On the Equilibria of Alternating Move Games

Aaron Roth
Carnegie Mellon University
Computer Science Department
alroth@cs.cmu.edu

Maria Florina Balcan
Georgia Institute of Technology
ninamf@cc.gatech.edu

Adam Kalai
Microsoft Research
adum@microsoft.com

Yishay Mansour
Google Research and Tel Aviv University
mansour@tau.ac.il

October 10, 2009

Abstract

We consider computational aspects of alternating move games, repeated games in which players take actions at alternating time steps rather than playing simultaneously. We show that alternating move games are more tractable than simultaneous move games: we give an FPTAS for computing an ϵ-approximate equilibrium of an alternating move game with any number of players. In contrast, it is known that for $k \geq 3$ players, there is no FPTAS for computing Nash equilibria of simultaneous move repeated games unless $P = PPAD$. We also consider equilibria in memoryless strategies, which are guaranteed to exist in two player games. We show that for the special case of $k = 2$ players, all but a negligible fraction of games admit an equilibrium in *pure* memoryless strategies that can be found in polynomial time. Moreover, we give a PTAS to compute an ϵ-approximate equilibrium in pure memoryless strategies in any 2 player game that admits an exact equilibrium in pure memoryless strategies.

1 Introduction

Consider a pair of gas stations situated across the street from each other, competing for customers: each has the benefit of observing the price set by the other, and can update its own price at any time. Such a scenario can be modeled as an infinitely repeated game in which players *alternate* making moves, rather than playing simultaneously. Although the game is now one of perfect information, equilibrium behavior is still not obvious: Despite the fact that myopically undercutting the opponent's price is a best response in the short term, because the game of selling gas is repeated indefinitely, the gas stations may wish to instead try and optimize long term revenue and avoid setting off a price war. Such scenarios motivate the study of equilibria in alternating move games. In fact, in the absence of a synchronizing mechanism, many repeated games are better modeled with alternating moves than with simultaneous moves. For example, ad-word auctions are repeated auction games in which players may (directly or indirectly) observe the actions of their opponents, and then update their bids at any time. When Yahoo used a first price auction, price cycling behavior was observed, indicating that players were iteratively changing their actions with knowledge of their opponent's bids [11]. As we shall show, our model captures asynchrony in games more generally: it is strategically equivalent to *random-move* models.

In this paper, we study (infinitely) repeated k-player n-action games.[1]. Each player i has a utility function over actions u_i, and the players take turns playing actions, producing a sequence of action vectors $a_1, \ldots, a_t$ where $a_{t'} = (a_i^1, \ldots, a_1^k)$. Player i only changes his action on time steps congruent to $i \bmod k$. At every time t, every player player i receives utility $u_i(a_t)$. Players wish to maximize their own *limit average* payoff as $t \to \infty$.

The main conceptual result of this work is the existence of an FPTAS for computing simple ϵ-approximate equilibria in any alternating move game. This provides a formal separation between the alternating move model and the simultaneous move model, showing that finding approximate equilibria in alternating move games is strictly easier than in simultaneous move repeated games: it is known that there does not exist an FPTAS for computing approximate equilibria in simultaneous move games with $k \geq 3$ players unless $P = PPAD$

[1] In section 5 we consider the special case of games with $k = 2$ players

[5]. This result is technically simple: we show that the folk theorem yields efficiently constructable approximate equilibria in the alternating move model, unlike in the simultaneous move model.

We then also consider the special case of 2-player games. In this case, we show that the alternating move model is equivalent to more general random-move models of asynchrony. Although the strategy space in a repeated game can in principle be a complicated function of the entire history of game play, these games are guaranteed to have a particularly simple set of equilibrium strategies in which each player plays a *stationary strategy*: a unit-memory strategy that only depends on the most recent action of his opponent. We note that the large literature on best-response dynamics in games (see e.g. [14, 2, 22]) can be seen as studying a particular (possibly) non-equilibrium pair of stationary strategies in alternating move games. Here, we consider the problem of finding pairs of stationary strategies that *are* at equilibrium.

Even the set of *pure* stationary strategy strategies for each player is extremely large: in an $n \times n$ bi-matrix game, each player has n^n possible pure stationary strategies. Hence, infinitely repeated alternating move games can alternatively be viewed as concisely represented simultaneous move games. This gives some pause, since even when a pure strategy Nash equilibrium is guaranteed to exist, computing even an α-approximate pure Nash equilibrium in concisely represented games is PLS-complete, for any computable α [23].The problem of finding equilibria in stationary strategies is less difficult, however: we give a PTAS for computing pure stationary strategies that constitute an ϵ-approximate pure equilibrium in any alternating move game that admits a pure equilibrium in stationary strategies. We further show that all but a negligible fraction of alternating move games admit pure exact equilibria in stationary strategies which can be found in polynomial time.

1.1 Related Work

1.1.1 The Complexity of Simultaneous Move Games It is standard in game theory to study the Nash equilibria of simultaneous move games. Unfortunately, the complexity results in this area have been almost uniformly negative. A series of papers has shown that it is PPAD complete to compute Nash equilibria, even in 2 player games, even when payoffs are restricted to lie in $\{0, 1\}$ [9, 8, 1]. It was recently shown that this hardness carries over to the repeated setting, even though the prevalence of equilibria in repeated games guaranteed by the "Folk Theorem" might suggest otherwise: it is PPAD hard to compute equilibria (even approximate ϵ-equilibria for sufficiently small ϵ) for repeated games with $n \geq 3$ players [5]. There has also been a body of work studying approximate equilibria, giving algorithms which achieve reasonably large constant approximations (see, e.g. [10, 24, 6, 7]). No PTAS is known, and no FPTAS exists unless $P = PPAD$.

Barany et al. considered two player games with *randomly* chosen payoff matrices, and showed that with high probability, such games have Nash equilibria with small support [3]. Their result implies that in random 2x2 games, Nash equilibria can be computed in expected polynomial time. We show a result that is similar in spirit: random 2 player alternating-move games admit stationary equilibria in *pure* strategies that can be found in polynomial time, except with negligibly small probability.

A notable structural result of Lipton, Markakis, and Mehta shows that there always exist ϵ-approximate equilibria with support over at most $O(\log n/\epsilon^2)$ strategies: this gives a subexponential-time algorithm for computing ϵ-approximate equilibria [20]. This result is similar in spirit to our own structural result: we show that any game that admits an equilibrium in pure stationary strategies admits an ϵ-approximate equilibria in pure stationary strategies which leads into a cycle of actions of length at most $O(1/\epsilon)$. Together with a reduction to a pair of zero sum games, this gives us a PTAS for computing ϵ-approximate equilibria for alternating move games.

