Games and Economic Behavior 85 (2014) 19–31

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

On the communication complexity of approximate Nash equilibria ☆

Paul W. Goldberg [a,*], Arnoud Pastink [b]

[a] University of Oxford, Dept. of Computer Science, Wolfson Building, Parks Road, Oxford OX1 3QD, UK
[b] Utrecht University, Department of Information and Computing Science, PO Box 80089, 3508TB Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 25 February 2013
Available online 3 February 2014

JEL classification:
C7
D83

Keywords:
Two player
Mixed strategy
Approximate equilibria
Efficient algorithms

ABSTRACT

We study the problem of computing approximate Nash equilibria of bimatrix games, in a setting where players initially know their own payoffs but not the other player's. In order to find a solution of reasonable quality, some amount of communication is required. We study algorithms where the communication is substantially less than the size of the game. When the communication is polylogarithmic in the number of strategies, we show how to obtain ϵ-approximate Nash equilibrium for $\epsilon \approx 0.438$, and for well-supported approximate equilibria we obtain $\epsilon \approx 0.732$. For one-way communication we show that $\epsilon = \frac{1}{2}$ is the best approximation quality achievable, while for well-supported equilibria, no value of $\epsilon < 1$ is achievable. When the players do not communicate at all, ϵ-Nash equilibria can be obtained for $\epsilon = \frac{3}{4}$; we also provide a corresponding lower bound of slightly more than $\frac{1}{2}$ on the smallest constant ϵ achievable.

1. Introduction

Algorithmic game theory is concerned not just with properties of a solution concept, but also how that solution can be obtained. It is considered desirable that the outcome of a game should be "easy to compute", which is typically formalized as polynomial-time computability, in the algorithms community. In that respect the PPAD-completeness results of Daskalakis et al. (2009a) and Chen and Deng (2006) are interpreted as a "complexity-theoretic critique" of Nash equilibrium. Following those results, a line of work addressed the problem of computing ϵ-Nash equilibrium, where $\epsilon > 0$ is a parameter that bounds a player's incentive to deviate, in a solution. Thus, ϵ-Nash equilibrium imposes a weaker constraint on how players are assumed to behave, and an exact Nash equilibrium is obtained for $\epsilon = 0$. The main open problem is to find out what values of ϵ admit a polynomial-time algorithm. Below we summarize some of the progress in this direction.

Beyond the existence of a fast algorithm, it is also desirable that a solution should be obtained by a process that is simple and decentralized, since that is likely to be a better model for how players in a game may eventually reach a solution. In that respect, most of the known efficient algorithms for computing ϵ-Nash equilibria are not entirely satisfying. They take as input the payoff matrices and output the approximate Nash equilibrium. If we try to translate such an algorithm into real life, it would correspond to a process where the players pass their payoffs to a central authority, which returns to them some mixed strategies that have the "low incentive to deviate" guarantee. In this paper we aim to model a setting where

☆ A preliminary version of this paper appeared in the Proceedings of the 5th SAGT. The first author was supported by EPSRC Grant EP/G069239/1 "Efficient Decentralised Approaches in Algorithmic Game Theory".

* Corresponding author.
E-mail addresses: Paul.Goldberg@cs.ox.ac.uk (P.W. Goldberg), A.J.Pastink@uu.nl (A. Pastink).

http://dx.doi.org/10.1016/j.geb.2014.01.009
0899-8256/

players perform individual computations and exchange some limited information. We revisit the question of what values of ϵ are achievable, subject to this restriction to more "realistic" algorithms.

There are various ways in which one can try to model the notion of a decentralized algorithm; here we consider a general approach that has previously been studied in Conitzer and Sandholm (2004) and Hart and Mansour (2010) in the context of computing exact Nash equilibria. The players begin with knowledge of their own payoffs but not the payoffs of the other players; this is often called an *uncoupled* setting (see Section 1.2.4 for an overview). An algorithm involves communication in addition to computation; to find a game-theoretic solution, a player usually has to know something about the other players' matrices, but hopefully not all of that information. We study the computation of ϵ-Nash equilibria in this setting, and the general topic is the trade-off between the amount of communication that takes place, and the value of ϵ that can be obtained. In uncoupled settings, there are natural dynamic processes that converge to correlated equilibria, but the results are less positive for exact and approximate Nash equilibria. This paper aims to contribute to the general goal of evaluating the merits of approximate Nash equilibrium as a solution concept, as opposed to (for example) exact or approximate correlated equilibrium.

1.1. Definitions

We consider 2-player games, with a *row player* and a *column player*, who both have n *pure strategies*. The game (R, C) is defined by two $n \times n$ *payoff matrices*, R for the row player, and C for the column player. The pure strategies for the row player are his rows and the pure strategies of the column player are her columns. If the row player plays row i and the column player plays column j, the *payoff* for the row player is R_{ij}, and C_{ij} for the column player. For the row player a *mixed strategy* is a probability distribution $\mathbf{x}$ over the rows, and a mixed strategy for the column player is a probability distribution $\mathbf{y}$ over the columns, where $\mathbf{x}$ and $\mathbf{y}$ are column vectors and $(\mathbf{x}, \mathbf{y})$ is a *mixed strategy profile*. The payoffs resulting from these mixed strategies $\mathbf{x}$ and $\mathbf{y}$ are $\mathbf{x}^{\mathrm{T}} R\mathbf{y}$ for the row player and $\mathbf{x}^{\mathrm{T}} C\mathbf{y}$ for the column player.

A *Nash equilibrium* is a pair of mixed strategies $(\mathbf{x}^*, \mathbf{y}^*)$ where neither player can get a higher payoff by playing another strategy assuming the other player does not change his strategy. Because of the linearity of a mixed strategy, the largest gain can be achieved by defecting to a pure strategy. Let $\mathbf{e}_i$ be the vector with a 1 at the i-th position and a 0 at every other position. Thus a Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies

$$\forall i = 1 \dots n \quad \mathbf{e}_i^{\mathrm{T}} R\mathbf{y}^* \leqslant \left(\mathbf{x}^*\right)^{\mathrm{T}} R\mathbf{y}^* \quad \text{and} \quad \left(\mathbf{x}^*\right)^{\mathrm{T}} C\mathbf{e}_i \leqslant \left(\mathbf{x}^*\right)^{\mathrm{T}} C\mathbf{y}^*.$$

We assume that the payoffs of R and C are between 0 and 1, which can be achieved by affine transformations. An ϵ*-approximate Nash equilibrium* (or, ϵ-Nash equilibrium) is a strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ such that each player can gain at most ϵ by unilaterally deviating to a different strategy. Thus, it is $(\mathbf{x}^*, \mathbf{y}^*)$ satisfying

$$\forall i = 1 \dots n \quad \mathbf{e}_i^{\mathrm{T}} R\mathbf{y}^* \leqslant \left(\mathbf{x}^*\right)^{\mathrm{T}} R\mathbf{y}^* + \epsilon \quad \text{and} \quad \left(\mathbf{x}^*\right)^{\mathrm{T}} C\mathbf{e}_i \leqslant \left(\mathbf{x}^*\right)^{\mathrm{T}} C\mathbf{y}^* + \epsilon.$$

We say that the *regret* of a player is the difference between his payoff and the payoff of his best response.

The *support* of a mixed strategy $\mathbf{x}$, denoted $\mathrm{Supp}(\mathbf{x})$, is the set of pure strategies that are played with non-zero probability by $\mathbf{x}$. An *approximate well-supported Nash equilibrium* strengthens the requirements of an approximate Nash equilibrium. For a mixed strategy $\mathbf{y}$ of the column player, a pure strategy $i \in [n]$ is an ϵ*-best response* for the row player if, for all pure strategies $i' \in [n]$ we have: $\mathbf{e}_i^{\mathrm{T}} R\mathbf{y} \geqslant \mathbf{e}_{i'}^{\mathrm{T}} R\mathbf{y} - \epsilon$. We define ϵ-best responses for the column player analogously. A mixed strategy profile $(\mathbf{x}, \mathbf{y})$ is an ϵ*-well-supported Nash equilibrium* (ϵ-WSNE) if every pure strategy in $\mathrm{Supp}(\mathbf{x})$ is an ϵ-best response against $\mathbf{y}$, and every pure strategy in $\mathrm{Supp}(\mathbf{y})$ is an ϵ-best response against $\mathbf{x}$.

