Distributed Resource Allocation for Self-Organizing Small Cell Networks: A Game Theoretic Approach

by

Lakshika Prabodini Semasinghe

A Thesis submitted to The Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg

Abstract

Future wireless networks are expected to be highly heterogeneous and ultra dense with different types of small cells underlaid with traditional macro cells. In the presence of hundreds of different types of small cells, centralized control and manual intervention in network management will be inefficient and expensive. In this case, self-organization has been proposed as a key feature in future wireless networks. In a self-organizing network, the nodes are expected to take individual decisions on their behavior. Therefore, individual decision making in resource allocation (i.e., Distributed Resource Allocation) is of vital important. The objective of this thesis is to develop a distributed resource allocation framework for self-organizing small cell networks.

Game theory is a powerful mathematical tool which can model and analyze interactive decision making problems of the agents with conflicting interests. Therefore, it is a well-appropriate tool for modeling the distributed resource allocation problem of small cell networks. In this thesis, I consider three different scenarios of distributed resource allocation in self-organizing small cell networks i.e., i). Distributed downlink power and spectrum allocation to ensure fairness for a small cell network of base stations with bounded rationality, ii). Distributed downlink power control for an ultra dense small cell network of base stations with energy constraints, iii). Distributed joint uplink-downlink power control for a small cell network of possibly deceitful nodes with full-duplexing capabilities. Specifically, I utilize evolutionary games, mean field games, and repeated games to model and analyze the three aforementioned scenarios. I also use stochastic geometry, which is a very powerful mathematical tool that can

model the characteristics of the networks with random topologies, to design the payoff functions for the formulated evolutionary game and the mean field game.

Acknowledgements

First and foremost, I wholeheartedly thank my adviser Prof. Ekram Hossain for giving me the opportunity to work with him. I was able to push my-self beyond my expectations with his excellent supervising, dedication and encouragement. Secondly, I would like to thank my committee members Prof. Sherif Sherif and Prof. Pourang Irani for the valuable feedback given throughout.

I acknowledge University of Manitoba and government of Manitoba for providing me with financial support.

I would like to pay gratitude to the three greatest inspirations and pillars of my life, my mother Wansha Semasinghe, my father, Indrarathna Semasinghe and my husband, Vikum Liyanage. Without my parents love, care, sacrifices and support I would never have been the person who I am today. Without my loving husband's constant encouragement, moral support and caring, both this thesis and my life would never have been completed. I dedicate this thesis to my parents and my husband. I would also like to thank my lovely family including my in-laws, and my dear friends for the constant encouragement given.

I finally thank all my lab mates and the staff of department of electrical and computer engineering, University of Manitoba for proving a peaceful and productive working environment.

Table of Contents

List of Figures vi

List of Tables viii

List of Abbreviations ix

Publications xi

1 Introduction 1
- 1.1 Small Cell Networks . . . 1
- 1.2 Self-Organizing Small Cell Networks . . . 4
- 1.3 Distributed Resource Allocation for Self-organizing Small Cell Networks 7
  - 1.3.1 Overview . . . 7
  - 1.3.2 Challenges in Designing Distributed Resource Allocation Schemes for Self-Organizing Small Cells Networks . . . 8
- 1.4 Motivations and Objectives . . . 10
  - 1.4.1 Motivations . . . 10
  - 1.4.2 Objectives . . . 12
- 1.5 Contributions and Scope of the Thesis . . . 14
- 1.6 Organization of the Thesis . . . 16

2 Game Theory for Self-Organizing Small Cell Networks 17
- 2.1 Fundamentals of Game Theory . . . 17
- 2.2 Types of Games . . . 19
- 2.3 Design of Payoff Functions . . . 21
- 2.4 Game Models for Distributed Resource Allocation in Small Cell Networks 27
  - 2.4.1 Related Work . . . 27
  - 2.4.2 Stochastic Geometry for Payoff Function Design . . . 28

3 An Evolutionary Game for Distributed Resource Allocation 32
- 3.1 Introduction . . . 33
  - 3.1.1 Overview . . . 33
  - 3.1.2 Contribution . . . 34

3.2 Related Work . . . 35
3.3 System Model and Assumptions . . . 36
3.4 Evolutionary Game Formulation . . . 39
3.4.1 Game Formulation . . . 41
3.4.2 Dynamics of Adaptation of Transmission Alignment . . . 44
3.4.3 Delay in Replicator Dynamics . . . 45
3.5 Analysis of Average SINR and Achievable Rate . . . 46
3.5.1 Average Received SINR . . . 47
3.5.2 Average Achievable Rate . . . 52
3.6 Evolutionary Equilibrium and Stability Analysis . . . 53
3.6.1 Evolutionary Equilibrium and It's Existence . . . 53
3.6.2 Stability of Evolutionary Equilibrium . . . 55
3.6.3 Stability of Delayed Replicator Dynamics . . . 58
3.7 EGT-Based Resource Allocation Algorithm . . . 59
3.7.1 Algorithm . . . 59
3.7.2 Comparison with Centralized Resource Allocation . . . 60
3.8 Numerical Results and Discussion . . . 61
3.8.1 Simulation Parameters for EGT based Resource Allocation Scheme . . . 61
3.8.2 Validation of the Theoretical Results . . . 62
3.8.3 Convergence of the Proposed EGT Algorithm . . . 63
3.8.4 Evolution of the Population . . . 66
3.8.5 Comparative Performance Evaluation . . . 66
3.8.6 Impact of Information Exchange Delay on Convergence . . . 68
3.9 Chapter Summary . . . 70

4 A Mean Field Game for Energy-Aware Power Control in Ultra-Dense Small Cell Networks 72
4.1 Introduction . . . 73
4.1.1 Overview . . . 73
4.1.2 Contribution . . . 75
4.2 Related Work . . . 75
4.3 System Model and Assumptions . . . 76
4.3.1 Network and Propagation Model . . . 76
4.3.2 Cost Function of an SBS . . . 78
4.3.3 State, Action Space, and Control Policy of an SBS . . . 81
4.4 Differential Game Formulation . . . 82
4.5 Formulation of Mean Field Game . . . 86
4.5.1 Assumptions . . . 86
4.5.2 Deriving the Cost Function . . . 87
4.5.3 Mean Field Equations . . . 92
4.6 Solution of the Mean Field Game: Mean Field Equilibrium . . . 93

4.6.1 Mean Field Equilibrium (MFE) . . . 93
4.6.2 Uniqueness of the MFE . . . 98
4.7 Numerical Results and Discussion . . . 102
4.7.1 Validating the Expressions Derived by Stochastic Geometry Analysis . . . 103
4.7.2 Behavior of the Mean Field at Equilibrium . . . 103
4.7.3 Power Control Policy at the Mean Field Equilibrium . . . 106
4.7.4 Comparison With Uniform Transmit Power Policy . . . 109
4.8 Chapter Summary . . . 110

5 A Repeated Game for Cheat-Proof Distributed Power Control in Full-Duplex Small Cell Networks 114
5.1 Introduction . . . 115
5.1.1 Overview . . . 115
5.1.2 Contribution . . . 116
5.2 System Model and Assumptions . . . 118
5.3 Stage Game and Analysis of Nash Equilibrium . . . 121
5.3.1 Stage Game, $\mathcal{G}_s$ . . . 121
5.3.2 Analysis of Nash Equilibrium . . . 125
5.4 Repeated Game and Analysis of Public Perfect Equilibrium . . . 127
5.4.1 Repeated Game, $\mathcal{G}_r$ . . . 127
5.4.2 Analysis of Perfect Public Equilibrium Set . . . 129
5.5 Learning Phase: Finding a PPE Operating Point . . . 131
5.5.1 Learning Algorithm . . . 131
5.5.2 Theoretical Analysis of the Learning Phase . . . 133
5.6 Operation Phase: Implementing an equilibrium strategy . . . 136
5.6.1 Steps in the Operation Phase . . . 136
5.6.2 Step 1: Detection of the Point of Change of Interference . . . 137
5.6.3 Step 2: Punishment . . . 140
5.7 Distributed Power Control Algorithm . . . 143
5.8 Numerical Results and Discussion . . . 145
5.8.1 Phase 1: Learning Pareto Optimal Solution . . . 145
5.8.2 Phase 2: Operation Phase . . . 149
5.9 Chapter Summary . . . 152

6 Conclusion and Future Direction 154
6.1 Conclusion . . . 154
6.1.1 Implementation of Proposed Algorithms in Practice . . . 156
6.2 Future Research Directions . . . 156
References . . . 159

A Appendix A **173**
A.1 Laplace transform of the aggregate interference 173
A.2 Average SINR of a Generic User 174

B Appendix A **176**
B.1 Proof of Lemma 5.5.1 176
B.2 Proof of Theorem 5.5.1 176

List of Figures

1.1 Traditional cellular grid model. 2
1.2 Small cells underlaying a macrocell. 3

3.1 A cluster of small cells considered for resource allocation. 37
3.2 Effect of SBS intensity $\lambda_s = A\lambda_m$ on SIR (for $\lambda_m = 4$ base stations/km^2, $x_1 = 0.4$, $x_2 = 0.6$). 63
3.3 Convergence of the algorithm to the evolutionary equilibrium (for $N = 2$, $L = 1$, $\lambda_s = 60\lambda_m$). 64
3.4 Convergence of the algorithm to the evolutionary equilibrium (for $N = 5$, $L = 2$, $\lambda_s = 60\lambda_m$). 65
3.5 CDF of the number of iterations to converge (for $N = 2$, $L = 1$, $\lambda_s = 60\lambda_m$). 65
3.6 CDF of the number of iterations to converge for $\mathcal{G}^1$. 66
3.7 CDF of the number of iterations to converge for $\mathcal{G}^2$. 66
3.8 Trajectories of proportions of population (for $N = 2$, $L = 1$, $\lambda_s = 40\lambda_m$). 66
3.9 Performance of the EGT algorithm with varying number of SBSs (for game $\mathcal{G}^1$ with $N = 5$, $L = 1$). 67
3.10 Performance of the EGT algorithm with varying number of SBSs (for game $\mathcal{G}^2$ with $N = 5$, $L = 1$). 68
3.11 Performance of the EGT algorithm with varying number of SBSs (for game $\mathcal{G}^2$ with $N = 5$, $L = 1$). 68
3.12 Effect of delay on the convergence of algorithm $\mathcal{G}^1$ (for $N = 3$, $L = 1$, $\lambda_s = 40\lambda_m$). 69
3.13 Effect of delay on the convergence of algorithm $\mathcal{G}^2$ (for $N = 3$, $L = 1$, $\lambda_s = 40\lambda_m$). 69
3.14 CDF of the required number of iterations for convergence of $\mathcal{G}^1$. . . . 70
3.15 CDF of the required number of iterations for convergence of $\mathcal{G}^2$. . . . 70

4.1 Average interference experienced by a generic small cell user (for $\lambda_s = 50\lambda_m$). 103
4.2 Variation of the distance to the closest edge of an SBS with λ_m. . . . 104

4.3 Mean field at the equilibrium for $c(t, e)$ with uniform initial energy distribution. 104
4.4 Cross-section of the mean field at equilibrium for $c(t, e)$. 105
4.5 Mean field at the equilibrium for $\hat{c}(t, e)$ with uniform initial energy distribution. 106
4.6 Cross-section of the mean field at equilibrium for $\hat{c}(t, e)$. 107
4.7 Equilibrium power policy for $\hat{c}(t, e)$ with uniform initial energy distribution. 108
4.8 Equilibrium power policy for $\hat{c}(t, e)$ with uniform initial energy distribution. 109
4.9 Cross-section of the power policy for $\hat{c}(t, e)$. 110
4.10 Transmit power variation of SBSs with different initial energy for $\hat{c}(t, e)$. 112
4.11 Variation of SINR at the receiver of a generic user with SBS density. . 112
4.12 Variation of SINR at the receiver of a generic user with SBS density. . 113

5.1 Distributed power control mechanism. 145
5.2 Phase 1: Payoff for the toy model. 147
5.3 Phase 1: Aggregate interference at the macro user for the toy model. 147
5.4 Phase 1: Payoff for a system with two small cells and one macro cell. 148
5.5 Phase 1: Payoff for a system with 8 small cells. 148
5.6 Effect of change-threshold on the performance of Page-Hinkly test. . . 149
5.7 Punishment rule: Average punishment duration. 150
5.8 Punishment rule: Payoff for the toy model. 152

6.1 Summary of the resource allocation techniques proposed in this thesis. 155
6.2 Implementation of Algorithms. 156

List of Tables

2.1 Game models for self organizing small cell networks 29

3.1 Chapter 3: Symbols 40
3.2 Chapter 3: Simulation parameters 62

4.1 Chapter 4: Symbols 79
4.2 Chapter 4: Simulation parameters 102
4.3 Comparison of the power control algorithms for two-tier networks . . 111

5.1 Chapter 5: Symbols 119
5.2 Chapter 5: Simulation parameters 146
5.3 Game matrix 146

List of Abbreviations

AP	Access point
BPP	Binomial point process
CAPEX	Capital expenditure
CCE	Coarse correlated equilibrium
CDF	Cumulative distribution function
CE	Correlated equilibrium
CSI	Channel state information
DSL	Digital subscriber line
EGT	Evolutionary game theory
FPK	Fokker-Planck-Kolmogorov
HCPP	Hard core point process
HetNet	Heterogeneous network
HJB	Hamilton-Jacobi-Bellman
IP	Internet protocol
KKT	Karush-Kuhn-Tucker
MFG	Mean field game
MFE	Mean field equilibrium
NE	Nash Equilibrium
OPEX	Operational expenditure

OFDMA	Orthogonal frequency-division multiple access
PCP	Poisson cluster process
PDE	Partial differential equation
PDF	Probability density function
PMF	Probability mass function
PPE	Perfect public equilibrium
PPP	Poisson point process
RC	Recursive core
RG	Repeated game
SBS	Small cell base station
SINR	Signal to interference plus noise ratio
SON	Self-organizing network
SPE	Sub-game perfect equilibrium
QoS	Quality of service
TDMA	Time division multiple access
WCDMA	Wideband code division multiple access
w.r.t.	With respect to

Publications

- Journal Publications:

1. **Prabodini Semasinghe**, Setareh Maghsudi, and Ekram Hossain "Game Theoretic Mechanisms for Resource Management in Massive Wireless IoT Systems" submitted to *IEEE Communications Magazine*.
2. **Prabodini Semasinghe**, Ekram Hossain, and Setareh Maghsudi, "Cheat-Proof Distributed Power Control in Full-Duplex Small Cell Networks: A Repeated Game with Imperfect Public Monitoring" submitted to *IEEE Transactions on Mobile Computing*.
3. 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," submitted to *IEEE Transactions on Wireless Communications*.
4. **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*, vol. 15, no. 2, pp. 350–363, Feb. 2016.
5. **Prabodini Semasinghe**, Ekram Hossain, and Kun Zhu, "An Evolutionary Game for Distributed Resource Allocation in Self-Organizing Small

Cells," *IEEE Transactions on Mobile Computing*, vol. 14, no. 2, pp. 274–287, Feb. 2015.

- Conference Publications:

  1. **Prabodini Semasinghe** and Ekram Hossain, "A cheat-proof power control policy for self-organizing full-duplex small cells," to be presented in Proc. of *IEEE Int. Conf. on Communications (ICC 2016)*, Kuala Lumpur, Malaysia, 23-27 May 2016.

  2. **Prabodini Semasinghe**, Kun Zhu, and Ekram Hossain, "Distributed resource allocation for self-organizing small cell networks: An evolutionary game approach," in Proc. of Workshop on Heterogeneous and Small Cell Networks (HetSNets) in conjunction with *IEEE Global Communications Conference (Globecom 2013)*, Atlanta, USA, 9-13 December 2013.

- Book Chapters:

  1. **Prabodini Semasinghe**, Kun Zhu, Ekram Hossain, and Alagan Anpalagan, "Game Theory and Learning Techniques for Self-Organization in Small Cell Networks," book chapter in *Design and Deployment of Small Cell Networks*. (Eds. A. Anpalagan, M. Bennis, and R. Vannithamby), Cambridge University Press, 2016.

Chapter 1

Introduction

1.1 Small Cell Networks

A simple communication system is composed of a transmitter, receiver and a communication channel. When it comes to wireless communications, electromagnetic radio spectrum acts as the communication channel. History of wireless communications runs back to pre-industrial age; however, early stages of wireless communications did not provide much system capacity due to the inefficient usage of radio spectrum [1]. Later with the development of the revolutionary idea of cellular concept by AT&T Bell Laboratories, capacity of the wireless communication systems increased substantially. Cellular concept takes the advantage of the property of electromagnetic signals that their power decay exponentially with the propagation distance. Due to this property, same frequencies can be spatially reused without having significant interference. A cellular network is composed of a set of fixed transceivers (which are called base stations) to which the several mobile devices in the vicinity are connected. The spatial coverage area of a base station is called a cell. After the development of this cell concept, by far cellular wireless has become one of the most fastest grown industries

in the history.

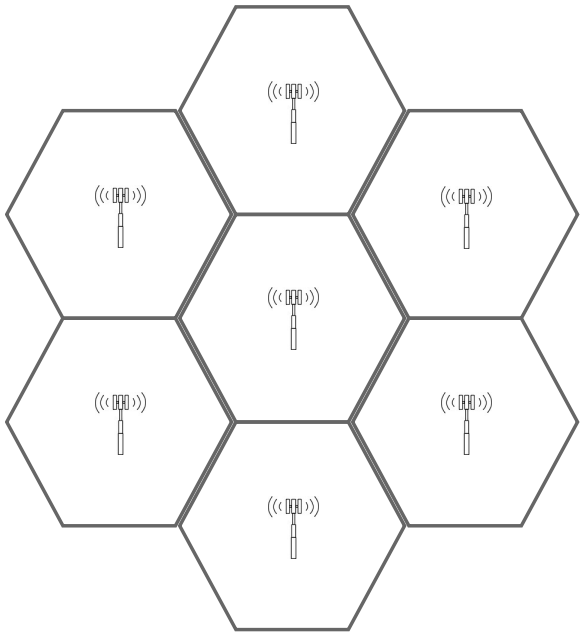


Figure 1.1: Traditional cellular grid model.

Traditional grid cellular model is as shown in Fig. 1.1. The actual network deployment may not be composed of ideal hexagon shaped cells as shown in the figure, however the idea is similar. This type of a traditional cellular network is called as a homogeneous network as the characteristics (transmit power, coverage area etc.) of all base stations are almost identical.

However, the recent evolution of smart phones and various other mobile devices have caused a significant enhancement in wireless traffic demand due to the expansion of bandwidth craving mobile applications such as video streaming, video chatting, and online gaming. This tremendous increase in wireless traffic demand has posed enormous challenges to the design of future wireless networks. Traditional homogenous cells (i.e., macro cells) will not be able to provide the increasing demand. Therefore, in-addition to the homogeneous traditional cellular networks which has been used so far, deploying different types of small cells (e.g., pico, micro, and femto) has been

proposed as an efficient and cost effective solution to support this constantly rising demand. Due to the reduced distance between the transmitter and the receiver, link quality would be higher in small cells. They would also provide more efficient spatial reuse [2].

Small cells can also deliver some other benefits such as offloading the macro network (traditional cellular network) traffic, providing service to coverage holes and the regions with poor signal reception (e.g., macro cell edges). Following this trend, future wireless networks [3, 4] are expected to be composed of hundreds of interconnected heterogeneous small cells. Fig. 1.2 shows how different types of small cells can be underlaid with a traditional macro cell. In addition to the small cells, as shown in the figure, there also can be D2D communications. Moreover, the network is no longer homogeneous; hence this type of networks are identified as heterogeneous networks (HetNets).

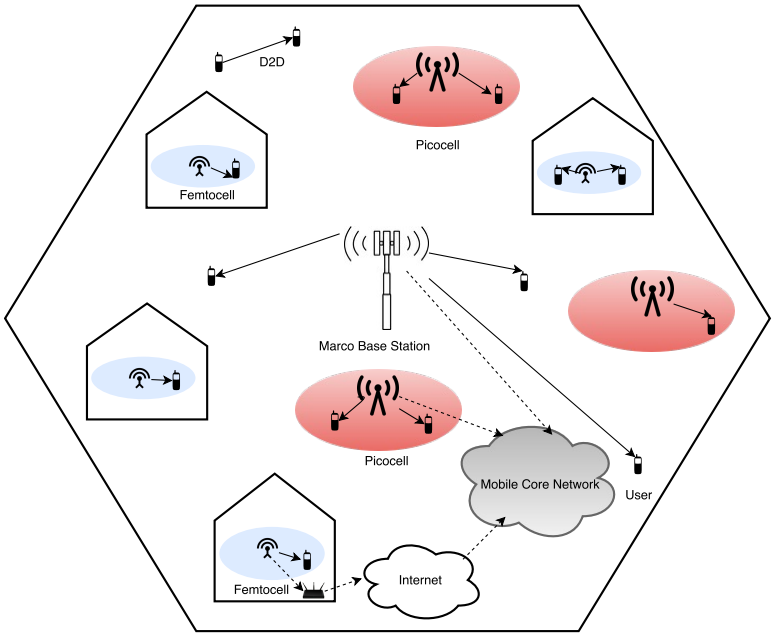


Figure 1.2: Small cells underlaying a macrocell.

Different from the cautiously planned traditional homogeneous networks, the ar-

chitecture of a HetNet is random and unpredictable due to the increased density of small cells and their impromptu way of deployment. In this case, manual intervention and centralized control used in traditional network management will be highly inefficient and expensive. Instead, **self-organization** has been proposed as an essential feature for future heterogeneous small cell networks. [5, 6]. Next section will provide a comprehensive introduction to self-organizing small cell networks.

1.2 Self-Organizing Small Cell Networks

The basis of any self-organizing system is its *autonomous and intelligent adaptivity*, i.e., the ability to respond to external environmental changes. Many literature in the context of wireless networks suggest that a self-organizing network should *learn* the environmental dynamics and adapt accordingly [7–9]. Specifically, for small cell networks, detecting the environmental dynamics can be done based on local interactions with other nodes and/or through spectrum sensing. In [10], the authors explain that the adaptive behavior of each member of a self-organizing set should also lead the whole system to form a global pattern which is called as the *emergent behavior*. Each network node [1] is expected to take individual decisions on their behavior, i.e., *distributed control*.

Based on the above notions, the basic cornerstones of a self-organizing network are identified as follows:

- Autonomous and intelligent adaptivity
- Ability to learn from the environment
- Emergent behavior

[1] A network node can be a base station or a user equipment.

- Distributed control.

Specifically, when it comes to HetNets composed of thousands of small cells, self-organization is preferred due to the following reasons.

- Numerous network devices with different characteristics are expected to be interconnected in HetNet. Nonetheless, these devices are expected to be with the '*plug and play*' capability. Therefore, the initial pre-operational configuration has to be done with minimum involvement of expertise.

- The spatio-temporal dynamics of the networks is now more unpredictable than traditional systems due to the unplanned nature of the small cell deployment. For example, the small cells can be deployed by the customer and in that case the locations of those small cells are unknown to the operator. Therefore, intelligent self adaptation by the network nodes is necessary.

- Centralized control will be highly inefficient and time consuming for a dense network due to the high computational power and the huge amount of information exchange required. Instead, small cell base stations (SBSs) should be capable of taking individual decisions on resource allocation based on local interactions.

- Self-organization of the network will also prevent possible human mistakes in configuration and network management which can drastically degrade the performance of the network and can result the extensively long recovery times.

- Self-organization of the network could also reduce a considerable amount of operational and capital expenditure (OPEX/CAPEX) due to all aforementioned reasons.

Self-Organizing Functionalities

In general, the self-organizing process of a small cell network can be divided into three phases, i.e., pre-operational phase, operational phase, and failure recovery phase. These three phases are called as *self-configuration*, *self-optimization*, and *self-healing* respectively. These are also referred to as **self-X** functionalities [5, 11, 12].

A brief overview of the operation and associated functions of each phase of self-organization is given below.

- ***Self-configuration***: Self-configuration is performed in the pre-operational process. During this phase, the small cell base stations connect to the network and execute their initialization algorithms automatically. This functionality is composed of basic set-up of the base station and the initialization of network parameters.

- ***Self-optimization***: The main task of self-optimization is to adapt automatically with the network dynamics for the optimal performance. In order to perform self-optimization, the network nodes need to measure certain network parameters (e.g., number of users, traffic patterns, and traffic load) and collect the information about the network conditions (e.g., channel gains). User admission, scheduling and resource allocation are some of the notable functions of self-optimization phase.

- ***Self-healing***: The network's ability to detect, diagnose, compensate, and recover from failures is identified as self-healing. The self-healing process is mainly composed of three functions [13], i.e., fault detection, fault diagnosis, and fault recovery.

1.3 Distributed Resource Allocation for Self-organizing Small Cell Networks

1.3.1 Overview

As explained above, self-organizing network nodes are expected to take individual decisions on their behavior including the decisions on resource allocation. Consequently, *distributed resource allocation* is a primary requirement for self-organizing small cell networks. Resource allocation belongs to self-optimization and self-healing phases of the above explained self-x functionalities. Significant and necessary features of a resource allocation scheme for a self-organizing small cell network are explained below.

- *Stability*: Stability of the algorithm is defined as its ability to converge within acceptable iterations.
- *Robustness*: Robustness is the ability of an algorithm to reach back to a stable state within a bounded duration of time in case of an unexpected change in the system or environment which makes the system deviate from a stable sate.
- *Scalability*: The complexity of self-organizing algorithms should not increase in an unbounded manner with the increase of network size. Less complex algorithms which occupy less computation resource could make the network more scalable. Also, the amount of information exchange should not increase unbounded with increase in the number of network nodes.
- *Less computation cost*: Some SBSs may not have high processing power as that of traditional macro base stations. In this case, complex algorithms which require high computation power may not be suitable for small cells.

- *Withstand imperfect information*: With certain self-organizing algorithms, the SBSs are expected to exchange information with nearby nodes (i.e., local interactions). However, this information can be distorted due to the noisy backhaul and can be delayed due to the time taken in processing and transmission. Channel State Information (CSI) can also be distorted or temporally unavailable due to the fading experienced by feedback channels. Moreover, if the status of each channel is estimated by spectrum sensing, the sensing result can be inaccurate.

- *Limited backhaul*: Unlike macro base stations which have a separate backhaul, SBSs such as femto base stations connect to the core network via a IP-based backhaul such as DSL. The same backhaul link may also be used for inter-cell coordination and periodic information exchange required by self-organizing algorithms.

1.3.2 Challenges in Designing Distributed Resource Allocation Schemes for Self-Organizing Small Cells Networks

Designing of distributed resource allocation schemes for HetNets self-organizing small cell networks which are with above explained features is challenging mainly due to the following reasons:

- Increased interference

- Spatio-temporal interference dynamics are more random than in traditional marcocellular networks as a result of the unpredictable deployment of small cells.

- Increased amount of information has to be exchanged among network nodes.

- Complexity of the resource allocation algorithms increases with the density of network nodes.

- Interconnection of different types of small cells.

Moreover, there are only a very few mathematical tools which are capable of modeling and analyzing distributed resource allocation problems; nevertheless implementing all aforementioned properties may not be possible. Moreover, different networks have different features and limitations; for example, some networks are ultra-dense, some may composed of network nodes with limited energy (battery-powered base stations) and some network nodes may have very limited backhaul capacity. Also, network operators are also with different objectives such as ensuring fairness, minimizing total cost, or prevent cheating. Some network operators may interested in resource allocation considering the instantaneous profit and on the other hand some network operators may consider maximizing the profit over a certain period of time. To this end, the mathematical tool modeling the resource allocation problem has to be selected considering the characteristics of the network and the requirements of the network operator.

Designing the properly fitting cost/payoff function (i.e., objective function) considering all these aspects is also challenging. Specifically, defining the cost functions based on the network performance metrics (e.g., achievable data rate, delay, and transmit power), modeling the network dynamics (e.g., randomness of the wireless channel, randomness of the user locations and base station deployment, and mobility of the users), meeting the requirements defined by the standards and realizing of the self-organizing network characteristics have to be considered within the scope of the mathematical tool used to model the resource allocation problem.

