Multi-user wireless access techniques for massive Machine-Type Communications

Amani Benamor

Ph.D. Thesis

Université de Limoges

University of Limoges

ED 653 - Sciences et Ingénierie (SI)

XLIM: Systèmes et Réseaux Intelligents (SRI)

A thesis submitted to University of Limoges
in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Sciences and Technologies of Information and Communication

Presented and defensed by

Amani BENAMOR

On November 7, 2023

Multi-user wireless access techniques for massive Machine-Type Communications

Thesis supervisor:

Ms.	**Ines KAMMOUN**	Professor at ENIS, Tunisia
Mr.	**Oussama HABACHI**	Professor at UCA, France
Mr.	**Jean-Pierre CANCES**	Professor at University of Limoges, France

JURY:

President of jury

Mr.	**Mohamed JMAIEL**	Professor at ENIS, Tunisia

Reporters

Mr.	**Yezekael HAYEL**	Professor at UCA, France
Ms.	**Soumaya HAMOUDA**	Professor at FSB, Tunisia

Examiners

Ms.	**Catherine DOUILLARD**	Professor at IMT Atlantique, France

To my greatest source of inspiration,
to my father Mondher and my mother Yasmina,
I owe you the success of my studies,
and this thesis is also yours.

Le combat que tu mènes aujourd'hui t'aidera à développer la force dont tu auras besoin demain, alors n'abandonne jamais...

Abstract

The drastic expansion in the number of wireless devices and the scarcity of available spectrum resources give rise to unprecedented challenges for the future sixth generation (6G) of wireless communication systems. More precisely, the operators have to cope with the continuous prosperity of the Internet of things (IoT) along with the ever-increasing deployment of Machine Type Devices (MTDs). In this regard, due to its compelling features, the Non-Orthogonal Multiple Access (NOMA) technique has sparked off significant interest as a sophisticated technology to meet the above-mentioned challenges. Particularly, NOMA schemes are mainly categorized into two fundamentals categories, namely the Power-Domain NOMA (PD-NOMA) and the Code-Domain NOMA (CD-NOMA). In this dissertation, we consider PD-NOMA based systems and more specifically Hybrid NOMA scenarios, wherein the MTDs are divided into different groups, each of which is allocated an orthogonal Resource Block (RB), so that the members of each group share a given RB to non-orthogonally transmit their signals. Interestingly, the performance of the Hybrid NOMA system primordially depends on how users are arranged into groups and then assigned distinct power levels in order to deliver distinguishable signals to the Base Station (BS). With this in mind, we investigate the user grouping problem intertwined with the power allocation issue. Thereby, we aim to solve these combined problems through a joint optimization. To do so, we need to adopt advanced techniques with an eye toward thoroughly splitting devices into multiple groups and efficiently partitioning power among them.

First, we consider a relatively dense network and propose a game theoretical framework based on a Bi-level game in order to perform a joint channel and power selection for MTDs. Indeed, the proposed approach consists in modeling the behaviors of the devices with the aid of a non-cooperative power control game underlying a cooperative Hedonic game. More precisely, the latter is adopted to lay out the coalitional formation process while enabling the MTDs to self-arrange into multiple groups. Then, within each group, the non-cooperative game is invoked to allow the MTDs to decide autonomously the appropriate transmit power to use to transmit their packets in a

non-orthogonal manner. In doing so, we derive two low-complexity algorithms that enable us to obtain a Nash-Stable partition. Furthermore, we numerically demonstrate that the proposed Bi-level game is advantageous in terms of the energy consumption and the packet transmission success rate compared to other existing techniques in the literature.

Thereafter, we consider a very massive MTC scenario and we adopt the Mean Field Game (MFG) framework with the aim of modeling the power allocation through a simplified model. In this vein, we propose a MFG-based approach that considers the effect of the collective behavior of players. Then, we derive a distributed power control algorithm that enables the MTDs to appropriately regulate their power levels according to brief information received from the BS. Specifically, the proposed approach is governed by the two fundamental Hamilton-Jacobi-Bellman (HJB) and Fokker-PlanckKolmogorov (FPK) equations, which considerably reduce the mathematical complexity of the solution. Numerical results are presented to illustrate the equilibrium properties of the proposed power control algorithm and to emphasize the performance gains of the MFG-based approach.

Finally, we focus on designing a novel resource allocation approach based on a Bi-level learning by extending the proposed MFG using Reinforcement Learning (RL) techniques. Particularly, we derive decision-making algorithms in which the RL approach is on the top of the MFG framework in order to jointly optimize resource allocation and power control problems. In this way, the combined approaches allow the devices to autonomously join the appropriate groups and then adjust their transmit power while taking into consideration the effect of the mass behavior of the players within each group. The simulation results illustrate the convergence of the proposed learning process and underline its robustness against interference impacts.

Résumé

L'expansion radicale du nombre de dispositifs sans fil et la rareté des ressources spectrales disponibles posent des défis inédits pour la future sixième génération (6G) de systèmes de communication sans fil. En effet, les opérateurs doivent faire face à la prospérité continue du Internet of Things (IoT) ainsi qu'au déploiement toujours plus important de Machine Type Devices (MTDs). À cet égard, en raison de ses caractéristiques convaincantes, le technique Non-Orthogonal Multiple Access (NOMA) a suscité un intérêt considérable en tant que technologie sophistiquée permettant de relever les défis susmentionnés. Plus précisément, les techniques NOMA se répartissent en deux catégories fondamentales, à savoir le Power-Domain NOMA (PD-NOMA) and le Code-Domain NOMA (CD-NOMA). Dans cette thèse, nous considérons les systèmes basés sur le PD-NOMA et plus particulièrement les scénarios de Hybrid NOMA, dans lequel les MTDs sont divisés en différents groupes tout en leur attribuant des Resource Block (RB) orthogonaux, de sorte que les membres de chaque groupe partagent un RB donné pour transmettre leurs signaux de manière non orthogonale. Il est intéressant de noter que les performances du système Hybrid NOMA dépendent essentiellement de la manière dont les utilisateurs sont organisés en groupes et dont les niveaux de puissance leur sont attribués afin de transmettre des signaux distincts à la Base Station (BS). Dans cette optique, nous étudions le problème du regroupement des utilisateurs lié à celui de l'allocation de puissance. Nous cherchons ainsi à résoudre ces problèmes combinés par une optimisation conjointe. Pour ce faire, nous devons adopter des techniques avancées afin de diviser minutieusement les MTDs en plusieurs groupes et leur attribuer efficacement des niveaux de puissance.

Tout d'abord, nous considérons un réseau relativement dense et nous proposons un jeu à deux niveaux en utilisant la théorie des jeux afin de résoudre conjointement les problèmes de sélection de canaux et d'allocation de puissance entre les MTDs. Plus précisément, l'approche proposée consiste à modéliser les comportements des dispositifs à l'aide d'un jeu non coopératif sous-jacent à un jeu Hedonic coopératif. Premièrement, le jeu Hedonic est adopté pour mettre en place le processus de formation de coalitions tout en permettant aux MTDs de s'arranger en plusieurs groupes.

Ensuite, au sein de chaque groupe, le jeu non coopératif est utilisé pour permettre à chaque MTD de décider de manière autonome du niveau de puissance le plus approprié à utiliser pour transmettre ses paquets. Ce faisant, nous proposons deux algorithmes qui nous permettent d'obtenir une partition Nash-Stable des coalitions. En outre, nous démontrons numériquement que le jeu à deux niveaux proposé est avantageux en termes de consommation d'énergie et de taux de réussite de la transmission des paquets par rapport aux autres techniques existantes dans la littérature.

Par la suite, nous nous intéressons à un scénario de MTC très massif. Dans le but de modéliser l'allocation de puissance à travers un modèle simplifié, nous adoptons le Mean Field Game (MFG) qui consiste à étudier l'effet du comportement collectif des joueurs. En outre, nous développons un algorithme distribué pour effectuer le contrôle de la puissance afin que les MTDs soient en mesure de réguler de manière appropriée leurs niveaux de puissance en fonction des informations limitées reçues de la BS. En effet, l'approche proposée permet de réduire considérablement la complexité mathématique de l'analyse du jeu qui se base principalement sur deux équations fondamentales de Hamilton-Jacobi-Bellman (HJB) et de Fokker-Planck-Kolmogorov (FPK). Des résultats numériques sont présentés en considérant une population dense, afin de démontrer les propriétés d'équilibre de l'algorithme développé et de mettre en évidence les gains de performance de l'approche proposée basée sur le MFG.

Enfin, nous nous concentrons sur l'application des outils de Reinforcement Learning (RL) à l'approche MFG proposée de sorte que nous aboutissons à une nouvelle technique d'allocation des ressources basée sur un apprentissage à deux niveaux. Précisément, nous dérivons des algorithmes qui permettent aux MTDs de prendre leurs décisions en invoquant premièrement la technique de RL pour résoudre le problème d'allocation des ressources, puis l'approche MFG pour optimiser le contrôle de la puissance. De cette manière, les MTD peuvent choisir de manière autonome les groupes les plus appropriés auxquels il est préférable d'appartenir. Une fois que chaque utilisateur a rejoint le groupe qu'il a choisi, il peut ajuster son niveau de puissance en tenant compte de l'effet du comportement de masse au sein de son groupe. Finalement, les résultats de la simulation montrent la convergence du processus d'apprentissage proposé et soulignent sa robustesse face aux impacts des interférences.

Preface

The work in this dissertation has been summarized and presented in the following conferences and journals:

1. A. Benamor, O. Habachi, I. Kammoun and J.P. Cances, "Game Theoretical Framework for Joint Channel Selection and Power Control in Hybrid NOMA," in *IEEE International Conference on Communications (ICC), 2020. Rank: B.*

2. A. Benamor, O. Habachi, I. Kammoun and J.P. Cances, "Mean Field Game-Theoretic Framework for distributed Power Control in Hybrid NOMA", in *IEEE Transactions on Wireless Communications, 2022. Quartile: Q1, Impact Factor: 8.346.*

3. A. Benamor, O. Habachi, I. Kammoun and J.P. Cances, "NOMA-based Power Control for Machine-Type Communications: A Mean Field Game Approach", in *IEEE International Performance, Computing, and Communications Conference (IPCCC), 2022. Rank: B.*

4. A. Benamor, O. Habachi, I. Kammoun and J.P. Cances, "Multi-Armed Bandit Framework for Resource Allocation in Uplink NOMA Networks", accepted in *IEEE Wireless Communications and Networking Conference (WCNC), 2023. Rank: B.*

5. A. Benamor, O. Habachi, I. Kammoun and J.P. Cances, "Reinforcement learning for Mean Field Game-based Resource Allocation in NOMA Networks: a Multi-Armed Bandit Approach", submitted to *IEEE Internet of Things Journal. Quartile: Q1, Impact Factor: 9.471.*

Acknowledgments

First and foremost, it is by the grace and bounty of the Most Gracious and the Most Merciful, Allah, that this thesis was undertaken and completed. I am grateful to Him, to whom all praise and thanks are due for giving me the strive, the strength and the perseverance to make my dream come true.

I am deeply grateful to my supervisor Ms. Inès Kammoun-Jemal for giving me the chance to carry out this thesis and affording me the opportunities and the good conditions to succeed. It is the witness of my deep appreciation, for her kindness, precious support and encouragement. Even during the most challenging moments throughout my PhD, she believed in me and gave me insightful advice.

I am also extremely grateful to my supervisor Mr. Jean Pierre Cances for having faith in me and for his tremendous support. I also thank him for seeing a potential in me and for his continual encouragement and notably guidance. His enthusiasm and positivism made working on this thesis a wonderful experience.

I wish to express my deepest thanks and sincere appreciation to my supervisor Mr. Oussama Habachi, whose professional suggestions and fruitful ideas helped me to better carry on my research work. Special thanks and gratitude go to him, without his valuable support, precious guidance, and constructive comments, this thesis would not have been possible.

I would express my warmest thanks to the members of the jury. I thank Mr. Mohamed Jmaiel, professor at National School of Engineers of Sfax, Tunisia, for accepting to chair my jury. I would like to thank Ms. Soumaya Hamouda, professor at Faculty of Sciences of Bizerte, Tunisia and Mr. Yezekael Hayel professor at University of Avignon, France for their interest in my thesis and for accepting to be the scientific reviewers of this work. I also thank Ms. Catherine DOUILLARD, Professor at IMT Atlantique, France, for accepting being my thesis examiner.

It is extremely pleasant to thank Mr. Vahid Meghdadi for his joyful and enthusiastic mood and the whole RUBIH team, with whom I spent those years and shared wonderful things.

No words can express my warmest gratitude to all my beloved family members: my parents, my grandmother, my brothers, my uncles and my aunties, for their unconditional love, invaluable support and sincere prayers.

I am deeply indebted to my parents for their sacrifice along the path of my academic pursuits and their extraordinary efforts they have exerted to enable me reach this important life milestone. Their endless belief in me and their immeasurable support make my PhD journey possible. I owe them the success of my studies, and this thesis is also theirs.

A heartfelt thank goes to my dearest fiance Nabil, who never stopped standing by my side when time gets hard and always being there when there is no one.

Last but not least, special thanks to my friends, for being part of my life and my side whenever I need them.

Table of contents

List of Figures

Acronyms

3GPP	**T**hird **G**eneration **P**artnership **P**roject
1G	**F**irst **G**eneration
2G	**S**econd **G**eneration
4G	**F**ourth **G**eneration
5G	**F**ifth **G**eneration
6G	**S**ixth **G**eneration
BS	**B**ase **S**tation
CDMA	**C**ode **D**ivision **M**ultiple **A**ccess
CD-NOMA	**C**ode-**D**omain **N**on-**O**rthogonal **M**ultiple **A**ccess
FDMA	**F**requency **D**ivision **M**ultiple **A**ccess
IoT	**I**nternet **o**f **T**hings
LTE	**L**ong **T**erm **E**volution
MAB	**M**ulti-**A**rmed **B**andits
MFG	**M**ean **F**ield **G**ame
MFI	**M**ean **F**ield **I**nterference
mMTC	massive **M**achine **T**ype **C**ommunications
MTC	**M**achine **T**ype **C**ommunications
MTD	**M**achine **T**ype **D**evice communications
MUST	**M**ulti-**U**ser **S**uperposition **T**ransmission
NOMA	**N**on-**O**rthogonal **M**ultiple **A**ccess
OFDM	**O**rthogonal **D**ivision **M**ultiple
OFDMA	**O**rthogonal **D**ivision **M**ultiple **A**ccess
OMA	**O**rthogonal **M**ultiple **A**ccess
PD-NOMA	**P**ower-**D**omain **N**on-**O**rthogonal **M**ultiple **A**ccess
RL	**R**einforcement **L**earning
SC	**S**uperposition **C**oding
SIC	**S**uccessive **I**nterference **C**ancellation
SINR	**S**ignal-to-**I**Interference-plus-**N**oise **R**atio
SNR	**S**ignal-to-**N**oise **R**atio
TDMA	**T**ime **D**ivision **M**ultiple **A**ccess
UCB	**U**pper **C**onfidence **B**ounds

Part I

Introduction and background

Chapter 1

Introduction

Contents

1.1 Overview and Motivation

The upcoming era of wireless networks is evolving in a completely revolutionary way, thanks to the Internet of Things (IoT), which represents a paramount technological trend that transforms everyday objects into information sources and allows them to be connected to the Internet. Particularly, driven by the pervasiveness and inclusiveness of IoT services, our world has witnessed a steady escalation of demands for a massive wireless access and an evolutionary myriad of applications [1]. Meanwhile, Machine Type communications (MTC), also known as Machine-to-Machine communications (M2M) represents as an attractive dominant scenario in the Fifth Generation (5G) of wireless communications that can support the ubiquitous of IoT systems and enable it to be part of future communication networks [2,3]. Indeed, a recent forecast report from Cisco [4] predicts that coming future systems will witness a tremendous growth in the worldwide network of interconnected devices, reaching up 29.3 billion by 2023 compared to 18.4 billion connected devices in 2018, wherein MTC will represent half of the world's Internet connections, about 14.7 billion by 2023, compared to 6.1 billion connected devices in 2018. This rapid growth of number of devices will gradually overwhelm the connectivity capability of 5G networks. Recently, the future Sixth Generation (6G) networks have become a focal point of interest for both the academia and industry since it is foreseen to provide a connection density 10 times higher than that in 5G [5].

Generally, MTC refer to a type of data communication that does not require a specific human intervention and can autonomously occur among machines such as devices, sensors, equipment, drones. These connections between machines are fundamentally different from the traditional Human-Type Communications (HTC) that focus on achieving high data rates for downlink transmissions of large-sized packets usually. As pointed out by the Third Generation Partnership Project (3GPP) [6], MTC networks are primarily dominated by uplink transmissions of short packets at low data rates and characterized by establishment of connections among a massive number of connected devices. Meanwhile, MTC often occur through a sporadic communication scenario which means that devices have sparse activities, so only a small portion of users can transmit simultaneously at the same time [7]. Thus, MTC lead to a transition from downlink-dominated traffic in HTC to predominantly uplink traffic and change from a transmission of large packets with high bit rate requirements to a completely

different task based on a sporadic transmission of small packets with low data rates. This clearly results in a shift in the way MTC can be studied, which needs to be envisaged from other perspectives than those taken to investigate HTC. On the other hand, even if each device can generate a bursty and sparse traffic, the one generated by a massive MTC (mMTC) scenario is rather challenging. Hence, to deal with this proliferation of wireless machines and manage the resulting access loads, innovative and robust access technologies are required.

Along the aforementioned characteristics, the design of appropriate Multiple Access (MA) techniques is one of the most significant keys to support mMTC networks. Basically, MA techniques are categorized into two main classes, namely Orthogonal Multiple Access (OMA) and Non-Orthogonal Multiple Access (NOMA). In particular, OMA pertains to an orthogonal allocation in which different resources are assigned to distinct users. By contrast, NOMA relates to a non-orthogonal resource assignment which means that several devices are able to access the same resources to transmit their signals. On the other hand, OMA encompasses various schemes, which have evolved with consecutive generations of cellular communication systems. Starting with Frequency Division Multiple Access (FDMA) in the First Generation (1G), then Time Division Multiple Access (TDMA) in the Second Generation (2G). Subsequently, Code Division Multiple Access (CDMA) has been emerged in the Third Generation (3G), and then Orthogonal Frequency Division Multiple Access (OFDMA) has been invoked in the Forth Generation(4G) networks. These user multiplexing schemes have been designed while considering the orthogonality allocation objective, which means that each resource (frequency channel or time slot or signature code or resource block) is exclusively allocated to at most one device at a time.

Worryingly, these schemes are not able to scale to comply with the unprecedented demands for the massive access and the network traffic. Indeed, serving users in an orthogonal manner is overwhelmed by the scarce radio resources. On the other hand, driven by the sporadic nature of MTC, scheduling an entire resource for each connection is both inefficient and unfeasible. In an effort to overcome these shortcomings, academia, industrial communities and even standardization bodies heightened their interest in conducting research studies toward NOMA techniques [8–10]. Particularly, 3GPP has included NOMA in its release-13 (Rel-13), labeled as Multi-user Superposition Transmission (MUST) [11].

Interestingly, NOMA has been more an eye-catching technique to handle connectivity issues and improve the spectral efficiency for 5G mMTC systems. Besides, with the proliferation of the IoT, NOMA is considered as a promising scalable research candidate to support the ongoing development of mMTC toward future wireless networks. Specifically, NOMA techniques fall into two fundamentals categories. The first one is called the Power-Domain NOMA (PD-NOMA) and refers to the case where the superimposed signals of different users are distinguished via distinct power allocation coefficients [12–14]. Alternatively, when users' signals overlap while they are assigned different codes, the Code-Domain NOMA (CD-NOMA) is the considered scheme [15–17].

1.2 Problems Statement and Objectives

Notwithstanding that NOMA is a distinguished multiplexing scheme due to its appealing features to support multiple transmissions, one cannot turn a blind eye to its some shortcomings. In fact, allowing manifold users to access shared resources simultaneously is beneficial but comes at the expense of a co-channel interference that linearly increases with the number of users performing NOMA jointly. This results in an increased computational complexity of the interference canceller at the receiver side, which in turn seriously affects the overall performance of NOMA systems. Thereby, considering a large number of users communicating with the Base Station (BS) through a common Resource Block (RB) is highly challenging and even unrealistic [18, 19]. Thus, a Hybrid NOMA network represents an alternative approach, in which users are divided into multiple groups, each of which will be allocated a given RB. Obviously, establishing hybrid NOMA networks relies heavily on the user grouping strategy, which seeks to strike a meaningful trade-off between NOMA gains and interference effects.

In this vein, abundant research contributions have been conducted to investigate user pairing schemes in which two users are allowed to form NOMA groups and simultaneously access a particular RB. In doing so, it has been shown that interference effects are considerably mitigated and thus the interference canceller is implemented at an affordable complexity cost. Nevertheless, due to the increased demand for massive connectivity and the scarcity of available radio resources, assigning an orthogonal RB

to only two users is becoming a bottleneck for the NOMA performance. Thereby, it is meaningful to focus on how to fully leverage the benefits of NOMA by clustering more than two users in a given resource, but at an acceptable computational complexity. Therefore, we study the grouping formation process while considering the trade-off between the decoding complexity of the canceller and the NOMA group size.

On the other hand, the performance of a PD-NOMA scheme is particularly based on the way the power is partitioned among the users. More precisely, the BS can successfully decode and recover the interfering signals from different transmitters by exploiting the disparity in power levels among them. Consequently, an improper power assignment yields an important interference impact, which impairs the effectiveness of the interference canceller at the BS. Besides, a particular user with a limited energy budget may require much more power than it can handle in order to cope with the resulting interference effects. Therefore, this results in a high energy consumption and thus lower down the overall system performance. As a result, it is of the utmost significance to thoroughly focus on studying the power control problem to suitably deal with the inter-user interference and thus further boost the NOMA network gain.

Thesis objectives

In this thesis, we consider a PD-NOMA scheme in the context of a Hybrid NOMA network in which multiple devices share the available resources. We investigate power allocation policies to deal with critical power capacities and limited battery lifetimes of the devices and thus satisfy energy-efficiency requirements. Furthermore, we are mainly interested in finding efficient user selection strategies to form NOMA groups while mitigating the inter-user interference as much as possible.

In this regard, meticulously arranging users into several NOMA groups is a crucial ingredient of our thesis. Indeed, we aim to practically avail the advantages of NOMA in its true sense by considering more than two users in each group, with an eye toward implementing the interference canceller at an affordable cost. In this direction, we focus on investigating how to divide users into multiple groups while orthogonal RBs are assigned to these groups. The members of each group are served by a common RB so that they can simultaneously transmit their signals to the BS. In doing so, each user in a given group encounters only interference effects from the users belonging to

that group and thus the inter-group interference is mitigated. Consequently, at the receiver, the complexity of the decoder implementation is remarkably reduced.

Interestingly, we more precisely aim to solve the dilemma between the demand for the massive MTDs access and the scarcity of available resources by addressing the user grouping issue intertwined with the power control problem. In this way, dividing users into groups will be governed by the group formation process, while the power allocation strategy controls the way that the signals of users in each group are superimposed. Therewith, these combined problems need to be solved through a joint optimization in order to efficiently arranging users into several groups and then smartly designing power allocation policies among them. For this sake, sophisticated tools are required to cope with this burdensome task. Throughout this thesis, we invoke Game theory as well as Reinforcement Learning (RL) to jointly optimize the resource allocation and power control problems.

1.3 Road map

This dissertation encompasses two parts that are composed of 6 chapters. In the first part, we present some of the necessary theoretical foundations to pave the way for a better understanding of this thesis. In this vein, chapter 2 reviews the basics of game theory and its fundamental classes ranging from non-cooperative games to cooperative games and the advanced Mean Field Game (MFG) theoretic framework. In particular, we emphasize on the main notions, key components, and appropriate solution concepts that pertain to each class of game. In Chapter 3, we first provide an overview of the OMA techniques, and then we delve into the fundamentals of the NOMA concept by outlining its different categories. Specifically, we take a more in-depth look at the most important existing related works and the essential strides that have been performed in this research direction. Finally, we spotlight our major contributions in view of the detailed research gaps while addressing the resource management problem in NOMA systems.

In the second part, we detail the approaches proposed throughout this dissertation. Indeed, in chapter 4, we formulate the joint resource allocation and power control optimization problem with the aid of a game theoretic framework. More precisely, the behaviors of the devices are modeled by a Bi-level game consisting of a non-

cooperative power control game and a cooperative coalitional game. In chapter 5, we focus on investigating densely deployed networks which, in turn require a more sophisticated game in order to be appropriately addressed. Thereby, we turn our attention to the MFG in which each particular user makes its move by conforming only to the mass behavior of its opponents instead of caring about the individual action of each of them. Thus, we propose a distributed iterative algorithm based on the MFG approach that takes into account the interference effects in order to enable the devices to appropriately determine their power levels . In chapter 6, we extend the MFG formulated in chapter 5 by invoking a RL tool, namely the multi-armed Bandit (MAB) approach with the aim of proposing a resource allocation approach based on a bi-level learning. Particularly, we derive distributed MFG underlying MAB algorithms with the aim of making devices autonomous in determining their most appropriate power levels and groups. Finally, chapter 7 concludes this thesis by recapitulating the fundamental contributions of our work and presenting some research directions that can be worth conducting in the future.

Chapter 2

Fundamentals of game theory

Contents

2.1 Introduction

Game theory has emerged as an analytical framework that provides an increasingly appealing mathematical tool for adequately characterizing cooperative and competitive behaviors of multiple players. Essentially, game theory has been adopted to address various challenges in social sciences, namely in economics. Then, it has been extended to develop a robust mechanism in variety of disciplines such as biology, political science, philosophy and recently in computer science and wireless communication networks. Basically, the players of the game are rational in nature, which means that each of them aims to maximize the game's outcome from its own perspective. Thus, each individual makes its decision and takes its strategy based on its self-interests.

Attracted by its distinguishing features, numerous research contributions have been devoted to apply game theory to solve critical optimization problems in wireless communication systems. Particularly, game theory provides interesting solution concepts to efficiently deal with the selfish nature of wireless users and derive distributed algorithms that enable the users to reach their desired performance objectives.

In this chapter, we introduce the game formulation and illustrate the most relevant characteristics of game theory in order to pave the way for designing a distributed control among interacting players.

2.2 General game formulation

The most common representation of a game is defined by a normal or strategic form that basically consists of three elements: the set of players, the set of actions or strategies, and the utility function, given as:

- **Set of players** $\mathcal{N} = \{1 \dots N\}$.
- **Set of actions** $\{\mathcal{A}_i\}_{i\in\mathcal{N}}$: $a_i \in \mathcal{A}_i$ is the strategy of player i
- **Utility function** $\{U_i\}_{i\in\mathcal{N}}$: called also a payoff function, it reflects the desired goal of playing the game. The choice of the utility function is a pivotal focus of game theory.

For a strategic game, $a = (a_i, a_{-i})$ denotes a strategy profile where $a_{-i} = [a_j]_{j\in\mathcal{N}, j\neq i}$ consists of the strategies of all players excluding the i-th player. Upon playing a game,

each player seeks to find the optimal strategy to optimize its utility function.

It should be emphasized that the above-mentioned components represent the basis of the general game in a strategic form. Besides, each type of game has its own specific components, key properties, solution concepts and substantial goals. Particularly, the interaction between players is governed by their game objectives. In other words, with the aim of achieving the game's outcome, they can either cooperate or have a conflict of interest that leads them to compete against each other. Hence, in this thesis, we distinguish two types of games: cooperative games and non-cooperative games.

Roughly speaking, non-cooperative games characterize competitive scenarios, wherein the players are involved with conflicting interests over the outcome of the decision-making process. In contrast, in cooperative games, the players are committed to form an agreement with each other in order to mutually benefit from the cooperation. Although these types of games allow each player to make a decision only once time, in some other situations the notion of time is considered as a key ingredient of the game in order to deal with the time-varying environment. In this regard, a differential game can be applied to enable the players to dynamically make their decisions evolving with time. In the presence of a large population, the differential game can be extended into a MFG by considering the effect of the collective behavior of devices, which interestingly alleviates the mathematical complexity of the game analysis. In what follows, we will dive into each type of game starting with cooperative games.

2.3 Cooperative game

The cooperative game theory provides several analytical tools to suitably investigate the cooperation and the negotiation among players attempting to maximize the mutual benefit from achieving an agreement. This class of games includes two main branches: bargaining theory and coalitional games. When players seek to reach a desired agreement, but are in conflict about what the agreement will be and how to reach it, we are dealing with a bargaining situation. Meanwhile, a cooperative game can be considered as a coalitional game when players focus on how to form cooperative groups of players, each of which called a coalition. Particularly, the coalitional game theory has been widely adopted as an efficient and a robust tool to design a cooperative decision-making process in communication networks. Indeed, this branch

of cooperative games incorporates various algorithms and concepts that have been grouped into two main categories: coalition-formation games and canonical games. In the following section, we look in deep into the basic components, main features and key solution concepts of the coalitional games class.

2.3.1 Coalitional game: preliminaries

2.3.1.1 Basic notions

Definition 2.1. A coalitional game is defined as a pair $(\mathcal{N}, v)$, where $\mathcal{N}$ represents the set of players and v denotes the value function that qualifies the gain of each player upon cooperating and being part of a given coalition.

Definition 2.2. A set $S = \{S_1 \dots S_l\}$ is said to be a collection of coalitions when it consists of a given number of subsets of $\mathcal{N}$ that covers a set of players. If the set S covers the whole set of players $\mathcal{N}$, it is simply a partition of $\mathcal{N}$.

Depending on the definition of the value function, the game takes two different types and forms given as follows:

Definition 2.3. A coalitional game is identified to be a transferable utility (TU) game if the gain received by each coalition is a real number that can be apportioned among the members of that coalition.

Definition 2.4. A coalitional game is identified to be a non-transferable utility (NTU) game if the worth of each coalition is a payoff vector, where each element denotes the payoff that each member of this coalition can receive by being part of. Thus, the payoff of each player depends on both the strategy it has chosen and the actions taken jointly by the other members of the coalition.

Furthermore, coalitional games can be either in a characteristic form or a partition form based on how the value function associated with each coalition is quantified. Indeed, when the outcome of a given coalition is determined only according to its members, we are dealing with a coalitional game in a characteristic form. If the way that the other coalitions are structured and the remaining players are arranged has an impact on the value of a coalition, the game is in the partition form. In what follows, we explore in detail the key features of the two classes of the coalitional

games: canonical games and coalition-formation games.

2.3.2 Canonical games

In the canonical games, players have the incentive to cooperate and form a grand coalition that encompasses all the players in order to receive better benefits than they would get if they had acted alone or in smaller coalitions. This characteristic refers to the mathematical property of superadditivity. In addition, the canonical games focus on the optimal way to ensure the stability of the grand coalition and guarantee the fairness of the utility distribution among the players.

Canonical coalitional games turn out to be an interesting framework to address several problems in wireless and communication networks. Indeed, the application of canonical coalitional games covers a wide range of scenarios that aim to ensure cooperative transmissions and a fair payoff distribution between players. For instance, in [20] the authors have studied the fairness of the rate allocation among the set of users, whereas in [21], the authors have investigated the possibilities of arranging the receivers and the transmitters into the grand coalitions. The work in [22] has proposed an approach in which the canonical game is involved to model cooperative transmissions between nodes in order to reach an agreement on the packet forwarding over the networks. Generally, whenever the goal of players is to maximize their utilities by constructing the larger coalition and forming an agreement on a fairness criteria of how the game outcome will be distributed, canonical coalitional games are a powerful tool to handle such problems.

2.3.3 Coalition-formation games

While canonical games are primarily interested in forming a stable and large coalition, coalition-formation games study how to establish an optimal coalitional structure of the set of players. Indeed, such games investigate scenarios where players are willing to cooperate and form an appropriate structure in the aim of maximizing their utilities collectively. In this regard, the main focus of coalition-formation games is how should players be arranged into coalitions in order to achieve a joint benefit. To this end, unlike the canonical games that implicitly assume that forming coalitions is always profitable, the formation process in coalition-formation games mainly relies on a negotiation, information exchange and a set of rules, particularly in the context

of communications systems.

Interestingly, coalition-formation games are widely exploited as an effective tool to adequately characterize the formation of coalitions in wireless and communication networks. For instance, in [23], a coalition formation game has been adopted to the unmanned aerial vehicle (UAV) networks in order to analyze the task assignment and derive a distributed cooperative algorithm while the energy consumption is a major concern. The authors in [24] have investigated the coalition formation to enable the nodes to organize themselves into multiple coalitions and autonomously collaborate by exchanging information between them. The model in [25] has focused on the coalition formation process in order to propose a distributed resource allocation algorithm and handle the mutual interference between devices in device-to-device (D2D) communications. Briefly, whenever an autonomous and a distributed cooperative model is required, the coalition-formation game is a natural candidate to enable the players to cooperate and self-organize into an appropriate coalitional structure.

Generally, a coalition formation game explores how the coalitional structure can be formed to fit the proposed model and how players autonomously decide to join a coalition. Nevertheless, addressing these questions for such a game is not straightforward, especially when the formation process is achieved in a distributed manner. According to [26], there are generic rules and prominent solution concepts that can be exploited to allow the players to arrange themselves into coalitions in order to construct an optimal structure. In the following, we go through some coalition-formation algorithms, namely Hedonic games and merge and split approaches.

