Beyond the Nash Equilibrium Barrier

Robert Kleinberg[1] Katrina Ligett[1] Georgios Piliouras[2] Éva Tardos[1]
[1]Department of Computer Science, Cornell University, Ithaca NY 148537.
[2]Department of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30308.
Department of Economics, Johns Hopkins University, Baltimore, MD 21218.
[1] {rdk,katrina,eva}@cs.cornell.edu [2]georgios.piliouras@gmail.com

Abstract:
Nash equilibrium analysis has become the de facto standard for judging the solution quality achieved in systems composed of selfish users. This mindset is so pervasive in computer science that even the few papers devoted to directly analyzing outcomes of dynamic processes in repeated games (e.g., best-response or no-regret learning dynamics) have focused on showing that the performance of these dynamics is comparable to that of Nash equilibria. By assuming that equilibria are representative of the outcomes of selfish behavior, do we ever reach qualitatively wrong conclusions about those outcomes?
In this paper, we argue that there exist games whose equilibria represent unnatural outcomes that are hard to coordinate on, and that the solution quality achieved by selfish users in such games is more accurately reflected in the *disequilibrium* represented by dynamics such as those produced by natural families of on-line learning algorithms. We substantiate this viewpoint by studying a game with a unique Nash equilibrium, but where natural learning dynamics exhibit non-convergent cycling behavior rather than converging to this equilibrium. We show that the outcome of this learning process is optimal and has much better social welfare than the unique Nash equilibrium, dramatically illustrating that natural learning processes have the potential to significantly outperform equilibrium-based analysis.

Keywords: Nash Equilibria, Price of Anarchy, Learning Dynamics, Replicator Equation

1 Introduction

For the last fifty years, Nash equilibrium has been the *de facto* solution standard in game theory. From early on it was well understood that Nash equilibria, depending on the nature of the game at hand, can be rather inefficient from the perspective of social welfare. Analyzing the inefficiency of games has been a subject of extensive study in computer science, typically from the standpoint of analyzing the price of anarchy or stability: the ratio of solution quality achieved by Nash equilibria to that of the optimal solution. (See [21] for a general survey).

Nash equilibrium and its analysis, despite their prominent role, have been the subject of much criticism over the years within both economics and computer science. Nash equilibria are unlikely in general to be a realistic prediction of game outcomes: natural game play need not converge to Nash equilibria [7], it is unclear how players are expected to coordinate on a Nash equilibrium outcome in games with multiple equilibria, and even in games with unique equilibria finding a Nash equilibrium may require computation using global information about the game play, that users may not have access to. Finding Nash equilibria may also be computationally too hard in some games [6, 8].

Nevertheless, reasoning about Nash equilibria is so pervasive in algorithmic game theory that even the few papers that explicitly analyze the outcomes of natural dynamic processes in repeated games — e.g. best-response dynamics [11] or no-regret learning [3, 4, 16, 17, 20] or even specialized dynamics [2] — have focused on showing that the performance of these dynamics is comparable to that of Nash equilibria. Thus, the possibility that natural dynamic processes can lead to outcomes that are *much better* than any equilibrium of the game has gone unexplored.

In this paper we will show that ignoring this possibility can lead to qualitatively incorrect conclusions about the outcome of repeated selfish play in certain games. Specifically, we introduce a game whose unique Nash equilibrium requires rather unnatural coordination by the players, and thus equilibrium analysis may be of limited utility in understanding selfish play. In this setting, we show that various dynamic processes — including best-response dynamics and a natural on-line learning algorithm — predict a vastly better outcome than the unique Nash. Our results give the most dramatic evidence to date that the outcomes of natural learning dynamics can be superior to equilibrium outcomes, and they illustrate the potential of such approaches in providing us with insights that would be unattainable by standard Nash equilibrium analysis.

1) Game definition and results summary

We will consider an uneven variant of matching pennies played along the edges of a cycle on the players, which we call Asymmetric Cyclic Matching Pennies. There are three players numbered $1, 2, 3$, with two strategies each, H and T. The utility of player i depends only on his action and the action of player $i-1$, as shown in Figure 1 (here, and throughout the paper, player numbers are considered to be cyclical, so $0 \equiv 3$). If player i's strategy matches the strategy of player $i-1$, then i receives 0 payoff. If player i plays strategy H whereas player $i-1$ plays strategy T, then i receives a payoff of 1. Lastly, if player i plays strategy T whereas player $i-1$ plays strategy H, then i receives a payoff of $M \geq 1$. The unique Nash equilibrium of this game (when played on any odd cycle) is for all players to mix between H and T. The payoff for this Nash equilibrium is $\frac{M}{M+1} < 1$ for each player.

player i \ player $i-1$	H	T
H	0	1
T	M	0

Figure 1: The payoff matrix for player i, $i \in \{1, 2, 3\}$.

In contrast, we will consider the outcome when all three players employ a simple learning dynamics. The learning dynamic we consider is the replicator dynamics, the continuum limit of the multiplicative-weights update process as the multiplicative factor approaches 1 and time is renormalized accordingly (see [16]). Analyzing the limit of the replicator dynamics for Asymmetric Cyclic Matching Pennies is especially interesting, as the flow lines of the differential equation do not converge to a set of fixed points. In this context, we show that the social welfare of the players approaches $M+1$ (the optimum of the game), which is significantly higher than the total welfare of < 3 at the unique Nash equilibrium. This provides compelling evidence of the limitations of worst-case and equilibrium-based analysis, and shows the potential of directly analysing the outcome of natural adaptive play.

At a technical level, our analysis departs from prior work on the analysis of learning in games by deriving strong conclusions about the set of limit points of the learning process, without making use of a global potential function as in [7, 10, 16]. Our game lacks such a potential function, hence we must instead pursue a much more delicate line of attack that requires using different potential functions on different subsets of the interior of the phase space along with specialized arguments to control the system behavior near points where these potential functions become constant, namely, the Nash equilibrium and the boundary of the phase space.

2) Related work: Best-response dynamics.

Best-response dynamics is perhaps the simplest dynamic strategic update process: one at a time, players myopically shift from their current strategy to one that is a best response to the profile of opponents' strategies at that time. Goemans, Mirrokni, and Vetta, in the first computer science paper to directly analyze the outcome of natural dynamics in repeated games, introduced a randomized best-response dynamic whose stationary distributions are termed *sink equilibria*. The ratio of solution quality between the social optimum and the worst-case sink equilibrium is the *price of sinking*, and it was shown in [11] that this parameter can be vastly *greater* than the price of anarchy (even in games whose price of anarchy is 1) but that it is comparable to the price of anarchy in certain classes of games, including atomic weighted congestion games.

The question of whether sink equilibria can be dramatically *better* than all Nash equilibria was not considered in [11] or in subsequent papers on sink equilibria. Our work resolves this question in the affirmative using Asymmetric Cyclic Matching Pennies as a simple example. Analyzing the price of sinking in this game is trivial: it has a unique sink equilibrium consisting of strategy profiles at which the social welfare reaches its optimum value of $M+1$, whereas the unique Nash equilibrium has a social welfare less than 3. However, the example of best-response dynamics in the Asymmetric Cyclic Matching Pennies prompts an immediate follow-up question: "Is this good outcome the result of the extreme myopia implied by best-response dynamics, or do non-myopic dynamics also lead to the same good outcome?" The bulk of our paper is devoted to resolving that question. In so doing, we show that the good outcome predicted by best-response dynamics is a robust prediction and not a pathology of that model of selfish behavior.

3) Related work: Dynamics of no-regret learning.

In the context of worst-case outcomes, there has been significant progress in incorporating models of user behavior in evaluating the quality of outcomes in games. Blum et al. [4] introduced the *price of total anarchy*, the ratio of optimal solution quality to that achieved in the worst-case when all players use no-regret learning processes (also known as *Hannan-consistent algorithms* [12]). The regret of a player in a game after n repeated plays is the average difference between the payoff of the player and the payoff of the best single strategy in hindsight; game play has the no-regret property if the player's regret tends to 0 as n tends to infinity.[1] Modeling user behavior via no-regret learning in a repeated interaction has a long history in game

[1] In other words, a player achieving the no-regret property may switch his or her strategy in each round, but is required to do at least as well as the best single strategy would have done in hindsight.

theory, and has many advantages. The no-regret property is analogous to the notion of equilibrium (see, for example, the survey of Blum and Mansour [5]). The no-regret property can be achieved via simple and efficient strategies: examples include the weighted majority algorithm [1, 18], also known as Hedge [9], and regret matching [13]. If all players use no-regret algorithms, this results in an empirical distribution of play that converges to the coarse (weak) correlated equilibria, also known as the Hannan set [12]. The solution quality of worst-case outcomes in the Hannan set has been studied by a number of authors. Blum, Even-Dar, and Ligett [3] observed that in nonatomic congestion games, no-regret learning converges to Nash equilibria, and Blum et al. [4] and Roughgarden [20] have shown that in broad classes of games, the outcomes of any no-regret learning match the price of anarchy bound. These results focus on evaluating the *worst-case* no-regret dynamics, and hence can lead to overly pessimistic predictions when the worst case occurs on unnatural outcomes that are hard to coordinate on. Balcan, Blum and Mansour [2] consider learning models in which players adaptively decide between greedy behavior and following a proposed good but untrusted strategy, and show that in two classes of games, such a mixed strategy (when helped with good advice) can efficiently reach low-cost solutions.

Our interest is in understanding the quality of outcomes reached by players using natural learning algorithms without any outside coordination. We focus on the replicator dynamic as it is perhaps the simplest and most-studied no-regret dynamic, and is the continuum limit of one of the simplest no-regret algorithms (see for example, [15] for a simple and direct proof of the no-regret property of the replicator dynamic). Dynamical systems such as the replicator dynamic arise as the continuum limit of a discrete no-regret procedure. Such dynamical systems have been most closely studied in the context of evolutionary game theory (see the book of Hofbauer and Sigmund [14] for a summary). Restricting attention to a natural learning algorithm is consistent with our goal of modeling natural player behavior, and it is also necessary because within the class of *all* no-regret learning algorithms, one can find contrived algorithms whose distribution of play converges to an arbitrary (e.g., worst-case) correlated equilibrium of any game [19].

