Graduate Program in Logic, Algorithms and Computation $\mu\Pi\lambda\forall$
M.Sc. Thesis

Game Theoretic Models for Power Control in Wireless Networks

by
Aikaterini Nikolidaki

supervised by:
Dimitris Fotakis

National and Kapodistrian University of Athens
Mathematics Department,
Athens, 2 March 2015

Η παρούσα Διπλωματική Εργασία
εκπονήθηκε στα πλαίσια των σπουδών
για την απόκτηση του
Μεταπτυχιακού Διπλώματος Ειδίκευσης
στη
Λογική και Θεωρία Αλγορίθμων και Υπολογισμού
που απονέμει το
Τμήμα Μαθηματικών
του
Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών

Εγκρίθηκε την 2η Μαρτίου 2015 από Εξεταστική Επιτροπή
αποτελούμενη από τους:

Ονοματεπώνυμο	Βαθμίδα	Υπογραφή
1. Ζάχος Ευστάθιος	Ομότιμος Καθηγητής	
2. Παγουρτζής Αριστείδης	Αναπληρωτής Καθηγητής	
3. Φωτάκης Δημήτρης	Επίκουρος Καθηγητής	

Abstract

In recent years, the technology of mobile communications has evolved rapidly due to increasing requirements, such as access to Internet services via mobile phones and requirements better quality services. Nowadays, the devices use the Long Term Evolution (LTE), which called also as 4G networks. The fourth generation (4G) networks replace the third networks generation (3G) and offer to users improved services at higher speeds. Mobile devices to access the Internet, such as smartphones, tablet PCs and netbooks are in high demand in the market for it is an effort to develop in energy consumption level, that the user does not need recharge the device at regular time intervals. Game theory provides valuable mathematical tools that can be used to solve problems of wireless communication networks and can be applied to multiple layers of wireless networks.

In this thesis, we study power control issue and consider it at the physical layer of wireless networks. Specifically, we study game theoretic models for power control in wireless communication networks (CDMA & LTE). In the game theory, we have focused in the non-cooperative power control games and assumed that both transmitters and receivers are selfish and rational. In addition, we insert regret learning techniques and their connection with the game theory. Finally, we investigate the regret learning techniques applied to the problem of power control in the next generation networks.

Keywords

Game Theory, Non-Cooperative games, Cooperative Games, Power Control, Wireless Networks, Regret Learning Algorithms, No-External Regret, No-Internal Regret, No-Swap Regret, LTE/LTE-Advanced, Uplink Transmission

Acknowledgements

Initially, I would like to express my gratitude to my supervisor, Professor Dimitris Fotakis for his guidance and cooperation while working on this thesis. Also, I thank my supervisor for all I've learned from him and believing in me.

In addition, I would like to thank the other members of the comitee, namely Professors Stathis Zachos and Aris Pagourtzis for inspiring me with their teaching.

Finally, I am most grateful to my parents for supporting me all the years of my studies.

Contents

List of Figures

Chapter 1

Introduction

An important part of applied mathematics is the game theory. That means, game theory is an important study tool for our everyday life. Our actions or moves for a situation depends not only on what we do, but also on what other people do. Some examples or else ”games” in game theory are the negotiation of a price with a seller, the vote at a presidential election, the participation in an auction on the Internet, even for trying to find a seat in a bus. Other most known games are chess, football, monopoly, etc.

Therefore, people have interactions between them for a situation. Their decisions for an everyday problem depends on themselves and the actions from the other people, so they are interdependent. Then, a *game* consist of two players or multi-players and they can have conflicting or common interests. All the players can have a strategy in order to move for a situation. Game theory provides mathematical process for selecting an optimum response to player to face the opponent player. Therefore, a definition of *Game Theory* can be as follows: a set of tools developed to analyze the interactions among multiple agents to achieve their goals. Game theory has been applied to many fields such as biology, economics political science, law, sociology, phylosophy and computer science.

Recently, the game theory is used more and more in computer science and specific in artificial intelligence, cybernetics, and networks. Using the game theory, we can model scenarios in which there is no centralized entity with full/partial information network conditions. Thus, considerable interest has been observed mainly in solving communication and network issues. The most interest for research using the game theory in these issues are the congestion control, the routing, the power control, the flow control and the adaptive interference avoidance.

In continuous of this chapter, we introduce some historical points of game theory and its connection with communications networks and mainly with the power control in wireless networks. Also, we introduce the evaluation of the wireless networks from CDMA to LTE and LTE-A (4G) networks. The aim and the motivation of this thesis is the connection of the next generation of wireless networks (LTE-A) with the game theory. Finally in this section, we analyze the motivation and the scope of this thesis.

1.1 The Relation between Games and Wireless Networks

The game theory can be classified as cooperative, non-cooperative. In non cooperative games, the agents make decisions independently and can not coordinate their strategies. Each non-cooperative game consists of a set of players, a selfish utility function for each player and a set of feasible strategy space for each player. To solve a problem in this theory, we study the existence, the uniqueness, the stability under various strategies and optimality gap. The basic element of the non-cooperative game theory is the Nash equilibrium [1]. Thus, in the non-cooperative networks, the agents are involved in the non-cooperative games.

A cooperative game is a game, that groups of players may enforce to work together to maximize their payoffs. Then, in this case, there is a competition between coalitions of players, rather than between individual players. In the cooperative networks, the agents are involved in the cooperative games. In the cooperative game theory, the wireless agents (users) in a network have an agreement on how to fairly and efficiently share the available spectrum resources [1].

The optimization of Wireless networks is a vast area of research. The systems of Telecommunication such as cellular networks, Wireless Local Area Networks (WLANs), Long Term Evolution (LTE) systems and cognitive radio systems have been designed using layered architecture based models.

The layers of OSI and the connection with the corresponding application fields are as follows:

Layer	**Application Field**
Tranport	Call admission control, Load Control, Cell selection
Network	Routing
Data Link	Medium Access Control
Physical	Power control, Spectrum allocation, Cooperative communications, MIMO systems

The components of a wireless network can be represented as components of a game. Figure 1.1 shows this connection.

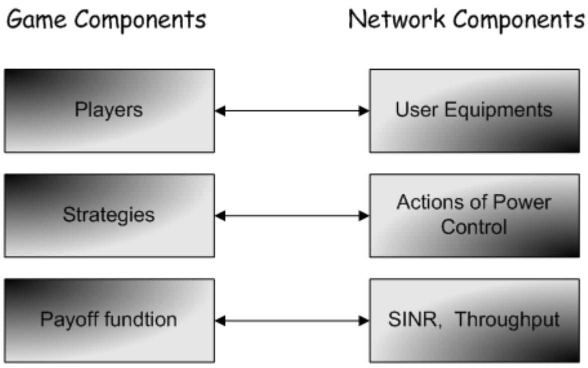

Figure 1.1: The connection between of the game theory and the wireless networks

1.2 Historical Points

1.2.1 Game Theory

Historically, the first published problems/games are developed by the competition of Cournot at the end of the 19th century. In the early of 20th century, Emil Borel worked systematically with games for two players with zero-sum. The zero-sum means that one of the two players gains and the other losses, then there is not cooperation between of these players. In the other hand, Borel could not to answer in the question about the existence of the solution for an arbitrary game. The first textbook about game theory published in 1944 with title "The Theory of Games and Economics Behavior" by John von Neumann and Oscar Morgenstern. In the last of 20th century (1928), John von Neumann proved the existence and uniqueness of a solution using the theorem of Brouwer's fixed point theorem. This theorem is called Minimax theorem. In 1950, John Nash introduced a generalization of the Minimax theorem. The thesis of Nash, contained the definition and properties of equilibrium for non-cooperative games and so called the "*Nash Equilibrium*". Since then, many others researchers have contributed to game theory and its applications to many fields.

1.2.2 Game theory in Communication Networks

Game theory has recently been applied to telecommunications. Indeed, this theory can be applied to analyze interactions between entities such as telecoms regulator, operators, manufactures and customers. The connection between game theory and communication problems is started by Benoit Mandelbrot in his thesis in 1952. From then, the interest of the game theory in the area of wireless communications has increased more and more.

In communications engineering, game theory is often used for distributed resource management algorithms. Some applications of game theory in wireless communications are transmission or power control, pricing, flow control, congestion control and load balancing.

More specific, *power control* in *cellular networks* has been extensively studied since the late 1980s as an important mechanism to control Signal-to-Interference Ratios (SIR), which in turn determine Quality-of-Service (QoS) metrics such as rate, outage, and delay [2].

In 1992, the work of Zander was one of the first studies about the power control techniques in cellular networks. The aim of Zander was to investigate the performance of transmitter power control algorithms and to find performance bounds and conditions of stability for all types of transmitter power control algorithms. He proposed a *centralized power control scheme* and proved that there is a unique solution. The unique solution is always feasible. This means that all the links converge to the same SIR, thus the system is in balanced.

An other study, Foschini and Miljanic (1993) [3] proposed an *iterative distributed protocol* in order to solve the problem of power control in cellular networks using the game theory.

The studies of Foschini, Miljanic and Zander were very important and useful for study and research from many other researchers. In chapter 4, we study and analyze the work and the outcomes of these authors. However, we will introduce additionally important studies for the power control.

In a later work, J. Dams, M. Hoefer, T. Kesselheim [4] proposed an other technique, the regret learning in non-cooperative networks, in order to solve drawbacks of the iteration of [3] such as lack of robustness, the adaption of power when the SINR is known.

In [5], the authors study algorithms in wireless networks where there are intereferences, using the Rayleigh model. For this reason, this model based on the SINR using stochastic propagation to address fading effects observed in reality. Also, they study the behavior of the external regret learning of some user at a time T. The authors apply the regret learning in ordrer to achieve the maximum capacity. In continuous, they proved that any no-regret learning algorithm, the number of successful transmissions needs to converge to a constant fraction of the non-fading optimum.

1.3 Motivation and Scope of this Thesis

In this thesis, we consider that the players of the games are user equipments such as mobile phones, base stations. Interaction between mobile phones is naturally present in wireless networks, since interference often exists and common resources must be shared.

Nowadays, in our smartphones and tablets is the fourth generation of mobile telecommunicaions technology, which called also Long Term evolution (LTE or 4G). LTE is an enhancement to the Universal Mobile Telecommunication System (UMTS) which will be introduced in 3rd Generation Partnership Project (3GPP) Release 10. One important feature of 3GPP LTE system, differentiating it from previous generations of cellular systems, is the distributed network architecture.

In particular, we assume that our network is compatible with the LTE Release 10 and beyond (LTE-Advanced) in 3rd Generation Partnership Project (3GPP) standard for wireless data communications and the wireless devices are in a non-cooperative network. Our goal is the minimum transmisstion power with the maximum throughput in a realistic environment based on economic incentives rules.

1.4 Organisation of this Thesis

The remainder of this thesis is organised as follows:

- In *chapter two*, we analyze the game theory, which can be classified as cooperative and non-cooperative games.
- In *chapter three*, we analyze the regret learning techniques.
- In *chapter four*, we analyze and discuss about some notations of the wireless netwotks, as well as the next generation wireless networks (LTE-A). Mainly, we focus on the related work for the power control in CDMA and LTE networks.
- In *chapter five*, we propose a network system model.

Chapter 2

Game Theory

Game theory is a mathematical tool for analyzing the strategic interactions between agents. Historically, the first publication of game theory and economic behavior was 1944 by Von Neumann and Oskar Morgenstern [6], [1]. During the 1950, John Nash developed the concept of Nash equilibrium. A Nash equilibrium consists of strategies of players in a game [1]. A Nash equlibrium is a stable point that no user can gain by unilaterally deviating, which means that no player has incentive to change his strategy [7].

The *Prisoner's Dilemma* is the most well-known and well-studied game in the literature of game theory [8]. The participants of this game interact or affect each other's outcomes. The description of this Dilemma is as follows: There is a crime and two prisoners, who are on trial. Each one of prisoners have to make a decision to confess or to non-confess. If they both do not confess, then the authorities have not elements against of them. Thus, we assume that the authorities will give 2 years to each of them. If only one of them do not remain silent then the authorities will give him one year and the other prisoner will have 5 years. If they both confess, then will give 4 years to each of them.

Therefore, these two prisoners have to choose between two possible strategies, to confess or to non-confess. The four outcomes are the costs incurred by the players in each situation. The strategies and the costs are presented in the following table. From the table, we can observe that the only stable solution in this game is that both prisoners confess, that at least one of the prisoners can switch from non-confess to confess and improve his own payoff. But, the better solution for both prisoners is when neither of them confesses. This solution is not a stable solution because of each of them prisoners would be tempted to defect and thereby serve less time.

	A confess	**A non-confess**
B confess	$(4,4)$	$(1,5)$
B non-confess	$(5,1)$	$(2,2)$

A strategic-form game model has three components: a finite set of players, denoted by N, a set of possible actions for each player i, denoted by A_i and a set of utility functions, denoted by $u_i : A \rightarrow \Re$. A Nash equilibrium is a strategic profile a ∈ A of actions such that for every player i ∈ N,

$$u(a_i, a_{-i}) \geq u(a_i^*, a_{-i})$$

However, the main types of representation of a game are:

- *strategic form* or else *normal form*. It is the most used form in the game theory and wireless networks. A strategy can be a *pure strategy* of a player, this means that the strategy assigns zero probability to all moves, except one. But, a strategy can be a *mixed strategy*, this means that there are more and different moves with different probabilities. A game with this form is usually represented by a matrix which shows the players, strategies, and payoffs such as we see in the previous example, the Prisoner's Dilemma.

- *extensive form*. A game in this form is represented as a tree, where the root of the tree is the beginning of the game. A sequence of moves defines a path on the tree and is referred to as the history of the game. The terminals nodes of the tree defines the outcome and is assigned the corresponding payoffs. Every extensive form can be transformed to a strategic form and reverse. This form is used to describe dynamic games, this means that the players have a sequential interaction. Thus, the players have some information about the choices of the other players. This form is more complete than the strategic form.

- *coalition form* or else *characteristic function form*. This form is used in non-cooperative games. The characteristic function describes how much collective payoff a set of players can gain by forming a coalition. Then the game is sometimes called a profit game.

The games can be classified according to their features mainly as *state* or *dynamic* games, *cooperative* or *non-cooperative* games. Therefore, the strategic form and the extensive form are used to describe non-cooperative games. On the other hand, the coalition form is used to describe cooperative games.

In this chapter, we discuss about *cooperative* and *non-cooperative* games. We analyze separately each area, given definitions and games/examples. Also, we discuss and analyze the fundamental concepts of game theory and the connection of this theory with a traditional wireless network.

2.1 Cooperative Games

In game theory, a cooperative game is a game where groups of players may enforce cooperative behaviour. A group of players is also called as a coalition. However, in this game exists a competition between coalitions of players, rather than between individual players. The cooperative games pay attention to the fairness and effectiveness.

An example is the following: we have two cars, which are running on the same narrow road head to head. Then, the drivers should choose a side to swerve in order to avoid the collision. If the drivers cooperate with each other and choose different sides to swerve, they can avoid the collision. If they choose the same side, they can not pass each other. In continuous, a table of this game is represented:

	Left	**Right**
Left	$(0, 0)$	$(5, 5)$
Right	$(5, 5)$	$(0, 0)$

The above table shows us the payoff of pass for one driver as **0** and then the payoff for the other driver is **5**, which means that there is a collision. If the drivers cooperate with each other, the payoff will be **5** to each other. We can observe, that there are two Nash equilibrium: $(5, 5)$ and $(5, 5)$, that they choose different side. Then, thay can have a Pareto efficiency solution.

In the cooperative networks, the agents are involved in the cooperative games. In the cooperative game theory, the wireless agents (users) in a network have an agreement on how to fairly and efficiently share the available spectrum resources [1]. The users may increase the Quality of Service (QoS) via cooperation [9]. In a cooperative communication, the single antenna mobiles in a multi-user environment can share their antennas such as these can generate a virtual multiple-antenna transmitter that allows them to achieve transmit diversity. Thus, in a cooperative communication system, each wireless user can transmit data as well as act as a cooperative agent for an another user, as illustrated in the Figure 2.1.

In this section, we study two types of cooperative games. Thus, we study the bargaining games and the coalitional games.

2.1.1 Bargaining Games

A known game for cooperation is the bargaining game. In the bargaining game, the users have the opportunity to reach a mutually beneficial agreement.

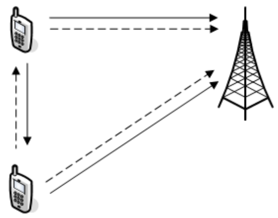

Figure 2.1: Cooperative communication

In [1] the authors considered two players and they are in bargaining game. This bargaining problem is a pair of $(U,(u_1^0,u_2^0))$, where $U \subset \Re_2$ is a compact and convex set and there is a set of possible utility pairs. Therfore, there is at least one utility $(U,(u_1,u_2))$ such that $u_1 > u_1^0$ and $u_2 > u_2^0$. A solution of this problem is (u_1^*,u_2^*)=$f(U,(u_1^0,u_2^0))$.

There is a list of axioms that must be satisfied:

- *Individual rationality*: $u_1^* > u_1^0$ and $u_2^* > u_2^0$.
- *Feasibility*: $(u_1^*,u_2^*) \in U$.
- *Pareto efficiency*: If $(u_1,u_2),(u_1',u_2') \in U$, $u_1 < u_1'$ and $u_2 < u_2'$ then $f(U,(u_1^0,u_2^0)) \neq (u_1,u_2)$.
- *Symmetry*: $(u_1,u_2) \in S \Leftrightarrow (u_2,u_1) \in S$ and $u_1^0 = u_2^0$. Then $u_1^* = u_2^*$.
- *Independence of irrelevant alternatives*: If $(u_1^*,u_2^*) \in U' \subset U$, then $f(U',(u_1^0,u_2^0)) = f(U,(u_1^0,u_2^0)) = (u_1^*,u_2^*)$.
- *Independence of linear transformations*: Let $u_1' = c_1u_1 + c_2$ and $u_2' = c_3u_2 + c_4$ with $c_1,c_3 > 0$. Then $f(U',(c_1u_1^0 + c_2, c_3u_2^0 + c_4)) = (c_1u_1^* + c_2, c_3u_2^* + c_4)$.

The basic element of the cooperative game theory is the **Nash Bargaining Solution (NBS)**. For *two-player* bargaining game, the NBS is as follows:

$$(u_1^*,u_2^*)=\text{arg max } (u_1 - u_1^0) \cdot (u_2 - u_2^0)$$

For *more players* the NBS is as follows:

$$s^*=\text{arg max } \prod_{i=1}^{N}(u_i - u_i^r)$$

where N is the players and u_i^r is the target utility value which the user will realize.

2.1.2 Coalitional Games

An other known game for cooperation is the coalitional game. In this game, a set of players can cooperate with others by forming cooperating groups. Let N is denoted the set of players, that $N = \{1, 2..., n\}$ and $S \subseteq N$, $S \neq \emptyset$ is a coalition. The set of all coalition is denoted by 2^N. The set N is also a coalition and is called as **grand coalition**. Also, there is the **empty coalition** for the empty set. The worth, the value or the power of coalition S is denoted as vS, that the players in S are in cooperation.

For a game with two players, i.e. n=2, then there is a set of 4 coalition $\{\emptyset, \{1\}, \{2\}, N\}$. For a game with three players, there is a set of 8 coalition $\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, N\}$. Then, for n players, we have 2^N coalitions. In continuous, the coaltitonal form is defined as [10]:

Definition 2.1.1 ***Coalitional Form*** *or* ***Characteristic Function Form****: The coalitional form of an n-person game is given by the pair (N, v), where $N = \{1, 2..., n\}$ is the set of players and v is a real-valued function. This form called the characteristic function of the game, defined on the set, 2^N, of all coalitions and satisfying*

- $v(\emptyset) = 0$*, i.e. the empty set has value equals to zero.*
- Superadditivity*: if S and T are disjoint coalitions ($S \cap T = \emptyset$), then $v(S) + v(T) \leq v(S \cup T)$, i.e. the value of two disjoint coalitions is at least as great when they work together as when they work apart.*

Definition 2.1.2 ***Core****: In a coalitional game $< N, v >$, its core is the set of feasible payoff profile $(x_i)_{i \in N}$ for which there is no coalition S and S-feasible payoff vector $(y_i)_{i \in S}$ such that $y_i > x_i$, $\forall i \in S$.*

$$\mathcal{C} = \{(x_i) : \sum_{i \in N} x_i = v(N) \text{ and } \sum_{i \in N} x_i \geq v(S), \forall S \subseteq N\}$$

The core is the set of payoff profiles that satisfy a system of weak linear inequalities, and thus is closed and convex. We can sovle a linear program in order to compute the core:

$$min_{(x_i)_{i \in N}} \sum_{i \in N} x_i$$
$$\sum_{i \in N} x_i = v(N)$$
$$\sum_{i \in N} x_i \geq v(S), \forall S \subseteq N$$

Definition 2.1.3 *A game with no transferable utility or NTU game is a function V that assigns every coalition S a set $V(S) \subset R^S$.*

Definition 2.1.4 *A game V is transferable utility or TU game if for a real-valued function* $v = (v(S))_{S \in C}$, $V(S) = \left\{ u^S \in R^S : \sum_i u_i^S \leq v(S) \right\}$.

The existence of the core depends on the feasibility of the linear program and is related to the *balanced property* of a game. A coalitional game *with transferrable* payoff is called *balanced* iff holds the above:

$$\sum_{S \subseteq N} \lambda_S v(S) \leq v(N), \forall \lambda = (\lambda_S)_{S \subseteq N},$$

where λ is a non-negative weight collection.

Theorem 2.1.1 *A coalitional game* with transferrable *payoff has a* non-empty core ***iff*** *it is* balanced.

Shapley proposed a solution concept, known as the Shapley value ψ. In each player in the game is assigned a unique payoff value. The value ψ_i denotes the payoff assigned to player i accordding to the Shapley value.

In continuous, there is a list of axioms that must be satisfied:

- *Symmetry*: If player i and player j are interchangeabl in v, i.e. $v(S \cup i) = v(S \cup j)$ for every coalition S that does not contain player i or j, then $\psi_i(v) = \psi_j(v)$.
- *Dummy player*: If player i is a dummy in v, i.e. $v(S) = v(S \cup i)$ for every coalition S, then $\psi_i(v) = 0$.
- *Additivity*: For any two games u and v, define the game $u + v$ by $(u + v)(S) = u(S) + v(S)$, then $\psi_i(u + v)(S) = \psi_i(u) + \psi_i(v)$ for all $i \in N$.
- *Efficiency*: $\sum_{i \in N} \psi_i(v) = v(N)$.

The Shapley value is the only value that satisfies all the above axioms, and is usually calculated as the expected marginal contribution of player i when joining the grand coalition given by

$$\psi_i(v) = \sum_{S \subseteq |N| - i} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [v(S \cup i) - v(S)].$$

2.2 Non-cooperative Games

In the non-cooperative networks, the agents are involved in the non-cooperative games. The basic element of the non-cooperative game theory is the Nash equilibrium [1]. A Nash equilibrium of a game is the strategy profile such that no agent can decrease his individual cost by unilaterally changing his strategy.

In this subsection, we will see the meanings of the strategic form, the Nash equilibrium and when a strategic game is pure strategy and when is mixed strategy. A game can have complete information or incomplete information. Also, a game can be *static* which means that the players have not more than one move. But a game can be *dynamic*, which means that the players can observe previous moves or data and after they make a new decision to move. Therefore, the dynamic games are general game models.

In continuous, we will introduce some conditions that can gurantee us the uniqueness of equilibrium. A game can have more than one equilibrium, then there are also some conditions in order to select one of these equilibrium. So, we will see the definition of *Pareto optimality*, which is widely used in the game theory.

2.2.1 Strategic Form

The components in a *strategic form game model* are the N set of players, the A_i set of actions for each player i and the u_i: $A \rightarrow \Re$, which is the payoff/utility function. The payoff function measures usually the outcome for player i of a stage. While, the utility is usually used for the outcome for a player i, which determined by the actions of all players (whole game). This strategy game can be represented as $G =< N, (A_i), (u_i) >$.

The utility or payoff u_i expresses the benefit of player i given the strategy profile s. However, the normal form is a matrix representation of a *static game*. A static game can be also called as *simultaneous game*. A static game is one in which all players make decisions without knowledge of the strategies that are being chosen by other players. But there is the case, that the players make decisions at different points in time, then the game is also simultaneous because each player has no information about the decisions of others.

A game with normal form is usually represented by a matrix which consists of the players, their strategies, and their payoffs. Thus, Let we have a game with two players, then one is the "row" player, and the other, the "column" player. Each rows or column represents a strategy and each box represents the payoffs to each player for every combination of strategies. Generally, such games are solved using the concept of a Nash equilibrium. The definition of the Nash equilibrium is refered in the next paragraph.