1.4 Motivations and Objectives

1.4.1 Motivations

Motivated by the significance of implementing self-organizing techniques for future heterogeneous networks, I aim at developing a rigorous scheme for distributed resource allocation of self-organizing small cell networks. Optimization is one possible tool that can be used to model resource allocation problems in wireless networks. However, once optimization is used, the problem is modeled in a centralized manner and resultant resource allocation schemes would be centralized; hence, it is not suitable for solving the resource allocation problems distributively at each node as needed by self-organizing networks. Therefore, I adopt game theory to model the distributed resource allocation problem as game theory provides a rich set of mathematical tools for modeling and analyzing distributed interactive decision making problems of agents with conflicting interests. Main motivations of using game theory to model the interactions of network nodes in self-organizing small cell networks are summarized below.

- The heterogeneous network nodes in small cell networks can be deployed by different operators/users and will be composed of multiple tiers (i.e., macro, femto, pico, micro). The performance of one tier could be easily affected by the behavior of other tiers. Therefore, modeling of interactive behavior is required. Game theory models provide a mathematical framework to analyze the competitive or cooperative interactions among the players in a multi-player system. As a result, game theory is a good candidate for devising resource allocation schemes for Self-Organizing Networks (SONs).

- Different network nodes could have different QoS requirements and can be self-

interested. Each node takes individual decisions to meet her own requirements rather than optimizing the system-wide performance. In this case, these nodes may have conflicting interests. Such self-interested behavior can be easily modeled by using game theory.

- The basic keystones of a self-organizing network are *ability to learn from environment, autonomous adaptivity, emergent behavior capability* and *distributed control*. In the context of game theory, the players could adapt their decisions to obtain a better payoff (i.e., autonomous adaptivity). Also, after several adaptation iterations, the game could reach the equilibrium (emergent behavior). In addition to that, players take the individual decisions based on the information they have (i.e., distributed control).

- Game theory provides a natural tool to develop distributed self-organizing algorithms as it allows local interactions and individual decision making. Local interactions will reduce the amount information exchange among the nodes and as a result the network becomes scalable and capable of operating with limited backhaul conditions.

Many types of games which are applicable in different settings have been developed so far in game theory (e.g., non-cooperative games, cooperative games, Stackelberg games, Bayesian games, differential games, evolutionary games, etc.). The notion of equilibrium can be different for different types of games. A comprehensive introduction to game theory is presented in next chapter.

Furthermore, I exploit stochastic geometry to design the payoff functions. Stochastic geometry is a very powerful tool to analyze networks with random topologies (e.g., to evaluate interference experienced at a network node). Network nodes are modelled via point processes that reflect their spatial locations.

1.4.2 Objectives

The main objective of this research is to develop a new distributed resource allocation framework for self organizing dense small cell networks underlaying macro-cellular networks. However, as different small cell networks can have different features and limitations, network operators may have different objectives for resource allocation. Therefore, in particular, I study several system models of small cell networks with different characteristics and different objectives. Throughout this work, game theory is used to model and analyze the interaction of the network nodes in each system considered. Subsequently, based on the game theoretic modeling and analysis, distributed resource allocation techniques are proposed for each system model. The research also analyze the convergence of the algorithms and effect of imperfect information/delay to ensure the stability and robustness.

Particularly, in this work, I address three scenarios with different resource allocation objectives as noted below.

1. **Downlink power-subcarrier allocation to ensure fairness among base stations with bounded-rationality** : Rational network nodes make rational choices by evaluating possible outcomes of every other node in the network. However, in a dense network there might be circumstances in which they are unable to make rational choices individually, since rationality requires the nodes to take the choices of all other nodes connected to the network into account. To this end, I first consider distributed downlink joint power-subcarrier allocation for a dense network with bounded-rational base stations who are not capable of tracking the moves of all the other base stations in the network. Moreover, for such a system, ensuring fairness is important as base stations may not individually try to maximize their performances due to bounded rationality.

2. **Downlink power control for ultra-dense networks of base stations with limited energy**: Due to irregular deployment and increased density, small cell base stations may not be connected to the grid power supply all the time. Instead, some small cell base stations may be battery operated, which necessitates frequent recharge, or they can be powered through energy harvesting, which is inherently opportunistic and random. For such a dense network of small cells with non-guaranteed and limited energy supply, rather than spending energy without considering future transmissions, it is important to consider maximizing each SBSs performance over a certain period of time considering the limited energy it owns. In this case, network nodes have to take individual rational decisions. Therefore, as the second scenario, I consider distributed energy aware downlink power control for an ultra dense small cell network where each SBS tries to maximize its performance over a given period of time.

3. **Joint uplink-downlink power control for a small cell network of possibly deceitful nodes with full-duplexing capabilities**: Full-duplex transmission is a newly emerging technology in wireless communications where users and base stations can receive and transmit simultaneously using the same frequency band. Moreover, there can be some situations where network nodes are not truthful to each other, but are deceitful and provide inaccurate information to obtain more benefit than other nodes. As the third scenario, I consider deriving a cheat-proof distributed power control technique for small cells with full-duplexing capabilities.

Chapters 3, 4, and 5 of this thesis present each of the aforementioned scenarios. Precise features of the system models and proposed resource allocation techniques will be explained in each chapter. I discuss the related work in each chapter. That is,

the related work for each system model and the corresponding game used to model the system is presented and the novelty of the research presented in this thesis is emphasized.

1.5 Contributions and Scope of the Thesis

In this thesis, I exploit different types of games to model small cell networks with different characteristics and different requirements. More specifically, three game models are adapted and extended to model three different distributed resource allocation scenarios for self-organizing small cell networks, which have not been addressed in the existing literature. In the following, a short discussion on the contributions of this thesis is presented considering the each type game model used.

1. Evolutionary game-based distributed downlink subcarrier and power allocation to ensure fairness for network nodes with bounded rationality

    - Distributed downlink power and subcarrier allocation problem in self-organizing small cell networks is formulated as an evolutionary game.
    - A stochastic geometry-based approach is used to analyze the utilities of the players in terms of average achievable SINR and average achievable rate.
    - A distributed algorithm of linear time complexity which ensures fairness among the base stations is proposed to reach the evolutionary equilibrium.
    - Stability of the equilibrium point and the effect of the delay in information exchange is analyzed.

2. Mean field game based energy aware power control for ultra-dense networks

- Downlink power control problem of a dense small cell network is formulated as a differential game and extended to a mean field game for an ultra-dense scenario.
- Using stochastic geometry-based analysis, two cost functions for the mean field game are derived in such a way that the mean field game setting becomes valid.
- Existence of the equilibrium is proven for the differential game and the sufficient conditions for the uniqueness of the mean field game are derived.
- An offline algorithm is proposed to obtain the mean field equilibrium, i.e., energy aware power control algorithm.

3. Repeated game based cheat-proof power control for full-duplex small cells

- Joint downlink-uplink power control problem for a full-duplex small cell networks is formulated as a non-cooperative repeated game with imperfect public monitoring.
- Theoretical characterization of perfect public equilibrium of the formulated repeated game is provided.
- A distributed learning algorithm with linear time complexity is provided to achieve perfect public equilibrium distributively.
- A deviation detection and punishment policy is implemented in order to prevent selfish network nodes deviating from the desired operating point for their own benefit.

A more detailed discussion on the contribution will be provided in latter chapters of the thesis.

1.6 Organization of the Thesis

In Chapter 2, I present a high level overview on the basics of the mathematical tools used in this thesis, i.e., i). an introduction to game theory and how different game models have been used to solve various problems in wireless communications, ii). a brief introduction to stochastic geometry. Then, Chapter 3, Chapter 4, and Chapter 5 present the core contribution of this thesis. Specifically, Chapter 3 explains the evolutionary game based distributed resource allocation scheme. Chapter 4 and 5, respectively, discuss the mean field game based power control for ultra-dense networks and repeated game based cheat-proof power control for full-duplex small cell networks. Chapter 6 summarizes and concludes the research presented in this thesis and explore some directions for future research. Symbols and notations used throughout the chapters are given in a table at the beginning of each chapter.

Chapter 2

Game Theory for Self-Organizing Small Cell Networks

2.1 Fundamentals of Game Theory

Game theory provides a strong set of mathematical tools for modeling and analyzing interactive decision making problems in which the interests of agents (i.e., players) may conflict with each other. It is a well developed area in applied mathematics and has been used primarily in economics to model competitions in markets. In recent years, game theory has also been widely adopted to solve many problems in the area of wireless communications [14,15]. A number of works have explored the applications of game theory for the analysis and optimization of various issues in wireless systems, in most cases to solve resource allocation problems in a competitive environment.

A game is a process in which the agents select certain strategies from their own strategy sets and obtain payoffs according to the strategies of all agents. The choice of a strategy can be made both simultaneously and non-simultaneously. In addition, an agent may make decisions multiple times according to the game rule. A game consists

of a set of players, a set of strategies available to those players, and a specification of payoffs for each combination of strategies.

1. *Set of players* $\mathcal{N}$: The set of decision makers involved in the game. The players are assumed to be rational or bounded rational depending on the type of the game.

2. *Set of strategies* $(\mathcal{S}_i)_{i\in\mathcal{N}}$: Strategies are the options that a player can select depending on the state of the game. Here $\mathcal{S}_i$ denotes the set of strategies of player $i \in \mathcal{N}$. A player's strategy could contain a single action, multiple actions, or probability distribution over multiple actions. As common in game theory, $\mathcal{S}_{-i}$ denotes the strategies of all players other than i. The state of a game depends on the strategies taken by all the players (i.e., $[\boldsymbol{s}_i, \boldsymbol{s}_{-i}]$). Note that different players could have different strategy sets.

3. *Payoff* π_i: The payoff represents the preference of each player under the current strategy profile. The payoff could be modelled as a cost function $c_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i})$, a utility function $u_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i})$, or a combination of both (e.g., in the form of equation (2.1)), where the cost function represents the cost of performing certain strategies (e.g., transmit power) which needs to be minimized, the utility function represents the gain (e.g., profit of service providers) which needs to be maximized.

$$\pi_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i}) = u(\boldsymbol{s}_i, \boldsymbol{s}_{-i}) - c(\boldsymbol{s}_i, \boldsymbol{s}_{-i}). \tag{2.1}$$

   It is straightforward to see that a player's payoff depends not only on her own strategy but also on the strategies of all other players.

2.2 Types of Games

As explained in Section 1.4, game theory is a well-fitting mathematical tool to model distributed resource allocation problems in wireless networks. Different game models (e.g., non-cooperative/cooperative, static/dynamic) can be used to address distributed resource allocation problems in small cell networks the choice of which depends on the characteristics of the network, applications, and also the objectives. Different game theory models may differ considerably in structure from many aspects, e.g., number of players, number of strategies, and payoffs. The number of players may vary in different games; i.e., finite or infinite. If a game has only one player, the game becomes an optimization problem. In different games, the number of strategies for players can be either finite (e.g., in a rock-scissor-paper game) or infinite (e.g., in a pricing game). The analysis of a finite strategy game and an infinite strategy game are different. The summation of payoffs of all players may also differ in different models. In general, this summation can be zero, a non-zero constant number, or any arbitrary value. The game process is an important aspect in the game structure. The players in a game may take actions simultaneously, in a certain order, or in a repeated fashion, according to which the game can be referred to as a *static game*, a *dynamic game*, and a *repeated game*, respectively.

Moreover, the assumptions of players' rationality are different. Most of the game theory models assume perfect rationality of players, while some models consider that the players are with limited rationality (i.e., bounded rationality). According to the above analysis, game models can be divided into the following categories.

Non-cooperative vs. cooperative games

Non-cooperative games are the most popular type of games. In non-cooperative games, the players are commonly considered to be rational and self-interested who have fully or partially conflicting interests. Each player selects the strategy to optimize her own payoff function. For non-cooperative games, the most commonly used solution concept is Nash Equilibrium the definition of which is given as follows:

Definition *Nash Equilibrium*: Let $\boldsymbol{s}_i \in \mathcal{S}_i$ and $\boldsymbol{s}_{-i} \in \mathcal{S}_{-i}$. Then the NE strategy profile $\left(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}^*\right)$ is defined as

$$\pi_i \left(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}^*\right) \geq \pi_i \left(\boldsymbol{s}_i, \boldsymbol{s}_{-i}^*\right) \tag{2.2}$$

for all $\boldsymbol{s}_i \in \mathcal{S}_i$ and for all $i \in \mathcal{N}$.

When the game reaches a Nash equilibrium, none of the players can improve her payoff by changing strategy unilaterally. There are also other solution concepts such as correlated equilibrium which can be considered as a generalized version of NE [16], evolutionary equilibrium and dominant-strategy equilibrium.

There can be situations that players may make agreements to cooperate. Cooperative game provides analytical tools to model and analyze the cooperative behavior of rational players who may form coalitions. In this case, the members of each coalition cooperate to maximize the coalition payoff and the competition is among coalitions instead of among individual players.

Static vs. dynamic games

A static game is one in which a single decision (time irrelevant but may contain multiple actions) is made by each player, and each player has no knowledge of the

decisions made by other players before making her own decision. Decisions are made simultaneously (or their order is irrelevant). A game is dynamic if the order in which the decisions are made is important or the strategy itself is time-dependent. For dynamic games, the dynamics can be abstracted from different aspects which lead to different types of dynamic games listed as follows:

(i) *Dynamic nature in games' play order*: The dynamic nature in games' play (decision) order leads to the development of multi-stage game (e.g., Stackelberg game). In this case, the decisions are made asynchronously and the games' play order is important. The players who move later can observe the decisions of the players who move first and then make the decisions accordingly. Note that if multiple players exist in one stage, the competition within this stage is usually formulated as a stage game.

(ii) *Dynamic nature in time dependency*: The dynamic nature in the time dependency leads to the development of differential game and evolutionary game. For differential game, the strategy of a player is time-dependent (i.e., function of time t). That is, the player seeks a best response strategy considering the entire time horizon. For evolutionary game, the players adapt their strategies according to the time-varying system state.

2.3 Design of Payoff Functions

Game theory was initially proposed and developed for economics and social sciences. Therefore, properly fitting game models in the context of wireless communication is challenging.

The payoff function should quantify the perceived preference or the satisfaction level of a player. In the context of wireless networks, the user satisfaction level may

depend on one or multiple performance metrics given as follows:

- Individual performance (e.g., rate, SINR, and delay)
- Global network performance
- Interference level caused to other network nodes
- Power/energy consumption
- User fairness

As self-organizing small cell technologies are still in its infancy, there is no well-defined framework for designing the payoff functions. However, after a comprehensive study on how game theory has been used in the existing literature to solve resource allocation problems in traditional cellular networks and small cell networks, I introduce some general approaches on how payoff functions can be designed for various applications and objectives in the context of self-organizing small cell networks.

A payoff function $\pi(x)$ is expected to satisfy the following criteria.

1. The non-stationary property: $\frac{d\pi(x)}{dx} > 0$, which states that the payoff increases with the preference or satisfaction.
2. The risk aversion property: $\frac{d^2\pi(x)}{dx^2} < 0$, which states that the payoff function is concave. In other words, the marginal payoff of satisfaction decreases with increasing level of satisfaction.

Depending on the objective, behavior, and rationality of the network nodes, different payoff functions are defined in the wireless communications literature. The payoff/utility functions which can be applied in the context of small cell networks are discussed below.

1. Payoff functions for power consumption

   Power/energy conservation is crucial in small cell networks as they might be operated in energy-limited environment (e.g., power supplied by a battery). [17] defines a simple energy aware payoff function as follows:

   $$\pi_i(e) = \frac{E_{tot}}{e_i}, \tag{2.3}$$

   where E_{tot} is the total energy available for each player and e_i is the energy required by player i for transmission. Players would try to achieve a higher payoff by reducing the transmission power.

2. Payoff functions for individual performance Instead of direct power minimization as that in equation (2.3), it is more appropriate for self-organizing algorithms to perform power control in such a way that the desired performance can be satisfied. The following logarithmic payoff function with individually perceived SINR as the input parameter can capture the self-interest of network nodes and is used for power control in [18, 19]:

   $$\pi_i(s_i, \boldsymbol{s}_{-i}) = \log(\gamma_i(s_i, \boldsymbol{s}_{-i})), \tag{2.4}$$

   where γ_i is the SINR of the i^{th} player. Such a logarithmic payoff function and its extensions are most popular payoff functions used in the context of resource allocation due to its simplicity and mathematical tractability [20]. For example, such form of payoff can be used for subcarrier allocation (in OFDMA networks) and joint power-subcarrier allocation as well.

   Another widely used payoff function is the Shannon capacity or the maximum

achievable rate which can be considered as an extended version of logarithmic function of SINR as shown below:

$$\pi_i(s_i, \boldsymbol{s}_{-i}) = \ln(1 + \gamma_i(s_i, \boldsymbol{s}_{-i})). \tag{2.5}$$

3. Fairness utility function

   One of the desired objectives of resource allocation is to provide fairness among users instead of obtaining the optimum performance. The most widely used payoff function which guarantees fairness is given below:

$$u(x) = \begin{cases} \frac{x^a}{a}, & \text{if } a < 0, \\ \log x, & \text{if } a = 0, \end{cases} \tag{2.6}$$

   where $a \leq 0$. By twice differentiation of (2.6) with respect to x the following can be obtained:

$$\frac{du(x)}{dx} = \begin{cases} x^{a-1}, & \text{if } a \neq 0, \\ \frac{1}{x}, & \text{if } a = 0, \end{cases} \tag{2.7}$$

   and

$$\frac{d^2u(x)}{dx^2} = \begin{cases} (a-1)x^{a-2}, & \text{if } a \neq 0, \\ \frac{-1}{x^2}, & \text{if } a = 0. \end{cases} \tag{2.8}$$

   It can be observed that the function given in equation (2.6) has both non-stationary and risk aversion properties for all $x > 0$ since $\frac{du(x)}{dx} > 0$ and $\frac{d^2u(x)}{dx^2} < 0$.

4. System payoff functions

   In self-organizing enabled small cell networks, a group of densely deployed small

cells could form a cluster and cooperate with each other to enhance the performance of the cluster [21]. In addition to that, cooperative games can also be formulated to design self-organizing algorithms for small cells. Accordingly, cooperative payoff functions, which reflect the overall network/cluster performance, are required.

The simplest and most intuitive cooperative payoff function would be the sum capacity/rate of the cluster/network as shown below:

$$\pi_i(\boldsymbol{s}) = \sum_{j \in \mathcal{N}_i} C_j(\boldsymbol{s}), \tag{2.9}$$

where $\mathcal{N}$ is the set of players in the i^{th} cluster who cooperates with each other and C_j is the capacity of the j^{th} player.

5. Multi-dimensional payoff function

   The payoff function can be designed considering multiple performance metrics. In such cases, these multiple metrics could appear in the payoff function (most case in a product form). One typical example is given as follows:

$$\pi_i = \pi_i^{\text{rate}} \pi_i^{\text{delay}}. \tag{2.10}$$

6. Payoff function with cost

   For a strategy adopted by a player, there could be a cost associated with it (e.g., cost of using bandwidth, power consumption) or it may affect the performance of other players (e.g., cause interference). This issue can be modelled by introducing certain cost functions into the payoff function. In particular, the payoff function (some may refer to this as *net utility*) can be defined to reflect

both the satisfaction of the player (modelled by utility function) and the cost (e.g., price per unit resource) as follows:

$$\pi_i(s_i, \boldsymbol{s}_{-i}) = u_i(s_i, \boldsymbol{s}_{-i}) - mx, \tag{2.11}$$

where $u_i(s_i, \boldsymbol{s}_{-i})$ is the utility based on the user satisfaction and m is the price paid for each resource x.

[22] uses a net utility function with logarithmic payoff as given below:

$$\pi_i(s_i, \boldsymbol{s}_{-i}) = a_i \log\left(1 + \gamma_i(s_i, \boldsymbol{s}_{-i})\right) - b_i m \gamma_i(s_i, \boldsymbol{s}_{-i}), \tag{2.12}$$

where γ_i is the SINR of the i^{th} user, a_i and b_i are weighting parameters and m is the cost for the received SINR. The gain of maximizing γ_i could be neutralized by the cost associated with the received SINR.

Guaranteeing the existence of equilibrium is one of the essential features of any game formulation. It is straightforward that the existence of equilibrium, convergence, and stability of the equilibrium is highly related to the payoff function and the structure of the game. Therefore, special payoff function can also be designed to fit the game model into special structures (e.g., super-modular, potential). Polynomial time computability is another important feature of a payoff function. Besides, when it comes to self-organizing small cell networks, the ability to compute with local information or with reduced information exchange is also highly desirable.

2.4 Game Models for Distributed Resource Allocation in Small Cell Networks

Eventhough game theory has been used widely to solve various problems in wireless networks, distributed resource allocation for self-organizing small cells is relatively a new and less addressed problem by the research community. In this section, I provide few examples on how various game models have been used to develop distributed resource allocation techniques in existing literature. A detailed discussion on related work for each scenario considered in this thesis and corresponding game is provided in each chapter.

2.4.1 Related Work

Game theory has been used to address the resource allocation and interference management problem in wireless networks [23–28] extensively. A two-tier resource allocation problem is formulated as a Stackelberg game in [25], where the macro base station (MBS) acts as the leader and the femto base stations are the followers. Then a distributed algorithm is proposed to achieve the equilibrium. In [29], game theory is used for resource allocation in an OFDMA-based small cell network. The authors investigate the performance of two games. The first is a non-cooperative game where the macro and the femto users try to improve their individual payoffs in a selfish manner. The second is a hierarchical game where the players try to improve the overall network efficiency. The hierarchical game model relies on the perfect knowledge of the channel state information at the leader side which is not realistic in a practical situation. Therefore, the authors propose a reinforcement learning process. In [30], the authors propose a reinforcement learning algorithm which converges to an

ϵ Nash equilibrium. The equilibrium is achieved through the smoothed best response (SBR) dynamics. The convergence of the SBR algorithm to an ϵ Nash equilibrium is guaranteed for a payoff function which depends on the sum rate of the entire network. [31] presents a non-cooperative potential game formulation for decentralized interference control in the downlink of an OFDMA network. Subcarrier allocation in this scheme is based on a game theoretic approach and power allocation is done by an optimization approach. This scheme also relies on the availability of all channel gains. Achieving Nash equilibria in a distributed manner based on fictitious play under certain conditions has been studied in [32] and [33].

A summary on how different game models have been used to solve several other resource allocation problems in small cells distributively is provided in Table 2.1.

2.4.2 Stochastic Geometry for Payoff Function Design

As explained in the previous chapter, network geometry of future wireless networks will be random due to the unplanned nature of the deployment of small cells. Network geometry has a substantial impact on the interference and hence on SINR experienced by each network node. Moreover, most of the performance metrics of the network such as maximum achievable rate, outage probability, energy efficiency and spectral efficiency are direct functions of SINR at network nodes. When applying game theory to model resource allocation problems in small cell networks, as explained in Section 2.3 payoff has to be defined as a function of one or few of these performance metrics. Thus, modeling and approximating the interference experienced at each network node is significant in designing payoff functions for self-organizing small cells. In this regard, I use stochastic geometry as a mathematical tool to model the network topologies and approximate the performance metrics such as interference, SINR and rate in order to

Table 2.1: Game models for self organizing small cell networks

Objective	**Game type**	**Player set**	**Strategy**	**Solution**
Power allocation [34]	Non-cooperative, Supermodular	SBSs	Downlink transmit power	NE
Power allocation [35]	Non-cooperative	SUs	Uplink transmit power	NE
Power allocation [36]	Non-cooperative	MUs & SUs	Uplink transmit power	NE
Resource block allocation [37]	Non-cooperative, Potential	SUs	OFDMA RBs for uplink	NE
Joint power-subchannel allocation [38]	Non-cooperative, Potential	SBSs	A composite of downlink transmit power and subcarrier	NE
Power allocation [39]	Non-cooperative	SBSs	Uplink transmit power	NE
Forming a group for interference draining [40]	Cooperative coalition formation	SBSs	Coalitions	RC
Joint power-subcarrier allocation [41]	Non-cooperative	SBSs	A composite of downlink transmit power and subcarrier (mixed strategy)	LE
Forming a cluster for joint beamforming [42]	Cooperative coalition formation	SBSs	Coalitions	RC
Joint power-spectrum allocation [16]	Non-cooperative	SBSs	Downlink power level and frequency band	ϵ-CCE
Coverage optimization [21]	Non-cooperative	SBSs	Downlink transmit power	NE
Resource block (RB) allocation in cognitive SBSs [43]	Non-cooperative	SBSs	Downlink RBs	CE

design payoff functions.

Stochastic geometry is a branch of mathematics which studies of random spatial point patterns. Recently, stochastic geometry has been used as a mathematical tool to model and analyze the networks with random topologies such as ad hoc networks and multi-tier HetNets [44]. In stochastic geometry analysis, based on the network topology, different point processes are used to model the positions of the network nodes.

Point Processes

A stochastic point process is a type of random process of which each realization is composed of a set of isolated points either in time or geographical space, or even in more general spaces. In stochastic geometry modeling, point processes are used to abstract the spatial deployment of network nodes. Different point process such as Poisson point process (PPP), Binomial point process (BPP), Hard core point process (HCPP) and Poisson cluster process (PCP) are used in literature to model different wireless networks. Among all this, in this thesis I use PPP to model the spatial distribution of network nodes. The formal definition of a Poisson point process is as follows.

Definition Poisson Point Process
let Φ be a point process in $\mathbb{R}^N$ with intensity λ. Φ is a Poisson point process if and only if following conditions are satisfied.

1. The number of point which belongs to Φ in any compact set $C \subset \mathbb{R}^N$ is a Poisson random variable.

2. Numbers of points in any two disjoint sets in $\mathbb{R}^N$ are independent.

Probability of number of points of Φ in any compact region of area A (i.e., $N(A)$) being equal to k is given by the following equation.

$$\Pr\left(N(A)=k\right)=\frac{\left(\lambda A\right)^{k}e^{-\lambda_{d}A}}{k!}. \tag{2.13}$$

It is reasonable enough to abstract the spatial distribution of small cell base stations and users by PPPs due to their independent and random way of deployment. Further, it has been shown that the results obtained using PPP assumption are within 1-2 dB accurate with the actual measurements of heterogeneous LTE networks [45]. PPP is also more analytically tractable than other point processes.

The mathematical tools introduced above will be used throughout this thesis. Next three chapters present the core contribution of the thesis.

Chapter 3

An Evolutionary Game for Distributed Resource Allocation

In this chapter, I propose an evolutionary game theory (EGT)-based distributed resource allocation scheme for small cells underlaying a macro cellular network. EGT is a suitable tool to address the problem of resource allocation in self-organizing small cells since it allows the players with bounded-rationality to learn from the environment and take individual decisions for attaining the equilibrium with minimum information exchange. EGT-based resource allocation can also provide fairness among users. I have shown how EGT can be used for distributed subcarrier and power allocation in orthogonal frequency-division multiple access (OFDMA)-based small cell networks while limiting interference to the macrocell users. Two game models are considered, where the utility of each small cell depends on average achievable signal-to-interference-plus-noise ratio (SINR) and data rate, respectively. For the proposed distributed resource allocation method, the average SINR and data rate are obtained based on a stochastic geometry analysis. Replicator dynamics is used to model the strategy adaptation process of the small cell base stations and an evolutionary equi-

librium is obtained as the solution. Based on the results obtained using stochastic geometry, the stability of the equilibrium is analyzed. I also extend the formulation by considering information exchange delay and investigate its impact on the convergence of the algorithm. Numerical results are presented to validate the theoretical findings and to show the effectiveness of the proposed scheme in comparison to a centralized resource allocation scheme.

3.1 Introduction

3.1.1 Overview

In this chapter, I utilize evolutionary game theory (EGT) [46] to develop a distributed subcarrier and power allocation scheme for downlink transmission in orthogonal frequency-division multiple access (OFDMA)-based small cell networks underlaying a macro cellular network. The SBSs are considered to be self-interested each of which aims to improve its own performance (e.g., in terms of signal-to-interference-plus-noise ratio [SINR] or average achievable rate) while limiting interference to the macrocell network. Each SBS takes individual decisions on the selection of a subcarrier and a power level for its downlink transmission. Since simplicity is an important requirement for distributed resource allocation in self-organizing networks, the SBSs are considered to be with only bounded-rationality[1]. Different from the traditional game models, in the EGT model, each player (i.e., an SBS) selects a strategy by replication and can adapt its selection for a better payoff (i.e., evolution). Accordingly, EGT focuses on the dynamics of the strategy adaptation in the population. A population is the set of players involved in the game. The behavior of the population

[1]This is different from the traditional rationality assumption used in noncooperative game models to obtain the Nash equilibrium. The rationality implies complete information and strong computation capability of each player to calculate the best response to other players' strategies.

can be described by the numbers of its members choosing each pure strategy.

