Evolution toward a Nash equilibrium

Ioannis Avramopoulos

In this paper, we study the dynamic behavior of Hedge, a well-known algorithm in theoretical machine learning and algorithmic game theory. The empirical average (arithmetic mean) of the iterates Hedge generates is known to converge to a minimax equilibrium in zero-sum games. We generalize that result to show convergence of the empirical average to Nash equilibrium in symmetric bimatrix games (that is bimatrix games where the payoff matrix of each player is the transpose of that of the other) in the sense that every limit point of the sequence of averages is an ϵ -approximate symmetric equilibrium strategy for any desirable ϵ . Our analysis gives rise to a symmetric equilibrium fully polynomial-time approximation scheme, implying $\mathbf{P} = \mathbf{PPAD}$.

1 Introduction

"You want forever, always or never."

— The Pierces

Game theory is a mathematical discipline concerned with the study of algebraic, analytic, and other objects that abstract the physical world, especially social interactions. The most important solution concept in game theory is the Nash equilibrium [Nash, 1950], a strategy profile (combination of strategies) in an N-player game such that no unilateral player deviations are profitable. The Nash equilibrium is an attractive solution concept, for example, as Nash showed, an equilibrium is guaranteed to exist in any N-person game. Over time this concept has formed a basic cornerstone of economic theory, but its reach extends beyond economics to the natural sciences and biology.

One of the limitations of Nash equilibrium as a plausible solution concept is that we did not have an efficient algorithm for computing one. In fact, it has been conjectured that Nash equilibrium computation is intractable as it is complete for the complexity class **PPAD** [Daskalakis et al., 2009, Chen et al., 2009]. This class, introduced by Papadimitriou [1994], contains a variety of related problems (such as computing Brouwer fixed points) that we didn't have efficient algorithms for. Thus, there was a gap between game theory and the theory of computing. In this paper, we take a step toward reconciling these disciplines using dynamical systems theory as an intermediate step.

The notion of equilibrium admits various definitions in the mathematical sciences. One such standard definition is as a fixed point of a dynamical system. Research in dynamical systems is hardly content with identifying the fixed points of a (continuous) flow or a (discrete) map. What is ultimately important in this mathematical branch is to understand the evolution of system trajectories whether near fixed points (or near, for example, limit cycles) or globally. The Nash equilibrium can also be understood as a fixed point, for example, of the best response correspondence of a game (and other dynamics). But dynamical systems (as, for example, studied in theory of learning in games [Fudenberg and Levine, 1998, Cesa-Bianchi and Lugosi, 2006] or evolutionary game theory [Weibull, 1995, Sandholm, 2010]) whose trajectories evolve toward a Nash equilibrium (generically, without restrictive assumptions on the payoff structure) eluded us. Since dynamical systems are algorithms, progress in this direction evidently informs algorithmic research.

There are classes of games where equilibrium computation was already known to be tractable: Zero-sum games are equivalent to linear programming and thus minimax equilibrium computation admits a polynomial-time algorithm. A bimatrix game is called symmetric if the payoff matrix of each player is the transpose of that of other. Every symmetric N-person game (and, thus every symmetric bimatrix game) admits a symmetric Nash equilibrium [Nash, 1951]. If the payoff matrix of the symmetric game is also symmetric, then the game is called doubly symmetric. Symmetric equilibrium computation in doubly symmetric games admits a fully polynomial time approximation scheme (FPTAS) [Ye, 1998]. In this paper, we consider the problem of computing a symmetric equilibrium in a symmetric bimatrix game (that is not necessarily doubly symmetric).

Chen et al. [2009] show that finding a Nash equilibrium in a 2-person game is a **PPAD**-complete problem and that an equilibrium FPTAS in these games (under either of "additive" or "multiplicative" notions of payoff approximation) implies $\mathbf{P} = \mathbf{PPAD}$. Avramopoulos [2018], drawing on [Jurg et al., 1992], shows that an FPTAS for a symmetric equilibrium in symmetric bimatrix games also implies that $\mathbf{P} = \mathbf{PPAD}$. Thus approximating an equilibrium in the class of games we consider is conjectured to be a hard problem. In this paper, we refute that belief. Our result was motivated by an elementary question concerning multiplicative weights dynamics.

An important result at the intersection of theoretical machine learning and game theory is that the *empirical average* of the strategies generated by a well-known multiplicative weights algorithm, namely, Hedge [Freund and Schapire, 1997, 1999], approaches the minimax strategy of a respective zero-sum game. In doubly symmetric bimatrix games, Hedge is a multiplicative version of *gradient ascent* (with an additional gradient exponentiation step). Gradient ascent is known to converge to critical points of nonlinear optimization problems (for example, see [Bertsekas, 1999, p. 48]) and symmetric equilibrium computation in a doubly symmetric game is a special case of a quadratic programming. It is natural to expect that by tuning Hedge's learning rate parameter, obtaining convergence in doubly symmetric games falls within the realm of possibilities of this algorithm.