2.3.3.1 Preference relation

Upon playing a coalition formation game, the players have to decide autonomously whether to leave or join coalitions with an effort to reach an optimal coalitional structure. Thus, the decisions made by the players reflect their preferences on their set of potential coalitions that they can be part of. In other words, each player should be able to compare the coalitions and choose which one it prefers to be a member of, using a preference relation defined as:

Definition 2.5. A preference relation, also called a preference order, $\triangleright$ serves to compare any two collections (or partitions) $A = \{A_1 \dots A_s\}$ and $B = \{B_1 \dots B_l\}$ of

the same set of players such that $\bigcup_{i=1}^{s} A_i = \bigcup_{j=1}^{l} B_j = H$. Thus, $A \triangleright B$ means that players in a subset H prefer to arrange themselves into A instead of B if the total coalition values achieved in A is strictly higher than that in B, i.e. $\sum_{i=1}^{s} v(A_i) > \sum_{i=1}^{l} v(B_i)$.

Generally, there are two categories of preference relations, namely coalition-value orders and individual value orders. Indeed, coalition-value orders perform the comparison between two collections by comparing the outcome of the coalitions belonging to these collections. One example of a coalition-value order is a utilitarian order, which states that it is preferred to partition a subset of players H by a collection A than by B, if the total payoff obtained by the players in A is better than that in B. Meanwhile, an individual value order is used to compare the utilities obtained individually by the players. In this regard, Pareto order is one of the most relevant examples of individual-value orders. Indeed, given two payoff allocation vectors a and b of two collections A and B, respectively, using Pareto order, A is preferred over B if and only if $a \geq b$, with at least one element a_i of a is strictly greater than one element b_i of b, such that $a_i > b_i$. Interestingly, the utilitarian order constitutes an important ingredient of TU games, whereas Pareto order is considered as an appropriate preference order for both TU and NTU games.

It is worth pointing out that the concept of preference orders can provide useful insights to any player $i \in \mathcal{N}$ to compare coalitions that it can be a member of. In this sense, let us now define a preference relation for any player $i \in \mathcal{N}$ as follows:

Definition 2.6. Consider a complete, reflexive, and transitive binary relation over the set of possible coalitions to which a player $i \in \mathcal{N}$ can belong, denoted by $\succcurlyeq_i$, i.e. over the set $\mathcal{N}_i = \{S_i \subseteq \mathcal{N} : i \in S_i\}$. Consequently, given two coalitions $S_1 \subseteq \mathcal{N}$ and $S_2 \subseteq \mathcal{N}$ such that $i \in S_1$ and $i \in S_2$, $S_1 \succcurlyeq_i S_2$ indicates that the user i prefers to be part of the coalition S_1 over being part of the coalition S_2, or at least i prefers both coalitions equally. Furthermore, $S_1 \succ_i S_2$ indicates that the player i strictly prefers being a member of S_1 over being a member of S_2.

2.3.3.2 Hedonic games

A Hedonic coalition formation game is one of the important classes of coalitional games. This type of game entails important properties and interesting solution con-

cepts that pave the way for modeling the formation of coalitions. A coalitional game is considered as a Hedonic game when it satisfies the following two requirements [27]:

- The gain of any player i resulting from playing the game is only impacted by the identity of the members of the coalition to which the player belongs.
- The coalition formation process takes place according to players' preferences among the set of all coalitions that they can possibly be a part of.

Definition 2.7. the Hedonic game is a pair $(\mathcal{N}, \succ)$ that meets the two conditions mentioned above and consists of the set of players $\mathcal{N}$ and a profile of preferences $\succ$ defined for all the players, i.e.$(\succcurlyeq_1 \cdots \succcurlyeq_N)$.

Basically, each player chooses to deviate from its current coalition and join another coalition if the latter allows the player to obtain a higher utility. In this regard, the player applies a switch rule to be able to make an autonomous decision. The switch rule can be expressed as:

Definition 2.8. Given a partition $\Pi = \{S_1 \dots S_M\}$ of the set of players $\mathcal{N}$, a player $i \in S_k$ decides to leave its current coalition S_k and join another coalition S_l where $k \neq l$, if and only if $S_l \cup \{i\} \succcurlyeq_i S_k$, resulting in a new partition formation $\Pi' = \{\Pi \setminus \{S_k, S_l\}\} \cup \{S_k \setminus \{i\}, S_l \cup \{i\}\}$.

According to the switch rule, each player acts unilaterally with an effort to maximize its utility. Thus, it selfishly makes the decision without considering its impact on the other players. Furthermore, when any player makes a single switch, the partition Π is clearly transformed into another partition Π'. Consequently, the coalitional formation in the Hedonic game is mainly based on successive switch operations that may lead to a stable final partition after a finite number of iterations. Thereby, the stability notion is viewed as one of the most pivotal concepts for the Hedonic game. Indeed, according to [27] four forms of the stability can be identified: the Nash stability, individual stability, contractual individual stability and core stability. The latter requires the immunity to coalition deviation, where in the other ones, a partition is stable if it is immune to an individual deviation. Specifically, a partition is Nash stable if no player can improve its utility by moving unilaterally to another existing coalition. A partition is individually stable if no player can improve its utility by moving to another existing

coalition S (by creating a new coalition eventually) without making the members of S worse off. Similarly, a partition is said to be contractually individually stable if no player can improve its utility by moving from a coalition S to another existing coalition $S^{'}$ (by creating a new coalition eventually) without making the members of S nor the members of $S^{'}$ worse off. Finally, a coalition $S \subseteq \mathcal{N}$ is said to be a block of the partition Π, if any player of the coalition S strictly prefers S to its current coalition in Π. If there is no blocking coalition, the partition Π is said to be core stable.

2.3.3.3 Merge and split algorithm

The merge and split algorithm is one of the most important frameworks for the coalition formation. It is composed essentially of two rules that are useful for forming and breaking coalitions, known as the merge and split rules, respectively [28]. These rules serve to model the coalition formation procedure since they govern the partition transformation at each merge and split iteration, defined as

Definition 2.9. (Merge rule) Given a set of players $\mathcal{N}$ arranged into disjoint coalitions $\{S_1 \dots S_K\}$, this set may decide to merge into a one single coalition $H = \bigcup_{i=1}^{K} S_i$, if the latter is preferred by the players.

Definition 2.10. (Split rule) A coalition $G = \bigcup_{i=1}^{L} S_i$ may decide to split and form smaller coalitions $\{S_1 \dots S_L\}$, if the resulting form is preferred by the members of G.

It is noteworthy that the decision to split or merge depends on the selected preference orders. For instance, if the Pareto order is invoked, coalitions agree to merge if there is at least one player that can improve its payoff without hurting the payoffs of remaining players. On the other hand, the split rule is made if no player in the coalition can be worse off by receiving a lower payoff in the new partition. Consequently, making the decision either to merge or split is equivalent to establishing an agreement among the players that allows at least one player to reach a higher payoff, in the newly formed partition, while maintaining the utilities of the other players.

2.4 Non-cooperative game theory

A non-cooperative game is the most investigated branch in game theory. It characterizes the competitive behavior of players having adversarial or conflicting interests over their decision-making process. Indeed, each player selfishly attempts to maximize its benefits from the game whatever will be the impact of its move on the other players. In the context of wireless and communications networks, non-cooperative games have been applied in several scenarios, such as, the resource allocation, the power control and the interference management. Here, we focus primarily on situations that pertain to power control issues.

Typically, the non-cooperative game in the strategic form can be analyzed by firstly identifying the players, their set of strategies and their utilities. Specifically, a two-player non-cooperative game can be illustrated in a matrix format, wherein each row corresponds to a strategy of the first player while each column represents a strategy of the second player, as shown in Table 2.1. Thus, the number of rows as well as the number of columns are equal to that of player' s strategies. In addition, all of the matrix represents the game outcome, i.e. the utility values of the two players when they choose these strategies. In what follows, we detail some situations where non-cooperative games can be exploited.

2.4.1 Nash equilibrium

The Nash equilibrium is a pivotal concept and one of the most important results in game theory. It has been introduced by J. Nash in [29] and represents a situation where no player has an incentive to deviate unilaterally. Indeed, even if every player aims to maximize its own utility selfishly, the game can settle at a point in which the players agree. Thus, we introduce the concept of the Nash equilibrium as:

Definition 2.11. Let $G = (\mathcal{N}, (a_i)_{i\in\mathcal{N}}, (u_i)_{i\in\mathcal{N}})$ be a non-cooperative game, where a strategy profile $a^* = (a_1^* \dots a_{\mathcal{N}}^*)$ is a Nash equilibrium of the game G if and only if

Table 2.1: Matrix representation of a non-cooperative game

	Strategy 1: b_1	Strategy 2: b_2
Strategy 1: a_1	$(U_1(a_1, b_1), U_2(a_1, b_1))$	$(U_1(a_1, b_2), U_2(a_1, b_2))$
Strategy 2: a_2	$(U_1(a_2, b_1), U_2(a_2, b_1))$	$(U_1(a_2, b_2), U_2(a_2, b_2))$

we have the following condition:

$$u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*), \forall a_i \in \mathcal{A}, \tag{2.1}$$

where $\mathbf{a}_{-i}^*$ is the vector of strategies of all players excluding the i-th player.

Once the Nash equilibrium is achieved, no player will be interested in unilaterally changing its current strategy in order to further increase its utility. Besides, the concept of the Nash equilibrium is related to another central characteristic of the non-cooperative game referred to as a best response function [30], defined as follows:

Definition 2.12. For a non-cooperative game, we define the best response operator $b_i(a_{-i})$ of a player i to the strategy profile of its opponents a_{-i} as:

$$b_i(a_{-i}) = \{a_i^{'} \in \mathcal{A}_i | u_i(a_i^{'}, a_{-i}) \geq u_i(a_i, a_{-i}), \forall a_i \in \mathcal{A}\} \tag{2.2}$$

Indeed, the best response for a player i is to find the set of strategies that represents the optimal actions against the strategies taken by all the other players a_{-i}. Every element of $b_i(a_{-i})$ is a best response of the player i to a_{-i}. In other words, if each opponent adheres to a_{-i}, player i will have an incentive to choose a strategy of $b_i(a_{-i})$. Interestingly, when every player plays its best response strategy to the strategies of its opponents, it has no an incentive to deviate to any other strategy. Hence, the Nash equilibrium is the strategy profile $a^* = (a_1^* \dots a_{\mathcal{N}}^*)$, each of which represents a best response to the remaining players' strategies, i.e. $a_i^* \in b_i(a_{-i}^*), \forall i \in \mathcal{N}$.

2.4.2 Example: The Prisoner's Dilemma

The Prisoner's Dilemma is one of the famous non-cooperative game examples, which captures a conflicting situation in which the individual outcome goes against the social welfare. In this game, two criminals are arrested in separate isolation cells after being suspected of a crime. Thus, they have no means of talking to each other or exchanging information. The police have a limited evidence and thus need a confession. Thus, they offer a deal to the two suspects: each one can either confess and cooperate with the police so testifying against the other or choose to keep silent. Thus, if neither of them confesses, each will go to the prison for 2 years. If both suspects confess, they

Table 2.2: Prisoner's Dilemma

	Confess (C)	Not confess (NC)
Confess (C)	(−4,−4)	(0,−5)
Not confess (NC)	(−5, 0)	(−2,−2)

will serve 4 years in the jail. Nevertheless, if one confesses and thus implicates the other prisoner while the other remains silent, the first prisoner will be released while the other will incur a full sentence of 5 years.

An example of the matrix representation of the Prisoner's Dilemma is shown in Table 2.2. Each prisoner has to choose between two strategies, either to confess (strategy C) or not confess (strategy NC). Once a player has decided whether or not to confess, he receives an utility value that represents the number of years that he will spend in the prison.

It is worth pointing out that each prisoner gets a better utility by confessing, i.e. choosing C, regardless of the other player's strategic choice. Indeed, each player cannot do better than confessing since by choosing to not confess, he will get a worse utility by serving a 5-year sentence if the other player has chosen to confess or if this latter has chosen to not confess, the former will get a sentence of 2 years. Thus unilateral deviations are not beneficial to either player. Therefore, confessing is the best-response function for each player, to any strategy taken by the other. Consequently, (C,C) is the Nash equilibrium for the game which results in an utility vector (−4,−4). Obviously, the dilemma faced by the suspects here is that a line of reasoning leads each to confess, whatever the other player does. However, the outcome obtained in this case, i.e. (−4,−4), is worse for each one than the outcome that would have been obtained if they both had remained silent i.e. (−2,−2).

2.4.3 Power control problem

The power control mechanism is concerned with looking for the optimal transmit power coefficient that can be assigned to a given player in order to satisfy its performance requirements. For wireless networks, the devices or the BS are deemed to be the players of the game. Particularly, a non-cooperative game has been considered as a convenient framework to deal with the power control problems in wireless networks. This is owing to the interesting similarity between the non-cooperative game and the

behaviors of network actors seeking to achieve an optimal power partition. Indeed, when each player makes its move selfishly upon invoking a non-cooperative game, its decision inevitably affects the payoff of its opponents. Meanwhile, wireless users are naturally involved in a competitive situation wherein, the transmit power of each of them represents a principal component of the interference level suffered by the other users, which in turn impacts their transmissions and their quality of service (QoS). Consequently, solving the non-cooperative games results in optimizing the power allocation problem among wireless users. In light of this, non-cooperative games have been broadly adopted in various wireless networks whenever the power control issue is a central challenge. For instance, in [31], a non-cooperative game-based approach has been proposed to optimize the power control problem and then assess the system performance in terms of the energy efficiency. The authors in [32] has adopted a non-cooperative game to model the competitive behaviors between different users with the aim of allocating the optimal power level to each user in Cognitive Radio Non-Orthogonal Multiple Access (CR-NOMA) systems. In [33], the authors have studied the interactions between multiple devices through a non-cooperative game. They have proposed a power control algorithm that enables each user to choose its own transmit power to improve its received signal while consuming less energy.

2.5 Differential games

The classes of games investigated in the preceding sections rely essentially on the assumption that each player makes its strategic choice only once and without any knowledge of what the other players have taken as actions. Nevertheless, in many situations, the notion of time has preponderant role in the decision-making process. Indeed, the players may need to adapt their strategies more than once in order to cope with the time-varying system state. In doing so, each player's utility is no longer determined solely by the action it has chosen, but also according to its current state at a given time t. On the other hand, the distribution of the states at any time t is affected by the previous states and the players' strategies. Consequently, in such a situation, the players have to react in response to the dynamics of the system state, while the players' moves in their turns impact the evolution of the system state. One way to deal with this situation is to adopt differential games.

2.5.1 Overview of differential games

Definition 2.13. Let $G = (\mathcal{N}, \{\mathcal{P}_i\}_{i\in\mathcal{N}}, \{\mathcal{S}_i\}_{i\in\mathcal{N}}, \{\mathcal{Q}_i\}_{i\in\mathcal{N}}, \{U_i\}_{i\in\mathcal{N}})$ be a differential game played during a time interval T, defined through the following components:

- **Player set** $\mathcal{N} = \{1 \dots N\}$.
- **Set of actions** $\{\mathcal{A}_i\}_{i\in\mathcal{N}}$: represents the set of strategies or actions available to a player i. The latter determines at time $t \in [0, T]$, the action $a_i(t) \in \{\mathcal{A}_i\}$ to be selected.
- **State space** $\{\mathcal{S}_i\}_{i\in\mathcal{N}}$: denotes the state of the player i evolving over time. The definition of the state is crucial in such a game, since it describes the dynamics of the system and impacts the player's decision making.
- **Control policy** $\{\mathcal{Q}_i\}_{i\in\mathcal{N}}$: characterizes for each player i the mapping from any state of its state space to an action from its set of actions. Besides, for each player, the control policy governs the evolution of the state with respect to time with the aim of maximizing its own utility.
- **Utility function** $\{U_i\}_{i\in\mathcal{N}}$: it measures the player satisfaction of participating in the game. Indeed, designing an appropriate utility allows each player to determine its strategy and the suitable decision to make in order to meet its objectives.

Typically, such a game studies the interaction between each player and every other players in the system. Thus, to reach an equilibrium, a large number of equations must be solved, which leads to an inherent mathematical complexity, especially for dense networks. In contrast, for a large-scale system, the effect of a particular player's action and decision on its opponents becomes negligible whereas the impact of the mass behavior on a single player is considerable and can be modelled as a collective effect, namely the mean-field [34]. For these reasons, the differential game can be approximated by an equivalent game so-called the MFG to cope with a large population of players.

2.5.2 Mean Field Game

The MFG has received a significant attention as an advanced tool for dealing with the presence of a large population, through a simplified model [34–36]. Indeed, the MFG

allows each player to determine its strategy by only conforming the mass behaviors (i.e. the mean field) instead of worrying about the specific strategy of every other player. In this way, the players can make their decisions distributively with limited information and a reduced control overhead. Henceforth, the conventional N-body non-cooperative game is turned into notably more tractable 2-body problem whose analysis is conducted through two fundamental equations, namely Hamilton-Jacobi-Bellman (HJB) and the Fokker-Planck-Kolmogorov (FPK) equations.

2.5.3 Mean field definition

The definition of the mean field is of prime interest to appropriately model the state evolution of the players. In fact, the mean field is defined as a probability distribution of the state over the set of players. Firstly, at any time $t \in [0, T]$, we express the proportion of players in state s as

$$M(t, s) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{s_i(t)=s\}}, \tag{2.3}$$

where the $\mathbb{1}$ is the indicator function, it is equal to 1 when $\{s_i(t) = s\}$ holds, otherwise 0. Then, according to [37], when the number of players tends to infinity, $M(t, s)$ converges to a mean field density $m(t, s)$ given by

$$m(t, s) = \lim_{N \to +\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{s_i(t)=s\}}. \tag{2.4}$$

2.5.4 Mean field game equations

The analysis of the presented framework is mainly driven through two coupled HJB and FPK equations. The former equation characterizes the interactions between the players and the mean field density, and then enables each player to determine its strategy, while the latter equation controls the dynamics of the mean field density in response to the players' decisions, as shown in Figure 2.1. These combined equations are known as the backward and forward equations respectively and can be expressed for any player $i \in \mathcal{N}$ as follows [35]:

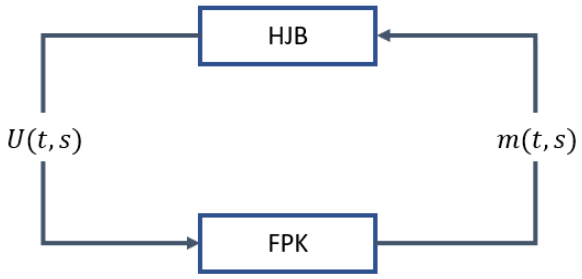

Figure 2.1: Mean Field Game with HJB and FPK equations

HJB equation:

$$-\frac{\partial v(t,s(t))}{\partial t} = \max_{p(t)}[U(t,s(t),p(t)) + \nabla_s v(t,s(t)).\frac{\partial s(t)}{\partial t}]. \tag{2.5}$$

FPK equation:

$$\frac{\partial m(t,s)}{\partial t} + \nabla_s(m(t,s).\frac{\partial s}{\partial t}) = 0. \tag{2.6}$$

The iterative solution of these coupled equations leads to the Mean Field Equilibrium (MFE) which can be represented by a convergent and stable set consisting of the optimal control strategy and the optimal mean field density.

2.6 Conclusion

Throughout this chapter, we have provided some insights about the game theoretical framework. Furthermore, we have taken a deeper look at the well-defined branches of game theory ranging from classical games such as non-cooperative games and cooperative games to differential games, namely the MFG. Particularly, we have highlighted the important notions, key characteristics and fundamental solution concepts that pertain to each type of game. Meanwhile, we have delved into the situations in which each class of games can be formulated.

In the next chapter, we introduce one of the breakthrough multiple access technologies in wireless networks that yields a meaningful improvement in the spectrum efficiency and the user connectivity due to its great potential benefits of smartly reusing the available radio resources, which is the Non Orthogonal Multiple Access.

Chapter 3

Non Orthogonal Multiple Access background

Contents

3.1 Introduction

Over the past decades, cellular communication and standardization have undergone a historical evolution in terms of multiple access techniques. The latter, commonly referred to as radio resources sharing or user multiplexing schemes, have always been identified as fundamental enabler and landmark of the consecutive generations of wireless communication systems from the 1G to the 5G [38]. Besides, the existing multiple access schemes can be classified under the umbrella of OMA or NOMA techniques depending on how the available resources are assigned to users. In the following, we first take more in-depth look at the different categories of multiple access schemes by presenting the significant strides that have been carried out in this direction. Second, we reveal our major contributions that constitute a bridge between the extremely limited RF resources and the stringent requirements of mMTC scenarios.

A general overview of the OMA approaches is given in the next section. The fundamental principles of the NOMA techniques are presented in section 3.3. We devote section 3.4 to first discussing related works that have investigated the joint channel selection and the power allocation problems for NOMA networks, and then presenting our principal contributions in view of the research gaps. Finally section 3.5 concludes the chapter.

3.2 Orthogonal Multiple Access

3.2.1 Principle of Orthogonality

As the oldest member of the multiple access family, FDMA has been essentially used in the 1G of cellular systems, namely the Advanced Mobile Phone System (AMPS). The FDMA scheme consists of splitting the available bandwidth into distinct non-overlapping frequency bands. Thus, each band is assigned to a single user in order to transmit its signal without inducing an interference to another user. Then, the world-wide migration to the digital technology led to the 2G systems in the early 1990s and its corresponding standardization known as Global System for Mobile Communications (GSM). Consequently, FDMA has been replaced by TDMA which has been considered as the dominant multiple access technique in the 2G GSM communica-

tion systems [39]. Unlike FDMA, in the TDMA users are served in different non-overlapping fractions of time or time slots. Thus, each user is allowed to transmit its information over the entire available bandwidth but only during its allocated time slot.

Since the digitization of wireless transmissions became widespread, the world have faced a growing demand for higher throughput requirements. To cope with this issue, the research community has turned the attention to an upgraded generation of cellular networks, i.e. the 3G. The evolutionary path from the 2G networks to the 3G has emerged new digital standards and sophisticated radio resource sharing technologies, most notably the CDMA [40]. Upon adopting this new multiple access scheme, the users are allowed to simultaneously access the same frequency band at the same time, while each user resorts to a user-specific signature known as a spreading code, in order to deliver its packets.

The continuously increasing number of wireless users poses challenging requirements for the 3G systems in terms of a high bandwidth efficiency and an attainable system's throughput. With the aim of addressing this challenges and ensuring sustainability of radio sharing technologies, the 3G has been shifted into the 4G of communications networks. In this vein, the Long-Term Evolution (LTE) has been standardized by 3GPP [41]. The pivotal concept of the LTE is to partition frequency and time domains to form a grid of resources, each of which referred to as a time-frequency RB [42]. Particularly, the time interval is divided into separate elements of time slots whereas the overall carrier bandwidth is subdivided into multiple sub-carriers. In order to provide a means for the multiple user access, Orthogonal Frequency Division Multiplexing (OFDM) has been proposed and specifically combined with the TMDA technique [43]. To this end, the sub-carriers within a particular time slot can only serve one specific user. In order to take one step further, multiple access schemes have further evolved into OFDMA which has been developed as a multi-user variant of the OFDM technique [44]. In contrast to aforementioned techniques, OFDMA schedules multiple users' transmissions via two-dimensional radio resources, i.e. time and frequency dimensions. In other words, unlike OFDM, the sub-carriers may be divided between several users, which means that OFDMA allows a dynamic allocation of sub-carriers among the users by exploiting the time and frequency domains, as shown in Figure 3.1. In this way, the available resources can be efficiently exploited

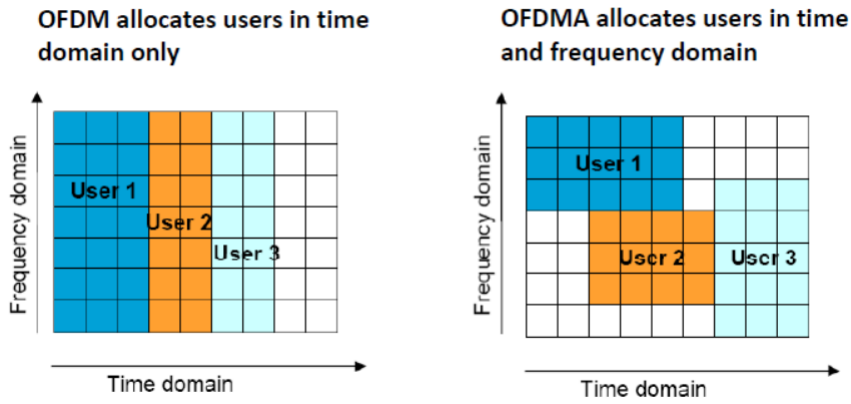

Figure 3.1: Difference between OFDM and OFDMA

by allocating the RBs to each user depending on when it needs to communicate and the amount of communication required.

Although the above-mentioned multiple access techniques deal differently with the assignment of shared radio resources to different users, they have been developed based on a common concern in mind: "orthogonality". This concept characterizes the way of separating the signals of different users by allocating distinct frequency bands, time slots, spreading codes or RBs to them.

3.2.2 OMA Limitations

Since the OMA approaches enable several users to access the network via an orthogonal division of system resources, the interference between adjacent resources is mitigated. Therefore, the signals of different users can be retrieved with a reasonable complexity by separating their messages with the aid of a simple user detection.

Nevertheless, due to the "orthogonality" characteristic, radio sharing technologies upper bound the number of served users by the number of available resources, bringing the network to its performance bottleneck. On the other hand, with the proliferation of the IoT, various unprecedented challenges have emerged for the 5G cellular communications in terms of demands for the high spectral efficiency and massive connectivity. Therefore, OMA techniques are considered as a serious impediment to the dramatic increase in the number of users and thus can no longer support the rapid

escalation of the wireless traffic. Therefore, sophisticated multiple access techniques should be designed to allow a large number of wireless users to be scheduled within scarce radio resources. Particularly, NOMA has gathered a wide attention from both the academia and industry as a revolutionary technology that serves as the fuel for the evolution of the cellular networks [8].

3.3 Key principles of NOMA

The core concept of NOMA is based on to the idea of designing a sophisticated multiple access technique with the aim of efficiently leveraging the available resources. To this end, NOMA accommodates several users within the same resources, while the power domain or code domain is adopted to deliver distinguishable signals to the receiver. Due to this non-orthogonality feature, NOMA has been ubiquitously recognized as an advantageous scheme that can fulfill the stringent requirements of high spectral efficiency and massive connectivity [45, 46]. Interestingly, NOMA not only allows multiple users to be served simultaneously, but ensures, through a certain power partition or code assignment, the fairness among them in terms of QoS requirements [47].

Basically, in conventional OMA systems, a user has to go first through an access-grant request to enable the BS to schedule its uplink transmission. In contrast, NOMA does not require in general a strict grant process and may allow multiple users to transmit their signals in a grant-free manner. Interestingly, the grant-free NOMA significantly reduces the cost of signaling and thus the transmission latency, which is particularly desirable for the mMTC scenarios.

3.3.1 Basic Technologies of NOMA

The non-orthogonal feature of NOMA is mainly driven by two advanced technologies, namely the Superposition Coding (SC) and the Successive Interference Cancellation (SIC). The former characterizes the transmission of combined signals at the transmitter side while the latter lays out the interference cancellation process at the receiver side.

3.3.1.1 Superposition Coding

Commonly, NOMA breaks the orthogonality by allocating to multiple users the same RB with the aid of the SC technique. This technique was first developed by Cover in 1972 [48], in an effort to constitute a superimposed mixture composed of different information signals from a one source to several receivers. Indeed, the SC is considered as one of the substantial building blocks of the proposed coding techniques proposed to approach the capacity of Gaussian Broadcast Channel (BC) [49]. Besides, it can be viewed as an efficient physical layer method that focuses on linearly overlapping messages signals in order to simultaneously transmit them over the shared resources. More particularly, the transmitter adopts the SC to encode the message for the weakest user, which has the worst channel condition, at a low rate. Then, it adds the information for the stronger user, which is associated with the better channel condition, and so on.

3.3.1.2 Successive Interference Cancellation

Once the transmitter employs the SC technique, it transmits the resulting superimposed signals within the same time-frequency resources. At the receiver side, Cover was first introduced the SIC procedure in order to detect and decode the signal of the different users [48]. Indeed, the SIC technique is invoked to split the combined signals and then retrieve the information of each user by taking advantage of the differences between the signals' strengths of users to determine their decoding orders. More precisely, when a signal of a particular user is decoded, it is removed from the overlapped signals before decoding the signal of the next user.

In this regard, the receiver processes the SIC to decode and recover the desired messages from the received superimposed signals resulting from the SC operation. Therefore, the way that the transmitter executes the SC scheme has a significant impact on the ability of the SIC to successfully decode the users' signals. Obviously, since each user faces a level of interference that depends on the direction of its transmission, the implementation of the SIC in the NOMA uplink transmission is different from that in the NOMA downlink transmission. In the following, we spotlight the main property of SIC process under downlink as well as uplink NOMA transmissions.

3.3.2 NOMA Transmission

3.3.2.1 Downlink NOMA Transmission

In downlink NOMA transmissions, the SC is implemented at the BS in order to send multiplexed user signals over a given RB. Indeed, upon applying the SC procedure, the BS generates superimposed signals x, which is transmitted then to the scheduled users presented in the NOMA network. Thus, the resulting overlapped signals x at the BS can be expressed as:

$$x = \sum_{n=1}^{N} \sqrt{q_n} z_n, \tag{3.1}$$

where q_n denotes the power allocation coefficient and z_n represents the message signal for a user n. Then, the received signal at the n-th user can be written as:

$$y_n = h_n x + r \tag{3.2}$$

where h_n denotes the channel gain coefficient between the BS and the n-th user and r is the additive noise with zero mean and variance σ^2. By invoking the SIC process, the closest user to the BS, having the best channel quality, successively detects the signal of the other users and then subtracts their data from the received superimposed signals y_n. Thereby, since now the received signal is decontaminated from the interference introduced by the other users, the receiver can recover its signal of interest and decode its own message. Subsequently, the user with the second best channel employs the same procedure, except that it treats the signal of the user with the best channel as a noise, and so on. Hence, the user with the weakest channel condition does not invoke the SIC and instead treats other users' signals as a noise. Thus, it proceeds directly to the recovery and decoding of its data.

3.3.2.2 Uplink NOMA Transmission

In an uplink NOMA scenario, several users communicate with the BS via one RB. Then, the BS resorts to the SIC process in order to separate the overlapped signals given by:

$$y = \sum_{n=1}^{N} h_n \sqrt{p_n s_n} + b_n, \tag{3.3}$$

where p_n and s_n denote the power allocation coefficient and transmit symbol of the user n, respectively. Indeed, the transmitters are ordered based on their signal strengths. First, the BS starts by decoding the strongest signal, considering the remaining users as an interference. Then, it subtracts the decoded signal from the composite received signals and decodes the signal of the next user, and so on. Besides, the outcome of the decoding process for each user n is determined by its Signal-to-Interference-plus-Noise Ratio (SINR) which is expressed as

$$\gamma_n = \frac{p_n |h_n|^2}{\sigma^2 + \sum_{m=1}^{n-1} |h_m|^2 p_m}. \tag{3.4}$$

Then its throughput (data rate) is written as:

$$R_n = \log(1 + \gamma_n) = \log(1 + \frac{p_n |h_n|^2}{\sigma^2 + \sum_{m=1}^{n-1} |h_m|^2 p_m}). \tag{3.5}$$

Typically, in order to process the SIC successfully, when considering weaker users as an interference, every user should have a SINR higher than the SINR threshold. It is worth pointing out that the SIC has a recursive nature which means that the BS proceeds to the decoding of the signal coming from a given source if and only if the signals from the previous users have been successfully decoded and subtracted from the received superimposed signals. In other words, upon invoking the SIC procedure, if the signal of a user n is decoded successfully, the BS subtracts it and moves to the user with the best channel gain among the remaining users; otherwise, the SIC algorithm stops and the remaining signals are ignored.

3.3.3 The existing NOMA techniques

Distinguished from the conventional OMA techniques, NOMA allows multiple users to share the available wireless resources, whereas a new dimension is adopted to perform multiplexing of different data of users. In other words, in order to superimpose the transmission signals over a common RB, an additional domain is exploited, namely the power domain or the code domain. Hence, the dominant NOMA schemes are broadly categorized into the PD-NOMA and the CD-NOMA. The main existing CD-NOMA

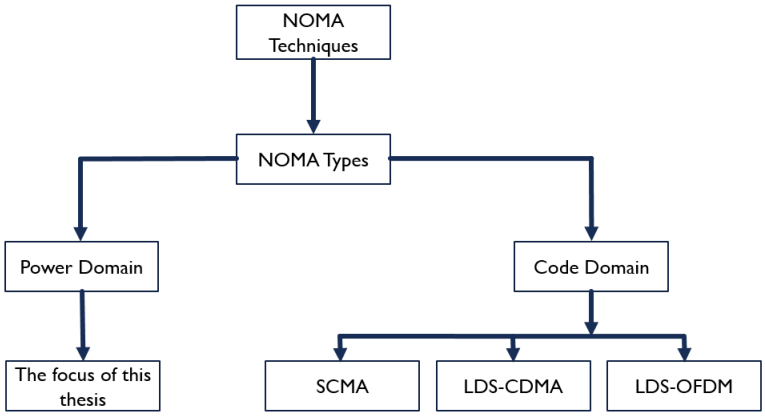

Figure 3.2: NOMA classification

and PD-NOMA techniques are summarised in the Figure 3.2.

