Proceedings of Machine Learning Research vol 313:1–38, 2026 37th International Conference on Algorithmic Learning Theory

Last-iterate Convergence for Symmetric, General-sum, 2×2 Games Under The Exponential Weights Dynamic

Guanghui Wang GWANG369@GATECH.EDU
College of Computing, Georgia Institute of Technology

Krishna Acharya KRISHNA.ACHARYA@GATECH.EDU
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

Lokranjan Lakshmikanthan LLAKSHMI3@GATECH.EDU
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

Juba Ziani JZIANI3@GATECH.EDU
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

Vidya Muthukumar VMUTHUKUMAR8@GATECH.EDU
School of Electrical and Computer Engineering, Georgia Institute of Technology
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

Editors: Matus Telgarsky and Jonathan Ullman

Abstract

We conduct a comprehensive analysis of the discrete-time exponential-weights dynamic with a constant step size on all *general-sum and symmetric* 2×2 normal-form games, i.e. games with 2 pure strategies per player, and where the ensuing payoff tuple is of the form $(A, A^\top)$ (where A is the 2×2 payoff matrix corresponding to the first player). Such symmetric games commonly arise in real-world interactions between "symmetric" agents who have identically defined utility functions — such as Bertrand competition and multi-agent performative prediction, and display a rich multiplicity of equilibria despite the seemingly simple setting. Somewhat surprisingly, we show through a first-principles analysis that the exponential weights dynamic, which is popular in online learning, converges in the last iterate for such games regardless of initialization with an appropriately chosen step size. For certain games and/or initializations, we further show that the convergence rate is in fact exponential and holds for any step size.

We illustrate our theory with extensive simulations and applications to the aforementioned game-theoretic interactions. In the case of multi-agent performative prediction, we formulate a new "mortgage competition" game between lenders (i.e. banks) who interact with a population of customers, and show that it fits into our framework.

Keywords: Exponential Weights, Symmetric Game, Last-iterate Convergence

1. Introduction

Game theory is a classical mathematical framework in which the formal study of multi-agent non-cooperative interaction is situated (Myerson, 2013; Karlin and Peres, 2017). In a "one-shot" normal-form game, if the players have full knowledge of the game, are selfish and rational, and no mediator is present, they are expected to play a *Nash equilibrium* (NE) strategy (Myerson, 2013) (see also Section 2 for a self-contained definition). Emerging applications of game theory abound in modern machine learning and artificial intelligence, such as the training of generative adversarial networks and robust models (Daskalakis et al., 2018; Madry et al., 2018), multi-agent reinforcement

© 2026 G. Wang, K. Acharya, L. Lakshmikanthan, J. Ziani & V. Muthukumar.

learning (Zhang et al., 2021), and situations where the pipeline for data may respond *strategically* to ML model deployment (commonly called *strategic classification* (Hardt et al., 2016) or more recently *performative prediction* (Perdomo et al., 2020)). Here, agents are actively *learning* about the game being played, often through repeated interaction. This raises the natural questions of: a) what learning algorithms should the agents use when interacting with one another, and b) do the ensuing *game dynamics* converge *in a day-to-day sense* to some Nash equilibrium?

These questions have a decades-long history in an interdisciplinary community spanning economics (Foster and Vohra, 1997, 1999; Hart and Mas-Colell, 2000; Hart, 2005), online learning (Hannan, 1957; Freund and Schapire, 1999; Cesa-Bianchi and Lugosi, 2006), control and dynamical systems (Hofbauer and Sigmund, 1998; Fox and Shamma, 2013; Bailey and Piliouras, 2018). In fact, game dynamics remain generally difficult to analyze and characterize, and most positive results in the literature are limited in scope. For example, it is known that a very specific class of dynamics satisfying *vanishing (external) regret*—initially introduced by Hannan (1957)[1]—played by one player, combined with best-response play by the other, converges to a Nash equilibrium in the special case of zero-sum games. When both players use no-regret strategies, the dynamics are only guaranteed to converge to a broader notion of equilibrium, a *coarse correlated equilibrium* (Cesa-Bianchi and Lugosi, 2006) (CCE), in general-sum games. Importantly, these convergence guarantees hold only in a *time-average* sense: i.e. the average of the strategies over time converges.

This provides no guarantee for the actual play at any given round. In fact, guarantees for convergence in a *last-iterate* sense—where the strategy at each individual time step t eventually approaches equilibrium—-are far more limited. Most results in this direction are negative, showing non-convergence or even chaotic behavior in broad classes of games. Our object of focus, the *Exponential Weights* dynamic (Cesa-Bianchi et al., 1997) (commonly called the replicator dynamics in continuous time as surveyed in Hofbauer and Sigmund (1998)), has been shown to be oscillatory for simple zero/constant-sum games (Bailey and Piliouras, 2018; Cheung, 2018) and more generally can even be chaotic, both for zero-sum and general-sum games (Cheung and Piliouras, 2019; Palaiopanos et al., 2017; Chotibut et al., 2021; Cheung and Piliouras, 2020; Cheung et al., 2021). While an optimistic variant of exponential weights (and/or, relatedly, an extragradient method) is commonly employed to mitigate the oscillations in the zero-sum case (Daskalakis and Panageas, 2019; Cai et al., 2022), a fix for non-zero-sum games remains elusive. In fact, there exist general-sum finite games for which *any* uncoupled, "first-order" dynamic would not converge (Hart and Mas-Colell, 2003).

In this paper, we establish a new set of surprisingly positive last-iterate convergence results for a previously overlooked class of games: two-player symmetric games with two actions per player, i.e. 2×2 *symmetric games*. Formally, these are normal-form games with 2 players, 2 actions, and where the payoff matrices of players 1 and 2, respectively A and B, satisfy[2] $B = A^\top$. Symmetric games describe many practical scenarios of interest. For example, applications such as road traffic and packet routing on networks can be modeled as congestion games—where strategies correspond to path choice, and "traffic" affects all participants equally. These games, in their atomic form, are known to correspond to symmetric games (Rosenthal, 1973). Symmetry also arises in Bertrand competition games, which inspire the canonical setting of multi-agent performative prediction (Narang et al., 2023) as well as our practically motivated case study to follow.

1. We discuss the implications of stronger notions of regret, such as internal/swap regret (Blum and Mansour, 2007; Greenwald and Jafari, 2003), in Section 5.
2. Note that this is a zero-sum game when A is a skew-symmetric matrix.

We center our study on the *Exponential Weights* (EW) algorithm, one of the most celebrated no-regret learning rules (Cesa-Bianchi et al., 1997). EW is widely used in online learning applications (both full-information and bandit learning) due to its simplicity and optimal (worst-case) regret guarantees. Importantly, we identify a significantly stronger form of convergence compared to the related literature: we are able to show not only that our dynamics i) converge in a *last-iterate* (as opposed to the standard time-average) sense in 2×2 symmetric games, but also that ii) they converge to specific Nash equilibria of our game, as opposed to the more general class of CCEs.

Our contributions: We conduct a systematic and comprehensive study of the exponential weights dynamic in the simple but non-trivial case of 2×2 symmetric games defined above. Note that even this simple setting admits a multiplicity of Nash and correlated equilibria — see Table 2 for a complete characterization. In particular, we show the following results:

- Section 3 shows that, surprisingly, the exponential weights dynamic always converges to either a pure or a strictly mixed NE for *any* "non-degenerate" initialization and on *any* "non-degenerate" 2×2 symmetric game[3]. In the cases where convergence to a pure NE is achieved, the step size/learning rate used in exponential weights can be any positive constant. When convergence to a strictly mixed NE is achieved, the step size needs to be smaller than a positive constant that depends on the parameters of the game. We obtain these results by conducting an intricate, first-principles, case-by-case analysis of the exponential weights dynamic.
- In Section 4, we interpret our results in the context of multi-agent performative prediction between 2 agents. We design a stylized, but canonical model of mortgage competition between two banks (or "learners"), where the strategies consist of interest rates and credit score thresholds, and the utilities of the banks depend on the distribution of the customers who choose each of the banks, as well as the probability of repaying their loans. Interestingly, our results in Section 3 imply that convergence to *asymmetric* pure NE can happen (despite the fact that a symmetric NE, either pure or mixed, always exists for symmetric games).

We next set up the formal framework for 2×2 symmetric games, and present our main results in Sections 3 and 4. We provide a detailed treatment of related work in Section 5.

2. Preliminaries

We begin by introducing the basic setting of 2-player, 2-action symmetric games. We then briefly review basic equilibrium concepts, and present (for completeness) the Nash and correlated equilibrium landscape of this setting.

Two-player-two-action symmetric games: We study a general two-player, two-action symmetric game. Each player $i \in \{1, 2\}$ selects one of two available actions (i.e. pure strategies), which we denote by θ_1 and θ_2. The corresponding payoffs are summarized in Table 1. Note that the 2×2 utility matrix A is parameterized by fixed constants $a, b, c, d \in \mathbb{R}$. We also define

$$\epsilon_1 := a - b \text{ and } \epsilon_2 = c - d \tag{1}$$

as two shorthand parameters—we will see that the signs of these parameters critically dictate both the equilibrium landscape and the behavior of exponential weights.

3. See Assumption 1 for formal assumptions. A "degenerate" initialization would be one for which the dynamics never move, and a "degenerate" game would be one for which any tuple of strategies constitutes a NE.

Player 1 utility matrix A

Player 2 \ Player 1	θ_1	θ_2
θ_1	a	b
θ_2	c	d

Player 2 utility matrix $B = A^\top$

Player 2 \ Player 1	θ_1	θ_2
θ_1	a	c
θ_2	b	d

Table 1: Utility matrices of Player 1 and Player 2

Condition	Pure NE	Strict Mixed-NE	CE
$\epsilon_1 < 0, \epsilon_2 < 0$	(θ_2, θ_2)	–	(θ_2, θ_2)
$\epsilon_1 > 0, \epsilon_2 > 0$	(θ_1, θ_1)	–	(θ_1, θ_1)
$\epsilon_1 < 0, \epsilon_2 > 0$	(θ_1, θ_2), (θ_2, θ_1)	$(p_{\text{SE}}, p_{\text{SE}})$	Given in (6)
$\epsilon_1 > 0, \epsilon_2 < 0$	(θ_1, θ_1), (θ_2, θ_2)	$(p_{\text{SE}}, p_{\text{SE}})$	Given in (7)
$\epsilon_1 = 0, \epsilon_2 < 0$	(θ_1, θ_1), (θ_2, θ_2)	–	Given in (8)
$\epsilon_1 = 0, \epsilon_2 > 0$	(θ_1, θ_1), (θ_2, θ_1), (θ_1, θ_2)	$(\theta_1, p), (p, \theta_1)$	Given in (9)

Table 2: NE under different conditions. Recall that $\epsilon_1 = a - b$, while $\epsilon_2 = c - d$. $p \in \Delta_2$ stands for any mixed strategy in the 2-dimensional simplex. We define $p_{\text{SE}} = \left(\frac{|\epsilon_2|}{|\epsilon_1|+|\epsilon_2|}, \frac{|\epsilon_1|}{|\epsilon_1|+|\epsilon_2|}\right)$.

For any pair of indices $m, n \in \{1, 2\}$, let $u_i(\theta_m, \theta_n)$ denote the payoff of player $i \in \{1, 2\}$ when player 1 chooses θ_m and player 2 chooses θ_n. Note that by symmetry of the game, we have $u_1(\theta_m, \theta_n) = u_2(\theta_n, \theta_m)$ for all $m, n \in \{1, 2\}$. We further define a pair of mixed strategies (p_1, p_2) where $p_i := (p_{i,1}, p_{i,2})$ denotes a probability distribution (on the 2-dimensional simplex denoted by Δ_2) over pure strategies θ_1 and θ_2 respectively for player $i \in \{1, 2\}$. Next, we recall the following basic equilibrium concepts.

Definition 1 (Nash equilibrium) *A pair of strategies* $(\theta_{m^*}, \theta_{n^*})$ *constitutes a* pure Nash equilibrium *(pure NE) iff the following conditions hold::*

$$u_1(\theta_{m^*}, \theta_{n^*}) \geq \max_{\theta' \in \{\theta_1, \theta_2\}} u_1(\theta', \theta_{n^*}), \quad \textit{and} \quad u_2(\theta_{m^*}, \theta_{n^*}) \geq \max_{\theta' \in \{\theta_1, \theta_2\}} u_2(\theta_{m^*}, \theta').$$

A pair of mixed strategies (p_1^*, p_2^*) *constitutes a* mixed Nash equilibrium *(mixed NE) iff the following two conditions hold:*

$$\mathbb{E}_{M \sim p_1^*, N \sim p_2^*}[u_1(\theta_M, \theta_N)] \geq \max_{p_1 \in \Delta_2} \mathbb{E}_{M' \sim p_1, N \sim p_2^*}[u_1(\theta_{M'}, \theta_N)],$$
$$\mathbb{E}_{M \sim p_1^*, N \sim p_2^*}[u_2(\theta_M, \theta_N)] \geq \max_{p_2 \in \Delta_2} \mathbb{E}_{M \sim p_1^*, N' \in p_2}[u_2(\theta_M, \theta_N)].$$

In Nash equilibria, players pick their actions simultaneously and independently of each other. While our central focus in this paper is convergence to NE, we also briefly consider the broader concept of *correlated equilibria* (CE) of a game. This concept was introduced by Aumann (1987) and, informally, uses a signaling device that recommends some action (θ_M, θ_N) to the players where (M, N) is now drawn from a *joint distribution* (for example, think of a traffic signal at an intersection). The signal induces implicit correlation between the players. At a correlated equilibrium, no player benefits from deviating from the recommended action in expectation. We provide a detailed review of the definition of CE, as well as characterize the set of CE for our game, in Appendix C.

Landscape of Nash and correlated equilibria: We now present the NE landscape of our canonical setting of symmetric 2×2 games in Table 2 under different conditions, dictated by the signs of $\epsilon_1 := a - b$ and $\epsilon_2 := c - d$. (Proving that the corresponding entries in the table are correct is a simple exercise in undergraduate game theory that we leave to the reader.) We make a few observations about the landscape of NEs:

- When the signs of ϵ_1 and ϵ_2 match (both strictly positive or negative), there is a unique, *symmetric*, pure NE. It is also easy to verify that these cases have a dominated strategy for both players (e.g. in the case where $\epsilon_1, \epsilon_2 < 0$, θ_1 is dominated by θ_2 for both players). Moreover, as we will subsequently see the set of CEs is also a singleton set comprising of the unique pure NE. Section 3 will demonstrate that these are the "easy" cases for convergence to NE.
- The cases where the signs of ϵ_1 and ϵ_2 *do not* match (including cases where $\epsilon_1 = 0$ or $\epsilon_2 = 0$) are much more interesting as they admit a multiplicity of pure NE as well as potentially *strictly* mixed NE. The set of CEs are also more complex in these cases, rendering a generic no-regret convergence analysis inadequate for our purposes. Proving convergence for these more complex cases is the main novelty in our analysis.
- We do not consider the case of $\epsilon_1 = \epsilon_2 = 0$ as this corresponds to a degenerate game where any tuple of mixed strategies would be a valid NE; see Assumption 1 for more details.

3. Last-iterate Convergence of Exponential Weights

Algorithm 1 Online Learning Dynamic through Exponential Weights

Input: Step size $\eta > 0$, Initialization $p_1^{(1)}, p_2^{(1)} \in \Delta_2$
for $t = 1, 2, 3 \ldots$ **do**
Decision: Each player $i \in \{1, 2\}$ plays θ_j with probability $p_{i,j}^{(t)}$ for $j \in \{1, 2\}$;
Update Rule: Players update their probability distributions to $p_1^{(t+1)}$, $p_2^{(t+1)}$ as follows:

$$p_{1,j}^{(t+1)} \propto p_{1,j}^{(t)} \cdot \exp\left(\eta \sum_{k=1}^{2} p_{2,k}^{(t)} \cdot u_1\left(\theta_j, \theta_k\right)\right), \quad \forall j \in \{1, 2\};$$

$$p_{2,j}^{(t+1)} \propto p_{2,j}^{(t)} \cdot \exp\left(\eta \sum_{k=1}^{2} p_{1,k}^{(t)} \cdot u_2\left(\theta_k, \theta_j\right)\right), \quad \forall j \in \{1, 2\}.$$

end for

We now characterize the last-iterate convergence of the exponential weights algorithm, formally described in Algorithm 1. We use $p_i^{(t)} \in \Delta_2$ to denote the mixed strategy used by player $i \in \{1, 2\}$ at time step $t \geq 1$. Thus, when both players play exponential weights, it induces a sequence of mixed strategy tuples $\{p_1^{(t)}, p_2^{(t)}\}_{t \geq 1}$. Our hope is that this sequence will eventually converge to some NE.

