BSc Thesis Applied Mathematics

Analysis of Nash equilibria for Kuhn poker and its extensions

Luuk van der Werf

Supervisor: R.P. Hoeksma

July, 2022

Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science

UNIVERSITY OF TWENTE.

Preface

I want to thank dr. Ruben Hoeksma for his supervision and honest and critical view, and everyone else who supported me.

Analysis of Nash equilibria for Kuhn poker and its extensions

Luuk van der Werf*

July, 2022

Abstract

With a proper poker strategy you could potentially win a lot of money, but poker is too large to solve analytically. That is where simplified games come into play. One of the most researched simplified poker games is Kuhn poker, a two-player game that has been solved analytically. This paper looks at an adjusted best response algorithm to find Nash equilibria of Kuhn poker. Multiple extensions have been made to Kuhn poker, such as adding a third player. This paper describes how three-player Kuhn poker could be extended to a repeated game, and shows that playing a Nash equilibrium in a single stage game can result in a negative expected value in the repeated version. Lastly, this paper shows how blinds could be added to two-player Kuhn poker, and how this influences the equilibria.

Keywords: Kuhn poker, Poker, Nash equilibrium, Best response

1 Introduction

Poker is a very popular family of card games, with an estimated number of over 100 million players worldwide [7, 16]. Since poker games are gambling games, a proper strategy could earn you a lot of money, so poker has been studied a lot. With the term *poker* we from now on refer to *No-limit Texas Hold 'em*, since this is the most popular variant of poker [21]. This variant is also played at the World Series Of Poker [8], a large poker tournament. Every player gets two cards at the start of a game of poker. These two cards are only to be seen by this player. There are five cards that can be seen by every player, the *community cards*, and they are revealed at specific moments in the game. Every player has to make the best *hand*; a five-card combination out of the two private cards and the five community cards. There is a specific order of the quality of hands [18]. Before and after a reveal of cards players can *bet* a certain amount of *chips*, disks that represent value such as money. The possible sizes of the bet depend on the rules used. Other players have to decide if they will go along with this bet, known as *calling*, or to opt out, known as *folding*. The round ends if one player remains, or if all community cards are revealed. If one player remains, he will end up with the *pot*, the total amount of chips bet during the round. If all community cards are revealed there will be one last betting round, after which every player still in the game reveals their private cards, and the player with the best hand wins the pot.

Since poker is played with 52 cards, multiple bet sizes, and up to eleven players, solving the game analytically is not feasible [20]. That is why simplified games have been introduced which can be solved analytically. The results of analysing these simplified games

*Email: l.vanderwerf-1@student.utwente.nl

can then be extrapolated to real poker. A widely studied simplified poker game is *Kuhn poker* [2, 6, 14, 20].

Kuhn poker is a very simplified version of poker, introduced by Kuhn in 1950 [10]. The game is played with two players and a deck of three cards. Each player has to put one chip in the pot before the round starts, which we refer to as *ante* one chip. The three cards have values 1, 2 and 3, which we refer to as *the 1*, *the 2* and *the 3*, respectively. Each player gets dealt one of the three cards. The value of the remaining card will stay unknown to the players. For clarity we call the two players Alice and Bob. The player with the card with the highest value wins, but just like in real poker, the players are allowed to bet. Alice has the first choice to bet one chip, or to *check*, which means passing on the turn to the next player. If she checks, Bob can choose whether he wants to bet or check. Only one player can bet in Kuhn poker, so this means you cannot *raise*, so betting after another player bet is not allowed. In Chapter 2 we show how this game can be solved analytically, such that we find the *Nash equilibria* of the game.

Definition 1.1 (Best response). A strategy **p** of Alice is a best response to a strategy **q** of Bob if $U_A(\mathbf{p}, \mathbf{q}) \geq U_A(\mathbf{p'}, \mathbf{q})$ for all possible strategies **p'** of Alice, where $U_A(\mathbf{p}, \mathbf{q})$ denotes the utility of Alice when playing strategy **p** versus strategy **q** of Bob.

Definition 1.2 (Nash equilibrium). A pair of strategies (**p**,**q**) is a Nash equilibrium if **p** is a best response to **q**, and **q** is a best response to **p** [15, 17].

In general finding Nash equilibria is not something that is trivial to do. It cannot just be placed in complexity class P or NP. P is the class of decision problems that can be solved in polynomial time, and NP is the class of decision problems of which answers can be verified in polynomial time [5]. Since Nash proved the existence of a Nash equilibrium in every finite game [15], there is always the answer ‘yes’ to the question ‘Does my (finite) game have a Nash equilibrium?’. That is why it is *very unlikely that NP-completeness can characterize the complexity of Nash* [3]. A complexity class in which the problem of finding Nash equilibria does fit is called PPAD, which is a special subclass of NP. Since finding Nash equilibria is a problem with such a special complexity, a lot of complex algorithms have been introduced to find Nash equilibria in specific cases. Fore example [13] for large games, and [22] for team Markov games. An intuitive algorithm for finding Nash equilibria was introduced by Li and Başar in 1987 [11]. This algorithm looks for the best response to the strategy of the other player, alternating between the players of the game. In their paper, they prove that this algorithm finds a Nash equilibrium if the game meets certain conditions. The conditions are that the utility functions need to be strongly convex and second order continuously Fréchet differentiable. In Chapter 3 we show that the utility functions of Kuhn poker do not meet one of the conditions. As a result, the algorithm does not find a Nash equilibrium for this game. To be able to find a Nash equilibrium of Kuhn poker, we slightly adjust the algorithm by Li and Başar. This leads us to our first research question: ‘Can we adjust the best response algorithm by Li and Başar such that it finds a Nash equilibrium of Kuhn poker?’. We discuss this question in Chapter 3.

