Int J Game Theory (2013) 42:1059–1083
DOI 10.1007/s00182-012-0362-6

On the approximation performance of fictitious play in finite games

Paul W. Goldberg · Rahul Savani ·
Troels Bjerre Sørensen · Carmine Ventre

Received: 10 September 2011 / Accepted: 20 November 2012 / Published online: 4 December 2012

Abstract We study the performance of Fictitious Play (FP), when used as a heuristic for finding an approximate Nash equilibrium of a two-player game. We exhibit a class of two-player games having payoffs in the range [0, 1] that show that FP fails to find a solution having an additive approximation guarantee significantly better than $1/2$. Our construction shows that for $n \times n$ games, in the worst case both players may perpetually have mixed strategies whose payoffs fall short of the best response by an additive quantity $1/2 - O(1/n^{1-\delta})$ for arbitrarily small δ. We also show an essentially matching upper bound of $1/2 - O(1/n)$.

Keywords Fictitious play · Approximate Nash equilibria · Decentralized dynamics

1 Introduction

We study the computation of *approximate Nash equilibria* of games, usually called ϵ-Nash equilibria, where ϵ is a measure of the quality of the approximation. For an

P. W. Goldberg · R. Savani
Department of Computer Science, University of Liverpool, Ashton Street, Liverpool L69 3BX, UK
e-mail: P.W.Goldberg@liverpool.ac.uk

R. Savani
e-mail: Rahul.Savani@liverpool.ac.uk

T. B. Sørensen
Department of Computer Science, University of Warwick, Gibbet Hill Rd, Coventry CV4 7AL, UK
e-mail: trold@dcs.warwick.ac.uk

C. Ventre (✉)
School of Computing, Teesside University, Borough Road, Middlesbrough TS1 3BA, UK
e-mail: C.Ventre@tees.ac.uk

approximate Nash equilibrium, we replace the "no incentive to deviate" criterion of a classical Nash equilibrium with "low incentive to deviate". In particular, a player can improve his payoff by at most ϵ by playing an alternative strategy. For positive ϵ, this is of course a less demanding objective than Nash equilibrium.

In the literature on equilibrium computation, a central issue is fast algorithms: in order for an outcome to be plausible, it should be possible to compute it in a reasonable amount of time. This leads to a focus on algorithms that run in polynomial time. On this topic, the PPAD-completeness results of Daskalakis et al. (2009b) and Chen et al. (2009) provide compelling evidence that there is no polynomial-time algorithm that can solve unrestricted normal-form games. These results have led to subsequent work on the polynomial-time computability of approximate Nash equilibria, and the study of how good an approximation can be computed in polynomial time. We continue by describing this line of work, followed by the motivation for studying the *Fictitious Play* procedure in this setting. We concentrate on two-player normal-form games.

1.1 Approximate Nash equilibria: details and background

An ϵ-Nash equilibrium is one where every player's strategy has a payoff of at most ϵ less than a best response; ϵ is called *regret* in this context. Formally, for two players with pure strategy sets M, N and payoff functions $u_i : M \times N \to \mathbb{R}$ for $i \in \{1, 2\}$, the mixed strategy σ is an ϵ-*best-response* against the mixed strategy τ, if for any $m \in M$, we have $u_1(\sigma, \tau) \geq u_1(m, \tau) - \epsilon$. A pair of strategies σ, τ is an ϵ-Nash equilibrium if they are ϵ-best responses to each other. For what values of $\epsilon > 0$ is it possible to compute ϵ-Nash equilibria in time polynomial in the size of a game? To make the question meaningful, we have to assume that the payoffs of the game lie in some limited range; by convention the range $[0, 1]$. (Without that assumption, we could, in a trivial way, boost the performance of an algorithm by scaling up (by a factor 10, say) the payoffs of a game of interest, so that a solution achieving value ϵ for the scaled-up game would achieve $\epsilon/10$ for the original game.)

In one of the first papers to consider this question, Daskalakis et al. (2009a) gave the following very simple algorithm that achieves an approximation guarantee of $\epsilon = \frac{1}{2}$. Suppose the row player allocates probability $\frac{1}{2}$ to some fixed pure strategy s^r, the column player plays a pure best response s^c to s^r, and the row player allocates the remaining $\frac{1}{2}$ to a best response to s^c. It is not hard to check that this guarantees an $\epsilon = \frac{1}{2}$ approximate Nash equilibrium, for any game. It would seem that more sophisticated algorithms ought to be able to perform much better (achieve much lower values of ϵ). However, progress in improving this approximation guarantee has been rather incremental. Daskalakis et al. (2007) gave an algorithm that achieves $\epsilon = 0.38197$, and an improvement due to Bosse et al. Bosse et al. (2007) attains $\epsilon = 0.36392$. The best known approximation guarantee, $\epsilon = 0.3393$, is due to Tsaknakis and Spirakis Tsaknakis and Spirakis (2008). These algorithms that have improved on $\epsilon = \frac{1}{2}$ work by solving one or more linear programs derived from the input game, and are in a sense centralized since it is necessary to integrate both players' payoffs into the LP. It is also known from Feder et al. (2007) that for $\epsilon < \frac{1}{2}$, an ϵ-Nash equilibrium may require the size of the support of the players' mixed strategies to be logarithmic in the number of pure strategies (in contrast with the algorithm of Daskalakis et al. (2009a), where only three pure strategies are used by the players). The main open question being addressed by this line of work is the possible existence of a *polynomial-time approximation scheme*, an algorithm that given any $\epsilon > 0$ finds an ϵ-Nash equilibrium in polynomial time.

The notion of approximation we work with here is distinct from the related notion of ϵ-*well-supported equilibria*, where a player must allocate probability 0 to any pure strategy whose payoff is more than ϵ less than optimal. The notion of well-supported equilibrium is more demanding than the ϵ-approximate equilibria considered here. Kontogiannis and Spirakis (2010) gave a polynomial-time algorithm that achieves $\epsilon = \frac{2}{3}$, and recently Fearnley et al. (2012) introduced a modification to the algorithm of Kontogiannis and Spirakis Kontogiannis and Spirakis (2010) that leads to a slight improvement of the approximation guarantee, but is still only slightly below $\frac{2}{3}$. Thus, less progress has been made in reducing ϵ for well-supported equilibria. Indeed, the mixed-strategy profiles that arise from applying FP to the games we construct here are not well-supported for any $\epsilon < 1$.

1.2 Fictitious play

FP (see Fudenberg and Levine 1998, Chapter 2 for more details) is a very simple iterative process for computing equilibria of games. When it converges, it necessarily converges to a Nash equilibrium. For two-player games, it is known to converge for zero-sum games (Robinson 1951), non-degenerate 2×2 games (Miyasawa 1961), or if one player has just two pure strategies (Berger 2005). On the other hand, Shapley exhibited a 3×3 game for which it fails to converge (Fudenberg and Levine 1998; Shapley 1964).

FP works as follows. Suppose that each player has a number of *actions*, or pure strategies. Initially (at iteration 1) each player starts with a single action. Thereafter, at iteration t, each player has a sequence of $t - 1$ actions that is extended with a t-th action chosen as follows. Each player makes a best response to a distribution that selects an opponent's action uniformly at random from his sequence. (To make the process precise, a tie-breaking rule should also be specified; however, in the games constructed here, we can avoid ties.) Thus the process generates a sequence of mixed-strategy profiles (viewing the sequences as probability distributions), and the hope is that they converge to a limiting distribution, which would necessarily be a Nash equilibrium.

1.3 FP and approximate equilibria

We consider the extent to which, after some number of steps, the empirical distributions obtained by the players are good (albeit not optimal) responses to each other. A recent paper of Conitzer (2009) shows that FP obtains an approximation guarantee of $\epsilon = (t + 1)/2t$ for two-player games, where t is the number of FP iterations, and furthermore, *if both players have access to infinitely many actions*, then FP cannot do better than this. The intuition behind the upper bound is that an action that appears most recently in a player's sequence has an ϵ-value close to 0 (at most $1/t$); generally an action that occurs a fraction γ back in the sequence has an ϵ-value of at most slightly more than γ (it is a best response to slightly less than $1-\gamma$ of the opponent's distribution), and the ϵ-value of a player's mixed strategy is at most the overall average, i.e., $(t+1)/2t$, which approaches $\frac{1}{2}$ as t increases.

However, as soon as the number of available actions is exceeded by the number of iterations of FP, various actions must get re-used, and this re-usage means, for example, that every previous occurrence of the most recent action has ϵ-value $\frac{1}{t}$. This appears to hold up the hope that FP may ultimately guarantee a significantly better additive approximation. Indeed, experiments suggest that it nearly always does better than the $\epsilon = \frac{1}{2}$ guaranteed by the results of Conitzer (2009). On Shapley's game (1964) mentioned above, FP obtains ϵ of about $\frac{1}{4}$. We show that unfortunately that is not what results in the worst case. Our hope is that this result may either guide the design of more "intelligent" dynamics having a better approximation performance, or alternatively generalize to a wider class of related algorithms, for example the ones discussed by Daskalakis et al. (2010).

1.4 Why FP?

A detailed motivation for studying the performance of FP is given by Conitzer (2009). Amongst others, we mention the following observations made by Conitzer (2009). FP has practical relevance. For example, FP is implemented and run to calculate (approximate) equilibria for several variants of poker (Dudziak 2006; Ganzfried and Sandholm 2008) and, in some cases, the solutions based on FP are provably better than those using alternative approaches (McMahan and Gordon 2007). Moreover, due to its simplicity, FP is more believable as a model of human behavior than more complicated algorithms. Finally, FP is appealing to different scientific communities with similar interests in games and complementary points of view: its practical relevance makes it relevant to the artificial intelligence/multiagent systems community, its appropriateness to describe rationality (and the fact that people may actually use it) appeals to the classical game theory community, and formal analysis of its approximation guarantee makes it interesting to the theoretical computer science community. In relation to the latter community, we also observe that FP has been studied as a heuristic for generic optimization (Lambert et al. 2005).