The most salient step in the exponential weights algorithm is the update from $t \to t + 1$, which proceeds as follows. At each time step t, each player $i \in \{1, 2\}$ randomly picks an action according to their current probability distribution $p_i^{(t)}$. Then, each player updates their weight on each action in a manner that is exponentially proportional to the expected utility derived by playing that action (which in turn depends on the mixed strategy of the other player). For example, considering round t and candidate action θ_j, player 1 obtains an expected utility of $\sum_{k=1}^{2} p_{2,k}^{(t)} \cdot u_1\left(\theta_j, \theta_k\right)$ since player

Utility Condition	Initialization	Nature of result	Requirement on η
r1. $\epsilon_1 < 0, \epsilon_2 < 0$	Any	Exponential convergence to pure NE	None
r2. $\epsilon_1 > 0, \epsilon_2 > 0$	Any	Exponential convergence to pure NE	None
r3. $\epsilon_1 < 0, \epsilon_2 > 0$	Opposite-sign	Exponential convergence to one of pure NEs	None
r4. $\epsilon_1 < 0, \epsilon_2 > 0$	Same sign	Asymptotic convergence to one of pure NEs	None
r5. $\epsilon_1 < 0, \epsilon_2 > 0$	Identical	Asymptotic convergence to strictly mixed NE	$< \frac{8}{\lvert\epsilon_1\rvert+\lvert\epsilon_2\rvert}$
r6. $\epsilon_1 > 0, \epsilon_2 < 0$	Same sign	Exponential convergence to one of the pure NEs	None
r7. $\epsilon_1 > 0, \epsilon_2 < 0$	Opposite signs	Asymptotic convergence to one of the NEs	$< \frac{8}{\lvert\epsilon_1\rvert+\lvert\epsilon_2\rvert}$
r8. $\epsilon_1 = 0, \epsilon_2 < 0$	Any	Exponential convergence to pure NE	None
r9. $\epsilon_1 = 0, \epsilon_2 > 0$	Identical	Asymptotic convergence to symmetric pure NE	None
r10. $\epsilon_1 = 0, \epsilon_2 > 0$	Non-identical	Asymptotic convergence to set of mixed NE	None

Table 3: Summary of our convergence results. This table is best read along with Table 2, which summarizes the NE characterization for each case.

2's probability distribution is given by $p_2^{(t)}$. The speed of this update is controlled by a *constant* step size $\eta > 0$. (As of now, our proofs, particularly in the more complex cases, do require both players to use the same step size; relaxing this requirement is an important future direction.)

Recall that we define the shorthand parameters $\epsilon_1 := a - b, \epsilon_2 := c - d$. Let us define the shorthand functional $\Delta_i^{(t)} = p_{i,1}^{(t)}\epsilon_1 + p_{i,2}^{(t)}\epsilon_2$. Before diving into the details, we first make the following very mild *non-degeneracy* assumption about the games and the algorithm's initialization:

Assumption 1 *Throughout the paper, we assume: 1) $|\epsilon_1| + |\epsilon_2| > 0$; 2) $p_i^{(1)}$ is not a pure strategy for either player $i \in \{1, 2\}$; 3) the initial functionals $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ are both non-zero.*

The assumptions above only exclude trivial cases that are straightforward to analyze. First, note that when $\epsilon_1 = \epsilon_2 = 0$, the algorithm does not update—since all strategy profiles in this case are a mixed NE, the algorithm in fact "converges" (in a single time step) to some NE. Regarding 2), if both $p_1^{(1)}$ and $p_2^{(1)}$ are pure strategies, then the algorithm does not update; if, for example, $p_1^{(1)}$ is a pure strategy but $p_2^{(1)}$ is not, then $p_1^{(t)}$ remains the same for all t, while $p_2^{(t)}$ converges to the corresponding best response—clearly, the resulting tuple need not be a NE.

Finally and regarding 3), if both $\Delta_1^{(1)} = \Delta_2^{(1)} = 0$, the algorithm does not update; if only one of them is 0, then by round $t = 2$ both become non-zero and automatically satisfy Assumption 1 thereafter. This is formalized in the following lemma, which is proved in Appendix B.

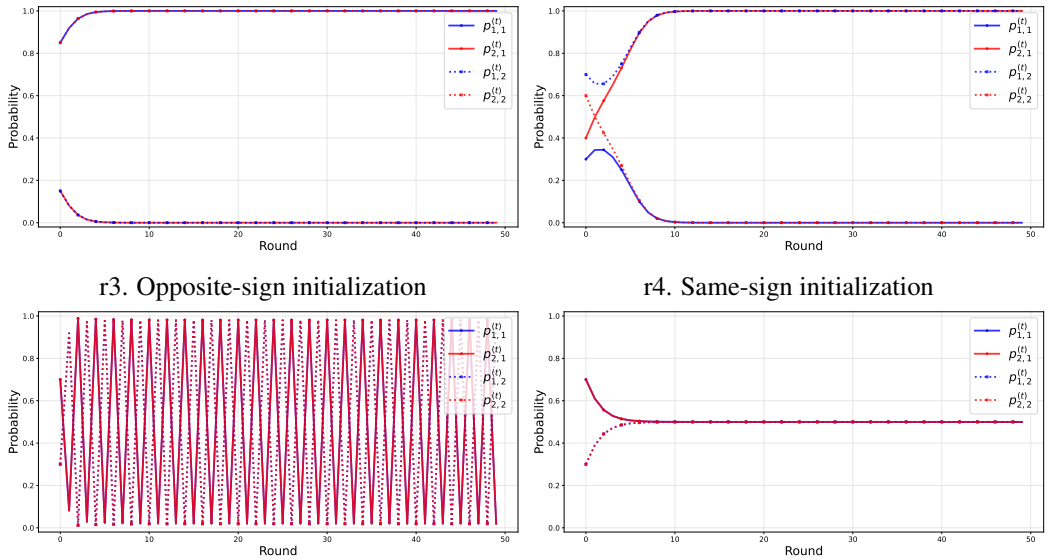
Lemma 2 *Suppose $|\epsilon_1| + |\epsilon_2| > 0$, and $p_1^{(1)}$ and $p_2^{(1)}$ are not pure strategies. We have the following:*

- *If $\Delta_1^{(1)} = \Delta_2^{(1)} = 0$, then $p_i^{(t)}$ remains the same for all $t \geq 1$ and $i \in \{1, 2\}$.*
- *If (without loss of generality) $\Delta_1^{(1)} = 0$ and $\Delta_2^{(1)} \neq 0$, then $\Delta_1^{(2)}, \Delta_2^{(2)} \neq 0$.*

3.1. Main result: convergence of the last iterate

Our main result, Theorem 1 (Parts I-IV), fully characterizes the convergence of Algorithm 1. Our results in a nutshell are informally summarized in Table 3.

We start with the most *naive* case, where ϵ_1 and ϵ_2 have the same sign. This case is naive because it implies the existence of a common dominating action for both players, which also determines the



r3. Opposite-sign initialization

r4. Same-sign initialization

r5. Identical initialization, with $\eta > \frac{8}{|\epsilon_1|+|\epsilon_2|}$

r5. Identical initialization, with $\eta < \frac{8}{|\epsilon_1|+|\epsilon_2|}$

Figure 1: Simulation results for Theorem 1 (Part II). The row in Table 3 to which each case corresponds is marked in the sub-captions. Corresponding simulations for Parts I, III and IV are in the appendices due to space limitations.

unique pure NE. The following result, which is proved in Appendix D, shows that in this case the exponential weights dynamic converges to the corresponding pure NE at an exponentially fast rate.

Theorem 1 (Part I) *Suppose ϵ_1, ϵ_2 are non-zero and have the same sign. Then, Algorithm 1 converges to a) the pure NE (θ_2, θ_2) in the case where $\epsilon_1, \epsilon_2 < 0$, and b) the pure NE (θ_1, θ_1) in the case where $\epsilon_1, \epsilon_2 > 0$, at an exponential rate for any step size $\eta > 0$ (see Theorem 4 in Appendix D for the explicit constant in the rate).*

Next, we consider the more interesting and non-trivial cases where ϵ_1 and ϵ_2 are non-zero and have opposite signs. We begin with the case where $\epsilon_1 < 0$ and $\epsilon_2 > 0$. Recall from Table 2 that this case admits two pure *asymmetric* NEs (θ_1, θ_2) and (θ_2, θ_1), and there is also a strictly mixed NE given by $(p_{\text{SE}}, p_{\text{SE}})$, where $p_{\text{SE}} = \left(\frac{|\epsilon_2|}{|\epsilon_1|+|\epsilon_2|}, \frac{|\epsilon_1|}{|\epsilon_1|+|\epsilon_2|}\right)$. The following result, which is proved in Appendix E, shows a rich variety of convergence behaviors of the exponential weights dynamic depending on the initialization. We also illustrate the results through simulation in Figure 1.

Theorem 1 (Part II) *Suppose $\epsilon_1 < 0$, and $\epsilon_2 > 0$. Then:*

- *(Opposite-sign initialization) If $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have different signs, then Algorithm 1 converges to one of the pure NEs (θ_1, θ_2) or (θ_2, θ_1) at an exponential rate (see Theorem 5 in Appendix E for the explicit constant in the rate);*
- *(Same-sign but not identical initialization) If $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have the same sign and $\Delta_1^{(1)} \neq \Delta_2^{(1)}$, then Algorithm 1 converges asymptotically to one of the pure NEs (θ_1, θ_2) or (θ_2, θ_1);*
- *(Identical initialization) If $\Delta_1^{(1)} = \Delta_2^{(1)}$, and $\eta \in \left(0, \frac{8}{|\epsilon_1|+|\epsilon_2|}\right)$, then Algorithm 1 converges to the strictly mixed NE $(\frac{\epsilon_2}{|\epsilon_1|+\epsilon_2}, \frac{|\epsilon_1|}{|\epsilon_1|+\epsilon_2})$ asymptotically.*

Interestingly, Theorem 1 (Part II) shows that the algorithm always converges to some pure NE when the initialization functionals $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ differ, but requires a sufficiently small step size $\eta \in \left(0, \frac{8}{|\epsilon_1|+|\epsilon_2|}\right)$ if the initialization is identical. Note that this requirement on the step size is, however, relatively mild: if we normalized the utilities to be in $[-1, 1]$, we have $|\epsilon_1|, |\epsilon_2| \leq 2$, and so the condition reduces to $\eta < 2$. For $\eta > \frac{8}{|\epsilon_1|+|\epsilon_2|}$, we also provide a simple example where the algorithm oscillates between two pairs of mixed strategies that are not mixed NE for such prohibitively large values of η. Moreover, in the special case of (2×2) congestion games, Palaiopanos et al. (2017, Section 4) provides counterexamples for specific large values of η, which we show exceed the threshold ($\eta > \frac{8}{|\epsilon_1|+|\epsilon_2|}$). Finally, Chotibut et al. (2021) establishes last-iterate convergence of the multiplicative weights algorithm for 2×2 congestion games, which constitute a special case of this class of games with $\epsilon_1 = -\epsilon_2$. Therefore, our results recover their conclusion for congestion games.

Next, we consider the case where $\epsilon_1 > 0$ and $\epsilon_2 < 0$. Recall that in this case the pure NEs are (θ_1, θ_1) and (θ_2, θ_2). There also exists a strictly mixed NE in the form of $(p_{\text{SE}}, p_{\text{SE}})$. The following result, which is proved in Appendix F, again shows a rich variety of convergence behaviors.

Theorem 1 (Part III) *Suppose $\epsilon_1 > 0$, and $\epsilon_2 < 0$. Then:*

- *(Same Sign initialization) If $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have the same sign, then Algorithm 1 converges to one of the pure NEs at an exponential rate (see Theorem 9 in Appendix F for the explicit constant in the rate);*
- *(Opposite Sign initialization) If $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have different signs, and $\eta \in \left(0, \frac{8}{|\epsilon_1|+|\epsilon_2|}\right)$, then Algorithm 1 converges to either a pure NE or strictly mixed-NE asymptotically.*

The above conclusion is similar, but not identical to its counterpart in Theorem 1 (Part II). On one hand, when $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have the same sign, the algorithm converges to a pure NE. On the other hand, when $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have different signs, the algorithm can converge either to the strictly mixed NE or pure NEs, but requires a sufficiently small step size (analogous to the case of identical initialization in Part II). Interestingly, the proofs of Theorems 1 (Part III) and 1 (Part II), proceed differently, despite the seeming symmetry in the cases $(\epsilon_1 < 0, \epsilon_2 > 0)$ and $(\epsilon_1 > 0, \epsilon_2 < 0)$.

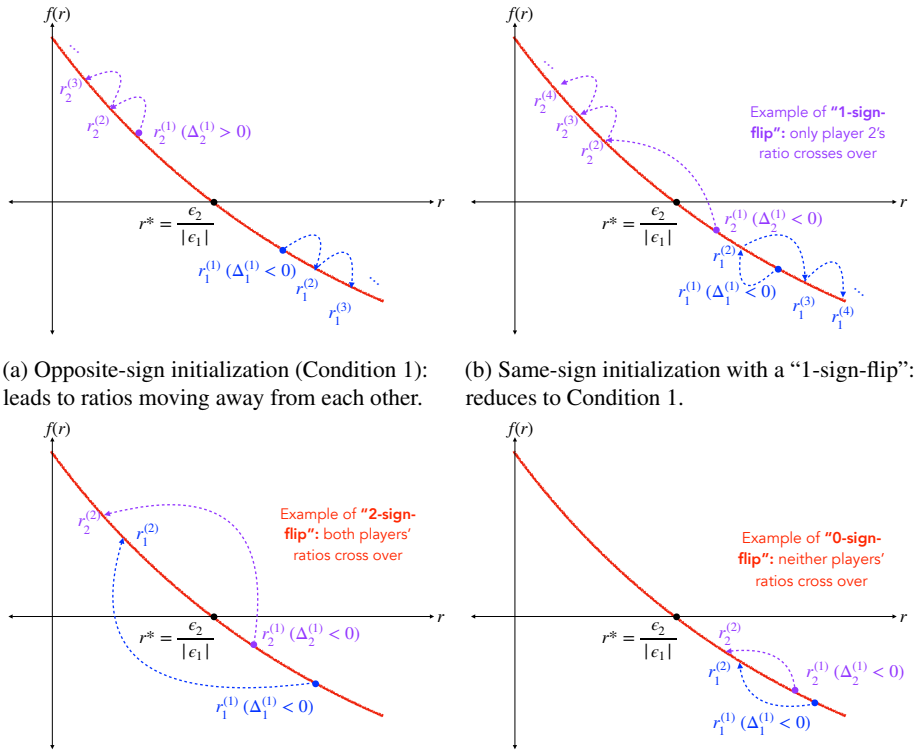
Finally, we consider the case where one of ϵ_1 and ϵ_2 is equal to zero. This may seem like a corner case (and is indeed a set of measure zero in the set of all 2×2 symmetric games), but we can handle it in our framework. Without loss of generality, we assume $\epsilon_1 = 0$. (Note that the other case ($\epsilon_2 = 0$) is purely symmetric to this case, by just switching θ_1 and θ_2.) Recall that in this case, the pure NEs are (θ_1, θ_1), (θ_1, θ_2) and (θ_2, θ_1), while the mixed NEs are (θ_1, p) and (p, θ_1) for any $p \in \Delta_2$. The following result, which is proved in Appendix G, shows convergence to different classes of NEs depending on the initialization (and does not possess any requirements on the step size).

Theorem 1 (Part IV) *Suppose $\epsilon_1 = 0$. Then:*

- *If $\epsilon_2 < 0$, then Algorithm 1 converges to the pure NE (θ_2, θ_2) at an exponential rate;*
- *If $\epsilon_2 > 0$ and $\Delta_1^{(1)} = \Delta_2^{(1)}$, then Algorithm 1 converges to the pure NE (θ_1, θ_1) asymptotically;*
- *If $\epsilon_2 > 0$ and $\Delta_1^{(1)} \neq \Delta_2^{(1)}$, then Algorithm 1 converges to a mixed NE. Moreover, there exist examples where the algorithm converges to a strictly mixed NE of the form $(\theta_1, [A, 1-A])$, where $A \in (0, 1)$ is some constant.*

Reconciliation of Theorem 1 with results showing oscillation/chaos: It is worth briefly clarifying, given the overwhelmingly negative results for exponential weights dynamics in games, why we are able to obtain convergence results of such a sweeping nature for our setting. The zero-sum games that induce oscillations are ones for which the *only* NE is strictly mixed (Bailey and Piliouras, 2018; Cheung, 2018) (the Matching Pennies game is a common example). As Table 2 shows, this never happens in a symmetric 2×2 game. Moreover, Cheung and Piliouras (2019) provide a sufficient and necessary condition for increasing volume of the flow of exponential weights (and thereby chaotic behavior) which, in the case of 2×2 symmetric games, can be verified to *never* occur (see (Cheung and Piliouras, 2019, Appendix G) for the detailed condition).

3.2. Proof sketch of Theorem 1



(a) Opposite-sign initialization (Condition 1): leads to ratios moving away from each other.

(b) Same-sign initialization with a "1-sign-flip": reduces to Condition 1.

(c) Same-sign initialization with a "2-sign-flip": happens finitely often.

(d) Same-sign initialization with a "0-sign-flip": happens finitely often.

Figure 2: Illustrations of how the signs of ratio/initialization parameters $\Delta_i^{(1)}$, and their potential "flips", affect analysis in the proof of Theorem 1 (Part II). r^* is the unique root of $f(\cdot)$.