Since poker is a very popular game, a lot of people claim to have found optimal strategies. A popular term for such strategies is *Game Theory Optimal* or *GTO*. The claim is that you cannot lose money, or that you are ‘unexploitable’ while playing such a GTO strategy [9]. For a very large game as poker, such claims are hard to verify, but we can verify these claims for simplified games. This leads us to our second research question: ‘Can GTO strategies in poker have a negative expected outcome against multiple players?’. This question is discussed in Chapter 4.

Kuhn poker is in a lot of ways very different from poker. One of the aspect of real poker that Kuhn poker omits is having so-called *blinds*. Blinds are bets that players in a certain position at the poker table are forced to make, much like antes. The difference is that every player has to put in the same ante, while blinds differ per position. Usually the game is played with two blinds, the *big blind* and the *small blind*. Generally the big blind is equal to the smallest bet size, and the small blind being half the big blind. In Kuhn poker the two players have no choice of opting out on playing, even when they have a bad hand. With the addition of blinds we can add the element of opting out to the game. This leads us to our third research question: 'How can blinds be added to Kuhn poker, and how does it change the Nash equilibria?'. This question is discussed in Chapter 5.

2 Solving Kuhn poker

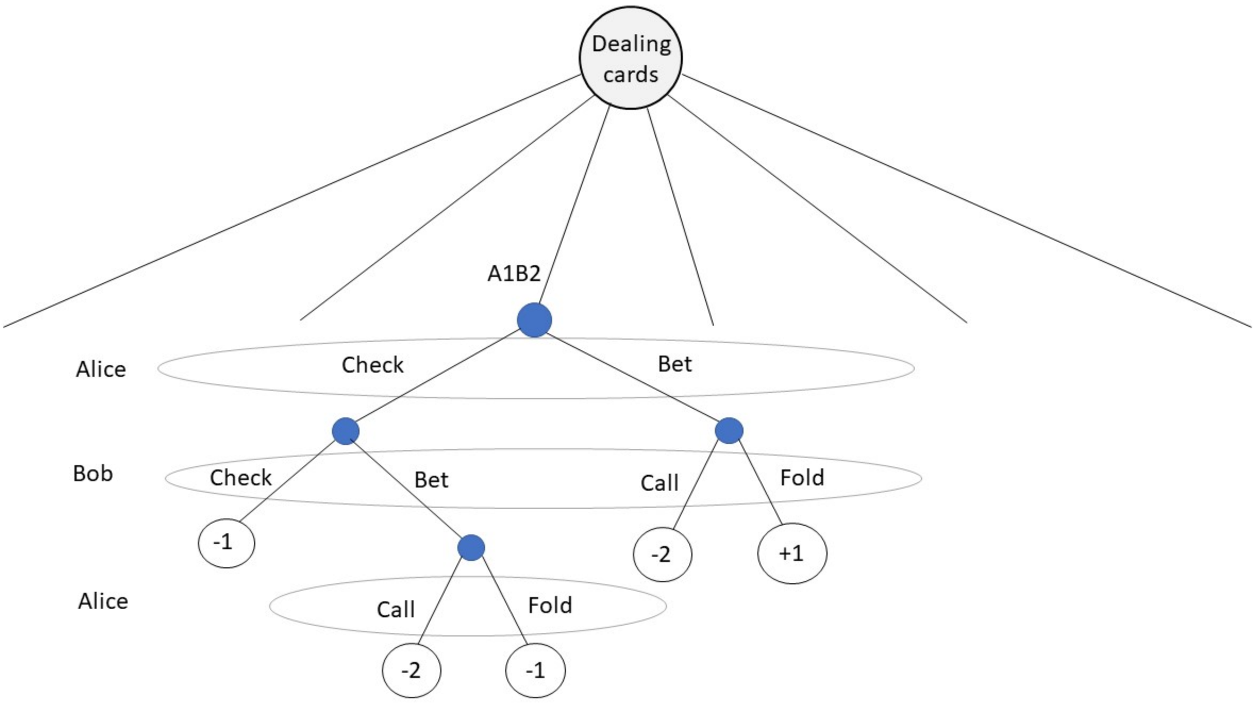

FIGURE 1: Partial game tree for A1B2 with the utility of Alice.

In this section we show a way to find the Nash equilibria of Kuhn poker introduced by Swanson in 2005 [19]. This proof is intuitive and illustrative, and results in the same equilibria as in the original paper by Kuhn [10]. To find the Nash equilibria of Kuhn poker, we need to know what decisions the players have to make. The partial game tree in Figure 1 depicts the decisions Alice and Bob can make in state *A1B2*, and what the outcomes of those decisions are for Alice. *A1B2* denotes Alice having the 1, and Bob having the 2. For every different card, a player has to decide in which fraction of games they will bet, and in which fraction they will call. For every decision we have a parameter, which we call the *strategy parameters*. The value of every strategy parameter will be between 0 and 1, since they are probabilities. Alice has to decide in what fraction of games she will bet with every different card. If she does not bet, she will check, so she will do that with a probability of one minus the probability she will bet. She can have three different cards, which gives a total of three strategy parameters for the decision of betting or checking. If Alice checked, Bob will have the option to bet. If he does, Alice has to make a decision if she will call or fold. Since Alice can have three different cards this again gives three strategy parameters, which makes the total number of parameters for Alice six. Similarly, Bob also has six strategy parameters. If Alice checks he has to choose how much percent of the time he will bet. If Alice bets he has to choose how much percent of the time he will call. Again with the three different cards this gives us six strategy parameters. Luckily, a lot of these strategy parameters are *strictly dominated* by a single value.