1.1.2 Two Player Zero Sum Games Two player alternating move bi-matrix games can be viewed as being played on a complete bipartite graph, in which player 1's action set A is identified with the vertex set on the right and player 2's action set B is identified with the vertex set on the left. Players take turns moving a piece between vertices: whenever the piece is moved from vertex a to b, player 1 receives payoff $u_1(a, b)$ and player 2 receives payoff $u_2(a, b)$. In a different context, under the name *cyclic games*, the special case of zero-sum games was considered in a more general model by Ehrenfeucht and Mycielski (on a bipartite graph that need not be complete) [12]. They show that zero-sum alternating move games have a value, and that they always have equilibria in pure stationary strategies, but give no algorithm for computing these strategies. Gurvich et al. [18] independently consider another variant of this zero-sum model (vertices are divided into two classes, but the graph need not be bipartite), under the name *mean-payoff games*, and also show the existence of a value and pure stationary strategies, using a general fixed-point theorem of Moulin about stochastic games [21]. They give an exponential time algorithm for computing these equilibria. Despite several mistaken claims of polynomial time algorithms for computing exact equilibria in the zero-sum case (See [25] for discussion), the best known results are due to Zwick and Patterson, who give a pseudo-polynomial time algorithm for computing pure-stationary equilibria in Gurvich et al.'s model, and Björklund and Vorobyov, who give a strongly subexponential time algorithm[25, 4]. The problem of computing exact equilibria even for the special case of zero-sum games remains open. An FPTAS for computing approximate equilibria in two-player zero sum games follows easily from [25], which we use as a key subroutine in our algorithms, but as far as we are aware, neither approximate equilibria nor general sum games have been considered in this literature.

1.1.3 Best Response Dynamics There is a large literature on best-response dynamics in games (see, e.g. [14, 2, 22]). Best response dynamics represent a particular pair of stationary strategies that can be played in alternating move bi-matrix games, but are usually not at equilibrium. It is often assumed that players will play according to best-response dynamics, and from this assumption, its convergence and social welfare properties are investigated. In this paper, we consider finding pairs of stationary strategies that are at equilibrium. It should be noted that best response dynamics *do* sometimes constitute an equilibrium[2]. It is an interesting question (one not considered in this paper) to study in which classes of games best response dynamics actually form an equilibrium in this model.

1.1.4 Other Work Fabrikant and Papadimitriou recently introduced the notion of *unit recall equilibria*, which are equilibria of the usual simultaneous-move repeated game when players are restricted to playing pure stationary strategies [13]. They show that not all games admit unit recall equilibria, but that random games admit unit recall equilibria with probability approaching 1. They conjecture that the problem of finding unit-recall equilibria when they exist is in P, but give no general algorithm either for exact computation or for approximation. Their model is the simultaneous move analogue of ours, but the equilibrium structure is quite different: for example, they show that the 2×2 zero sum game "matching pennies" has no unit recall equilibrium. In contrast, all zero sum games, and all 2×2 general sum games have equilibria in pure stationary strategies in our model. We take inspiration from their probabilistic result to show a similar result in our model: random $n \times n$ alternating move games have equilibria in pure stationary strategies except with negligible probability.

[2] For example, in the alternating move version of "Rock, Paper, Scissors" the best response dynamic that always responds with "Paper" to Rock, "Scissors" to Paper, and "Rock" to Scissors form an equilibrium

2 Model and Preliminaries

We consider finite k-player n-action games. Such games have finite action sets A_i for each player $i \in [k]$. The game also has k utility functions $u_i : A_1 \times \ldots \times A_k \to [-1, 1]$ for each player i. Note that any finite game can be rescaled so that utilities are bounded in $[-1, 1]$. Play proceeds for an infinite number of periods $t = 1, 2, \ldots$, in which players take turns choosing actions, without loss of generality in the order $1, 2, \ldots, k$. The game G is formally specified by $G = (\{A_i\}, \{u_i\})$.

Play proceeds in periods $t = 0, 1, 2, 3 \ldots$. On periods $t = i \bmod k$, player i chooses an action $a_t^i \in A_i$. For convenience, we say that $a_t^i = a_{t'}^i$ for $t \neq i \bmod k$, where t' is the most recent timestep $t' = i \bmod k$ that player i selected an action. We write $a_t = (a_t^1, \ldots, a_t^k)$ for the vector of actions at time t. The *limit-average payoff* for player i is

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i(a_t)$$

when this limit exists, otherwise it is *undetermined*.

The *history* at period t is $H_t = (a_0, a_1, a_2, \ldots, a_{t-1})$, the sequence of actions played so far. Denote the set of finite histories by $\mathcal{H} = ((A_1 \times \ldots \times A_k))^*$ A *pure strategy* for player i is a function $s_i : \mathcal{H} \to A_i$. A *mixed strategy* for player i is $\sigma_i : \mathcal{H} \to \Delta(A_i)$, where $\Delta(S)$ denotes the set of probability distributions over any finite set S. A pure strategy profile is $s = (s_1, s_2, \ldots, s_k)$, a mixed profile is $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_k)$. As is standard, we use σ_{-i} to denote the set of strategies for all players other than i.

The *expected average payoff until period* T for player i in profile σ is the expected payoff over the choices of a_t^j drawn from $\sigma_j(H_t)$ for $t \leq T$, $t = j \bmod k$:

$$P_{i,T}(\sigma) = \mathrm{E}_{a_t \leftarrow \sigma} \left[\frac{1}{T} \sum_{i=1}^{T} u_i(a_t) \right].$$

The *expected average payoff* for player i is defined to be $P_i(\sigma) = \lim_{T \to \infty} P_{i,T}(\sigma)$, when this limit exists, otherwise it is undetermined. A pure strategy is a special case of a mixed strategy and we extend the domain of P_i to pure strategies, as well. We are interested in pairs of strategies that form equilibria.

DEFINITION 2.1. *For any $\epsilon \geq 0$, a set of mixed strategies $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_k)$ constitute an* ϵ-equilibrium *if* *no player i can unilaterally deviate to some other strategy σ_i' so as to gain more than ϵ in the long run: for each $i \in [k]$ and any mixed strategy σ_i',*

$$\limsup P_{i,T}(\sigma_i', \sigma_{-i}) - P_{i,T}(\sigma) \leq \epsilon.$$

A Nash equilibrium is a 0-equilbrium.

Note that strategies at Nash equilibrium do not necessarily have to have a pair of well-defined limit average payoffs.

In sections 3 and 5, we will discuss the special case of stationary strategies in 2 player games. In a two player game, we will denote the action spaces of players 1 and 2 as A and B respectively. A stationary (mixed) strategy $\pi_1 : B \to \Delta(A)$ for player 1 is a function only of the most recent action of player B. With a slight abuse of notation, we view π_1 as a pure strategy where $\pi_1(a_0, b_1, \ldots, b_t) = \pi_1(b_t)$ for any $t = 1, 3, \ldots$, and similarly for player 2. For player 1, this must be specified together with an initial action a_0 in order to determine a strategy in the game. Moreover, a standard result from the literature on stochastic games can be adapted to show that there always exists an equilibrium in stationary strategies, and that it is independent of the initial action. This theorem is proven using Kakutani's fixed point theorem and is nonconstructive:

Theorem 2.1. (See e.g. [18, 21]) *For every game $G = (A, B, u_1, u_2)$, there exists $a_0 \in A$ and a pair of mixed stationary strategies $\pi_1 : B \to \Delta(A)$ and $\pi_2 : A \to \Delta(B)$ such that $(a, \pi_1), \pi_2$ constitutes an equilibrium.*

Remark 2.1. *Note that (a, π_1, π_2) constitute an equilibrium in the full strategy space, not just in the space of stationary strategies: even though both players are using stationary strategies, neither can improve even with a more complicated strategy.*

Note that a pair of deterministic (pure) stationary strategies and a starting point (a, π_1, π_2) will eventually lead into a cycle of actions $C(a, \pi_1, \pi_2) = (a_1, b_2, \ldots, a_k, b_k)$ of length $k \leq 2n$. In this case, $P_i(a, \pi_1, \pi_2)$ is simply equal to the average payoff to player i on this cycle.