Piliouras, Kleinberg and Tardos [16] consider the quality of solutions reached by the multiplicative weight algorithm and its continuum limit, the replicator dynamic, in repeated atomic congestion games. They show that if players use this learning algorithm to adjust their strategies, then in almost all such games (when congestion costs are selected at random independently on each edge), game play converges to a pure Nash equilibrium. This demonstrates that such dynamics can surpass the Price of Total Anarchy and also the Price of Anarchy for mixed Nash equilibria. The analysis of [16] used the fact that congestion games have a natural potential function that serves as a Lyapunov function of the dynamic system, and hence the dynamics converge to stable fixed points (which are a subset of Nash equilibria). However, this type of analysis is rather limited, as in many games, natural learning algorithms do not converge to a stable point. Daskalakis et al. [7] show that in some settings, the cumulative distributions of players produced by multiplicative weights algorithms with different learning rates actually drift away from the equilibrium. To analyze the quality of such outcomes, it is not useful to analyze stable points; one needs to work directly with the limit set of the process.

Gaunersdorfer and Hofbauer [10] analyze the limit behavior of the replicator dynamics for a few simple games, including an even matching pennies game played on a 3-cycle (the variant of our game with $M = 1$), and show convergence to the 6-cycle of best responses. The uneven version of this game that we consider here allows for a more interesting distinction between the welfare of the unique Nash equilibrium and the limit behavior of the replicator dynamic, but is much harder to analyze. For example, it is not hard to see that the replicator dynamic is monotone increasing in the social welfare for the Cyclic Matching Pennies game studied in [10], while this is not true in our game. The convergence proof of [10] is based on a potential function showing that the replicator dynamic converges to the boundary of the feasible region, while our analysis for Asymmetric Cyclic Matching Pennies is not based on a single potential function.

2 Preliminaries

In Asymmetric Cyclic Matching Pennies, in order to express a player i's mixed strategy, it suffices to express the probability with which player i chooses strategy H; we denote that probability as x_i. Consequently, a mixed strategy profile is represented by the vector $\vec{x} = (x_1, x_2, x_3)$, or equivalently, as a point in the unit cube. We will write $u_i(\vec{x})$ for the utility of player i when all players play according to $\vec{x}$. Also, let $u_i(H, \vec{x}_{-i})$ (or $u_i(T, \vec{x}_{-i})$, respectively) denote the expected utility of player i when he deviates from x_i to pure strategy H (T, respectively), but the other two players play mixed strategies according to $\vec{x}$.

We are interested in outcomes of repeated play of this game. In particular, we will consider the outcome when all three players employ a simple learning dynamics, the replicator dynamics. The replicator dynamics is defined as

$$\dot{x}_i = x_i(u_i(H, \vec{x}_{-i}) - u_i(\vec{x})).$$

Lemma 2.1. *The replicator dynamics in Asymmetric Cyclic Matching Pennies corresponds to the following* *system of differential equations:*

$$\begin{aligned}
\dot{x}_1 &= x_1(1-x_1)\big(1-(M+1)x_3\big) &(1)\\
\dot{x}_2 &= x_2(1-x_2)\big(1-(M+1)x_1\big) &(2)\\
\dot{x}_3 &= x_3(1-x_3)\big(1-(M+1)x_2\big) &(3)
\end{aligned}$$

Proof. We will prove the statement for player 1; the other proofs proceed analogously.

The replicator equation for player 1 corresponds to $\dot{x}_1 = x_1(u_1(H, \vec{x}_{-1}) - \hat{u}_1(\vec{x}))$. Since $u_1(H, \vec{x}_{-1}) = 1 - x_3$ and $u_1(\vec{x}) = x_1(1-x_3) + M(1-x_1)x_3$, this gives us

$$\begin{aligned}
\dot{x}_1 &= x_1(u_1(H, \vec{x}_{-1}) - \hat{u}_1(\vec{x}))\\
&= x_1\big(1 - x_3 - x_1(1-x_3) - M(1-x_1)x_3\big)\\
&= x_1\big(1-x_3)(1-x_1) - M(1-x_1)x_3\big)\\
&= x_1(1-x_1)\big(1-(M+1)x_3\big). \quad \square
\end{aligned}$$

We will write $x_i(t)$ for the strategy of player i at time t. We are interested in the social welfare, defined as the sum of utilities for all players, as one measure of the quality of a mixed strategy profile.

Observation 2.2. *In Asymmetric Cyclic Matching Pennies, the social welfare for a mixed strategy profile ($\vec{x}$) is equal to*

$$\begin{aligned}
SW(\vec{x}) &= x_1(1-x_3) + M(1-x_1)x_3 +\\
&+ x_2(1-x_1) + M(1-x_2)x_1 +\\
&+ x_3(1-x_2) + M(1-x_3)x_2\\
&= (M+1)\big(x_1 + x_2 + x_3 - x_1x_2 -\\
&- x_1x_3 - x_2x_3\big)\\
&= (M+1)\big(1 - x_1x_2x_3 -\\
&- (1-x_1)(1-x_2)(1-x_3)\big).
\end{aligned}$$

The standard benchmark in equilibrium analysis is the social welfare of Nash equilibria. Here, we see that the unique Nash equilibrium has social welfare substantially lower than the optimum social welfare.

Lemma 2.3. *The unique Nash equilibrium of the Asymmetric Cyclic Matching Pennies game is $(\frac{1}{M+1}, \frac{1}{M+1}, \frac{1}{M+1})$ and has social welfare $SW = 3\frac{M}{M+1}$. The optimum social welfare is $M+1$, approximately $M/3$ times larger, and can be achieved via a correlated equilibrium.*

Proof. Note that if any player plays a pure strategy, the unique best response of the next player is to play the opposite pure strategy. Since the cycle is of odd length and hence has no pure Nash equilibria, this implies that in any equilibrium each player i must play the mixed strategy that makes the next player $i+1$ indifferent between his two strategies: play H with probability $\frac{1}{M+1}$ and T with the remaining probability $\frac{M}{M+1}$. The utility of player $i+1$ is then $\frac{M}{M+1}$ for any strategy, and hence the social welfare of the unique mixed Nash equilibrium is $3\frac{M}{M+1}$, as claimed.

The social welfare of any play is at most $M+1$, which we get by two players matching and the third one using the opposite strategy. There are a number of correlated equilibria with this high social welfare. For example, two players playing T and the third one H (where the H player is selected uniformly at random) is a correlated equilibrium. $\square$

3 Analysis

The analysis proceeds in two steps. First, in Subsection 3.1, we show that the trajectory of the replicator dynamic converges to the faces of the cube, unless started on the diagonal $x_1(0) = x_2(0) = x_3(0)$. Then, in Subsection 3.2, we show that in fact the dynamics converges to the *6-cycle* of best responses, connecting the points $(0,1,0), (1,1,0), (1,0,0), (1,0,1), (0,0,1), (0,1,1)$, which are the 6 pure strategies with the maximum social welfare of $M+1$.

First note that this will establish our claim for the high social welfare of the outcome of the replicator dynamic. More formally, let $x_i(t)$ denote the value of x_i at time t, and $SW(t)$ denote the social welfare at time t. We will show in Theorem 3.10 that $\max_i x_i(t) \to 1$ and $\min_i x_i(t) \to 0$ as $t \to \infty$.

Lemma 3.1. *If $\max_i x_i(t) \to 1$ and $\min_i x_i(t) \to 0$ as $t \to \infty$ then we also have that the social welfare $SW(t) \to M+1$ as $t \to \infty$.*

Proof. First observe that when $\min_i x_i(t) = 0$ and $\max_i x_i(t) = 1$ then any choice of the third player results in payoffs $M, 1, 0$ to the three players in some order, which is the maximum social welfare of $M+1$. Now the lemma follows as social welfare is a continuous function of the vector x. $\square$

To prove that the trajectory converges to the boundary of the cube, we use a type of potential function argument in Subsection 3.1, but need to use different arguments in different parts of the cube (the outcome space). In Theorem 3.2 we show that when the social welfare is lower than the unique Nash equilibrium, i.e., $SW(\vec{x}) \leq \frac{3M}{(M+1)}$, then social welfare is increasing. However, social welfare is not monotone increasing throughout the whole trajectory, so we need to switch to a different potential function. Theorem 3.4 shows that when social welfare is above the Nash welfare (i.e., $SW(\vec{x}) > \frac{3M}{M+1}$), then the value $\prod_i x_i^M(1-x_i)$ is decreasing. This latter function is 0 at all faces of the cube (and non-negative inside the cube). We use this two-step analysis to show that the trajectory converges to the boundary of the cube.

Next, in Subsection 3.2, we show that $\max_i x_i(t) \to 1$ and $\min_i x_i(t) \to 0$ as $t \to \infty$, establishing the claim that the trajectory converges to the 6-cycle. To do this we consider the signs of the values $x_i(t) - \frac{1}{M+1}$, that is, we consider which values are below or above the unique Nash equilibrium value. We define σ as follows: When $\sigma_{i-1}(t)$ is 0, player i is indifferent between his two strategies; when $\sigma_{i-1}(t) = 1$, player i prefers his strategy T (and hence the replicator dynamic decreases x_i); when $\sigma_{i-1}(t) = -1$, player i prefers his strategy H (and hence the replicator dynamic increases x_i). So the sign vector at time t, which we call $\boldsymbol{\sigma}(t)$, indicates the direction of change, i.e, the sign of $\dot{x}_i(t)$. To show the claimed convergence, we first we argue in Lemma A.8 that there exists a time $t \geq 0$ such that $\boldsymbol{\sigma}(t)$ contains at least one occurrence of +1 and at least one occurrence of -1. Then we consider the sequence of times t_n when one of the coordinates of $\boldsymbol{\sigma}(t)$ is 0, and show that the minimum $\min_i x_i(t_n)$ is monotone decreasing and converges to 0 as $n \to \infty$. Finally, in Theorem 3.10 we extend the analysis to the times between t_n values and also show also that the maximum $\max_j\{x_j(t)\} \to 1$ as $t \to \infty$.

3.1 Convergence to the boundary

As we have seen, we can denote any mixed strategy profile as a point in a unit cube (x_1, x_2, x_3). We will prove that as long as the initial point $\vec{x}$ is not on the main diagonal ($x_1 = x_2 = x_3$), then repeated application of the replicator dynamics in Asymmetric Cyclic Matching Pennies will converge to the boundary of the unit cube.

We wish to show that given any fully mixed starting point of the replicator dynamics off the diagonal, for any $M \geq 7$, there is a time T such that for all $t > T$, $SW(\vec{x}) > \frac{3M}{(M+1)}$. We split the analysis in two steps. First, we show that if the initial point is off the diagonal then the dynamics will escape the region with $SW(\vec{x}) \leq \frac{3M}{(M+1)}$ and will never return to it. In the second step, we show that any trajectory that stays in the region with $SW(\vec{x}) > \frac{3M}{(M+1)}$ will converge to the boundary.

Region with social welfare less than or equal to Nash

We start by showing that if $SW(\vec{x}) \leq \frac{3M}{(M+1)}$ and $\vec{x}$ is not the Nash equilibrium, then the social welfare increases. The proof of this theorem and the accompanying lemmas are deferred to the appendix, for lack of space.