In this section, we provide some brief intuition for the rich variety of behavior that we observe across initialization and game environment. An initial observation — that simplifies the proof for all the cases — is that we can parameterize any mixed strategy in the 2-dimensional simplex through a single positive real number, i.e. the ratio $r_i := p_{i,1}/p_{i,2}$ for player i. Since $p_{i,1} + p_{i,2} = 1$, r_i uniquely identifies a mixed strategy $(p_{i,1}, p_{i,2})$ for player i. It is then easy to see that, for a given player $i \in \{1, 2\}$ and the other player $j \in \{1, 2\}$ (where $j \neq i$), the update for player i can be written

as a function of solely (r_i, r_j), as

$$r_i^{(t+1)} = r_i^{(t)} \cdot \exp(\eta f(r_j^{(t)})),$$

where $f(r_j^{(t)}) = \frac{\epsilon_1 r_j^{(t)} + \epsilon_2}{1 + r_j^{(t)}} = \Delta_j^{(t)}$. An initial question one might ask at this stage is what the *fixed points* of the above mapping would be, i.e. when $\bar{r} = \bar{r} \cdot \exp(\eta f(\bar{r}))$. It is easy to see that fixed points correspond either to $\bar{r} = 0$ (i.e. the pure strategy θ_1), $\bar{r} = \infty$ (i.e. the pure strategy θ_2) or, in the cases where the the signs of ϵ_1 and ϵ_2 do not match (Parts II and III) $\bar{r} = r^* = -\frac{\epsilon_2}{\epsilon_1}$ (i.e. the mixed strategy p_{SE}, which corresponds to the unique root of $f(r)$). The name of the game is then to show that each of $\{r_i\}_{i=1}^2$ converges to one of the above fixed points, which in turn corresponds to either a pure/strictly mixed NE in each of the relevant cases.

Proof sketch of Theorem 1 (Part I) The cases where ϵ_1 and ϵ_2 have the same sign (Part I) are easy. To see this, consider the case where $\epsilon_1, \epsilon_2 > 0$. In this case, $f(\cdot)$ is positive, leading to monotonically increasing updates $\{r_i^{(t)}\}_{t \geq 1}$ in t for each $i \in \{1, 2\}$, that will converge exponentially fast (at a rate depending on the initial value $f(r_i^{(1)})$) to the pure NE (θ_1, θ_1). (Similarly, if $\epsilon_1, \epsilon_2 < 0$, $f(\cdot)$ is instead negative, leading to exponential convergence to the pure NE (θ_2, θ_2).)

Proof sketch of Theorem 1 (Part II): The analysis of Parts II-IV is far trickier and heavily depends on the initialization of the players' strategies, controlled by $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$. Due to space limitations we illustrate our analysis for the proof of Theorem 1 (Part II), noting that Parts III and IV utilize similar (but not identical) ideas. A key initial observation is that $f(r)$ in Part II is monotonically *decreasing*, meaning that the sign of $\text{sign}(\Delta_i^{(t)}) = -\text{sign}(r_i^{(t)} - r^*)$ for all $t \geq 1$ (see Figure 2 for an illustration). The analysis is then divided into three conditions:

1. Here as well, there is an "easy" initialization in which $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have opposite signs — this type of initialization turns out to lead to similar monotonic behavior as in Part I. To see this, consider the case where $\Delta_1^{(1)} < 0$ and $\Delta_2^{(1)} > 0$ (Condition 1 in Appendix E). Because $r_1^{(t+1)} = r_1^{(t)} \exp(\eta \Delta_2^{(t)})$, the update leads to the ratio r_1 increasing which further *decreases* $\Delta_1^{(t)}$. Similarly, the update leads to the ratio r_2 *decreasing* which further *increases* $\Delta_2^{(t)}$. As illustrated in Figure 2a, in this case the parameters $\Delta_i^{(t)}$ move further and further away from each other, eventually converging to the asymmetric pure NE (θ_1, θ_2).
2. The more difficult case is when the signs of $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ *match* at initialization. In this case the above monotonicity argument no longer directly applies, due to the fact that the sign of $r_i^{(1)} - r^*$ for each player $i \in \{1, 2\}$ may change unpredictably over time. Clearly, if one of the players flips sign (what we call a "1-sign-flip"), the analysis reduces to the easy initialization condition above; see Figure 2b for an illustration. Remarkably, we are able to show that the other two bad events of a "0-sign-flip", meaning the signs stay the same from round $t \to t+1$, and "2-sign-flip", meaning the signs both change from round $t \to t+1$, can only happen *finitely many times*. (See Figures 2c and 2d for an illustration of these bad events). In fact, Lemma 7 shows via a potential-based argument that a "2-sign-flip" can happen at most a constant number of times (for an explicit constant that depends on the parameters of the game).
3. Interestingly, the above analysis critically uses the fact that $r_1^{(1)} \neq r_2^{(2)}$, which only occurs at unequal initialization. The case of identical initialization needs to be handled separately

and turns out to correspond to analyzing a 1-dimensional dynamical system on the variable $u^{(t)} := \ln\left(\frac{r^{(t)}}{r^*}\right)$ (where we dropped the index i as the ratios are identical). We show that the map $u \to T(u)$ is *contractive* for the specified threshold on η and appeal to Banach's fixed point theorem to show convergence to the unique fixed point, given by $\bar{u} = 0 \implies \bar{r} = r^*$, which corresponds to the unique strictly mixed NE.

4. Applications to multi-agent performative prediction

The recently proposed framework of Performative Prediction (Perdomo et al., 2020) studies interactions in which machine learning models are trained on data that is not i.i.d. In particular, it focuses on scenarios where, at each time step, the observed data distribution depends on the previously deployed model or decision. This dependence may arise because individuals strategically adapt their behavior in response to the decision rule, or because the deployed policy itself shapes the underlying environment from which new data is generated. Narang et al. (2023) extended Performative Prediction to multi-agent settings, where n different agents simultaneously deploy machine learning models and decision rules. They showed that simple dynamics (specifically, retraining models at each time step and gradient descent) converge in a last-iterate sense to a stable solution. However, these results i) do not cover our exponential weights dynamic, and ii) the stable solutions obtained may not coincide with equilibria of the actual ensuing game without additional assumptions such as strong monotonicity (which does not hold for finite normal-form games).

A special case of a lending game between two banks: In this section, we formulate a new model for Performative Prediction in a special but practically motivated case: that of a game between two banks, or learners, which need to decide on application thresholds and interest rates for lendees. Here, the banks aim to approve loan applications from customers. Each bank $i \in \{1, 2\}$ selects two parameters $(\tau_i, \gamma_i) := \theta_i$, where τ_i is the minimum credit score threshold for giving a loan to a customer and γ_i is the interest rate offered to approved customers.

Let $y_j \in [0, 1]$ be the normalized credit score of customer i, which is drawn from a distribution D_y. A customer with a credit score y_j is approved by Bank i if only if $y_j \geq \tau_i$. For simplicity, we assume that y_j directly represents the customer's probability of repaying the loan. That is, letting $Y_j \in \{0, 1\}$ be a random variable controlling whether the customer j repays their loans, with 0 meaning default and 1 repayment, we have $y_j = \mathbb{P}[Y_j = 1]$. A customer's (normalized) credit score can be interpreted as a noisy observation of y_j. In this case, if the customer chooses Bank i with parameter $\theta_i = (\tau_i, \gamma_i)$, the expected reward for this bank would be $(1 + \gamma_i) \cdot y_j - (1 - y_i)$: if the customer repays the loan, with probability y_j, the return is $1 + \gamma_i$; if not, the return is 0.

For given parameter choices $\theta_1 = (\tau_1, \gamma_1)$ and $\theta_2 = (\tau_2, \gamma_2)$, the allocation of customers and rewards for banks are determined by the following conditions:

1. **Exclusive allocation by threshold**: If $\tau_1 \leq y_j < \tau_2$, the customer goes to Bank 1, as the score qualifies for Bank 1 but not Bank 2. Conversely, if $\tau_2 \leq y < \tau_1$, the customer chooses Bank 2.

2. **Thresholds met by both banks**: If $\tau_1, \tau_2 \leq y_j$ and $\gamma_1 < \gamma_2$, the customer selects Bank 1 due to the lower interest rate. Conversely, if $\gamma_1 > \gamma_2$, the customer chooses Bank 2. If there is a tie and $\gamma_1 = \gamma_2$, the customer goes to each bank with probability 0.5.

3. **Rejection by both banks**: If $y_j < \tau_1$ and $y_j < \tau_2$, the customer is rejected by both banks.

Bank 1 / Bank 2	$\theta_1 = (\tau_\ell, \gamma_h)$	$\theta_2 = (\tau_h, \gamma_\ell)$
$\theta_1 = (\tau_\ell, \gamma_h)$	$a = \frac{1}{2}h(\gamma_h, \gamma_\ell, 1)$	$b = h(\gamma_\ell, \tau_h, 1)$
$\theta_2 = (\tau_h, \gamma_\ell)$	$c = h(\gamma_h, \tau_\ell, \tau_h)$	$d = \frac{1}{2}h(\gamma_\ell, \tau_h, 1)$

Table 4: Utility matrix of Bank 1, which exactly corresponds to the utility matrix of Player 1 in Table 1. Similar to Table 1, the utility matrix of Bank 2 is symmetric to that of Bank 1.

In particular, when considering two banks that learn their strategies over time, this is an exact example of performative prediction: the consumer data that each bank sees is a function of both of their deployed decision rules, since consumers strategically select which bank they get a loan from.

A 2x2 bank game: We focus on a simple case where there are two choices for each parameter: $0 \leq \tau_\ell < \tau_h \leq 1$ and $0 \leq \gamma_\ell < \gamma_h \leq 1$. We consider two possible actions per player: either (τ_l, γ_h), or (τ_h, γ_l); the rationale is that a bank with a lower threshold, i.e. which allows customers with worse credit, sets a higher rate with said customers[4]. Let

$$h_{D_y}(\gamma, \tau_a, \tau_b) = \int_{\tau_a}^{\tau_b} [(2+\gamma)y - 1]p(y)dy. \tag{2}$$

It is not hard to see that a) this is a symmetric 2×2 game that fits into our framework, and moreover, b) the payoff matrix of player 1 is then given by Table 4, where the values of a, b, c and d are explicitly defined. Therefore, Theorem 1 directly applies to this Bank Game and implies convergence to either pure symmetric NE, pure asymmetric NE or strictly mixed NE depending on the parameters of the Bank Game (i.e. the values of $\gamma_\ell, \gamma_h, \tau_\ell$ and τ_h, as well as the distribution on customers D_y). Our experiments section (Appendix H) provides simulations showing convergence under different values of ϵ_1 and ϵ_2. Overall, Theorem 1 implies rich consequences for lenders and lendees in this application: symmetric NE correspond to homogeneous bank behavior, but asymmetric pure NE correspond to heterogeneous behavior (e.g. one bank picking a high interest rate and low threshold, while the other bank does the opposite). Our results show that either type of behavior can be realized depending on the parameters of the game and initialization conditions.

5. Related work

We organize our discussion of related work under three verticals below.

Black-box results for "no-regret" dynamics: The design of algorithms that achieve vanishing external regret dates back to (Hannan, 1957; Blackwell, 1956), as well as the interpretation of such algorithms as solving repeated zero-sum games. A modern non-asymptotic result by Freund and Schapire (1999) shows that the time-average of the payoff of two agents playing *any* no-external-regret algorithms against each other will converge to the (unique) value of the zero-sum game. These results do not extend to arbitrary general-sum games; in fact, "uncoupled" dynamics that only use first-order information can be shown to never converge to NE for certain games (Hart and Mas-Colell, 2003), and "higher-order" dynamics are instead needed (Toonsi and Shamma, 2023). Instead, a

4. It is not hard to show that this is without loss of generality, as actions (τ_l, γ_l) and (τ_h, γ_h) are dominated under reasonable choices of parameters. See Appendix H.1 for a formal proof of this fact.

folklore result (see, e.g. (Cesa-Bianchi and Lugosi, 2006)) shows that the time-average of the joint strategy profile of the agents will converge to the polytope of *coarse correlated equilibria* (CCE). Such a result still allows for oscillation within the CCE polytope[5] (except in the special case where it is a singleton set, meaning that the set of all CCE coincides with the set of all NE). When the players' algorithms achieve the stronger guarantee of vanishing *internal/swap regret* (Blum and Mansour, 2007; Greenwald and Jafari, 2003) (see also the "adaptive heuristics" defined in the economics literature (Hart, 2005), as well as the notion of calibrated learning (Foster and Vohra, 1997)), the time-averaged convergence is specialized to the polytope of Aumann's *correlated equilibria* (CE) (Aumann, 1987). This result also allows for oscillations within the CE polytope. Table 2 showed that both the CE polytope and the CCE polytope are non-trivial in general even for 2×2 symmetric games. Achieving convergence to a specific NE then necessitated a specialized study of a specific dynamic, which we chose to be exponential weights in this paper. It is also worth noting that we analyzed exponential weights with a *universally constant* step size, while the no-external-regret manifestation of exponential weights requires a step size that either uniformly decays with rounds t or depends inversely on the horizon of learning T (Cesa-Bianchi et al., 1997) — so, strictly speaking, our dynamic is not even no-external-regret.

Interestingly, the *last iterate* of no-swap-regret dynamics was recently shown to converge to a specific symmetric Nash equilibrium for *zero-sum symmetric games* when the players initialize identically (Leme et al., 2024). This result is complementary to our work, which studies a specific dynamic in a different setting (still symmetric, but 2×2 and often not zero-sum).

Oscillation and chaos in last-iterate of game dynamics: The *last iterate* of game dynamics has seen intense recent study, motivated both by natural desiderata of behavioral stability (Hofbauer and Sigmund, 1998) and by applications in machine learning such as the training process of generative adversarial networks (Daskalakis et al., 2018). Results on the last-iterate, particularly for the standard exponential weights dynamic, are predominantly negative, frequently showing oscillations or some sort of chaotic behavior. Even for zero-sum games, Bailey and Piliouras (2018); Cheung (2018) showed that multiplicative weights with a constant step size oscillates towards the boundaries of pure strategies when the game only has strictly mixed NEs. Cheung and Piliouras (2019) further showed the advent of chaotic behavior for all algorithms within the "Follow-the-Regularized-Leader" (FTRL) family (which includes exponential weights) for a broad class of games, including graphical constant-sum games, certain symmetric games and arbitrary 2×2 general-sum (not necessarily symmetric) games. Palaiopanos et al. (2017) also discovers a difference between the exponential weights dynamic and the "linear" version of the multiplicative weights update, with the former leading to chaotic behavior with large enough step sizes, in congestion games. The papers that provide a general framework for oscillations/chaos conduct a volume analysis showing that if the volume of the "flow" of the dynamics (from some initialization set of non-zero Lebesgue measure) increases, it implies chaotic behavior — but the converse does not hold. Negative results exist even for the weaker requirement of *local asymptotic stability* near any NE — in particular, any mixed NE is shown to be unstable under the family of FTRL dynamics both in continuous and discrete time (Vlatakis-Gkaragkounis et al., 2020; Giannou et al., 2021). As we discussed in Sections 3 and 4,

5. It *is* possible to achieve convergence to a specific equilibrium point within the CCE polytope for certain games if one player plays a no-regret algorithm and the other enacts a series of "threats", or "grim-trigger" strategies that penalize both players if they deviate from a specific equilibrium (see, e.g. (Collina et al., 2023)). Such dynamics are asymmetric and highly specialized (due to the threat-based nature of the strategies), and do not constitute the standard modus operandi in multi-agent learning.

our positive convergence results for 2×2 symmetric games can be readily reconciled with these results. We showed that the mixed NE is only approached through specific directions (involving identical initialization), which turns out to be the only stable direction of deviation. Moreover, many of the oscillatory situations (e.g. 2×2 congestion games (Palaiopanos et al., 2017) turned out to only occur under a sufficiently large step size.

Positive convergence results for specific dynamics: A plethora of last-iterate convergence results have been shown for zero-sum games (most relevant to our setting are the ones for finite games, i.e. min-max optimization constrained to the probability simplices (Daskalakis and Panageas, 2019; Cai et al., 2022)) under modifications of exponential weights to include optimism and/or extragradient method, as well as a sufficiently small or decaying step size. In contrast, we allow for non-zero-sum games but require symmetry and the game to be 2×2, and we consider the original exponential weights dynamic; moreover, in many cases we can handle an arbitrarily large step size. Positive *average-iterate* convergence results have also been developed for the fictitious-play dynamic (Brown, 1951; Robinson, 1951) (which corresponds to "Follow-the-Leader in online learning) for zero-sum games with a diagonal payoff matrix (Abernethy et al., 2021) as well as zero-sum symmetric games of a generalized Rock-Paper-Scissors variety (Lazarsfeld et al., 2025). It is an intriguing open question whether fictitious play dynamics would converge in our setting, whether in the average or last iterate. Specific to congestion games (with any finite number of players/pure strategies), Kleinberg et al. (2009) showed that *randomized* variants of exponential weights can converge *with high probability* for a sufficiently small step size. Chotibut et al. (2021) established the last-iterate convergence of the multiplicative weights algorithm for 2×2 congestion games. We recover their conclusion in this special setting as a consequence of Theorem 1 (Part II).