The main result of this paper is an FPTAS to compute ϵ-approximate equilibria in any alternating move game, even for $k \geq 3$ players. These simple equilibria are in the spirit of the folk theorem, and are not stationary. We also give a PTAS for computing ϵ-approximate equilibria in pure stationary strategies in any game that admits an equilibrium in pure stationary strategies. This accounts for *almost all* alternating move games: we prove that a random game will admit an equilibrium in pure stationary strategies except with negligible probability.

2.1 Random Move Model for Two Player Games A possible alternative model of non-simultaneous game play involves choosing a player at random to move at each time step. If we suppose, for simplicity, that A and B are disjoint, then a history is specified by an arbitrary sequence in $(A \cup B)^*$. Again, a strategy for either player consists of a function from histories to actions. A stationary strategy π_i is now implemented only when play changes – if a player moves more than once in a row in the random order game then the player does not change her play. However, in order to specify a stationary strategy for player 2, we also need to specify b_0, the action player 2 would play if she is chosen first. Again the goal is to maximize average payoff $u_i(a_t, b_t)$, over $t = 1, 2, \ldots$. Since the definitions are quite similar to those given in the previous section, we don't formally define what an equilibrium is in this model.

The random-move model is strategically equivalent to the alternating move model. This robustness makes the alternating move model a particularly compelling model of asynchrony in games:

Theorem 2.2. *For any equilibrium in stationary strategies $(a_0, \pi_1), \pi_2$ of an alternating move game $\mathcal{G} = (A, B, u_1, u_2)$, there exists some $b_0 \in B$ such that $(a_0, \pi_1), (b_0, \pi_2)$ is an equilibrium of the random-order game $\mathcal{G}$, with the same expected payoff.*

Proof Sketch:
It is easy to see that the expected payoffs are the same.

The equilibrium (a_0, π_1, π_2) leads to a unique cycle of actions. Take any b_0 on that cycle. The claim is that (a_0, π_1) and (b_0, π_2) are mutual best responses. Suppose not. There are two cases, either player 1 can improve or player 2 can improve. First, say that strategy $\sigma_2 : (A \cup B)^* \to \Delta(B)$ is better against (a_0, π_1) than (b_0, π_2). There are two subcases. Either σ_2 is better when player 1 starts or when player 2 starts. Let us first consider when player 1 starts.

From σ_2, we construct a strategy in the alternating move game, call it σ_2'. This strategy works by simulation. Player 2 writes down on a piece of paper a simulated play of (a_0, π_1) against σ_2 in the random order move game. Her sequence of plays in the alternating move game is then chosen to be the subsequence of plays she makes just before play switches to player 1, in the simulation. For example, if the random move simulation were $\beta_0, \alpha_1, \alpha_2, \beta_2, \beta_3, \alpha_4, \beta_5, \alpha_6, \ldots$, her sequence of plays in the alternating move game would be $b_1 = \beta_0, b_3 = \beta_3, b_5 = \beta_5, \ldots$. Compare the following two pairs of random variables. Let (X, Y) be the payoffs on the first t plays of the alternating move game, however we start counting either at $t = 1$ or $t = 2$ at random. Let (X', Y') be the total payoffs on the pretend sequence during the first t' rounds, where t' is chosen so that the number of times control switches (the chosen player switches) is t. It is not difficult to see that in the case where player 1 starts in the random move model, $\mathrm{E}[t'] = 2t$, $\mathrm{E}[X'] = 2\mathrm{E}[X]$ and $\mathrm{E}[Y'] = 2\mathrm{E}[Y]$.

Since $\mathrm{E}[t']$ concentrates for sufficiently large t, this implies that σ_2' is better against (a_0, π_1) than π_2 is, hence a contradiction.

Finally, it is well-known that any strategy that works using additional state can be simulated without state (by each period resampling the state freshly, conditional on the observed play).

The remaining cases for the theorem are entirely similar.

3 Two Player Zero Sum Games

Zero sum games represent the special case when $u_1 = -u_2$. Hence we will refer simply to a single utility function $u(a, b)$, and a maximization player (who wishes to maximize the limit average payoff), and a minimization player (who wishes to minimize the limit average payoff). Without loss of generality, player 1 is the maximization player. Two player zero sum games can be reduced to *mean payoff games*, the class of games studied by the majority of the literature on alternating move games (see, e.g. [12, 18, 25, 4]). It is known that mean payoff games have a value, and that they always admit equilibria in *pure* stationary strategies. Moreover, it is known how to compute these strategies in time polynomial in the number of actions, but only pseudopolynomial in the utility function [25], and also in subexponential time (even with large utility) [4]. In this section, we review the results that will be needed for our main result.

A *mean payoff game* is defined by a directed graph $G = (V = \{V_1 \cup V_2\}, E)$ and an edge weighting function $w : E \to \{-W, -W+1, \ldots, W-1, W\}$, for integer W. Players take turns choosing edges $e_1, e_2, \ldots$ along a path: whenever play is at a vertex $v \in V_1$, the *maximization player* chooses the next edge. Whenever play is at a vertex $v \in V_2$, the *minimization player* chooses the next edge. The maximization player wishes to maximize the quantity $\lim_{t\to\infty} \sup \sum_{i=1}^{t} w(e_i)/t$, and the minimization player wishes to minimize the quantity $\lim_{t\to\infty} \inf \sum_{i=1}^{t} w(e_i)/t$.

FACT 3.1. (E.G. [12, 18]) *Mean payoff games have a* value $v(a)$ *for every vertex* $a \in V_1$ *and* $b \in V_2$*: There exists a pure stationary strategy* π_1 *such that starting at vertex* a*, player* 1 *can guarantee payoff at least* $v(a)$*, and there exists a pure stationary strategy* π_2 *such that starting at vertex* b*, player* 2 *can guarantee payoff at most* $v(b)$*.* (a, π_1, π_2) *constitutes an equilibrium of the game for any starting point* a*.*

THEOREM 3.1. (ZWICK AND PATTERSON 96[25]) *There exists an algorithm to compute the values of the game and pairs of equilibrium strategies* (π_1, π_2) *in time* $\widetilde{O}(|V|^4 \cdot |E| \cdot W)$*.*

We observe that (rational-valued) two player zero-sum alternating move games can be easily reduced to mean payoff games:

FACT 3.2. *Given a two player zero-sum alternating move game* $\mathcal{G} = (A, B, u(\cdot,\cdot))$*, the mean payoff game defined by* $V = A \cup B$*, edge set* $E = \{(a \to b), (b \to a) : a \in A, b \in B\}$*, and the weighting function that results from* rescaling u *from* $[-1, 1]$ *to have integer values has the same set of equilibrium strategies as* $\mathcal{G}$*. Since we construct a complete bipartite graph, every state has the same value.*

OBSERVATION 3.1. *There exists an FPTAS for computing pure* ϵ*-approximate stationary strategy equilibria in zero sum games.*

Proof. Given a game $\mathcal{G} = (A, B, u)$, we can discretize $u : A \times B \to [-1, 1]$ to take values that are multiples of ϵ. This only changes the payoff of any pair of strategies by at most ϵ,and the reduction from $\mathcal{G}$ to a mean-payoff game now yields $W \le 1/\epsilon$. Zwick and Patterson's algorithm then yields ϵ-approximate equilibria in time $\widetilde{O}(n^6/\epsilon)$.

