Universiteit Utrecht

UTRECHT UNIVERSITY

MSc THESIS

Aspects of communication complexity for approximating Nash equilibria

Author:
Arnoud PASTINK

Supervisors:
Prof. dr. Jan van LEEUWEN
Prof. dr. Paul W. GOLDBERG

June 28, 2012

Abstract

Since it was shown that finding a Nash equilibrium is PPAD-complete [6], even for 2-player normal form games [2], a lot of attention has been given to ϵ-approximate Nash equilibria. Almost all results on ϵ-approximate Nash equilibria assume full knowledge of the game that is played. This thesis will focus on ϵ-approximate Nash equilibria in an uncoupled setup, players only have knowledge of their own payoff matrix.

For an uncoupled setup a few lower bound results on the communication complexity are known [4] [10], but these results only apply to exact Nash equilibria.

In this thesis we will look in different ways at the communication complexity of ϵ-approximate Nash equilibria in an uncoupled setup. First we will look at small games, where each player can play only a few different actions. For these small games we derive lower- and upper bounds on the approximation in settings with no- or very limited communication. Next we show upper bounds on the communication complexity for general games and lower bounds on the communication complexity for reaching good approximations.

In the next sections we bound the communication that is allowed. For models with no communication we prove that any ϵ-approximate Nash equilibrium will have $\epsilon > 0.5$ for any algorithm, in the worst case. Results on one-way communication indicate that finding an ϵ-well-supported Nash equilibrium requires more information than finding an ϵ-approximate Nash equilibrium. In the last section we show a 0.432-approximate Nash equilibrium and a 0.732-WSNE with limited communication allowed. Next to the limited communication these algorithms also have a polynomial running time, which makes them comparable to existing polynomial-time algorithms with no bound on the communication.

Contents

Chapter 1

Introduction

The solution concept of a Nash equilibrium has already been a central notion in game theory for many years. Next to the properties of a solution concept, algorithmic game theory is also concerned with how a solution can be obtained. A solution should be relatively easy to compute. Since it was shown that finding a Nash equilibrium is PPAD-complete [6], even for 2-player normal form games [2], a lot of attention has shifted to other versions of the solution concept of Nash equilibrium to overcome the apparent difficulty of finding an exact Nash equilibrium. These concepts have replaced the requirement of no incentive to deviate with the requirement that the upper bound on the incentive to deviate is ϵ, where $\epsilon > 0$ is a parameter.

The most used concept for approximating a Nash equilibrium is ϵ-approximate Nash equilibrium. Next to this notion, the stronger notion of ϵ-well-supported Nash equilibrium (ϵ-WSNE) has been introduced. A number of results have been achieved on approximations for both notions. All these results assume full knowledge of the game that is played. The communication complexity in an uncoupled setup for exact Nash equilibria, players only have knowledge of their own payoff matrix, has been studied.

In this work the focus will be on a combination of these two subjects, the communication complexity of approximate Nash equilibria in an uncoupled setup. The analysis will be in two directions: By fixing the amount of communication allowed we will analyse what approximation is possible and by fixing the approximation we will look at the communication that is needed.

This introduction will introduce some key concepts for this thesis and an overview of the research that has already been done in this area.

1.1 Definitions

Consider *2-player games*, with a *row player* and a *column player*, who both have n *pure strategies*. The game $G(R, C)$ 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 his 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 belonging to these mixed strategies $\mathbf{x}$ and $\mathbf{y}$ are $\mathbf{x}^T R\mathbf{y}$ for the row player and $\mathbf{x}^T C\mathbf{y}$ 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 on the ith position and a 0 on every other position. A Nash equilibrium is a pair of mixed strategies $(\mathbf{x}^*, \mathbf{y}^*)$ such that :

$$\forall i = 1 \dots n: \quad \mathbf{e}_i^T R \mathbf{y}^* \leq (\mathbf{x}^*)^T R \mathbf{y}^*$$
$$\forall i = 1 \dots n: \quad (\mathbf{x}^*)^T C \mathbf{e}_i \leq (\mathbf{x}^*)^T C \mathbf{y}^*$$

The payoffs of R and C are normalised between 0 and 1. This is not a restriction on the model; if a pair of strategies is in equilibrium, then by multiplying all entries of a matrix by a constant or adding a constant to all entries, will not affect the equilibrium. So if a game does not have payoffs between 0 and 1, we can normalise the payoffs without changing the game.

An *ϵ-approximate 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 for some $\epsilon > 0$. Note that this approximation notion is additive. Additive ϵ-approximate Nash equilibrium is the dominant approximation method, but some research has been done on multiplicative approximation. The difference between the strategy of a player and his best response will also be called the regret of a player. A strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is an ϵ-approximate Nash equilibrium if the following holds:

$$\forall i = 1 \dots n: \quad \mathbf{e}_i^T R \mathbf{y}^* \leq (\mathbf{x}^*)^T R \mathbf{y}^* + \epsilon$$
$$\forall i = 1 \dots n: \quad (\mathbf{x}^*)^T C \mathbf{e}_i \leq (\mathbf{x}^*)^T C \mathbf{y}^* + \epsilon$$

Next to ϵ-approximate Nash equilibrium we will also define the notion of *well-supported ϵ-approximate Nash equilibrium*, ϵ-WSNE in short. The difference with a regular ϵ-approximate Nash equilibrium is that every pure strategy that is played with positive probability should be an approximate best-response. A strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is an ϵ-WSNE if the following holds:

$$\forall i : \mathbf{x}_i^* > 0 \Rightarrow \quad \mathbf{e}_i^T R \mathbf{y}^* \geq \mathbf{e}_j^T R \mathbf{y}^* - \epsilon \qquad \forall j = 1 \dots n$$
$$\forall i : \mathbf{y}_i^* > 0 \Rightarrow \quad (\mathbf{x}^*)^T C \mathbf{e}_i \geq (\mathbf{x}^*)^T C \mathbf{e}_j - \epsilon \quad \forall j = 1 \dots n$$

The notion of ϵ-WSNE is stronger than the notion of ϵ-approximate Nash equilibrium, every well supported ϵ-approximate Nash equilibrium is also an ϵ-approximate Nash equilibrium, but the converse does not have to be true. The weakness of ϵ-approximate Nash equilibrium is that it does not disallow the player to play pure strategies with positive probability that give him a very low payoff. This can be seen as irrational, playing an action that you know is not good for you. With an ϵ-WSNE this behaviour is not possible, every action that is played with positive probability must have a payoff that is within ϵ of the best response strategy.

The *support* of a mixed strategy $\mathbf{x}$, denoted by $\text{Supp}(\mathbf{x})$, is the set of pure strategies that are played with non-zero probability by $\mathbf{x}$.

The basic assumption we make is that of *uncoupledness*, the players only have knowledge of their own payoff matrix. The players are completely *cooperative*, we can ignore strategic and selfish behaviour by the players. We also assume that both players are *completely rational*, they make optimal choices regarding the objective.

We will analyse the *communication complexity* of ϵ-approximate Nash equilibrium. We will discuss the definition of communication complexity in the next chapter. We are only interested in the communication complexity of ϵ-approximate Nash equilibrium when each player knows his own payoff matrix. This means that we do not look at communication before the payoffs are known. So players can make arrangements about their strategies before the payoffs are known.

1.2 Existing results on communication complexity of Nash equilibria

There are a few results concerning the communication complexity of Nash equilibria. In [4] it is shown that a lower bound on the communication complexity for 2-player games of finding a pure Nash equilibrium is $\Omega(n^2)$, 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 method relies on finding if a pure Nash equilibrium exists, where the existence of a mixed Nash equilibrium is trivial[15].

In [10] the communication complexity of uncoupled equilibrium procedures is studied. They show that for reaching a pure Nash equilibrium, reaching a pure Nash equilibrium in a Bayesian setting and for reaching a mixed Nash equilibrium, the lower bound on the communication complexity is $\Omega(2^s)$, where s is the number of players. To show that reaching these equilibria is not just due to the complexity of the input, they also show that you can reach a correlated equilibrium in a polynomial number of steps. The methods they use cannot be extended to analysing the communication complexity of ϵ-approximate Nash equilibria. For pure Nash equilibria, their analysis is based on games that do not have a Nash equilibrium and for mixed strategy Nash equilibrium the analysis is based on equilibria that require a large description. Approximate Nash equilibria always exist and can have small descriptions, so the developed techniques do not work for ϵ-approximate Nash equilibria.

These previous results on communication complexity analyse the amount of communication needed to reach a certain state. This is not the problem that is solved generally in this thesis. For most of the proofs the amount of communication that is allowed is restricted and the analysis is on how well a Nash equilibrium can be approximated with the bound on the communication.

1.3 Support size of Nash equilibrium

Very closely related to the study of communication complexity of ϵ-approximate Nash equilibria is the research that has been done on the size of the support of strategies that are used in a Nash equilibrium, where the support is defined as the number of pure strategies that are played with positive probability. This research is closely related because in most algorithms the communication between the players will be about their strategy. In other words, if we can limit the support of the strategies that are used, we can also limit the communication that is needed to communicate this strategy.

It is shown in [14] that for every $\epsilon > 0$ there exists an ϵ-approximate Nash equilibrium with support logarithmic in the number of pure strategies. The proof works by showing that if we have a strategy pair $(\mathbf{x}, \mathbf{y})$ that is a Nash equilibrium of a game G, there exists a strategy pair $(\mathbf{x}', \mathbf{y}')$ such that $\mathbf{x}'$ and $\mathbf{y}'$ have logarithmic support and the payoff of both players is very close to the payoff the Nash equilibrium. This result extends to ϵ-WSNE [13]. The proof given by Lipton et al. [14] extends to a simple subexponential algorithm that finds an ϵ-approximate Nash equilibrium.

Next to this result, it is shown in [9] that to reach an ϵ-approximate Nash equilibrium with $\epsilon < 0.5$, the support of the strategies have to be at least of logarithmic size in the number of pure strategies. This result nicely shows a kind of optimality for the DMP-algorithm, which uses strategies with a support of at most 2 and reaches a 0.5-approximate Nash equilibrium.

There are no results of this kind for ϵ-WSNE.

1.4 Zero-sum Games

A special class of games are zero-sum games. A game is a zero-sum game if the sum of the payoffs of the players is zero for any choice of strategies. For two-player bimatrix games this means that for every entry of the payoff matrix, the sum of the entry of the row player and the column player is zero. In general a zero-sum game G will look like $G = (A, -A)$.

Zero-sum games have a few interesting properties with respect to Nash equilibria. Firstly, all Nash equilibria of a zero-sum game have the same payoff value.

Zero-sum games are different from general games in terms of computation because a Nash equilibrium of a zero-sum game can be computed in polynomial time. This is possible because of the connection of zero-sum games with linear programming [17] [5] [12]. This property makes this class of games interesting for finding Nash equilibria for general games. If you can find a zero-sum game that is related to the real game, the Nash equilibrium that you find in the zero-sum game could give an ϵ-approximate Nash equilibrium for the general game.

We will now look at a property of a zero-sum game when we only consider one player, say the row player with payoff matrix R. If he plays a zero-sum game with his payoff matrix, so $G = (R, -R)$, the other player (with $-R$) has a 'worst case' payoff matrix for the row player, every action that yields a high payoff for the row player has a very low payoff for the column player and every good action for the column player is bad for the row player. For the column player, the Nash equilibrium property is by definition:

$$\forall \mathbf{y}' \quad \mathbf{x}^*(-R)\mathbf{y}^* \geq \mathbf{x}^*(-R)\mathbf{y}'$$

where $\mathbf{y}'$ is a strategy. This leads to the following inequality:

$$\forall \mathbf{y}' \quad \mathbf{x}^* R \mathbf{y}^* \leq \mathbf{x}^* R \mathbf{y}'$$

So if the column player changes his strategy from $\mathbf{y}^*$ to some other strategy, the payoff of the row player will be at least the payoff of the Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$. As a consequence, the row player can play $\mathbf{x}^*$ as a safety strategy. The row player can compute $\mathbf{x}^*$, because it comes from his own zero-sum game, and it will give him a guarantee on his payoff.

The computation of a Nash equilibrium for a zero-sum game in an uncoupled setup requires no communication at all. Both players can compute their safety strategy and use this to play the game. Because all Nash equilibria of a zero-sum game have the same payoff, both players have no incentive to deviate.

1.5 Existing algorithms

In recent years a number of algorithms have been developed that guarantee an ϵ-approximate Nash equilibrium for a certain value of ϵ. This overview is by no means a complete overview of all existing algorithms. It merely gives an indication of the different methods that are used to solve this problem.

Apart from the DMP-algorithm, all algorithms for finding approximate Nash equilibria need full knowledge of both payoff matrices. This means that these algorithms are of no use in an uncoupled setup.

In one of the sections we will analyse the communication complexity of communicating complete payoff matrices. When a whole payoff matrix is communicated, these algorithms that need full knowledge can of course be used to get a good approximation of a Nash equilibrium.

DMP-algorithm
The easiest algorithm that achieves a 0.5-approximate Nash equilibrium is the DMP-algorithm [7]. The algorithm picks an arbitrary row for the row player, say row i. Let $j \in \arg\max_{j'} C_{ij'}$ and $k \in \arg\max_{k'} R_{k'j}$. 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.

Of all existing algorithms, the DMP-algorithm is the only algorithm that can be modelled as an algorithm with limited communication in an uncoupled setup. Adapt the algorithm such that instead of the row player picking an arbitrary row i, he always picks row 1. The column player will play a best response to row 1. The only communication that is needed is the column player communicating his pure strategy best response to the row player. The column player has n columns, so communicating one column number will have a communication complexity of $O(\log n)$. This gives us an upper bound of $\frac{1}{2}$-approximate Nash equilibrium with $O(\log n)$ communication and for models with one-way communication.

Other existing algorithms
The algorithm presented in [1] can be seen as a modification of the DMP-algorithm and achieves a 0.38197-approximate Nash equilibrium. Instead of a player playing an arbitrary strategy with some positive probability, the row player gives some probability to the strategy $\mathbf{x}$ belonging to the Nash equilibrium of the zero-sum game $(R-C, C-R)$. Note that this is algorithm will not work in an uncoupled setup, knowledge of both payoff matrices is required. If the strategy pair $(\mathbf{x}, \mathbf{y})$ belonging to the Nash equilibrium of $(R-C, C-R)$ gives a 0.38197-approximate Nash equilibrium, this strategy is played. Else the column player will play a best response $\mathbf{e}_j$ to $\mathbf{x}$ and the row player will play a mixture of $\mathbf{x}$ and $\mathbf{e}_i$, where $\mathbf{e}_i$ is a best response to the strategy $\mathbf{e}_j$ of the column player. In the remainder of the paper the worst-case performance is improved to a 0.36395-approximate Nash equilibrium.

The algorithm with the best approximation that is currently known, gives a 0.3393-approximate Nash equilibrium[16]. The method used is quite different from the other approximation algorithms, it relies on finding local minima. Let $f(\mathbf{x}, \mathbf{y})$ denote the maximum regret of either player. The algorithm starts at some random strategy profile $(\mathbf{x}, \mathbf{y})$ and finds a local minimum of $f(\mathbf{x}, \mathbf{y})$. This local minimum can be found in polynomial time and the value of f in this local minimum will be at most 0.3393 higher than in the global minimum. Where the global minimum is obviously a Nash equilibrium with a regret of zero.

A well-supported algorithm
Untill now only one algorithm is known [13] that gives a ϵ-WSNE for $\epsilon < 1$. The difficulty with ϵ-WSNE is that every pure strategy that is played with positive probability should be an approximate best response. The algorithm works in two stages. First it checks whether there is a combination of i, j such that $R_{ij} + C_{ij} \geq 2 - \alpha$, where α is a parameter. If this combination exists, $R_{ij} \geq 1 - \alpha$ and $C_{ij} \geq 1 - \alpha$ and the strategy $(\mathbf{e}_i, \mathbf{e}_j)$ gives a α-WSNE.

Otherwise a zero-sum game $(D, -D)$ is constructed with $D = R + X$ and $-D = C + X$. This solves to $D = 0.5(R - C)$ and $X = -0.5(R + C)$. Let $(\mathbf{x}, \mathbf{y})$ be a Nash equilibrium of $(D, -D)$. The idea is that deviating from this strategy in the game (R, C) can only give a limited improvement, because we know that every entry of X lies in the range $[-0.5(2 - \alpha), 0]$. To be more precise, if we let D_i be the ith row of D and suppose $\mathbf{x}_i > 0$, so i is in the support of $\mathbf{x}$:

$$\begin{aligned}
\forall j \quad & D_i y \geq D_j y \\
\forall j \quad & (R+X)_i y \geq (R+X)_j y \\
\forall j \quad & R_i y \geq R_j - (X_i - X_j) y
\end{aligned}$$

Because it was first checked if there was a combination of i, j such that $R_{ij} + C_{ij} \geq 2 - \alpha$, the range of $(X_i - X_j)$ is limited to $(X_i - X_j) \in [-0.5(2 - \alpha), 0.5(2 - \alpha)]$. By minimising α, this leads to a $\frac{2}{3}$-WSNE.

1.6 Overview

The remainder of this thesis is divided into six sections.

Chapter 2 gives an introduction to communication complexity.

Chapter 3 will be about small games where each player has only two or three different actions to choose from. This restriction on the number of actions leads to some natural lower bounds in models with no commmuncation. Next to this some algorithms are presented that reach an upper bound that is close to the proven lower bound. If we allow only a little bit of information, the approximations can improve further.

In chapter 4 we analyse the communication complexity when a whole payoff matrix is communicated. This gives an upper bound on the amount of communication for any model.

Chapter 5 analyses the communication complexity from a different perspective. When we look at what information can be useful to communicate, the two logical options are communicating a part of your payoff matrix or communicating a part of your strategy. This section shows that in both cases you need a considerable amount of communication to achieve good results.

Chapter 6 shows lower bounds on the approximation of Nash equilibrium when the allowed amount of communication is restricted. We first show that when no communication is allowed, the regret of a player is strictly larger than 0.5. Next to this bound, lower bounds are proven when the communication is restricted to one-way communication. These lower bounds are surprisingly high and show that an interaction between the players is needed to get good approximations.

In chapter 7 some algorithms are discussed that give an upper bound on the approximation of a Nash equilibrium. First a fairly straightforward adaption of the DMP-algorithm is presented for models where no communication is allowed. The other two algorithms need some communication, the communcation of a mixed strategy is needed. These algorithms cleverly use some properties of zero-sum games to guarantee approximations to both ϵ-approximate Nash equilibria and ϵ-WSNE that are close to the best results that are known.