An additional motivation behind this work can be found in the fact that FP is an uncoupled dynamics, i.e., it is a dynamics in which the decision of a player on the action to play next does not depend on the payoffs of the other players. In other words, in an uncoupled dynamics the decisions of a player are functions of his own payoff and the history only. Hart and Mas-Colell (2003, 2006) try to demarcate the border between those classes of dynamics for which convergence to "good" approximate Nash equilibria can be obtained and those for which it cannot. Our result is another step in this direction and complements the results of Hart and Mas-Colell in that FP is a discrete-time deterministic uncoupled dynamics; Hart and Mas-Colell study deterministic, continuous-time uncoupled dynamics (Hart and Mas-Colell 2003) and stochastic, discrete-time uncoupled dynamics (Hart and Mas-Colell 2006).

1.5 Our contribution

Like most of the work on approximate Nash equilibrium computation, we focus on two-player normal-form games. Our main result is a class of $n \times n$ games for which we show that FP fails to obtain any constant $\epsilon < \frac{1}{2}$, thus extending Conitzer (2009) result from the infinite to the finite case. The result holds without any constraints on the number of iterations; we show that a kind of cyclical behavior persists.

In Sect. 2 we give our main result, the lower bound on ϵ of $1/2 - O(1/n^{1-\delta})$ for any $\delta > 0$, and in Sect. 3 we give the corresponding upper bound of $1/2 - O(1/n)$.

Our main result thus highlights a kind of weakness or limitation of FP. We noted above that Shapley (1964) already exhibited a 3×3 game where FP does not converge, which is perhaps the first result showing a weakness of FP. However, on Shapley's example, FP finds an ϵ-Nash equilibrium for $\epsilon \approx \frac{1}{4}$ in only a few iterations, and thereafter the approximation never gets much worse. Thus, it raises the question of whether the short-term or long-term behavior of FP are ever as bad as Conitzer's (2009) upper bound of $\frac{1}{2}$. Our example shows that both the short-term and the long-term approximation quality may indeed be arbitrarily close to the upper bound of $\frac{1}{2}$. Our result does require the initial (pure) strategies at iteration 0 to be fixed; it may be seen that if both players begin by using their last rather their first strategies, that immediately we have a very good approximate equilibrium. Thus, it may still be that certain variants of FP can in fact perform better. This issue is discussed further in Sect. 4.

1.6 Other related work

We conclude this introduction by mentioning some further negative results for FP, and how they relate to the results presented here.

Brandt et al. (2010) consider classes of games where FP is known to converge, such as $2 \times n$ games, or potential games (Monderer and Shapley 1996), or constant-sum games. They exhibit examples of these in which FP may take exponentially many steps (as a function of the size of the representation of the game) for the players to select some strategy that must be played with non-zero probability in any Nash equilibrium of the game. This represents an alternative critique of FP, that it is slow in finding Nash equilibrium strategies. In their games however, the mixed-strategy profiles obtained by FP in the short term, are good when viewed as approximate Nash equilibria.

Fudenberg and Levine (1995) develop a "cautious" version of FP, in which an agent's mixed strategy is not excessively sensitive to the exact relative frequencies of actions played by the other player(s). It is designed to address the well-known problem that FP is not *safe*, in that since it is deterministic, a player could be "ripped off" at each step by an adversary, and his realised payoff can be less than his minmax payoff. Indeed, it is useful to consider the following "battle of the sexes" type game, discussed in detail in Fudenberg and Kreps (1993, p. 339). It is worth revisiting here since it has some features in common with the more complicated games we construct below, in particular that since it is symmetric, FP will let the players use the same sequence (*top* corresponding with *left* and *bottom* with *right*).

1																			
β	1																		
β	β	1																	
β	β	β	1																
β	β	β	β	α	1														
β	β	β	β	β	α	1													
β	β	β	β	β	β	α	1												
β	β	β	β	β	β	β	α	1											
β	β	β	β	β	β	β	β	α	1										
β	β	β	β	β	β	β	β	β	α	1					β	β	β	β	α
β	β	β	β	β	β	β	β	β	β	α	1					β	β	β	β
β	β	β	β	β	β	β	β	β	β	β	α	1					β	β	β
β	β	β	β	β	β	β	β	β	β	β	β	α	1					β	β
β	β	β	β	β	β	β	β	β	β	β	β	β	α	1					β
β	β	β	β	β	β	β	β	β	β	β	β	β	β	α	1				
β	β	β	β	β	β	β	β	β	β		β	β	β	β	α	1			
β	β	β	β	β	β	β	β	β	β			β	β	β	β	α	1		
β	β	β	β	β	β	β	β	β	β				β	β	β	β	α	1	
β	β	β	β	β	β	β	β	β	β					β	β	β	β	α	1

Fig. 1 The game $\mathcal{G}_5$ belonging to the class of games used to prove the lower bound

$$\begin{pmatrix} (0,0) & (2,1) \\ (1,2) & (0,0) \end{pmatrix}$$

Suppose that the players begin with a prior distribution on their opponents' distribution, having probabilities in the ratio $1 : \sqrt{2}$.[1] Thereafter, a player's estimate of the opponent's distribution is the average of the prior with the observed sequence of actions. By symmetry, both players will copy each other, and will repeatedly alternate between $(top, left)$ and $(bottom, right)$. As a result, the players' realised payoffs (of zero) fall short of the expected payoffs they could have obtained with their (mixed) minimax strategies. In this paper, the quantities we are comparing are different; we compare the payoff arising from randomizing over the players' observed behavior (which, depending on the game, may be either higher or lower than the realised payoff) with the payoff resulting from best-responding to the opponent's observed behavior. In that respect, FP performs well on the above game, converging quickly to the mixed equilibrium.

2 Lower bound

We specify a class of symmetric games with parameter n, whose general idea is conveyed by Fig. 1, which shows the row player's payoff matrix for $n = 5$[2]; since the game is symmetric, the column player's matrix is its transpose. A blank entry indicates

[1] An irrational number is used in order to guarantee that there is never an issue with tie-breaking.

[2] It is notationally convenient to let the n-th game be of size $4n \times 4n$; our bounds in terms of n use the big-O notation, so this is safe.

a payoff of zero; let $\alpha = 1 + \frac{1}{n^{1-\delta}}$ and $\beta = 1 - \frac{1}{n^{2(1-\delta)}}$ for $\delta > 0$. We assume both players start FP at action 1 (top left).

Generally, let $\mathcal{G}_n = (R, R^{\mathrm{T}})$ be a $4n \times 4n$ game in which the row player's payoff matrix R is specified as follows. Let $[a : b]$, $a \leq b$, be the set of integers between a and b. For $i, j \in [1 : 4n]$ we have

- If $i \in [2 : n]$, $R_{i,i-1} = 1$. If $i \in [n+1 : 4n]$, $R_{i,i} = 1$.
- If $i \in [n+1 : 4n]$, $R_{i,i-1} = \alpha$. Also, $R_{2n+1,4n} = \alpha$.
- Otherwise, if $i > j$ and $j \leq 2n$, $R_{i,j} = \beta$.
- Otherwise, if $i > j$ and $i - j \leq n$, $R_{i,j} = \beta$. If $j \in [3n+1 : 4n]$, $i \in [2n+1 : j-n]$, $R_{i,j} = \beta$.
- Otherwise, $R_{ij} = 0$.

For ease of presentation we analyze FP on $\mathcal{G}_n$; the obtained results can be seen to apply to a version of $\mathcal{G}_n$ with payoffs rescaled into [0, 1] (cf. the proof of Theorem 2).

The sub-game given by the first $n \times n$ actions is similar to Conitzer's lower bound game (Conitzer 2009) in that when players start from action 1, they will play each action only once. However, differently from Conitzer's game, in our case players' best responses are unique.

2.1 Overview

In this section, we give a general overview and intuition on how our main result works, before embarking on the technical details. Number the actions $1, \ldots 4n$ from top to bottom and left to right, and assume that both players start at action 1. FP proceeds in a sequence of steps, which we index by the positive integer t, so that step t consists of both players adding the tth element to their sequences of length $t-1$. We have the following observation:

Observation 1 *Since the game is symmetric and both players start with the same action, there is an FP sequence in which both players play the same action in every step.*

This simplifies the analysis since we can analyze a single sequence of numbers (the shared indices of the actions chosen by the players). In fact, the FP sequence is unique as multiple best responses never arise.

A basic insight into the behavior of FP on the games in question is provided by Lemma 1, which tells us a great deal about the structure of the players' sequence. Let s_t be the action played at step t. We set $s_1 = 1$.

Lemma 1 *For any time step t, if $s_t \neq s_{t+1}$ then $s_{t+1} = s_t + 1$ (or $s_{t+1} = 2n + 1$ if $s_t = 4n$).*

Proof It is not hard to check that for $t \leq n$, at step t each player plays action t. (As we noted, the sub-game given by the first $n \times n$ actions is similar to Conitzer's (2009) construction, in which FP behaves in this way.) For step $t > n$, suppose the players play $s_t \neq 4n$ (by Observation 1, we assume that the two players play the same action). s_t is a best response at step t, and since $R_{s_t+1,s_t} > R_{s_t,s_t} > R_{j,s_t}$

($j \notin \{s_t, s_t + 1\}$), action $s_t + 1$ is the only other candidate to become a better response after s_t is played. Thus, if $s_{t+1} \neq s_t$, then $s_{t+1} = s_t + 1$. A similar argument applies to the case $s_t = 4n$. □

Lemma 1 implies that the FP sequence consists of a block of consecutive 1's followed by some consecutive 2's, and so on through all the actions in ascending order until we get to a block of consecutive $4n$'s. The blocks of consecutive actions then cycle through the actions $\{2n + 1, \ldots, 4n\}$ in order, and continue to do so repeatedly.