The key features of this EGT-based solution are its simplicity and reduced amount of information exchange among the network nodes. Also, the EGT-based resource allocation can provide fairness among all SBSs. The proposed algorithm does not rely on the perfect knowledge of real-time channel state information, which makes it suitable for densely deployed small cell networks.

3.1.2 Contribution

The main contributions of this chapter can be summarized as follows:

1. I formulate the distributed power and subcarrier allocation problem in self-organizing small cell networks as an evolutionary game with the SBSs as the players. The strategy adaptation process of the SBSs is modelled by replicator dynamics and the evolutionary equilibrium is obtained as the solution. Specifically, I formulate two games based on two different utility functions for the players.

2. A stochastic geometry-based approach is used to analyze the utilities of the players in terms of average achievable SINR and average achievable rate. The accuracy of the analytical results is validated by simulations.

3. A distributed algorithm is proposed to reach the evolutionary equilibrium and its performance is compared with that of an optimization-based centralized resource allocation scheme.

4. Based on the results obtained from stochastic geometry-based analysis, the stability of the equilibrium is analyzed for the case when the utility of the users is given by their achieved average SINR.

5. I extend the formulation by considering information exchange delay and study its impact on the convergence of the algorithm.

3.2 Related Work

Recently, EGT (Evolutionary Game Theory) has been adopted to solve wireless communications and networking problems. Application of evolutionary coalitional game theory to solve various problems in wireless networking can be found in [47]. The paper also explains the open issues and trends in the field. [48] proposes a reinforcement learning-based distributed mechanism for strategy and payoff learning in wireless networks. The stability of the learning algorithm is discussed based on evolutionary game dynamics. An EGT-based method is used in [49] to solve the problem of network selection in an environment where multiple networks are available. In [50], the authors present an EGT-based framework to analyze the interaction of mobile users in a WCDMA system. In [51], the service selection in small cell networks is modeled and analyzed by using evolutionary game theory.

EGT-based resource allocation for small cell networks is proposed in [52]. The problem of subcarrier and power allocation for small cell networks underlaying a macro network is formulated as an evolutionary game. The authors propose two techniques to attain the evolutionary equilibrium. One is based on replicator dynamics and the other is based on reinforcement learning. They also study the 'replication by imitation' approach to reach the equilibrium. In [53], the authors propose an inter-cell interference coordination technique inspired by evolutionary games. Interference mitigation is done by power allocation and the proper selection of subcarriers for each base station. They also compare the performance of EGT-based resource allocation with different reinforcement learning algorithms. In this chapter, I develop an EGT-

based distributed resource allocation scheme for self-organizing small cell networks with a more rigorous and detailed analysis. Different from the other related work in the literature, I use a stochastic geometry-based approach to analyze the stability of the equilibrium of the evolutionary game. I also investigate the effect of information exchange delay on the convergence performance of the distributed resource allocation algorithm.

3.3 System Model and Assumptions

I consider the downlink transmission of an OFDMA-based two-tier cellular network composed of macrocells and a set of underlaying self-organizing small cells. The macro network is considered to be infinite and modeled by an infinite homogeneous Poisson point process (PPP) Φ_m in $\mathbb{R}^2$. The density of Φ_m is given by λ_m. For resource allocation, I focus on a generic small cell cluster (see Fig. 3.1) which is located inside a macrocell. I assume that the cluster is highly dense and the propagation environment has a high path-loss exponent. Due to the high path-loss exponent and the low transmit power of the small cell base stations, the interference is significant only from transmitters located in a small area around the receiver. The interference coming from the further small cell base stations is negligible. Therefore, for a generic user inside the small cell cluster, the total aggregate interference due to the SBSs (inside and outside of the cluster) would be approximately same as the interference only due to the base stations inside of the cluster. In other words, due to the low transmit power and high path-loss, the small cell cluster has the similar effect as a large network to a generic user inside the cluster. The spatial distribution of a large number of points which are randomly distributed over a large area can be well approximated by an infinite homogeneous Poisson point process (PPP) [54, 55]. Following this, to analyze

the effect of interference, the small cell cluster is also approximated by an infinite homogeneous PPP denoted by Φ_s. The density of Φ_s is given by λ_s.

Moreover, I assume that the average channel gain from a generic SBS inside the cluster to the macro user(s) within the same macrocell receiving on the same subcarrier is known. For simplicity, it is also assumed that each SBS serves only one user at a time which is at a distance r_s (equal to the small cell radius) from its serving SBS.

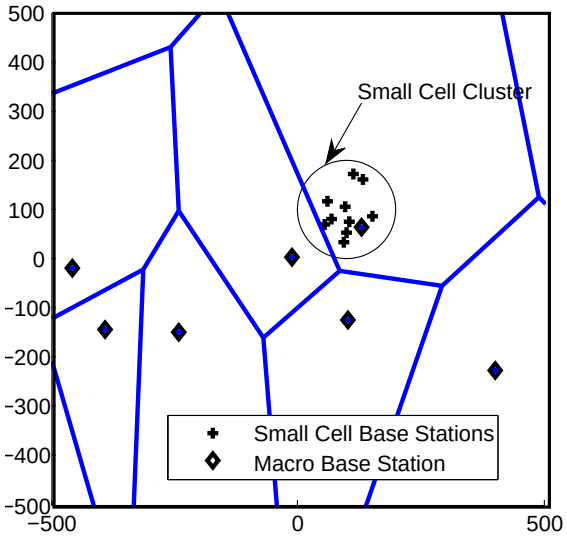


Figure 3.1: A cluster of small cells considered for resource allocation.

The macrocell and the small cells share the same set $\mathcal{N} = \{1, 2, ..., N\}$ of orthogonal subcarriers for downlink transmission. In the context of self-organizing small cell networks, each SBS selects one subcarrier from the available subcarriers. [2] They are also capable of selecting a transmit power level from a finite set of values. Let $\mathfrak{L} = \{1, 2, ..., L\}$ denote the set of power levels [30].

Each SBS should select a suitable subcarrier-power combination which is referred to as the "transmission alignment" of that SBS. For each subcarrier n, there is a threshold for the maximum aggregate interference that can be caused by the entire

[2]Considering the case where the SBSs select multiple subcarriers will require us to consider a larger strategy set in the game formulation. This issue will be further discussed in Section 3.4.

small cell cluster to the macro users. The maximum allowable aggregate interference depends on the user's QoS constraint (e.g., SINR constraint) and the channel gain between the base station and the user. This threshold is a pre-defined value denoted by $T^{(n)}$. An SBS may receive a penalty or reward depending on whether the interference constraint for macro users is violated or not. The penalty or reward is modeled in the payoff function which will be explained in Section 3.4.

For SBS k, the transmit power vector is defined by $\mathbf{P}_k = (p_k^{(1)}, p_k^{(2)}, ..., p_k^{(N)})$, where each element $p_k^{(n)}$ is the transmit power level of the k^{th} SBS over subcarrier n. If subcarrier n is not used by SBS k, then $p_k^{(n)} = 0$. Since I assume that each SBS selects only one subcarrier for its transmission, only one element of $\mathbf{P}_k$ will be non-zero.

The received SINR of the small cell user served by a generic SBS k over subcarrier n with power level l can be written as follows:

$$\text{SINR}_l^{(n)} = \frac{g_{k,k}^{(n)} p_l}{N_0 + \sum_{m \in \mathcal{I}_k^{m,(n)}} g_{m,k}^{(n)} p_m + \sum_{j \in \mathcal{I}_k^{s,(n)}} g_{j,k}^{(n)} p_j^{(n)}}, \tag{3.1}$$

where $g_{j,k}^{(n)}$ is the channel gain between the transmitter j and receiver k over subcarrier n, and N_0 is the variance of the additive white Gaussian noise at the receiver. The subscript m denotes the macro base station (MBS)/macro user. $\mathcal{I}_k^{m,(n)}$ and $\mathcal{I}_k^{s,(n)}$ denote the set of interfering MBSs and SBSs, respectively, on subcarrier n. Note that $g_{j,k}^{(n)}$ incorporates the path-loss and fading for which Rayleigh fading is considered. p_m denotes the transmit power of the MBS and p_l denotes the transmit power of an SBS which uses power level l. It is also assumed that at any given time instant, the MBSs transmit on all subcarriers with the same power level p_m.[3]

[3]This can be considered as the maximum transmit power at a macro BS. That is, for simplicity, I do not consider downlink power control at the macro BSs, which would depend on the subchannel and power allocation methods adopted by the macro BSs as well as the locations of the macro users

The interferer set $\mathcal{I}_k^{m,(n)}$ can also be modelled by the point process Φ_m. Specifically, Φ_m is modelled by a Poisson Point Process (PPP). Denoted by $\lambda_s^{(n)}$ the intensity of the small cell interferer set $\mathcal{I}_k^{s,(n)}$. The value of $\lambda_s^{(n)}$ is given by

$$\lambda_s^{(n)} = \lambda_s \times \left(\sum_{l=1}^{L} x_l^{(n)} \right), \tag{3.2}$$

where λ_s is the intensity of the SBSs and $x_l^{(n)}$ is the proportion of SBSs who transmit on subcarrier n with power level l.

It is also assumed that the small cell users have a perfect delay-free feedback channel to their base stations. Later in this chapter, I will relax this assumption by extending the formulation to explicitly consider the delay due to information exchange between users and the corresponding SBSs.

All the symbols that are used in the system model and the rest of the chapter are listed in Table 3.3.

3.4 Evolutionary Game Formulation

In this section, I first give the evolutionary game formulation for distributed resource allocation in self-organizing small cell networks. The replicator dynamics is used to model the strategy adaptation process. The evolutionary equilibrium is considered to be the solution of the formulated game and its stability is analyzed. I also extend the formulation by considering information exchange delay.

and their channel conditions.

Table 3.1: Chapter 3: Symbols

Symbol	**Description**
α	Path-loss exponent
Φ_m	PPP which represents the spatial distribution of MBSs
Φ_s	PPP which represents the spatial distribution of SBSs
λ_m	Density of Φ_m
λ_s	Density of Φ_s
π_a	Payoff of an SBS selecting action a
τ	Delay in information exchange
$\mathcal{A}$	Set of transmission alignments (actions)
A_c	Area of the small cell cluster
$d_{i,k}$	Distance between the BS i and user k
$F_l^{(n)}$	Number of SBSs on subcarrier n and power level l
$g_{j,k}^{(n)}$	Channel gain between BS j and user k on subcarrier n
$\mathcal{I}_k^{m,(n)}$	Set of interfering MBSs for user k on subcarrier n
$\mathcal{I}_k^{s,(n)}$	Set of interfering SBSs for user k on subcarrier n
K	Number of SBSs
k_a	Number of SBSs selecting action a
$\mathfrak{L}$	Set of power levels
$\mathcal{L}_{f(x)}(v)$	Laplace transform of $f(x)$
$\mathcal{N}$	Set of subcarriers
N_0	Noise variance
$p_k^{(n)}$	Transmit power of SBS k on subcarrier n
p_l	Transmit power of an SBS using power level l
p_m	MBS transmit power
r_s	Radius of a small cell
$r_l^{(n)}$	Achievable rate of an SBS on subcarrier n, power level l
$T^{(n)}$	Interference threshold for macro users on subcarrier n
w_1, w_2	Biasing factors
x_a	Proportion of SBSs using action a
$x_l^{(n)}$	Proportion of SBSs using subcarrier n, power level l

3.4.1 Game Formulation

The adaptive subcarrier-power allocation among bounded-rational SBSs can be formulated as an evolutionary game as follows:

Set of players ($\mathcal{K}$): The set of small cell base stations $\mathcal{K} = \{1, 2, ..., K\}$ denote the set of players of the game.

Set of actions ($\mathcal{A}$): The SBSs (i.e., the players) are interested in selecting a suitable transmission alignment (i.e., subcarrier-power level combination). According to the system model, each player has $C = |\mathfrak{L}| \times |\mathcal{N}|$ transmission alignments. Accordingly, I define the action set for each player as $\mathcal{A} = \{a_1, ..., a_C\}$ which includes all possible transmission alignments.

Note that it is also possible to consider the case where each SBS selects multiple subcarriers. Then the number of actions available for a player will be equal to $\sum_{r=1}^{N} \binom{N}{r} |\mathfrak{L}|^r$. For simplicity of analysis, in this work I assume that each SBS selects only one subcarrier for transmission.

Population: In the context of an evolutionary game, the set of players also constitutes the population. Denote by K and k_a the total number of SBSs in the cluster and the number of SBSs selecting action $a \in \mathcal{A}$, respectively. Then the frequency of action a used in the population is given by

$$x_a = \frac{k_a}{K}, \tag{3.3}$$

where the frequency x_a is also referred to as the population share of action a. The population share of all actions constitute the population state denoted by a vector $\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_{|\mathcal{A}|}]^T$. Note that $x_l^{(n)} = x_a$, where action a corresponds to the selection of subcarrier n and power level l. In the remaining of this chapter, I use

$x_l^{(n)}$ and x_a interchangeably.

Payoff function: The payoff of an SBS quantifies its satisfaction level of selecting action a for which two components are considered. The first component measures the utility of an SBS when certain transmission alignment is used. The second component is a penalty (or reward) term depending on whether the aggregate interference from the small cell cluster to the macro user receiving on the selected subcarrier exceeds the threshold or not. In this case, the protection to the macro users can be provided. Specifically, the payoff of an SBS selecting action a is defined as

$$\pi_a = \pi_l^{(n)} = w_1 \left(\mathcal{U}(\mathrm{SINR}_l^{(n)}) \right) - w_2 \left(I_m^{(n)} - T^{(n)} \right), \tag{3.4}$$

where w_1 and w_2 are biasing factors which can be determined based on which network (i.e., macro or small cell network) should be given priority in resource allocation. $\mathcal{U}(\mathrm{SINR}_l^{(n)})$ is a utility function measuring the achieved performance. $T^{(n)}$ is the interference threshold for subcarrier n. $I_m^{(n)}$ is the aggregate interference caused by the small cell cluster which needs to be estimated. Denote by $F_l^{(n)}$ the number of SBSs transmitting using subcarrier n and power level l. $I_m^{(n)}$ can be approximated by the following expression:

$$I_m^{(n)} = \sum_{k=1}^{K} p_k^{(n)} \bar{g}_{h,m}^{(n)}, \tag{3.5}$$

or equivalently,

$$I_m^{(n)} = \sum_{l=1}^{L} F_l^{(n)} \; p_l \; \bar{g}_{h,m}^{(n)} = \bar{g}_{h,m}^{(n)} \; A_c \lambda_s \; \sum_{l=1}^{L} x_l^{(n)} \; p_l, \tag{3.6}$$

where A_c is the area of the small cell cluster, $\bar{g}_{h,m}^{(n)}$ is the average channel gain from a generic SBS to the macro user receiving on subcarrier n, and $x_l^{(n)}$ is the fraction of

SBSs transmitting on subcarrier n with power level l.

Specifically, I consider two utility functions as follows[4]:

$$\mathcal{U}_1(\mathrm{SINR}_l^{(n)}) = \mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right], \tag{3.7}$$

and

$$\mathcal{U}_2(\mathrm{SINR}_l^{(n)}) = \mathbf{E}\left[r_l^{(n)}\right]. \tag{3.8}$$

Accordingly, for an SBS selecting action a, corresponding payoff functions for each of the above utilities $\pi_a^{(1)}$ and $\pi_a^{(2)}$ can also be written as follows:

$$\pi_a^{(1)} = (\pi^{(1)})_l^{(n)} = w_1^{(1)}\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] - w_2^{(1)}\left(I_m^{(n)} - T^{(n)}\right), \tag{3.9}$$

and

$$\pi_a^{(2)} = (\pi^{(2)})_l^{(n)} = w_1^{(2)}\mathbf{E}\left[r_l^{(n)}\right] - w_2^{(2)}\left(I_m^{(n)} - T^{(n)}\right). \tag{3.10}$$

With the above definitions, two evolutionary games $\mathcal{G}^1$ and $\mathcal{G}^2$ are defined as follows:

$$\mathcal{G}^1 = \left(\mathcal{K}, \mathcal{A}, \pi_a^{(1)})\right), \tag{3.11}$$

and

$$\mathcal{G}^2 = \left(\mathcal{K}, \mathcal{A}, \pi_a^{(2)})\right). \tag{3.12}$$

Hereafter, superscripts "(1)" and "(2)" will be used in payoff (π_a), utility (u), and population share (x_a) to denote games $\mathcal{G}^1$ and $\mathcal{G}^2$, respectively.

[4]Later in this chapter I will show that these two utility functions result in different convergence behaviors of the proposed resource allocation algorithm and also different amount of average interference caused to macro users. Also, these functions can correspond to two different types of application requirements.

3.4.2 Dynamics of Adaptation of Transmission Alignment

In the context of the evolutionary game for transmission alignment selection, each SBS will adapt its strategy according to its received payoff. This is referred to as the evolution of the game during which the strategy adaptation of SBSs will change the population share, and therefore, the population state will evolve over time.

In this case, the population share x_a and population state $\mathbf{x}$ are functions of time t which can be denoted as $x_a(t)$ and $\mathbf{x}(t)$, respectively. The strategy adaptation process and the corresponding population state evolution can be modeled and analyzed by *replicator dynamics*, which is a set of ordinary differential equations defined as follows:

$$\dot{x}_a(t) = x_a(t)\left(\pi_a(t) - \bar{\pi}(t)\right), \tag{3.13}$$

for all $a \in \mathcal{A}$, with initial population state $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{X}$, where $\mathbb{X}$ is the state space which contains all possible population distributions. π_a is the payoff of each SBS choosing transmission alignment a and $\bar{\pi}$ is the average payoff of the entire population. The average payoff $\bar{\pi}$ is given by

$$\bar{\pi} = \sum_{a \in \mathcal{A}} \pi_a x_a. \tag{3.14}$$

The replicator dynamics governs the rate of strategy (i.e., transmission alignment) adaptation of the population. As the game (either $\mathcal{G}^1$ or $\mathcal{G}^2$) is repeated, each SBS observes its own payoff and compares it with the average payoff of the system. Then, in the next period, the SBS randomly selects another strategy if its payoff is less than the average. According to the replicator dynamics, the number of SBSs selecting transmission alignment a will increase if the corresponding payoff is higher than the average (i.e., $\pi_a > \bar{\pi}$).

Lemma 3.4.1. *With the replicator dynamics defined in (3.13), the summation of all population shares of all strategies in $\mathcal{A}$ can be guaranteed to be equal to one during the evolution of transmission alignment selection. Also, each population share x_a strictly falls in [0,1].*

Proof. The summation of the variation of rate of all population shares can be obtained as

$$\sum_{a}^{|\mathcal{A}|} \dot{x}_a(t) = \sum_{a}^{|\mathcal{A}|} x_a(t)\pi_a(t) - \sum_{a}^{|\mathcal{A}|} x_a(t)\bar{\pi}(t). \tag{3.15}$$

According to the definition, $\bar{\pi}(t) = \sum_{a=1}^{|\mathcal{A}|} x_a(t)\pi_a(t)$. Therefore, with the initial condition $\sum_{a}^{|\mathcal{A}|} x_a(0) = 1$, we have $\sum_{i}^{|\mathcal{A}|} \dot{x}_a(t) = 0$ which indicates that the summation of all population shares does not vary with time. Also, we have $\dot{x}_a(t) \leq 0$ with $x_a(t) = 1$, and $\dot{x}_a(t) \geq 0$ with $x_a(t) = 0$. Therefore, $x_a(t) \in [0, 1]$ can be guaranteed for all $t \in [0, \infty)$. □

3.4.3 Delay in Replicator Dynamics

In the evolutionary game of transmission alignment selection, the payoff of an SBS at a particular time depends on the strategies of all players in the population (i.e., the population state) at that time. However, in practical networks, the information about the population state and average payoff may be delayed due to the latency in the signaling channels. Therefore, there is a gap between the time that a population state arises and the time that the population state impacts the payoffs of the SBSs. In that case, the SBSs have to make their decisions based on the previously received data. Denote by τ the delay in information exchange. Then the decisions of SBSs on transmission alignment selection at time t are made according to the population

state at $t - \tau$. This delay impacts the strategy adaptation process (i.e., replicator dynamics). Specifically, the rate of variation of population share adopting a particular strategy is proportional to the difference between the delayed payoff for that strategy and the delayed average payoff of the population. In this case, the replicator dynamics with delayed information can be written as a set of delayed differential equations as shown below:

$$\dot{x}_a = x_a(t-\tau)\left(\pi_a(t-\tau) - \bar{\pi}(t-\tau)\right). \tag{3.16}$$

With delayed replicator dynamics, I investigate the impact of information exchange delays on the convergence of strategy adaptation process to the equilibrium.

To quantify the players' payoff functions in the game formulation presented above, in the following section, I use stochastic geometry-based analysis to obtain average SINR and average transmission rate for a user served by a generic SBS with a certain transmission alignment.

3.5 Analysis of Average SINR and Achievable Rate

At each step of the population evolution of the game $\mathcal{G}^1$ and $\mathcal{G}^2$, due to the bounded rationality of the players, the selection of transmission alignment is random. During the population evolution of the game, the penalty term in the payoff function affects how many SBSs selects a particular transmission alignment. However, the spatial distribution of the SBSs transmitting on each subcarrier is still random. Therefore, a PPP can be used to model their spatial distribution.

According to Slivnyak's theorem [56], the statistics for a PPP is independent of the test location. Therefore, the analysis holds for any generic small cell user located at a generic location.

By explicitly incorporating the distance gain, the instantaneous SINR for a generic

user served by SBS k in (3.1) can be rewritten as

$$\mathrm{SINR}_l^{(n)} = \frac{h_{k,k}^{(n)} r_s^{-\alpha} p_l}{N_0 + \sum_{i \in \mathcal{I}_k^{(n)}} h_{i,k}^{(n)} d_{i,k}^{-\alpha} p_i^{(n)}}, \tag{3.17}$$

where $h_{i,k}^{(n)}$ and $d_{i,k}$ are the channel gains due to fading and the distance between the base station i and user of SBS k, respectively, α is the path-loss exponent, and $\mathcal{I}_k^{(n)}$ denotes the union of the two interferer sets. That is, the interfering set $\mathcal{I}_k^{(n)}$ is composed of two sets, i.e., the set of macro interferers $\mathcal{I}_k^{m,(n)}$ and the set of small cell interferers $\mathcal{I}_k^{s,(n)}$.

3.5.1 *Average Received* SINR

The expected average SINR of a user attached to SBS k, receiving on subcarrier n with power level l, is given by

$$\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \mathbf{E}_{h_{k,k}^{(n)}, i \in \mathcal{I}_k^{(n)}} \left[\frac{h_{k,k}^{(n)} r_s^{-\alpha} p_l}{N_0 + \sum_{i \in \mathcal{I}_k^{(n)}} h_{i,k}^{(n)} d_{i,k}^{-\alpha} p_i^{(n)}} \right]. \tag{3.18}$$

Since the expectation of any positive random variable X is given by $\mathbf{E}[X] = \int_0^\infty \Pr\,(x > t)\,\mathrm{d}t$,

$$\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \mathbf{E}_{i \in \mathcal{I}_k^{(n)}} \left[\int_{t=0}^{\infty} \Pr\left(h_{k,k}^{(n)} > \frac{r_s^{\alpha} \left(N_0 + I_k^{(n)} \right) t}{p_l} \right) \mathrm{d}t \right], \tag{3.19}$$

where $I_k^{(n)} = \sum_{i \in \mathcal{I}_k^{(n)}} h_{i,k}^{(n)} d_{i,k}^{-\alpha} p_i^{(n)}$ is the aggregate interference at the receiver served by SBS k on subcarrier n.

Assuming Rayleigh fading, which results in exponentially distributed channel power gains (i.e., $h_{k,k}^{(n)} \sim \exp(\mu)$),

$$\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \mathbf{E}_{i\in\mathcal{I}_k^{(n)}}\left[\int_{t=0}^{\infty}\exp\left(-\mu\frac{r_s^{\alpha}\left(N_0+I_k^{(n)}\right)t}{p_l}\right)\mathrm{d}t\right]. \tag{3.20}$$

By averaging over all possible interferers, the following can be derived.

$$\begin{aligned}\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] &= \int_{t=0}^{\infty}\exp\left(-\frac{\mu t r_s^{\alpha}N_0}{p_l}\right)\\ &\int_{i_k^{s,(n)}\in\mathcal{I}_k^{s,(n)}}\exp\left(-\frac{\mu t r_s^{\alpha}i_k^{s,(n)}}{p_l}\right)f_{I_k^{s,(n)}}(i_k^{s,(n)})\mathrm{d}i_k^{s,(n)}\\ &\int_{i_k^{m,(n)}\in\mathcal{I}_k^{m,(n)}}\exp\left(-\frac{\mu t r_s^{\alpha}i_k^{m,(n)}}{p_l}\right)f_{I_k^{m,(n)}}(i_k^{m,(n)})\mathrm{d}i_k^{m,(n)}\ \mathrm{d}t.\end{aligned} \tag{3.21}$$

Let $v = \frac{\mu t r_s^{\alpha}}{p_l}$. Then

$$\begin{aligned}&\int_{i_k^{m,(n)}\in\mathcal{I}_k^{m,(n)}}\exp\left(-\frac{\mu t r_s^{\alpha}i_k^{m,(n)}}{p_l}\right)f_{I_k^{m,(n)}}(i_k^{m,(n)})\mathrm{d}i_k^{s,(n)}\\ &= \int_{i_k^{m,(n)}\in\mathcal{I}_k^{m,(n)}}\exp\left(-vi_k^{m,(n)}\right)f_{I_k^{m,(n)}}(i_k^{m,(n)})\mathrm{d}i_k^{s,(n)}\\ &= \mathcal{L}_{I_k^{m,(n)}}(v),\end{aligned}$$

where $\mathcal{L}_{I_k^{m,(n)}}(v)$ is the Laplace transform of the Probability Density Function (PDF) of aggregate interference[5] coming from macro network (i.e., cross-tier interference) at a generic user k receiving on subcarrier n.

Similarly, integration over the set $\mathcal{I}_k^{s,(n)}$ in (3.21) can be reduced as $\mathcal{L}_{I_k^{s,(n)}}(v)$ which is the Laplace transform of the interference caused by the small cell network (i.e., co-

[5]For notational convenience, this is simply referred to as the Laplace transform of the aggregate interference.

tier interference) at the user k on subcarrier n. Substituting into equation (3.21), the following can be derived.

$$\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \int_{t=0}^{\infty} \exp\left(-vN_0\right) \mathcal{L}_{I_k^{m,(n)}}(v) \mathcal{L}_{I_k^{s,(n)}}(v)\, \mathrm{d}t. \tag{3.22}$$

The derivation of the Laplace transform of the aggregate interference has been well explained in several previous work [57], [58]. However, since these works assume constant transmit power, the derived expressions in these work cannot be directly used for the analysis in this case. For this system model, the expectation of the Laplace transform has to be taken with respect to the point process for the interferers, the Rayleigh fading gain, and also the transmit power.

Recall that $\mathcal{I}_k^{s,(n)}$ denotes the point process representing the spatial distribution of the interfering small cell base stations. Then $\mathcal{L}_{I_k^{s,(n)}}(v)$ can be expressed as follows:

$$\begin{aligned}
\mathcal{L}_{I_k^{s,(n)}}(v) &= \mathbf{E}_{\mathcal{I}_k^{s,(n)}}\left[\exp\left(-vI_k^{s,(n)}\right)\right] \\
&= \mathbf{E}_{\mathcal{I}_k^{s,(n)}} \mathbf{E}_{h_{i,k}^{(n)}} \mathbf{E}_{p_i}\left[\exp\left(-v \sum_{i \epsilon \mathcal{I}_k^{s,(n)}} p_i d_{i,k}^{-\alpha} h_{i,k}^{(n)}\right)\right] \\
&= \mathbf{E}_{\mathcal{I}_k^{s,(n)}}\left[\Pi_{i \in \mathcal{I}_k^{s,(n)}} \mathbf{E}_{h_{i,k}^{(n)}} \mathbf{E}_{p_i}\left[\exp\left(-vp_i d_{i,k}^{-\alpha} h_{i,k}^{(n)}\right)\right]\right],
\end{aligned} \tag{3.23}$$

where p_i denotes the transmit power of an SBS ($p_i \in \{p_1, p_2, \cdots, p_L\}$).