But every symmetric bimatrix game is the sum of a doubly symmetric game and a symmetric zero-sum game in the sense that every (payoff) matrix C can be decomposed as follows:

$$C = \frac{1}{2}(C + C^T) + \frac{1}{2}(C - C^T).$$

Since the iterates of Hedge are expected to converge to a symmetric equilibrium in a doubly symmetric game (which would imply that the empirical average of the iterates would converge likewise) and the empirical average of the iterates is known to converge to such an equilibrium in symmetric zero-sum games, is it then possible that the empirical average of iterates converges to a symmetric equilibrium in the general class of symmetric bimatrix games? Our answer is affirmative.

1.1 Our results and techniques

We show in Theorem 1 that under a fixed learning rate parameter α every limit point of the sequence of empirical averages is an ϵ -approximate symmetric equilibrium strategy for any desirable ϵ . We achieve that by tuning parameter α . A corollary is that a symmetric equilibrium always exists in symmetric bimatrix games. We may thus obtain alternative (constructive) proofs of existence of equilibria and fixed points in a variety of related problems (such as N-person games). Our analysis gives an equilibrium fully polynomial time approximation scheme, which implies **P=PPAD**.

The basis of our analysis is a formula for the equilibrium approximation error (cf. Lemma 6), which we derive directly from the equation defining Hedge. We pierce through this formula in Lemma 7 using an inductive proof that makes use of the "multiplicative weights convexity lemma." (Lemma 7 was the hardest part of the analysis required to obtain equilibrium approximation bounds.) The multiplicative weights convexity lemma states that the composition of the

relative entropy function with Hedge is a convex function of α a result appears for the first time in [Avramopoulos, 2016]. Using this lemma we obtain in Lemma 7 a key recursive relationship using logarithms that we make use of in our inductive proof. Our use of the logarithmic function is related to a proof that the time average of the trajectory of the replicator dynamic in a symmetric bimatrix game with an interior equilibrium converges to that equilibrium [Schuster et al., 1981] (see also [Weibull, 1995, p. 91]). That the relative entropy function (also known as Kullback-Leibler divergence) facilitates analyzing multiplicative weights (and its continuous approximation, namely, the replicator dynamic) is well known (for example, see [Weibull, 1995, Freund and Schapire, 1999]).

1.2 OTHER RELATED WORK

There is a significant amount of work on Nash equilibrium computation especially in the setting of 2-player games. We simply mention a boundary of those results. The Lemke-Howson algorithm for computing an equilibrium in a bimatrix game is considered by many to be the state of the art in exact equilibrium computation but it has been shown to run in exponential time in the worst case [Savani and von Stengel, 2004]. There is a quasi-polynomial algorithm for additively approximate Nash equilibria in bimatrix games due to Lipton et al. [2003] (based on looking for equilibria of small support on a grid). Prior to our work, the best polynomial-time approximation algorithm for a Nash equilibrium achieves a 0.3393 approximation [Tsaknakis and Spirakis, 2009].

1.3 Outline of the rest of the paper

Section 2 starts off with preliminaries in equilibrium computation discussing bimatrix and symmetric bimatrix games, defining Hedge, and proving a key relationship used in our main results as an implication of the multiplicative weights convexity lemma. Section 3 contains our main results on asymptotic convergence to an equilibrium as well as on equilibrium approximation. Our main result that $\mathbf{P} = \mathbf{PPAD}$ is obtained via an equilibrium fully polynomial-time approximation scheme.

2 Equilibrium computation background and preliminary results

2.1 Symmetric bimatrix games

A 2-player (bimatrix) game in normal form is specified by a pair of $n \times m$ matrices A and B, the former corresponding to the row player and the latter to the column player. A mixed strategy for the row player is a probability vector $P \in \mathbb{R}^n$ and a mixed strategy for the column player is a probability vector $Q \in \mathbb{R}^m$. The payoff to the row player of P against Q is $P \cdot AQ$ and that to the column player is $P \cdot BQ$. Let us denote the space of probability vectors for the row player by \mathbb{P} and the corresponding space for the column player by \mathbb{Q} . A Nash equilibrium of a 2-player game (A, B) is a pair of mixed strategies P^* and Q^* such that all unilateral deviations from these strategies are not profitable, that is, for all $P \in \mathbb{P}$ and $Q \in \mathbb{Q}$, we simultaneously have that

$$P^* \cdot AQ^* \ge P \cdot AQ^* \tag{1}$$

$$P^* \cdot BQ^* > P^* \cdot BQ. \tag{2}$$

If $B = A^T$, where A^T is the transpose, (A, B) is called *symmetric*. Let (C, C^T) be a symmetric bimatrix game. We call an equilibrium (P^*, Q^*) symmetric if $P^* = Q^*$. If (X^*, X^*) is a symmetric equilibrium, we call X^* a symmetric equilibrium strategy. C denotes the symmetric game (C, C^T) .

¹In this manuscript, I erroneously believed to have shown P = PPAD. The error is in Lemma 10.