4 Alternating Move Games and the Folk Theorem

In this section, we prove our main result: a separation between the complexity of computing approximate equilibria of simultaneous move games, and of alternating move games. We give an FPTAS for computing ϵ-equilibria of alternating move games, for an arbitrary number of players. No FPTAS exists for computing equilibria of repeated games with $k \ge 3$ players in the simultaneous move model unless $P = PPAD$ [5].

We show that in any k player alternating move game, we can find ϵ-approximate (non-stationary) pure equilibria in time polynomial in the input size (n^k) and $1/\epsilon$ using the "Folk Theorem". The equilibria that we construct are not stationary, but will nevertheless be simple and easy to implement: Each player i will play according to a stationary "safety strategy" σ_i, which will guarantee that he obtains utility at least v_i, his value of a residual zero-sum mean-payoff game, in which all other players are adversarial. If any player i deviates from σ_i, all other players j switch to playing according to π_i^j, a "punishment" strategy for player i which will guarantee that player i cannot obtain utility greater than v_i. Thus, at any given time, every player is playing according to a stationary strategy, and requires only an additional $\log k$ bits of memory to keep track of which, if any, player has deviated. This will constitute an equilibrium in the spirit of the folk theorem: every player will continue to play a simple stationary strategy, for fear of punishment that will force a lower utility.

THEOREM 4.1. *There is an FPTAS for computing pure equilibria in any k player n action alternating move game.*

Before we describe the algorithm, we define some convenient notation. Let $\mathcal{A}_i = A_{i+1} \times \ldots \times A_k \times A_1 \times \ldots \times A_{i-1}$ be the cross product of all action sets of players other than i. This will represent the domain of a 'history' of the last $k-1$ actions taken when it is time for player i to move. We will construct a series of mean-payoff games with the maximization player representing player i, with vertex set $V_1 = \mathcal{A}_i$, and the minimization player representing a combination of all other players, with vertex set $V_2 = \bigcup_{j \neq i} \mathcal{A}_j$. The edge set in these games will represent legal game trajectories: there will be an edge e from a vertex $(a_{i+1}, \ldots, a_k, a_1, \ldots, a_{i-1}) \in \mathcal{A}_i$ to vertex $(a'_{i+2}, \ldots, a'_k, a'_1, \ldots, a'_i) \in \mathcal{A}_{i+1}$ if and only if $(a_2, \ldots, a_{i-1}, a_{i+2}, \ldots, a_k) = (a'_2, \ldots, a'_{i-1}, a'_{i+2}, \ldots, a'_k)$. The weight of edge e will be $w(e) = u_1(a_1, \ldots, a_k)$.

Given a strategy π_2 for the minimization player, denote $\pi_2^{\mathcal{A}_j}$ to be the restriction of π_2 to the vertex set $\mathcal{A}_j$. Call the mean payoff game constructed in this manner $\mathcal{G}_i$.

Algorithm 1 An FPTAS for finding ϵ-approximate equilibria in any alternating move game.

FindApproximateEquilibria$(\epsilon, \mathcal{G})$:

for $i = 1$ to k **do**
  Construct game $\mathcal{G}_i$. Solve $\mathcal{G}_i$ to accuracy $\epsilon/2$ using the FPTAS for mean-payoff games. Denote the maximization player's strategy π_1, and the minimization player's strategy π_2.
  Let $\sigma_i = \pi_1$ and let $\pi_j^i = \pi_2^{\mathcal{A}_j}$.
end for
Let $f_i = \begin{cases} \pi_i^j, & \text{If player } j \text{ has deviated from } f_j; \\ \sigma_i, & \text{If no player } j \text{ has deviated from } f_j. \end{cases}$
return $(\{f_i\}_{i=1}^k)$

Proof. First note that the input matrix is of size n^k, which is the parameter we wish to be polynomial with respect to. Our algorithm runs in time poly$(n^k, 1/\epsilon)$. To see this, note that our algorithm solves k mean-payoff games to accuracy $\epsilon/2$, each with vertex size $|V| = k \cdot n^{k-1}$ and edge size $|E| = kn^k$. By observation 3.1 and theorem 3.1, this takes time $O(k^6 n^{5k-4}/\epsilon)$. It remains to show that strategies $\{f_i\}_{i=1}^k$ constitute an ϵ-approximate equilibrium. By fact 3.1, each game $\mathcal{G}_i$ has a value v_i such that if player i plays according to σ_i^*, he can guarantee himself payoff at least v_i (independent of starting action). Since the game $\mathcal{G}_i$ is the reduction from $\mathcal{G}$ to a mean-payoff game with edge costs equivalent to player i's utility, if player i plays in $\mathcal{G}$ according to σ_i, he achieves payoff at least $v_i - \epsilon/2$, since σ_i is an $\epsilon/2$ approximation to the optimal strategy for player i. Similarly, any deviation from σ_i for player i will result in *all* other players j switching to playing according to their punishment strategies π_j^i. When all players play according to π_j^i, they are playing consistently with the punishment strategy π_2 computed for $\mathcal{G}_i$, by construction. Since this is an $\epsilon/2$ approximation, when all other players are playing as such, no strategy of player i can achieve better utility than $v_i + \epsilon/2$. Therefore, for each player i, no deviation from f_i can achieve more than an ϵ gain in utility, which completes the proof.