Due to the random selection of transmission alignments at each step of the population evolution, $\mathcal{I}_k^{s,(n)}$ can also be modeled by a PPP with the intensity given by (3.2). The Probability Generating Functional (PGFL) of a 2D PPP Φ can be written

as follows [57]:

$$\mathbf{E}_{\Phi}[\Pi_{x\in\Phi} f(x)] = \exp\left(-\lambda \int_{\mathbb{R}^2} (1 - f(x))\, \mathrm{d}x\right), \tag{3.24}$$

where λ is the intensity of the point process. Therefore, $\mathcal{L}_{I_k^{s,(n)}}(v)$ can be written follows:

$$\begin{aligned}\mathcal{L}_{I_k^{s,(n)}}(v) &= \exp\left(-\lambda_s^{(n)} \int_{\mathbb{R}^2} \left(1 - \mathbf{E}_{h_{i,k}^{(n)}} \mathbf{E}_{p_i}[\exp(-v p_i d_{i,k}^{-\alpha} h_{i,k}^{(n)})]\right)\right) \\ &= \exp\left(-2\pi\lambda_s^{(n)} \int_{d_{i,k}=0}^{\infty} \left(1 - \mathbf{E}_{h_{i,k}^{(n)}} \mathbf{E}_{p_i}[\exp(-v p_i d_{i,k}^{-\alpha} h_{i,k}^{(n)})]\right)\, d_{i,k}\, \mathrm{d}d_{i,k}\right),\end{aligned} \tag{3.25}$$

where $\lambda_s^{(n)}$ is the intensity of interfering SBSs on subcarrier n. Note that although the lower limit of the integration over $d_{i,k}$ is considered as zero in this work, there can be a minimum separation between the base stations in practical networks. However, in that case the expression in (3.25) will have an incomplete gamma function and hence the expression for average SINR will not be simplified into a closed form.

By simplifying the expression given in (3.25) (see Appendix A.1), the Laplace transform of the aggregate interference caused by the small cell network is given by equation (3.26) as follows:

$$\mathcal{L}_{I_k^{s,(n)}}(v) = \quad \exp\left\{-\pi\lambda_s^{(n)}\, \mathbf{E}_{p_s}[p_s^{\frac{2}{\alpha}}] \mathbf{E}_{h_{i,k}^{(n)}}[{h_{i,k}^{(n)}}^{\frac{2}{\alpha}}] v^{\frac{2}{\alpha}} \Gamma\left(1 - \frac{2}{\alpha}\right)\right\}, \tag{3.26}$$

where $\Gamma(z)$ is the complete Gamma function given by, $\Gamma(z) = \int_{t=0}^{\infty} t^{z-1} e^{-t} \mathrm{d}t$, and p_s is the transmit power of a generic SBS in the interferer set.

Recall that all MBSs transmit on all subcarriers with uniform power p_m. By

following a similar derivation as above the following can be obtain

$$\mathcal{L}_{I_k^{m,(n)}}(v) = \exp\left\{-\pi\lambda_m \; p_m^{\frac{2}{\alpha}} \mathbf{E}_{h_{m,k}^{(n)}}[{h_{m,k}^{(n)}}^{\frac{2}{\alpha}}] v^{\frac{2}{\alpha}} \Gamma\left(1-\frac{2}{\alpha}\right)\right\}. \tag{3.27}$$

By substituting the expressions obtained in (3.26) and (3.27) into equation (3.22), equation (3.28) can be obtained as follows:

$$\begin{aligned}\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = &\int_{t=0}^{\infty} \exp\left(-vN_0\right) \\ &\exp\left\{-\pi\lambda_m \; p_m^{\frac{2}{\alpha}} \mathbf{E}_{h_{m,k}^{(n)}}[{h_{m,k}^{(n)}}^{\frac{2}{\alpha}}] v^{\frac{2}{\alpha}} \Gamma\left(1-\frac{2}{\alpha}\right)\right\} \\ &\exp\left\{-\pi\lambda_s^{(n)} \mathbf{E}_{p_s}[p_s^{\frac{2}{\alpha}}] \mathbf{E}_{h_{i,k}^{(n)}}[{h_{i,k}^{(n)}}^{\frac{2}{\alpha}}] v^{\frac{2}{\alpha}} \Gamma\left(1-\frac{2}{\alpha}\right)\right\} \, \mathrm{d}t.\end{aligned} \tag{3.28}$$

Since I consider a dense network, the network will be interference limited. Hence, the effect of noise can be considered negligible compared to the interference. I also assume $\alpha = 4$ for analytical simplicity. Then by simplifying the expression given in (3.28) the following can be obtained (see Appendix A.2 for derivation):

$$\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \frac{8p_l}{A^2\left(\lambda_m\sqrt{p_m} + \lambda_s^{(n)}\mathbf{E}\left[\sqrt{p_s}\right]\right)^2}, \tag{3.29}$$

where $A = \pi^2 r_s^2$.

The Probability Mass Function (PMF) of the transmit power of any interferer (i.e., p_s in (3.29)) can be directly obtained from the proportions of the population selecting each action. For transmission alignment corresponding to subcarrier n and

power level l, the PMF of the transmit power of a generic interferer is given as follows:

$$\Pr\left(p_s = p_j\right) = \begin{cases} \frac{k_j^{(n)}}{\sum_{t=1}^{L} k_t^{(n)} - 1}, & \text{if } j \neq l, \\ \frac{k_j^{(n)} - 1}{\sum_{t=1}^{L} k_t^{(n)} - 1}, & \text{if } j = l, \end{cases}$$

or equivalently,

$$\Pr\left(p_s = p_j\right) = \begin{cases} \frac{x_j^{(n)}}{\sum_{t=1}^{L} x_t^{(n)} - \frac{1}{K}}, & \text{if } j \neq l, \\ \frac{x_j^{(n)} - \frac{1}{K}}{\sum_{l=t}^{L} x_t^{(n)} - \frac{1}{K}}, & \text{if } j = l, \end{cases}$$

where $k_j^{(n)}$ is the number of players selecting subcarrier n and power level j and $x_j^{(n)} = \frac{k_j^{(n)}}{K}$. Therefore, for a user receiving on subcarrier n and power level l, $\mathbf{E}[\sqrt{p_s}]$ can be calculated as follows:

$$\begin{aligned} \mathbf{E}[\sqrt{p_s}] = & \frac{p_1 x_1^{(n)}}{\sum_{t=1}^{L} x_t^{(n)} - \frac{1}{K}} + \frac{p_2 x_2^{(n)}}{\sum_{t=1}^{L} x_t^{(n)} - \frac{1}{K}} + ... + \\ & \frac{p_l \left(x_l^{(n)} - \frac{1}{K}\right)}{\sum_{l=t}^{L} x_t^{(n)} - \frac{1}{K}} + ... + \frac{p_L x_L^{(n)}}{\sum_{t=1}^{L} x_t^{(n)} - \frac{1}{K}}. \end{aligned} \tag{3.30}$$

The value of $\lambda_s^{(n)}$ is given by (3.2).

3.5.2 Average Achievable Rate

The average achievable rate of a generic small cell user for transmission on subcarrier n using power level l is given by

$$\mathbf{E}[r_l^{(n)}] = \ln\left(1 + \mathrm{SINR}_l^{(n)}\right). \tag{3.31}$$

Same as in Section 3.5.1, the expectation has to be taken w.r.t. the channel gain and interfering nodes. Therefore,

$$\begin{aligned}\mathbf{E}[r_l^{(n)}] &= \mathbf{E}_{h_{k,k}^{(n)}, i\in\mathcal{I}_k^{(n)}}\left[\ln\left(1+\frac{h_{k,k}^{(n)} r_s^{-\alpha} p_l}{N_0+\sum_{i\in\mathcal{I}_k^{(n)}} h_{i,k}^{(n)} d_{i,k}^{-\alpha} p_i^{(n)}}\right)\right] \\ &= \mathbf{E}_{i\in\mathcal{I}_k^{(n)}}\left[\int_{t=0}^{\infty}\Pr\left(h_{k,k}^{(n)} > \frac{r_s^{\alpha}(e^t-1)(N_0+I_k^{(n)})}{p_l}\right)\mathrm{d}t\right].\end{aligned}$$

Since $h_{k,k}^{(n)} \sim \exp(\mu)$, by following the same steps as in the analysis of SINR in Section 3.5.1, I obtain

$$\mathbf{E}\left[r_l^{(n)}\right] = \int_{t=0}^{\infty}\exp(-vN_0)\mathcal{L}_{I_k^{m,n}}(v)\mathcal{L}_{I_k^{s,n}}(v)\mathrm{d}t, \tag{3.32}$$

where $v = \frac{\mu r^{\alpha}(e^t-1)}{p_l}$. By simplifying the above expression for an interference-limited network and $\alpha = 4$, the following can be obtained:

$$\mathbf{E}\left[r_l^{(n)}\right] = \int_{t=0}^{\infty}\exp\left(\frac{-A}{2\sqrt{p_l}}\left(\lambda_m\sqrt{p_m}+\lambda_s^{(n)}\mathbf{E}\left[\sqrt{p_s}\right]\right)\sqrt{e^t-1}\right)\,\mathrm{d}t, \tag{3.33}$$

where $A = \pi^2 r_s^2$.

In the next section, I will use the derived expressions in (3.29) and (3.33) to analyze the stability of the evolutionary equilibrium of the game model.

3.6 Evolutionary Equilibrium and Stability Analysis

3.6.1 Evolutionary Equilibrium and It's Existence

Evolutionary equilibrium, which is defined as the fixed points of the replicator dynamics [49], is considered to be the solution of the evolutionary games $\mathcal{G}^1$ and $\mathcal{G}^2$ for transmission alignment selection. That is, at the equilibrium point, the population state will not change, which implies that the rate of strategy adaptation will be be

zero ($\dot{x}_l^{(n)} = 0$). Once the evolutionary equilibrium is reached, no player has the willingness to change its strategy since all SBSs have the same payoff as the average payoff of the population. In this way, the evolutionary equilibrium can also provide fairness among the SBSs.

Using the expressions given in (3.29) and (3.33), the replicator dynamics $\dot{x}_a^{(1)}$ and $\dot{x}_a^{(2)}$ of the games $\mathcal{G}^1$ and $\mathcal{G}^2$, respectively, can be expressed as in (3.34) and (3.35), for all $a \in \mathcal{A}$.

$$\begin{aligned}\dot{x}_a^{(1)} &= (\dot{x}^{(1)})_l^{(n)} \\ &= (x^{(1)})_l^{(n)} \left(\frac{w_1^{(1)} 8 p_l}{A^2 \left(\lambda_m \sqrt{p_m} + \lambda_s^{(n)} \mathbf{E}\left[\sqrt{p_s}\right] \right)^2} - w_2^{(1)} \left(\bar{g}_{h,m}^{(n)} A_c \lambda_s \sum_{q=1}^{L} (x^{(1)})_q^{(n)} p_q - T^{(n)} \right) - \right. \\ &\quad \left. \sum_{n \in \mathcal{N}} \sum_{q \in \mathcal{L}} (x^1)_q^{(n)} \pi_q^{(n)} \right). \end{aligned} \tag{3.34}$$

$$\begin{aligned}\dot{x}_a^{(2)} &= (\dot{x}^2)_l^{(n)} \\ &= (x^{(2)})_l^{(n)} w_1^{(2)} \int_{t=0}^{\infty} \exp\left(\frac{-A}{2\sqrt{p_l}} \left(\lambda_m \sqrt{p_m} + \lambda_s^{(n)} \mathbf{E}\left[\sqrt{p_s}\right] \right) \sqrt{e^t - 1} \right) \, \mathrm{d}t \\ &\quad - (x^{(2)})_l^{(n)} w_2^{(2)} \left(\bar{g}_{h,m}^{(n)} A_c \lambda_s \sum_{q=1}^{L} (x^{(2)})_q^{(n)} p_q - T^{(n)} \right) - (x^{(2)})_l^{(n)} \sum_{n \in \mathcal{N}} \sum_{q \in \mathcal{L}} (x^{(2)})_q^{(n)} \pi_q^{(n)}. \end{aligned} \tag{3.35}$$

According to the definition, obtaining the evolutionary equilibrium is equivalent to solving a system of algebraic equations given by setting the left hand side of replicator dynamics (3.34) and (3.35) to zero (i.e., let $\dot{x}_l^{(n)} = 0$).

Note that in the replicator dynamics defined in (3.34) and (3.35), two types of

evolutionary equilibrium, namely, boundary evolutionary equilibrium and interior evolutionary equilibrium exist. The boundary evolutionary equilibrium corresponds to the case where there exists a population share $x_a = 1$, while $x_b = 0$ for all $b \neq a \in S$. The interior equilibrium $\mathbf{x}^*$ corresponds to the case where $x_a \in (0, 1), \forall a \in S$. The boundary equilibria are not stable in the sense that any small perturbation will make the system deviate from the equilibrium state.

3.6.2 Stability of Evolutionary Equilibrium

To evaluate the stability of the interior evolutionary equilibrium, the eigenvalues of the Jacobian matrix corresponding to the replicator dynamics need to be evaluated. The system is stable if all eigenvalues have a negative real part [59]. With the analytical expressions obtained from the stochastic geometry-based analysis, the stability of the equilibrium is analytically tractable in some cases. In the following, I analyze the stability of the equilibrium of game $\mathcal{G}^1$ for a system with two subcarriers and uniform power as an example. The stability of the game $\mathcal{G}^2$ is not analyzed since $\mathcal{U}_2(\text{SINR}_l^{(n)})$ is not available in closed form. However, the results on the stability (i.e., convergence) of $\mathcal{G}^2$ will be obtained by simulations in Section 3.8.

Specifically, I consider a system with two subcarriers and all SBSs transmitting on the same power level. In this case there are two transmission alignments. Denote the two transmission alignments by a_1 and a_2, which correspond to the selection of subcarrier 1 and subcarrier 2, respectively. $\pi_{a_1}^{(1)}$ and $\pi_{a_2}^{(1)}$ are the corresponding payoffs and $x_{a_1}^{(1)}$ and $x_{a_2}^{(1)}$ are the corresponding population shares in $\mathcal{G}^1$.

For the case where $x_{a_1}^{(1)}, x_{a_2}^{(1)} \neq 0$, from (3.2),

$$\lambda_s^{(n)} = \lambda_s x_{a_n}^{(1)}. \tag{3.36}$$

Also, $\mathbf{E}\left[\sqrt{p}\right]$ will be simplified to $\sqrt{p}$ since I assume uniform power for all SBSs. By substituting the values for $\mathbf{E}\left[\sqrt{p}\right]$ and $\lambda_s^{(n)}$ into (3.29), we obtain

$$\pi_{a_1}^{(1)} = \frac{w_1^{(1)} 8p}{A^2\left(\lambda_m\sqrt{p_m} + \lambda_s\sqrt{p}(x_{a_1}^{(1)})\right)^2} - w_2^{(1)}\left(\bar{g}_{h,m}A_c\lambda_s x_{a_1}^{(1)}p - T^{(n)}\right), \quad (3.37)$$

and

$$\pi_{a_2}^{(1)} = \frac{w_1^{(1)} 8p}{A^2\left(\lambda_m\sqrt{p_m} + \lambda_s\sqrt{p}(x_{a_2}^{(1)})\right)^2} - w_2^{(1)}\left(\bar{g}_{h,m}A_c\lambda_s x_{a_2}^{(1)}p - T^{(n)}\right). \quad (3.38)$$

For any subcarrier with zero population will result in zero payoff. Then

$$\pi_i^{(1)}{}_{\left(\text{for } x_i^{(1)}=1\right)} = \frac{w_1^{(1)} 8p}{A^2\left(\lambda_m\sqrt{p_m}\lambda_s\sqrt{p}\right)^2} - w_2^{(1)}\left(\bar{g}_{h,m}A_c\lambda_s p - T^{(n)}\right). \quad (3.39)$$

For the stability of the system, the following Theorem can be stated.

Theorem 3.6.1. *For a network with two orthogonal subcarriers and one transmit power level, the interior evolutionary equilibrium in game $\mathcal{G}^1$ is asymptotically stable.*

Proof. The replicator dynamics can be derived as follows:

$$\begin{aligned}
\dot{x}_{a_1}^{(1)} &= x_{a_1}^{(1)}\left(\pi_{a_1}^{(1)} - x_{a_1}^{(1)}\pi_{a_1}^{(1)} - x_{a_2}^{(1)}\pi_{a_2}^{(1)}\right) \\
&= x_{a_1}^{(1)}\left(\pi_{a_1}^{(1)}\left(1 - x_{a_1}^{(1)}\right) - x_{a_2}^{(1)}\pi_{a_2}^{(1)}\right) \\
&= x_{a_1}^{(1)}x_{a_2}^{(1)}\left(\pi_{a_1}^{(1)} - \pi_{a_2}^{(1)}\right) \\
&= x_{a_1}^{(1)}\left(1 - x_{a_1}^{(1)}\right)\left(\pi_{a_1}^{(1)} - \pi_{a_2}^{(1)}\right),
\end{aligned} \quad (3.40)$$

where the 3^{rd} step follows from the fact that $(x^{(1)})_{a_1} + (x^{(1)})_{a_2} = 1$.

For the system to have a stable equilibrium point, all eigenvalues of the Jacobian of the system of equations should have a negative real part. As there is one system equation, an equivalence to this condition is that the Jacobian should be negative definite [60]. To this end, I first denote by f the right hand side of (3.40). Accordingly,

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}x_{a_1}^{(1)}} &= x_{a_1}^{(1)}\left(1-x_{a_1}^{(1)}\right)\times\left(\frac{\mathrm{d}\pi_{a_1}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}}-\frac{\mathrm{d}\pi_{a_2}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}}\right)+\left(\pi_{a_1}^{(1)}-\pi_{a_2}^{(1)}\right)\left(1-2x_{a_1}^{(1)}\right)\\ &= x_{a_1}^{(1)}\left(1-x_{a_1}^{(1)}\right)\times\left(\frac{\mathrm{d}\pi_{a_1}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}}-\frac{\mathrm{d}\pi_{a_2}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}}\right)+\left(\pi_{a_1}^{(1)}-\pi_{a_2}^{(1)}\right)\left(x_{a_2}^{(1)}-x_{a_1}^{(1)}\right). \quad (3.41)\end{aligned}$$

According to the definition, $\pi_{a_2}^{(1)}=\pi_{a_1}^{(1)}$ at the equilibrium point. Hence, $\left(\pi_{a_1}^{(1)}-\pi_{a_2}^{(1)}\right)\left(x_{a_2}^{(1)}-x_{a_1}^{(1)}\right)=0$. Then $\frac{\mathrm{d}\pi_{a_1}^{(1)}}{dx_{a_1}^{(1)}}$ and $\frac{\mathrm{d}\pi_{a_2}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}}$ can be calculated as in equations (3.42) and (3.43).

$$\frac{\mathrm{d}\pi_{a_1}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}}=\frac{-w_1^{(1)}16p^{\frac{3}{2}}\lambda_s}{A^2\left(\lambda_m\sqrt{p_m}+\lambda_s\sqrt{p}x_1^{(1)}\right)^3}-w_2^{(1)}\left(\bar{g}_{h,m}A_c\lambda_s p\right). \quad (3.42)$$

$$\begin{aligned}\frac{\mathrm{d}\pi_{a_2}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}} &= \frac{\mathrm{d}\left(\frac{w_1^{(1)}8p}{A^2\left(\lambda_m\sqrt{p_m}+\lambda_s\sqrt{p}x_{a_2}^{(1)}\right)^2}-w_2^{(1)}\bar{g}_{h,m}A_c\lambda_s x_{a_2}^{(1)}p+w_2^{(1)}T^{(n)}\right)}{\mathrm{d}x_{a_1}^{(1)}}\\ &= \frac{\mathrm{d}\left(\frac{w_1^{(1)}8p}{A^2\left(\lambda_m\sqrt{p_m}+\lambda_s\sqrt{p}(1-x_{a_1}^{(1)})\right)^2}-w_2^{(1)}\bar{g}_{h,m}A_c\lambda_s(1-x_{a_1}^{(1)})p+w_2^{(1)}T^{(n)}\right)}{\mathrm{d}x_{a_1}^{(1)}}\\ &= \frac{w_1^{(1)}16p^{\frac{3}{2}}\lambda_s}{A^2\left(\lambda_m\sqrt{p_m}+\lambda_s^{(2)}\sqrt{p}(1-x_{a_1}^{(1)})\right)^3}+w_2^{(1)}\left(\bar{g}_{h,m}A_c\lambda_s p\right). \quad (3.43)\end{aligned}$$

Note that $w_1^{(1)}$ and $w_2^{(1)}$ are positive values. From (3.42) and (3.43), it is obvious

that for any $x_{a_1}^{(1)}, x_{a_2}^{(1)} > 0$, I have $\frac{\partial \pi_{a_1}^{(1)}}{\partial x_{a_1}^{(1)}} > 0$ and $\frac{\partial \pi_{a_2}^{(1)}}{\partial x_{a_2}^{(1)}} < 0$. For any $x_i = 0$, $\pi_i = 0$ and the other π_j where $j \neq i$ is a constant as in equation (3.39). Hence, $\left(\frac{d\pi_{a_1}^{(1)}}{dx_{a_1}^{(1)}} - \frac{\mathrm{d}\pi_{a_2}^{(1)}}{\mathrm{d}x_{a_1}^{(1)}} \right) < 0$. And from (3.41), $\frac{\mathrm{d}f}{\mathrm{d}x_{a_1}^{(1)}} < 0$ for any non-zero $x_{a_i}^{(1)}$. This proves that $\frac{\mathrm{d}f}{\mathrm{d}x_{a_1}^{(1)}}$ is strictly negative for any non-zero value of $x_{a_1}^{(1)}$ and $x_{a_2}^{(1)}$. Hence $\dot{x}_{a_1}^{(1)}$ evaluated at any interior equilibrium point is negative. This completes the proof. □

Note that the stability analysis for more than two transmission configurations can be done following a similar procedure. However, this would be computationally more involved.

3.6.3 Stability of Delayed Replicator Dynamics

It is worth noting that the stability of the original replicator dynamics does not always hold for its delayed counterpart due to the impact of delay. Specifically, the evolutionary equilibrium is stable for small delays. However, when the delay is larger than a certain bifurcation point, the evolutionary equilibrium will not be stable under the delayed strategy adaptation. Analytical quantification of this bifurcation point is done in [51] for systems with linear delayed replicator dynamics by analyzing the characteristic equations of the system. However, a similar technique cannot be used to analyze this system due to the high non-linearity of the delayed replicator dynamics. To analyze the stability of delayed differential equations, there are several methods used in the literature such as Lyapunov method [61] and those using fuzzy models [62]. In the first method, characterization of the stability is done in terms of Lyapunov function. The second approach is based on modeling the nonlinear delay system as TakagiSugeno (TS) fuzzy sub-models with time delay. However, the numerical evaluation of the bifurcation point is always possible with extensive simulations. In Section 3.8, I will show the effect of delay on the convergence of the

proposed EGT-based resource allocation algorithm numerically.

3.7 EGT-Based Resource Allocation Algorithm

3.7.1 Algorithm

The evolutionary equilibrium is achieved through strategy adaptation. To this end, all players initially play a randomly selected strategy. At each iteration, each player compares her own payoff with the average payoff of the population and selects a different strategy if the current payoff is less than the population average.

Based on the replicator dynamics of the EGT, I develop a resource allocation scheme to attain the evolutionary equilibrium. The following steps summarize the algorithm.

- *Step 1* (*Initialization*): The SBSs choose a transmission alignment randomly and set $i = 1$.
- **loop**
  - *Step 2* (*Exploitation*): Each SBS transmits on the selected transmission configuration and observes the received utility. The utility and the transmission alignment information are then sent to a central controller.
  - *Step 3* (*Learning*): The central controller calculates the average payoff of the population and the population state and broadcasts it to all SBSs.
  - *Step 4* (*Update*): Each SBS compares its own payoff with the average payoff of the population. If the payoff is less than the average, the SBS randomly selects another subcarrier and power level for transmission.
  - Set $i = i + 1$.

- if $i \geq \text{Max}_\text{i}$, **end loop**; *otherwise* **go to** *Step 2*.

Max_i is the maximum number of iterations that the algorithm can execute which depends on the length of the transmission interval (e.g., a time slot) and the processing power of the SBS.

Note that the strategy adaptation process in the proposed EGT-based algorithm does not rely on the knowledge of the strategy selection of the other players. To calculate the payoff, a player who played action a needs to know x_a and for the evolution it requires the average payoff of the entire cluster of small cells. Therefore, the amount of information exchange is reduced. The central controller (e.g., a femto gateway in a two-tier macrocell-femtocell network) is only expected to calculated the average payoff of the system while the decisions on subcarrier and power level selection are taken distributively at each SBS. The proposed algorithm provides an identical payoff to each SBS at the equilibrium point. However, note that both the games $\mathcal{G}^1$ and $\mathcal{G}^2$ can have multiple interior equilibrium points. Solution refinement techniques are not included in the EGT algorithm as it requires the algorithm to run several times and hence increases the complexity and also the convergence time.

3.7.2 Comparison with Centralized Resource Allocation

For comparison purpose, I consider a centralized resource allocation in the same setting. The performance of the proposed EGT-based algorithm is compared with a centralized resource allocation algorithm based on optimization. Under the same system model, the optimization problem formulation to maximize the sum payoff (equivalent to $\mathcal{G}^1$) of the small cell network is given as follows:

$$\text{Maximize} \quad \sum_{k=1}^{K}\sum_{n=1}^{N}\sum_{l=1}^{L} x_{k,n,l}\frac{w_1^{(1)} g_{k,k}^{(n)} p_l}{\sum_{m\in\mathcal{I}_m^{(}n)} g_{m,k}^{(n)} p_m + \sum_{s=1,s\neq k}^{K}\sum_{t=1}^{L} g_{s,k}^{(n)} p_t} - w_2^{(1)}\left(\bar{g}_{h,k}^{(n)}\left(\sum_{l=1}^{L}\left(\sum_{k=1}^{K} x_{k,n,l}\right) p_l\right) - T^{(n)}\right) \tag{3.44}$$

subject to

$$\sum_{n=1}^{N}\sum_{l=1}^{L} x_{k,n,l} = 1, \forall k, \tag{3.45}$$

$$x_{k,n,l} \in \{0,1\}, \forall n, \forall k, \forall l, \tag{3.46}$$

where $x_{k,n,l} = 1$ if SBS k is transmitting on subcarrier n with power level l. The objective function is given by (3.44) which is the same as the payoff function of game $\mathcal{G}^1$. The constraint in (3.45) guarantees that each small cell base station selects no more than one subcarrier and one power level. The solution for the above optimization problem gives an upper bound for the sum payoff of the small cell network. An equivalent optimization problem for the game $\mathcal{G}^2$ can be obtained by replacing (3.44) with (3.10). The performance of the EGT-based algorithm will be compared with this upper bound.

3.8 Numerical Results and Discussion

3.8.1 Simulation Parameters for EGT based Resource Allocation Scheme

This section presents simulation results to validate the theoretical findings and to evaluate the performance of the proposed algorithm. Some of the important simula-

Table 3.2: Chapter 3: Simulation parameters

Parameter	**Value**
λ_m	4 base stations/km^2
Small cell cluster radius	200 m
Small cell radius, r_s	10 m
MBS Tx power, p_m	45 dBm
SBS Tx power, p_s	5, 7, 10 dBm
Path-loss exponent, α	4
$w_1^{(1)}, w_2^{(1)}, w_1^{(2)}, w_2^{(2)}$	1

tion parameters are given in Table II. Simulations are averaged over 3000 iterations.

