P_i$, we similarly find that $y_i^{t+s} - y_j^{t+s} > 1$ for all $s \geq 1$. Thus (37) holds for both cases. Now at coordinate $i+1$, observe that that

$$y_i^{t+s} - y_{i+1}^{t+s} \ \geq \ y_i^t - y_{i+1}^t - s \cdot \eta \cdot a_i \ > \ y_i^t - y_{i+1}^t - (y_i^t - y_{i+1}^t - 1) \ = \ 1$$

for all $1 \leq s \leq \tau - 1 < \frac{y_i^t - y_{i+1}^t - 1}{\eta \cdot a_i}$. Thus expression (37) also holds for coordinate $i+1$, which means $y^{t+s} \in P_i$ for all $1 \leq s \leq \tau - 1$.

To prove the second part of the proposition, observe that the first part establishes at time $t + \tau$ that $y_i^{t+\tau} = y_i^t + \tau \cdot \eta A_{ii} = y_i^t$. Also, from expression (39), we conclude that $y_{i-1}^{t+\tau} < y_{i-1}^t - 2\tau$. Moreover, at time $t + \tau$ we also have

$$\begin{aligned} y_i^{t+\tau} - y_{i+1}^{t+\tau} \ &\leq \ y_i^t - y_{i+1}^t - \tau \cdot \eta \cdot a_i \\ &\leq \ y_i^t - y_{i+1}^t - (y_i^t - y_{i+1}^t - 1) \ = \ 1 \ . \end{aligned}$$

If $y_i^{t+\tau} - y_{i+1}^{t+\tau} < -1$, then it follows that from the previous arguments that also $y_{i+1}^{t+\tau} - y_j^{t+\tau} > 1$ for all $j \in [n] \setminus \{i, i+1\}$. This establishes that $y^{t+\tau} \in P_{i+1}$.

If instead $1 \geq y_i^{t+\tau} - y_{i+1}^{t+\tau} \geq -1$, then it still remains that $y_i^{t+\tau} - y_j^{t+\tau} \geq 1$ for all other $j \in [n] \setminus \{i, i+1\}$. In this case, as $y_{i+1}^{t+\tau} \geq y_i^{t+\tau} - 1$, we can then further verify for all such j that $\frac{y_i^{t+\tau} + y_{i+1}^{t+\tau}}{2} - y_j^{t+\tau} \geq \frac{1}{2}$, and thus $y^{t+\tau} \in P_{i \sim (i+1)}$. ■

Proposition 30 *Fix $i \in [n]$ and suppose $y^t \in P_{i \sim (i+1)}$ such that $y_{i+1}^{t+\tau} - y_j^{t+\tau} > 1$ for all $j \in [n] \setminus \{i, i+1\}$. Assume $\eta > 2/a_{\min}$. Then at time $t+1$, we have $y_{i-1}^{t+1} < y_{i-1}^t - 2c$ for some positive constant $c > 0$, and one of three scenarios occurs: if $y_{i+1}^t - y_{i+2}^t \geq 1 + \eta a_{i+1}$, then (i) $y^{t+1} \in P_{i+1}$. Otherwise, either: (ii) $y^{t+1} \in P_{i+2}$ or (iii) $y^{t+1} \in P_{(i+1) \sim (i+2)}$.*

Proof First, recall from Proposition 15 that since $y^t \in P_{i \sim (i+1)}$, we have $\operatorname{supp}(x^t) = \{i, i+1\}$. Then the coordinates of $\Delta y^t = y^{t+1} - y_i^t = \eta A x^t$ are specified by

$$\begin{aligned} \Delta y_{i-1}^t \ &= -\, \eta \cdot a_{i-1} \cdot x_i^t \\ \Delta y_i^t \ &= -\, \eta \cdot a_i \cdot x_{i+1}^t \\ \Delta y_{i+1}^t \ &= +\, \eta \cdot a_i \cdot x_i^t \\ \Delta y_{i+2}^t \ &= +\, \eta \cdot a_{i+1} \cdot x_{i+1}^t \\ \Delta y_j^t \ &= 0 \quad \text{for all other } j \ . \end{aligned} \tag{40}$$

We start by showing the behavior of the differences $y_{i+1}^{t+1} - y_j^t$ for all $j \in [n] \setminus \{i+1, i+2\}$.

- At coordinate i, observe that:

$$y_{i+1}^{t+1} - y_i^{t+1} \ = \ y_{i+1}^t - y_i^t + \eta a_i(x_{i+1}^t + x_i^t) \ \geq \ y_{i+1}^t - y_i^t + 2 \ ,$$

  where the inequality follows from the assumptions that $\eta > 2/a_{\min}$ and $\mathrm{supp}(x^t) = \{i, i{+}1\}$. As $y_{i+1}^t - y_i^t \geq -1$ by assumption of $y^t \in P_{i\sim(i+1)}$, we have $y_{i+1}^{t+1} - y_i^{t+1} > 1$. Moreover, if $y_{i+1}^t - y_i^t > 0$, it follows that $y_{i+1}^{t+1} \geq y_i^t$. Similarly, if $-1 \leq y_{i+1}^t - y_i^t < 0$, then we still have $y_{i+1}^{t+1} > y_{i+1}^t + 1$, and thus also $y_{i+1}^{t+1} \geq y_i^t$.

- At coordinate $i-1$, we have $y_{i-1}^{t+1} = y_{i-1}^t - \eta a_{i-1} x_i^t < y_{i-1}^t - 2c$, where $c = x_i^t > 0$ is some positive constant. Moreover, the inequality holds given $\eta > 2/a_{\min}$. Note that this proves the first claim of the proposition.

  Observe further that we have $y_{i+1}^{t+1} - y_{i-1}^{t+1} \geq y_i^t - y_{i-1}^t + 2c > 1$, where, the final inequality comes from the assumption that $y_i^t - y_{i-1}^t \geq 1$.

- For all other $j \neq i+2$, we have we have $\Delta y_j^t = 0$. It follows that $y_{i+1}^{t+1} - y_j^{t+1} \geq y_i^t - y_j^t > 1$.

We now examine the difference $y_{i+1}^{t+1} - y_{i+2}^{t+1}$. Observe from expression (40) that

$$\begin{aligned} y_{i+1}^{t+1} - y_{i+2}^{t+1} \ &= \ y_{i+1}^t - y_{i+2}^t + \eta a_i x_i^t - \eta a_{i+1} x_{i+1}^t \\ &= \ y_{i+1}^t - y_{i+2}^t + \eta a_i (1 - x_{i+1}^t) - \eta a_{i+1} x_{i+1}^t \\ &\geq \ y_{i+1}^t - y_{i+2}^t - \eta a_{i+1} \ , \end{aligned}$$

where the inequality holds given that $x_{i+1}^t \leq 1$. If $y_{i+1}^t - y_{i+2}^t > 1 + \eta a_{i+1}$, then $y_{i+1}^{t+1} - y_{i+2}^t > 1$, and claim (i) of the proposition holds by combining the preceding arguments.