The prisoner' dilemma is an example of a static game. An other example is the "**Rock-Paper-Scissor**" game. Both of the players can make a decision at the same time, randomly. Each of the players have no information about the decisions of the other. Therefore, there is a game with two players and each of them has 3 different strategies to make decision. Then, we have a 3×3 table because of the combination of their strategy profiles forms. Thus, the Player 1's strategies are denoted as rows and Player 2's strategies as columns. In each cell, the first number represents the payoff to the Player 1 and the second number represents the payoff to the Player 2.

p_1/ p_2	**Rock**	**Paper**	**Scissor**
Rock	$(0,0)$	$(-1,1)$	$(1,-1)$
Paper	$(1,-1)$	$(0,0)$	$(-1,1)$
Scissor	$(-1,1)$	$(1,-1)$	$(0,0)$

In this examplee, we have the following:

- Players: $N = 1, 2$
- Actions: $A_1 = A_2 = Rock, Paper, Scissor$
- Payoffs for Player 1:
  $r_1(Rock, Scissor) = r_1(Scissor, Paper) = r_1(Paper, Rock) = 1,$
  $r_1(Scissor, Rock) = r_1(Paper, Scissor) = r_1(Rock, Paper) = -1,$
  $r_1(Rock, Rock) = r_1(Paper, Paper) = r_1(Scissor, Scissor) = 0,$
- Payoffs for Player 2: $r_2(a) = -r_1(a)$, for each a.

However, this game is a zero-sum game because of the property:

$$\sum_i r_i(a) = 0.$$

Which means, that one player's gain is exactly the other players' loss.

Definition 2.2.1 *A game in strategic form is defined as* ***finite*** *if (i) N the set of players is finite and (ii) the A_i for each $i \in N$ are finite.*

Mixed Strategies

In the *pure strategies*, the players clearly decide on one behavior or another [6]. While in the *mixed strategies*, the players can decide to play each of these pure strategies with different probabilities. Below is defined the mixed strategy [6].

Definition 2.2.2 Mixed Strategy: *The mixed strategy $\sigma_i(a_i)$ or else σ_i of player i is a probability distribution over his pure strategies $a_i \in A_i$.*

Let Σ_i is the mixed strategy space of player i, where $\sigma_i \in \Sigma_i$. The notion of profile is chacterized by the probability distribution assigned by each player to his pure strategies, i.e. $\sigma = \sigma_1, ..., \sigma_N$. The strategy profile of the opponents players is denoted as σ_{-i}. The utility to profile σ for the player i is defined as

$$u_i(\sigma) = \sum_{a_i \in A_i} \sigma_i(a_i)\, u_i(a_i, \sigma_{-i})$$

2.2.2 Extensive Form

In a game, the players can have a **sequential interaction**, which means that the move of one player is conditioned by the move of the other player. These games are called **dynamic games** and can be represented in an *extensive form*. An extensive form game is a **game tree**. This tree is a rooted tree where each non-terminal node represents a choice that a player must make, and each terminal node gives payoffs for all players. A game in extensive form can be analyzed directly or can be converted into an equivalent strategic form. The extensive form can seperate into two categories, **the extensive form with perfect information** and **the extensive form with imperfect information**.

In an extensive game with perfect information, every player is at any point aware of the previous choices of all other players. The players have a sequential interactions, which means that only one player moves at a time. These games can be analyzed by **backward induction**. This technique solves the game by first considering the last possible choices in the game.

2.2.3 Nash Equilibrium & Existence

Definition 2.2.3 Nash Equilibrium (NE): *A Nash equilibrium of a strategic game with components N, A_i and u_i is a profile $a' \in A$ of actions such that $\forall i \in N$ then*

$$u_i(a'_i, a'_{-i}) \geq u_i(a_i, a'_{-i}) \;\forall a_i \in A_i$$

where a_i is the strategy of player i and a_{-i} is the strategies of all players other than player i.

The Nash equilibrium defines the best response strategy of each player. Thus, no player can improve his payoff by a unilateral deviation from the NE, given that the other players adopt the NE.

Theorem 2.2.1 Kakutani's Fixed Point Theorem: *Let X be a compact convex subset of* $\Re^n$ *and let* $f : X \to X$ *be a set-valued function such that the set* $f(x)$ *is nonempty and convex for all* $x \in X$ *and the graph of* f *is closed. Then, there exists* $x* \in X$ *such that* $x* \in f(x*)$.

Definition 2.2.4 Closed Graph: *A set-valued function* $f : X \to X$ *is said to have a closed graph if the set* $\{(x, y) \,|y \in f(x)\}$ *is a closed subset of* $X \times Y$ *in the product topology i.e. for all sequences* $\{x_n\}$ *and* $\{y_n\}$, $n \in N$ *such that* $x_n \to x$, $y_n \to y$ *and* $y_n \in f(x_n), \forall n \in N$, *we have* $y \in f(x)$.

From the Kakutani's Fixed Point Theorem is denoted the next theorem for the existence a Nash equilibrium in a stretegic game [1].

Theorem 2.2.2 *A strategic game with components N,* A_i *and* u_i *has NE if* $\forall i \in N$, *the set* $A_i \neq \varnothing$ *is a compact convex subset of a Euclidian space and the payoff function* u_i *is continuous and quasi-concave on* A_i.

Definition 2.2.5 Mixed Strategy NE: *The mixed strategy NE of a strategic game is a NE where players' strategies are non-deterministic but are regulated by probabilistic rules.*

Theorem 2.2.3 [**Nash 1951**]*: Every finite* ***strategic game*** *has a mixed strategy Nash equilibrium.*

Theorem 2.2.4 [**Kuhn 1953**]*: Every finite* ***extensive game*** *of perfect information has a pure strategy Nash equilibrium.*

2.2.4 Uniqueness of an Equilibrium

From the above theorem, we have the existence property of a mixed strategy Nash equilibrium in games. An other property is the uniqueness of an equilibrium [1]. If the feasible region and the payoff function of each user are convex shapes then there is a unique equilibrium in the game.

The NE gives the best strategy given that all the other players persist to their equilibrium strategy too. There is an aspect to find the NE. If the players adjust their strategies iteratively based on accumulated observations as the game unroll then the process could converge to some equilibrium point. A such case that can guarantee us the convergence to the NE is the *potential game*. The idea of potential games was first proposed by Monderer and Shapley (1996) [11].

In game theory, a game with components N, A_i and u_i is said to be a **potential game** if there is a potential function $P : A \to \Re$ such that one the following conditions holds:

- **Exact potential game**: $P(a_i, a_{-i}) - P(a'_i, a_{-i}) = u(a_i, a_{-i}) - u(a'_i, a_{-i})$, $\forall i \in N, a \in A, a'_i \in A_i$.

- **Ordinal potential game**:
  $sgn\,(P(a_i, a_{-i}) - P(a'_i, a_{-i})) = sgn\,(u(a_i, a_{-i}) - u(a'_i, a_{-i}))$,
  $\forall i \in N, a \in A, a'_i \in A_i$, where $sgn(\cdot)$ is the sign function.

Theorem 2.2.5 *A strategic game is an exact potential game with a potential function P():*

- *iff* $\frac{\partial^2 u_i}{\partial u_i \partial u_j} = \frac{\partial^2 u_j}{\partial u_i \partial u_j}$ $\forall i, j \in N$

- *iff there are functions Po:* $A \to \Re$ *and Pi:*$A_{-i} \to \Re$ *such that* $u(a_i, a_{-i}) = Po(a_i, a_{-i}) + Pi(a_{-i})$, $\forall i \in N$, *where* $P(a_i, a_{-i}) = Po(a_i, a_{-i})$

- *if there exist functions* $P_{ij} : A_i \times A_j \to \Re$ *and* $P_i : A_i \to \Re$ *such that* $P_{ij}(a_i, a_j) = P_{ji}(a_j, a_i)$ *and* $u_i(a) = \sum\limits_{j \in N} P_{ij}(a_i, a_j) - Pi(a_i)$, $\forall i, j \in N$ *and* $a \in A$.

This game is known as ***bilateral symmetric game****:*

$$P(a) = \sum_{i \in N} \sum_{j=1}^{i-1} P_{ij}(a_i, a_j) - \sum_{i \in N} P_i(a_i)$$

An other case that can guarantee us the convergence to the uniqueness of equilibrium is the *standard function.* In [12], proposed an interference function $I(\mathbf{p})$ in order to reduced the problem of the uplink power control in cellular networks. Yates defined the inequality $p_i \geq I(\mathbf{p})$, where $\mathbf{p} = (p_1, ..., p_n)$ is the power vector of the N users, $I_i(\mathbf{p}) = (I_1(\mathbf{p}), ..., I_n(\mathbf{p}))$ is the interference of other users that user i must overcome.

Definition 2.2.6 [**YATES**] : *An Interference function* $I(p) = (I_1(\boldsymbol{p}), ..., I_n(\boldsymbol{p}))$ *is* ***standard*** *if for all* $p \geq 0$, *the following properties are satisfied:*

- *Positivity:* $I(p) > 0$, *if* $p > 0$

- *Monotonicity: if* $p \geq p'$ *then* $I(p) \geq I(p')$

- *Scalabity:* $\forall a > 1$, $a \cdot I(p) > I(a \cdot p)$

Theorem 2.2.6 [**YATES**] : *If* $I(p)$ *is feasible, then for any initial power vector* p, *the standard power control algorithm converges to a unique fixed point* $p*$.

2.2.5 More than one Equlibrium?

A non-cooperative game can have more than one equilibrium. In such a game, we wonder if some points outperform others and if there is an optimal one. Then, we have to face multi-objective optimization problems, that these problems are not easy to define the optimaity. So, we can reduce this multi-dimension problem into one-dimension one and it is achieved from the comparison of the weighted sum of the individual payoffs.

Pareto Optimality

The Pareto optimality or Pareto efficiency is a payoff profile that no player can improve its own utility without reducing the utility any other player. Therefore, if there are more than one equilibrium points, usually the optimal ones in the Pareto sense are preffered [1]. The term is named after Vilfredo Pareto, an Italian economist who used the concept in his studies of economic efficiency and income distribution [13]. The Pareto optimality has been used in economics, social sciences and also in wireless networks.

Definition 2.2.7 Pareto-optimal/efficient: *Given a set of choices and a way of valuing them, the Pareto frontier or Pareto set or Pareto front is the set of choices that are Pareto efficient. The utility* $\boldsymbol{u}$ *belongs to a set* $U \in \Re^N$ *is Pareto efficient if there is no* $\boldsymbol{u}' \in U$ *for which* $u_i' > u_i$, $\forall i \in N$. *The utility* $\boldsymbol{u}$ *belongs to a set* $U \in \Re^N$ *is strongly Pareto efficient if there is no* $\boldsymbol{u}' \in U$ *for which* $u_i' \geq u_i$, $\forall i \in N$ *and* $u_i' > u_i$ *for some* $i \in N$.

Equilibrium Refinement

Sometimes, we have to face multiple Nash equilibria in a game. From these equilibrium solutions, we may have non desirable or non reasonable outcomes, then it is necessary to *refine* them. Therefore, an **equilibrium refinement** provides a way of *selecting* one or a few equilibria from among many in a game. Each refinement attempts to define some equilibria as more likely, more rational or more robust to deviations by players than others. For example, if one equilibrium Pareto dominates another, then it may be viewed as more likely to be chosen by the players.

Firstly, the concept of equilibrium refinement is proposed by Selten (1975), Myerson (1978). Selten introduced the concept of a perfect equilibrium. To define it, the mixed extension of a finite strategic game $G = (N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N})$ is considered. Also, there exists at least one perfect equilibrium in any finite game [14].

Definition 2.2.8 ***ϵ-Perfect Equilibrium***: *An ϵ-Perfect Equilibrium G is a strictly mixed strategy x^ϵ, such that, for each player i, $x_i^\epsilon \in argmax_{pi} u_i(p_i, x_{-i}^\epsilon)$ subject to $x_i^\epsilon(a_i) \geq \epsilon(a_i)$ for some $\epsilon(a_i)$ where $0 < \epsilon(a_i) < \epsilon$.*

Definition 2.2.9 ***Perfect Equilibrium***: *A Perfect Equilibrium is any limit of ϵ-constraint equilibria as ϵ goes to zero.*

Let see an example/game with two players:

player 1/ player 2	x_1	x_2
y_1	$(1,1)$	$(2,0)$
y_2	$(0,2)$	$(2,2)$

We can observe that the matrix game has two Nash equiliibrium (y_1, x_1) and (y_2, x_2). The second equilibriumm solution is perfect. It can be confirmed that for the first equilibrium, if Player 1 plays y_1 with probability $1-\epsilon$ and y_2 with probability ϵ, the player 2 has no interest in deviating from his equilibrium action x_1. Respectively process for the player 2 is applied. On the other hand, if player 1 plays y_2 with probability $1-\epsilon$ and y_1 with probability ϵ, the player 2 has a better expected utilty by deviating from his equilibrium action x_2. Respectively process for the player 2 is applied.

We have to note that the meaning of perfect information is different from the meaning of complete information, respectively the concepts of imperfect and incomplete information. In games with complete information, it is assumed that the data of the game is common knowledge. Considering a strategic form game, which means that the actions available to the players and the utility functions are common knowledge. Every player knows the data of the game and all players know that the opponent players know the data of the game. The games with incomplete information are known as Bayesian games in game theory. The players have only partial information about the game. But in a game with perfect information, all the players have perfect knowledge of the history in the game, and imperfect information otherwise [6].

2.2.6 Examples/Games

In this subsection, we introduce some non-cooperative games [6]. These games are applied in a wireless network. The players are the users controlling their devices, so we denote their devices as players. We assume that the players are rational, they try to maximize their payoffs or to minimize their costs according to their strategies.

Static Games

1. Forwarder's Dilemma Game

Forwarder's Dilemma is a *symmetric nonzero-sum game*, beacause the players can mutually increase their payoffs by cooperating. This game can be classified in the *network layer*. The Forwarder's Dilemma is regarded as the Prisoner's Dilemma in classical game theory. We assume that there is a two-player game. Let p_1, p_2 are two players. Each of them has a packet and want to transmit to his receiver, r_1, r_2, correspondigly. The communication between p_i and his r_i is possible if the other player p_j $(j \neq i)$ forwards the packet. The payoff is equal to the difference of the reward and the cost. The cost c is fixed, $0 < c << 1$. In this example c is the energy and computation spent for the forwarding action. This action is able to enable the communication between p_2 and r_2, which gives p_2 a reward of 1. Then the dilemma is that each player can be found in temptation to drop the packet that it should forward to the corresponding receiver, as this would save some of his resources. Let p_2 the player, that should forward the packet but he drop it. Then, the packet that the p_1 wanted to be relayed will be dropped. The solution of this problem is that they could do by mutually relaying each other's packet. We can see the network scenario in this game, as follows:

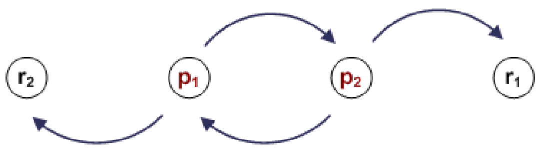

Figure 2.2: Forwarder's Dilemma Game

This game can be represented by the following table, whch is a strategic-form representation. In this table, the corresponding columns and rows are the payoffs of the actions of players p_1, p_2. The strategy options of players are **F** forward packet of the other player and **D** drop packet.

p_1/ p_2	**F**	**D**
F	$(1-c, 1-c)$	$(-c, 1)$
D	$(1, -c)$	$(0, 0)$

This game can be solved in several ways, such as *strict dominance*, *iterated strict dominance*. Strict dominance is the strictly best strategy for any strategy of the other players. Let strategy s_i is the strategy for player i and is said to be strictly dominates if $u_i(s_i', s_{-i}) < u_i(s_i, s_{-i})$, $\forall s_{-i} \in S_{-i}$, $\forall s_i' \in S_i$.

We can observe, that the strategy **D** strictly dominates the strategy **F**. In continuous, we solve the game by iterated strict dominance. From the first row of the matrix, we can see, that the player p_1 will never choose the **F** strategy. From the first column of the matrix, we can see, that the player p_2 will never choose the **F** strategy. Then, the solution of the game is (D, D) and the payoff is $(0, 0)$. This is paradox, as the pair (F, F) would have led to a better payoff for each of the players.

According to the definition of Nash equilibrium, that none of the players can unilaterally change its strategy to increase its payoff/utility. Therefore, the pair (D, D) is the Nash equilibrium.

2. Joint Packet Forwarding Game

The players in the Joint Packet Forwarding Game are not in a symmetric situation. Also, this game can be classified in the *network layer*. We assume that there is a two-player game. Let s is a sender and wants to send a packet to the receiver r in each time slot. In order to transmit this packet, sender needs players/devices p_1 and p_2 to forward. The cost c is fixed, $0 < c << 1$, if a player forwards the packet of the sender. If both players p_1 and p_2 forward, then they each receive a reward of 1. We can see the network scenario in this game, as follows:

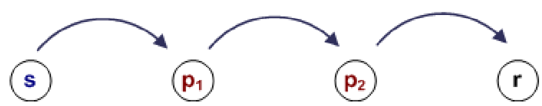

Figure 2.3: Joint Packet Forwarding Game

This game can be represented by the following table, that the corresponding columns and rows are the payoffs of the actions of players p_1, p_2. The strategy options of players are **F** forward packet of the other player and **D** drop packet.

p_1/ p_2	**F**	**D**
F	$(1-c, 1-c)$	$(-c, 0)$
D	$(0, 0)$	$(0, 0)$

We can observe, that none of the strategies of any player strictly dominates the other. In order to overcome this problem, is defined the meaning of *weak dominance*. Weak dominance is the strictly better strategy for at least one opponent strategy. Let strategy s'_i is the strategy for player i and is said to be weakly dominated by strategy s_i if $u_i\left(s_i', s_{-i}\right) \leq u_i\left(s_i, s_{-i}\right)$, $\forall s_{-i} \in S_{-i}$, $\forall s_i' \in S_i$.

If player p_2 uses the strategy **D**, then it is weakly dominated by the the strategy **F**. Using the elimination based on iterated weak dominance, which results in the strategy profile (F, F). An important notation is that the result of the iterative weak dominance is not unique in general. While the solution of the iterated strict dominance is unique.

In addition, the pairs (F, F) and (D, D) are Nash equilibria.

3. Multiple Access Game

Multiple Access Game is a *nonzero-sum game.* In this game, the players have to share a common resource, the wireless medium. This example can be classified in the *medium access layer.* We assume again that there is a two-player game. Let p_1 and p_2 players, who want to send some packets to their receivers r_1, r_2, correspondigly. This transmission can be done through a shared medium. The players have a packet to send in each time slot. Each player have one move in each time slot. So, they can decide to transmit this packet or not. The transmission cost c for a player p_1 is fixed, $0 < c << 1$. If p_2 does not transmit in a time slot and the p_1 send a packet to r_1 in that time slot, then the transmission of this packet is successful. Then the p_1 gets a reward of 1. On the other hand, if both players want to transmit in the same time slot then there is a collision.

This game can be represented by the following table, that the corresponding columns and rows are the payoffs of the actions of players p_1, p_2. The strategy options of players are **T** transmit packet to the other player and **NT** not transmit packet.

p_1/ p_2	**T**	**NT**
T	$(0, 0)$	$(0, 1-c)$
NT	$(1-c, 0)$	$(-c, -c)$

We can observe that there is not a strictly dominating strategy. In order to give us a solution this game, is using the concept of best response, i.e. $b(s_{-i}) = \arg\max_{s_i \in S_i} u_i\left(s_i, s_{-i}\right)$. The best response is not unique in general. There are two Nash equilibria, the pair (NT, T) and the pair (T, NT). This game has a mixed-strategy Nash equilibrium, the pair $(p = 1-c, q = 1-c)$. We denote p the probability that player p_1 transmit the packet and q the probability that player p_2 transmit the packet. The payoff of player p_1: $u_1 = p\left(1-c-q\right)$ and the payoff of player p_2: $u_2 = q\left(1-c-p\right)$.

4. Jamming Game

On the other hand, Jamming Game is a *zero-sum game*. This example can be classified in the *physical layer*. We assume that the wireless medium is split into two channels CH_1 and CH_2, accordind to the Frequency Division Multiple Access (FDMA). We assume again that there is a two-player game. Let the player p_1 wants to send a packet in each time slot to his receiver r_1. The player p_2 is a *malicious player*. His objective is to prevent p_1 from a successful transmission by transmitting on the same channel in the given time slot. The aim of p_1 is to have a successful transmission in spite of the presence of p_2. The p_1 gets a payoff of 1 if p_2 can not jam his transmission. He gets a payoff of -1 if p_2 can jam the transmission of the packet. The payoffs for the p_2 are the opposite of those of p_1. For transmission cost c, then each payoff would be 1-c and -1-c. Therefore, we have a zero-sum game:

$$\sum_{i \in N} (reward_i - cost_i) = 0$$

This game can be represented by the following table, that the corresponding columns and rows are the payoffs of the actions of players p_1, p_2. The strategy options of players are CH_1 channel 1 and CH_2 channel 2.

p_1/ p_2	CH_1	CH_2
CH_1	$(-1,1)$	$(1,-1)$
CH_2	$(1,-1)$	$(-1,1)$

The Jamming game can not be solved by iterated strict dominance. Also, this game has not pure-Nash equilibrium. There are only a mixed-strategy Nash equilibrium, that the players play a uniformly random distribution strategy. For example, each one of the player select one of the channel with 1/2 probability.

Dynamic Games

1. Sequential Multiple Access Game

Sequential Multiple Access Game is a modified version of the Multiple Access Game. Which means that the moves of players are not in synchronized. Thus, this game is characterized by the extensive form. So, these players have a sequential interaction. We assume again that there is a two-player game. Let p_1 and p_2 players, who want to send some packets to their receivers. The player p_1 always moves first played one of his strategies and the player p_2 observes the move of p_1 before making his own move. The strategy options of players are **T** transmit packet to the other player and **NT** not transmit packet. The set of possible strategies for the p_1 is the $\{T, NT\}$, the set of possible strategies for the p_2 is the $\{TT, TNT, NTT, NTNT\}$. The strategy TNT means that p_2 transmits if p_1 transmits and remains in the state of non transmit if p_1 does not transmit. The transmission cost c for a player p_1 is fixed, $0 < c << 1$. The player gets a reward of 1 if there is a successful transmission. This game can be represented by the following tree:

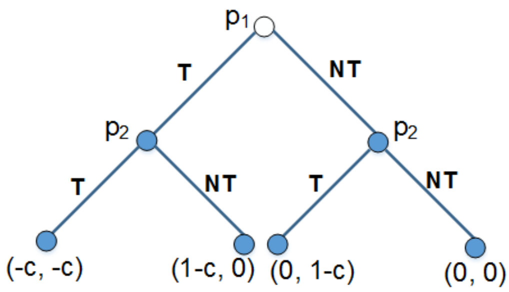

Figure 2.4: Sequential Multiple Access Game

There are three pure Nash equilibria, the pairs (T, NTT), $(T, NTNT)$ and (NT, TT). We can observe that if p_2 plays the strategy TT, then the best response of the p_1 is to paly NT. The move TT from the p_2 is a destruction, as this strategy is not the best strategy for the player p_2 because of the choice of the player p_1, i.e. if p_1 plays T in the first round. This destruction is called as *incredible threat*.

Then, the technique of *Backward Induction* is able to avoid the destructive equilibrium. This game with the backward induction technique can be represented by the following tree:

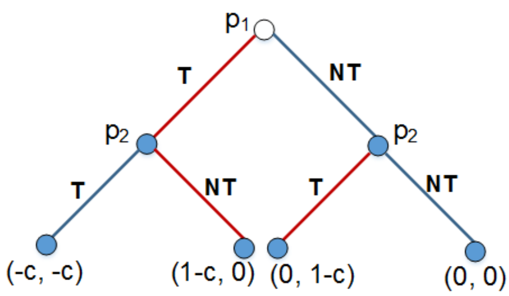

Figure 2.5: The backward induction solution of the Sequential Multiple Access Game

This game is finite and belongs to the case of complete information. The player p_2 knows that he will play the last move. Let that the history h in the game is T, then the p_2 notes that the move NT is the best payoff for him to play in the last stage. The move T is the best for the p_2, if the p_1 has NT. The best choices of the game are the red lines of the tree, that is represented in Figure 2.5. Given all the best moves of p_2 in the last stage, then the p_1 finds his best moves as well. Using the backward induction technique, we arrive at the root of the tree.

We conlclude that this game is perfect information, as each player knows at which node he is when he makes his decision. In addition, it belongs to the game of complete information. We also assume that the players are reliable. Then, this method will give a unique prediction. Therefore, the backward induction solution is the $(1-c, 0)$, as $h = \{T, NT\}$.

Chapter 3

Regret Learning Techniques

In the real life, we are making decisions in order to face some difficulties such as the decision for the optimal route to drive from home to work each day, as in the Figure 3.1. Thus, in this example, we have to face a repeated play of a game against an opponent with an unknown strategy. So at each time step, the algorithm probabilistically chooses an action and then incurs the loss for its action chosen, such as how long its route took. In the next day, all this process is repeated. Then, we have to solve a dynamic system when there are multiple players, all adjusting their behavior in such a way. The regret analysis is an important technique to try to solve such as analyzing problems. The space of study consists of a finite number of actions.

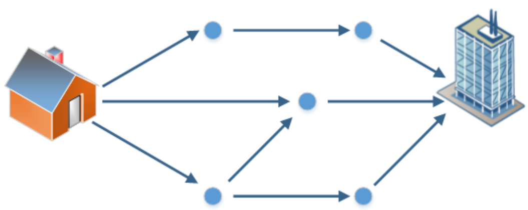

Figure 3.1: A decision is picked for the optimal route from home to work.