Definition 2.1 (Strictly dominated)**.** In a game with n players, a strategy s_i is said to be strictly dominated by strategy s_i^* if s_i^* always gives a strictly higher utility than s_i, independent of the strategies of the other players. So, if $U_i(s_1, ..., s_i*, ..., s_n) > U_i(s_1, ..., s_i, ..., s_n)$ for any strategy $s_i \neq s_i*$ and any combination of strategies of the opponents $(s_1, ..., s_{i-1}, s_{i+1}, ..., s_n)$, we call s_i* the strictly dominant strategy and s_i the inferior strategy.

Since every inferior strategy is strictly dominated, a rational player will always play the dominant strategy. The first two strategy parameters that have a strictly dominant strategy are the parameters for calling with the 1, both for Alice and Bob. When a player has the 1, the opponent will always have a higher valued card, and thus calling a bet always gives you a negative payoff. This means the dominant strategy is setting these parameters to 0.

Proof. Let us assume a player decides to call with the 1 with probability $p \in (0, 1]$. We know that every time a player calls with the 1 he loses an extra chip, so $U(p) = -1 \cdot p = -p$, which is independent of the strategy of the opponent. Now for $p^* = 0$, $U(p^*) = U(0) = -1 \cdot 0 = 0$, and thus $U(p^*) > U(p)$, so $p^* = 0$ strictly dominates $p > 0$. □

Two other parameters that have a strictly dominant strategy are the parameters for folding with the 3, again both for Alice and Bob. When a player has this card, he knows he will always have the highest card, so he will never fold a bet. Also, Alice and Bob will both never bet with the 2. This might be less obvious to see, but think from the perspective of the opponent. When Alice has the 2, Bob has either the 1 or the 3. Bob will always fold with the 1 and will always call with the 3, so Alice will never win anything when she bets with the 2. The last parameter that has a strictly dominant value is only for Bob. Bob will always bet with the 3 when gets the opportunity, this is when Alice does not bet. Bob can only improve his payoff with this, so he will always do it. On the contrary, Alice will not always bet with the 3. If she does not bet with the 3, Bob might bet with a lower valued card and Alice could have a higher payoff in this way. This leaves us with five parameters to consider. We call the parameters as described in Table 1.

p_1	Alice bets with the 1
p_2	Alice calls with the 2
p_3	Alice bets with the 3
q_1	Bob bets with the 1
q_2	Bob calls with the 2

TABLE 1: Names of the strategy parameters.

Since the game is played with three cards and two players, the cards can be distributed in $3! = 6$ ways to consider. Let us consider the distribution *A1B2*. Alice loses a chip with probability $p_1 \cdot q_2$, which means she bets with the 1 and Bob calls her bet. She wins two chips with probability $p_1 \cdot (1-q_2)$, which means she bets and Bob does not call her bet and thus folds. If we do this for all the six hands we get the total probability of Alice winning or losing chips, depending on the five strategy parameters. The total utility for Alice is

$$U_A(\mathbf{p}, \mathbf{q}) = \frac{1}{6}[p_1(1-3q_2) + p_2(3q_1 - 1) + p_3(q_2 - q_1) - q_1] \ . \tag{1}$$

From Equation 1 we see that if $1 - 3q_2 = 0$, $3q_1 - 1 = 0$ and $q_2 - q_1 = 0$, the strategy $\mathbf{p} = (p_1, p_2, p_3)$ does not influence Alice's utility. This is the case for $q_1 = q_2 = \frac{1}{3}$ [19], which is the Nash equilibrium strategy for Bob. Alice's utility can be rewritten as

$$U_A(\mathbf{p}, \mathbf{q}) = \frac{1}{6}[q_1(3p_2 - p_3 - 1) + q_2(p_3 - 3p_1) + (p_1 - p_2)] \ . \tag{2}$$

From Equation 2 we see that if $3p_2 - p_3 - 1 = 0$ and $p_3 - 3p_1 = 0$, the strategy $\mathbf{q} = (q_1, q_2)$ does not influence the utility of Bob. This is the case for the family of solutions $p_1 = \frac{p_3}{3}$, $p_2 = \frac{p_3}{3} + \frac{1}{3}$ and $0 \leq p_3 \leq 1$, which is the family of Nash equilibrium solutions for Alice. This is the same as in the original article by Kuhn [10].

3 Adjusting the iterative best response algorithm

In [11] Li and Başar present an iterative best response algorithm to find Nash equilibria. They formulate the algorithm as

$$u_{k+1} = \text{argmin}_{u \in U} J^1(u, v_k) \ ,$$
$$v_{k+1} = \text{argmin}_{v \in V} J^2(u_{k+1}, v) \ .$$

Here J^1 and J^2 are the cost functions of player 1 and player 2, and U and V are the strategy spaces of player 1 and player 2, respectively. Cost functions can be seen as an equivalent of utility functions, only the cost function is minus the utility function. So in general, players want to minimize their cost, and thus maximise their utility.

For the sequence $\{u_k, v_k\}$ to converge to a Nash equilibrium, the cost functions have to be strongly convex and need to be second order continuously Fréchet differentiable. If we look at Kuhn poker, we see that the utility functions are linear functions, so this means that they are convex, as well as concave [1]. Since they are convex as well as concave, they are by definition not strictly convex, while a strongly convex function needs to be strictly convex. This means the utility functions of Kuhn poker are not strongly convex. So, the convergence condition does not hold. This means that the iterative best response algorithm might not converge to a Nash equilibrium. We now show how to apply the algorithm of Li and Başar on Kuhn poker, and that in practice it does indeed not converge.

If we fill in our utility functions (1) and (2) from Chapter 2 in the iterative best response algorithm we can interpret this as the following linear programmes (LPs).