If on the other hand $1 \geq y_{i+1}^{t+1} - y_{i+2}^{t+1}$, then either $y_{i+2}^{t+1} - y_{i+1}^{t+1} > 1$, and claim (ii) of the proposition holds by combining the preceding arguments, or $1 \geq y_{i+2}^{t+1} - y_{i+1}^{t+1} \geq -1$, and claim (iii) of the proposition holds. ■

13.2.2. ENERGY GROWTH BOUNDS

In the following lemma, we establish lower and upper bounds on the energy growth when the dual iterates move between the regions from Definition 14:

Lemma 31 *Fix $i \in [n]$, and let $\Delta\phi^*(y^t) = \phi^*(y^{t+1}) - \phi^*(y^t)$. Then the following hold:*

(i) If $y^t \in P_i$ and $y^{t+1} \in P_i$, then $\Delta\phi^*(y^t) = 0$.

(ii) If $y^t \in P_i$ and $y^{t+1} \in P_{i+1}$, then $1 < \Delta\phi^*(y^t) < \eta a_{\max}$.

(iii) If $y^t \in P_i$ and $y^{t+1} \in P_{i\sim(i+1)}$, then $0 \leq \Delta\phi^*(y^t) \leq 1$.

(iv) If $y^t \in P_{i\sim(i+1)}$ and $y^{t+1} \in P_{i+1}$ or $y^{t+1} \in P_{i+2}$, then $0 \leq \Delta\phi^*(y^t) \leq \frac{1}{4}(\eta a_{\max})^2$.

(v) If $y^t \in P_{i\sim(i+1)}$ and $y^{t+1} \in P_{(i+1)\sim(i+2)}$ then $0 \leq \Delta\phi^*(y^t) \leq \eta a_{\max} + \frac{5}{4}$.

Moreover, in cases (iii) through (v), if $\phi^*(y^t) > 0$, then there exists an absolute constant $c > 0$ such that $\phi^*(y^t) \geq c$.

Proof We prove each part independently:

- *Part (i)*: By Proposition 15, $x^t = x^{t+1} = e_i$, and thus by Proposition 25: $\Delta\phi^*(y^t) = y_i^{t+1} - y_i^t$. As $y_i^{t+1} = y_i^t + \eta(Ae_i)_i = y_i^t$, we then have $\Delta\phi^*(y^t) = 0$.

- *Part (ii)*: Using Proposition 25 and expression (40), for $y^t \in P_i$, we have $\phi^*(y^{t+1}) = y_{i+1}^{t+1} - \frac{1}{2} = y_{i+1}^t + \eta a_i + \frac{1}{2}$, and $\phi^*(y^{t+1}) = y_i^{t+1} - \frac{1}{2} = y_i^t - \frac{1}{2}$. Thus

$$\Delta\phi^*(y^t) = y_{i+1}^t - y_i^t + \eta a_i < -1 + \eta a_{i+1} < \eta a_i \ ,$$

  where we use the fact that $y_{i+1}^t - y_i^t < -1$ for $y^t \in P_i$. On the other hand, $y_{i+1}^{t+1} - y_i^{t+1} > 1$ by definition of P_{i+1}, and thus $\Delta\phi^*(y^t) > 1$.

- *Part (iii)*: As $y^t \in P_i$, we have $\Delta y_i^t = 0$, and thus $y_i^{t+1} = y_i^t$. Then by Proposition 25, the energy at times t and $t+1$ are given by

$$\begin{aligned} \phi^*(y^{t+1}) &= \frac{1}{4}(y_i^{t+1} - y_{i+1}^{t+1})^2 + \frac{1}{2}(y_{i+1}^{t+1} + y_i^{t+1}) - \frac{1}{4} \\ \phi^*(y^t) &= y_i^t - \frac{1}{2} = y_i^{t+1} - \frac{1}{2} \\ \text{and} \quad \Delta\phi^*(y^t) &= \frac{1}{4}(y_i^{t+1} - y_{i+1}^{t+1})^2 + \frac{1}{2}(y_{i+1}^{t+1} - y_i^{t+1}) + \frac{1}{4} \ . \end{aligned}$$

  By definition of $P_{i\sim(i+1)}$, we have $|y_i^{t+1} - y_{i+1}^{t+1}| \leq 1$, and thus by convexity, we have $\Delta\phi^*(y^{t+1}) \leq 1$. If $y_i^{t+1} - y_{i+1}^{t+1} = 0$, then $\Delta\phi^*(y^{t+1}) = 0$. On the other hand, recall that $y_{i+1}^{t+1} = y_{i+1}^t + \Delta y_{i+1}^t$, and that $\Delta y_{i+1}^t = \eta a_i$ is an absolute constant. Thus if $\Delta\phi^*(y^{t+1}) > 0$, then we must have $\Delta\phi^*(y^{t+1}) \geq c$ for some absolute constant $c > 0$.

- *Part (iv)*: By Proposition 25, the energy at times t and $t+1$ are given by

$$\begin{aligned} \phi^*(y^{t+1}) &= y_{i+1}^{t+1} - \frac{1}{2} \\ \phi^*(y^t) &= \frac{1}{4}(y_i^t - y_{i+1}^t)^2 + \frac{1}{2}(y_{i+1}^t + y_i^t) - \frac{1}{4} \ . \end{aligned}$$

  As $y^t \in P_{i\sim(i+1)}$, by expression (36) we have $y_{i+1}^{t+1} = y_{i+1}^t + \eta a_i x_i^t$. Then using the closed-form expression of x_i^t in Proposition 25, we have $x_i^t = \frac{1}{2}(y_i^t - y_{i+1}^t + 1)$. It follows that

$$\Delta\phi^*(y^t) = \frac{1}{2}(1 - \eta a_i)(y_{i+1}^t - y_i^t) - \frac{1}{4}(y_i^t - y_{i+1}^t)^2 + \frac{\eta a_i}{2} - \frac{1}{4} \ .$$

  By convexity, we find $\Delta\phi^*(y^t) \leq \frac{1}{4}(\eta a_i)^2$. If $y_{i+1}^t - y_i^t = 1$, then $\Delta\phi^*(y^t) = 0$. On the other hand, if $-1 \leq y_{i+1}^t - y_i^t < 1$, then $\Delta\phi^*(y^t) > 0$, and the there must exist an aboslute constant $c > 0$ such that $\Delta\phi^*(y^t) \geq c$.