3.3.3.1 Code-Domain NOMA

The CD-NOMA represents one of the fundamental implementations of NOMA. Inspired by the conventional CDMA concept, CD-NOMA enables multiple users to occupy shared time-frequency resources while using spreading codes to spread the users' data. The major difference to CDMA crystallizes in the choice of the spreading sequences which have to be either sparse or non-orthogonal with low cross-correlation codes. Besides, the CD-NOMA is primarily based on the codebook design in which the spreading sequences serve to generate the codewords known by both transmitter and receiver sides. CD-NOMA has been proposed in several techniques such as Low-Density Spreading CDMA (LDS-CDMA), Low-Density Spreading OFDM (LDS-OFDM), Sparse Code Multiple Access (SCMA), which will be detailed in the following.

LDS-CDMA:
LDS-CDMA introduced in [50], constitutes an advanced version of the CDMA approach. Basically, CDMA combines the transmitted data of several users through or-

thogonal spreading sequences, which requires only a simple-user detection with a low complexity to decode each user's signal. Nevertheless, such a system can only serve a small number of users simultaneously, constrained by the number of chips. With the aim of accommodating much more users, non-orthogonal sparse spreading sequences are used by LDS-CDMA instead of the conventional dense signatures in CDMA. Due to the sparse nature of codes, the interference within each chip is further reduced. In this regard, the suitable design of spreading codes results in significantly mitigating the interference impact, which in turn yields in system performance improvements. At the receiver, the Message Passing Algorithm (MPA) is invoked as a sophisticated multi-user detection to recover the original signals.

LDS-OFDM:
LDS-OFDM [51] can be viewed as a combined scheme of LDS-CDMA and OFDMA, in which LDS-CDMA is applied in a multi-carrier form. As in the case of LDS-CDMA, low-density sequences are invoked at the transmitters, while the MPA-based detection is applied at the receiver side. Firstly, the information symbols are mapped to the low density spreading sequences. Then, the output of the signature is spread across a set of subcarriers using an OFDM modulator. Unlike LDS-CDMA, LDS-OFDM relies on the multi-carrier communications, which enables it to be a highly compatible approach with wideband channels. In addition, the interplay between the OFDM for the modulation mapping and LDS for the multiple access, results in a greater flexibility in the resource allocation [52].

SCMA:
As an enhanced variant of CD-NOMA techniques, SCMA was proposed in 2013 by H. Nikopour and H. Baligh [53]. Similar to the aforementioned schemes, the SCMA concept handles the challenge of multiple users accessing the available RBs with the aid of sparse spreading codes. Specifically, SCMA uses multi-dimensional constellations to generate distinctive codebooks in order to distinguish users. Indeed, the information bits of a given user are directly mapped to a sparse vector, so-called a codeword, drawn from its predefined codebook. Then, the resulting mapped codewords of all users are superimposed over shared resources allowing SCMA to achieve higher overloading factor than LDS-OFDM. Due to the sparsity feature of codebooks, the receiver side can adopt a low-complexity detection algorithm, namely the MPA,

to efficiently detect and retrieve the users' signals.

3.3.3.2 Power-Domain NOMA

The elegant idea of the PD-NOMA is to allow several users' signals to be superimposed within the same time-frequency block by assigning to them different power coefficients in order to enable the BS to differentiate their messages. Indeed, the BS performs the SIC while leveraging the power difference among users with the aim of separating the combined signals and recover the information of each user. In this way, the power control represents a pivotal research issue for PD-NOMA systems. In fact, since the users' signals are distinguished via distinct power levels, an improper power allocation leads to increased interference levels, which in turn results in inefficient detection of the interfering signals at the receiver side. Furthermore, in order to ensure a successful transmission with the lowest energy consumption, the power coefficients must be properly assigned to the multiplexed users on each RB so that their QoS requirements and power constraints are taken into account. In this thesis, we investigate the PD-NOMA system and focus on the careful optimization of the power allocation in order to control the interference of superimposed signals and maximize the overall system performance.

3.3.4 Hybrid NOMA

3.3.4.1 NOMA limitations

Despite the aforementioned features, we can not turn a blind eye to the limitations of NOMA. Indeed, multiple users are admitted to access a certain RB simultaneously, but at the cost of an emerging co-channel interference, which in turn may spoil the system performance. Although the use of the SIC procedure can alleviate some of the interference effects, the complexity of the decoding and the cancelling other users' signals, especially in downlink transmissions, increases with the number of users. Consequently, designing a NOMA network in a massive connectivity scenario is a challenging and even a burdensome task. To cope with this issue while availing the above-mentioned features of NOMA, a Hybrid NOMA scheme has been proposed as an alternative approach [18].

3.3.4.2 Hybrid NOMA principle

In NOMA systems, the power domain is extensively exploited as a new dimension to multiplex several users' signals over the available time-frequency resources. Henceforth, NOMA can be recognized as an "add-on" technique to the conventional OMA schemes which means that NOMA has the potential to harmoniously co-exit with any OMA technique, thus forming a Hybrid NOMA system.

Typically, in Hybrid NOMA systems, the users shift their attention from jointly performing NOMA toward meticulously forming multiple groups. Orthogonal RBs are assigned to these groups so that the members of each group share a particular RB in order to non-orthogonally transmit their messages. What is so beautiful about this alternative scheme is the amalgam between OMA and NOMA schemes such that the orthogonality concerns only the inter-groups whereas the non-orthogonality feature is invoked within each group. Herewith, users belonging to the same group communicate simultaneously with the BS through the resource assigned to that group. Therefore, each user perceives an interference level only from the members of its group, while the inter-group interference is avoided which considerably reduces the complexity of the SIC implementation at the receiver.

Ding et al. have investigated in [18], a Hybrid NOMA approach, wherein users are split up into multiple groups. They have proven that in the proposed Hybrid NOMA system, the transmit power allocation is optimized and the attainable network performance is boosted. Nevertheless, they have underlined that the performance of the Hybrid NOMA network fundamentally depends on how the users are divided into different groups and how the transmit power levels are allocated to the members of each group. Therewith, the joint optimization of the user grouping and the resource allocation is a non-trivial task that needs more impactful tools to be solved. All this drives the researchers to investigate the resource management problem for the Hybrid NOMA network with an eye toward intelligently designing power allocation strategies and efficiently arranging the users into multiple groups.

3.4 Resource management in NOMA systems

In this dissertation, we are interested in jointly solving the resource allocation and power control problems in a Hybrid NOMA scenario. Before diving deeply into the

contributions of this thesis, in the following, we provide interesting insights into two pivotal research directions in the context of NOMA, namely the power control and the user grouping.

3.4.1 Power Control for NOMA

Usually, wireless communication networks are characterized by an important level of interference encountered by each user. More precisely, the power control is a critical pillar in the design of PD-NOMA systems since assigning an inappropriate power coefficient may increase the energy consumption of a given transmitter. Meanwhile, at the receiver side, an improper power allocation results in an inefficient detection performed by the SIC process, which in turn affects the overall system performance. In an effort to mitigate the interference effects, the power allocation has been broadly investigated in NOMA-based communication systems.

Among these research activities, NOMA-based D2D networks have gained a considerable interest as one of the most challenging scenarios that requires efficient power control policies to tackle the emerging interference problem [54–59]. For instance, the authors in [54] have designed a centralized power control algorithm with the aim of maximizing the system performance by optimizing the trade-off between achieving a maximum sum rate of D2D pairs and ensuring minimum rate requirements of NOMA users. In [55, 56], NOMA technology has been applied to a D2D group in order to maximize the energy efficiency by alleviating the interference impact. Moreover, the authors have formulated a joint assignment of subchannels and power levels with the objective of maintaining the SINR requirements for the different users. The difference between these works is that [55] is concerned with maximizing the total system rate, while [56] focuses on minimizing the total transmit power.

Meanwhile, a plenty of researches have been devoted to model the power control problem in NOMA aided Heterogeneous Networks (HetNet) [60–62]. Theses contributions seek at improving the spectral efficiency and enhancing the overall system performance. More precisely, in [60], NOMA scheme is invoked within each small cell in order to boost the spectrum efficiency while taking into account the fairness issue. The work of [61] studied how to incorporate the NOMA technique in both macrocells and small-cells with the objective of maximizing the system throughput under the constraints of users' QoS requirements. The target of the proposed NOMA-based

HetNet approach in [62] is to optimize the energy efficiency of small-cells as well as macro cells through an effective joint subchannel selection and power control.

Furthermore, the power control issue has been one of the major concerns in a multi-carrier NOMA system [13,63–67] in which the total bandwidth is divided into multiple subcarriers and allocated to users to improve the spectrum utilization. For instance, the authors in [63] have addressed the joint power and subcarrier allocation problems to enhance the weighted system throughput using a monotonic optimization. In [64], NOMA is amalgamated with the well-known slotted ALOHA protocol to form Aloha-based NOMA (NM-ALOHA) scheme where each user is assumed to select a sub-carrier and a power level independently and uniformly in a random way. Zhu et al. have investigated in [13] a downlink power allocation problem in an effort to maximize the overall system throughput. The work in [66] has been interested in improving the individual energy efficiency through a resource allocation optimisation in order to ensure the fairness among users.

Although the above-mentioned contributions have significantly improved the performance of NOMA systems, to some extent, none of them have focused on studying the massive access scenario. However, the massive connectivity represents the primary requirement of IoT and the central properties of the mMTC. Although, the MTC have an infrequent transmission nature of small packets with a low data rate, the traffic generated by the massive number of MTDs is rather challenging. Therefore, providing the massive connectivity is overwhelmed by the scarcity of resources. Consequently, supporting a massive number of devices while ensuring the stringent requirements imposed by the MTC is not a burdensome task. To this end, several research efforts have been conducted to avail the ability of NOMA in smartly reusing the available spectrum in order to support a wide range of MTDs [68–73]. These works have designed multiple access techniques with one common goal in mind: how to comply with the demand for the massive access of MTDs. Nevertheless, they have not taken into account the specific characteristics of MTDs. Indeed, MTC traffic is often characterized by the sporadic transmission which means that the MTD activity is sparse. Thus, only a small portion of devices are in a data transmission state and may therefore transmit simultaneously in the same time slot [7]. In fact, the MTDs transmit their signals on a periodic basis and switch to the idle mode to prolong their battery lifetime. Such behavior needs an efficient access technique in order to optimize

the engagement of massive devices in the network, thereby reducing the transmission latency and saving the signaling overhead. In doing so, MTDs having limited energy budgets, can satisfy rigorous energy consumption requirements.

Motivated by the aforementioned challenges, we seek to address the dilemma between the demand for the massive connections of MTDs and the scarcity of available resources. For this sake, in this dissertation, we investigate the PD-NOMA concept to pave the way for an efficient power control that firstly manages the access of the MTDs to the network and then differentiates their signals through distinct power levels.

3.4.2 User grouping Based Hybrid NOMA Networks

As well as the power control, the user grouping constitutes a fundamental pillar for the design of NOMA schemes. Ideally, multiple users can share a particular RB to reach a high spectral efficiency using NOMA. In [74], the authors have proposed a NOMA scheme that allows any number of devices to be served on an RB signal with no restriction on this number. However, accommodating such a huge number of users comes at the expensive of an increased computational complexity of the SIC at the receiver side. Thereafter, it is neither feasible nor efficient to jointly superimpose all the users' signals using one RB. Hence, practical implementations have considered multiple groups of users where NOMA is carried out for a fewer number of users within each group.

Ultimately, significant research efforts have been exerted to investigate the NOMA pairing concept where NOMA is implemented for two users per group [18,75–82]. Typically, most of user pairing schemes arrange users into two groups consisting of near and far users. Then, one near user is paired with one far user to be non-orthogonally multiplexed over the same resources while different power levels are allocated to them. Indeed, the group division is performed on the basis of either channel conditions or a distance criterion. For instance, [75, 76, 78, 79] have proposed user pairing techniques with the goal of forming a pair of users whose channel conditions are significantly distinctive. The authors in [18] have investigated the impact of the user pairing on the NOMA downlink system. They have proven that the larger the difference between the channel coefficients of paired users, the better is the execution of the SIC process and thus the higher is the achieved rate performance. However, the performance of the

NOMA network is affected when the channel gain difference in a user pair decreases. Conversely, other existing pairing studies [80–82] have performed the non-orthogonal multiplexing while using a criteria relied on the distance gap between the users. In fact, they have showed that it is preferable to pair users having a certain distance difference between them in order to boost the system throughput. Nevertheless, with the increase in the demand of massive connectivity, considering only two users sharing a given RB is a bottleneck of the NOMA system performance [83].

In contrast, only a few contributions have been conducted to leverage maximum advantages of NOMA in its true sense by clustering more than two devices within each RB. In [84, 85], two to three users are admitted to form a NOMA group and share common resources in order to maximize the sum-rate of the grouped users. The authors in [86] have proposed a user clustering scheme in which two and four users are multiplexed using the NOMA concept in order to improve the total throughput of the users under minimum rate constraints. On the other hand, variable cluster sizes have been considered in [83, 87–90]. The authors in [83] have developed a dynamic user clustering approach depending on the distinctness among channel gains of users. This work has been extended in [87] with the goal of maximizing the sum rate of all the users while taking into account a rate constrained scenario. However, both contributions have assumed the same number of users per cluster which is an unrealistic assumptions for NOMA transmissions. The approach proposed in [88] seeks to approximate the optimal solution for the user grouping problem by performing an exhaustive search. However, this approach is not affordable for the practical NOMA scenario due to its high computational complexity especially when a large number of users are involved. The common concept of these studies [83, 87, 88] is that they treat the group size as a given parameter. To overcome this issue, in [89], a maximum feasible cluster size has been analytically derived and a distributed grouping scheme has been designed accordingly. A larger number of users per cluster has been investigated in [90] where the authors have proposed a user clustering scheme using a genetic algorithm-based heuristic approach. However, the performance gains of this scheme are restrained by the complexity of the SIC procedure, which linearly increases with respect to the group size.

In the light of the above, finding an optimal NOMA grouping is greatly challenging since the attainable performance of the NOMA system is highly related to the group

size, the selection of devices that it is preferable to form a certain group together and the targeted objectives of the grouping formation process. It has been shown that when only two users are considered to form a cluster, interference effects can be mitigated by applying the SIC process at an affordable complexity cost. However, the available RBs are assumed to be orthogonally allocated between the NOMA groups. Due to the scarcity of resources, serving only two users via a single orthogonal RB is not anymore suitable for the massive IoT scenarios. Therefore, it is crucial to investigate the user selection strategy of more than two users with an eye toward satisfying specific objectives such as maximizing the throughput, reducing the energy consumption, minimizing the latency. Consequently, how to design an effective user grouping while providing a significant trade-off between the NOMA performance gains and the decoding complexity of the SIC remains an ongoing research topic.

3.4.3 Joint user grouping and power control

In view of the above-mentioned challenges, this thesis studies the user grouping concept along with the power allocation policy for the Hybrid NOMA system. Unlike what is typically assumed in the NOMA-based user clustering, we treat the group size as an unknown design parameter that we seek to investigate while taking into account the efficiency of the SIC procedure. Particularly, we are interested in addressing the optimal user clustering problem intertwined with the power allocation issues to determine the way that users are superposed and share the associated RB. Usually such combined problems are solved through a joint optimization. For this sake, firstly we model the aforementioned optimization problems with the aid of a Bi-level game composed of a non-cooperative power control game underlying a cooperative Hedonic game. In fact, the users invoke the Hedonic game to organize themselves into multiple groups. Then, within each group, they autonomously decide the appropriate transmit power to use upon applying the non-cooperative game. Secondly, in the presence of a large population, we turn our attention to the MFG theoretic framework in order to cope with densely deployed networks. In fact, the proposed MFG enables the users to interact with the collective behavior of the other players, whatever the specific strategy of each one. By doing so, the users can appropriately regulate their transmit power according to the limited feedback received from the BS which considerably simplifies the resolution of the game. Finally, we resort to a RL tool, namely the MAB

approach to jointly optimize the resource allocation and the power control problem. In fact, the MAB approach is applied on the top of the MFG framework to derive distributed decision-making algorithms that enable the users to arrange themselves into groups. Once each user has joined its selected group, it determines the transmit power to use in order to deliver its packets.

3.5 Conclusion

In this chapter, we have spotlighted some key ideas of this dissertation. Firstly, we have reviewed the background of NOMA schemes by outlining their main properties and presenting their operating principles. Besides, due to its ability to be concordant with a conventional OMA technique, we have introduced a Hybrid NOMA scheme as an alternative approach that represents the amalgam between OMA and NOMA scheme wherein the users are meticulously divided into multiple NOMA groups. Secondly, we have presented the potential resource management challenges posed to the NOMA network performance and the research activities that have been conducted in this direction. Furthermore, we have delved into our principal contributions to fill the void in the literature while dealing with the resource allocation problem in Hybrid NOMA systems.

In the following part, we dive deeply into the study of the Hybrid NOMA network with a primary focus on the joint optimization problems of the user grouping and the power control. More precisely, in the next chapter, we model the behavior of MTDs using a Bi-level game composed of a non-cooperative power control game and a cooperative coalitional game.

Part II

A Game theory Framework for Resource Allocation in NOMA-based Networks

Chapter 4

Game Theoretical Framework for Joint Channel Selection and Power Control in Hybrid NOMA

Contents

4.1 Introduction

Towards the future 6G of wireless communication systems, mMTCs targeting seamless and ubiquitous connections, have captured a great attention in recent years. Indeed, the MTC have been considered as an enabling technology that is able to support the proliferation of the IoT and provide the communication infrastructure for its emerging applications. Besides, one of the substantial objectives of mMTC is to set up connections among ever-increasing number of MTDs and to cope with the immense data traffic volume generated by them. However, the requirements for massive connectivity and a higher spectral efficiency put the current cellular systems under severe constraints. In fact, the orthogonal channel access, in which at most one device in each time slot can transmit on each RB, is becoming a bottleneck for MTC applications. Hence, a paramount shift in wireless technologies is required and it is even an urgent task to be tackled.

Specifically, the design of appropriate multiple access techniques is one of the most convenient approaches to address the dilemma between the scarcity of resources and the massive connectivity. For this sake, wireless communication systems are strongly trending towards NOMA-based access techniques, wherein multiple users perform simultaneous transmissions in common resources, resulting in overlapped signals. Then, the BS executes the SIC procedure to extract and decode the desired signal for each user.

An efficient way to leverage the features of NOMA is to arrange the devices into multiple groups and then assign orthogonal RBs to them. In doing so, a Hybrid NOMA system is established, wherein distinct power levels have to be allocated to the devices in each group allowing them to conduct the transmission over the shared RB. Indeed, non-orthogonally multiplexed signals of a large number of devices result in a higher SIC computational complexity. Hence, the user grouping is considered as one of the most practical implementations that can significantly lower down the complexity of the SIC decoding by accommodating a fewer number of MTDs within each group. On the other hand, the SIC process is invoked in order to decode the interfering signals by exploiting the disparity of power coefficients among users. In this way, the assignment of inappropriate power allocation affects the efficiency of the SIC and increases the energy consumption of the users, which in turn deteriorates the

overall system performance. Therefore, it is of the utmost importance to carefully study the power control problem in order to mitigate the interference between users sharing the same resources and thus further enhance the gain of Hybrid NOMA systems. Henceforth, the power control as well as the user grouping constitute two interacting components that have to be jointly optimized to set up a NOMA scheme. Despite the significant research efforts that have been exerted in studying the resource allocation and the user grouping for NOMA-based networks, there are still interesting questions that need to be answered: How to arrange users into groups without treating the group size as a given parameter? How to leverage the benefits of incorporating more than two users without spoiling the system performance? And how to make each device autonomous to determine its most suitable transmit power level that comply with its energy consumption requirements?

Meanwhile, game theory is ubiquitously adopted in NOMA networks as a robust mathematical framework to provide flexible solutions to critical optimization problems such as the power management, the user grouping and the wireless channel allocation. Particularly, game theory encompasses a set of impactful analytical tools that can deal with the selfish nature of wireless users and appropriately derive a distributed decision-making process that enables the users to autonomously react against their opponents [91–100]. In [91], a Hybrid NOMA system has been investigated in a downlink scenario where the users are arranged into multiple coalitions by the aid of a cooperative game in the partition formation. Time slots are assigned to each group, wherein the users communicate through a NOMA-based scheme. Sung et al have proposed in [93] a distributed power control algorithm using the game theoretic framework for a two-cell NOMA system in order to minimize the power consumption under data rate constraints. In [95], the authors have applied game theory to uplink NOMA-based ad-hoc wireless networks in order to derive a distributed approach that maximizes the energy-efficiency of the users while meeting their QoS requirements. Furthermore, Hybrid NOMA assisted cognitive radio networks (CRNs) have been studied in [97,98]. The authors have adopted a cooperative coalitional game to maximize the energy-efficiency of CRNs by designing a user clustering scheme that allows the users to choose between NOMA and OMA to transmit their packets. Besides, the author of [100] has extended the work in [64] by formulating a game-theoretic approach for the NM-ALOHA scheme where the NOMA technique is employed for the random

access and the uplink transmissions are not coordinated by the BS. Thus, the author has proposed a NM-ALOHA game whose purpose is to enable users to more effectively choose the transmission parameters, namely the probability of transmissions. On the other hand, under a NOMA-based MTC context, the work in [99] has focused on applying a non-cooperative game theory to derive optimal power allocation strategies in order to improve the communication reliability.

Unlink the above-mentioned works, the present chapter use game theory in order to model a Hybrid NOMA scenario, wherein the users are divided into multiple groups and NOMA is implemented for the devices in each group. More precisely, orthogonal RBs are assigned to the groups, whereas the members of each group share the assigned resource to non-orthogonally transmit their packets. Thereby, we formulate the resource allocation problem as a Bi-level game theoretical framework with the aim of meticulously forming groups of devices and suitably designing power control strategies within each group. The main contributions of this chapter are depicted as follows:

- Unlike what is usually studied in NOMA based techniques, we derive two low-complexity algorithms to jointly address the resource allocation and the power control problem in an uplink PD-NOMA scheme. The proposed approach enables multiple MTDs to organize themselves into coalitions and autonomously decide the transmit power to use.

- We use a Bi-level game composed of a non-cooperative game underlying a cooperative Hedonic game. Firstly, the MTDs apply the Hedonic game to create a partition of coalitions. Then, within each coalition, the MTDs use the non-cooperative power control game to determine by themselves the appropriate transmit power in a decentralized manner.

- We consider that MTDs' transmissions may be probabilistic in order to enable the BS to allocate all the MTDs when the number of users in the system increases.

- Simulation results show that the proposed Bi-level game allows the devices to achieve a high successful packet transmission rate while consuming less energy.

The rest of this chapter is organized as follows. In the next section, we introduce the

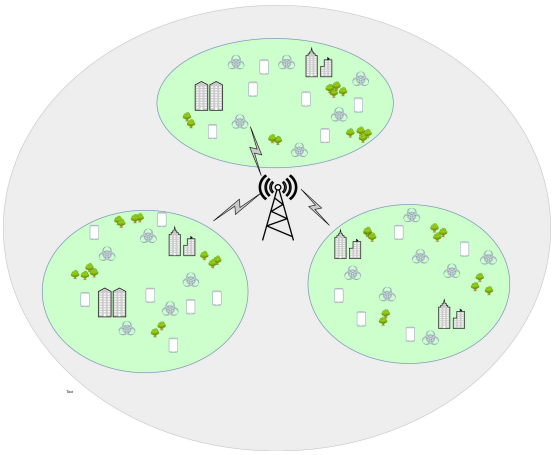

Figure 4.1: System model.

network model and focus on the PD-NOMA to perform the signal multiplexing. Section 4.3 introduces the Bi-level game and proposes two low complexity algorithms to solve the joint channel selection and power allocation problems for NOMA networks. To illustrate the performance of the proposed techniques, simulation results are given in section 4.4. The chapter is concluded in section 4.5.

4.2 System model and assumptions

4.2.1 Network model

Consider an uplink NOMA network, as depicted in Figure 4.1, where a single BS is located at the centre whereas a set of N MTDs are independently scattered throughout the coverage area. The positions of the devices are modeled using the homogeneous Poisson point processes (PPP) Φ_N with density ω_N [101]. It is noteworthy that the PPP model can conveniently abstract the network in which the MTDs are randomly distributed and each device generates its own traffic according to its position without any need to a particular human intervention. Throughout this thesis, we consider a Hybrid NOMA scenario composed of multiple groups, so-called coalitions wherein the available bandwidth is split up into K sub-carriers. The latter are orthogonally assigned to the groups whereas each group of MTDs use one sub-carrier to non-

orthogonally communicate with the BS. More precisely, when a device i is a part of a given coalition, to which the k-th sub-carrier is allocated, the channel coefficient between this device and the BS is represented by $h_{k,i} = \frac{g_{k,i}}{l_{k,i}}$, where $g_{k,i}$ and $l_{k,i}$ denote the Rayleigh fading and the path loss, respectively. We adopt the Free-Space path loss model [102] to define the path loss. Thus $l_{k,i} = \left(\frac{\lambda_k \sqrt{G_l}}{4\pi r_i}\right)$, where G_l is the product of the transmit and receive antenna field radiation patterns in the Line-Of-Sight (LOS) direction, λ_k is the signal weave-length over the k-th sub-carrier and r_i is the distance between MTD i and the BS.

In this way, the devices belong to each coalition transmit their messages through the associated sub-carrier k. Hence, the received overlapped signals at the BS from the k-th group is expressed as:

$$y_k = \sum_{i=1}^{N_k} h_{k,i} \sqrt{p_{k,i}} s_{k,i} + b_k, \tag{4.1}$$

where $s_{k,i}$ and $p_{k,i}$ represent the transmit symbol and the power coefficient of the device i through the k-th sub-carrier respectively. The transmit power of the user i is constrained by the maximum transmit power P^{max}. In addition, b_k denotes the additive noise of variance σ^2 over the sub-carrier k. Once the BS receives the superimposed signals, it applies the SIC procedure to detect and recover each user' s signal.

4.2.2 SIC process

Since in our work we exploit the PD-NOMA scheme, multiple devices are allowed to simultaneously access the same sub-carrier and are thus multiplexed using distinct power allocation coefficients in order to enable the BS to distinguish their signals. Thus, at the BS side, the SIC is invoked to split the combined signals and alleviate the co-channel interference.

Interestingly, in the Hybrid NOMA scenarios there is no interference between MTDs belonging to different NOMA coalitions, instead each device perceives only an interference level from the members of the same coalition. According to the SIC principle, the BS decodes the strongest MTD's signal received on a given RB considering all the other MTDs of the cluster as an interference, subtracts the signal of the strongest

device from the received superimposed signals and decodes the signal of the next MTD, and so on. Hence, in order to process the SIC successfully, when considering weaker MTDs as an interference, every device should have a SINR higher than the desired SINR threshold defined as γ_{th}.

4.2.3 Assumptions

Consider a coalition $\mathbb{C}$ of devices sharing one sub-carrier. Since the devices have to use different power levels to non-orthogonally transmit their packets, an appropriate power allocation should be performed. On the other hand, MTDs having limited power capacities generally require high reliable connections with low data rate requirements. Note that, achieving a high SINR leads the device to transmit at a high power level, which is not beneficial for its battery life. Therefore, each device seeks to successfully transmit its packets while reducing its power consumption as much as possible. Consequently, the central objective of this chapter is to develop an efficient power control strategy with the aim of ensuring the transmission reliability by keeping the SINR of each MTD beyond an acceptable level while meeting its stringent requirement of the energy consumption. With this in mind, we provide a novel power allocation approach that focuses on assigning distinct power allocation coefficients to different users based on their minimum QoS requirements in terms of SINR.

In what follows, we investigate the SINR requirement and develop a power allocation technique. In fact, we are interested in finding an optimal power policy that enables a given user to achieve the SINR threshold γ_{th}. Interestingly, we take advantage of the closed relationship between the signal-to-noise ratio (SNR) and transmit power expressions to design our power control approach, in which the devices aim to be distinguished in terms of received SNR values rather than power levels. The latter will be then determined once the users have different SNR values. Let us now set the SNR levels at the BS to guarantee a successful decoding and thus maximize the capacity of the network. For the sake of notation simplicity, we omit the sub-carrier index.

Proposition 4.1. Consider that the L MTDs of coalition $\mathbb{C}$ are sorted using their channel gains in a decreasing order. Let $\Gamma = \{\gamma_1, \gamma_2, \cdots, \gamma_L\}$ be a target SNR vector

defined as follows:

$$\begin{aligned} \gamma_1 &= \gamma_{th} \\ \gamma_2 &= \gamma_{th} \times (1 + \gamma_1) \\ \gamma_3 &= \gamma_{th} \times (1 + \gamma_1 + \gamma_2) \\ &\vdots \end{aligned} \tag{4.2}$$

If every MTD targets a distinct SNR value, the BS is able to process SIC successfully.

Proof. Note that each device seeks to satisfy the SINR requirement while consuming less energy, which means that its SINR has to be equal to γ_{th}. Firstly, we assume that the BS has decoded successfully the messages of the $L-1$ strongest MTDs. Since the weakest user (MTD 1) has a target SNR of γ_{th}, i.e. $\gamma_1 = \gamma_{th}$, by construction, the BS can decode its message. Consider now MTD 2 and assume that the BS has decoded all the message of stronger MTDs. Since the target SNR of MTD 1 is γ_{th}, we have $p_1|h_1|^2 = \sigma^2\gamma_{th}$. Then, the $SINR_2$ of MTD 2 is

$$SINR_2 = \frac{p_2|h_2|^2}{(\sigma^2 + p_1|h_1|^2)} = \frac{p_2|h_2|^2}{(\sigma^2(1+\gamma_{th}))}. \tag{4.3}$$

Since the SNR of MTD 2 is $\gamma_2 = \gamma_{th}(1+\gamma_{th})$, we have

$$SINR_2 = \frac{\gamma_{th}(1+\gamma_{th})}{(1+\gamma_{th})} = \gamma_{th}. \tag{4.4}$$

We end up with the SINR of MTDs 1 and 2 that are both equal to γ_{th} and thus the BS can decode them. By induction, it can be easily proven that using the proposed construction, the BS can perform the SIC successfully. □

Theorem 4.1. Consider L MTDs transmitting on the same RB using a power allocation scheme according to the Proposition 4.1 and denote by Γ the target SNR vector. Then, if we allocate a new MTD with a target SNR not higher than $\max\{\Gamma\}$ the SIC will fail.

Proof. Assume that the m-th device joins the coalition and targets an SNR γ_m, and $\exists\gamma_i \in \Gamma$ such that $\gamma_{i-1} \leq \gamma_m \leq \gamma_i$. Using the Proposition 4.1, the SINR of the i-th

user is given as:

$$\frac{p_i|h_i|^2}{(\sigma^2 + \sum_{j=1}^{i-1} |h_j|^2 p_j)} = \gamma_{th}, \tag{4.5}$$

but, since the m-th MTD joins the coalition, we obtain

$$\sigma^2 + \sum_{j=1}^{i-1} p_j|h_j|^2 < \sigma^2 + p_m|h_m|^2 + \sum_{j=1}^{i-1} p_j|h_j|^2, \tag{4.6}$$

and then the SINR of MTD i becomes strictly lower than γ_{th}. Thus, the decoding of the signal of the MTD i and all the weaker MTDs will fail. Assume now that there exists a device i that can reduce its target SNR γ_i', then we have two cases. If $\gamma_i' \leq \gamma_{i-1}$, then the SIC will fail according to the first part of the proof. In fact, this case is similar to a new user who joins a coalition composed of MTDs $\{1, \cdots, i-1\}$ with a target SNR lower than $\max\{\gamma_1, \cdots, \gamma_{i-1}\}$. Let us focus now on the case where $\gamma_{i-1} < \gamma_i' < \gamma_i$. We have

$$\frac{p_i|h_i|^2}{\sigma^2 + \sum_{j=1}^{i-1} p_j|h_j|^2} = \gamma_{th} \Rightarrow \sigma^2 + \sum_{j=1}^{i-1} p_j|h_j|^2 = \frac{p_i|h_i|^2}{\gamma_{th}} \Rightarrow \frac{p_i^{'}|h_i|^2}{\sigma^2 + \sum_{j=1}^{i-1} p_j|h_j|^2} = \gamma_{th}\frac{p_i^{'}}{p_i} < \gamma_{th}. \tag{4.7}$$

Hence, the SINR of MTD i is lower than γ_{th} and the SIC fails. □

Since the transmit power of the users are constrained by P^{max}, SNR values are also upper bonded by $\gamma_{\alpha_{max}}$ where α_{max} is the number of eligible devices within a coalition $\mathbb{C}$ that can transmit simultaneously, defined as follows:

Corollary 4.1. Consider a coalition $\mathbb{C}$ of MTDs transmitting on a sub-carrier, we have the following capacity upper bound, for all $i \in \{1, 2, \ldots, |\mathbb{C}|\}$:

$$\alpha_{max} = min\{\ max\{\ \alpha_{max}^i\},\ |\mathbb{C}|\}, \tag{4.8}$$

where α_{max}^i verifies

$$\gamma_{\alpha_{max}^i} \leq \frac{P^{max}|h_i|^2}{\sigma^2} \leq \gamma_{\alpha_{max}^i+1}. \tag{4.9}$$

Here, $\gamma_{\alpha_{max}^i}$ is the highest SNR value that the i-th device can target while $\gamma_{\alpha_{max}}$ is

the highest SNR value that any device can target in the coalition $\mathbb{C}$.

Consequently, α_{max} is the maximum number of MTDs that can transmit simultaneously in the coalition $\mathbb{C}$.