2.1.1 Some further notation

We denote the space of symmetric bimatrix games by \mathbb{C} . If the payoff entries lie in the range [0,1], we denote the corresponding space by $\hat{\mathbb{C}}$. Given $C \in \mathbb{C}$, we denote the corresponding set of pure strategies by $\mathcal{K}(C) = \{1, \ldots, n\}$. Pure strategies are denoted either as i or as E_i , a probability vector whose mass is concentrated in position i. $\mathbb{X}(C)$ is the probability simplex (space of mixed strategies) corresponding to $C \in \mathbb{C}$. We denote the (relative) interior of $\mathbb{X}(C)$ by $\mathbb{X}(C)$ (every pure strategy in $\mathbb{X}(C)$ has probability mass). Let $X \in \mathbb{X}(C)$. We define the support or carrier of X by

$$C(X) \equiv \{i \in \mathcal{K}(C) | X(i) > 0\}.$$

2.1.2 Approximate equilibria

Conditions (1) and (2) simplify as follows for a symmetric equilibrium strategy X^* :

$$\forall X \in \mathbb{X}(C) : (X^* - X) \cdot CX^* \ge 0.$$

An ϵ -approximate symmetric equilibrium satisfies:

$$\forall X \in \mathbb{X}(C) : (X^* - X) \cdot CX^* \ge -\epsilon.$$

We may equivalently write the previous expression as

$$(CX^*)_{\max} - X^* \cdot CX^* \le \epsilon,$$

where

$$(CX^*)_{\max} = \max\{Y \cdot CX^* | Y \in \mathbb{X}(C)\}.$$

If $Y^+ \cdot CX^* = (CX^*)_{\text{max}}$, then Y^+ is called a best response to X^* .

2.2 The convexity Lemma of multiplicative weights and implications

Hedge [Freund and Schapire, 1997, 1999] induces the following map in our setting:

$$T_i(X) = X(i) \cdot \frac{\exp\left\{\alpha E_i \cdot CX\right\}}{\sum_{i=1}^n X(j) \exp\left\{\alpha E_i \cdot CX\right\}}, \quad i = 1, \dots, n,$$
(3)

where C is the payoff matrix of a symmetric bimatrix game, n is the number of pure strategies, E_i is the probability vector corresponding to pure strategy i, and X(i) is the probability mass of pure strategy i. Parameter α is called the *learning rate*, which has the role of a *step size* in our equilibrium computation setting. In this paper, we do not study the behavior of the iterates that T generates per se but instead the sequence $\{\bar{X}^K\}_{K=0}^{\infty}$ of empirical averages of the iterates that T generates starting from an interior to the probability simplex strategy. The empirical average \bar{X}^K at iteration $K=0,1,2,\ldots$ is a simple arithmetic mean

$$\bar{X}^K = \frac{1}{K+1} \sum_{k=0}^K X^k.$$

2.2.1 The multiplicative weights convexity lemma

Let us now give some preliminary results on Hedge dynamics. Part of our analysis of Hedge relies on the relative entropy function between probability distributions (also called *Kullback-Leibler divergence*). The relative entropy between the $n \times 1$ probability vectors P > 0 (that is, for all i = 1, ..., n, P(i) > 0) and Q > 0 is given by

$$RE(P,Q) \doteq \sum_{i=1}^{n} P(i) \ln \frac{P(i)}{Q(i)}.$$

However, this definition can be relaxed: The relative entropy between $n \times 1$ probability vectors P and Q such that, given P, for all $Q \in \{Q \in \mathbb{X} | \mathcal{C}(P) \subset \mathcal{C}(Q)\}$, is

$$RE(P,Q) \doteq \sum_{i \in \mathcal{C}(P)} P(i) \ln \frac{P(i)}{Q(i)}.$$

We refer to [Weibull, 1995, p.96] from well-known properties of the relative entropy function. With this background in mind, we state the following lemma (which we refer to as the multiplicative weights convexity lemma) generalizing [Freund and Schapire, 1999, Lemma 2].

Lemma 1 ([Avramopoulos, 2018]). Let T be as in (3). Then

$$\forall X \in \mathring{\mathbb{X}}(C) \ \forall Y \in \mathbb{X}(C) : RE(Y, T(X)) \ is \ a \ convex \ function \ of \ \alpha.$$

The next lemma is shown in [Avramopoulos, 2018]. We repeat the proof for completeness. In the proof, we use the following "secant inequality" for a convex function $F(\cdot)$ and its derivative $F'(\cdot)$:

$$\forall b > a : F'(a) \le \frac{F(b) - F(a)}{b - a} \le F'(b).$$
 (4)

Lemma 2. Let $C \in \hat{\mathbb{C}}$. Then, for all $Y \in \mathbb{X}(C)$ and for all $X \in \mathring{\mathbb{X}}(C)$, we have that

$$\forall \alpha > 0 : RE(Y, T(X)) \le RE(Y, X) - \alpha(Y - X) \cdot CX + \alpha(\exp\{\alpha\} - 1).$$