  If $y^{t+1} \in P_{i+2}$, we find by similar calculations that $\Delta\phi^*(y^t) \leq \frac{1}{4}(\eta a_{i+1})^2$, and the same arguments for the lower bound also apply.

- *Part (iv)*: As $y^t \in P_{i\sim(i+1)}$ we have from expression (36) that $y_{i+1}^{t+1} - y_{i+1}^t = \eta a_i x_i^t \leq \eta a_i$ and $y_{i+2}^{t+1} - y_{i+2}^t = \eta a_{i+1} x_{i+1}^t \leq \eta a_{i+1}$. Then using the bounds from Lemma (27), we have

$\phi^*(y^{t+1}) \le y_{i+1}^{t+1} + \frac{1}{2}$ and $\phi^*(y^t) \ge y_{i+1}^t - \frac{3}{4}$, and thus $\Delta\phi^*(y^t) \le y_{i+1}^{t+1} - y_{i+1}^t + \frac{5}{4} \le \eta a_i + \frac{5}{4}$. Recall from Proposition 8 that in general $\Delta\phi^*(y^t) \ge 0$. On the other hand, as Δy_{i+1}^t and Δy_i^t are absolute constants, if $\Delta\phi^*(y^t) > 0$, then there must be an absolute constant $c \ge 0$ such that $\Delta\phi^*(y^t) \ge c$.

■

13.2.3. PROOF OF LEMMA 28

We now prove each of the three parts of Lemma 28:

Proof of Part (i). By the convergence property of Lemma 16, if $x^0 \in \Delta_n$ satisfies $\gamma(x^0) > 0$, then $x^1 = e_i$ for some $i \in [n]$, By the equivalence from Proposition 15, $y^1 \in P_i$. Moreover, we must have $\phi^*(y^1) = \Theta(1)$. Now by the cycling order established in Propositions 29 and 30, for any $i \in [n]$, the dual iterates are restricted to transitioning only between the regions of Definition 14 in the following manner:

(a) $P_i \to P_{i+1}$ (b) $P_i \to P_{i\sim(i+1)}$ (c) $P_{i\sim(i+1)} \to P_{(i+1)}$

(d) $P_{i\sim(i+1)} \to P_{i+2}$ (e)$P_{i\sim(i+1)} \to P_{(i+1)\sim(i+2)}$.

We show that there exists t_0 with $t_0 \le O(n)$, such that for all $t \ge t_0$, transitions (d) and (e) never occur. Noting the structure of the transitions (a), (b), and (c), and recalling the primal-dual map of Proposition 30, observe that this is sufficient for proving part (i) of the lemma.

Now assume $y^t \in P_{i\sim(i+1)}$, and observe by the statement of Proposition 30 that transitions (d) and (e) cannot occur when $y_{i+1}^t - y_{i+2}^t \ge 1 + \eta a_{i+1}$. By Lemma 27, note also that if $y^t \in P_{i\sim(i+1)}$, then $y_{i+1}^t \ge \phi^*(y^t) - \frac{1}{2}$. Moreover, among the types of transitions that can occur according to Propositions 29 and 30, and using the velocity definition from expression (40), observe that coordinate $(i+2)$ of the dual iterate can increase only in the regions $P_{(i)\sim(i+1)}$, P_{i+1}, and $P_{(i+1)\sim(i+2)}$, and it must be decreasing in the regions $P_{(i+2)\sim(i+3)}$, P_{i+3}, and $P_{(i+3)\sim(i+4)}$. Due to the transition orders of Propositions 29 and 30, it follows that between the most recent time coordinate $(i+2)$ has decreased and the next time the coordinate increases, at least $\Omega(n)$ rounds have elapsed.

We assume without loss of generality that within these $\Omega(n)$ rounds, the energy $\phi^*(y^t)$ is strictly increasing when transitioning between phases (otherwise, the energy remains constant for these rounds, and a similar argument can be applied once the energy has sufficiently increased). By Proposition 31, the energy increase between each region is an absolute constant, and thus the total energy of the dual iterates increases by at least $\Omega(n)$ before the next time coordinate $(i+2)$ increases.

Putting these arguments together, suppose $y^t \in P_{i\sim(i+1)}$, where $1 \le t = cn$ for some constant $c \ge 1$. Then for sufficiently large c, we must have $y_{i+2}^t \le d \cdot \phi^*(y^t)$ for some absolute constant $0 < d < 1$. Recalling the lower bound on y_{i+1}^t from above, it follows that

$$y_{i+1}^t - y_{i+2}^t \;\ge\; \phi^*(y^t) - \frac{1}{2} - d \cdot \phi^*(y^t) \;\ge\; \phi^*(y^t)(1-d) - \frac{1}{2} \,.$$

As we assumed that the energy has strictly increased by an absolute constant at each subsequent time the dual iterate re-enters $P_{i\sim(i+1)}$, the energy $\phi^*(y^t)$ is strictly increasing with c. As ηa_{i+1} is a fixed absolute constant, then for c sufficiently large, we have

$$\phi^*(y^t)(1-d) - \frac{1}{2} > 1 + \eta a_{i+1} \,.$$

Proposition 29 further implies that upon the next time $t' > t$ where $y^{t'}$ re-enters $P_{i\sim(i+1)}$, then also $y^t_{i+1} - y^t_{i+2} > 1 + \eta a_{i+1}$. Repeating this argument for all $i \in [n]$, we set t_0 as the maximum time among all such t used. It follows that for any $i \in [n]$, and for any subsequent $t' > t_0$ where $y^{t'} \in P_{i\sim(i+1)}$, we must have $y^{t'}_{i+1} - y^{t'}_{i+2} \geq 1 + \eta a_{i+1}$, and thus transitions of types (d) and (e) cease to occur. ■