I also select the interference threshold values (T^n) to assure an SINR of 5dB at the macro receiver. Recall that the values of the weighting factors in (3.9) and (3.10) have to be decided considering the network parameters. I convert SINR, aggregate interference, and interference threshold into decibels before calculating the payoff in order to bring all the terms into the same range. The weighting factors are then set to 1 in order to give equal significance to both components in the payoff function.

3.8.2 Validation of the Theoretical Results

I first validate the stochastic geometry-based analysis for the average signal-to-interference ratio (SIR) of a user of a generic small cell, as given in equation (3.29). For this purpose, I consider a system with two subcarriers. The user association to the MBSs is assumed to be based on the maximum average received power. The macrocell boundaries are defined by the Voronoi tessellation [63]. For simulations, only the interference caused by the small cells inside the cluster is taken into account. The simulation environment is as shown in Fig. 3.1.

I compare the actual values of SIR obtained by simulations (averaged over 3000

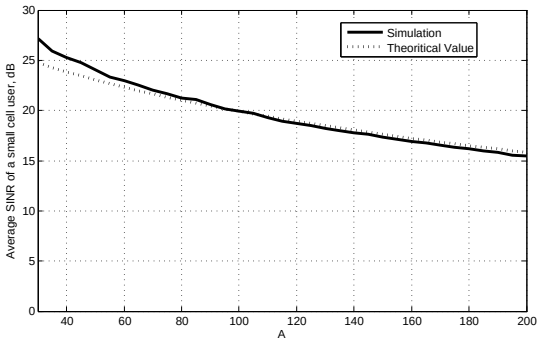


Figure 3.2: Effect of SBS intensity $\lambda_s = A\lambda_m$ on SIR (for $\lambda_m = 4$ base stations/km^2, $x_1 = 0.4$, $x_2 = 0.6$).

realizations) with the theoretical value in (3.29). The values obtained by simulations are quite close to the theoretical values (which are based on an infinite network) specially for the higher values of A, i.e., for a dense small cell network. This is due to the fact that the PPP assumption to model the spatial distribution of interfering SBSs inside the cluster holds only for a dense network. For a sparse network, the effects of the small cells outside the cluster can be significant. Therefore, the average SINR obtained from simulation is smaller than that obtained from analysis.

3.8.3 Convergence of the Proposed EGT Algorithm

To analyze the convergence of the algorithm, I first consider a system with two subcarriers and all SBSs transmit using the same power level of 10 dBm. Fig. 3.3 plots the payoff received by an SBS selecting each transmission configuration, i.e., each subcarrier in this case. It can be observed that the system converges to the equilibrium after several iterations. The figure also indicates that at the equilibrium all SBSs achieve the same payoff which shows the fairness of the proposed resource allocation scheme.

For a network with 5 subcarriers and 2 power levels, i.e., $N = 5$, $L = 2$ and the set

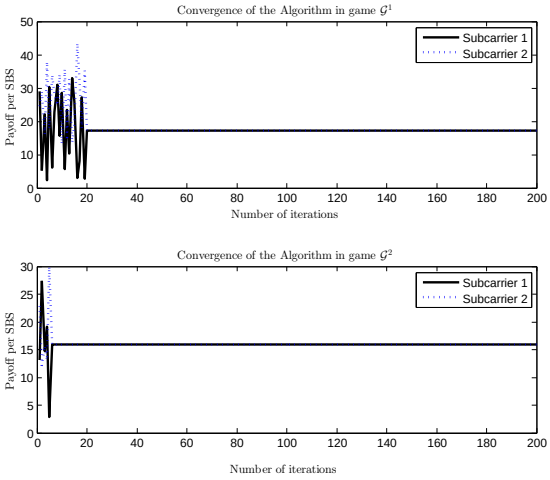


Figure 3.3: Convergence of the algorithm to the evolutionary equilibrium (for $N = 2$, $L = 1$, $\lambda_s = 60\lambda_m$).

of possible power levels given by $\mathcal{L} = \{10 \text{ dBm}, 7 \text{ dBm}\}$, Fig. 3.4 shows the variation of the payoff of an SBS with the number of iterations (obtained by simulations). It can be seen that the algorithm converges within several iterations even for a higher number of transmission configurations.

It can be observed from Figs. 3.3-3.4 that game $\mathcal{G}^2$ converges faster than $\mathcal{G}^1$. In fact, the number of iterations required for convergence depends also on the initial state of the population as well as the intermediate states it follows during the iterations until convergence. Hence, by observing one instance, it is impossible to decide under which game (i.e., $\mathcal{G}^1$ or $\mathcal{G}^2$) the algorithm is more efficient. Therefore, in Fig. 3.5, I plot the Cumulative Distribution Function (CDF) of the required number of iterations for convergence of both games. It is evident from the figure that $\mathcal{G}^2$ converges faster than $\mathcal{G}^1$.

In order to analyze the effect of the number of transmission configurations on the convergence of the algorithm, I plot the CDF of the required number of iterations for convergence under different transmission alignments for both $\mathcal{G}^1$ and $\mathcal{G}^2$. Figs.

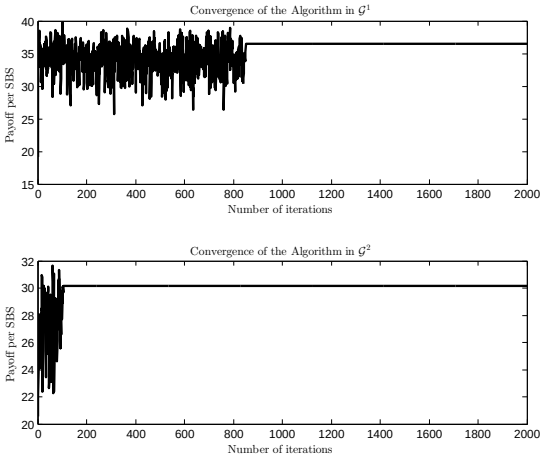


Figure 3.4: Convergence of the algorithm to the evolutionary equilibrium (for $N = 5$, $L = 2$, $\lambda_s = 60\lambda_m$).

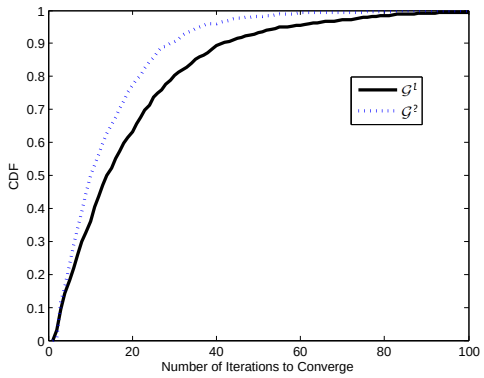


Figure 3.5: CDF of the number of iterations to converge (for $N = 2$, $L = 1$, $\lambda_s = 60\lambda_m$).

3.6-3.7 show the results for $\mathcal{G}^1$ and $\mathcal{G}^2$, respectively, for systems with two subcarriers and five subcarriers. The number of SBSs is 30 for both the cases. It is evident from the figures that the larger the number of transmission configurations, the higher is the number of iterations it takes to converge.

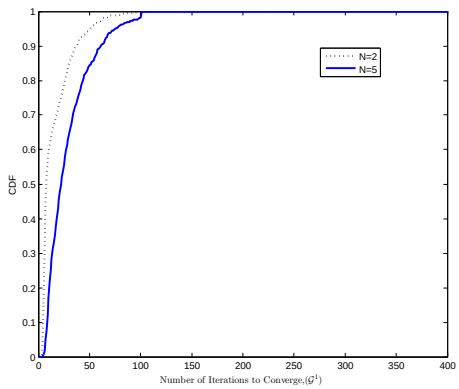


Figure 3.6: CDF of the number of iterations to converge for $\mathcal{G}^1$.

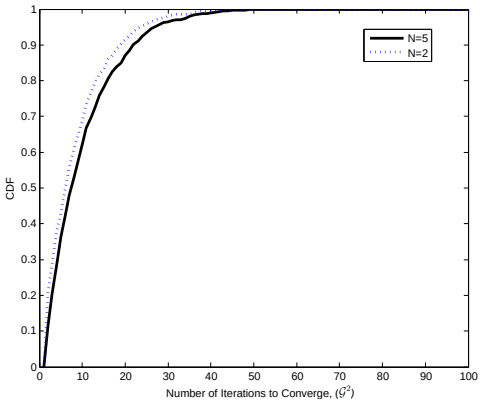


Figure 3.7: CDF of the number of iterations to converge for $\mathcal{G}^2$.

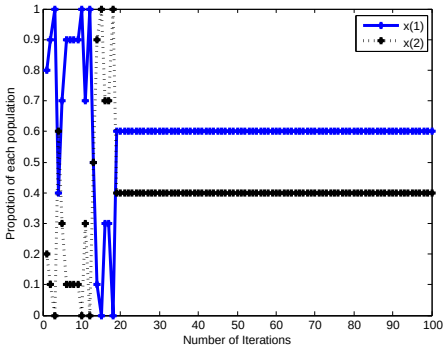


Figure 3.8: Trajectories of proportions of population (for $N = 2$, $L = 1$, $\lambda_s = 40\lambda_m$).

3.8.4 Evolution of the Population

The trajectories of the proportion of SBSs selecting each strategy for $\mathcal{G}^1$ are illustrated in Fig. 3.8. It can be seen that the proportions eventually converge to the equilibrium. However, I do not obtain smooth trajectories due to the randomness of the strategy adaptation process (see ***step 5*** of the EGT-based algorithm given in Section 3.7).

3.8.5 Comparative Performance Evaluation

In Figs. 3.9-3.10, I compare the results with the upper bound obtained by solving the centralized optimization problem in Section 3.7. It can be seen from these figures

that the gap between the maximum payoff and the payoff obtained by the EGT-based algorithm increases in both games with the number of base stations. Due to the full corporation among base stations in the centralized scheme, it results in better payoffs than that of the distributed EGT algorithm. However, the main advantage of using EGT algorithm in this regard is its simplicity and less information exchange. Specifically, EGT-based algorithm has a linear time complexity, i.e., $\mathcal{O}(Max_i)$. On the other hand, the complexity of solving the optimization problem grows exponentially with the number of SBSs in the network. The optimization problem is a binary integer programming problem. I solve the problem by an exhaustive search. In that case, there are $(N \times \mathfrak{L})^K$ number of distinct patterns of transmission configuration selection by SBSs. In addition, to solve the optimization problem, the central controller should have information about the channel gains between all the users and their interfering and serving base stations.

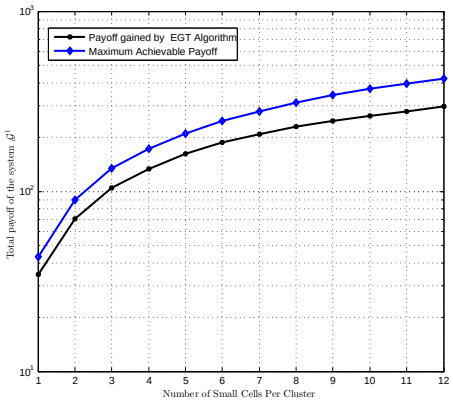


Figure 3.9: Performance of the EGT algorithm with varying number of SBSs (for game $\mathcal{G}^1$ with $N = 5$, $L = 1$).

I also investigate the effect of the weighting factors on the performance of the macro network. Fig. 3.11 shows the variation of the average interference caused by the small cell cluster at a macro user with $w_2^{(1)}$. The simulation results are obtained

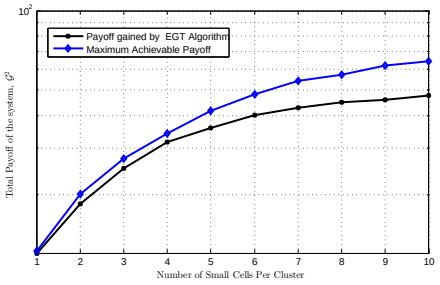


Figure 3.10: Performance of the EGT algorithm with varying number of SBSs (for game $\mathcal{G}^2$ with $N = 5$, $L = 1$).

for a network with $N = 5$, $L = 1$, and $\lambda_s = 40\lambda_m$. For both the games, the interference caused to the macro network decreases with increasing $w_2^{(i)}$. It is also interesting to note that $\mathcal{G}^1$ provides more protection to the macro user than $\mathcal{G}^2$ even though $\mathcal{G}^2$ has better convergence performance.

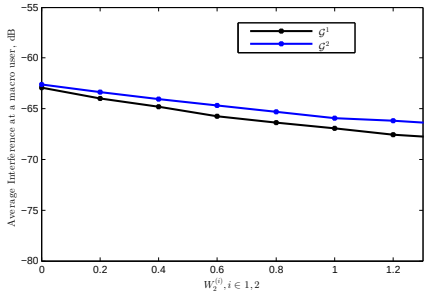


Figure 3.11: Performance of the EGT algorithm with varying number of SBSs (for game $\mathcal{G}^2$ with $N = 5$, $L = 1$).

3.8.6 Impact of Information Exchange Delay on Convergence

I also investigate the impact of information delay on the convergence of the replicator dynamics. The delay τ indicates that at iteration n, all base stations have the information corresponding to iteration $n - \tau$ (see equation (3.16)). The information at $n = 0$ is used to compute the values corresponding to $n < 0$. Delay is assumed

to be constant throughout all the iterations. As it can be seen from Figs. 3.12-3.13, the larger the delay, the larger is the number of iterations it takes to converge and the system becomes less stable. Under a small delay, the system can still converge to the equilibrium. However, when the delay is larger than a certain bifurcation point, the system will diverge. Another important observation is that even-though game $\mathcal{G}^1$ converges to the same equilibrium point with $\tau = 2$, $\mathcal{G}^2$ converges to different equilibrium points. Hence, with delayed information exchange, for system with multiple equilibrium points, there is no guarantee that the system will converge to the same equilibrium point as that for the delay-free system.

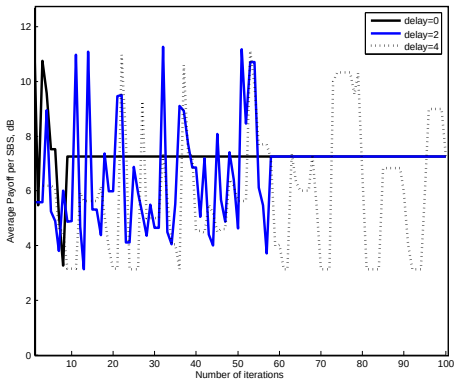


Figure 3.12: Effect of delay on the convergence of algorithm $\mathcal{G}^1$ (for $N = 3$, $L = 1$, $\lambda_s = 40\lambda_m$).

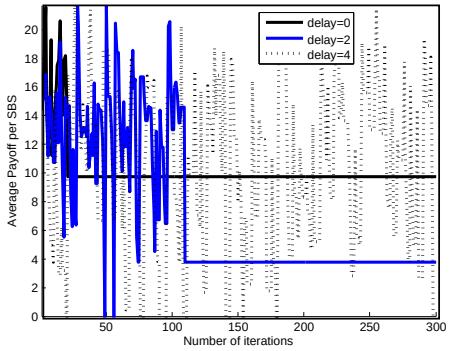


Figure 3.13: Effect of delay on the convergence of algorithm $\mathcal{G}^2$ (for $N = 3$, $L = 1$, $\lambda_s = 40\lambda_m$).

In order to further characterize the effect of delay, I plot the CDF of the number of iterations required for the convergence of the algorithm in Figs. 3.14-3.15 for a network with two subcarriers and one power level. It can be observed that the higher the delay, the higher is the number of iterations required for convergence in both the games. As the resource allocation algorithm is expected to converge within a certain period of time, the delayed information can lead to a very poor network performance.

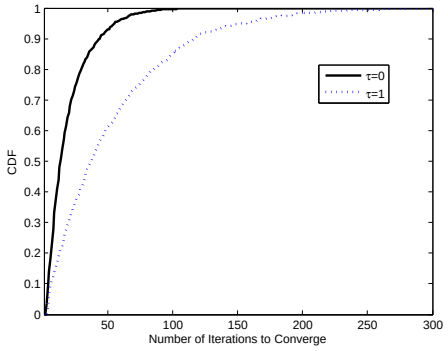


Figure 3.14: CDF of the required number of iterations for convergence of $\mathcal{G}^1$.

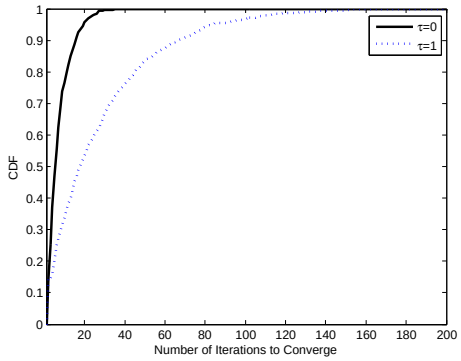


Figure 3.15: CDF of the required number of iterations for convergence of $\mathcal{G}^2$.

3.9 Chapter Summary

In this chapter, I have presented an evolutionary game theory-based algorithm to solve the problem of distributed resource allocation in self-organizing small cell networks underlaying macrocell networks. The available set of subcarrier-power level combinations at each small cell base station has been defined as the strategy set and the replicator dynamics is used to model the strategy adaptation process of the small cell base stations. The evolutionary equilibrium, which is the fixed point of the replicator dynamics, has been considered to be the solution of the formulated evolutionary game. Based on the average achievable SINR of a small cell user derived based on stochastic geometry analysis, the stability of the equilibrium point has been analytically proven for a system with two subcarriers. For larger system configurations, the stability of the equilibrium point has been shown by simulations. I have also considered the impact of delayed information exchange on the convergence of the proposed algorithm. Simulation results have been presented to validate the analytical developments and illustrate the performance of the proposed algorithm when compared to a centralized resource allocation algorithm. Also, the effect of information delay on the

equilibrium has been investigated by simulations.

The resource allocation scheme presented in this chapter ensures fairness among the small cells however; it assumes players are bounded rational and consider instantaneous payoff when taking decisions. Despite that, some network operators may interested in resource allocation considering maximizing the payoff or minimizing the cost (i.e., network nodes are rational) for a certain period of time. Moreover, small cell base stations may be battery operated, hence with energy limitations. Considering these aspects, in next chapter, I present a distributed energy aware power control scheme for ultra dense small cell networks which consider maximizing the payoff over a pre-defined period of time.

Chapter 4

A Mean Field Game for Energy-Aware Power Control in Ultra-Dense Small Cell Networks

In this chapter, a novel energy-aware distributed power control paradigm is proposed for dense small cell networks co-existing with a traditional macrocellular network. The power control problem is first modelled as a differential game and the existence of the Nash Equilibrium is proven. Then I extend the formulated stochastic game to a mean field game (MFG) considering a highly dense network. An MFG is a special type of differential game which is ideal for modeling the interactions among a large number of entities. I also analyze the performance of two different cost functions for the mean field game formulation. Both of these cost functions are designed using stochastic geometry analysis in such a way that the cost functions are valid for the MFG setting. A finite difference algorithm is then developed based on the Lax-Friedrichs scheme and Lagrange relaxation to solve the corresponding MFG. Each small cell base station can independently execute the proposed algorithm offline, i.e.,

prior to data transmission. The output of the algorithm shows how each small cell base station should adjust its transmit power in order to minimize the cost over a predefined period of time. Moreover, sufficient conditions for the uniqueness of the mean field equilibrium for a generic cost function are also given. The effectiveness of the proposed algorithm is demonstrated via numerical results.

4.1 Introduction

4.1.1 Overview

Since classical games have to model the interaction of each player with every other player, analysis of a system with a large number of players can be complex. Therefore, when it comes to a dense network of interconnected base stations, solving the power control problem based on classical game theory becomes very hard and sometimes impossible due to the large number of players. In this context, the theory of mean field game (MFG) [64–66], which has been used for solving a variety of problems in different research areas [67–71], can be used.

MFGs can be considered as a special form of differential games applicable for a system with a large number of players. While classical game theory models the interaction of a single player with all the other players of the system, an MFG models the individual's interaction with the effect of the collective behavior (mass) of the players. This collective behavior is reflected in the mean field. Individual player's interaction with the mean field is modeled by a Hamilton-Jacobi-Bellman (HJB) equation. The motion of the mass according to the players' actions is modelled by a Fokker-Planck-Kolmogorov (FPK) equation [64]. These coupled FPK and HJB equations are also called backward and forward equations, respectively. The solution

of an MFG can be obtained by solving these two equations. When modelled as an MFG, since the system can be completely defined by two equations (which are also called the mean field equations), analysis of the system becomes much easier. Moreover, solutions to the MFGs can be obtained distributively and behaviors of all the players can be described by one control. In addition, MFGs can take the stochastic nature of the system into account. All of the aforementioned properties make MFG appropriate for modeling the power control problem for dense self-organizing small cell networks. However, modeling the collective effect of the players (i.e., the effect of mass/mean field) has to be done in a realistic way. Accurate modeling of the effect of the mass is a major challenge when adopting MFGs to solve problems in wireless communications.

In this chapter, I formulate the downlink power control problem of a dense small cell network underlaying a macrocellular network as an MFG. The small cell base stations (SBSs), when battery-operated, are assumed to be constrained by a finite energy. To model the mass (or mean field), I adopt a stochastic geometry approach. Specifically, I consider minimizing a cost function under certain constraints over a pre-defined period of time. The cost function is derived by using a stochastic geometry approach in such a way that it reflects the signal-to-interference-plus-noise ratio (SINR) at the receivers and the interference caused to the macro cellular network. I propose a finite difference technique to solve the mean field equations for the formulated MFG. The key feature of the proposed algorithm is that it can be executed offline. By executing the algorithm, each base station can obtain a power policy which depends on the initial energy distribution among the SBSs. The SBSs can then use that power control policy for data transmission for a pre-defined period of time. Another advantage of the algorithm is, it minimizes the cost over a certain

period of time instead of taking decisions only based on the instantaneous cost.

4.1.2 Contribution

The contributions of this chapter can be summarized as follows.

1. The downlink power control problem of a small cell network underlaying a traditional macro network (i.e., for a system model consisting of multiple transmitters and multiple receivers) is formulated as a differential game and extended to a mean field game for a dense scenario.

2. The existence of a Nash Equilibrium for the formulated differential game is proven.

3. Using stochastic geometry-based analysis, two cost functions for the mean field game are derived in such a way that the mean field game setting becomes valid. In this way, it combines the theory of MFG with that of stochastic geometry.

4. The forward and backward equations of the mean field game are solved based on the finite difference technique proposed in [72].

5. An algorithm is proposed to obtain the mean field equilibrium for the formulated game.

6. The sufficient conditions are given for the uniqueness of the mean field equilibrium for a generic cost function.

4.2 Related Work

Implementing most of the existing game theory based distributed resource allocation algorithms in the literature for a ultra-dense network would require an extensive

amount of information exchange among the base stations. Recently, mean field game has gained the attention of the research community as a tool to model dense heterogeneous networks (or small cell networks). In [70, 73–75], the power control problem is modelled as mean field games for scenarios where multiple transmitters transmit to a single receiver (e.g., uplink transmissions in a cellular network). The problem is first formulated as a stochastic differential game and then its convergence to a mean field game is shown for a very large number of transmitters. In [70], the authors show the power control policy obtained at the mean field equilibrium. [73] and [74] present the sufficient conditions for the uniqueness of the respective games formulated in these papers. The work presented in [68] formulates the power control problem in a cognitive radio network as a hierarchical mean field game. The mean field game formulations in all of the aforementioned papers consider scaled interference (i.e., interference at the receiver is normalized by the number of transmitters) only which may not be valid for a large-scale small cell network. In addition, none of the above papers presents any technique for solving the mean field equations, which is also very challenging.

4.3 System Model and Assumptions

4.3.1 Network and Propagation Model

I consider an infinite small cell network underlaying an infinite macrocell network. In practice, the spatial distribution of SBSs is random and independent of each other, as most of the time the SBSs are deployed opportunistically in an unplanned manner. Also, the deployment of macro base stations in practical LTE networks is closer to a random deployment than the hexagonal grid model [76]. A network with an infinite

number of nodes, which are randomly and independently distributed in a large area, can be abstracted by a Poisson Point Process (PPP) [77, 78]. Therefore, I model the spatial distributions of the SBSs and MBSs by two independent PPPs denoted by Φ_s and Φ_m, respectively. The intensities of Φ_s and Φ_m are given by λ_m and λ_s, respectively. Users are connected to the base station from which they receive the highest average pilot signal power. The transmit powers of the pilot signals of SBSs and MBSs are given by $p_{s,pilot}$ and $p_{m,pilot}$, respectively. When users are associated to the base station from which they receive the highest average pilot signal power, the cell boundaries can be shown by a weighted Voronoi tessellation [79]. The pilot powers of SBSs are lesser than those of MBSs. Although several users may be associated with a base station, I assume that each base station serves only one user at a particular time instant. For downlink transmission, scheduling of users in a small cell can be performed by using schemes such as proportional fair or round-robin scheduling schemes.

I consider the problem of downlink transmit power control at the SBSs. Co-channel deployment is considered, i.e., both the MBSs and SBSs transmit on the same channel. Power control is done in order to minimize the average cost of each SBS over a given finite time horizon T. I will define the cost function later. Each SBS k is assumed to be with a finite amount of energy, denoted by $E_{k,max}$, to spend within the given period of time, T. In this way, I can take the heterogeneity of the small cell base stations into account, i.e., different types SBSs can have different initial energy levels or energy constraints. To consider any SBS k with an infinite amount of energy in the same setting, $E_{k,max}$ can be set to a large value, i.e., $E_{k,max} >> p_{max}T$, where p_{max} is the maximum allowable transmit power for an SBS. The users served by each SBS k have a minimum SINR requirement denoted by Γ_k. In this way, this

system model captures heterogeneity of the mobile users in-terms of their QoS (i.e., different users can have different SINR requirements). [1] The channels between all the transmitters and all the receivers are assumed to experience i.i.d. Rayleigh fading.

The SINR at the user served by SBS k at time t is given by

$$\mathrm{SINR}_k(t) = \frac{p_k(t) g_{k,k}(t) r_{k,k}(t)^{-\alpha}}{I_{s,k}(t) + I_{m,k}(t) + N_0}, \tag{4.1}$$

where $I_{s,k}(t) = \sum_{l \in \mathcal{K}, l \neq k} p_l(t) g_{l,k}(t) r_{l,k}(t)^{-\alpha}$ and $I_{m,k}(t) = \sum_{\forall m \in \Phi_m} p_m g_{m,k}(t) r_{m,k}(t)^{-\alpha}$ denote the interference caused by small cell and macro cell networks, respectively. $g_{l,k}$ is the fading power gain between transmitter l and receiver k (i.e., user served by the base station k), $r_{l,k}$ is the distance between the transmitter l and the receiver k , N_0 is the noise power and α is the path-loss exponent. The following inequality should hold for any SBS to satisfy its QoS constraint:

$$\frac{p_k(t) g_{k,k}(t) r_{k,k}(t)^{-\alpha}}{I_{s,k}(t) + I_{m,k}(t) + N_0} \geq \Gamma_k, \qquad \forall\, k \in \mathcal{K}, \tag{4.2}$$

which can be expressed as follows:

$$p_k(t) g_{k,k}(t) r_{k,k}(t)^{-\alpha} - \Gamma_k \left(I_{s,k}(t) + I_{m,k}(t) + N_0 \right) \geq 0, \quad \forall k \in \mathcal{K}. \tag{4.3}$$

The major symbols that are used throughout this chapter are given in Table 4.1

4.3.2 Cost Function of an SBS

The cost function of SBS k at time t is composed of two components as follows:

[1] Later, it can be seen that both user and base station heterogeneity can be considered in the differential game formulation. However, the mean field game formulation can only take the heterogeneity of the base stations into account.