Chapter 2

Communication Complexity

The "classical" setting of communication complexity is based on the model introduced by Yao in [18]. We will follow the representation in [8]. We have two agents[1], one holding the input $\mathbf{x} \in \{0,1\}^n$ and the other holding the 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 will 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 $(\mathbf{x},\mathbf{y}) \in \{0,1\}^n \times \{0,1\}^n$ the protocol 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 $\mathcal{P}$. We will denote this number by $CC(\mathcal{P},f,\mathbf{x},\mathbf{y})$. The communication complexity of a protocol $\mathcal{P}$ for a function f is defined as the worst case communication complexity over all possible inputs for $(\mathbf{x},\mathbf{y}) \in \{0,1\}^n \times \{0,1\}^n$, we will denote this with $CC(\mathcal{P},f)$:

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

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

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

2.1 ϵ-approximate Nash equilibrium procedures

The questions we can answer with the function f as defined are decision problems, only two outcomes are possible. We will be looking at the communication complexity of ϵ-approximate Nash equilibrium. The existence question of ϵ-approximate Nash equilibria is trivial, every game has ϵ-approximate Nash equilibria. So we cannot use the standard model to find the communication complexity of ϵ-approximate Nash equilibrium, because it relies on decision problems. To make the communication complexity suitable for our problems, we will loosely follow the model introduced in [10].

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

An ϵ-approximate Nash equilibrium procedure outputs an ϵ-approximate Nash equilibrium, for a given ϵ, for the game they are playing. The output of the procedure will be reached by communication between the agents. The input of an ϵ-approximate Nash equilibrium procedure $\mathcal{P}$ is a game G in the family of games $\mathcal{G}$ and a $\epsilon < 1$. On termination of $\mathcal{P}$, each player outputs his own strategy such that this set of strategies together is an ϵ-approximate Nash equilibrium. If the execution of $\mathcal{P}$ has not terminated, the protocol specifies what message should be send next, based on the input G and the communication so far.

The communication complexity of an ϵ-approximate Nash equilibrium procedure $\mathcal{P}$ applied to a game G with a given ϵ is the number of bits communicated before $\mathcal{P}$ terminates when the input is (G, ϵ), which we will denote as $CC(\mathcal{P}, G, \epsilon)$. The communication complexity of a procedure $\mathcal{P}$ applied to a family of games $\mathcal{G}$ and a fixed ϵ, denoted as $CC(\mathcal{P}, \mathcal{G}, \epsilon)$ is the worst case communication complexity over all games in that family:

$$CC(\mathcal{P}, \mathcal{G}, \epsilon) = \max_{G \in \mathcal{G}} CC(\mathcal{P}, G, \epsilon)$$

We can use this notion of communication complexity to compute the communication complexity of different families of games.

2.2 Size of messages

Until now we haven't specified the maximum size of a message. A common approach [10] is to let the size of the message be at most logarithmic in the number of actions that can be played, so with n actions, the maximum message size is $\log n$. This corresponds usually to the model where the communication takes place by playing the game, the communication of an action has a communication complexity of $O(\log n)$. We can also limit the size of a message to a constant number of bits. This corresponds to a model where the players communicate over some private channel and only play the game after they have agreed on a strategy. We will limit the size of a message to 1 bit for this model. If a single message is larger than 1 bit, it can be send in multiple consecutive rounds. Unless specified, we will use the model where the size of a message is limited to 1 bit.

2.3 Asymmetric and symmetric rounds of communication

A round of communication can be sent in two ways, the messages of both agents can be send simultaneously (symmetric), or one agent can first send his bit of information to the other player and after the other player received this bit, he responses with his bit of information (asymmetric). The asymmetric communication is clearly stronger than the symmetric communication, the other player can use the bit of information he received in the message he sends. We can bound the number of rounds of the symmetric communication by two times the number of rounds of asymmetric communication.

Theorem 1. *For any procedure $\mathcal{P}$ the number of symmetric communication rounds can be bounded by: number of symmetric rounds $\leq 2 \times$ number of asymmetric rounds*

Proof:
A round of communication in an asymmetric model consists of two parts: agent 1 communicating his bit and the response of agent 2. Agent 2 can use the information of agent 1 within 1 round.

We can model the asymmetric communication with $ac_{\mathcal{P}} = (m_{11}, m_{12}, m_{21}, m_{22}, \ldots, m_{t1}, m_{t2})$, where m_{i1} denotes the message of agent 1 in the ith round and m_{i2} the message of agent 2 in the ith round. In a symmetric model agent 2 cannot use the information of the message of agent 2 that was sent in the same round. We can model this with $sc_{\mathcal{P}} = (r_1, r_2, \ldots, r_t)$ where r_i denotes the ith round of communication. We can map the messages of the asymmetric model to the symmetric model as follows:

$$\forall i; 1 \leq i \leq t : r_{2i-1} = m_{iA}$$
$$\forall i; 1 \leq i \leq t : r_{2i} = m_{iB}$$

This mapping ensures that each message in both models will have the same information prior to sending the message. With this mapping the number of rounds in the symmetric model is clearly twice as large as the asymmetric model, this proves the theorem.

2.4 Useful bounds on the communication complexity

Knowledge of all payoff matrices is sufficient to compute a Nash equilibrium of a game. A payoff matrix has dimensions $n \times n$, so n^2 entries. The communication complexity of communicating a payoff matrix is therefore $O(n^2)$. This means that any algorithm will not have a communication complexity that is higher than $O(n^2)$, with $O(n^2)$ all the knowledge that is necessary can be available.

The smallest bound on the communication complexity that could be useful is obviously no communication at all. The next bound that could be interesting is $O(\log n)$. Since a player has n pure strategies, indicating a pure strategy has a communication complexity of $O(\log n)$. By communicating a strategy the other player can anticipate on what the other player does and this can obviously improve an approximation.

A mixed strategy is a $1 \times n$ vector, which can be represented by n numbers. Therefore it seems that the next useful bound on the communication complexity is $O(n)$. This bound would allow the communcation of a full strategy profile. However it was shown [14] that for every Nash equilibrium, there exists an ϵ-approximate Nash equilibrium for any $\epsilon > 0$ with support size logarithmic in the number of pure strategies. So every strategy profile can be approximated with a strategy profile that uses only a logarithmic number of pure strategies. This observation makes $O(\log^2 n)$ a useful bound on the communication complexity, it would allow that complete strategy profiles are communicated. A bound of $O(n)$ could also be used to communicate all the payoffs of a pure strategy, but it seems unlikely that this is useful.

Chapter 3

Small Games

For games with only a few actions, good bounds on the approximation of a Nash equilibrium can be achieved, even with very little communication. This chapter will first prove some lower bounds for small games with no communication. Then we will show various upper bounds for 2×2-games, first in a setting where no communication is allowed and then in a setting where a few bits of communication are allowed.

3.1 Lower bound on ϵ for 2×2-games with no communication

The set of possible payoff matrices for the row player and the column player is given by $\mathcal{M}$. The row player has a function f_r and the column player a function f_c that maps a matrix of $\mathcal{M}$ to a mixed strategy:

$$\begin{aligned} f_r :& \mathcal{M} \to \Delta_{n_1} \\ f_c :& \mathcal{M} \to \Delta_{n_2} \end{aligned}$$

For 2×2-games, we can simplify this to $f_r : \mathcal{M} \to \mathbb{R}$ for the row player and $f_c : \mathcal{M} \to \mathbb{R}$ for the column player, where f_r is the probability given to row 1 and f_c the probability given to column 1. The general form of the payoff matrices is:

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

Suppose the payoff matrix of the row player is one of the following:

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

The payoff matrix of the column player is given by:

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

For this set of payoff matrices, the players cannot improve over a $\frac{1}{4}$-approximate Nash equilibrium. Suppose the row player has matrix R^1. This means that to achieve a $\frac{1}{4}$-approximate Nash equilibrium $f_r(R^1) \geq \frac{3}{4}$. Similarly if the row player has matrix R^2, $f_r(R^2) \leq \frac{1}{4}$. The row player can either put $f_c(C) \leq \frac{1}{2}$ or $f_c(C) \geq \frac{1}{2}$.

Theorem 2. *For 2×2-games with no communication the players cannot improve over a $\frac{1}{4}$-approximate Nash equilibrium.*

Proof:
Case 1: $f_c(C) \leq \frac{1}{2}$
Suppose the row player has payoff matrix R^1, so $f_r(R^1) \geq \frac{3}{4}$. The payoff of the row player is equal to:

$$f_r(R^1) \cdot C_{11} \cdot f_c(C) + (1 - f_r(R^1)) \cdot C_{22} \cdot (1 - f_c(C)) = f_r(R^1) \cdot f_c(C) + (1 - f_r(R^1)) \cdot (1 - f_c(C))$$

Because $f_r(R^1) > (1 - f_r(R^1))$, the column player will get a higher payoff if he gives a higher probability to column 1. Since $f_c(C) \leq \frac{1}{2}$, the heighest payoff is reached when $f_c(C) = \frac{1}{2}$. The payoff will be:

$$f_r(R^1) \cdot \frac{1}{2} + (1 - f_r(R^1)) \cdot \frac{1}{2} = \frac{1}{2}$$

So the highest payoff that can be achieved is $\frac{1}{2}$, while the optimal strategy is when $f_c(C) = 1$, the payoff would have been $f_r(R^1) \cdot 1 = f_r(R^1) \geq \frac{3}{4}$. So the incentive to deviate is larger than or equal to: $f_r(R^1) - \frac{1}{2} \geq \frac{3}{4} - \frac{1}{2} \geq \frac{1}{4}$.

Case 2: $f_c(C) \geq \frac{1}{2}$
Suppose the row player has payoff matrix R^2, so $f_r(R^2) \leq \frac{1}{4}$. The payoff of the row player is equal to:

$$f_r(R^2) \cdot C_{11} \cdot f_c(C) + (1 - f_r(R^2)) \cdot C_{22} \cdot (1 - f_c(C)) = f_r(R^2) \cdot f_c(C) + (1 - f_r(R^2)) \cdot (1 - f_c(C))$$

Because $(1 - f_r(R^2)) > f_r(R^2)$, the column player will get a higher payoff if he gives a higher probability to column 2. Since $f_c(C) \geq \frac{1}{2}$, the heighest payoff is reached when $f_c(C) = \frac{1}{2}$. The payoff will be:

$$f_r(R^2) \cdot \frac{1}{2} + (1 - f_r(R^2)) \cdot \frac{1}{2} = \frac{1}{2}$$

So the highest payoff that can be achieved is $\frac{1}{2}$, while the optimal strategy is when $f_c(C) = 0$, the payoff would have been $f_r(R^2) \cdot 1 = f_r(R^2) \geq \frac{3}{4}$. So the incentive to deviate is larger or equal to: $f_r(R^2) - \frac{1}{2} \geq \frac{3}{4} - \frac{1}{2} \geq \frac{1}{4}$. This concludes the proof. □

3.2 Lower bound on ϵ for 3×3-games with no communication

We can extend our approach for 2×2-games to a lower bound of $\epsilon = \frac{1}{3}$ for 3×3-games. Consider a game G where the row player has one of the following payoff matrices:

$$R^1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad R^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad R^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

The payoff matrix of the column player is given by:

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Let $f_c(C) = (c_1, c_2, c_3)$ be the probability distribution of the column player and $f_r(R^k) = (r_1^k, r_2^k, r_3^k)$ the probability distribution of the row player.

Theorem 3. *For 3×3-games with no communication the players cannot improve over a $\frac{1}{3}$-approximate Nash equilibrium.*

Proof:
It is easy to see that if the row player has matrix R^k, $r_k^k \geq \frac{2}{3}$. Assume $c_1 \leq c_2, c_3$ and consider the row player has matrix R^1, so $r_1^1 \geq \frac{2}{3}$. The obvious best response for the column player would be to play the pure strategy c_1, which would give a payoff of r_1^1. The incentive to deviate will be:

$$\begin{aligned} r_1^1 - (r_1^1 c_1 + r_2^1 c_2 + r_3^1 c_3) =& r_1^1(1 - c_1) - c_2 r_2^1 - r_3^1 c_3 \\ \geq& \frac{2}{3}(1 - c_1) - c_2 r_2^1 - c_3 r_3^1 \end{aligned}$$

The inequality holds because $r_1^1 \geq \frac{2}{3}$. Because c_2 and c_3 cannot be distinguished we can assume $c_2 \geq c_3$, this means that the inequality above is minimised if r_2^1 is as large as possible, so $r_2^1 = \frac{1}{3}$

$$\begin{aligned} \frac{2}{3}(1 - c_1) - \frac{1}{3}c_2 =& \frac{2}{3}(1 - (\frac{1}{3} - \epsilon)) - \frac{1}{3}(\frac{1}{3} + \epsilon) \\ =& \frac{4}{9} + \frac{2}{3}\epsilon - \frac{1}{9} - \frac{1}{3}\epsilon \\ =& \frac{1}{3} + \frac{1}{3}\epsilon \geq \frac{1}{3} \end{aligned}$$

The smallest probability of the column player is given to c_1, this can be as low as $\frac{1}{3} - \epsilon$ with $0 \leq \epsilon \leq \frac{1}{3}$. The biggest probability is given to c_2, so that can be as big as $\frac{1}{3} + \epsilon$. So the incentive to deviate is at least $\frac{1}{3}$. A similar analysis can be given when the row player has R^2 or R^3. This concludes the proof. □

3.3 Upper bound on ϵ for 2×2-games with no communication

Even with no communication at all we can bound the approximation to a Nash equilibrium.

It is easy to see that you can achieve a $\frac{1}{2}$-approximate Nash equilibrium on 2×2-games when you play a fully mixed strategy, so both pure strategies with probability $\frac{1}{2}$. We can achieve a better bound if we assign a higher probability to the pure strategy that seems more promising.

Lemma 1. *For 2×2-games with no communication, the players can achieve a 0.38197-approximate Nash equilibrium.*

We will assign a probability $0 \leq p \leq \frac{1}{2}$ to each of the pure strategies and the remainder of the probability $(1 - 2p)$ to the strategy that has the highest expected payoff against a fully mixed strategy. In case of a tie in expected payoff, we randomly assign the remaining probability to one of the two pure strategies. So one action has probability p and the other $p + (1 - 2p) = 1 - p$, where $0 \leq p \leq \frac{1}{2}$, so $1 - p \geq p$.

The row player has matrix R and the column player has matrix C. Assume w.l.o.g. $f_r(R) = (1 - p)$ and $f_c(C) = (1 - p)$. We will now look at the incentive to deviate for the row player:

The row player could deviate to row 1 or to row 2. A deviation to row 1 would mean that row 1 has a higher payoff than row 2, which is also what he "expected". If he has an incentive to deviate to row 2, his choice of allocating a higher probability to row 1 was not a good choice.

Deviation to row 1: The row player played this action with probability $1-p$, so his incentive to deviate is smaller than or equal to:

$$1 - (1 - p) = p$$

Deviation to row 2: This action was played with probability p. The row player played row 1 with a higher probability because his expected payoff was at least as high as for row 2. So row 1 is also partial a best response. The probability of this best response is at least $(p-1)p$, because you play row 1 with probablity $1-p$ and the column player plays each action with at least probability p. So the incentive to deviate for the row player is less or equal to:

$$1 - (p + (1 - p)p)$$

The incentive to deviate is smaller or equal to: $\max\{p, 1-(p+(1-p)p)\}$. We want to find the minimum value of p. This is the case when both terms are equal.

$$\begin{aligned} p &= 1 - (p + (1 - p)p) \\ p^2 - 3p + 1 &= 0 \\ p &= \frac{3-\sqrt{5}}{2} \vee p = \frac{3+\sqrt{5}}{2} \end{aligned}$$

Because $0 \le p \le \frac{1}{2}$, $p = \frac{3-\sqrt{5}}{2} \approx 0.38197$. So we have a 0.38197-approximate Nash equilibrium for 2×2-games with no communication. □

3.3.1 A better bound on 2×2-games

A better approximation is possible when we allow each player to have three different probability distributions instead of two.

Theorem 4. *For 2×2-games with no communication allowed, the players can achieve a 0.271-approximate Nash equilibrium.*

The probability distribution that is chosen only depends on the payoff matrix of the player. The payoff matrices are:

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

We will have the following distributions, with $p \ge \frac{1}{2}$ and $\alpha \ge 0$ parameters.

- $f_r(R) = p$ if $R_{11} + R_{12} \ge R_{21} + R_{22} + \alpha$
- $f_r(R) = 1 - p$ if $R_{11} + R_{12} + \alpha \le R_{21} + R_{22}$
- $f_r(R) = \frac{1}{2}$ otherwise

The row player and column player both have 3 different probability distributions, this gives us: $2 \cdot 3 \cdot 3 = 18$ different cases. Because of symmetry we can eliminate a lot of these cases. This leaves us with the following four cases:

$$\text{case } R^1: \begin{matrix} & 1-p & p \\ \begin{matrix} p \\ 1-p \end{matrix} & \multicolumn{2}{c}{\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}} \end{matrix} \quad \text{case } R^2: \begin{matrix} & \frac{1}{2} & \frac{1}{2} \\ \begin{matrix} p \\ 1-p \end{matrix} & \multicolumn{2}{c}{\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}} \end{matrix}$$

$$\text{case } R^3: \begin{array}{c} \\ \frac{1}{2} \\ \frac{1}{2} \end{array} \begin{array}{c} \begin{array}{cc} 1-p & p \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \end{array} \quad \text{case } R^4: \begin{array}{c} \\ \frac{1}{2} \\ \frac{1}{2} \end{array} \begin{array}{c} \begin{array}{cc} \frac{1}{2} & \frac{1}{2} \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \end{array}$$

Case R^1
The payoff for the row player is equal to:

$$p(1-p)R_{11} + p^2R_{12} + (1-p)^2R_{21} + p(1-p)R_{22}$$

If he would deviate to row 1, his payoff would be $(1-p)R_{11} + pR_{12}$. His incentive to deviate is:

$$\begin{aligned} &((1-p)R_{11} + pR_{12}) - (p(1-p)R_{11} + p^2R_{12} + (1-p)^2R_{21} + p(1-p)R_{22}) \\ =&((1-p)^2R_{11} + p(1-p)R_{12}) - ((1-p)^2R_{21} + p(1-p)R_{22}) \\ \le&(1-p)^2R_{11} + p(1-p)R_{12} \\ \le&(1-p)^2 + p(1-p) \\ =&(1-p) \end{aligned}$$

The last inequality holds because all payoffs are in $[0, 1]$. This property will be used in all other cases.

His incentive to deviate to row 2 is:

$$\begin{aligned} &((1-p)R_{21} + pR_{22}) - (p(1-p)R_{11} + p^2R_{12} + (1-p)^2R_{21} + p(1-p)R_{22}) \\ =&(p(1-p)R_{21} + p^2R_{22}) - (p(1-p)R_{11} + p^2R_{12}) \\ \le&(p(1-p)R_{21} + p^2R_{22}) - p(1-p)(R_{11} + R_{12}) \\ \le&(p(1-p)R_{21} + p^2R_{22}) - p(1-p)(R_{21} + R_{22} + \alpha) \\ =&(2p^2 - p)R_{22} - \alpha p(1-p) \\ \le&(2p^2 - p) - \alpha p(1-p) \end{aligned}$$

The second inequality holds because $R_{11} + R_{12} \ge R_{21} + R_{22} + \alpha$.