Theorem 3.2. *For any fully mixed strategy profile $\vec{x}$ such that $SW(\vec{x}) \leq \frac{3M}{(M+1)}$, we have that $\frac{dSW(\vec{x})}{dt} \geq 0$. In fact, $\frac{dSW(\vec{x})}{dt} = 0$ if and only if $\vec{x}$ is the Nash equilibrium $(\frac{1}{M+1}, \frac{1}{M+1}, \frac{1}{M+1})$.*

We then complete the argument that there is a time T, such that for all $t > T$, $SW(\vec{x}) > \frac{3M}{(M+1)}$.

Theorem 3.3. *For any starting point of the replicator dynamics off the diagonal, for any $M > 5$, there is a time T, such that for all $t > T$, $SW(\vec{x}) > \frac{3M}{(M+1)}$*

Proof. Theorem 3.2 states that the social welfare is strictly increasing as long as the social welfare is less than or equal to $\frac{3M}{(M+1)}$ (unless we are at the NE). We will examine the following cases:

A)The social welfare to converges to $\frac{3M}{(M+1)}$: This implies that the replicator dynamics converges to the Nash equilibrium, since by Theorem 3.2 all other possible asymptotes have $SW > \frac{3M}{(M+1)}$. It is easy to check that the main diagonal $x = y = z$ is a invariant for the replicator dynamics. So, starting from a point off the diagonal the only way to converge to Nash is via a sequence of points all of which lie off the diagonal. However, this is impossible since the NE is a saddle point whose single attracting direction is the diagonal (1,1,1).

B)The social welfare does not converge to $\frac{3M}{(M+1)}$: In conjunction with Theorem 3.2 this implies that in finite time T, we reach a point (other than Nash) with social welfare equal to $\frac{3M}{(M+1)}$. At this point the derivative of the social welfare is strictly positive, so there exists a $\delta > 0$ such that for all $t \in (T, T+\delta)$, $SW > \frac{3M}{(M+1)}$. Now, let's assume that there exists $t' \geq T + \delta$ such that $SW \leq \frac{3M}{(M+1)}$. Since, by Theorem 3.2 the replicator dynamics is now "trapped" in the region with $SW \geq \frac{3M}{(M+1)}$, any such point has to be a local minimum of the social welfare. The only such candidate is the Nash equilibrium, but this violates our assumption. ☐

Region with social welfare greater than Nash

In the region of the strategy space where $SW(\vec{x}) > \frac{3M}{M+1}$, we will prove that $\prod_i x_i^M(1-x_i)$ is a Lyapunov function of the dynamics. Furthermore, we will show that it actually converges to 0. This, in conjunction to Theorem 3.3, implies that starting from any fully mixed strategy profile off the main diagonal, the replicator dynamics will converge to the boundary on the unit cube. For lack of space, the proof is deferred to the appendix.

Theorem 3.4. *If $SW(\vec{x}) > \frac{3M}{M+1}$ and $\vec{x}$ is not on a face of the unit cube, then $\prod_i x_i^M(1-x_i)$ is decreasing, that is, $\left(\prod_i x_i^M(1-x_i)\right)' < 0$. Furthermore, $\left(\prod_i x_i^M(1-x_i)\right)$ converges to 0.*

Corollary 3.5. *For any starting point of the replicator dynamics off the diagonal, for any $M > 5$, the dynamics converge to the boundary of the unit cube.*

3.2 Convergence to the 6-cycle

The analysis in this section also consists of a two-step argument. Having already proved convergence of the dynamics to the boundary, next we will establish that the trajectories of the dynamic indeed cycle indefinitely around (a restricted neighborhood of) the boundary. In the second step, we will utilize new potential functions to establish convergence to the 6-cycle of best responses, connecting the points $(0,1,0),(1,1,0),(1,0,0),(1,0,1),(0,0,1),(0,1,1)$, which are the 6 pure strategies with the maximum social welfare of $M+1$.

Cycling behavior

Let us consider a partition of the cube into regions based on the sign pattern of the derivatives $\dot{x}_i(t)$: one region of the cube consisting of strategies from which all three players decrease their values, another region where player 1 increases his value but the other players decrease theirs, and so on. To do so, we define at each time t a sign vector $\boldsymbol{\sigma}(t) \in \{-1,0,+1\}^{\mathbb{Z}}$ by specifying that

$$\sigma_i(t) = \operatorname{sgn}\left(x_i(t) - \tfrac{1}{M+1}\right) = -\operatorname{sgn}\left(\dot{x}_{i+1}(t)\right). \quad (4)$$

Each region of interest is then identified with its sign vector $\boldsymbol{\sigma}(t)$; notice that this partition into regions occurs along axis-parallel planes at the Nash equilibrium value ($\frac{1}{M+1}$). Our goal now is to examine the successive hitting points of the trajectory of the replicator dynamics with these planes. Specifically, we will argue that after some time t_0, these hitting points define a discrete set that partitions the trajectory into intervals of finite length. Obviously any such hitting point will have at least one coordinate $x_i(t) = \frac{1}{M+1}$. Further, the signs of the values $x_i(t) - \frac{1}{M+1}$ will be central to our proof, since they will help us characterize the nature of the cycling behavior and therefore apply the final potential argument in the second step.

We say that $\boldsymbol{\sigma}(t)$ is *mixed* if it contains at least one occurrence of +1 and at least one occurrence of -1. We say that a *zero-crossing* occurs at time t if $\boldsymbol{\sigma}(t)$ contains at least one occurrence of 0 and at least one occurrence of a nonzero sign (i.e., if it is a hitting point other than the fully mixed Nash).

Keeping in mind our goal of proving convergence to the 6-cycle, notice that each zero-crossing on the 6-cycle has exactly one coordinate equal to $\frac{1}{M+1}$ and the other two equal to 0 and 1 respectively. These points are mixed zero-crossings. So, intuitively, a minimal condition that our proof must imply is that any trajectory of the replicator dynamics is partitioned into intervals of finite length by a countable set of points which are mixed zero-crossings. We in fact prove this statement and use it as a stepping stone for our potential function arguments.

The formal analysis consists of a long sequence of technical lemmas characterizing the evolution of $\boldsymbol{\sigma}(t)$ as a function of t, and can be found in the appendix. Here, we encapsulate the main essence of these lemmas in two arguments and provide the intuition behind the proofs.

Lemma 3.6. *Unless $x_1(0) = x_2(0) = x_3(0)$, there exists a finite time $t \geq 0$ such that a mixed zero-crossing occurs at time t_0.*

Sketch. Subsection 3.1 implies that the trajectory will reach a zero-crossing (but not necessarily a mixed one) in finite time. Starting from such a point, we will argue that the replicator dynamics will indeed reach a *mixed* zero-crossing in finite time. Suppose that we reach a zero-crossing, that is, a point where some but not all players play their Nash strategy. There must be some player i playing a Nash strategy such that player $i-1$ is not playing a Nash strategy. If we take an infinitesimal step forward, then player $i-1$ is essentially pushing player i to move to a strategy even further away from his own. So, we will reach a mixed $\boldsymbol{\sigma}(t)$.

If $\boldsymbol{\sigma}(t)$ is a zero-crossing, the proof is complete. If not, we show that there exists $t_0 > t$ such that we reach a mixed zero-crossing at time t_0. This time t_0 is merely

$$t_0 = \sup\{t'' > t \mid \boldsymbol{\sigma}(u) \text{ is constant on } [t, t'')\},$$

and by continuity it can be shown to correspond to a mixed zero-crossing. The trickier part is to establish that we reach it in finite time. This is shown by bounding the derivative of a specific measure of the distance from the set of zero-crossings, away from zero. Thus, we reach a mixed zero-crossing in finite time. □

Lemma 3.6 is proven formally in the appendix as a combination of Lemmas A.6, A.7, A.8 and A.9.

We are a little less than halfway there. We have proven that we will reach one mixed zero-crossing, but now we need to argue that the trajectory visits infinitely many isolated mixed zero-crossings. The following lemma and corollary help us establish that by characterizing the set T of all $t > t_0$ such that a (mixed) zero-crossing occurs at time t. The full proof appears in the appendix.

Lemma 3.7. *If $\boldsymbol{\sigma}(t_0)$ is mixed, then $\boldsymbol{\sigma}(t)$ is mixed for all $t > t_0$. Furthermore, the set T is unbounded and has no accumulation point.*

Sketch. First, we argue by contradiction that the set T has no accumulation points. Indeed, if the set T has an accumulation point t^* then applying the Mean Value Theorem and continuity, we can show this implies that this point is the Nash equilibrium. But an application of the uniqueness theorem for first order ODE's implies that in order to reach the Nash in finite time, we need to start from the Nash equilibrium, implying a contradiction.

Next, we argue that $\boldsymbol{\sigma}(t)$ is mixed for all $t > t_0$. We suppose by way of contradiction that there exists a finite

t' such that

$$t' = \inf\{t > t_0 \mid \boldsymbol{\sigma}(u) \text{ is not mixed}\}.$$

We then perform a case analysis on $\boldsymbol{\sigma}(t')$, to see it cannot exist.

Finally, it is straightforward to show that T is unbounded: for any $t > t_0$, the set T contains a point $t' \geq t$ because either t itself is in T, or else, as we have argued in Lemma 3.6, t implies the existence of a $t' > t$ belonging to T. □

Via standard analytic arguments, we can then derive the following corollary (proof in the appendix):

Corollary 3.8. *There is an order-preserving one-to-one correspondence between the positive integers and the set T of zero-crossings occurring after t_0.*

Let us number the elements of T as $t_1, t_2, t_3, \ldots$ using the one-to-one correspondence defined in Corollary 3.8, and let $\boldsymbol{\sigma}_n$ denote $\boldsymbol{\sigma}(t_n)$ for all $n > 0$. Representing a sign vector $\boldsymbol{\sigma}$ by its three components $(\sigma_1, \sigma_2, \sigma_3)$, we see that each of the sign vectors $\boldsymbol{\sigma}_n$ is mixed, and is therefore represented by one of the ordered triples

$$\begin{array}{lll} (0,-1,+1), & (-1,0,+1), & (-1,+1,0), \\ (0,+1,-1), & (+1,0,-1), & (+1,-1,0). \end{array} \tag{5}$$

In fact, inspection of the proofs of Lemmas A.7 and A.9 reveals that the sequence $(\boldsymbol{\sigma}_n)$ cycles through these six sign patterns in the order specified above. We may assume without loss of generality (by replacing t_0 with a later time if necessary) that σ_1 has the sign pattern represented by $(0, -1, +1)$. In light of Theorem 3.2 we may also assume without loss of generality that $SW(t) > \frac{3M}{M+1}$ and $\prod_i x_i^M(1-x_i)$ is decreasing for all $t > t_0$. We are now in a position to argue about convergence of the dynamics to the 6-cycle.