Table 4.1: Chapter 4: Symbols

Symbol	**Description**
α	Path-loss exponent
Γ_k	SINR requirement for the users of SBS k
Φ_m	PPP which represents the spatial distribution of MBSs
Φ_s	PPP which represents the spatial distribution of SBSs
λ_m	Density of Φ_m
λ_s	Density of Φ_s
$c_k(t)$	Value of the cost at SBS k at time t
$\mathcal{E}_k$	State space (i.e., possible energy levels) of SBS k
$\boldsymbol{e}_k$	State of the system at time t
$e_k(t)$	Available energy of SBS k at time t
$E_{k,max}$	Maximum available energy of SBS k
$g_{k,l}$	Fading channel gain between transmitter k and receiver l
$I^m{}_k(t)$	Interference caused to the nearest macro user by SBS k at time t
$I_{s,k}(t), I_{m,k}(t)$	Interferences caused by small cell network and macro cell network at the user served by SBS k at time t
N_0	Variance of noise power
$\mathcal{K}$	Set of SBSs
$\mathcal{P}_k$	Set of all possible transmit powers of SBS k
$p_k(t)$	Transmit power of SBS k at time t
p_m	MBS transmit power
p_{max}	Maximum transmit power for SBSs
$p_{m,pilot}$,$p_{s,pilot}$	Pilot signal power of MBSs and SBSs
$r_{k,l}$	Distance between transmitter k and receiver l
T	Time period during which power control is done
$u_k(t)$	Value function of SBS k at time t
$v(t,e)$	Lagrange multiplier at time t and energy e
w_1, w_2	Biasing factors
X, Y	Number of discretization levels in time axis and energy axis, respectively

- The cost associated with the satisfaction of the QoS constraint (Γ_k), denoted by $f_k^{(1)}(t)$, and

- The cost associated with the interference caused to the nearest macro user, denoted by $f_k^{(2)}(t)$.

Based on (4.3), $f_k^{(1)}(t)$ is defined as follows:

$$f_k^{(1)}(t) = \left(\Gamma_k \left(I_{s,k}(t) + I_{m,k}(t) + N_0\right) - p_k(t) g_{k,k}(t) r_{k,k}(t)^{-\alpha}\right)^2. \tag{4.4}$$

Minimizing f_1 will attempt to satisfy the QoS constraint, but it will also discourage further increase of transmit power after satisfying the QoS constraint. A similar cost function is also used in [80] and [81] for uplink power control. On the other hand, $f_k^{(2)}(t)$ is defined as the interference caused at the nearest macro user at time t, which is given by ${I^m}_k(t)$, as follows:

$$f_k^{(2)}(t) = {I^m}_k(t) = p_k(t) g_{k,m} r_{k,m}^{-\alpha}. \tag{4.5}$$

Accordingly, the cost function of SBS k at time t (i.e., $c_k(t)$) is defined as a linear combination of above two functions:

$$c_k(t) = w_1 f_k^{(1)}(t) + w_2 f_k^{(2)}(t), \tag{4.6}$$

where w_1 and w_2 are biasing factors which bring the above two terms into one scale. The network operator has the freedom to set these biasing factors.

Note that for the formulation of the mean field game, in Section 4.5, I will generalize the cost function for any generic SBS such that interchangeability (or permutation) of the states among the SBSs does not affect the outcome of the game.

4.3.3 State, Action Space, and Control Policy of an SBS

The *state* of SBS k at time t is defined by the amount of available energy at that time, which is given by, $e_k(t)$. Therefore, the state space $\mathcal{E}_k$ of SBS k can be written as follows:

$$\mathcal{E}_k = [0, e_k(0)] = \{e_k(t) \in \mathbb{R} | 0 \leq e_k(t) \leq e_k(0)\}, \tag{4.7}$$

where $e_k(0)$ is the available energy of SBS k at time 0.

I also define the *state of the system* at time t, $\boldsymbol{e}(t)$ as follows:

$$\boldsymbol{e}(t) = [e_k(t)_{\forall k}]^T. \tag{4.8}$$

The set of *actions* for SBS k includes all possible transmit powers as follows:

$$\mathcal{P}_k = [0, p_{max}], \tag{4.9}$$

where p_{max} is the maximum allowable transmit power of any SBS. The transmit power of SBS k at time t is denoted by $p_k(t)$.

The evolution of the state (in this case, available energy) over time is decided by a *control*, which in this case corresponds to the transmit power given by $p_k(t) \in [0, p_{max}]$. Consequently, the *state equation* of the system is defined as follows.

Definition (State equation): The state of SBS k is given by the random variable $e_k(t) \in [0, e_k(0)]$ whose evolution is defined by the following differential equation:

$$de_k(t) = -p_k(t)dt, \quad 0 \leq t \leq T. \tag{4.10}$$

The *control policy* is a mapping of the state to an action. This is defined over the given period of time, T. I denote the control policy of player k over the time period

T by $p_k\,(0 \rightarrow T)$. An optimal power control policy, $p_k^*\,(0 \rightarrow T)_{\forall k}$ should minimize the average cost of each player k over the given finite time horizon, T. Therefore, I write $p_k^*\,(0 \rightarrow T)$ as follows:

$$p_k^*\,(0 \rightarrow T) = \arg \min_{p_k(0 \rightarrow T)} \mathbf{E}\left[\int_0^T c_k(t)dt + c_k(T)\right],$$

where $c_k(T)$ is the terminal cost (i.e., cost at the end of time period T).

The objective is to obtain the optimal power control policy distributively at each SBS in order to minimize the average cost over time interval T. This can be seen as an optimal control problem [82], but with several controllers (each SBS is a controller in this case). Such a problem can be formulated as a differential game [14]. Differential games can be seen as a generalization of the optimal control problems for the cases where there are more than one controller. A mean field game is an extention to a differential game when the system has a large number of players. In the next two sections, I show the differential game formulation and its extension to a mean field game (denoted by $\mathcal{G}_s$ and $\mathcal{G}_m$, respectively). The set of SBSs $\mathcal{K} = \{1, 2, ..., K\}$ is the set of players in these game models.

4.4 Differential Game Formulation

In this section, I formulate the differential game to model the downlink transmit power control problem for the system model described above. To formulate the differential game denoted by $\mathcal{G}_s$, I define the value function $u_k(t)$ as follows:

$$u_k(t) = \min_{p_k(t \rightarrow T)} \mathbf{E}\left[\int_t^T c_k(\tau)d\tau + c_k(T)\right], \qquad t \in [0, T] \tag{4.11}$$

where $c_k(T)$ is the terminal cost.

According to Bellman's principle of optimality [83], an optimal control policy should have the property that whatever the initial state and initial decision are, the remaining decisions must form an optimal policy with regard to the state resulting from the first decision [72, 84]. Accordingly, the optimal power control policy can then be defined in-terms of the value function as follows.

Definition (Optimal control): The power profile $p_k^*(t \to T)$ is the optimal power control policy for SBS k if for any $t \in [0, T]$, $\mathbf{E}\left[\int_t^T c_k(p_k^*(\tau))d\tau + c_k(T)\right] = u_k(t), \quad t \in [0, T]$.

This value function should satisfy a partial differential equation which is in the form of a Hamilton-Jacobi-Bellman (HJB) equation [85]. The HJB equation corresponding to the optimal control problem given in equation (4.11) satisfying the state equation (4.10) can be written as follows:

$$\frac{\partial u_k(t)}{\partial t} + \min_{p_k(t)} \left(c_k\left(p_k(t)\right) - p_k(t)\frac{\partial u_k(t)}{\partial e} \right) = 0, \tag{4.12}$$

where $H\left(e_k(t), \frac{\partial u_k(t)}{\partial e}\right) = \min_{p_k(t)} \left(c_k(t) - p_k(t)\frac{\partial u_k(t)}{\partial e}\right)$ is called the Hamiltonian. Now, the Nash equilibrium of the game $\mathcal{G}_s$ is defined as follows.

Definition (Nash equilibrium of game $\mathcal{G}_s$):
A power profile

$$\boldsymbol{p}^* = [p_1^*(0 \to T), p_2^*(0 \to T), ..., p_k^*(0 \to T), ..., p_K^*(0 \to T)]$$

is a Nash equilibrium of the game $\mathcal{G}_s$ if and only if

$$p_k^*(0 \to T) = \arg \min_{p_k(0 \to T)} \mathbf{E}\left[\int_0^T c_k\left(p_k(t), \boldsymbol{p}_{-k}^*\right) dt + c_k(T)\right], \qquad \forall k \tag{4.13}$$

subject to

$$de_k(t) = -p_k(t)dt \quad (0 \leq t \leq T), \qquad \forall k \tag{4.14}$$

where $\boldsymbol{p}^*_{-k}$ denotes the transmit power vector of the SBSs except SBS k.

When the above condition is satisfied, none of the players can have a lesser cost by deviating unilaterally from the current power control policy. Hence, it is equivalent to the Nash equilibrium of game $\mathcal{G}_s$. The Nash equilibrium of the above differential game can be obtained by solving the HJB equation associated with each player given in equation (4.12) [86]. I state following theorem on the existence of the Nash equilibrium for $\mathcal{G}_s$.

Theorem 4.4.1. *There exists at least one Nash equilibrium for the differential game $\mathcal{G}_s$.*

Proof. Existence of a solution to the HJB equation in (4.12) ensures the existence of the Nash equilibrium for the game $\mathcal{G}_s$. It is known that there exists a solution to the HJB equation if the Hamiltonian is smooth [73, 87]. The Hamiltonian for equation (4.12) can be written as in (4.15).

$$\begin{aligned} & H\left(e_k(t), \frac{\partial u_k(t)}{\partial e}\right) = \min_{p_k(t)} \left(c_k(t) - p_k(t)\frac{\partial u_k(t)}{\partial e}\right) \\ = & \min_{p_k(t)} \Big[w_1 \left(\Gamma_k \left(I_{s,k}(t) + I_{m,k}(t) + N_0\right) - p_k(t) g_{k,k}(t) r_{k,k}(t)^{-\alpha}\right)^2 \\ + & \; w_2 \left(p_k(t) g_{k,m} r_{k,m}^{-\alpha}\right) - p_k(t)\frac{\partial u_k(t)}{\partial e} \Big]. \end{aligned} \tag{4.15}$$

The first, second, and third derivatives of the Hamiltonian w.r.t. $p_k(t)$ can be

written as follows:

$$\frac{\partial H}{\partial p_k(t)} = -2w_1 g_{k,k} r_{k,k}^{-\alpha} \left(\Gamma_k \left(I_{s,k}(t) + I_{m,k}(t) + N_0\right)\right.$$
$$\left. - p_k(t) g_{k,k}(t) r_{k,k}(t)^{-\alpha}\right) + w_2 g_{k,m} r_{k,m}^{-\alpha} - \frac{\partial u_k}{\partial e}, \tag{4.16}$$

$$\frac{\partial^2 H}{\partial p_k(t)^2} = 2w_1 \left(g_{k,k} r_{k,k}^{-\alpha}\right)^2, \tag{4.17}$$

$$\frac{\partial^3 H}{\partial p_k(t)^3} = 0. \tag{4.18}$$

For any $n > 3$, $\frac{\partial^n H}{\partial p_k(t)^n} = 0$. The function has derivatives of all orders, hence it is smooth. Therefore, it can be concluded that there exists at least one Nash equilibrium for the differential game $\mathcal{G}_s$.

□

Obtaining the equilibrium for game $\mathcal{G}_s$ for a system with K players involves solving K simultaneous partial differential equations (PDEs). However, for a dense small cell network, obtaining the Nash equilibrium by solving $\mathcal{G}_s$ would be difficult (if not impossible) due to the large number of simultaneous PDEs. Therefore, for modeling and analysis of a dense small cell network, I propose a mean field game formulation where the system can be defined solely by two coupled equations. In the next section, I show the extension of game $\mathcal{G}_s$ to the mean field game $\mathcal{G}_m$.

4.5 Formulation of Mean Field Game

4.5.1 Assumptions

First, the mean field is defined as follows:

Definition (Mean field)

$$m(e,t) = \lim_{K\to\infty} \frac{1}{K} \sum_{\forall k \in \mathcal{K}} \mathbb{1}_{\{e_k(t)=e\}}, \tag{4.19}$$

where $\mathbb{1}$ denotes an indicator function which returns 1 if the given condition is true and zero otherwise.

For a given time instant, mean field is the probability distribution of the states over the set of players.

The general setting of mean field games is based on the following four assumptions [88]:

1. Rationality of the players,
2. The existence of a continuum of the players (i.e., continuity of the mean field),
3. Interchangeability of the states among the players (i.e., permutation of the states among the players would not affect the outcome of the game), and
4. Interaction of the players with the mean field.

The first assumption is generally applied in any type of game to ensure that the players can take logical decisions. The presence of a large number of SBSs in the system model ensures the existence of the continuum of the players. I derive the cost function (which will be shown in next subsection) in order to ensure the interchangeability of

the actions among the players. The idea of the fourth assumption is that each player interacts with the mean field instead of interacting with all the other players.

4.5.2 Deriving the Cost Function

A cost function, which depends only on control (and/or state) and mean field, would ensure that the third assumption of the mean field game setting is valid. To derive such a cost function for $\mathcal{G}_m$, I follow a stochastic geometry-based approach. In this case, for simplicity, I assume an interference-limited network setting (i.e., $N_0 = 0$). This assumption can be justified due the fact that the network is highly dense. It is also assumed that all SBSs have the same QoS constraint given by γ.

By taking the spatial averages over the point process, I generalize $f_k^{(1)}(t)$ and $f_k^{(2)}(t)$ for any generic player which transmits with power $p(t)$ at time t as follows. I denote the new functions by $f^{(1,mean)}(t)$ and $f^{(2,mean)}(t)$. The function $f^{(1,mean)}(t)$ is given by

$$\begin{aligned} f^{(1,mean)}(t) &= \Bigg(- p(t)\mathbf{E}\left[g_{k,k}(t)\right]\mathbf{E}\left[r_{k,k}(t)\right]^{-\alpha}\Gamma\mathbf{E}_{\mathcal{I}_s,g_{l,k}(t),p_l(t)}\left[\sum_{l\in\mathcal{I}_s} p_l(t)g_{l,k}(t)r_{l,k}(t)^{-\alpha}\right] \\ &+ \Gamma\mathbf{E}_{\mathcal{I}_m,g_{m,k}(t),p_m(t)}\left[\sum_{m\in\mathcal{I}_m} p_m(t)g_{m,k}(t)r_{m,k}(t)^{-\alpha}\right]\Bigg)^2, \end{aligned} \tag{4.20}$$

where $\mathcal{I}_s$ and $\mathcal{I}_m$ are the sets of interfering SBSs and MBSs, respectively. The function $f^{(2,mean)}(t)$ for any generic SBS is given by

$$f^{(2,mean)}(t) = p(t)\mathbf{E}_{\phi_s}\left[g_{k,m}\right]\mathbf{E}_{\phi_m}\left[r_{k,m}\right]^{-\alpha}. \tag{4.21}$$

Then the cost function for a generic SBS can be written as follows:

$$c(t) = w_1 f^{1,mean}(t) + w_2 f^{2,mean}(t). \tag{4.22}$$

Derivation of $f^{(1,mean)}(t)$

Derivation of $\mathbf{E}\left[I_s(t)\right]$ and $\mathbf{E}\left[I_m(t)\right]$:

I derive $\mathbf{E}\left[I_s(t)\right] = \mathbf{E}_{\mathcal{I}_s}\left[\sum_{l\in\mathcal{I}_s} p_l(t) g_{l,k}(t) r_{l,k}(t)^{-\alpha}\right]$ for a generic SBS k at the origin. According to Slivnyak's theorem [89], the statistics for a PPP is independent of the test location. Therefore, the analysis holds for any small cell user at a generic location. Since the channel gains and the transmit powers of the interferes are independent of the point process Φ_s,

$$\mathbf{E}\left[I_s(t)\right] = \mathbf{E}[p_k(t)]\mathbf{E}[h_{k,k}(t)]\mathbf{E}_{\Phi_s}\left[\sum_{l\in\mathcal{I}_s} r_{l,k}(t)^{-\alpha}\right]. \tag{4.23}$$

For Rayleigh fading, assuming $h_{l,k} \sim \exp(1)$ for $\forall k, l \in \Phi_s$, by using Campbell's theorem [90], we have the following:

$$\mathbf{E}\left[I_s(t)\right] = \mathbf{E}\left[p_k(t,t)\right] \int_{\mathcal{R}^2} r_{l,k}(t)^{-\alpha} d(\mathcal{R}). \tag{4.24}$$

Since the received power cannot be larger than transmit power, the path-loss is assumed to be 1 when $r_{l,k}(t) < 1$. Then, the average interference at a generic user at the origin can be derived as follows:

$$\begin{aligned} \mathbf{E}\left[I_s(t)\right] &= \mathbf{E}[p_k(t)] 2\pi\lambda_s \left[\int_0^1 r dr + \int_1^\infty r^{-\alpha} r dr\right], \\ &= 2\pi\lambda_s \mathbf{E}[p_k(t)] \left(\frac{1}{2} + \frac{1}{\alpha - 2}\right). \end{aligned} \tag{4.25}$$

By following similar steps, the average interference caused from the macro network can also be derived as follows:

$$\mathbf{E}\left[I_m(t)\right] = 2\pi\lambda_m p_m \left(\frac{1}{2} + \frac{1}{\alpha - 2}\right). \tag{4.26}$$

Derivation of $\mathbf{E}\left[r_{k,k}(t)\right]$:

Each small cell user is assumed be connected to the nearest SBS. It is also known that the distance to the nearest base station from any generic point is Rayleigh distributed [91]. Therefore, the probability density function (PDF) of $r_{k,k}(t)$ can be written as: $f_{r_{k,k}}(r) = 2\pi\lambda_s r e^{-\pi\lambda_s r^2} dr$. Therefore, the average distance is the given by:

$$\mathbf{E}\left[r_{k,k}(t)\right] = \frac{1}{2\sqrt{\lambda_s}}. \tag{4.27}$$

Derivation of $f^{(2,mean)}$

In order to determine $f^{(2,mean)}$, I need to determine PDF of the distance to the nearest possible macro user (i.e., $r_{k,m}$) from any generic small cell user. The nearest macro user can be just beyond the edge of coverage area of the small cell. In practice, cell edges can be created both due to MBSs and SBSs. However, due to the limited transmit power of the small cell base stations, the coverage area of small cells may not overlap. For analytical tractability, in this case I assume that the small cells do not overlap with each other. Therefore, the edges of the small cells are formed only due to the MBSs.

PDF of $r_{k,m}$ can be derived as follows. Considering the cell edge between an SBS and an MBS, we can write,

$$p_{s,pilot} R^{-\alpha} = p_{m,pilot} (X - R)^{-\alpha}, \tag{4.28}$$

where R is the distance from SBS to the closest cell edge and X is the distance to the nearest MBS. Since the distribution of MBSs follows a PPP with intensity λ_m, the cumulative distribution function (CDF) $F_{R_{k,m}}(r_{k,m})$ and PDF $f_{R_{k,m}}(r_{k,m})$ of $r_{k,m}$ are given by

$$\begin{aligned} F_{R_{k,m}}(r_{k,m}) &= 1 - e^{-\lambda_m \pi b^2 r^2}, \\ f_{R_{k,m}}(r_{k,m}) &= 2\pi\lambda_m r b^2 e^{-\lambda_m \pi b^2 r^2}, \end{aligned} \tag{4.29}$$

where $b = \left[1 + \frac{p_{m,pilot}}{p_{s,pilot}}^{\frac{1}{\alpha}}\right]$.

The above equations imply that $r_{k,m}$ is Rayleigh distributed and the expected value is given by

$$\mathbf{E}_{m\in\phi_m}\left[r_{k,m}\right] = \frac{1}{2\sqrt{\lambda_m}\left[1 + \left(\frac{p_{m,pilot}}{p_{s,pilot}}\right)^{\frac{1}{\alpha}}\right]}. \tag{4.30}$$

I assume that the path-loss exponent α is equal to 4. By substituting the values from equations (4.25), (4.26), (4.27), and (4.30) for $f^{(1,mean)}$ and $f^{(2,mean)}$ in expression (4.22), the cost function of a generic SBS transmitting with power $p(t)$ at time t can be written as follows:

$$c(t) = w_1\left(-16\lambda_s^2 p(t) + 2\pi\Gamma\left[p_m\lambda_m + \mathbf{E}\left[p_l(t)\right]\lambda_s\right]\right)^2 + w_2 16 p(t)\lambda_m^2\left[1 + \left(\frac{p_{m,pilot}}{p_{s,pilot}}\right)^{\frac{1}{4}}\right]^4 \tag{4.31}$$

(4.32)

For a generic cost function, the control (i.e., transmit power) of time t would only depend on the state of the SBS. Hence, the expectation of the transmit power over all interfering SBSs can be written in terms of the mean field. Then, the above equation

can be re-written as follows:

$$\begin{aligned} c(t,e) &= w_1 \left(-16\lambda_s^2 p(t,e) + 2\pi\Gamma \left[p_m \lambda_m + \lambda_s \int_{e\in\mathcal{E}} p(t,e)m(t,e)de \right] \right)^2 \\ &+ w_2 16 p(t,e)\lambda_m^2 \left[1 + \left(\frac{p_{m,pilot}}{p_{s,pilot}} \right)^{\frac{1}{4}} \right]^4 . \end{aligned} \tag{4.33}$$

For comparison purpose, I also introduce another cost function[2] denoted by $\hat{c}(t,e)$, which is similar to that in Chapter 3, as follows:

$$\begin{aligned} \hat{c}(t,e) &= -\hat{w}_1 \mathbf{E}_{\phi_s,\phi_m} \left[\mathrm{SINR}_k(p(t,:),m(t,:)) \right] \\ &\quad + \hat{w}_2 p_k(t) \mathbf{E}_{\phi_s} \left[g_{k,m} \right] \mathbf{E}_{\phi_m} \left[r_{k,m} \right]^{-\alpha} , \end{aligned} \tag{4.34}$$

where $\hat{w}_1$ and $\hat{w}_2$ are weighting factors.

This cost function does not take the QoS constraint into account. The SBSs can increase their transmit powers even after satisfying the QoS constraint. A performance comparison of these two cost functions will be shown in Section 4.7.

The first term of $\hat{c}(t,e)$ is derived using stochastic geometry analysis in previous chapter (in (3.29)) and then $\hat{c}(t,e)$ can be written as follows:

$$\begin{aligned} \hat{c}(t,e) &= -\hat{w}_1 \frac{8p(t,e)}{A^2 \left(\lambda_m \sqrt{p_m} + \lambda_s \int_{\forall \bar{e}\in\mathcal{E}} \sqrt{p(t,\bar{e})} m(t,\bar{e})\, de \right)^2} \\ &+ \hat{w}_2 16 p(t,e)\lambda_m^2 \left[1 + \left(\frac{p_{m,pilot}}{p_{s,pilot}} \right)^{\frac{1}{4}} \right]^4 . \end{aligned} \tag{4.35}$$

[2]We will see later in the chapter that these two cost functions result in different power control policies.

4.5.3 Mean Field Equations

Since the cost functions now only depend on the mean field and the control, the optimal control problem given in equation (4.11) is similar for all the players in the system. The HJB in equation (4.12) can then be modified as follows [67]:

$$\frac{\partial u(t,e)}{\partial t} + \min_{p(t,e)} \left(c\left(p(t,e), m(t,e)\right) - p(t,e)\frac{\partial u(t,e)}{\partial e} \right) = 0, \tag{4.36}$$

where $\min_{p(t,e)} \left(c\left(p(t,e), m(t,e)\right) - p(t,e)\frac{\partial u}{\partial e} \right)$ is the Hamiltonian, generally denoted by $H\left(e, m(t,e), \frac{\partial u(t,e)}{\partial e}\right)$, and $u(t,e)$ is the value function. The same equations are applicable for $\hat{c}(t,e)$ as well. The HJB equation models an individual player's interaction with the mass (i.e., mean field). This is also called the *backward equation*.

The motion of the mean field corresponds to a Fokker-Planck-Kolmogorov (FPK) equation which is called as the *forward equation*. The forward equation of game $\mathcal{G}_m$ is given as

$$\frac{\partial m(t,e)}{\partial t} + \frac{\partial}{\partial e}\left(m(t,e)\frac{\partial H}{\partial z} \right) = 0, \tag{4.37}$$

where $z = \frac{\partial u}{\partial e}$.

It was proven that $\frac{\partial H}{\partial z}$ can be replaced by the control [72], which is in this case $p(t,e)$. Hence, the modified FPK equation can be written as

$$\frac{\partial m(t,e)}{\partial t} - \frac{\partial}{\partial e}\left(m(t,e)p(t,e) \right) = 0. \tag{4.38}$$

The mean field equilibrium (MFE) can be obtained by solving the two coupled PDEs given in equations (4.36) and (4.38).

4.6 Solution of the Mean Field Game: Mean Field Equilibrium

4.6.1 Mean Field Equilibrium (MFE)

The solution of the MFG, namely, the mean field equilibrium (MFE) can be obtained by solving the mean field equations. There is no general technique to solve the mean field equations. In this section, I propose a finite difference technique to obtain the MFE based on the method proposed in [72]. The coupled equations (4.36) and (4.38) are iteratively solved until the equilibrium is achieved. The convergence point of the algorithm is guaranteed to be the optimal solution (i.e., MFE) if the objective function of the optimal control problem, $\mathbf{E}\left[\int_{t=0}^{T} c(t,e)dt + c(T)\right]$ is convex.

As I propose a finite difference method, the time axis $[0,T]$ and the state space $[0,E_{max}]$ are discretized into $X \times Y$ spaces. Hence, there are $X+1$ points in time and $Y+1$ points in state space. I also define

$$\delta t := \frac{T}{X} \quad \text{and} \quad \delta e := \frac{E_{max}}{Y}.$$

Solution to the forward equation

The forward equation is solved using the Lax-Friedrichs scheme to guarantee the positivity of the mean field. The Lax-Friedrichs scheme is first order accurate in both space and time [92]. By applying the Lax-Friedrichs scheme to equation (4.38), we have

$$\begin{aligned} M(i+1,j) &= \frac{1}{2}\left[M(i,j-1) + M(i,j+1)\right] \\ &\quad + \frac{\delta t}{2(\delta e)}\left[P(i,j+1)M(i,j+1) - P(i,j-1)M(i,j-1)\right], \end{aligned} \tag{4.39}$$

where $M(i,j)$ and $P(i,j)$ denote, respectively, the values of the mean field and power at time instant i and energy level j in the discretized grid.

Solution to the backward equation

The existing finite difference techniques to solve partial differential equations can not be applied directly to solve the HJB equation due to the Hamiltonian. Therefore, I reformulate the problem by writing the HJB equation as its corresponding optimal control problem with the forward equation as a constraint. The reformulated problem is as follows:

$$\min_{p(t,e),m(t,e)} \mathbf{E}\left[\int_{t=0}^{T} c(t)dt + c(T)\right],$$

subject to

$$\frac{\partial m(t,e)}{\partial t} - \frac{\partial}{\partial e}\left(m(t,e)p(t,e)\right) = 0, \;\; \forall (t,e) \in [0,T] \times [0, E_{max}]$$

and

$$\int_{e\in\mathcal{E}} m(t,e)de = 1, \;\; \forall t \in [0,T]\,. \tag{4.40}$$

The second constraint is to guarantee that the mean field gives the PDF of the state distribution over SBSs at each time instant.

Then, the Lagrangian $L\left(m(t,e), p(t,e), v(t,e)\right)$ for the above problem with the

Lagrange multiplier $v(t,e)_{\forall t,e}$ can be written as follows:

$$\begin{aligned} & L\left(m(t,e),p(t,e),u(t,e)\right) = \mathbf{E}\left[\int_{t=0}^{T} c(t,e)dt\right] + \\ & \int_{t=0}^{T}\int_{e=0}^{E_{max}} v(t,e)\left[\frac{\partial m(t,e)}{\partial t} - \frac{\partial\left(m(t,e)p(t,e)\right)}{\partial e}\right] dedt \\ = & \int_{t=0}^{T}\int_{e=0}^{E_{max}} m(t,e)c(t,e)de \;\; dt \\ + & \int_{t=0}^{T}\int_{e=0}^{E_{max}} v(t,e)\left[\frac{\partial m(t,e)}{\partial t} - \frac{\partial\left(m(t,e)p(t,e)\right)}{\partial e}\right] dedt, \end{aligned} \tag{4.41}$$

where I have assumed the terminal cost $c(T)$ to be equal to zero.