Proof of Part (ii). Using the cycling order from Part (i) of the lemma, recall that between times t_k and t_{k+1}, the dual iterates transition either from $P_{i+k \text{ (mod } n)}$ to $P_{i+k+1 \text{ (mod } n)}$ directly, or from $P_{i+k \text{ (mod } n)}$ to $P_{(i+k \text{ (mod } n))\sim(i+k+1 \text{ (mod } n))}$ to $P_{i+k+1 \text{ (mod } n)}$. For consecutive iterates within the region $P_{i+k+1 \text{ (mod } n)}$, the energy growth is zero. Moreover, recall from Part (i) of the lemma that the dual iterate spends at most a single round in $P_{(i+k \text{ (mod } n))\sim(i+k+1 \text{ (mod } n))}$. Then using the cases of Lemma 31, we can bound in the worst case

$$\phi^*(y^{t_{k+1}}) - \phi^*(y^{t_k}) \ \leq \ \eta a_{\max} + 1 + \frac{1}{4}(\eta a_{\max})^2 \ ,$$

which is at most an absolute constant. ■

Proof of Part (iii). Observe Part (i) of the lemma establishes that from time t_0 onward, the dual iterates perpetually cycle through the regions $\{P_k\}$ in the order

$$\dots P_i \to P_{i+1} \to \dots \to P_n \to P_1 \dots \ .$$

Moreover, between any two consecutive regions P_i and P_{i+1}, the only other region the dual iterate ever enters is $P_{i\ \sim(i+1)}$. From Proposition 30, under the large stepsize assumptions, the dual iterate spends at most a single timestep in any such $P_{i\sim(i+1)}$. Thus τ_k is at least as large as the number of steps spent in P_i before exiting.

Now for any $\ell \geq 0$, Part (i) of the lemma establishes that $y^{t_\ell} \in P_{i+\ell \text{ (mod } n)}$. Then by Proposition 29, observe that

$$\tau_\ell \ \geq \ \frac{y^{t_\ell}_{i+\ell \text{ (mod } n)} - y^{t_\ell}_{i+\ell+1 \text{ (mod } n)} - 1}{\eta \cdot a_{i+\ell \text{ (mod } n)}} \ . \tag{41}$$

Recall further from Lemma 27 that if $y^t \in P_i$, then $y^t_i \geq \phi^*(y^t)$ Thus we can further write

$$\tau_\ell \ \geq \ \frac{\phi^*(y^{t_\ell}) - y^{t_\ell}_{i+\ell+1 \text{ (mod } n)}}{\eta \cdot a_{i+\ell \text{ (mod } n)}} - \frac{1}{\eta \cdot a_{i+\ell \text{ (mod } n)}} \ . \tag{42}$$

As the second term in (42) is a positive absolute constant, it suffices to show that

$$\frac{\phi^*(y^{t_\ell}) - y^{t_\ell}_{i+\ell+1 \text{ (mod } n)}}{\eta \cdot a_{i+\ell \text{ (mod } n)}} \ \geq \ \alpha'_\ell \cdot \phi^*(y^{t_\ell}) - \beta'_\ell \tag{43}$$

for positive constants $\alpha'_\ell > 0$ and $\beta'_\ell > 0$ for all $\ell = 0, \dots, K$.

We prove this claim by induction following similarly to the proof of Lemma 13 for Fictitious Play. The key property to note is that, for any fixed coordinate i, following the most recent time the dual iterate was in the region P_{i+1}, and until the dual iterate again enters either $P_{(i-2)\sim(i-1)}$ or $P_{(i-1)}$, the velocity $\Delta y^t_i = 0$ at all intermediate times t. Together, with Propositions 29 and 30, this key property implies that within the first $\ell = 0, \dots, n$ phases, at least $\Omega(n)$ rounds have elapsed between the most recent time prior to t_0 that each coordinate i was *decreasing*, and the next time it *increases* within the first n phases. Moreover, $t_0 \leq O(n)$ by definition, and along with the energy growth bounds of Lemma 31 we have $\phi^*(y^0) = O(n)$. By a similar argument as in Part (i) of the lemma, for each $\ell = 0, \dots, n$, it follows that $y^{t_\ell}_{i+\ell+1 \pmod n} \leq b_\ell \cdot \phi^*(y^{t_\ell})$ for some positive constant $0 < b_\ell < 1$. Thus expression (43) holds for these initial n phases.

Now assume the claim holds through Phase $k-1$. To prove the claim also holds for Phase k, we can use the same calculations as in the proof of Lemma 13 for Fictitious Play. Specifically, using the cycling property of Part (i) of the lemma: $y^{t_k}_{i+k+1 \pmod n} \leq y^{t_{k+3-n}}_{i+k+1 \pmod n} + d_k$ for some constant $d_k > 0$. Moreover, by the energy growth bound of Part (ii) and also Lemma 27, we have $y^{t_{k+2-n}}_{i+k+1 \pmod n} \leq \phi^*(y^{t_{k+2-n}}) + \rho$, where $\rho > 0$ is an absolute constant. Then following identical calculations as in expressions (17) through (21) in the proof of Lemma 13, we conclude that

$$\begin{aligned} \frac{\phi^*(y^{t_k}) - y^{t_k}_{i+k+1 \pmod n}}{\eta \cdot a_{i+k \pmod n}} &\geq \frac{\phi^*(y^{t_k}) \cdot a_{i+k+1 \pmod n} \cdot \alpha_{k+2-n}}{a_{i+k \pmod n}} - \frac{\beta_{k+2-n} + \rho + d_k}{\eta \cdot a_{i+k \pmod n}} \\ &= \alpha'_k \cdot \phi^*(y^{t_k}) - \beta'_k \end{aligned}$$

for positive constants α'_k and β'_k. By expression (43), this proves the claim. ■

13.3. Proof of Theorem 17

Using the machinery of Lemma 28, we prove the regret bound of Theorem 17:

Proof The proof follows similarly to that of Theorem 11 for Fictitious Play. First, let $x^0 \in \Delta_n$ be such that $\gamma(x^0) > 0$ (for γ as defined in expression (7)). As established in the proof of Lemma 16, the set of points $x \in \Delta_n$ with $\gamma(x) > 0$ has full Lebesgue measure. Then under the large step size assumptions of the theorem, $x^1 = e_i$ for some $i \in [n]$, and the cycling, energy growth, and phase length bounds of Lemma 28 apply.