5 Stationary Equilibria in Two Player Games

In this section, we consider the special case of 2-player games and the problem of finding equilibria in stationary strategies. We give a PTAS for computing pairs of stationary strategies that form approximate equilibria. The result will follow from a structural lemma, and a reduction to the FPTAS for two player zero-sum games. First, we show that any game that admits a pure-strategy stationary strategy equilibrium admits a pure-strategy stationary strategy ϵ-equilibrium (a, π_1, π_2) that traverses a cycle $C = C(a, \pi_1, \pi_2)$ of actions of length at most $O(1/\epsilon)$. We then show that given the cycle C traversed by some ϵ-equilibrium, we can complete a pair of strategies (π'_1, π'_2) consistent with that cycle: $C = C'(a', \pi'_1, \pi'_2)$, such that (a', π'_1, π'_2) is also an ϵ-equilibrium. We do this by considering two residual zero-sum games: in game i, player i wishes to minimize the payoff to player $-i$, but is restricted to playing a strategy π'_i that is consistent with cycle C. If C is part of some ϵ-equilibrium, then the values of these two residual games must be within ϵ of the values to each player (respectively) of playing along C: otherwise, some player would have a strategy that guaranteed an ϵ-improvement in payoff. Therefore, given such a cycle, we can complete a pair of equilibrium strategies by solving the residual zero sum games. We now prove the main result of this section:

Theorem 5.1. *There is a PTAS to compute ϵ approximate pure stationary strategy equilibria in games that admit a pure stationary strategy equilibrium.*

The algorithm follows. First we define some convenient notation: Given a cycle $C = (a_0, b_1, a_2, \ldots, a_{k-1}, b_k)$, say that edge $(a_{2i} \to b_{2i+1}) \in C$ if (a_{2i}, b_{2i+1}) are consecutive actions in C. Similarly, $(b_{2i-1}, a_{2i}) \in C$ if (b_{2i-1}, a_{2i}) are consecutive actions in C (this includes $(b_k \to a_0)$). We abuse notation and write the utility of a cycle to player i as $u_i(C)$, to denote its average value to player i:

$$u_i(C) \equiv \frac{1}{k+1}(u_i(a_0, b_1) + u_i(a_2, b_1) + u_i(a_2, b_3) + \ldots + u_i(a_{k-1}, b_k) + u_i(a_0, b_k))$$

We will also refer to intervals $I = (a_0, b_1, a_2, \ldots, a_{k-1}, b_k)$ which are sequences of actions that are not closed into a cycle $(b_k \to a_0) \notin I$. Similarly, we may write $u_i(I)$:

$$u_i(I) \equiv \frac{1}{k}(u_i(a_0, b_1) + u_i(a_2, b_1) + u_i(a_2, b_3) + \ldots + u_i(a_{k-1}, b_k))$$

Our algorithm will create and approximately solve pairs of zero-sum (mean-payoff) games that correspond to game $\mathcal{G}$. The vertex sets of both of these games correspond to the action sets of $\mathcal{G}$: $V = (A \cup B)$. The edge set corresponding to the *unrestricted* action set of player 1 is $E_1 = \{(b \to a) : a \in A, b \in B\}$. Similarly, the unrestricted action set of player 2 is represented by $E_2 = \{(a \to b) : a \in A, b \in B\}$. We will also want to restrict players to playing strategies that are consistent with particular cycles C: A strategy π_i is consistent with C if there is some π_{-i} such that for some $a \in C$, $C = C(a, \pi_i, \pi_{-i})$. The C-restricted action sets of the players are represented by edge sets:

$$E_1^C = \{(b \to a) : (b, a) \in C\} \cup \{(b \to a) : \forall a \in A, (a, b) \notin C\}$$

$$E_2^C = \{(a \to b) : (a, b) \in C\} \cup \{(a \to b) : \forall b \in B, (a, b) \notin C\}$$

$\mathcal{G}_1^C$ will be the mean payoff game in which player 1 wishes to maximize his utility with an unrestricted action set, and player 2 wishes to minimize player 1's utility while being restricted to playing consistently with C: $\mathcal{G}_1^C$ is the mean payoff game $(V, E_1 \cup E_2^c, u_1)$. Similarly, $\mathcal{G}_2^C$ is the mean payoff game in which player 2 wishes to maximize his utility and is unrestricted, and player 1 is the minimization player with a restricted action set: $\mathcal{G}_2^C$ is the mean payoff game $(V, E_1^C \cup E_2, -u_2)$.

Lemma 5.1. *Fixing a stationary strategy of one player, we can compute a best response for the other player in polynomial time.*

Proof. This has been observed before (see e.g. [25, 13]). The best response *stationary strategy* will be one that traverses the maximum average weight cycle of the bipartite action graph, fixing one player's edge choices. We can find this in polynomial time using an algorithm of Karp [19]. Since when we fix the strategy of one player, the best-response problem for the other player is the solution to an MDP, the best response stationary strategy will be a best response.

We are now ready to present the PTAS.

Algorithm 2 A PTAS for computing ϵ-approximate stationary strategy equilibria in two player games.

FindApproximateEquilibrium$(\epsilon, \mathcal{G} = (A, B, u_1, u_2))$:
- **for all** action cycles C of length at most $12/\epsilon$ **do**
  - **Construct** mean payoff games $\mathcal{G}_1^C$ and $\mathcal{G}_2^C$.
  - **Compute** the $\epsilon/2$-approximate value $v_1(a)$ in $\mathcal{G}_1$ and let π_2 be the corresponding stationary strategy for player 2, for some action a traversed by C.
  - **Compute** the $\epsilon/2$-approximate value $v_2(a)$ of vertex a in $\mathcal{G}_2$ and let π_1 be the corresponding stationary strategy for player 1.
  - **Compute** best responses for player 1 to π_2 and for player 2 to π_1, and let v_1^* and v_2^* be the corresponding utilities.
  - **if** $v_1^* \leq u_1(C) + \epsilon$ and $v_2^* \leq u_2(C) + \epsilon$ **then**
    - **return** (a, π_1, π_2)
  - **end if**
- **end for**
- **return** $\mathcal{G}$ has no equilibrium in pure stationary strategies.

Note that our algorithm clearly runs in time polynomial in $n^{O(1/\epsilon)}$: It examines at most $n^{12/\epsilon}$ cycles, and for each, solves a pair of mean payoff games to accuracy $\epsilon/2$ which takes time $O(n^6/\epsilon)$. The proof of correctness begins with our main structural lemma:

Lemma 5.2. *Any game that admits a pure strategy equilibrium admits an ϵ-equilibrium on a cycle of length $6/\epsilon$.*

Proof. Consider some pure stationary strategy equilibrium (a, π_1, π_2) that leads into the play cycle $C = C(a, \pi_1, \pi_2) = (a_0, b_1, a_2, b_3, \ldots, a_{L-2}, b_{L-1})$. If $L \leq 6/\epsilon$, we are done. Therefore, we may assume that $L \geq 6/\epsilon$. Say that the payoff of (a, π_1, π_2) to player 1 is v_1, and the payoff to player 2 is v_2: $v_1 = u_1(C)$, $v_2 = u_2(C)$. We will consider intervals of C of length $L_1 = 3/\epsilon$ completed to form cycles of length $L_1 + 1$. For $r \in [0, L-1]$, define the interval:

$$I(r) = (a_r, b_{r+1 \bmod L}, a_{r+2 \bmod L}, b_{r+3 \bmod L},$$

$\ldots, a_{r+L_1-1 \bmod L}, b_{r+L_1 \bmod L})$

to be the subinterval of C of length L_1 beginning at a_r. Let $I'(r) = C - I(r)$ be the remaining interval of C, and let $C_2(r)$ be the cycle that is formed by "closing" $I(r)$ (by the single action deviation of player 1 that swaps a single edge in C from $(b_{r+L_1 \bmod L} \to a_{r+L_1+1 \bmod L})$ to $e^*(r) \equiv (b_{r+L_1 \bmod L} \to a_r)$). Call the strategy resulting from this single-edge deviation by player 1 $\pi'_1(r)$. Say that $I'(r)$ has length L_2 so that $L_1 + L_2 = L$. Note that if player 1 unilaterally makes this edge-deviation, he achieves payoff $u_1(C_2(r))$, and we have the inequalities:

$$v_1 \geq u_1(C_2(r)) \geq u_1(I(r)) \cdot \left(\frac{L_1}{L_1+1}\right)$$

where the first inequality follows from the fact that (π_1, π_2) is an equilibrium, and the second follows since edge e^* constitutes only a $1/(L_1+1)$ fraction of cycle $C_2(r)$. We recall that the average value to player 1 of cycle C is the weighted average of the value to player 1 of interval $I(r)$ and interval $I'(r)$. With this fact and the inequality above, we obtain:

$$\begin{aligned} v_1 &= \frac{L_1 u_1(I(r)) + L_2 u_1(I'(r))}{L} \\ &\leq \frac{L_1 v_1 (1 + 1/L_1) + L_2 u_1(I'(r))}{L} \\ &= \frac{v_1(1 + L - L_2) + L_2 u_1(I'(r))}{L} \end{aligned}$$