In [15],[16] are referred that the regret minimizing algorithms become known from the author Hannan [17]. Hannan was the first, who presented the *repeated two-player games* 60 years ago. The *regret* is defined as a measure of the quality of a sequence of actions. According to [16], the regret consists of two categories, the *external regret* and the *internal regret*. The *external regret* is also known as best *expert problem*. This category compares the performance an online algorithm to the best of a finite number of actions. On the other hand, the *internal regret* compares the loss of an online algorithm to the loss of a modified online algorithm, which consistently replaces one action by another. The notion of internal regret introduced first by Foster and Vohra [18]. In continuous, Blum and Mansour introduced the notion of *swap regret*. So, this algorithm is a simple online algorithm,which can efficiently convert any low external regret algorithm into a low internal regret algorithm. This means that the swap regret algorithm is stronger than internal regret algorithm, as well the former allows simultaneously swap multiple pairs of actions. The internal and the swap regret are tight connencted to the *correlated equilibria*.

Also, the authors in [16] give us a full information model and a partial information model, as follows:

Definition 3.0.10 [Blum and Mansour]: *Let N available actions X= {1,...,N}, an algorithm H selects a distribution* p^t *over the N actions,* $l_i^t \in [0,1]$ *is the loss of the i-th action at time t and a loss vector* $l^t \in [0,1]^N$*. In the* ***full information model****, the online algotithm H receives the loss vector* l^t *and experience a loss* $l_H^t = \sum_{i=1}^N p_i^t l_i^t$.

Definition 3.0.11 [Blum and Mansour]: *Let N available actions X= { 1,...,N }, an algorithm H selects a distribution* p^t *over the N actions,* $l_i^t \in [0,1]$ *is the loss of the i-th action at time t and a loss vector* $l^t \in [0,1]^N$*. In the* ***partial information model****, the online algotithm H receives* $(l_{k^t}^t, k^t)$*, where* k^t *is distributed according to* p^t *and* $l_H^t = l_{k^t}^t$ *is its loss. The loss of the ith action during the first T time steps is* $L_i^t = \sum_{t=1}^T l_i^t$ *and the loss of H is* $L_H^t = \sum_{t=1}^T l_H^t$.

In continuous, are given the most known definitions [19] about no-external regret learning and no-internal regret learning. Specifically, a no-external regret via a multiplcative updating scheme is achieved by Foster and Schapire. In 1997, Foster and Vohra prosposed the no-internal regret learning, that depends on complete payoff information at all times t and including also information about strategies, that are not employed at time t.

Definition 3.0.12 [Freund and Schapire (96)]: *No-external-regret learning converges to the set of minimax equilibria, in zero-sum games.*

Definition 3.0.13 [Foster and Vohra (97)]: *No-internal-regret learning converges to the set of correlated equilibria, in general-sum games.*

In 1974, Robert J. Aumann [20] was the first, who proposed the concept of correlated equilibrium. This new concept of correlated equilibrium is more general than Nash equilibrium. Thus, a strategy profile is chosen randomly according to a certain distribution. When a strategy is recommended then the users have to conform with this strategy. So, the distribution is called the correlated equilibrium [21]. The definition of correlated equilibrium is as follows [1]:

Definition 3.0.14 CE: *A correlated equilibrium of a strategic game with components: a finite set of players, denoted by N, a set of possible actions for each player i, denoted by A_i and a set of utility functions, denoted by $u_i : A \to \mathbb{R}$ consists of:*

- *a finite probability space (Ω, π), where Ω is a set of states and π is a probability measure on Ω,*
- *an information partition $\mathcal{P}_i$ of Ω_i, $\forall i \in N$,*
- *a function $\sigma_i : \Omega \to A_i$, which represents strategy of player i and maps an observed state to an action, with $\sigma_i(\omega) = \sigma_i(\omega')$, $\omega, \omega' \in P_i$, for some $P_i \in \mathcal{P}_i$,*

such that

$$\sum_{\omega \in \Omega} \pi(\omega) u_i(\sigma_{-i}(\omega), \sigma_i(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega) u_i(\sigma_{-i}(\omega), \tau_i(\omega))$$

for all $i \in N$ and any strategy function $\tau_i()$.

In [19], is defined a general class of no-regret learning algorithms referred as Φ-no-regret learning algorithms. Also, they showed that the no-external-regret and the no-internal-regret are special cases of Φ-no-regret.

Definition 3.0.15 [Greenwald]: *Let Φ be a finite of the set of linear maps $\phi : \Delta(A) \to \Delta(A)$: $\forall \alpha, 0 \leq \alpha \leq 1$, $\forall q_1, q_2 \in \Delta(A)$*

$$\phi(\alpha q_1 + (1-\alpha) q_2) = \alpha \phi(q_1) + (1-\alpha)\phi(q_2)$$

Each $\phi \in \Phi$ converts one nondeterministic action for an agent into another.

Proposition 3.0.1 [Greenwald]: *If learning algorithm A satisfies no-internal-regret, then A also satisfies Φ-no-regret, for all finite subsets Φ of the set of stochastic matrices.*

In continuous of this chapter, we will see fundamentals concepts such as one-shot game, transforamtion, the definition of regret game that depends on these previous concepts. Also, we will see the external regret minimization and its connection with the game theory, the interrnal regret, the swap regret and a generic reduction from swap to external regret. Finally, we will give examples to each of them and some bounds.

3.1 More Preliminaries

3.1.1 One-Shot Game

Definition 3.1.1 ***One-Shot Game***: *The real-valued, one shot game can be represented as a triple* $\Gamma = \langle N, < A_i >_{i \in N}, < r_i >_{i \in N} \rangle$, *where* N *is a finite set of players,* A_i *is the set of actions for each i player and* $r_i : \prod_j A_j \rightarrow \mathbb{R}$ *is the reward function for player i.*

Thus, each player i independently selects an action from A_i and receives a reward (payoff/utility) according to its reward function. That is, if each player j plays action $a_j \in A_j$, then player i obtains reward r_i. This one-shot game can be represented in a table. An example is the "**Rock-Paper-Scissor**" game. This game is analyzed in the previous chapter.

3.1.2 Transformations

The concept of an *action transformation* or else *transformation* serves as the basis for our definitions of both equilibria and regret. An action transformation is denoted as ϕ, which is a measurable function from a set of actions A to itself, $\phi : A \rightarrow A$ [22]. Measurability is defined with respect to the σ-algebra associated with the action set.

A σ-algebra or sigma-algebra or σ-field or sigma-field is an important concept in mathematical analysis and in probability theory. Thus, a σ-algebra on a set X is a collection of subsets of X that is closed under countable-fold set operations. Let σ-algebra $\mathcal{F}$ of subsets of X that are satisfied the following conditions:

- $\emptyset \in \mathcal{F}$.
- If $B \in \mathcal{F}$ then its complement B^c is also in $\mathcal{F}$.
- If $B_1, B_2, ...$ is a countable collection of sets in $\mathcal{F}$ then their union $\cup_{n=1}^{\infty} B_n$.

Let A denote an action set and a function $\phi : A \rightarrow \Delta(A)$ denote an action transformation, where $\Delta(A)$ is denoted as the set of probability distributions over the set A. Then, there are several sets of action transformations such as the set of swap transformations $\Phi_{SWAP}(A)$, the set of external transformations $\Phi_{EXT}(A)$, the set of internal transformations $\Phi_{INT}(A)$, the set of σ transformations $\Phi_{\sigma}(A)$. Thus, the $\Phi_{SWAP}(A)$ is the set of all action transformations that map actions to pure strategies. An external action transformation is simply a constant transformation, so for $a \in A$, is defined as

$$\Phi_{EXT}^{(a)}(x) = \delta_a, \forall x \in A$$

An internal transformation act as the identity, except on one particular input, so for $a, b \in A$, is defined as

$$\Phi_{INT}^{(a,b)}(x) = \begin{cases} \delta_b & , x = a \\ \delta_x & , \text{otherwise} \end{cases}$$

The external and the internal action transformations are subsets of Φ_{SWAP}. Also, we have that $|\Phi_{SWAP}(A)| = |A|^{|A|}$, $|\Phi_{INT}(A)| = |A|^2$ and $|\Phi_{EXT}(A)| = |A|$. An action transformation can be extended as a strategy transformation. Let $[\phi] : \Delta(A) \to \Delta(A)$ is the linear transformation and is defined as

$$[\phi](q) = \sum_a q \cdot \phi(a)$$

Let a measurable set $S \subset A_i$ and an action $a \in A_i$ then the set of σ transformations $\Phi_\sigma^{S \to a}$ is defined as

$$\Phi_\sigma^{S \to a}(x) = \begin{cases} \delta_a & , x \in S \\ \delta_x & , \text{otherwise} \end{cases}$$

Note that $\Phi_{INT}(A_i) \subseteq \Phi_\sigma(A_i)$ for any A_i and a measurable set S. Also, the set of σ transformations is as powerful as Φ_{SWAP}.

In introduction of this chapter, we give the definition of the *correlated equilibrium.* Below, we give an other definition, Φ-Equilibrium [19].

Definition 3.1.2 ***Φ-Equilibrium***: *Given a game and a collection of sets of transformations $\Phi = < \Phi_i >_{i \in N}$, a probabilty distribution q over the set of all possible actions A is called a Φ-equilibrium iff*

$$\mathbb{E}\left[r_i(\phi(a_i), a_{-i}) - r_i(a)\right] \leq 0,$$

for all players i and for all $\phi_i \in \Phi_i$.

3.1.3 Repeated Game

There are the cases that the players interact repeatedly over time. Thus, the repeated game model is an extensive-form game in which the same stage game is played at each date for some duration of T rounds. The rounds may be finite or infinite. The repetition of the same game might foster cooperation. The case of infinitely repeated game is more interesting, because of the players can care about their future payoff except of their current payoff. So, a current behavior of a player can affect the other players in the future.

Definition 3.1.3 ***Infinitely Repeated Game***: *A strategic game can be represented as a triple* $< N, (A_i), (u_i) >$. *A* δ*-discounted infinitely repeated game is an extensive-form game with perfect information and simultaneous moves, having the following:*

- *N, that is the set of players.*
- a^t *that is the chosen action, which depends on the history* $(a^1, a^2, ..., a^{t-1})$.
- A_i, *that is the set of actions available to any player i, regardless of any history.*
- *The payoff function for player i is the discounted average of immediate payoffs from each round of the repeated game,*

$$u_i(a^1, a^2, ..., a^{t-1}, ...) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(a^t).$$

where, the discount factor δ *measures how much the players value the future payoff over the current payoff.*

Theorem 3.1.1 *For any feasible and strictly payoff profile v such that* $v_i > v_i^N$, *for all* $i \in N$ *and* v_i^N *being the payoff of the stage-game Nash equilibrium, there exists* $\underline{\delta} \in (0, 1)$, *such that for all* $\delta \in [\underline{\delta}, 1]$, *there exists a repeated-game strategy profile which is a* ***subgame perfect equilibrium*** *of the repeated game and yields the expected payoff profile v.*

3.1.4 Vector Game

A generalization to vector payoffs started to study by Nieuwenhuis and Corley. They defined the concept of vector maximization (or efficiency or Pareto optimality) and the vector of minimization [23], as follows:

Definition 3.1.4 ***Vector Maximization:*** *Let* $u = (u_1, ..., u_n)$, $v = (v_1, ..., v_n) \in D \subset R^n$. *If* $u_i \leq v_i, i = 1, ..., n$ *and* $u_j < v_j$ *for some j. The point* $u \in D$ *is said to be a vector maximum of D, which means that* $u \in v \max D$, *if* $u \nleq v$, $\forall v \in D$.

In continuous, we consider a two-person bimatrix vector and we will give the definition of the vector game or else the vector-valued game.

Definition 3.1.5 ***Vector Game:*** *A vector game is a tuple* $< A, A', V, r >$ *where:*

- *A is the set of actions available to the first player,*
- *A' is the set of actions available to the opponent player,*
- *V is a real HIlbert space,*
- *r is the opponent's reward function,* $r : A \times A' \rightarrow V$

A real Hilbert space is a vector space over $\mathbb{R}$ with an inner product. The space must be complete with respect to the norm $||x|| = \sqrt{< x, x >}$.

3.1.5 A Learning Model

A *single agent learning model* consists of a set of actions $N = \{1, ..., n\}$, a mixed action vector $q^t \in Q$ for all times t, a pure action vector $a^t = e_i$ for some pure action i for all times t and a reward vector $r^t = (r_1, ..., r_n) \in [0, 1]^n$.

A *learning algorithm* **H** is a sequence of functions $q^t : History^{t-1} \rightarrow Q$, where a History is a sequence of action-reward pairs $(a^1, r^1), (a^1, r^1), ...$ Thus, we gave the definition of a learning algorithm for an informed repeated game.

3.1.6 Price of Anarchy

The notion of the *Price of Anarchy* (PoA) is introduced by Koutsopias and Papadimitriou in 1999 [24]. The PoA is a measure of the effect of selfishness in games. Therefore, the *Price of Anarchy* is the ratio between the social welfare of the optimum solution and that of the worst Nash equilibrium.

Initially, we consider that $\mathcal{A}_i$ is the set of pure strategies for the player i. The set of mixed strategies for the player i is denoted as $\mathcal{S}_i$. Then, in each game there is an associated social utility function $u : \mathcal{A} \rightarrow \mathbb{R}$. The individual utility function for the player i is denoted as $a_i : \mathcal{A} \rightarrow \mathbb{R}$. Therefore, the function $\bar{u} : \mathcal{S} \rightarrow \mathbb{R}$ is the expected social utility over randomness of the players. The function $\bar{a} : \mathcal{S} \to \mathbb{R}$ is the expected value of the utility of a strategy profile to player i.

The social value of the socially optimum strategy profile is defined as $OPT = \max_{S \in \mathbb{S}} \bar{u}(S)$ in maximazation problems. Similarly in the minimazation problems, we have $OPT = \min_{S \in \mathbb{S}} \bar{u}(S)$.

Definition 3.1.6 ***Price of Anarchy***: *Consider an instance of a maximization game, then the Price of Anarchy is defined to be $\frac{OPT}{\bar{u}(S)}$, where S is the worst Nash equilibrium for the game. On the other hand, the Price of Anarchy for an instance of a minimization game is defined to be $\frac{\bar{u}(S)}{OPT}$, where S is the worst Nash equilibrium for the game.*

3.2 Regret

Regret can be characterized as a feeling of remorse over something that has happened, particularly as a result of one's own actions [25]. In the game theory, we consider a player i with strategy s_i. Then, the concept of *regret* is defined as the difference between the payoffs obtained by utilizing strategy s_i and the payoffs that could have been achieved had some other strategy $\bar{s}_i$, been played instead. So, the regret is denoted as a measure of performance and is closely related to equilibrium concept.

Definition 3.2.1 ***Regret***: *The average payoff that player i would have obtained if that player had adopted strategy $\bar{s}_i$ every time in the past instead of the s_i is the Regret value. Let $u_i^t(\bar{s}_i, s_{-i})$ is the payoff of player i at time t by taking $\bar{s}_i$ against the other player taking s_{-i}. Therefore, the formulating of regret value is:*

$$\Re_i^T = \max \left\{ \frac{1}{T} \sum_{t=1}^{T} \left[u_i^t(\bar{s}_i, s_{-i}) - u_i^t(s_i, s_{-i}) \right], 0 \right\}$$

Let a repeatedly game in an uncertain environment that the players have to make decisions in order to choose their next action. We need to choose a route of N possible routes, everyday from the home to the work. Let that the traffic is different each day. Thus, an algorithm can probabilistically choose an action (a route), at each time step. The algorithm incurs the loss of its action chosen (the duration time). Therefore, we are interested to study and analyse learning algorithms that the resulting behavior relates to the game-theoretic equilibria, when all players in that system are simulataneously adapting in such a manner.

The *no regret* can be described as follows: "*a sequence of plays is optimal if there is no regret for playing the given strategy sequence rather than playing* *any other possible sequence of strategies*" [19]. Thus, we want algorithms that minimize the regret.

Let that we have a game with cost and N players. If the strategy vector $s = (s_1, ..., s_n)$ is palyed then the cost for player i is $c_i(s)$. This game is played T days, with $s^t = (s_1^t, ..., s_n^t)$ being the strategy vector used on day t. The cost of player i is $\sum_{t=1}^{T} c_i(s^t)$. Then, the overall cost is $\sum_{t=1}^{T} \sum_{i} c_i(s^t)$. The formulation of regret is defined as:

Definition 3.2.2 ***Regret or Time-averaged Regret****: A sequence of strategy vectors $(s^1, ..., s^T)$ is regret for player i with respect to action s if:*

$$\frac{1}{T}\left[\sum_{t=1}^{T} c_i(s^t) - \sum_{t=1}^{T} c_i(s)\right]$$

The formulation of no-regret is defined as [26]:

Definition 3.2.3 ***No-Regret****: A sequence of strategy vectors $(s^1, ..., s^T)$ is no-regret for player i if:*

$$\sum_{t=1}^{T} c_i(s^t) \leq \min_x \sum_{t=1}^{T} c_i(x, s_{-i}^t)$$

where $\sum_{t=1}^{T} c_i(x, s_{-i}^t)$ is the cost of strategy x.

Definition 3.2.4 ***Vanishing Regret****: A sequence of strategy vectors $s^T, ..., s^T)$ has vanishing reret for player i if, assuming that $0 \leq c_i(s) \leq 1$:*

$$\lim_{T\to\infty} \sup \tfrac{1}{T} \sum_{t=1}^{T} c_i(s^t) - \min_x \tfrac{1}{T} \sum_{t=1}^{T} c_i(x, s_{-i}^t) \leq 0$$

Definition 3.2.5 **(λ, μ)*-smooth****: A cost game is (λ, μ)-smooth if for any strategy vectors s, s_*, we have:*

$$\sum_i c_i(s_i^*, s_{-i}) \leq \lambda \sum_i c_i(s_i^*) + \mu \sum_i c_i(s)$$

The above definitions give us the next theorem that is related to the Price of Anarchy.

Theorem 3.2.1 *If a cost game is (λ, μ)-smooth and all players have no regret on a sequence $s^1, ..., s^T$ of plays, then we have the following:*

$$\sum_{t=1}^{T}\sum_{i} c_i(s^t) \leq \frac{\lambda}{1-\mu} T \min_{x} \sum_{i} c_i(s)$$

where the ratio $\frac{\lambda}{1-\mu}$ *is the same bound as Price of Anarchy.*

Proof. *Initially, we start from the definition of* No-regret*, as follows:*

$$\sum_{t=1}^{T} c_i(s^t) \leq \min_{x} \sum_{t=1}^{T} c_i(x, s_{-i}^t)$$

In continous, we consider s^* *that is the min cost vector, which means that for any player i,* s_i^* *is no-regret. Therefore:*

$$\min_{x} \sum_{t=1}^{T} c_i(x, s_{-i}^t) \leq \sum_{t=1}^{T} c_i(s_i^*, s_{-i}^t)$$

For all the players i and from the two above inequalities we have:

$$\sum_{t=1}^{T}\sum_{i} c_i(s^t) \leq \sum_{t=1}^{T}\sum_{i} c_i(s_i^*, s_{-i}^t)$$

Then, we use the definition of (λ, μ)*-smooth and is applied in the right hand of the inequality, as follows:*

$$\sum_{t=1}^{T}\sum_{i} c_i(s_i^*, s_{-i}^t) \leq \lambda \sum_{t=1}^{T}\sum_{i} c_i(s^*) + \mu \sum_{t=1}^{T}\sum_{i} c_i(s^t) = \lambda T \sum_{i} c_i(s^*) + \mu \sum_{t=1}^{T}\sum_{i} c_i(s^t)$$

Therefore, we conclude that:

$$\sum_{t=1}^{T}\sum_{i} c_i(s^t) \leq \lambda T \sum_{i} c_i(s^*) + \mu \sum_{t=1}^{T}\sum_{i} c_i(s^t) \Rightarrow$$
$$(1-\mu) \sum_{t=1}^{T}\sum_{i} c_i(s^t) \leq \lambda T \sum_{i} c_i(s^*) \Rightarrow$$
$$\sum_{t=1}^{T}\sum_{i} c_i(s^t) \leq \frac{\lambda}{1-\mu} T \sum_{i} c_i(s^*).$$

□

There is an other form of regret that sometimes called *distribution regret*. The previous definitions are referred to the actions that the players actually plays. The distribution regret is calculated with respect to the player's mixed strategy. Let now q_{-i}^t the probability distribution, which serves as a model of the environment at time, then the expected regret felt by player i at time t is the difference between the expected payoff of strategy $\bar{s}_i$ and strategy s_i [27]. The formula of this is as follows:

Definition 3.2.6 ***Expected Regret:***

$$\mathbb{E}\left[R^t_{s_i \to \bar{s}_i}\right] = \mathbb{E}\left[r_i(\bar{s}_i, q^t_{-i}) - r_i(s_i, q^t_{-i})\right]$$

In continuous, is given the definition of the cumulative expected regret through time T. It is defined the feeling by player i from strategy s_i towards strategy $\bar{s}_i$ is the summation over the instantenous values of expected regret, whenever strategy s_i is played rather than strategy $\bar{s}_i$, as follows:

Definition 3.2.7 ***Cumulative Expected Regret:***

$$\mathbb{E}\left[R^T_{s_i \to \bar{s}_i}\right] = \sum_{t=1}^{T} \mathbb{1}_{\left\{s^t_i = s_i\right\}} \mathbb{E}\left[R^t_{s_i \to \bar{s}_i}\right]$$

where, the indicator function $\mathbb{1}_{\left\{s^t_i = s_i\right\}}$

Following, is defined the Φ-*Regret Game*. Thus, is constructed a vector game such that the rewards obtained in the vector game correspond to the regret experienced in a repeated game.

Definition 3.2.8 Φ***-Regret Game:*** *Given a one shot game* $\Gamma =< N, < A_i >_{i \in N}, < r_i >_{i \in N} >$, *a player* i *and a set of action transformations* $\Phi \subseteq \Phi^{SWAP}(A_i)$, *the* Φ*-regret game for player* i *the vector game*

$$< A_i, A_{-i}, \mathbb{R}^{\Phi}, \rho^{\Phi}_i >$$

where $\rho^{\Phi}_i : A_i \times A_{-i} \to \mathbb{R}^{\Phi}$ *is defined as*

$$\rho^{\Phi}_i(a_i, a_{-i}) = r_i(\phi(a_i), a_{-i}) - r_i$$

3.2.1 External Regret

Initially, the concept of the external regret is introduced by Hannan (1957). In 1995, the no-external-regret property is denoted as *universal consistency* by Fudenberg & Levine. The aim for the external regret setting is to design an online algorithm that will be able to approach the performance of the best algorithm from a given class of algorithms $\mathcal{G}$. The external regret are also called as the *best expert* problem. One important application of external regret is a general methodology for developing online algorithms whose performance matches that of an optimal static offline algorithm, by modeling the possible static solutions as different actions.

At each time step t, an online algorithm H selects a disrtibution p^t over the N actions. In continous, the adversary selects a loss vector $l^t \in [0, 1]^N$.

The loss of the i-th action at time t is $l_i \in [0,1]$. The loss of the i-th action during the first T times steps is $l_i^T = \sum_{t=1}^{T} l_i^t$ and the loss of online algorithm H is $l_H^T = \sum_{t=1}^{T} l_H^t$. The aim is to minimize the external regret $R_{\mathcal{G}} = L_{\mathcal{H}}^T - L_{\mathcal{G},\min}^T$, where $\mathcal{G}$ is a class of algorithms and is called the *comparison class*.

Definition 3.2.9 *Given a sequence of plays $\{s^t\}$ of length T, the sequence of plays $\{s_i^t\}$ for player i is said to exhibit* ***no-external regret*** *if and only if $\forall \epsilon > 0$, $\forall \bar{s}_i \in S_i$,*

$$ExternalRegret_{S_i \to \bar{s}_i}^T < \epsilon T,$$

where $ExternalRegret_{S_i \to \bar{s}_i}^T = \sum_{s_i \in S_i} ExternalRegret_{s_i \to \bar{s}_i}^T$

3.2.2 Internal Regret

In 1997-1999, the authors Foster and Vohra were the first, who introduced the concept of internal regret algorithms. Other authors in this procedure were Hart and Mas-Collel (2000), Cesa-Bianchi and Lugoci (2003), and Blum and Mansour (2005). *Internal regret* allows us to modify an online action sequence by changing every occurrence of a given action i to an other action j. Therefore, *internal regret* compares the loss of an online algorithm to the loss of a modified online algorithm and changes one action by another [28]. The *no-internal regret* criterion is a refinement of the no-external regret, in which the only substitutions that are considered are those which are preferable. Thus, the regret is positive, when one strategy is considered in place of another.

In introduction of this chapter, we see the definition of Foster and Vohra that it refer us in a general game. Thus, the importance of internal regret in game theory is due to the fact that *in a general game, if each player has sublinear internal regret then the empirical frequencies converge to a correlated equilibrium* [28].