As it stands, Lemma 1 makes no promise about the lengths of these blocks, and indeed it does not itself rule out the possibility that one of these blocks is infinitely long (which would end the cycling process described above and cause FP to converge to a pure Nash equilibrium). The subsequent results say more about the lengths of the blocks. They show that in fact the process never converges (it cycles infinitely often) and furthermore, the lengths of the blocks increase in a geometric progression. The parameters α and β in $\mathcal{G}_n$ govern the ratio between the lengths of consecutive blocks. We choose a ratio large enough to ensure that the n most recently played actions occupy all but an exponentially small fraction of the FP sequence. At the same time, the ratio is small enough that the corresponding probability distribution does not allocate much probability to any individual action.

In more detail, Lemma 3 analyses the first pass through actions $n+1, \ldots, 3n$ made by the players, and shows that the number of times that each action i in this set is played is somewhat greater than the number of times $i - 1$ was played. The ratio between these two numbers is slightly more than 1. Lemma 4 obtains a very similar result, for actions $3n + 1, \ldots, 4n$. Consequently by the time action $4n$ is reached, we note that

- most of the empirical probability is placed on the more recently-played actions, and
- the amount of empirical probability on the first actions $1, \ldots, 2n$ is very low indeed, an exponentially small fraction.

Lemma 5 makes precise the second of the above two points: the effect of the initial play of actions $1, \ldots, 2n$ is by now negligible. By this stage then, we can effectively ignore actions $1, \ldots, 2n$ and focus on the sub-game that uses actions $2n + 1, \ldots, 4n$. The general point next is that from now on, there is a "current action" i being played, and i and its recent predecessors in the sequence $2n + 1, \ldots, 4n$ have the highest empirical probability. (The predecessor of $2n + 1$ is $4n$, in this context.) Theorem 1 gives a precise statement of this, that the ratio between empirical probabilities of consecutive actions is approximately 1 plus an inverse polynomial in n. Thus Theorem 1 gives a fairly precise description of the empirical distributions obtained by the players. Theorem 2 deduces that at any point in time, the best response by either player is worth approximately 1, while the average payoff of his current empirical distribution is approximately $1/2$.

As an aside, we conclude this section with the following observation, which is not hard to check from the structure of the game.

Observation 2 *The game has a mixed Nash equilibrium in which both players use the uniform distribution over actions* $\{2n + 1, \ldots, 4n\}$. *The equilibrium has payoff* *approximately $\frac{1}{2}$ to each player. There are no pure Nash equilibria, although if both players use the same action in the range $\{n+1, \ldots, 4n\}$ then they would receive payoff 1. Recall that $\alpha > 1$, so a payoff of 1 to each player does not imply an equilibrium.*

2.2 The proof

We now identify some properties of the probabilities that FP assigns to actions. Let $\ell_t(i)$ be the number of times that action i is played by the players until time step t. Let $p_t(i)$ be the corresponding probability assigned to action i at step t, and for a subset of actions S let $p_t(S)$ denote the total probability of elements in S. So it is immediate to observe that

$$p_t(i) = \frac{\ell_t(i)}{\sum_{j=1}^{4n} \ell_t(j)} = \frac{\ell_t(i)}{t}.$$

The next fact follows easily from the FP rule.

Lemma 2 *For all actions $i \leq n$, $p_t(i) = \frac{1}{t}$ and therefore $\ell_t(i) = 1$ for any time step $t \geq i$.*

Proof At step 1, each player sets $p_1(1) = 1$ and $p_1(i) = 0$ for $i > 1$. For $t \leq n$ the sequence chosen by both players is $(1, 2, \ldots, t)$, so $p_t(i) = \frac{1}{i}$ for $i \leq t$ and 0 otherwise. Lemma 1 implies that none of the first n actions will be a best response subsequently, which implies the claim. □

By Lemma 1, each action is played a number of consecutive times, in order, until action $4n$ is played; at this point, this same pattern repeats but only for the actions in $\{2n + 1, \ldots, 4n\}$. We let $t^\star$ be the length of the longest sequence containing all the actions in ascending order, that is, $t^\star$ is the last step of the first consecutive block of $4n$'s. We also let t_i be the last time step in which i is played during the first $t^\star$ steps, i.e., t_i is such that $\ell_{t_i}(i) = \ell_{t_i-1}(i) + 1$ and $\ell_t(i) = \ell_{t^\star}(i)$ for $t \in \{t_i, \ldots, t^\star\}$.

Lemma 3 *For all actions $n + 1 \leq i \leq 3n$ and all $t \in \{t_i, \ldots, t^\star\}$,*

$$\frac{\alpha - \beta}{\alpha - 1} p_t(i-1) \leq p_t(i) \leq \frac{1}{t} + \frac{\alpha - \beta}{\alpha - 1} p_t(i-1),$$

and therefore

$$\frac{\alpha - \beta}{\alpha - 1} \ell_t(i-1) \leq \ell_t(i) \leq 1 + \frac{\alpha - \beta}{\alpha - 1} \ell_t(i-1).$$

Proof By definition of t_i, action i is played at step t_i. This means that i is a best response for the players given the probability distributions at step $t_i - 1$. In particular, the expected payoff of i is better than the expected payoff of $i + 1$, that is,

$$\beta \sum_{j=1}^{i-2} p_{t_i-1}(j) + \alpha p_{t_i-1}(i-1) + p_{t_i-1}(i)$$
$$\geq \beta \sum_{j=1}^{i-2} p_{t_i-1}(j) + \beta p_{t_i-1}(i-1) + \alpha p_{t_i-1}(i).$$

Since $\alpha > 1$, the above implies that $p_{t_i-1}(i) \leq \frac{\alpha-\beta}{\alpha-1} p_{t_i-1}(i-1)$. By explicitly writing the probabilities, we get

$$\frac{\ell_{t_i-1}(i)}{t_i-1} \leq \frac{\alpha-\beta}{\alpha-1} \frac{\ell_{t_i-1}(i-1)}{t_i-1} \iff \ell_{t_i}(i) - 1 \leq \frac{\alpha-\beta}{\alpha-1} \ell_{t_i}(i-1) \iff \tag{1}$$
$$\frac{\ell_{t_i}(i)}{t_i} \leq \frac{1}{t_i} + \frac{\alpha-\beta}{\alpha-1} \frac{\ell_{t_i}(i-1)}{t_i} \iff p_{t_i}(i) \leq \frac{1}{t_i} + \frac{\alpha-\beta}{\alpha-1} p_{t_i}(i-1). \tag{2}$$

At step $t_i + 1$ action i is not a best response to the opponent's strategy. Then, by Lemma 1, $i+1$ is the unique best response and so the expected payoff of $i+1$ is better than the expected payoff of i given the probability distributions at step t_i, that is,

$$\beta \sum_{j=1}^{i-2} p_{t_i}(j) + \alpha p_{t_i}(i-1) + p_{t_i}(i) \leq \beta \sum_{j=1}^{i-2} p_{t_i}(j) + \beta p_{t_i}(i-1) + \alpha p_{t_i}(i).$$

Since $\alpha > 1$, the above implies that

$$p_{t_i}(i) \geq \frac{\alpha-\beta}{\alpha-1} p_{t_i}(i-1), \tag{3}$$

and then that

$$\ell_{t_i}(i) \geq \frac{\alpha-\beta}{\alpha-1} \ell_{t_i}(i-1). \tag{4}$$

By definition of t_i action i will not be played again until time step $t^\star$. Similarly, Lemma 1 shows that $i-1$ will not be a best response twice in the time interval $[1, t^\star]$ and so will not be played until step $t^\star$. Therefore, the claim follows from (1), (2), (3) and (4). □

Lemma 4 *For all actions* $i \in \{3n+1, \ldots, 4n-1\}$ *and all* $t \in \{t_i, \ldots, t^\star\}$,

$$\frac{\alpha-\beta}{\alpha-1} p_t(i-1) \leq p_t(i) \leq \frac{1}{t} + \frac{\alpha-\beta}{\alpha-1} p_t(i-1) + \frac{\beta}{\alpha-1} p_t(i-n),$$

and therefore

$$\frac{\alpha-\beta}{\alpha-1} \ell_t(i-1) \leq \ell_t(i) \leq 1 + \frac{\alpha-\beta}{\alpha-1} \ell_t(i-1) + \frac{\beta}{\alpha-1} \ell_t(i-n).$$

Proof By definition of t_i, action i is played at time step t_i. This means that i is a best response for the players after $t_i - 1$ steps. In particular, the expected payoff of i is better than the expected payoff of $i + 1$, that is,

$$\beta\left(\sum_{j=1}^{2n} p_{t_i-1}(j) + \sum_{j=i-n}^{i-2} p_{t_i-1}(j)\right) + \alpha p_{t_i-1}(i-1) + p_{t_i-1}(i)$$
$$\geq \beta\left(\sum_{j=1}^{2n} p_{t_i-1}(j) + \sum_{j=i-n+1}^{i-2} p_{t_i-1}(j)\right) + \beta p_{t_i-1}(i-1) + \alpha p_{t_i-1}(i).$$