Case R^2
The payoff of the row player is equal to:

$$\frac{1}{2}pR_{11} + \frac{1}{2}pR_{12} + (\frac{1}{2} - \frac{1}{2}p)R_{21} + (\frac{1}{2} - \frac{1}{2}p)R_{22}$$

If he would deviate to row 1, his payoff would be $\frac{1}{2}R_{11} + \frac{1}{2}R_{12}$. His incentive to deviate is:

$$\begin{aligned} &(\frac{1}{2}R_{11} + \frac{1}{2}R_{12}) - (\frac{1}{2}pR_{11} + \frac{1}{2}pR_{12} + (\frac{1}{2} - \frac{1}{2}p)R_{21} + (\frac{1}{2} - \frac{1}{2}p)R_{22}) \\ =&((\frac{1}{2} - \frac{1}{2}p)R_{11} + (\frac{1}{2} - \frac{1}{2}p)R_{12}) - ((\frac{1}{2} - \frac{1}{2}p)R_{21} + (\frac{1}{2} - \frac{1}{2}p)R_{22})) \\ \le&(\frac{1}{2} - \frac{1}{2}p)R_{11} + (\frac{1}{2} - \frac{1}{2}p)R_{12} \\ \le&(\frac{1}{2} - \frac{1}{2}p) + (\frac{1}{2} - \frac{1}{2}p) \\ =&1 - p \end{aligned}$$

If he would deviate to row 2 his payoff would be $\frac{1}{2}R_{21} + \frac{1}{2}R_{22}$. His incentive to deviate is:

$$\begin{aligned}
&(\frac{1}{2}R_{21} + \frac{1}{2}R_{22}) - (\frac{1}{2}pR_{11} + \frac{1}{2}pR_{12} + (\frac{1}{2} - \frac{1}{2}p)R_{21} + (\frac{1}{2} - \frac{1}{2}p)R_{22}) \\
=&(\frac{1}{2}pR_{21} + \frac{1}{2}pR_{22}) - (\frac{1}{2}pR_{11} + \frac{1}{2}pR_{12}) \\
=&(\frac{1}{2}pR_{21} + \frac{1}{2}pR_{22}) - \frac{1}{2}p(R_{11} + R_{12}) \\
\leq&(\frac{1}{2}p(R_{21} + R_{22})) - \frac{1}{2}p(R_{21} + R_{22}) \\
=&0
\end{aligned}$$

Where the inequality holds because $R_{11} + R_{12} \geq R_{21} + R_{22}$

Case R^3
The payoff of the row player is equal to:

$$(\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} + (\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22}$$

Deviation to row 2 is symmetric to deviation to row 1, so we will only analyse row 1. If the row player deviates to row 1, his payoff would be $(1 - p)R_{11} + pR_{12}$. This gives an incentive to deviate of:

$$\begin{aligned}
&((1 - p)R_{11} + pR_{12}) - ((\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} + (\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22}) \\
=&(\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} - ((\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22}) \\
\leq&(\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} - (\frac{1}{2} - \frac{1}{2}p)(R_{21} + R_{22}) \\
\leq&(\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} - (\frac{1}{2} - \frac{1}{2}p)(R_{11} + R_{12} - \alpha) \\
=&(\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} - (\frac{1}{2} - \frac{1}{2}p)R_{11} - (\frac{1}{2} - \frac{1}{2}p)R_{12} + \alpha(\frac{1}{2} - \frac{1}{2}p) \\
=&(p - \frac{1}{2})R_{12} + \alpha(\frac{1}{2} - \frac{1}{2}p) \\
\leq&(p - \frac{1}{2}) + \alpha(\frac{1}{2} - \frac{1}{2}p)
\end{aligned}$$

Where the second inequality holds because $R_{11} + R_{12} - \alpha \leq R_{21} + R_{22}$.

Case R^4
The payoff of the row player is equal to:

$$\frac{1}{4}R_{11} + \frac{1}{4}R_{12} + \frac{1}{4}R_{21} + \frac{1}{4}R_{22}$$

Like the previous case, deviation to row 2 is symmetric to deviation to row 1, so we will only analyse row 1. If the row player deviates to row 1, his payoff would be $\frac{1}{2}R_{11} + \frac{1}{2}R_{12}$. His incentive to deviate is:

$$
\begin{aligned}
&(\frac{1}{2}R_{11} + \frac{1}{2}R_{12}) - (\frac{1}{4}R_{11} + \frac{1}{4}R_{12} + \frac{1}{4}R_{21} + \frac{1}{4}R_{22}) \\
=&(\frac{1}{4}R_{11} + \frac{1}{4}R_{12}) - (\frac{1}{4}R_{21} + \frac{1}{4}R_{22}) \\
=&(\frac{1}{4}R_{11} + \frac{1}{4}R_{12}) - \frac{1}{4}(R_{21} + R_{22}) \\
\leq&\frac{1}{4}(R_{11} + R_{12}) - \frac{1}{4}(R_{11} + R_{12} - \alpha) \\
\leq&\frac{1}{4}(R_{11} + R_{12}) - \frac{1}{4}(R_{11} + R_{12}) + \frac{1}{4}\alpha \\
=&\frac{1}{4}\alpha
\end{aligned}
$$

Where the first inequality holds because $R_{11} + R_{12} - \alpha \leq R_{21} + R_{22}$.

For case R^1, a deviation to row 1 will always gives a $(1-p)$-approximate Nash equilibrium, regardless the value of α. The same is true for deviation to row 1 or row 2 in case R^2. For R_1 with deviation to row 2 and for the cases R^3 and R^4, the incentive to deviate depends on α. To get a $(1-p)$-approximate Nash equilibrium, the following should hold:

$$
\begin{aligned}
(2p^2 - p) - \alpha p(1-p) &\leq (1-p) \\
(p - \frac{1}{2}) + \alpha(\frac{1}{2} - \frac{1}{2}p) &\leq (1-p) \\
\frac{1}{4}\alpha &\leq (1-p) \\
\frac{1}{2} &\leq p \leq 1 \\
\alpha &\geq 0
\end{aligned}
$$

Solving this set of inequalities gives:

$$
\begin{aligned}
0 \leq \alpha \leq \frac{1}{4}(7 - \sqrt{33}) \quad &\text{and} \quad \frac{1}{2} \leq p \leq \frac{\sqrt{\frac{\alpha^2+4\alpha+8}{(\alpha+2)^2}}\alpha + 2\sqrt{\frac{\alpha^2+4\alpha+8}{(\alpha+2)^2}} + \alpha}{2(\alpha+2)} \\
\frac{1}{4}(7 - \sqrt{33}) < \alpha < 2 \quad &\text{and} \quad \frac{1}{2} \leq p \leq \frac{\alpha - 3}{\alpha - 4} \\
\alpha = 2 \quad &\text{and} \quad p = \frac{1}{2}
\end{aligned}
$$

We want to maximise p, this will give the lowest $(1-p)$-approximate Nash equilibrium. The function $p(\alpha) = \frac{\sqrt{\frac{\alpha^2+4\alpha+8}{(\alpha+2)^2}}\alpha+2\sqrt{\frac{\alpha^2+4\alpha+8}{(\alpha+2)^2}}+\alpha}{2(\alpha+2)}$ is monotone increasing for $\alpha \geq 0$, so the value of p is the highest when $\alpha = \frac{1}{4}(7 - \sqrt{33})$. This gives $p(\frac{1}{4}(7 - \sqrt{33})) \approx 0.729$.

The function $p(\alpha) = \frac{\alpha-3}{\alpha-4}$ is monotone decreasing for $0 \leq \alpha \leq 4$, and the difference in payoff cannot be larger than 2. Because the function is monotone decreasing, α should be as low as possible, so $\alpha = \frac{1}{4}(7 - \sqrt{33}) + \epsilon$. This gives $p(\frac{1}{4}(7 - \sqrt{33} + \epsilon)) \approx 0.729$.

So the maximum value we get is $p \approx 0.729$, which gives a 0.271-approximate Nash equilibrium. Earlier a lower bound of $\frac{1}{4}$ was shown, so a big improvement on this algorithm cannot be found. □

3.4 Upper bound on ϵ with very limited communication

If we allow some communication between the players, we can improve over the procedures with no communication. We will try to find procedures with a good bound on ϵ and only a few bits of communication.

3.4.1 A $\frac{1}{4}$ approximation with 1 round of communication for 2×2-games

This algorithm will only use 1 bit of communication per player. We assume the communication to be symmetric, so both players send their 1 bit at the same time. The information that a player will send is which strategy he will give a probability of at least $\frac{1}{2}$, this is enough to ensure a $\frac{1}{4}$-approximate Nash equilibrium.

Theorem 5. *For 2×2-games with one round of communication, the players can achieve a $\frac{1}{4}$-approximate Nash equilibrium.*

We will show the procedure for the row player, the procedure for the column player is similar. Like the procedure in the previous section, with $p > \frac{1}{2}$, the players can play three different strategies:

- $f_r(R) = p$
- $f_r(R) = (1 - p)$
- $f_r(R) = \frac{1}{2}$

The row player tells the column player if $R_{11}+R_{12} \geq R_{21}+R_{22}$. Assume $R_{11}+R_{12} \geq R_{21}+R_{22}$. By giving this information to the column player, the row player assures the column player that $f_r(R) \geq \frac{1}{2}$. The row player receives the same information from the column player. Assume w.l.o.g. the column player sends $f_c(C) \leq \frac{1}{2}$.

The row player can play two different strategies: $f_r(R) = p$ or $f_r(R) = \frac{1}{2}$. The options for the column player are restricted to $f_c(C) = 1 - p$ and $f_c(C) = \frac{1}{2}$. This leads to the following cases:

$$\text{case } R^1 : \begin{matrix} & 1-p & p \\ p & R_{11} & R_{12} \\ 1-p & R_{21} & R_{22} \end{matrix} \qquad \text{case } R^2 : \begin{matrix} & \frac{1}{2} & \frac{1}{2} \\ p & R_{11} & R_{12} \\ 1-p & R_{21} & R_{22} \end{matrix}$$

$$\text{case } R^3 : \begin{matrix} & 1-p & p \\ \frac{1}{2} & R_{11} & R_{12} \\ \frac{1}{2} & R_{21} & R_{22} \end{matrix} \qquad \text{case } R^4 : \begin{matrix} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & R_{11} & R_{12} \\ \frac{1}{2} & R_{21} & R_{22} \end{matrix}$$

After receiving the bit of communication from the column player the row player considers to change his strategy from $f_r(R) = p$ to $f_r(R) = \frac{1}{2}$. He starts with analysing what would happen if he does not change his strategy and keeps $f_r(R) = p$. When the column player would play $f_c(C) = \frac{1}{2}$, so case R^2, the best response for the row player is row 1. His incentive to deviate to this row is smaller or equal to $1 - p$, as we have seen in the previous section.

When the column player plays $f_c(C) = 1 - p$, so case R^1, the row player can have an incentive to deviate to either row. If he has an incentive to deviate to row 1, his incentive to deviate is smaller or equal to $1-p$ as we have seen in the previous section. If this holds after he received the bit of communication from the column player he does not change his strategy. If, however, his best response is row 2, the row player will make the decision to change his strategy to $f_r(R) = \frac{1}{2}$. This brings us to the cases R^3 and R^4.

When the column player sticks to $f_c(C) = 1-p$, we get case R^3. The probability distribution of the column player didn't change, so the best response of the row player is still row 2. The incentive to deviate in this case is:

$$\begin{aligned}
&(1-p)R_{21} + pR_{22} - ((\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12} + (\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22}) \\
=&(\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22} - ((\frac{1}{2} - \frac{1}{2}p)R_{11} + \frac{1}{2}pR_{12}) \\
\leq&(\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22} - (\frac{1}{2} - \frac{1}{2}p)(R_{11} + R_{12}) \\
\leq&(\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22} - (\frac{1}{2} - \frac{1}{2}p)(R_{21} + R_{22}) \\
=&(\frac{1}{2} - \frac{1}{2}p)R_{21} + \frac{1}{2}pR_{22} - (\frac{1}{2} - \frac{1}{2}p)R_{21} - (\frac{1}{2} - \frac{1}{2}p)R_{22} \\
=&(p - \frac{1}{2})R_{22} \\
\leq&(p - \frac{1}{2})
\end{aligned}$$

Where the second inequality holds because $R_{11} + R_{12} \geq R_{21} + R_{22}$.

If the column player also changes his mixed-strategy profile to $f_c(C) = \frac{1}{2}$, we get case R^4. Before we look at the incentive to deviate in this situation, we will first look at some properties of the values of the payoffs. We assumed $R_{11} + R_{12} \geq R_{21} + R_{22}$. The best response for the row player was row 2, so $(1-p)R_{11} + pR_{12} \leq (1-p)R_{21} + pR_{22}$. Those two observations lead to the following:

$$\begin{aligned}
(1-p)R_{11} + pR_{12} \leq&(1-p)R_{21} + pR_{22} \\
R_{11} + \frac{p}{1-p}R_{12} \leq&R_{21} + \frac{p}{1-p}R_{22} \\
R_{11} + \frac{p}{1-p}R_{12} - \frac{2p-1}{1-p}R_{22} \leq&R_{21} + R_{22} \leq R_{11} + R_{12} \\
R_{11} + \frac{p}{1-p}R_{12} - \frac{2p-1}{1-p}R_{22} \leq&R_{11} + R_{12} \\
\frac{p}{1-p}R_{12} - \frac{2p-1}{1-p}R_{22} \leq&R_{12} \\
\frac{p}{1-p}R_{12} - \frac{1-p}{1-p}R_{12} \leq&\frac{2p-1}{1-p}R_{22} \\
\frac{2p-1}{1-p}R_{12} \leq&\frac{2p-1}{1-p}R_{22} \\
R_{12} \leq&R_{22}
\end{aligned}$$

The best response for the row player can only be row 1, because $R_{11} + R_{12} \geq R_{21} + R_{22}$. With this information we can compute the incentive to deviate for the row player:

$$\begin{aligned}
&(\frac{1}{2}R_{11} + \frac{1}{2}R_{12}) - (\frac{1}{4}R_{11} + \frac{1}{4}R_{12} + \frac{1}{4}R_{21} + \frac{1}{4}R_{22}) \\
=&\frac{1}{4}R_{11} + \frac{1}{4}R_{12} - \frac{1}{4}R_{21} - \frac{1}{4}R_{22} \\
\leq&\frac{1}{4}R_{11} + \frac{1}{4}R_{12} - \frac{1}{4}R_{21} - \frac{1}{4}R_{12} \\
=&\frac{1}{4}R_{11} - \frac{1}{4}R_{21} \leq \frac{1}{4}R_{11} \leq \frac{1}{4}
\end{aligned}$$

So the incentive to deviate is smaller or equal to $\frac{1}{4}$.

We have covered all the possible situations, the conditions we have to meet are:

$$\begin{aligned}
p - \frac{1}{2} &\leq 1 - p \\
\frac{1}{4} &\leq 1 - p
\end{aligned}$$

When we maximise this, it clearly leads to $p = \frac{3}{4}$. This means that we have proven a $\frac{1}{4}$-approximate Nash equilibrium with only 1 symmetric round of communication. □

3.4.2 Breaking the $\frac{1}{4}$ barrier for 2×2-games with a bit more communication

In the previous section we described a procedure that achieved a $\frac{1}{4}$-approximate Nash equilibrium. We achieved this bound by changing the strategy of a player if his best response was to the pure strategy that he gave a low probability. We changed the probability of that strategy from $1 - p$ to $\frac{1}{2}$. By choosing $\frac{1}{2}$, we ensured that if both players changed their strategy the incentive to deviate couldn't be larger than $\frac{1}{4}$. If only one player changed his strategy, his incentive to deviate would be limited by $p - \frac{1}{2}$ and the other player could not be harmed by this change of strategy.

If we want to improve the $\frac{1}{4}$-approximate Nash equilibrium, we have to address both the case where one of the players changes his strategy and the case where both players change their strategy. Both the $\frac{1}{4}$ and the $p - \frac{1}{2}$ bounds are tight and lead to an incentive to deviate of $\frac{1}{4}$. The bound on the case where both players change their strategy seems hard to improve. Therefore we will try to find a procedure where at most one of the players is allowed to change his strategy. If only one on the players changes his strategy to a strategy that is close to a fully mixed strategy, the other player will only have a small incentive to deviate when he sticks to his original strategy. This change in the model requires a little bit more information, we have to make sure that at most one of the players changes his strategy.

Theorem 6. *For 2×2-games with two rounds of communication, the players can achieve a 0.211-approximate Nash equilibrium.*

The bound of $p - \frac{1}{2}$ can be lowered by assigning a higher probability to the pure strategy that is the best response for the player. However, we have to be careful with changing this probability, because it could change the best response of the other player. So we will have to find a value that lowers the $p - \frac{1}{2}$ bound and keeps the incentive to deviate of the other player also low enough. If the row player has an incentive to deviate to the row that has a probability of $1-p$, he will now play this row with probability $1-a$ and the other row with probability a.

Because only one of the players is allowed to change his original strategy, we have only three cases, with $a \leq \frac{1}{2}$ and $p \geq \frac{1}{2}$.

$$R^1 : \begin{array}{c} \\ p \\ 1-p \end{array} \begin{array}{c} \begin{array}{cc} 1-p & p \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \end{array} \quad R^2 : \begin{array}{c} \\ p \\ 1-p \end{array} \begin{array}{c} \begin{array}{cc} 1-a & a \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \end{array} \quad R^3 : \begin{array}{c} \\ a \\ 1-a \end{array} \begin{array}{c} \begin{array}{cc} 1-p & p \end{array} \\ \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \end{array}$$

When we have case R^1 and the best response of the row player is to row 1, the incentive to deviate will be smaller than or equal to $1-p$. When the best response of the row player is row 2 and column 2 for the column player, we will get case R^2. If the best response for the column player is column 1 and row 1 for the row player we will get case R^3. When the best response of the row player is row 2 and the best response of the column player is column 1, we will still get case R^2 or R^3, via some communication the players will decide which player may change his strategy profile.