A recent yet relatively extensive line of work studies performative prediction (Perdomo et al. (2020); Miller et al. (2021); Brown et al. (2022); Bechavod et al. (2021); Mendler-Dünner et al. (2022); Jagadeesan et al. (2022); Zrnic et al. (2021); Shavit et al. (2020) to only name a few), which looks at single-agent machine learning settings where the training data on a given round depends on the previously deployed model(s)—capturing situations where deployed models induce strategic behavior or distribution shifts. The case of performative prediction in *multi-agent settings*, however, has received significantly less attention. In particular, Narang et al. (2023) study convergence of learning dynamics in specially structured games induced by these machine-learning-oriented problems. In these works, multiple players iteratively train their own machine learning models, and each player's data distribution—and hence their utility—depends on the set of models deployed on a given round. This literature shows that both a form of *repeated empirical risk minimization* as well as gradient descent (when used by all players in the game) can converge in a last-iterate sense to a stable outcome. However, two limitations remain: (i) said dynamics can converge to a "sub-optimal" point, which is not an equilibrium of the underlying game between learners, and (ii) convergence guarantees rely on strong assumptions both on the players' loss functions — especially strong monotonicity, which does not hold for normal-form games, and on the sensitivity of outcomes to players' actions.

The works (Vlatakis-Gkaragkounis et al., 2020; Giannou et al., 2021) show that for any general-sum game, a particular NE is locally asymptotically stable under FTRL dynamics if and only if it is pure. Note that these results only imply *local convergence* of the dynamic, i.e. if it is initialized sufficiently close to a pure Nash Equilibrium. On the other hand, we utilize the special structure of 2×2 symmetric games to provide a *global convergence* result, i.e. for a broader set of initializations that could be arbitrarily far from any NE.

6. Discussion

Overall, our results paint a more positive than expected picture for the exponential weights dynamic in the special case of 2×2 symmetric games. The most pressing open question is whether this story extends in any way to symmetric games with more than 2 actions. Preliminary simulations suggest convergence behavior in many (but not all) cases, but characterizing the ensuing dynamical systems is an important future direction. Even in the symmetric 2×2 case, interesting open questions remain — such as whether we can prove fast non-asymptotic rates of convergence for all the cases in Theorem 1 (not just the easy ones), and whether we can handle unequal step sizes for different players.

Acknowledgments

GW was supported by the Apple Scholars in AI/ML PhD fellowship by Apple and ARC-ACO fellowship provided by Georgia Tech. VM was supported by the NSF (through award CCF-2239151 and award IIS-2212182), an Adobe Data Science Research Award and an Amazon Research Award. KA and JZ were supported by the NSF (through award IIS-2504990 and award IIS-2336236).

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.

Robert J Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica: Journal of the Econometric Society*, pages 1–18, 1987.

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.

Yahav Bechavod, Katrina Ligett, Steven Wu, and Juba Ziani. Gaming helps! Learning from strategic interactions in natural dynamics. In *International Conference on Artificial Intelligence and Statistics*, pages 1234–1242. PMLR, 2021.

David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(6), 2007.

Gavin Brown, Shlomi Hod, and Iden Kalemaj. Performative prediction in a stateful world. In *International conference on artificial intelligence and statistics*, pages 6045–6061. PMLR, 2022.

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

Yang Cai, Argyris Oikonomou, and Weiqiang Zheng. Finite-time last-iterate convergence for learning in multi-player games. *Advances in Neural Information Processing Systems*, 35:33904–33919, 2022.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

Nicolo Cesa-Bianchi, Yoav Freund, David Haussler, David P Helmbold, Robert E Schapire, and Manfred K Warmuth. How to use expert advice. *Journal of the ACM (JACM)*, 44(3):427–485, 1997.

Yun Kuen Cheung. Multiplicative weights updates with constant step-size in graphical constant-sum games. *Advances in Neural Information Processing Systems*, 31, 2018.

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*, 2021.

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

Natalie Collina, Eshwar Ram Arunachaleswaran, and Michael Kearns. Efficient Stackelberg Strategies for Finitely Repeated Games. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, pages 643–651, 2023.

Constantinos Daskalakis and Ioannis Panageas. Last-Iterate Convergence: Zero-Sum Games and Constrained Min-Max Optimization. *10th Innovations in Theoretical Computer Science*, 2019.

Constantinos Daskalakis, Andrew Ilyas, Vasilis Syrgkanis, and Haoyang Zeng. Training GANs with Optimism. In *International Conference on Learning Representations*, 2018.

Dean P Foster and Rakesh Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 29(1-2):7–35, 1999.

Dean P Foster and Rakesh V Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(589):40–55, 1997.

Michael J Fox and Jeff S Shamma. Population games, stable games, and passivity. *Games*, 4(4): 561–583, 2013.

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

Angeliki Giannou, Emmanouil Vasileios Vlatakis-Gkaragkounis, and Panayotis Mertikopoulos. Survival of the strictest: Stable and unstable equilibria under regularized learning with partial information. In *Conference on Learning Theory*, pages 2147–2148. PMLR, 2021.

Amy Greenwald and Amir Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003. Proceedings*, pages 2–12. Springer, 2003.

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

Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. Strategic classification. In *Proceedings of the 2016 ACM conference on innovations in theoretical computer science*, pages 111–122, 2016.

Sergiu Hart. Adaptive heuristics. *Econometrica*, 73(5):1401–1430, 2005.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

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

Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

Meena Jagadeesan, Tijana Zrnic, and Celestine Mendler-Dünner. Regret minimization with performative feedback. In *International Conference on Machine Learning*, pages 9760–9785. PMLR, 2022.

Anna R Karlin and Yuval Peres. *Game theory, alive*, volume 101. American Mathematical Soc., 2017.

Robert Kleinberg, Georgios Piliouras, and Éva Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 533–542, 2009.

John Lazarsfeld, Georgios Piliouras, Ryann Sim, and Andre Wibisono. Fast and furious symmetric learning in zero-sum games: Gradient descent as fictitious play. *arXiv preprint arXiv:2506.13086*, 2025.

Renato Leme, Georgios Piliouras, and Jon Schneider. Convergence of no-swap-regret dynamics in self-play. *Advances in Neural Information Processing Systems*, 37:76210–76231, 2024.

Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations*, 2018.

Celestine Mendler-Dünner, Frances Ding, and Yixin Wang. Anticipating performativity by predicting from predictions. *Advances in Neural Information Processing Systems*, 35:31171–31185, 2022.

John P Miller, Juan C Perdomo, and Tijana Zrnic. Outside the echo chamber: Optimizing the performative risk. In *International Conference on Machine Learning*, pages 7710–7720. PMLR, 2021.

Roger B Myerson. *Game theory*. Harvard University Press, 2013.

Adhyyan Narang, Evan Faulkner, Dmitriy Drusvyatskiy, Maryam Fazel, and Lillian J Ratliff. Multiplayer performative prediction: Learning in decision-dependent games. *Journal of Machine Learning Research*, 24(202):1–56, 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.

Juan Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *International Conference on Machine Learning*, pages 7599–7609. PMLR, 2020.

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

Robert W Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

Yonadav Shavit, Benjamin Edelman, and Brian Axelrod. Causal strategic linear regression. In *International Conference on Machine Learning*, pages 8676–8686. PMLR, 2020.

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

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

Kaiqing Zhang, Zhuoran Yang, and Tamer Basar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of Reinforcement Learning and Control*, pages 321–384, 2021.

Tijana Zrnic, Eric Mazumdar, Shankar Sastry, and Michael Jordan. Who leads and who follows in strategic classification? *Advances in Neural Information Processing Systems*, 34:15257–15269, 2021.

Appendix A. Notation and Preliminaries for proofs

In this section, we introduce some notation that will help simplify the exposition of the proof of Theorem 1. In all that follows, recall that we defined $\epsilon_1 := a - b$ and $\epsilon_2 := c - d$. Specifically, ϵ_1 denotes the change in utility of player 1 when their strategy changes from θ_1 to θ_2 when player 2 plays θ_1, and ϵ_2 denotes the change in utility of player 1 when their strategy changes from θ_1 to θ_2 when player 2 plays θ_2. Also, for each player $i \in \{1, 2\}$, recall that we defined the functional

$$\Delta_i^{(t)} = p_{i,1}^{(t)}\epsilon_1 + p_{i,2}^{(t)}\epsilon_2. \tag{3}$$

We now introduce our new notation below:

- For each player $i \in \{1, 2\}$, let $r_i^{(t)} = \frac{p_{i,1}^{(t+1)}}{p_{i,2}^{(t+1)}}$ be the ratio of probabilities allocated to pure strategy θ_1 over pure strategy θ_2. Then, for any pair of players $i \neq j \in \{1, 2\}$, we have

$$r_i^{(t+1)} = \frac{p_{i,1}^{(t)}}{p_{i,2}^{(t)}} \exp\left(\eta(p_{j,1}^{(t)}a + p_{j,2}^{(t)}c - p_{j,1}^{(t)}b - p_{j,2}^{(t)}d)\right) = r_i^{(t)} \exp(\eta\Delta_j^{(t)}). \tag{4}$$

  Note that, since $p_{i,1}^{(t)} + p_{i,2}^{(t)} = 1$ and $p_{i,1}^{(t)} = p_{i,2}^{(t)} r_i^{(t)}$, we have $p_{i,2}^{(t)} = \frac{1}{1+r_i^{(t)}}$, and $p_{i,1}^{(t)} = \frac{r_i^{(t)}}{1+r_i^{(t)}}$. Clearly, the mapping from $r \to p$ is a bijection.

- Recall that we defined the function $f(r) := \frac{\epsilon_1 r + \epsilon_2}{1+r}$ in Section 3.2. The function $f(r)$ allows us to rewrite our update step in the following simplified form, solely in terms of r. Namely, for any pair of players $i \neq j \in \{1, 2\}$, we have

$$r_i^{(t+1)} = r_i^{(t)} \exp(\eta f(r_j^{(t)})), \tag{5}$$

  noting that $f(r_j^{(t)}) = \frac{\epsilon_1 r_j^{(t)} + \epsilon_2}{1+r_j^{(t)}} = p_{j,1}^{(t)}\epsilon_1 + p_{j,2}^{(t)}\epsilon_2 = \Delta_j^{(t)}$. Further, it is easy to verify that $f'(r) = \frac{\epsilon_1 - \epsilon_2}{(1+r)^2}$.

- Finally, let $r^* = \frac{|\epsilon_2|}{|\epsilon_1|}$. If $\epsilon_1 \cdot \epsilon_2 < 0$ (which is the case for Parts II and III of Theorem 1), then $f(r^*) = 0$.

Appendix B. Proof of Lemma 2

In this section, we prove Lemma 2. We first consider the first bullet point of Lemma 2, i.e. what happens when $\Delta_1^{(1)} = \Delta_2^{(1)} = 0$. From Equation (4), it is easy to see that the ratio $r_i^{(t)}$ of each player $i \in \{1, 2\}$ does not change for all $t \geq 1$. Since the map $r \to p$ is a bijection, the mixed strategies also do not change.

Next, we consider the second bullet point of Lemma 2, i.e. the case when $\Delta_1^{(1)} = 0$ but $\Delta_2^{(1)} \neq 0$. Since $\Delta_1^{(1)} = 0$, Equation (4) tells us that $r_2^{(2)} = r_2^{(1)}$, meaning that $\Delta_2^{(2)} = f(r_2^{(2)}) = f(r_2^{(1)}) = \Delta_2^{(1)} \neq 0$. Moreover, if $\Delta_1^{(1)} = 0$, the parameters of the game must satisfy $\epsilon_1 \neq \epsilon_2$. (This is because,

if we instead had $\epsilon_1 = \epsilon_2 = \epsilon$, then we would obtain $f(r) = \epsilon$ for all values of r, meaning that $0 = \Delta_1^{(1)} = f(r_1^{(1)}) = \epsilon$, leading to $\epsilon_1 = \epsilon_2 = 0$. This is a contradiction with the assumption in the lemma that $|\epsilon_1| + |\epsilon_2| > 0$.)

Because $\epsilon_1 \neq \epsilon_2$, the sign of $f'(r)$ is solely determined by the sign of $\epsilon_1 - \epsilon_2$ for all values of r, implying that the function f is strictly monotone. On the other hand, since $\Delta_2^{(1)} \neq 0$, Equation (4) again implies that $r_1^{(2)} \neq r_1^{(1)}$. Since f is strictly monotone and $\Delta_1^{(1)} = f(r_1^{(1)}) = 0$, we have $\Delta_1^{(2)} = f(r_1^{(2)}) \neq f(r_1^{(1)}) = 0$. This ultimately yields $\Delta_1^{(2)} \neq 0$, which completes the proof of the lemma.

Appendix C. Characterization of correlated equilibria (CEs)

In this section, we first review the definition of correlated equilibrium (CE), and then characterize the set of CEs for our game. Formally, we define $\nu = (\nu_{1,1}, \nu_{1,2}, \nu_{2,1}, \nu_{2,2}) \in \Delta_4$ as a joint probability distribution over pairs of strategies $\{(\theta_m, \theta_n) : (m, n) \in \{1, 2\} \times \{1, 2\}\}$.

Definition 3 (Correlated Equilibrium (CE)) *A joint probability distribution ν over pairs of strategies $\{(\theta_p, \theta_q) : (p, q) \in \{1, 2\} \times \{1, 2\}\}$ is a correlated equilibrium iff the following inequalities hold:*

$$\sum_{(m,n)\in\{1,2\}\times\{1,2\}} \nu_{m,n}\, u_1(\theta_m, \theta_n) \geq \sum_{(m,n)\in\{1,2\}\times\{1,2\}} \nu_{m,n} u_1(\theta_{m'}, \theta_n), \;\; \forall m' \in \{1, 2\},$$

$$\sum_{(m,n)\in\{1,2\}\times\{1,2\}} \nu_{m,n}\, u_2(\theta_m, \theta_n) \geq \sum_{(m,n)\in\{1,2\}\times\{1,2\}} \nu_{m,n}\, u_2(\theta_m, \theta_{n'}), \;\; \forall n' \in \{1, 2\}.$$

In general, the set of CEs is a polytope (in our case, a subset of Δ_4) and is bigger than the set of NEs, which may be disjoint and non-convex. We summarized the landscape of CEs along with the landscape of NEs in Table 2; we provide the calculation of the CEs here for completeness.

Based on Definition 3, the condition for a strategy profile to be a CE can be written as:

$$\nu_{1,1}\epsilon_1 \geq \nu_{1,2}(-\epsilon_2), \;\; \nu_{2,1}(-\epsilon_1) \geq \nu_{2,2}\epsilon_2, \;\; \nu_{1,1}\epsilon_1 \geq \nu_{2,1}(-\epsilon_2), \;\; \nu_{1,2}(-\epsilon_1) \geq \nu_{2,2}\epsilon_2.$$

When $\epsilon_1 < 0$ and $\epsilon_2 > 0$, this condition reduces to

$$\max\left\{\frac{|\epsilon_1|}{\epsilon_2}\nu_{1,1}, \frac{\epsilon_2}{|\epsilon_1|}\nu_{2,2}\right\} \leq \min\{\nu_{1,2}, \nu_{2,1}\}. \tag{6}$$

Similarly, when $\epsilon_1 > 0$ and $\epsilon_2 < 0$, the condition reduces to

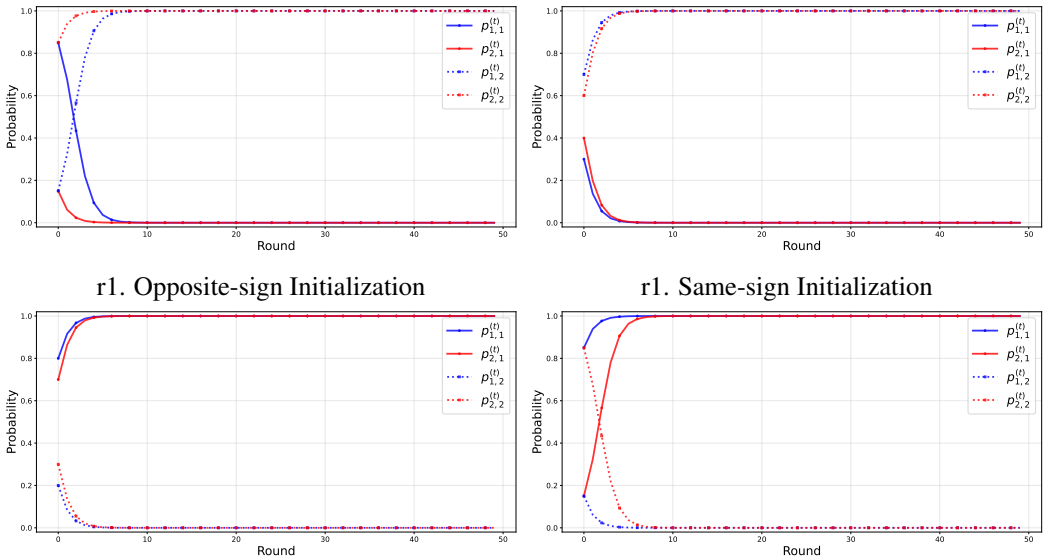
$$\max\left\{\frac{|\epsilon_1|}{\epsilon_2}\nu_{2,1}, \frac{\epsilon_2}{|\epsilon_1|}\nu_{1,2}\right\} \leq \min\{\nu_{1,1}, \nu_{2,2}\}. \tag{7}$$

When $\epsilon_1 = 0$ and $\epsilon_2 < 0$, we have

$$\nu_{1,2} = \nu_{2,1} = 0. \tag{8}$$

When $\epsilon_1 = 0$, $\epsilon_2 > 0$, we have

$$\nu_{2,2} = 0. \tag{9}$$



r1. Opposite-sign Initialization

r1. Same-sign Initialization

r2. Same-sign Initialization

r2. Opposite-sign initialization

Figure 3: Simulation results for Theorem 1 (Part I). The row in Table 3 to which each case corresponds is marked in the sub-captions.