The no-regret regret can be written as mathematical formula. Let

$$InternalRegret_{s_i \to \bar{s}_i}^T = \left(InternalRegret_{s_i \to \bar{s}_i}^T\right)^+,$$

where $X^+ = max\{X, O\}$.

Definition 3.2.10 *Given a sequence of plays $\{s^t\}$ of length T, the sequence of plays $\{s_i^t\}$ for player i is said to exhibit* ***no-internal regret*** *if and only if $\forall \epsilon > 0$, $\forall \bar{s}_i \in S_i$,*

$$InternalRegret^T_{S_i \to \bar{s}_i} < \epsilon T,$$

where $InternalRegret^T_{S_i \to \bar{s}_i} = \sum_{s_i \in S_i} InternalRegret^T_{s_i \to \bar{s}_i}$

Theorem 3.2.2 *No-internal regret implies no-external regret.*

Proof

We have that $\forall T$, $\forall \bar{s}_i \in S_i$, $ExternalRegret^T_{S_i \to \bar{s}_i} \leq InternalRegret^T_{S_i \to \bar{s}_i}$. *From the no-internal regret, it follows that* $\forall \bar{s}_i \in S_i$, *for arbitrary* $\epsilon > 0$,

$$\mathbb{E}\left[InternalRegret^T_{S_i \to \bar{s}_i}\right] < \epsilon T$$

Thus, from the previous inequality it follows that $\mathbb{E}\left[ExternalRegret^T_{S_i \to \bar{s}_i}\right] < \epsilon T$. *Therefore, no-internal regret implies no-external regret.*

□

An addition work about these algorithms was by Greenwald and Jafari (2003) [19]. They introduced a general class, which spans the spectrum from no-internal-regret learning to no-external-regret learning, and beyond. Greenwald and Jafari used the approachability of Blackwell and the generalization of this approachability by Jafari (2003) in order to prove the next proposition. Before of this proposition are useful the next definition, lemma and its corollary.

Definition 3.2.11 *A* Φ*-no-regret learning algorithm is one that* ρ_Φ*-approaches* $\mathbb{R}^\Phi$.

Lemma 3.2.1 *If learning algorithm A satisfies* Φ*-no-regret, then A also satisfies* Φ'*-no-regret, for all finite subsets* $\Phi' \subseteq SCH(\Phi)$, *the* super convex hull *of* Φ, *defined as follows:*

$$SCH(\Phi) =$$
$$\left\{\sum_{i=1}^{k+1} a_i\phi_i \mid \phi_i \in \Phi,\ for\ 1 \leq i \leq k, \phi_{k+1} = I, a_i \geq 0,\ for\ 1 \leq i \leq k, a_{k+1} \in \mathbb{R}\ and\ \sum_{i=1}^{k+1} a_i = 1\right\}$$

Corollary 3.2.1 *If learning algorithm A satisfies* Φ*-no-regret, then A also satisfies* Φ'*-no-regret, for all finite subsets* $\Phi' \subseteq CH(\Phi)$, *the* convex hull *of* Φ.

Proposition 3.2.1 [**Greenwald** & **Jafari**] : *If learning algorithm A satisfies no-internal-regret, then A also satisfies* Φ*-no-regret, for all finite subsets* Φ *of the set of stochastic matrices.*

Proof *For the proof are used a lemma and proposition*

□

3.3 External Regret Minimization

We referred in the previous section the concept of the external regret. The main goal is the minimization of the external regret, $R_{\mathcal{G}} = L^T_{\mathcal{H}} - L^T_{\mathcal{G},\min}$, where $\mathcal{G}$ is a class of algorithms and $\mathcal{H}$ is an online algorithm. Startly in this section, we will see some results that are not guaranteed a low regret with respect to the overall optimal sequence of decisions in hindsight. The next theorem show us a very large regret, as we used all possible functions. The set of all possible function is denoted as $\mathcal{G}_{all}$.

Theorem 3.3.1 *For any online algorithm H there is a sequence of T loss vectors such that regret $R_{\mathcal{G}}$ is at least $T(1 - 1/N)$.*

Proof *At each time t, the action i_t of lowest probability p_i^t gets a loss of 0. The other actions get a loss of 1. Then $min_i\{p_i^t\} \leq 1/N$. Thus in T time steps, the loss of algorithm H is at least $T(1 - 1/N)$. There exists a function g, which belongs to $\mathcal{G}_{all}$, $g\left(t\right) = i_t$ with a total loss of 0.*

□

In continuous, we study some regret minimization algorithms. Initially, we describe a Greedy algorithm, in which all losses are assumed to be either 0 or 1. It is proved that this algorithm is weak. The loss is at most an $O\left(N\right)$ factor from the best action. Thereafter, we go to the Randomized Greedy (RG) algorithm, that is an improvement version of the first algorithm. In the RG algorithm is assigned a uniform distribution over all those actions with minimum total loss. It achieved to be the loss at most an $O\left(\log N\right)$ factor from the best action. This algorithm is also weak. However, a Randomized Weighted Majority (RWM) algorithm and a Polynomial Weights (PW) algorihm are designed in order to found a better bound for the total loss and are analyzed below.

3.3.1 Greedy and Randomized-Greedy Algorithms

A regret minimization algorithm is developed, a Greedy algorithm. Let $L_i^t = \sum_{\tau=1}^{t} l_i^\tau$ is the cumulative loss at time for the action i. The algorithm, at each time t, selects action $x^t = \arg\min_{i \in X} L_i^{t-1}$. The pseudo-code of Greedy algorithm is presented in the following table.

Theorem 3.3.2 *The Greedy algorithm, for any sequence of losses has*

$$L^T_{Greedy} \leq N L^T_{min} + (N - 1).$$

Algorithm 1 Greedy Algorithm

Initialization:
$x^1 = 1$
Procedure at time t:
Let $L_{\min}^{t-1} = \min_{i \in X} L_i^{t-1}$ and $S^{t-1} = \{i : L_i^{t-1} = L_{\min}^{t-1}\}$.
Let $x^t = \min S^{t-1}$.

Proof *At each time t, the Greedy gets a loss of 1 and the L_{min}^T does not increase, at least one action is removed form S^T. This can occur at most N times, i.e. the algorithm get loss at most N, before L_{min}^T increases by 1. Therefore, we use induction to prove that $L_{Greedy}^t \leq N - |S^t| + NL_{min}^t$.*

□

The pseudo-code of Randomized Greedy algorithm is presented in the following table.

Algorithm 2 Randomized Greedy Algorithm

Initialization:
$p_i^1 = 1/N$, for $i \in \{1, ..., N\}$
Procedure at time t:
Let $L_{\min}^{t-1} = \min_{i \in X} L_i^{t-1}$ and $S^{t-1} = \{i : L_i^{t-1} = L_{\min}^{t-1}\}$.
Let $p_i^t = 1/|S^{t-1}|$ for $i \in S^{t-1}$, otherwise $p_i^t = 0$.

Theorem 3.3.3 *The Randomized Greedy algorithm, for any sequence of losses has*

$$L_{RG}^T \leq \ln N + (1 + \ln N) L_{min}^T.$$

3.3.2 Randomized Weighted Majority Algorithm (RWM)

The idea of Randomized Weighted Majority Algorithm is introduced by Littlestone and Warmuth. The RWM algorithm is also known as Hedge algorithm. In the previous algorithm, we can observe that the losses are greatest when the sets $S^t = \{i : L_i^t = L_{min}^t\}$ are small, which means that the online loss is inversely proportional to $|S^t|$. Then, the authors Littlestone and Warmuth proposed to give weights to the actions which are near best at the present time, in order to overcome this weackness.

The weight for an action i is denoted as $w_i = (1 - \epsilon)^{L_i}$, where L_i is the total loss for an action i and ϵ is a parameter, a small constant. The total weight

Algorithm 3 Randomized Weighted Majority (RWM) Algorithm

Initialization:
$w_i^1 = 1$, for $i \in X$
$p_i^1 = 1/N$, for $i \in X$
Procedure at time t:
if $l_i^{t-1} = 1$, let $w_i^t = w_i^{t-1}(1-\epsilon)$;
else $l_i^{t-1} = 0$, let $w_i^t = w_i^{t-1}$;
Let $p_i^t = \frac{w_i^t}{W^t} = \frac{w_i^t}{\sum\limits_{i \in X} w_i^t}$.

is denoted as W^T. The losses are assumed that belong to the set $\{0, 1\}$. The pseudo-code of RWM algorithm is presented in the following table.

Theorem 3.3.4 *For $\epsilon \leq 1/2$, the loss of RWM algorithm on any sequence of binary $\{0, 1\}$ losses satisfies:*

$$L_{RWM}^T \leq (1+\epsilon)L_{min}^T + \frac{lnN}{\epsilon},$$

for $\epsilon = min\{\sqrt{lnN/T}, 1/2\}$, we have $L_{RWM}^T \leq L_{min}^T + 2\sqrt{L_{min} lnN}$

Proof *In [16], they showed that any time the online algorithm has significant expected loss, the W^t must drop substantially. Also, they considered that $W^{T+1} \geq \max_i w_i^{T+1} \Longrightarrow W^{T+1} \geq (1-\epsilon)^{L_{min}^T}$.*

Let $W^t = \sum_i w_i^t$, $W^1 = N$. The fraction $\dfrac{\sum_i w_i^t}{W^t}$ is the expected loss of RWM algorithm at time t and is denoted as F^t. So, each of the actions experiencing a loss of 1 has its weight multiplied by $(1-\epsilon)$ while the rest are unchanged. Then

$$W^{t+1} = W^t - \epsilon F^t W^t = W^t(1-\epsilon F^t) = W^1 \prod_{t=1}^{T}(1-\epsilon F^t) = N \cdot \prod_{t=1}^{T}(1-\epsilon F^t)$$

In continuous, the inequality $(1-\epsilon)^{L_{min}^T} \leq W^{T+1}$ is used and then we have

$$(1-\epsilon)^{L_{min}^T} \leq N \cdot \prod_{t=1}^{T}(1-\epsilon F^t) \Rightarrow$$

$$\ln\left((1-\epsilon)^{L_{min}^T}\right) \leq \ln\left(N \cdot \prod_{t=1}^{T}(1-\epsilon F^t)\right) \Rightarrow$$

$$L_{min}^{T} \ln(1-\epsilon) \leq (\ln N) + \sum_{t=1}^{T} \ln(1-\epsilon F^{t}) \xrightarrow[ln(1-x)\leq -x]{\text{using the Bernoulli's inequality:}}$$

$$L_{min}^{T} \ln(1-\epsilon) \leq (\ln N) - \sum_{t=1}^{T} \epsilon F^{t} \Rightarrow$$

$$L_{min}^{T} \ln(1-\epsilon) \leq (\ln N) - \epsilon L_{RWM}^{T} \Rightarrow$$

$$\epsilon L_{RWM}^{T} \leq (\ln N) - L_{min}^{T} \ln(1-\epsilon) \Rightarrow$$

$$L_{RWM}^{T} \leq \frac{(\ln N)}{\epsilon} - \frac{L_{min}^{T} \ln(1-\epsilon)}{\epsilon} \xrightarrow[\text{for } 0\leq x\leq \frac{1}{2}]{-ln(1-x)\leq x+x^{2}}$$

$$L_{RWM}^{T} \leq \frac{(\ln N)}{\epsilon} + (1+\epsilon) L_{min}^{T}$$

□

3.3.3 Polynomial Weights Algorithm (PW)

An extension of the RWM algorithm to losses in the closed interval [0,1] is the Polynomial Weights (PW) algorithm.

Let $W^t = \sum_i w_i^t$, $W^1 = N$. The term of weight for the action i at time t is $w_i^t = w_i^{t-1}(1-\epsilon l_i^{t-1})$ The fraction $\dfrac{\sum_i w_i^t}{W^t}$ is the expected loss of PW algorithm at time t and is denoted as F^t. Then

$$W^{t+1} = W^t - \epsilon F^t W^t = W^t(1-\epsilon F^t) = W^1 \prod_{t=1}^{T}(1-\epsilon F^t) = N \cdot \prod_{t=1}^{T}(1-\epsilon F^t)$$

The pseudo-code of PW algorithm is presented in the following table.

Algorithm 4 Polynomial Weights (PW) Algorithm

Initialization:
$w_i^1 = 1$, for $i \in X$
$p_i^1 = 1/N$, for $i \in X$
Procedure at time t:
Let $w_i^t = w_i^{t-1}(1-\epsilon l_i^{t-1})$;
Let $p_i^t = \frac{w_i^t}{W^t} = \frac{w_i^t}{\sum\limits_{i\in X} w_i^t}$.

Theorem 3.3.5 *The Polynomial Weights (PW) algorithm, using $\eta \leq 1/2$, for any [0,1]-valued loss sequence and for any k has,*

$$L_{PW}^T \le L_k^T + \epsilon \cdot Q_k^T + \frac{lnN}{\epsilon}$$

then for $\epsilon = min\{\sqrt{lnN/T}, 1/2\}$ *and* $Q_k^T \le T$*, where* $Q_k^T = \sum_{t=1} T(l_k^t)^2$ *, we have* $L_{PW}^T \le L_{min}^T + 2\sqrt{TlnN}$.

Proof *From the analysis of RWM, we have*

$$W^{t+1} = W^t\left(1 - \epsilon F^t\right) \Rightarrow W^{T+1} = N\prod_{t=1}^{T}\left(1 - \epsilon F^t\right)$$

where $W^t = N$ *and* F^t *is PW's loss at time t. In continuous, we have*

$$\ln W^{T+1} = \ln N + \sum_{t=1}^{T}\ln\left(1 - \epsilon F^t\right) \le \ln N - \epsilon\sum_{t=1}^{T}F^t = \ln N - \epsilon L_{PW}^T.$$

In continuous, we analyze the lower bound for $\ln W^{T+1}$*, as follows*

$$\ln W^{T+1} \ge \ln w_k^{T+1} = \sum_{t=1}^{T}\ln\left(1 - \epsilon l_k^t\right) \xrightarrow[\text{for } 0 \le x \le \frac{1}{2}]{-ln(1-x) \le x + x^2}.$$

$$\ln W^{T+1} \ge -\sum_{t=1}^{T}\epsilon l_k^t - \sum_{t=1}^{T}\left(\epsilon l_k^t\right)^2 \Rightarrow$$

$$\ln W^{T+1} \ge -\epsilon L_k^T - \epsilon^2 Q_k^T$$

Therefore, the upper and the lower bounds on $\ln W^{T+1}$ *are combined, as follows:*

$$-\epsilon L_k^T - \epsilon^2 Q_k^T \le \ln N - \epsilon L_{PW}^T \Rightarrow$$

$$L_{PW}^T \le L_k^T + \epsilon \cdot Q_k^T + \frac{lnN}{\epsilon}$$

□

In continuous, we will give two theorems, which are showed us that the regret bound is near optimal. In the first theorem is proved that one cannot hope to get sublinear regret when $T < \log_2 N$. In the second theorem is proved that one cannot hope to achieve regret $o(\sqrt{T})$ even when the actions is two.

Theorem 3.3.6 *Let* $T < \log_2 N$*. There exists a stochastic generation of losses such that, for any online algorithm* R_1*, we have* $\mathbb{E}[L_{R_1}^T] = T/2$ *and* $L_{min}^T = 0$.

Proof *Let at time* $t = 1$*, a random subset of* $N/2$ *actions get a loss of* 0 *and the rest get a loss of* 1*. Let at time* $t = 2$*, a random subset of* $N/4$ *actions* *get a loss of* 0 *and these actions have loss* 0 *at time* $t = 1$. *The rest subset of actions get a loss of* 1. *We repeat this process and we can observe that at each time step, a random subset of half of the actions that have received loss* 0 *so far get a loss of* 0. *The rest subset of actions get a loss of* 1. *Any online algorithm incurs an expected loss of* 1/2 *at each time step. At each time step* t, *the expected fraction of probability mass* p_i^t *on actions that receive a loss of* 0 *is at most* 1/2. *Let* $T < \log_2 N$, *then there will alwaays be some action with total loss of zero.*

□

3.4 Regret Matching Algorithms

The regret matching algorithms are considered as a general class of learning algorithms for a repeated game setting. These algorithms are parameterized by a set Φ of transformations over the set of actions and a link function $f : \mathbb{R}^{\Phi} \to \mathbb{R}_{+}^{\Phi}$. Firstly, a Φ-regret vector is analyzed by comparing the average reward obtained by an agent over some finite sequence of rounds to the average reward that could have been obtained had the agent instead played each transformations of its sequence of actions [22]. A property closely related to Blackwell's condition for approachability is satisfied by the regret matching algorithms.

The regret matching property is defined as:

Definition 3.4.1 *Given a finite set of action transformations* $\Phi \subset \Phi_{ALL}(A)$ *and a function* $f : \mathbb{R}^{\Phi} \to \mathbb{R}_{+}^{\Phi}$, *a learnig algorithm* $\mathcal{A}$ *is called an* (f, Φ)*-regret matching algorithm if for all reward functions* r, *for all times* T, *for all histories* $h \in H^{T-1}$,

$$f(R_{t-1}^{\Phi}(h)) \cdot \mathbb{E}_{a \sim \mathcal{A}_t(h)}[\rho_t^{\Phi}(a, r)] \leq 0$$

Definition 3.4.2 *Let* $f : \mathbb{R}^{\Phi} \to \mathbb{R}_{+}^{\Phi}$, $f' : \mathbb{R}^{\Phi} \to \mathbb{R}_{+}^{\Phi}$ *be link functions. If there exists a function* $\psi : \mathbb{R}^{\Phi} \to \mathbb{R}_{+}^{\Phi}$ *such that* $\psi(x) f(x) = f'(x)$ *and* $||f(x)|| > 0 \Rightarrow \psi(x) > 0$, $\forall x \in \mathbb{R}^{\Phi}$, *then a learning algorithm is an* (f, Φ)*- regret-matching algorithm if and only if it is an* (f', Φ)*- regret-matching algorithm.*

3.5 Regret Minimization and Game Theory

In this section, we will discuss and analyze the connection between regret minimization and the fundamentals concepts of the game theory. If we have a two-player game and a player with external regret sublinear in T then will have an average payoff that is at least the value of the game minus a vanishing error term. If we have a general game, that all the players use procedures with sublinear *swap-regret* then they will converge to an approximate *correlated equilibrium*. Therefore, we will define a game G and we will give the definitons of ε*-Correlated Equilibrium* and ε*-Dominated*. The last definition means that for a player who minimizes swap-regret, the frequency of playing dominated actions is vanishing [16].

3.5.1 Correlated Equilibria and Swap Regret

Let a game $G =< M, (A_i), s_i >$ has a finite set M of m players. Player i has a set A_i of N actions and a loss function $s_i : A_i \times (\times_{j \neq i} A_j) \rightarrow [0, 1]$ that maps the action of player i and the actions of the other players to a real number.

A correlated equilibrium is a distribution P over the joint action space and means that if for each player it is the best response to play the suggesed action.

Definition 3.5.1 ***ε-Correlated Equilibrium****: A joint probability distribution P over $\times A_i$ is an ε-correlated equilibrium if for every player j and for any function $F : A_i \rightarrow A_j$, then we have:*

$$E_{a \sim P}[s_j(a_j, a^{-j})] - E_{a \sim P}[s_j(F(a_j), a^{-j})] \leq \varepsilon$$

where a^{-j} denotes the joint actions of the other players.

In continuous, we will see an important theorem, which relates the swap regret to the distance from equilibrium. An other result is that the payoff of each player is its payoff in some approxiamate correlated eqyilibrium. Also, we can observe that the algorithm converge to the set of CE if the average swap regret vanishes [16].

Theorem 3.5.1 *Let a game $G =< M, (A_i), s_i >$. Let that for T times steps, the strategy for every player has* swap regret *of at most $R(T, N)$. Which means that the empirical distribution Q of the joint actions played by the players is an R/T-correlated equilibrium.*

Proof *We fix a function $F : A_i \rightarrow A_i$ for player i. The player i has swap regret at most R. Then, we have $L^T \leq L^T_F + R$, where L^T is the loss of player i. The empirical distribution Q assigns to every P^t a probability $1/T$. Using the definition of the regret, we have that:*

$$L^T - L_F^T = \sum_{t=1}^{T} \mathbb{E}_{x^t \sim P^t}[s_i(x^t)] - \sum_{t=1}^{T} \mathbb{E}_{x^t \sim P^t}[s_i(F(x_i^t), x_{-i}^t)] =$$
$$= \sum_{t=1}^{T} \mathbb{E}_{x^t \sim P^t}[s_i(x^t) - s_i(F(x_i^t), x_{-i}^t)] =$$
$$= \sum_{t=1}^{T} \mathbb{E}_{x^t \sim P^t}[\mathcal{R}_i(x^t, F)] =$$
$$= T \cdot \mathbb{E}_{x^t \sim Q}[\mathcal{R}_i(x^t, F)] \leq R \Rightarrow$$
$$\Rightarrow \mathbb{E}_{x^t \sim Q}[\mathcal{R}_i(x^t, F)] \leq \frac{R}{T}$$

□

3.6 Approachability

Definition 3.6.1 ***Blackwell Instnance****[29]: A Blackwell instnance is a tuple $(\mathcal{X}, \mathcal{Y}, u, S)$ with $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$ compact and convex, $u : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}^d$ biaffine and $S \subset \mathbb{R}^d$ convex and closed . Then, for any Blackwell instance, we have that*

- *S is satisfiable if $\exists x \in \mathcal{X} \forall y \in \mathcal{Y} such that u(x, y) \in S$.*
- *S is response-satisfiable if $\forall y \in \mathcal{Y} \exists x \in \mathcal{X} such that u(x, y) \in S$.*
- *S is halfspace-satisfiable if, for any halfspace $H \supseteq S$, H is satisfiable.*

Definition 3.6.2 *Approachable: Given a Blackwell instance $(\mathcal{X}, \mathcal{Y}, u, S)$, S is approachable if there exists some algorithm $\mathcal{A}$ which selects points in $\mathcal{X}$, such that for any sequence $y_1, y_2, ... \in \mathcal{Y}$, we have that*

$$dist(\tfrac{1}{T}) \sum_{t=1}^{T} u(x_t, y_t), S) \to 0 \text{ as } T \to \infty,$$

where $x_t \leftarrow A(y_1, y_2, ..., y_{t-1})$.

Theorem 3.6.1 *(Blackwell's Approachability Theorem (1956)) For any Blackwell instance $(\mathcal{X}, \mathcal{Y}, u, S)$, S is approachable iff it is response-satisfiable.*

3.7 From Swap to External Regret

Startly, we consider a "*black box reduction*" showing that from the problem of designing a no-swap regret algorithm we can design a no-external regret algorithm. The "*black box reduction*" is presented in the Figure 3.2.

We consider, that the number of actions is denoted as N and so the different no-external algorithms are denoted as $A_1, A_2, ..., A_N$. Each of these algorithms give us a probability distribution q_i^t at each time step t. We note that an algorithm A_i will be responsible for ensuring against profitable deviations from action i to other actions. In the following theorem, we will see that the outcome of the combination of the no-external algorithms is the same of that of the no-swap algorithm. So, if each algorithm A_i provide a no-external-regret guarantee then we take a no-swap-regret guarantee. The last algorithm is also called as "*master algorithm*" and is denoted as H. At each time step t, the algorithm H works as follows:

- **receives** distributions q_i^t, $\forall i \in \{1, ..., N\}$ over actions from the algorithms $A_1, A_2, ..., A_N$.
- **compute** a single distribution p^t and we will take the p^t as a stationary distribution of the Markov chain.
- **receives** a loss vector l^t from the adversary.
- **returns** to each A_i the loss vector $p_i^t \cdot l^t$.

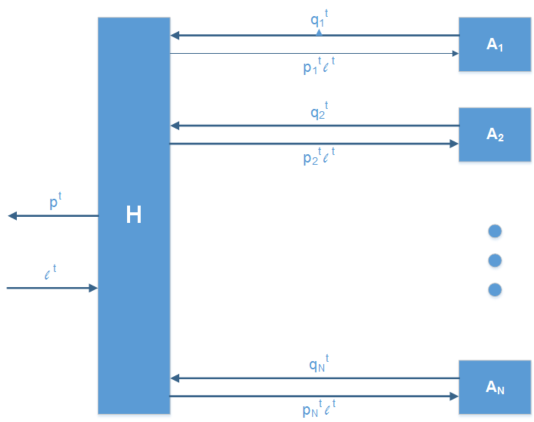

Figure 3.2: The structure of the swap regret reduction.