Since $\alpha > 1$, the above implies that $p_{t_i-1}(i) \leq \frac{\alpha-\beta}{\alpha-1} p_{t_i-1}(i-1) + \frac{\beta}{\alpha-1} p_{t_i-1}(i-n)$. Similarly to the proof of Lemma 3 above this can be shown to imply

$$p_{t_i}(i) \leq \frac{1}{t_i} + \frac{\alpha-\beta}{\alpha-1} p_{t_i}(i-1) + \frac{\beta}{\alpha-1} p_{t_i}(i-n), \tag{5}$$
$$\ell_{t_i}(i) \leq 1 + \frac{\alpha-\beta}{\alpha-1} \ell_{t_i}(i-1) + \frac{\beta}{\alpha-1} \ell_{t_i}(i-n). \tag{6}$$

At time step $t_i + 1$ action i is not a best response to the opponent's strategy. By Lemma 1, $i + 1$ is the unique best response and so the expected payoff of $i + 1$ is better than the expected payoff of i, thus implying that

$$p_{t_i}(i) \geq \frac{\alpha-\beta}{\alpha-1} p_{t_i}(i-1) + \frac{\beta}{\alpha-1} p_{t_i}(i-n) \geq \frac{\alpha-\beta}{\alpha-1} p_{t_i}(i-1), \tag{7}$$
$$\ell_{t_i}(i) \geq \frac{\alpha-\beta}{\alpha-1} \ell_{t_i}(i-1) + \frac{\beta}{\alpha-1} \ell_{t_i}(i-n) \geq \frac{\alpha-\beta}{\alpha-1} \ell_{t_i}(i-1). \tag{8}$$

Similarly to Lemma 3, the claim follows from (5), (6), (7) and (8), the definition of t_i and the fact that, by Lemma 1, there are no two disjoint sub-intervals of $[1, t^\star]$ such that an action belonging to $\{3n + 1, \ldots, 4n - 1\}$ is a best response. □

The next lemma shows that we can "forget" about the first $2n$ actions at the cost of paying an exponentially small addend in the payoff function.

Lemma 5 *For any* $\delta > 0, \alpha = 1 + \frac{1}{n^{1-\delta}}$ *and* $\beta = 1 - \frac{1}{n^{2(1-\delta)}}$, $\sum_{j=1}^{2n} p_{t^\star}(j) \leq 2^{-n^\delta}$.

Proof We first rewrite and bound from above the sum of the probabilities we are interested in:

$$\sum_{j=1}^{2n} p_{t^\star}(j) = \sum_{j=1}^{2n}\left[\frac{\ell_{t^\star}(j)}{\sum_{j=1}^{4n} \ell_{t^\star}(j)}\right] = \frac{\sum_{j=1}^{2n} \ell_{t^\star}(j)}{\sum_{j=1}^{4n} \ell_{t^\star}(j)}$$
$$= \frac{1}{\sum_{j=2n+1}^{4n} \ell_{t^\star}(j)} \leq \frac{1}{\ell_{t^\star}(4n-1)}.$$

Note that by Lemmas 2, 3 and 4 we have that

$$\ell_{t^\star}(4n-1) \geq \frac{\alpha-\beta}{\alpha-1}\ell_{t^\star}(4n-2) \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^2 \ell_{t^\star}(4n-3)$$
$$\geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^{3n-1} \ell_{t^\star}(n) = \left(\frac{\alpha-\beta}{\alpha-1}\right)^{3n-1}.$$

By plugging in the values of α and β given in the hypothesis we have that

$$\sum_{j=1}^{2n} p_{t^\star}(j) \leq \frac{1}{\left(\left(1+\frac{1}{n^{1-\delta}}\right)^{n^{1-\delta}}\right)^{3n^\delta - \frac{1}{n^\delta}}} \leq \frac{1}{2^{3n^\delta - \frac{1}{n^\delta}}} \leq \frac{1}{2^{n^\delta}},$$

where the penultimate inequality follows from the observation that the function $(1+1/x)^x > 2$ for $x > 2$. □

The theorem below generalizes the above arguments to the cycles that FP visits in the last block of the game, i.e., the block which comprises actions $S = \{2n+1, \ldots, 4n\}$. Since we focus on this part of the game, to ease the presentation, our notation uses circular arithmetic on the elements of S. For example, the action $j+2$ will denote action $2n+2$ for $j = 4n$ and the action $j-n$ will be the action $3n+1$ for $j = 2n+1$. Note that under this notation $j-2n = j+2n = j$ for each action j in the block.

Theorem 1 *For any $\delta > 0$, $\alpha = 1 + \frac{1}{n^{1-\delta}}$ and $\beta = 1 - \frac{1}{n^{2(1-\delta)}}$, n sufficiently large, any $t \geq t^\star$ we have*

$$\frac{p_t(i)}{p_t(i-1)} \geq 1 + \frac{1}{n^{1-\delta}} \quad \textit{for all } i \in S \textit{ with } i \neq s_t, s_t+1,$$
$$\frac{p_t(i)}{p_t(i-1)} \leq 1 + \frac{3}{n^{1-\delta}} \quad \textit{for all } i \in S.$$

Proof The proof proceeds by induction on t.
Base For the base of the induction, consider $t = t^\star$ and note that at that point $s_{t^\star} = 4n$ and $s_{t^\star} + 1 = 2n+1$. Therefore we need to show the lower bound for any action $i \in \{2n+2, \ldots, 4n-1\}$. From Lemmas 3 and 4 we note that for $i \neq 4n, 2n+1$,

$$\frac{p_{t^\star}(i)}{p_{t^\star}(i-1)} \geq \frac{\alpha-\beta}{\alpha-1} = 1 + \frac{1}{n^{1-\delta}}.$$

As for the upper bound, we first consider the case of $i \neq 4n, 2n+1$. Lemma 3 implies that for $i = 2n+2, \ldots, 3n$,

$$\frac{p_{t^\star}(i)}{p_{t^\star}(i-1)} \leq \frac{1}{t^\star} + \frac{\alpha-\beta}{\alpha-1},$$

while Lemma 4 implies that for $i = 3n+1, \ldots, 4n-1$,

$$\frac{p_{t^\star}(i)}{p_{t^\star}(i-1)} \leq \frac{1}{t^\star} + \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t^\star}(i-n)}{p_{t^\star}(i-1)} = \frac{1}{t^\star} + \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{\ell_{t^\star}(i-n)}{\ell_{t^\star}(i-1)}. \tag{9}$$

To give a unique upper bound for both cases, we only focus on (9) (since it is weaker) and next focus on the ratio $\frac{\ell_{t^\star}(i-n)}{\ell_{t^\star}(i-1)}$. We use Lemmas 3 and 4 and get

$$\ell_{t^\star}(i-1) \geq \frac{\alpha-\beta}{\alpha-1}\ell_{t^\star}(i-2) \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^2 \ell_{t^\star}(i-3) \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^{n-1} \ell_{t^\star}(i-n).$$

By setting α and β as in the hypothesis and noticing that $t^\star \geq n \geq n^{1-\delta}$ we then obtain that

$$\frac{p_{t^\star}(i)}{p_{t^\star}(i-1)} \leq 1 + \frac{2}{n^{1-\delta}} + \left(1+\frac{1}{n^{1-\delta}}\right)^{1-n} \frac{n^{2(1-\delta)}-1}{n^{1-\delta}}. \tag{10}$$

We end this part of the proof by showing that the last addend on the right-hand side of (10) is bounded from above by $\frac{1}{4n^{1-\delta}}$. To do so we need to prove

$$\left(1+\frac{1}{n^{1-\delta}}\right)^{1-n} \leq \frac{1}{4n^{1-\delta}}\frac{n^{1-\delta}}{n^{2(1-\delta)}-1}, \tag{11}$$

which is equivalent to

$$\left(\left(1+\frac{1}{n^{1-\delta}}\right)^{n^{1-\delta}}\right)^{n^\delta - \frac{1}{n^{1-\delta}}} \geq 4(n^{2(1-\delta)}-1).$$

We now bound from below the left-hand side of the latter inequality:

$$\left(\left(1+\frac{1}{n^{1-\delta}}\right)^{n^{1-\delta}}\right)^{n^\delta - \frac{1}{n^{1-\delta}}} > \frac{2^{n^\delta}}{2^{\frac{1}{n^{1-\delta}}}} > \frac{2^{n^\delta}}{2},$$

where the first inequality follows from the fact that the function $(1+1/x)^x$ is greater than 2 for $x > 2$ and the second one follows from the fact that $2^{\frac{1}{n^{1-\delta}}} < 2$ for $n^{1-\delta} > 1$. Then, since for $n \geq \sqrt[2(1-\delta)]{4}$, $5n^{2(1-\delta)} \geq 4(n^{2(1-\delta)}-1)$, to prove (11) it is enough to show

$$2^{n^\delta} \geq 2(5n^{2(1-\delta)}) \iff n^\delta \geq 2(1-\delta)\,\log_2(10n).$$

To prove the latter, since $\delta > 0$, it is enough to observe that the function n^δ is certainly bigger than the function $2\log_2(10n) > 2(1-\delta)\log_2(10n)$ for n large enough (e.g., for $\delta = 1/2$, this is true for $n > 639$).

The following claim, whose proof is given in an appendix, concludes the proof of the base of the induction.