Proof. Since, by Lemma 1, RE(Y,T(X)) - RE(Y,X) is a convex function of α , we have by the aforementioned secant inequality that, for $\alpha > 0$,

$$RE(Y, T(X)) - RE(Y, X) \le \alpha \left(RE(Y, T(X)) - RE(Y, X)\right)' = \alpha \cdot \frac{d}{d\alpha} RE(Y, T(X))$$
 (5)

where it can be readily computed that

$$\frac{d}{d\alpha}RE(Y,T(X)) = \frac{\sum_{j=1}^{n} X(j)(CX)_{j} \exp\{\alpha(CX)_{j}\}}{\sum_{j=1}^{n} X(j) \exp\{\alpha(CX)_{j}\}} - Y \cdot CX.$$

Using Jensen's inequality in the previous expression, we obtain

$$\frac{d}{d\alpha}RE(Y,T(X)) \le \frac{\sum_{j=1}^{n} X(j)(CX)_j \exp\{\alpha(CX)_j\}}{\exp\{\alpha X \cdot CX\}} - Y \cdot CX.$$
 (6)

Note now that

$$\exp\{\alpha x\} \le 1 + (\exp\{\alpha\} - 1)x, x \in [0, 1],$$

an inequality used in [Freund and Schapire, 1999, Lemma 2]. Using the latter inequality, we obtain from (6) that

$$\frac{d}{d\alpha}RE(Y,T(X)) \le \frac{X \cdot CX}{\exp\{\alpha X \cdot CX\}} - Y \cdot CX + (\exp\{\alpha\} - 1)\frac{\sum_{j=1}^{n} X(j)(CX)_{j}^{2}}{\exp\{\alpha X \cdot CX\}}$$

and since $\exp\{\alpha X \cdot CX\} \ge 1$ again by the assumption that $C \in \hat{\mathbb{C}}$, we have

$$\frac{d}{d\alpha}RE(Y,T(X)) \le X \cdot CX - Y \cdot CX + (\exp\{\alpha\} - 1)\sum_{j=1}^{n} X(j)(CX)_{j}^{2}.$$

Noting that $\sum X(j)(CX)_j^2 \leq 1$ and combining with (5) yields the lemma.

As a corollary to the previous lemma, we obtain the following lemma:

Lemma 3. Let $C \in \hat{\mathbb{C}}$. Then, for all $i \in \{1, ..., n\}$ and for all $X \in \mathring{\mathbb{X}}(C)$, we have that

$$\forall \alpha > 0 : \ln(T_i(X)) \ge \ln(X(i)) + \alpha(E_i - X) \cdot CX - \alpha(\exp\{\alpha\} - 1).$$

Proof. This lemma is a simple implication of Lemma 2 noting that $RE(E_i, X) = -\ln(X(i))$.

We note that (5) can be obtained from Slater's inequality as follows: Let $\hat{X} \equiv T(X)$. Then

$$\frac{\hat{X}(i)}{X(i)} = \frac{\exp\{\alpha(CX)_i\}}{\sum_{j=1}^n X(j) \exp\{\alpha(CX)_j\}}.$$

Slater's inequality (cf. [Dragomir, 2003]) gives

$$\sum_{j=1}^{n} X(j) \exp\{\alpha(CX)_j\} \le \exp\left\{\alpha \frac{\sum_{j=1}^{n} X(j)(CX)_j \exp\{\alpha(CX)_j\}}{\sum_{j=1}^{n} X(j) \exp\{\alpha(CX)_j\}}\right\}.$$

Combining the previous inequalities, we obtain

$$\frac{\hat{X}(i)}{X(i)} \ge \exp\left\{\alpha\left((CX)_i - \frac{\sum_{j=1}^n X(j)(CX)_j \exp\{\alpha(CX)_j\}}{\sum_{j=1}^n X(j) \exp\{\alpha(CX)_j\}}\right)\right\}.$$

Taking logarithms

$$\ln(\hat{X}(i)) - \ln(X(i)) \ge \alpha \left((CX)_i - \frac{\sum_{j=1}^n X(j)(CX)_j \exp\{\alpha(CX)_j\}}{\sum_{j=1}^n X(j) \exp\{\alpha(CX)_j\}} \right).$$

Multiplying both sides with Y(i) and summing over i = 1, ..., n

$$RE(Y,X) - RE(Y,\hat{X}) \ge \alpha \left(Y \cdot CX - \frac{\sum_{j=1}^{n} X(j)(CX)_j \exp\{\alpha(CX)_j\}}{\sum_{j=1}^{n} X(j) \exp\{\alpha(CX)_j\}} \right).$$

Rearranging

$$RE(Y, \hat{X}) - RE(Y, X) \le \alpha \left(\frac{\sum_{j=1}^{n} X(j)(CX)_j \exp\{\alpha(CX)_j\}}{\sum_{j=1}^{n} X(j) \exp\{\alpha(CX)_j\}} - Y \cdot CX \right),$$

and, thus,

$$RE(Y, \hat{X}) - RE(Y, X) \le \alpha \frac{d}{d\alpha} RE(Y, T(X)).$$

3 Convergence under a fixed learning rate and P=PPAD

3.1 Our asymptotic convergence result

Theorem 1. Let $C \in \hat{\mathbb{C}}$ and $X^k \equiv T^k(X^0)$, where $X^0 \in \mathring{\mathbb{X}}(C)$ is the uniform distribution. Assume the learning rate $\alpha > 0$ is constant. Then considering the sequence of empirical averages

$$\left\{ \frac{1}{K+1} \sum_{k=0}^{K} X^k \equiv \bar{X}^K \right\}_{K=0}^{\infty},$$

every limit point of this sequence is an $n(\exp{\{\alpha\}}-1)$ -approximate symmetric equilibrium strategy of C, where n is the number of pure strategies of C.