The LP for Alice is

Maximize $\frac{1}{6}[p_1(1-3q_2) + p_2(3q_1 - 1) + p_3(q_2 - q_1) - q_1]$

subject to $0 \leq p_i \leq 1$ for $i \in \{1, 2, 3\}$,

where $\mathbf{q} = (q_1, q_2)$ is Bob's strategy.

The LP for Bob is

Minimize $\frac{1}{6}[p_1(1-3q_2)+p_2(3q_1-1)+p_3(q_2-q_1)-q_1]$

subject to $0 \leq p_i \leq 1$ for $i \in \{1,2,3\}$,

where $\mathbf{p}=(p_1,p_2,p_3)$ is Alice's strategy.

For the above to be proper LPs, the strategies of the opponents need to be fixed. We start with fixing initial values for Bob's strategy. Using an LP solver we can calculate the best response for Alice versus the initial value. This response of Alice gives us an LP for Bob's strategy. Solving this LP gives us again a strategy of Bob. We can repeat this process until we are satisfied with the outcomes. A stopping criterion for finding a Nash equilibrium could be

STOP if $|\mathrm{NE}_{\mathbf{p}}-\mathbf{p}|<\epsilon$ and $|\mathrm{NE}_{\mathbf{q}}-\mathbf{q}|<\epsilon$ for $\epsilon>0$,

where $\mathrm{NE}_{\mathbf{p}}$ and $\mathrm{NE}_{\mathbf{q}}$ denote Nash equilibrium solutions for Alice and Bob, respectively. The choice of ϵ determines how accurate the outcome will be.

A problem that arises is that it immediately starts looping between strategies that only consist of zeroes and ones, while the Nash equilibria we seek to find consist of mixed strategies. For example, if we start with initial values $q_1=q_2=0$ we get the LP

Maximize $\frac{1}{6}(p_1-p_2)$

subject to $0 \leq p_i \leq 1$ for $i \in \{1,2,3\}$.

The optimal solution for this LP is $p_1=1, p_2=0$ while the value of p_3 does not change the value of the objective function. The LP solver assigns the value 0 to a parameter that does not influence the objective function, and that is not in a constraint together with another variable. This is the case for p_3, so it will be 0. This means that even filling in the Nash equilibrium for Bob as initial value will not keep the system in equilibrium. This can be explained by the fact that if you fill in the Nash equilibrium for Bob, $q_1=q_2=\frac{1}{3}$, the objective function will just be $-\frac{1}{18}$, which is clearly not dependent on p_i. Also, the parameters are not together in a constraint. This means the solver will just set $p_1=p_2=p_3=0$. When Bob responds to this, he will deviate from his Nash equilibrium strategy since he can exploit the strategy of Alice.

To overcome the looping behaviour described in the previous paragraph, we slightly modify the algorithm of Li and Başar. We include a part of the last played strategy by a player as a part of their new strategy. This means only a fraction of the best response will be part of the actual response. We denote this fraction by α. The algorithm then changes to

$$u_{k+1}=\alpha\cdot\mathrm{argmin}_{u\in U}J^1(u,v_k)+(1-\alpha)\cdot u_k \ ,$$
$$v_{k+1}=\alpha\cdot\mathrm{argmin}_{v\in V}J^2(u_{k+1},v)+(1-\alpha)\cdot v_k \ .$$

The looping behaviour of the original algorithm changes in the adjusted algorithm if we set the ratio $\alpha > 0.1$, for example $\alpha = 0.5$, but is not fixed yet. With initial values $q_1 = q_2 = 0$ we see a looping behaviour between four strategies pairs, with their rounded values in Table 2, starting at iteration $k = 55$.

p_1	p_2	p_3	q_1	q_2
$\frac{4}{15}$	$\frac{1}{3}$	$\frac{11}{15}$	$\frac{2}{3}$	$\frac{4}{5}$
$\frac{2}{15}$	$\frac{2}{3}$	$\frac{13}{15}$	$\frac{1}{3}$	$\frac{2}{5}$
$\frac{1}{15}$	$\frac{1}{3}$	$\frac{14}{15}$	$\frac{2}{3}$	$\frac{1}{5}$
$\frac{8}{15}$	$\frac{2}{3}$	$\frac{7}{15}$	$\frac{1}{3}$	$\frac{3}{5}$

Table 2: The four looping strategy pairs.

Lowering the ratio α makes the looping behaviour a lot less clear, and the values get closer to a known Nash equilibrium. For example, if we set $\alpha = 0.01$ and start with initial value $q_1 = q_2 = 0$, we see that after 1000 iterations we get the strategy pair

$$\mathbf{p} = (0.2148, 0.5588, 0.6822) \text{ versus } \mathbf{q} = (0.3438, 0.3393) .$$

This is quite close to a Nash equilibrium. If we take $p_3 = 0.6822$, then $|p_1 - \frac{p_3}{3}| = |0.2148 - \frac{0.6822}{3}| = 0.0126$ and $|p_2 - (\frac{p_3}{3} + \frac{1}{3})| = |0.5588 - (\frac{0.6822}{3} + \frac{1}{3})| = 0.0019$. Also $|\mathbf{q} - (\frac{1}{3}, \frac{1}{3})| = |(0.3438, 0.3393) - (\frac{1}{3}, \frac{1}{3})| = (0.0105, 0.0060)$. As we see, the differences between the found values and an actual Nash equilibrium are of order 10^{-2}.

4 Extending three-person Kuhn to a repeated game

Multiple extensions to Kuhn poker have been made, such as a variant with cheating [14], adding more cards [2], or removing the limits on the bet sizes [6]. One of the largest and most interesting extensions has been made by adding one extra player and thus also one extra card [20]. This gives us four cards, with values 1, 2, 3 and 4. Most rules of two-player Kuhn poker still apply in the same sense. Only one player can bet, and if it comes down to a showdown, the player with the card with the highest value wins.