We note that the case with $\epsilon_2 = 0$ and $\epsilon_1 \neq 0$ is exactly symmetric to the case with $\epsilon_1 = 0$ and $\epsilon_2 \neq 0$, simply by switching θ_1 and θ_2. For simplicity, we omit this case.

Observe that in all of the instances, excluding the ones where $\text{sign}(\epsilon_1) = \text{sign}(\epsilon_2)$, the set of CEs forms a non-trivial polytope, and there exist infinitely many points that are CE but not mixed or pure NEs. This means that even average-iterate convergence to a specific NE would *not* follow by directly plugging in regret bounds, let alone last-iterate convergence.

Appendix D. Proof of Theorem 1 (Part I)

In this section, we prove Theorem 1 (Part I). Without loss of generality, we consider the case where $\epsilon_1, \epsilon_2 < 0$. (Note that the other case is symmetric: it suffices to switch $\theta^{(1)}$ and $\theta^{(2)}$ to obtain the result.) Moreover, since $\epsilon_1, \epsilon_2 < 0$, Equation (3), tells us that for each player $i \in \{1, 2\}$, we have $\Delta_i^{(t)} < 0$ for all t. This implies that the sequence $\{r_i^{(t)}\}_{t=1}^{\infty}$ is strictly decreasing. We first provide a formal convergence result for this case.

Theorem 4 *Suppose Assumption 1 holds, and the game satisfies $\epsilon_1, \epsilon_2 < 0$. Then, Algorithm 1 with any step size $\eta > 0$ satisfies the following:*

- *If $|\epsilon_1| < |\epsilon_2|$, then for each player pair $i \neq j \in \{1, 2\}$, we have $p_{i,1}^{(t+1)} \leq r_i^{(1)} \exp(\eta t \Delta_j^{(1)})$, where $\Delta_j^{(1)} < 0$. This implies that the algorithm converges to the symmetric pure NE (θ_2, θ_2) exponentially fast;*

- *If $|\epsilon_1| \geq |\epsilon_2|$, then for each player pair $i \neq j \in \{1, 2\}$, we have $p_{i,1}^{(t+1)} \leq r_i^{(1)} \exp(\eta t \epsilon_2)$, where $\epsilon_2 < 0$. This implies that the algorithm converges to the symmetric pure NE (θ_2, θ_2) exponentially fast.*

We now provide the proof of Theorem 4, considering each condition listed above.

Condition 1: $|\epsilon_2| > |\epsilon_1|$**.** Under this condition, we have $f'(r) > 0$, meaning that $f(\cdot)$ is strictly increasing. Since $\{r_i^{(t)}\}_{t=1}^{\infty}$ is decreasing, we know $r_i^{(t)} < r_i^{(1)}$, so $f(r_i^{(t)}) < f(r_i^{(1)})$, and as a consequence, we have

$$r_i^{(t+1)} = r_i^{(1)} \exp\left(\eta \sum_{k=1}^{t} f\left(r_j^{(k)}\right)\right) \leq r_i^{(1)} \exp(\eta t \Delta_j^{(1)}).$$

Therefore, we have $p_{i,1}^{(t+1)} \leq p_{i,2}^{(t+1)} r_i^{(1)} \exp(\eta t \Delta_j^{(1)}) \leq r_i^{(1)} \exp(\eta t \Delta_j^{(1)})$. Note that $\Delta_j^{(1)} < 0$ is a strictly negative constant that depends on the algorithm's initialization. In conclusion, $p_{i,1}^{(t+1)} \leq r_i^{(1)} \exp(\eta t \Delta_j^{(1)})$, while $p_{i,2}^{(t+1)} \geq 1 - r_i^{(1)} \exp(\eta t \Delta_j^{(1)})$. Therefore, the algorithm converges to the symmetric pure NE (θ_2, θ_2).

Condition II: $|\epsilon_2| \leq |\epsilon_1|$**.** In this case, $f'(r) \leq 0$, so f is non-increasing. Since $r \geq 0$ (i.e. the ratio is always non-negative), we have $f(r) \leq f(0) = \epsilon_2 < 0$. So

$$r_i^{(t+1)} = r_i^{(1)} \exp\left(\eta \sum_{k=1}^{t} f(r_j^{(k)})\right) \leq r_i^{(1)} \exp(\eta t \epsilon_2),$$

which gives us $p_{i,1}^{(t+1)} \leq p_{i,2}^{(t+1)} r_i^{(1)} \exp(\eta t \epsilon_2) \leq r_i^{(1)} \exp(\eta t \epsilon_2)$. Note that $\epsilon_2 < 0$ is a strictly negative constant. In conclusion, we have $p_{i,1}^{(t+1)} \leq r_i^{(1)} \exp(\eta t \epsilon_2)$, while $p_{i,2}^{(t+1)} \geq 1 - r_i^{(1)} \exp(\eta t \epsilon_2)$. Therefore, the algorithm again converges to the symmetric pure NE (θ_2, θ_2).

Appendix E. Proof of Theorem 1 (Part II)

In this section, we prove Theorem 1 (Part II). In this case, we have $\epsilon_1 < 0$ and $\epsilon_2 > 0$. This implies that $f'(r) < 0$, i.e. f is monotonically decreasing. Moreover, we define $r^* = \frac{\epsilon_2}{|\epsilon_1|}$ to be the unique root of the function $f(\cdot)$. We consider different initialization conditions one by one.

E.1. Exponential Convergence With Opposite-Sign Initialization

We first provide a formal description of exponential convergence in the "easy" cases where the signs of the initialization functionals $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ are opposite.

Theorem 5 *Suppose Assumption 1 holds, and $\epsilon_1 < 0, \epsilon_2 > 0$. Then, Algorithm 1 with any step size $\eta > 0$ satisfies the following:*

- *If $\Delta_1^{(1)} < 0$, $\Delta_2^{(1)} > 0$, then we have $p_{1,2}^{(t)} \leq \frac{1}{r_1^{(1)}} \exp(-\eta t \Delta_2^{(1)})$, while $p_{2,1}^{(t)} \leq r_2^{(1)} \exp(\eta t \Delta_1^{(1)})$. This implies that the algorithm converges to the asymmetric pure NE (θ_1, θ_2) exponentially fast;*
- *If $\Delta_1^{(1)} > 0$,, $\Delta_2^{(1)} < 0$, then we have $p_{1,1}^{(t)} \leq r_1^{(1)} \exp(\eta t \Delta_2^{(1)})$, while $p_{2,2}^{(t)} \leq \frac{1}{r_2^{(1)}} \exp(-\eta t \Delta_1^{(1)})$. This implies that the algorithm converges to the asymmetric pure NE (θ_2, θ_1) exponentially fast.*

We now provide the proof of Theorem 5, considering each condition listed above.

Condition 1: $\Delta_1^{(1)} < 0$, $\Delta_2^{(1)} > 0$ (Basic Case 1). In this case, it can be easily shown by induction that the ratio of player 1, $\{r_1^{(t)}\}_{t=1}^{\infty}$, is monotonically increasing, while the ratio of player 2, $\{r_2^{(t)}\}_{t=1}^{\infty}$ is monotonically decreasing. To see this, first note that $f(r_1^{(1)}) = \Delta_1^{(1)} < 0$, and $f(r_2^{(1)}) = \Delta_2^{(1)} > 0$. Equation (5) then tells us that $r_1^{(2)} = r_1^{(1)} \exp(\eta f(r_2^{(1)})) > r_1^{(1)}$, while $r_2^{(2)} = r_2^{(1)} \exp(\eta f(r_1^{(1)})) < r_2^{(1)}$. Since f is decreasing, $f(r_1^{(2)})$ becomes smaller (i.e. more negative) while $f(r_2^{(2)})$ becomes larger (i.e. more positive). Therefore, we have

$$r_1^{(t+1)} = r_1^{(1)} \exp\left(\eta \sum_{k=1}^{t} f(r_2^{(k)})\right) \geq r_1^{(1)} \exp(\eta t f(r_2^{(1)})) = r_1^{(1)} \exp(\eta t \Delta_2^{(1)}),$$

and

$$r_2^{(t+1)} = r_2^{(1)} \exp\left(\eta \sum_{k=1}^{t} f(r_1^{(k)})\right) \leq r_2^{(1)} \exp(\eta t f(r_1^{(1)})) = r_2^{(1)} \exp(\eta t \Delta_1^{(1)}).$$

From the above, we get $p_{1,2}^{(t)} \leq \frac{1}{r_1^{(1)}} \exp(-\eta t \Delta_2^{(1)})$, while $p_{2,1}^{(t)} \leq r_2^{(1)} \exp(\eta t \Delta_1^{(1)})$. Noting that $\Delta_1^{(1)} < 0$ and $\Delta_2^{(1)} > 0$ are constants, this proves the desired exponential rate of convergence to the asymmetric pure NE (θ_1, θ_2).

Condition 2: $\Delta_1^{(1)} > 0$, $\Delta_2^{(1)} < 0$ (Basic Case 2). This case is exactly symmetric to Condition 1. Similar to Condition 1, we can show (using induction and based on the fact that f is decreasing) that the ratio of player 1, $r_1^{(t)}$, is decreasing while the ratio of player 2, $r_2^{(t)}$, is increasing. As a consequence, we have $p_{1,1}^{(t)} \leq r_1^{(1)} \exp(\eta t \Delta_2^{(1)})$, while $p_{2,2}^{(t)} \leq \frac{1}{r_2^{(1)}} \exp(-\eta t \Delta_1^{(1)})$. Noting again that $\Delta_1^{(1)} > 0$ and $\Delta_2^{(1)} < 0$ are again constants, this proves the desired exponential rate of convergence to the asymmetric pure NE (θ_2, θ_1).

This argument further implies that, if there exists a finite round t_0 at which $\Delta_1^{(t)}$ and $\Delta_2^{(t)}$ have different signs, the algorithm thereafter converges exponentially fast to the asymmetric pure NE.

E.2. Asymptotic Convergence in Same-Sign Initialization

Next, we consider the more challenging initialization in which $\Delta_1^{(1)}$, $\Delta_2^{(1)}$ have the same sign, and $\Delta_1^{(1)} \neq \Delta_2^{(1)}$.

Condition 3: $\Delta_1^{(1)}$, $\Delta_2^{(1)}$ have the same sign, and $\Delta_1^{(1)} \neq \Delta_2^{(1)}$. Note that since f is strictly monotone and $f(r) = \Delta$, the map $r \to \Delta = f(r)$ is one-to-one. Thus, $\Delta_1^{(1)} \neq \Delta_2^{(1)}$ implies that $r_1^{(1)} \neq r_2^{(1)}$. Assume without loss of generality that $r_i^{(1)} < r_j^{(1)}$, for a pair of players $i \neq j \in \{1, 2\}$. In this case, we will show that $\lim_{t\to\infty} r_j^{(t)} = \infty$ and $\lim_{t\to\infty} r_i^{(t)} = 0$, that is, player j converges to θ_1 while player i converges to θ_2.

Recall that $\epsilon_1 < 0$, $\epsilon_2 > 0$ implies that $f' < 0$, i.e. $f(\cdot)$ is monotonically decreasing. Moreover, since $f(r^*) = 0$, we have $\text{Sign}(r_i^{(t)} - r^*) = -\text{Sign}(\Delta_i^{(t)})$ for both i and j. To show convergence to the pure NE, it is sufficient to show that $r_i^{(t)} - r^*$ and $r_j^{(t)} - r^*$ have opposite signs after a constant, finite number of time steps, as the analysis then reduces to one of Basic Case 1 or Basic Case 2 above.

We first show the following fundamental result, which indicates that the relation $r_j^{(t)} > r_i^{(t)}$ is preserved for all t. This lemma will be critical for our analysis.

Lemma 6 *Assume that $r_j^{(1)} > r_i^{(1)}$. Then, we have $r_j^{(t)} > r_i^{(t)}$ for all $t \geq 1$. Moreover, we have $\ln \frac{r_j^{(t)}}{r_i^{(t)}} \geq \ln \frac{r_j^{(1)}}{r_i^{(1)}} > 0$.*

Proof We prove this lemma by induction. The relation holds for the base case ($t = 1$) by assumption, so we only need to prove the inductive step. Suppose that $r_j^{(t)} > r_i^{(t)}$. Then, we have

$$\ln r_j^{(t+1)} - \ln r_i^{(t+1)} = \ln \frac{r_j^{(t+1)}}{r_i^{(t+1)}} = \ln \frac{r_j^{(t)}}{r_i^{(t)}} + \eta[f(r_i^{(t)}) - f(r_j^{(t)})] \geq \ln r_j^{(t)} - \ln r_i^{(t)} > 0,$$

where the first inequality is based on the induction hypothesis, the positivity of η, and the fact that $f(\cdot)$ is monotonically decreasing. This completes the proof of the lemma. ■

The above lemma implies that $\lim_{t\to\infty} \frac{r_j^{(t)}}{r_i^{(t)}} = \infty$. However, it does not directly prove convergence of the algorithm, as it is possible for both the ratios $r_j^{(t)}$ and $r_i^{(t)}$ to oscillate, while *their* ratio diverges to infinity. We first rule out the following corner cases, where one of $r_j^{(t+1)} - r^*$ and $r_i^{(t+1)} - r^*$ is equal to zero (note that based on Lemma 6 they cannot be simultaneously equal to zero):

- $r_i^{(t+1)} = r^* < r_j^{(t+1)}$: In this case, we have $r_j^{(t+2)} = r_j^{(t+1)} > r^*$, while $r_i^{(t+2)} = r_i^{(t+1)} \exp(\eta f(r_j^{(t+1)})) < r_i^{(t+1)} = 0$. Thus, $r_j^{(t+2)} - r^*$ and $r_i^{(t+2)} - r^*$ have different signs at round $t + 2$.
- $r_i^{(t+1)} < r^* = r_j^{(t+1)}$: A similar argument yields $r_i^{(t+2)} < r^*$, while $r_j^{(t+2)} > r^*$. Again, $r_j^{(t+2)} - r^*$ and $r_i^{(t+2)} - r^*$ have different signs at round $t + 2$.

In summary, these corner cases, if they occur, reduce the problem to one of the Basic Cases above, after which the algorithm would converge to the pure NE.

We now provide the argument when such corner cases do not manifest. In other words, we consider $r_j^{(1)} - r^*$ and $r_i^{(1)} - r^*$ to both be non-zero and have the same sign. Suppose that $r_j^{(t)} - r^*$ and $r_i^{(t)} - r^*$ have the same sign at round t. Then from round t to round $t + 1$, the sign of $r_j^{(t)} - r^*$ and $r_i^{(t)} - r^*$ can change in only three ways (excluding the corner cases):

- Event 1 ("2-sign-flip"): $r_j^{(t+1)} - r^*$ and $r_i^{(t+1)} - r^*$ still have the same sign, but different from sign of $r_j^{(t)} - r^*$ and $r_i^{(t)} - r^*$;
- Event 2 ("0-sign-flip"): $r_j^{(t+1)} - r^*$ and $r_i^{(t+1)} - r^*$ still have the same sign, and also the same as the sign of $r_j^{(t)} - r^*$ and $r_i^{(t)} - r^*$;
- Event 3 ("1-sign-flip"): $r_j^{(t+1)} - r^*$ and $r_i^{(t+1)} - r^*$ have different signs. Based on Lemma 6, we must then have $r_j^{(t+1)} > r^* > r_i^{(t+1)}$.

Note that, once Event 3 happens, the problem again reduces to the Basic Cases. Therefore, it is sufficient to show that the first two "bad events" only happen finitely many times, and Event 3 must eventually happen. To do so, we state the following key lemma, which shows that 2-sign flips can only happen a finite number of times. The proof of Lemma 7 is deferred to Appendix E.4.

Lemma 7 *Suppose $\epsilon_1 < 0$, $\epsilon_2 > 0$. Let n be the number of times that a "2-sign flip" (Event 1) happens. Define the universal constants*

$$\beta = \eta \cdot \max\{-\epsilon_1, \epsilon_2\},$$
$$C = \eta(\epsilon_2 - \epsilon_1) \min\left\{\frac{r^* \exp(-\beta)}{(1 + r^* \exp(-\beta))^2}, \frac{r^* \exp(\beta)}{(1 + r^* \exp(\beta))^2}\right\} > 0.$$

Further, let $W_1 = \ln \frac{r_j^{(1)}}{r_i^{(1)}} > 0$. Then, we have

$$n \leq \frac{1}{\ln(1 + C)} \ln \frac{2\beta}{W_1}.$$

To show that Event 2 ("0-sign-flip") happens finitely often, we *disprove* the following claim. We can now finish the proof by *disproving* the following claim.