Claim 1 *The upper bound holds at time step $t^\star$ for $i = 4n, 2n+1$.*

Induction step Now we assume the claim is true until time step $t-1$ and we show it for time step t. By the inductive hypothesis, the following is true, with $j \neq s_{t-1}, s_{t-1}+1$

$$1 + \frac{1}{n^{1-\delta}} \leq \frac{p_{t-1}(j)}{p_{t-1}(j-1)} \leq 1 + \frac{3}{n^{1-\delta}}, \tag{12}$$

$$\frac{p_{t-1}(s_{t-1})}{p_{t-1}(s_{t-1}-1)} \leq 1 + \frac{3}{n^{1-\delta}},$$

$$\frac{p_{t-1}(s_{t-1}+1)}{p_{t-1}(s_{t-1})} \leq 1 + \frac{3}{n^{1-\delta}}. \tag{13}$$

We first consider the case in which $s_t \neq s_{t-1}$. By Lemma 1, the action played at time t is $s_{t-1}+1$, i.e., $s_t = s_{t-1}+1$. Let $s_{t-1} = i$ and then we have $s_t = i+1$. By the inductive hypothesis, for all the actions $j \neq i, i+1, i+2$ we have

$$\frac{\alpha - \beta}{\alpha - 1} = 1 + \frac{1}{n^{1-\delta}} \leq \frac{p_t(j)}{p_t(j-1)} \leq 1 + \frac{3}{n^{1-\delta}}. \tag{14}$$

Indeed, for these actions j, $\ell_{t-1}(j) = \ell_t(j)$ and $\ell_{t-1}(j-1) = \ell_t(j-1)$. Therefore the probabilities of j and $j-1$ at time t are simply those at time $t-1$ rescaled by the same amount and the claim follows from (12). The upper bound on the ratio $\frac{p_t(i+2)}{p_t(i+1)}$ easily follows from the upper bound in (12) as $\ell_{t-1}(i+2) = \ell_t(i+2)$ and $\ell_{t-1}(i+1) < \ell_t(i+1) = \ell_{t-1}(i+1)+1$. However, as $s_t = i+1$ here we need to prove lower and upper bound also for the ratio $\frac{p_t(i)}{p_t(i-1)}$ and the upper bound for the ratio $\frac{p_t(i+1)}{p_t(i)}$.

Claim 2 *The following inequalities hold:* $1 + \frac{1}{n^{1-\delta}} \leq \frac{p_t(i)}{p_t(i-1)} \leq 1 + \frac{3}{n^{1-\delta}}$.

The proof of Claim 2 is in the Appendix.

Claim 3 *It holds that* $\frac{p_t(i+1)}{p_t(i)} \leq 1 + \frac{3}{n^{1-\delta}}$.

The proof of Claim 3 is in the Appendix.

Finally, we consider the case in which $s_{t-1} = s_t$. In this case, for the actions $j \neq s_t, s_t+1$ it holds $\ell_{t-1}(j) = \ell_t(j)$ and $\ell_{t-1}(j-1) = \ell_t(j-1)$. Therefore, similarly to the above, for these actions j the claim follows from (12). The upper bound for the ratio $\frac{p_t(s_t+1)}{p_t(s_t)}$ easily follows from (13) as $\ell_{t-1}(s_t+1) = \ell_t(s_t+1)$ and $\ell_{t-1}(s_t) < \ell_t(s_t) = \ell_{t-1}(s_t)+1$. The remaining case to analyze is the upper bound on the ratio $\frac{p_t(s_t)}{p_t(s_t-1)}$. To prove this we can use *mutatis mutandis* the proof of the upper bound contained in Claim 2 with $s_t = i$. □

The claimed performance of FP, in terms of the approximation to the best response that it computes, follows directly from the following theorem.

Theorem 2 *For any $\delta > 0$ and any time step t, FP returns an ϵ-Nash equilibrium with $\epsilon \geq \frac{1}{2} - O\left(\frac{1}{n^{1-\delta}}\right)$.*

Proof For $t \leq n$ the result follows since the game is similar to Conitzer (2009). In detail, for $t \leq n$ the payoff associated with the best response, which in this case is $s_t + 1$, is bounded from above by 1. On the other hand, the payoff associated with the current strategy prescribed by FP is bounded from below by $\frac{\beta}{i^2}\sum_{j=0}^{i-1} j$ where $i = s_t$. Therefore, the regret of either player normalized to the [0, 1] interval satisfies $\epsilon \geq \frac{1}{\alpha} - \frac{\beta}{\alpha}\frac{i-1}{2i}$. Since $\frac{i-1}{2i} < 1/2$, the fact that $1 - \frac{\beta}{2} - \frac{\alpha}{2} + \frac{\alpha}{n^{1-\delta}} \geq 0$ (which is true given the values of α and β) yields the claim. For $t \leq t^\star$ the result follows from Lemmas 3 and 4; while the current action s_t (for $t \leq t^*$) has payoff approximately 1, the players' mixed strategies have nearly all their probability on the recently played actions, but with no action having very high probability, so that some player is likely to receive zero payoff; by symmetry each player has a payoff of approximately $\frac{1}{2}$. This is made precise below, where it is applied in more detail to the case of $t > t^\star$.

We now focus on the case $t > t^\star$. Recall that for a set of actions S, $p_t(S) = \sum_{i\in S} p_t(i)$. Let S_t be the set $\{2n+1,\ldots,s_t\} \cup \{s_t+n,\ldots,4n\}$ if $s_t \leq 3n$, or the set $\{s_t - n,\ldots,s_t\}$ in the case that $s_t > 3n$. Let $S'_t = \{2n+1,\ldots,4n\} \setminus S_t$. Also, let $s_t^{\max} = \arg\max_{i\in\{2n+1,\ldots,4n\}}(p_t(i))$; note that by Theorem 1, $s_t^{\max}$ is equal to either s_t or s_t^-, where $s_t^- = s_t - 1$ if $s_t > 2n$, or $4n$ if $s_t = 2n$.

We start by establishing the following claim:

Claim 4 *For sufficiently large n, $p_t(S_t) \geq 1 - \frac{2n}{2^{n^\delta}}$.*

Proof To see this, note that for all $x \in S'_t$, by $p_t(s_t^{\max}) \geq p_t(s_t^{\max} - 1)$ and Theorem 1 we have

$$\frac{p_t(s_t^{\max})}{p_t(x)} = \frac{p_t(s_t^{\max})}{p_t(s_t^{\max}-1)}\,\frac{p_t(s_t^{\max}-1)}{p_t(s_t^{\max}-2)}\cdots\frac{p_t(x+1)}{p_t(x)} \geq \left(1 + \frac{1}{n^{1-\delta}}\right)^{k-1},$$

where k is the number of factors on the right-hand side of the equality above, i.e., the number of actions between x and $s_t^{\max}$. Thus, as $k \geq n$,

$$p_t(x) \leq \frac{p_t(s_t^{\max})}{\left(1+\frac{1}{n^{1-\delta}}\right)^{k-1}} \leq \left(1 + \frac{1}{n^{1-\delta}}\right)^{1-k} \leq \left(1 + \frac{1}{n^{1-\delta}}\right)^{1-n} \leq 4^{(1-n)/(n^{1-\delta})}.$$

Hence $p_t(S'_t) \leq (2n)4^{(1-n)/(n^{1-\delta})} = \frac{n4^{1/n^{1-\delta}}}{2^{n^\delta}} \leq \frac{2n-1}{2^{n^\delta}}$, where the last inequality follows from the fact that, for large n, $4^{1/n^{1-\delta}} < 2 - \frac{1}{n}$. Then $p_t(S_t) \geq 1 - p_t(S'_t) - p_t(\{1,\ldots,2n\})$, which establishes the claim, since Lemma 5 establishes a strong enough upper bound on $p_t(\{1,\ldots,2n\})$. □

Claim 5 *The current best response at time t, s_t, has payoff at least $\beta\left(1 - \frac{2n}{2^{n^\delta}}\right)$.*

Proof The action s_t receives a payoff of at least β when the opponent plays any action from S_t; the claim follows using Claim 4. □

Let E_t denote the expected payoff to either player that would result if they both select an action from the mixed distribution that allocates to each action x the probability $p_t(x)$. The result will follow from the following claim:

Claim 6 *For sufficiently large n, $E_t \leq \frac{\alpha}{2} + \frac{6}{n^{1-\delta}} + \alpha \frac{2n}{2^{n^\delta}}$.*

Proof The contribution to E_t from actions in $\{1, \dots, 2n\}$, together with actions in S'_t, may be bounded from above by α times the probability that any of those actions get played. By Lemma 5 and Claim 4, this probability is exponentially small, namely $2n/2^{n^\delta}$.

Now we consider the contribution to E_t from the actions in S_t. If they play different actions, their total payoff will be at most α, since one player receives payoff 0. If they play the same action, they both receive payoff 1. We continue by bounding from above the probability that they both play the same action. This is bounded from above by the largest probability assigned to any single action, namely $p_t(s_t^{\max})$.

Suppose for contradiction that $p_t(s_t^{\max}) > 6/n^{1-\delta}$. At this point, note that by Theorem 1, for any action $s \in S_t$, we have

$$\frac{p_t(s_t^{\max})}{p_t(s)} \leq \left(1 + \frac{3}{n^{1-\delta}}\right)^k,$$

where k is the number of actions whose index is in between the indices of s and $s_t^{\max}$; note that as in the proof of Theorem 1 we are using circular arithmetic for the indices of actions in S_t. Therefore, denoting $r = \left(1 + \frac{3}{n^{1-\delta}}\right)^{-1}$, we obtain

$$p_t(S_t) = \sum_{s \in S_t} p_t(s) = p_t(s_t) + \sum_{i=s_t-n}^{s_t-1} p_t(i) \geq p_t(s_t^{\max}) \sum_{k=0}^{n-1} r^k.$$

Applying the standard formula for the partial sum of a geometric series we have

$$p_t(S_t) \geq \frac{6}{n^{1-\delta}} \left(\frac{1-r^n}{1-r}\right).$$

Noting that $1 - r^n > \frac{1}{2}$ we have $p_t(S_t) > \frac{6}{n^{1-\delta}} \cdot (\frac{1}{2}) \cdot \left(\frac{n^{1-\delta}}{3}\right)$ which is greater than 1, a contradiction.