Since we now have four cards, this give us $4! = 24$ distributions to consider. This means that if you would extend the game to n players and $n + 1$ cards, you have $(n + 1)!$ distributions to consider, which gets unreasonably big very fast. The claim is that three-player Kuhn poker is the largest game with more than two players to be solved analytically [20]. Next to $(n + 1)!$ distributions, the number of strategy parameters also increases if n increases. The number of strategy parameters can be calculated by multiplying the number of players with the number of cards and the number of decisions every player has per card. For two players and three cards, every player has two decisions per card. For three players and four cards this amount is doubled to four decisions per card. This pattern continues, so the number of strategy parameters $S(n)$ can be calculated by $S(n) = n \cdot (n+1) \cdot 2^{n-1}$ where n is again the number of players. This number does not nearly grow as fast as the number of card distributions. Actually, if n gets large, the number of parameters grows with a factor close to 2. This is since $S(n)/S(n + 1) = \frac{n(n+1) \cdot 2^{n-1}}{(n+1)(n+2) \cdot 2^n} = \frac{n}{2(n+2)}$, and $\lim_{n \to \infty} \frac{n}{2(n+2)} = \frac{1}{2}$. To find the utilities for the players we need to combine the different card distributions with the parameters. This then explains why three-player Kuhn poker is claimed to be the largest game with more than two players that has been solved analytically [20].

For this game again a lot of parameters do not need to be considered, since they are strictly dominated. This leaves us with 21 of the 48 parameters to consider. A family of Nash equilibria for these 21 parameters has been found by [20].

One of the most interesting results of the extension of Kuhn poker to three players is that one player can transfer utility from one opponent to the other without departing from equilibrium. In fact, the utilities for Alice, Bob, and third player Carol for the family of equilibria are $U_1 = -\frac{1}{24}(\frac{1}{2} + \beta)$, $U_2 = -\frac{1}{48}$ and $U_3 = \frac{1}{24}(1 + \beta)$, respectively. Here $\beta = max\{b_{11}, b_{22}\}$, where b_{11} and b_{22} are two parameters that can be chosen by Bob only. The value of β can be between 0 and $\frac{1}{4}$ in the family of equilibria. In Figure 2 we see how β influences U_1 and U_3.

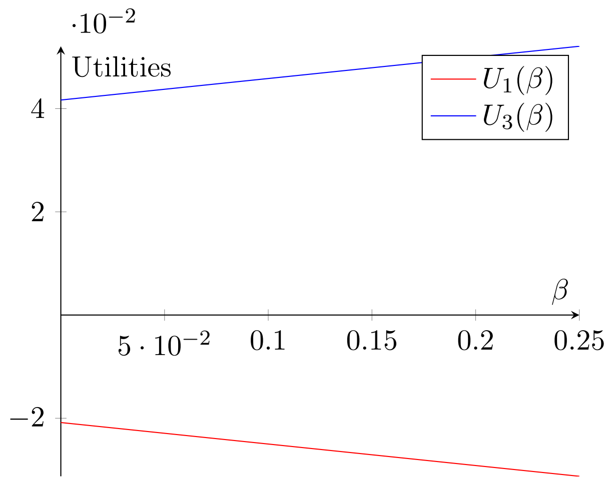

FIGURE 2: The utilities of players in position 1 and 3, as a function of β.

The fact that Bob can transfer utility from Alice to Carol means that he could favour one opponent above the other. If we regard the game as a single stage game, so only one round of the game, rational players want to maximize their utility of that single stage. We know that in the found equilibrium family the utility for Bob is $U_2 = -\frac{1}{48}$, so independent of his choice of β. This means he can choose β without changing his own utility, so no logical choice for the value of β exists. This fact changes if we extend this single stage game to a repeated game. To extend this single stage game to a repeated game we need to make some assumptions:

Assumption 1: All the three players play a strategy according to the found equilibrium family of the single stage game. This means the only choice that influences this repeated game is the choice of β for the player in position 2.

Assumption 2: After each round the player in position 2 will reveal his choice of β, since this value might not be clear from the actions played in the past round.

Assumption 3: If a player in position 1 or 3 is *disfavoured* by the player in position 2, he will try to disfavour this player as well.

Definition 4.1 (Disfavour)**.** The player in position 2 is said to disfavour the player in position 1 if $\frac{1}{8} < \beta \leq \frac{1}{4}$, while he disfavours the player in position 3 if $0 \leq \beta < \frac{1}{8}$. If $\beta = \frac{1}{8}$ neither of the players is said to be disfavoured.

Assumption 4: The repeated game will have an infinite number of rounds.

Assumption 5: The players will base their choice of β on the previous two rounds.

Since this is now a repeated game, Alice, Bob and Carol will change from position after every round. The first round of every game will start with Alice in position 1, Bob in position 2 and Carol in position 3. We denote this as $A_1B_2C_3$. The round after this we will get the position $C_1A_2B_3$, and the round after that $B_1C_2A_3$.