As I use a finite difference scheme to solve the forward equation (i.e., first constraint in the reformulated optimization problem), I also discretize the Lagrangian to solve the above given optimal control problem. The discretized Lagrangian L_D is given as in (4.42), where $V(i,j)$ and $C(i,j)$ denote the Lagrange multiplier and the value of the cost function at point (i,j) on the discretized grid.

$$\begin{aligned} L_D \;\; = \;\; & \delta e\;\delta t \sum_{i=1}^{X+1}\sum_{j=1}^{Y+1}\left[M(i,j)C(i,j) + V(i,j)\left(\frac{M(i+1,j) - 0.5\left(M(i,j+1)+M(i,j-1)\right)}{\delta t}\right)\right. \\ & \left. -V(i,j)\left(\frac{P(i,j+1)M(i,j+1) - P(i,j-1)M(i,j-1)}{2\delta e}\right)\right]. \end{aligned} \tag{4.42}$$

The optimal decision variables (given by P^*, M^*, V^*) must satisfy the Karush-Kuhn-Tucker (KKT) conditions. For an arbitrary point $(\bar{i},\bar{j})$ in the discretized grid, by evaluating and re-arranging the KKT condition, $\frac{\partial L_D}{\partial M(i,j)} = 0$, I deduce the following

equation to update V:

$$\begin{aligned} V(\bar{i}-1,\bar{j}) &= 0.5\left[V(\bar{i},\bar{j}-1)+V(\bar{i},\bar{j}+1)\right]-\delta t C(\bar{i},\bar{j})-\delta t\sum_{j=1}^{Y+1}\left(M(\bar{i},j)\frac{\partial C(\bar{i},j)}{\partial M(\bar{i},\bar{j})}\right) \\ &+ \frac{\delta t P(\bar{i},\bar{j})}{2\delta e}\left[V(\bar{i},\bar{j}-1)-V(\bar{i},\bar{j}+1)\right]. \end{aligned} \tag{4.43}$$

If $V(N+1,:)$ is known, the values of the Lagrange multipliers can be updated iteratively using the above equation.

Assume an optimization problem whose objective function is given by $f(\boldsymbol{x})$ and has l equality constraints each denoted by $h_i(\boldsymbol{x})_{i\in\{1,2,\ldots,l\}}$. It is known that the following relationship exists at the optimal solution [93]:

$$\bigtriangledown f(\boldsymbol{x}^*)=\sum_{i=1}^{l} v_i \bigtriangledown h_i(\boldsymbol{x}^*), \tag{4.44}$$

where $\boldsymbol{x}^*$ is the optimal solution and v_i is the Lagrange multiplier corresponding to h_i.

Let $(p^*(t,e),m^*(t,e))_{\forall(t,e)\in[0,T]\times[0,E_{max}]}$ denote the solution for the optimal control problem given in (4.40). Now, consider the optimal control problem given below for any arbitrary e' at time T:

$$\min_{p(T,e'),m(T,e')} f_T(p(T,e'),m(T,e'))=\mathbf{E}\left[\int_{t=T}^{T} c(t)dt+c(T)\right],$$

subject to

$$\frac{\partial m(T,e')}{\partial t}-\frac{\partial}{\partial e}\left(m(T,e')p(T,e')\right)=0. \tag{4.45}$$

According to Bellman's principle of optimality [83], it can be concluded that the

optimal solution to the above problem (4.45) is given by $p^*(T, e')$. As $c(T) = 0$, $\bigtriangledown f_T(p(T, e'), m(T, e')) = 0$. Assuming that the derivative of the first constraint is non-zero at the optimal point and from equation (4.44), it can be concluded that $v(T, e) = 0$ for all e. Hence, by setting $V(T, e) = 0, \forall e \in \mathcal{E}$, and then using the expression in equation (4.43) the values of the Lagrange multipliers can be updated.

Next, I consider the KKT condition, $\frac{\partial L_D}{\partial P(\bar{i},\bar{j})} = 0$ for any arbitrary point $(\bar{i}, \bar{j})$ in the discretized grid. Then

$$\sum_{j=1}^{Y+1} \left(M(\bar{i}, j) \frac{\partial C(\bar{i}, j)}{\partial P(\bar{i}, \bar{j})} \right) - \frac{M(\bar{i}, \bar{j})}{2\delta e} \left[V(\bar{i}, \bar{j} - 1) - V(\bar{i}, \bar{j} + 1) \right] = 0. \tag{4.46}$$

Equation (4.46) has to be solved for $P(\bar{i}, \bar{j})$ to obtain the transmit power at point $(\bar{i}, \bar{j})$.

Obtaining the MFE

The equations (4.39), (4.43), and (4.46) can be solved iteratively until the convergence point is obtained. The complete algorithm to obtain the converging point is given in **Algorithm 1**. I state the following theorem regarding the convergence point assuming that the Lax-Friedrichs scheme is accurate for $c(t, e)$,

Theorem 4.6.1. *The convergence point of the given algorithm is the mean field equilibrium of game $\mathcal{G}_m$ with cost function, $c(t, e)$.*

Proof. The Hessian w.r.t. $P(\bar{i}, \bar{j})$ and $M(\bar{i}, \bar{j})$ of the discretized version of the objective function of the optimization problem given in equation (4.40) can be proven to be positive for any arbitrary $(\bar{i}, \bar{j})$. Hence, the problem given in equation (4.40) is a convex optimization problem. Since the KKT conditions are necessary and sufficient conditions for the optimal solution of a convex optimization problem, the convergence

point of the algorithm is equivalent to the MFE of game $\mathcal{G}_m$ with cost function, $c(t, e)$. □

Algorithm 1 Computing the mean field equilibrium

Initialization: Initialize $M(0,:)$, V(N+1,:), $iteration = 1$
repeat
  for all $i = 1 : 1 : X$ **do**
    for all $j \in \{1, ..., Y\}$ **do**
      Calculate $M(i+1, j)$ using equation (4.39)
    end for
  end for
  if $P(i, M+1) = 0$ **then**
    $M(i+1, Y+1) = M(i, Y+1)$
  else
    $M(i+1, Y+1) = 0$
  end if
  $\forall i$, Normalize M
  for all $i = X+1 : -1 : 1$ **do**
    for all $j \in \{1, ..., Y+1\}$ **do**
      Update $V(i-1, j)$ using equation (4.43)
    end for
  end for
  for all $i = 1 : 1 : X+1$ **do**
    for all $j \in \{1, ..., Y+1\}$ **do**
      Update $P(i, j)$ using equation (4.46)
    end for
  end for
  $iteration = iteration + 1$
until $iteration \geq Iter_{max}$

4.6.2 Uniqueness of the MFE

In the following theorem, I state the **sufficient conditions** for $\mathcal{G}$ to have a unique solution.

Theorem 4.6.2. *The game $\mathcal{G}$ has a unique solution if the following conditions are satisfied:*

1. $\frac{\partial}{\partial m}H\left(p,z,m\right)>0$,

2. $\frac{\partial}{\partial z}\left(mp\right)>0$,

3. $\frac{\partial}{\partial z}H\left(p,z,m\right)>0$,

where $z=\frac{\partial u}{\partial e}$.

Proof. Assume that $(m_0(t,e),u_0(t,e))$ and $(m_1(t,e),u_1(t,e))$ are two different solutions for the game $\mathcal{G}$. Here I use the notation $x(t,e)$ to denote a continuous function of $t\in[0,T]$ and $e\in[0,E_{max}]$. Consider the following integration:

$$I(1)=\frac{d}{dt}\int_{e\in\mathcal{E}}\left(u_1(t,e)-u_0(t,e)\right)\left(m_1(t,e)-m_0(t,e)\right)de. \tag{4.47}$$

The above integration is rearranged as follows:

$$\begin{aligned} I(1) &= \int_{e\in\mathcal{E}}\left(\frac{\partial u_1(t,e)}{\partial t}-\frac{\partial u_0(t,e)}{\partial t}\right)\left(m_1(t,e)-m_0(t,e)\right)de \\ + &\int_{e\in\mathcal{E}}\left(u_1(t,e)-u_0(t,e)\right)\left(\frac{\partial m_1(t,e)}{\partial t}-\frac{\partial m_0(t,e)}{\partial t}\right)de. \end{aligned}$$

After substituting equations (4.36) and (4.38), we obtain

$$
\begin{aligned}
I(1) &= \int_{e\in\mathcal{E}} H\left(e, m_0(t,e), \frac{\partial u_0(t,e)}{\partial e}\right)(m_1(t,e)-m_0(t,e))\,de \\
&- \int_{e\in\mathcal{E}} H\left(e, m_1(t,e), \frac{\partial u_1(t,e)}{\partial e}\right)(m_1(t,e)-m_0(t,e))\,de \\
&+ \int_{e\in\mathcal{E}} \frac{\partial}{\partial e}\left(m_1(t,e)p_1(t,e)\right)(u_1(t,e)-u_0(t,e))\,de \\
&- \int_{e\in\mathcal{E}} \frac{\partial}{\partial e}\left(m_0(t,e)p_0(t,e)\right)(u_1(t,e)-u_0(t,e))\,de \\
&= \int_{e\in\mathcal{E}} H\left(e, m_0(t,e), \frac{\partial u_0(t,e)}{\partial e}\right)(m_1(t,e)-m_0(t,e))\,de \\
&- \int_{e\in\mathcal{E}} H\left(e, m_1(t,e), \frac{\partial u_1(t,e)}{\partial e}\right)(m_1(t,e)-m_0(t,e))\,de \\
&+ \int_{e\in\mathcal{E}} \left(m_0(t,e)p_0(t,e)\right)\left(\frac{\partial}{\partial e}u_1(t,e)-\frac{\partial}{\partial e}u_0(t,e)\right)de \\
&- \int_{e\in\mathcal{E}} \left(m_1(t,e)p_1(t,e)\right)\left(\frac{\partial}{\partial e}u_1(t,e)-\frac{\partial}{\partial e}u_0(t,e)\right)de.
\end{aligned}
$$

Let $\forall(t,e),\ m_\theta(t,e) = m_0(t,e) + \theta\left(m_1(t,e)-m_0(t,e)\right)$ and $u_\theta(t,e) = u_0(t,e) + \theta\left(u_1(t,e)-u_0(t,e)\right)$.

Consider the intergral

$$
\begin{aligned}
I(\theta) &= \int_{e\in\mathcal{E}} \left[H\left(e, m_0(t,e), \frac{\partial u_0(t,e)}{\partial e}\right) - H\left(e, m_\theta(t,e), \frac{\partial u_\theta(t,e)}{\partial e}\right)\right](m_\theta(t,e)-m_0(t,e))\,de \\
&+ \int_{e\in\mathcal{E}} \left(\frac{\partial}{\partial e}u_\theta(t,e)-\frac{\partial}{\partial e}u_0(t,e)\right)\left(m_0(t,e)p_0(t,e)-m_\theta(t,e)p_\theta(t,e)\right)de.
\end{aligned}
$$

Next, we write

$$\begin{aligned}\frac{I(\theta)}{\theta} &= \int_{e\in\mathcal{E}} H\left(e, m_0(t,e), \frac{\partial u_0(t,e)}{\partial e}\right)(m_1(t,e)-m_0(t,e))\,de \\ &- \int_{e\in\mathcal{E}} H\left(e, m_\theta(t,e), \frac{\partial \theta(t,e)}{\partial e}\right)(m_1(t,e)-m_0(t,e))\,de \\ &+ \int_{e\in\mathcal{E}} \left(\frac{\partial}{\partial e}u_1(t,e)-\frac{\partial}{\partial e}u_0(t,e)\right) m_0(t,e)p_0(t,e)de \\ &- \int_{e\in\mathcal{E}} \left(\frac{\partial}{\partial e}u_1(t,e)-\frac{\partial}{\partial e}u_0(t,e)\right) m_\theta(t,e)p_\theta(t,e)de.\end{aligned}$$

Using the chain rule, we have

$$\frac{d\frac{I(\theta)}{\theta}}{d\theta} = \frac{\partial I(\theta)}{\partial m_\theta}\frac{\partial m_\theta}{\partial \theta} + \frac{\partial I(\theta)}{\partial z}\frac{\partial z}{\partial \theta},$$

where $z = \frac{\partial u_\theta}{\partial e}$.

By evaluating $\frac{d\frac{I(\theta)}{\theta}}{d\theta}$, we can write

$$\frac{d\frac{I(\theta)}{\theta}}{d\theta} = \int_{e\in\mathcal{E}} \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} c & d \\ e & f \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} de, \tag{4.48}$$

where $a = m_1(t,e) - m_0(t,e)$, $b = \frac{\partial}{\partial e}u_1(t,e) - \frac{\partial}{\partial e}u_0(t,e)$, $c = -\frac{\partial}{\partial m_\theta}H\left(e, m_\theta(t,e), \frac{\partial u_\theta(t,e)}{\partial e}\right)$, $f = -\frac{\partial}{\partial \frac{\partial u}{\partial e}}\left(m_\theta(t,e)p_\theta(t,e)\right)$, $e = -\frac{\partial}{\partial z}H\left(e, m_\theta(t,e), \frac{\partial u_\theta(t,e)}{\partial e}\right)$, and $d = -p_\theta(t,e)$.

It can also be deduced that $\frac{I(\theta)}{\theta}\Big|_{\theta=0} = 0$. If $\frac{dI(\theta)}{d\theta} \leq 0$, $\frac{I(\theta)}{\theta}\Big|_{\theta=1} \leq 0$ and hence $I(1) \leq 0$. From equation (4.47),

$$\frac{d}{dt}\int_{e\in\mathcal{E}} (u_1(t,e)-u_0(t,e))\,(m_1(t,e)-m_0(t,e))\,de \leq 0.$$

According to the definition $m_1(0,:) = m_0(0,:)$ and $u_1(T,:) = u_0(T,:)$. Assuming

Table 4.2: Chapter 4: Simulation parameters

Parameter	Value
λ_m	0.00005 base stations/m^2
λ_s	$50\lambda_m$
p_m, $p_{m,pilot}$	43 dBm
p_{max}	10 dBm
$p_{s,pilot}$	13 dBm
w_1 , w_2	1000, 50000
T	0.5 s, 1 s

that m and u are monotone functions, then

$$\frac{d}{dt}\int_{e\in\mathcal{E}} \left(u_1(t,e)-u_0(t,e)\right)\left(m_1(t,e)-m_0(t,e)\right)de = 0.$$

Therefore, if $\begin{pmatrix} c & d \\ e & f \end{pmatrix}$ is negative all the time $u_1(t,e) = u_0(t,e)$ and $m_1(t,e) = m_0(t,e)$, $\forall (t,e) \in [0,T]\times[0,E_{max}]$. Hence, the solution is unique.

□

4.7 Numerical Results and Discussion

This section presents numerical results on the performance of the proposed algorithm. I also validate the stochastic geometry-based expressions derived in Section 4.5.2. The values of the main simulation parameters are given in Table 4.2.

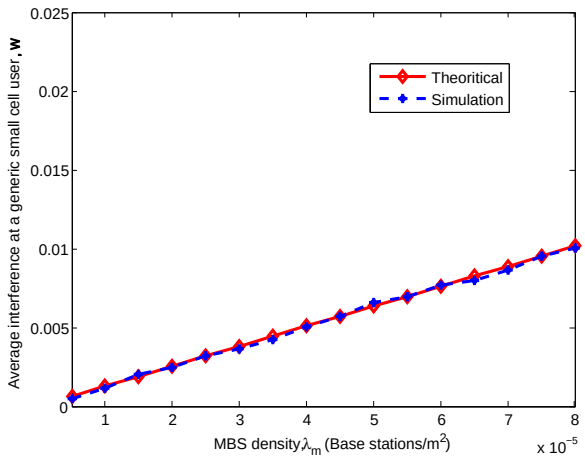


Figure 4.1: Average interference experienced by a generic small cell user (for $\lambda_s = 50\lambda_m$).

4.7.1 Validating the Expressions Derived by Stochastic Geometry Analysis

First, I validate the expressions derived by stochastic geometry analysis. To validate the average interference given in equation (4.25), I only consider the interference caused due to the SBSs (i.e., only one PPP is considered for simulation). The same result would hold for the interference caused by the macro network. A comparison of the simulation results with those obtained based on the expression in (4.25) is shown in Fig. 4.1. In Fig. 4.2, I validate the expression for the average distance to the closest possible macro user given in equation (4.30). The exact match of the theoretical and simulation results validates the accuracy of the derived expressions.

4.7.2 Behavior of the Mean Field at Equilibrium

In this section I observe the behavior of the mean field at the equilibrium. First, I set $E_{max} = 0.1J$, $p_{max} = 0.01W$, and $T = 0.5$ (i.e., 50 LTE frames). The initial energy

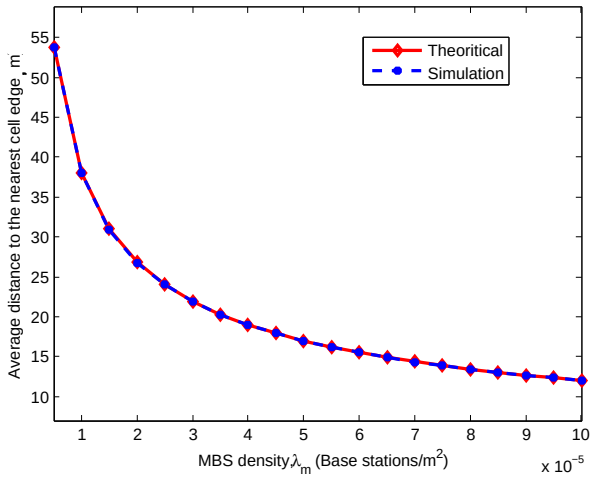


Figure 4.2: Variation of the distance to the closest edge of an SBS with λ_m.

distribution $m(0,:)$ is assumed to be uniform. The mean field at the equilibrium for cost function $c_{t,e}$ is shown in Fig. 4.3.

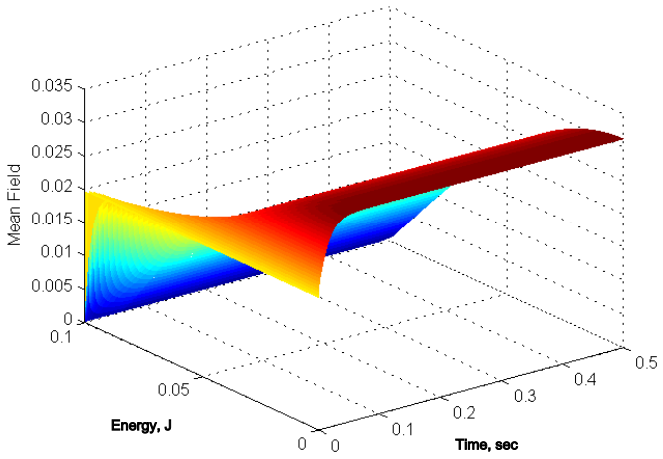


Figure 4.3: Mean field at the equilibrium for $c(t,e)$ with uniform initial energy distribution.

It can be seen from the figure that the number of SBSs with higher energy levels decreases with time. The probability of base stations having zero energy increases at the beginning of the time frame and later settles to a constant. This means,

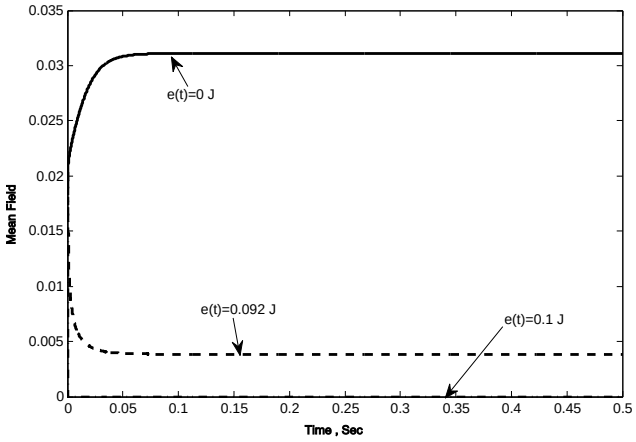


Figure 4.4: Cross-section of the mean field at equilibrium for $c(t, e)$.

although some SBSs empty their battery while transmission, all SBSs do not empty their batteries. This is because, the quadratic term (i.e., $f^{(1,mean)}$ in (4.20)) of the cost function c discourages the SBSs to increase their transmit power after satisfying the QoS constraint. Therefore, the SBSs which start transmission with higher initial energy do not empty their batteries throughout the transmission.

For illustration, I also plot several cross-sections of the mean field in Fig. 4.4, which shows the variation of the probability distribution of SBSs having a certain energy with time. Since the initial distribution is uniform, the initial probabilities are similar for all energy levels. After the transmission starts, there is no SBS with full energy as everybody transmits with non-zero power. Therefore, the probability of SBSs with maximum energy (i.e., 0.1 J) drops to zero right after the start of the transmission. The probability of SBSs having zero energy increases for sometime, as SBSs who had smaller initial energy would eventually empty their batteries.

In Fig. 4.5 and Fig. 4.6, I show the MFE considering $\hat{c}(t, e)$ in equation (4.35). $\hat{w}_1$ and $\hat{w}_2$ are set to 1. Unlike in the previous cost function $c(t, e)$, this cost function does not discourage the SBSs to increase transmit power after satisfying the QoS

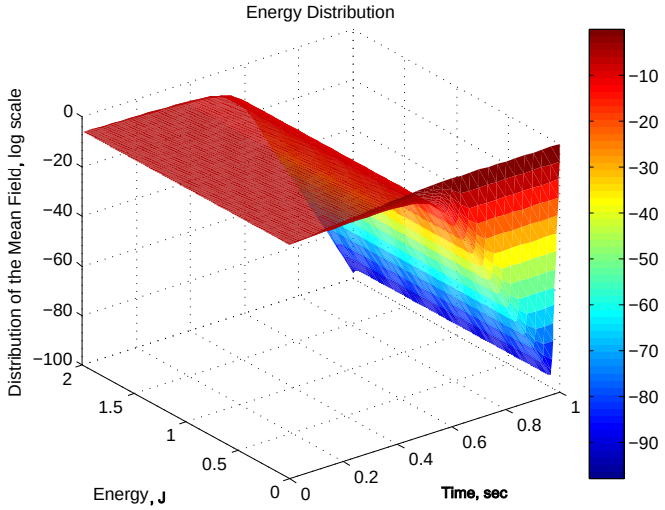


Figure 4.5: Mean field at the equilibrium for $\hat{c}(t, e)$ with uniform initial energy distribution.

constraint. Therefore, the SBSs tend to use more energy during T and result in a different mean field behavior. (Note that in this case, there are SBSs with a higher energy than the previous case as $E_{max} = 2J$.) It can be seen in Fig. 4.5 that, the probability of an SBS having zero energy is equal to one at the end of the time period T (i.e., $m(T; 0) = 1$). This means all the SBSs have emptied their energy allowance during the transmission and have zero available energy at the end of the considered time frame T. Therefore, it can be concluded that the cost function $c(t, e)$ performs better than the cost function $\hat{c}(t, e)$ in terms of energy saving.

4.7.3 Power Control Policy at the Mean Field Equilibrium

I show the transmit power policies for the game $\mathcal{G}_m$ with both cost functions $c(t, e)$ and $\hat{c}(t, e)$. Once the power policy is calculated, an SBS can decide on the transmit power based on its available energy at each time instant. Re-computation of the power policy is needed at the beginning of each time interval T (i.e., 0, T, $2T$, $3T$, ...) only if the probability distribution of allowable energy changes.

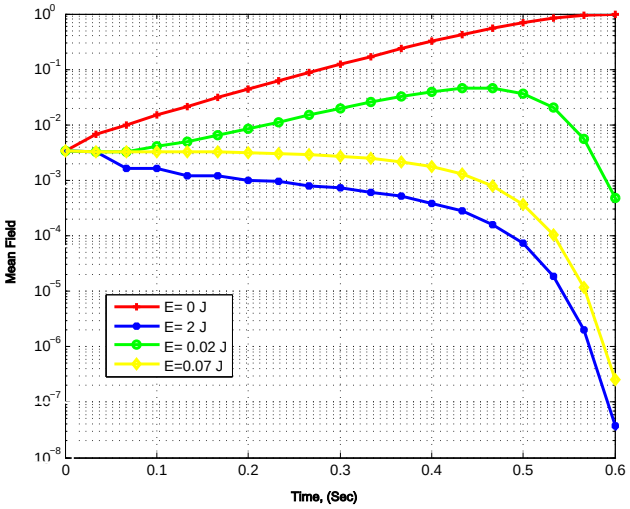


Figure 4.6: Cross-section of the mean field at equilibrium for $\hat{c}(t, e)$.

Fig. 4.7 shows the equilibrium power policy for the cost function $c(t, e)$. A uniform initial energy distribution is considered. All SBSs start transmission at low power levels. The SBSs with lower energy may empty their batteries after sometime decreasing the average interference caused to the other users. Then, the SBSs, which have sufficient energy to transmit throughout T, increase their transmit power. As the cost function $c(t, e)$ discourage the SBSs to increase power after satisfying the QoS constraint, the transmit power remains almost constant. However, the cost function $\hat{c}(t, e)$ results a different system behavior.

Fig. 4.8 shows the transmit power policy at the equilibrium for cost function $\hat{c}$. I also consider a uniform distribution of initial energy. This figure also shows that, the SBSs with higher energy start transmitting with maximum allowable transmit power while the SBSs with lower energy start with lower power. However, the SBSs with lower energy tend to increase their transmit power after some time. A cross-section of the power policy plot is shown in Fig. 4.9 for energy levels $2J, 0.2J, 0.05J$, and $0J$. The figure shows that the SBSs with higher energy start transmitting with maximum

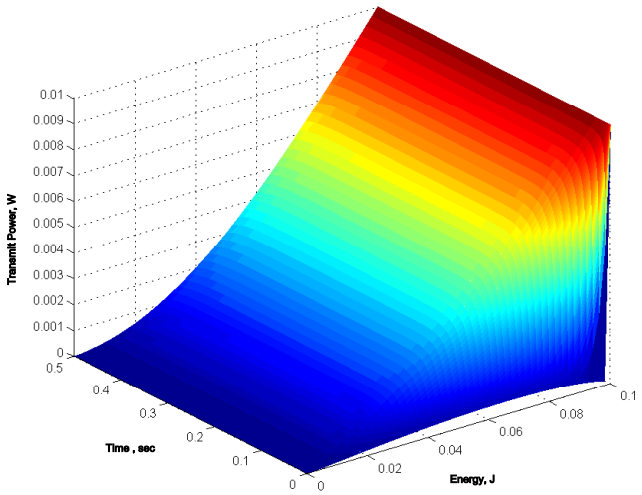


Figure 4.7: Equilibrium power policy for $\hat{c}(t, e)$ with uniform initial energy distribution.

allowable transmit power while the SBSs with lower energy start with lower power. However, the SBSs with lower energy tend to increase their transmit power after some time.

The above phenomenon is illustrated more in Fig. 4.10 where I show the transmit power policies with three different initial energy levels. The SBSs who start the game with an initial energy of $0.05J$ do not transmit at higher power at the beginning of the time period T. They increase the transmit power later in the time slot. By that time, the SBSs who started the game with higher energy have spent most of their energy and lowered their transmit power. The SBSs with less initial energy can have a better cost by increasing their transmit power later in time period T due to reduced interference. In an actual implementation, the pre-defined time T is composed of a certain number of LTE frames. In simulations I have used $T = 0.5s$ and $T = 1s$ which is equivalent to 50s and 100 LTE frames respectively. The proposed power control algorithm determines the transmit power over T for each base station in such a way

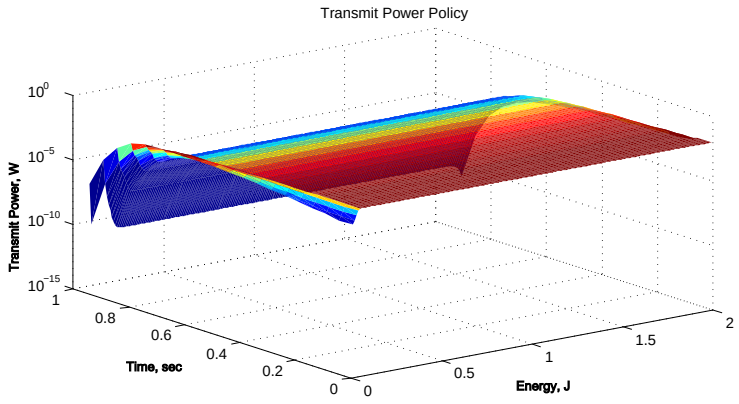


Figure 4.8: Equilibrium power policy for $\hat{c}(t, e)$ with uniform initial energy distribution.

that it minimizes the total cost for the time period T. Therefore, after the execution of the algorithm, the transmit power for each LTE frame in time T is known. When several users are connected to the base station, each user may get a certain number LTE frames out of the total number of frames in time T. The number and the order of the frames given to a user depends on the scheduling scheme used. The scheduled user at each LTE frame is served on the corresponding transmit power decided by the proposed algorithm for that frame. When a new user is admitted to the base station, service for that user can be started in the next T time period.