Claim 1 *Let t_0 be the smallest round such that all "2-sign-flips" have finished (as per Lemma 7, t_0 will be finite). Moreover, suppose that $r_j^{(t_0)} - r^*$ and $r_i^{(t_0)} - r^*$ still have the same sign. Then, both $r_j^{(t)} - r^*$ and $r_i^{(t)} - r^*$ do not change their sign for all $t > t_0$.*

Suppose that the claim above is true. Lemma 6 tells us that $r_j^{(t)} > r_i^{(t)}$ for all t. First, suppose that $r_j^{(t)} > r_i^{(t)} > r^*$ for all $t > t_0$. Since $f(\cdot)$ is strictly decreasing, we have for any player $m \in \{1, 2\}$ that $f(r_m^{(t)}) < f(r^*) = 0$. Equation (5) would then give us

$$r_m^{(t+1)} = r_m^{(t)} \exp\big(\eta f(r_n^{(t)})\big),$$

for any player pair $m \neq n \in \{1, 2\}$. This implies that the ratio $\{r_m^{(t)}\}_{t \geq 1}$ would be decreasing for each player $m \in \{1, 2\}$. Moreover, $r_m^{(t)} \geq 0$ since it is a probability ratio. Therefore, based on the monotone convergence theorem, we have $\lim_{t \to \infty} r_m^{(t)} = \ell_m$ for some constant $\ell_m \geq 0$. Further, since $f(\cdot)$ is continuous and monotone, we have

$$f(\ell_m) = f\left(\lim_{t \to \infty} r_m^{(t)}\right) = \lim_{t \to \infty} f(r_m^{(t)}) = \frac{1}{\eta} \lim_{t \to \infty} \left(\ln r_n^{(t+1)} - \ln r_n^{(t)}\right) = 0,$$

i.e., $f(\ell_m) = 0$. Since f is strictly monotone and $f(r^*) = 0$, we have $\ell_m = r^*$ for both players $m \in \{1, 2\}$. Thus, we have

$$\lim_{t \to \infty} \ln \frac{r_j^{(t)}}{r_i^{(t)}} = \lim_{t \to \infty} \ln r_j^{(t)} - \ln r_i^{(t)} = 0.$$

However, Lemma 6 implies that $\ln r_j^{(t)} - \ln r_i^{(t)} \geq W_1 > 0$, which contradicts the above equation.

We similarly handle the other possible situation, where $r_i^{(t)} < r_j^{(t)} < r^*$ for all $t > t_0$. Since $f(\cdot)$ is strictly decreasing, we have for any player $m \in \{1, 2\}$ that $f(r_m^{(t)}) > f(r^*) = 0$. Equation (5) would then give us

$$r_m^{(t+1)} = r_m^{(t)} \exp\big(\eta f(r_n^{(t)})\big)$$

for any player pair $m \neq n \in \{1, 2\}$. This implies that the ratio $\{r_m^{(t)}\}_{t\geq 1}$ would be increasing for each player $m \in \{1, 2\}$. Moreover, under the claim that the signs do not change for either player, we would have $r_m^{(t)} \leq r^*$. Therefore, based on the monotone convergence theorem, we have $\lim_{t\to\infty} r_m^{(t)} = \ell_m$ for some constant $\ell_i \leq r^*$. Further, since f is continuous and monotone, we have

$$f(\ell_m) = f\left(\lim_{t\to\infty} r_m^{(t)}\right) = \lim_{t\to\infty} f(r_m^{(t)}) = \frac{1}{\eta} \lim_{t\to\infty} \big(\ln r_n^{(t+1)} - \ln r_n^{(t)}\big) = 0,$$

i.e., $f(\ell_m) = 0$. Since f is monotone and $f(r^*) = 0$, we have $\ell_m = r^*$ for both players $m \in \{1, 2\}$. Thus, we have

$$\lim_{t\to\infty} \ln \frac{r_j^{(t)}}{r_i^{(t)}} = \lim_{t\to\infty} \ln r_j^{(t)} - \ln r_i^{(t)} = 0.$$

However, Lemma 6 implies that $\ln r_j^{(t)} - \ln r_i^{(t)} \geq W_1 > 0$, which contradicts the above equation.

This completes the *disproof* of Claim 1. Combined with Lemma 2, this means that a 1-sign-flip (Event 3) must happen at some finite round, which completes the proof of Theorem 1 (Part II) in this case.

E.3. Convergence and Divergence in the Identical Initialization Case

Finally, we consider the case of equal initialization, where $\Delta_1^{(1)} = \Delta_2^{(1)}$.

Condition 4: $\Delta_1^{(1)} = \Delta_2^{(1)}$**,** $\eta < \frac{8}{|\epsilon_1|+|\epsilon_2|}$. For convenience, we define $u_i^{(t)} = \ln \frac{r_i^{(t)}}{r^*}$. We can then define a 1-dimensional dynamical system for each player, given by

$$u_i^{(t+1)} = \ln \frac{r_i^{(t+1)}}{r^*} = u_i^{(t)} + \eta\, f(r^* \exp(u_j^{(t)})).$$

Since $\Delta_1^{(1)} = \Delta_2^{(1)}$ and both players use the same update rule (5), we observe that $r_1^{(t)} = r_2^{(t)}$ for all $t \geq 1$, and the same holds for $u_i^{(t)}$. Therefore, we can denote a common variable $u^{(t)} := u_i^{(t)}$ and write its update in the following form:

$$u^{(t+1)} = T(u^{(t)}),$$

where we define

$$T(u) = u + \eta\, \frac{\epsilon_1 r^* \exp(u) + \epsilon_2}{1 + r^* \exp(u)}.$$

We now have

$$T'(u) = 1 + \eta\, \frac{(\epsilon_1 - \epsilon_2) r^* \exp(u)}{(1 + r^* \exp(u))^2} = 1 - \eta\, \frac{\Gamma r^* \exp(u)}{(1 + r^* \exp(u))^2}.$$

where we denote $\Gamma = |\epsilon_1| + \epsilon_2 > 0$ as shorthand. Also note that $T(0) = 0$ is a fixed point. Because T is a continuous function, we can use the mean value theorem to write

$$|T(u) - T(0)| = |T'(u')||u - 0|$$

for any value u, where u' is some point between u and 0. Under our assumption on the step size, we have $\eta\Gamma \in (0, 8)$. In this case, based on the fact that $\max_{z \geq 0} \frac{z}{(1+z)^2} = \frac{1}{4}$, it can be easily seen that $|T'(u')| < 1$ for all u'. Putting these together, we see that T is a *contraction map*. Applying Banach's fixed point theorem then implies that $u^{(t)} \to 0$.

Condition 5: $\Delta_1^{(1)} = \Delta_2^{(1)}$**,** $\eta > \frac{8}{|\epsilon_1|+|\epsilon_2|}$. Here, we construct a simple example in which the algorithm oscillates. For simplicity, let $\eta = 1$ and $|\epsilon_1| = \epsilon_2$. This means that $r^* = 1$, and the map T simplifies to

$$T(u) = u + \epsilon_1 \frac{\exp(u) - 1}{1 + \exp(u)} = u + \epsilon_1 \tanh\left(\frac{u}{2}\right).$$

Since $\tanh(u)$ and u are both odd functions in u, T is also an odd function. Formally, we have $T(-u) = -T(u)$. Let $a > 0$ be any constant, and

$$\epsilon_1 = -2a\coth\left(\frac{a}{2}\right), \quad \epsilon_2 = 2a\coth\left(\frac{a}{2}\right).$$

Note that $a\coth(\frac{a}{2}) > 2$, so $\eta(|\epsilon_1| + |\epsilon_2|) > 8$ in this case. Therefore, we have

$$T(a) = a + \varepsilon_1 \tanh\left(\tfrac{a}{2}\right) = a - 2a\coth\left(\tfrac{a}{2}\right)\tanh\left(\tfrac{a}{2}\right) = a - 2a\left(\tanh \cdot \coth\right) = -a,$$

and because T is an odd function we have $T(-a) = -T(a) = -(-a) = a$. Therefore, the algorithm will oscillate if initialized at a or $-a$. Note that such a game can be defined for any positive constant a.

E.4. Proof of Lemma 7

We complete this section by providing the proof of Lemma 7. Recall that, without loss of generality, we index a player pair $i \neq j \in \{1, 2\}$ such that $r_j^{(1)} > r_i^{(1)}$. We also define $u_m^{(t)} = \ln \frac{r_m^{(t)}}{r^*}$ for each player $m \in \{1, 2\}$. Since $r_i^{(1)} < r_j^{(1)}$, we have $u_i^{(1)} < u_j^{(1)}$. Moreover, for each player $m \in \{1, 2\}$, $u_m^{(t)}$ and $r_m^{(t)} - r^*$ have the same sign. Reproducing the main argument in the proof of Lemma 6 below, we also have

$$u_j^{(t+1)} - u_i^{(t+1)} = \ln \frac{r_j^{(t+1)}}{r_i^{(t+1)}} = \ln \frac{r_j^{(t)}}{r_i^{(t)}} + \eta[f(r_i^{(t)}) - f(r_j^{(t)})] > \ln \frac{r_j^{(t)}}{r_i^{(t)}} = u_j^{(t)} - u_i^{(t)} > 0. \quad (10)$$

Next, we show that, if Event 1 (i.e. "2-sign-flip") happens at round t, then $|u_m^{(t)}|$ must be small for both players $m \in \{1, 2\}$.

Lemma 8 *Suppose $u_i^{(t)}$ and $u_j^{(t)}$ have the same sign, and the sign of both $u_i^{(t+1)}$ and $u_j^{(t+1)}$ flip. Then, we must have* $\max\{|u_i^{(t)}|, |u_j^{(t)}|\} \leq \beta = \eta \cdot \max\{-\epsilon_1, \epsilon_2\}$.

Proof First, note that f is monotonically decreasing, $f(0) = \epsilon_2 > 0$, and $f(\infty) = \epsilon_1 < 0$. This means that $\epsilon_1 \leq f(r) \leq \epsilon_2$ for any ratio r.

Next, consider the case where $u_m^{(t)} < 0$, and $u_m^{(t+1)} > 0$ for both players $m \in \{1, 2\}$. We would like to show that in this case that $\max\{-u_1^{(t)}, -u_2^{(t)}\} \leq \beta$. We will prove this statement by contradiction. Since we have $-u_i^{(t)} > -u_j^{(t)}$, we suppose instead that $-u_i^{(t)} > \beta > \eta\epsilon_2$, i.e., $u_i^{(t)} < -\beta < -\eta\epsilon_2$. Then, we have

$$u_i^{(t+1)} = u_i^{(t)} + \eta f(r^* \exp(u_j^{(t)})) < -\eta\epsilon_2 + \eta\epsilon_2 = 0,$$

where the inequality above substitutes our supposition $u_i^{(t)} < -\eta\epsilon_2$ and uses the fact that $f(r) \leq \epsilon_2$ for any value of r. Overall, this yields $u_i^{(t+1)} < 0$, which contradicts our assumption of "2-sign-flip" which would instead yield $u_i^{(t+1)} > 0$. Similarly, consider the case where $u_m^{(t)} > 0$, and $u_m^{(t+1)} < 0$ for both players $m \in \{1, 2\}$. We will again prove the statement by contradiction. Suppose instead that $u_j^{(t)} > \beta > -\eta\epsilon_1$. Then, we have

$$u_j^{(t+1)} = u_j^{(t)} + \eta f(r^* \exp(u_i^{(t)})) > -\eta\epsilon_1 + \eta\epsilon_1 = 0,$$

where the inequality above substitutes our supposition $u_j^{(t)} > -\eta\epsilon_1$ and uses the fact that $f(r) \geq \epsilon_1$ for any value of r. Overall, this yields $u_j^{(t+1)} > 0$, which contradicts our assumption of "2-sign-flip" which would instead yield $u_j^{(t+1)} < 0$. This completes the proof of this lemma. ■

Armed with Lemma 8, we are ready to prove Lemma 7. Define the potential function $W_t = u_j^{(t)} - u_i^{(t)}$. We know from Lemma 6 that W_t is monotonically increasing. Then, we have

$$W_{t+1} = u_j^{(t+1)} - u_i^{(t+1)} = u_j^{(t)} - u_i^{(t)} + \eta(g(u_i^{(t)}) - g(u_j^{(t)})) = (1 - \eta g'(\delta_t))(u_j^{(t)} - u_i^{(t)}),$$

where we defined $g(u) := f(r^* \exp(u))$ as shorthand. As it is easy to see that $g(u)$ is monotone and continuous, the last equality above is based on the mean value theorem (where $\delta^{(t)} \in [u_i^{(t)}, u_j^{(t)}]$). Note that for any $|u| \leq \beta$, we have

$$1-\eta g'(u) = 1+\eta\frac{(\epsilon_2 - \epsilon_1)r^* \exp(u)}{(1 + r^* \exp(u))^2} \geq 1+\eta(\epsilon_2-\epsilon_1) \min\left\{\frac{r^* \exp(-\beta)}{(1 + r^* \exp(-\beta))^2}, \frac{r^* \exp(\beta)}{(1 + r^* \exp(\beta))^2}\right\}.$$

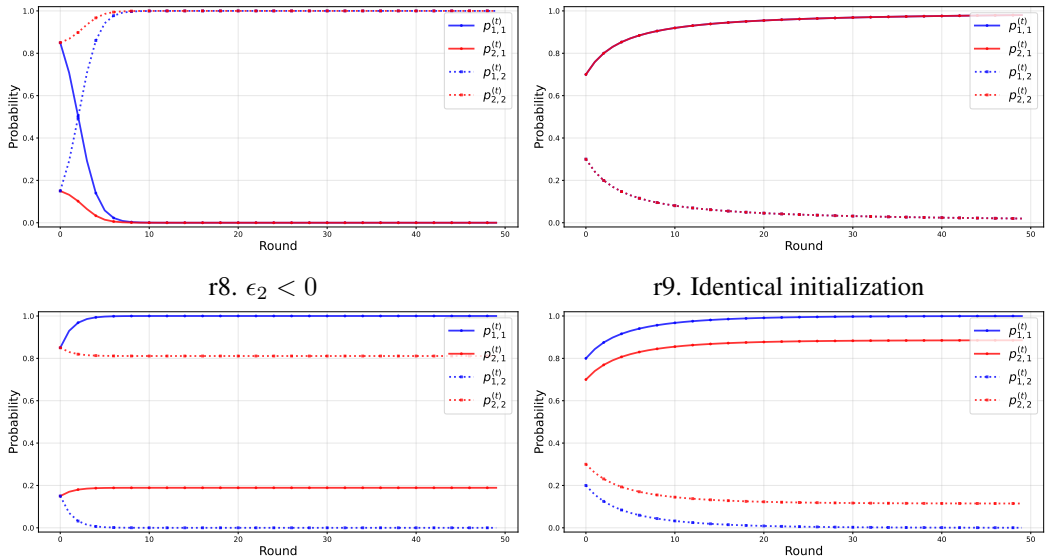
Let $\{t_1, \ldots, t_n\}$ denote the rounds on which "2-sign-flip" happens. Using the monotonically increasing property of $\{W_t\}_{t\geq 1}$, we have

$$W_{t_n} \geq (1 + C)W_{t_{n-1}} \ldots \geq (1 + C)^n W_1,$$

where n is the total number of rounds where Event 1 happens, and

$$C = \eta(\epsilon_2 - \epsilon_1) \min\left\{\frac{r^* \exp(-\beta)}{(1 + r^* \exp(-\beta))^2}, \frac{r^* \exp(\beta)}{(1 + r^* \exp(\beta))^2}\right\} > 0.$$

On the other hand, Lemma 8 tells us that $W_{t_n} \leq 2\beta$. Combining the upper and lower bounds on W_{t_n} completes the proof of Lemma 7.



r8. $\epsilon_2 < 0$ r9. Identical initialization

r9. Opposite-Sign initialization r9. Same-sign initialization

Figure 4: Simulation results for Theorem 1 (Part IV). The row in Table 3 to which each case corresponds is marked in the sub-captions.

Appendix F. Proof of Theorem 1 (Part III)

In this section, we prove Theorem 1 (Part III). In this case, we have $\epsilon_1 > 0$ and $\epsilon_2 < 0$. The corresponding simulation results are given in Figure **??** (we did not include them in the main paper due to space limitations). We first recall/introduce some basic conclusions. Recall that $f'(r) = \frac{\epsilon_1 - \epsilon_2}{(1+r)^2} > 0$, which means that $f(\cdot)$ is strictly increasing for the case where $\epsilon_1 > 0$ and $\epsilon_2 < 0$. Moreover, $r^* = \frac{|\epsilon_2|}{\epsilon_1}$ is the unique root of $f(\cdot)$. Recall that for both players $i \in \{1, 2\}$, we can write the following dynamical system

$$u_i^{(t+1)} = \ln \frac{r_i^{(t+1)}}{r^*} = u_i^{(t)} + g(u_j^{(t)}),$$

where we defined $g(u) = \eta f(r^* \exp(u))$ as shorthand. Note that, since f is strictly increasing and $f(r^*) = 0$, if $r - r^* > 0$, $\Delta = f(r) > 0$, and also $u = \ln \frac{r}{r^*} > 0$. On the other hand, if $r - r^* < 0$, $\Delta = f(r) < 0$, and also $u = \ln \frac{r}{r^*} < 0$. Thus, we have $\text{Sign}(\Delta) = \text{Sign}(r - r^*) = \text{Sign}(u)$. Next, we have

$$g'(u) = \frac{\eta(\epsilon_1 - \epsilon_2) r^* \exp(u)}{(1 + r^* \exp(u))^2}.$$

For all $u \in \mathbb{R}$, we have $g'(u) \leq \frac{\eta(\epsilon_1 - \epsilon_2)}{4}$. Therefore, we have $g'(u) \in (0, 2)$ when $\eta \in \left(0, \frac{8}{|\epsilon_1| + |\epsilon_2|}\right)$. Next, we provide our convergence result for different choices of initialization.