Potential arguments for convergence to the 6-cycle

We define for each $i \in \mathbb{Z}$ the function

$$y_i(t) = \ln\left(\frac{x_i(t)}{1 - x_i(t)}\right) = \ln(x_i(t)) - \ln(1 - x_i(t)). \tag{6}$$

From the equation $\dot{x}_i = x_i(1-x_i)(1-(M+1)x_{i-1})$ we easily obtain

$$\dot{y}_i = \frac{\dot{x}_i}{x_i} + \frac{\dot{x}_i}{1 - x_i} = \frac{\dot{x}_i}{x_i(1 - x_i)} = 1 - (M+1)x_{i-1}. \tag{7}$$

We first prove monotonic behavior of a simple function of the player strategies (namely, $w_n = \min\{x_1(t_n), x_2(t_n), x_3(t_n)\}$), which we then use as a key step to showing convergence to the six-cycle. The proof of this lemma is deferred to the appendix for lack of space. The key insight is to break down the analysis of the cyclic behavior into odd (1,3,5,7,...) and even (2,4,6,8,...) steps. The even steps are easy because w_n is decreasing continually throughout those time intervals. In the odd steps, however, w_n increases initially but then a different player's variable becomes the minimum and w_n decreases again; despite this, we need to show that at the time of the next zero-crossing, the new w_n is smaller than the old one. This is done using a carefully constructed linear combination of y_i's that serves as a Lyapunov function during the odd-numbered interval.

Lemma 3.9. *Assuming $M \geq 7$, the sequence $w_n = \min\{x_1(t_n), x_2(t_n), x_3(t_n)\}$ is monotonically decreasing in n, and it converges to zero as $n \to \infty$.*

Finally, we arrive at our main theorem, which, in conjunction with Lemma 3.1, demonstrates that $SW(t) \to M + 1$.

Theorem 3.10. *Unless $x_1(0) = x_2(0) = x_3(0)$, for Asymmetric Cyclic Matching Pennieswith $M \geq 7$, the vector $\vec{x}(t)$ converges to the 6-cycle spanned by the off-diagonal vertices of the cube. In other words, $\min_j\{x_j(t)\} \to 0$ and $\max_j\{x_j(t)\} \to 1$ as $t \to \infty$.*

Sketch. Lemma 3.9 establishes that $w_n = \min_j\{x_j(t_n)\}$ converges to zero as $n \to \infty$. As a part of that proof, we also show that $\min_j\{x_j(t)\}$ decreases monotonically from time t_n to t_{n+1} when n is even. So, to prove that $\min_j\{x_j(t)\} \to 0$ we only need to show that $\min_j\{x_j(t)\}$ does not grow too large in the middle of an interval (t_n, t_{n+1}), when n is odd. Note that for n odd, the functions $x_n(t), x_{n+1}(t), x_{n+2}(t)$ have the following behavior on the interval (t_n, t_{n+1}): x_n starts at $\frac{1}{M+1}$ and decreases, x_{n+1} starts below $\frac{1}{M+1}$ and increases to $\frac{1}{M+1}$, x_{n+2} starts above $\frac{1}{M+1}$ and increases. Thus, the quantity $\min_j\{x_j(t)\}$ is maximized on the interval $t_n \leq t \leq t_{n+1}$ at the unique time r_n in that interval satisfying $x_n(r_n) = x_{n+1}(r_n)$. Our objective is thus to show that $x_n(r_n) \to 0$ as $n \to \infty$, which again requires a detailed case analysis.

The proof that $\max_j\{x_j(t)\} \to 1$ is similar in spirit to the one demonstrating $\min_j\{x_j(t)\} \to 0$. Once again, we break down the steps into odd and even cases and we show that linear combinations of the $y_i(t)$ can be employed as Lyapunov functions. □

Corollary 3.11. *In Asymmetric Cyclic Matching Pennieswith $M \geq 7$, so long as the initial player strategies are off-diagonal, the replicator dynamics achieves $SW(t) \to M + 1$ as $t \to \infty$.*

4 Conclusions

Despite our community's many successes in analyzing equilibria and their properties, it is important to be aware of the limitations of equilibria, particularly the limitations of their predictive power. In this paper, we have shown that in some games, natural dynamic processes can lead to outcomes that are *much better* than any equilibrium. These results underscore significant drawbacks of equilibrium-based analysis as a tool for understanding the outcomes of selfish behavior in games—limiting ourselves to equilibria as a reference point could lead us to qualitatively incorrect conclusions about system behavior.

The time has come to shift our field's perspective on games from one that attempts to cast dynamic behavior in terms of static limit points to one with more sophisticated, nuanced views and techniques. Our work, both at a conceptual and at a technical level, highlights the importance of this shift.

Acknowledgments

This work was supported by NSF grants AF-0910940, CCF-0325453, CCF-0643934, CCF-0729006, CNF-0937060, DMS-1004416, and IIS-0905467; AFOSR Project FA9550-09-1-0420; ONR grants N00014-98-1-0589 and N00014-09-1-0751; an Alfred P. Sloan Foundation Fellowship; a Microsoft Research New Faculty Fellowship; and a Yahoo! Research Alliance Grant.

A Omitted proofs

A.1 Convergence to the boundary

We begin with the following helpful lemmas:

Lemma A.1. *For any probabilities x_1, x_2, x_3 we have that $x_1x_2+x_1x_3+x_2x_3 \leq \frac{(x_1+x_2+x_3)^2}{3}$, where equality holds if and only if $x_1 = x_2 = x_3$.*

Proof. Note that $(x_1-x_2)^2+(x_1-x_3)^2+(x_2-x_3)^3 \geq 0$, with equality if and only if $x_1 = x_2 = x_3$. Multiplying out, dividing by two, and rearranging, this gives us

$$\begin{aligned} x_1x_2 + x_1x_3 + x_2x_3 &\leq x_1^2 + x_2^2 + x_3^2 \\ &= (x_1 + x_2 + x_3)^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3. \end{aligned}$$

Combining terms and dividing through by three gives the desired result. ☐

Lemma A.2. *If $x_1+x_2+x_3-x_1x_2-x_1x_3-x_2x_3 \leq \frac{3M}{(M+1)^2}$, then either $x_1+x_2+x_3 \leq \frac{3}{M+1}$ or $x_1+x_2+x_3 \geq \frac{3M}{M+1}$.*

Proof. Given the assumption and an application of Lemma A.1, we get

$$\begin{aligned} \frac{3M}{(M+1)^2} &\geq x_1 + x_2 + x_3 - x_1x_2 - x_1x_3 - x_2x_3 \\ &\geq x_1 + x_2 + x_3 - \frac{(x_1 + x_2 + x_3)^2}{3}. \end{aligned}$$

Denoting $x_1 + x_2 + x_3$ by a, this may be rewritten as $a - \frac{a^2}{3} \leq \frac{3M}{(M+1)^2}$. Solving this equation, we see that either $a \leq \frac{3}{M+1}$ or $a \geq \frac{3M}{M+1}$, as desired. ☐

Lemma A.3. *If $x_1 + x_2 + x_3 \leq \frac{3}{M+1}$, then $x_1 + x_2 + x_3 \geq (M+1)(x_1x_2+x_1x_3+x_2x_3)$, where the equality holds if and only if $x_1 = x_2 = x_3 = \frac{1}{M+1}$ or $x_1 = x_2 = x_3 = 0$.*

Proof. By application of Lemma A.1 and the assumption, we get

$$\begin{aligned} 3(x_1x_2 + x_1x_3 + x_2x_3) &\leq (x_1 + x_2 + x_3)^2 \\ &\leq \frac{3}{M+1}(x_1 + x_2 + x_3). \end{aligned}$$

For the first inequality to holds as equality, it must be the case the $x_1 = x_2 = x_3$ (Lemma A.1). For the second, it must be the case that either $x_1 + x_2 + x_3 = 0$ or $x_1 + x_2 + x_3 = \frac{3}{M+1}$. Combining these two requirements, implies that either $x_1 = x_2 = x_3 = \frac{1}{M+1}$ or $x_1 = x_2 = x_3 = 0$, as desired. ☐

Theorem A.4 (Theorem 3.2)**.** *For any fully mixed strategy profile $\vec{x}$ such that $SW(\vec{x}) \leq \frac{3M}{(M+1)}$, we have that $\frac{dSW(\vec{x})}{dt} \geq 0$. In fact, $\frac{dSW(\vec{x})}{dt} = 0$ if and only if $\vec{x}$ is the Nash equilibrium $(\frac{1}{M+1}, \frac{1}{M+1}, \frac{1}{M+1})$.*

Proof. By Lemma A.2 and the hypothesis that $SW(\vec{x}) \leq \frac{3M}{(M+1)}$, we have that either $x_1 + x_2 + x_3 \leq \frac{3}{M+1}$ or $x_1 + x_2 + x_3 \geq \frac{3M}{M+1}$.

We begin with the case when $x_1 + x_2 + x_3 \geq \frac{3M}{M+1}$. For $M > 5$, we have that $x_1 + x_2 + x_3 > 5/2$, which implies that $x_1, x_2, x_3 > 1/2$. We have that

$$SW(\vec{x}) = (M+1)\left(x_1(1 - x_3) + x_2(1 - x_1) + x_3(1 - x_2)\right).$$

Hence, we can derive that

$$\begin{aligned} &\frac{dSW(\vec{x})}{dt} = \\ &= (M+1)[\dot{x}_1(1 - x_3) - \dot{x}_3x_1 + \dot{x}_2(1 - x_1) - \\ &\quad - \dot{x}_1x_2 + \dot{x}_3(1 - x_2) - \dot{x}_2x_3] \\ &= (M+1)(\dot{x}_1(1 - x_2 - x_3) + \dot{x}_2(1 - x_1 - x_3) + \\ &\quad + \dot{x}_3(1 - x_1 - x_2)) \\ &= (M+1)\big[x_1(1 - x_1)\,(1 - (M+1)x_3)\,(1 - x_2 - x_3) \\ &\quad + x_2(1 - x_2)\,(1 - (M+1)x_1)\,(1 - x_1 - x_3) \\ &\quad + x_3(1 - x_3)\,(1 - (M+1)x_2)\,(1 - x_1 - x_2)\big]. \end{aligned}$$

The last summation is strictly greater than zero, since if $1/2 < x_1, x_2, x_3 < 1$ and $M > 5$, it is straightforward to show that all summands are strictly positive.