Example 1: At position $A_1B_2C_3$, the choice of β is left to Bob. Let us say he chooses $\beta = \frac{1}{4}$. This is the value of β that is the most disadvantageous for the player in position 1, Alice. The next round is $C_1A_2B_3$, so Alice will choose the value of β. Since she was disfavoured by Bob, she will now disfavour Bob, so she will play $\beta = 0$. Next round is $B_1C_2A_3$. Carol was favoured by both players, which gives him a positive utility for the repeated game. This makes that the only logical choice for Carol is to play $\beta = \frac{1}{8}$, which does not favour any of the two opponents. The pattern of Alice and Bob 'punishing' each other continues to all the rounds after, since they constantly reply to each other. This gives the following utilities per three rounds for the repeated game for Alice, Bob and Carol, respectively:

$$U_A = U_1(\tfrac{1}{4}) + U_2 + U_3(\tfrac{1}{8}) = -\tfrac{1}{32} - \tfrac{1}{48} + \tfrac{3}{64} = -\tfrac{1}{192} \ ,$$

$$U_B = U_1(\tfrac{1}{8}) + U_2 + U_3(0) = -\tfrac{5}{192} - \tfrac{1}{48} + \tfrac{1}{24} = -\tfrac{1}{192} \ , \text{ and}$$

$$U_C = U_1(0) + U_2 + U_3(\tfrac{1}{4}) = -\tfrac{1}{48} - \tfrac{1}{48} + \tfrac{5}{96} = \tfrac{1}{96} \ ,$$

where U_1, U_2, U_3 denote the utilities of players in position 1, 2 and 3, respectively. Since Carol is the only player that does not get disfavoured by any opponent, he is the one with a positive utility. Alice and Bob end up with a negative utility, and since every round is a zero-sum game $U_A + U_B + U_C = 0 \implies U_C = -U_A - U_B$.

From this example we see that playing anything different than $\beta = \frac{1}{8}$ as Bob in the first round will lead to a negative utility for the repeated game. The same goes for the other players in any other round when they have to choose β. So while per single round the value of β does not matter for the player in position 2, on the long run only $\beta = \frac{1}{8}$ does not lead to a negative utility.

The underlying assumption of Example 1 is that all the players act individually and do not cooperate. Let us now assume that players can cooperate.

Example 2: Again, the first round is $A_1B_2C_3$. Bob plays $\beta = \frac{1}{8}$, since he does not want to disfavour any of his opponents. Now, instead of playing individually, Alice and Bob cooperate. When in position 2, they will both favour each other at the expense of Bob. So at $C_1A_2B_3$ Alice plays $\beta = 0$, and at $B_1C_2A_3$ Carol plays $\beta = \frac{1}{4}$. Then we get back at $A_1B_2C_3$. Bob has been disfavoured by Alice, but also by Carol. The best he can do to 'punish' them back is playing $\beta = \frac{1}{8}$. Any other value of β would favour an opponent that disfavoured Bob. This gives us the utilities per three rounds $U_A = U_1(\frac{1}{8}) + U_2 + U_3(\frac{1}{4}) = \frac{1}{192}$, $U_B = U_1(\frac{1}{4}) + U_2 + U_3(0) = -\frac{2}{192}$, $U_C = U_1(0) + U_2 + U_3(\frac{1}{8}) = \frac{1}{192}$. In this way by cooperating, Alice and Carol both get a positive utility, while Bob cannot do anything to stop them from exploiting him. You might think that Bob could try playing $\beta = 0$ or $\beta = \frac{1}{4}$ to change the strategy of either Alice or Carol, but in both cases the utility of either Alice or Carol goes to 0, which is not worse than when every player plays $\beta = \frac{1}{8}$, so neither

Alice nor Carol has a reason to deviate from their cooperating strategy.

A Nash equilibrium strategy is something that would be considered a Game Theory Optimal strategy. In Example 2 we showed that playing a Nash equilibrium strategy of the single stage game can lead to a negative utility in multiple rounds, which goes against the claims of a lot of poker 'experts' who claim that GTO strategies are never unprofitable [9].

5 Adding blinds to Kuhn poker

In this section we show how blinds could be added to Kuhn poker. Let us call the size of the small blind X, and the size of the big blind Y, with $X < Y$. The bet size for the rest of the game will be Y as well. The new game will go as follows. Bob will have to put in a blind of X, and Alice will have to put in a blind of Y. Then we add an extra decision to the game. Bob has to decide if he wants to put in the same amount as Alice, so Y in total. If he does so, the game continues the same as before, only with Y as bet size. If Bob decides to fold, the game stops, and Bob loses his blind X to Alice. This gives us two new extra strategy parameters to consider, called q_3 and q_4. q_3 is the probability that Bob calls Y while having the card with value 1, and q_4 is the probability that Bob calls Y while having the card with value 2. If Bob has the card with value 3 he will always call, so we do not have to consider this. For every different hand we now have to add the probabilities of Bob calling or folding the hand in first place. This gives us the utility for Alice

$$\begin{aligned} U_A(\mathbf{p}, \mathbf{q}) = \frac{1}{6}Y[&p_1(-3q_2q_4 + 2q_4 - 1) + p_2(3q_1q_3 - 1) \\ &+ p_3(-q_1q_3 + q_2q_4) + 2(q_3 - 1) + q_1q_3] + \frac{1}{6}X[3 - 2q_4 - q_3] \ . \end{aligned} \tag{3}$$

To verify if this corresponds with Equation 1, we fill in $q_3 = q_4 = 1$. This means that Bob will always call big blind Y, so this leaves us with the original game, only with bet size Y. Filling in $q_3 = q_4 = 1$ leaves us with

$$U_A(\mathbf{p}, (q_1, q_2, 1, 1)\ ) = \frac{1}{6}Y[p_1(1 - 3q_2) + p_2(3q_1 - 1) + p_3(q_2 - q_1) - q_1] \ . \tag{4}$$

As expected X is gone from the equation, since always calling big blind Y never gives us X as outcome anymore. Equation 4 is exactly the same as Equation 1, only multiplied by Y, as expected. If we set $q_3 = q_4 = 0$ we get some other interesting behaviour. Equation 3 becomes

$$U_A(\mathbf{p}, (q_1, q_2, 0, 0)\ ) = \frac{1}{6}Y[-(p_1 + p_2) - 2] + \frac{1}{2}X \ . \tag{5}$$

Equation 5 is only dependent on strategy parameters of Alice. Since Alice wants to maximize her utility, she will then play $p_1 = p_2 = 0$ and we are left with $U_A = -\frac{1}{3}Y + \frac{1}{2}X$. If $Y > \frac{3}{2}X$ this strategy gives Alice a negative and thus Bob a positive utility, and Alice cannot do anything against it.