Proof. Since $\gamma_{\alpha_{max}}$ is the highest SNR that a device in the coalition can target, the vector Γ defined in Proposition 4.1 becomes $\Gamma = \{\gamma_1, \cdots, \gamma_{\alpha_{max}}\}$. Meanwhile, we assume that the devices have to target distinct SNR values from Γ, so we have almost α_{max} devices that can target different SNR and thus can be allocated. Since no MTD can reduce its target SNR as proved in Theorem 4.1, we cannot allocate more than α_{max} devices in coalition $\mathbb{C}$. □

In fact, when a user i chooses to target its highest SNR value $\gamma_{\alpha^i_{max}}$, it has to transmit with its highest power level $P_i^{max} = (\gamma_{\alpha^i_{max}}\sigma^2)/|h_i|^2 \leq P^{max}$ in order to achieve a successful transmission. But, the user can target an SNR value $\in \{\gamma_1, \cdots, \gamma_{\alpha^i_{max}}\}$, so that it can transmit with a lower power level. Thus, the device i has α^i_{max} distinct SNR values that it can target, each of which requires a different transmit power coefficient to be achieved. Meanwhile, according to Proposition 4.1, the BS can successfully separate and decode the superimposed signals when the devices target distinct SNR values. Besides, based on Corollary 4.1, there are α_{max} different SNR values, which means that only α_{max} devices can transmit simultaneously to the BS using the sub-carrier associated with the coalition $\mathbb{C}$. Therefore, the members of this coalition cannot all transmit simultaneously at the same time. Consequently, by being a part of the coalition $\mathbb{C}$, there are two constraints: the first constraint faced by each user i, i.e. α^i_{max}, which depends mainly on its channel coefficient and determines its maximum transmit power to be used P_i^{max} whereas the second constraint is imposed to the devices sharing the same coalition, i.e. α_{max}, which represents the coalition reuse factor.

In this regard, upon changing the coalition, α^i_{max} and α_{max} take another values so each user may have more or less chance to transmit. As a result, it may have an incentive to deviate and join another coalition. Thereby, each user has to be able to decide which coalition is preferred to be part of and then determine the transmit power to be used in order to ensure a successful transmission. For this purpose, we propose a Bi-level game theoretical framework composed of a coalition formation game and

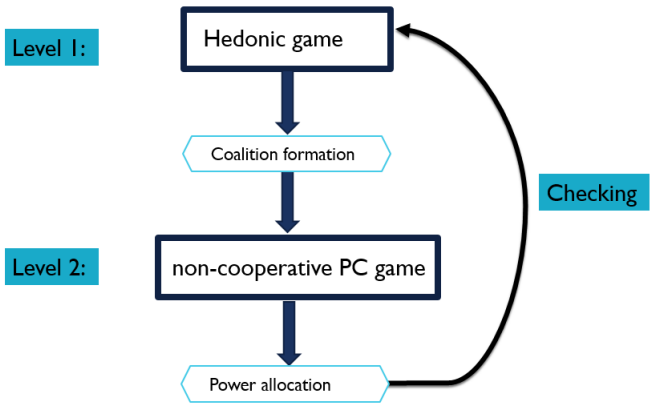

Figure 4.2: Bi-level game illustration

a non-cooperative game to model the behavior of the N devices competing over the K sub-carriers with an eye toward attractively achieving the trade-off between the successful packet transmission rate and the power consumption.

4.3 Bi-level game theoretical framework

In this section, we formulate the resource allocation problem as a Bi-level game by firstly invoking the Hedonic game as a coalition formation game to lay out the partition formation process and then adopting a non-cooperative game to model the power control within each coalition. Particularly, the devices arrange themselves into multiple coalitions using the Hedonic game framework, then the transmit power levels are determined using a non-cooperative game. A graphic illustration of the proposed Bi-level game is depicted in Figure 4.2.

4.3.1 Non-cooperative NOMA-based power control game

We start by modeling the power allocation problem using a non-cooperative game. The aim of this game is to ensure the disparity in transmit power coefficients among users sharing the same group so that they can simultaneously communicate with the BS. The latter can then avail the power difference between the users to decode their signals.

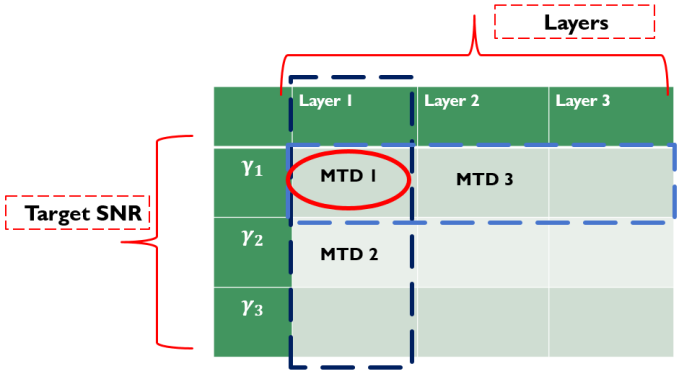

Figure 4.3: Layered system illustration

Definition 4.1. Let $G = (\mathcal{N}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \{U_i\}_{i\in\mathcal{N}})$ be a non-cooperative power control game for the proposed Hybrid NOMA scheme, where

- **Player set** $\mathcal{N} = \{1 \dots N\}$: denotes the player set of MTDs.
- **Set of actions** $\{\mathcal{A}_i\}_{i\in\mathcal{N}} = P_i \times Pt_i$: represents the set of strategies available to a user $i \in \mathcal{N}$. Indeed, a strategy is a couple $\{p_i, pt_i\}$ where $p_i \in [0, P_i^{max}]$ is the transmission power of the device i, and $pt_i \in [0, 1]$ is its transmission probability. In fact, the user i transmits its packets in the current time slot with a transmission probability pt_i. The set of the transmit power of all devices is denoted by $\mathcal{P}$ and the set of their transmission probabilities denoted by $\mathcal{P}t$.
- **Utility function** $\{U_i\}_{i\in\mathcal{N}}$: As we have shown in Chapter 2, the design of the utility function is of prime importance in game theory, as it reflects the desired goal of the game. In this work, the objectives of the devices is to satisfy the SINR requirement while consuming less energy. Thus, we need to define an utility function that can conveniently capture this trade-off. In what follows, we will go into the details of the proposed utility function.

4.3.1.1 The transmit power matrix

We have shown in Corollary 4.1 that in order to enable the BS to perform the SIC successfully, only α_{max} users can transmit simultaneously by targeting distinct SNR values. Thereby, it may not be possible for all the users in the coalition to target different SNR values. Keeping this in mind, we propose a layered system in which we define multiple layers for each target SNR value which means that it is possible

for more than one user to target the same SNR but they have to be in different layers. In this direction, we define the target SNR allocation $\mathbf{\Lambda}$ as a matrix whose rows correspond to the target SNR vector and columns represent the layers, as shown in Figure 4.3. In this case, the MTDs, e.g. MTD 1 and MTD 2, which belong to the same layer, must target different SNR values but they can target the same SNR if they are in different layers, e.g. MTD 1 and MTD 3. Hence, at each layer only one MTD can join the group of devices targeting the same SNR, e.g. MTD 1.

Thereafter, we need to manage the access of the users that target the same SNR value while belonging to different layers. To this end, we introduce the transmission probability for each MTD i targeting a SNR value γ_l as follows

$$pt(i) = \frac{1}{\|\Lambda(l,:)\|_0},$$

where $\|\Lambda(l,:)\|_0$ stands for the number of users targeting γ_l. By doing so, if there is only one user that targets a certain SNR value, its transmission probability $pt(i) = 1$, e.g. MTD 2 in Figure 4.3, otherwise the users targeting the same SNR value, have the same transmission probability, e.g. the transmission probability of MTD 1 and MTD 3 is 0.5.

4.3.1.2 Utility function

In order to pose the power control problem under non-cooperative game settings, we first have to suitably define a utility function. In fact, each user aims to meet the SINR requirement, i.e. reaching the SINR threshold γ_{th}, while consuming less energy. This phenomenon can be concisely quantified by the following utility function

$$U_i(p_i) = \frac{\sum_{k=1}^{K} f(\gamma_{k,i})}{\sum_{k=1}^{K} p_{k,i}}, \tag{4.10}$$

where $\gamma_{k,i}$ is the SINR of the device i on the sub-carrier k and $f(\cdot)$ represents the efficiency function. The latter reflects the probability that a packet is successfully transmitted and it is assumed to be an increasing, continuous, and S-shaped function. Besides, we require that $f(0) = 0$ and $f(\gamma_{th}) = f(\infty) = 1$ to ensure that when $p = 0$

or $\gamma = 0$ the efficiency is null. On the other hand, if the SINR is higher than γ_{th}, the packet is successfully transmitted and then $f(\gamma_{k,i}) = 1$. It is noteworthy that any function that meets the above-mentioned conditions can be adopted. Throughout this thesis, we consider the efficiency function $f(x) = (1-e^{-x})^N$, well-known in the power control games where $N = 100$ is the block length. This utility function, that has bits per joule as units, perfectly captures the trade-off between the SINR requirement and the battery life and is particularly suitable for applications where the energy efficiency is crucial. A more detailed discussion about the utility function and the efficiency function can be found in [103].

In our proposed scheme, we consider a Hybrid NOMA system where each sub-carrier is associated to a NOMA group and every device is allocated to only one coalition. On the other hand, we have introduced above the transmission probability via the layered system in order to manage the activity of users and handle their access to the same sub-carrier. Consequently, we can reformulate the utility function defined in (4.10) as:

$$U_i(p_i, \mathbf{p}_{-i}, pt_i, \mathbf{pt}_{-i}) = pt_i \frac{f(\gamma_i)}{p_i}, \tag{4.11}$$

where $\mathbf{p}_{-i}$ and $\mathbf{pt}_{-i}$ are the transmit power and probability of MTDs other than the device i respectively. Note that there is no utility for the devices if the BS fails in decoding their signals when performing the SIC. Although every device selfishly seeks to find the optimal strategy that gives it the highest utility, the game can reach a stable strategy profile $\{\mathbf{p}^*, \mathbf{pt}^*\}$ in which no device can maximize its utility by individually deviating from its current strategy. This stable profile known as a Nash Equilibrium and represents the solution concept for a non-cooperative game. Based on the Definition (2.1), $\{\mathbf{p}^*, \mathbf{pt}^*\}$ is a Nash equilibrium for the proposed game if

$$U_i(p_i^*, \mathbf{p}_{-i}^*, pt_i^*, \mathbf{pt}_{-i}^*) \geq U_i(p_i, \mathbf{p}_{-i}^*, pt_i, \mathbf{pt}_{-i}^*),\ \forall\{p_i, pt_i\},\ i \in \{1, 2, \ldots, |\mathbb{C}|\}. \tag{4.12}$$

In the following, we detail the algorithm for the proposed non-cooperative power control game and derive the corresponding Nash equilibrium.

4.3.1.3 Power control algorithm

In this section, we derive a power control algorithm to determine the transmit power vector $\mathbf{p}$ as well as the transmission probability vector $\mathbf{pt}$ of the users in each coalition.

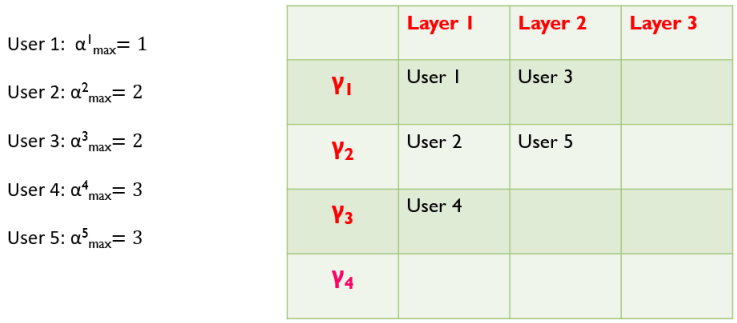

Figure 4.4: Layered system example

Algorithm 1: Power Control Algorithm

Input: the set of devices in the coalition $\mathbb{C}$ who are transmitting on the RB k, and their channels $\mathbf{H}(k,\mathbb{C}) = (h_{k,1},\ldots,h_{k,|\mathbb{C}|})$.

Output: a power vector $\mathbf{p}$ and a transmission probability vector $\mathbf{pt}$

Initialization: $\Gamma^{\mathbb{C}} = \{\gamma_1, \gamma_2, \cdots, \gamma_{\alpha_{max}}\}$.

Sort $(h_{k,1},\ldots,h_{k,|\mathbb{C}|})$ in an increasing order obtaining a vector of the sorted elements and a vector of the arrangement of these elements Π.

$\mathbf{p} = 0_{1\times|\mathbb{C}|}$, $\mathbf{pt} = 0_{1\times|\mathbb{C}|}$ and $\mathbf{\Lambda} = 0_{\alpha_{max}\times|\mathbb{C}|}$ is target SNR allocation.

for $pos = 1 : |\mathbb{C}|$ **do**
- $i = \Pi(pos)$, $l = 1$
- **while** $(\mathbf{p}(i) = 0)$ **do**
  - $m = \min\{\ n \mid n \leq \alpha^i_{max}$ and $\|\Lambda(n,:)\|_0 < l\ \}$
  - **if** $(m \neq \varnothing)$ **then**
    - $\Lambda(m,l) = i$
    - $\gamma_m = \Gamma^{\mathbb{C}}(m)$
    - $\mathbf{p}(i) = \frac{\gamma_m \times \sigma^2}{|h_{k,i}|^2}$
  - **else**
    - $l = l + 1$

for $n = 1 : \alpha_{max}$ **do**
- **for** $l = 1 : |\mathbb{C}|$ **do**
  - **if** $(\Lambda(n,l)! = 0)$ **then**
    - $i = \Lambda(n,l)$
    - $\mathbf{pt}(i) = \frac{1}{\|\Lambda(n,:)\|_0}$

where α_{max} and α^i_{max} are defined in Corollary 4.1 and $\|.\|_0$ is the L_0 norm, which is the number of non-zero elements.

Indeed, the proposed approach enables each user to find the appropriate target SNR which ensures that its packets will be successfully transmitted with the lowest energy consumption. The proposed algorithm is outlined in Algorithm 1 and an example of the target SNR allocation matrix is given in Figure 4.4.

The power control Algorithm 1 starts by sorting the MTDs according to their channel gains so that the device which has the weakest gain is the first one that targets the lowest SNR in the first layer. Then, each device i seeks to target the lowest SNR that is not already taken in the current layer l. If it does not find a SNR value m that gives $\gamma_m \leq \gamma_{\alpha^i_{max}}$ and is not allocated to another device, the user i selects the next level and checks if it is possible to transmit on this layer by targeting a lower SNR value which is not higher than $\gamma_{\alpha^i_{max}}$. Once the user determines its target SNR, it calculates its transmit power level and its transmission probability. By doing so, each device requires only a small amount of energy to successfully transmit its signals. Hence, we ensure that the proposed power control approach is energy efficient.

Proposition 4.2. The power allocation scheme, resulting from Algorithm 1, is a Nash equilibrium for the proposed power control NOMA-based game.

Proof. We assume that every MTD in the coalition $\mathbb{C}$ targets the lowest SNR γ_m in the current layer l ($m = \min\{n | n \leq \alpha^i_{max}$ and $\|\Lambda(n,:)\|_0 \leq l\},\ i \in \mathbb{C}$). If a device i aims to increase its utility by changing its target SNR γ_m, it may chooses a higher SNR target $\gamma_p > \gamma_m$ which in turn results in a higher transmit power and thus a lower utility value. On the other hand, since the user has chosen the lowest target SNR value γ_m that has not been taken by a weaker user, if it decides to choose a lower SNR target $\gamma_q < \gamma_m$, it has to join the $\{l+1\}$-th layer and then share the transmission time with the devices in this layer. Thus it gets a lower transmission probability which results in decreasing its utility. □

4.3.2 Hedonic game coalition formation game

In this section, we focus on how to construct an appropriate coalitional structure of the set of players with the aid of the Hedonic game [27]. Such a game is one of the most prominent coalition-formation games that can provide an autonomous and a distributed cooperative model allowing the players to self-organize into an optimal partition.

4.3.2.1 Game settings

Definition 4.2. *A Hedonic game* $< \mathcal{N}, \succeq_i >$ is formulated by a finite set of devices $\mathcal{N}$, and a set of preference profiles $\{\succeq_1, \succeq_2, \cdots, \succeq_\mathcal{N}\}$. For each player $i \in \mathcal{N}$, $\succeq_i$ specifies its preference relation and is defined as a reflexive, complete and transitive binary relation on the set of coalitions that this player can possibly be a part of, i.e. $\mathcal{N}_i = \{\mathbb{C} \subseteq \mathcal{N} : i \in \mathbb{C}\}$.

We say that a coalition $\mathbb{C}$ is preferred to the MTD i than a coalition $\mathbb{C}'$ if $U_i(\mathbb{C} \cup \{i\}) \geq U_i(\mathbb{C}' \cup \{i\})$. Note that the utility value of each device is related to its target SNR that depends on the strategies that the other users in the coalition take jointly. In doing so, it evaluates its utility over its coalition and not on the whole structure. Thus, according to Definition (2.4), the considered coalition formation game is called the Hedonic game with a Non-Transferable Utility (NTU).

Definition 4.3. *A partition*, denoted $\Psi = \{\mathbb{C}_1 \dots \mathbb{C}_k\}$, is a set of disjoint subsets of $\mathcal{N}$ involving all devices of $\mathcal{N}$.

Definition 4.4. A partition $\Psi = \{\mathbb{C}_1 \dots \mathbb{C}_k\}$ is Nash stable if no MTD can improve its utility by moving unilaterally to another existing coalition.

Particularly, the Nash stability is the strongest stability's notion in the Hedonic game. Interestingly, similar to the Nash Equilibrium concept in the non-cooperative games, the Nash stability consists in ensuring that no player has an incentive to deviate from its current coalition.

4.3.2.2 Generating Nash stable coalition structures

Let Ψ be a partition of a set of K coalitions where each coalition is allocated an orthogonal RB. Each member of the partition Ψ represents a coalition. Each device belongs to only one coalition in a given time slot. Thus, we have the following steps:

- Firstly, each user i sorts its channel gains over the K RBs in a decreasing order. Then, it joins the cluster corresponding to its best channel.
- Second, the device determines its power allocation coefficient and its transmission probability by executing the power allocation algorithm (Algorithm 1). Then, it calculates its own utility value U_i.

Algorithm 2: Nash stable clustering (NSC)

Input: the set of devices in the cell $\mathbb{M}$ and their channels $\mathbf{H}$, the K RBs.
Output: a partition Ψ.
Initialization: $\mathbf{\Psi} = 0_{K\times|\mathbb{M}|}$, $U = 0_{1\times|\mathbb{M}|}$.
Every MTD sorts $(\mathbf{H}(:,i)) = (h_{1,i},\ldots,h_{K,i})$ using the first coordinate in a decreasing order obtaining the sorted elements of $(\mathbf{H}(:,i))$ and the arrangement of these elements in Π_i.
for $i = 1 : |\mathbb{M}|$ **do**
    find first j such that $\mathbf{\Psi}(\Pi_i(1),:) = 0$
    $\mathbf{\Psi}(\Pi_i(1),j) = i$
for $k = 1 : K$ **do**
    $[\mathbf{p},\mathbf{pt}] = PC(\ \mathbf{\Psi}(k,:)\ ,\ \mathbf{H}(k,\mathbf{\Psi}(k,:))\ )$
    for $i = 1 \in \mathbf{\Psi}(k,:)$ **do**
        $U(i) = \mathbf{pt}(i)\frac{f(\gamma_{th})}{\mathbf{p}(i)}$
isFinal=False
while *!isFinal* **do**
    isFinal=True
    for $i = 1 : |\mathbb{M}|$ **do**
        for $k = 1 : K$ **do**
            if $i \notin \mathbf{\Psi}(k,:)$ **then**
                $\Psi_{tmp}(k,:) = \Psi(k,:) \cup \{i\}$
                $[\mathbf{p},\mathbf{pt}] = PC(\ \Psi_{tmp}(k,:),\ \mathbf{H}(k,\mathbf{\Psi_{tmp}}(k,:))$
                if $\mathbf{pt}(i)\frac{f(\gamma_{th})}{\mathbf{p}(i)} > U(i)$ **then**
                    $U(i) = pt(i)\frac{f(\gamma_{th})}{p(i)}$
                    $\mathbf{\Psi} = \mathbf{\Psi}\backslash\{i\}$
                    find first j such that $\mathbf{\Psi}(k,:) = 0$
                    $\mathbf{\Psi}(k,j) = i$
                    isFinal=False

- Since each user aims to maximize its utility, it invokes the Algorithm 1 for each RB k in order to determine if it is better to deviate from its current coalition. If this deviation gives it a better utility, the MTD joins the new coalition.
- Finally, if the devices are no longer interested to leave their coalitions, the algorithm terminates.

Indeed, upon playing the Hedonic game, the players are able to autonomously decide whether to leave or join coalitions with an effort to maximize their utility values. Thereby, the proposed approach lays out the partition formation process and enables the players to reach an optimal coalitional structure. A summary of the developed algorithm based on the Hedonic game is given in Algorithm 2. Let us focus now on the stability of the proposed game.

Theorem 4.2. The partition resulting from Algorithm 2 is Nash Stable.

Proof. Since the algorithm terminates when there is no incentive for any user to leave its current coalition, we deduce from Definition (4.4) of the stability that the proposed partition is Nash Stable. □

4.4 Simulation Results

In this section, we present a comprehensive Matlab-based simulation of the proposed Bi-level game, described in the previous sections. We compare the proposed Bi-level theoretical framework with the NM-ALOHA scheme proposed in [64] and the game-theoretic approach NM-ALOHA game, introduced in [100]. We consider a scenario composed of $N = 500$ MTDs independently deployed according to the homogeneous PPP process of density $\omega_N = 0.01$.

4.4.1 Packet success rate

In this section, we focus on the probability that a packet is successfully transmitted using the proposed Bi-level game theoretical technique versus the NM-ALOHA and the NM-Aloha game. In Figure 4.5, we show the packet transmission rate with success as a function of the number of sub-carriers K and the number of MTDs N. As we can clearly observe, the packet success rate decreases drastically when the

number of devices increases for both NM-ALOHA and NM-ALOHA game, while the proposed technique still achieves acceptable performances. For example, when $K = 9$ sub-carriers and $N = 320$ MTDs, our proposed game boosts the packet success rate to 0.9185 compared to 0.0450 when NM-ALOHA is applied and 0.2402 when NM-ALOHA game is played. This owing to the fact that as the number of the users in the system N increases, the network becomes denser which in turn results in higher interference effects that both NM-ALOHA and NM-ALOHA game can not handle even with the use of SIC. By contrast, our proposed approach is able to efficiently face the resulting interference impacts and thus allowing much more devices to achieve a successful packet transmission, reaching up $335,088\%$ of improvement compared to NM-ALOHA game when $N = 500$ and $K = 9$. This good performance can be explained by the fact that the proposed power allocation approach has appropriately taken into account the SINR requirement of the users which leads to successful transmissions of an important number of them.

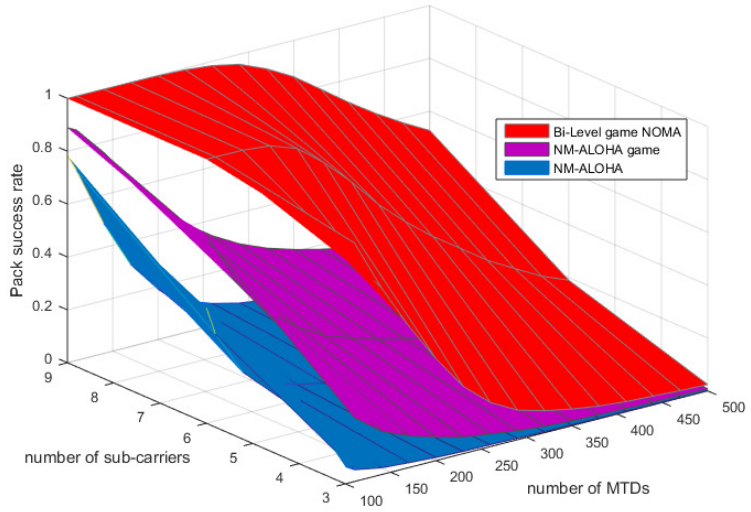

Figure 4.5: Packet transmission rate with success for different N and K.

4.4.2 Average throughput

This section is devoted to the performance comparison in terms of the average throughput between our proposed Bi-level game, the NM-ALOHA and NM-Aloha game. Figure 4.6 illustrates the average throughput as a function of the number of sub-carriers

K and the number of MTDs N. As it was expected, the average throughput of NM-ALOHA and NM-ALOHA game is higher than the proposed technique when there is a low number of MTDs in the system. In fact, Figure 4.7 shows that for $K = 10$ sub-carriers, when $N < 170$ MTDs, NM-ALOHA and NM-ALOHA game achieve a higher throughput, otherwise our game significantly outperforms them. More precisely, our proposed Bi-level game yields a significant performance improvement in the throughput of $268,502\%$ against the NM-Aloha game when $N = 500$. This is mainly due to the fact that when the system is sparse, NM-ALOHA and NM-ALOHA game allocate to MTDs more than the required capacity. In addition, when $N < 350$ MTDs, the average throughput of our game increases as N increases while it decreases for the case of NM-ALOHA and NM-ALOHA game.

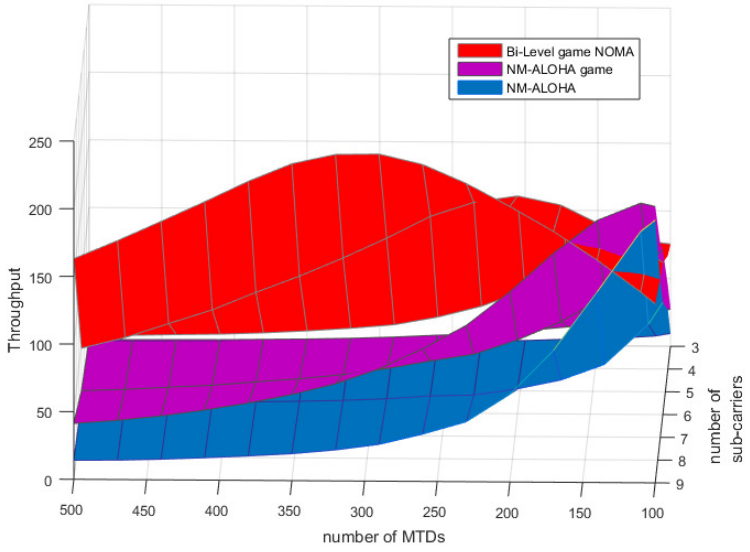

Figure 4.6: Average throughput of MTDs for different N and K.

4.4.3 Average utility

The proposed energy efficiency-based utility function has units of bits/joule and then measures the total number of reliable bits transmitted per joule of the energy consumed. We investigate now the effect of the number of sub-carriers K on the energy efficiency of the system. Indeed, Figure 4.8 shows the average utility as a function of the number of sub-carriers K for $N = 500$ MTDs. It can be observed from this figure that as K increases, the average utility increases too. As expected, our pro-

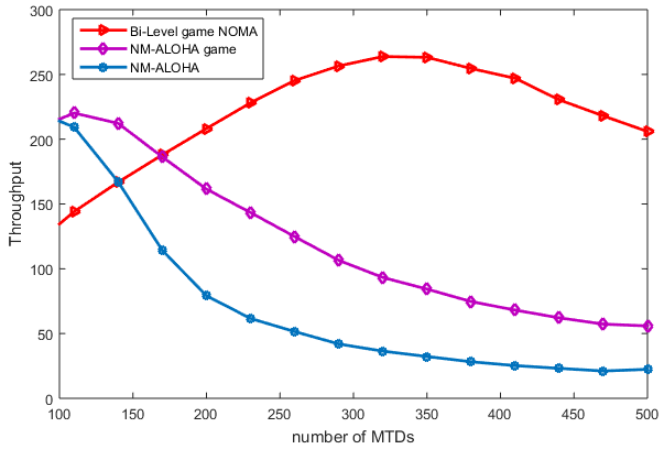

Figure 4.7: Average throughput versus N with $K = 10$.

posed approach results in a significant improvement in the average utility, and thus in the energy efficiency, compared to the NM-ALOHA and the NM-ALOHA game. Interestingly, this result is somehow intuitive since the proposed Bi-level game technique allows each device to deliver its packets with the lowest power that enables it to perform a successful transmission, while the other schemes may allocate to the user a higher transmit power level which in turn yields in a reduced utility compared to our formulated game. Particularly, the achieved result is highly revealed when there are $N = 500$ devices sharing $K = 10$ sub-carriers. In this case, our proposed approach greatly enhances the average utility reaching up $427,265\%$ compared to the NM-ALOHA game. This is owing to the fact that when the devices are divided into a larger number of coalitions, the system becomes sparser, resulting in a higher transmission probability and a lower transmit power for each user. Hence, the proposed Bi-level game enables the devices to satisfy the SINR requirement with the lowest power levels and thus achieve higher utility values.

4.4.4 Energy consumption

Now, we highlight the energy consumption of the proposed Bi-level game compared to the other existing techniques, the NM-ALOHA and the NM-ALOHA game, under different scenarios composed of different numbers of devices N and sub-carriers K. To this end, we adopt the model of [104] to evaluate the energy consumption. According

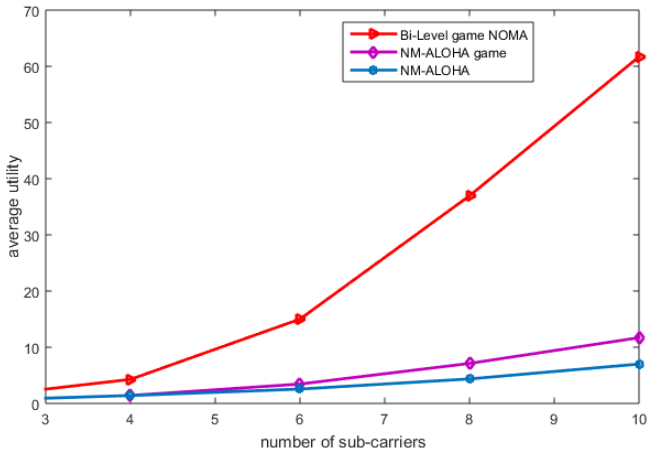

Figure 4.8: Average utility versus K with $N = 500$.

to this model, a device i consumes the following energy to transmit a L-bit message:

$$\begin{cases} L \times E_{elec} + L \times \epsilon_{fs} r_i^2 & \text{if } r_i < d_0 \\ L \times E_{elec} + L \times \epsilon_{amp} r_i^4 & \text{if } r_i \geq d_0 \end{cases} \tag{4.13}$$

Depending on whether the distance r_i between MTD i and the BS is higher than threshold distance r_0 or not, either the free space ϵ_{fs} (r_i^2 power loss) or the multi path fading ϵ_{amp} (r_i^4 power loss) channel model is used. For making this dissertation self-contained, we review here the above parameters from [104] in Table 4.1. Figure 4.9 illustrates the energy consumption for the different techniques. As we can clearly see, the devices consume far less energy when using the proposed technique than the other schemes. In fact, the energy consumption gain achieved upon adopting the Bi-level game for $N = 500$ users competing on $K = 9$ RBs is about $94,198\%$ and $96,693\%$ compared to the NM-ALOHA game and NM-ALOHA respectively. This owing to the fact that our Bi-level game seeks to allocate to the users the appropriate transmit power, i.e. the lowest power levels that enable them to successfully transmit their messages. For this purpose, the proposed game allows each user to deviate from its coalition if this deviation yields in a better utility allowing it to conveniently address the SINR requirement with a lower energy consumption. Consequently, it can be concluded from this figure that the proposed power control approach is energy efficient.

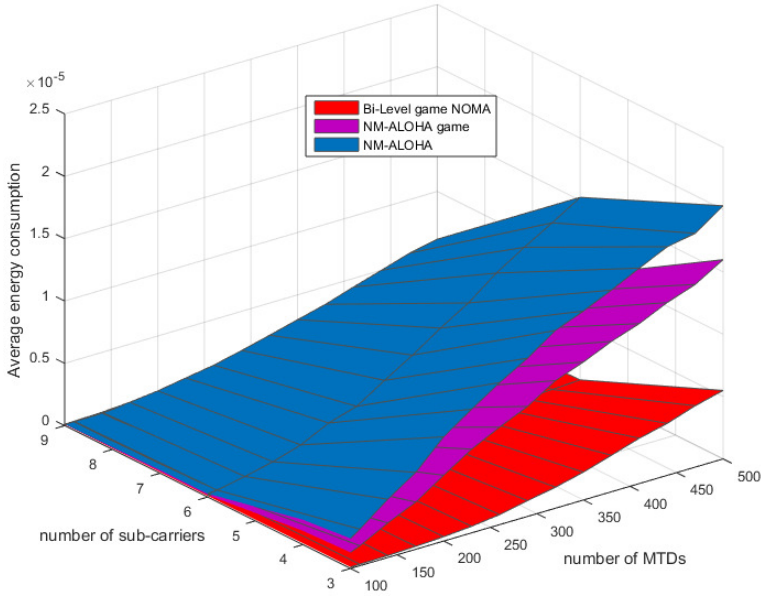

Figure 4.9: Energy consumption for different N and K.

Table 4.1: Energy Consumption Parameters

Symbol	Description	Value
d_0	Threshold distance (m)	87
E_{elec}	Electronic energy (nJ/bit)	50
ϵ_{fs}	Power amplification for the free space (pJ/bit/m2)	10
ϵ_{amp}	Multi-path fading power amplification (pJ/bit/m2)	0.0013

4.5 Conclusion

In this chapter, we have proposed a Bi-level game framework to derive two algorithms for MTDs in Hybrid NOMA networks. First, the MTDs self-arrange into coalitions using the Hedonic game framework. Then, the members of each coalition invoke the non-cooperative power control game to determine the power levels to be used for sending their packets over one RB and in a non-orthogonal manner. Simulation results demonstrate that our proposed scheme enables MTDs to meet the SINR requirement and thus achieve good packet success rates with the lowest power levels compared to the other existing techniques NM-ALOHA and NM-ALOHA game. Hence, our

proposed Bi-level game has efficiently solved the joint problems of the user grouping and the power allocation while striking an attractive trade-off between the successful packet transmission rate and the energy consumption.