The communication model: Each player $q \in \{r, c\}$ has an algorithm $\mathcal{A}_q$ whose initial input data is q's $n \times n$ payoff matrix. Communication proceeds in a number of rounds, where in each round, each player may send a single bit of information to the other player. During each round, each player may also carry out a polynomial (in n) amount of computation. (A natural variant of the model would omit the restriction to polynomial computation. Indeed, our lower bounds on communication requirement do not depend on computational limits.) At the end, each player q outputs a mixed strategy $\mathbf{x}_q$. We aim to design (pairs of) algorithms $(\mathcal{A}_r, \mathcal{A}_c)$ that output ϵ-Nash strategy profiles $(\mathbf{x}_r, \mathbf{x}_c)$, and are economical with the number of rounds of communication. This is similar to the *mixed Nash equilibrium procedure* of Hart and Mansour (2010), here applied to approximate rather than exact equilibria.

Notice that given $\Theta(n^2)$ rounds of communication, we can apply any centralized algorithm $\mathcal{A}$ by getting (say) the row player to pass additive approximations of all his payoffs to the column player, who applies $\mathcal{A}$ and passes to the row player the mixed strategy obtained by $\mathcal{A}$ for the row player. (The quality of the ϵ-Nash equilibrium is proportional to the quality of the additive approximations used.) For this reason we focus on algorithms with many fewer rounds, and we obtain results for logarithmic or polylogarithmic (in n) rounds.

We also consider a restriction to *one-way communication*, where one player may send but not receive information.

1.2. Related work

We start by reviewing some algorithms that we adapt to the communication-bounded setting. Then we review the background work on communication complexity, and related work in computing Nash equilibria, including learning of equilibria in *uncoupled* settings.

1.2.1. Algorithms for approximate equilibria

In recent years a number of algorithms (Kontogiannis et al., 2006; Daskalakis et al., 2009b, 2007; Bosse et al., 2007; Tsaknakis and Spirakis, 2007) have been developed that compute (in polynomial time) ϵ-Nash equilibria for various values of ϵ. Of these, Tsaknakis and Spirakis (2007) obtain the best (smallest) value of ϵ, of approximately 0.3393; for the special case of symmetric games, ϵ-values approaching $\frac{1}{3}$ can be obtained (Kontogiannis and Spirakis, 2011). The more demanding criterion of *well-supported* ϵ-Nash equilibrium, disallows a player from allocating positive probability to any pure strategy whose payoff is more than ϵ worse than the best response. Progress on polynomial-time algorithms for this solution concept has been more limited; at this time the lowest ϵ that can be guaranteed by a polynomial-time algorithm is only slightly less than $\frac{2}{3}$ (Fearnley et al., 2012), obtained via a modification of a $\frac{2}{3}$-approximation algorithm of Kontogiannis and Spirakis (2010). Prior to that, Daskalakis et al. (2009b) gave a $\frac{5}{6}$-approximation algorithm, that is contingent on a graph-theoretic conjecture. In this context, our 0.732-approximation algorithm substantially improves on the result of Daskalakis et al. (2009b), both in terms of approximation quality and a more demanding model (communication-bounded algorithms). However, we do not know how to obtain the better approximation quality of Kontogiannis and Spirakis (2010) and Fearnley et al. (2012) in the communication-bounded setting. Next we discuss two of the earlier algorithms in the literature whose ideas we use here.

DMP-algorithm: The DMP-algorithm (Daskalakis et al., 2009b) works as follows to achieve a $\frac{1}{2}$-approximate Nash equilibrium. The algorithm picks an arbitrary row for the row player, say row i. Let $j \in \arg\max_{j'} C_{ij'}$. Let $k \in \arg\max_{k'} R_{k'j}$. So j is a pure-strategy best response for the column player to row i and k is a best response strategy for the row player to column j. The strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ will now be $\mathbf{x}^* = \frac{1}{2}\mathbf{e}_i + \frac{1}{2}\mathbf{e}_k$ and $\mathbf{y}^* = \mathbf{e}_j$. With this strategy pair the row player plays a best response with probability $\frac{1}{2}$ to a pure strategy of the column player and the column player has a pure strategy that is with probability $\frac{1}{2}$ a best response.

The DMP-algorithm is well-adapted to the limited-communication setting. Suppose the row player uses $i = 1$ as his initial choice of row. The column player needs to tell the row player her value of j, a communication of $O(\log n)$ bits. No further communication is needed. Notice moreover that the communication is all one-way; the row player does not need to tell the column player anything.

Subsequent algorithms for computing ϵ-Nash equilibria cannot so easily be adapted to a limited-communication setting, but we can use some of the ideas they develop, to obtain values of ϵ below $\frac{1}{2}$ in this setting.

An algorithm of Bosse et al. (2007): The algorithm presented in Bosse et al. (2007) can be seen as a modification of the DMP-algorithm and achieves a 0.38197-approximate Nash equilibrium. Instead of a player allocating some probability to some arbitrary pure strategy, the algorithm starts with the row player allocating some probability to the row-player strategy $\mathbf{x}$ belonging to the Nash equilibrium of the zero-sum game $(R - C, C - R)$. In solving the zero-sum game efficiently we apply the connection of zero-sum games with linear programming (Neumann, 1928; Dantzig, 1963; Karmarkar, 1984). If the (mixed) strategy profile $(\mathbf{x}, \mathbf{y})$ that constitutes a Nash equilibrium of $(R - C, C - R)$ gives a 0.38197-approximate Nash equilibrium for (R, C), this solution is used. Otherwise, the column player plays a best response $\mathbf{e}_j$ to $\mathbf{x}$ and the row player plays a mixture of $\mathbf{x}$ and $\mathbf{e}_k$, where $\mathbf{e}_k$ is a best response to the strategy $\mathbf{e}_j$ of the column player. (Bosse et al., 2007, goes on to improve the worst-case performance to a 0.36395-approximate Nash equilibrium.)

Notice that this algorithm cannot be adapted in a straightforward way to our communication-bounded setup, since it requires a computation using knowledge of both matrices. The starting-point of our algorithms of Section 4 is the players separately solving $(R, -R)$ and $(-C, C)$.

1.2.2. Communication complexity

The "classical" setting of communication complexity is based on the model introduced by Yao (1979). We will follow the representation in Kushilevitz (1997). We have two agents,[1] one holding an input $\mathbf{x} \in \{0, 1\}^n$ and the other holding an input $\mathbf{y} \in \{0, 1\}^n$. The objective is to compute $f(\mathbf{x}, \mathbf{y}) \in \{0, 1\}$, a joint function of their inputs. The computation of $f(\mathbf{x}, \mathbf{y})$ is done via a communication protocol $\mathcal{P}$. During the execution of the protocol, the agents send messages to each other. While the protocol has not terminated, the protocol specifies what message the sender should send next, based on the input of the protocol and the communication so far. If the protocol terminates, it will output the value $f(\mathbf{x}, \mathbf{y})$. A communication protocol $\mathcal{P}$ computes f if for every input pair $(\mathbf{x}, \mathbf{y}) \in \{0, 1\}^n \times \{0, 1\}^n$, it terminates with the value $f(\mathbf{x}, \mathbf{y})$ as output.

The communication complexity of a communication protocol $\mathcal{P}$ for computing $f(\mathbf{x}, \mathbf{y})$ is the number of bits sent during the execution of $\mathcal{P}$, which we denote by $CC(\mathcal{P}, f, \mathbf{x}, \mathbf{y})$. The communication complexity of a protocol $\mathcal{P}$ for a function f is

[1] We use agents instead of players to avoid confusion, the communication does not have to be between the players of the game.

defined as the worst case communication complexity over all possible inputs for $(\mathbf{x}, \mathbf{y}) \in \{0,1\}^n \times \{0,1\}^n$, which we denote by $CC(\mathcal{P}, f)$:

$$CC(\mathcal{P}, f) = \max_{(\mathbf{x},\mathbf{y}) \in \{0,1\}^n \times \{0,1\}^n} CC(\mathcal{P}, f, \mathbf{x}, \mathbf{y}).$$

The communication complexity of a function f is the minimum over all possible protocols:

$$CC(f) = \min_{\mathcal{P}} CC(\mathcal{P}, f).$$

1.2.3. Existing results on communication complexity of Nash equilibria

There are a few results concerning the communication complexity of Nash equilibria. Conitzer and Sandholm (2004) show a lower bound on the communication complexity for 2-player games of finding a pure Nash equilibrium of $\Omega(n^2)$, where n is the number of pure strategies for each player. They also show a simple algorithm that finds a pure Nash equilibrium (if it exists) in $O(n^2)$. They do not extend their analysis to mixed Nash equilibria; their focus is on searching for a pure Nash equilibrium (if one exists), in contrast with the existence of a mixed Nash equilibrium, which is guaranteed (Nash, 1951). For unrestricted bimatrix games, it can be seen that the communication complexity of finding an exact equilibrium is $\Omega(n^2)$.[2] That observation leads to the question addressed here, of whether approximate equilibria have lower communication complexity.