If case R^2, the player can have an incentive to deviate to either row. A deviation to row 1 obviously leads to an incentive to deviate smaller or equal to $1-p$. The incentive to deviate to row 2 is:

$$\begin{aligned}
&(1-a)R_{21} + aR_{22} - (1-a)pR_{11} - apR_{12} - (1-a)(1-p)R_{21} - a(1-p)R_{22} \\
=&(1-a)pR_{21} + apR_{22} - (1-a)pR_{11} - apR_{12} \\
\leq&(1-a)pR_{21} + apR_{22} - ap(R_{11} + R_{12}) \\
\leq&(1-a)pR_{21} + apR_{22} - ap(R_{21} + R_{22}) \\
=&(1-2a)pR_{21} \leq (1-2a)p
\end{aligned}$$

If R^3, the best response of the row player is row 2, if his best response was row 1 he would not have changed his strategy. The incentive to deviate for this case is:

$$\begin{aligned}
&(1-p)R_{21} + pR_{22} - a(1-p)R_{11} - apR_{12} - (1-a)(1-p)R_{21} - (1-a)pR_{22} \\
=&(a-ap)R_{21} + apR_{22} - a(1-p)R_{11} - apR_{12} \\
\leq&(a-ap)R_{21} + apR_{22} - a(1-p)(R_{11} + R_{12}) \\
\leq&(a-ap)R_{21} + apR_{22} - a(1-p)(R_{21} + R_{22}) \\
=&(2ap-a)R_{22} \leq 2ap - a
\end{aligned}$$

Since we want to improve over the $\frac{1}{4}$-approximate Nash equilibrium of the previous section, we obviously want $p \geq \frac{3}{4}$. This leads to the following set of inequalities we want to solve:

$$\begin{aligned}
(1-2a)p &\leq 1-p \\
2ap - a &\leq 1-p \\
\frac{3}{4} &\leq p \leq 1 \\
0 &\leq a \leq 1
\end{aligned}$$

When we solve this we get:

$$\begin{aligned}
a = \frac{1}{3} &\quad \text{and} \quad p = \frac{3}{4} \\
\frac{1}{3} < a \le \frac{1}{2}(\sqrt{3}-1) &\quad \text{and} \quad \frac{3}{4} \le p \le -\frac{1}{2(a-1)} \\
\frac{1}{2}(\sqrt{3}-1) < a < \frac{1}{2} &\quad \text{and} \quad \frac{3}{4} \le p \le -\frac{a+1}{2a+1} \\
a = \frac{1}{2} &\quad \text{and} \quad p = \frac{3}{4}
\end{aligned}$$

We want to find a maximal value of p, this is reached when $a = \frac{1}{2}(\sqrt{3}-1)$ with $p \approx 0.789$. This leads to an incentive to deviate smaller or equal to $1-p = 1-0.789 = 0.211$. So we have found a procedure that gives us a 0.211-approximate Nash equilibrium. □

The procedure for achieving a 0.211-approximate Nash equilibrium

We can achieve a 0.211-approxmate Nash equilibrium with only a few bits of communication. When we use an asymmetric communication protocol we only need three bits of communcation. First player 1 (it is agreed beforehand who is player 1), tells player 2 if $R_{11} + R_{12} \ge R_{21} + R_{22}$. With this information player 2 can decide which strategy he should play: $f_c(C) = 0.789$, $f_c(C) = 0.211$, $f_c(C) = \frac{1}{2}(\sqrt{3}-1)$ or $f_c(C) = 1 - \frac{1}{2}(\sqrt{3}-1)$. Player 2 will communicate which strategy he will use, this will cost 2 bits. Player 1 now knows which strategy player 2 will use. If player 2 will play $f_c(C) = \frac{1}{2}(\sqrt{3}-1)$ or $f_c(C) = 1 - \frac{1}{2}(\sqrt{3}-1)$, player 1 has to play $f_r(R) = 0.789$ if $R_{11} + R_{12} \ge R_{21} + R_{22}$ or else $f_r(R) = 0.211$. If player 2 plays $f_c(C) = 0.789$ or $f_c(C) = 0.211$, player 1 can choose between all four strategies.

When we use a symmetric communication protocol, we need 2 rounds of communcation. The first round of communication is the same as in the previous section. In the second round of communication they tell each other if they want to change their strategy. If the both want to change their strategy, only player 1 is allowed to do so (they have agreed beforehand who player 1 is).

Chapter 4

Upper bounds on the communication complexity

We will look at some procedures where the whole payoff matrix is communicated to the other player. When you know all the entries of both payoff matrices, all the information of a game is available. So these models will give a certain upper bound on the amount of communication of any procedure.

For the computation of an approximate Nash equilibrium we use the subexponential algorithm developed by Lipton et al. [14]. Next to a subexponential running time, this algorithm guarantees strategies for both players with a support size logarithmic in the number of pure strategies. This allows efficient communication of the strategy of the player who computes the approximate Nash equilibrium to the other player.

The algorithms presented in this chapter give an ϵ-approximate Nash equilibrium instead of exact Nash equilibria for two reasons. First the payoff entries from a payoff matrix have to be communicated. The exact values cannot be communicated, the values are rounded. Second, the algorithm by Lipton et al. gives an approximate Nash equilibrium instead of on exact Nash equilibrium. Together these two factors give an ϵ-approximate Nash equilibrium for each algorithm.

The algorithm by Lipton et al. guarantees an ϵ-approximate Nash equilibrium when the support size k is at least $\frac{12 \ln n}{\epsilon^2}$.

4.1 $n \times n$-games

Given a $n \times n$ game $G = (R, C)$ where the entries of R and C are in $[0, 1]$ and $\delta = \frac{1}{2}\epsilon$.

Algorithm 1

1. Discretize $[0, 1]$ into the set $\mathcal{V} = \{\delta - \frac{1}{2}\delta, 2\delta - \frac{1}{2}\delta, \ldots, k\delta - \frac{1}{2}\delta\}$ where $k\delta - \frac{1}{2}\delta \leq 1$ and $(k+1)\delta - \frac{1}{2}\delta > 1, k \in \mathbb{N}$
2. Define a new game G' where each R'_{ij} and C'_{ij} gets a value from the set $\mathcal{V}$ such that:
   $\forall i, j : |R'_{ij} - R_{ij}| \leq \delta/2$
   $\forall i, j : |C'_{ij} - C_{ij}| \leq \delta/2$
3. Let the column player communicate C' to the row player

4. The row player computes a Nash equilibrium strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ of the game $G' = (R', C')$ with the Lipton et al. algorithm that uses a support size $k = \frac{12 \ln n}{\delta^2}$
5. The row player communicates $\mathbf{y}^*$ to the column player

Theorem 7. *Algorithm 1 is an ϵ-approximate Nash equilibrium of G.*

Proof:
The strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is a δ-approximate Nash equilibrium of G', because $\delta = \frac{1}{2}\epsilon$ this is equal to a $\frac{1}{2}\epsilon$-approximate Nash equilibrium. If it is also a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G we are done.

The payoff entries of R are rounded, so it could be that $(\mathbf{x}^*, \mathbf{y}^*)$ is not a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G. However, the incentive to deviate is bounded because the payoff entries of R' are close to the entries of R. Because mixed strategies are a convex combination of pure strategies, the deviation to a pure strategy will be the biggest. As mentioned before the difference between any payoff in the payoff matrices R and R' is not bigger than $\delta/2$. So for any pure strategy $\mathbf{e}_i$:

$$\begin{aligned} \mathbf{e}_i^T R\mathbf{y}^* - (\mathbf{x}^*)^T R\mathbf{y}^* &= |\mathbf{e}_i^T R\mathbf{y}^* - \mathbf{e}_i^T R'\mathbf{y}^*| + |(\mathbf{x}^*)^T R\mathbf{y}^* - (\mathbf{x}^*)^T R'\mathbf{y}^*| \\ &\leq \delta/2 + \delta/2 \leq \delta \end{aligned}$$

So the extra incentive to deviate because the row player has R instead of R' is at most δ, where $\delta = \frac{1}{2}\epsilon$. The strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G'. Together this guarantees a ϵ-approximate Nash equilibrium for the row player.

The same argument can be used for the column player:

$$\begin{aligned} \mathbf{e}_i^T C\mathbf{y}^* - \mathbf{x}^* C\mathbf{y}^* &= |\mathbf{e}_i^T C\mathbf{y}^* - \mathbf{e}_i^T C'\mathbf{y}^*| + |(\mathbf{x}^*)^T C\mathbf{y}^* - (\mathbf{x}^*)^T C'\mathbf{y}^*| \\ &\leq \delta/2 + \delta/2 \leq \delta \end{aligned}$$

The extra incentive to deviate because the column player has C instead of C' is also at most δ. The strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G'. Together this guarantees a ϵ-approximate Nash equilibrium for the column player. □

Theorem 8. *$CC(\text{Algorithm 1}, \mathcal{G}, \epsilon) = O(n^2 \log 2/\epsilon + \log^2 n)$, where $\mathcal{G}$ is the family of $n \times n$-games.*

Proof:
All the entries in the payoff matrix are communicated. There are $2/\epsilon$ different values for these entries. So sending one entry has a communication complexity of $O(\log 2/\epsilon)$. The payoff matrix has n^2 entries, so the total communication complexity of communicating a payoff matrix is $O(n^2 \log 2/\epsilon)$.

The row player sends a mixed strategy to the column player. The support size of this mixed strategy is logarithmic in the number of pure strategies. Communicating a pure strategy has a communication complexity of $O(\log n)$, so the communication complexity of a mixed strategy is $O(\log^2 n)$. □

4.2 $2 \times n$-games

The row player has 2 strategies and the column player has n strategies. We want to compute an ϵ-approximate Nash equilibrium of this game. Like the previous section we will compute an approximate Nash equilibrium using the Lipton et al. algorithm. The computation of this approximate Nash equilibrium will be on a reduced game that is smaller than the original game.

The observation we have to make is that strategies of the column player that are very similar, i.e. have almost the same payoffs, will give similar approximate Nash equilibria. After rounding the payoff matrix of the column player duplicate columns are removed, such that every payoff combination occurs at most one time in C'.

Given a $2 \times n$ game $G = (R, C)$ where the entries of R and C are in $[0, 1]$, $\epsilon < 1$ and $\delta = \frac{1}{2}\epsilon$.

Algorithm 2

1. Discretize $[0, 1]$ into the set $\mathcal{V} = \{\delta - \frac{1}{2}\delta, 2\delta - \frac{1}{2}\delta, \ldots, k\delta - \frac{1}{2}\delta\}$ where $k\delta - \frac{1}{2}\delta \leq 1$ and $(k+1)\delta - \frac{1}{2}\delta > 1, k \in \mathbb{N}$
2. Define a new game G' where each R'_{ij} and C'_{ij} gets a value from the set $\mathcal{V}$ such that:
   $\forall i, j : |R'_{ij} - R_{ij}| \leq \delta/2$
   $\forall i, j : |C'_{ij} - C_{ij}| \leq \delta/2$
3. Remove any duplicate columns from C', so $|C'| \leq (1/\delta)^2$
4. Column player: $\forall j$: Communicate C'_{1j} and C'_{2j} and the column number (in C) to the row player.
5. The row player computes a Nash equilibrium strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ of the game $G' = (R', C')$ with the Lipton et al. algorithm that uses a support size $k = \frac{12 \ln n}{\delta^2}$
6. The row player communicates $\mathbf{y}^*$ to the column player

Theorem 9. *Algorithm 2 gives an ϵ-approximate Nash equilibrium.*

Proof:
The row player rounds his payoffs and has the same number of strategies in G' as in G. So the proof is the same as the proof for algorithm 1.

The column player rounded his payoffs and removed a number of strategies. For strategies that were in C and C' the proof is the same as the proof in used in algorithm 1. If a strategy was not in C', there was another strategy that was rounded to the same value and this strategy was in C'. The values of that strategy cannot be distinguished anymore from the strategy that was used in C' so the same argument holds as in algorithm 1. This proves the theorem. □

Theorem 10. *$CC(\text{Algorithm 2}, \mathcal{G}, \epsilon) = O((2/\epsilon)^2 \cdot (\log n + 2\log 2/\epsilon) + \log^2 n)$, with $\mathcal{G}$ the family of $2 \times n$-games.*

Proof:
The information that has to be sent for one column is which column it is and the two rounded values of that column. Sending the column number has a communication complexity of $O(\log n)$ and sending a value a communication complexity of $O(\log 2/\epsilon)$ bits. Because $|C'| \leq (2/\epsilon)^2$ the total communication complexity for communicating the payoff matrix is $O((2/\epsilon)^2(\log n + 2\log 2/\epsilon))$. The row player sends a mixed strategy $\mathbf{y}^*$ to the column player, which has a communication complexity of $O(\log^2 n)$. □

4.2.1 Improving the communication complexity bound

In the previous algorithm the size of C' (and R') was bounded by $(2/\epsilon)^2$. If we look at the values of the payoff entries in C' we can see that there could be dominated strategies in C'. If we remove all weakly dominated strategies, the size of C' will only be $O(2/\epsilon)$. All the strategies consist of two payoffs (C_{1j}, C_{2j}). Every value v of a payoff can occur at most once, in combination with the highest value for the other payoff in C'. All other combinations with v will be weakly dominated by this strategy. Since a payoff can only have $2/\epsilon$ different values, $|C'| \leq 2/\epsilon$.

It could be that one of the remaining strategies in C' is strictly dominant. In this case the column player communicates this strategy to the row player and the row player computes a best response. This will be a Nash equilibrium of G'. Else we will follow algorithm 2 and compute a Nash equilibrium. Because $|C'| \leq 2/\epsilon$ the communication complexity will be $O((2/\epsilon) \cdot (\log n + 2\log 2/\epsilon) + \log^2 n)$.

4.3 $d \times n$-games

Algorithm 2 can be generalised to the situtation where the row player has a small number of strategies, say d strategies.

Given a $d \times n$ game $G = (R, C)$ where the entries of R and C are in $[0, 1]$, $\epsilon < 1$ and $\delta = \frac{1}{2}\epsilon$.

Algorithm 3

1. Discretize $[0, 1]$ into the set $\mathcal{V} = \{\delta - \frac{1}{2}\delta, 2\delta - \frac{1}{2}\delta, \ldots, k\delta - \frac{1}{2}\delta\}$ where $k\delta - \frac{1}{2}\delta \leq 1$ and $(k+1)\delta - \frac{1}{2}\delta > 1, k \in \mathbb{N}$
2. Define a new game G' where each R'_{ij} and C'_{ij} gets a value from the set $\mathcal{V}$ such that:
   $\forall i, j : |R'_{ij} - R_{ij}| \leq \delta/2$
   $\forall i, j : |C'_{ij} - C_{ij}| \leq \delta/2$
3. Remove any duplicate or weakly dominated columns from C' so $|C'| \leq (1/\delta)^{d-1}$
4. Column player: $\forall j$: Communicate $C'_{1j} \ldots C'_{dj}$ and the column number (in C) to the row player.
5. The row player computes a Nash equilibrium strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ of the game $G' = (R', C')$ with the Lipton et al. algorithm that uses a support size $k = \frac{12 \ln n}{\delta^2}$
6. The row player communicates $\mathbf{y}^*$ to the column player

Theorem 11. *Algorithm 3 gives an ϵ-approximate Nash equilibrium of the game $G = (R, C)$.*

Proof:
The strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is a δ-approximate Nash equilibrium, so a $\frac{1}{2}\epsilon$-approximate Nash equilibrium, of G'. If it is also a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G we are done.

The payoff entries of R are rounded, so it could be that $(\mathbf{x}^*, \mathbf{y}^*)$ is not a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G. However, the incentive to deviate is bounded becasue the payoff entries of R' are close to the entries of R. Because mixed strategies are a convex combination of pure strategies, the deviation to a pure strategy will be the biggest. As mentioned before the difference between any payoff in the payoff matrices R and R' is not bigger than $\delta/2$. So for any pure strategy $\mathbf{e}_i$:

$$\begin{aligned} \mathbf{e}_i^T R\mathbf{y}^* - (\mathbf{x}^*)^T R\mathbf{y}^* &= |\mathbf{e}_i^T R\mathbf{y}^* - \mathbf{e}_i^T R'\mathbf{y}^*| + |(\mathbf{x}^*)^T R\mathbf{y}^* - (\mathbf{x}^*)^T R'\mathbf{y}^*| \\ &\leq \delta/2 + \delta/2 \leq \delta \end{aligned}$$

So the extra incentive to deviate because the row player has R instead of R' is at most δ, where $\delta = \frac{1}{2}\epsilon$. The strategy pair $(\mathbf{x}^*, \mathbf{y}^*)$ is a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G'. Together this guarantees an ϵ-approximate Nash equilibrium for the row player.

The column player rounded his payoffs and removed a number of strategies. The pure strategy $\mathbf{e}_i$ of G to which the column player has the largest incentive to deviate can be of three different kinds:

- $\mathbf{e}_i$ was a strategy of C'.
- $\mathbf{e}_i$ was removed from C' because it was a duplicate column.
- $\mathbf{e}_i$ was removed from C' because it was weakly dominated.

If $\mathbf{e}_i$ was in both C and C', the proof is similar to the proof of the row player. Because the difference in payoffs between C and C' cannot be larger than $\delta/2$, the incentive to deviate cannot be larger than:

$$\begin{aligned}\mathbf{e}_i^T C\mathbf{y}^* - \mathbf{x}^* C\mathbf{y}^* &= |(\mathbf{e}_i^T C\mathbf{y}^{\mathbf{x}} - \mathbf{e}_i^T C'\mathbf{y}^*)| + |((\mathbf{x}^*)^T C\mathbf{y}^* - (\mathbf{x}^*)^T C'\mathbf{y}^*)| \\ &\leq \delta/2 + \delta/2 \leq \delta\end{aligned}$$

Since $\delta = \frac{1}{2}\epsilon$ and $(\mathbf{x}^*, \mathbf{y}^*)$ is a $\frac{1}{2}\epsilon$-approximate Nash equilibrium of G', this gives a ϵ-approximate Nash equilibrium for the column player.

If the strategy belonging to $\mathbf{e}_i$ was removed from C' because it was a duplicate column, there is another strategy c_j that was rounded to the same value and it was not removed from C' as a duplicate column. If not c_j was used and $\mathbf{e}_i$ was used, it was rounded to the same value and the outcome of the equilibrium would be the same. So after the rounding you cannot distinguish these two strategies anymore. So the same argument holds as in the previous case.

The last case, where the strategy belonging to $\mathbf{e}_i$ was removed from C' because it was weakly dominated by another strategy, we take such a weakly dominated strategy c_j in C'. This means that the strategy $\mathbf{e}_i$ in C' had at least one payoff that was lower than c_j and the rest of the payoffs were not higher. Assume we add some (positive) constant to the payoff entry of $\mathbf{e}_i$ that was lower than the payoff entry of c_j in C'. The two strategies are now the same in C'. This means we can use the argument of the previous case to show that the incentive to deviate cannot be larger than δ. Because we only added a postive constant to the original rounded $\mathbf{e}_i$, the original rounded $\mathbf{e}_i$ also cannot have a larger payoff, so the incentive to deviate will be bounded by δ. This concludes the proof. □

Theorem 12. *CC(Algorithm 3, $\mathcal{G}, \epsilon$) $= O((2/\epsilon)^{d-1}(\log n + d \log 2/\epsilon) + \log^2 n)$, with $\mathcal{G}$ the family of $d \times n$-games.*

Proof:
The communication complexity of communicating one payoff is $O(\log 2/\epsilon)$. Every column has d payoffs and communication of the column number has a communication complexity of $O(\log n)$. So the communication complexity of communicating one column is $O(\log n + d \log 2/\epsilon)$. After the removal of all duplicate columns, $|C'| \leq (2/\epsilon)^d$. After removal of the weakly dominated strategies $|C'| \leq (2/\epsilon)^{d-1}$. This can be explained by fixing $d-1$ payoffs for a column. The last payoff can have any value. The column with the highest payoff in the last column will weakly dominate all the other strategies, because the payoffs are the same for the first $d-1$ and the last payoff has a higher payoff. So the removal of weakly dominated strategies will decrease the number of strategies by a factor d. Because $|C'| \leq (2/\epsilon)^{d-1}$ we will get a total communication complexity of $O((2/\epsilon)^{d-1}(\log n + d \log 2/\epsilon))$ for communicating a payoff matrix. The row player communicates a mixed strategy to the column player, this has a communication complexity of $O(\log^2 n)$. This proves the theorem. □

Chapter 5

Lower bounds for good approximations

In this section we will look at a lower bound on the communication complexity when we want to reach a certain good approximation. First we will look at a special kind of 2×2-games and after that at $n \times n$-games.