Solving for $u_1(I'(r))$ we find:

$$u_1(I'(r)) \geq (1 - \frac{1}{L_2})v_1$$

since all utilities are in the range $[-1, 1]$. Similarly, since player 1 can equally well deviate to play along a cycle consisting of $I'(r)$ and two additional edges, and player 2 can make a symmetric deviation, we can go through the same derivations to obtain:

$$u_1(I(r)) \geq (1 - \frac{2}{L_1})v_1 \tag{5.1}$$

$$u_2(I(r)) \geq (1 - \frac{1}{L_1})v_2 \tag{5.2}$$

So, $u_1(C_2(r)) \geq v_2 \cdot (1 - 2/L_1)(L_1)/(L_1+1) = v_2 \cdot (L_1 - 2)/(L_1+1)$. Therefore, for *any* r, $\pi'_1(r)$ is a $3/(L_1+1) < \epsilon$ best response to π_2 for $L_1 = 3/\epsilon$. It remains to show that there exists some r so that π_2 is also an ϵ-best response to $\pi'_1(r)$.

We consider r selected from $[0, L-1]$ uniformly at random. By the linearity of expectation, we have that $\mathrm{E}_r[u_2(I'(r))] = v_2$ since $I'(r)$ is a randomly selected interval of C. Note that player 2 can at any time deviate from (π_1, π_2) by playing along edge $e'(r) = (a_r \to b_{r+L_1 \bmod L})$ (e^* in reverse) to yield cycle $I'(r) \cup \{e'(r)\}$. Since (π_1, π_2) is an equilibrium, it must be that $E_r[u_2(e'(r))] \leq v_2$, since otherwise cycle $I'(r) \cup \{e'(r)\}$ would be a beneficial deviation for player 2 in expectation, a contradiction.

Let $\pi'_2(r)$ be a best response to $\pi'_1(r)$ that leads into a single cycle $D(r)$ (player 2 always has a best response that leads into a single cycle). Let $v_2^* = E_r[u_2(D(r))]$. If $v_2^* \leq u_2(C_2(r)) + \epsilon$, then there must exist some r such that $u_2(D(r)) \leq u_2(C_2(r)) + \epsilon$: that is, that π_2 is an ϵ-best response to $\pi'_1(r)$. Otherwise, we have $v_2^* > u_2(C_2(r)) + \epsilon \geq v_2$. The last inequality follows from equation 5.2, which gives us:

$$u_2(C_2(r)) \geq u_2(I(r)) \frac{L_1}{L_1+1} \geq$$

$$v_2 \frac{L-1}{L+1} \geq (1-\epsilon)v_2 \geq v_2 - \epsilon$$

We have two cases to consider: In the first case, there exists an r such that $D(r)$ does not contain edge $e'(r)$. In this case, $(a, \pi_1, \pi'_2(r))$ also yields cycle $D(r)$ (since π_1 and $\pi'_1(r)$ are identical except for edge $e'(r)$). Since (a, π_1, π_2) is an equilibrium, in this cases, $u_2(D(r)) \leq v_2$, and so $(a, \pi'_1(r), \pi'_2(r))$ is an ϵ-equilibrium. Therefore we can restrict our attention to the second case: for each r, $D(r)$ contains edge $e'(r)$. Recall that $E_r[u_2(e'(r))] \leq v_2$, and so it must be that $E_r[u_2(D - \{e'(r)\})] > v_2$. Now consider the path $D'(r) = D(r) - \{e'(r)\} \cup I'(r)$. Since $E_r[u_2(I'(r))] = v_2$ we have $E_r[u_2(D'(r))] > v_2$, and in particular, this must hold for some fixed r. But since $D'(r)$ does not include edge $e'(r)$, player 2 has some deviation f_2^* from the original equilibrium (a, π_1, π_2) such that (a, π_1, f_2^*) leads to path $D'(r)$. But this is a contradicts the fact that (a, π_1, π_2) is an equilibrium[3]. This completes the proof.

Proof. [Proof of Theorem] By our structural lemma, if $\mathcal{G}$ admits an equilibrium in pure stationary strategies, the algorithm will consider some cycle C^* that is part of an $\epsilon/2$ approximate equilibrium in pure stationary strategies. It remains to show that the resulting strategies and starting action a (a, π_1, π_2) that result from approximately solving the residual zero-sum mean payoff games $\mathcal{G}_1, \mathcal{G}_2$ form an $\epsilon/2$-approximate equilibrium. By assumption, there exists some pair of $\epsilon/2$-equilibrium strategies (π_1^*, π_2^*) such that $C^* = C(a', \pi_1^*, \pi_2^*) = C(a, \pi_1^*, \pi_2^*)$ (Since C traverses action a). That is, for all π'_1, π'_2, $P_1(a, \pi'_1, \pi_2^*) \leq u_1(C^*) + \epsilon/2$, and

[3] Note that $D'(r)$ is not necessarily a cycle (it may repeat some edges), and f_2^* is not necessarily a stationary strategy. However, the equilibrium condition holds in the space of all strategies, not just stationary strategies.

$P_2(a, \pi_1^*, \pi_2') \leq u_2(C^*) + \epsilon_2$. Since π_1^* and π_2^* both are consistent with cycle C^*, they represent possible strategies in the residual mean-payoff games $\mathcal{G}_2$ and $\mathcal{G}_1$ respectively. Therefore, by the definition of value, we have that the values of $\mathcal{G}_1$ and $\mathcal{G}_2$ satisfy $v_1(a) \leq u_1(C) + \epsilon/2$, and $v_2(a) \leq u_2(C) + \epsilon/2$. Since we compute $\epsilon/$ approximations to $v_1(a)$ and $v_2(a)$, v_1^* and v_2^* respectively, we have that $v_1(a) \leq u_1(C) + \epsilon$, and $v_2(a) \leq u_2(C) + \epsilon$. We therefore have that no player has an ϵ-improving deviation from $C^* = (a, \pi_1, \pi_2)$, and so (a, π_1, π_2) define an ϵ-approximate equilibrium by definition.