Recall by Proposition 6 that for constant $\eta > 0$, to bound the regret $\mathrm{Reg}(T)$, it suffices to obtain an upper bound on the energy $\phi^*(y^{T+1})$. By Lemma 28, the energy ϕ^* increases by at most a constant between each phase $k = 0, \dots, K$. On the other hand, depending on the magnitude of η, and due to certain boundary cases between regions P_i and $P_{i\,\sim(i+1)}$, the energy may not be *strictly* increasing between every phase (e.g., see cases (iii), (iv) of Lemma 31). Thus similar to the proof of Theorem 11, for each phase $k = 0, 1, \dots, K$, we define the indicator variable $c_k \in \{0, 1\}$ as follows:

$$c_k := \begin{cases} 0 & \text{if } \phi^*(y^{t_k}) = \phi^*(y^{t_{k-1}}) \\ 1 & \text{if } \phi^*(y^{t_k}) > \phi^*(y^{t_{k-1}}) \end{cases} .$$

Lemmas 31 and 28 imply that if $c_k = 1$, then $\phi^*(y^{t_k}) - \phi^*(y^{t_{k-1}}) = \Theta(1) = \Theta(c_k)$. Thus it holds that

$$\phi^*(y^{T+1}) = \Theta\Big(\sum_{k=1}^{K} c_k\Big) . \tag{44}$$

Now recall that for readability we use the notation $f \lesssim g$ and $f \gtrsim g$ to indicate $f = O(g)$ or $f = \Omega(g)$ respectively. Then using the phase length bound in Part (iii) of Lemma 28, together with the fact that $\phi^*(y^{t_k}) - \phi^*(y^{t_{k-1}}) = \Theta(c_k)$, we have

$$\tau_k \;\geq\; \alpha_k \cdot \phi^*(y^{t_k}) - \beta_k \gtrsim \sum_{i=1}^{k} c_i - \beta_k \tag{45}$$

for positive constants $\alpha_k, \beta_k > 0$

From here, we repeat identical calculations as in expressions (22) through (26) from the proof of Theorem 11. As a consequence, we similarly obtain

$$\Big(\sum_{k=0}^{K} c_k\Big)^2 \lesssim T \quad \implies \quad \sum_{k=0}^{K} c_k \lesssim \sqrt{T}\,.$$

By expression (44), this in turn gives $\phi^*(y^{T+1}) \lesssim O(\sqrt{T})$. By Proposition 7, we then conclude that

$$\text{Reg}(T) \;\leq\; \frac{2\phi^*(y^{T+1})}{\eta} + \frac{2M}{\eta} \;\leq\; O(\sqrt{T})\,.$$

Here, the last inequality comes from the fact that $M = \max_{x \in \Delta_n} \frac{\|x\|_2^2}{2} \leq 1$ and η is an absolute constant. ■

14. Boundary Behavior of GD in Small Stepsize Regime

14.1. Overview of Results

In this section, we present several auxiliary results related to the behavior of Gradient Descent under different stepsize regimes. We start by presenting an overview of these results: First, recall that in the large stepsize regime, Lemmas 16 and 28 together imply that after at most a single iteration, every primal iterate of Gradient Descent will lie on the boundary of Δ_n. In the following theorem, we show that a similar *boundary invariance* property holds for any $\eta > 0$: if the energy ever exceeds a (game-dependent) constant value, every subsequent primal iterate of Gradient Descent must lie on the boundary of Δ_n.

Theorem 32 *Let $\{x^t\}$ and $\{y^t\}$ be the iterates of* (GD Primal-Dual) *on an n-dimensional RPS matrix A with $\eta > 0$, and from any initial $x^0 \in \Delta_n$. Then there exists a constant D (depending on the entries of A) such that, if the energy $\phi^*(y^{t'}) > D$ for some $t' \geq 0$, then x^t is on the boundary of Δ_n at all subsequent iterations $t > t'$.*

Theorem 32 extends the boundary invariance result of Bailey and Piliouras (2019a) for 2×2 games to the present n-dimensional symmetric setting. Note that while boundary invariance was the key driver in establishing $O(\sqrt{T})$ regret in the large stepsize regime (specifically, the cycling between vertices of Δ_n as seen in the right plot of Figure 2), this property alone appears insufficient for establishing sublinear regret bounds using smaller constant stepsizes: as seen in the middle plot of Figure 2, even for moderately small constant η, the primal iterates of Gradient Descent may transition between different boundary faces of Δ_n in an irregular manner, making the energy growth harder to control. Because of this, for smaller constant η, simply ensuring that the primal iterates are

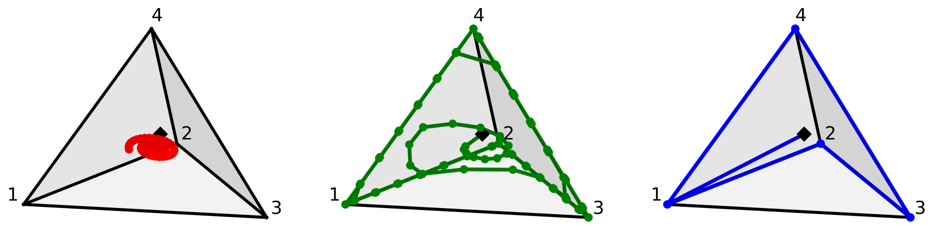


Figure 2: The primal iterates of Gradient Descent in $n = 4$ unweighted RPS run for $T = 100$ steps, and initialized at $x^0 = [0.2, 0.2, 0.25, 0.35]$. From left to right, using stepsize $\eta = \{1/\sqrt{T}, 0.3, 10\}$.

on the boundary of Δ_n is insufficient for establishing a similar phase behavior and regret guarantee as in Lemma 28 and Theorem 17. Obtaining $O(\sqrt{T})$ regret bounds for Gradient Descent using *any* constant stepsize thus remains open.

At the other extreme, while using a small time-horizon-dependent stepsize (e.g., with $\eta = \Theta(1/\sqrt{T})$) might result in all primal iterates remaining *interior* (as seen in the left plot of Figure 2), note that in this case, Proposition 7 (which establishes $\mathrm{Reg}(T) \le O(\phi^*(y^{T+1})/\eta)$) still only guarantees a worst case regret bound scaling with $\sqrt{T}$. In particular, we prove the following guarantee:

Lemma 33 *Let $\{x^t\}$ and $\{y^t\}$ be the iterates of* (GD Primal-Dual) *on an n-dimensional RPS matrix A with $\eta = \frac{1}{\sqrt{T}}$, and from an interior $x^0 \in \Delta_n$. Then if x^t is interior for all $t \ge 1$, then*

$$\phi^*(y^{T+1}) \le \frac{na_{\max}^2}{2} \quad \text{and} \quad \mathrm{Reg}(T) \le \sqrt{T} \cdot \left(\frac{na_{\max}^2}{2} + 1\right) .$$

In the following subsections, we develop the proofs of Theorem 32 and Lemma 33, but we begin by first establishing a certain low-dimensional representation of the dual iterates of Gradient Descent on RPS games.