5.1 Two actions per player

We will show that for certain games a player needs to know the approximate value of the payoffs of the other player to compute an ϵ-approximate Nash equilibrium. This corresponds to a communication complexity of $\Omega(\log 1/\epsilon)$. To prove this for 2×2-games we will look at following game, suggested in [10] with $\frac{1}{2} \leq M \leq 1$:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

This game has a unique Nash equilibrium, with $f_r(R) = (\frac{1}{2}, \frac{1}{2})$ and $f_c(C) = (\frac{M}{M+1}, \frac{1}{M+1})$. To reach this unique Nash equilibrium the column player needs to know the exact value of M, which is only known to the row player. What we will prove is that to reach an ϵ-approximate Nash equilibrium, the column player has to know a good approximation of M. If we can prove this, it means that the row player had to communicate the approximate value of M, which would take $\Omega(\log 1/\epsilon)$ bits.

Theorem 13. *For every ϵ-approximate Nash equilibrium, the strategy of the column player, f_c, has to lie in the range $[\frac{M}{M+1} - 2\epsilon, \frac{M}{M+1} + 2\epsilon]$ with $0 < \epsilon < \frac{1}{12}$.*

In the exact Nash equilibrium, the strategy of the row player is $f_r = \frac{1}{2}$. Assume that for some particular ϵ-approximate Nash equilibrium, the strategy of the row player is $f_r = \frac{1}{2} + \kappa$. The value of κ cannot just be any value in the interval $\kappa \in [-\frac{1}{2}, \frac{1}{2}]$ for values of ϵ smaller than $\frac{1}{12}$. We will restrict the interval of κ to $\kappa \in [-\frac{1}{6}, \frac{1}{6}]$. We will first proof that the restriction on κ does not weaken the proof of theorem 10.

Lemma 2. *For every ϵ-approximate Nash equilibrium, κ should at least be larger than $-\frac{1}{6}$ and smaller than $\frac{1}{6}$.*

Suppose $\kappa > \frac{1}{6}$. This implies $f_r \geq \frac{2}{3}$. It also means that the column player has an incentive to deviate to column 2, his payoff will be at least $\frac{2}{3}$ if he plays the pure strategy $f_c = 0$. To ensure an ϵ-approximate Nash equilibrium, his initial support for column 2 should be high. Let a be the strategy of the column player, the probability that he assigns to pure strategy 1. The incentive to deviate for the column player will be:

$$\begin{aligned}
\frac{1}{2} + \kappa - ((\frac{1}{2} + \kappa)(1-a) + (\frac{1}{2} - \kappa)a) &\leq \epsilon \\
\frac{1}{2} + \kappa - (\frac{1}{2} - \frac{1}{2}a + \kappa - a\kappa + \frac{1}{2}a - a\kappa) &\leq \epsilon \\
2a\kappa &\leq \epsilon \\
\frac{1}{3}a &\leq \epsilon \\
a &\leq 3\epsilon
\end{aligned}$$

If we return to the row player, the payoff matrix will look like:

$$R = \begin{array}{c} \\ \geq \frac{2}{3} \\ \leq \frac{1}{3} \end{array} \begin{array}{c} \begin{array}{cc} \leq 3\epsilon & \geq 1-3\epsilon \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix} \end{array}$$

Since $\frac{1}{2} \leq M \leq 1$, a deviation to row 2 is clearly preferred. we can minimize this incentive to deviate by making the support for row 2 as large as possible, so $\frac{1}{3}$ and M as small as possible. This still gives an incentive to deviate of:

$$\begin{aligned}
(1-3\epsilon)\frac{1}{2} - (\frac{2}{3} \cdot 3\epsilon + \frac{1}{3} \cdot (1-3\epsilon) \cdot \frac{1}{2}) &= \\
\frac{1}{2} - \frac{3}{2}\epsilon - 2\epsilon - \frac{1}{6} + \frac{1}{2}\epsilon &= \\
\frac{1}{3} - 3\epsilon &
\end{aligned}$$

By definition, the incentive to deviate cannot be larger than ϵ:

$$\begin{aligned}
\frac{1}{3} - 3\epsilon &\leq \epsilon \\
\frac{1}{3} &\leq 4\epsilon \\
\frac{1}{12} &\leq \epsilon
\end{aligned}$$

So for values of $\epsilon < \frac{1}{12}$, κ cannot be larger than $\frac{1}{6}$.

Suppose $\kappa < -\frac{1}{6}$. This implies $f_r \leq \frac{1}{3}$. It also means that the column player has an incentive to deviate to column 1, his payoff will be at least $\frac{2}{3}$ if he plays the pure strategy $f_c = 1$. To ensure an ϵ-approximate Nash equilibrium, his initial support for column 1 should be high. The incentive to deviate for the column player will be:

$$\begin{aligned}
\frac{1}{2} - \kappa - ((\frac{1}{2} + \kappa)(1-a) + (\frac{1}{2} - \kappa)a) &\le \epsilon \\
\frac{1}{2} - \kappa - (\frac{1}{2} - \frac{1}{2}a + \kappa - a\kappa + \frac{1}{2}a - a\kappa) &\le \epsilon \\
-2\kappa + 2a\kappa &\le \epsilon \\
\frac{1}{3} - \frac{1}{3}a &\le \epsilon \\
-\frac{1}{3}a &\le \epsilon - \frac{1}{3} \\
a &\ge 1 - 3\epsilon
\end{aligned}$$

If we return to the row player, the payoff matrix will look like:

$$R = \begin{array}{c} \\ \le \frac{1}{3} \\ \ge \frac{2}{3} \end{array} \begin{array}{c} \begin{array}{cc} \ge 1-3\epsilon & \le 3\epsilon \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix} \end{array}$$

A deviation to row 1 is clearly preferred. we can minimize this incentive to deviate by making the support for row 1 as large as possible, so $\frac{1}{3}$ and M as large as possible. This still gives an incentive to deviate of:

$$\begin{aligned}
(1-3\epsilon) \cdot 1 - (\frac{1}{3} \cdot (1-3\epsilon) + \frac{2}{3} \cdot 3\epsilon) &= \\
1 - 3\epsilon - \frac{1}{3} + \epsilon - 2\epsilon &= \\
\frac{2}{3} - 4\epsilon &
\end{aligned}$$

By definition, the incentive to deviate cannot be larger than ϵ:

$$\begin{aligned}
\frac{2}{3} - 4\epsilon &\le \epsilon \\
\frac{2}{3} &\le 5\epsilon \\
\frac{2}{15} &\le \epsilon
\end{aligned}$$

So for values of $\epsilon < \frac{2}{15}$, κ cannot be smaller than $-\frac{1}{6}$. This together with the first bound that $\epsilon < \frac{1}{12}$ proves theorem 12. □

The strategy of the row player in the Nash equilibrium is $f_c = \frac{M}{M+1}$, assume his strategy to be $f_c = a$. This will give the following game:

$$R = \begin{array}{c} \\ \frac{1}{2}+\kappa \\ \frac{1}{2}-\kappa \end{array} \begin{array}{c} \begin{array}{cc} a & 1-a \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix} \end{array} \quad C = \begin{array}{c} \\ \frac{1}{2}+\kappa \\ \frac{1}{2}-\kappa \end{array} \begin{array}{c} \begin{array}{cc} a & 1-a \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

The payoff of the row player will be:

$$\text{Payoff}(R) = (\frac{1}{2} + \kappa)a + (\frac{1}{2} - \kappa)(1-a)M$$

In any ϵ-approximate Nash equilibrium, the incentive to deviate should be smaller than or equal to ϵ. For a deviation to row 1, this is:

$$\begin{aligned}
a - ((\frac{1}{2} + \kappa)a + (\frac{1}{2} - \kappa)(1-a)M) &\leq \epsilon \\
a - \frac{1}{2}a - \kappa a - \frac{1}{2}M + \frac{1}{2}aM + \kappa M - \kappa aM &\leq \epsilon \\
\frac{1}{2}a - \kappa a + \frac{1}{2}aM - \kappa aM &\leq \frac{1}{2}M - \kappa M + \epsilon \\
a - 2\kappa a + aM - 2\kappa aM &\leq M - 2\kappa M + 2\epsilon \\
a(1 - 2\kappa + M - 2\kappa M) &\leq M - 2\kappa M + 2\epsilon \\
a &\leq \frac{M - 2\kappa M}{1 - 2\kappa + M - 2\kappa M} + \frac{2\epsilon}{1 - 2\kappa + M - 2\kappa M} \\
a &\leq \frac{(1-2\kappa)M}{(1-2\kappa)(M+1)} + \frac{2\epsilon}{(1-2\kappa)(M+1)} \\
a &\leq \frac{M}{M+1} + \frac{2\epsilon}{(1-2\kappa)(M+1)}
\end{aligned}$$

We would like to see $(1-2\kappa)(M+1) \geq 1$, this would mean that $\frac{2\epsilon}{(1-2\kappa)(M+1)} \leq 2\epsilon$. Since $\kappa \leq \frac{1}{6}$ and $M \geq \frac{1}{2}$ we get: $(1-2\kappa)(M+1) \geq (1 - 2 \cdot \frac{1}{6})(\frac{1}{2}+1) = 1$ This gives us a restriction of $a \leq \frac{M}{M+1} + 2\epsilon$. We can do the same kind of analysis for a deviation to row 2:

$$\begin{aligned}
(1-a)M - ((\frac{1}{2} + \kappa)a + (\frac{1}{2} - \kappa)(1-a)M) &\leq \epsilon \\
\frac{1}{2}M - \frac{1}{2}aM - \frac{1}{2}a - \kappa a + \kappa M - \kappa aM &\leq \epsilon \\
-\frac{1}{2}aM - \frac{1}{2}a - \kappa a - \kappa aM &\leq -\frac{1}{2}M - \kappa M + \epsilon \\
aM + a + 2\kappa a + 2\kappa aM &\geq M + 2\kappa M - 2\epsilon \\
a(1 + M + 2\kappa + 2\kappa M) &\geq M + 2\kappa M - 2\epsilon \\
a &\geq \frac{M + 2\kappa M}{1 + M + 2\kappa + 2\kappa M} - \frac{2\epsilon}{1 + M + 2\kappa + 2\kappa M} \\
a &\geq \frac{(1+2\kappa)M}{(1+2\kappa)(M+1)} - \frac{2\epsilon}{(1+2\kappa)(M+1)} \\
a &\geq \frac{M}{M+1} - \frac{2\epsilon}{(1+2\kappa)(M+1)}
\end{aligned}$$

We again would like to see that $(1+2\kappa)(M+1) \geq 1$. We can ensure this because $(1+2\kappa)(M+1) \geq (1 + 2 \cdot -\frac{1}{6})(\frac{1}{2}+1) = 1$

Together with the first bound on κ we found that $\frac{M}{M+1} - 2\epsilon \leq a \leq \frac{M}{M+1} + 2\epsilon$. So the strategy of the column player should be close to $f_c = \frac{M}{M+1}$, which would require the communication of the approximate value of M before a ϵ-approximate Nash equilibrium could be reached.

It is given that $\frac{1}{2} \leq M \leq 1$. In the Nash equilibrium the column player should play $\frac{M}{M+1}$, so the column player already knows that $\frac{1}{3} \leq f_c \leq \frac{1}{2}$. The strategy of the column player is still an ϵ-approximate Nash equilibrium if the strategy is 2ϵ away from the optimal solution. So the strategy of the column player should be an element of the following set: $\{\frac{1}{3}, \frac{1}{3}+4\epsilon, \frac{1}{3}+8\epsilon, \ldots, \frac{1}{2}\}$. Since only one element will lead to an ϵ-approximate Nash equilibrium, the row player will have to inform the column player which strategy is the correct one. This means the row player has to communicate $\Omega(\log 1/\epsilon)$ bits before an ϵ-approximate Nash equilibrium can be reached. □

5.2 n actions per player

We will show that for every $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium, at least $\log n$ bits of communication are required for $\epsilon > 0$.

Theorem 14. *For every $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium, at least $\log n$ bits of communication are required.*

The payoff matrix of the row player is the identity matrix:

$$\forall i, j : \quad R_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

Consider the following set of column player matrices $C^1, \ldots, C^n$ where C^k has a payoff entry of 1 for every entry of the kth column and a 0 on every other place:

$$\forall i, j : \quad C^k_{ij} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{else} \end{cases}$$

Suppose the column player has payoff matrix C^k. To reach a $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium the column player should assign a probability of at least $\frac{2}{3} + \epsilon$ to column k. The best response for the row player to this strategy is to play a pure strategy k, which would give a payoff of at least $\frac{2}{3} + \epsilon$. So any $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium should at least have a payoff of $\frac{2}{3} + \epsilon - (\frac{1}{3} - \epsilon) = \frac{1}{3} + 2\epsilon$.

The column player has to assign at least $\frac{2}{3} + \epsilon$ to column k, the remaining $\frac{1}{3} - \epsilon$ could be assigned to other columns. The payoff of the row player on the remainder of this probability is obviously at most $\frac{1}{3} - \epsilon$, which is too low for a $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium.

The row player only has one possible payoff matrix, which is known to everybody. So we will look at communication from the column player to the row player. Information on the strategy of the column player can be useful for the row player, mainly which column is played with a probability of at least $\frac{2}{3} + \epsilon$ by the column player. The mixed strategy of the column player will only be a function of his own payoff matrix, while the mixed strategy $\mathbf{x}$ of the row player will be a function of his own payoff matrix and the bit string b he receives from the column player:

$$\begin{aligned} f_c :& C \to \Delta_n \\ f_r :& (R, b) \to \Delta_n \end{aligned}$$

Given any deterministic communication protocol CO, all communication depends on the payoff matrix of the player and the communication so far, assume that less than $\lceil \log n \rceil$ bits of communication have been sent, say $\lceil \log n \rceil - 1$ bits. Without knowledge of which column is column k the row player cannot lower his regret, so in the optimal scenario all bits of communication have been used to identify this column.

Since the column player has n columns, the protocol CO causes the column player to send the same bit sequence b, for two distinct column matrices C_k and C_l. The mixed strategy of the row player depends on the bit string b and his own payoff matrix R. The bit sequence is the same for C_k and C_l, so the strategy of the row player will also be the same.

The payoff of the row player should at least be $\frac{1}{3} + 2\epsilon$. If the column player has payoff matrix C^k, k is played with at least $\frac{2}{3} + \epsilon$ probability. To reach a payoff of at leat $\frac{1}{3} + 2\epsilon$, the row player has to assign a certain minimum probability to row k. This probability on row k by the row player should be:

$$\mathbf{x}[k] \geq \frac{\frac{1}{3} + 2\epsilon}{\frac{2}{3} + \epsilon} > 0.5$$

Any strategy that is a $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium for the row player when the column player has payoff matrix C^k should assign a probability higher than 0.5 to row k. Any strategy that is a $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium for the row player when the column player has payoff C^l should assign a probability higher than 0.5 to row l. Given b the strategy for the row player is the same for C^k and C^l. To guarantee a $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium, $\mathbf{x}[k] > 0.5$ and $\mathbf{x}[l] > 0.5$ should hold. This cannot hold, since $\mathbf{x}$ is a probability distribution. Therefore no $(\frac{1}{3} - \epsilon)$-approximate Nash equilibrium can be reached with less than $\lceil \log n \rceil$ bits of communication. $\square$

Chapter 6

Lower bounds on the approximation of Nash equilibria

For $n \times n$ games we can get lower bounds on an approximate Nash equilibrium if we fix the amount of communication allowed.

In a previous chapter a bound on $\frac{1}{3}$ was shown for 3×3-games when no communication is allowed. One could wonder whether the approach that is used to show lower bounds in this case can be used to show even better lower bounds for larger games. Unfortunately this is not the case, the set of payoff matrices that is used allows no lower bound higher than $\frac{1}{3}$. With a different analysis, better lower bounds can be proven.

First a lower bound is proven for a model where no communication is allowed. This lower bound gives $\epsilon > 0.5$.

For one-way communication the lower bound that is achieved is $0.5-o(\frac{1}{\sqrt{n}})$. The DMP-algorithm can be implemented as an algorithm with one-way communication and gives a 0.5-approximate Nash equilibrium. Therefore the lower bound we achieve is strict. For well-supported Nash equilibria, the lower bound for one-way communcation is a 1-WSNE. This shows that this stricter notion of approximation needs more information to give any non-trivial result.

6.1 A lower bound larger than 0.5 without communication

In this section we will show a lower bound strictly larger than 0.5 for models with no communication, to show that 0.5 is not the answer we are looking for.

We will first introduce a commitment measure τ that measures the amount of freedom a player has with choosing his strategy.

The distance between two probability distributions $\mathbf{x}$ and $\mathbf{y}$ is the sum of all positive differences between the two distributions, also known as the variation distance.

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

Let Ω^{R} denote the set of strategies the row player may use and Ω^{C} the set of strategies the column player may use. For each player we define his center strategy. For the row player the strategy $\mathbf{c}^{\mathrm{R}}$ is the probability distribution such that the maximum distance between $\mathbf{c}^{\mathrm{R}}$ and any strategy $\boldsymbol{\omega} \in \Omega^{\mathrm{R}}$ is minimised.

$$\mathbf{c}^{\mathrm{R}} = \arg\min_{\mathbf{c}^{\mathrm{R}}} \sup_{\boldsymbol{\omega} \in \Omega^{\mathrm{R}}} d(\mathbf{c}^{\mathrm{R}}, \boldsymbol{\omega})$$

The center distribution $\mathbf{c}^{\mathrm{C}}$ of the column player is defined in a similar way. The commitment τ^{R} of the row player is defined as:

$$\tau^{\mathrm{R}} = 1 - (\sup_{\boldsymbol{\omega} \in \Omega^{\mathrm{R}}} d(\mathbf{c}^{\mathrm{R}}, \boldsymbol{\omega}))$$

The commitment τ^{C} of the column player is defined similarly. This commitment measure τ^{R} will be a value in $[0, 1]$ that indicates the spread of strategies a player considers.