5.1 The Existence of Pure Stationary Strategy Equilibria We have given a PTAS to compute approximate equilibria in pure stationary strategies in games that admit equilibria in pure stationary strategies. How common are such games? It has long been known that all zero-sum games admit equilibria in pure stationary strategies. Moreover, in general sum games, Gurvich showed that for every integer k, all $2 \times k$ games admit equilibria in pure stationary strategies [17]. We include a new simple proof for the 2×2 case in appendix A, for intuition. Unfortunately, in subsequent work, Gurvich exhibited a 3×3 game that only admits equilibria in mixed stationary strategies [16]. Fortunately, pure stationary strategy equilibria are *common*. We show that a *random* $n \times n$ matrix game admits pure stationary strategy equilibria except with negligible probability.

THEOREM 5.2. *An $n \times n$ bi-matrix game with payoffs chosen independently and uniformly at random from the unit interval $[0, 1]$ admits an equilibrium in pure stationary strategies with probability at least $1 - 1/n^c$ for any constant c.*

Proof. We show that with high probability over a random game $\mathcal{G} = (A, B, u_1 \in_r [0, 1]^{n^2}, u_2 \in_r [0, 1]^{n^2})$, there exists a simple length-1 cycle consisting of a pair of actions $c^* = (a^*, b^*)$ that has utility for each player higher than the value of the residual zero-sum game outside of c^*. In particular, there exists a simple equilibrium if two events occur:

1. There exists an $a^* \in A$ and $b^* \in B$ such that $u_1(a^*, b^*) \geq 1 - \log(n)/4n$ and $u_2(a^*, b^*) \geq 1 - \log(n)/4n$.
2. There exists an $a^P \in A$ and $b^P \in B$ such that for all $a \in A$, $u_1(a, b^P) \leq 1 - \log(n)/2n$ and for all $b \in B$, $u_2(a^P, b) \leq 1 - \log(n)/2n$.

If both events occur, then the following strategies together with the starting point (a^*, b^*) constitute an equilibrium:

$$\pi_1(b) = \begin{cases} a^*, & b = b^*; \\ a^P, & \text{otherwise.} \end{cases} \qquad \pi_2(a) = \begin{cases} b^*, & a = a^*; \\ b^P, & \text{otherwise.} \end{cases}$$

To see this, note that if neither player deviates, then he achieves payoff at least $1 - \log(n)/4n$ because of event 1. Consider the deviations available to player 1: he can either deviate to play into another length 1 cycle (a', b^P) for some a' and achieve payoff at most $1 - \log(n)/2n < 1 - \log(n)/4n$ (by condition 2), or he can deviate to play into some length 4 cycle (a^*, b^*, a', b^P), and achieve payoff at most: $(1+1+(1-\log(n)/2n)+(1-\log(n)/2n))/4 \leq 1 - \log(n)/4n$. The analysis for player 2 is identical, and so neither player has any beneficial deviations given conditions 1 and 2. We now bound the probability of both events occurring:

First, we consider the probability of event 1, $Ev1$. Let $G(a, b)$ be the indicator random variable denoting the event that action pair $(a, b) \in A \times B$ satisfies the conditions of event 1. We therefore have $\mathrm{E}[G(a, b)] = \log^2(n)/(16n^2)$. Note that each of the n^2 variables $G(a, b)$ are independent. We therefore have:

$$\Pr[\overline{Ev1}] \leq \left(1 - \frac{\log^2(n)}{16n^2}\right)^{n^2} \leq$$

$$\exp\left(-\frac{\log^2(n)}{16}\right) = \frac{1}{n^{\log n/16}}$$

Next, we consider the probability of event 2, $Ev2$. Let $R(a)$ be the indicator random variable denoting the event that $a \in A$ satisfies the condition on a^P of event 2, i.e., $R(a)$ is the event that for every $b \in B$ we have $u_2(a, b) \leq 1 - \log(n)/2n$. Let $C(b)$ denote the event that $b \in B$ satisfies the conditions on b^P. For all $a \in A$ and $b \in B$,

$$\mathrm{E}[R(a)] = \mathrm{E}[C(b)] = (1 - \frac{\log(n)}{2n})^n$$

$$\geq \frac{1}{e \cdot \sqrt{2}}^{\log(n)/2} > \frac{1}{\sqrt{n}}$$

Therefore, $\mathrm{E}[\sum_{a \in A} R(a)] = \mathrm{E}[\sum_{b \in B} C(b)] > \sqrt{n}$, and so by a pair of Chernoff bounds and a union bound:

$$\Pr[\overline{Ev1}] \leq \Pr[\sum_{a \in A} R(a) = 0] + \Pr[\sum_{b \in B} C(b) = 0] \leq 2e^{-\sqrt{n}/2}$$

Since both events occur except with negligible probability, the theorem follows from a union bound.

REMARK 5.1. *In contrast, the probability of there existing a pure strategy simultaneous move equilibrium in a random game tends to $1 - 1/e$ as n tends to infinity [15]. Moreover, we believe that the above theorem actually substantially* underestimates *the probability of a random game admitting a pure equilibrium in stationary strategies. In searching for counterexamples to the* *conjecture that all games admit such equilibria, we generated 500,000 random 3x3 games and 100,000 random 4x4 games and found that* all *of them had pure equilibria in stationary strategies.*

We note that in the proof of theorem 5.2, we actually show that a randomly generated $n \times n$ bimatrix game admits a particularly *simple* stationary strategy equilibrium except with negligible probability: one that is parameterized by only 4 actions. An equilibrium of this type can be found efficiently simply by considering all combinations of 4 actions, and checking whether they can be used to form a pair of equilibrium strategies of this type. We therefore get the following corollary:

COROLLARY 5.1. *A pure stationary strategy equilibrium can be found in polynomial time, given an $n \times n$ bi-matrix game with payoffs chosen independently and uniformly at random from the unit interval* $[0, 1]$ *except with probability at most $1/n^c$ for any constant c.*

6 Discussion

Not all repeated interactions are best modeled as synchronized, simultaneous move repeated games. When agents take actions without a synchronizing mechanism and with knowledge of the past (as they do in models in which best response dynamics are studied), their interaction might be better modeled as an *alternating move* repeated game.

In this paper, we have shown that the problem of computing ϵ-approximate Nash equilibria in the alternating move model is strictly easier than in the simultaneous move model, because there exists an FPTAS to find equilibria in the spirit of the folk theorem, even for $k \geq 3$ players. Is the problem of computing exact equilibria easier? We note that even in the two-player zero-sum case (in which exact equilibria can be computed in the simultaneous move model via linear programming), we do not know of a polynomial time algorithm to compute *exact* equilibria in the alternating move model.

We also propose stationary strategy equilibria as a compelling solution concept for two-player general sum alternating move games. Unlike in simultaneous move games, in which pure strategy Nash equilibria fail to exist with constant probability even in randomly generated games, *pure* stationary strategy equilibria exist and can be found in polynomial time in randomly generated two player games, except with negligible probability. We note that best response dynamics constitute a specific, easy to compute set of stationary strategies. In what classes of games are they at equilibrium? We suggest that the alternating move game model provides a nice framework in which to ask this question.