14.2. Lower-dimensional Representation of Dual Iterates

Here, we establish the following lower dimensional representation of the dual iterates $\{y^t\}$ of Gradient Descent on a fixed $(n-1)$-dimensional subspace. This representation is leveraged in the proof of Theorem 32, and it follows by the existence of an interior equilibrium x^* for every RPS matrix A from Definition 3. Stated formally:

Lemma 34 *Let A be an n-dimensional RPS matrix from Definition 3, and let $x^* \in \Delta_n$ denote an interior NE for A. Let $\{x^t\}$ and $\{y^t\}$ denote the iterates of* (GD Primal-Dual) *on A using any stepsize $\eta > 0$ and from any $x^0 \in \Delta_n$. Then there exist positive constants $\alpha_1, \dots, \alpha_{n-1} > 0$ (depending only on A) such that, for all t:*

$$y_n^t \;=\; -\big(\alpha_1 \cdot y_1^t + \cdots + \alpha_{n-1} \cdot y_{n-1}^t\big) .$$

Proof Recall from Proposition 8.2, that $Ax^* = 0$. Then by the skew-symmetry of A, observe that for any $x \in \Delta_n$: $\langle x^*, Ax\rangle = \langle x, A^\top x^*\rangle = -\langle x, Ax^*\rangle = 0$. Now by definition of (GD Primal-Dual), $y^t = \eta \sum_{k=0}^{t-1} Ax^k$, and thus $\langle x^*, y^t\rangle = \eta \sum_{k=0}^{t-1}\langle x^*, Ax^k\rangle = 0$. It follows that we can write

$$y_n^t x_n^* \;=\; -(x_1^* y_1^t + \cdots + x_{n-1}^* y_{n-1}^t) .$$

Rearranging, and letting $\alpha_i := (x_i^*/x_n^*) > 0$ for all $i = 1, \dots, n-1$ (where positivity of each α_i follows from the fact that x^* is interior), then completes the proof. ■

14.3. Sufficient Energy Growth Implies Primal Iterates on Boundary

In this section, we give the proof of Theorem 32, which we restate here:

Theorem 32 *Let $\{x^t\}$ and $\{y^t\}$ be the iterates of* (GD Primal-Dual) *on an n-dimensional RPS matrix A with $\eta > 0$, and from any initial $x^0 \in \Delta_n$. Then there exists a constant D (depending on the entries of A) such that, if the energy $\phi^*(y^{t'}) > D$ for some $t' \geq 0$, then x^t is on the boundary of Δ_n at all subsequent iterations $t > t'$.*

Proof Let $\bar{P} \subseteq \mathbb{R}^n$ be the set defined by:

$$\bar{P} \;=\; \left\{ y \in \mathbb{R}^n : y_i - \frac{1}{n}\big(\langle y, \mathbf{1}\rangle\big) + \frac{1}{n} \geq 0 \text{ for all } i \in [n] \right\} . \tag{46}$$

By the closed-form expression for the primal iterates of Gradient Descent from Proposition 25, observe by definition of $\bar{P}$ that $\mathrm{supp}(x^t) = [n]$ if and only if $y^t \in \bar{P}$.

Now recall by Lemma 34 that there exist positive constants $\alpha_1, \ldots, \alpha_{n-1} > 0$ such that $y^t_n = -\Big(\sum_{i=1}^{n-1} \alpha_i y^t_i \Big)$ for all times t. In other words, under (GD Primal-Dual), the dual iterates $\{y^t\}$ all lie on an $(n-1)$-dimensional subspace U of $\mathbb{R}^n$, where

$$U \;=\; \left\{ y \in \mathbb{R}^n : y_n = -\big(\textstyle\sum_{i=1}^{n-1} \alpha_i y_i \big) \right\} \tag{47}$$

Now by slight abuse of notation, for a vector $y \in \mathbb{R}^n$, let $[y]_n \in \mathbb{R}^{n-1}$ denote its first $(n-1)$ coordinates. Then we define the subset $\bar{L} \subset \mathbb{R}^{n-1}$ as the intersection of $\bar{P}$ and the subspace U:

$$\bar{L} \;=\; \left\{ z \in \mathbb{R}^{n-1} : z = [y]_n, \text{ for } y \in \bar{P} \right\} . \tag{48}$$

For any $z \in \mathbb{R}^{n-1}$, let $y \in \mathbb{R}^n$ be the vector $\Gamma(z) := \big(z_1, \ldots, z_{n-1}, -(\sum_{i=1}^{n-1} \alpha_i z_i)\big) \in \mathbb{R}^n$. It follows that $z \in \bar{L} \iff \Gamma(z) \in \bar{P}$. Moreover, by construction $\bar{L}$ is compact.

Now for $z \in \mathbb{R}^{n-1}$ (again by slight abuse of notation) let $H(z)$ be the function $\phi^*(\Gamma(z))$ (e.g., using the variable substitution $y_n = -(\sum_{i=1}^{n-1} \alpha_i z_i)$ in the energy function $\phi^*(y)$). It follows that H is continuous over $\mathbb{R}^n$, and by the compactness of $\bar{L}$, it follows that H has a maximum value $D = \max_{z \in \bar{L}} H(z)$ over $\bar{L}$. Observe that D is a constant depending only on $\{\alpha_i\}$, which depend only on the matrix A.

For the dual iterates $\{y^t\}$ of (GD Primal-Dual), let z^t be the corresponding iterates $z^t = [y^t]_n$. Recall by Proposition 8 that $\phi^*(y^{t+1}) - \phi^*(y^t) \geq 0$ for all t, which also means that $H(z^{t+1}) - H(z^t) \geq 0$ for all t. Then by the arguments above, suppose at some time $t' > 0$ that $\phi^*(y^{t'}) = H(z^{t'}) > D$. It follows that $z^{t'} \notin \bar{L}$, which implies $y^{t'} \notin \bar{P}$ and thus $\mathrm{supp}(x^{t'}) \neq [n]$ (meaning $x^{t'}$ is on the boundary of Δ_n). As H is non-decreasing over time, for every subsequent iterate z^t for $t \geq t'$, we must also have $H(z^t) > D$. By the same reasoning, x^t is on the boundary of Δ_n, which yields the statement of the theorem. ■

14.4. Regret Bound for Time-Vanishing Stepsize

In this section, we give the proof of Lemma 33, restated below:

Lemma 33 *Let $\{x^t\}$ and $\{y^t\}$ be the iterates of* (GD Primal-Dual) *on an n-dimensional RPS matrix A with $\eta = \frac{1}{\sqrt{T}}$, and from an interior $x^0 \in \Delta_n$. Then if x^t is interior for all $t \geq 1$, then*

$$\phi^*(y^{T+1}) \leq \frac{na_{\max}^2}{2} \quad \textit{and} \quad \mathrm{Reg}(T) \leq \sqrt{T} \cdot \left(\frac{na_{\max}^2}{2} + 1\right) .$$