Theorem 15. *In models with no communication, the players cannot improve over a 0.501-approximate Nash equilibrium.*

The proof will be a case analysis on commitment. It will prove a lower bound of $0.5 + z$ with $z = 0.001$. We will show that for all commitment values, the regret of a player is at least 0.501. We identify three cases:

- A player has a high commitment: $\tau^{\mathrm{R}} \geq 0.501$ or $\tau^{\mathrm{C}} \geq 0.501$
- A player has a low commitment: $\tau^{\mathrm{R}} \leq 0.05$ or $\tau^{\mathrm{C}} \leq 0.05$
- Both players have intermediate commmitment: $0.05 < \tau^{\mathrm{R}} < 0.501$ and $0.05 < \tau^{\mathrm{C}} < 0.501$

6.1.1 Case 1: A player has high commitment

Let $\tau^{\mathrm{R}} \geq 0.501$. Consider the following set of row player payoff matrices $R^1, \ldots, R^n$ where R^k has a payoff of 1 for every entry in the kth row and a 0 on every other place:

$$\forall i, j: \quad R^k_{ij} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{else} \end{cases}$$

The row player has a center distribution $\mathbf{c}^{\mathrm{R}}$ as defined before. Let k be the smallest entry in the vector $\mathbf{c}$:

$$k = \min_{i=n} \mathbf{c}^{\mathrm{R}}[i]$$

Since $\mathbf{c}^{\mathrm{R}}$ is a probability distribution (sums to 1), $k \leq \frac{1}{n}$. Let the row player have payoff matrix R^k. Since $\tau^{\mathrm{R}} \geq 0.501$, the distance to any strategy in Ω^{R} is at most 0.499. Let $\mathbf{x} \in \Omega^{\mathrm{R}}$ be the strategy that assigns the highest probability to row k. It holds that $d(\mathbf{x}, \mathbf{c}^{\mathrm{R}}) \leq 0.499$, so it also holds that $\mathbf{x}[k] - \mathbf{c}^{\mathrm{R}}[k] \leq 0.499$. Since $\mathbf{c}^{\mathrm{R}}[k] \leq \frac{1}{n}$ it holds that $\mathbf{x}[k] \leq 0.499 + \frac{1}{n}$.

Row k is the only row that yields a positive payoff for the row player and this row is played with a probability of at most $0.499 + \frac{1}{n}$. The best response for the row player obviously has a payoff of 1. This gives a regret of at least $1 - (0.499 + \frac{1}{n}) = 0.501 - \frac{1}{n}$.

6.1.2 Case 2: A player has low commitment

Let $\tau^{\mathrm{C}} \leq 0.05$. This low commitment implies that the column player has at least a few strategies that are very different.

For the column player, take an arbitrary strategy $\mathbf{s}_1 \in \Omega^{\mathrm{C}}$. Because $\tau^{\mathrm{C}} \leq 0.05$, there must be some strategy $\mathbf{s}_2$ with $d(\mathbf{s}_1, \mathbf{s}_2) \geq 0.95$, else $\mathbf{s}_1$ could be the center strategy $\mathbf{c}$ with $\tau^{\mathrm{C}} \geq 0.05$.

Now consider the strategy $\mathbf{s}_{12} = \frac{\mathbf{s}_1 + \mathbf{s}_2}{2}$, which is of course closer to $\mathbf{s}_1$ and $\mathbf{s}_2$. For this strategy to not be the center strategy $\mathbf{c}$ there must be some strategy $\mathbf{s}_3 \in \Omega^{\mathrm{C}}$ with $d(\mathbf{s}_{12}, \mathbf{s}_3) \geq 0.95$. Because $\mathbf{s}_1$ is half of the strategy $\mathbf{s}_{12}$, it holds that $d(\mathbf{s}_1, \mathbf{s}_3) \geq 0.90$ and $d(\mathbf{s}_2, \mathbf{s}_3) \geq 0.90$.

$$\begin{aligned} d(\mathbf{s}_1, \mathbf{s}_2) &\geq 0.95 \\ d(\mathbf{s}_1, \mathbf{s}_3) &\geq 0.90 \\ d(\mathbf{s}_2, \mathbf{s}_3) &\geq 0.90 \\ d(\mathbf{s}_{12}, \mathbf{s}_3) &\geq 0.95 \end{aligned}$$

The next step is to construct a payoff matrix R of the row player. This payoff matrix R will be a $(3 \times n)$-matrix, the row player will only have 3 pure strategies. The construction of the rows will be such that row i is a best response to s_i and a bad 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 $R_{2j} = 0$, $R_{3j} = 0$. If $\mathbf{s}_2[j]$ is the largest, $R_{2j} = 1$ and $R_{1j} = 0$, $R_{3j} = 0$. If $\mathbf{s}_3[j]$ is the largest, $R_{3j} = 1$ and $R_{1j} = 0$, $R_{2j} = 0$. 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 a column i such that $\mathbf{s}_1[i] > 0$, but $R_{1i} = 0$ and $R_{2i} = 1$ so $\mathbf{s}_2[i] > \mathbf{s}_1[i]$. This could hold for a number of columns of R, but the total probability assigned by $\mathbf{s}_1$ to these columns is bounded by 0.05. If the probability on these columns was higher than 0.05, it would hold that $d(\mathbf{s}_1, \mathbf{s}_2) < 0.95$. Similarly we can bound the probability assigned by $\mathbf{s}_1$ to columns such that $\mathbf{s}_1[j] > 0$ with $R_{1i} = 0$ and $R_{3i} = 1$. Since $d(\mathbf{s}_1, \mathbf{s}_3) \geq 0.9$ this probability at most 0.1. This leads to at most 0.15 of the probability of $\mathbf{s}_1$ that could give a payoff of 0 for row 1. The remaining 0.85 probability of $\mathbf{s}_1$ will have a 1 on the corresponding columns in 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.

For row 3 we use $d(\mathbf{s}_{12}, \mathbf{s}_3) \geq 0.95$. Consider a column i such that $\mathbf{s}_3[i] > 0$, but $R_{3i} = 0$ and $R_{1i} = 1$ or $R_{2i} = 1$. If $\mathbf{s}_{12}[i] > \mathbf{s}_3[i]$ it will also hold that $\mathbf{s}_1[i] > \mathbf{s}_3[i] \vee \mathbf{s}_2[i] > \mathbf{s}_3[i]$. In that case the total probability on these columns of $\mathbf{s}_3$ is at most 0.05. On the other hand there could be a column j such that $\mathbf{s}_{12}[j] \leq \mathbf{s}_3[j]$ but $\mathbf{s}_1[j] > \mathbf{s}_3[j] \vee \mathbf{s}_2[j] > \mathbf{s}_3[j]$, this would also give a loss in the payoff matrix R in row 3. This situation implies that $\mathbf{s}_3[j] > 0$ and it would add $\mathbf{s}_3[j]$ to the commitment of the column player. Since his commitment is at most 0.05, at most 0.05 of the probability of $\mathbf{s}_3$ could be on these columns. Because $\mathbf{s}_{12} = \frac{1}{2}\mathbf{s}_1 + \frac{1}{2}\mathbf{s}_2$ this probability could be at most 0.10 on $\mathbf{s}_1$ and $\mathbf{s}_2$ combined, at most 0.10 of the probability of $\mathbf{s}_3$ could be on columns such that $\mathbf{s}_3[i] > 0$, $R_{3i} = 0$ and $R_{1i} = 1$ or $R_{2i} = 1$. The payoff for row 3 if the column player plays $\mathbf{s}_3$ is therefore at least 0.9.

- 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 could 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 could 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 could give him a payoff of at most 0.1 and playing row 1 a payoff of at most 0.1. Next to this the sum of payoffs of $\mathbf{s}_1$ and $\mathbf{s}_2$ is at most 0.1.

The strategy of the row player is $f_r = (r_1, r_2, r_3)$. Assume $r_1 \leq r_2, r_3$, so $r_1 \leq \frac{1}{3}$ and the column player plays strategy $\mathbf{s}_1$. The best response strategy $f_r = (1, 0, 0)$ has a payoff of $a \in [0.85, 1]$. Because row 1 clearly gives the highest payoff, the regret is minimised when this row is played with as much probability as possible, so $\frac{1}{3}$. Because the probability on row 1 was defined as the lowest probability, the probability on the other to rows is also $\frac{1}{3}$. This gives a regret of at least:

$$\begin{aligned} a - (\frac{1}{3}a + \frac{1}{3} \cdot 0.05 + \frac{1}{3} \cdot 0.1) &= \frac{2}{3}a - 0.05 \\ &\geq \frac{2}{3} \cdot 0.85 - 0.05 \\ &\approx 0.517 \end{aligned}$$

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

Assume $r_3 \leq r_1, r_2$ 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:

$$\begin{aligned} a - (\frac{1}{3}a + \frac{1}{3} \cdot 0.1) &= \frac{2}{3}a - \frac{1}{30} \\ &\geq \frac{2}{3} \cdot 0.9 - \frac{1}{30} \\ &\approx 0.567 \end{aligned}$$

So regardless the strategy of the row player, the regret of the row player is always larger than 0.501 when the commitment of the other player is at most 0.05.

6.1.3 Case 3: Both players have intermediate commitment

Both players have intermediate commitment: $\tau^{\mathrm{R}}, \tau^{\mathrm{C}} \in [0.05, 0.501]$. Assume $\tau^{\mathrm{R}} \geq \tau^{\mathrm{C}}$.

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 lth column and a 0 on every other place:

$$\forall i, j : \quad C^l_{ij} = \begin{cases} 1 & \text{if } j = l \\ 0 & \text{else} \end{cases}$$

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

The construction of the payoff matrix R of the row player will depend on the center strategy $\mathbf{c}^{\mathrm{R}}$ of the row player. Take the $(n - \sqrt{n})$ rows of R which have the highest value of $\mathbf{c}^{\mathrm{R}}$ corresponding to these rows. These rows will all give a payoff of 0 in R. For the construction of the remaining $\sqrt{n}$ rows of R we look at $\mathbf{c}^{\mathrm{C}}$, the center distribution of the column player. We select the $(n - \sqrt{n})$ columns of R with the highest value in $\mathbf{c}^{\mathrm{C}}$. If row i is one of the rows with one of the $\sqrt{n}$ smallest entries for $\mathbf{c}^{\mathrm{R}}$ and column j is a column with one of the $(n-\sqrt{n})$ highest entries for $\mathbf{c}^{\mathrm{C}}$, then $R_{ij}=1$. The payoff entries in R that are still open can be seen as a $(\sqrt{n}\times\sqrt{n})$-sub-matrix. This sub-matrix will be constructed such that each row and every column of this sub-matrix has exactly one payoff entry of 1 and all other payoffs are 0, a permutation of the identity matrix. When we rearrange the rows and columns of R on descending value of $\mathbf{c}^{\mathrm{R}}$ and $\mathbf{c}^{\mathrm{C}}$, the payoff would look like Table 6.1.

Let the column player have payoff matrix C^l such that $\mathbf{c}^{\mathrm{C}}[l]$ is one of the $\sqrt{n}$ smallest entries of $\mathbf{c}^{\mathrm{C}}$.

The sum of the $\sqrt{n}$ lowest entries of $\mathbf{c}^{\mathrm{R}}$ is at most $\frac{\sqrt{n}}{n}=\frac{1}{\sqrt{n}}$. Since $\tau^{\mathrm{R}}>0.05$ and $d(\mathbf{c}^{\mathrm{R}},\boldsymbol{\omega})<0.95$ for all $\boldsymbol{\omega}\in\Omega^{\mathrm{R}}$, the total probability on these rows given by any strategy $\boldsymbol{\omega}$ is at most $0.95+\frac{1}{\sqrt{n}}$. This gives at least a probability of $0.05-\frac{1}{\sqrt{n}}$ to the remaining $(n-\sqrt{n})$ rows. A similar argument holds for the column player. The total probability on the columns that correspond with the $\sqrt{n}$ lowest entries of $\mathbf{c}^{\mathrm{C}}$ will be played with at most $0.95+\frac{1}{\sqrt{n}}$ probability by any $\boldsymbol{\omega}\in\Omega^{\mathrm{C}}$. The remaining $(n-\sqrt{n})$ columns are played with a probability of at least $0.05-\frac{1}{\sqrt{n}}$ by any strategy of the column player.

The $(n-\sqrt{n})$ rows with highest $\mathbf{c}^{\mathrm{R}}$ values will give a payoff of 0 in R. The payoff of the remaining $\sqrt{n}$ rows can be divided into two parts: the part that corresponds with the $(n-\sqrt{n})$ columns with highest $\mathbf{c}^{\mathrm{C}}$ values and the remaining $\sqrt{n}$ columns. The part corresponding to the $(n-\sqrt{n})$ columns with highest $\mathbf{c}^{\mathrm{C}}$ is played by the column player with a probability of at least $0.05-\frac{1}{\sqrt{n}}$ and by the row player with at most $0.95+\frac{1}{\sqrt{n}}$. This part of R is filled with 1's and will give a payoff of $(0.95+\frac{1}{\sqrt{n}})\cdot(0.05-\frac{1}{\sqrt{n}})\cdot 1=0.0475-\frac{0.9}{\sqrt{n}}-\frac{1}{n}$. The second part ($\sqrt{n}$ rows and $\sqrt{n}$ columns) is played by both players with at most $0.95+\frac{1}{\sqrt{n}}$ probability. One of the columns in this part, column l, is played with probability of at least 0.499 by the column player. The value of $\mathbf{c}^{\mathrm{C}}$ is very low for these $\sqrt{n}$ columns, the row player is uninformed about which columns are played with positive probability. Because this part of the matrix was a permutation of the identity matrix, the payoff for this part will be $o(\frac{1}{n})$. This gives a total payoff of $0.0475-\frac{0.9}{\sqrt{n}}-\frac{1}{n}+o(\frac{1}{n})$ for the row player.

The pure strategy best response of the row player will be a row i such that $R_{il}=1$, since column l is played with at least 0.499 probability. The payoff of this best response will be at least $(0.05-\frac{1}{\sqrt{n}})+0.499=0.549-\frac{1}{\sqrt{n}}$. This leads to a regret of at least:

$$\begin{aligned}(0.549-\frac{1}{\sqrt{n}})-(0.0475-\frac{0.9}{\sqrt{n}}-\frac{1}{n}+o(\frac{1}{n})) &= 0.5015-\frac{1}{\sqrt{n}}+\frac{0.9}{\sqrt{n}}+\frac{1}{n}-o(\frac{1}{n})\\ &= 0.5015-(\frac{0.1}{\sqrt{n}}-\frac{1}{n}+o(\frac{1}{n}))\end{aligned}$$

For large values of n it holds that $\frac{0.1}{\sqrt{n}}-\frac{1}{n}+o(\frac{1}{n})<0.0005$. This gives a regret of at least 0.501. □

$$R = \begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \end{matrix}$$

Table 6.1: R for $n = 9$ after rearranging the rows and columns on descending value of $\mathbf{c}^{\mathrm{R}}$ and $\mathbf{c}^{\mathrm{C}}$

6.2 A $\frac{1}{2} - o(\frac{1}{\sqrt{n}})$ bound for one-way communication

In a one-way communication model, an unlimited amount of communication is allowed from player 1 to player 2, but no communication is allowed from player 2 to player 1. A simple upper bound for this model is the $\frac{1}{2}$-approximate Nash equilibrium that is given by the simple DMP-algorithm. This DMP-algorithm requires only $\log n$ bits of communication. We will show that is algorithm is optimal for models with one-way communication by giving a $\frac{1}{2} - o(\frac{1}{\sqrt{n}})$ lower bound.

To prove this lower bound we will look at a model where there is only communication allowed from the row player to the column player.

Theorem 16. *In models with one-way communication, the players cannot improve over a $(0.5 - o(\frac{1}{\sqrt{n}}))$-approximate Nash equilibrium.*

We define a game $G = (R, C)$, where R and C are payoff matrices with dimensions $\binom{n}{k} \times n$, with $k \ll n$. Consider the following set of column player payoff matrices $C^1, \ldots, C^n$, where C^l has a payoff of 1 for every entry in the lth column and a 0 on every other place:

$$\forall i, j: \quad C^l_{ij} = \begin{cases} 1 & \text{if } j = l \\ 0 & \text{else} \end{cases}$$

The row player has 1 matrix with $\binom{n}{k}$ rows, where a row consists of k 1's and $(n-k)$ 0's. Every row is a different combination, so the $\binom{n}{k}$ rows are all distinct combinations of k 1's in a row of length n. For example, the payoff matrix of the row player witn $n = 5$ and $k = 2$ will look like:

$$R = \begin{matrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{matrix}$$

The strategy of the row player is given by the probability distribution D^r. Since no communication is allowed from the column player to the row player, the strategy of the row player can only depend on his payoff matrix. Unlimited communication from the row player to the column player is allowed, so we can assume the column player knows everything of the row player. Since the strategy of the row player only depends on his payoff matrix, the strategy of the column player will only have to depend on his own payoff matrix and the payoff matrix of the row player. For every possible payoff matrix of the column player, he will have a different strategy:

$$D^r : R \to \Delta_{\binom{n}{k}}$$
$$D^c_l : (C^l, R) \to \Delta_n$$

We will show that for this class of games, you cannot improve over a $(\frac{1}{2} - o(\frac{1}{\sqrt{n}}))$-approximate Nash equilibrium. This implies for large values of n approximately a $\frac{1}{2}$-approximate Nash equilibrium.

The idea behind the proof is that the information of the row player will not have a big influence on the strategy of the column player, he needs to assign a probability of at least $\frac{1}{2} + o(\frac{1}{\sqrt{n}})$ to a specific column. The strategy of the row player on the other hand is only defined by his own payoff matrix, which is insufficient to reach a good approximation.

During the proof we will search for a lower bound of $\frac{1}{2} - z$, where the value of z still has to be determined.

First observe that a best response for the column player is $D^c_l(C^l) = (0, \ldots, 0, 1, 0, \ldots, 0)$, where the 1 is the lth entry. This corresponds to playing the pure strategy of column l, when the column player has C^l. So to reach a $(\frac{1}{2} - z)$-approximate Nash equilibrium, D^c_l should have on the lth entry at least a probability of $(\frac{1}{2} + z)$.

The row player has one matrix R with all different combinations of k 1's in a row of length n. Another way we can look at this matrix is looking at the columns of this matrix. By definition each column of this matrix consists of $\frac{k}{n} \cdot \binom{n}{k}$ 1's and $(1 - \frac{k}{n}) \cdot \binom{n}{k}$ 0's.

D^r assigns a probability to each row of R. Because every column of R has the same number of 1's, we can also look at D^r as an unnormalised distribution Φ (summing to k) over the columns. This distribution Φ assigns to each column j a value $\Phi(j)$, which gives the probability that a 1 will be in this column given D^r. This value $\Phi(j)$ will be at most 1, when every row that is played with positive probability has a 1 for that column number. Because every row consists of k 1's, the sum of over all values will sum to k:

$$\sum_{j=1}^{n} \Phi(j) = k$$

We define the column with the lowest value $\Phi(j)$ to be m:

$$m = \arg\min_{j} \Phi(j)$$

Consider the column player has payoff matrix C^m. Remember that the sum over all these values $\Phi(j)$ is k and there are n columns, so m is at most $\frac{k}{n}$. This means that column m, which is played at least $\frac{1}{2} + z$ of the time by the column player, gives a payoff of 0 at least with a probability of at least $1 - \frac{k}{n}$.

Consider the difference in payoff between the response mentioned above dan an improved response D^*. The improved response distribution D^* will differ from D^r in the following way. For every row we will look if there is a 1 on the mth entry. If this is the case, we do not change anything. If there is a 0 on the mth entry we do the following: We look at the positions where there is a 1 in this row. Of all the entries where there is a 1, we select the entry to which the column player gives the lowest probability, entry a. Now we move all the probability from this row to the row that has a 0 on entry a and a 1 on entry m.