The assumption X^0 is the uniform distribution is to simplify notation. It is a simple exercise to lift that assumption. Note that the definition of the empirical average implies the recursive relationship

$$\bar{X}^{K+1} = \frac{1}{K+2} X^{K+1} + \frac{K+1}{K+2} \bar{X}^K,$$

which we make use of in the sequel.

Lemma 4. Suppose X^0 is the uniform distribution. Then

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le -\frac{1}{\alpha(K+1)} \sum_{j=1}^n \bar{X}^K(j) \ln (X^{K+1}(j))$$

Proof. Let $T(X) \equiv \hat{X}$. Then straight algebra gives

$$\frac{\hat{X}(i)}{\hat{X}(j)} = \frac{X(i)}{X(j)} \exp\{\alpha((CX)_i - (CX)_j)\}\$$

and taking logarithms on both sides we obtain

$$\ln\left(\frac{\hat{X}(i)}{\hat{X}(j)}\right) = \ln\left(\frac{X(i)}{X(j)}\right) + \alpha((CX)_i - (CX)_j).$$

We may write the previous equation as

$$\ln\left(\frac{X^{k+1}(i)}{X^{k+1}(j)}\right) = \ln\left(\frac{X^k(i)}{X^k(j)}\right) + \alpha((CX^k)_i - (CX^k)_j)$$

Summing over k = 0, ... K, we obtain

$$\ln\left(\frac{X^{K+1}(i)}{X^{K+1}(j)}\right) = \ln\left(\frac{X^0(i)}{X^0(j)}\right) + \alpha \sum_{k=0}^K ((CX^k)_i - (CX^k)_j)$$

and dividing by K+1 and rearranging, we further obtain

$$\frac{1}{\alpha(K+1)} \ln \left(\frac{X^{K+1}(i)}{X^{K+1}(j)} \right) = \frac{1}{\alpha(K+1)} \ln \left(\frac{X^0(i)}{X^0(j)} \right) + (E_i - E_j) \cdot C\bar{X}^K.$$

Under the assumption X^0 is the uniform distribution,

$$\frac{1}{\alpha(K+1)}\ln\left(\frac{X^{K+1}(i)}{X^{K+1}(j)}\right) = (E_i - E_j) \cdot C\bar{X}^K,$$

which implies

$$\frac{1}{\alpha(K+1)} \ln \left(\frac{X^{K+1}(i_{\max})}{X^{K+1}(j)} \right) = (C\bar{X}^K)_{\max} - (C\bar{X}^K)_j$$

and further implies

$$(C\bar{X}^K)_{\max} - (C\bar{X}^K)_j \le -\frac{1}{\alpha(K+1)} \ln (X^{K+1}(j)).$$

Multiplying both sides with $\bar{X}^K(j)$ and summing over $j = 1, \ldots, n$, we finally obtain

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le -\frac{1}{\alpha(K+1)} \sum_{j=1}^n \bar{X}^K(j) \ln (X^{K+1}(j))$$

as claimed.

Lemma 5. Suppose X^0 is the uniform distribution. Then

$$(C\bar{X}^K)_{\max} - \bar{X}^{K+1} \cdot C\bar{X}^K \le -\frac{1}{\alpha(K+1)} \sum_{j=1}^n \bar{X}^{K+1}(j) \ln (X^{K+1}(j))$$

Proof. The proof is directly analogous to the proof of Lemma 4.

Lemma 6. Under the assumptions of Theorem 1,

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le -\frac{1}{\alpha(K+1)} \sum_{j=1}^n \bar{X}^{K+1}(j) \ln \left(X^{K+1}(j) \right) + \frac{1}{K+1}.$$

Proof. Since

$$\bar{X}^{K+1} = \frac{1}{K+2} X^{K+1} + \frac{K+1}{K+2} \bar{X}^K$$

we have

$$\bar{X}^{K+1} \cdot C\bar{X}^K = \frac{1}{K+2} X^{K+1} \cdot C\bar{X}^K + \frac{K+1}{K+2} \bar{X}^K \cdot C\bar{X}^K$$

and, therefore,

$$\bar{X}^K \cdot C\bar{X}^K = \frac{K+2}{K+1}\bar{X}^{K+1} \cdot C\bar{X}^K - \frac{1}{K+1}X^{K+1} \cdot C\bar{X}^K,$$

which, using the assumption $C \in \hat{\mathbb{C}}$ so that $X^{K+1} \cdot C\bar{X}^K \leq 1$, implies

$$\bar{X}^K \cdot C\bar{X}^K \geq \frac{K+2}{K+1} \bar{X}^{K+1} \cdot C\bar{X}^K - \frac{1}{K+1} \geq \bar{X}^{K+1} \cdot C\bar{X}^K - \frac{1}{K+1},$$

which, combined with Lemma 5, implies the lemma.