The expected payoff E_t to either player is, by symmetry, half the expected total payoff, so we have $E_t \leq \left(1 - \frac{2n}{2^{n^\delta}} - \frac{6}{n^{1-\delta}}\right) \frac{\alpha}{2} + \frac{6}{n^{1-\delta}} + \frac{2n}{2^{n^\delta}} \alpha$, which yields the claim. □

We now show that FP never achieves an ϵ-value better than $\frac{1}{2} - O\left(\frac{1}{n^{1-\delta}}\right)$. From the last two claims the regret of either player normalized to [0, 1] is

$$\begin{aligned}
\epsilon &\geq \frac{\beta}{\alpha}\left(1 - \frac{2n}{2^{n^\delta}}\right) - \frac{1}{2} - \frac{6}{\alpha n^{1-\delta}} - \frac{2n}{2^{n^\delta}} \\
&= \left(1 - \frac{n^{1-\delta}+1}{n^{2(1-\delta)} + n^{1-\delta}}\right)\left(1 - \frac{2n}{2^{n^\delta}}\right) - \frac{1}{2} - \frac{6}{n^{1-\delta}+1} - \frac{2n}{2^{n^\delta}}
\end{aligned}$$

$$= \frac{1}{2} - \frac{n^{1-\delta}+1}{n^{2(1-\delta)}+n^{1-\delta}} - \frac{n^{1-\delta}+1}{n^{2(1-\delta)}+n^{1-\delta}} \frac{2n}{2^{n^\delta}} - \frac{6}{n^{1-\delta}+1} - \frac{2n}{2^{n^\delta}}$$
$$= \frac{1}{2} - O\left(\frac{1}{n^{1-\delta}}\right).$$

This concludes the proof. □

3 Upper bound

In this section, we consider an $m \times n$ game. Let a, b denote the FP sequences of actions of length t, for players 1 and 2 respectively. For two iterations $k, l \in \{1, \ldots, t\}$ with $k \leq l$, let $a_{[k:\ell]}$ denote the subsequence of actions $a_k, \ldots, a_\ell$. We overload notation and use a to also denote the mixed strategy that is uniform on the corresponding sequence of actions.

An action that appears most recently in a player's sequence has an ϵ-value close to 0 (at most $1/t$); generally an action that occurs a fraction γ back in the sequence has an ϵ-value of at most slightly more than γ (it is a best response to slightly less than $1-\gamma$ of the opponent's distribution). When the number of available actions is exceeded by the number of iterations of FP, various actions must get re-used, and this re-usage means, for example, that every previous occurrence of the most recent action has ϵ-value $\frac{1}{t}$. Thus, to derive an upper bound on ϵ, we consider the most recent occurrences of each action in a. For $k \in \{1, \ldots, t\}$, let $f_k(a)$ denote the most recent occurrence of a_k in the sequence a, that is,

$$f_k(a) := \max_{\ell \in \{1,\ldots,t\},\ a_\ell = a_k} \ell.$$

For brevity, we sometimes write f_k in place of $f_k(a)$. Let z be a best response against b, and let ϵ denote the smallest ϵ for which a is an ϵ-best-response against b. We have the following.

$$\epsilon = u_1(z, b) - u_1(a, b) \frac{1}{t} \sum_{i=1}^{t} (u_1(z, b) - u_1(a_i, b))$$
$$= \frac{1}{t} \sum_{i=1}^{t} \left[\frac{f_i - 1}{t} \big(u_1(z, b_{[1:f_i-1]}) - u_1(a_i, b_{[1:f_i-1]})\big) \right.$$
$$\left. + \frac{t - f_i + 1}{t} \big(u_1(z, b_{[f_i:t]}) - u_1(a_i, b_{[f_i:t]})\big) \right]$$
$$\leq \frac{1}{t} \sum_{i=1}^{t} \left[\frac{t - f_i + 1}{t} \big(u_1(z, b_{[f_i:t]}) - u_1(a_i, b_{[f_i:t]})\big) \right] \tag{15}$$
$$\leq \frac{1}{t} \sum_{i=1}^{t} \frac{t - f_i + 1}{t} \tag{16}$$

$$= 1 + \frac{1}{t} - \frac{1}{t^2} \sum_{i=1}^{t} f_i. \tag{17}$$

Inequality (15) holds since, by definition, a_i is a best response against $b_{[1:f_i-1]}$. Inequality (16) holds since payoffs are in the range [0, 1]. To provide a guarantee on the performance of FP, we find a sequence that maximizes the RHS of (17), i.e., that minimizes $\sum_{i=1}^{t} f_i(a)$.

Definition 1 For an FP sequence a, let $S(a) := \sum_{i=1}^{t} f_i(a)$ and let $\hat{a} \in \arg\min_a S(a)$.

The following three lemmas allow us to characterize $\hat{a}$.

Lemma 6 *The entries of $\hat{a}$ take on exactly m distinct values.*

Proof The entries of an FP sequence of player 1 can take on at most m distinct values, since player 1 has only m actions. Suppose for the sake of contradiction that the entries of $\hat{a}$ take on strictly less than m distinct values. Then there is an action, say z, that does not appear in $\hat{a}$ and an action z' that appears more than t/m times. Obtain a from $\hat{a}$ by replacing a single occurrence of z' in $\hat{a}$ with z. Then $S(a) < S(\hat{a})$, a contradiction. □

We now define a reordering of the actions in an FP sequence a to give new sequence a' so that $S(a') < S(a)$ if $a \neq a'$. The reordering groups together occurrences of actions and orders them according to their last occurrences in a.

Definition 2 Suppose the entries of a take on d distinct values. We define $x_1, \ldots, x_d$ to be the last occurrences, $\{f_i(a) \mid i \in [t]\}$, in ascending order. Formally, let $x_d := a_t$ and for $k < d$ let $x_k := a_i$ be such that

$$i := \arg\max_{j=1,\ldots,t} \{a_j \mid a_j \notin \{x_{k+1}, \ldots, x_d\}\}.$$

For $i = 1, \ldots, d$, let

$$\#(x_i) := |\{a_j \mid j \in [t], a_j = x_i\}|,$$

which is the number of occurrences of x_i in a. Define a' as

$$a' := \underbrace{x_1, \ldots, x_1}_{\#(x_1)}, \underbrace{x_2, \ldots, x_2}_{\#(x_2)}, \ldots, \underbrace{x_d, \ldots, x_d}_{\#(x_d)}.$$

Lemma 7 *For any FP sequence a, let a' be as in Definition 2. If $a' \neq a$ then $S(a') < S(a)$.*

Proof For all $i = 1, \ldots, t$ we have $f_i(a') \leq f_i(a)$, and since $a' \neq a$ there is at least one i such that $f_i(a') < f_i(a)$.

Lemma 8 *Let $m, t \in \mathbb{N}$ be such that $m|t$. Let a be a sequence of length t of the form*

$$a = \underbrace{1, \ldots, 1}_{c_1}, \underbrace{2, \ldots, 2}_{c_2}, \ldots, \underbrace{m, \ldots, m}_{c_m}.$$

Then $S(a)$ is minimized if and only if $c_1 = \cdots = c_m = t/m$.

Proof We refer to the maximal length subsequence of entries with value $u \in \{1, \dots m\}$ as *block* u. Consider two adjacent blocks u and $u+1$, where block u starts at i and ends at $j-1$ and block $u+1$ starts at j and finishes at k. The contribution of these two blocks to $S(a)$ is

$$\sum_{i}^{j-1}(j-1) + \sum_{j}^{k} k = j^2 - (1+i+k)j + (i+k+k^2).$$

If $1+i+k$ is even, this contribution is minimized when $j = \frac{1+i+k}{2}$. If $1+i+k$ is odd, this contribution is minimized for both values $j = \lfloor \frac{1+i+k}{2} \rfloor$ and $j = \lceil \frac{1+i+k}{2} \rceil$.

Now suppose for the sake of contradiction that $S(a)$ is minimized when $c_1 = \cdots = c_m = t/m$ does not hold. There are two possibilities. Either there are two adjacent blocks whose lengths differ by more than one, in which case we immediately have a contradiction. If not, then it must be the case that all pairs of adjacent blocks differ in length by at most one. In particular, there must be a block of length $t/m+1$ and another of length $t/m-1$ with all blocks in between of length t/m. Flipping the leftmost of these blocks with its right neighbor will not change the sum $S(a)$. Repeatedly doing this until the blocks of lengths $t/m+1$ and $t/m-1$ are adjacent, does not change $S(a)$. Then we have two adjacent blocks that differ in length by more than one, which contradicts the fact that $S(a)$ was minimized. □

Lemma 9 *For $m|t$ and FP strategies (a, b) we have that a is an ϵ^*-best-response against b. where*

$$\epsilon^* = \frac{1}{2} + \frac{1}{t} - \frac{1}{2m}.$$

Proof We show that $\hat{a}$ is an ϵ^*-best-response against b. Applying Lemma 6, Lemma 7 and Lemma 8, we have that

$$\hat{a} = \underbrace{z_1, \dots, z_1}_{t/m}, \underbrace{z_2, \dots, z_2}_{t/m}, \dots, \underbrace{z_n, \dots, z_n}_{t/m},$$

where $z_1, \dots, z_m$ is an arbitrary labeling of player 1's actions. Using (17), we have that

$$\begin{aligned}\epsilon &\leq 1 + \frac{1}{t} - \frac{1}{t^2}\sum_{i=1}^{t} f_i(\hat{a}) = 1 + \frac{1}{t} - \frac{1}{t^2}\frac{t}{m}\sum_{i=1}^{m}\left(\frac{i \cdot t}{m}\right) \\ &= 1 + \frac{1}{t} - \frac{m+1}{2m} = \frac{1}{2} + \frac{1}{t} - \frac{1}{2m}.\end{aligned}$$

This concludes the proof. □

Combining Lemma 9 with the analogous result for player 2's sequence b, we have the following.