F.1. Exponential Convergence with Same Sign Initialization

We first provide a formal description of exponential convergence in the "easy" cases where the signs of the initialization functionals $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ are the same.

Theorem 9 *Suppose Assumption 1 holds, $\epsilon_1 > 0$ and $\epsilon_2 < 0$. Then, Algorithm 1 with any step size $\eta > 0$ satisfies the following:*

- *If $\Delta_1^{(1)} > 0$, $\Delta_2^{(1)} > 0$, then we have $p_{i,2}^{(t+1)} \leq \frac{1}{r_i^{(1)}} \exp\left(-\eta \Delta_j^{(1)}\right)$ for any player pair $i \neq j \in \{1, 2\}$. This implies that the algorithm converges to the pure symmetric NE (θ_1, θ_1) exponentially fast;*
- *If $\Delta_1^{(1)} < 0$, $\Delta_2^{(1)} < 0$, then we have $p_{i,1}^{(t+1)} \leq r_i^{(1)} \exp\left(\eta \Delta_j^{(1)}\right)$ for any player pair $i \neq j \in \{1, 2\}$. This implies that the algorithm converges to the pure symmetric NE (θ_2, θ_2) exponentially fast.*

We now provide the proof of Theorem 9, considering each condition listed above.

Condition 1: $\Delta_1^{(1)} > 0$, $\Delta_2^{(1)} > 0$. Recall that for any player pair $i \neq j \in \{1, 2\}$, we have $r_i^{(2)} = r_i^{(1)} \exp(\Delta_j^{(1)})$. Based on induction, it is easy to see that the ratios for both players, $r_1^{(t)}$ and $r_2^{(t)}$, are increasing in t. Moreover, for each player $i \in \{1, 2\}$, the functional $\Delta_i^{(t)} = f(r_i^{(t)})$ is also increasing in t (since $f(\cdot)$ is increasing in its argument and $r_i^{(t)}$ is increasing in t). Therefore, we have

$$r_i^{(t+1)} = r_i^{(1)} \exp\left(\eta \sum_{k=1}^{t} \Delta_j^{(k)}\right) \geq r_i^{(1)} \exp\left(\eta t \Delta_j^{(1)}\right)$$

for any player pair $i \neq j \in \{1, 2\}$. Noting that $\Delta_j^{(1)} > 0$ is a positive constant, this proves the desired exponential rate of convergence to the symmetric pure NE (θ_1, θ_1).

Condition 2: $\Delta_1^{(1)} < 0$, $\Delta_2^{(1)} < 0$. Using a similar inductive argument, it is easy to see that the ratios for both players $r_1^{(t)}$ and $r_2^{(t)}$ are now decreasing in t. Because $f(\cdot)$ is increasing, the functional $\Delta_i^{(t)} = f(r_i^{(t)})$ is also decreasing in t (since $f(\cdot)$ is increasing in its argument and $r_i^{(t)}$ is decreasing in t). Therefore, we have

$$r_i^{(t+1)} = r_i^{(1)} \exp\left(\eta \sum_{k=1}^{t} \Delta_j^{(k)}\right) \leq r_i^{(1)} \exp\left(\eta t \Delta_j^{(1)}\right)$$

for any player pair $i \neq j \in \{1, 2\}$. Noting that $\Delta_j^{(1)} < 0$ is a negative constant, this proves the desired exponential rate of convergence to the symmetric pure NE (θ_2, θ_2).

F.2. Convergence with Opposite Sign Initialization

Finally, we consider the case where $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have different signs.

Condition 3: $\Delta_1^{(1)}$ and $\Delta_2^{(1)}$ have different signs. Suppose that we have $\Delta_i^{(t)} > 0$ and $\Delta_j^{(t)} < 0$ for some player pair $i \neq j \in \{1, 2\}$ for some $t \geq 1$. We start by excluding some corner cases for the update to round $t + 1$, i.e. where one or both of the $\Delta_i^{(t+1)}$ becomes equal to 0.

- Suppose that $\Delta_i^{(t+1)} = 0$ and $\Delta_j^{(t+1)} > 0$. Then, we have $\Delta_j^{(t+2)} = \Delta_j^{(t+1)} > 0$, and $r_i^{(t+2)} = r_i^{(t+1)} \exp(\eta \Delta_j^{(t+1)}) > r_i^{(t+1)}$. Because $f(\cdot)$ is increasing, we have $\Delta_i^{(t+2)} > \Delta_i^{(t+1)} = 0$.

This means that $\Delta_i^{(t+2)} > 0$ for both players $i \in \{1, 2\}$, and the situation reduces to Condition 1.

- Suppose that $\Delta_i^{(t+1)} = 0$ and $\Delta_j^{(t+1)} < 0$. A similar argument to the first corner case yields $\Delta_i^{(t+2)} < 0$ for both players $i \in \{1, 2\}$, and the situation reduces to Condition 2.
- Finally, suppose that $\Delta_j^{(t+1)} = \Delta_i^{(t+1)} = 0$. Then, since f is strictly monotone and $f(r^*) = 0$, we have that $r_i^{(t+1)} = r_j^{(t+1)} = r^*$. Moreover, since $r_i^{(t+1)} = r_i^{(t)} \exp(\Delta_j^{(t)})$, the algorithm will stay at $r_i^{(t+1)} = r_j^{(t+1)} = r^*$, which is the unique *strictly* mixed NE of the game.

Therefore, we can exclude the corner cases above. Then, as in the proof of Theorem 1 (Part II), three events can transpire going from $t \to t+1$:

- Event 1 ("2-sign-flip"): $\Delta_i^{(t+1)} < 0$ and $\Delta_j^{(t+1)} > 0$;
- Event 2 ("1-sign-flip"): $\Delta_i^{(t+1)}$ and $\Delta_j^{(t+1)}$ have the same sign;
- Event 3 ("0-sign-flip"): $\Delta_i^{(t+1)} > 0$ and $\Delta_j^{(t+1)} < 0$;

We first focus on Event 1. In particular, suppose Event 1 happens at round t. Define the potential function $V_t = |u_1^{(t)}| + |u_2^{(t)}|$. Because the signs of $u_i^{(t)}$ and $\Delta_i^{(t)}$ match for both players $i \in \{1, 2\}$, we have

$$V_t = u_i^{(t)} - u_j^{(t)},$$

and therefore we can write the potential function update as

$$V_{t+1} = u_j^{(t+1)} - u_i^{(t+1)} = u_j^{(t)} - u_i^{(t)} + (g(u_i^{(t)}) - g(u_j^{(t)})) = (g'(\delta^{(t)}) - 1)V_t, \tag{11}$$

where the last equality is based on the mean value theorem for some $\delta^{(t)} \in [u_j^{(t)}, u_i^{(t)}]$. From this, we can infer the following:

- If $\eta \in \left(0, \frac{4}{|\epsilon_1|+|\epsilon_2|}\right)$, then $g'(u) < 1$. Thus, (11) implies that $V_{t+1} < 0$, which contradicts that $V_{t+1} \geq 0$ by definition. Thus, if $\eta \in \left(0, \frac{4}{|\epsilon_1|+|\epsilon_2|}\right)$, Event 1 cannot ever happen.
- If instead $\eta \in \left[\frac{4}{|\epsilon_1|+|\epsilon_2|}, \frac{8}{|\epsilon_1|+|\epsilon_2|}\right)$, we have $g'(u) - 1 < 1$. In this case, V_{t+1} will shrink by a constant every time Event 1 happens.

To summarize, if $\eta \in \left(0, \frac{8}{|\epsilon_1|+|\epsilon_2|}\right)$, one of the two outcomes will transpire:

- *If Event 1 a finite number of times:* i.e., there exists a large enough t_0 such that Event 1 will not happen after t_0. Now, consider all rounds $t > t_0$. If Event 2 happen at some round t, then the problem reduces to Conditions 1 or 2. Otherwise, only Event 3 happens at each round. Based on the monotone convergence theorem and monotonicity of $f(\cdot)$, the algorithm converges to the strictly mixed NE in this case.
- *If Event 1 happens infinitely many times:* In this case, $\lim_{t\to\infty} V_t = 0$, which means that both $|u_i^{(t)}|$ and $|u_j^{(t)}|$ converge to 0. In this case, the algorithm converges to the strictly mixed NE.

In conclusion, for $\eta \in \left(0, \frac{8}{|\epsilon_1|+|\epsilon_2|}\right)$, the algorithm will converge to either the pure NE or the strictly mixed NE. This completes the proof of Theorem 1 (Part III).

F.3. An example of oscillation with large step size under opposite-sign initialization

In this section, we provide a simple example of a game in which the algorithm oscillates if the step size is instead too large, i.e. $\eta > \frac{8}{|\epsilon_1|+|\epsilon_2|}$, with a very specific initialization. This eample is analogous and similar to the one provided in Condition 5 in the proof of Theorem 1 (Part II). For simplicity, we consider $\eta = 1$ and $\epsilon_1 = |\epsilon_2| = \epsilon$, where ϵ will be a parameter that we set later. Note that this implies that the unique root of $f(\cdot)$ is $r^* = 1$.

We now consider *equal and opposite* initialization, i.e. $\Delta_1^{(1)} = -\Delta_1^{(2)} = \Delta^{(1)}$. It is easy to verify that in this case, $r_2^{(1)} = 1/r_1^{(1)}$. Moreover, we can show via induction that $r_2^{(t)} = 1/r_1^{(t)}$. Therefore, it suffices to keep track of a common ratio $r^{(t)} := r_1^{(t)}$ and the corresponding common variable $u^{(t)} = \ln \frac{r^{(t)}}{r^*} = \ln r^{(t)}$. We can then define a 1-dimensional dynamical system on $u^{(t)}$, where the update $t \to t+1$ is given by

$$\begin{aligned} u^{(t+1)} &= u^{(t)} - \Delta^{(t)} \\ &= u^{(t)} - f(r^{(t)}) \\ &= u^{(t)} - f(\exp(u^{(t)})). \end{aligned}$$

Recalling the definition $f(r) = \frac{\epsilon_1 r + \epsilon_2}{1+r}$ and noting that we are considering a game where $\epsilon_1 = -\epsilon_2 = \epsilon$, we have $f(r) = \epsilon \frac{r-1}{1+r}$, we have

$$\begin{aligned} u^{(t+1)} &= u^{(t)} - \epsilon \frac{\exp(u^{(t)}) - 1}{1 + \exp(u^{(t)})} \\ &= u^{(t)} - \epsilon \tanh\left(\frac{u^{(t)}}{2}\right) =: T(u^{(t)}). \end{aligned}$$

Ultimately, we have the map $u \to T(u) := u - \epsilon \tanh\left(\frac{u}{2}\right)$. Analogous to the example provided in Condition 5 of the proof of Theorem 1 (Part II), we now consider $a > 0$ to be any constant, and set

$$\epsilon = 2a \coth\left(\frac{a}{2}\right).$$

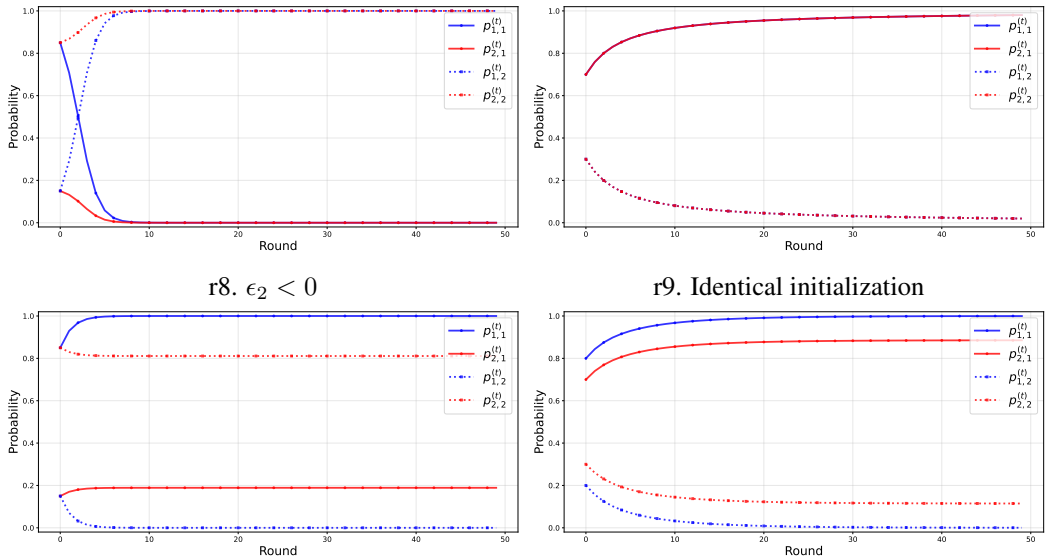
As previously argued, this choice satisfies $\eta > \frac{8}{|\epsilon_1|+|\epsilon_2|}$. Therefore, we have

$$\begin{aligned} T(a) &= a - \epsilon \tanh\left(\frac{a}{2}\right) \\ &= a - 2a \coth\left(\frac{a}{2}\right) \tanh\left(\frac{a}{2}\right) = -a. \end{aligned}$$

Similarly, noting that $T(u)$ is again an even function, we have $T(-a) = a$. Therefore, the algorithm will oscillate if initialized at a or $-a$. Note that such a game can be defined for any positive constant a.

Appendix G. Proof of Theorem 1 (Part IV)

In this section, we prove Theorem 1 (Part IV). In this case, we have (without loss of generality) that $\epsilon_1 = 0$ and $\epsilon_2 \neq 0$. The corresponding simulation results are given in Figure 1 (we did not include them in the main paper due to space limitations). We study the convergence of the algorithm under different conditions on the game, listed below.



r8. $\epsilon_2 < 0$ r9. Identical initialization

r9. Opposite-Sign initialization r9. Same-sign initialization

Figure 5: Simulation results for Theorem 1 (Part IV). The row in Table 3 to which each case corresponds is marked in the sub-captions.

Condition 1: $\epsilon_1 = 0$, $\epsilon_2 < 0$. Recall from Equation (5) that we have $r_i^{(t+1)} = r_i^{(t)} \exp\left(\eta f(r_j^{(t)})\right)$ for any player pair $i \neq j \in \{1, 2\}$, where $f(r) = \frac{\epsilon_2}{1+r} < 0$, and $f'(r) = -\frac{\epsilon_2}{(1+r)^2} > 0$. Since $f(r) < 0$, we can argue through induction that the ratio $\{r_i^{(t)}\}_{t=1}^{\infty}$ is non-increasing for both players $i \in \{1, 2\}$. This means that $r_i^{(t)} \leq r_i^{(1)}$, and, because $f(\cdot)$ is increasing, we have $f(r_i^{(t)}) \leq f(r_i^{(1)})$. As a consequence, we have for any player pair $i \neq j \in \{1, 2\}$:

$$r_i^{(t+1)} = r_i^{(1)} \exp\left(\eta \sum_{k=1}^{t} f(r_j^{(k)})\right) \leq r_i^{(1)} \exp(\eta t \Delta_j^{(1)}).$$

This implies that $p_{i,1}^{(t+1)} \leq p_{i,2}^{(t+1)} r_i^{(1)} \exp(\eta t \Delta_j^{(1)}) \leq r_i^{(1)} \exp(\eta t \Delta_j^{(1)}) = r_i^{(1)} \exp(\eta t p_j^{(1)} \epsilon_2)$, and $\epsilon_2 < 0$. Therefore, $p_{i,1}^{(t)}$ converges to 0 at an exponentially fast rate under this condition.

Condition 2: $\epsilon_1 = 0$, $\epsilon_2 > 0$, $\Delta_1^{(1)} = \Delta_2^{(1)}$. In this case, we have $f(r) > 0$. A similar inductive argument as in Condition 1 then tells us that $\{r_i^{(t)}\}_{t=1}^{\infty}$ is strictly increasing. Moreover, we have

$$\ln r_i^{(t+1)} = \ln r_i^{(t)} + \eta \frac{\epsilon_2}{1 + r_i^{(t)}}.$$

Suppose that$\lim_{t\to\infty} r_i^{(t)} = \ell_i < \infty$ for some $\ell_i > 0$. Taking the limit as $t \to \infty$ on both sides, we have

$$\ln \ell_i = \ln \ell_i + \eta \frac{\epsilon_2}{1 + \ell_i}.$$

The above equation is solved at $\ell_i = \infty$, which contradicts with the assumption that $\ell_i < \infty$. Thus, $\lim_{t\to\infty} r_i^{(t)} = \infty$ for $i \in \{1, 2\}$.

Condition 3: $\epsilon_1 = 0$, $\epsilon_2 > 0$, $\Delta_1^{(1)} \neq \Delta_2^{(1)}$. For this case, we first show that at least one player will converge to θ_1, which means that the algorithm will at least converge to a mixed NE. Specifically, we have

$$\ln r_i^{(t+1)} = \ln r_i^{(t)} + \eta\epsilon_2 \frac{1}{1 + r_j^{(t)}}.$$

Since the ratio is always non-negative, i.e. $r \geq 0$, we know both $r_i^{(t)}$ and $r_j^{(t)}$ are monotonically increasing. Based on the monotone convergence theorem, $r_i^{(t)}$ (and $r_j^{(t)}$) either converge to some $0 < \ell_i < \infty$ ($0 < \ell_j < \infty$), or go to infinity. Suppose both $r_i^{(t)}$ and $r_j^{(t)}$ converge to some constants. Let $\lim_{t\to\infty} r_i^{(t)} = \ell_i$ and $\lim_{t\to\infty} r_j^{(t)} = \ell_j$. Then, we have

$$\ell_i = \lim_{t\to\infty} \ln r_i^{(t+1)} = \lim_{t\to\infty} \ln r_i^{(t)} + \lim_{t\to\infty} \eta\epsilon_2 \frac{1}{1 + r_j^{(t)}} = \ell_i + \frac{\eta\epsilon_2}{1 + \ell_j},$$

The above equation can only be solved at $\ell_j = \infty$, which leads to a contradiction. Therefore, at least one of $r_i^{(t)}$ and $r_j^{(t)}$ will go to infinity.