Lemma 7. Under the assumptions of Theorem 1, the sequence $\{X^k(j)\}$ of probability masses corresponding to pure strategy $j \in \{1, ..., n\}$ satisfies the following relation

$$-\frac{1}{\alpha(K+1)}\bar{X}^{K}(j)\ln\left(X^{K}(j)\right) \leq -\frac{1}{\alpha(K+1)}\left(\frac{1}{K+1}\sum_{k=0}^{K}X^{k}(j)\ln\left(X^{k}(j)\right)\right) + (\exp\{\alpha\} - 1) + \rho^{K},$$

where $\rho = 1/2$.

Proof. Our proof is by induction. The basis of the induction (K = 0) is straightforward. Suppose now the relation is true for iteration K > 0. We then have for the next iteration

$$-\frac{1}{\alpha(K+2)}\bar{X}^{K+1}(j)\ln(X^{K+1}(j)) =$$

$$= -\frac{1}{\alpha(K+2)} \left(\frac{1}{K+2} X^{K+1}(j) + \frac{K+1}{K+2} \bar{X}^K(j) \right) \ln \left(X^{K+1}(j) \right) =$$

Using straight algebra we obtain

$$= -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{1}{\alpha(K+2)} \frac{K+1}{K+2} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) = -\frac{1}{\alpha(K+2)} \frac{1}{K+2} \frac{1}{K+2} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) = -\frac{1}{\alpha(K+2)} \frac{1}{K+2} \frac{1}{K+2} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) = -\frac{1}{\alpha(K+2)} \frac{1}{K+2} \frac{1}{K+2} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) = -\frac{1}{\alpha(K+2)} \frac{1}{K+2} \frac{1}{K+2} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) = -\frac{1}{\alpha(K+2)} \frac{1}{K+2} \frac{1}{K+2} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) = -\frac{1}{\alpha(K+2)} \frac{1}{K+2} \frac$$

$$= -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{(K+1)^2}{(K+2)^2} \cdot \frac{1}{\alpha(K+1)} \bar{X}^K(j) \ln \left(X^{K+1}(j) \right) \tag{7}$$

We may now make use of the relation

$$\ln(X^{k+1}(j)) \ge \ln(X^k(j)) + \alpha((CX^k)_j - X^k \cdot CX^k) - \alpha(\exp\{\alpha\} - 1)$$

which is obtained from Lemma 3. Since $(CX^k)_j - X^k \cdot CX^k \ge -1$ by the assumption $C \in \hat{\mathbb{C}}$, using straight algebra we obtain

$$\ln(X^{k+1}(j)) \ge \ln(X^k(j)) - \alpha \exp\{\alpha\}$$

Combining the previous inequality with (7), we obtain

$$\leq -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{(K+1)^2}{(K+2)^2} \cdot \frac{1}{\alpha(K+1)} \bar{X}^K(j) \ln \left(X^K(j) \right) + \\ + \frac{(K+1)^2}{(K+2)^2} \frac{1}{\alpha(K+1)} \alpha \exp\{\alpha\} \bar{X}^K(j)$$

Using the induction hypothesis we further obtain

$$\leq -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{(K+1)^2}{(K+2)^2} \frac{1}{\alpha(K+1)} \left(\frac{1}{K+1} \sum_{k=0}^K X^k(j) \ln \left(X^k(j) \right) \right) + \frac{(K+1)^2}{(K+2)^2} \frac{1}{\alpha(K+1)} \alpha \exp\{\alpha\} \bar{X}^K(j) + \frac{(K+1)^2}{(K+2)^2} (\exp\{\alpha\} - 1 + \rho^K)$$

which implies (since $0 < \bar{X}^K(j) < 1$)

$$\leq -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{(K+1)^2}{(K+2)^2} \frac{1}{\alpha(K+1)} \left(\frac{1}{K+1} \sum_{k=0}^K X^k(j) \ln \left(X^k(j) \right) \right) + \frac{(K+1)^2}{(K+2)^2} \frac{1}{K+1} \exp\{\alpha\} + \frac{(K+1)^2}{(K+2)^2} (\exp\{\alpha\} - 1 + \rho^K)$$

which further implies

$$\leq -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{1}{\alpha(K+2)} \left(\frac{1}{K+2} \sum_{k=0}^{K} X^{k}(j) \ln \left(X^{k}(j) \right) \right) + \frac{K+1}{K+2} \exp\{\alpha\} - \left(\frac{K+1}{K+2} \right)^{2} (1 - \rho^{K})$$

and since

$$\frac{K+1}{K+2} \ge 1 - \rho, K = 0, 1, \dots$$

we obtain

$$\leq -\frac{1}{\alpha(K+2)} \frac{1}{K+2} X^{K+1}(j) \ln \left(X^{K+1}(j) \right) - \frac{1}{\alpha(K+2)} \left(\frac{1}{K+2} \sum_{k=0}^{K} X^{k}(j) \ln \left(X^{k}(j) \right) \right) + \frac{K+1}{K+2} (\exp{\{\alpha\}} - 1 - \rho - \rho^{K} + \rho^{K+1})$$

which finally implies

$$\leq -\frac{1}{\alpha(K+2)} \left(\frac{1}{K+2} \sum_{k=0}^{K+1} X^k(j) \ln \left(X^k(j) \right) \right) + (\exp\{\alpha\} - 1) + \rho^{K+1}$$

completing the proof of the lemma.