In the next chapter, a denser network is envisaged which in turn requires an advanced game framework in order to be suitably addressed. In such a model, we consider that each user is called upon to conform the collective behavior of its opponents instead of participating in one-to-one interactions. Keeping this in mind, we adopt the mean field theory as an effective optimization tool to deal with large-scale networks.

Chapter 5

Mean Field Game-Theoretic Framework for Distributed Power Control in Hybrid NOMA

Contents

5.1 Introduction

Driven by the scarcity of resources and the ubiquity of IoT systems, the 6G networks have to handle the heterogeneous and the ever-increasing number of devices [105]. MTC represents an attractive paradigm to enable the IoT to be part of future cellular networks. Worryingly, although the generated traffic by each device is relatively small and sparse, the BS may face a huge traffic when the number of MTDs increases. Hence, providing massive connectivity and ensuring a higher spectral efficiency for large-scale MTDs networks is very challenging.

Attracted by its appealing features, numerous research efforts have been exerted to investigate NOMA in different scenarios [13, 54, 68, 106–112]. Nevertheless, almost all of them have either considered a limited number of users in the system and per group or investigated semi-grant free scenarios in which the BS allocates resources to devices at the cost of scalability. The main reason for this limitation is that having more users in each group results in a higher interference level and challenges the way the BS assigns the resources to these users. Hence, our goal in this chapter is to design an NOMA-based approach for a large number of devices without increasing the complexity at the BS side.

In the previous chapter, we have formulated a Bi-level game theoretical framework composed of a non-cooperative game underlying a cooperative coalitional game, namely the Hedonic game and we have derived two algorithms which lead to a Nash-Stable partition. However, the classical game generally focuses on characterizing the interaction between each device and every other devices in the system. Besides, as it has been pointed out in [113], the Nash equilibrium analysis usually requires solving a large number of equations, which results in an inherent mathematical complexity. By doing so, applying such a game can barely be useful when a dense network is envisaged. To overcome this issue, we turn our attention in this chapter to the Mean Field Theory.

Particularly, in the MFG, each device is called upon to focus on how to deal effectively with the collective behavior of its opponents, rather than being concerned with the specific individual strategy of each. Here, the collective effect faced by the devices represents the mean field and stands for the distribution of the system state over the user set [35, 36, 113, 114]. Typically, the analysis of a MFG is mainly conducted

through two coupled equations, namely the HJB and the FPK equations. The former equation characterizes the interactions between the players and the mean field, and then allows each player to make its own decision, whereas the latter equation rules the evolution of the mean field based on the players' decisions.

MFG has sparked a considerable interest in suitably designing a distributed power control for densely deployed wireless networks [115–120]. Meanwhile, some contributions have mainly focused on the interplay between the NOMA approach and the mean field theory [121–124]. In [121], MFG has been exploited in order to meet the trade-off between the QoS requirements and the energy consumption for the CD-NOMA scheme in a mMTC scenario. [122] and [123] have proposed a MFG approach to model the collective behavior of multi-user scenarios in mobile edge computing systems. In these works, the users are divided into NOMA clusters based on the difference in their channel gains, then the resource allocation problem is formulated as a MFG. The authors have resorted to a deep RL algorithm to solve the game and obtain the Mean Field Equilibrium (MFE). Whilst in [124], the authors have adopted the MFG to derive a distributed power control policy for NOMA-assisted unmanned aerial vehicle networks.

In this chapter, we aim to propose a mean field game-theoretic framework in the context of a dense Hybrid NOMA system with an eye toward efficiently modeling power allocation between devices in different NOMA clusters. The main contributions can be summarized as follows:

- We consider a Hybrid NOMA scenario in which the coalitions of MTDs will be established and allocated orthogonal RBs so that the members of each coalition use one RB to transmit their packets in a non-orthogonal manner.

- We consider that the BS is not concerned with the power control and broadcasts only a limited feedback to users. Thereby, we are able to mitigate the performance drop observed with almost all existing grant-free approaches, especially in dense scenarios.

- We first formulate the power control problem of MTDs underlying uplink NOMA networks as a differential game. Then, we extend the proposed game into the MFG which interestingly alleviates the mathematical complexity of the scheme.

- We conduct the analysis of the formulated MFG through the HJB and FPK equations and derive a distributed iterative algorithm in an effort to approach the MFE.
- Unlike what is usually examined in classical games, the proposed MFG allows the devices to adjust their transmit power in response to brief information received from the BS, instead of worrying about the actions of their opponents.
- We analyze the equilibrium and prove the convergence of the proposed power control algorithm.
- We provide numerical simulations that assess the potential performance gains offered by the proposed approach with several scenarios of number of devices and number of sub-carriers.

The rest of this chapter is organized as follows. The next section is devoted to the description of the system model and the assumptions that we consider throughout this chapter. Section 5.3 formulates a differential game to address the corresponding power control problem. Modeling the differential game as a MFG is depicted in Section 5.4. Then, in Section 5.5, we propose a finite difference method to solve the HJB and FPK equations, and derive a distributed algorithm to iteratively reach the MFE. To illustrate the performance of the proposed technique, simulation results are given in Section 5.6. Finally, the chapter is concluded in Section 5.7.

5.2 System Model

5.2.1 Network model

Consider the same uplink NOMA network introduced in Chapter 4, composed of N devices attempting to communicate with a BS across K available RBs. We consider that the available bandwidth is divided into K orthogonal sub-carriers. Specifically, we investigate in this chapter a dense scenario in which $N >> K$.

Let us focus on a coalition $\mathbb{C}$ of devices that transmit their packets non-orthogonally using one sub-carrier. We assume that each MTD joins the coalition corresponding to its best channel. Note that under a Hybrid NOMA scenario, each device experiences interference effects from its coalition members. Thus, the inter-group interference is

not considered when the BS invokes the SIC to split the superimposed signals received from each group and then retrieve the desired information of each user within that coalition. For the sake of notation simplicity, we omit the sub-carrier index in the following.

In Chapter 4, we have investigated the upper bound of the allocation capacity. In particular, we have proven that if each device uses a distinct SNR target, there are almost α_{max} devices in a given coalition that can access the associated RB to simultaneously transmit their packets. Meanwhile, with the aim of managing the activity of the devices, we propose in the following an access probability denoted by p_t.

Proposition 5.1. Assuming that the BS successfully decodes a MTD's signal with a probability P_s, this probability can be maximized when $p_t = \frac{\alpha_{max}}{N}$.

Proof. Since P_s is the probability of successful decoding of device's signal, it can be written as

$$P_s = p_t(1 - \frac{p_t}{\alpha_{max}})^{N-1}.$$

Let us now derive this probability with respect to p_t. Then, we obtain

$$(1 - \frac{p_t}{\alpha_{max}} - \frac{N-1}{\alpha_{max}}p_t)(1 - \frac{p_t}{\alpha_{max}})^{N-2} = 0.$$

Thus, we end up with $p_t = \frac{\alpha_{max}}{N}$. Consequently, if $p_t = \frac{\alpha_{max}}{N}$, P_s is maximized. □

In the remainder of this chapter, we refer to α instead of α_{max} for the sake of notation simplicity. At each time slot, the user activity is controlled by the access probability $p_t = \frac{\alpha}{N}$. Thereby, we have in average α active devices. Now, we consider a dense deployment scenario in which a large number of users are involved. Thus, it becomes increasingly challenging to ensure the successful transmission of different devices, especially at a high network load. Therefore, in an effort to handle the massive connectivity of the devices, we formulate the resource allocation problem as a MFG framework. To this end, we need first to pose our proposed scenario in the context of the differential game theory. We assume that the number of users in each cluster is known by the users and is given as a feedback from the BS. Actually, we consider that this information can be learned from the interference feedback or sent back by the BS. As, we are studying very dense scenarios, this value can be estimated by the

BS and fed back to the users less frequently than the value of the interference level, which will be introduced in section 5.4.

5.3 Differential game model for power control

In order to pave the way for the MFG framework, we start in this section by adopting the differential game theoretical framework to model the power allocation problem.

Definition 5.1. The differential power control game for the proposed approach is defined by $G = (\mathcal{N}, \{\mathcal{P}_i\}_{i\in\mathcal{N}}, \{\mathcal{S}_i\}_{i\in\mathcal{N}}, \{\mathcal{Q}_i\}_{i\in\mathcal{N}}, \{\ U_i\}_{i\in\mathcal{N}})$, where $\mathcal{N} = \{1 \dots N\}$ is the set of players which are the MTDs in our case. The set of transmit power levels of a device i is represented by $\mathcal{P}_i$ which can be used to transmit its information. In addition, S_i is the state space of the user i while $\mathcal{Q}_i$ is the power control policy to be determined by the device i in order to maximize its own utility denoted by U_i.

5.3.1 State space

We aim in this chapter to establish a Hybrid NOMA system by arranging the users into coalitions and making each of them able to react in response to the collective behavior of the other players sharing the same group. Thus, we are interested in taking into consideration the proportion of users in each coalition. To this end, we define the state of each player i as the combination of its channel coefficient h_i and the sub-carrier k_i associated with its coalition. Therefore, the state of device i at each time t is given as $s_i(t) = \{h_i(t), k_i\}$.

Proposition 5.2. In order to simulate the fading process in an efficient way, we adopt the Jakes' channel model for Rayleigh fading channels, given in [125], as follows:

$$g(t) = \begin{pmatrix} g_R(t) \\ g_I(t) \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{B}} \sum_{n=1}^{B} a_n \cos\left(\beta_n t + \phi\right) \\ \frac{2}{\sqrt{B}} \sum_{n=1}^{B} b_n \cos\left(\beta_n t + \phi\right) \end{pmatrix}, \tag{5.1}$$

where $a_n = \cos(\psi_n)$, $b_n = \sin(\psi_n)$ and $\beta_n = \beta_d \cos\omega_n$. Besides, $\omega_n = (2\pi n - \pi + \theta)/(4B)$, where B is the number of sinusoids and ψ_n, θ and ϕ are independent random variables with a uniform distribution over $[-\pi, \pi)$ for all n.

Let us now focus on the state evolution governed by the state equation given as follows:

Definition 5.2. (State equation): For every device, the evolution law of its state is defined by the following differential equation:

$$ds(t) = dh(t) = \frac{dg(t)}{l} = \begin{pmatrix} -\frac{2}{l\sqrt{B}} \sum_{n=1}^{B} a_n \beta_n \sin(\beta_n t + \phi) dt \\ -\frac{2}{l\sqrt{B}} \sum_{n=1}^{B} b_n \beta_n \sin(\beta_n t + \phi) dt \end{pmatrix} = \begin{pmatrix} -q w_{1,n}(t) dt \\ -q w_{2,n}(t) dt \end{pmatrix}, \tag{5.2}$$

where $q = \frac{2}{l\sqrt{B}}$, $w_{1,n}(t) = \sum_{n=1}^{B} a_n \beta_n \sin(\beta_n t + \phi)$ and $w_{2,n}(t) = \sum_{n=1}^{B} b_n \beta_n \sin(\beta_n t + \phi)$.

5.3.2 Utility function

In game theory, the design of the utility function is crucial, as it catches how satisfied a user is when playing the game. Indeed, a packet is successfully decoded when the device achieves an SINR higher than γ_{th}. On the other hand, in the context of MTDs having a limited number of packets to transmit periodically, if a given device reaches a high SINR, it obviously consumes a lot of energy uselessly. In this regard, in our work, the objectives of the players are to meet SINR requirement and to reduce the power consumption as much as possible. With this in mind, we adopt the following utility function to adequately address the above trade-off:

$$U_i(p_i, \mathbf{p}_{-i}) = \frac{f(\gamma_i)}{p_i}, \tag{5.3}$$

where $\mathbf{p}_{-i}$ is the transmit power of all the MTDs except i. The efficiency function, which is represented by $f(\cdot)$, reflects the packet success rate. It is an increasing and continuous function that has a sigmoidal shape. For more details on the efficiency function please refer to Chapter 4.

5.3.3 Optimal control problem

Since each device aims to maximize its own utility defined by (5.3), it tries to find the optimal power control policy $\mathcal{Q}_i^*(t)$ to be used at time $t \in [0, T]$. Thus, for each device $i \in \mathcal{N}$, we formulate the general optimal control problem as follows

$$\mathcal{Q}_i^*(t) = \underset{p_i(t)}{\operatorname{argmax}} \, \mathrm{E}\left[\int_0^T U_i(t, s_i(t))\, dt\right], \tag{5.4}$$

then, the maximum utility is determined by the following value function

$$v_i(t, s_i(t)) = \max_{p_i(t)} \mathrm{E}\left[\int_t^T U_i(\tau, s_i(\tau))\, d\tau\right], \quad t \in [0, T]. \tag{5.5}$$

Based on the Bellman's optimality principle and the optimal control theory [113], the above value function is the solution to a partial differential equation, namely a HJB equation. Hence, solving the HJB equation yields obtaining the optimal power control policy $\mathcal{Q}_i^*(t)$ that gives the maximum utility for each device i. We have expressed the HJB equation as (2.5) in chapter 2 and here we derive the Hamiltonian as follows

$$\begin{gathered} H(s_i(t), p_i(t), \nabla_s v_i(t, s_i(t))) = \\ \max_{p_i(t)}[U_i(t, s_i(t), p_i(t)) + \nabla_s v_i(t, s_i(t)) \frac{\partial s_i(t)}{\partial t}]. \end{gathered} \tag{5.6}$$

5.3.4 Nash equilibrium

Let us introduce the Nash equilibrium in the context of differential game as follows

Definition 5.3. A power control profile $Q^*(t) = [Q_1^*(t), \ldots, Q_N^*(t)]$, where $Q_i^*(t) = p_i^*(t)$, $i \in \mathcal{N}$, is the Nash equilibrium of the differential game G if and only if $Q_i^*(t)$ is the optimal feedback for the control problem

$$Q_i^*(t) = \underset{p_i(t)}{\operatorname{argmax}} \, \mathrm{E}\left[\int_0^T U_i(p_i(t), \mathbf{p}_{-i}^*(t))\, dt\right], \tag{5.7}$$

subject to:

$$s_i(t) = \{h_i(t), k_i\} \text{ and } ds_i(t) : \text{ defined in Definition (5.2).} \tag{5.8}$$

At the Nash equilibrium, no device still has an incentive to unilaterally change its current power control policy with the aim of further improving its utility. According to [126], the Nash equilibrium exists for the differential game if the HJB equation (2.5) related to each device can be solved.

Proposition 5.3. The Nash equilibrium exists for the proposed differential game G.

Proof. According to [127], the HJB equation has a solution if the Hamiltonian is smooth. Given the utility function defined in (5.3), we can write the first derivative of the Hamiltonian function (5.6) with respect to $p_i(t)$ as:

$$\begin{aligned}\frac{\partial H}{\partial p_i(t)} &= \frac{\frac{\partial \gamma_i(t)}{\partial p_i(t)} p_i(t) f'(\gamma_i(t)) - f(\gamma_i(t))}{p_i(t)^2} \\ &= c\frac{f'(\gamma_i(t))}{p_i(t)} - \frac{f(\gamma_i(t))}{p_i(t)^2} \\ &= Z(p_i(t)),\end{aligned} \tag{5.9}$$

where $c = \frac{\partial \gamma_i(t)}{\partial p_i(t)}$. We have $f(x) = (1-e^{-x})^L$, so it is infinitely differentiable as well as $f'(x) = L(1-e^{-x})^{(L-1)}$. On the other hand, as aforementioned, when p_i the transmit power of the device i is equal to 0, the efficiency is null so that $U_i = 0$ and there is no transmission, hence the device can determine its utility function as long as $p_i > 0$. Consequently, $Z(p_i(t))$ is infinitely differentiable. Therefore, for any $n > 1$, $\frac{\partial H^n}{\partial p_i(t)^n}$ exists and the Hamiltonian function has derivatives of all orders, i.e. its smoothness is ensured. As a result, it can be concluded that the Nash equilibrium exists for the proposed game. □

Although it is possible to prove that the Nash equilibrium exists, finding the equilibrium requires solving N HJB equations simultaneously, which is complicated and even untractable for a dense network. On the other hand, when a large number of users are involved, the effect of a single user's action on the other players becomes negligible but the impact of the mass on each player is important and can be modelled as the collective effect or the mean field. Consequently, in the presence of a large population, the differential game can be approximated by an equivalent game so-called MFG.

5.4 Mean Field Game analysis for power control

Typically, in the presence of a large number of devices, the mean field framework establishes when the fulfillment of the following assumptions is guaranteed [114]:

- Rationality of the devices: each of which is assumed to act rationally and independently while optimizing its own utility.

- Interaction between a player and the mean field: this assumption is based on the manner the interaction among devices is investigated. In the context of the MFG, each player is called upon to interact with the mean field rather than participate in one-to-one interactions.

- Continuum of the number of players: in the presence of a large number of MTDs, we can model such a large population as a continuum of players.

- Indistinguishability (or Interchangeability of the states) [128]: this assumption relies on players' anonymity which means that the game's outcome is not impacted by any permutation of states between the players and the state evolution of the players does not depend on any particular user. Hereafter, we can omit the subscript i which referred to the devices.

5.4.1 Mean field interference

Basically, in the context of MFG, each device interacts with the collective behavior of its opponents in order to make its own decision. In our case, this mass behavior is captured by the Mean Field Interference (MFI). Thus, similar to [116], we can define the MFI as the weighted sum of the players whose are competing over the sub-carriers. In addition, since the activity of the users is controlled by the access probability $p_t = \frac{\alpha}{N}$, the interference perceived by any device can be given as follows:

$$I_i(t) = \frac{\alpha}{N} \sum_{j \neq i} |h_j(t)|^2 p_j(t). \tag{5.10}$$

Considering the mean field definition in (2.4), we can reformulate the aggregated interference to any device as follows:

$$\begin{aligned} I_i(t) &= \frac{\alpha}{N}\sum_{j=1}^{N}|h_j(t)|^2 p_j(t) - \frac{\alpha}{N}|h_i(t)|^2 p_i(t) \\ &= \alpha\int |h(t)|^2 p(t) M(t,s)dh - \frac{\alpha}{N}|h_i(t)|^2 p_i(t). \end{aligned} \tag{5.11}$$

Since $p_i(t)$ has reasonably low values compared to N, we may assume that

$$\lim_{N\to+\infty}\frac{\alpha}{N}|h_i|^2 p_i(t) = 0. \tag{5.12}$$

Therefore, the mean interference term can be expressed as:

$$I^{mean}(t) = \lim_{N\to+\infty} I_i(t) = \alpha\int |h(t)|^2 p(t) m(t,h)dh. \tag{5.13}$$

Indeed, by attempting to access the channel, each player uploads implicitly its local information, including its state and its transmit power, to the BS. This latter calculates the mean field (2.4) as well as the MFI (5.13) based on the received information from the participating devices. Then, the BS broadcasts the obtained value of the MFI, on a given channel, to the devices of the corresponding coalition. Once each device receives feedback information, it estimates its interference level from its standpoint as follows:

$$\tilde{I}_i(t) = \left(1 - \frac{r_i}{R}\right) I^{mean}(t), \tag{5.14}$$

where r_i is the distance from the device i to the BS and R is the radius of the cell. Since, in the game under consideration, we are dealing with the average effects of the players, each MTD perceives an interference level from all its opponents belonging to the same coalition. Meanwhile, thanks to the PD-NOMA concept, the BS can perform the SIC procedure to cancel part of the interference before decoding the signal of each user. Intuitively, based on the distance from the BS, each user can estimate the interference level at the BS when the latter decodes its signal. Indeed, since the signal from the closest device, which has the best channel gain, is decoded first, it is subject to interference from all the devices sharing the same coalition. In addition, at a distance $r_i = R/2$, as the device is classified in the middle of the vector of devices, half of the interfering users can be removed. Whereas, for the farthest one

at a distance $r_i = R$ which has the lowest channel gain, since the signals of all the stronger devices are already decoded before decoding its own signal, it perceives a very low interference level.

Finally, each device can determine its SINR value and its utility in response to the estimated interference as:

$$\gamma^{mean}(t,s) = \frac{p(t)|h(t)|^2}{\sigma^2 + \tilde{I}_i(t)}, \tag{5.15}$$

$$U^{mean}(t,s) = \frac{f(\gamma^{mean}(t,s))}{p(t)}. \tag{5.16}$$

5.4.2 Mean field game equations

Commonly, solving a MFG relies on the two combined equations HJB and FPK given in Chapter 2 as (2.5) and (2.6). The former equation rules the decision making process of each individual user in response to the mass behavior of its opponents, while the latter equation controls the mean field's evolution according to the players' decisions.

Proposition 5.4. Using the state equation given by (5.2), we write the HJB and FPK equations as (5.18) and (5.17) respectively:

$$\frac{\partial m(t,h)}{\partial t} - qw_{1,n}(t)\nabla_{h_R} m(t,h) - qw_{2,n}(t)\nabla_{h_I} m(t,h) = 0. \tag{5.17}$$

$$\frac{\partial v(t,h)}{\partial t} = \max_{p(t)}[U^{mean}(t,s,p(t)) - qw_{1,n}(t)\nabla_{h_R} v(t,h) - qw_{2,n}(t)\nabla_{h_I} v(t,h)]. \tag{5.18}$$

Proof. By invoking the expression of the Nabla operator, the divergence of a vectorial quantity u is given by: $\nabla u = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} = \nabla u_x + \nabla u_y + \nabla u_z$. In our work, the evolution of the system is governed by the evolution of the state $s(t) = \{h(t), k\} = \{h_R(t), h_I(t), k\}$. Therefore, the components of a vector are along h_R, h_I and k axes and thus we have $\nabla_s u = \frac{\partial u_{h_R}}{\partial h_R} + \frac{\partial u_{h_I}}{\partial h_I} + \frac{\partial u_k}{\partial k}$. On the other hand, in our case, we have $u = (m(t,s)\frac{\partial s}{\partial t})$. Hence

$$\begin{aligned}\nabla_s(m(t,s)\frac{\partial s}{\partial t}) &= \frac{\partial(m(t,h_R)\frac{\partial h_R}{\partial t})}{\partial h_R} + \frac{\partial(m(t,h_I)\frac{\partial h_I}{\partial t})}{\partial h_I} \\ &= \nabla_{h_R}(m(t,h_R)\frac{\partial h_R}{\partial t}) + \nabla_{h_I}(m(t,h_I)\frac{\partial h_I}{\partial t}).\end{aligned}$$

Using the state equation defined as

$$ds(t) = dh(t) = \frac{dg(t)}{l} = \begin{pmatrix} -qw_{1,n}(t)dt \\ -qw_{2,n}(t)dt \end{pmatrix} = \begin{pmatrix} \frac{\partial h_R(t)}{\partial t}dt \\ \frac{\partial h_I(t)}{\partial t}dt \end{pmatrix}.$$

in FPK equation, given as (2.6), we end up with the following FPK equation:

$$\frac{\partial m(t,h)}{\partial t} - qw_{1,n}(t)\nabla_{h_R} m(t,h_R) - qw_{2,n}(t)\nabla_{h_I} m(t,h_I) = 0.$$

In the same way, we have derived the HJB equation. □

Definition 5.4. The formulated MFG is expressed as a combination of two fundamental equations, namely the HJB and the FPK given as (5.18) and (5.17) respectively.

Based on the mean field theory, the HJB equation (5.18) rules the computation of the value function $v(t,h)$ and the optimal power control policy $Q^*(t)$ to be adopted in response to any given mean field. On the other hand, the FPK equation (5.17) controls the evolution of the density of devices $m(t,h)$ in order to match the power control policy returned by the HJB equation. The iterative resolution of these coupled equations leads to the MFE.

Definition 5.5. The ping-pong interaction between the value function $v(t,h)$ and the mean field $m(t,h)$ converges to a stable combination of $v^*(t,h)$ and $m^*(t,h)$, which corresponds to MFE.

The solutions of the HJB (5.18) and FPK (5.17) equations, at time t and state s, are the value function $v(t,h)$ and the mean field $m(t,h)$ respectively. Once, the value function is obtained, the power control policy $Q(t)$ can be deduced. Furthermore, the evolution of the system density $m(t,h)$ depends on what the HJB equation can return as a value function $v(t,h)$ whereas, $m(t,h)$ in its turn impacts the computation of $v(t,h)$. Thereby, this interaction leads to the convergence point, namely the MFE which can be achieved by invoking the finite difference method.

5.5 Algorithm design of Mean Field Game

In this section, we focus on the finite difference to solve the set of mean field equations for the proposed game. According to the finite difference method [129], we discretize the time interval $[0, T]$, and the channel state spaces $[0, h_R^{max}]$, $[0, h_I^{max}]$ into $X \times Y \times Z$ spaces. Since, the mean field's evolution takes place in three-dimensional space of time and state components, we define the iteration steps of time and state spaces as

$$\delta_t = \frac{T}{X}, \delta_{h_R} = \frac{h_R^{max}}{Y}, \delta_{h_I} = \frac{h_I^{max}}{Z},$$

respectively. Therefore, we have a set of the time points given as: $t = \{0, 1 * \delta_t, 2 * \delta_t, \ldots, X * \delta_t\}$ and sets of state components defined as $h_R = \{0, 1 * \delta_h, 2 * \delta_h, \ldots, Y * \delta_h\}$ and $h_I = \{0, 1 * \delta_h, 2 * \delta_h, \ldots, Y * \delta_h\}$. On the other hand, the value function $V(t, h)$ as well as the mean field $m(t, h)$ can be rewritten as $V(x, y, z)$ and $m(x, y, z)$ respectively, where $0 \leq x \leq X$, $0 \leq y \leq Y$ and $0 \leq z \leq Z$.

5.5.1 Solution to the FPK equation

In order to solve the forward equation, the Lax-Friedrichs method [130] is used. Thus, the FPK equation derived in (5.17) can be extended and depicted as

$$M(x + 1, y, z) = \frac{1}{2}\Psi + \frac{q\delta_t}{2\delta_{h_R}}\Phi + \frac{q\delta_t}{2\delta_{h_I}}\Omega, \tag{5.19}$$

where

$$\Psi = M(x, y - 1, z) + M(x, y + 1, z) + M(x, y, z - 1) + M(x, y, z + 1), \tag{5.20}$$

$$\Phi = M(x, y + 1, z)W_{1,n}(x, y + 1, z) - M(x, y - 1, z)W_{1,n}(x, y - 1, z), \tag{5.21}$$

$$\begin{aligned} \Omega = {} & M(x, y, z + 1)W_{1,n}(x, y, z + 1)W_{2,n}(x, y, z + 1) \\ & - M(x, y, z - 1)W_{1,n}(x, y, z - 1)W_{2,n}(x, y, z - 1), \end{aligned} \tag{5.22}$$

$M(x, y, z)$, $W_{1,n}(x, y, z)$ and $W_{2,n}(x, y, z)$ denote the modified values of $m(t, h)$, $w_{1,n}$ and $w_{2,n}$ at time x with the channel states y and z in the discretized grid, respectively.

5.5.2 Solution to the HJB equation

Upon adopting the discretization method, the complicated derivative expressions $\frac{\partial v(t,h)}{\partial t}$, $\nabla_{h_R} v(t,h)$ and $\nabla_{h_I} v(t,h)$ can be reformulated as

$$\frac{\partial v(t,h)}{\partial t} = \frac{v(x,y,z) - v(x-1,y,z)}{\delta_t}, \tag{5.23}$$

$$\nabla_{h_R} v(t,h) = \frac{v(x,y,z) - v(x,y-1,z)}{\delta_{h_R}}, \tag{5.24}$$

$$\nabla_{h_I} v(t,h) = \frac{v(x,y,z) - v(x,y,z-1)}{\delta_{h_I}}. \tag{5.25}$$

Each device has to solve the HJB equation to determine its transmit power in any state and at any time. Thus, for any arbitrary point (x,y,z) in the discretized grid, by substituting (5.23), (5.24) and (5.25) into HJB equation (5.18), the value function can be updated as follows:

$$\begin{aligned}\max_{p(x,y,z)} [U^{mean}(p(x,y,z), m(x,y,z)) - qw_{1,n}\frac{v(x,y,z) - v(x,y-1,z)}{\delta_{h_R}} \\ - qw_{2,n}\frac{v(x,y,z) - v(x,y,z-1)}{\delta_{h_I}}] + \frac{v(x,y,z) - v(x-1,y,z)}{\delta_t} = 0.\end{aligned} \tag{5.26}$$

In addition, at point (x,y,z) in the discretized grid, the optimal power control is derived as:

$$\begin{aligned}p(x,y,z) = \underset{p(x,y,z)}{\operatorname{argmax}}[U^{mean}(p(x,y,z), m(x,y,z)) \\ -qw_{1,n}\frac{v(x,y,z) - v(x,y-1,z)}{\delta_{h_R}} - qw_{2,n}\frac{v(x,y,z) - v(x,y,z-1)}{\delta_{h_I}}].\end{aligned} \tag{5.27}$$

Then, we obtain the optimal power strategy to be used in response to the collective behavior modelled by I^{mean},

$$p(x,y,z) = \gamma^* \frac{\tilde{I}(x) + \sigma^2}{|h(x,y,z)|^2}, \tag{5.28}$$

where γ^* is the solution to

$$\gamma f'(\gamma) - f(\gamma) = 0. \tag{5.29}$$

5.5.3 Proposed algorithm of mean field game

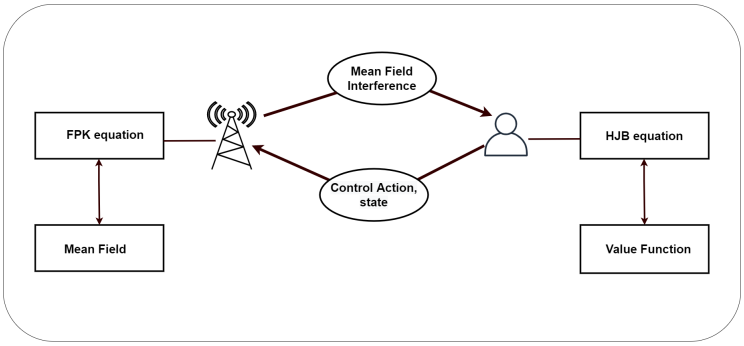

Figure 5.1: Interaction process between HJB and PFK equations.

Algorithm 3: Distributed power control policy for MFG solution

Initialisation: $M(0,:,:)$:=joint mean field distribution, initial power levels; $I^{mean}(0) := 0$, I^{mean} value for $t = 0$.

Output: optimal transmit power and mean field

for $N = 1 : nbIter$ **do**
  for $x = 1 : 1 : X$ **do**
    for $y = 1 : 1 : Y$ **do**
      for $z = 1 : 1 : Z$ **do**
        At each device:
        Estimate interference level $\tilde{I}(x)$ as (5.14)
        Calculate power level $p_{temp}(x, y, z)$ according to (5.28) with the probability p_t
        if $(p_{temp}(x, y, z) \leq P_{max})$ **then**
          Update power level $p(x, y, z) = p_{temp}(x, y, z)$
          Calculate utility value $U^{mean}(x, y, z)$ as (5.16)
        else
          $p(x, y, z) = 0$
          $U^{mean}(x, y, z) = 0$
        At the BS:
        Update mean field $M(x + 1, y, z)$ using the update of (5.19)
        Update MFI $I^{mean}(x + 1)$ according to (5.13)

In this section, following the above derivations, we propose a distributed algorithm to achieve iteratively the MFE for the formulated MFG through a power policy adaptation. We assume, in the proposed algorithm, that each device can only belong to one coalition at a time and that it chooses to join the group corresponding to its best

channel. Initially for $t = 0$, there is no interference, each device starts by calculating its transmit power. Then, after transmitting, the BS becomes aware of the transmit power of each device, its chosen channel gain and the corresponding sub-carrier. Once, the BS collects the information from different MTDs, it calculates the mean field value and the MFI for $t+1$. Then, the BS broadcasts the calculated value of the aggregated interference as feedback information to the set of devices. Each of them, estimates its interference level $\tilde{I}$ from its perspective based on its location and then determines its power level as well as its utility in response to the estimated interference. This interaction between the BS and every device is shown in Figure 5.1. A summary of the developed algorithm is provided in Algorithm 3 and its effectiveness is proven numerically by simulations in the next section.

5.6 Simulation Results

In this section, we present comprehensive Matlab-based simulations of the proposed MFG-based power control, described in the previous sections. We consider a Hybrid NOMA scenario composed of N devices which are independently distributed according to the homogeneous PPP process of density ω_N. As explained above, the devices are divided into K coalitions, where the members of each coalition use one sub-carrier to transmit their packets. The main simulation parameters are summarized in Table 5.1.

Table 5.1: System Parameters

Parameter	Value
System effective bandwidth BW	5.4MHz
Density of homogeneous PPP, ω_N	0.1
Cell range, R	200m
Bandwidth of a sub-carrier	180KHz
Number of available sub-carriers K	20
Maximal frequency reuse, α	5
Time interval, T	0.3 s (i.e. 30 LTE frames)
SINR value satisfying (5.29), γ^*	6.4 (or 8.1 dB)

In the first part of this section, we evaluate the performance of the proposed MFG and illustrate the equilibrium properties of the proposed power control approach.

Subsequently, we provide a comparison with other existing techniques in the second part.