By playing $\mathbf{q} = (\frac{1}{3}, \frac{1}{3}, 1, 1)$ Bob can guarantee himself a utility of $\frac{1}{18}Y$, but as we have seen, by playing $\mathbf{q} = (q_1, q_2, 0, 0)$ Bob can guarantee himself a utility of $\frac{1}{3}Y - \frac{1}{2}X$, where the choice of q_1 and q_2 does not matter. If we set these utilities equal to each other, we get $\frac{1}{18}Y = \frac{1}{3}Y - \frac{1}{2}X$. This equality holds if $X = \frac{5}{9}Y$, so for this ratio the utilities of the two strategies are the same. If $X \leq \frac{5}{9}Y$ then $\frac{1}{3}Y - \frac{1}{2}X \geq \frac{1}{18}Y$, so Bob would prefer playing $q_3 = q_4 = 0$ over $q_3 = q_4 = 1$. On the contrary, Bob would prefer playing $q_3 = q_4 = 1$ over $q_3 = q_4 = 0$ if $X \geq \frac{5}{9}Y$.

With the strategy $\mathbf{q} = (q_1, q_2, 0, 0)$ Bob can guarantee himself a utility of $\frac{1}{3}Y - \frac{1}{2}X$, which is positive if $\frac{2}{3}Y > X$. Even though with this strategy Bob can guarantee himself a positive utility, it is not a Nash equilibrium.

Proof. From Equation 5 we saw that the best response for Alice to $\mathbf{q} = (q_1, q_2, 0, 0)$ is $\mathbf{p} = (0, 0, p_3)$ for any $p_3 \in [0, 1]$. If we fill in this best response in Equation 3 we get

$$U_A(\ (0,0,p_3), \mathbf{q}) = \frac{1}{6}Y[p_3(-q_1q_3 + q_2q_4) + 2(q_3 - 1) + q_1q_3] + \frac{1}{6}X[3 - 2q_4 - q_3]\ . \quad (6)$$

Since p_3 does not influence the value of Equation 5, we can w.l.o.g. set $p_3 = 0$, as well as choosing $X \leq \frac{5}{9}Y$, for example $X = \frac{4}{9}Y$. Equation 6 then becomes

$$U_A(\ (0,0,0), \mathbf{q}) = \frac{1}{6}Y[2(q_3 - 1) + q_1q_3] + \frac{1}{24}Y[3 - 2q_4 - q_3]\ . \quad (7)$$

Since Bob wants to minimize Equation 7, his best response is $\mathbf{q}^* = (1, 0, 0, 1)$. With this strategy $U_A(\ (0,0,0),\ (1,0,0,1)\ ) = -\frac{7}{24}$, while $U_A(\ (0,0,0),\ (q_1,q_2,0,0)\ ) = -\frac{5}{24}$. Since $-\frac{7}{24} < -\frac{5}{24}$ Bob deviates from the strategy $\mathbf{q} = (q_1, q_2, 0, 0)$ to $\mathbf{q}^* = (1, 0, 0, 1)$, so $\mathbf{q}$ is not a Nash equilibrium. □

6 Discussion

In the previous chapters we looked at properties of simplified poker games, but the question is what this actually tells us about real poker. A popular research topic on simplified games is finding Nash equilibria; to test algorithms [4], as well as to use the results to say something about real poker [2]. Since real poker is such an extensive game, we are not close to finding a Nash equilibrium for this game [12]. Even if we are able to find a Nash equilibrium, it is not easily used outside mathematics. Real poker would have at least thousands of strategy parameters. If a player would try to play a Nash equilibrium, he would need to remember all the values of the parameters, and also randomise his action according to the parameters to actually play the Nash equilibrium.

In Chapter 3 we discussed the question 'Can we adjust the best response algorithm by Li and Başar such that it finds a Nash equilibrium of Kuhn poker?'. The short answer to this question is 'yes'. We showed how to modify the best response algorithm by Li and Başar, such that it approximates a Nash equilibrium for Kuhn poker. The modified algorithm has not yet been tested on other games, nor have we proved or disproved that this method can always find a Nash equilibrium in certain games. We showed that the strong convexity condition does not hold, but it might also be the case that the differentiability condition does not hold, due to the utility functions being constrained. This was also brought up in the discussion of the paper of Li and Başar, and they see an extension possible to 'Constrained Nash games'. It has been shown that the original algorithm does not work on Kuhn poker when used in practice. Testing the modified algorithm on more games and proving it can or cannot find Nash equilibria in general would be an interesting topic for future research.

In Chapter 4, we discussed the question 'Can GTO strategies in poker have a negative expected outcome against multiple players?'. To be able to answer this question, we extended three-player Kuhn poker to a repeated game. To say something about the strategies of the players for this repeated game, we assumed that the players can only play strategies in the family of Nash equilibria that were found for the single stage game. This led to interesting results, but the assumption might have some implications. If a player has a negative utility for the repeated game it might be better to deviate from the equilibrium strategy for the single stage game. Future research could show if this is actually the case. To come back to the research question, it can be argued that the answer to this question is 'yes'. We now know that for three-player Kuhn poker there exist strategies that fall under the denominator GTO, while having a negative expected value in the long run. Since real poker is way more extensive, we can argue that the probability is high that there exist strategies that could be called GTO, while having a negative expected value in the long run.