Theorem 3.7.1 *If there is a no-external regret algorithm, then there is a no-swap regret algorithm.*

Proof *The time averaged expected loss of the H algorithm is*

$$L_H = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} p_i^t \cdot l_i^t$$

The time averaged expected loss of the H algorithm under a switching function $\delta : \{1, ..., N\} \rightarrow \{1, ..., N\}$ *is*

$$L_{H,\delta} = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} p_i^t \cdot l_{\delta(i)}^t$$

Our ***goal*** *is to prove that the swap regret of H is at most* $N\mathcal{R}$, *where* $\mathcal{R}$ *is an external regret algorithm:*

$$L_H \leq L_{H,\delta} + NR$$

The actions of an A_j *algorithm are being chosen according to its recommended distributions* $q_j^1, ..., q_j^T$ *and that the true loss vectors are* $p_j^1 \cdot l^1, ..., p_j^T \cdot l^T$. *The time averaged expected loss of the algorithm* A_j *is*

$$L_{A_j} = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} q_{j,i}^t \left(p_j^t \cdot l_i^t\right)$$

In continuous, we fix an action $k \in 1, ..., N$ *to the loss* l_i^t *and also the* A_j *is a no-regret algorithm, then we take the following inequality*

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} q_{j,i}^t \left(p_j^t \cdot l_i^t\right) \leq \frac{1}{T}\sum_{t=1}^{T} p_j^t \cdot l_k^t + \mathcal{R}_j$$

where the term $\mathcal{R}_j$ *goes to* 0 *as* $T \rightarrow \infty$.

Thereafter, we fix a switching function δ, *that the above inequality is witten as*

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} q_{j,i}^t \left(p_j^t \cdot l_i^t\right) \leq \frac{1}{T}\sum_{t=1}^{T} p_j^t \cdot l_{\delta(i)}^t + \mathcal{R}_j \xRightarrow[\text{over all } j=1,2,...N]{\text{Summing the two sides}}$$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} q_{j,i}^t p_j^t \cdot l_i^t \leq \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{N} p_j^t \cdot l_{\delta(i)}^t + \sum_{j=1}^{N} \mathcal{R}_j$$

where the term $\sum_{j=1}^{N} \mathcal{R}_j$ *goes to* 0 *as* $T \to \infty$ *and from them we can observe that the right-hand side of this inequality is equal to the time-averaged expected loss under a function* δ *of the master algorithm. From this observation, we choose the splitting of the loss vector* l^t *amongst the no-external regret algorithms* $A_1, ..., A_N$ *to guarantee this property.*

Therefore, we have to prove that the left-hand side of this inequality is equal to the time-averaged expected loss of the master algorithm:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} p_i^t \cdot l_i^t = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} q_{j,i}^t p_j^t \cdot l_i^t \xrightarrow[\text{to prove}]{\text{We want}}$$

$$p_i^t = \sum_{j=1}^{N} q_{j,i}^t p_j^t$$

Using the definition of the stationary of a Markov chain, we want to show that the above equality is satisfied. For this reason, we have to show that p^t *is a stationary distribution. The set of states is* $\{1, ..., N\}$*. The distribution* $q_{j,i}^t$ *is a transition probability from the state* $j \in \{1, ..., N\}$ *to the state* $i \in \{1, ..., N\}$*. If the* p^t *is a probability distribution then it is the stationary distribution of this Markov chain. The Markov of chain is showed in the Figure 3.3.*

□

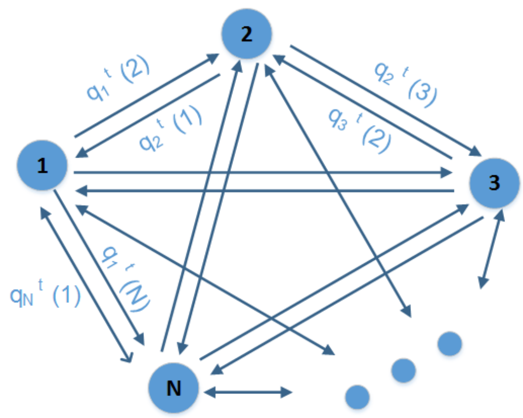

Figure 3.3: The Markov chain.

Therefore, we proved the next theorem. From this theorem is derived the following corollary:

Theorem 3.7.2 *: An $\mathcal{R}$ external regret proceduce H guarantees that for any sequence of T losses l^t and for every function delta : $1, ..., N \to 1, ..., N$, we have*

$$L_H = L_{H,\delta} + N \cdot \mathcal{R}$$

Corollary 3.7.1 *There exists an online algorithm H such that for every function $F : \{1, ..., N\} \to \{1, ..., N\}$,*

$$L_H \leq L_{H,F} + O\left(N\sqrt{T \log N}\right),$$

Therefore, the concept of this corollary is that the swap regret of H is at most $O\left(N\sqrt{T \log N}\right)$.

3.8 The Hierarchy of Equilibrium Concepts

There are three relaxations of Pure Nash Equilibrium (PNE), that each of them has more permissive and more computationally tractable than PNE. These three equlirium concepts are the set mixed Nash equilibrium (MNE), Correlated Equilibria (CE) and Coarse Correlated Equilibria (CCE). We have note that there are games such as Rock-Paper-Scissors game, atomic selfish routing games with multiple players that need not have PNE, even with only 2-players and quadratic cost functions [30]. Thus, the sets CCE, CE and MNE are guaranteed us the existence equilibrium in such finite games. The hierarchy of the equilibrium concepts is represented in the Figure 3.4.

From the Figure 3.4, we can observe that the biggest set is CCE. That set is a *quite tractable* set of equilibria, and hence a relatively plausible prediction of realized play. In the previous section, we see that there are simple learning procedures, which are computationally efficient and converge quickly to the set of CCE. Such as, the no-external regret algorithm that converges to CCE.

The next set of hierarchy is the CE and is *tractable*. In the previous section, we see that there are learning procedures that converge fairly quickly to CE [31]. Such as, the no-internal regret algorithm and the no-swap regret algorithm, which converge to CE.

We have note that the CE is tractable in the same strong as CCE, because of the proof of Theorem 3.7.1: "if there is a no-external regret algorithm, then there is a no-swap-regret algorithm". This reduction preserves computational efficiency.

The next set of hierarchy is mixed Nash equilibrium (MNE). However, MNE is a computationally *intractable* set of equilibria. The set MNE is guaranteed us the existence of a point but it is hard to compute. When we have 2-player zero-sum games then the no-external regret converges quickly to an approximate MNE.

We conclude that the sets CCE and CE are tractable in general games. The set MNE is tractable in 2-player zero-sum games. And the last set PNE is tractable in symmetric routing/congestion games [32].

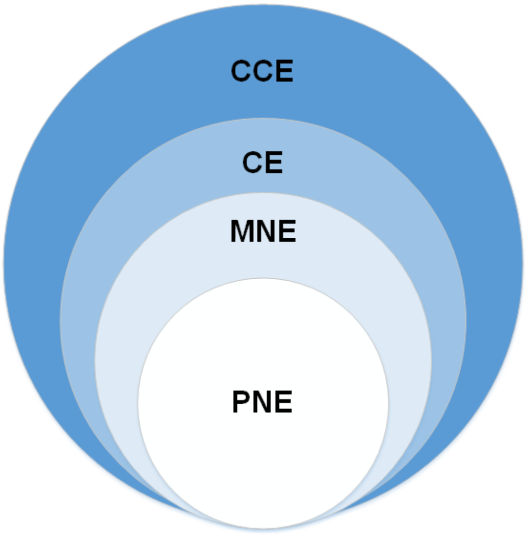

Figure 3.4: The hierarchy of the equilibrium concepts.

Chapter 4

Power Control in Wireless Networks

In this chapter, we study game theoretic models for power control in wireless networks. Initially, we introduce the evolution of wireless networks, from the first generation (1G) to the fourth generation (4G) wireless networks. In the next section, we analyze with more details the evolution of the 4G and mailnly the physical layer. At the physical layer analysis, the transmit power of wireless nodes has a huge impact. If there are two users and they want to communicate among them, then the ideal is to exist a sufficient power. If the transmit power is too high, significant interference will be generated to other users which will degrade other users performance. We have note that a mobile device has limited energy and hence cannot afford high power consumption. Thus, it is necessary to exist a sufficient transmit power in order to achieve better Quality-of-Service (QoS), to minimize the interference to the other users, to maxomize the battery life of the mobile devices and general to maximize the throughput of the wireless system. Therefore in the last section, we study some game theoretic algorithms, which have designed to in order to control the power of each user and hence the QoS in a given wireless environment.

4.1 Introduction in Wireless Networks

The first operational cellular communication system, that called as 1G, was deployed in the Norway in 1981 and was followed by similar systems in the US and UK. The 1G provided voice transmissions by using frequencies around 900 MHz and analogue modulation.

In continuous, the second generation (2G) of the wireless mobile network is developed. The 2G was based on low-band digital data signaling. The most popular of this generation is known as Global Systems for Mobile Communications (GSM). The first GSM systems used a 25MHz frequency spectrum in the 900MHz band [33]. GSM systems operate in the 900MHz and 1.8 GHz bands throughout the world with the exception of the Americas where they operate in the 1.9 GHz band. The GSM uses time-division multiple access (TDMA) and Frequency-division multiple access (FDMA) for user and cell separation. Within Europe, the GSM technology made possible the seamless roaming across all countries. Simultaneously, an other technology was developed in North America, which is called CDMAOne (Code Division Multiple Access).

So, FDMA is a method that the total bandwidth available is devided into a series of frequency bands. These frequencies are not overlapping among them and each of them carry a seperate signal. On the other hand, TDMA is an other method that the total time is devided into time slots of fixed length, one for each sub-channel. Two of four fundamental types of channel access schemes, as shown in Figures 4.1:

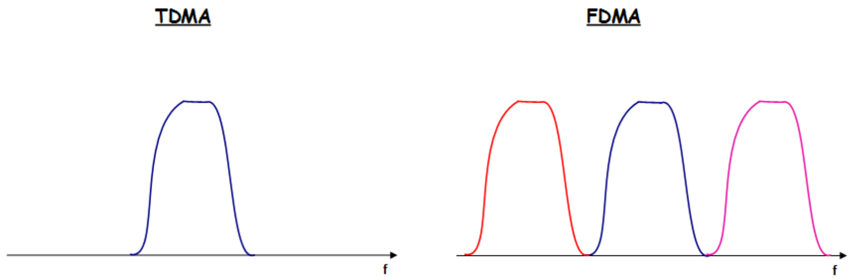

Figure 4.1: FDMA-TDMA

In third generation (3G) networks, there is Universal Mobile Telecommunication System (UMTS) from 3GPP standard, which based on Wideband Code Division Multiple Access (W-CDMA). Also in 3G, there is CDMA-2000 from Qualcomm standard. CDMA200 is also known as C2K or IMT MultiCarrier(IMT-MC). These technologies use code-division multiple access (CDMA). CDMA distinguishes between multiple transmissions carried simultaneously on a single wireless signal. It carries the transmissions on that signal, freeing network room for the wireless carrier and providing interference-free calls for the user. The 3G telecommunication networks support services that provide an information transfer rate of at least 200 kbit/sec.

The next generation is the fourth generation (4G) networks, which developed the Long Term Evolution (LTE) from 3GPP standard (in 2007) and the Mobile WiMAX from IEEE standard (in 2009). WiMAX and LTE use orthogonal frequency division multiple access (OFDM). OFDM is a FDM scheme used as a digital multi-carrier modulation method as well orthogonal sub-carrier signal are used to carry data. So, each sub-carrier can be modulated as a QAM (Quadrature amplitude modulation) or phase-shift keying at a low symbol rate.

The other two fundamental types of channel access schemes, as shown in Figure 4.2:

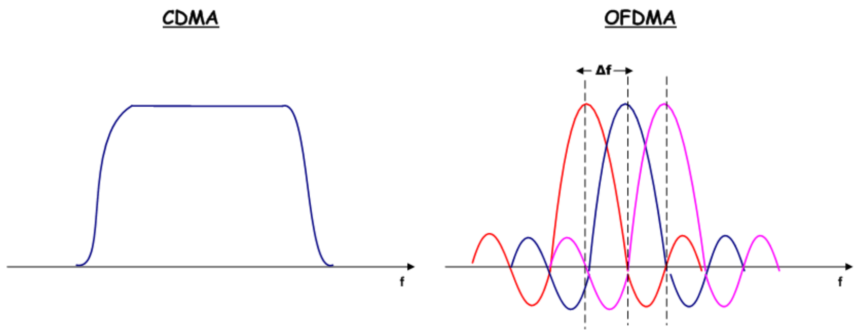

Figure 4.2: CDMA-OFDMA

In 2010, LTE Advanced was to be standardized as part of Release 10 of the 3GPP specification. LTE Advanced based on the existing LTE specification Release 10. In Figure 4.4 is represented the evolution of wireless networks from 1G to 4G LTE and we can observe that the last technology has achieved faster and better mobile broadband [34]. However, the evolution of wireless networks is continued.

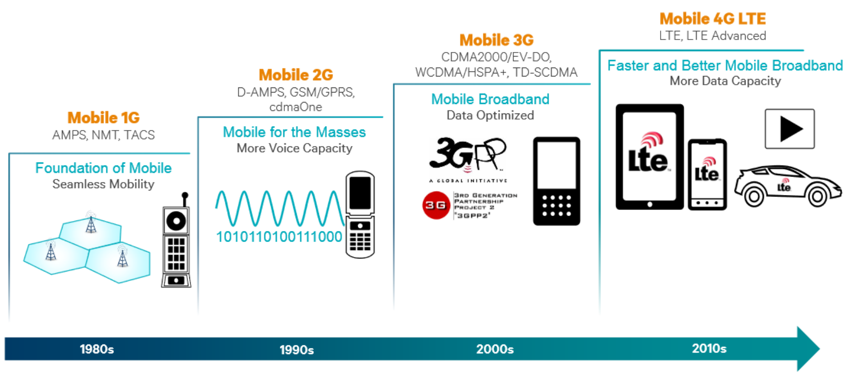

Figure 4.3: Aiming to faster and better services

In the following table, the evolution of mobile standards is presented with additional details as

Standards	3GPP	Qualcomm	China	IEEE
Carriers	AT&T T-Mobile US	Sprint Verizon	China Mobile	Sprint
2G digital & data services	GSM: 2G GPRS: 2.5G EDGE: 2.75G	CDMAOne		
3G	Rel. 4: UMTS 3G Rel. 5: HSDPA 3.5G Rel. 6: HSUPA 3.5G Rel. 7: HSPA 3.5G Rel. 8/9: LTE 3.9G	CDMA2000 EVDO rev A EVDO rev C Wimax, 3GPP LTE	TD-SCFDMA	WiMAX 3.9G
4G	Rel. 10: LTE-dvanced		TD-LTE	WiMAX 4G

Figure 4.4: Aiming to faster and better services

4.2 LTE Release 10 & beyond (LTE-Advanced)

The Long Term Evolution (LTE) is proposed by the Third Generation Partnership Project (3GPP) [35], [36]. The LTE cellular system is the evolution of the Third Generation (3G) Universal Mobile Telecommunication System (UMTS), which achieve a higher data rate (100 Mbps for downlink and 50 Mbps for uplink), a reduced latency and and a maximized in capacity to support the recent rapidly growing demand for high-speed multimedia applications such as video streaming, online games, Voice over Internet Protocol (VoIP), internet surfing [37]. The LTE standard is specified in 3GPP Release 8 [35]. An enhanced version of LTE is LTE-Advanced (LTE-A), which is introduced more specifications into Radio Resource Management (RRM) by the International Mobile Telecommunications-Advanced (IMT -Advanced) [36], [38]. The LTE-A standard is specified in 3GPP Release 10 and beyond and this technology is known as Fourth Generation (4G) [38].

The architecture of the LTE/LTE-A cellular system called as System Architecture Evolution (SAE), which consists of two parts: the Evolved Packet Core (EPC) and the Evolved Universal Terrestrial Radio Access Network (E-UTRAN), as illustrated in Figure 4.5 [35], [36]. The EPC consists of the mobility management entity (MME), the serving gateway (S-GW) and the packet data network gateway (P-GW) [35]. The E-UTRAN consists of user

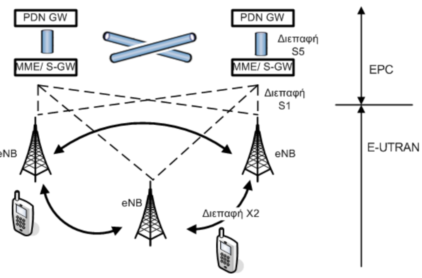

Figure 4.5: LTE/LTE-A cellular system

equipments (UE) and evolved NodeBs (eNodeBs or eNBs). The eNBs are the LTE base stations. The user equipment is a device and can be a smart phone or a laptop. The connection between eNBs is established by X2 interface and the connection between EPC and E-UTRAN becomes by the S1 interface [35], [36], [38].

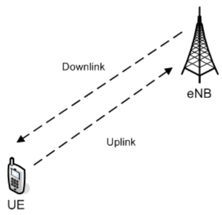

Figure 4.6: Downlink/Uplink Transission

The bandwidht of an LTE system is from 1.4 MHz to 20 MHz, while the bandwidht of an LTE-A system is up to 100 MHz. The technology of orthogonal frequency division multiple access (OFDMA) is used in the LTE/LTE-A for the downlink transmission, where the data are transmited from the eNB to the UE, as illustrated in Figure 4.6 [39], [38]. In this technology, the channel bandwidth is divided into small radio resources, which are called physical resource blocks (PRBs). The OFDMA is based on the use of the Discrete Fourier Transform (DFT) and the Inverse Discrete Fourier Transform (IDFT). The uplink/downlink transmissions in an LTE/LTE-A are organized in radio frames. The LTE/LTE-A use the Time Division Duplex (TDD)and the Frequency Division Duplex (FDD). Each PRB consists of 12 consecutive subcarriers and 7 OFDM symbols. One time slot consists of 7 OFDM symbols, where one time slot is 0.5 msec [39], [37]. A PRB always consists of 180 kHz in frequency. One subframe consists of 2 consecutive slots. One frame is 10 msec, which equals to 10 subframes, as illustrated in Figure 4.7 [39], [38]. The technology of Single-carrier frequency division multiple access (SC-FDMA) is used in the LTE/LTE-A for the uplink transmission, where the data are transmited from the UE to the eNB, as illustrated in Figure 4.6 [39], [38]. Using the SC-FDMA in the uplink is achieved a reduction to the peak-to-average power ratio of UEs, a reduction to their power consumption [37].

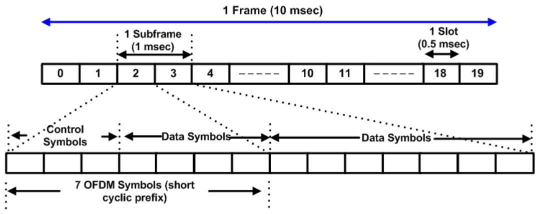

Figure 4.7: LTE/LTE-A frame structure

The E-UTRA radio interface between the eNB and UE is composed of the layer 1 (Physical Layer-PHY), the layer 2 (Medium Access Control-MAC) and the layer 3 (Radio Resource Control-RRC) [36]. The PHY modulates symbols over the radio interface, the MAC controls shared access to the radio interface across different UEs. The RRC handles radio configuration control and radio resource management with the purpose of broadcasting system information, paging and maintenance or establishment of a connection between the UE and E-UTRAN [39], [36].

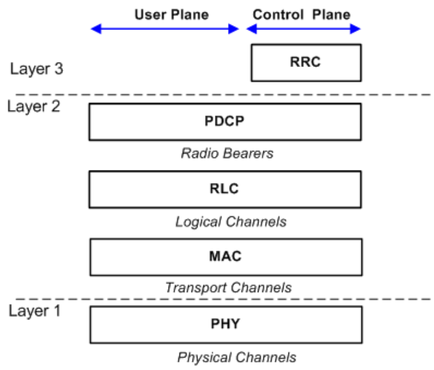

Figure 4.8: LTE/LTE-A channels and protocol layers

The physical layer (PHY) provide services to the MAC layer as shown in Figure 4.8. The provision of the services is done from the transport channels [39]. Therefore, the PHY is characterized from the coding, the modulation, multiantenna processing and from the mapping of the signal to physical time-frequency resources [39], [40]. The channels are divided into two parts: the uplink channels and the downlink channels. The types of physical channels are [39]:

Downlink Physical Channels	
PDSCH Physical Downlink Shared Channel	It carriers data of the user for very high data rates.
PDCCH Physical Downlink Control Channel	It is used for downlink scheduling decisions.
PMCH Physical Multicast Channel	It is used for Multicast/Broadcast single frequency network operation.
PBCH Physical Broadcast Channel	It is used for system information in order to the access the network from a user.
PCFICH Physical Control Format Indicator Channel	It informs the UE for the symbols that were used for the PDCCH.
PHICH Physical Hybrid ARQ Indicator Channel	It carries the Hybrid ARQ acknowledgement to indicate to the UE whether the data should be retransmitted or not.

Figure 4.9: Downlink Physical Channels

Uplink Physical Channels	
PUSCH Physical Uplink Shared Channel	It carriers data of the user for very high data rates.
PUCCH Physical Uplink Control Channel	It is used for uplink control information and to send the answers ACK (acknowledgements) or NACK (not acknowledgements) to the eNodeB whether the downlink transmission was successfully received or not.
PRACH Physical Random Access Channel	It is used for random access.

Figure 4.10: Uplink Physical Channels

In LTE, power control has open loop in the downlink and closed loop in the uplink. The power control in uplink is necessary, that the interference at the eNB is reduced. The algorithm of closed loop power control based on average power over a SC-FDMA symbol. The received power is estimated at the UEs with the reference signal received power. The power control for the PUSCH channel [36], [41], [42] is as follows:

$$P_{PUSCH} = min\{P_{max}, 10 \cdot log_{10} \cdot M + P_0 + \alpha \cdot PL + \delta_{TF} + f\} \text{ [dBm]}$$

where,

- P_{max} is the maximum allowed transmit power of the UE class.
- M is the number of physical resource blocks (PRBS).
- P_0 is the noise power adjustment and is used to control SNR target.
- PL is the downlink path loss and is fractionally compensated up to the factor α with range [0,1].

- δ_{TF} is UE specific power offset. It is cell/UE specific modulation and coding scheme defined in the 3GPP specifications for LTE.
- f is a function, which is a UE specific correction value (an absolute or an accumulated) depends on δ_{TF}.

The power control for the PUCCH channel [42] is as follows:

$$P_{PUCCH} = min\{P_{max}, h(n_{CQI}, n_{HARQ}) + P_0 + PL + \delta + g\}$$

where,

- P_{max} is the maximum allowed transmit power of the UE class.
- $h(n_{CQI}, n_{HARQ})$ is a PUCCH format dependent value, that it depends on n_{CQI} and n_{HARQ}.
- n_{CQI} is the number of bits for CQI information.
- n_{HARQ} is the number of bits for HARQ ACK or HARQ NACK information.
- P_0 is the noise power adjustment and is used to control SNR target.
- PL is the full path loss.
- δ is UE specific power offset. It is cell/UE specific modulation and coding scheme defined in the 3GPP specifications for LTE.
- g is a function, which is a UE specific correction value and depends on δ.

4.3 Related work - Game Theoretic Models for Power Control

The issue of the power control has been extensively studied since the late 1980s, especially for CDMA systems. An appropriate and useful mathematical tool that has applied in wireless communications systems is the game theory. Firstly, the game theory has been applied in the CDMA systems. Specifically, there is a literature that the game theory is applied to the power control in a wireless network. The power control is continued to employ many authors. In the rest of this section, we see some studies, some game theoretic models for power control in CDMA and LTE networks.

We study the problem of power control that can be modeled as a non-cooperative game under the aspects of the game theory. The users in a network are considered as individual players. The aim of the players is to maximize their degree of satisfication that it is expressed as utility function. A non-cooperative power control game consists of the next basic components:

- A set of players, $\mathcal{N} = \{1, 2, ..., N\}$. The players are the users in the network.
- A set of possible actions for each player $i \in \mathcal{N}$, A_i. We have note, that $A = A_1 \times A_2 \times ... \times A_N$ is the strategy space. It is determined with respect to mobile terminal's physical limitations on the resources a_i that user i controls. Such as the transmission power, that $0 \leq p_i \leq p_i^{max}$, when $a_i = p_i$. An other example is the control of $0 \leq p_i \leq p_i^{max}$ and of the transmission rate $R_i \leq R_i^{max}$, when $a_i = (p_i, R_i)$.
- A set of utility functions mapping action profiles into the real numbers, $U_i : A \to \mathbb{R}$ for each player $i \in \mathcal{N}$. In the network, U_i are usually functions that represent the number of bits successfully transmitted per unit of battery energy, as in [43].

Therefore, a non-cooperative game or a non-cooperative network is usually represented as $G =< N, \{A_i\}, \{U_i\} >$. The aim of each user i is to select its strategy (a transmission power) from the set A_i in order to maximize its own uility. The choice of an utility function determines the nature of the game as well the actions of the users.

The power control in cellular networks has been extensively studied as an important mechanism to control Signal-to-Interference Ratios (SIR). The SIR is also known as CIR (Carrier-to-interference ratio or C/I), which in turn determine Quality-of-Service (QoS) metrics such as rate, outage, and delay [2]. We have note, that the reducing of the transmitter power can affect negative a certain link, because that link will also be more vulnerable to interference. The SIR at user i is denoted as follows:

$$SIR_i = \frac{G_{ii} P_i}{\sum_{j \neq i} G_{ij} P_j}$$

where, G_{ii} is the path gain of the signal path in cell i, P_i is the transmitter power used by the base station in cell i. The product $G_{ij} P_j$ is the interference power.