Next we will consider the second case, where $x_1 + x_2 + x_3 \leq \frac{3}{M+1}$. Here, we will consider the equivalent definition of social welfare derived in Observation 2.2, that $SW(\vec{x}) = (M +$

$1)\,(1-x_1x_2x_3-(1-x_1)(1-x_2)(1-x_3))$. Specifically, it suffices to show that under the theorem hypothesis, the function $x_1x_2x_3+(1-x_1)(1-x_2)(1-x_3)$ decreases. We have that

$$\begin{aligned}
&\Big(x_1x_2x_3+(1-x_1)(1-x_2)(1-x_3)\Big)' =\\
&= x_1x_2x_3\cdot\\
&\quad\Big((1-x_1)(1-(M+1)x_3)+\\
&\quad\quad+(1-x_2)(1-(M+1)x_1)+\\
&\quad\quad+(1-x_3)(1-(M+1)x_2)\Big)\\
&\quad+(1-x_1)(1-x_2)(1-x_3)\cdot\\
&\quad\Big(-x_1(1-(M+1)x_3)-\\
&\quad\quad-x_2(1-(M+1)x_1)-\\
&\quad\quad-x_3(1-(M+1)x_2)\Big)\\
&= x_1x_2x_3\cdot\\
&\quad\Big(3-(M+2)(x_1+x_2+x_3)+\\
&\quad\quad+(M+1)(x_1x_2+x_1x_3+x_2x_3)\Big)\\
&\quad+(1-x_1)(1-x_2)(1-x_3)\cdot\\
&\quad\Big(-(x_1+x_2+x_3)+\\
&\quad\quad+(M+1)(x_1x_2+x_1x_3+x_2x_3)\Big)\\
&\le x_1x_2x_3\Big(3-(M+1)(x_1+x_2+x_3)\Big)\\
&\quad+(1-x_1)(1-x_2)(1-x_3)\cdot\\
&\quad\Big(-(x_1+x_2+x_3)+\\
&\quad\quad+(M+1)(x_1x_2+x_1x_3+x_2x_3)\Big)
\end{aligned}$$

by Lemma A.3. Distributing terms, we see this is equal to

$$\begin{aligned}
=&\frac{3x_1x_2x_3}{x_1+x_2+x_3}\cdot\\
&\Big((x_1+x_2+x_3)-\frac{M+1}{3}(x_1+x_2+x_3)^2\Big)\\
&-(1-x_1)(1-x_2)(1-x_3)\cdot\\
&\Big((x_1+x_2+x_3)-(M+1)(x_1x_2+x_1x_3+x_2x_3)\Big)
\end{aligned}$$

The last line by Lemma A.1 is at most

$$\begin{aligned}
\le&\left(\frac{3x_1x_2x_3}{x_1+x_2+x_3}-(1-x_1)(1-x_2)(1-x_3)\right)\cdot\\
&((x_1+x_2+x_3)-(M+1)(x_1x_2+x_1x_3+x_2x_3))\,.
\end{aligned}$$

In Lemma A.3 we argued that $(x_1+x_2+x_3)-(M+1)(x_1x_2+x_1x_3+x_2x_3)\ge 0$. Here, we will show that for $M>5$, $\frac{3x_1x_2x_3}{x_1+x_2+x_3}<(1-x_1)(1-x_2)(1-x_3)$. By hypothesis, we have that $x_1+x_2+x_3\le\frac{3}{M+1}$. Without loss of generality let us assume that $x_1\ge x_2\ge x_3$. This implies that $x_1\le\frac{3}{M+1}, x_2\le\frac{3}{2(M+1)}$ and $x_3\le\frac{1}{M+1}$. As a result,

$$\begin{aligned}
\frac{3x_1x_2x_3}{x_1+x_2+x_3} &\le 3x_2x_3\\
&\le \frac{9}{2(M+1)^2}\\
&< \frac{(M-2)^3}{(M+1)}\cdot\frac{1}{(M+1)^2},
\end{aligned}$$

by our assumption that $M>5$. Finally, since $(1-x_1)\ge\frac{M-2}{M+1}$, $(1-x_2)\ge\frac{2M-1}{2(M+1)}$,, and $x_3\ge\frac{M}{M+1}$, we know that

$$(1-x_1)(1-x_2)(1-x_3)\ge\frac{M-2}{M+1}\cdot\frac{2M-1}{2(M+1)}\cdot\frac{M}{M+1},$$

and each of these terms is at least $\frac{M-2}{M+1}$.

In order to have $\big(x_1x_2x_3+(1-x_1)(1-x_2)(1-x_3)\big)'=0$, it must be the case that the inequalities in Lemmas A.1 and A.3 holds as equalities, but this happens only if $\vec{x}$ is the Nash equilibrium $(\frac{1}{M+1},\frac{1}{M+1},\frac{1}{M+1})$. $\square$

Theorem A.5 (Theorem 3.4)**.** *If $SW(\vec{x})>\frac{3M}{M+1}$ and $\vec{x}$ is not on a face of the unit cube, then $\prod_i x_i^M(1-x_i)$ is decreasing, that is, $\big(\prod_i x_i^M(1-x_i)\big)'<0$. Furthermore, $\big(\prod_i x_i^M(1-x_i)\big)$ converges to 0.*

Proof. Consider $\frac{d}{dt}\left(\prod_i x_i^M(1-x_i)\right)$.

$$\begin{aligned}
&\left(\prod_i x_i^M(1-x_i)\right)'\\
&\quad=\sum_i\left(x_i^M(1-x_i)\right)'\prod_{j\neq i}x_j^M(1-x_j)\\
&\quad=\sum_i\left(\dot{x}_i(Mx_i^{M-1}(1-x_i)-x_i^M)\right)\prod_{j\neq i}x_j^M(1-x_j)\\
&\quad=\sum_i\left(\dot{x}_ix_i^{M-1}\left(M-(M+1)x_i\right)\right)\prod_{j\neq i}x_j^M(1-x_j)\\
&\quad=\sum_i\left(x_i(1-x_i)(1-(M+1)x_{i-1})x_i^{M-1}(M-(M+1)x_i)\right)\\
&\quad\quad\cdot\prod_{j\neq i}x_j^M(1-x_j),
\end{aligned}$$

where the last line is by the definition of $\dot{x}_i$. Then, rearranging, we get

$$\begin{aligned}
&\left(\prod_i x_i^M(1-x_i)\right)' = \\
&= \sum_i \left(x_i^M(1-x_i)(1-(M+1)x_{i-1})\cdot\right. \\
&\qquad \left.(M-(M+1)x_i)\right)\prod_{j\neq i} x_j^M(1-x_j) \\
&= \left(\sum_i (1-(M+1)x_{i-1})(M-(M+1)x_i)\right)\cdot \\
&\qquad \left(\prod_i x_i^M(1-x_i)\right) \\
&= (M+1)^2\left(\prod_i x_i^M(1-x_i)\right)\cdot \\
&\qquad \left(\sum_i\left(\frac{1}{M+1}-x_{i-1}\right)\left(\frac{M}{M+1}-x_i\right)\right) \\
&= (M+1)\left(\prod_i x_i^M(1-x_i)\right)\left(\frac{3M}{(M+1)}-SW(\vec{x})\right)
\end{aligned}$$

where the last equality follows from the formulation of social welfare we derived in Observation 2.2. By assumption, $(M+1)$ is strictly positive and the final term is strictly negative; the middle term is strictly positive so long as no component of $\vec{x}$ is 0 or 1.

Since $\Pi_i x_i^M(1-x_i) \geq 0$, the process will converge to an asymptote with the property that $\left(\prod_i x_i^M(1-x_i)\right)' = 0$. However, this implies that either $SW = \frac{3M}{M+1}$ or that $\Pi_i x_i^M(1-x_i) = 0$. The first one is impossible by our assumption about social welfare, whereas the second one implies that the process converges to the boundary. □

A.2 Convergence to the 6-cycle

In this section, for notational convenience, we extend the sequence of functions $x_1(t), x_2(t), x_3(t)$ to a doubly infinite sequence of functions $\ldots, x_{-1}(t), x_0(t), x_1(t), \ldots$ with period 3; in other words, $x_{i+3}(t) = x_i(t)$ for all i, t.

The first lemma proves that if the dynamics reaches the unique Nash equilibrium, it will never leave it.

Lemma A.6. *If there exists a time t_0 such that $\boldsymbol{\sigma}(t_0)$ is the zero vector, then for all t, $\boldsymbol{\sigma}(t)$ is the zero vector.*

Proof. The hypothesis of the lemma is equivalent to the assertion that $x_1(t_0) = x_2(t_0) = x_3(t_0) = \frac{1}{M+1}$. The uniqueness theorem for first-order ODE's implies that the differential equation (1)-(3) has a unique solution satisfying $x_1(t_0) = x_2(t_0) = x_3(t_0) = \frac{1}{M+1}$, namely the constant solution in which $x_j(t) = \frac{1}{M+1}$ for all j, t. This implies $\boldsymbol{\sigma}(t)$ is the zero vector for all t. □

The second lemma gives us a handle on situations where some, but not all, of the players play their equilibrium strategies, and the situations where this can happen.

Lemma A.7. *If a zero-crossing occurs at time t, then there exists an open interval $I = (t-\delta, t+\delta)$ containing t such that t is the only zero-crossing in I, and $\boldsymbol{\sigma}(u)$ is mixed for every $u \in (t, t+\delta)$.*

Proof. Let j be an index such that $\sigma_j(t) = 0$ and $\sigma_{j-1}(t) \neq 0$. By the continuity of x_{j-1} it follows that there is an open interval $I = (t-\delta, t+\delta)$ containing t such that $x_{j-1}(u) - \frac{1}{M+1}$ has constant sign throughout I; if $\sigma_{j+1}(t) \neq 0$ then we may likewise assume $x_{j+1}(u) - \frac{1}{M+1}$ has constant sign throughout I. To prove that t is the only zero-crossing in I, — i.e. that $\sigma_j(u), \sigma_{j+1}(u)$ are nonzero on $I \setminus \{t\}$ — we argue by contradiction. If there were to exist $u \in I \setminus \{t\}$ such that $\sigma_j(u) = \sigma_j(t) = 0$, it would imply that $x_j(u) = x_j(t)$. By the Mean Value Theorem, that would imply $\dot{x}_j(t') = 0$ for some t' lying strictly between t and u. This is impossible, since $\operatorname{sgn}(\dot{x}_j(t')) = -\sigma_{j-1}(t')$, which is nonzero by construction. As for the possibility that $\sigma_{j+1}(u) = 0$ for $u \in I \setminus \{t\}$, this can be excluded by a two-case argument. If $\sigma_{j+1}(t) \neq 0$ then by construction we have chosen I such that $\sigma_{j+1}(u) = \sigma_{j+1}(t)$ for all $u \in I$. If $\sigma_{j+1}(t) = 0 = \sigma_{j+1}(u)$, then another application of the Mean Value Theorem implies the existence of a time t' lying strictly between t and u such that $\sigma_j(t') = 0$, contradicting the fact that σ_j is nonzero on $I \setminus \{t\}$.