Theorem 3 *If $m|t$ and $n|t$, the FP strategies (a, b) are an ϵ^*-equilibrium, where*

$$\epsilon^* = \frac{1}{2} + \frac{1}{t} - \frac{1}{2\max(m, n)}.$$

For a square $n \times n$ game and t superlinear in n, we achieve a $(1/2 - O(1/n))$-approximate Nash equilibrium.

4 Conclusions

FP is a classical decentralized algorithm to solve games. In this paper, we analyze the quality of strategy profiles output by FP in terms of their approximation guarantee as ϵ-Nash equilibria. We exhibit a class of two-player n-action games with payoffs in the range [0, 1] for which FP fails to guarantee an approximation better than $1/2 - O(1/n^{1-\delta})$ for any $\delta > 0$. We complement this result by showing that FP returns $(1/2 - O(1/n))$-approximate Nash equilibria thus proving that our lower bound is essentially tight.

Our result suggests that further work on the topic might address the question of whether $\frac{1}{2}$ is a fundamental limit to the approximation performance obtainable by certain types of algorithms that are in some sense simple or decentralized. We noted in the introduction that $\epsilon = \frac{1}{2}$ is achieved by the simple algorithm of Daskalakis et al. (2009a), but smaller values of ϵ seem to require more complex algorithms that solve an LP. The question of specifying appropriate classes of algorithms is itself challenging; it is also considered in Daskalakis et al. (2010) in the context of algorithms that (without computational complexity theoretic assumptions) provably fail to find Nash equilibria. This research agenda is likely to be challenging, as evidenced by the fact that we had to construct a highly-artificial class of games to produce the worst-case performance for FP. A more general result would seemingly have to rely on a general construction of worst-case input games for unspecified algorithms belonging to some class.

The alternative direction is to aim for more positive results, saying that some variant of FP can do rather better. We noted in the introduction that in our games, if FP started at a different initial pure-strategy profile, we might immediately have a good approximation. It may be that this, or some other simple "smarter" version of FP could in fact get improved approximation guarantees. For that purpose, we hope that the class of games exhibited here can help to indicate the kind of obstacles that may need to be overcome.

Acknowledgements This work was supported by EPSRC grants EP/G069239/1 and EP/G069034/1 "Efficient Decentralised Approaches in Algorithmic Game Theory."

Appendix A: Proof of Claim 1

Proof We first show the claim for $i = 4n$. At time step $t^\star$, FP prescribes to play $4n$. This in particular means that the action $4n$ achieves a payoff which is at least as much as that of action $2n + 1$ after $t^\star - 1$ time steps. We write down the inequality given by this fact focusing only on the last $2n$ actions (we will consider actions $1, 2, \ldots, n$ below) and obtain:

$$p_{t^\star-1}(4n) + \alpha p_{t^\star-1}(4n-1) + \beta p_{t^\star-1}(3n) \geq \alpha p_{t^\star-1}(4n) \\ + p_{t^\star-1}(2n+1) + \beta p_{t^\star-1}(4n-1) \tag{18}$$

and then since $\alpha > 1$

$$\frac{p_{t^\star-1}(4n)}{p_{t^\star-1}(4n-1)} \leq \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t^\star-1}(3n)}{p_{t^\star-1}(4n-1)} - \frac{1}{\alpha-1}\frac{p_{t^\star-1}(2n+1)}{p_{t^\star-1}(4n-1)}.$$

Similarly to the proof of Lemma 3 above this can be shown to imply

$$\begin{aligned}\frac{p_{t^\star}(4n)}{p_{t^\star}(4n-1)} &\leq \frac{1}{t^\star} + \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t^\star}(3n)}{p_{t^\star}(4n-1)} - \frac{1}{\alpha-1}\frac{p_{t^\star}(2n+1)}{p_{t^\star}(4n-1)} \\ &\leq \frac{1}{t^\star} + \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t^\star}(3n)}{p_{t^\star}(4n-1)}.\end{aligned} \tag{19}$$

We now bound from above the ratio $\frac{\beta}{\alpha-1}\frac{p_{t^\star}(3n)}{p_{t^\star}(4n-1)}$. By repeatedly using Lemmas 3 and 4 we have that

$$\begin{aligned}p_{t^\star}(4n-1) &\geq \frac{\alpha-\beta}{\alpha-1}p_{t^\star}(4n-2) \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^2 p_{t^\star}(4n-3) \\ &\geq \cdots \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^{n-1} p_{t^\star}(3n).\end{aligned}$$

This yields

$$\frac{\beta}{\alpha-1}\frac{p_{t^\star}(3n)}{p_{t^\star}(4n-1)} \leq \frac{\beta}{\alpha-1}\left(\frac{\alpha-1}{\alpha-\beta}\right)^{n-1} \leq \frac{1}{4n^{1-\delta}},$$

where the last inequality has been proved above (see 11). Therefore, since $t \geq n^{1-\delta}$, (19) implies

$$\frac{p_{t^\star}(4n)}{p_{t^\star}(4n-1)} \leq 1 + \frac{2}{n^{1-\delta}} + \frac{1}{4n^{1-\delta}}.$$

To conclude this part of the proof we must now consider the contribution to (18) of the actions $1, \ldots, 2n$ that are not in the last block. However, Lemma 5 shows that all those actions are played with probability $1/2^{n^\delta}$ at time $t^\star$. Thus the overall contribution of these actions is bounded from above by $\frac{1}{2^{n^\delta}}(\alpha-\beta) \leq \frac{1}{2^{n^\delta}}$. Similarly to the above, we observe that, for n sufficiently large, $n^\delta \geq \log_2(4n) \geq (1-\delta)\log_2(4n)$, which implies that $\frac{1}{2^{n^\delta}} \leq \frac{1}{4n^{1-\delta}}$. This concludes the proof of the upper bound at time $t^\star$ for $i = 4n$.

Consider now the case $i = 2n+1$. At time step $t^\star + 1$, $4n$ is not played by FP, which means that $4n$ is not a best response after $t^\star$ time steps. By Lemma 1, the best response is $2n+1$; then, in particular, the payoff of $2n+1$ is not smaller than the payoff of $4n$ at that time. We write down the inequality given by this fact focusing only on the last $2n$ actions (we will consider the actions in $\{1, 2, \ldots, 2n\}$ below) and obtain

$$p_{t^\star}(4n) + \alpha p_{t^\star}(4n-1) + \beta p_{t^\star}(3n) \leq \alpha p_{t^\star}(4n) + p_{t^\star}(2n+1) + \beta p_{t^\star}(4n-1)$$

and then since $\alpha > 1$

$$\frac{p_{t^\star}(4n)}{p_{t^\star}(4n-1)} \geq \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t^\star}(3n)}{p_{t^\star}(4n-1)} - \frac{1}{\alpha-1}\frac{p_{t^\star}(2n+1)}{p_{t^\star}(4n-1)}. \tag{20}$$

We next show that $\frac{\beta p_{t^\star}(3n) - p_{t^\star}(2n+1)}{(\alpha-1)p_{t^\star}(4n-1)} \geq -\frac{1}{4n^{1-\delta}}$ or equivalently that $\frac{p_{t^\star}(3n)}{p_{t^\star}(2n+1)} \geq \frac{1}{\beta} - \frac{(\alpha-1)p_{t^\star}(4n-1)}{4\beta n^{1-\delta} p_{t^\star}(2n+1)}$. To prove this it is enough to show that $\frac{p_{t^\star}(3n)}{p_{t^\star}(2n+1)} \geq \frac{1}{\beta}$. We observe that

$$\frac{p_{t^\star}(3n)}{p_{t^\star}(2n+1)} = \frac{p_{t^\star}(3n)}{p_{t^\star}(3n-1)}\frac{p_{t^\star}(3n-1)}{p_{t^\star}(3n-2)} \cdots \frac{p_{t^\star}(2n+2)}{p_{t^\star}(2n+1)} \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^{n-1} \geq \frac{1}{\beta},$$

where the first inequality follows from Lemma 3 and the second inequality follows from the observation (similar to the above) that for n sufficiently large $n^\delta \geq 2\log_2(2n)$. Then to summarize, for α and β as in the hypothesis, (20) implies that

$$\frac{p_{t^\star}(4n)}{p_{t^\star}(4n-1)} \geq 1 + \frac{1}{n^{1-\delta}} - \frac{1}{4n^{1-\delta}}.$$

As above we consider actions $1, \ldots, 2n$ and observe that their contribution to the payoffs is bounded from above by $\frac{1}{4n^{1-\delta}}$. Now to conclude the proof of the claim for the case $i = 2n+1$ we simply notice that the above implies $p_{t^\star}(4n-1) < p_{t^\star}(4n)$ and Lemmas 3 and 4 imply that $p_{t^\star}(2n+1) < p_{t^\star}(4n-1)$ which together prove the claim. □

Appendix B: Proof of Claim 2

Proof To prove the claim we first focus on the last block of the game, i.e., the block in which players have actions in $\{2n+1, \ldots, 4n\}$. Recall that our notation uses circular arithmetic on the number of actions of the block.