Proof Recall Proposition 25 and expression (33) that under (GD Primal-Dual), if x^t is interior (meaning $\mathrm{supp}(x^t) = [n]$), then the energy of the corresponding dual iterate y^t is given by:

$$\phi^*(y^t) \ = \ \frac{\|y^t\|_2^2}{2} + \frac{\langle y^t, \mathbf{1}\rangle}{2n} - \frac{(\langle y^t, \mathbf{1}\rangle)^2}{2n} - \frac{1}{2n} \,, \tag{49}$$

which is 1-smooth. As x^t is interior for all $t \geq 0$ by assumption, the energy function $\phi^*(y^t)$ is as given in (49) for all $t \geq 1$. Then by smoothness, we have for all $t \geq 1$ that

$$\begin{aligned}
\phi^*(y^{t+1}) - \phi^*(y^t) &\leq \langle \nabla\Psi(y^t), y^{t+1} - y^t\rangle + \frac{1}{2}\|y^{t+1} - y^t\|_2^2 \\
&= \langle \nabla\phi^*(y^t), \eta A \nabla\phi^*(y^t)\rangle + \frac{\eta^2}{2}\|A\nabla\phi^*(y^t)\|_2^2 \\
&= \frac{\eta^2}{2}\|A\nabla\phi^*(y^t)\|_2^2 \\
&\leq \frac{\eta^2 \cdot na_{\max}^2}{2} \,,
\end{aligned}$$

where the second equality follows by skew-symmetry of A, and the final inequality follows by the fact that $\nabla\phi^*(y^t) = x^t$ is a probability distribution. By definition, $\phi^*(y^0) = \phi^*(0) = -1/2n$, and it follows that

$$\phi^*(y^{T+1}) \ = \ \sum_{t=0}^{T} \phi^*(y^{t+1}) - \phi^*(y^t) \ \leq \ T \cdot \frac{\eta^2 \cdot na_{\max}^2}{2} \ = \ \frac{na_{\max}^2}{2} \,,$$

where the last inequality follows from the setting of $\eta = \frac{1}{\sqrt{T}}$. This proves the first claim of the lemma.

On the other hand, by Proposition 7, we find (recalling that $\max_{x\in\Delta_n}(\|x\|_2^2)/2 = 1$) that

$$\mathrm{Reg}(T) \ \leq \ \frac{2 \cdot \phi^*(y^{T+1})}{\eta} + \frac{2}{\eta} \ \leq \ \sqrt{T} \cdot \left(\frac{na_{\max}^2}{2} + 1\right) ,$$

which proves the second statement of the lemma. ■

15. Experimental Results

Throughout the plots in this paper, we use (◆) to denote the *initial condition* of the dynamics.

15.1. Behavior of Primal Iterates

Comparing the behavior of GD in three and four-dimensional RPS. In Figure 3 we show the behavior of the primal iterates of Gradient Descent on unweighted $n = 3$ RPS, analogous to plots for unweighted $n = 4$ RPS from Figure 2. Comparing the $n = 3$ and $n = 4$ experiments, we observe that the small ($\eta = 1/\sqrt{T}$) and large ($\eta = 10$) stepsize regimes result in qualitatively similar behavior: the GD iterates stay interior in the small stepsize regime, and they converge to the boundary in one step in the large stepsize regime.

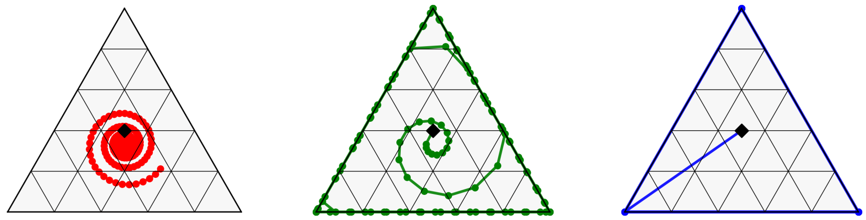
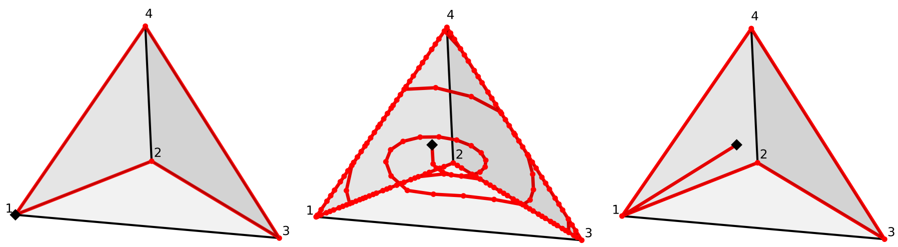


Figure 3: Primal iterates of GD in $n = 3$ unweighted RPS initialized at $x^0 = [0.3, 0.4, 0.3]$ and run for $T = 100$ steps. From left to right, we show the iterates using stepsizes $\eta = \{1/\sqrt{T}, 0.3, 10\}$.



(a) FP (b) GD ($\eta = 0.3$) (c) GD ($\eta = 10$)

Figure 4: Primal iterates of FP initialized at $x^0 = [1, 0, 0, 0]$ and GD initialized at $x^0 = [0.35, 0.05, 0.21, 0.39]$, run for $T = 200$ steps.

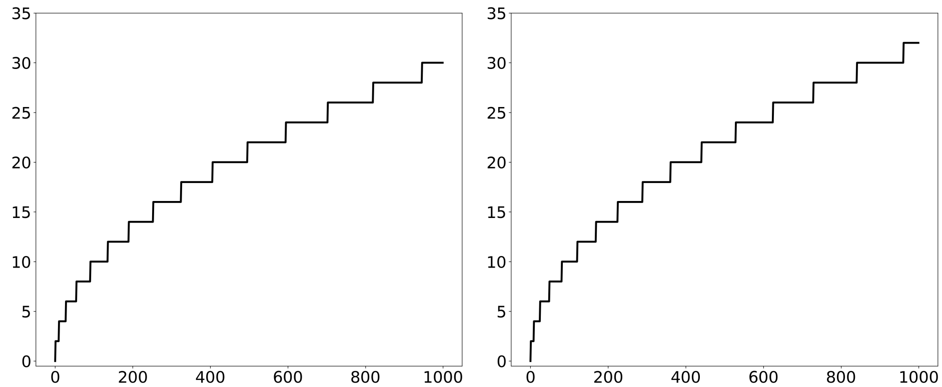
Comparing the behavior of FP and GD in four-dimensional RPS. In Figure 4, we plot the primal iterates of Fictitious Play and Gradient Descent in $n = 4$ unweighted RPS. The left plot shows Fictitious Play initialized at a vertex, and its subsequent cycling between the four vertices. The right plot shows Gradient Descent initialized from an interior distribution, with large stepsize $\eta = 10$. As expected from Lemma 16, the first step of the dynamic goes to vertex e_1, and the iterates subsequently cycle similarly to the FP dynamic thereafter. In the middle plot, using a moderate constant setting of $\eta = 0.3$, the iterates of Gradient Descent eventually converge and remain on the boundary, but in (an initially) less regular manner compared to when $\eta = 10$.