5.6.1 Performance metrics

In order to highlight the properties of our approach, we rely on the following metrics:

- Packet success rate: is calculated as the ratio between the number of users that have successfully transmitted and the number of active users (that have decided to transmit in the current time slot).
- Average transmission rate: is determined as the ratio between the number of users that have successfully transmitted and the total number of users in the system.
- Average utility: is the ratio of the sum of the utility values of devices whose packets have been successfully decoded by the BS to the total number of users in the system.
- Average energy: similarly to the average utility, the average energy consumption is calculated based on the energy consumed by the devices which have successfully delivered their packets.

5.6.2 Behavior of the game at the equilibrium

Throughout this section, the depicted results correspond to different numbers of MTDs, i.e. $N = 2000$, $N = 4000$, $N = 6000$, $N = 8000$ and $N = 10000$, sharing $K = 20$ sub-carriers in a Hybrid NOMA network.

In Figure 5.2, we illustrate the evolution of the packet success rate over the time. As we can clearly see, this rate converges at $t = 10$ ms to about 0.68 (68% of success rate). It is substantially interesting to observe that regardless of the number of devices playing the game, the system achieves the same packet success rate and becomes steady at the same point. This is mainly due to the fact that the devices adapt their transmission strategies to feedback received from the BS. Roughly speaking, we have the same average number of devices that simultaneously would transmit their packets. Consequently, each user has the same chance to successfully deliver its packet regardless of the total number of its opponents in the system.

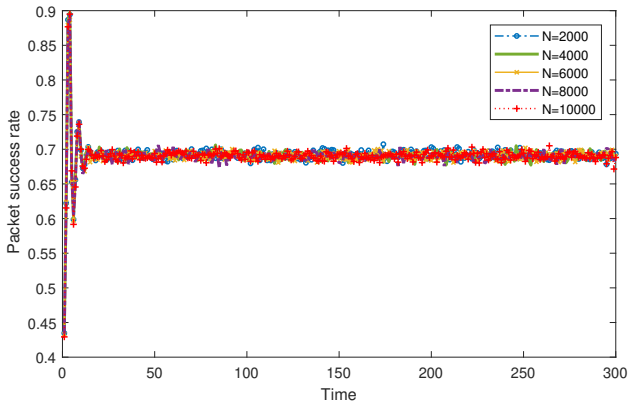

Figure 5.2: Packet success rate versus T when $K = 20$.

It is worth noting that the evaluation of our game's performance in terms of the packet success rate and the illustration of its variation with respect to time, as shown in Figure 5.2, perfectly reflects the variation of the MFI over the time. In fact, initially, there is no interference, so a large number of users choose to transmit, which nevertheless has an impact on the transmission success of each of them since the BS may fail in the decoding procedure while applying the SIC to separate the superimposed signals. Hence, the packet success rate is about 0.43. Then, the BS determines the MFI based on the received information from the participating devices and broadcasts it to the set of users. Now, the latter are facing a considerable interference level, wherefore some devices may not be able to regulate their transmit power to achieve γ^* while dealing with the received MFI. Thus, few users would choose to transmit their messages during this time slot. Subsequently, the BS is able to decode most of them successfully which results in an interesting value of packet success rate, about 0.88. After a certain time, i.e. $t = 10$ ms, this "ping-pong" interaction between the MFI and the users' responses leads the packet success rate to reach its convergence point.

Figure 5.3 shows the average transmission rate over the time slots. It can be observed that this rate stagnates at $t = 10$ ms, however the convergence values for the four cases are different, the one reached by $N = 2000$ devices is greater than those achieved by the other cases. This result is somehow intuitive since when we have $N = 2000$ devices, the system is the most sparse compared to the other cases. Therefore, the

devices perceive much less interference than that encountered by 10000 for example.

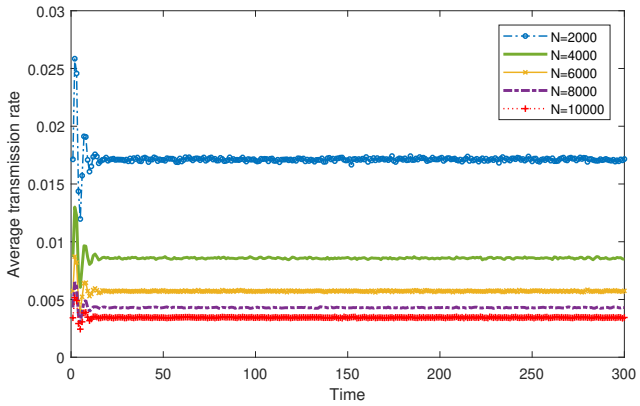

Figure 5.3: Average transmission rate versus T when $K = 20$.

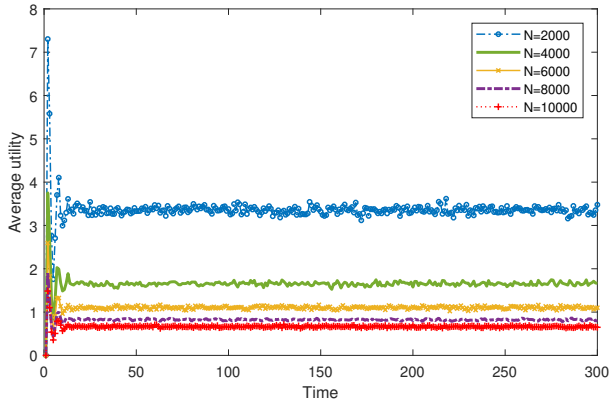

Figure 5.4: Average utility versus T when $K = 20$.

Figures 5.4 and 5.5 depict respectively the average utility and the average energy consumption over the time T for the devices that have successfully delivered their packets. It can be concluded from these figures that users manage to reach a stable point at $t = 10$ ms in terms of the utility and the energy consumption. Interestingly, the behavior of figures 5.4 and 5.5 follow that of Figure 5.3 and they converge at the same time slot, which illustrate the equilibrium achieved by our proposed MFG. In other words, our approach settles down at the point where the players that have

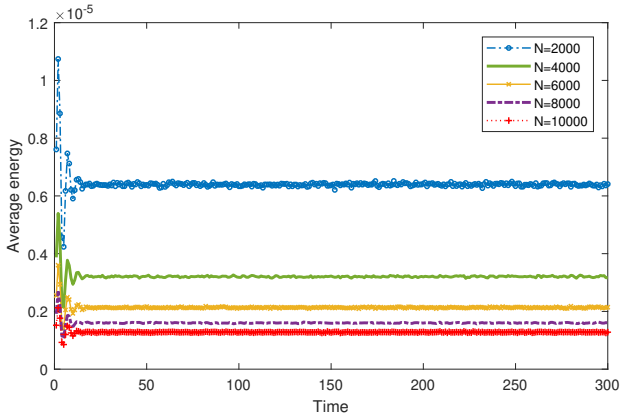

Figure 5.5: Average energy consumption versus T when $K = 20$.

succeeded in transmitting, become satisfied with both their energy consumption and utility. It is noteworthy that for the case when $N = 2000$, the devices may reach the highest utility value while consuming more energy than the other cases. This is owing to the fact that these performances are measured with respect to the ratio of the number of devices that successfully transmitted to the total number of devices presented in the system, illustrated in Figure 5.3.

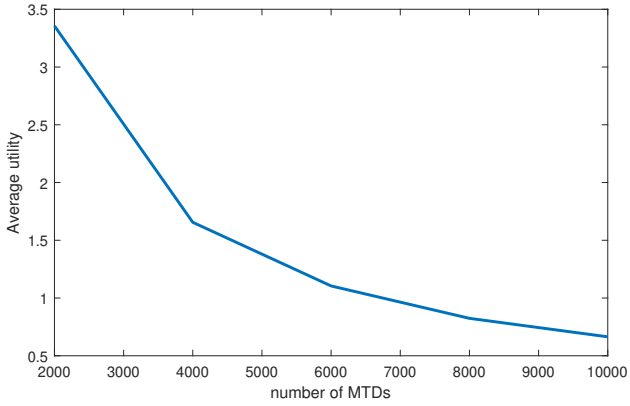

Figure 5.6: Average utility versus N when $K = 20$.

Figure 5.6 and 5.7 depict respectively the average utility and average energy consumption as a function of the number of devices N. It can be observed in these figures that the average utility value as well as energy consumption decrease as N increases. This is mainly due to the interference effects which become more significant as the system

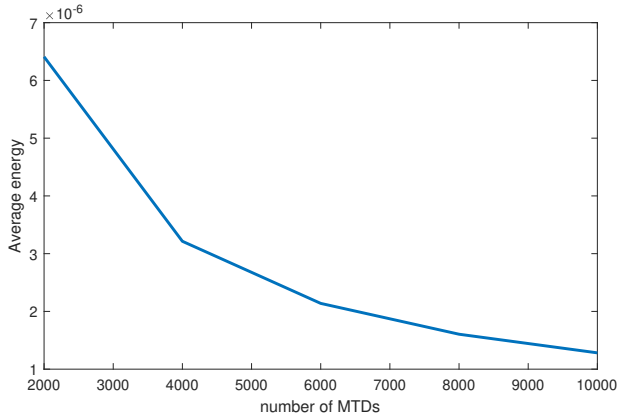

Figure 5.7: Average energy consumption versus N when $K = 20$.

becomes denser. Consequently, the transmission success of the devices becomes more challenging.

5.6.3 Comparison

This section is devoted to the comparison of the proposed approach in this chapter with the Bi-level theoretical framework investigated in Chapter 4 and with the game-theoretic approach NM-ALOHA game given in [100]. Throughout this section, we consider different scenarios composed of $200 \leq N \leq 1600$ MTDs competing over $10 \leq K \leq 20$ sub-carriers.

5.6.3.1 Packet success rate

Figure 5.8 displays the packet success transmission rate with respect to the number of sub-carriers K and the number of devices N. It can be observed from this figure that as the number of devices increases, the packet success rate decreases drastically for both the NM-ALOHA game and the Bi-level game NOMA due to the interference effects, while the proposed MFG enables no matter how many devices in the system to achieve the same rate. As explained above, this is due to the fact that the devices adapt their transmission strategies to feedback received from the BS. It can be concluded from this figure that the benefit of adopting our approach is particularly highlighted when the network is made up of a large number of devices, that is to say $N = 1600$, since in this case the packet success rate reaches up 0.688 when $K = 20$

while it is about 0.1508 for the Bi-level game NOMA and 0.0124 for the NM-ALOHA game. This means that the packet success rate is improved by $356,233\%$ over our previous work on the Bi-level game. Thus, this figure significantly reveals the robustness of our proposed game against the interference impact since it still achieves acceptable performances. Consequently, our approach is well appropriate to deal with densely deployed networks.

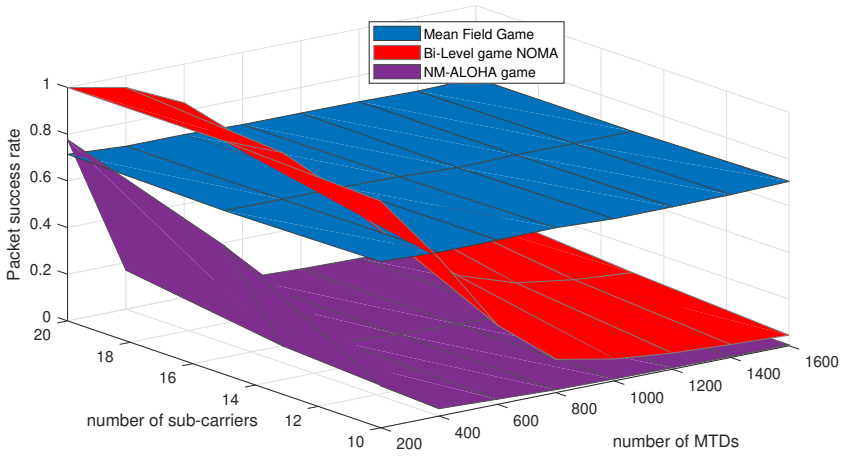

Figure 5.8: Packet transmission rate with success for different N and K.

5.6.3.2 Total throughput

We illustrate in Figure 5.9 the total throughput of our proposed scheme versus the Bi-level game NOMA and the NM-ALOHA game as a function of the number of sub-carriers K and the number of MTDs N. As we can clearly see, when the system is relatively sparse, the highest throughput is achieved with Bi-level game NOMA. However, as the number of devices increases and the network becomes denser, the total throughput decreases for both Bi-level game NOMA and NM-ALOHA game whereas, our proposed technique still provides an acceptable throughput. In fact, in Figure 5.10, we show that the total throughput is not influenced by the increasing number of users N, on the contrary it achieves a steady value for the different scenarios, resulting in an important improvement of $78,335\%$ compared to the Bi-level game when $N = 1600$. Unsurprisingly, this result is somehow expected due to two reasons. Firstly, the throughput is calculated with respect to the devices which have

successfully transmitted. Secondly, we demonstrated, in Figure 5.8, that when the devices are competing over the same number of sub-carriers, e.g, $K = 20$, they have almost the same chance to successfully deliver their messages. Hence, our proposed framework outperforms other techniques in terms of robustness to interference effects, allowing the MTDs, regardless of their number, to achieve an acceptable throughput.

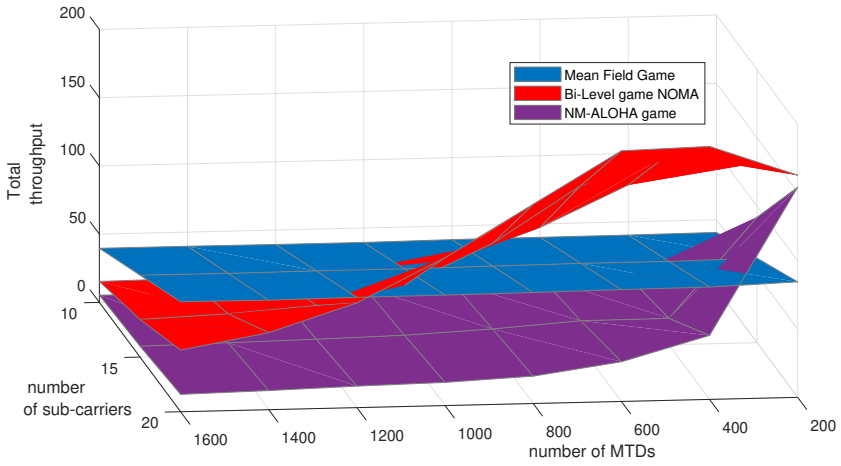

Figure 5.9: Total throughput of devices for different N and K.

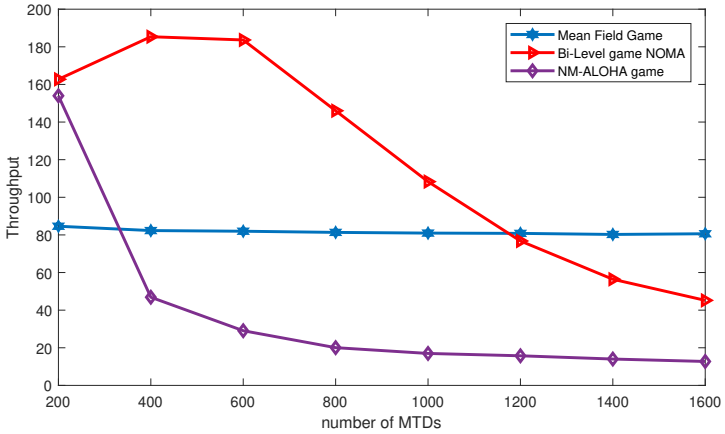

Figure 5.10: Total throughput versus N when $K = 20$.

5.6.3.3 Average utility

In our game, the proposed utility function, which has units of bits/joule, serves to measure the total number of reliable bits transmitted per joule of energy consumed so that it can be used to assess the energy efficiency in the NOMA network. Therefore, we focus now on the effect of the number of sub-carriers K on the energy efficiency of a scenario composed of $N = 1600$ users. As can be seen clearly in Figure 5.11, the average utility is increasing as K increases. In fact, the more the network is divided into coalitions, the fewer the devices are in each coalition, the less interference they face. As expected, our proposed approach yields an interesting performance enhancement in the average utility and even surpasses the Bi-level game NOMA and the NM-ALOHA game as the number of sub-carriers increases, reaching up 4.72 against 3.15 for the Bi-level game NOMA and 0.10 for the NM-ALOHA game when $K = 20$. More precisely, the proposed MFG can greatly boost the average utility by $49,841\%$ over the Bi-level game NOMA. Hence, we clearly notice the benefit of our MFG in terms of the average utility. This is owing to the fact that the proposed power strategy takes into account the MFI, so that each device can respond effectively to the collective behavior by regulating its transmit power and then determining its utility value.

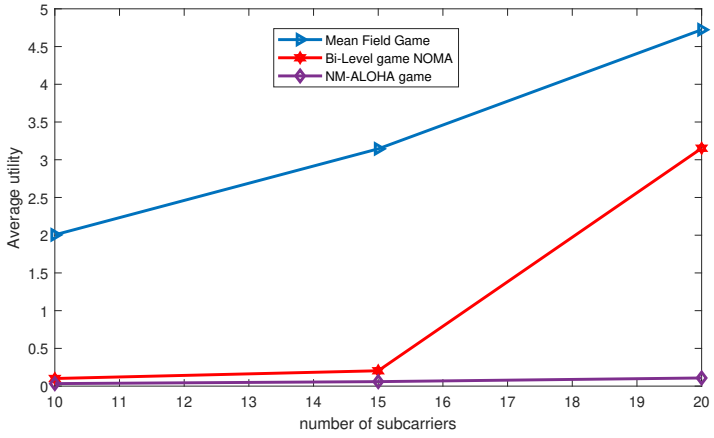

Figure 5.11: Average utility versus K when $N = 1600$.

5.6.3.4 Energy consumption

Now, we investigate the energy consumption of the proposed MFG under several scenarios of N users and K sub-carriers. For this purpose, we adopt the same model provided in Chapter 4 in order to assess the energy consumption. Figure 5.12 depicts the average energy consumption for the different techniques. As we have mentioned above, we have considered the energy consumption of the devices whose packets have been successfully decoded by the BS. Indeed, as we have shown in Figure 5.8, the proposed scheme enables the users to have a good packet success rate, especially for a dense network, compared to the Bi-level game NOMA and the NM-ALOHA game, which means that our game yields the highest average number of users that successfully transmitted for almost all scenarios. Meanwhile, the proposed game, as illustrated in Figure 5.12, enables the devices to reduce their energy consumption compared to the other schemes. Indeed, Figure 5.13 mainly spotlights the comparison between the proposed MFG and the NM-ALOHA game and clearly shows that with the aid of the MFG the energy consumption is minimized by $73,945\%$.

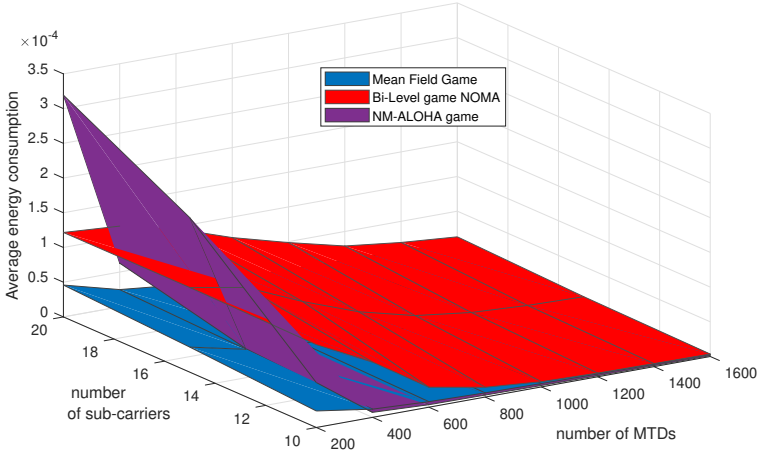

Figure 5.12: Average energy consumption for different N and K.

5.7 Conclusion

In this chapter, we have investigated the effect of the aggregated interference perceived by each device on the power control in ultra-dense networks. We have considered that BS does not perform the resource allocation and only broadcasts a feedback in order

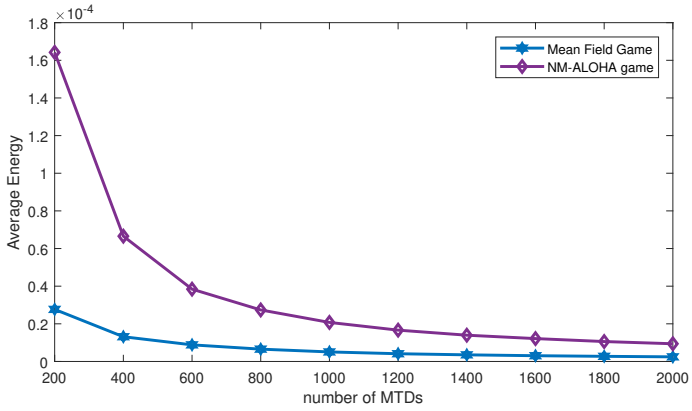

Figure 5.13: Average energy consumption versus N when $K = 10$.

to alleviate the performance drop associated with almost all grant-free techniques in dense scenarios. We have first proposed a differential game to address the power allocation problem, then we have turned our attention to an equivalent tractable MFG which drastically reduced the mathematical complexity of the proposed scheme since it can be characterized by the two combined equations HJB and FPK. More precisely, the substantial idea of the presented framework is to model the mass behavior of the active devices as the mean field and allow each of them to react in response to the MFI received from the BS as feedback information, thus considerably simplifying the resolution of the game. An iterative algorithm is derived to approach the MFE by paving the way for a distributed control in which the devices adjust autonomously their transmit power levels according to the MFI. Simulation results are presented to illustrate the equilibrium behaviors of the game under consideration and demonstrate the robustness of the formulated MFG compared to other existing approaches in the literature, especially in the presence of a dense population.

In the next chapter, we extend the proposed MFG approach with the aid of a RL tool to make the devices able to choose the appropriate coalition while dealing with the effect of the collective behavior captured by the MFI.

Chapter 6

Reinforcement learning for Mean Field Game-based Resource Allocation in NOMA Networks: a Multi-Armed Bandit Approach

Contents

6.1 Introduction

Usually, wireless communication networks are characterized by an important level of interference encountered by each user. In an effort to alleviate the interference effects, numerous research contributions have been devoted to modeling the power control problem under game theory setting [31, 91, 97, 99, 100, 131]. Furthermore, game theory is applied to suitably derive distributed algorithms that aim to optimize the resource allocation for densely deployed IoT networks [113]. For instance, the MFG has increasingly gained attention in wireless communication networks, namely when a large number of users are involved. Interestingly, the MFG simplifies the resolution of the power control by drastically reducing the mathematical complexity of the problem to a two-body complexity rooted in two tractable combined HJB and FPK equations. Afterward, the MFE is obtained by iteratively solving these coupled equations. Conventionally, the finite difference method is invoked to approach the MFE [129]. Nevertheless, when the game is characterized by large state and action spaces, applying the finite difference method requires a higher computational burden.

As an alternative method, RL techniques have been exploited as a sophisticated tool to solve the MFG [122, 123, 132, 133]. Particularly, the MAB framework [134], which represents an important class of RL algorithms, has been specifically adopted to optimize the resource allocation problems in the context of wireless networks. Indeed, the amalgam of MFG and RL algorithms has garnered a substantial attention as it provides useful insights into how to effectively deal with the resource management problems. The authors, in [132], have designed an RL-based MFG algorithm with the intention of maximizing the sum-rate among users in the context of unmanned aerial vehicle-enabled mmWave systems. Shi et al. have considered in [133] a cooperative multi-access edge computing framework and have resorted to the deep RL tool to learn the optimal policy in order to achieve the Nash equilibrium of the MFG. [123] and [122] have first applied the MFG framework to model the collective behavior of multi-user NOMA scenarios in mobile edge computing systems. Then, deep RL algorithms have been proposed to solve the game and optimize the resource allocation issues between different users in NOMA clusters. Concomitantly, within the MAB framework, there are some contributions have examined the NOMA approach, such as [135–137]. In fact, the authors of [135] have been interested in organizing the transmissions of users as well as their power allocation strategies by invoking the

MAB framework. In [136], a distributed MAB algorithm has been proposed in order to handle the channel access and power control issues, whereas in [137] a MAB learning approach has been conceived to address the scheduling problem for fast-grant MTDs.

In this chapter, we propose a RL approach based on the mean field theory in order to jointly solve the resource allocation and power control problems in a Hybrid NOMA scenario. We more specifically investigate the MAB algorithm to model the competitive behaviors of the players over the set of arms, i.e. set of available RBs, with an eye toward maximizing their rewards.

At the time of writing, although the literature provides some contributions that have combined the MFG and the RL techniques and others that have applied the MAB algorithm in NOMA-based networks, there is no published literature that has investigated the combination of the MFG framework and the MAB approach underlying NOMA networks. To the best of our knowledge, our proposed approach is the first work that focuses on jointly solving the channel selection and power control problems using MFG-based MAB approaches for Hybrid NOMA scenarios. The main contributions of this work can be given as follows:

- We examine a Hybrid NOMA-based network composed of several coalitions of devices. The members belonging to each group share a single sub-carrier to transmit their signals.
- We extend the MFG approach proposed in the previous chapter by invoking a RL-based scheme. Thus, unlike what we have investigated in Chapter 5, we aim to solve the formulated MFG using MAB frameworks instead of the traditional finite difference method.
- We propose a resource allocation technique based on a bi-level learning with the aim of jointly optimizing the user clustering and power control problems. In fact, we derive distributed MFG underlying MAB algorithms in which the MAB technique is invoked to enable the MTDs to self-organize into coalitions. Then, the MFG is applied such that MTDs can adjust their transmit power levels based on the received MFI.
- The two developed MFG-based MAB algorithms are designed with the aid of the ϵ-decreasing greedy and the Upper Confidence Bounds (UCB) methods in

order to address the action selection problem for the MAB approach and thus allow the devices to decide which coalition it is better to belong to.

- Regret analysis is presented to evaluate the performance of the proposed MFG-based MAB techniques. We show that the regret incurred during the learning process evolves logarithmically.
- We provide numerical simulations that underline the features of the combined MFG and MAB frameworks under several scenarios made up of different numbers of devices and sub-carriers.

In the light of the above, we construct the rest of this chapter as follows. We discuss the problem formulation of this chapter in the next section. Section 6.3 is devoted to derive two distributed MFG-aided MAB algorithms in order to address the joint problems of the sub-carrier selection and power allocation. To assess the performance of the proposed approaches, numerical results are provided in Section 6.4. Finally, Section 6.5 concludes the present chapter.

6.2 Problem Formulation

Similar to Chapter 5, we consider a dense NOMA scenario in which a large number of users are involved. The uplink power control is formulated as a MFG where the HJB and FPK equations model the mass behavior of the devices as a MFI that each user has to interact with in order to make its decision. Particularly, This MFI is given as:

$$I^{mean}(t) = \alpha \int |h(t)|^2 p(t) m(t,h) dh. \tag{6.1}$$

Then, the BS broadcast the obtained value of the MFI as a unique feedback to the devices participating in the game. Once each of them receives this feedback, it estimates its interference level from its perspective as:

$$\tilde{I}_i(t) = \left(1 - \frac{r_i}{R}\right) I^{mean}(t). \tag{6.2}$$

Upon invoking the finite difference method, we have obtained the expression of the optimal power control as:

$$p(t) = \gamma^* \frac{\tilde{I}(t) + \sigma^2}{|h(t)|^2} \tag{6.3}$$

The aforementioned approach relies on the fact that each user, at each time t, joins the cluster that corresponds to its best channel. Now, if a device aims to deviate from its current coalition, it needs to choose which coalition is preferable to be part of. Such a decision requires usually more information about the other coalitions. In other words, in this chapter we seek to make the devices autonomous in their choice of groups while modeling their behavior in each group as the mean field, so that each device can regulate its transmit power in response to the MFI. In this case, we are dealing with the joint optimization problems of the user grouping and the power control. Thereby, applying the finite difference method to solve these combined problems is not practically affordable, especially for high-dimensional state and action spaces. With this in mind, we spotlight RL-based approaches, particularly MAB framework in the following.

6.3 Multi-armed bandit framework

In this section, we formulate the joint user grouping and power allocation problems in a MFG framework underlying a multi-user MAB approach. Firstly, the devices invoke the MAB tool to arrange themselves into multiple NOMA coalitions. Then, within each coalition, the MTDs apply the MFG approach to autonomously adjust their transmit power based on limited feedback received from the BS.

We propose two MFG-based MAB algorithms using the ϵ-decreasing greedy and UCB techniques in order to enable each user to make a move upon selecting an arm with the aim of maximizing its own utility. In this direction, we define the set of devices as the set of learners and the available RBs as the arms to be chosen by each learner. Let $\mathcal{A}_i = \{a_1 \dots a_K\}$ denotes the set of possible arms for each device i. Indeed, at time slot t, the player i first pulls an arm a_i, then it joins the coalition corresponding to this chosen RB. Secondly, it determines the appropriate transmit power $p_i(t, a_i)$ to be used by being part of the chosen coalition while addressing the MFI received from the BS. After transmission, the BS informs each user whether its packet was received and decoded successfully or not by sending back a reward value $r_i(t, a_i)$, allowing it in turn to determine its utility value $U_i(t, a_i)$. In fact, we assume that upon selecting its arm a_i and calculating the corresponding power coefficient $p_i(t, a_i)$, the player transmits its message to the BS. The latter applies the SIC procedure to separate the

superimposed signals. Thus, if the packet of the player i is successfully decoded, it receives $r_i(t, a_i) = 1$ and its utility is calculated as in the equation (5.16). Otherwise, it receives $r_i(t, a_i) = 0$ which implies that the user has no utility by choosing the arm a_i at time t.

Algorithm 4: ϵ-decreasing greedy

Input: ϵ_0, $Q(j, :)$, T
for $t = 1 : T$ **do**
  $\epsilon_t = \min(1, \frac{\epsilon_0}{t})$ Generate a random number $x \in [0, 1]$ **if** $x < \epsilon_t$ **then**
    select a_j randomly
  else
    $a_j = \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, Q_t(j, a)$

6.3.1 ϵ-decreasing greedy

The ϵ-greedy method is widely used as one of the most prominent solution concepts for the arm selection problem in the MAB framework [134]. It allows users to explicitly manage an exploration-exploitation trade-off with an exploration rate ϵ. Indeed, at each time slot, each device decides either to explore or exploit. In other words, it arbitrarily picks an arm with a probability ϵ or it selects with a probability of $1 - \epsilon$, the optimal arm which gives it the highest average reward $Q_t(i, :)$ considering the past observations. Nevertheless, if the exploration parameter ϵ is constant for the entire-process, we end up with a sub-optimal allocation and a linear regret, which in turn affect the overall system performance. In order to overcome this issue, the exploration coefficient ϵ has to be adjusted over the time. Thus, in our work, we apply the ϵ-decreasing proposed by [138] whose key idea is outlined in the Algorithm 4. In fact, we define the adaptation of time-dependent exploration parameter as follows:

$$\epsilon_t = \min(1, \frac{\epsilon_0}{t}), \tag{6.4}$$

where $\epsilon_0 > 0$ is the initial exploration parameter. In this way, at the beginning of the learning process, more exploration is performed, allowing each user to discover the arm space as much as possible. Then, ϵ_t is dynamically regulated as a function of the learning time. Thus, the user can now properly select its best arm according to its acquired experience.

Algorithm 5: Parameters update

for $t = 2 : T$ **do**
  for $j = 1 : N$ **do**
    $s_t(j, a_j) = s_{t-1}(j, a_j) + r_j(t, a_j)$;
    $n_t(j, a_j) = n_{t-1}(j, a_j) + 1$;
    $Q_t(j, a_j) = \frac{s_{t-1}(j,a_j)}{n_{t-1}(j,a_j)}$

Algorithm 6: ϵ-decreasing MFG-based MAB method for joint channel selection and power control: ϵ-decreasing MFG

for $t = 1 : T$ **do**
  for $j = 1 : N$ **do**
    Select an arm a_j using Algorithm (4) of ϵ-decreasing greedy
    Estimate the interference level $\tilde{I}(t, a_j)$ as (6.2)
    Calculate the power level $p_j(t, a_j)$ according to (6.3).
    Update parameters using Algorithm (5).
  At the BS:
  Update the mean field and then the MFI $I^{mean}(t+1)$ according to (6.1).

6.3.2 Upper Confidence Bounds algorithm

UCB algorithm was first proposed by [138], and broadly adopted to deal with the action selection problem in MAB setting. Unlike ϵ-decreasing greedy method, UCB implicitly distinguishes between exploration and exploitation phases by selecting the arm associated with the highest average reward given the past observations. This arm is known as the UCB index and given by the following equation for each user $i \in \mathcal{N}$:

$$a_i(t) = \underset{a \in \mathcal{A}}{\operatorname{argmax}} \left[Q_t(i, a) + \sqrt{\frac{2 \log(t)}{n_t(i, a)}} \right] \tag{6.5}$$

where $n_t(i, a_i)$ is the number of times the arm a_i has been played during the previous time slots. In fact, the UCB index at time slot t gathers two components, an upper confidence bias $\psi_t(a_i) = \sqrt{\frac{2 \log(t)}{n_t(i,a_i)}}$ and the average reward $Q_t(i, :)$ of playing the arm a_i up to time t. Particularly, $\psi_t(a_i)$ which depends on $n_t(i, a_i)$, is used to encourage the exploration and serves as an interval around the average reward. Thus, the more the arm is played, the more this interval is shrunken, which in turn reduces the probability of discarding this arm in the future observations. Consequently, UCB concept tends

to effectively meet the trade-off between exploration and exploitation.