Thus, the aim is to adjust the power of each user for a given channel allocation, such that the interference levels at the receiver locations are minimized. We have note that this problem have studied and analyzed in several works and remain one of the most important problem in wireless networks. In continuous of this section, we will study some algorithms, which have designed in order to mitigate that problem. We can observe that most studies have been based on the work of Zander and on the work of Foschini and Miljanic.

Iterative power control algorithms invented since the early 1990s find a transmit power vector so as to ensure that each user attains the target SIR [44] while the overall power consumption is minimized. So, the aim of Zander was to investigate the performance of transmitter power control algorithms and to find performance bounds and conditions of stability for all types of transmitter power control algorithms [45]. Zander tried to minimize the outage probability, which is the probability that some randomly chosen link is subject to excessive interference. Startly, he considers a TDMA/FDMA scheme and m links. He proposed a centralized power scheme.

A minimum SIR, γ_0, is considered in the transmission system. This value is a threshold and is also called as *protection ratio*. The outage probability is defined as

$$P(\gamma_0) = \Pr\{SIR \le \gamma_0\} = \frac{1}{N}\sum_{j=1}^{N} \Pr\{SIR_j \le \gamma_0\}$$

It is shown that the maximum achievable SIR, SIR^*, is the following

$$SIR^* = \frac{1}{\lambda^* - 1}$$

where $\lambda^*$is the largest eigenvalue of the matrix $\mathbf{Z} = [Z_{ij}]$. In order to prove that there is a unique maximum achievable SIR was used the Perron-Frobenius theorem. That theorem is analyzed in the appendix A.

He proposed the *Stepwise Removal Algorithm* (SRA), which is a dynamic algorithm. In each step exists only one eigenvalue computation. At each time step, one cell is removed until all the CIR in all remaining cells are larger than γ_0. Then, the goal is to maximize the lower bound for the γ^* of the next matrix $\mathbf{Z}'$. Drawback of the SRA algorithm is its centralized nature because of the full information of the link gain matrix.

The pseudo-code of Stepwise Removal Algorithm is presented in the following table.

In 1993, Foschini and Miljanic [3] proposed an iterative distributed protocol in order to solve the problem of power control in cellular networks. A n×n matrix C=$\frac{\beta \cdot G_{i,j}}{G_{i,i}}$ for $i \neq j$, otherwise C=0 and a vector $\mathbf{u}$= $\frac{\beta \cdot v_i}{G_{i,i}}$ are defined. Therefore, the linear equation is P* = C$\cdot$ P^* + u, where the P^* is a unique

Algorithm 5 Stepwise Removal Algorithm (SRA)

Step 1:
Determine γ^* corresponding $\mathbf{Z}$.
If $\gamma^* \geq \gamma_0$
use the eigenvector $\mathbf{P}^*$
Else set $N' = N$
Step 2:
Remove the cell k for which the maximum of the row and column sums
$r_k = \sum_{j=1}^{N} Z_{kj}$ and $r_k^T = \sum_{j=1}^{N} Z_{jk}$
is maximixed and form the $(N'-1) \times (N'-1)$ matrix $\mathbf{Z}'$.
Determine γ^* corresponding to $\mathbf{Z}'$.
If $\gamma^* \geq \gamma_0$
use the eigenvector $\mathbf{P}^*$
Else set $N' = N' - 1$ and repeat the step 2.

vector. The iteration is $P[t+1] = C \cdot P[t] + u$, where t is the time step. In [3] proved that $P[t]$ congerges to P*, where P^* is a Nash equilibrium.

Specifically in [3], they assumed a cellular network with N links. Each link is a pair of a transmitter and a receiver. The channel gains are fixed. The SIR for the user i is denoted as

$$SIR_i = \frac{G_{ii} P_i}{\sum_{j \neq i} G_{ij} P_j + v_i}$$

where $G_{i,j}$ is the cross channel gain from the jth transmitter to the ith receiver, P_i is the power of the ith transmitter and v_i is the thermal noise power at the ith receiver.

In continuous, each SIR_i is constrained by a positive constant β_i. The aim is to minimize the total power such that $SIR_i \geq \beta_i$. Then, this constaint can be represented in matrix form as

$$(\mathbf{I} - \mathbf{C})\mathbf{P} \geq \mathbf{u}$$

where $\mathbf{P} = (P_1, ..., P_N)^T$ is the power vector, $\mathbf{u}$ is the vector of noise powers and is formulated as

$$\mathbf{u} = \left(\frac{\beta_1 v_1}{G_{11}}, \frac{\beta_2 v_2}{G_{22}}, ..., \frac{\beta_N v_N}{G_{NN}} \right)^T$$

The matrix $\mathbf{C}$ is denoted as

$$\mathbf{C}_{ij} = \begin{cases} \dfrac{\beta_i G_{ij}}{G_{ii}}, & i \neq j \\ 0, & i = j \end{cases}$$

where $i, j \in 1, ..., N$.

In continuous, Foschini and Miljanic showed that the iterative power control algorithm $\mathbf{P}[t+1] = \mathbf{C} \cdot \mathbf{P}[t] + \mathbf{u}$ congerges to $\mathbf{P}*$ when the Perron-Frobenius eigenvalue of $\mathbf{C}$, $\rho_C < 1$. Thus, the iterative power control algorithm, which can referred also as Foschini and Miljanic (FM) algorithm or Distributed Power Conttrol (DPC) algorithm can be rewritten as

$$\mathbf{P}_i[t+1] = \frac{\beta_i}{SIR_i[t]} \mathbf{P}_i[t], \forall i$$

We have note that the above work has studied on distributed power control algorithm (DPC) for wireless networks with fixed channels. Then, the authors of the paper [46] showed that the outcomes of the DPC algorithm do not accurately capture the dynamics of a time varying channel. So, the authors in [46] based on the algorithm of Foschini and Miljanic [3] and they added a new assumption that the channel gains G_{ij} are allowed to vary with time. In continuous, they showed that their power control algorithm converges to the optimal power allocation in a random channel environment.

Therefore, the distributed power control of [46] is formulated as

$$p\left[t+1\right] = C\left[t\right] \cdot p\left[t\right] + v\left[t\right]$$

In [46], they shown that the power $p\left[t\right]$ converges in distribution to a well defined random variable if and only if the Lyapunov exponent $\lambda_C < 0$ and it is defined as

$$\lambda_C = \lim_{t \longrightarrow \infty} \frac{1}{t} \log ||C[1]C[2]...C[t]||$$

In addition, it is proved that $p[t]$ converges weakly to a limit random variable $p(\infty)$ if $\lambda_C < 0$ and $\mathbb{E}\left[\log(1 + ||u[t]||)\right] < \infty$. Moreover, the next $\lim_{t \longrightarrow \infty} \mathbb{E}\left[\log(SIR_i[t])\right] = \log \beta_i$, $\forall i$ is valid. On the other hand, if $\lambda_C > 0$ then $p\left[t\right] \to \infty$, as $t \to \infty$.

It is noticed that the target SIR have changed in a random channel environment, rather than being limited $SIR_i = \beta_i$. The random version of the FM algorithm is limited by $\mathbb{E}[\log SIR_i] = log\beta_i$, where $\log \mathbb{E}[SIR_i] \geq \mathbb{E}[\log SIR_i]$.

Thus, the power updates of the random version of FM algorithm are unlikely to provide a minimum expected power solution, when the gain matrices

$G[t]$ are independent identically distributed. This algorithm does not sufficiently track any information on the random channel, for this reason they insert a new QoS criterion. They showed that the optimal power allocation can be found through a stochastic approximation algorithm. This stochastic process yielded an optimal fully distributed on-line algorithm for controlling transmitter powers in wireless networks.

An other study for the uplink power control is came from Yates in [12]. Specifically, Yates proposed an interference function $I\left(\mathbf{p}\right)$ in order to reduced the problem of the uplink power control in cellular networks. Yates defined the inequality $p_i \geq I\left(\mathbf{p}\right)$, where $\mathbf{p} = (p_1, ..., p_n)$ is the power vector of the N users, $I_i\left(\mathbf{p}\right) = \left(I_1\left(\mathbf{p}\right), ..., I_n\left(\mathbf{p}\right)\right)$ is the interference of other users that user i must overcome. Thus in a network, it is considered that there are M base stations and N users. The above inequality can be rewritten as:

$$p_i \geq I\left(\mathbf{p}\right) = \frac{\frac{\beta_i}{G_{i,j}}}{\sum_{i \neq j} G_{i,j} \cdot P_j + v_i}$$

where, $G_{m,i}$ is the channel gain of user i to base station m and v_m is the noise power at base station m.

Definition 4.3.1 (YATES) *An Interference function* $I(\boldsymbol{p}) = \left(I_1\left(\boldsymbol{p}\right), ..., I_n\left(\boldsymbol{p}\right)\right)$ *is standard if for all* $\boldsymbol{p} \geq 0$*, the following properties are satisfied:*

- *Positivity:* $I(\boldsymbol{p}) > 0$*, if* $\boldsymbol{p} > 0$
- *Monotonicity: if* $\boldsymbol{p} \geq \boldsymbol{p}'$ *then* $I(\boldsymbol{p}) \geq I(p')$
- *Scalabity:* $\forall a > 1$*,* $a \cdot I(\boldsymbol{p}) > I(a \cdot \boldsymbol{p})$

Therefore, when $I\left(\mathbf{p}\right)$ is a standard interference function then the iteration $\mathbf{p}[t+1] = I(\mathbf{p}[t])$ is called a standard power control algorithm. From the above properties, we can examine the convergence of a problem. Note that, if the standard power control algorithm has a fixed point then that fixed point is unique.

The idea of standard function has been used in some works such as in [47]. The authors in [47] considered the best strategy as a standard function. Finally, they showed that their non-cooperative power control game (NPG) has a unique equilibrium. The NPG is expressed as

$$\max_{p_j \subset P_j} U_j\left(p_j, p_{-j}\right), \; \forall j \in \mathcal{N}$$

where the utility U_j is denoted as

$$U_j\left(p_j, p_{-j}\right) = \frac{LR}{Mp_j} f(SIR_j), \text{ bits/joule}$$

where L is the information bits that a user transmits, R is the rate, M are bits that $M > L$ and p_j is the power of user j.

An other important point about the the convergence is came from the reference [48]. In [48], they study the *rate of convergence* for cellular radio systems using iterative power control algorithm. They showed that the FM algorithm converges to the fixed point at a *geometric rate*. An observation from [48] is that the increasing of the number of users is corresonding to slower speed of convergence in a cellular network.

Let N users, M base stations, $\mathbf{p} = [p_1, ...p_N]^T$ is the power vector of the system and $\lambda = \rho(\mathbf{G})$ is denoted the Perron-Forbenius eigenvalue of the matrix $\mathbf{G}$.

Theorem 4.3.1 *[48] Let $\lambda < 1$ in the FM algorithm. Starting from any initial power vector $p[0]$, the sequence $p[t]$ converges geometrically to the fixed point, such that $||p[t] - p^*|| \leq a^t ||p[0] - p^*||$, for some $a \in [0, 1)$.*

In [43], the authors proposed a non-cooperative power control game. They used the case of an AWGN channel, where the BER expressions for Binary Phase Shift Keying (BPSK), Differential Phase Shift Keying (DPSK), Coherent Frequency-Shift Keying (Coherent-FSK), Non Coherent Frequency-Shift Keying (Non-coherent FSK) modulation techniques. In [43], they used SIR and utility functions in order to achieve an efficient power control.

The utility function is the number of bits successfully transmitted per unit of battery energy, thus the utility for the use i by power level p_i, as follows

$$u_i(p_i, p_{-i}) = \tfrac{R}{p_i} \cdot f(\gamma_i), \quad \text{[bits/joule]}$$

where R is the rate that each user transmits information i.e bits/sec, γ_i is the SIR of user i and the function f is the efficiny function that depends on the BER (the bit error rate), i.e. $f(\gamma_i) = (1 - 2 \cdot BER(\gamma_i))^L$.

In continuous, they consider a game Γ_c with components N, $\{P_i\}$ and $\{u_i^c(\cdot)\}$. This game is called as non-cooperative power control game (NPGP). The utilities of that game are $u_i^c(p) = u_i(p) - c_i(p_i, p_{-i})$, where c_i is the pricing function. They take account a linear pricing scheme as $c_i(p_i, p_{-i}) = ca_i p_i$. The value of c is a constant.

Therefore, the multiobjective optimization problem that game solves and is as follows

$$\begin{gathered} \max p_i u_i^c(p_i, p_{-i}) \\ \text{s.t. } p_i \in P_i, \forall i \end{gathered}$$

The next interest move of the authors are the using of supermodularity. In supermodula games, each player's want to increase its strategy increases with an increase in other players' strategies. Therefore, they modified the NPGP with a parameter c and it is proved that the new game is a supermodular game. The Nash equilibrium of the new game is unique.

In [7], the authors present us a detailed example that the power control problem in a CDMA system can be solved by game theory. In this problem, each user's utility is increasing in his SINR and decreasing in his power level. In a real communication environment, when a user raises his transmission power then the interference by other user is increased. The authors in [7] define the SINR of user j as follows:

$$SINR_j = \gamma_j = \frac{W}{R} \frac{h_j \cdot p_j}{\sum_{i \neq j} h_i \cdot p_i + \sigma^2}$$

where h_j is the path gain from user j to the base station, p_j is the power transmitted by user j, W is the bandwidth, R is the rate that users transmit information, L is the bit packes and σ^2 is the additive white Gaussian noise (AWGN).

The utility function of user j has the unit of bits/J and is such that [7], [43]:

$$u_j = \frac{R}{p_j} \left(1 - 2BER(\gamma_j)\right)^L$$

where, BER(γ_j) is the bit error rate. In the utility function is applying non-coherent frequency shift keying (FSK) in an AWGN channel, then is rewritten as follows:

$$u_j = \frac{R}{p_j} \left(1 - e^{-0.5\gamma_j}\right)^L$$

In continuous of [7], a refereed game and a repeated power control game are considered. In the first game, they showed that if the base station is a "referee", then it could achieve a solution which is a Pareto improvement over the Nash Equilibrium of the simple pricing game. In their repeated game, the model is a discrete time system, where in each time slot, every user transmits one packet. Also, it is assumed that every user knows the received power of all transmissions in the previous time slots. Finally, these two games operates with the same way, when there are not cheats.

In a later work, J. Dams, M. Hoefer, T. Kesselheim in [4] proposed an other technique, the regret learning in non-cooperative networks, in order to solve drawbacks of the iteration of [3] such as lack of robustness, the adaption of power when the SINR is known. An important observation is that the iteration system of [4] converges from each *starting point* to a fixed point if that point exists. In the Figure 4.11 is represented the difference of these schemes. However, the iteration $p^t = C^t \cdot p^0 + \sum_{k=0}^{t-1} C^k \eta$ can not approach the fixed point. Thus. the time T is bounded until each transmission is "almost" feasible. The result is that from the starting point $p^0 = 0$ and after $t \geq \frac{\log \delta}{\log \lambda_{max}} \cdot n \cdot \log(3n)$ rounds implies that $(1-\delta)p^* \leq p^{\mathbf{t}} \leq p^*$, for all p^t.

Theorem 4.3.2 *If the starting point is an arbitrary* p^0 *and after* $\boldsymbol{t} \geq T = \dfrac{\log \delta - \log \max_i \left|\frac{p_i^0}{p_i^*} - 1\right|}{\log \max_i \left|1 - \frac{\eta_i}{p_i^*}\right|}$ *implies that* $(1-\delta)p^* \leq p^{\mathbf{t}} \leq (1+\delta)p^*$.

Proof *Let that the* p^* *is used as weights, then* $||x|| = \max_i \frac{x_i}{p_i^*}$ *In continuous, the matrix* C *is denoted as matrix norm,* $||C|| = \max_i \frac{1}{p_i^*} \sum_j C_{i,j} p_j^*$. *The term* $C \cdot p_i^*$ *is replaced by* $p_i^* - \eta_i$. *The the matrix norm of* C *is rewritten as* $||C|| = \max_i \left|1 - \frac{\eta_i}{p_i^*}\right|$. *Therefore, the distance* p^t *from the point* p^* *is defined as*

$$||p^t - p^*|| \leq ||C||^t \max_i \left|\frac{p_i^0}{p_i^*} - 1\right| \leq \delta$$

□

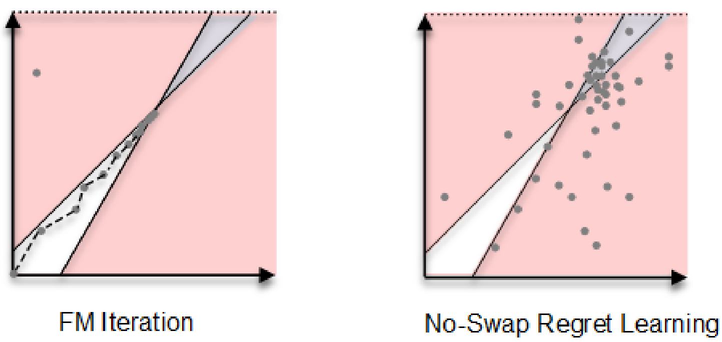

Figure 4.11: The comparison of FM iteration and No-swap Regret Learning.

Thus, the model of [4] is considered as a normal form game, where each sender i picks a transmission power as a strategy. Each user i chooses his power out of an interval from 0 to the maximal power level of user i, $[0, P_i^{max}]$. The $\Phi - regret$ user i is as follows:

$$R_i^{\Phi}(T) = sup_{\phi \in \Phi} \sum_{t=1}^{T} u_i(\phi(p_i^t), p_{-i}^t) - u_i(p_i^t, p_{-i}^t)$$

where Φ is a set of measurable functions, u_i is a utility function. The authors in [4] considered two cases for the set Φ, the external regret and the swap-regret.

The no-regret sequences can be computed in a distributed way. These algorithms are randomized. In the previous literatures [16], the space of actions is finite and the expected regret of a user after T rounds is at most $O(\sqrt{TNlogN})$. The authors in [4] study an algorithm $\boldsymbol{A}$, which can be used to construct an algorithm for power control on infinite action spaces achieving swap regret at most $O(T^{\frac{a+b}{1+b}})$. An other important point is that they prove that all no-swap-regret sequences converge to the optimal power vector p*. But, a no-external-regret sequence might make only 2 of n links successful. Finally, they use the definition of ε - correlated equilibrium in order to bound the probability that user i has successfully transmission. Bellow, we see the results of paper [4] with more details.

Initially, they consider the following utility function:

$$u_i(p) = \begin{cases} f_i(p_i), & if\ user\ i\ is\ successful\ with\ p_i\ against\ p_{-i} \\ 0, & otherwise \end{cases}$$

where $f_i : [0, p_i^{max}] \to [0, 1]$ is a continuous and strictly decreasing function for each $i \in [n]$. The power for user i is denoted as $p_i \in [0, p_i^{max}]$. The powers for all users except the user i is denoted as p_{-i}.

The following theorem show us that no-swap regret sequences can be computed in a distributed way. It is applied no-swap regret algorithm for finite action spaces on a suitable finite subset of the powers. The finite subset is constructed by the following method: the set of powers are divided into intervals of equal length and is used the right borders as the input action set for the algorithm. The discretization is chosen in an iteratively refined to guarantee that the no-swap regret property holds.

Theorem 4.3.3 *Let A be any no-swap regret algorithm for arbitrary finite action spaces, whose swap regret after T rounds in case of N actions is at most $O(T^a \cdot N^b)$, where a and b are suitable constants with $0 \leq a < 1$, $b \geq 0$. Then A can be used to construct an algorithm for power control on infinite action spaces achieving swap regret at most $O(T^{\frac{a+b}{1+b}})$.*

Proof *Let the utility function $u_i(\cdot, p_{-i})$ for the user i. Let $[0, p_i^{max}]$ is the set of strategies. That set is devided into N intervals of equal lenght. So, the* *utility at the right border of each interval is at most* $\frac{S_i \cdot p_i^{max}}{N}$ *worse than the maximum in the respective interval, where* $S_i = \max_{p_i,h} \frac{f_{p_i} - f_{p_i+h}}{h}$. *Then, we have*

$$u_i(x, p_i) \leq u_i(\tfrac{(k+1)p_i^{max}}{N}, p_i) + \tfrac{S_i \cdot p_i^{max}}{N}, \quad \forall x \in \left[\tfrac{kp_i^{max}}{N}, \tfrac{(k+1)p_i^{max}}{N}\right]$$

Let T is the number of steps. Then, the intervals are set $N = \lceil T^{\frac{1-a}{1+b}} \rceil$. *The algorithm use the finite strategy set of size N:* $\left\{\frac{p_i^{max}}{N}, \frac{2p_i^{max}}{N}, ..., p_i^{max}\right\}$.

Therefore, the regret is at most $O(T^a \cdot N^b) + T \cdot S_i p_i^{max} = O(T^{\frac{a+b}{1+b}})$. *This bound is produced from:*

- *Due to the restriction to the optimal strategies in the above finite set, the swap regret is at most* $O(T^a \cdot N^b)$.
- *Due to the restriction to the finite set, at T steps are lost at most* $T \cdot S_i p_i^{max}$.

□

In continuous, we see the analysis of the convergence of no-swap regret sequences to the optimal power vector p^* and the fraction of rounds in which each link is successful converges to 1. However, the next theorem gives us a lower bound for the number of steps that there is a successful transmission in a link. The power vector belongs to the closed interval $[(1-\delta)p_i^*, (1+\delta)p_i^*]$. The bound depends on the utility function and the fraction of rounds that we have a successful transmission. Finally, from the theorem we can conclude a bound converging to 1 as the swap regret per step approaches 0.

Theorem 4.3.4 *For every sequence* $p^1, ..., p^T$ *with swap regret at most* $\epsilon \cdot T$ *and for every* $\delta > 0$ *the fraction of steps in which user i sends successfully is at least*

$$Q \cdot \frac{f_i((1+\delta)p_i^*)}{f_i((1-\delta)p_i^*)} - \frac{\epsilon}{f_i((1-\delta)p_i^*)}$$

where Q denotes the fraction of rounds in which a power vector p with $(1-\delta)p_i^* \leq p \leq (1+\delta)p_i^*$ *is chosen.*

Using the definition of ε-correlated equilibrium, the above theorem can be rewritten as follows:

Proposition 4.3.1 *For every* ϵ*-correlated equilibium* π *and for every* $\delta > 0$ *the probability that user i sends successfully is at least*

$$Q \cdot \frac{f_i\left((1+\delta)\,p_i^*\right)}{f_i\left((1-\delta)\,p_i^*\right)} - \frac{\epsilon}{f_i\left((1-\delta)\,p_i^*\right)}$$

where $Q = \Pr_{p\sim\pi}\left[(1-\delta)\,p_i^* \leq p \leq (1+\delta)\,p_i^*\right]$.

Proof

$$\mathbb{E}_{p\sim\pi}\left[u_i\left((1+\delta)\,p_i^*, p_{-i}\right)|\mathcal{E}\right] - \mathbb{E}_{p\sim\pi}\left[u_i\left(p\right)|\mathcal{E}\right] \leq \frac{\epsilon}{\Pr_{p\sim\pi}[\mathcal{E}]} \quad (1)$$

- $\mathbb{E}_{p\sim\pi}\left[u_i\left((1+\delta)\,p_i^*, p_{-i}\right)|\mathcal{E}\right] \geq f_i\left((1+\delta)\,p_i^*\right)\cdot \Pr_{p\sim\pi}\left[p_{-i} \leq (1+\delta)\,p_{-i}^* |\mathcal{E}\right] \Rightarrow$
  $$\mathbb{E}_{p\sim\pi}\left[u_i\left((1+\delta)\,p_i^*, p_{-i}\right)|\mathcal{E}\right] \geq f_i\left((1+\delta)\,p_i^*\right)\cdot \frac{\Pr_{p\sim\pi}\left[(1-\delta)p_i^* \leq p \leq (1+\delta)p_i^*\right]}{\Pr_{p\sim\pi}[E]} \quad (2)$$
- $\mathbb{E}_{p\sim\pi}\left[u_i\left(p\right)|\mathcal{E}\right] \leq f_i\left((1-\delta)\,p_i^*\right)\cdot \Pr_{p\sim\pi}\left[\mathcal{S}|\mathcal{E}\right] \Rightarrow$
  $$\mathbb{E}_{p\sim\pi}\left[u_i\left(p\right)|\mathcal{E}\right] \leq f_i\left((1-\delta)\,p_i^*\right)\cdot \frac{\Pr_{p\sim\pi}[\mathcal{S}]}{\Pr_{p\sim\pi}[E]} \quad (3)$$

From the inequalities (1), (2) *and* (3), *we have the next inequality:*

$$f_i\left((1+\delta)\,p_i^*\right)\cdot \frac{\Pr_{p\sim\pi}\left[(1-\delta)p_i^* \leq p \leq (1+\delta)p_i^*\right]}{\Pr_{p\sim\pi}[E]} - f_i\left((1-\delta)\,p_i^*\right)\cdot \frac{\Pr_{p\sim\pi}[\mathcal{S}]}{\Pr_{p\sim\pi}[E]} \leq \frac{\epsilon}{\Pr_{p\sim\pi}[\mathcal{E}]} \Rightarrow$$

$$f_i\left((1+\delta)\,p_i^*\right)\cdot \Pr_{p\sim\pi}\left[(1-\delta)\,p_i^* \leq p \leq (1+\delta)\,p_i^*\right] - f_i\left((1-\delta)\,p_i^*\right)\cdot \Pr_{p\sim\pi}\left[\mathcal{S}\right] \leq \epsilon$$

□

In order to bound the probability $\Pr_{p\sim\pi}\left[(1-\delta)\,p_i^* \leq p \leq (1+\delta)\,p_i^*\right]$ is devided in the case of $\Pr_{p\sim\pi}\left[p > (1+\delta)\,p_i^*\right]$ and in the case of $\Pr_{p\sim\pi}\left[p > (1-\delta)\,p_i^*\right]$. Then

- $\Pr_{p\sim\pi}\left[p > (1+\delta)\,p_i^*\right] \leq \epsilon\left(\frac{n}{\delta}\max_i \frac{2}{s_i\eta_i} + 2\right)^{T+1}$
  $$\text{where, } T = \frac{\log\frac{\delta}{4} - \log\max_i\left|\dfrac{p_i^{max}}{(1+\frac{\delta}{2})p_i^*}\right|}{\log\max_i\left|1 - \dfrac{\eta_i}{p_i^*}\right|}$$
- $\Pr_{p\sim\pi}\left[p > (1-\delta)\,p_i^*\right] \leq \left(\frac{\epsilon}{1-r} + \Pr_{p\sim\pi}\left[p \nleq (1+\delta)\,p_i^*\right]\right)\left(\frac{n}{r}\right)^{T'+1}$
  $$\text{where, } T' = \frac{\log\delta}{\log\max_i\left|1 - \dfrac{\eta_i}{p_i^*}\right|} \text{ and } r \leq \min_i f_i\left((1-\delta)\,p_i^*\right)$$

In [5], the authors study algorithms in wireless networks where there are intereferences, using the Rayleigh model. For this reason, this model based on the SINR using stochastic propagation to address fading effects observed in reality. Also, they study the behavior of the external regret learning of some user at a time T. The authors apply the regret learning in ordrer to achieve the maximum capacity. In continuous, they proved that any no-regret learning algorithm, the number of successful transmissions needs to converge to a constant fraction of the non-fading optimum.