Finally, the fact that $x_j(t) - \frac{1}{M+1} = 0$ and $\operatorname{sgn}(\dot{x}_j) = -\sigma_{j-1}(t)$ implies that $\operatorname{sgn}\left(x_j(u) - \frac{1}{M+1}\right) = -\sigma_{j-1}(t)$ for all $u \in (t, t+\delta)$. As $\sigma_{j-1}(u) = \sigma_{j-1}(t)$ for all such u, we may conclude that $\boldsymbol{\sigma}(u)$ is mixed for all such u, as claimed. □

The third lemma asserts that as long as the initial point is off the diagonal, the dynamics will at some point reach a point where some player biases more towards H and another more towards T than at the unique equilibrium.

Lemma A.8. *Unless $x_1(0) = x_2(0) = x_3(0)$, there exists a time $t \geq 0$ such that $\boldsymbol{\sigma}(t)$ is mixed.*

Proof. If $\boldsymbol{\sigma}(0)$ is mixed, there is nothing to prove. If there is a zero-crossing at any time $t \geq 0$ then Lemma A.7 implies that $\boldsymbol{\sigma}(u)$ is mixed for all u such that $u - t$ is a sufficiently small positive number. If there exists $t \geq 0$ such that $\boldsymbol{\sigma}(t)$ is the zero vector, then Lemma A.6 implies that $\boldsymbol{\sigma}(0) = \vec{0}$, violating the hypothesis of the lemma.

It remains for us to exclude the cases that $\sigma_j(t) = -1$ for all $j, t \geq 0$ or that $\sigma_j(t) = +1$ for all $j, t \geq 0$. Let $a = \min_j\{x_j(0)\}$ and $b = \max_j\{x_j(0)\}$. When $\sigma_j(t) = -1$ for all $j, t \geq 0$, it means that $x_j(t)$ is less than $\frac{1}{M+1}$ and increasing — hence remains in the interval $\left[a, \frac{1}{M+1}\right]$ — for all j. Similarly, when $\sigma_j(t) = +1$ for all j, it means that $x_j(t)$ is greater than $\frac{1}{M+1}$ and decreasing — hence remains in the interval $\left[\frac{1}{M+1}, b\right]$ for all j. In both cases, this contradicts the fact that $\vec{x}(t)$ converges to the boundary of the cube when $\vec{x}(0)$ lies off the diagonal. □

We then see that such a situation will cause one of the players to pass through a point where he employs the equilibrium strategy.

Lemma A.9. *If $\boldsymbol{\sigma}(t)$ is mixed and there is no zero-crossing at t, then there is a zero-crossing at some time $t' > t$ such that $\boldsymbol{\sigma}(t')$ is mixed and the function $\boldsymbol{\sigma}(u)$ is constant on the half-open interval $[t, t')$.*

Proof. Since there is no zero-crossing at t, the continuity of the functions x_i implies that there is an open interval containing t on which the function $\boldsymbol{\sigma}(u)$ is constant. Hence, if we define t' by

$$t' = \sup\{t'' > t \,|\, \boldsymbol{\sigma}(u) \text{ is constant on } [t, t'')\},$$

then $t' > t$ and $\boldsymbol{\sigma}(u)$ is constant on $[t, t')$. Furthermore, if t' is finite, then it follows that at least one component of $\boldsymbol{\sigma}(t')$ is zero because the other two possibilities (that $\boldsymbol{\sigma}(t') = \boldsymbol{\sigma}(t)$ or that the two sign vectors differ by reversing a nonzero sign) both violate continuity. Lemma A.6 and our hypothesis that $\boldsymbol{\sigma}(t)$ is mixed preclude the possibility that $\boldsymbol{\sigma}(t') = \vec{0}$. Hence there is a zero-crossing at t'.

To prove that t' is finite, we will show that on the interval (t, t'), the distance from $x_i(u)$ to $\frac{1}{M+1}$ is monotonically increasing in u for exactly two indices $i \in \{1, 2, 3\}$, and for the remaining value of i the distance from $x_i(u)$ to $\frac{1}{M+1}$ is monotonically decreasing at a rate bounded away from zero. Recall that the derivative $\dot{x}_i(u)$ has sign $-\sigma_{i-1}(u)$. Hence the distance from x_i to $\frac{1}{M+1}$ is increasing at time u if and only $\sigma_i(u) = -\sigma_{i-1}(u)$. The equation

$$(\sigma_0\sigma_1)(\sigma_1\sigma_2)(\sigma_2\sigma_3) = \sigma_1^2\sigma_2^2\sigma_3^2 = 1$$

implies that the relation $\sigma_i(u) = -\sigma_{i-1}(u)$ is satisfied by an even number of indices $i \in \{1, 2, 3\}$, and this number must be exactly 2 since $\boldsymbol{\sigma}(t)$ is mixed. Letting i denote the unique index in $\{1, 2, 3\}$ such that $\sigma_i(u) = \sigma_{i-1}(u)$ for all $u \in (t, t')$, we know that $|\dot{x}_i(u)| = x_i(u)(1 - x_i(u))|1 - (M + 1)x_{i-1}(u)|$. Having already established that the function $|1 - (M + 1)x_{i-1}(u)| = (M + 1)\left|x_{i-1}(u) - \frac{1}{M+1}\right|$ is monotonically increasing on (t, t'), we see that

$$\begin{aligned}|\dot{x}_i(u)| &= x_i(u)(1 - x_i(u))|1 - (M + 1)x_{i-1}(t)| \\ &\geq \min\left\{x_i(t)(1 - x_i(t)), \tfrac{M}{(M+1)^2}\right\} \cdot \\ &\quad |1 - (M + 1)x_{i-1}(t)|, \qquad (8)\end{aligned}$$

where the second inequality is justified by the fact that $x_i(u)$ lies strictly between $x_i(t)$ and $\frac{1}{M+1}$ for all $u \in (t, t')$, and the function $x(1 - x)$, being concave, assumes its minimum value on this interval at one of the endpoints. Equation (8) establishes that the rate of decrease of $|x_i(u) - \frac{1}{M+1}|$ is bounded away from zero on the interval (t, t'), and consequently t' is finite.

Finally, we must show that $\boldsymbol{\sigma}(t')$ is mixed. Since $|x_{i-1} - \frac{1}{M+1}|$ and $|x_{i+1} - \frac{1}{M+1}|$ are monotonically increasing on (t, t'), it follows that $\sigma_{i-1}(t') = \sigma_{i-1}(t)$ and $\sigma_{i+1}(t') = \sigma_{i+1}(t)$. Our choice of i ensures that $\sigma_{i-1}(t)\sigma_{i+1}(t) = -1$, so $\sigma_{i-1}(t')\sigma_{i+1}(t') = -1$ as well, establishing that $\boldsymbol{\sigma}(t')$ is mixed. □

Lemma A.10 (Lemma 3.7)**.** *If $\boldsymbol{\sigma}(t_0)$ is mixed, then $\boldsymbol{\sigma}(t)$ is mixed for all $t > t_0$. Furthermore, the set T is unbounded and has no accumulation point.*

Proof. If the set T has an accumulation point t^*, then there is some i such that for all $\delta > 0$, the relation $x_i(t) - \frac{1}{M+1} = 0$ holds infinitely often in the interval $(t^* - \delta, t^* + \delta)$. The Mean Value Theorem implies that $\dot{x}_i(t) = 0$ infinitely often in the interval $(t^* - \delta, t^* + \delta)$, hence $x_{i-1}(t) - \frac{1}{M+1} = 0$ infinitely often in that interval. Applying the Mean Value Theorem once more we see that $x_{i-2}(t) - \frac{1}{M+1} = 0$ infinitely often in $(t^* - \delta, t^* + \delta)$ as well. By continuity, we may conclude that $(x_{i-2}(t^*), x_{i-1}(t^*), x_i(t^*)) = \left(\frac{1}{M+1}, \frac{1}{M+1}, \frac{1}{M+1}\right)$. Lemma A.6 now implies that $\boldsymbol{\sigma}(t) = \vec{0}$ for all t, violating our assumption that $\boldsymbol{\sigma}(t_0)$ is mixed.

Consequently, T has no accumulation point. Our next objective is to prove that $\boldsymbol{\sigma}(t)$ is mixed for all $t > t_0$. Assume, by way of contradiction, that $\{t > t_0 \,|\, \boldsymbol{\sigma}(t) \text{ is not mixed}\}$ is nonempty, and let t' denote the infimum of this set. If there is no zero-crossing at t' then continuity implies that $\boldsymbol{\sigma}(u)$ is constant for u in an open interval containing t', but this violates our definition of t'. Consequently, we may assume $t' \in T$. As T has no accumulation point, there is a positive ε such that the interval $(t' - \varepsilon, t')$ contains no zero-crossings. By our definition of t', we know that $\boldsymbol{\sigma}(t' - \varepsilon/2)$ is mixed, and Lemma A.9 now implies that $\boldsymbol{\sigma}(t')$ is mixed. Then Lemma A.7 implies that $\boldsymbol{\sigma}(u)$ is mixed for all u in an open interval $(t' - \delta, t' + \delta)$, contradicting our definition of t'.

Finally, it is easy to show that T is unbounded: for any $t > t_0$, the set T contains a point $t' \geq t$ because either t itself is in T, or else Lemma A.9 implies the existence of a $t' > t$ belonging to T. □

Corollary A.11 (Corollary 3.8)**.** *There is an order-preserving one-to-one correspondence between the positive integers and the set T of zero-crossings occurring after t_0.*

Proof. For each $t \in T$, let $n(t)$ denote the cardinality of the set $T \cap [t_0, t]$. We know that $T \cap [t_0, t]$ is finite because any infinite subset of $[t_0, t]$ has an accumulation point whereas T does not. Thus, $t \mapsto n(t)$ defines a function from T to the positive integers. It is clearly one-to-one and order-preserving because if $s < t$ are elements of T then $T \cap [t_0, t]$ has at least one more element than $T \cap [t_0, s]$, namely the element t. Finally, to show that every positive integer is equal to $n(t)$ for some $t \in T$, we argue as follows. If t is any element of T then $\{s \in T \mid s > t\}$ is nonempty since T is unbounded. Letting t' denote the infimum of this set, we have $t' \in T$ (as otherwise t' would be an accumulation point of T) and t' is, by construction, the unique element of $T \cap [t_0, t']$ that does not belong to $[t_0, t]$. It follows that $n(t') = n(t) + 1$. A similar argument establishes that $\inf(T)$ is an element of T and that $n(\inf(T)) = 1$. Hence, the image of the function $t \mapsto n(t)$ contains 1 and is closed under the successor operation, implying that it contains every positive integer. ☐

Lemma A.12 (Lemma 3.9)**.** *Assuming $M \geq 7$, the sequence $w_n = \min\{x_1(t_n), x_2(t_n), x_3(t_n)\}$ is monotonically decreasing in n, and it converges to zero as $n \to \infty$.*

Proof. We initially focus on proving monotonicity, deferring convergence to zero until the final paragraph of the proof.