The probability on entry a is defined as the smallest of all the entries where this row has a 1. We can bound the probability that was given to this entry by the column player. A probability of at least $\frac{1}{2} + z$ is given to column m, so a probability of $\frac{1}{2} - z$ can be distributed over the remaining columns. The column belonging to entry a has the smallest probability of at least k columns, so the probability given to column a is at most: $\frac{1/2-z}{k}$.

The result of this change in the probability distribution is that every row that is played with positive probability will have a 1 on the mth entry. There is a probability at least $(1-\frac{k}{n})$ that a row sampled from D^r didn't have a 1 on the mth entry. This means that the increase in payoff to the improved response D^* is at least:

$$(1-\frac{k}{n})\cdot(\frac{1}{2}+z)-(1-\frac{k}{n})\cdot\frac{1/2-z}{k}=(1-\frac{k}{n})\cdot(\frac{1}{2}+z-\frac{\frac{1}{2}-z}{k})$$

We will show that this increase in payoff is close to $\frac{1}{2}$ for well chosen k and z. Assume that z is chosen such that $z=\frac{\frac{1}{2}-z}{k}$:

$$z=\frac{\frac{1}{2}-z}{k}$$
$$z=\frac{1}{2k+2}$$

This will make the gap:

$$(1-\frac{k}{n})\cdot(\frac{1}{2}+z-z)=\frac{1}{2}-\frac{k}{2n}$$

The row player has a gap of $\frac{1}{2}-\frac{k}{2n}$ and the column player has a gap of $\frac{1}{2}-z$, with $z=\frac{1}{2k+2}$ We can use these two observations to find the value of k such that the increase in payoff is the same for the row player and column player:

$$\frac{1}{2}-\frac{k}{2n}=\frac{1}{2}-\frac{1}{2(k+1)}$$
$$\frac{k}{2n}=\frac{1}{2(k+1)}$$
$$k=\frac{1}{2}(\sqrt{4n+1}-1)\vee k=\frac{1}{2}(-\sqrt{4n+1}-1)$$

Since k should be greater than 0, only the first solution is feasible. So we have $k=\frac{1}{2}(\sqrt{4n+1}-1)$ and $z=\frac{\frac{1}{2}(\sqrt{4n+1}-1)}{2n}$, which is $o(\frac{1}{\sqrt{n}})$. We have proven now that for general games with one-way communication you cannot do better than a $(\frac{1}{2}-o(\frac{1}{\sqrt{n}}))$-approximate Nash equilibrium. □

6.3 A 1-WSNE bound for one-way communication

For ϵ-WSNE the bound for models with no communication and one-way communication is the same, no non-trivial bound on ϵ can be given. The prove given is for models with one-way communication, but obviously also holds for models with no communication.

Theorem 17. *In models with one-way communication, the players cannot improve over a 1-WSNE.*

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.

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

It is easy to see that if the column player wants to improve over a 1-WSNE and he has matrix C^1, he should play $f_c = 1$. If the column player has payoff matrix C^2, he should play $f_c = 0$ to avoid a 1-WSNE. Although communication is allowed from the row player, the strategy of the column player is already defined by his payoff matrix alone.

No communication is allowed from the column player to the row player, the strategy of the row player can only depend on his payoff matrix: $f_r : R \to \mathbb{R}$. Assume the row player assigns positive probability to row 1 and the column player has payoff matrix C^2. Since the column player plays a pure strategy, the row player assigns positive probability to a strategy that gives a payoff of 0, while row 2 guarantees a payoff of 1. This leads to 1-WSNE. If the row player assigns positive probability to row 2 and the column player has payoff matrix C^1, the row player will also have a 1-WSNE. Since the strategy of the column player does not depend on information of the column player, the row player cannot improve over a 1-WSNE. $\square$

Chapter 7

Upper bounds on the approximation of Nash equilibria

Depending on the amount of communication that is allowed, we get different approximations of a Nash equilibrium. Almost all known algorithms that prove upper bounds on the approximation have no restrictions on the communication complexity, except for the simple DMP-algorithm [7] which has a logarithmic communication complexity. It can be seen that if more communication is allowed, a better approximation is possible.

7.1 A $\frac{3}{4}$-approximate Nash equilibrium with no communication

When no communication is allowed we could imagine the following procedure: The row player picks row 1. Let $j \in \arg\max_{j'} C_{1j'}$. Let the column player pick column 1. Now let $i \in \arg\max_{i'} R_{i'1}$. So, j is a best-response column for the column player to row 1 and i is a best-response row for the row player to column 1. The approximate Nash equilibrium will be $\mathbf{x}^* = \frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_i$ and $\mathbf{y}^* = \frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_j$.

Theorem 18. *The strategy pair* $(\mathbf{x}^*, \mathbf{y}^*)$ *is a* $\frac{3}{4}$*-approximate Nash equilibrium.*

Proof: Let i' be a best pure strategy response of the row player to $\mathbf{y}^*$. Then his incentive to deviate is:

$$\begin{aligned}(\frac{1}{2}R_{i'1} + \frac{1}{2}R_{i'j}) - (\frac{1}{4}R_{11} + \frac{1}{4}R_{1j} + \frac{1}{4}R_{i1} + \frac{1}{4}R_{ij}) &\leq (\frac{1}{4}R_{i'1} + \frac{1}{2}R_{i'j}) - (\frac{1}{4}R_{11} + \frac{1}{4}R_{1j} + \frac{1}{4}R_{ij}) \\ &\leq \frac{1}{4}R_{i'1} + \frac{1}{2}R_{i'j} \\ &\leq \frac{1}{4} + \frac{1}{2} = \frac{3}{4}\end{aligned}$$

Where the first equality holds because i was a best response to column 1 and the second inequality holds because all payoffs are in $[0, 1]$. The same kind of argument holds for the column player. This proves the theorem. □

7.1.1 A $(\frac{3}{4} - \frac{1}{4n})$-approximate Nash equilibrium

Instead of playing row 1 with probability $\frac{1}{2}$, the row player assigns to each row a probability of $\frac{1}{2n}$ and assigns a probality of $\frac{1}{2}$ to a best response to a fully mixed strategy of the column player. The column player plays a similar strategy.

Like the previous algorithm, with probability $\frac{1}{2}$ a pure strategy best response is played. The difference with the previous algorithm is that we can guarantee that this pure best response is played with a probability higher than $\frac{1}{2}$. The remainder of the probability (also $\frac{1}{2}$) is fully mixed over all strategies, this leads to a probability of $\frac{1}{2n}$ on each row. This implies that the pure strategy best response also gets this $\frac{1}{2n}$ probability. Together with the $\frac{1}{2}$ probability, this leads to a total probability on a best response strategy of $\frac{1}{2} + \frac{1}{2n}$. This strategy is a best response to the fully mixed strategy of the other player, which is played with probability $\frac{1}{2}$. This implies that a best response is played with a probability of at least $(\frac{1}{2} + \frac{1}{2n}) \cdot \frac{1}{2} = \frac{1}{4} + \frac{1}{4n}$. This leads to a $(\frac{3}{4} - \frac{1}{4n})$-approximate Nash equilibrium. The same kind of argument holds for the column player.

7.2 A 0.438-approximate Nash equilibrium

This section will provide a 0.438-approximate Nash equilibrium where the communication is restricted to both players communicating a mixed strategy to the other player. The algorithm will use some properties of zero-sum games to get the approximation. We will try to find a α-approximate Nash equilibrium first and later on we will find the optimal value for α.

Theorem 19. *When the row player and the column player are allowed to communicate one mixed strategy profile, a 0.438-approximate Nash equilibrium can be found.*

First the row player will find 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 games, we know that the payoff value for their Nash equilibrium will be unique and can be computed in polynomial time. Both players compare this payoff value with α. We distinguish two cases, the payoff value at the Nash equilibrium of both players is lower than α (Case 1) or at least one of the players has a value equal to or higher than α for a Nash equilibrium (Case 2). With $O(1)$ communication, the case that holds can be identified.

Case 1:
First consider the case where both players have a Nash equilibrium with value smaller than α. The row player has a Nash equilibrium strategy pair $(\mathbf{x}^*_R, \mathbf{y}^*_R)$ and the column player a Nash equilibrium 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 $(\mathbf{x}^*_C, \mathbf{y}^*_R)$.

Since $\mathbf{y}^*_R$ was a Nash equilibrium strategy in the zero-sum game $(R, -R)$ and the row player still plays with payoff matrix R, by definition of a Nash equilibrium, 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. This leads to an α-approximate Nash equilibrium for the row player. 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 argue that when the row player has strategy $\mathbf{x}^*_C$, the column player has a α-approximate Nash equilibrium. This concludes case 1.

Case 2:
If at least one of the players has a value of at least α for the zero-sum game, he can get a payoff of at least α if he plays this strategy regardless the strategy of the other player. Assume w.l.o.g. that the row player has a payoff of at least α in his zero-sum game. He communicates this strategy $\mathbf{x}_R^*$ to the column player. The column player will compute a pure strategy best response $\mathbf{e}_j$ to the strategy of the row player and communicate this strategy to the row player.

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 strategy is a 0-approximate Nash equilibrium. The row player can guarantee a payoff of α. His best response to $\mathbf{e}_j$ could have a value of β, with $\beta \leq 1$. At this point the row player has a $(\beta - \alpha)$-approximate Nash equilibrium. We are searching for an $\alpha < \frac{1}{2}$ and β could be as high as 1. We next deal with the possibility that $\beta - \alpha > \alpha$.

The column player has a 0-approximate Nash equilibrium, while we are only looking for a α-approximate Nash equilibrium and the row player has a strategy that could be not good enough for a α-approximate Nash equilibrium. To change this, we use a method first used in [3], 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 allowed, as long as the regret while playing $\mathbf{e}_j$ is at most α. Consider the row player shifting $\frac{1}{2}\alpha$ of his probability to a best response strategy. The payoff the column player gets could be $\frac{1}{2}\alpha$ lower because of this move. The payoff of another pure strategy could be as much as $\frac{1}{2}\alpha$ higher because of this shift. The strategy $\mathbf{e}_j$ was a 0-approximate Nash equilibrium, so by the shift of $\frac{1}{2}\alpha$ of the row players' probability, the regret of the column player is at most $\frac{1}{2}\alpha + \frac{1}{2}\alpha = \alpha$, which leads to a α-approximate Nash equilibrium, which is allowed for the column player.

The row player is allowed to change the allocation of $\frac{1}{2}\alpha$ of his probability allocated to strategies with the worst payoff. Since we rearrange the probability that was allocated to the rows with the lowest payoffs, the remainder of his probability, $1 - \frac{1}{2}\alpha$ has a payoff of at least α. The probability is shifted to his best response with a value of β, with $\alpha \leq \beta \leq 1$. A regret of α will give the row player an α-approximate Nash equilibrium, a payoff value of $\beta - \alpha$ is high enough to give this guarantee. This leads to the following inequality:

$$(1 - \frac{1}{2}\alpha)\alpha + \frac{1}{2}\alpha\beta \geq \beta - \alpha\ , \quad 0 \leq \alpha \leq \beta \leq 1$$

The solutions to this inequality are:

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

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 $0 \leq \alpha \leq 1$ 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 $\alpha \leq \beta \leq 1$. 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. The strategy pair belonging to ths 0.438-approximate Nash equilibrium is $(0.781\mathbf{x}_{\mathrm{R}}^* + 0.219\mathbf{e}_i, \mathbf{e}_j)$, where $\mathbf{e}_i$ is a best response strategy for the row player to $\mathbf{e}_j$. □

7.2.1 Lowering the communication complexity

We can improve the communication complexity to $O(\log^2 n)$ if we send an approximation of the strategies. Instead of looking at a minimal payoff of 0.438, we search for a payoff of at least $0.438 + \delta$, for any $\delta > 0$. By approximating the strategies, the upper bound is weakened by 2δ.

Theorem 20. *A $(0.438+2\delta)$-approximate Nash equilibrium can be found with a communication complexity of $O(\log^2 n)$.*

Like the previous algorithm we distinguish two cases: the case where a player has a high payoff for his zero-sum game and the case where neither playe has a high payoff for his zero-sum game.

A player has a high payoff for his zero-sum game
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 a Nash equilibrium of the zero-sum game. The column player will play a pure best response to the strategy of the row player, regardless the support of the strategy of the row player. So we will mainly look at the row player.

The zero-sum game $(R, -R)$ gives a strategy pair $(\mathbf{x}_{\mathrm{R}}^*, \mathbf{y}_{\mathrm{R}}^*)$. 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 $\mathbf{x}_{\mathrm{R}}^*$. Let $\mathbf{x}_{\mathrm{R}}'$ be the mixed strategy for the row player with a probability of $\frac{1}{k}$ for every member of A. We want the distribution $\mathbf{x}_{\mathrm{R}}'$ to have a payoff close to the payoff of $\mathbf{x}_{\mathrm{R}}^*$. This corresponds to the following event:

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

As mentioned in [14] the expression $((\mathbf{x}_{\mathrm{R}}')^T R \mathbf{y}_{\mathrm{R}}^*)$ is essentialy a sum of k independent random variables each of expected value $((\mathbf{x}_{\mathrm{R}}^*)^T R \mathbf{y}_{\mathrm{R}}^*)$, where every random variable has a value in $[0, 1]$. This means we can bound the probability that ϕ does not hold, which we will call ϕ^c. When we apply a standard tail inequality [11] to bound the probability of ϕ^c, we get:

$$Pr[\phi^c] \leq e^{-2k\delta^2}$$

With $k = \frac{\ln n}{\delta^2}$, this gives $Pr[\phi^c] \leq \frac{1}{n^2}$ and $Pr[\phi] \geq 1 - \frac{1}{n^2}$. If $\mathbf{x}_{\mathrm{R}}'$ is a wrong distribution, we sample again.

The strategy $\mathbf{x}_{\mathrm{R}}'$ 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 $\mathbf{x}_{\mathrm{R}}'$ of $O(\log^2 n)$.

The column player computes a pure strategy best response to $\mathbf{x}_{\mathrm{R}}'$ 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 to his best response strategy, 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 at most $0.219 + 0.219 = 0.438$, this leads to a 0.438-approximate Nash equilibrium.

Both players have a low payoff for the zero-sum game
If both players do not have a payoff of at least $(0.438 + \delta)$ for their zero-sum game both players have to communicate a strategy with logarithmic support. For the row player, the zero-sum game $(R, -R)$ gives a strategy pair $(\mathbf{x}^*_{\mathrm{R}}, \mathbf{y}^*_{\mathrm{R}})$. Form a multiset A by sampling k times from the set of pure strategies of the column player, independently at random according to the distribution $\mathbf{y}^*_{\mathrm{R}}$. Let $\mathbf{y}'_{\mathrm{R}}$ be the mixed strategy for the row player with a probability of $\frac{1}{k}$ for every member of A. For the column player, the zero-sum game $(-C, C)$ gives a strategy pair $(\mathbf{x}^*_{\mathrm{C}}, \mathbf{y}^*_{\mathrm{C}})$. Form a multiset B by sampling k times from the set of pure strategies of the row player, independently at random according to the distribution $\mathbf{x}^*_{\mathrm{C}}$. Let $\mathbf{x}'_{\mathrm{C}}$ be the mixed strategy for the row player with a probability of $\frac{1}{k}$ for every member of B.

Together this will give a strategy pair $(\mathbf{x}'_{\mathrm{C}}, \mathbf{y}'_{\mathrm{R}})$. We have to ensure that $\mathbf{y}'_{\mathrm{R}}$ cannot give the row player a payoff higher than $(0.438 + 2\delta)$. Likewise $\mathbf{x}'_{\mathrm{C}}$ should be a strategy such that the column player is unable of getting a payoff higher than $(0.438 + 2\delta)$. To reach this we only have to look at deviations to pure strategies. The following events represent this:

$$\forall i = 1 \dots n: \ \psi_{a,i} = \{\mathbf{e}_i^T R \mathbf{y}'_{\mathrm{R}} - \mathbf{e}_i^T R \mathbf{y}^*_{\mathrm{R}} < \delta\}$$
$$\forall i = 1 \dots n: \ \psi_{b,i} = \{(\mathbf{x}'_{\mathrm{C}})^T R \mathbf{e}_i - (\mathbf{x}^*_{\mathrm{C}})^T R \mathbf{e}_i < \delta\}$$

Note that the events ψ_a and ψ_b are independent, they represent the strategies of the zero-sum game of the player. If the events hold for all pure strategies, the payoff of the players cannot be higher than $0.438 + 2\delta$. The complement of these events, $\psi^c_{a,i}$ and $\psi^c_{b,i}$, can again be bounded by a standard tail inequality [11]:

$$Pr[\psi^c_{a,i}] \leq e^{-2k\delta^2}$$
$$Pr[\psi^c_{b,i}] \leq e^{-2k\delta^2}$$

These are the probabilities for one pure strategy, we have to sum over all strategies to get the bound:

$$Pr[\psi^c_a] = \sum_{i=1}^{n} Pr[\psi^c_{a,i}] = n \cdot \frac{1}{n^2} = \frac{1}{n}$$

Again the probability of a wrong distribution is small, we sample until we have a good distribution. The strategy $\mathbf{y}'_{\mathrm{R}}$ will be communicated to the column player and $\mathbf{x}'_{\mathrm{C}}$ is communicated to the row player. The players will play the game with the strategy profile $(\mathbf{x}'_{\mathrm{C}}, \mathbf{y}'_{\mathrm{R}})$, where both cannot get a payoff equal to or higher than $(0.438{+}2\delta)$ which leads to a $(0.438{+}2\delta)$-approximate Nash equilibrium.

All cases have been covered and we get a $(0.438 + 2\delta)$-approximate Nash equilibrium. □

7.3 A 0.732-WSNE

The algorithm of the previous section can be adjusted such that it produces an ϵ-well-supported Nash equilibrium. The bound will be weaker than the bound for ϵ-approximate Nash equilibrium. Like the previous algorithm, we will first search for an α-approximate Nash equilibrium and later find the optimal value for α.

Theorem 21. *When the row player and the column player are allowed to communicate one mixed strategy profile, a 0.732-approximate Nash equilibrium can be found.*

The algorithm starts with both players computing a Nash equilibrium of zero-sum games. The row player solves the zero-sum game $(R, -R)$ and the column player computes $(-C, C)$. We distinguish two cases, the payoff value of the Nash equilibrium of both players is lower than α (Case 1) or at least one of the players has a value equal to or higher than α for the Nash equilibrium (Case 2). With $O(1)$ communication, the case that holds can be identified.

Case 1:
First consider the case where both players have a Nash equilibrium with value smaller than α. The row player has a strategy pair $(\mathbf{x}^*_\text{R}, \mathbf{y}^*_\text{R})$ and the column player a strategy pair $(\mathbf{x}^*_\text{C}, \mathbf{y}^*_\text{C})$. The row player communicates $\mathbf{y}^*_\text{R}$ to the column player and the column player sends $\mathbf{x}^*_\text{C}$ to the row player. They will now play the game with the strategy pair $(\mathbf{x}^*_\text{C}, \mathbf{y}^*_\text{R})$. If they play according to these strategies, they both cannot get a payoff of α or more, so the strategy is an α-WSNE.