Also related to this paper, Hart and Mansour (2010) study the communication complexity of uncoupled equilibrium procedures (discussed in more detail below in Section 1.2.4) in the context of multiplayer, binary action games. The emphasis is on lower bounds on the communication requirement. Analogously to the $\Omega(n^2)$ communication needed for pure or mixed Nash equilibrium that we noted above, they obtain a lower bound of $\Omega(2^s)$ (where s is the number of players) on the communication needed to find an exact mixed equilibrium, or determine the existence of a pure one. (Note that in their setting, each player has a payoff matrix of size 2^s, so that essentially all the payoffs may need to be communicated.) On the other hand, they obtain a polynomial upper bound on the communication required to find a *correlated equilibrium*, discussed further below. Their methods do not seem to be applicable in an obvious way to approximate equilibria. For example, the lower bound for computing a mixed equilibrium involves a game whose solution requires probabilities having exponentially large descriptions, which would not be needed in the context of approximate equilibria.

1.2.4. Uncoupled learning of equilibria

An extensive literature studies *uncoupled* procedures for finding game-theoretic solutions. The terminology "uncoupled" is introduced in Hart and Mas-Colell (2003); it refers to settings where each player knows his own (but not the others') utility function. Then, there is a sequence of rounds (a.k.a. time steps, or periods), in which each player plays a strategy, and receives the payoff resulting from the entire strategy profile. Our setting of communication complexity is related to this, in that each player can use his choice of action (in a round) to transmit information. The main difference is that here, we do not assume a "rational" choice of action where a player tries to maintain his payoff over time by predicting the choices of his opponents. In our set-up, players communicate some information over a (hopefully short) sequence of rounds, and afterwards promise to use certain mixed strategies. Our interest is in both upper and lower bounds on the required length of the sequence. As noted in Conitzer and Sandholm (2004), lower-bound type results generally ignore strategic considerations, which perhaps helps to justify our own inattention to rationality in this paper.

In the context of uncoupled search for Nash equilibrium, Hart and Mas-Colell (2003) show that when players do not remember the history of play, it may be impossible to reach Nash equilibrium. Note that the obstacle is informational rather than due to rationality of the players. A subsequent paper of Hart and Mas-Colell (2006) analyzes how much of the history of play needs to be recalled by the players. In the case of mixed (approximate) Nash equilibria, the approach is to test many probability distributions in a search for one that constitutes an approximate equilibrium; a large number of rounds is required to achieve this. Foster and Young (2006) show how this can be achieved in a "radically uncoupled" setup, where a player does not directly observe the opponents' behavior, but observes it indirectly via the payoffs he obtains. Again, a very large number of rounds are required to find an approximate equilibrium. Daskalakis et al. (2010) study negative results, namely failure to converge to Nash equilibrium, for standard multiplicative weights update algorithms, in the context of bimatrix games. Their results consider three variants of uncoupled dynamics.

There are more natural learning algorithms that converge (in various senses) to the weaker solution concept of correlated equilibrium (e.g. Foster and Vohra, 1997; Hart and Mas-Colell, 2000). When we relax our objective from approximate Nash equilibrium to approximate correlated equilibrium, then learning can take place with a sublinear number of rounds, from a straightforward application of no-regret learning algorithms. The idea is applied in Theorem 30 of Hart and Mansour (2010). In particular, we equip each player[3] with a no-regret algorithm, and suppose that at each round it duly selects (and

[2] Consider a game where there is a unique, fully-mixed Nash equilibrium. If the payoffs are perturbed slightly, the resulting equilibrium, for (say) the row player, will be affected in a non-trivial way by all the perturbations of the column player's payoffs. This immediately results in the requirement of $\Omega(n^2)$ communication.

[3] Indeed, there may be any number of players, not just 2.

outputs) a pure strategy, which requires $\log(n)$ bits to output. Indeed, Theorem 17 of Hart and Mansour (2010) shows how *exact* correlated equilibrium may be found in a polynomial number of rounds.

Foster and Young (2006) point out as motivation for uncoupled learning rules, that uncoupledness prevents a learning rule from behaving like a centralized algorithm and just constituting a theory of equilibrium selection. In this paper we similarly avoid the possibility of implementing a centralized algorithm, though restricting to a sublinear number of rounds of communication, so that it is impossible for one player to reveal all (or even a large fraction) of his payoffs to the other player.

1.3. Overview of our results

For general $n \times n$ games we show the following bounds on the obtainable quality of an approximate Nash equilibrium if we fix the amount of communication allowed. We start by considering a version where no communication is allowed. Theorem 1 gives a simple way to find a $\frac{3}{4}$-Nash equilibrium, in this setting. Theorem 2 identifies a corresponding lower bound of slightly more than $\frac{1}{2}$. For one-way communication we exhibit (Theorem 3) a lower bound of $\frac{1}{2} - o(\frac{1}{\sqrt{n}})$. The DMP-algorithm can be implemented as an algorithm with one-way communication and gives a $\frac{1}{2}$-approximate Nash equilibrium. Therefore the constant $\frac{1}{2}$ in the lower bound of Theorem 3 is tight, in this context. In Section 4.1 we show how to compute a 0.438-Nash equilibrium using polylogarithmic communication. In Section 5 we discuss the significance of the results, along with open problems.

2. Approximate Nash equilibria with no communication

The simplest way to restrict communication is to disallow it entirely.[4] That means that for each player $q \in \{r, c\}$, we must find a function f_q from q's payoff matrix to a mixed strategy, such that for all pairs of matrices (R, C), we have that $(f_r(R), f_c(C))$ is an ϵ-Nash equilibrium. In this section we show that the achievable value of ϵ lies somewhere between 0.501 and $\frac{3}{4}$. The $\frac{3}{4}$ upper bound is achieved via a simple algorithm (differing from the $\frac{3}{4}$-approximation algorithm of Kontogiannis et al., 2006, in terms of the solution it finds). Theorem 2 presents the lower bounds of 0.501.

Theorem 3 in Section 3 furnishes a lower bound of $\frac{1}{2}$, even when one-way communication is permitted, and has a simpler proof (the proof is similar to Case 2 in the proof of Theorem 2). This raises the question: why bother to include a complicated proof (specific to the communication-free setting) whose result is only a small improvement (over the one-way communication setting)? The reason is that we rule out the possibility that $\frac{1}{2}$ is in fact the answer, and as we discuss in the conclusions (Section 5), $\frac{1}{2}$ seems to arise frequently as a barrier to progress in the study of algorithms for approximate Nash equilibria, so it is informative to rule out that possibility. Our lower bound of 0.501 could be increased slightly by tweaking the parameters of the proof, but we believe that the resulting progress would be incremental.

Theorem 1. *It is possible to guarantee a $\frac{3}{4}$-approximate Nash equilibrium, even if there is no communication between the players.*

Proof. Each player allocates probability $\frac{1}{2}$ to his first pure strategy, and $\frac{1}{2}$ to his best response to the other player's first pure strategy. In detail, let $i \in \arg\max_{i'} R_{i'1}$ and let $j \in \arg\max_{j'} C_{1j'}$. The approximate Nash equilibrium will be $f_r(R) = \frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_i$ and $f_c(C) = \frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_j$.

Let i' be a best pure strategy response of the row player to $f_c(C)$. Then his incentive to deviate is

$$\left(\frac{1}{2}R_{i'1} + \frac{1}{2}R_{i'j}\right) - \left(\frac{1}{4}R_{11} + \frac{1}{4}R_{1j} + \frac{1}{4}R_{i1} + \frac{1}{4}R_{ij}\right)$$
$$\leqslant \left(\frac{1}{4}R_{i'1} + \frac{1}{2}R_{i'j}\right) - \left(\frac{1}{4}R_{11} + \frac{1}{4}R_{1j} + \frac{1}{4}R_{ij}\right) \leqslant \frac{1}{4}R_{i'1} + \frac{1}{2}R_{i'j} \leqslant \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

where the first inequality holds because i was a best response to column 1 (so $R_{i1} \geqslant R_{i'1}$) and the next inequalities hold because payoffs lie in $[0, 1]$. The same kind of argument holds for the column player. This proves the theorem. □

The following lemma provides a construction that is used in Theorems 2 and 3.

Definition 1. Let M_n be a matrix with n columns and $\binom{n}{k}$ rows, where $k = \lfloor\sqrt{n}\rfloor$ and a row consists of k 1's and $(n-k)$ 0's. Every row is distinct, so the $\binom{n}{k}$ rows are all the possible sequences with k 1's in a row of length n.

[4] This is to some extent inspired by earlier work of the first author (Goldberg, 2006) that studied an approach to pattern classification in which the set of observations of each class must be processed by an algorithm that proceeds independently of the corresponding algorithms that receive members of the other classes.