They used an utiliy function that depends on the success probability Q_i of link i. The utility for the user i is as follows:

$$u_i(q_1, ..., q_n) = \begin{cases} 2 \cdot Q_i(q_1, ..., q_n) - 1, & if \;\; q_i = 1 \\ 0, & if \;\; q_i = 0 \end{cases}$$

The success probability $Q_i(q_1, ..., q_n)$ of link i interpreted as follows: let each sender s_i transmits with probability q_i to the receiver r_i and the propagation is applied in a Rayleigh-fading environment, then the successful probability is defined as

$$Q_i(q_1, ..., q_n) = q_i \cdot e^{\frac{\beta\nu}{\bar{S}_{i,i}}} \prod_{j \neq i} \left(1 - \frac{\beta q_j}{\beta + \frac{\bar{S}_{i,i}}{\bar{S}_{j,i}}} \right)$$

where β is the threshold of the SINR, ν is a constant for the ambient noise and $\bar{S}_{j,i}$ is the received signal strength.

In continuous, it is considered a sequence $q^1, ..., q^t$ of action vectors that exhibits external regret $\epsilon \cdot T$ for each user $i = 1, ..., n$ and they showed that the average number of successful transmission is in $\Omega(OPT - \epsilon \cdot n)$. The term OPT is the size of the largest feasible set in the non-fading model under uniform transmission powers.

In [49], the authors designed a non-cooperative power control game (NCPCG) and a non-cooperative throughput game (NCTG). Also, they proposed an optimal complex centralized algorithm and is developed as a performance bound. Studied the social behavior of individual users in the proposed system model and the authors in [49] tried to enhance the overall system performance. The proposed schemes converge to the near-optimal solutions, compared with the optimal solutions from the centralized scheme.

Specifically, the NCPCG is denoted as $\max\limits_{P_i \leq P_{max}} u_i (P_i, P_{-i}, v_i)$, where P_i is the transmitted power level of user i and v_i is the assigned value function that depends on the throughput T_i and the bit error rate BER of user i. Thus, the value function is $v_i = \ln\left(\left(2^{T_i} - 1\right) / c_3^i\right) + 1$, where $c_3^i = -c_2^i / \ln\left(BER_i / c_1^i\right)$, $c_1 \approx 0.2$ and $c_2 \approx 1.5$, as well they used MQAM modulation scheme and the $BER \approx c_1 \exp^{c_2\left(\Gamma_i / 2^{T_i - 1}\right)}$. The utility function is represented as

$$u_i = P_i\,(v_i - \ln\Gamma_i)$$

where $\ln\Gamma_i$ is denoted the cost.

The main goal is to assign a v_i by the NCTG in order to maximize the throughput of the overall system with the constraint of $P_i \le P_{max}$ for all users.

Therefore, the NCTG is denoted as $\max_{T_i} u'_i$, where the utility function $u'_i = T_i \cdot L$ and L is an indication function for system feasibility. Note that the game starts from any feasible initaial value.

In [21], the authors study the behavior of individual distributed secondary user to control its rate when the prime user is absent. The aim of each secondary user is to maximize its rates over different channels. The authors in [21] proposed a distributed protocol based on an adaptive learning algorithm for multiple secondary users using only local information. The proposed learning algorithm in [21] based on no-regret learning and converges to a set of correlated equilibria with probability one. Finally, they showed that the optimal correlated equilibria has better fairness and better performances than the Nash equilibrium.

Specifically, the authors considered that there are N channels, M primary users and K secondary users in a wireless network. These channels are shared among M and K. They defined an interference matrix L, that the adjacent secondary users can interference with each others. Also, they defined a channel availability matrix A(t), that each user can transmit over a specific channel with a set of variety rates. For each available channel, a secondary user can select L+1 discrete rates $Y = \{0, v_1, ..., v_L\}$. The strategy space Ω_i for secondary user i is on the available channels and can be denoted as:

$$\Omega_i = \prod_{n=1}^{N} Y^{A_{in}}$$

The action of user is $r_i^n = v_l$ representing user i occupies channel n by rate u_l. The strategy profile is defined as $r^n = (r_1^n, r_2^n, ..., r_K^n)'$. The utility function is defined as the maximum achievable rate for the secondary users over all the available channels as follows:

$$U_i = \sum_{n=1}^{N} A_{in} R_i(r_i^n, r_{-i}^n)$$

where $R_i(r_i^n, r_{-i}^n)$ is the outcome of resource competition for user i and the other users.

They proposed two refinements. The maximum sum correlated equilibrium that maximize the sum of utilities of the secondary users. And the maximum fair correlated equilibrium that seeks to improve the worst case situation. In continous is presented the linear programming problem:

$$max_p \sum_{i \in K} E_p(U_i) \text{ or } max_p min_i E_p(U_i)$$
$$\text{s.t.}$$
$$p(r_i, r_{-i})[U_i(r_i', r_{-i}) - U_i(r_i, r_{-i})] \leq 0$$
$$\forall r_i, r_i' \in \Omega_i, \forall i \in K.$$

where $E_p(.)$ is the expectation over p. The constraints guarantee the solution is within the correlated equilibrium set.

They used the regret-matching learnig algorithm. For any two distinct actions $r_i \neq r_i'$ in Ω_i and at every time T, the regret of user i at time T for not playing r_i' is

$$\mathbb{R}_i^T(r_i, r_i') := max\{D_i^T(r_i, r_i'), 0\}$$

where D_i^T is the average payoff that the user i would have obtained, if it had played action r_i' every time in the past instead of choosing the action r_i.

The adaptive learning algorithm can guarantee us that the relative frequency of users's action r converges almost sure to a set of correlated equilibrium.

The relative frequency of users's action r that the use play till T periods of time is defined as

$$z_T(r) = \frac{1}{T}\#\{t \leq T : r_t = r\}$$

The pseudo-code of Regret-Matching Learning Algorithm is presented in the following table.

The authors in [50]propose a Coalitional Game Theoretical mechanism and a Correlated equilibrium Game Theoretical mechanism, in order to achieve Co-Channel Interference (CCI) mitigation in a distributed manner. They considered two types of cognitive base stations in a Long Term Evolution (LTE) network, i.e. the macro cell evolved NodeB (eNB) and the femtocell Home evolved NodeBs (HeNB). The UMTS LTE proposes a distributed network architecture. In the case of Correlated Equilibrium Games, the authors compare the Nash equilibria (NE) and the correlated equilibria, that the latter has more advantages than the NE. Then, the authors propose that the resource block (RB) selection of eNBs in the downlink of an LTE system can be formulated as a correlated equilibrium game. The solution of this game can be taken by using the non-regret learning algorithm.

An other study is the combination of power control (PC) and the Inter-cell interference control (ICIC) based on game theory [51]. This combination had

Algorithm 6 Regret-Matching Learning Algorithm by [21]

Initialization:
For each user i assign arbitrarily probability $p_i^1(r_i)$
For $t = 1$ to T
Step 1:
Find $D_i^T(r_i, r_i')$.
Step 2:
Find average regret $\mathbb{R}_i^T(r_i, r_i')$.
Step 3:
Let $r_i \in \Omega_i$ be the strategy last chosen by i, i.e. $r_i^t = r_i$.
Then, $p_i^{t+1}(r_i') = \frac{1}{\mu}\mathbb{R}_i^T(r_i, r_i')$, $\forall r_i' \neq r_i$
Then, $p_i^{t+1}(r_i) = 1 - \sum\limits_{r_i' \neq r_i} p_i^{t+1}(r_i')$
where μ is a certain constant that is sufficient large.

as a result to achieve a better performance. Also, the authors study their model in a LTE network. In wireless networks, the SINR in the center of the cell is different from the one in the outage zones. This difference in a cell give not equal service quality to the users. They proposed a model that based on defining different roles of the users within a cell. The roles of a user is determined by its activity ($\sum\limits_{j=1}^{k} UEA_j$), distance from eNB ($UEP0s_i$), load of the system, type (UET_i). So, the role of a user i can be formulated as

$$Role_i(PCA_i) = \left\{ UET_i, \sum_{j=1}^{k} UEA_j, UEP0s_i, \max[uf_i(ST_i, p_i, P_i)] \right\}$$

where, uf_i is the utility function of user i that depends on the service type such as voice, data, text, image, video, also depends on the power level p_i and the overall power limit in the cell P_i. Then, the utility measures the quantity of information that is received and the throughput that is achieved by consuming a basic unit of energy. The utility is the total number of correct bits that a user can transmit per unit of its battery energy.

In paper [52], they study a no-regret learning algorithm for simulateneous power control and channel allocation in cognitive radio networks. Specifically, they tried to find an algorithm for an exact potential game that allows cognitive radio pairs to update their transmission powers and frequencies simulateneously. This algorithm converges to a pure Nash equilibrium. They observed through simulations that the no-regret learning algorithm can achieve the same performance as the traditional potential game. In the no-regret learning algorithm, the player knows the strategy in the current round and the utilities of all possible strategies.

The authors studied the game $\Gamma = \left\{N, \{S_i\}_{i\in N}, \{U_i\}_{i\in N}\right\}$. Thus, N is the finite set of players. The players are transmitting-receiving pairs. In continuous, S_i is the set of strategies associated with player i and $U_i : S \to R$. They used the utility function from the paper [53], as follows:

$$U_i(s_i, s_{-i}) = T_i^1(s_i, s_{-i}) + T_i^2(s_i, s_{-i}) + T_i^3(s_i, s_{-i})$$

where the term T_i^1 express the impact that other players have on the interference sensed by the receiver in the pair i, as follows:

$$T_i^1(s_i, s_{-i}) = -\sum_{j\neq i, j=1}^{N} p_j h_{ji} I(j,i)$$

The term T_i^2 give the impact of a potential action for player i on the interference observed by all other users, as follows:

$$T_i^2(s_i, s_{-i}) = -\sum_{j\neq i, j=1}^{N} p_i h_{ij} I(i,j)$$

The term T_i^3 depends only on the action selected by player i and provides an incentive for individual players to increase their power levels, as follows:

$$T_i^3(s_i, s_{-i}) = a\log(1 + p_i h_{ii}) + \frac{\beta}{p_i}$$

The $\frac{\beta}{p_i}$ term takes account the utility associated with longer battery life.

The game Γ with the above utility has proved that is an exact potential game [53].

In each iteration, the player examine all the various possibilities of the action space, by calculating the corresponding weights of actions of the player. Each player has $|C| \times |P|$ weights. Therefore, each player catch the action pair with the highest weight. The weight assigned to action pair $s_i = (c_i, p_i) \in S_i$, at time $t+1$, is denoted as q_i^{t+1}. Also, it is used a parameter γ, which is defined as $0 < \gamma < 1$.

After each iteration, the above no-regret algorithm updates weights associated with each action based the cumulative utility function, $U_i^t(s_i) = \sum_{j=1}^{t} U_i(s_i, s_{-i}^j)$ and calculates again the weights associated with the set of its strategies for the next round. We have note in this algorithm that, the players update their actions simultaneously, instead to the traditional potential game, that the players update their actions sequentially. Finally, they observed through simulations that the no-regret learning algorithm is slower than a traditional potential game. However, it is noticed that this learning algorithm is very useful for the evolving of wireless networks.

The pseudo-code of No-regret algorithm is presented in the following table [52].

Algorithm 7 No-Regret Algorithm by [52]

Initialization:
For each user i assign randomly a power level and a channel.
Procedure:
For $t = 1$ to T
For $i = 1$ to N
For $k = 1$ to $|P|$
For $kk = 1$ to $|C|$
Calculate U_i as $U_i\left(s_i, s_{-i}\right) = T_i^1\left(s_i, s_{-i}\right) + T_i^2\left(s_i, s_{-i}\right) + T_i^3\left(s_i, s_{-i}\right)$.
Calculate the weight $q_i^{t+1} = \dfrac{(1+\gamma)^{U_i^t(s_i)}}{\sum\limits_{s_i'} (1+\gamma)^{U_i^t\left(s_i'\right)}}$.
End for kk
End for k
Select the largest weight for all the users,
Assign the power level and channel corresponding to the largest weight to all the users.
End for i
End for t

Chapter 5

Proposed System Model

We assume that our network is compatible with the LTE Release 10 and beyond (LTE-Advanced) in 3rd Generation Partnership Project (3GPP) standard for wireless data communications. We follow the parameters settings agreed in 3GPP such as Bandwidth, Modulation Coding Schemes (MCS), Carrier Frequency, Path Loss, Thermal Noise [39]. In our scheme, the wireless devices are in a non-cooperative network. Our goal is the minimum transmission power with the maximum throughput in a realistic environment. Therefore, our object is to design a framework for power control using regret learning algorithm.

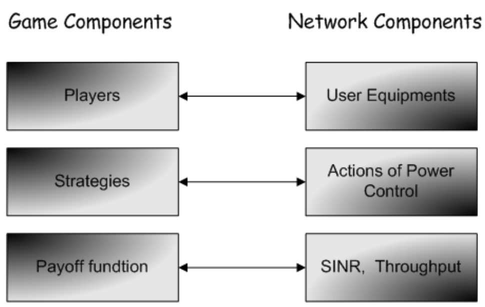

Figure 5.1: The connection between of the game theory and the wireless networks

5.1 Network Model

5.1.1 Signal Model-SINR

We consider that our network consists of i receivers and j senders (i, j = 1,...,N). Each sender is a selfish agent and use a power level to transmit packets with success to the receiver, so the sender j transmit the signal at power p_j multiplied by the gain $G_{i,j}$ and by the $F_{i,j}$ model Rayleigh fading. In particular, the gain represent the distance between sender and receiver. In the Rayleigh fading environment there are many objects that scatter the radio signal before it arrives at the receiver and the received signal power has an exponential distribution.

The received signal in Rayleigh fading channel is given from the follow:

$$\text{Y=H}\cdot\text{X+Z},$$

where Z is AWGN at tha base station, X is the transmitted signal by user i and H is the channel gain.

Therefore, the received signal at each input is given by

$$r_i = \sum_{j=1}^{N} \sqrt{G_{ij} \cdot F_{ij} \cdot P_j} b_j s_j + n_i$$

where, b_j are data bits taking on values of ± 1 with equal probability, s_j is the fixed k-dimensional spreading sequence of user j with elements taking values $\pm 1/\sqrt{k}$, n_i is assumed to be additive white Gaussian noise (AWGN) with zero mean and covariance $\sigma^2 I$.

The SINR for user i is as follows:

$$SINR_i = \frac{G_{i,i} \cdot F_{i,i} \cdot P_i}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i} \geq \beta_i$$

The transmission of packet is successful if the above conastraint of $SINR_i$ is satisfied. The SINR in dB for user i is as follows:

$$SINR_i = 10 log_{10} 10 \ \frac{G_{i,i} \cdot F_{i,i} \cdot P_i}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i}$$

where $v_i = \sigma^2 I$ is the noise at the receiver i and β_i is a threshold of SINR. The communication can become more reliable when we use the ratio bit energy (E_b) per noise spectral density (N_0) such that

$$\frac{E_b}{N_0} = \frac{BW}{R} \ \frac{G_{i,i} \cdot F_{i,i} \cdot P_i}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i}$$

where BW is the channel bandwidth and R is the channel data rate.

5.1.2 Rayleigh Fading

The Rayleigh distribution has a probability density function (pdf) [54] given by

$$f(z) = \begin{cases} \frac{z}{\sigma^2} e^{\frac{-z^2}{2\sigma^2}}, & 0 \leq z \leq \infty \\ 0, & z < 0 \end{cases}$$

where z is a random variable and σ^2 is the fading envelope of the Rayleigh distribution.

5.1.3 Shadowing

The wireless signals are blocked by many objects such as high building and mountains [54]. This often hapend in large urban areas.

$$p(s) = \frac{e^{-((lns - m_s)^2/2\sigma_s^2)}}{s\sigma_s\sqrt{2\pi}}$$

5.1.4 Path Loss

Path loss represents signal attenuation between the effective transmitted power and the receiver power. Path loss is measured on dB.

$$pathloss = \frac{g}{d_{ij}^n}$$

where, n is the path loss exponent for different propagation environments depending on the characteristics of the communication medium, d_{ij} is the distance between transmitter of the jth link to receiver of the ith link and g is a constant equals to 1.

5.1.5 Outage Probability

The outage probability $Pout_i$ of the ith receiver/transmitter pair is given by

$$Pout_i = Prob(SINR_i \leq \beta_i) = Prob(G_{ii} \cdot F_{ii} \cdot P_i \leq \beta_i \cdot [\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i])$$

The outage probability $Pout$ of the system is given by

$$Pout = \max_i Pout_i$$

which means that $Pout$ is the worst outage probability of the ith receiver/transmitter pairs.

Theorem 5.1.1 *The outage probability $Pout_i$ of the ith receiver/transmitter pair in a Rayleigh fading is given by*

$$Pout_i = 1 - e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \frac{G_{ii} \cdot P_i}{G_{ii} \cdot P_i + \beta_i \cdot G_{i,j} \cdot P_j}$$

Proof. We assume that all the received powers $x_1, ..., x_n$ are independent exponentially distributed, with means $\mathbf{E}[x_i] = 1/\lambda_i$. The value $x_1 = G_{ii} \cdot F_{ij} \cdot P_i$ is the received power from the desired user and the received mean value of $\mathbf{E}[x_1] = \mathbf{E}[G_{ij} \cdot F_{ij} \cdot P_j] = G_{ij} \cdot P_j$ [55], [5]. The outage probability for the user can be expressed as

$$Pout_i = Pr[SINR_i \leq \beta_i] = Pr\left[\frac{G_{ii} \cdot F_{ii} \cdot P_i}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i} \leq \beta_i\right]$$

$$= Pr\left[G_{ii} \cdot F_{ii} \cdot P_i \leq \beta_i \cdot (\textstyle\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i)\right]$$

$$= 1 - Pr\left[G_{ii} \cdot F_{ii} \cdot P_i > \beta_i \cdot (\textstyle\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i)\right]$$

$$Pr\left[x_1 > \beta_i \cdot (\textstyle\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i)\right] = \mathbf{E}\left\{\exp\left[-\frac{\beta_i \cdot (\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i)}{\mathbf{E}[x_1]}\right]\right\}$$

$$= e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \int_{t_2=0}^{\infty} \cdots \int_{t_n=0}^{\infty} e^{-\lambda_1(t_2+...+t_n)} \prod_{i=2}^{n} \lambda_i e^{-\lambda_i t_i} dt_2 ... dt_n$$

$$= e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i=2}^{n} \int_{t_i=0}^{\infty} \lambda_i e^{-(\lambda_1+\lambda_i)t_i} dt_i$$

$$= e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i=2}^{n} \frac{\lambda_1}{\lambda_1 + \lambda_i}$$

5.1.6 Modulation Schemes

For the numerical results, we have considered a system employing Quadrature phase-shift keying (QPSK), 16 Quadrature amplitude modulation ($16QAM$), 64 Quadrature amplitude modulation ($64QAM$), for which the Bit Error Rate (BER) in Additive White Gaussian Noise ($AWGN$) is given respectively by [56], [57]:

$$BER_{QPSK} = \frac{1}{k} \frac{2}{\sqrt{\pi}} \int_x^{\infty} \quad e^{-(\frac{E_b \cdot k}{2 \cdot N_0})} \quad d(\sqrt{\frac{E_b \cdot k}{2 \cdot N_0}})$$

$$BER_{16QAM} = \frac{3}{2k} \frac{2}{\sqrt{\pi}} \int_x^{\infty} \quad e^{-(\frac{E_b \cdot k}{10 \cdot N_0})} \quad d(\sqrt{\frac{E_b \cdot k}{10 \cdot N_0}})$$

$$BER_{64QAM} =$$
$$\frac{7}{4k}\frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-(\frac{E_b \cdot k}{42 \cdot N_0})} d(\sqrt{\frac{E_b \cdot k}{42 \cdot N_0}}) - \frac{49}{64} \cdot (\frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-(\frac{E_b \cdot k}{42 \cdot N_0})} d(\sqrt{\frac{E_b \cdot k}{42 \cdot N_0}}))^2$$

where $k = log_2 M$.

5.2 Power Control in Uplink

5.2.1 Power Control Game

Theorem 5.2.1 *A strategic game G with components N, A_i and u_i has a NE if $\forall i \in N$, the set $A_i \neq \varnothing$ is a compact convex subset of a Euclidian space and the payoff function u_i is continuous and quasi-concave on A_i [1].*

Definition 2. [Dams, Hoefer and Kesselheim (12)] In a non-cooperative game G, we assume that each link i chooses his power out of an interval P_i =[0,p_i^{max}], which is a strategy set for the ith user.We define p_i^{max} as the maximum power level user i can use.

Utility Let $u_i : A \to \Re$ is the utility function fot the user i, each user selects a power level p_i from the set P_i. The power set of all the users except the user is is denoted by p_{-i}. The utility level for the ith user is:

$$u(p_i, p_{-i}) = f_{1i}(SINR(p_i, p_{-i})) - f_{2i}(p_i)$$

where, $f_{1i}(\cdot)$ is a function of the SINR and $f_{2i}(\cdot)$ is the energy price, as follows

$$f_{1i}(SINR(p_i, p_{-i})) = Throughput(i) \cdot S$$

$$f_{2i}(p_i) = c_i \cdot p_i$$

where, S is a constant of our system.

Throughput In this model, we take a function to compute the throughput as follows [58]:

$$Throughput(i) = BW \cdot log_2(1 + \frac{SINR_i}{d_{i,j}^n})$$

where, BW is the channel bandwidth, d_{ij} is the distance between transmitter of the jth link to receiver of the ith link.

Theorem 5.2.2 *In this power control game G exists a Nash equilibrium.*

Proof. To prove the existence, we have to show that the condition $\frac{\partial^2 u(P_i,P_{-i})}{\partial^2 P_i} < 0$ is valid.

The first order partial derivative function of the utility function with respect to P_i is as follows:

$$\frac{\partial u(P_i, P_{-i})}{\partial P_i} =$$
$$BW \cdot S \cdot \left(1 + \frac{G_{i,i} \cdot F_{i,i} \cdot P_i}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i}\right)^{-1} \cdot \frac{G_{i,i} \cdot F_{i,i}}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i} - c_i$$

The second order partial derivative function of the utility function with respect to P_i is as follows:

$$\frac{\partial^2 u(P_i, P_{-i})}{\partial^2 P_i} =$$
$$-BW \cdot S \cdot \frac{G_{i,i}^2 \cdot F_{i,i}^2}{\left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)^2} \cdot \left(1 + \frac{G_{i,i} \cdot F_{i,i} \cdot P_i}{\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i}\right)^{-2} < 0$$

Theorem 5.2.3 *In this power control game G the Nash equilibrium is unique.*

Proof. To prove the uniqueness, we take an interference function $I(p) = (I_1(p), I_2(p), ..., I_N(p))$ such that $p_i \geq I\left(\mathbf{p}\right)$ and the proporties of the definition in [12] must be satisfied.