If n is even, inspection of the sign patterns in (5) reveals that $w_n = x_{n-1}(t_n)$ and $w_{n+1} = x_{n-1}(t_{n+1})$. Furthermore, the sign of $\dot{x}_{n-1}(t)$ for $t \in (t_n, t_{n+1})$ is given by $-\sigma_{n-2}(t) = -\sigma_{n+1}(t)$, which is seen to be negative again by inspecting the sign patterns in (5). It follows that $w_{n+1} < w_n$ when n is even.

If n is odd, assume without loss of generality that $n \equiv 1 \pmod 6$, so that $\sigma_1(t_n) = 0$ and $x_1(t_n) = \frac{1}{M+1}$; the other two cases $n \equiv 3, 5 \pmod 6$ are the same up to cyclic symmetry. We have $w_n = x_2(t_n)$ and $w_{n+1} = x_1(t_{n+1})$, so we must show that $x_1(t_{n+1}) < x_2(t_n)$. Recall the functions $y_i(t)$ defined in (6) and their derivatives given by (7). We have

$$\begin{aligned} \frac{d}{dt}(y_1 + 2y_2) &= 3 - (M+1)(2x_1 + x_3) \\ &< 3 - (M+1)x_3. \end{aligned} \tag{9}$$

We now divide the argument into three cases.

Case 1: $x_3(t_n) \geq \frac{3}{M+1}$**.**
Since x_3 is increasing on the interval (t_n, t_{n+1}), the right side of (9) is negative for $t \in (t_n, t_{n+1})$, from which we may conclude that $y_1(t_{n+1}) + 2y_2(t_{n+1}) < y_1(t_n) + 2y_2(t_n)$. We have $y_1(t_n) = y_2(t_{n+1}) = \ln(1/M)$, and upon substituting these values we find that

$$y_1(t_{n+1}) < 2y_2(t_n) + \ln(M) < y_2(t_n) + \ln\left(\tfrac{M}{M+1}\right)$$

which implies $x_1(t_{n+1}) < x_2(t_n)$ since $x \mapsto \ln(x) - \ln(1-x)$ is a monotonically increasing function of x.

Case 2: $x_3(t_{n+1}) \leq \frac{M}{M+1}$**.**
During the time interval (t_n, t_{n+1}), x_3 increases while remaining bounded above by $\frac{M}{M+1}$. The function $x^M(1-x)$ is monotonically increasing on the interval $\left(0, \frac{M}{M+1}\right)$, hence

$$x_3(t_{n+1})^M(1 - x_3(t_{n+1})) > x_3(t_n)^M(1 - x_3(t_n)).$$

On the other hand, we have assumed that $\prod_i x_i^M(1-x_i)$ is monotonically decreasing for all $t > t_0$, so we may conclude that

$$\begin{aligned} &x_1(t_{n+1})^M(1 - x_1(t_{n+1}))x_2(t_{n+1})^M(1 - x_2(t_{n+1})) \\ &\quad < x_1(t_n)^M(1 - x_1(t_n))x_2(t_n)^M(1 - x_2(t_n)). \end{aligned}$$

Again using the fact that $x_1(t_n) = x_2(t_{n+1}) = \ln(1/M)$ to cancel terms from both sides, we obtain $x_1(t_{n+1})^M(1 - x_1(t_{n+1})) < x_2(t_n)^M(1 - x_2(t_n))$, which implies $x_1(t_{n+1}) < x_2(t_n)$ using the monotonicity of $x^M(1-x)$ on the interval $\left(0, \frac{M}{M+1}\right)$.

Case 3: $x_3(t_n) < \frac{3}{M+1}$ **and** $x_3(t_{n+1}) > \frac{M}{M+1}$**.**
By continuity, we know there are times $u, v \in (t_n, t_{n+1})$ such that $x_3(u) = \frac{3}{M+1}$, $x_3(v) = \frac{M-2}{M+1}$. Now,

$$\begin{aligned} y_3(u) - y_3(t_n) &< \ln\left(\frac{3}{M-2}\right) - \ln\left(\frac{1}{M}\right) \\ &= \ln(M) - \ln\left(\frac{M-2}{3}\right) \\ &< y_3(t_{n+1}) - y_3(v) \end{aligned}$$

and

$$\begin{aligned} &\int_{t_n}^{u} 1 - (M+1)x_2(t)\,dt \\ &\quad < \int_{v}^{t_{n+1}} 1 - (M+1)x_2(t)\,dt, \end{aligned} \tag{10}$$

where the final line follows from the fact that $\dot{y}_3 = 1 - (M+1)x_2$. Now, x_2 is increasing on the interval (t_n, t_{n+1}) so the integrand $1 - (M+1)x_2$ is decreasing. Letting $c = 1 - (M+1)x_2(u)$, this means that the left side of (10) is greater than $(u - t_n)c$ while the right side is less than $(t_{n+1} - v)c$, hence (10) implies

$t_{n+1} - v > u - t_n$. Now, using (9),

$$\begin{aligned}
&y_1(t_{n+1}) + 2y_2(t_{n+1}) - y_1(t_n) - 2y_2(t_n) \\
&\quad < \int_{t_n}^{t_{n+1}} 3 - (M+1)x_3(t)\,dt \\
&\quad = \int_{t_n}^{u} 3 - (M+1)x_3(t)\,dt \\
&\qquad + \int_{u}^{v} 3 - (M+1)x_3(t)\,dt \\
&\qquad + \int_{v}^{t_{n+1}} 3 - (M+1)x_3(t)\,dt \\
&\quad < \int_{t_n}^{u} 2\,dt + \int_{u}^{v} 0\,dt + \int_{v}^{t_{n+1}} (-2)\,dt \\
&\quad = 2(u - t_n - t_{n+1} + v) \\
&\quad < 0,
\end{aligned}$$

where we have used the fact that $3 - (M+1)x_3(t)$ is bounded above by 2 on the interval $[t_n, u]$, by 0 on the interval $[u, v]$, and by -2 on the interval $[v, t_{n+1}]$. (This last upper bound is the one that requires $M \geq 7$.) Having thus established that $y_1(t_{n+1}) + 2y_2(t_{n+1}) < y_1(t_n) + 2y_2(t_n)$ we finish the argument, as in Case 1, by using the fact that $y_2(t_{n+1}) = y_1(t_n) = \ln(1/M)$ to conclude that $y_1(t_{n+1}) < 2y_2(t_n) + \ln(M) < y_2(t_n) + \left(\frac{M}{M+1}\right)$ and consequently $x_1(t_{n+1}) < x_2(t_n)$.

Finally, we prove that $w_n \to 0$ as $n \to \infty$ or, equivalently, that $\ln(w_n) - \ln(1 - w_n) \to -\infty$. Recalling that the function $x \mapsto \ln(x) - \ln(1-x)$ is monotonically increasing in x, we see that the sequence $q_n = \ln(w_n) - \ln(1 - w_n)$ is monotonically decreasing. Furthermore, the proofs given above in Cases 1 and 3 above have shown that $q_{n+1} < q_n - \ln\left(\frac{M}{M+1}\right)$ for every n satisfying the hypotheses of one of those cases. If there are infinitely many such n, then $q_n \to -\infty$ and we are done. Otherwise, there are infinitely many n satisfying the hypotheses of Case 2; let S denote the set of all such n. Considering that $\vec{x}(t)$ converges to the boundary of the cube as $t \to \infty$ and that the sequence $(t_n)_{n \in S}$ is unbounded, we see that the sequence $(\vec{x}(t_n))_{n \in S}$ converges to the boundary of the cube. However, for every $n \in S$ we have $\max_j\{x_j(t_n)\} = x_3(t_n) \leq x_3(t_{n+1}) \leq \frac{M}{M+1}$. So the only way $(\vec{x}(t_n))_{n \in S}$ could converge to the boundary of the cube is if $w_n = \min_j\{x_j(t_n)\}$ converges to 0, as claimed. □

Theorem A.13 (Lemma 3.10). *Unless $x_1(0) = x_2(0) = x_3(0)$, the vector $\vec{x}(t)$ converges to the 6-cycle spanned by the off-diagonal vertices of the cube. In other words, $\min_j\{x_j(t)\} \to 0$ and $\max_j\{x_j(t)\} \to 1$ as $t \to \infty$.*

Proof. We have already seen that $w_n = \min_j\{x_j(t_n)\}$ converges to zero as $n \to \infty$, and that $\min_j\{x_j(t)\}$ decreases monotonically from time t_n to t_{n+1} when n is even, so to prove that $\min_j\{x_j(t)\} \to 0$ we only need to show that $\min_j\{x_j(t)\}$ does not grow too large in the middle of an interval (t_n, t_{n+1}), when n is odd. Note that for n odd, the functions $x_n(t), x_{n+1}(t), x_{n+2}(t)$ have the following behavior on the interval (t_n, t_{n+1}): x_n starts at $\frac{1}{M+1}$ and decreases, x_{n+1} starts below $\frac{1}{M+1}$ and increases to $\frac{1}{M+1}$, x_{n+2} starts above $\frac{1}{M+1}$ and increases. Thus, the quantity $\min_j\{x_j(t)\}$ is maximized on the interval $t_n \leq t \leq t_{n+1}$ at the unique time r_n in that interval satisfying $x_n(r_n) = x_{n+1}(r_n)$. Our objective is thus to show that $x_n(r_n) \to 0$ as $n \to \infty$.

As before, we assume that $n \equiv 1 \pmod 6$, since the other cases $n \equiv 3, 5 \pmod 6$ are handled using the same argument, up to cyclic symmetry. (Henceforth, when we write $n \to \infty$, we are referring to the subsequence defined by setting $n = 6k + 1$ as $k \to \infty$.) Let s_n denote the earliest time in the interval $[t_n, t_{n+1}]$ when at least one of the relations $x_1 = x_2$ or $x_3 \geq \frac{2}{M+1}$ holds; note, in particular, that $s_n \leq r_n$. Let $v_n = x_2(s_n)$. We first argue that $v_n \to 0$ as $n \to \infty$. If $s_n = t_n$ then $v_n = w_n$, which converges to zero by Lemma 3.9. If $s_n > t_n$ then $v_n = x_2(s_n) \leq x_1(s_n) \leq x_3(s_n) \leq \frac{2}{M+1}$, and the convergence of v_n to zero now follows from the convergence of $\vec{x}(t)$ to the boundary of the cube.