Lemma 1. *Suppose we have a bimatrix game where the row player's payoff matrix is M_n (as in Definition 1). Let $\mathbf{x}$ be a mixed strategy for the row player. Then, there exists a column of M_n such that if the column player uses any mixed strategy $\mathbf{y}$ that allocates probability p to that column, then the row player's regret is at least $p - O(1/\sqrt{n})$.*

Proof. The rows of M_n contain 1's in a fraction $\frac{k}{n}$ of their entries. By symmetry, so do the columns, thus every column contains $\frac{k}{n} \cdot \binom{n}{k}$ 1's and $(1 - \frac{k}{n}) \cdot \binom{n}{k}$ 0's (recall $k = \lfloor\sqrt{n}\rfloor$).

$\mathbf{x}$ assigns a probability to each row of M_n. Define an unnormalized probability distribution Φ over the columns as follows. Let $\Phi(j)$ be the probability that a 1 will be in column j of M_n, given a row sampled from $\mathbf{x}$, thus $\Phi(j)$ is the expected value of the entry of the j-th column, for rows sampled from $\mathbf{x}$. Note that $\Phi(j) \leqslant 1$, with equality when every row that is played with positive probability has a 1 in column j. Because every row contains k 1's, the $\Phi(j)$ values will sum to k: $\sum_{j=1}^{n} \Phi(j) = k$.

We define column m to be one with a lowest value of Φ: $m \in \arg\min_j \Phi(j)$. m is the column that minimizes the row player's payoff. We choose m to be the special column in the statement of the lemma, and we suppose that the column player allocates probability p to m.

Since the sum over all values $\Phi(j)$ is k and there are n columns, this means that $\Phi(m) \leqslant \frac{k}{n}$. When column m is played (and we assume it is played with probability p) it gives the row player a payoff of 0 with a probability of at least $1 - \frac{k}{n}$.

We now consider the row player's strategy $\mathbf{x}$ and construct an improved response $\mathbf{x}^*$ as follows. $\mathbf{x}^*$ will differ from $\mathbf{x}$ in the following way. For every row i we see if its m-th entry is a 1. If this is the case, we do not change anything. If instead its m-th entry is a 0, we do the following: look at the entries where there is a 1 in row i. Of all the entries where there is a 1, we select the one to which the column player's distribution $\mathbf{y}$ gives the lowest probability, say entry a (i.e. choose column $a \in \arg\min_{j:\ M_n[i,j]=1} \mathbf{y}[j]$). Now we move all the probability allocated to row i by $\mathbf{x}$, to the row of M_n that instead has a 0 in entry a and a 1 in entry m, and is otherwise the same as i.

The probability on entry a is defined as the smallest among all the entries where row i has a 1. We can bound the probability that is allocated to this entry by distribution $\mathbf{y}$. A probability at least p is given to column m, so a probability of $1 - p$ can be distributed over the remaining columns. The column containing entry a has the smallest probability among at least k columns, so the probability given to column a is at most $\frac{1-p}{k}$.

The result of this construction of $\mathbf{x}^*$ from $\mathbf{x}$ is that every row that is played with positive probability by $\mathbf{x}^*$ will have a 1 in the m-th entry. There is a probability at least $(1 - \frac{k}{n})$ that a row sampled from $\mathbf{x}$ does not have a 1 in the m-th entry. This means that the increase in payoff from replacing $\mathbf{x}$ with $\mathbf{x}^*$ is at least

$$\left(1 - \frac{k}{n}\right) \cdot p - \left(1 - \frac{k}{n}\right) \cdot \frac{1-p}{k} = \left(1 - \frac{k}{n}\right) \cdot \left(p - \frac{1-p}{k}\right) \geqslant p - \frac{1}{k}.$$

Noting that $k = \lfloor\sqrt{n}\rfloor$ gives us the desired result. □

We use the following technical extension in the proof of Theorem 2. It is a straightforward corollary of Lemma 1.

Corollary 1. *Suppose we have a bimatrix game where the row player's payoff matrix R has M_n (as in Definition 1) as a submatrix. Suppose furthermore that all rows of R that do not intersect M_n pay the row player zero, and for all rows of R that intersect M_n, all columns that are not columns of M_n pay the row player 1.*

Let $\mathbf{x}$ be any mixed strategy for the row player that allocates probability at least p_r to rows that do not intersect M_n. Let $\mathbf{y}$ be a mixed strategy for the column player, that allocates at least p_c to columns that do not intersect M_n, but allocates probability p to some column ℓ intersecting M_n. Then, there exists a choice of column ℓ such that the row player's regret is at least $p - O(1/\sqrt{n}) + p_r p_c$.

Proof. Suppose $\mathbf{x}$ is modified as follows. For rows that intersect M_n, modify their probabilities according to Lemma 1. For other rows, set their probability to 0, and transfer their probability to an arbitrary row that has payoff 1 when column ℓ is played.

This change increases by $p - O(1/\sqrt{k})$, the payoffs to the row player resulting from the column player playing columns containing M_n. Note that this gain is not conditioned on the column player playing columns intersecting M_n; it is an absolute gain. In more detail, for rows intersecting M_n, a fraction $1 - \frac{k}{n}$ of them (w.r.t. probability measure $\mathbf{x}$) have their payoff raised by at least $p - \frac{1-p}{k}$. For rows not intersecting M_n, their payoffs are raised by at least p.

There is an additional increase to the row player's payoff due to the transfer of probability from rows not intersecting M_n to rows intersecting M_n, in the event that the column player plays a column not containing M_n. In this case the payoffs increase from 0 to 1, resulting in an additional payoff to the row player of (at least) $p_r.p_c$. □

In the communication-free setting, each player q computes a function f_q from his payoff matrix to a mixed strategy. We will first introduce a "radius" measure ρ^q that measures the variability of mixed strategies that may be selected by q, i.e. the image of the set of all payoff matrices under f_q. Essentially ρ^q is the radius of a smallest enclosing sphere containing q's available strategies, under the variation distance metric. We continue by giving a precise definition and notation.

The variation distance between two probability distributions $\mathbf{x}$ and $\mathbf{x}'$ over $[n]$, is half the sum of all positive differences between the two distributions, i.e.

$$d(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{n} \frac{1}{2}\big|\mathbf{x}[i] - \mathbf{x}'[i]\big|.$$

For $n \times n$ games, let Ω_n^r denote the set of strategies the row player may use (i.e. the image of f_r) and Ω_n^c the set of strategies the column player may use. For each player we define his "center strategy". For the row player let $\mathbf{c}_n^r$ be some probability distribution such that the maximum distance between $\mathbf{c}_n^r$ and any strategy $\boldsymbol{\omega} \in \Omega_n^r$ is minimized.

$$\mathbf{c}_n^r \in \arg\min_{\mathbf{c}} \sup_{\boldsymbol{\omega}\in\Omega_n^r} d(\mathbf{c},\boldsymbol{\omega}).$$

The center distribution $\mathbf{c}_n^c$ of the column player is defined in a similar way. The *radius* ρ_n^r of the row player is defined as

$$\rho_n^r = \sup_{\boldsymbol{\omega}\in\Omega_n^r} d\big(\mathbf{c}_n^r,\boldsymbol{\omega}\big).$$

The radius ρ_n^c of the column player is defined similarly. This radius measure ρ_n^r is a value in $[0,1]$ that indicates the variability of strategies a player may use, and is low when the player always plays strategies that are close to some "central" strategy.

Theorem 2. *For bimatrix games with payoffs in the range* $[0,1]$, *if each player independently computes a mixed strategy based on his own payoff matrix, then it is impossible to guarantee an* ϵ*-approximate Nash equilibrium for* $\epsilon < 0.501$.

Proof. The proof will be a case analysis on radius. In the proof, our analysis is with respect to an arbitrary fixed value of n, so we drop the subscript n from the radius values ρ_n^r and ρ_n^c, also the center probability vectors $\mathbf{c}_n^r$ and $\mathbf{c}_n^c$. We will show that for all n, the worst-case regret of a player is at least 0.501. We identify two cases:

1. A player has a high radius: $\rho^r \geqslant 0.95$ or $\rho^c \geqslant 0.95$.
2. Neither player has a high radius: $\rho^r < 0.95$ and $\rho^c < 0.95$.

Case 1: *A player has high radius*

Assume the column player has high radius, thus $\rho^c \geqslant 0.95$. We use this high radius to identify a set of strategies that are quite far apart from each other, under variation distance.

For the column player, take an arbitrary strategy $\mathbf{s}_1 \in \Omega^c$. Because $\rho^c \geqslant 0.95$, there must be some strategy $\mathbf{s}_2$ with $d(\mathbf{s}_1,\mathbf{s}_2) \geqslant 0.95$, otherwise $\mathbf{s}_1$ could be the center strategy $\mathbf{c}$ with $\rho^c < 0.95$.