The fact that action $i+1$ is better than action i after $t-1$ time steps implies that

$$p_{t-1}(i) + \alpha p_{t-1}(i-1) + \beta p_{t-1}(i-n) \leq \alpha p_{t-1}(i) + p_{t-1}(i+1) + \beta p_{t-1}(i-1),$$

and then since $\alpha > 1$,

$$\frac{p_{t-1}(i)}{p_{t-1}(i-1)} \geq \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t-1}(i-n)}{p_{t-1}(i-1)} - \frac{1}{\alpha-1}\frac{p_{t-1}(i-2n+1)}{p_{t-1}(i-1)}. \tag{21}$$

We next show that $\frac{\beta p_{t-1}(i-n)-p_{t-1}(i-2n+1)}{(\alpha-1)p_{t-1}(i-1)} \geq -\frac{1}{4n^{1-\delta}}$ or equivalently

$$\frac{p_{t-1}(i-n)}{p_{t-1}(i-2n+1)} \geq \frac{1}{\beta} - \frac{(\alpha-1)p_{t-1}(i-1)}{4\beta n^{1-\delta} p_{t-1}(i-2n+1)}.$$

To prove this it is enough to show that $\frac{p_{t-1}(i-n)}{p_{t-1}(i-2n+1)} \geq \frac{1}{\beta}$. We observe that

$$\frac{p_{t-1}(i-n)}{p_{t-1}(i-2n+1)} = \frac{p_{t-1}(i-n)}{p_{t-1}(i-n-1)} \cdots \frac{p_{t-1}(i-2n+2)}{p_{t-1}(i-2n+1)} \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^{n-1} \geq \frac{1}{\beta},$$

where the first inequality follows from inductive hypothesis (we can use the inductive hypothesis as all the actions involved above are different from i and $i+1$) and the second inequality follows from the aforementioned observation that, for sufficiently large n, $n^{\delta} \geq 2\log_2(2n)$. Then to summarize, for α and β as in the hypothesis, (21) implies that

$$\frac{p_t(i)}{p_t(i-1)} = \frac{p_{t-1}(i)}{p_{t-1}(i-1)} \geq 1 + \frac{1}{n^{1-\delta}} - \frac{1}{4n^{1-\delta}},$$

where the first equality follows from $\ell_{t-1}(i) = \ell_t(i)$ and $\ell_{t-1}(i-1) = \ell_t(i-1)$, which are true because $s_t = i+1$.

Since action $i+1$ is worse than action i at time step $t-1$ we have that

$$p_{t-1}(i) + \alpha p_{t-1}(i-1) + \beta p_{t-1}(i-n) \geq \alpha p_{t-1}(i) + p_{t-1}(i+1) + \beta p_{t-1}(i-1)$$

and then since $\alpha > 1$

$$\frac{p_{t-1}(i)}{p_{t-1}(i-1)} \leq \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_{t-1}(i-n)}{p_{t-1}(i-1)} - \frac{1}{\alpha-1}\frac{p_{t-1}(i-2n+1)}{p_{t-1}(i-1)}.$$

Similarly to the proof of Lemma 3 above this can be shown to imply

$$\begin{aligned}\frac{p_t(i)}{p_t(i-1)} &\leq \frac{1}{t} + \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_t(i-n)}{p_t(i-1)} - \frac{1}{\alpha-1}\frac{p_t(i-2n+1)}{p_t(i-1)} \\ &\leq \frac{1}{t} + \frac{\alpha-\beta}{\alpha-1} + \frac{\beta}{\alpha-1}\frac{p_t(i-n)}{p_t(i-1)}. \end{aligned} \tag{22}$$

We now bound from above the ratio $\frac{\beta}{\alpha-1}\frac{p_t(i-n)}{p_t(i-1)}$. By repeatedly using the inductive hypothesis (14) we have that

$$p_t(i-1) \geq \frac{\alpha-\beta}{\alpha-1}p_t(i-2) \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^2 p_t(i-3) \geq \left(\frac{\alpha-\beta}{\alpha-1}\right)^{n-1} p_t(i-n).$$

(Note again that we can use the inductive hypothesis as none of the actions above is i or $i+1$.) This yields

$$\frac{\beta}{\alpha-1}\frac{p_t(i-n)}{p_t(i-1)} \leq \frac{\beta}{\alpha-1}\left(\frac{\alpha-1}{\alpha-\beta}\right)^{n-1} \leq \frac{1}{4n^{1-\delta}},$$

where the last inequality is proved above (see 11). Therefore, since $t \geq n^{1-\delta}$, (22) implies that

$$\frac{p_t(i)}{p_t(i-1)} \leq 1 + \frac{2}{n^{1-\delta}} + \frac{1}{4n^{1-\delta}}.$$

To conclude the proof we must now consider the contribution to the payoffs of the actions $1, \ldots, 2n$ that are not in the last block. However, Lemma 5 shows that all those actions are played with probability $1/2^{n^\delta}$ at time $t^\star$. Since we prove above (see Lemma 1) that these actions are not played anymore after time step $t^\star$ this implies that $\sum_{j=1}^{2n} p_t(j) \leq \sum_{j=1}^{2n} p_{t^\star}(j) \leq 2^{-n^\delta}$. Thus the overall contribution of these actions is bounded from above by $\frac{1}{2^{n^\delta}}(\alpha-\beta) \leq \frac{1}{2^{n^\delta}} \leq \frac{1}{4n^{1-\delta}}$ where the last bound follows from the aforementioned fact that, for n sufficiently large, $n^\delta \geq (1-\delta)\log_2(4n)$. This concludes the proof of Claim 2. □

Appendix C: Proof of Claim 3

Proof From (21) (and subsequent arguments) we get $p_t(i) > p_t(i-1)$ and from (14) we get $p_t(i-1) > p_t(i-2) > \ldots > p_t(i-2n+1) = p_t(i+1)$. Therefore, $p_t(i) > p_t(i+1)$ thus proving the upper bound. □

References

Berger U (2005) Fictitious play in $2 \times n$ games. J Econ Theory 120:139–154

Bosse H, Byrka J, Markakis E (2007) New algorithms for approximate Nash equilibria in bimatrix games. In: Proceedings of the 3rd international workshop on internet and network economics (WINE), pp 17–29

Brandt F, Fischer F. Harrenstein P (2010) On the rate of convergence of fictitious play. In: Proceedings of the 3rd symposium on algorithmic game theory (SAGT), pp 102–113

Chen X, Deng X, Teng S-H (2009) Settling the complexity of computing two-player Nash equilibria. J ACM 56(3):1–57

Conitzer V (2009) Approximation guarantees for fictitious play. In: Proceedings of the 47th annual Allerton conference on communication, control, and computing, pp 636–643

Daskalakis C, Mehta A, Papadimitriou CH (2007) Progress in approximate Nash equilibria. In: Proceedings of the 8th ACM conference on electronic commerce (EC), pp 355–358

Daskalakis C, Mehta A, Papadimitriou CH (2009a) A note on approximate Nash equilibria. Theor Comput Sci 410(17):1581–1588

Daskalakis C, Goldberg PW, Papadimitriou CH (2009b) The complexity of computing a Nash equilibrium. SIAM J Comput 39(1):195–259

Daskalakis C, Frongillo R, Papadimitriou CH, Pierrakos G, Valiant G (2010) On learning algorithms for Nash equilibria. In Proceedings of the 3rd symposium on algorithmic game theory (SAGT), pp 114–125

Dudziak W (2006) Using fictitious play to find pseudo-optimal solutions for full-scale poker. In: Proceedings of the 2006 international conference on artificial intelligence (ICAI), pp 374–380

Fearnley J, Goldberg PW, Savani R, Sørensen TB (2012) Approximate well-supported Nash equilibria below two-thirds. arXiv tech report 1204.0707. In: Proceedings of the 5th symposium on algorithmic game theory (SAGT)

Feder T, Nazerzadeh H, Saberi A (2007) Approximating Nash equilibria using small-support strategies. In Proceedings of the 8th ACM conference on electronic commerce (EC), pp 352–354

Fudenberg D, Kreps DM (1993) Learning mixed equilibria. Games Econ Behav 5:320–367

Fudenberg D, Levine DK (1995) Consistency and cautious fictitious play. J Econ Dyn Control 19(5–7):1065–1089

Fudenberg D, Levine DK (1998) The theory of learning in games. MIT Press, Cambridge, MA

Ganzfried S, Sandholm T (2008) Computing an approximate jam/fold equilibrium for 3-player no-limit Texas Hold'em tournaments. In: Proceedings of the 7th international joint conference on autonomous agents and multi-agent systems (AAMAS), pp 919–925

Hart S, Mas-Colell A (2003) Uncoupled dynamics do not lead to Nash equilibrium. Am Econ Rev 93(5):1830–1836

Hart S, Mas-Colell A (2006) Stochastic uncoupled dynamics and Nash equilibrium. Games Econ Behav 57(2):286–303

Kontogiannis S, Spirakis PG (2010) Well supported approximate equilibria in bimatrix games. Algorithmica 57(4):653–667

Lambert TJ, Epelman MA, Smith RL (2005) A fictitious play approach to large-scale optimization. Oper Res 53(3):477–489

McMahan HB, Gordon GJ (2007) A fast bundle-based anytime algorithm for poker and other convex games. In: Proceedings of the eleventh international conference on artificial intelligence and statistics (AISTATS)

Miyasawa K (1961) On the convergence of the learning process in a 2×2 nonzero sum two-person game. Research Memo 33. Princeton University, Princeton

Monderer D, Shapley LS (1996) Potential games. Games Econ Behav 14(1):124–143

Robinson J (1951) An iterative method of solving a game. Ann Math 54(2):296–301

Shapley L (1964) Some topics in two-person games. In: Dresher M, Shapley LS, Tucker AW (eds) Advances in game theory, vol 52. Annals of Mathematics Studies. Princeton University Press, Princeton, pp 1–29

Tsaknakis H, Spirakis Paul G (2008) An optimization approach for approximate Nash equilibria. Internet Math 5(4):365–382