To complete the proof that $x_1(r_n) \to 0$ as $n \to \infty$, we consider two cases. First, for those n such that $s_n = r_n$, we have $x_1(r_n) = x_2(r_n) = x_2(s_n) = v_n$, which converges to zero. Second, for those n such that $s_n < r_n$, we have $x_3(t) \geq \frac{2}{M+1}$ for all $t \in [s_n, t_{n+1}]$. We may now use the relation $\dot{y}_1 + \dot{y}_2 = 2 - (M+1)(x_1 + x_3)$ to conclude that $\frac{d}{dt}[y_1(t) + y_2(t)] < 2 - (M+1)x_3 \leq 0$ on the interval $[s_n, t_{n+1}]$, which includes r_n. Consequently,

$$\begin{aligned}
y_1(r_n) &= \frac{1}{2}\left(y_1(r_n) + y_2(r_n)\right) \\
&\leq \frac{1}{2}\left(y_1(s_n) + y_2(s_n)\right) \\
&\leq \frac{1}{2}\left(-\ln(M) + \ln(v_n) - \ln(1 - v_n)\right). \quad (11)
\end{aligned}$$

The fact that $v_n \to 0$ implies that the right side of (11) tends to $-\infty$, implying that $y_1(r_n)$ tends to $-\infty$ and that $x_1(r_n)$ tends to zero.

Now we turn to the proof that $\max_j\{x_j(t)\} \to 1$ as $t \to \infty$. We begin by considering the intervals $[t_n, t_{n+1}]$ such that n is odd. On any such interval, x_n starts at $\frac{1}{M+1}$ and decreases, x_{n+1} starts at $w_n < \frac{1}{M+1}$ and increases, x_{n+2} starts above $\frac{1}{M+1}$ and increases. Hence when n is odd, $\max_j\{x_j(t)\} = x_{n+2}(t)$ for all $t \in [t_n, t_{n+1}]$. Our first objective is to prove that $x_{n+2}(t_{n+1}) \to 1$ for odd n tending to infinity. (Of course, this doesn't establish that $\inf_{t_n \leq t \leq t_{n+1}}\{x_{n+2}(t)\}$ tends to 1 for odd n tending to infinity, but we will return to this issue at the end of the proof.) To prove that $x_{n+2}(t_{n+1}) \to 1$, we first introduce the parameter

$$z_i = y_{i+1} - y_i - (M+1)\ln(1-x_i),$$

which satisfies

$$\begin{aligned}
\dot{z}_i &= [1-(M+1)x_i] - [1-(M+1)x_{i-1}] + \\
&\quad + \frac{(M+1)\dot{x_i}}{1-x_i} \\
&= (M+1)(x_{i-1}-x_i) + \\
&\qquad + (M+1)x_i[1-(M+1)x_{i-1}] \\
&= (M+1)x_{i-1}[1-(M+1)x_i].
\end{aligned}$$

In particular, since $x_{n+1}(t) < \frac{1}{M+1}$ for all $t \in (t_n, t_{n+1})$ when n is odd, we have $\dot{z}_{n+1} > 0$ and therefore,

$$\begin{aligned}
&z_{n+1}(t_{n+1}) - z_{n+1}(t_n) > 0 \\
&y_{n+2}(t_{n+1}) - y_{n+2}(t_n) \\
&\quad > y_{n+1}(t_{n+1}) - y_{n+1}(t_n) + \\
&\qquad + (M+1)[\ln(1-x_{n+1}(t_{n+1})) - \\
&\qquad\quad - \ln(1-x_{n+1}(t_n))] \\
&\quad = -\ln(M) - \ln(w_n) + \ln(1-w_n) + \\
&\qquad + (M+1)\left[\ln\left(1-\tfrac{1}{M+1}\right) - \ln(1-w_n)\right] \\
&\quad > -\ln(M) - \ln(w_n) + (M+1)\ln\left(1-\tfrac{1}{M+1}\right).
\end{aligned}$$

Combining this with the relation $y_{n+2}(t_n) \geq -\ln(M)$, we see that

$$\begin{aligned}
&y_{n+2}(t_{n+1}) \\
&\quad > -2\ln(M) - \ln(w_n) + (M+1)\ln\left(1-\tfrac{1}{M+1}\right),
\end{aligned}$$

and consequently $y_{n+2}(t_{n+1})$ tends to infinity for odd n tending to infinity, implying that $x_{n+2}(t_{n+1})$ tends to 1 as claimed.

To deal with intervals $[t_n, t_{n+1}]$ where n is even, we begin by observing that on any such interval, x_n starts at $\frac{1}{M+1}$ and increases, x_{n+1} starts above $\frac{1}{M+1}$ and decreases, x_{n-1} starts at w_n and decreases. Hence, the quantity $\max_j\{x_j(t)\}$ is minimized at the unique time r_n satisfying $x_n(r_n) = x_{n+1}(r_n)$. By taking n sufficiently large, we may assume $(M+1)w_n < 1/2$. Now we find that for $t \in (t_n, t_{n+1})$,

$$\begin{aligned}
\dot{y}_n = 1-(M+1)x_{n-1} &> 1-(M+1)w_n > 1/2 \\
\dot{y}_{n+1} = 1-(M+1)x_n &> 1-(M+1) = -M \\
\frac{d}{dt}(2My_n + y_{n+1}) &> 2M\cdot(1/2) - M = 0.
\end{aligned}$$

Consequently,

$$\begin{aligned}
y_n(r_n) &= \frac{1}{2M+1}\left(2My_n(r_n) + y_{n+1}(r_n)\right) \\
&> \left(\frac{2M}{2M+1}\right) y_n(t_n) + \left(\frac{1}{2M+1}\right) y_{n+1}(t_n) \\
&= \left(\frac{2M}{2M+1}\right)\ln(1/M) + \left(\frac{1}{2M+1}\right) y_{n+1}(t_n).
\end{aligned}$$

Since $n-1$ is odd, the preceding paragraph established that the quantity $y_{n+1}(t_n)$ on the right side tends to infinity for even n tending to infinity. Thus, we also have that $y_n(r_n) \to \infty$ and $x_n(r_n) \to 1$ for even n tending to infinity.

We have shown for even n that $\inf_{t_n \leq t \leq t_{n+1}}\{\max_j x_j(t)\} \to 1$ as $n \to \infty$; to conclude the proof we show the same for odd n. This is easily done, since we know that $\max_j x_j(t) = x_{n+2}(t)$ and that $x_{n+2}(t)$ is a monotonically increasing function of t for n odd and $t \in [t_n, t_{n+1}]$. Thus,

$$\begin{aligned}
\inf_{t_n \leq t \leq t_{n+1}}\{\max_j x_j(t)\} &= x_{n+2}(t_n) \\
&= \max_j x_j(t_n) \\
&\geq \inf_{t_{n-1} \leq t \leq t_n}\{\max_j x_j(t)\}.
\end{aligned}$$

We have already seen that the right side tends to infinity for even $n-1$ tending to infinity, so the left side tends to infinity as well. $\square$

References

[1] S. Arora, E. Hazan, and S. Kale. The multiplicative weights update method: A meta-algorithm and applications. Preprint available for download at http://www.cs.princeton.edu/∼arora/publist.html.

[2] N. Balcan, A. Blum, and Y. Mansour. Circumventing the Price of Anarchy: Leading Dynamics to Good behavior. *Innovations in Computer Science* (ICS) 2010.

[3] A. Blum, E. Even-Dar, K. Ligett. Routing Without Regret: On Convergence to Nash Equilibria of Regret-Minimizing Algorithms in Routing Games. In the *Symposium on Principles of Distributed Computing*, 2006.

[4] A. Blum, M. Hajiaghayi, K. Ligett, and A. Roth. Regret Minimization and the Price of Total anarchy. In the *40th ACM Symposium on Theory of Computing*, 373-382, 2008.

[5] A. Blum and Y. Mansour. Learning, Regret Minimization and Equilibria, in *Algorithmic Game Theory* N. Nisan, T. Roughgarden, E. Tardos and V. Vazirani (eds), Cambridge University Press, 2007.

[6] X. Chen, X. Deng, Settling the Complexity of Two-Player Nash Equilibrium. In the *Proceedings of the IEEE Symposium on Foundations of Computing*, 2006.

[7] C. Daskalakis, R. Frongillo, C. Papadimitriou, G. Pierrakos, and G. Valiant. On Learning Algorithms for Nash Equilibria. *International Symposium on Algorithmic Game Theory*, 2010.

[8] C. Daskalakis, P. W. Goldberg and C. H. Papadimitriou, The Complexity of Computing a Nash Equilibrium, In the *38th ACM Symposium on Theory of Computing*, 2006.

[9] Y. Freund, and R. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29:79-103, (1999).

[10] A. Gaunersdorfer and J. Hofbauer. Fictitious Play, Shapley Polygons, and the Replicator Equation. *Games and Economic Behaviour* 11 (1995), 279-303.

[11] M. Goemans, V. Mirrokni, and A. Vetta. Sink Equilibria and Convergence. In the *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 142–154, 2005.

[12] J. Hannan, Approximation to Bayes risk in repeated plays. In M. Dresher, A. Tucker, and P. Wolfe eds, *Contributions to the Theory of Games*, vol 3, pp 79-139, Princeton University Press.

[13] S. Hart, and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127-1150, 2000.

[14] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics.* Cambridge Univ, Press, 1998.

[15] J. Hofbauer, S. Sorin, and Y. Viossat. Time Average Replicator and Best-Reply Dynamics. *Mathematics of Operations Research* 34/2 (2009), 263-269.

[16] R. Kleinberg and G. Piliouras and E. Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In the *41st ACM Symposium on Theory of Computing*, (2009).

[17] R. Kleinberg, G. Piliouras, and Tardos Load balancing without regret in the billboard model. In *Proceedings of the 28th Symposium on Principles of Distributed Computing* (PODC 2009)

[18] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–260, 1994.

[19] G. Piliouras. *A learning theoretic approach to game theory.* PhD Thesis, Cornell 2010.

[20] T. Roughgarden. Intrinsic Robustness of the Price of Anarchy. In the *41st ACM Symposium on Theory of Computing*, (2009).

[21] T. Roughgarden and É. Tardos. Introduction to the Inefficiency of Equilibria, in *Algorithmic Game Theory* eds. N. Nisan, T. Roughgarden, E. Tardos and V. Vazirani, Cambridge University Press, 2007.