Now consider the strategy $\mathbf{s}_{12} = \frac{\mathbf{s}_1+\mathbf{s}_2}{2}$, thus $d(\mathbf{s}_{12},\mathbf{s}_1) = d(\mathbf{s}_{12},\mathbf{s}_2) = \frac{1}{2}d(\mathbf{s}_1,\mathbf{s}_2) \leqslant \frac{1}{2}$. For this strategy not to be a center strategy $\mathbf{c}$ contradicting $\rho^c \geqslant 0.95$, there must be some strategy $\mathbf{s}_3 \in \Omega^c$ with $d(\mathbf{s}_{12},\mathbf{s}_3) \geqslant 0.95$. Because $\mathbf{s}_1$ constitutes half of the strategy $\mathbf{s}_{12}$, it holds that $d(\mathbf{s}_1,\mathbf{s}_3) \geqslant 0.90$ and similarly $d(\mathbf{s}_2,\mathbf{s}_3) \geqslant 0.90$. We have

$$d(\mathbf{s}_1,\mathbf{s}_2) \geqslant 0.95; \qquad d(\mathbf{s}_1,\mathbf{s}_3) \geqslant 0.90; \qquad d(\mathbf{s}_2,\mathbf{s}_3) \geqslant 0.90; \qquad d(\mathbf{s}_{12},\mathbf{s}_3) \geqslant 0.95.$$

The next step is to construct an $n \times n$ payoff matrix R of the row player. Only the first 3 rows of R will contain non-zero entries. The construction of rows 1, 2, 3 will be such that for $i, j \in \{1,2,3\}$, row i is a best response to s_i and a poor response to s_j $(j \neq i)$.

For every column j of R determine the maximum of $\mathbf{s}_1[j]$, $\mathbf{s}_2[j]$ and $\mathbf{s}_3[j]$. If $\mathbf{s}_1[j]$ is the largest, $R_{1j} = 1$ and $R_{2j} = R_{3j} = 0$. If $\mathbf{s}_2[j]$ is the largest, $R_{2j} = 1$ and $R_{1j} = R_{3j} = 0$. If $\mathbf{s}_3[j]$ is the largest, $R_{3j} = 1$ and $R_{1j} = R_{2j} = 0$. In case of a tie in the comparison of $\mathbf{s}_1[j]$, $\mathbf{s}_2[j]$ and $\mathbf{s}_3[j]$, all the entries corresponding to the tie get a 1.

Consider columns i for which $R_{2i} = 1$, so that $\mathbf{s}_2[i] > \mathbf{s}_1[i]$. The total probability assigned by $\mathbf{s}_1$ to these columns is bounded by 0.05. If the probability on these columns were higher than 0.05, it would follow that $d(\mathbf{s}_1,\mathbf{s}_2) < 0.95$. Similarly we can bound the probability assigned by $\mathbf{s}_1$ to columns i such that $R_{3i} = 1$. Since $d(\mathbf{s}_1,\mathbf{s}_3) \geqslant 0.9$ this probability is at most 0.1. From these observations, we have that at most 0.15 of the probability distribution $\mathbf{s}_1$ is allocated to columns that could give a payoff of 0 for row 1. Since each column of R contains at least one 1, the remaining 0.85 probability of $\mathbf{s}_1$ will be allocated to columns that have a 1 in the corresponding entry of row 1. The payoff for row 1 if the column player plays $\mathbf{s}_1$ is therefore at least 0.85. We can use a similar argument to claim that when the column player plays $\mathbf{s}_2$, the row player can get a payoff of at least 0.85 by playing pure strategy row 2, and at most 0.05 for row 2, and at most 0.1 for row 3.

For row 3 we use $d(\mathbf{s}_{12},\mathbf{s}_3) \geqslant 0.95$. Consider columns i for which $R_{3i} = 0$, so that either $R_{1i} = 1$ or $R_{2i} = 1$. A column i having this property, contributes $\geqslant \frac{1}{2}\mathbf{s}_3[i]$ to the overlap between $\mathbf{s}_3$ and $\mathbf{s}_{12}$. Indeed, if both $R_{1i} = 1$ and $R_{2i} = 1$, it contributes $\mathbf{s}_3[i]$ to the overlap. So we can deduce that with respect to columns selected using $\mathbf{s}_3$, Pr[row 1 pays 1] + Pr[row 2 pays 1] $\leqslant 0.1$. Again, since each column of R contains at least one 1, the remaining 0.9 probability of $\mathbf{s}_3$ will be allocated to columns that have a 1 in the corresponding entry of row 3. The payoff for row 3 if the column player plays $\mathbf{s}_3$ is therefore at least 0.9, while the payoffs to rows 1 and 2 sum to at most 0.1. To summarize:

- If the column player plays $\mathbf{s}_1$, the row player gets a payoff of at least 0.85 by playing row 1. Playing row 2 would give him a payoff of at most 0.05 and playing row 3 a payoff of at most 0.1.
- If the column player plays $\mathbf{s}_2$, the row player gets a payoff of at least 0.85 by playing row 2. Playing row 1 would give him a payoff of at most 0.05 and playing row 3 a payoff of at most 0.1.
- If the column player plays $\mathbf{s}_3$, the row player gets a payoff of at least 0.9 by playing row 3. Playing row 2 would give him a payoff of at most 0.1 and playing row 3 a payoff of at most 0.1. Moreover, the sum of payoffs of row 1 and row 2 is at most 0.1.

Given the row player's strategy, let (r_1, r_2, r_3) be the probabilities with which he plays rows 1, 2, 3. Assume $r_1 \leqslant r_2, r_3$, so $r_1 \leqslant \frac{1}{3}$ and suppose the column player plays strategy $\mathbf{s}_1$. The best response strategy $(1, 0, 0)$ has a payoff of $a \in [0.85, 1]$. Because row 1 clearly gives the highest payoff, the regret is minimized when this row is played with as much probability as possible, so $r_1 = \frac{1}{3}$. Because the probability on row 1 was defined as the lowest probability, the probability on the other two rows is also $\frac{1}{3}$. This gives a regret of at least

$$a - \left(\frac{1}{3}a + \frac{1}{3}(0.05) + \frac{1}{3}(0.1)\right) = \frac{2}{3}a - 0.05 \geqslant \frac{2}{3}(0.85) - 0.05 \approx 0.517.$$

The analysis for $r_2 \leqslant r_1, r_3$ where the column player plays $\mathbf{s}_2$ is similar.

Assume $r_3 \leqslant r_1, r_2$ and the column player plays $\mathbf{s}_3$. The best response to $\mathbf{s}_3$ has a payoff of at least 0.9 and row 1 and 2 combined can have a payoff of at most 0.1. This gives a regret of at least

$$a - \left(\frac{1}{3}a + \frac{1}{3}(0.1)\right) = \frac{2}{3}a - \frac{1}{30} \geqslant \frac{2}{3}(0.9) - \frac{1}{30} \approx 0.567.$$

So regardless of the strategy of the row player, the worst-case regret of the row player is always larger than 0.501 when the radius of the other player is at least 0.95.

Case 2: *Neither player has high radius*

Suppose both players have radius $\rho^r, \rho^c < 0.95$. Consider the following set of payoff matrices for the column player: $C^1, \ldots, C^n$ where C^ℓ has a payoff of 1 for every entry in the ℓ-th column and a 0 elsewhere:

$$\forall i, j: \quad C^\ell_{ij} = 1 \quad \text{if } j = \ell; \quad 0 \quad \text{otherwise}.$$

To achieve a 0.501-approximate Nash equilibrium, when the column player has payoff matrix C^ℓ, the column player should assign at least 0.499 to column ℓ.

The construction of the payoff matrix R of the row player will depend on the center strategy $\mathbf{c}^r$ of the row player. Take the $(n - \sqrt{n})$ rows of R which have the highest values $\mathbf{c}^r[i]$, where $\mathbf{c}^r[i]$ is the i-th entry of $\mathbf{c}^r$. We construct matrix R for which these rows are all zero. For the construction of the remaining $\sqrt{n}$ rows of R we consider $\mathbf{c}^c$, the center distribution of the column player. We select the $(n - \sqrt{n})$ columns j of R having the highest values $\mathbf{c}^c[j]$. If row i is one of the rows with one of the $\sqrt{n}$ smallest entries for $\mathbf{c}^r$ and column j is a column with one of the $(n - \sqrt{n})$ highest entries for $\mathbf{c}^c$, then we set $R_{ij} = 1$. The payoff entries in R that are still undefined can be seen as a $(\sqrt{n} \times \sqrt{n})$-sub-matrix.