The outage probability of the ith receiver/transmitter pair is given in the previous theorem. We want to minimize the outage probability of the ith user in order to reduced his power. Then, $Pout_i \leq C_i$

$$1 - e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \frac{G_{ii} \cdot P_i}{G_{ii} \cdot P_i + \beta_i \cdot G_{i,j} \cdot P_j} = 1 - e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \frac{1}{1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}} =$$
$$1 - e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]^{-1}$$

$$Pout_i \leq C_i \qquad \Longrightarrow$$

$$1 - e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]^{-1} \leq C_i \qquad \Longrightarrow$$

$$1 - C_i \leq e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]^{-1} \qquad \Longrightarrow$$

$$e^{\frac{v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right] \leq \frac{1}{1 - C_i} \qquad \Longrightarrow$$

$$log \left(e^{\frac{v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right] \right) \leq log \left(\frac{1}{1 - C_i} \right) \qquad \Longrightarrow$$

$$\frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{P_i \cdot \sum_{i \neq j} log \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right]}{log\left(\frac{1}{1-C_i}\right)} \leq P_i$$

The interference function is given by:

$$I(p) = \frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{P_i \cdot \sum_{i \neq j} log \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right]}{log\left(\frac{1}{1-C_i}\right)}$$

This function is positive because of the $G_{ij} > 0$, $G_{ii} > 0$, $P_i > 0$, $P_j > 0$, $v_i > 0$, $\beta_i > 0$ and $0 < C_i < 1$. Then the property "$I(p) > 0$, if $p > 0$ "is satisfied.

The property of scalabity "$\forall a > 1$, $a \cdot I(p) > I(a \cdot p)$" is satisfied, as follows:

$$I(a \cdot p) = \frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{a \cdot P_i \cdot \sum_{i \neq j} log \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot a \cdot P_j}{G_{ii} \cdot a \cdot P_i} \right]}{log\left(\frac{1}{1-C_i}\right)}$$

$$= \frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{a \cdot P_i \cdot \sum_{i \neq j} log \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right]}{log\left(\frac{1}{1-C_i}\right)}$$

$$< a \cdot \left[\frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{P_i \cdot \sum_{i \neq j} log \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right]}{log\left(\frac{1}{1-C_i}\right)} \right]$$

$$= a \cdot I(p)$$

The property of monotonicity "if $p \geq p'$ then $I(p) \geq I(p')$" is satisfied, as follows: The interference function $I(p) = \frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{P_i \cdot \sum_{i \neq j} log \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} \right]}{log\left(\frac{1}{1-C_i}\right)}$ can be written as:

$$I(p) = \frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{\frac{G_{ii} \cdot P_i}{\beta_i} \cdot \sum_{i \neq j} log\left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]}{\frac{G_{ii}}{\beta_i} \cdot log\left(\frac{1}{1-C_i}\right)}$$

$$I(p) = \frac{v_i \cdot \beta_i}{G_{ii} \cdot log\left(\frac{1}{1-C_i}\right)} + \frac{\sum_{i \neq j} log\left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]^{\frac{G_{ii} \cdot P_i}{\beta_i}}}{\frac{G_{ii}}{\beta_i} \cdot log\left(\frac{1}{1-C_i}\right)}$$

Firstly, if $P_j \geq P'_j$ for $j \neq i$ then $I(P_j) \geq I(P'_j)$ is satisfied for $j \neq i$. Proving the monotonicity of P_i, we consider $a = G_{i,j} \cdot P_j$ as a positive constant. The known function $f(x) = log(1 + \frac{a}{x})^x$ for $a > 0$ is monotonic. Then I(p) is also monotonic.

Therefore, the term $I(\mathbf{p})$ is denoted as standard interference function, as well the properties of Yates [12] are satisfied. Thus, the game converges to the fixed and unique solution p^*, where $p^* = I(\mathbf{p}^*)$.

But, we want the minimum outage probability. So, we can redefined the standard interference function as $A(p) = \min I(p)$. We can observe that the game converges to p^* as well the solution $p^* = A(p^*)$.

In continuous, we can observe that we can find an upper and a lower bound for the SINR, as well $\beta_i \leq SINR \leq \frac{1}{\beta_i} \log\left(\frac{1}{1-C_i}\right)$. The parameter C_i is the threshold of the outage probability of the ith user. The upper bound about SINR is analyzed as follows:

$$Pout_i \leq C_i \quad \Longrightarrow$$

$$1 - e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]^{-1} \leq C_i \quad \Longrightarrow$$

$$1 - C_i \leq e^{\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i}} \prod_{i \neq j} \left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right]^{-1} \quad \Longrightarrow$$

$$\log(1 - C_i) \leq \frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i} - \sum_{i \neq j} \log\left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right] \quad \Longrightarrow$$

It is known that $\log(1 + z) \leq z$, for $z \geq 0$. From this, we take that $\sum_{i \neq j} \log(1 + z) \leq \sum_{i \neq j} z$, for $z \geq 0$. Then:

$$\frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i} - \sum_{i \neq j} \log\left[1 + \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i}\right] \leq \frac{-v_i \cdot \beta_i}{G_{ii} \cdot P_i} - \sum_{i \neq j} \frac{\beta_i \cdot G_{i,j} \cdot P_j}{G_{ii} \cdot P_i} = -\beta_i \cdot SINR$$

Therefore, we have that

$$\log(1 - C_i) \leq -\beta_i \cdot SINR \quad \Longrightarrow$$

$$\beta_i \cdot SINR \leq \log(1 - C_i)^{-1} = \log\left(\frac{1}{1 - C_i}\right)$$

We conclude that

$$SINR \leq \frac{1}{\beta_i} \log\left(\frac{1}{1 - C_i}\right)$$

Finally, from the upper and the lower bounds of the SINR, we can observe that a power of ith user p_i can belong to the next closed interval

$$p_i \in \left[\frac{\beta_i \cdot \left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)}{G_{ii}}, \frac{\left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)}{\beta_i} \log\left(\frac{1}{1 - C_i}\right)\right].$$

5.2.2 Power Control via Regret Learning

In this section, we see that the convergence to the optimal power p^* is guaranteed, as well we use the correlated equilibrium from game theory. An correlated equilibrium (CE) behaves similar with that of mixed equilibrium (MNE) of the game theory because of the probability distribution over strategy vectors. From the chapter 3, we know that the set of MNE is a subset of CE.

Below, we give an example, in which there is CE. Initially, we will prove that this example has MNE.

Let that we have two players P_I, P_{II}. In the below matrix, these players with their differents actions C and D are represented. If the both players choose the action C then they have the best overall benefit. The action C gives us the choice of the cooperation and in the case of wireless networks is the transmission with the less aggressively. But, one of the two players can become more aggressive. In that case he will choose the action D, while the other user will remain in the action C. Then, the first player will obtain better utility function. But, the other player will obtain lower utility function. Then the overall benefit is decreased. However, the both players can become aggressive (D, D) then the utility function of the both users will be very low.

	C	**D**
C	$(6, 6)$	$(2, 7)$
D	$(7, 2)$	$(0, 0)$

Initially, we observe, in the above matrix, that (D, C) and (C, D) are Pure Nash equilibrium (PNE).

In the case of mixed strategies, the user P_I will follow a probability distribution p over action C and a probability distribution $1-p$ over action D. Accordingly for the user P_I, which will follow a probability distribution q over action C and a probability distribution $1-q$ over action D. If the user P_I follows the action C, then the expected profit of P_I is equal to $6p+2(1-p)$. If the user P_I follows the action D, then the expected profit of P_I is equal to $7p+0(1-p)$. We have note that this game is symmetric. Similarly, the accordance outcomes for player P_I are calculated. We can conclude that the users P_I, P_{II} choose C and D with probabilities 2/3 and 1/3, respectively. Finally, the expexted utility function for each player i is equal to $\overline{u}_i = 4\frac{2}{3}$.

The case of correlated equilibrium is applied as follows: Let that there is a trusted party, which tells each player what to do based on the outcomes (C, D), (D, C) and (C, C). Each of them outcome has probability 1/3. We have note that the game has correlated equilibrium if no user wants to deviate from the trusted party's instruction. Let that the trusted party tells player P_I to choose the action D, then P_I has no incentive to deviate because of his payoff = 7. This payoff is the highest for him. Thus, the player P_I knows the outcome (C, D) and that P_I will obey the instruction to remain in the action C.

Let that the trusted party tells player P_I to choose the action C. P_I knows that the outcome must have been either (D, C) or (C, C), each happening with equal probability 1/2. Then, the expected utility of P_I on playing C (no deviate from the trusted party's instruction) is equal to $\overline{u}_C = 6\cdot1/2+2\cdot1/2 = 4$, where 6 is the payoff from P playing C and the number 2 is the payoff from P playing D. If the player P_I decides to deviate from the trusted party's instruction playind the action D instead of C, then the expected utility of P_I on playing D is equal to $\overline{u} = 7\cdot1/2+0\cdot1/2 = 3.5 < 4$. We can observe that the last expected utility is lower than the expected utility from the trusted party's instruction. So, P_I has no incentive to deviate. Finally, the expected utility function for each player i is equal to $u_i = 7\cdot1/3+6\cdot1/3+2\cdot1/3 = 5 > 4\frac{2}{3}$. We conclude that the expected utility of CE is higher than the expected utility of MNE. Therefore, the set of MNE is a subset of CE.

Howeever, in a game there is a case to exist more than one CE. In order to adjust the strategies of players and converge to the set of CE, they can track a set of "regret" values for strategy update. Remember, the regret is the expected utility function that each user i would have obtained, if that player had adopted an other action a' every time in the past instead of the action a.

In this section, we use the definition of ϵ-correlated equilibrium [28] and is defined as follows:

Definition 5.2.1 **ϵ-CE:** *A joint probability distribution π over the set of power* *vectors* $P_1 \times ... \times P_n$*, where* $P_i = [0, p_i^{max}]$*, is an* ϵ*-correlated equilibrium if for every user i and for any function* $\phi : P_i \to P_i$*, we have*

$$\mathbb{E}_{p\sim\pi}\left[u_i\left(\phi_i\left(p_i\right), p_{-i}\right)\right] - \mathbb{E}_{p\sim\pi}\left[u_i\left(p_i, p_{-i}\right)\right] \leq \epsilon$$

where p_{-i} *denotes the joint actions of the other users.*

It means that π is an ϵ-correlated equilibrium if each user can increase its expected utility at most ϵ. From the paper of [4], we have that a game has swap regret at most RT, for T times steps each user follow a strategy. Therefore, each sequence $p^1, ..., p^T$ of swap regret at most R corresponds to an R/T-correlated equilibrium.

Moreover in [4], the power p_i of ith user belongs to the interval $[(1-\delta)\,p^*, (1+\delta)\,p^*]$. The swithcing operation $\phi_i\left(p_i\right)$ is equal to $(1+\delta)\,p^*$, as well the user i could always choose $(1+\delta)\,p^*$.

Therefore, the expected utility due to the above switching operation, using the utility function of section 5.2.1, can be written as follows:

$$\mathbb{E}_{p\sim\pi}\left[u_i\left((1+\delta)\,p_i^*, p_{-i}\right)\right] =$$

$$\mathbb{E}_{p\sim\pi}\left[BW \cdot log_2\Big(1 + \frac{\frac{G_{i,i}\cdot F_{i,i}\cdot(1+\delta)p_i^*}{\sum_{i\neq j} G_{i,j}\cdot F_{i,j}\cdot p_j + v_i}}{d_{i,j}^n}\Big) - c_i\,(1+\delta)\,p_i^*\right]$$

The expected utility without the switching operation, using the utility function of section 5.2.1, can be written as follows:

$$\mathbb{E}_{p\sim\pi}\left[u_i\left(p\right)\right] = \mathbb{E}_{p\sim\pi}\left[BW \cdot log_2\Big(1 + \frac{\frac{G_{i,i}\cdot F_{i,i}\cdot p_i}{\sum_{i\neq j} G_{i,j}\cdot F_{i,j}\cdot p_j + v_i}}{d_{i,j}^n}\Big) - c_i p_i\right]$$

In [4], a bound of the probability that user i sends successfully was found. Therefore, we can rewritten this bound using the utility function of section 5.2.1, as follows:

For every π, which is an ϵ-correlated equilibrium, and for every $\delta > 0$ the the probability that user i sends successfully is at least

$$\Pr_{p\sim\pi}\left[(1-\delta)\,p_i^* \leq p \leq (1+\delta)\,p_i^*\right] \frac{BW\cdot log_2(1+\frac{\frac{G_{i,i}\cdot F_{i,i}\cdot(1+\delta)p_i^*}{\sum_{i\neq j} G_{i,j}\cdot F_{i,j}\cdot p_j+v_i}}{d_{i,j}^n})-c_i(1+\delta)p_i^*}{BW\cdot log_2(1+\frac{\frac{G_{i,i}\cdot F_{i,i}\cdot(1+\delta)p_i^*}{\sum_{i\neq j} G_{i,j}\cdot F_{i,j}\cdot p_j+v_i}}{d_{i,j}^n})-c_i(1+\delta)p_i^*} -$$

$$\frac{\epsilon}{BW\cdot log_2(1+\frac{\frac{G_{i,i}\cdot F_{i,i}\cdot(1-\delta)p_i^*}{\sum_{i\neq j} G_{i,j}\cdot F_{i,j}\cdot p_j+v_i}}{d_{i,j}^n})-c_i(1-\delta)p_i^*}$$

An other idea for the finding of an ϵ-correlated equilibrium is to take account an other switching operation. In the previous section, we found an other interval, that $p_i \in \left[\frac{\beta_i \cdot \left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)}{G_{ii}}, \frac{\left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)}{\beta_i} \log\left(\frac{1}{1-C_i}\right)\right]$.

Therefore, the swithcing operation $\phi_i\left(p_i\right)$ is equal to $\frac{\left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)}{\beta_i} \log\left(\frac{1}{1-C_i}\right)$, as well the user i could always choose this operation. The expected utility due to the above switching operation, using the utility function of section 5.2.1, can be written as follows:

$$\mathbb{E}_{p \sim \pi}\left[u_i\left(\frac{\left(\sum_{i \neq j} G_{i,j} \cdot F_{i,j} \cdot P_j + v_i\right)}{\beta_i} \log\left(\frac{1}{1-C_i}\right), p_{-i}\right)\right].$$

Now, we assume that each user i can choose a power level p_i from the set of strategies. Let that each p_i has a weight q_i [52]. Then, in each iteration t, the user i chooses that power level p_i with the largest weight q_i. The weight can be calculated as follows:

$$q_i^{t+1} = \frac{(1+\gamma)^{u_i^t(p_i)}}{\sum\limits_{p_i'} (1+\gamma)^{u_i^t(p_i')}}$$

where u_i^t is the utility function for the user i at time t of the section 5.2.1 and the parameter γ is $0 < \gamma < 1$.

Appendix A

Perron-Frobenius Theorem

Let a non-negative matrix squqre Z, which is an irreducible matrix [59]. Then:

- The matrix Z has a positive (real) eigenvalue λ_{max} such that all other eigenvalues of Z satisfy $|\lambda| \leq \lambda_{max}$.
- The eigenvalue λ_{max} has algebraic multiplicity one and has an eigenvector $x > 0$.
- Any non-negative eigenvector is a multiple of x.
- If $y > 0$ is a vector and μ is a number, where $Zy \leq \mu y$ then $\mu \geq \lambda_{max}$. For $\mu = \lambda_{max}$ iff y is a multiple of x.
- If $0 \leq S \leq Z$ and $Z \neq S$ then every eigenvalue s of S satisfies $|s| < \lambda_{max}$.
- All the diagonal components Z_{ii} of Z have eigenvalues all of which have absolute value $< \lambda_{max}$.
- If Z is primitive matrix, then all other eigenvalues of Z satisfy $|\lambda| < \lambda_{max}$.

Bibliography

[1] B. Wang, Y. Wu, and K. R. Liu, "Game Theory for Cognitive Radio Networks: An Overview," *Elsevier, Computer Networks*, pp. 2537–2561, 2010.

[2] P. Hande, S. Rangan, M. Chiang, and X. Wu, "Distributed uplink power control for optimal sir assignment in cellular data networks," *Networking, IEEE/ACM Transactions on*, vol. 16, pp. 1420–1433, Dec 2008.

[3] G. Foschini and Z. Miljanic, "A Simple Distributed Autonomous Power Control Algorithm and Its Convergence," *IEEE Transactions on Vehicular Technoly*, vol. 42, no. 4, pp. 641–646, 1993.

[4] J. Dams, M. Hoefer, and T. Kesselheim, "Convergence time of power-control dynamics," *IEEE Journal On Selected Areas In Communications*, vol. 30, no. 11, pp. 2231–2237, Dec. 2012.

[5] J. Dams, M. Hoefer, and T. Kesselheim, eds., *Scheduling in Wireles Networks with Rayleigh-Fading Interference*, Proceedings of the Twenty-fourth Annual ACM Symposium on Parallelism in Algorithms and Architectures, SPAA '12, 2012.

[6] M. Felegyhazi and J. Hubaux, "Game Theory in Wireless Networks: A Tutorial," tech. rep., 2006.

[7] A. B. MacKenzie and S. B. Wicker, eds., *Game Theory in Communications: Motivation, Explanation and Application to Power Control*, Proceedings to the IEEE Global Telecommunications Conference, 2001.

[8] E. T. N. Nisan, T. Roughgarden and V. V. Vazirani, *Algorithmic Game Theory.* Cambridge University Press, 2007.

[9] A. Nostratinia and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, Oct. 2004.

[10] T. S. Ferguson, *Game Theory.* University of California at Los Angeles, 2014.

[11] D. Monderer and L. Shapley, "Potential games," *Games and Economic Behavior 14 (1), Academic Press, Inc.*, pp. 124–143, 1996.

[12] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications 13*, pp. 1341–1348, Sep. 1995.

[13] "Pareto efficiency," *Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc.*, 2015.

[14] R. Selten, "Reexamination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, vol. 4, pp. 25–55, 1975.

[15] J.-P. Calliess, "On fixed convex combinations of no-regret learners," in *Machine Learning and Data Mining in Pattern Recognition* (P. Perner, ed.), vol. 5632 of *Lecture Notes in Computer Science*, pp. 494–504, Springer Berlin Heidelberg, 2009.

[16] A. Blum and Y. Mansour, *Learning, Regret minimization, and Equilibria.* Cambridge University Press, 2007.

[17] J. Hannan, *Contributions to the Theory of Games.* Princerton University Press, 1957.

[18] D. Foster and R. Vohra, "Asymptotic calibration," *Biometrika*, p. 379390, 1998.

[19] A. Greenwald and A. Jafir, eds., *A General Class of No-Regret Algorithms and Game-Theoretic Equilibria*, Proceedings of the 2003 Computational Learning Theory Conference, Aug. 2003.

[20] R. J. Aumann, "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, vol. 1, no. 1, pp. 67 – 96, 1974.

[21] Z. Han, C. Pandana, and K. Liu, eds., *Distributed Opportunistic Spectrum Access for Cognitive Radio using Correlated Equilibrium and No-Regret Learning*, IEEE, Wireless Communications and Networking Conference, WCNC, 2007.

[22] A. Greenwald, Z. Li, and C. Marks, "Bounds for regret matching algorithms," tech. rep., 2006.

[23] H. W. Corley, "Games with vector payoffs," *Journal of Optimization Theory and Applications*, vol. 47, no. 4, pp. 491–498, 1985.

[24] A. Blum, M. Hajiaghyi, K. Ligett, and A. Roth, eds., *Regret Minimization and the Price of Total Anarchy*, Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing, STOC'08, 2008.

[25] D. Foster and R. Vohra, "Regret in the on-line decision problem," *Games and Economic Behavior*, p. 4055, 1999.

[26] E. Tardos, "Cs 6840: Algorithmic game theory: No-regret learning." Cornell University, Lecture 11, 2012.

[27] A. R. Greenwald, *Learning to Play Network Games: Does Rationality Yield Nash Equilibrium?* New York University, 1999.

[28] A. Blum and Y. Mansour, "From external to internal regret," *Journal of Machine Learning Research*, vol. 8, pp. 1307–1324, Dec. 2007.

[29] J. Abernethy, P. Bartlett, and E. Hazan, eds., *Blackwell Approachability and No-Regret Learning are Equivalent*, Workshop and Conference Proceedings, JMLR, 2011.

[30] T. Roughgarden, "Cs 364a: Algorithmic game theory: Potential games; a hierarchy equilibria." Stanford University, Lecture 13, 2013.

[31] T. Roughgarden, "Cs 364a: Algorithmic game theory: From external regret to swap regret and the minimax theorem." Stanford University, Lecture 18, 2013.

[32] T. Roughgarden, "Cs 364a: Algorithmic game theory: Pure nash equilibria and pls-completeness." Stanford University, Lecture 19, 2013.

[33] V. Pereira and T. Sousa, "Evolution of Mobile Communications: from 1G to 4G," April 2004.

[34] "The evolution of mobile technologies 1g to 2g to 3g to 4g lte." Qualcomm, June 2014.

[35] I. F. Akyildiz, D. M. Gutierrez-Estevez, and E. C. Reyes, "The evolution to 4G cellular systemms: LTE-Advanced," *Elsevier, Physical Communication*, pp. 217–244, Dec. 2010.

[36] M. Rinne and O. Tirkkonen, "LTE, the radio technology path towards 4G," *Elsevier, Computer Communications*, pp. 1894–1906, Oct. 2010.

[37] J. Zyren and D. W. McCoy, "Overview of the 3GPP Long Term Evolution Physical Layer," *White Paper, No. 3GPPEVOLUTIONWP*, 2007.

[38] "3Gpp LTE-Advanced," *http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced*, 2013.

[39] E. Dalman, S. Parkvall, and J. Skold, "4G LTE/LTE - Advanced for mobile broadband," *Academic Press, Oxford, UK*, 2011.

[40] *LTE, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical channels and modulation.* 3 GPP Technical Specification, 3gpp ts 36.213 version 8.8.0 release 8 ed., 2011.

[41] B. Muhammad and A. Mohammed, eds., *Uplink Closed Loop Power Control for LTE System*, Emerging Technologies (ICET), 6th International Conference, Oct. 2010.

[42] *LTE, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical layer procedures.* 3 GPP Technical Specification, 3gpp ts 36.211 version 10.0.0 release 10 ed., 2009.

[43] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Pareto efficiency of pricing-based power control in wireless data networks," *Wireless Communications and Networking Conference, WCNC IEEE*, pp. 231–235, 1999.

[44] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 41, pp. 57–62, Feb 1992.

[45] V. Douros and G. Polyzos, "Review of some fundamental approaches for power control in wireless networks," *Elsevier Computer Communications*, vol. 34.

[46] T. Holliday, N. Bambos, A. Goldsmith, and P. Glynn, eds., *Distributed power control for time varying wireless networks: optimality and convergence*, Proceedings of Allerton Conference on Communication, Control and Computing, Oct. 2003.

[47] C. Saraydar, N. Mandayam, and D. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 291–303, 2002.

[48] C. Y. Huang and R. Yates, "Rate of convergence for minimum power assignment algorithms in cellular radio systems," *Wireless Networks*, vol. 4, no. 3, p. 223231, June 1998.

[49] C. Z. Han and K. Liu, "Noncooperative power-control game and throughput game over wireless networks," *IEEE, Transactions on Communications*, vol. 53.

[50] A. Attar, V. Krishnamurthy, and O. N. Gharehshiran, "Interference management using cognitive base-stations for umts lte," *IEEE Communications Magazine*, vol. 49, pp. 152–159, Aug. 2011.

[51] V. Poulkov, P. Koleva, O. Asenov, and G. Iliev, "Combined power and inter-cell interference control for lte based on role game approach," in *Telecommunications and Signal Processing (TSP), 2011 34th International Conference on*, pp. 213–217, Aug 2011.

[52] B. Latifa, Z. Gao, and S. Liu, eds., *No-Regret Learning for Simultaneous Power Control and Channel Allocation in Cognitive Radio Networks*, IEEE, Computing, Communications and Applications Conference, ComComAp, 2012.

[53] M. Bloem, T. Alpcan, and T. Basar, eds., *A Stackelberg Game for Power Control and Channel Allocation in Cognitive Radio Networks*, In Proc. Workshop on Game Theory in Communication Netwoks, GameComm, Oct. 2007.

[54] J. W. Stokes and J. A. Ritchey, eds., *A General Method For Evaluating Outage Probabilities Using Pade Approximations*, vol. 3, IEEE, Global Telecommunications Conference, GLOBECOM, The Bridge to Global Integration, Nov. 1998.

[55] G. L. Stuber, *Principles of Mobile Communication.* New York, Boston, Dordrecht, London, Moscow: Kluwer Academic Publishers, 2nd ed., 2002.

[56] V. Panwar and S. Kumar, "Bit error rate (ber) analysis of rayleigh fading channels in mobile communication," *International Journal of Modern Engineering Research (IJMER)*, vol. 2.

[57] A. S. Babu and K. S. Rao, "Evaluation of ber for awgn, rayleigh and dr. rician fading channels under various modulation schemes," *International Journal of Computer Applications*, vol. 26, no. 9, pp. 23–28, July 2011.

[58] A. S. Tanenbaum and D. J. Wetherall, *Computer Networks*. Prentice Hall, 1996.

[59] S. Sternberg, *The Perron-Frobenius theorem*. Harvard University, May 2011.