Lemma 8. Under the assumptions of Theorem 1, the sequence of empirical averages

$$\left\{ \frac{1}{K+1} \sum_{k=0}^{K} X^k \equiv \bar{X}^K \right\}_{K=0}^{\infty},$$

satisfies the following relation

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le \frac{n\exp\{-1\}}{\alpha(K+1)} + \frac{1}{K+1} + \left(\frac{K+2}{K+1}\right)n(\exp\{\alpha\} - 1) + \left(\frac{K+2}{K+1}\right)n\rho^K.$$
(8)

Proof. Lemma 6 implies that

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le -\frac{1}{\alpha(K+1)} \sum_{j=1}^n \bar{X}^{K+1}(j) \ln\left(X^{K+1}(j)\right) + \frac{1}{K+1}.$$
 (9)

Lemma 7 further implies that

$$-\frac{1}{\alpha(K+2)}\bar{X}^{K+1}(j)\ln\left(X^{K+1}(j)\right) \le -\frac{1}{\alpha(K+2)}\left(\frac{1}{K+2}\sum_{k=0}^{K+1}X^k(j)\ln\left(X^k(j)\right)\right) + (\exp\{\alpha\} - 1) + \rho^K.$$

Using elementary calculus we obtain $-x \ln(x) \le \exp\{-1\}, x \in [0, 1]$ and, therefore,

$$-\frac{1}{\alpha(K+2)}\bar{X}^{K+1}(j)\ln\left(X^{K+1}(j)\right) \le \frac{\exp\{-1\}}{\alpha(K+2)} + (\exp\{\alpha\} - 1) + \rho^K.$$

Thus,

$$-\frac{1}{\alpha(K+1)} \sum_{j=1}^{n} \bar{X}^{K+1}(j) \ln \left(X^{K+1}(j) \right) \le \frac{n \exp\{-1\}}{\alpha(K+1)} + \left(\frac{K+2}{K+1} \right) n \left(\exp\{\alpha\} - 1 \right) + n \left(\frac{K+2}{K+1} \right) \rho^{K+1}(j)$$

and combining with (9) we obtain

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \leq \frac{n \exp\{-1\}}{\alpha(K+1)} + \frac{1}{K+1} + \left(\frac{K+2}{K+1}\right) n(\exp\{\alpha\} - 1) + n\left(\frac{K+2}{K+1}\right) \rho^K$$

as claimed. \Box

Proof of Theorem 1. (8) implies that

$$\lim_{K \to \infty} \left\{ (C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \right\} = n(\exp\{\alpha\} - 1).$$

Therefore, if \bar{X} is any limit point of $\{\bar{X}^K\}$, then

$$(C\bar{X})_{\max} - \bar{X} \cdot C\bar{X} \le n(\exp\{\alpha\} - 1),$$

implying it is an $n(\exp{\{\alpha\}} - 1)$ -approximate equilibrium strategy and the theorem follows.

3.2 A FULLY POLYNOMIAL-TIME APPROXIMATION SCHEME

Deriving an equilibrium fully polynomial-time approximation scheme from relation (8) is straightforward as shown next. First we need a lemma:

Lemma 9. Under the assumptions of Theorem 1, for all $\theta > 0$, in

$$K = \left\lfloor \frac{n \exp\{-1\} + 1 + \alpha}{\alpha \theta} \right\rfloor \tag{10}$$

iterations, we have that

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le \left(\frac{K+2}{K+1}\right) n(\exp\{\alpha\} - 1) + \left(\frac{K+2}{K+1}\right) n\rho^K + \theta.$$

Proof. Let us repeat (8) for convenience:

$$(C\bar{X}^K)_{\max} - \bar{X}^K \cdot C\bar{X}^K \le \frac{n \exp\{-1\}}{\alpha(K+1)} + \frac{1}{K+1} + \left(\frac{K+2}{K+1}\right)n(\exp\{\alpha\} - 1) + \left(\frac{K+2}{K+1}\right)n\rho^K.$$

Letting

$$\theta = \frac{n\exp\{-1\}}{\alpha(K+1)} + \frac{1}{K+1}$$

implies that K should be the first integer such that

$$K+1 \ge \frac{n \exp\{-1\}}{\alpha \theta} + \frac{1}{\theta} = \frac{n \exp\{-1\} + 1 + \alpha}{\alpha \theta}.$$

Since

$$\left| \frac{n \exp\{-1\} + 1 + \alpha}{\alpha \theta} \right| + 1 \ge \frac{n \exp\{-1\} + 1 + \alpha}{\alpha \theta}$$

the lemma follows.