We choose this submatrix to contain a submatrix $M_{n'}$ as in Definition 1, where (for $k' = \lfloor \sqrt{n'} \rfloor$) $\binom{n'}{k'} = \sqrt{n}$. The extra columns of the submatrix have all their payoffs set to 1. The entire matrix R now satisfies the conditions of Corollary 1. We see next that any row-player strategy $\mathbf{x}$ also satisfies the conditions of Corollary 1 with a p_r value approximately 0.05. Let S be the set of non-zero rows; by construction $\sum_{i \in S} \mathbf{c}^r[i] \leqslant \frac{1}{\sqrt{n}}$. Since $d(\mathbf{x}, \mathbf{c}^r) < 0.95$, we have $\sum_{i \in S} \mathbf{x}[i] < 0.05 + \frac{1}{\sqrt{n}}$, so $\sum_{i \notin S} \mathbf{x}[i] > 0.05 - \frac{1}{\sqrt{n}}$. Similarly $\mathbf{y}$ has measure $> 0.05 - \frac{1}{\sqrt{n}}$ on columns not intersecting $M_{n'}$.

The values of p_r and p_c in Corollary 1 are $0.05 - \frac{1}{\sqrt{n}}$, and the value of p is 0.499 (column ℓ may be freely chosen by an adversary), so we get a regret of at least $0.499 - O(\frac{1}{\sqrt{n}}) + (0.05 - O(\frac{1}{\sqrt{n}}))^2 = 0.5015 - O(\frac{1}{\sqrt{n}})$. $\square$

3. One-way communication

We noted in Section 1.2.1 that $\epsilon = \frac{1}{2}$ can be achieved with one-way communication, by a simple implementation of the DMP-algorithm, using logarithmic communication. The following result gives a matching lower bound of $\frac{1}{2}$. It thus also furnishes a slightly simpler lower-bound result for the communication-free setting of the previous section, but of course the lower bound itself is necessarily weaker.

Theorem 3. *It is impossible to guarantee to find an ϵ-Nash equilibrium, for any constant $\epsilon < \frac{1}{2}$, with unlimited one-way communication.*

Proof. We consider games $G = (R, C)$, where R and C are payoff matrices with dimensions $\binom{n}{k} \times n$, with $k \approx \sqrt{n}$. Consider the following set of column player payoff matrices $C^1, \dots, C^n$, where C^ℓ has a payoff of 1 for every entry in the ℓ-th column and a 0 elsewhere:

$$\forall i, j: \quad C^\ell_{ij} = 1 \quad \text{if } j = \ell; \quad 0 \quad \text{otherwise}.$$

The row player has matrix $R = M_n$ with M_n as in Definition 1.

Let $\mathbf{x}$ be the strategy of the row player, resulting from matrix R. Let $\mathbf{y}_\ell$ be the strategy of the column player resulting from matrices R and C^ℓ; note that with unlimited one-way communication we can assume that the row player communicates all of R (and indeed, $\mathbf{x}$) to the column player.

We will show that for this class of games, one cannot do better than a $(\frac{1}{2} - o(\frac{1}{\sqrt{n}}))$-approximate Nash equilibrium.

We search for a lower bound of $\frac{1}{2} - z$, and we identify that a value of z of $\frac{1}{\sqrt{n}}$ applies.

First observe that a best response for the column player having matrix C^ℓ is $\mathbf{e}_\ell$, the pure strategy of column ℓ. Column ℓ has payoff 1 and other columns have payoff 0. So to reach a $(\frac{1}{2} - z)$-approximate Nash equilibrium, $\mathbf{y}_\ell$ must allocate a probability at least $(\frac{1}{2} + z)$ to column ℓ.

So, an arbitrary column ℓ can be required to have probability at least $\frac{1}{2} - z$. Lemma 1 says that the row player's regret is at least $\frac{1}{2} - z - O(\frac{1}{\sqrt{n}})$. Put $z = \frac{1}{\sqrt{n}}$ and we find that for all $\mathbf{x}$, ℓ may be chosen such that in order for the column player to have regret less than $\frac{1}{2} - O(\frac{1}{\sqrt{n}})$, the row player must have regret at least $\frac{1}{2} - O(\frac{1}{\sqrt{n}})$. $\square$

Theorem 4. *It is impossible to guarantee to find an ϵ-well-supported Nash equilibrium, for any constant $\epsilon < 1$, with unlimited one-way communication.*

Proof. To prove this theorem we will only have to look at 2×2 games. The row player has the identity matrix and the column player has one of two different column matrices. Communication is only allowed from the row player to the column player.

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad C^1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{or} \quad C^2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

In any ϵ-well-supported Nash equilibrium for $\epsilon < 1$, the column player must play pure strategy column j, given payoff matrix C^j. That is necessary regardless of the information she receives from the row player.

No communication is allowed from the column player to the row player, so the row player's strategy is determined by matrix R. Let $f_r(R)$ be the row player's strategy. If $f_r(R)$ allocates positive probability to row i, then we fail to have an ϵ-well-supported Nash equilibrium (for any $\epsilon < 1$) when the column player has matrix C^{3-i}, since when that happens, row i pays the row player 0 while the other row pays 1. $\square$

4. Communication-bounded algorithms

This section presents the main positive results, algorithms that compute approximate Nash equilibria that are limited to polylogarithmic communication. Section 4.1 gives the main result for ϵ-Nash equilibria, and Section 4.3 gives a variation of the algorithm that computes ϵ-well-supported Nash equilibrium for $\epsilon \approx 0.732$.

4.1. A 0.438-approximate Nash equilibrium procedure with limited communication

This section provides a 0.438-approximate Nash equilibrium procedure where the amount of communication between the players is polylogarithmic in n. We present the algorithm as an α-approximate Nash equilibrium procedure first and then optimize α. At various points the algorithm uses the operation of communicating a mixed strategy (a probability distribution over $[n]$) from one player to the other; the details of this operation are given in Section 4.2. The general idea is to communicate a sample of size $O(\log n)$ from the distribution and argue that the corresponding empirical distribution is a good enough estimate for our purposes.

First the row player finds a Nash equilibrium for the zero-sum game $(R, -R)$ and the column player computes a Nash equilibrium for the zero-sum game $(-C, C)$. Since both games are zero-sum, we know that the payoff values for their Nash equilibria will be unique. Both players compare this payoff value with α. We distinguish two cases,

1. neither player can ensure himself a payoff more than α, or
2. at least one of the players can ensure a payoff more than α.

With $O(1)$ communication, the case that holds can be identified.

Case 1: *The value of both zero-sum games is $\leqslant \alpha$ to each player*

The row player finds a strategy pair $(\mathbf{x}_r^*, \mathbf{y}_r^*)$ as solution to $(R, -R)$, while the column player finds a strategy pair $(\mathbf{x}_c^*, \mathbf{y}_c^*)$ as solution to $(-C, C)$. The row player communicates $\mathbf{y}_r^*$ to the column player (as described in Section 4.2) and the column player sends $\mathbf{x}_c^*$ to the row player. They now play the game (R, C) using strategy pair $(\mathbf{x}_c^*, \mathbf{y}_r^*)$. Since $\mathbf{y}_r^*$ is a Nash equilibrium strategy in the zero-sum game $(R, -R)$ and the row player still plays with payoff matrix R, by the minimax theorem, the row player has no strategy that can give him a payoff of α or higher. The row player has a best response with a value of at most α, so his regret is also at most α. The strategy $\mathbf{x}_c^*$ was a Nash equilibrium strategy in the zero-sum game $(-C, C)$ and the column player still has payoff matrix C. So we can use the same argument for the column player to claim that when the row player plays strategy $\mathbf{x}_c^*$, the column player has regret at most α. So, we have an α-approximate Nash equilibrium. This concludes Case 1.

Case 2: *One or both players can guarantee a payoff $> \alpha$*

If at least one of the players has a value of more than α for his zero-sum game, he can get a payoff of more than α if he plays this strategy, regardless of the strategy of the other player. Assume w.l.o.g. that it is the row player who has a payoff greater than α in his zero-sum game. He communicates this strategy $\mathbf{x}_r^*$ to the column player (again, as described in Section 4.2). The column player identifies a pure strategy best response $\mathbf{e}_j$ to $\mathbf{x}_r^*$ and communicates $\mathbf{e}_j$ to the row player (using $\log n$ bits).

At this point in the algorithm we have the strategy pair $(\mathbf{x}_r^*, \mathbf{e}_j)$. The column player has a best response strategy, so at this point his regret is 0. The row player's strategy $\mathbf{x}_r^*$ is paying him more than α. Let $\beta \leqslant 1$ be the value of his best response to $\mathbf{e}_j$. So at this point the row player has a regret of at most $\beta - \alpha$. We next deal with the possibility that $\beta - \alpha > \alpha$.