Case 2:
Assume w.l.o.g. the row player, has a Nash equilibrium with at least a payoff of α for his zero-sum game. Let the row player communicate this strategy $\mathbf{x}^*_\text{R}$ to the column player. The column player computes a pure strategy best response $\mathbf{e}_j$ to the strategy of the row player 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 the pure strategy of the column player.

At this point in the algorithm we have a strategy pair $(\mathbf{x}^*_\text{R}, \mathbf{e}_j)$. The strategy of the column player is a best response to $\mathbf{x}^*_\text{R}$, so his strategy is a 0-WSNE. When we look at the row player, all we know is that his payoff is at least α. We cannot give a bound on his approximation. It could be that a row in his strategy gives a payoff of 0, while his best response has a value of 1. This would give a 1-WSNE for the row player.

Like the previous algorithm we allow the row player to shift some of his probability to his best response to $\mathbf{e}_j$. Again we are allowed to shift $\frac{1}{2}\alpha$ of the probability of the row player, this ensures the column player a α-WSNE.

The row player could have rows in his support for which it holds that the payoff of this row is more than α lower than the best response of the row player. The total probability of these rows is bounded. The total payoff for the row player is at least α, so some rows in the support of $\mathbf{x}^*_\text{R}$ give a high payoff. Let the best response of the row player have a value of β. The amount of probability on rows that are not α-well-supported while the total payoff is at least α is maximised if all these rows have a payoff slightly below $\beta - \alpha$ and the remainder of the probability of the row player is on the best response row. The total probability of rows that are not α-well supported can be $\frac{1}{2}\alpha$. This leads to the following payoff:

$$\beta(1 - \frac{1}{2}\alpha) + (\beta - \alpha - \epsilon)\frac{1}{2}\alpha$$

The total payoff is by definition at least α. If we let $\epsilon \to 0$, we get the following inequality:

$$\begin{aligned}
\beta(1 - \frac{1}{2}\alpha) + (\beta - \alpha)\frac{1}{2}\alpha &\geq \alpha \\
\beta - \frac{1}{2}\alpha\beta + \frac{1}{2}\alpha\beta - \frac{1}{2}\alpha^2 &\geq \alpha \\
\beta - \frac{1}{2}\alpha^2 &\geq \alpha
\end{aligned}$$

This solves to:

$$0 < \alpha \leq \sqrt{3} - 1, \quad \frac{1}{2}(\alpha^2 + 2\alpha) \leq \beta \leq 1$$
$$\alpha = \sqrt{3} - 1, \quad \beta = 1$$

The value of α is maximised when $\beta = 1$, so when the best response has a value of 1. This gives $\alpha = \sqrt{3} - 1$. So the algorithm described gives a $(\sqrt{3} - 1)$-WSNE for both players, approximately a 0.732-WSNE. The strategy pair belonging to ths 0.732-WSNE is $(0.644\mathbf{x}^*_{\mathrm{R}} + 0.366\mathbf{e}_i, \mathbf{e}_j)$, where $\mathbf{e}_i$ is a best response strategy for the row player to $\mathbf{e}_j$. □

7.3.1 Lowering the communication complexity

Like the previous algorithm, the communication complexity of this algorithm can be improved to $O(\log^2 n)$ without weakening the approximation bound a lot. This can be done by not sending the complete strategy of the Nash equilibrium of the zero-sum game, but only an approximation of this strategy with logarithmic support. We will use elements of the result by [14] that for every value of ϵ there exists an ϵ-approximate Nash equilibrium with logarithmic support. This will give us a $(\sqrt{3} - 1 + 2\delta)$-WSNE for any $\delta > 0$.

Theorem 22. *A $(0.732 + 2\delta)$-approximate Nash equilibrium can be found with a communication complexity of $O(\log^2 n)$.*

Again we start with solving the zero-sum games $(R, -R)$ and $(-C, C)$, but we search now for a payoff value of at least $(\sqrt{3} - 1 + \delta)$. At least one of the players can have a payoff of at least $(\sqrt{3} - 1 + \delta)$ or neither player has such a payoff.

A player has a high payoff for his zero-sum game
We first look at the case where one of the players, assume w.l.o.g. the row player, has a high payoff in the Nash equilibrium of his zero-sum game. The column player will play a best pure response to the strategy of the row player, regardless the support of the strategy of the row player.

The zero-sum game $(R, -R)$ gives a strategy pair $(\mathbf{x}^*_{\mathrm{R}}, \mathbf{y}^*_{\mathrm{R}})$. 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 $\mathbf{x}^*_{\mathrm{R}}$. Let $\mathbf{x}'$ be the mixed strategy for the row player with a probability of $\frac{1}{k}$ for every member of A. The distribution $\mathbf{x}'_{\mathrm{R}}$ should have a payoff close to the payoff of $\mathbf{x}^*_{\mathrm{R}}$. This corresponds to the following event:

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

As mentioned in [14] the expression $((\mathbf{x}'_{\mathrm{R}})^T R \mathbf{y}^*_{\mathrm{R}})$ is essentialy a sum of k independent random variables each of expected value $((\mathbf{x}^*_{\mathrm{R}})^T R \mathbf{y}^*_{\mathrm{R}})$, 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 ϕ^c. When we apply a standard tail inequality [11] to bound the probability of ϕ^c, we get:

$$Pr[\phi^c] \leq e^{-2k\delta^2}$$

With $k = \frac{\ln n}{\delta^2}$, this gives $Pr[\phi^c] \leq \frac{1}{n^2}$. By definition $Pr[\phi] = 1 - Pr[\phi^c]$, so $Pr[\phi] \geq 1 - \frac{1}{n^2}$.If $\mathbf{x}'_{\mathrm{R}}$ is a wrong distribution, we sample again.

The strategy $\mathbf{x}'_{\mathrm{R}}$ has a guaranteed payoff of at least $(\sqrt{3}-1)$. 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 $\mathbf{x}'_{\mathrm{R}}$ of $O(\log^2 n)$.

The column player computes a pure strategy best response $\mathbf{e}_j$ to $\mathbf{x}'_{\mathrm{R}}$ and communicates this strategy to the row player. Some rows in the strategy of the row player might not be well supported at this point in the algorithm. As we have seen before, the total probability of these rows is at most $\frac{1}{2}(\sqrt{3}-1)$. The probability assigned to these rows will be redistributed like before.

This change in strategy of the row player can decrease the payoff of the column player by as much as $\frac{1}{2}(\sqrt{3}-1)$ and increase another pure strategy by as much as $\frac{1}{2}(\sqrt{3}-1)$. His strategy was a best response, a 0-WSNE, so his strategy is now a $(\sqrt{3}-1)$-WSNE.

Both players have a low payoff for the zero-sum game
If both players do not have a payoff of at least $(\sqrt{3}-1+\delta)$ for their zero-sum game, both players have to communicate a strategy with logarithmic support.

For the row player we have the zero-sum game $(R,-R)$ which gives a strategy pair $(\mathbf{x}^*_{\mathrm{R}},\mathbf{y}^*_{\mathrm{R}})$. Form a multiset A by sampling k times from the set of pure strategies of the column player, independently at random according to the distribution $\mathbf{y}^*_{\mathrm{R}}$. Let $\mathbf{y}'_{\mathrm{R}}$ be the mixed strategy for the column player of the zero-sum game with a probability of $\frac{1}{k}$ for every member of A. For the column player we have the zero-sum game $(-C,C)$, which gives a strategy pair $(\mathbf{x}^*_{\mathrm{C}},\mathbf{y}^*_{\mathrm{C}})$. Form a multiset B by sampling k times from the set of pure strategies of the row player, independently at random according to the distribution $\mathbf{x}^*_{\mathrm{C}}$. Let $\mathbf{x}'_{\mathrm{C}}$ be the mixed strategy for the row player with a probability of $\frac{1}{k}$ for every member of B.

Together this gives a strategy pair $(\mathbf{x}'_{\mathrm{C}},\mathbf{y}'_{\mathrm{R}})$. We have to ensure that $\mathbf{y}'_{\mathrm{R}}$ cannot give the row player a payoff higher than $(\sqrt{3}-1+2\delta)$. Likewise $\mathbf{x}'_{\mathrm{C}}$ should be a strategy such that the column player is unable of getting a payoff higher than $(\sqrt{3}-1+2\delta)$. To reach this we only have to look at deviations to pure strategies. The following events represent this:

$$\forall i = 1\ldots n: \ \psi_{a,i} = \{\mathbf{e}_i^T R\mathbf{y}'_{\mathrm{R}} - \mathbf{e}_i^T R\mathbf{y}^*_{\mathrm{R}} < \delta\}$$
$$\forall i = 1\ldots n: \ \psi_{b,i} = \{(\mathbf{x}'_{\mathrm{C}})^T R\mathbf{e}_i - (\mathbf{x}^*_{\mathrm{C}})^T R\mathbf{e}_i < \delta\}$$

Note that the events ψ_a and ψ_b are independent, they represent the strategies of the zero-sum game of the player. If the events hold for all pure strategies, the payoff of the players cannot be higher than $(\sqrt{3}-1+2\delta)$. Again we can bound the complement of the events described above [11]:

$$Pr[\psi^c_{a,i}] \leq e^{-2k\delta^2}$$
$$Pr[\psi^c_{b,i}] \leq e^{-2k\delta^2}$$

These are the probabilities for one pure strategy, we have to sum over all strategies to get the bound:

$$Pr[\psi^c_a] = \sum_{i=1}^{n} Pr[\psi^c_{a,i}] = n \cdot \frac{1}{n^2} = \frac{1}{n}$$

The probability of a wrong distribution is small, only $\frac{1}{n}$. We sample until we have a good distribution.

The strategy $\mathbf{y}'_{\mathrm{R}}$ will be communicated to the column player and $\mathbf{x}'_{\mathrm{C}}$ is communicated to the row player. The players will play the game with the strategy profile $(\mathbf{x}'_{\mathrm{C}}, \mathbf{y}'_{\mathrm{R}})$, where both cannot get a payoff equal to or higher than $(\sqrt{3} - 1 + 2\delta)$ which leads to a $(\sqrt{3} - 1 + 2\delta)$-WSNE.

All cases have been covered and we get a $(\sqrt{3} - 1 + 2\delta)$-WSNE. □

7.3.2 Probabilistic part

In both cases of this algorithm, zero-sum game has a high payoff or no high payoff, we have to find a strategy of logarithmic size that approximates the strategy of a Nash equilibrium. This part of the algorithm is not deterministic, but the probability that a good solution will be found quickly is very large. For the first case, a high payoff zero-sum game exists, the probability of a wrong solution in a round is $\frac{1}{n^2}$. The chances of not having a good solution after n tries is $1 - n^{-2n}$, which is very close to 1. For the second case the probability of a bad solution in a round is $\frac{1}{n}$, which gives a probability of a wrong answer after n rounds of $1 - n^{-n}$, this is also very close to 1.

Chapter 8

Conclusion

8.1 Results

It was already known that with $O(\log n)$ communication a 0.5-approximate Nash equilibrium could be reached and that with full information of a game a 0.3393-approximate Nash equilibrium is possible. This work gets some more results considering the trade off between approximation of a Nash equilibrium and the amount of knowledge of a player. We have shown that with no communication, a $(0.75 - \frac{1}{4n})$-approximate Nash equilibrium can be reached and that no algorithm can guarantee a 0.501-approximate Nash equilibrium or better without communication. This thesis also proves that with limited communication, $O(\log^2 n)$, an approximation with $\epsilon < 0.5$ can be achieved, a 0.438-approximate Nash equilibrium is possible. As a last result on ϵ-approximate Nash equilibria, it is shown that one-way communication cannot guarantee good bounds on the approximations: the lower bound for models with one-way communication is $0.5 - \frac{1}{\sqrt{n}}$, which shows the optimality of the DMP-algorithm for models with one-way communication. For a good approximation, interaction between the players is important. Apart from the results where the amount of communication is fixed, it is also shown that to reach a good approximation of a Nash equilibrium a certain amount of communication is necessary. This can even hold for games with only two actions per player.

Next to results on ϵ-approximate Nash equilibria, we get results for ϵ-WSNE. It is shown that one-way communication cannot guarantee any bound for ϵ-WSNE. For finding an ϵ-approximate Nash equilibrium one-way communication could give us a bound of $0.5 - \frac{1}{\sqrt{n}}$, but for any $\epsilon < 1$ communication in both directions is needed for ϵ-WSNE. Apart from this negative result we also show a 0.732-WSNE with limited communication. This shows that it is possible to get an approximation on well-supported Nash equilibria with limited communication. The bound of 0.732 is also quite close to the 0.667-WSNE, the best approximation known.

The dynamics for small games (each player only has two actions) are very different from general games with more actions. Good upper bounds are shown for small games with no communication and these bounds can be improved when we allow only a little bit of communication. This shows that the limitation in options combined with the knowledge of your own payoff matrix is enough to get good approximations for ϵ-approximate Nash equilibria. This power is quickly lost when the payoff matrices become bigger. The results are different for ϵ-WSNE, this is expected since we know that we need communication in both directions to get any approximation at all.

Regarding upper bounds on the communication complexity this thesis shows that when a Nash equilibrium only has to be approximated, it is enough to know an approximation of the payoff matrix of the other player. Even stronger, if two strategies are approximately the same only one of the strategies is needed to find an approximate Nash equilibrium.

8.2 Future research

All communication protocols in this thesis assume that the information that is communicated is reported truthfully, but truthfullness is never forced. Future work could address the case where players have an incentive to report their information truthfully.

In this thesis we have found strong lower bounds on the approximation when we allowe no communication or one-way communication. A question that is still open is what the lower bound on the approximation is when a communication complexity of $O(\log n)$ is allowed. As can be seen from this thesis, the communication of strategies seems valuable information for approximating a Nash equilibrium. Communication of a pure strategy has a communication complexity of $O(\log n)$ so when we bound the communication complexity on $O(\log n)$ only a constant number of pure strategies can be communicated. Since a lower bound of 0.5-approximate Nash equilibrium is known for strategies with constant support size, it seems that with $O(\log n)$ communication the lower bound will also be a 0.5-approximate Nash equilibrium.

In section 7 we use zero-sum games in a new way: the Nash equilibrium strategy $\mathbf{y}^{\mathrm{R}}$ belonging to the zero-sum game of the row player is used by the column player to guarantee a low regret for the row player. We could ask the question whether this technique could be used in other algorithms to ensure a low regret for some special cases.

The technique first introduced in [3] that allows a player to switch a part of his probability to a best response strategy had already been used to improve ϵ-approximate Nash equilibrium algorithms. This thesis uses it for the first time to get an ϵ-WSNE. It could be that this technique could be used to improve the 0.667-WSNE algorithm developed in [13].

For models with no communications a lower bound of 0.501-approximate Nash equilibrium is proven. This lower bound can probably be improved. This 0.501 lower bound shows that 0.5 is not the answer for the lower bound. The commitment measure that is introduced to prove the 0.501 lower bound has some nice properties. With the current set of payoff matrices the column player always has a dominant strategy. A set of matrices such that neither player has a dominant strategy might be needed to get a better lower bound. A natural lower bound for games with no communication could be close to 0.75, the best upper bound algorithm gets a bound of 0.75-$o(\frac{1}{n})$ and it seems unlikely that this upper bound can be improved significantly.

The support size of a strategy is important in an uncoupled setup with limited communication. For ϵ-approximate Nash equilibria it is known that for $\epsilon < 0.5$ the support size needs to be at least logarithmic in the number of pure strategies. This implies that the very simple DMP-algorithm is optimal for strategies with constant support size. No easy algorithms exist for ϵ-WSNE for any $\epsilon < 1$. This can lead us to think that for every $\epsilon < 1$ the support size for a ϵ-WSNE should be at least logarithmic in the number of pure strategies. It would be interesting to research this.

This thesis gives algorithms that give a 0.438-approximate Nash equilibrium and a 0.732-WSNE, where the player has no knowledge of the payoff matrix of the other player. Can these bounds be improved, while keeping the restriction of limited knowledge of the other payoff matrix?

Bibliography

[1] Hartwig Bosse, Jaroslaw Byrka, and Evangelos Markakis. New algorithms for approximate nash equilibria in bimatrix games. In *Proceedings of the 3rd international conference on Internet and network economics*, WINE'07, pages 17–29, Berlin, Heidelberg, 2007. Springer-Verlag.

[2] Xi Chen and Xiaotie Deng. Settling the complexity of two-player nash equilibrium. In *Foundations of Computer Science, 2006. FOCS '06. 47th Annual IEEE Symposium on*, pages 261 –272, oct. 2006.

[3] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *J. ACM*, 56:14:1–14:57, May 2009.

[4] Vincent Conitzer and Tuomas Sandholm. Communication complexity as a lower bound for learning in games. In *Proceedings of the twenty-first international conference on Machine learning*, ICML '04, pages 24–, New York, NY, USA, 2004. ACM.

[5] George B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1963.

[6] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, STOC '06, pages 71–78, New York, NY, USA, 2006. ACM.

[7] Constantinos Daskalakis, Aranyak Mehta, and Christos Papadimitriou. A note on approximate nash equilibria. In Paul Spirakis, Marios Mavronicolas, and Spyros Kontogiannis, editors, *Internet and Network Economics*, volume 4286 of *Lecture Notes in Computer Science*, pages 297–306. Springer Berlin / Heidelberg, 2006. 10.1007/11944874_27.

[8] Eyal and Kushilevitz. Communication complexity. volume 44 of *Advances in Computers*, pages 331 – 360. Elsevier, 1997.

[9] Tomas Feder, Hamid Nazerzadeh, and Amin Saberi. Approximating nash equilibria using small-support strategies. In *Proceedings of the 8th ACM conference on Electronic commerce*, EC '07, pages 352–354, New York, NY, USA, 2007. ACM.

[10] Sergiu Hart and Yishay Mansour. How long to equilibrium? the communication complexity of uncoupled equilibrium procedures. *Games and Economic Behavior*, 69(1):107 – 126, 2010.

[11] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):pp. 13–30, 1963.

[12] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, STOC '84, pages 302–311, New York, NY, USA, 1984. ACM.

[13] Spyros Kontogiannis and Paul Spirakis. Well supported approximate equilibria in bimatrix games. *Algorithmica*, 57:653–667, 2010. 10.1007/s00453-008-9227-6.

[14] Richard J. Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM conference on Electronic commerce*, EC '03, pages 36–41, New York, NY, USA, 2003. ACM.

[15] John Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):pp. 286–295, 1951.

[16] Haralampos Tsaknakis and Paul Spirakis. An optimization approach for approximate nash equilibria. In Xiaotie Deng and Fan Graham, editors, *Internet and Network Economics*, volume 4858 of *Lecture Notes in Computer Science*, pages 42–56. Springer Berlin / Heidelberg, 2007. 10.1007/978-3-540-77105-0_8.

[17] J. v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928. 10.1007/BF01448847.

[18] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing(preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, STOC '79, pages 209–213, New York, NY, USA, 1979. ACM.