Theorem 2. Under the assumptions of Theorem 1, to attain an equilibrium approximation error that is at most ϵ , we need at most

$$K = \left| \frac{n \exp\{-1\} + 1 + \ln(1 + \epsilon'/(3n))}{(\epsilon'/3) \ln(1 + \epsilon'/(3n))} \right|$$

iterations, for some $\epsilon' < \epsilon$, which can be readily computed a priori.

Proof. The number of iterations in the statement of the theorem is obtained by letting

$$\alpha = \ln\left(1 + \frac{\epsilon'}{3n}\right) \qquad \theta = \frac{\epsilon'}{3}$$

in (10). Since the first integer K' such that

$$n\rho^{K'} \le \frac{\epsilon'}{3}$$

is

$$K' = \left\lceil \ln \left(\frac{\epsilon'}{3n} \right) / \ln(\rho) \right\rceil$$

which is smaller than K, these values imply an equilibrium approximation error of at most

$$\left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3}.$$

We are looking for ϵ' such that

$$\left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} \leq \epsilon.$$

To that end, we choose a lower bound \hat{K} on the number of iterations, for example,

$$\hat{K} = \left\lfloor \frac{n \exp\{-1\} + 1 + \ln(1 + \epsilon/(3n))}{(\epsilon/3) \ln(1 + \epsilon/(3n))} \right\rfloor$$

and we solve

$$\left(\frac{\hat{K}+2}{\hat{K}+1}\right)\frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left(\frac{\hat{K}+2}{\hat{K}+1}\right)\frac{\epsilon'}{3} \le \epsilon.$$

for ϵ' . Since

$$\left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} \le \left(\frac{\hat{K}+2}{\hat{K}+1}\right)\frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left(\frac{\hat{K}+2}{\hat{K}+1}\right)\frac{\epsilon'}{3}$$

such a choice of ϵ' automatically satisfies

$$\left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left(\frac{K+2}{K+1}\right)\frac{\epsilon'}{3} \le \epsilon$$

and this completes the proof.

Corollary 1. P = PPAD

Proof. [Avramopoulos, 2018, Theorem 15] shows that symmetric equilibrium approximation in fully polynomial time in symmetric games implies that P = PPAD. Hence the corollary.

Acknowledgments

This paper has benefitted from my interaction with my YouTube account and I thank those that are responsible for the configuration of the content in that account.

References

- I. Avramopoulos. Multiplicative weights, equalizers, and P=PPAD. arXiv eprint 1609.08934 (cs.GT), 2016.
- I. Avramopoulos. On incremental deployability. arXiv eprint 1805.10115 (cs.GT), 2018.
- D. P. Bertsekas. Nonlinear Programming. Athena Scientific, Belmont, Mass., second edition, 1999.
- N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, New York, NY, 2006.
- X. Chen, X. Deng, and S. Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 56(3), 2009.
- C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a Nash equilibrium. SIAM J. Comput., 39(1):195–259, 2009.
- S. S. Dragomir. A survey on Cauchy-Bunyakovsky-Schwarz type discrete inequalities. *Journal of Inequalities in Pure and Applied Mathematics*, 4(3), 2003.
- Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.
- Y. Freund and R. E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29:79–103, 1999.

- D. Fudenberg and D. K. Levine. The Theory of Learning in Games. MIT Press, 1998.
- A. P. Jurg, M. J. M. Jansen, J. A. M. Potters, and S. H. Tijs. A symmetrization for finite two-person games. ZOR Methods and Models for Operations Research, 36:111-123, 1992.
- R. Lipton, E. Markakis, and A. Mehta. Playing large games using simple strategies. In *Proc. EC'03*, pages 36–41, 2003.
- J. F. Nash. Equilibrium points in n-person games. PNAS, 36(1), Jan. 1950.
- J. F. Nash. Non-cooperative games. The Annals of Mathematics, Second Series, 54(2):286–295, Sept. 1951.
- C. H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Sciences*, 48(3):498–532, 1994.
- W. H. Sandholm. Population Games and Evolutionary Dynamics. MIT Press, 2010.
- R. Savani and B. von Stengel. Exponentially many steps for finding a Nash equilibrium in a bimatrix game. In *Proc.* 45th Annual IEEE Symposium on Foundations of Computer Science, pages 258–267, 2004.
- P. Schuster, K. Sigmund, J. Hofbauer, and R. Wolff. Selfregulation of behaviour in animal societies. Biological Cybernetics, 40:1–8, 1981.
- H. Tsaknakis and P. G. Spirakis. An optimization approach for approximate Nash equilibria. *Internet Mathematics*, 5(4):365–382, 2009.
- J. W. Weibull. Evolutionary Game Theory. MIT Press, 1995.
- Y. Ye. On the complexity of approximating a KKT point of quadratic programming. *Mathematical Programming*, 80:195–211, 1998.