7 Acknowledgements

The authors thank Greg Valiant for stimulating discussions. The first author would like to thank Vladimir Gurvich for pointing us to [17, 16] and graciously disabusing us of the conjecture that all games admit equilibria in pure stationary strategies.

References

[1] T. Abbott, D. Kane, and P. Valiant. On the complexity of two-player win-lose games. In *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, pages 113–122, 2005.

[2] B. Awerbuch, Y. Azar, A. Epstein, V.S. Mirrokni, and A. Skopalik. Fast convergence to nearly optimal solutions in potential games. In *Proceedings of the 9th ACM conference on Electronic commerce*, pages 264–273. ACM New York, NY, USA, 2008.

[3] I. Barany, S. Vempala, and A. Vetta. Nash equilibria in random games. *Random Structures and Algorithms*, 31(4), 2007.

[4] H. Björklund and S. Vorobyov. A combinatorial strongly subexponential strategy improvement algorithm for mean payoff games. *Discrete Applied Mathematics*, 155(2):210–229, 2007.

[5] C. Borgs, J. Chayes, N. Immorlica, A.T. Kalai, V. Mirrokni, and C. Papadimitriou. The myth of the folk theorem. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 365–372. ACM New York, NY, USA, 2008.

[6] H. Bosse, J. Byrka, and E. Markakis. New algorithms for approximate Nash equilibria in bimatrix games. *LECTURE NOTES IN COMPUTER SCIENCE*, 4858:17, 2007.

[7] P. Briest, P.W. Goldberg, and H. Roeglin. Approximate Equilibria in Games with Few Players. *Arxiv preprint arXiv:0804.4524*, 2008.

[8] X. Chen and X. Deng. Settling the complexity of two-player Nash equilibrium. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 261–272. IEEE Computer Society Washington, DC, USA, 2006.

[9] C. Daskalakis, P.W. Goldberg, and C.H. Papadimitriou. The complexity of computing a Nash equilibrium. 2009.

[10] C. Daskalakis, A. Mehta, and C. Papadimitriou. Progress in approximate Nash equilibria. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 355–358. ACM New York, NY, USA, 2007.

[11] B. Edelman and M. Ostrovsky. Strategic bidder behavior in sponsored search auctions. *Decision support systems*, 43(1):192–198, 2007.

[12] A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8(2):109–113, 1979.

[13] A. Fabrikant and C.H. Papadimitriou. The complexity of game dynamics: BGP oscillations, sink equilibria, and beyond. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 844–853. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2008.

[14] M. Goemans, V. Mirrokni, and A. Vetta. Sink equilibria and convergence. In *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, pages 142–151, 2005.

[15] K. Goldberg, AJ Goldman, and M. Newman. The probability of an equilibrium point. *Journal of Research of the National Bureau of Standards*, 72:93–101, 1968.

[16] V. Gurvich. A stochastic game with perfect information that has no Nash equilibrium in pure stationary strategies. *Russian Mathematical Surveys*, 43(2):171–172, 1988.

[17] V. Gurvich. The existence theorem for Nash equilibrium in pure stationary strategies for ergodic extensions of (2xk)-bimatrix games. *Russian Mathematical Surveys*, 45(4):170–171, 1990.

[18] VA Gurvich, AV Karzanov, and LG Khachivan. Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Computational Mathematics and Mathematical Physics*, 28(5):85–91, 1990.

[19] R.M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete mathematics*, 23(3):309–311, 1978.

[20] R.J. Lipton, E. Markakis, and A. Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM conference on Electronic commerce*, pages 36–41. ACM New York, NY, USA, 2003.

[21] H. Moulin. Extensions of two person zero-sum games. *Mémoires de la Société Mathématique de France*, 45:5–111, 1976.

[22] T. Roughgarden. Intrinsic Robustness of the Price of Anarchy. In *Proceedings of the 40th annual ACM symposium on Theory of computing.* ACM New York, NY, USA, 2009.

[23] A. Skopalik and B. Vöcking. Inapproximability of pure Nash equilibria. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 355–364. ACM New York, NY, USA, 2008.

[24] H. Tsaknakis and P.G. Spirakis. An optimization approach for approximate Nash equilibria. *Lecture Notes in Computer Science*, 4858:42, 2007.

[25] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1-2):343–359, 1996.

A Pure Stationary Strategy Equilibria

In this section, we write down games in matrix form. Given a matrix A, if $A(a,b) = (x,y)$ this denotes that $u_1(a,b) = x$ and $u_2(a,b) = y$.

THEOREM A.1. *All general sum games with $|A| = |B| = 2$ have an equilibrium in pure stationary strategies.*

Proof. Consider a generic two person matrix game $\mathcal{G}$:

	L	R
U	(a,b)	(c,d)
D	(e,f)	(g,h)

We first observe that if $\mathcal{G}$ has a pure strategy (simultaneous move) Nash equilibrium, then it also has a pure strategy alternating move stationary strategy equilibrium: without loss of generality, assume that (U,L) is a pure strategy Nash equilibrium: then the constant function strategies $\pi_1 \equiv U$, $\pi_2 \equiv L$ constitute an equilibrium of the alternating move game: they induce the cycle (U,L), and any deviation by the first (row) player either induces the same cycle, or cycle (D,L). Similarly, the second player can only deviate to induce cycle (U,R). However, by the assumption that (U,L) is a Nash equilibrium of the simultaneous move game, neither of these deviations yield improvements. Therefore, we may assume that $\mathcal{G}$ has no pure strategy Nash equilibria in the simultaneous move game, since otherwise we are done. So, without loss of generality, we have the following inequalities on the values in the payoff matrix:

$$e > a \qquad h > f \qquad c > g \qquad b > d$$

Consider the pair of strategies that induces the cycle (D,R,U,L) for which player 1 receives utility $(a+e+g+c)/4$ and player 2 receives utility $(b+f+h+d)/4$: $\pi_1(L) = D, \pi_1(R) = U, \pi_2(U) = L, \pi_2(D) = R$. If neither player has a beneficial deviation, we are done. Otherwise, assume player 1 has some beneficial deviation. Player 1 can deviate to either cycle (U,L) or (D,R). Without loss of generality, assume that the unique improving cycle is (U,L) (recall that it cannot be that both (U,L) and (D,R) improve over (D,R,U,L) by the inequalities above). In this case, the pair of strategies $\pi_1(L) = U, \pi_1(R) = U, \pi_2(U) = L, \pi_2(D) = R$ which leads into cycle (U,L) giving player 1 utility a and player 2 utility b is a stationary strategy equilibrium. This is because player 1 can only deviate to cycle (D,R,U,L) or cycle (D,R), both of which are not improvements, and player 2 can only deviate to cycle (U,R), but since $b > d$, this is not an improvement for player 2.

Unfortunately, there exists a 3x3 game that has no equilibrium in pure stationary strategies. Gurvich showed that the following game has no pure stationary strategy equilibria for ϵ sufficiently small [16]

	b_1	b_2	b_3
a_1	(0,1)	(0,0)	(1,0)
a_2	(ϵ,0)	(0,1)	(0,0)
a_3	(0,1 − ϵ)	(ϵ,0)	(0,1)