15.2. Empirical Regret

To corroborate our regret bounds from Sections 4 and 5, in this section we empirically examine the regret of Fictitious Play and Gradient Descent on RPS matrices.

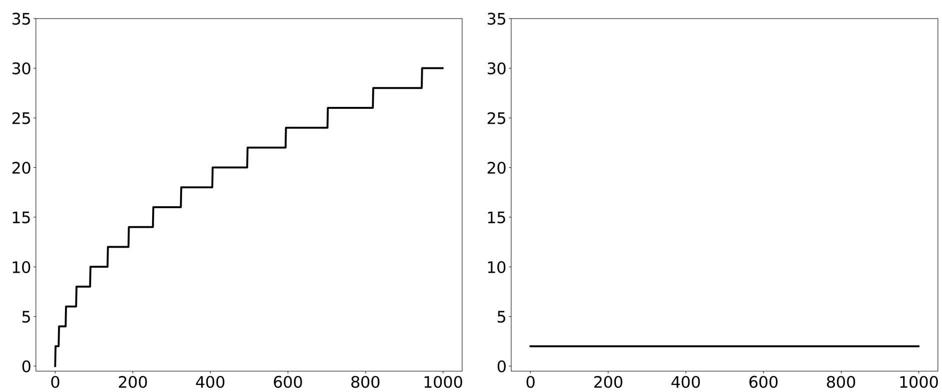
Fictitious Play. Figure 5 shows the total regret of FP in $n = 3$ and 4 RPS scaling like $O(\sqrt{T})$, corroborating Theorem 11. In Figure 6, we also corroborate the result of Theorem 24, showing that when using the ‘tournament’ tiebreaking rule of Definition 23, FP achieves constant regret.

Gradient Descent. In Figure 7, we plot the regret obtained by GD for stepsizes $\eta = \{\frac{1}{\sqrt{T}}, 0.3, 10\}$ and run for $T = 1000$ steps. Figure 7(c) corroborates the $O(\sqrt{T})$ bound of Theorem 17. Finally, in Figure 8, we compare the regret of Gradient Descent with *time-decreasing* stepsizes (specifically, $\eta_t = 1/\sqrt{t}$ for $t = 1, \ldots, T$) to Gradient Descent with large, constant stepsize ($\eta = 10$). Surprisingly, in both Figures 7 and 8, we observe that the large stepsize setting obtains lower empirical regret than the time-horizon dependent/decreasing stepsize setting. Investigating this behavior more thoroughly is an interesting direction for future work.



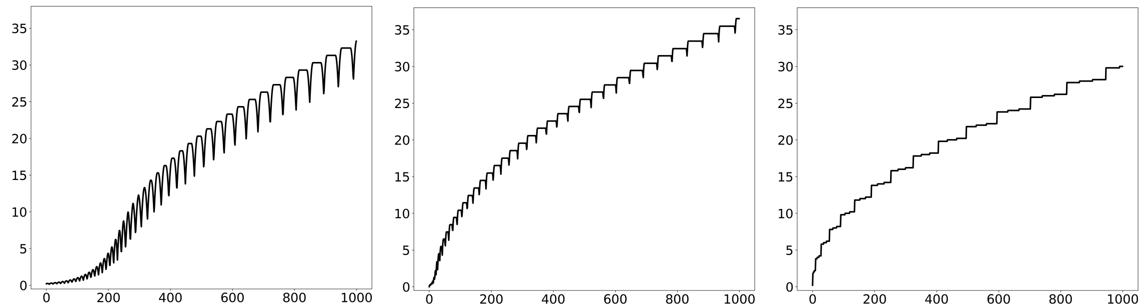
(a) $n = 3$ RPS (b) $n = 4$ RPS

Figure 5: Total regret vs iteration for FP initialized at $x^0 = [1, 0, 0, 0]$ in $n = 3$ and 4 RPS, run for $T = 1000$ steps.



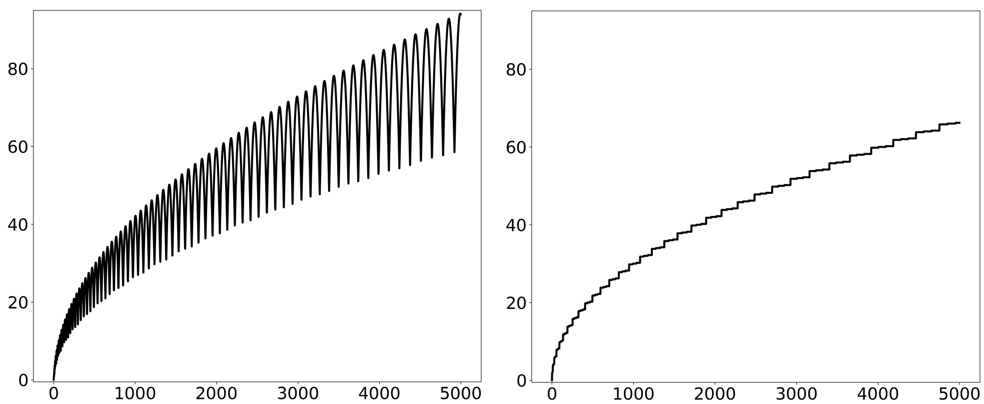
(a) Lexicographic tiebreaking (b) Tournament tiebreaking

Figure 6: Total regret vs iteration for FP with different tiebreaking rules initialized at $x^0 = [1, 0, 0, 0]$ in $n = 3$ RPS, run for $T = 1000$ steps.



(a) $\eta = 1/\sqrt{T}$ (b) $\eta = 0.3$ (c) $\eta = 10$

Figure 7: Total regret vs iteration for Gradient Descent initialized at $x^0 = [0.3, 0.4, 0.3]$ in $n = 3$ RPS with stepsizes $\eta = \{1/\sqrt{T}, 0.3, 10\}$, run for $T = 1000$ steps.



(a) $\eta = 1/\sqrt{t}$ (b) $\eta = 10$

Figure 8: Total regret vs iteration for Gradient Descent initialized at $x^0 = [0.3, 0.4, 0.3]$ in $n = 3$ RPS with stepsizes $\eta = \{1/\sqrt{t}, 10\}$, run for $T = 5000$ steps.