At this stage the column player has regret 0 while we are only looking for regret to be bounded by α; meanwhile the row player has a strategy that might not be good enough for an α-approximate Nash equilibrium. To change this, we use a method used in Chen et al. (2009, Lemma 3.2), which allows the row player to shift some of his probability to his best response to $\mathbf{e}_j$. By shifting some of his probability, it could be that $\mathbf{e}_j$ no longer is a best response strategy for the column player. This is acceptable, as long as the column player's regret while playing $\mathbf{e}_j$ is at most α. Suppose the row player shifts $\frac{1}{2}\alpha$ of his probability to a best response strategy. The payoff the column player gets with $\mathbf{e}_j$ could be $\frac{1}{2}\alpha$ lower because of this move. The payoff of some other column(s) could go as much as $\frac{1}{2}\alpha$ higher because of this shift. The strategy $\mathbf{e}_j$ had regret 0, so by the shift of $\frac{1}{2}\alpha$ of the row player's probability, the regret of the column player is at most $\frac{1}{2}\alpha + \frac{1}{2}\alpha = \alpha$, which constitutes an α-approximate Nash equilibrium, for the column player.

The row player is allowed to change the allocation of $\frac{1}{2}\alpha$ of his probability that was allocated to strategies having the lowest payoff. The remainder of his probability, $1 - \frac{1}{2}\alpha$, had already at least an average payoff of α. The probability is shifted to his best response with a value of β, with $\alpha \leqslant \beta \leqslant 1$. The following inequality is a sufficient condition for the row player's regret to be at most α:

$$\left(1 - \frac{1}{2}\alpha\right)\alpha + \frac{1}{2}\alpha\beta \geqslant \beta - \alpha, \quad 0 \leqslant \alpha \leqslant \beta \leqslant 1.$$

The solutions to this inequality are

$$\begin{aligned} &0 < \alpha \leqslant \frac{1}{2}(5 - \sqrt{17}), && \alpha \leqslant \beta \leqslant \frac{\alpha^2 - 4\alpha}{\alpha - 2}, \\ &\frac{1}{2}(5 - \sqrt{17}) < \alpha < 1, && \alpha \leqslant \beta \leqslant 1, \\ &\alpha = 0, \quad \beta = 0, && \alpha = 1, \quad \beta = 1 \end{aligned}$$

where it holds that if $\alpha = \frac{1}{2}(5 - \sqrt{17})$ then $f(\alpha) = \frac{\alpha^2 - 4\alpha}{\alpha - 2} = 1$ and for $0 \leqslant \alpha \leqslant 1$ this function is monotone increasing. This procedure will give an α-approximate Nash equilibrium, so α should be as low as possible. Next to this it should also hold for every β with $\alpha \leqslant \beta \leqslant 1$. The lowest α such that this condition hold is when $f(\alpha) = 1$, thus $\alpha = \frac{1}{2}(5 - \sqrt{17}) \approx 0.438$.

So if the row player rearranges $\frac{1}{2} \cdot 0.438 = 0.219$ of his probability to his best response row, both players have a strategy that guarantees them a 0.438-approximate Nash equilibrium.

4.2. Communicating mixed strategies

We describe how to communicate an approximation of the mixed strategies that are computed, using $O(\log^2 n)$ bits. We ultimately obtain an ϵ of $0.438 + \delta$, for any $\delta > 0$.

We first look at the case where one of the players, assume w.l.o.g. the row player, has a payoff higher than α in the Nash equilibrium of his zero-sum game $(R, -R)$. The column player plays a pure best response to the strategy of the row player, regardless of the support of the strategy of the row player. So we mainly consider the row player.

The zero-sum game $(R, -R)$ gives a strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$. Fix $k = \frac{\ln n}{\delta^2}$ and form a multiset A by sampling k times from the set of pure strategies of the row player, independently at random according to the distribution $\mathbf{x}^*$. Let $\mathbf{x}'$ be the mixed strategy for the row player with a probability of $\frac{1}{k}$ for every member of A. We want the distribution $\mathbf{x}'$ to have a payoff close to the payoff of $\mathbf{x}^*$. This corresponds to the following event:

$$\phi = \{((\mathbf{x}')^{\mathrm{T}} R\mathbf{y}^*) - ((\mathbf{x}^*)^{\mathrm{T}} R\mathbf{y}^*) < -\delta\}.$$

As noted in Lipton et al. (2003) the expression $((\mathbf{x}')^{\mathrm{T}} R\mathbf{y}^*)$ is essentially a sum of k independent random variables each of expected value $((\mathbf{x}^*)^{\mathrm{T}} R\mathbf{y}^*)$, where every random variable has a value between 0 and 1. This means we can bound the probability that ϕ does not hold, which we will call $\bar{\phi}$. When we apply a standard tail inequality (Hoeffding, 1963) to bound the probability of $\bar{\phi}$, we get:

$$\Pr[\bar{\phi}] \leqslant e^{-2k\delta^2}.$$

With $k = \frac{\ln n}{\delta^2}$, this gives $\Pr[\bar{\phi}] \leqslant \frac{1}{n^2}$ and $\Pr[\phi] \geqslant 1 - \frac{1}{n^2}$. If $\mathbf{x}'$ does not give payoffs close enough to $\mathbf{x}^*$, we sample again.

The strategy $\mathbf{x}'$ has a guaranteed payoff of $0.438 + \delta - \delta = 0.438$. This strategy is communicated to the column player. The support of this strategy is logarithmic and all probabilities are rational (multiples of $\frac{1}{k}$). Communication of one pure strategy has a communication complexity of $O(\log n)$. This will give a communication complexity for $\mathbf{x}'$ of $O(\log^2 n)$.

The column player computes a pure strategy best response to $\mathbf{x}'$ and communicates this strategy in $O(\log n)$ to the row player. The strategy of the row player might not yet lead to a 0.438-approximate Nash equilibrium, his payoff could be too low. As we have seen before, if the row player redistributes at most 0.219 of his probability, he is guaranteed to have a strategy that leads to a 0.438-approximate Nash equilibrium.

This change in strategy of the row player can decrease the payoff of the column player by as much as 0.219 and increase another pure strategy by as much as 0.219. His strategy was a best response, a 0-approximate Nash equilibrium, and the improvement to another pure strategy is maximal $0.219 + 0.219 = 0.438$, this leads to a 0.438-approximate Nash equilibrium.

In the alternative case, where both players have a low ($< \alpha$) payoff in their zero-sum games, the technique is essentially the same: each player samples k times from the opposing distribution, checks that it limits his own payoff to at most $\alpha + \delta$, re-samples as necessary, and communicates the k-sample.

4.3. A 0.732-well-supported Nash equilibrium procedure with limited communication

We give a variant of the algorithm of the previous section, that produces an ϵ-well-supported Nash equilibrium for $\epsilon = \sqrt{3} - 1$. Like the previous algorithm, we will first search for an α-approximate Nash equilibrium and later find the optimal value for α.

The algorithm starts in the same way as in Section 4.1 with both players computing the Nash equilibrium of zero-sum games. The row player solves the zero-sum game $(R, -R)$ and the column player solves $(-C, C)$. The two cases that arise are also the same; Case 1 proceeds as in Section 4.1 while Case 2 requires a variation to the algorithm.

Case 1: *The value of both zero-sum games is $\leqslant \alpha$ to each player*

First consider the case where both players have a Nash equilibrium with value smaller than α. The row player has a strategy pair $(\mathbf{x}_r^*, \mathbf{y}_r^*)$ and the column player a strategy pair $(\mathbf{x}_c^*, \mathbf{y}_c^*)$. The row player communicates $\mathbf{y}_r^*$ to the column player and the column player sends $\mathbf{x}_c^*$ to the row player. They will now play the game with the strategy pair $(\mathbf{x}_c^*, \mathbf{y}_r^*)$. If they play according to these strategies, then no pure strategy yields a payoff of α or more, so note that the strategy profile is an α-well-supported Nash equilibrium.

Case 2: *One or both players can guarantee a payoff $> \alpha$*

Suppose that a player, assume w.l.o.g. the row player, has a payoff more than α in the Nash equilibrium of his zero-sum game $(R, -R)$. Let the row player communicate this strategy $\mathbf{x}_r^*$ to the column player. The column player computes a pure strategy best response $\mathbf{e}_j$ to $\mathbf{x}_r^*$ and communicates this strategy to the row player. Because the row player had a payoff of at least α in the game $(R, -R)$, he also has a payoff of at least α against $\mathbf{e}_j$.

At this point in the algorithm we have a strategy pair $(\mathbf{x}_r^*, \mathbf{e}_j)$. The strategy of the column player is a best response to $\mathbf{x}_r^*$, so his strategy has regret 0. We have no guarantee on the performance of the row player's strategy, in the context of a well-supported Nash equilibrium.

As in the previous algorithm we allow the row player to shift some of his probability to his best response to $\mathbf{e}_j$. Note that if we shift $\frac{1}{2}\alpha$ of the probability of the row player, this ensures the column player's payoffs vary by at most α.