Algorithm 7: UCB MFG-based MAB method for joint channel selection and power control: UCB-based MFG

for $t = 1 : T$ **do**
  for $j = 1 : N$ **do**
    $a_j(t) = \underset{a \in \mathcal{A}}{\operatorname{argmax}} \left[Q_t(j,a) + \sqrt{\frac{2\log(t)}{n_t(j,a)}} \right]$
    Estimate the interference level $\tilde{I}(t, a_j)$ as in (6.2)
    Calculate the power level $p_j(t, a_j)$ according to (6.3).
    Update parameters using Algorithm (5).
  At the BS:
  Update the mean field and then the MFI $I^{mean}(t+1)$ according to (6.1).

6.3.3 Distributed learning algorithms with Multi-Armed Bandit

In this section, we derive two distributed MFG-based MAB algorithms to solve the joint problems of the user grouping and power control in a Hybrid NOMA network. The first algorithm, illustrated in Algorithm (6) , adopts the ϵ-decreasing greedy method whereas the second algorithm, depicted in Algorithm (7) , resorts to the UCB method with the aim of efficiently performing the decision-making process. For the two proposed methods, we assume that a device can only belong to one coalition at a time. At each time slot t, each learner j pulls an arms a_j that represents the sub-carrier to use in order to deliver its packets. Thus, it joins the coalition associated with this arm.

When the user attempts to access the channel, it implicitly uploads to the BS information about its state and the selected arm. Subsequently, the BS broadcasts feedback information about the MFI $I^{mean}(t)$ expressed in (6.1). Then, from its perspective, each device can estimate its interference level $\tilde{I}(t, a_j)$ as in (6.2), according to its distance from the BS and calculates its power level $p_j(t, a_j)$ in response to the estimated interference. After the transmission, once the BS processes the SIC procedure, each user receives the reward $r_j(t, a_j)$ that allows it to update its parameters as indicated in the Algorithm 5. At the end of time slot t, the BS updates the mean field m as well as the MFI I^{mean} for time slot $t+1$. This interaction between each user and the BS is illustrated in Figure 6.1.

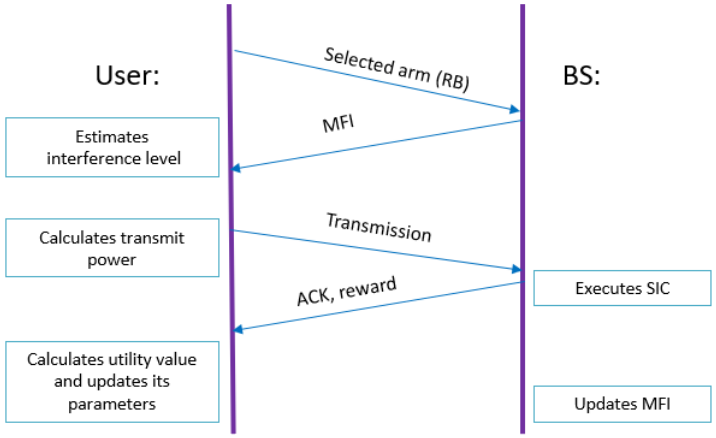

Figure 6.1: Interaction process between any device and the BS.

6.3.4 Regret analysis

The performance measure of the MAB approach is commonly based on the calculation of the total expected regret incurred during the learning process. Generally, it is defined as the difference between the actually obtained reward and the one that would have been obtained if the optimal arm had been selected. The expected regret over a period of T' time slots can be expressed as

$$R_i = T' r_i^* - \sum_{t=1}^{T'} \mathbb{E}[r_i(a_i(t))]. \tag{6.6}$$

Since in the approach under consideration we have α users that can simultaneously transmit at a given time, the total expected regret is given by:

$$R_{MAB} = T' \sum_{i=1}^{\alpha} r_i^* - \sum_{i} \sum_{t=1}^{T'} \mathbb{E}[r_i(a_i(t))]. \tag{6.7}$$

where $r_i(a_i(t))$ is the reward of the i-th device acquired at time t and r_i^* is its optimal reward.

6.3.4.1 Regret of ϵ-decreasing MFG algorithm

Now, we analyze the regret incurred when the ϵ-decreasing MFG Algorithm (6) is invoked. In doing so, we start by showing that learning the best arm can be performed in finite time as Lemma 1:

lemma 1. The proposed ϵ-decreasing MFG algorithm identifies an ϵ-best arm with at least probability $1-\delta$ when an arm sampling is carried out l times, where:

$$l = \frac{1}{2\epsilon^2}\log(\frac{2K}{\delta}). \tag{6.8}$$

with $\delta \in [0,1]$ denotes the probability of failure.

Proof. Denote by ϵ'-best arm a' an arm whose reward r' is different from the best reward r^* by less than ϵ, that is: $|r^* - r'| \leq \epsilon'$. Indeed, the user needs to sample each arm l times in order to obtain ϵ'-best arm with a probability $1-\frac{\delta}{K}$. Thus, we have

$$P(|r^* - r'| > \epsilon') \leq \frac{\delta}{K}. \tag{6.9}$$

On the other hand, according to the Hoeffding inequality, we obtain:

$$P(|r^* - r'| > \epsilon') \leq 2e^{-2l\epsilon^{2'}}. \tag{6.10}$$

Consequently, we end up with:

$$l = \frac{1}{2\epsilon^{2'}}\log(\frac{2K}{\delta}). \tag{6.11}$$

Henceforth, when a device samples an arm l times, ϵ-best reward is obtained with a probability $1-\frac{\delta}{K}$. □

lemma 2. All the devices can learn their best arms with a high probability, at least $1-\delta$, by adopting the proposed ϵ-decreasing MFG algorithm T^* rounds, where

$$T^* = \frac{l}{\alpha(1-\frac{1}{N})^{(N-1)}}. \tag{6.12}$$

Proof. We have shown, in Chapter 5, that the probability of successfully decoding

user's packet can be defined as $P_s = \frac{\alpha}{N}(1-\frac{1}{N})^{(N-1)}$. Therefore, the collision probability P_c of the set of N players over K sub-carriers can be given as:

$$\begin{aligned} P_c &= 1 - \sum_{n=1}^{N}\sum_{k=1}^{K}\frac{\alpha}{N}(1-\frac{1}{N})^{(N-1)} \\ &= 1 - N.K.\frac{\alpha}{N}(1-\frac{1}{N})^{(N-1)} \\ &= 1 - \alpha K(1-\frac{1}{N})^{(N-1)}. \end{aligned} \tag{6.13}$$

Hence, the number of successful samples for a given arm over a period of time t is expressed as follows:

$$N_s = (1-P_c)\frac{t}{K} = \alpha t(1-\frac{1}{N})^{(N-1)}, \tag{6.14}$$

For a period of T^* time slots, we have $N_s = l$ which in turn results in:

$$T^* = \frac{l}{\alpha(1-\frac{1}{N})^{(N-1)}}. \tag{6.15}$$

□

lemma 3. The expected regret incurred over a horizon T by N devices employing the proposed the ϵ-decreasing MFG algorithm over K arms is upper bounded as follows:

$$R_{\epsilon-decreasingMFG} = \mathcal{O}(\log T). \tag{6.16}$$

Proof. The regret accumulated during a period of time t can be analyzed as the sum of the regret incurred during the two phases, i.e. the exploitation phase $R_1(t)$ and the exploration phase $R_2(t)$. Once a given user pulls a random variable that allows it to explore and choose its best arm learned during a period of t time slots, it will not regret, i.e. $R_2(t) = 0$. On the other hand, the exploration probability for our proposed approach at time t is given as $\epsilon_t = \min(1, \frac{\epsilon_0}{t})$. Subsequently, for each device, the expected regret accumulated during the exploration phase over a period of time t can be given as

$$R_{2,i}(t) \leq \sum_{t'=1}^{t}\epsilon_{t'} = \epsilon_0 + \sum_{t'=\epsilon_0+1}^{t}\epsilon_{t'}. \tag{6.17}$$

The discrete sum can be approximated using an integral:

$$R_{2,i}(t) \leq \epsilon_0 + \epsilon_0 \int_{\epsilon_0}^{t-1} \frac{1}{x}\, dx = \epsilon_0 + \epsilon_0 \log(\frac{t-1}{\epsilon_0}). \tag{6.18}$$

Then, the total expected regret incurred by all MTDs is bounded by:

$$R_2(t) \leq N\epsilon_0 + N\epsilon_0 \log(\frac{t-1}{\epsilon_0}) \leq N\epsilon_0 + N\epsilon_0 \log(t) = \mathcal{O}(\log t). \tag{6.19}$$

□

6.3.4.2 Regret of UCB-based MFG algorithm

Based on [138], the expected regret accumulated by invoking the UCB-based Algorithm (7) over a time period of T' is upper bounded by:

$$R_{UCB-basedMFG} \leq 8N \log(T') \sum_{j=2}^{K} \frac{1}{\Delta_j} + 4N \sum_{j=2}^{K} \Delta_j \tag{6.20}$$

where $\Delta_j = r^* - r_j(a_j)$ is the deviation function that measures the instantaneous loss of playing an arm a_j by a player j.

In the next section, we present some numerical results to analyze the equilibrium behaviors of the proposed MAB approaches and demonstrate their effectiveness against the interference impact.

6.4 Simulation Results

In this section, we use extensive Matlab-based simulations to validate the proposed algorithms. Particularly, we consider a Hybrid NOMA system made up of N devices occupying K sub-carriers. At each time t, each user belongs to only one coalition and communicates with the BS via the sub-carrier assigned to that coalition. Simulation parameters introduced in Table 5.1 of Chapter 5.

Firstly, we assess the performance of the proposed ϵ-decreasing MFG Algorithm (6) by illustrating its convergence properties. Then, we provide comparisons between the two proposed approaches, i.e. the ϵ-decreasing MFG Algorithm (6) and the UCB-based MFG Algorithm (7), and other existing techniques in the literature.

6.4.1 Performance metrics

In order to spotlight the features of our MFG-based MAB techniques, we adopt the following metrics:

- Number of active devices per coalition: is the average number of devices that can transmit simultaneously in each group at each time slot.
- Packet success rate: is the ratio between the number of MTDs whose packets have been successfully decoded and the number of active MTDs that decided to transmit.
- Average transmission rate: is calculated as the ratio of the number of users whose packets have been successfully decoded to the total number of MTDs in the system.
- Average utility: we assume that the utility function is calculated only when the user satisfies the SINR requirement which means when its SINR is higher than the SINR threshold γ_{th}. In another word, the user has a utility value if the BS succeed in decoding its signals upon executing the SIC, otherwise it has no utility. Thereby, the average utility is the ratio between the utility of the users whose signals have been successfully retrieved by the BS and the total number of MTDs.
- Average energy: Similar to the calculation of the average utility, the average energy is calculated by considering the energy consumed when a device achieves a successful transmission.

6.4.2 Behavior of the ϵ-decreasing MFG approach at the equilibrium

Throughout this section, we evaluate our proposed ϵ-decreasing MFG technique for multiple Hybrid NOMA scenarios in which we have N users: $N = 2000$, $N = 4000$, $N = 6000$, $N = 8000$ and $N = 10000$, transmitting over $K = 20$ sub-carriers within a time period $T = 0.3$ s (i.e. 30 LTE frames). The ϵ-decreasing MFG algorithm is executed with the exploration parameter $\epsilon_0 = 20$.

We start by showing the packet success rate over time slots in Figure 6.2. It is significantly interesting to observe that the rate settles at $t = 10$ ms to about 0.78 (78% of success rate). Furthermore, we can clearly see that this rate stagnates at

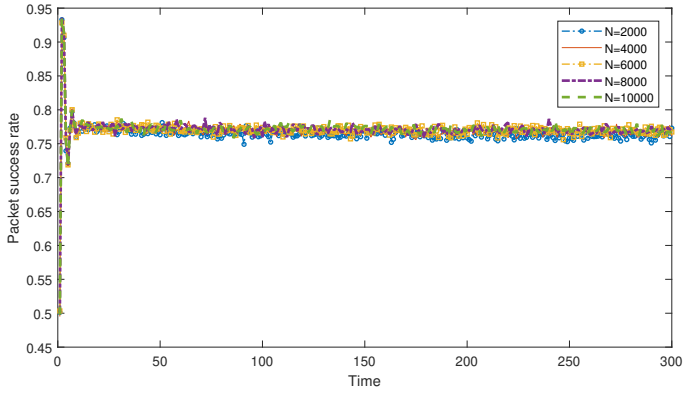

Figure 6.2: Packet success rate with respect to T with $K = 20$.

the same value for the different cases. Thus, each MTD has the same chance of successfully sending its message regardless of the network size. Interestingly, these results reflect a significant interference management achieved through the appropriate expression of the proposed access probability $p_t = \frac{\alpha}{N}$ which ensures that the same average number of devices are active in a given time slot. Consequently, our proposed scheme can reduce the performance drop witnessed by almost all existing grant-free schemes, especially in dense scenarios.

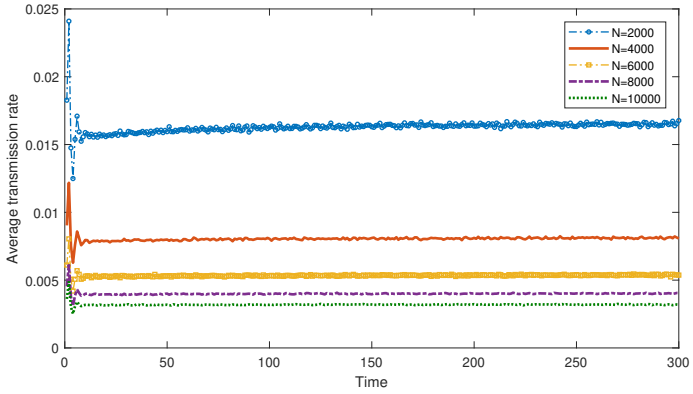

Figure 6.3: Average transmission rate with respect to T with $K = 20$.

In Figure 6.3, we depict the average transmission rate with respect to the time slots. In contrast to Figure 6.2, this rate stagnates at different values for the different network sizes, since it depends on the total number of users in the system. Thus, the highest

value is reached when the network is the most sparse, i.e., $N = 2000$. Then, the average transmission rate decreases as the network becomes denser. This is mainly due to the fact that the interference effects become more challenging in the denser network.

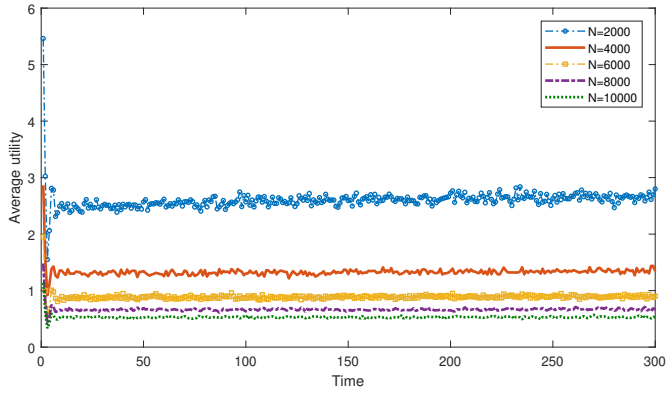

Figure 6.4: Average utility with respect to T with $K = 20$.

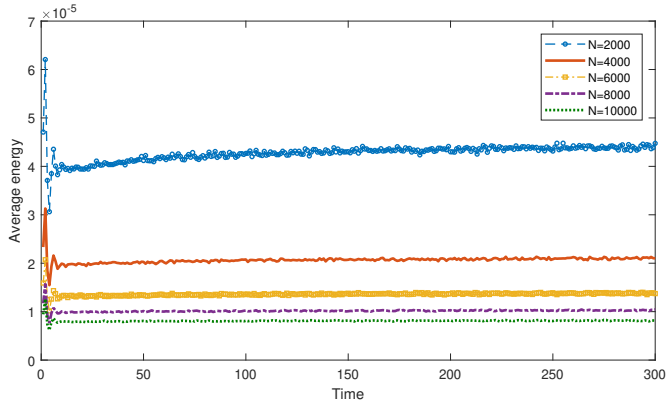

Figure 6.5: Average energy with respect to T with $K = 20$.

Now, we measure the average utility as well as the average energy in Figure 6.4 and 6.5 respectively. It is interesting to note that these figures have the similar equilibrium behavior to what we have shown in Figure 6.3, which emphasizes the convergence of the ϵ-decreasing MFG approach. Hence, the proposed technique settles down at the point where the players which achieve successful transmissions, meet their desired goal of maximizing their utilities with less energy consumption. Interestingly, the highest

utility value is achieved when $N = 2000$, but this results also in a higher energy consumption than what can be observed in the other cases. These behaviors are achieved since the average utility and the average energy consumption are respectively obtained by calculating the utility values and the consumed energy of the MTDs that have successfully transmitted, and then averaging over the total number of users in the system. It is worth noting that we have used the same model of Chapter 4 to evaluate the energy consumption of the proposed approach.

6.4.3 Comparison

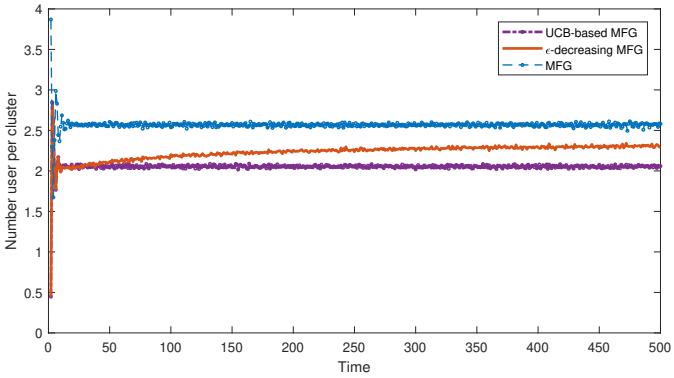

Figure 6.6: Number of active users per cluster with respect to T with $K = 20$.

Now, we provide a comparison between the ϵ-decreasing MFG algorithm, the second proposed learning approach, i.e. UCB-based MFG algorithm, and the basic deterministic algorithm of MFG approach developed in Chapter 5. The simulation results obtained for this comparison, are devoted to the case of $N = 2000$ devices sharing $K = 20$ sub-carriers over a training period duration of $T = 500$ time slots.

In Figure 6.6, we illustrate the average number of active devices that can transmit simultaneously per group with respect to the time slots. Upon comparing the proposed ϵ-decreasing MFG and UCB-based MFG algorithms, we observe that the former yields a higher value than the latter. However, this number reaches its highest value when the MFG is adopted. This is mainly due to the fact that the MFG is a deterministic approach, which means that each device has no other option than to join the cluster corresponding to its best channel. In contrast, our proposed approaches allow each device to choose its coalition according to the ϵ-decreasing greedy or UCB algorithms.

Indeed, after making its choice, each user first estimates its interference level $\tilde{I}$ according to the equation (6.2) based on the MFI received from the BS for the chosen coalition. Then, the user calculates its power level as in the equation (6.3), in response to the estimated interference. Since the MFG technique requires each user to join the coalition that corresponds to its highest channel gain, its transmit power is likely to be less than the maximum transmit power, meaning that the user is able to cope with the estimated interference level $\tilde{I}$ by having an acceptable power level. On the other hand, by invoking the MFG-based MAB algorithms, the user may join a coalition that corresponds to a lower channel gain. In fact, facing an important interference level while having a lower channel gain may require much more power than the device can handle, i.e. a power level higher than the maximum transmit power. Therefore, it withdraws to play this arm. Consequently, the proposed algorithms result in a lower number of active devices per coalition than the MFG approach.

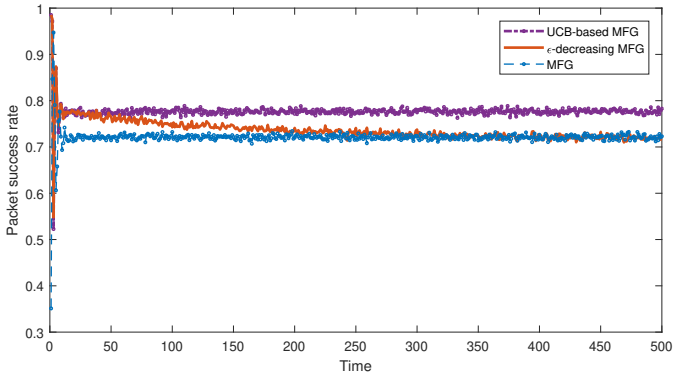

Figure 6.7: Packet success rate with respect to T with $K = 20$.

Figure 6.7 and 6.8 display respectively the packet success rate and the average transmission rate for the different techniques. Interestingly, as shown in Figure 6.7, UCB-based MFG achieves the highest value of the packet success rate, about 78% of success compared to the other techniques. Nevertheless, the MFG outperforms the MFG-based MAB algorithms in terms of the average transmission rate, as depicted in Figure 6.8. Indeed, when a user transmits its packet, it has more chance to achieve a successful transmission by adopting the UCB-based MFG algorithm than invoking the other approaches. But, the MFG allows more users to successfully transmit their packets than the proposed approaches.

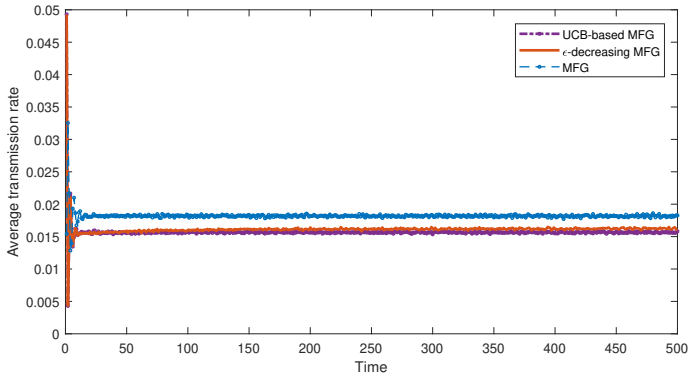

Figure 6.8: Average transmission rate with respect to T with $K = 20$.

These results are somehow intuitive, since the packet success rate is calculated as the ratio of the number of devices that successfully transmitted to the number of users that transmitted. By contrast, the average transmission rate is measured as the ratio between the number of devices that have succeeded in transmitting and the total number of devices in the system. Indeed, as shown in Figure 6.6, the number of active users per cluster reaches its highest value when MFG is invoked, which means that we have more devices playing MFG than the other techniques, allowing it to achieve the highest value in terms of average transmission rate as in Figure 6.8. On the other hand, the ϵ-decreasing MFG results in a higher number of active users per cluster than the UCB-based MFG algorithm, as represented in Figure 6.6, but the latter guarantees a higher value in terms of the packet success rate than the ϵ-decreasing MFG algorithm, as shown in Figure 6.7.

Now, we are interested in comparing the different approaches in terms of the average utility and the average energy consumption depicted in Figures 6.9 and 6.10 respectively. It can be concluded from these figures that although the MFG approach outperforms the two proposed MFG-based MAB algorithms in terms of the average utility, it requires much more energy consumption to reach this higher utility. Unsurprisingly, upon comparing the two MFG-based MAB algorithms, we clearly observe that the UCB-based MFG achieves a higher average utility and a lower average energy consumption than the ϵ-decreasing MFG algorithm.

The latest simulation results reveal the performance comparison of the proposed UCB-based MFG approach with the MFG technique as well as the Bi-level theoretical

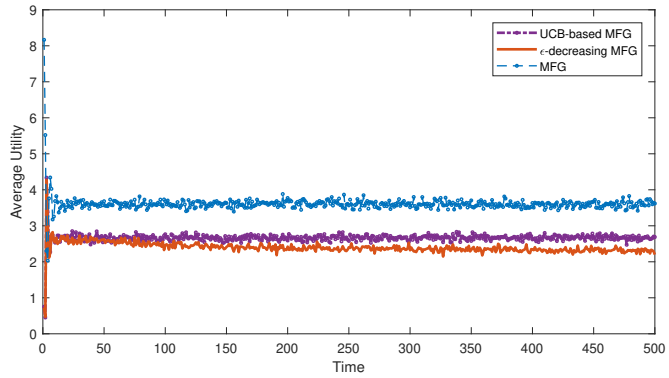

Figure 6.9: Average utility with respect to T with $K = 20$.

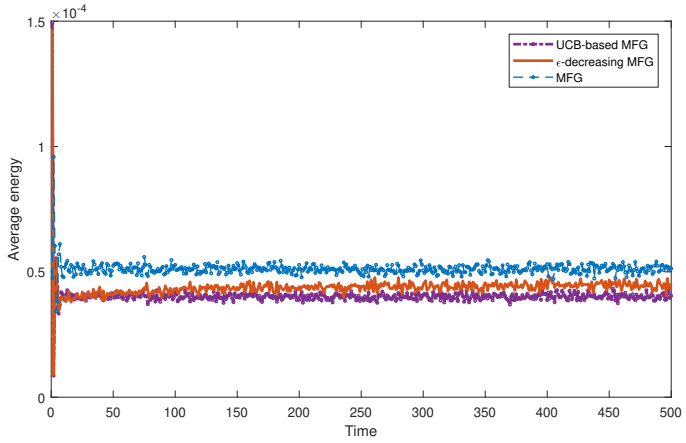

Figure 6.10: Average energy consumption with respect to T with $K = 20$.

framework developed in Chapter 4 and the NM-ALOHA game investigated in [100]. In Figure 6.11, we display the packet success rate as a function of the number of RBs K and the number of MTDs N. Clearly, this rate decreases as the network becomes denser for both the Bi-level game and the NM-ALOHA game, while it remains stable for the UCB-based MFG algorithm and the MFG framework. As explained above, by using the proposed access probability p_t, we are able to achieve an effective interference management that in turn results in mitigating the performance drop faced by almost all the proposed grant-free techniques, especially in very dense network. Besides, this performance comparison in terms of the packet success rate is highly spotlighted in Figure 6.12 which represents the case where the N devices share $K = 20$ RBs. As

we can clearly observe, the packet success rate is considerably enhanced with the proposed UCB-based MFG algorithm, reaching up $6,73\%$ of improvement compared to the MFG technique. Consequently, it is interesting to highlight that facing the interference effects, the UCB-based MFG technique provides much more robustness than the MFG approach, which accentuates the benefit of adopting the proposed MAB-based approach.

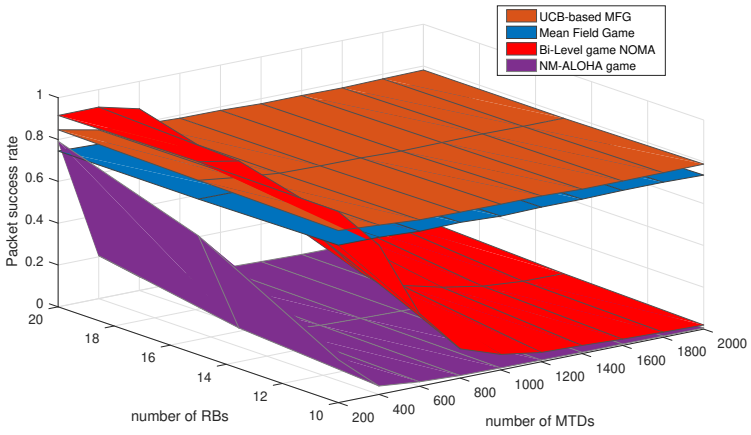

Figure 6.11: Packet success rate for different N and K.

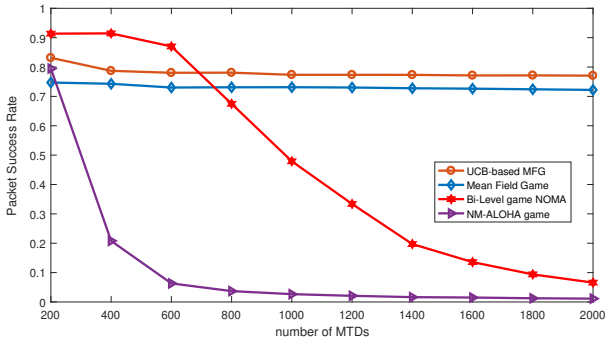

Figure 6.12: Packet success rate versus N when $K = 20$.

Finally, Figure 6.13 is devoted to illustrating the variation of the average utility for the scenario of $N = 2000$ devices as the number of sub-carriers K increases. As we can clearly see, the proposed UCB-based MFG algorithm improves the average utility by $24,434\%$ and $265,4\%$ against the Bi-level game and the NM-ALOHA game,

respectively. Nevertheless, the MFG approach outperforms all other techniques. The reason behind this is mainly related to the choice of the coalition. Indeed, it has been investigated in [31] that the utility function for a given user is maximized when it transmits on its best channel. Since MFG approach requires each user to join the coalition corresponding to its best channel and then transmit over the associated sub-carrier, it is unsurprisingly that the MFG approach achieves a higher average utility than the UCB-based MFG algorithm, wherein each device can choose another sub-carrier rather than its best sub-carrier.

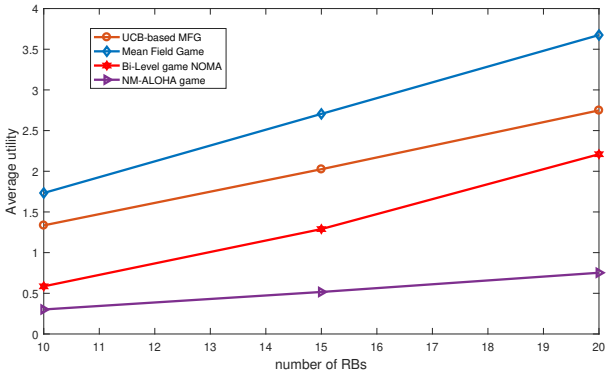

Figure 6.13: Average utility versus K when $N = 2000$.

6.5 Conclusion

In this chapter, a Hybrid NOMA network has been investigated in a dense deployment context in which a large population of MTDs is split up into independent coalitions. We derived a bi-level learning to jointly address the user grouping and the power control problem. Firstly, we modeled dense scenarios using a MFG framework while taking into consideration the effect of the collective behavior of devices. Then, we exploited the RL-based MAB approach with the aim of paving the way for an autonomous decision-making process of the players participating in the formulated MFG. Thereafter, we derived two MFG underlying MAB algorithms that allow the MTDs to arrange themselves into coalitions and regulate their power levels based on a brief feedback received from the BS. Our simulation results emphasize the equilibrium behaviors of proposed MFG-based MAB approaches and demonstrate their effective-

ness against the interference effects, resulting from the densely deployed networks.

Chapter 7

Conclusion and perspectives

7.1 Conclusion

Facing the exponential demand for massive connectivity and the scarcity of available resources, next-generation wireless networks have to meet very challenging performance targets in terms of providing massive access and ensuring a higher spectral efficiency. In this vein, designing a sophisticated multiple access technique has been considered as one of the key solutions to deal with the drastic increase in the number of MTDs. Specifically, NOMA has intrigued researchers as an emerging technology to meet the above-mentioned challenges. Nevertheless, serving a large number of users simultaneously on the same resources comes at the cost of a high inter-user interference that may deteriorate the performance gains of NOMA networks. Thereupon, it is highly challenging to ask all users to jointly communicate with the BS in a non-orthogonal manner through common resources. Alternatively, a Hybrid NOMA approach has been proposed by availing the ability of NOMA to be concordant with a conventional OMA scheme. Indeed, hybrid NOMA networks require that users be arranged into orthogonal groups and that the members of each group share the same RB in order to boost the overall system performance.

In this regard, this dissertation strives to find novel NOMA scenarios while addressing the resource management problems. To this end, we have focused on how to meticulously form multiple user groups and smartly develop power allocation strategies with the goal of achieving a meaningful trade-off between the benefits offered by adopting the NOMA technique and the emerging interference effects resulting from simultaneous user transmissions. Since the user grouping is intertwined with the power control

optimization, impactful tools are needed to solve the combined problems. For this sake, we have firstly invoked the game theoretic framework, then we have exploited RL algorithms to further enhance the network performance by managing more efficiently the resource allocation among devices in different NOMA clusters.

We have started by addressing the above-mentioned optimization problems using a Bi-level game consisting of a cooperative Hedonic game on top of a non-cooperative power control. Indeed, Hedonic game has been invoked to lay out the coalitional formation process by allowing devices to self-organize into several coalitions. Once the coalitions are formed, the users belonging to the same coalition use the non-cooperative game to ensure the disparity in power levels among them so that they can simultaneously deliver their packets over the associated RB. We have numerically proven that the proposed Bi-level game significantly increases the NOMA capacity and strikes an attractive trade-off between the successful transmission rate and the energy consumption compared to other existing techniques in the literature.

Thereafter, we have shifted our interest to the MFG framework to deal with dense scenarios, wherein the impact of a particular player' move on the overall system is negligible but the effect of the collective behavior of a large population on each user is significant. Furthermore, we have proposed a distributed power control algorithm that takes into account the interference effects to enable each user to effectively make its own decision. In this way, the devices can appropriately determine their power levels according to the limited feedback received from the BS, which greatly simplifies the game resolution. Indeed, the analysis of the game has been governed by the combination of the HJB and FPK equations and then the finite difference method has been invoked to solve the MFG game. Thus we have shown that the ping-pong interaction between the HJB and FPK equations leads to the MFE. Our simulation results have highlighted the equilibrium behaviors of the formulated game and demonstrated the effectiveness of the proposed approach compared to other works, especially when a dense network is envisaged.

Finally, we have extended the proposed MFG by adopting the RL tool, namely the MAB approach with the aim of making the devices able to suitably determine by themselves their coalitions. As a result, we have derived a bi-level learning in which the MAB technique is on the top of the MFG framework, with an eye toward jointly

solving the user grouping and power control problems. Particularly, we have developed decision-making techniques using both the ϵ-decreasing greedy and UCB algorithms to deal with the strategy selection problem for the MAB approach. In fact, the devices can self-organize into multiple groups with the aid of MAB algorithms. Once each player joins its cluster, it autonomously regulates its transmit power according to the received MFI from the BS using the MFG. Simulation results have been given to evaluate the potential performance gains of the proposed learning process by illustrating its convergence and then comparing it to the other approaches proposed in this thesis.