Example for converging to strictly mixed NE Here, we construct an instance where $r_j^{(t)}$ converges to a constant, while $r_i^{(t)}$ goes to infinity.

Theorem 10 *For any player pair $i, j \in \{1, 2\}$, $i \neq j$, let $r_j^{(1)}$ be a constant, and let A be any constant such that $A > r_j^{(1)}$. Let*

$$r_i^{(1)} \geq \frac{\eta\epsilon_2}{\left(1 - \exp\left(-\frac{\eta\epsilon_2}{1+A}\right)\right) \ln \frac{A}{r_j^{(1)}}} > 0.$$

Then, we have $\lim_{t\to\infty} r_j^{(t)} = A' \in (0, A]$, and $\lim_{t\to\infty} r_i^{(t)} = \infty$. This implies that $\lim_{t\to\infty} p_{q,1}^{(t)} = \frac{A'}{1+A'}$, while $\lim_{t\to\infty} p_{p,1}^{(t)} = \infty$.

Proof We prove the theorem by induction. Note that $r_j^{(1)} \leq A$ by assumption. Assuming that $r_j^{(t)} \leq A$,we have

$$r_i^{(t+1)} = r_i^{(t)} \exp\left(\eta\epsilon_2 \frac{1}{1 + r_j^{(t)}}\right) \geq r_i^{(1)} \exp\left(\eta\epsilon_2 \frac{t}{1 + A}\right)$$

for all $t \geq 1$. Therefore, we have

$$\begin{aligned} \ln r_j^{(t+1)} &= \ln r_j^{(1)} + \eta\epsilon_2 \sum_{k=1}^{t} \frac{1}{1 + r_i^{(k)}} \leq \ln r_j^{(1)} + \eta\epsilon_2 \sum_{k=1}^{t} \frac{1}{r_i^{(k)}} \\ &\leq \ln r_j^{(1)} + \frac{\eta\epsilon_2}{\left(1 - \exp\left(-\frac{\eta\epsilon_2}{1+A}\right)\right)} \cdot \frac{1}{r_i^{(1)}} \leq \ln A, \end{aligned} \tag{12}$$

and thus $r_j^{(t+1)} \leq A$. To summarize, by induction, we showed that $r_j^{(t)} \leq A$ for all t. Since the ratio $\{r_j^{(t)}\}_{t\geq 1}$ is also increasing in t, we have $\lim_{t\to\infty} r_j^{(t)} = A'$ for some $A' \in (0, A]$ by the monotone

convergence theorem. Moreover, because $r_i^{(t)} \geq r_i^{(1)} \exp\left(\frac{\eta\epsilon_2 t}{1+A}\right)$, we have $\lim_{t\to\infty} r_i^{(t)} = \infty$. This completes the proof of Theorem 1 (Part IV). ■

Appendix H. Experimental Result For The Bank Game

We now illustrate the convergence of our framework through experiments based on the bank setup introduced in Section 4.

H.1. The Bank Game

	(τ_ℓ, γ_ℓ)	(τ_ℓ, γ_h)	(τ_h, γ_ℓ)	(τ_h, γ_h)
(τ_ℓ, γ_ℓ)	$\frac{1}{2}h(\gamma_\ell, \tau_\ell, 1)$	0	$\frac{1}{2}h(\gamma_\ell, \tau_h, 1)$	0
(τ_ℓ, γ_h)	$h(\gamma_\ell, \tau_\ell, 1)$	$\frac{1}{2}h(\gamma_h, \tau_\ell, 1)$	$h(\gamma_\ell, \tau_h, 1)$	$\frac{1}{2}h(\gamma_h, \tau_h, 1)$
(τ_h, γ_ℓ)	$h(\gamma_\ell, \tau_\ell, \tau_h) + \frac{1}{2}h(\gamma_\ell, \tau_h, 1)$	$h(\gamma_h, \tau_\ell, \tau_h)$	$\frac{1}{2}h(\gamma_\ell, \tau_h, 1)$	0
(τ_h, γ_h)	$h(\gamma_\ell, \tau_\ell, 1)$	$h(\gamma_h, \tau_\ell, \tau_h) + \frac{1}{2}h(\gamma_h, \tau_h, 1)$	$h(\gamma_\ell, \tau_h, 1)$	$\frac{1}{2}h(\gamma_h, \tau_h, 1)$

Table 5: Utility matrix for bank 1. The rows are the strategies for bank 2, while the columns are the strategies for bank 1.

In this section, we provide more details about the bank game defined in Section 4. More specifically, for each player, there are four actions: $(\tau_\ell, \gamma_\ell), (\tau_\ell, \gamma_h), (\tau_h, \gamma_\ell)$ and (τ_h, γ_h). The utilities of both banks under different pair of decisions are governed by (2). We provide a full characterization of the utility of Bank 1 in Table 5.

Thresholds selection For fixed interest rates $0 < \gamma_\ell < \gamma_h < 1$, we always set the corresponding thresholds as

$$\tau_\ell = \frac{1}{1+\gamma_h}, \text{ and } \tau_h = \frac{1}{1+\gamma_\ell}.$$

These choices for the thresholds are natural from Equation (2). Indeed, it is a dominant strategy for a *rational* bank with interest rate γ to set a threshold of $\tau^*(\gamma) := \frac{1}{2+\gamma}$. Setting a higher threshold $\tau > \tau^*(\gamma)$ leads to ignoring customers in the range $y \in [\tau^*(\gamma), \tau]$ that provide a utility of $(2+\gamma)y - 1 > 0$, while setting a lower threshold $\tau < \tau^*(\gamma)$ leads to potentially incorporating some customers in the range $y < \tau^*(\gamma)$ that would yield a utility of $(2+\gamma)y - 1 < 0$ as a result of low probability of repayment; both cases cannot improve total utility.

Next, we show that the two decisions (τ_ℓ, γ_ℓ) and (τ_h, γ_h) are always dominated decisions, so the weights assigned to these actions by the exponential weights algorithm will decrease exponentially fast. We firstly introduce several easy to verify properties of h.

Lemma 11 *Assume* $\text{Supp}(D_y) = [0, 1]$. *We have the following:*

1. *Let* $\tau_a < \tau_b < \tau_c$. *Then* $h_{D_y}(\gamma_1, \tau_a, \tau_c) = h_{D_y}(\gamma_1, \tau_a, \tau_b) + h_{D_y}(\gamma_1, \tau_b, \tau_c)$.
2. $h(\gamma, \tau_a, \tau_b)$ *is monotonically increasing with respect to* γ.

3. *We have* $h(\gamma, \tau_a, \tau_b) \in [-1, 2]$.

4. *Let* $y_\gamma = \frac{1}{2+\gamma}$. *If* $y_\gamma \in (\tau_a, \tau_b)$, *then* $h(\gamma, \tau_a, y_\gamma) < 0$, *and* $h(\gamma, y_\gamma, \tau_b) > 0$.

Next, based on these properties, it can be easily observe that $u_1((\tau_h, \gamma_\ell), \theta) > u_1((\tau_\ell, \gamma_\ell), \theta)$ for all $\theta \in \{(\tau_\ell, \gamma_\ell), (\tau_\ell, \gamma_h), (\tau_h, \gamma_\ell), (\tau_h, \gamma_h)\}$. Excluding (τ_ℓ, γ_ℓ), we have $u_1((\tau_\ell, \gamma_h), \theta) > u_1((\tau_h, \gamma_h), \theta)$ for $\theta \in \{(\tau_\ell, \gamma_h), (\tau_h, \gamma_\ell), (\tau_h, \gamma_h)\}$.Therefore, (τ_ℓ, γ_ℓ) and (τ_h, γ_h) are dominated decisions.

Credit score distribution We model customer credit scores using two types of distributions:

- Truncated Gaussian (μ, σ): This is a standard normal distribution conditioned to lie within the normalized credit score range $y \in [0, 1]$, and models customers with credit scores symmetrically distributed around a mean value $\mu \in [0, 1]$.

- Piecewise Uniform Distribution: This distribution models β_1 fraction of customers with credit scores uniformly distributed in $[0, \tau_l]$, a fraction β_2 with scores in (τ_l, τ_h), and the remainder in $[\tau_h, 1]$.

$$y \sim \begin{cases} \frac{\beta_1}{\tau_l}, & y \in [0, \tau_l] \\ \frac{\beta_2}{\tau_h - \tau_l} & y \in (\tau_l, \tau_h) \\ \frac{1-(\beta_1+\beta_2)}{1-\tau_h}, & y \in [\tau_h, 1] \end{cases} \tag{13}$$

Interest Rates and Credit Thresholds We ran experiments across a wide range of interest rates, where $0 < \gamma_l < \gamma_h < 1$, and the corresponding thresholds To simulate different conditions for ϵ_1 and ϵ_2, we consider the following parameter configurations (Here, $(+, +)$ means $\epsilon_1 > 0$, $\epsilon_2 > 0$):

- (+, +): $\gamma_l = 0.4$, $\gamma_h = 0.8$, with a truncated Gaussian distribution $(\mu, \sigma) = (0.3, 0.1)$.

- (-, -): $\gamma_l = 0.4$, $\gamma_h = 0.8$, with a truncated Gaussian distribution $(\mu, \sigma) = (0.1, 0.3)$.

- (+, -): $\gamma_l = 0.4$, $\gamma_h = 0.8$, with a truncated Gaussian distribution $(\mu, \sigma) = (0.1, 0.2)$.

- (-, +): $\gamma_l = 0.6$, $\gamma_h = 0.7$, with piecewise uniform distribution $\beta_1 = 0.01, \beta_2 = 0.95$.

Algorithm Configuration We set the step size η in Algorithm 1 to be 0.1. To make the setting more general, we consider applying the exponential weights algorithm to the more general setting where we have four decisions: $(\tau_\ell, \gamma_\ell), (\tau_\ell, \gamma_h), (\tau_h, \gamma_\ell)$ and (τ_h, γ_h). As discussed in Section 4, the first and last action are dominated and the weights will converge to 0 fast.

The experimental results are give in Figures 6–9. It can be seen that the convergence results matches our theorem. We start with the '- -' and '+ +' cases in Figures 6 and 7 (corresponding to Theorem 1 (Part I)), and see convergence to the unique pure symmetric NE.

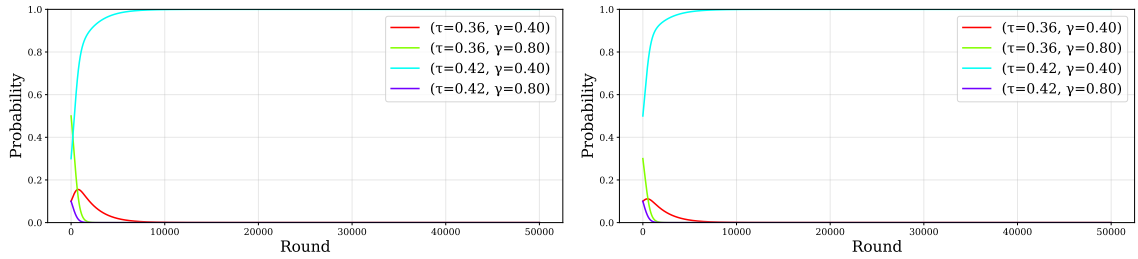


Figure 6: **Case - -** Exponential weights dynamics for Bank1 (left) and Bank2 (right), converging to $((\tau_h, \gamma_l), (\tau_h, \gamma_l))$. $y \sim$ truncated Gaussian ($\mu = 0.1, \sigma = 0.3$), $\gamma_l = 0.4$, $\gamma_h = 0.8$. Initial weights: Bank1 $(0.1, 0.5, 0.3, 0.1)$, Bank2 $(0.1, 0.3, 0.5, 0.1)$.

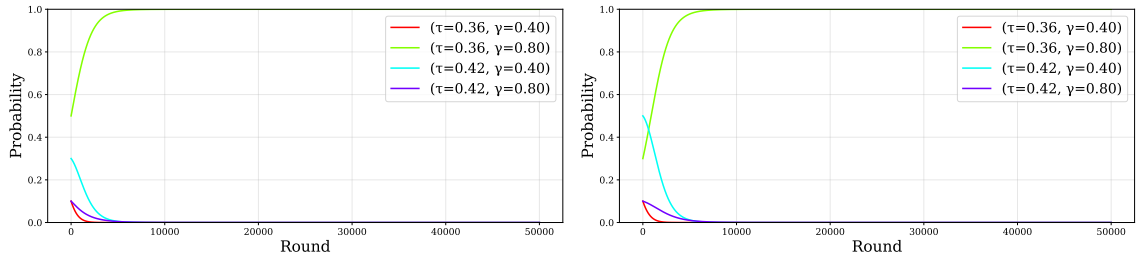


Figure 7: **Case + +** Exponential weights dynamics for Bank1 (left) and Bank2 (right), converging to $((\tau_l, \gamma_h), (\tau_l, \gamma_h))$. $y \sim$ truncated Gaussian ($\mu = 0.3, \sigma = 0.1$), $\gamma_l = 0.4$, $\gamma_h = 0.8$. Initial weights: Bank1 $(0.1, 0.5, 0.3, 0.1)$, Bank2 $(0.1, 0.3, 0.5, 0.1)$.

We now present two representative figures for the '$-+$' case (corresponding to Theorem 1 (Part II)). Figure 8 illustrates convergence to the pure asymmetric NE $((\tau_l, \gamma_h), (\tau_h, \gamma_l))$.

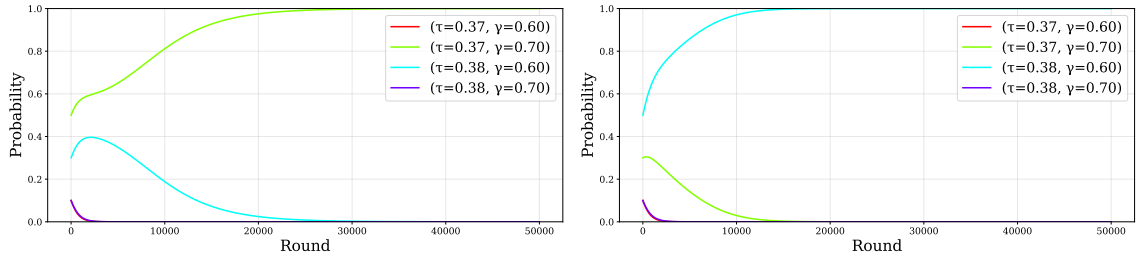


Figure 8: **Case - +** Exponential weights dynamics for Bank1 (left) and Bank2 (right) converging to the pure NE $((\tau_l, \gamma_h), (\tau_h, \gamma_l))$. Parameters: $\epsilon_1 < 0, \epsilon_2 > 0$, $y \sim$ piecewise uniform distribution, $\gamma_l = 0.6$, $\gamma_h = 0.7$. Initial weights: Bank1 $(0.1, 0.5, 0.3, 0.1)$, Bank2 $(0.1, 0.3, 0.5, 0.1)$.

Finally, we present three figures for the '$+-$' case. Figure 9 shows convergence to the pure symmetric NE $((\tau_h, \gamma_l), (\tau_h, \gamma_l))$.

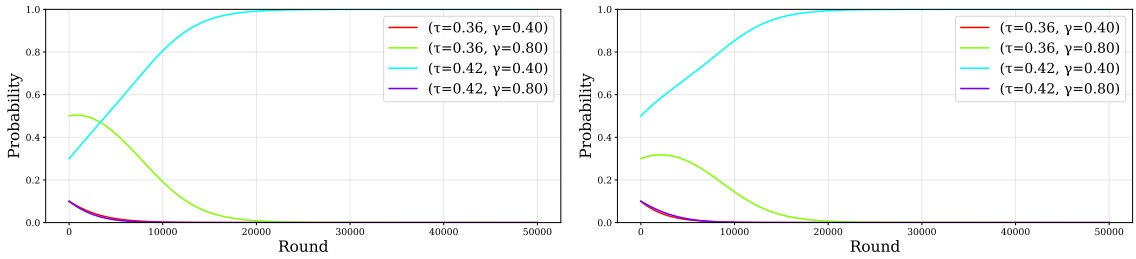


Figure 9: **Case + -** Exponential weights dynamics for Bank1 (left) and Bank2 (right), converging to the pure NE $((\tau_h, \gamma_l), (\tau_h, \gamma_l))$. $y \sim$ truncated Gaussian ($\mu = 0.1, \sigma = 0.2$), $\gamma_l = 0.4$, $\gamma_h = 0.8$. Initial weights: Bank1 $(0.1, 0.5, 0.3, 0.1)$, Bank2 $(0.1, 0.3, 0.5, 0.1)$

.