Let the best response of the row player to $\mathbf{e}_j$ have value $\beta \geqslant \alpha$. The row player's payoff is a random variable x that takes values in $[0, 1]$ with expectation $E[x] \geqslant \alpha$, since $\mathbf{x}_r^*$ is the security strategy for payoff matrix R. The maximum value x can take is β. The algorithm takes all strategies for which the row player's payoff is less than $\beta - \alpha$, and replaces any probability allocated to them by $\mathbf{x}_r^*$, to any strategy whose payoff is at least $\beta - \alpha$, thus satisfying the conditions for the row player to also have an α-well-supported Nash equilibrium.

We upper bound the probability $\Pr[x \leqslant \beta - \alpha]$ as follows. Subject to $E[x] \geqslant \alpha$ and $\max(x) = \beta$, this is maximized when x takes values β or $\beta - \alpha$. Let $p = \Pr[x \leqslant \beta - \alpha]$. Then

$$E[x] < p(\beta - \alpha) + (1 - p)\beta = -\alpha p + \beta.$$

We have $E[x] \geqslant \alpha$. Plugging that into the above,

$$\alpha \leqslant -\alpha p + \beta, \quad \text{i.e. } p \leqslant \frac{\beta - \alpha}{\alpha}.$$

To ensure that the amount of probability shifted is at most p, is suffices to let $\frac{1}{2}\alpha < \frac{\beta-\alpha}{\alpha}$, i.e. $\alpha^2 + 2\alpha - 2\beta \geqslant 0$. This is satisfied by $\alpha = -1 + \sqrt{1 + 2\beta}$, so that the worst case value of β is 1, resulting in the claimed value of $\sqrt{3} - 1 \approx 0.732$.

5. Conclusions

Our results raise some open problems, such as how good an approximation should be achievable in the communication-free setting, and how well we can do in the setting of limited (two-way) communication. Our communication-bounded algorithms are also based on algorithms that compute approximate equilibria *in polynomial time*, and it would be very interesting if further upper bounds on the communication complexity could be obtained for algorithms whose computational time was not known to be polynomial. Pastink (2012) considers some related topics, including the communication required for approximate equilibria of games of fixed size. It may be that future work should address the issue of communication protocols where the players have an incentive to report their information truthfully.

We believe that the communication-limited algorithm for 0.438-approximate Nash equilibria is significant, also the 0.501 lower bound in the communication-free setting, since in the context of searching for ϵ-approximate Nash equilibria, $\epsilon = 0.5$ frequently seems to arise as a limit on what is achievable. For example, if we search for approximate equilibria of constant support, the DMP-algorithm (Daskalakis et al., 2009b) achieves this for $\epsilon = 0.5$, however, Feder et al. (2007) show that for $\epsilon < 0.5$, the support size may need to be logarithmic in n. (The corresponding logarithmic upper bound on the support size that may be needed, is due to Lipton et al., 2003.) In a similar way, while Fictitious Play is known to guarantee to find ϵ-approximate equilibria for ϵ approaching 0.5 (Conitzer, 2009), it has also been established that $\epsilon = 0.5$ is, in the worst case, a lower bound on the approximation quality attainable (Goldberg et al., 2013). And, as we find in Theorem 3, 0.5 is also the best approximation that can be guaranteed when there is a restriction to one-way communication. Finally, Fearnley et al. (2013) show that to find ϵ-Nash equilibria with $\epsilon \geqslant \frac{1}{2}$, a strictly smaller fraction of the payoffs of the game need to be checked, than is needed for certain smaller positive values of ϵ.

Acknowledgment

The first author thanks Sergiu Hart for useful discussions during the iAGT workshop in May 2011.

References

Bosse, H., Byrka, J., Markakis, E., 2007. New algorithms for approximate Nash equilibria in bimatrix games. In: Proceedings of the 3rd WINE. In: LNCS, vol. 4858, pp. 17–29.
Chen, X., Deng, X., 2006. Settling the complexity of two-player Nash equilibrium. In: Proceedings of the 47th FOCS. IEEE, pp. 261–272.
Chen, X., Deng, X., Teng, S.H., 2009. Settling the complexity of computing two-player Nash equilibria. J. ACM 56. 14, 57 pp.
Conitzer, V., 2009. Approximation guarantees for fictitious play. In: Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing, pp. 636–643.
Conitzer, V., Sandholm, T., 2004. Communication complexity as a lower bound for learning in games. In: Proceedings of the 21st ICML, pp. 24–32.
Dantzig, G.B., 1963. Linear Programming and Extensions. Princeton Univ. Press.
Daskalakis, C., Mehta, A., Papadimitriou, C.H., 2007. Progress in approximate Nash equilibria. In: Proc. of 8th ACM-EC, pp. 355–358.
Daskalakis, C., Goldberg, P.W., Papadimitriou, C.H., 2009a. The complexity of computing a Nash equilibrium. SIAM J. Comput. 39 (1), 195–259.
Daskalakis, C., Mehta, A., Papadimitriou, C.H., 2009b. A note on approximate Nash equilibria. Theoretical Computer Sci. 410 (17), 1581–1588.
Daskalakis, C., Frongillo, R., Papadimitriou, C.H., Pierrakos, G., Valiant, G., 2010. On learning algorithms for Nash equilibria. In: Proceedings of 3rd SAGT, pp. 114–125.
Fearnley, J., Goldberg, P.W., Savani, R., Sørensen, T.B., 2012. Approximate well-supported Nash equilibria below two-thirds. In: Proceedings of the 5th International Symposium on Algorithmic Game Theory, pp. 108–119.
Fearnley, J., Gairing, M., Goldberg, P.W., Savani, R., 2013. Learning equilibria of games via payoff queries. In: Proceedings of EC'13: 14th ACM Conference on Electronic Commerce, pp. 397–414.
Feder, T., Nazerzadeh, H., Saberi, A., 2007. Approximating Nash equilibria using small-support strategies. In: Proceedings of ACM Conference on Electronic Commerce, pp. 288–294.
Foster, D., Vohra, R.V., 1997. Calibrated learning and correlated equilibrium. Games Econ. Behav. 21 (1), 40–55.
Foster, D.P., Young, H.P., 2006. Regret testing: learning to play Nash equilibrium without knowing you have an opponent. Theoretical Econ. 1, 341–367.
Goldberg, P.W., 2006. Some discriminant-based PAC algorithms. J. Mach. Learn. Res. 7, 283–306.
Goldberg, P.W., Sørensen, T.B., Savani, R., Ventre, C., 2013. On the approximation performance of fictitious play in finite games. Int. J. Game Theory 42 (4), 1059–1083. http://dx.doi.org/10.1007/s00182-012-0362-6.
Hart, S., Mansour, Y., 2010. How long to equilibrium? The communication complexity of uncoupled equilibrium procedures. Games Econ. Behav. 69 (1), 107–126.
Hart, S., Mas-Colell, A., 2000. A simple adaptive procedure leading to correlated equilibrium. Econometrica 68 (5), 1127–1150.

Hart, S., Mas-Colell, A., 2003. Uncoupled dynamics do not lead to Nash equilibrium. Amer. 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.
Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. J. Amer. Statistical Assoc. 58 (301), 13–30.
Karmarkar, N., 1984. A new polynomial-time algorithm for linear programming. In: Proceedings of the 16th STOC. ACM, pp. 302–311.
Kontogiannis, S.C., Spirakis, P.G., 2010. Well supported approximate equilibria in bimatrix games. Algorithmica 57 (4), 653–667.
Kontogiannis, S.C., Spirakis, P.G., 2011. Approximability of symmetric bimatrix games and related experiments. In: ACM Symp. on Experimental Algorithms (SEA). In: LNCS, vol. 6630, pp. 1–20.
Kontogiannis, S., Panagopoulou, P., Spirakis, P., 2006. Polynomial algorithms for approximating Nash equilibria of bimatrix games. In: ECCC. TR06-081.
Kushilevitz, E., 1997. Communication complexity. In: Advances Comput., vol. 44. Elsevier, pp. 331–360.
Lipton, R.J., Markakis, E., Mehta, A., 2003. Playing large games using simple strategies. In: Proceedings of the 4th ACM-EC, pp. 36–41.
Nash, J.F., 1951. Non-cooperative games. Ann. of Math. 54 (2), 286–295.
Neumann, J.v., 1928. Zur Theorie der Gesellschaftsspiele. Math. Ann. 100, 295–320.
Pastink, A., 2012. Aspects of communication complexity for approximating Nash equilibria. MSc dissertation. Utrecht University.
Tsaknakis, H., Spirakis, P., 2007. An optimization approach for approximate Nash equilibria. In: Proceedings of the 3rd WINE. In: LNCS, vol. 4858, pp. 42–56.
Yao, A.C.C., 1979. Some complexity questions related to distributive computing (preliminary report). In: Proceedings of the 11th STOC. ACM, pp. 209–213.