7.2 Perspectives

7.2.1 Interplay between NOMA and Multiple Antennas Techniques

Throughout this thesis, our proposed approaches involved a single antenna. However, the Multi-Input Multi-Output (MIMO) schemes are of a substantial importance to bring an additional dimension to the NOMA technique via the spatial domain, which can further improve the system performance. Indeed, combining MIMO systems with NOMA approaches can significantly enhance the spectrum utilization by exploiting the spatial diversity gain of MIMO and the features of NOMA in allowing multiple transmissions [139].

More precisely, to model a MIMO-aided NOMA system, we can assume that the BS is equipped with a number M of antennas to serve N single-antenna users, where $N > M$, through several beams. Actually, the proposed Hybrid NOMA scenario in this thesis mainly relies on idea of arranging the users into several coalitions and then each which uses one sub-carrier to deliver the signals of its members. In this sense, the application of MIMO can greatly comply with the Hybrid NOMA approach by forming M NOMA groups and then assigning beams to these groups so that the users falling in each group share a single beam in the PD-NOMA basis [9]. Thus, the interplay between MIMO and NOMA techniques is an interesting approach that deserves a deep dive to appropriately form a cluster-based structure.

Worryingly, additional challenges are emerged once NOMA is amalgamated with MIMO [140]. Indeed, assigning a large number of users to different groups and beams in the MIMO-assisted NOMA approach is a tedious task. Thereby, the user grouping

along with the beamforming or the precoding must be carefully designed. Meanwhile, since the power control is one of the critical pillars in the design of NOMA systems, the MIMO-NOMA scheme also requires a special attention to the power allocation problem among users sharing a NOMA group in order to ensure successful transmissions.

7.2.2 Security Provisioning for NOMA

Security issues in wireless transmission systems is a great challenge due to the broadcast characteristic of the wireless medium that makes communications between users vulnerable to eavesdropper attacks. Specifically, when NOMA is invoked, the application of the SIC process presents a high security risk since one user is able to decode the signals of the other users. On the other hand, the communications between the BS and the users can be intercepted and overheard by wiretapper and unauthorized users. Hence, it is of the utmost significance to thoroughly study the security provisioning issues for NOMA-based networks in order to take advantage of the NOMA features of simultaneously serving multiple users without spoiling the transmission security.

Traditionally, cryptographic protocols have been used and designed in the upper layers at the expensive of increased overhead and computational complexity. Alternatively, the Physical Layer Security (PLS) has been proposed as a leading research protocol to enhance the confidentiality of wireless communications. Besides, the amalgam between the PLS and the NOMA technique has been considered as a promising approach that sparked a considerable interest of researchers in an effort to combat the eavesdropping [141, 142]. The idea of PLS is to provide confidential transmissions by exploiting the characteristics of users' channel gains in the physical layer in order to distinguish the legitimate from eavesdropping receivers. One popular solution used by PLS is to generate an artificial noise at the transmitter to degrade the eavesdropper channel and then ensure a high channel capacity for the desired receiver. Thus, it is of great significance to study the security issues of the approaches proposed in this thesis and examine the application of the PLS technique to the considered Hybrid NOMA networks.

7.2.3 Imperfect Channel Estimation

Throughout this thesis, we have assumed that the BS can acquire the perfect channel state information (CSI) for all users. In fact, the availability of the CSI at the BS is principally required to perform the order of users, the power control and then execute the SIC process in order to correctly detect and decode signals of different users. Nevertheless, this assumption is very strong and cannot be readily achieved in real-time NOMA networks, as the continuous availability of the perfect CSI requires a pilot based training process, resulting in increased overhead and complexity and then leading to a delayed feedback. Hence, practical NOMA systems should operate under channel uncertainties and imperfect CSI scenarios. But this in turn leads to an ambiguous order of users and an inaccurate power allocation that affect the accuracy of the decoding procedure using the SIC. Hence, studying the impact of the imperfect CSI on NOMA networks results in a shift in the way the resource allocation issues related to NOMA are addressed, including the user selection, the user grouping and the power control. Consequently, it is of paramount importance to develop robust resource allocation strategies while keeping in mind the imperfect CSI acquisition in order to suitably deal with the challenges of adopting the NOMA concept under practical scenarios.

Bibliography

[1] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE communications surveys & tutorials*, 17(4):2347–2376, 2015.

[2] Vangelis Gazis. A survey of standards for machine-to-machine and the internet of things. *IEEE Communications Surveys & Tutorials*, 19(1):482–511, 2016.

[3] Shree Krishna Sharma and Xianbin Wang. Toward massive machine type communications in ultra-dense cellular iot networks: Current issues and machine learning-assisted solutions. *IEEE Communications Surveys & Tutorials*, 22(1):426–471, 2019.

[4] CISCO. *Cisco Annual Internet Report 2018–2023*, 2020.

[5] Zhengquan Zhang, Yue Xiao, Zheng Ma, Ming Xiao, Zhiguo Ding, Xianfu Lei, George K Karagiannidis, and Pingzhi Fan. 6g wireless networks: Vision, requirements, architecture, and key technologies. *IEEE Vehicular Technology Magazine*, 14(3):28–41, 2019.

[6] 3GPP. Service requirements for machine-type communications. *Tech. Rep.*, 2014.

[7] Zhu Han, Husheng Li, and Wotao Yin. *Compressive sensing for wireless networks*. Cambridge University Press, 2013.

[8] Linglong Dai, Bichai Wang, Yifei Yuan, Shuangfeng Han, I Chih-Lin, and Zhaocheng Wang. Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends. *IEEE Communications Magazine*, 53(9):74–81, 2015.

[9] Yuanwei Liu, Zhijin Qin, Maged Elkashlan, Zhiguo Ding, Arumugam Nallanathan, and Lajos Hanzo. Non-orthogonal multiple access for 5g and beyond. *Proceedings of the IEEE*, 105(12):2347–2381, 2017.

[10] Aamina Akbar, Sobia Jangsher, and Farrukh A Bhatti. Noma and 5g emerging technologies: A survey on issues and solution techniques. *Computer Networks*, 190:107950, 2021.

[11] Heunchul Lee, Sungsoo Kim, and Jong-Han Lim. Multiuser superposition transmission (must) for lte-a systems. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016.

[12] SM Riazul Islam, Nurilla Avazov, Octavia A Dobre, and Kyung-Sup Kwak. Power-Domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges. *IEEE Communications Surveys & Tutorials*, 19(2):721–742, 2016.

[13] Jianyue Zhu, Jiaheng Wang, Yongming Huang, Shiwen He, Xiaohu You, and Luxi Yang. On optimal power allocation for downlink non-orthogonal multiple access systems. *IEEE Journal on Selected Areas in Communications*, 35(12):2744–2757, 2017.

[14] Omar Maraqa, Aditya S Rajasekaran, Saad Al-Ahmadi, Halim Yanikomeroglu, and Sadiq M Sait. A survey of rate-optimal power domain noma with enabling technologies of future wireless networks. *IEEE Communications Surveys & Tutorials*, 22(4):2192–2235, 2020.

[15] B. Di, L. Song, Y. Li, and S. Zhang. Trellis coded modulation for Code-Domain non-orthogonal multiple access Networks. In *2018 IEEE International Conference on Communications*, pages 1–6, May 2018.

[16] M. Gan, J. Jiao, L. Li, S. Wu, and Q. Zhang. Performance analysis of uplink uncoordinated Code-Domain NOMA for SINs. In *International Conference on Wireless Communications and Signal Processing*, 2018.

[17] Zilong Liu and Lie-Liang Yang. Sparse or dense: A comparative study of code-domain noma systems. *IEEE Transactions on Wireless Communications*, 20(8):4768–4780, 2021.

[18] Zhiguo Ding, Pingzhi Fan, and H Vincent Poor. Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions. *IEEE Transactions on Vehicular Technology*, 65(8):6010–6023, 2015.

[19] Mohammed Elbayoumi, Walaa Hamouda, and Amr Youssef. A hybrid noma/oma scheme for mtc in ultra-dense networks. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pages 1–6. IEEE, 2020.

[20] Richard J La and Venkat Anantharam. A game-theoretic look at the gaussian multiaccess channel. *DIMACS series in discrete mathematics and theoretical computer science*, 66:87–106, 2004.

[21] Suhas Mathur, Lalitha Sankar, and Narayan B Mandayam. Coalitions in cooperative wireless networks. *IEEE Journal on Selected areas in Communications*, 26(7):1104–1115, 2008.

[22] Zhu Han and H Vincent Poor. Coalition games with cooperative transmission: a cure for the curse of boundary nodes in selfish packet-forwarding wireless networks. *IEEE transactions on communications*, 57(1):203–213, 2009.

[23] Heyu Luan, Yitao Xu, Dianxiong Liu, Zhiyong Du, Huiming Qian, Xiaodu Liu, and Xiaobing Tong. Energy efficient task cooperation for multi-uav networks: A coalition formation game approach. *IEEE Access*, 8:149372–149384, 2020.

[24] Tianyu Wang, Lingyang Song, Zhu Han, and Walid Saad. Distributed cooperative sensing in cognitive radio networks: An overlapping coalition formation approach. *IEEE Transactions on Communications*, 62(9):3144–3160, 2014.

[25] Jinming Hu, Wei Heng, Yaping Zhu, Gang Wang, Xiang Li, and Jing Wu. Overlapping coalition formation games for joint interference management and resource allocation in d2d communications. *IEEE Access*, 6:6341–6349, 2018.

[26] Krzysztof R Apt and Andreas Witzel. A generic approach to coalition formation. *International game theory review*, 11(03):347–367, 2009.

[27] Anna Bogomolnaia, Matthew O Jackson, et al. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

[28] Robert J Aumann and Jacques H Dreze. Cooperative games with coalition structures. *International Journal of game theory*, 3(4):217–237, 1974.

[29] John F Nash Jr. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

[30] Drew Fudenberg and Jean Tirole. *Game theory*. MIT press, 1991.

[31] M. Haddad, P. Wiecek, O. Habachi, and Y. Hayel. On the two-user multi-carrier joint channel selection and power control game. *IEEE Transactions on Communications*, 64(9):3759–3770, 2016.

[32] Shaima' S. Abidrabbu and Hüseyin Arslan. Energy-efficient resource allocation for 5g cognitive radio noma using game theory. In *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–5, 2021.

[33] Yousef Ali Al-Gumaei, Nauman Aslam, Ahamed Mohammed Al-Samman, Tawfik Al-Hadhrami, Kamarul Noordin, and Yousef Fazea. Non-cooperative power control game in d2d underlying networks with variant system conditions. *Electronics*, 8(10):1113, 2019.

[34] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266. Springer, 2011.

[35] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[36] Yves Achdou and Italo Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010.

[37] Hamidou Tembine and Minyi Huang. Mean field difference games: Mckean-vlasov dynamics. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 1006–1011. IEEE, 2011.

[38] Yunlong Cai, Zhijin Qin, Fangyu Cui, Geoffrey Ye Li, and Julie A McCann. Modulation and multiple access for 5g networks. *IEEE Communications Surveys & Tutorials*, 20(1):629–646, 2017.

[39] Raymond Steele and Lajos Hanzo. *Mobile radio communications: Second and third generation cellular and WATM systems: 2nd*. IEEE Press-John Wiley, 1999.

[40] Klein S Gilhousen, Irwin M Jacobs, Roberto Padovani, Andrew J Viterbi, Lindsay A Weaver, and Charles E Wheatley. On the capacity of a cellular cdma system. *IEEE transactions on vehicular technology*, 40(2):303–312, 1991.

[41] Erik Dahlman, Stefan Parkvall, and Johan Skold. *4G: LTE/LTE-advanced for mobile broadband*. Academic press, 2013.

[42] David Astély, Erik Dahlman, Anders Furuskär, Ylva Jading, Magnus Lindström, and Stefan Parkvall. Lte: the evolution of mobile broadband. *IEEE Communications magazine*, 47(4):44–51, 2009.

[43] Richard van Nee and Ramjee Prasad. *OFDM for wireless multimedia communications*. Artech House, Inc., 2000.

[44] Junyi Li, Xinzhou Wu, and Rajiv Laroia. *OFDMA mobile broadband communications: A systems approach*. Cambridge University Press, 2013.

[45] Yuya Saito, Yoshihisa Kishiyama, Anass Benjebbour, Takehiro Nakamura, Anxin Li, and Kenichi Higuchi. Non-orthogonal multiple access (noma) for cellular future radio access. In *2013 IEEE 77th vehicular technology conference (VTC Spring)*, pages 1–5. IEEE, 2013.

[46] Muhammad Basit Shahab and Soo Young Shin. A time sharing based approach to accommodate similar gain users in noma for 5g networks. In *2017 IEEE 42nd Conference on Local Computer Networks Workshops (LCN Workshops)*, pages 142–147. IEEE, 2017.

[47] Stelios Timotheou and Ioannis Krikidis. Fairness for non-orthogonal multiple access in 5g systems. *IEEE signal processing letters*, 22(10):1647–1651, 2015.

[48] Thomas Cover. Broadcast channels. *IEEE Transactions on Information Theory*, 18(1):2–14, 1972.

[49] Thomas M Cover. Thomas. elements of information theory. *Wiley Series in Telecommunications*, 1991.

[50] Reza Hoshyar, Ferry P Wathan, and Rahim Tafazolli. Novel low-density signature for synchronous cdma systems over awgn channel. *IEEE Transactions on Signal Processing*, 56(4):1616–1626, 2008.

[51] Reza Hoshyar, Razieh Razavi, and Mohammad Al-Imari. Lds-ofdm an efficient multiple access technique. In *2010 IEEE 71st Vehicular Technology Conference*, pages 1–5. IEEE, 2010.

[52] AL-Imari Mohammed, Muhammad Ali Imran, and Rahim Tafazolli. Low density spreading for next generation multicarrier cellular systems. In *2012 International Conference on Future Communication Networks*, pages 52–57. IEEE, 2012.

[53] Nikopour Hosein and Baligh Hadi. Sparse code multiple access. In *Proc. IEEE 24th Int. Symp. Pers. Indoor Mobile Radio Commun.(PIMRC)*, pages 332–336, 2013.

[54] Yijin Pan, Cunhua Pan, Zhaohui Yang, and Ming Chen. Resource allocation for d2d communications underlaying a noma-based cellular network. *IEEE Wireless Communications Letters*, 7(1):130–133, 2017.

[55] Jingjing Zhao, Yuanwei Liu, Kok Keong Chai, Yue Chen, and Maged Elkashlan. Joint subchannel and power allocation for noma enhanced d2d communications. *IEEE Transactions on Communications*, 65(11):5081–5094, 2017.

[56] Taehyun Yoon, Tien Hoa Nguyen, Xuan Tung Nguyen, Daeseung Yoo, Byungtae Jang, et al. Resource allocation for noma-based d2d systems coexisting with cellular networks. *IEEE Access*, 6:66293–66304, 2018.

[57] Hanyu Zheng, Shujuan Hou, Hai Li, Zhengyu Song, and Yuanyuan Hao. Power allocation and user clustering for uplink mc-noma in d2d underlaid cellular networks. *IEEE Wireless Communications Letters*, 7(6):1030–1033, 2018.

[58] Mahmoud M Selim, Mohamed Rihan, Yatao Yang, Lei Huang, Zhi Quan, and Junxian Ma. On the outage probability and power control of d2d underlaying noma uav-assisted networks. *IEEE Access*, 7:16525–16536, 2019.

[59] Shanshan Yu, Wali Ullah Khan, Xiaoqing Zhang, and Ju Liu. Optimal power allocation for noma-enabled d2d communication with imperfect sic decoding. *Physical Communication*, 46:101296, 2021.

[60] Jingjing Zhao, Yuanwei Liu, Kok Keong Chai, Arumugam Nallanathan, Yue Chen, and Zhu Han. Spectrum allocation and power control for non-orthogonal multiple access in hetnets. *IEEE Transactions on Wireless Communications*, 16(9):5825–5837, 2017.

[61] Dadong Ni, Li Hao, Quang Thanh Tran, and Xiaomin Qian. Power allocation for downlink noma heterogeneous networks. *IEEE Access*, 6:26742–26752, 2018.

[62] Fang Fang, Julian Cheng, and Zhiguo Ding. Joint energy efficient subchannel and power optimization for a downlink noma heterogeneous network. *IEEE Transactions on Vehicular Technology*, 68(2):1351–1364, 2018.

[63] Yan Sun, Derrick Wing Kwan Ng, Zhiguo Ding, and Robert Schober. Optimal joint power and subcarrier allocation for mc-noma systems. In *2016 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2016.

[64] Jinho Choi. NOMA-based random access with multichannel ALOHA. *IEEE Journal on Selected Areas in Communications*, 35(12):2736–2743, 2017.

[65] Lou Salaün, Marceau Coupechoux, and Chung Shue Chen. Joint subcarrier and power allocation in noma: Optimal and approximate algorithms. *IEEE Transactions on Signal Processing*, 68:2215–2230, 2020.

[66] Alemu Jorgi Muhammed, Zheng Ma, Panagiotis D Diamantoulakis, Li Li, and George K Karagiannidis. Energy-efficient resource allocation in multi-carrier noma systems with fairness. *IEEE Transactions on Communications*, 67(12):8639–8654, 2019.

[67] Yaru Fu, Lou Salaün, Chi Wan Sung, and Chung Shue Chen. Subcarrier and power allocation for the downlink of multicarrier noma systems. *IEEE Transactions on Vehicular Technology*, 67(12):11833–11847, 2018.

[68] Zhaohui Yang, Wei Xu, Hao Xu, Jianfeng Shi, and Ming Chen. Energy efficient non-orthogonal multiple access for machine-to-machine communications. *IEEE communications letters*, 21(4):817–820, 2016.

[69] Mahyar Shirvanimoghaddam, Massimo Condoluci, Mischa Dohler, and Sarah J Johnson. On the fundamental limits of random non-orthogonal multiple access in cellular massive iot. *IEEE Journal on Selected Areas in Communications*, 35(10):2238–2252, 2017.

[70] Shujun Han, Xiaodong Xu, Xiaofeng Tao, and Ping Zhang. Joint power and sub-channel allocation for secure transmission in noma-based mmtc networks. *IEEE Systems Journal*, 13(3):2476–2487, 2019.

[71] Shujun Han, Xiaodong Xu, Zilong Liu, Pei Xiao, Klaus Moessner, Xiaofeng Tao, and Ping Zhang. Energy-efficient short packet communications for uplink noma-based massive mtc networks. *IEEE Transactions on Vehicular Technology*, 68(12):12066–12078, 2019.

[72] Zhaohui Yang, Yijin Pan, Wei Xu, Rui Guan, Yinlu Wang, and Ming Chen. Energy efficient resource allocation for machine-to-machine communications with noma and energy harvesting. In *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 145–150. IEEE, 2017.

[73] Ali Shahini and Nirwan Ansari. Noma aided narrowband iot for machine type communications with user clustering. *IEEE Internet of Things Journal*, 6(4):7183–7191, 2019.

[74] Zeyu Zhang, Yanzhao Hou, Qiang Wang, and Xiaofeng Tao. Joint sub-carrier and transmission power allocation for mtc under power-domain noma. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2018.

[75] Jingrui Guo, Xiangyang Wang, Jingwen Yang, Jianhai Zheng, and Bingqiang Zhao. User pairing and power allocation for downlink non-orthogonal multiple access. In *2016 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6. IEEE, 2016.

[76] Wei Liang, Zhiguo Ding, Yonghui Li, and Lingyang Song. User pairing for downlink non-orthogonal multiple access networks using matching algorithm. *IEEE Transactions on communications*, 65(12):5319–5332, 2017.

[77] Xinyi Zhang, Jun Wang, Jintao Wang, and Jian Song. A novel user pairing in downlink non-orthogonal multiple access. In *2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5. IEEE, 2018.

[78] Phúc Hũu, Mohamed Amine Arfaoui, Sanaa Sharafeddine, Chadi M Assi, and Ali Ghrayeb. A low-complexity framework for joint user pairing and power control for cooperative noma in 5g and beyond cellular networks. *IEEE Transactions on Communications*, 68(11):6737–6749, 2020.

[79] Fan Jiang, Zesheng Gu, Changyin Sun, and Rongxin Ma. Dynamic user pairing and power allocation for noma with deep reinforcement learning. In *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2021.

[80] Hina Tabassum, Md Shipon Ali, Ekram Hossain, Md Hossain, Dong In Kim, et al. Non-orthogonal multiple access (noma) in cellular uplink and downlink: Challenges and enabling techniques. *arXiv preprint arXiv:1608.05783*, 2016.

[81] Muhammad Basit Shahab, Mohammad Irfan, Md Fazlul Kader, and Soo Young Shin. User pairing schemes for capacity maximization in non-orthogonal multiple access systems. *Wireless Communications and Mobile Computing*, 16(17):2884–2894, 2016.

[82] Salifou Mounchili and Soumaya Hamouda. Pairing distance resolution and power control for massive connectivity improvement in noma systems. *IEEE Transactions on Vehicular Technology*, 69(4):4093–4103, 2020.

[83] Md Shipon Ali, Hina Tabassum, and Ekram Hossain. Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems. *IEEE access*, 4:6325–6343, 2016.

[84] Jie Mei, Lei Yao, Hang Long, and Kan Zheng. Joint user pairing and power allocation for downlink non-orthogonal multiple access systems. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016.

[85] Soumendra Nath Datta and Suresh Kalyanasundaram. Optimal power allocation and user selection in non-orthogonal multiple access systems. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6. IEEE, 2016.

[86] Lipeng Zhu, Jun Zhang, Zhenyu Xiao, Xianbin Cao, and Dapeng Oliver Wu. Optimal user pairing for downlink non-orthogonal multiple access (noma). *IEEE Wireless Communications Letters*, 8(2):328–331, 2018.

[87] Yuh-Ren Tsai and Hsuan-An Wei. Quality-balanced user clustering schemes for non-orthogonal multiple access systems. *IEEE Communications Letters*, 22(1):113–116, 2017.

[88] Ken Long, Pengyu Wang, Wei Li, and Dejian Chen. Spectrum resource and power allocation with adaptive proportional fair user pairing for noma systems. *IEEE Access*, 7:80043–80057, 2019.

[89] Abdulkadir Celik, Ming-Cheng Tsai, Redha M Radaydeh, Fawaz S Al-Qahtani, and Mohamed-Slim Alouini. Distributed user clustering and resource allocation for imperfect noma in heterogeneous networks. *IEEE Transactions on Communications*, 67(10):7211–7227, 2019.

[90] Hanliang You, Zhiwen Pan, Nan Liu, and Xiaohu You. User clustering scheme for downlink hybrid noma systems based on genetic algorithm. *IEEE Access*, 8:129461–129468, 2020.

[91] Kaidi Wang, Zhiguo Ding, and Wei Liang. A game theory approach for user grouping in hybrid non-orthogonal multiple access systems. In *International Symposium on Wireless Communication Systems*, 2016.

[92] Mohammed W Baidas. A game-theoretic approach to distributed energy-efficiency maximization in uplink noma relay ad-hoc networks. In *2020 23rd International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pages 1–6. IEEE, 2020.

[93] Chi Wan Sung and Yaru Fu. A game-theoretic analysis of uplink power control for a non-orthogonal multiple access system with two interfering cells. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2016.

[94] Kashif Mehmood, Muhammad Tabish Niaz, and Hyung Seok Kim. Game-theoretic power control for energy constrained machine type communications. In *2019 15th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*, pages 99–106. IEEE, 2019.

[95] Mohammed W Baidas. Distributed energy-efficiency maximization in energy-harvesting uplink noma relay ad-hoc networks: Game-theoretic modeling and analysis. *Physical Communication*, 43:101188, 2020.

[96] Kaidi Wang, Wei Liang, Yi Yuan, Yuanwei Liu, Zheng Ma, and Zhiguo Ding. User clustering and power allocation for hybrid non-orthogonal multiple ac-

cess systems. *IEEE Transactions on Vehicular Technology*, 68(12):12052–12065, 2019.

[97] Ashok Kumar and Krishan Kumar. A game theory based hybrid noma for efficient resource optimization in cognitive radio networks. *IEEE Transactions on Network Science and Engineering*, 8(4):3501–3514, 2021.

[98] Ashok Kumar and Krishan Kumar. Energy-efficient resource optimization using game theory in hybrid noma assisted cognitive radio networks. *Physical Communication*, 47:101382, 2021.

[99] Kang Kang, Zhenni Pan, Jiang Liu, and Shigeru Shimamoto. A game theory based power control algorithm for future MTC NOMA Networks. In *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pages 203–208. IEEE, 2017.

[100] Jinho Choi. A game-theoretic approach for NOMA-ALOHA. In *IEEE European Conference on Networks and Communications*, pages 54–9, 2018.

[101] Henning Thomsen, Carles Navarro Manchón, and Bernard Henri Fleury. A traffic model for machine-type communications using spatial point processes. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–6. IEEE, 2017.

[102] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, New York, NY, USA, 2005.

[103] Farhad Meshkati, H Vincent Poor, Stuart C Schwartz, and Narayan B Mandayam. An energy-efficient approach to power control and receiver design in wireless data networks. *IEEE transactions on communications*, 53(11):1885–1894, 2005.

[104] Zhetao Li, YuXin Liu, Ming Ma, Anfeng Liu, Xiaozhi Zhang, and Gungming Luo. Msdg: A novel green data gathering scheme for wireless sensor Networks. *Computer Networks*, 142:223–239, 2018.

[105] Amy Nordrum. The internet of fewer things [news]. *IEEE Spectrum*, 53(10):12–13, 2016.

[106] Zheng Yang, Zhiguo Ding, Pingzhi Fan, and Naofal Al-Dhahir. A general power allocation scheme to guarantee quality of service in downlink and uplink noma systems. *IEEE transactions on wireless communications*, 15(11):7244–7257, 2016.

[107] M. S. Ali, H. Tabassum, and E. Hossain. Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems. *IEEE Access*, 4:6325–6343, 2016.

[108] N. Zhang, J. Wang, G. Kang, and Y. Liu. Uplink nonorthogonal multiple access in 5G systems. *IEEE Communications Letters*, 20(3):458–461, March 2016.

[109] Yaru Fu, Yi Chen, and Chi Wan Sung. Distributed power control for the downlink of multi-cell NOMA systems. *IEEE Transactions on Wireless Communications*, 16(9):6207–6220, 2017.

[110] Z. Ding, P. Fan, and H. V. Poor. Impact of user pairing on 5g nonorthogonal multiple-access downlink transmissions. *IEEE Transactions on Vehicular Technology*, 65(8):6010–6023, Aug 2016.

[111] Marie-Joséph a Youssef, Joumana Farah, Charbel Abdel Nour, and Catherine Douillard. Resource allocation in noma systems for centralized and distributed antennas with mixed traffic using matching theory. *IEEE Transactions on Communications*, 68(1):414–428, 2019.

[112] Ming Zeng, Wanming Hao, Octavia A Dobre, Zhiguo Ding, and H Vincent Poor. Power minimization for multi-cell uplink noma with imperfect sic. *IEEE Wireless Communications Letters*, 9(12):2030–2034, 2020.

[113] Zhu Han, Dusit Niyato, Walid Saad, Tamer Başar, and Are Hjørungnes. *Game theory in wireless and communication Networks: theory, models, and applications*. Cambridge university press, 2012.

[114] Olivier Guéant. A reference case for mean field games models. *Journal de mathématiques pures et appliquées*, 92(3):276–294, 2009.

[115] François Mériaux and Samson Lasaulce. Mean-field games and green power control. In *International Conference on Network Games, Control and Optimization (NetGCooP 2011)*, pages 1–5. IEEE, 2011.

[116] François Mériaux, Samson Lasaulce, and Hamidou Tembine. Stochastic differential games and energy-efficient power control. *Dynamic Games and Applications*, 3(1):3–23, 2013.

[117] Prabodini Semasinghe and Ekram Hossain. Downlink power control in self-organizing dense small cells underlaying macrocells: A mean field game. *IEEE Transactions on Mobile Computing*, 15(2):350–363, 2015.

[118] Chungang Yang, Jiandong Li, Prabodini Semasinghe, Ekram Hossain, Samir M Perlaza, and Zhu Han. Distributed interference and energy-aware power control for ultra-dense D2D Networks: A mean field game. *IEEE Transactions on Wireless Communications*, 16(2):1205–1217, 2016.

[119] Yue Zhang, Chungang Yang, Jiandong Li, and Zhu Han. Distributed interference-aware traffic offloading and power control in ultra-dense Networks: Mean field game with dominating player. *IEEE Transactions on Vehicular Technology*, 68(9):8814–8826, 2019.

[120] Xiaohu Ge, Haoming Jia, Yi Zhong, Yong Xiao, Yonghui Li, and Branka Vucetic. Energy efficient optimization of wireless-powered 5G full duplex cellular Networks: A mean field game approach. *IEEE Transactions on Green Communications and Networking*, 3(2):455–467, 2019.

[121] Charles Bertucci, Spyridon Vassilaras, Jean-Michel Lasry, Georgios S Paschos, Mérouane Debbah, and Pierre-Louis Lions. Transmit strategies for massive machine-type communications based on mean field games. In *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–5. IEEE, 2018.

[122] Lixin Li, Qianqian Cheng, Xiao Tang, Tong Bai, Wei Chen, Zhiguo Ding, and Zhu Han. Resource allocation for NOMA-MEC systems in ultra-dense Networks: A learning aided mean-field game approach. *IEEE Transactions on Wireless Communications*, 2020.

[123] Qianqian Cheng, Lixin Li, Yan Sun, Dawei Wang, Wei Liang, Xu Li, and Zhu Han. Efficient resource allocation for NOMA-MEC system in ultra-dense Network: A mean field game approach. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2020.

[124] Zihe Zhang, Lixin Li, Xiaomin Liu, Wei Liang, and Zhu Han. Matching-based resource allocation and distributed power control using mean field game in the NOMA-based UAV Networks. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 420–426. IEEE, 2018.

[125] Yahong Rosa Zheng and Chengshan Xiao. Simulation models with correct statistical properties for rayleigh fading channels. *IEEE Transactions on communications*, 51(6):920–928, 2003.

[126] Halil Mete Soner. Controlled markov processes, viscosity solutions and applications to mathematical finance. In *Viscosity solutions and applications*, pages 134–185. Springer, 1997.

[127] Lawrence C Evans. Partial differential equations and monge-kantorovich mass transfer. *Current developments in mathematics*, 1997(1):65–126, 1997.

[128] Hamidou Tembine, Quanyan Zhu, and Tamer Basar. Risk-sensitive mean-field stochastic differential games. *IFAC Proceedings Volumes*, 44(1):3222–3227, 2011.

[129] Martin Burger and Jan Michael Schulte. Adjoint methods for hamilton-jacobibellman equations. In *Westfälische Wilhelms-Universität Münster*. 2010.

[130] Rakib Ahmed. Numerical schemes applied to the burgers and buckley-leverett equations. *University of Reading*, 2004.

[131] Sima Sobhi-Givi, Mahrokh G Shayesteh, and Hashem Kalbkhani. Energy-efficient power allocation and user selection for mmwave-noma transmission in m2m communications underlaying cellular heterogeneous networks. *IEEE Transactions on Vehicular Technology*, 69(9):9866–9881, 2020.

[132] Lixin Li, Huan Ren, Qianqian Cheng, Kaiyuan Xue, Wei Chen, Mérouane Debbah, and Zhu Han. Millimeter-wave networking in the sky: A machine learning and mean field game approach for joint beamforming and beam-steering. *IEEE Transactions on Wireless Communications*, 19(10):6393–6408, 2020.

[133] Dian Shi, Hao Gao, Li Wang, Miao Pan, Zhu Han, and H Vincent Poor. Mean field game guided deep reinforcement learning for task placement in cooperative

multiaccess edge computing. *IEEE Internet of Things Journal*, 7(10):9330–9340, 2020.

[134] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[135] Marie-Joséph a Youssef, Venugopal V Veeravalli, Joumana Farah, Charbel Abdel Nour, and Catherine Douillard. Resource allocation in noma-based self-organizing networks using stochastic multi-armed bandits. *IEEE Transactions on Communications*, 69(9):6003–6017, 2021.

[136] Mohamed Ali Adjif, Oussama Habachi, and Jean-Pierre Cances. Joint channel selection and power control for noma: A multi-armed bandit approach. In *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, pages 1–6. IEEE, 2019.

[137] Manal El Tanab and Walaa Hamouda. Fast-grant learning-based approach for machine-type communications with noma. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[138] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2):235–256, 2002.

[139] Zhiguo Ding, Fumiyuki Adachi, and H Vincent Poor. The application of mimo to non-orthogonal multiple access. *IEEE Transactions on Wireless Communications*, 15(1):537–552, 2015.

[140] Shipon Ali, Ekram Hossain, and Dong In Kim. Non-orthogonal multiple access (noma) for downlink multiuser mimo systems: User clustering, beamforming, and power allocation. *IEEE access*, 5:565–577, 2016.

[141] Zhijin Qin, Yuanwei Liu, Zhiguo Ding, Yue Gao, and Maged Elkashlan. Physical layer security for 5g non-orthogonal multiple access in large-scale networks. In *2016 IEEE international conference on communications (ICC)*, pages 1–6. IEEE, 2016.

[142] Yuanwei Liu, Zhijin Qin, Maged Elkashlan, Yue Gao, and Lajos Hanzo. Enhancing the physical layer security of non-orthogonal multiple access in large-scale

networks. *IEEE Transactions on Wireless Communications*, 16(3):1656–1672, 2017.

width=!,height=!,