In Chapter 5 we discussed the question 'How can blinds be added to Kuhn poker, and how does it change the Nash equilibria?'. We showed that adding a small blind and a big blind can give us situations where there exists a strategy with a higher utility than the Nash equilibria of Kuhn poker without blinds. We did prove that this strategy is actually not a Nash equilibrium. Future research could be aimed at finding a Nash equilibrium of this game. Another topic for future research could be to look at more elements that could be added to Kuhn poker, to make it resemble real poker better. For example, the blinds could be added to three-player Kuhn poker.

7 Conclusion

In this paper we showed that a very simple game as Kuhn poker can lead to a lot of interesting results. In Chapter 3 we showed how to modify a best response algorithm such that it finds Nash equilibria for Kuhn poker. In Chapter 4 we showed that if we extend three-player Kuhn poker to a repeated game, there exist equilibrium strategies that lead to a negative expected value. It might very well be the case that this is also true for real poker, and that GTO strategies do not need to be profitable. In Chapter 5 we showed a way to add blinds to Kuhn poker. We showed that for certain ratios between the small and big blind, new strategies arise with a higher utility than the former Nash equilibria.

References

[1] Dimitris Bertsimas and John N. Tsitsiklis. *Introduction to linear optimization.* Athena Scientific, 1997, p. 15.

[2] John Billingham. *Equilibrium solutions of three player Kuhn poker with N>3 cards: A new numerical method using regularization and arc-length continuation.* 2018. DOI: 10.48550/ARXIV.1802.04670. URL: https://arxiv.org/abs/1802.04670.

[3] Constantinos Daskalakis, Paul Goldberg, and Christos Papadimitriou. "The Complexity of Computing a Nash Equilibrium". In: *SIAM J. Comput.* 39 (Feb. 2009), pp. 195–259. DOI: 10.1137/070699652.

[4] Sam Ganzfried, Austin Nowak, and Joannier Pinales. "Successful Nash Equilibrium Agent for a 3-Player Imperfect-Information Game". In: *CoRR* (2018). arXiv: 1804.04789. URL: http://arxiv.org/abs/1804.04789.

[5] Oded Goldreich. *P, NP and NP-completeness: The Basics of Computational Complexity.* Cambridge University Press, 2010.

[6] John Hawkin, Robert Holte, and Duane Szafron. "Automated Action Abstraction of Imperfect Information Extensive-Form Games". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 25.1 (Aug. 2011), pp. 681–687. URL: https://ojs.aaai.org/index.php/AAAI/article/view/7880.

[7] *How many people play online poker?* Mar. 2022. URL: https://www.globalbrandsmagazine.com/how-many-people-play-online-poker/.

[8] *How To Play | Texas Holdem.* URL: https://www.wsop.com/poker-games/texas-holdem/#.

[9] *How you should think about poker (but probably don't).* Dec. 2019. URL: https://upswingpoker.com/gto-poker-game-theory-optimal-strategy/.

[10] H.W. Kuhn. "A Simplified Two-Person Poker". In: *Contributions to the Theory of Games* 1 (Dec. 1950), pp. 97–103.

[11] Shu Li and Tamer Başar. "Distributed algorithms for the computation of noncooperative equilibria". In: *Automatica* 23.4 (1987), pp. 523–533. ISSN: 0005-1098. DOI: https://doi.org/10.1016/0005-1098(87)90081-1. URL: https://www.sciencedirect.com/science/article/pii/0005109887900811.

[12] Viliam Lisý and Michael H. Bowling. "Eqilibrium Approximation Quality of Current No-Limit Poker Bots". In: *CoRR* abs/1612.07547 (2016). arXiv: 1612.07547. URL: http://arxiv.org/abs/1612.07547.

[13] Stephen McAleer et al. "Pipeline PSRO: A Scalable Approach for Finding Approximate Nash Equilibria in Large Games". In: *CoRR* abs/2006.08555 (2020). arXiv: 2006.08555. URL: https://arxiv.org/abs/2006.08555.

[14] Amanda Metzner and Daniel Zwillinger. "Kuhn Poker with Cheating and Its Detection". In: *CoRR* abs/2011.04450 (2020). arXiv: 2011.04450. URL: https://arxiv.org/abs/2011.04450.

[15] John F. Nash. "Equilibrium points in n-person games". In: *Proceedings of the National Academy of Sciences* 36.1 (1950), pp. 48–49. DOI: 10.1073/pnas.36.1.48. URL: https://www.pnas.org/doi/abs/10.1073/pnas.36.1.48.

[16] *Online poker sector overview for 2021: Stats, key drivers and more.* Oct. 2021. URL: https://fastoffshore.com/2021/10/online-poker-sector-overview-2021.

[17] Hans Peters. *Game theory a multi-leveled approach.* 2nd ed. Springer, 2015, pp. 38–39.

[18] *Poker Hand Rankings, The Best Texas Hold'em Poker Hands.* Sept. 2021. URL: https://upswingpoker.com/poker-hands-rankings/.

[19] Jason Swanson. *Game theory and poker.* URL: swansonsite.com/W/instructional/game_theory.

[20] Duane Szafron, Richard G. Gibson, and Nathan R Sturtevant. "A Parameterized Family of Equilibrium Profiles for Three-Player Kuhn Poker". In: *AAMAS* (Dec. 2013).

[21] *Texas Hold'em Poker.* URL: https://www.pokerstars.eu/poker/games/texas-holdem/.

[22] Xiaofeng Wang and Tuomas Sandholm. "Reinforcement Learning to Play an Optimal Nash Equilibrium in Team Markov Games". In: *Advances in Neural Information Processing Systems.* Vol. 15. MIT Press, 2002. URL: https://proceedings.neurips.cc/paper/2002/file/f8e59f4b2fe7c5705bf878bbd494ccdf-Paper.pdf.