4.7.4 Comparison With Uniform Transmit Power Policy

Several works in the literature proposed solutions to the downlink power control problem for two-tier cellular networks. However, providing a numerical comparison (by simulation) of the proposed mean field game based solution with existing solutions would be unrealistic as different works have used different system models and different network parameters. However, a qualitative comparison among the different schemes

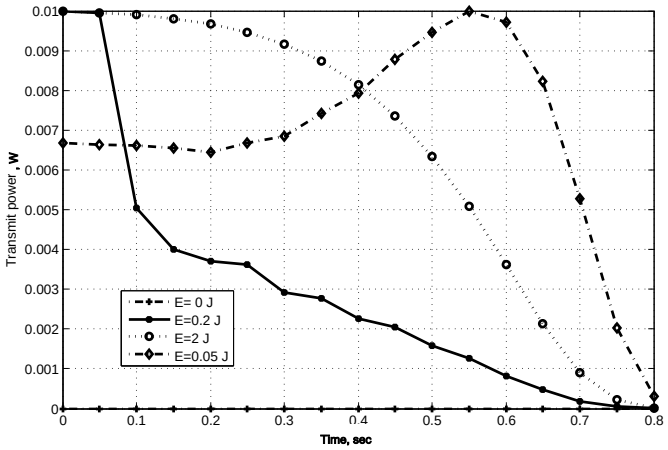


Figure 4.9: Cross-section of the power policy for $\hat{c}(t, e)$.

is provided in Table 4.3.

To illustrate the performance of the proposed algorithm numerically, I use uniform transmit power setting as a benchmark algorithm. In this case, the uniform transmit power p_k of an SBS k with initial energy $e_k(0)$ equals to $\frac{e_k(0)}{T}$. Fig. 4.11 plots the variation of average SINR with λ_s for both uniform transmit power setting and the proposed algorithm for cost function $\hat{c}(t, e)$. The results show that the transmit power policy given by the proposed algorithm performs better when the network becomes more dense. The variation of average SINR over T with λ_s for $c(t, e)$ is compared with the uniform transmit power policy in Fig. 4.12. Also, in this case the proposed algorithm outperforms the uniform power policy. However, the SINR does not increase after satisfying the QoS constraint.

4.8 Chapter Summary

I have proposed an energy-aware distributed power control algorithm for self-organizing small cell networks. The power control problem for a small cell network underlaying a macrocell network (i.e., in a co-channel deployment scenario) is first

Table 4.3: Comparison of the power control algorithms for two-tier networks

	Performance Metric	Mathematical model	Distributed or centralized	Impact on macro-tier	Remarks
[94]	Individual rate	Stackelberg game	Distributed	No	Perfect knowledge of all the channel gains is required.
[95]	Total minimum spectral efficiency	Optimization	Distributed	No	Distributed algorithm is suboptimal
[96]	Sum-rate of small cells	Optimization	Semi-distributed	No	Small cells are clustered and inside the cluster resource allocation is done
[97]	Individual throughput-power tradeoff	Optimization	Distributed	Yes	Admission control is also done for small cell users
[98]	Individual capacity	Robust Stackelberg game	Distributed	No	Imperfect channel state information is considered
[99]	Individual SINR	Evolutionary game	Distributed	Yes	Less information exchange, linear complex
Proposed Algorithm	Individual SINR	Mean field game	Distributed	Yes	Considers minimizing cost over a pre-defined period of time, energy aware,executed offline

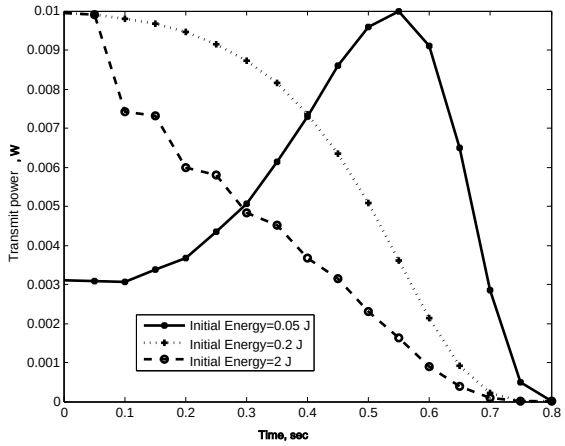


Figure 4.10: Transmit power variation of SBSs with different initial energy for $\hat{c}(t, e)$.

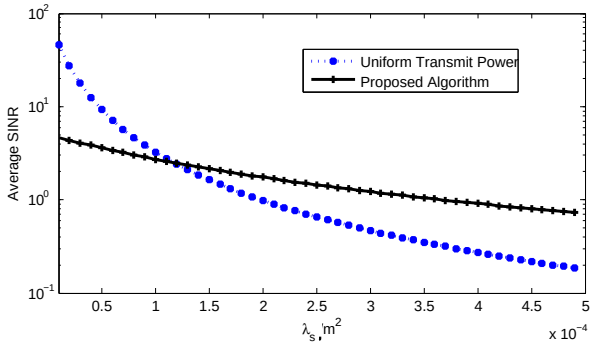


Figure 4.11: Variation of SINR at the receiver of a generic user with SBS density.

formulated as a differential game. The differential game for power control is then extended to a mean field game for a dense network. An iterative finite difference technique is proposed to solve the mean field equations based on Lax-Friedrichs scheme and Lagrange relaxation. I also have shown the sufficient conditions for the uniqueness of the mean field equilibrium. The performance of the algorithm has been analyzed for two cost functions. The main advantage of the proposed algorithm is that it can be distributively executed offline. Also, the algorithm considers minimizing the

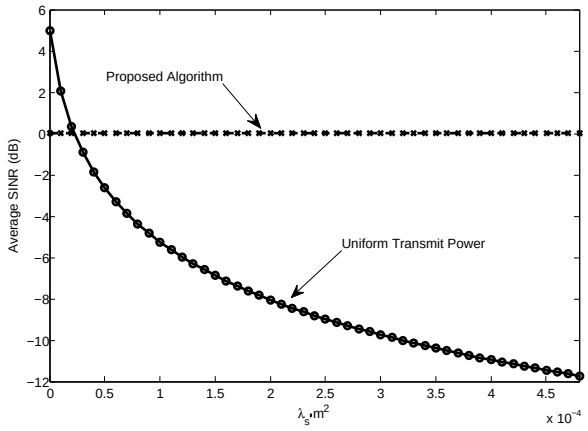


Figure 4.12: Variation of SINR at the receiver of a generic user with SBS density.

cost over a pre-defined period of time, instead of minimizing the running cost. Numerical results have been presented to demonstrate the performance of the proposed algorithm.

In this chapter and in Chapter 3, I have considered half-duplex small cells. However, full-duplex transmission technology is also a recently emerging technology in wireless communications. In next chapter, I consider distributed power control for small cells with full-duplexing capabilities. Moreover, both evolutionary game and mean field game based resource allocation techniques proposed in this thesis by far, assume that the network nodes are truthful to each other. Both these resource allocation algorithms fail if the network nodes are deceitful and send inaccurate information to other network nodes to manipulate the resource allocation scheme in such a way that deceiving nodes can obtain more benefit than others. In next chapter, I also address this problem and derive a cheat-proof distributed power control technique for small cells with full-duplexing capabilities.

Chapter 5

A Repeated Game for Cheat-Proof Distributed Power Control in Full-Duplex Small Cell Networks

In this chapter, I address the problem of distributed power control in a two-tier cellular network, where full-duplex small cells underlay a macro cell in a co-channel deployment scenario. I first formulate the distributed power control problem as a non-cooperative game and then extend it to a repeated game with imperfect public monitoring. The repeated game formulation prevents deceitful small cells from deviating from the social optimal solution for their own benefit. I establish the existence and uniqueness of the Nash equilibrium in the formulated non-cooperative game. I also characterize the set of perfect public equilibrium for the repeated game. A two-phase distributed algorithm is proposed to achieve and enforce a Pareto optimal transmit power profile. The solution obtained by this algorithm is also social optimal. Phase 1 of the algorithm is a fully-distributed learning phase based on Perturbed Markov Chains, where each base station individually learns a Pareto optimal operating point.

Phase 2 is composed of two rules: i) Detection rule based on Page-Hinckley test to detect cheating, and ii) Punishment rule to motivate cheating base stations to cooperate. Through theoretical analysis, I prove that the proposed distributed power control mechanism achieves a public perfect equilibrium point of the formulated repeated game. The power control algorithm is also cheat-proof and needs only a small amount of information exchange among network nodes. The effectiveness of the algorithm is shown through numerical analysis. The proposed model, algorithm, and analysis are also valid for a half-duplex system as a special case.

5.1 Introduction

5.1.1 Overview

Full-duplex transmission (i.e., transmitting and receiving at the same time in the same frequency band)[1] is an emerging technology which has a potential to significantly increase the spectral efficiency and hence the network capacity. Recent studies show that full-duplex technology works better for low-power transmission nodes, since self-interference can be reduced to the noise power level [100]. Therefore, deployment of full-duplex small cells is expected to be an important feature of next generation cellular networks. Most of the existing power control schemes including two methods proposed in previous chapters of this thesis, are either for uplink or downlink resource allocation, i.e., applicable only for half-duplex systems. When it comes to full-duplex systems, uplink and downlink transmissions cannot be analyzed separately, since uplink and downlink transmissions are mutually dependent due to self-interference. Consequently, the existing solutions are not applicable for a full-

[1]Simultaneous transmission and reception in the same frequency band is referred to as in-band full-duplexing and simultaneous transmission and reception in different frequency bands is referred to as out-of-band full-duplexing. In this work, I consider in-band full-duplex transmission.

duplex system unless perfect self-interference cancellation is possible. To date, only a small number of research works have addressed the problem of distributed resource allocation for full-duplex small cells. In [101], the authors propose a stable matching game for subcarrier allocation in full-duplex single-tier small cells. However, no theoretical analysis is performed. In [102], joint uplink-downlink resource allocation problem for a full-duplex single cell is modelled as a non-cooperative game and solved for NE. Most importantly, a great majority of the existing distributed resource allocation mechanisms are based on different types of non-cooperative *one-shot* games; that is, the resource allocation game is played only once. However, the resource allocation processes are inherently repeated. Taking this fact into account, it is natural to model the resource allocation problems by repeated games [103] instead of one-shot games. In a repeated game, a base game (called the *stage game*) is played a finite number of times or infinitely many times. In such a setting, every player is able to take decisions based on previous joint action profile of players as well as previous outcomes. Therefore, decisions produced by repeated games differ fundamentally from those of one-shot games. Power control and spectrum sharing problems of wireless systems are modelled as repeated games in [104], [105], [106] and [107], among others. However, those works rely on the availability of perfect information of each player's history (also called *Perfect Monitoring*). In [108], repeated games with imperfect public monitoring are used to develop a TDMA-based spectrum sharing scheme for a cognitive radio network.

5.1.2 Contribution

In this chapter, I study the problem of distributed and cheat-proof joint downlink-uplink power control for a two-tier network with multiple small cells underlaying a

macro cell. This game model, developed based on a new payoff function, is applicable to full-duplex systems in which uplink and downlink transmissions are coupled. Also, in this model, I relax the assumption of perfect monitoring and only assume imperfect monitoring of a public signal (*Imperfect Public Monitoring*). In order to solve the formulated game, I propose a distributed power control algorithm that consists of two phases. In Phase 1, each network node (either a small cell base station (SBS) or a small cell user) distributively learns the equilibrium. In Phase 2, the following two steps are executed: i) *Cheating detection*, which is performed by the macro cell base station (MBS) to monitor the system and to detect any deviation from the learned equilibrium point, and ii) *Punishment*, which is implemented to motivate the cheating nodes to cooperate with the system (incentive-compatibility). The main contributions of this chapter can be summarized as follows:

- I consider the joint downlink-uplink power control problem in a co-channel deployed two-tier network with full-duplex small cells underlaying a macro cell. I model the problem by a non-cooperative game, which accommodates the power control for both macro and small cell tiers.

- I establish the existence and uniqueness of the NE of the formulated non-cooperative stage game.

- I extend the formulated one-stage non-cooperative game to a repeated game with imperfect public monitoring to obtain a solution which Pareto dominates the NE of the stage game.

- Through theoretical analysis, I characterize the perfect public equilibrium (PPE)[2] payoff set of the formulated repeated game.

[2]The perfect public equilibrium (PPE) is equivalent to the concept of subgame perfect equilibrium

- I develop a distributed learning algorithm with linear time complexity to find the social optimal operating point. The convergence of the proposed learning model to the social optimal is proven theoretically. Moreover, it is also shown that the convergence point of the learning model is Pareto optimal.

- I implement a deviation detection and punishment policy in order to prevent selfish network nodes from acting in their own benefit and deviating from the social optimal operating point. The deviation detection algorithm is based on the Page-Hinckley test. Moreover, the following punishment strategies are studied: i) grim-trigger (punish forever) strategy and ii) punish and forgive strategy.

- I design a distributed power control scheme based on the aforementioned learning algorithm, deviation detection algorithm, and punishment policy. The convergence of the power control scheme to a PPE point of the formulated repeated game is proven theoretically.

- I also evaluate the performance of the proposed power control technique through extensive numerical analysis.

The main symbols that are used throughout this chapter are listed in Table 5.1.

5.2 System Model and Assumptions

I consider a network consisting of K small cells. In each small cell, an SBS and a user communicate with each other in a full-duplex manner using the same channel. A small cell can have multiple users in its coverage area; however, in a particular

in repeated games with perfect monitoring. The formal definition of PPE will be given in Section IV.

Table 5.1: Chapter 5: Symbols

Symbol	Description
α	Path-loss exponent
θ_k	Weighting factor of small cell k's payoff function
δ	Discount factor for repeated game $\mathcal{G}_r$
η	Interference measurement error at the macro user
ν	Learning rate of Phase 1 of power control algorithm
π_k	Payoff of player k in game $\mathcal{G}_s$
$\boldsymbol{a}_k$	Action of player k
$\boldsymbol{a}^{SO}$	Pareto optimal payoff of game $\mathcal{G}_s$
$\mathfrak{C}(.)$	Self-interference function
$\mathcal{K}$	Set of base stations
$\boldsymbol{h}(t)$	History at step t of game $\mathcal{G}_r$
I_{max}	Interference threshold at the marco user
$Iter_{max}$	Maximum number of iterations of Phase 1
$m(t)$	Public message at step t
$\mathfrak{m}_k(t)$	Mood of player k at step t
N_0	Variance of noise power
p_k^{UL}, p_k^{DL}	Uplink, downlink transmit power of base station k
$r_{i,j}$	Distance between node i and j
$\boldsymbol{s}_k$	Public strategy of player k
$\text{SINR}_k^{UL}, \text{SINR}_k^{DL}$	Uplink, downlink SINR of base station k
v_k	Payoff of player k in game $\mathcal{G}_r$

time interval (or time step), only one user communicates with the corresponding SBS. All small cells underlay a macrocell in which the MBS serves a macro user in the downlink in a time interval using half-duplex communication mode[3]. Each base station and the user served by this base station during a certain transmission interval are considered as *one entity*. The set of cells is denoted by $\mathcal{K} = \{0, 1, 2, ..., K\}$, where 0 stands for the macro cell and $1, 2, ..., K$ are used to denote the small cells. Co-channel deployment is included in the model; that is, all base stations including the MBS transmit through the same channel. Downlink and uplink transmit powers of each base station $k \in \{0, 1, 2, ..., K\}$ are denoted by p_k^{UL} and p_k^{DL}, respectively. Small cells can operate in any of the following three modes: i) full-duplex mode when $p_k^{UL}, p_k^{DL} > 0$, ii) half-duplex mode when either p_k^{UL} or p_k^{DL} is zero, and iii) OFF mode when $p_k^{UL} = p_k^{DL} = 0$. I assume that the link between the MBS and the macro user is half-duplex and consider only downlink transmission from the MBS to the macro user (i.e., $p_0^{UL} = 0$). For simplicity, only long-term signal attenuation due to path-loss is considered. However, it is straightforward to include fading in the model, and the analyses presented in the later sections remain valid.

The downlink signal-to-interference-plus-noise ratio (SINR) at the user and the uplink SINR at the base station of a generic cell k are respectively given by

$$\text{SINR}_k^{DL} = \frac{p_k^{DL} r_{k_b,k_u}^{-\alpha}}{N_0 + I_{k_u}^{UL} + I_{k_u}^{DL} + \mathfrak{C}(p_k^{UL})}, \tag{5.1}$$

and

$$\text{SINR}_k^{UL} = \frac{p_k^{UL} r_{k_u,k_b}^{-\alpha}}{N_0 + I_{k_b}^{UL} + I_{k_b}^{DL} + \mathfrak{C}(p_k^{DL})}, \tag{5.2}$$

[3]Uplink transmissions from the macro users to the MBS are not considered here. For a communication link between a base station and a user, in a repeated game set up, a complete *transmission interval* consists of multiple time steps during which the power control algorithm is executed until it converges.

where N_0 is variance of noise power and α is the path-loss exponent. Moreover, k_u and k_b denote the user and base station of cell k, respectively. In addition, $r_{i,j}$ denotes the distance between network nodes[4] i and j. Also,

- $I_l^{UL} = \sum_{j \in \mathcal{K}-\{l\}} p_j^{UL} r_{j_u,l}^{-\alpha}$ is the interference caused by the uplink transmissions of all the other users at the network node l,
- $I_l^{DL} = \sum_{j \in \mathcal{K}-\{l\}} p_j^{DL} r_{j_b,l}^{-\alpha}$ is the interference caused by the downlink transmissions of the other base stations at the network node l, and
- $\mathfrak{C}(p)$ is the self-interference at a network node which transmits at power p.

The above equations are valid for the macro cell as well. In particular, $\text{SINR}_0^{UL} = 0$ and $\mathfrak{C}(p_k^{UL}) = 0$, since according to this system model $p_k^{UL} = 0$. It is also assumed that each network node selects its transmit power from a pre-defined finite set.

The MBS is capable of obtaining a noisy measurement of the interference caused by small cells to the macro user being served during a transmission interval. After each transmission step, the MBS broadcasts a public message based on the measured interference to all SBSs on a delay-free channel. I also assume that all users have a perfect delay-free feedback channel to their base stations. Thus, users can update the base stations about their performances at each step. Moreover, the initial location information of the small cell network nodes is available at the MBS.

5.3 Stage Game and Analysis of Nash Equilibrium

5.3.1 Stage Game, $\mathcal{G}_s$

In the following, I first formulate the stage game $\mathcal{G}_s$.

[4]A network node is either a user or a base station.

Set of players, $\mathcal{K}$: Each base station and its associated user are considered as *one player* of $\mathcal{G}_s$. Thus the set of players is denoted by $\mathcal{K} = \{0, 1, 2, ..., K\}$.

Set of actions for player k, $\mathcal{A}_k$: Downlink and uplink transmit powers of each base station-user pair $k \in \mathcal{K} - \{0\}$ is drawn from a finite set of power levels given by $\mathcal{P}_k^{UL} = \{0, p_k^{UL,1}, p_k^{UL,2}, ..., p_k^{UL,max}\}$ and $\mathcal{P}_k^{DL} = \{0, p_k^{DL,1}, p_k^{DL,2}, ..., p_k^{DL,max}\}$,

respectively. Note that for player $k = 0$ (i.e., the MBS according to the system model), $\mathcal{P}_0^{UL} = \{0\}$. $\mathcal{P}_0^{DL}$ is similar to other players, as given above. The action of any player k (denoted by $\boldsymbol{a}_k$) is the combination of uplink and downlink transmit powers. This is also called the transmit power profile (i.e., $\boldsymbol{p}_k$) of player k. The action of player k thus yields

$$\boldsymbol{a}_k = \boldsymbol{p}_k = \begin{pmatrix} p_k^{UL} & p_k^{DL} \end{pmatrix}. \tag{5.3}$$

In addition, $\boldsymbol{a}_{-k}$ represents the actions of all players other than k. The joint action profile of the system is thus given by $\boldsymbol{a} = \boldsymbol{p} = [\boldsymbol{a}_0\ \boldsymbol{a}_1\ ...\ \boldsymbol{a}_K] \in \mathcal{A} = \mathcal{A}_0 \times \mathcal{A}_1 \times ... \times \mathcal{A}_K$. The size of the action set of each player is given by: $|\mathcal{A}_k| = |\mathcal{P}_k^{UL}||\mathcal{P}_k^{DL}|$. I will use the symbols $\boldsymbol{a}$ and $\boldsymbol{p}$ interchangeably in the rest of this chapter.

Payoff function of player k, π_k: For small cells underlaying a macro cell network, the following payoff function for a small cell $k \in \mathcal{K} - \{0\}$ is generally used ([109,110]):

$$\text{Payoff}_k = \mathcal{F}_1\left(\text{SINR}_k^{UL}, \text{SINR}_k^{DL}\right) - \mathcal{F}_2\left(\sum_{j \in \mathcal{K}-\{k\}} p_k^{UL} r_{k_u,j}^{-\alpha} + \sum_{j \in \mathcal{K}-\{k\}} p_k^{DL} r_{k_b,j}^{-\alpha}\right). \tag{5.4}$$

Note that the argument in the second term of the right-hand-side of (5.4) is the interference caused by small cell k (i.e., player k, where $k \neq 0$) to the macro user and other small cell nodes. Thus, the payoff function consists of two terms: i) $\mathcal{F}_1(\cdot)$ that corresponds to the gained utility and depends on the SINR at receivers, and ii) $\mathcal{F}_2(\cdot)$

that corresponds to the cost or penalty for causing interference to the other network nodes. Although including such penalty factor might assist theoretical analysis, it is virtual and thus does not cause any actual performance reduction to small cells. As a result, dishonest small cell nodes may continue increasing their transmit powers in order to increase the real positive utility, neglecting the increase in the virtual penalty. This might result in a performance reduction for the macro network as well as other small cells.

Typically, the following two types of dishonest behavior might be observed:

1. *Small cell network nodes are dishonest to each other, but all of them are truthful to the MBS:*

   In this case, the MBS can select the maximum transmit powers for all nodes in the small cell network (i.e., $p_k^{UL,max}, p_k^{DL,max}, \forall k \in \{1, 2, ..., K\}$) in such a way that the interference level at the macro user is below a certain pre-defined threshold, denoted by I_{max}. Let $\mathcal{A}' = \mathcal{A}_1 \times ... \times \mathcal{A}_K$ be the set of all joint transmit power profiles of small cells. Moreover, let $\boldsymbol{p}_{small} = [\boldsymbol{p}_1 \; ... \; \boldsymbol{p}_K] \in \mathcal{A}'$ denote the transmit power vector (i.e., joint transmit power profile) of all nodes in the small cell network. Then, before execution of the algorithm, the MBS solves the following optimization problem, in order to decide maximum transmit power of each node in the network:

$$\begin{aligned} \underset{\boldsymbol{p}_{small} \in \mathcal{A}'}{\text{Maximize}} \quad & \boldsymbol{p}_{small} \\ \text{subject to} \quad & \sum_{k=1}^{K} \left(r_{0_u,k_b} p_k^{DL} + r_{0_u,k_u} p_k^{UL} \right) \leq I_{max} \\ & p_k^{UL}, p_k^{UL} \geq 0 \; , \quad \forall k \in \{1, 2, ..., K\}. \end{aligned}$$

   Note that only the initial location information of the small cell network nodes

(which is available at the MBS) is required to solve the optimization problem.

2. *Small cell network nodes are dishonest to each other and the MBS:*

   In this case, there is no benefit in solving the above optimization problem and setting the maximum transmit power for each node in the small cell network. Thus, there is no minimum interference guarantee and the MBS has to accept the interference at the operating point.

For both the aforementioned cases, instead of considering a payoff function in the form of (5.4), I formulate a payoff function which is only based on the utility gained due to the SINR value at the receiver. In particular, the payoff of player k is formulated as

$$\pi_k = \theta_k \frac{\log\left(1+\mathrm{SINR}_k^{UL}\right)}{\log\left(1+\mathrm{SINR}_k^{UL,max}\right)} + (1-\theta_k)\frac{\log\left(1+\mathrm{SINR}_k^{DL}\right)}{\log\left(1+\mathrm{SINR}_k^{DL,max}\right)}, \tag{5.5}$$

where $\theta_k \in [0,1]$ is a weighting factor and $\mathrm{SINR}_k^{DL,max}$ and $\mathrm{SINR}_k^{UL,max}$, respectively, denote the maximum possible downlink and uplink SINR values that can be achieved in small cell k. The maximum possible SINR is achieved when the transmitter is transmitting at the maximum possible transmit power and the interference at the receiver (caused by other nodes and self-interference) is zero. Therefore, $\mathrm{SINR}_k^{UL,max}$ and $\mathrm{SINR}_k^{DL,max}$ can be calculated as

$$\mathrm{SINR}_k^{UL,max} = \frac{p_k^{UL,max}}{N_0}, \quad \text{and} \tag{5.6}$$

$$\mathrm{SINR}_k^{DL,max} = \frac{p_k^{DL,max}}{N_0}. \tag{5.7}$$

It is also worth noting that π_k has a value between 0 and 1 and equivalent to the aggregate normalized uplink and downlink rate for Gaussian interference. Note

that by changing the weighting factor θ_k, the priority for uplink and downlink can be changed. In other words, the uplink SINR can be prioritized by increasing θ. These biasing factors can be set based on the system requirement prior to the execution of the power control algorithm.

5.3.2 Analysis of Nash Equilibrium

In the following theorem, I comment on the existence and the uniqueness of the NE of the above stage game $\mathcal{G}_s$.

Theorem 5.3.1. *Stage game $\mathcal{G}_s$ has a unique NE when the following condition is satisfied:*

- *For all small cells, the self-interference cancellation technique is strong enough such that the minimum average uplink received power at the SBS is greater than the self-interference induced by the maximum downlink transmit power, i.e.,*

$$p_k^{UL,min} r_{k_u,k_b}^{-\alpha} > \mathfrak{C}\left(p_k^{DL,max}\right), \forall k \in \{1,2,...,K\}. \tag{5.8}$$

Moreover, the equilibrium power profile is given by $\boldsymbol{p}_{k,NE} = \left(p_k^{UL,max}, p_k^{DL,max}\right)$, $\forall k \in \mathcal{K}$.

Proof. Following payoff function can be written for any small cell $k \in \{1,2,...,K\}$:

$$\pi_k = \theta \frac{\log\left(1 + \frac{p_k^{DL} r_{k_b,k_u}^{-\alpha}}{N_0 + I_{k_u}^{UL} + I_{k_u}^{DL} + \mathfrak{C}\left(p_k^{UL}\right)}\right)}{\log\left(1 + \text{SINR}_k^{DL,max}\right)} + (1-\theta) \frac{\log\left(1 + \frac{p_k^{UL} r_{k_u,k_b}^{-\alpha}}{N_0 + I_{k_b}^{UL} + I_{k_b}^{DL} + \mathfrak{C}\left(p_k^{DL}\right)}\right)}{\log\left(1 + \text{SINR}_k^{UL,max}\right)}. \tag{5.9}$$

For player 0, the payoff is given by

$$\pi_0 = \frac{\log\left(1 + \frac{p_0 r_{0_b,0_u}^{-\alpha}}{N_0 + I_{0_u}^{UL} + I_{0_u}^{DL}}\right)}{\log\left(1 + \text{SINR}_0^{DL,max}\right)}. \tag{5.10}$$

By Nash's existence theorem [111], a game with a finite number of players and a finite number of pure strategy actions for each player has at least one (mixed strategy) NE. Moreover, from (5.6) and (5.7), the maximum achievable uplink and downlink SINRs are constants. If the condition given in (5.8) is satisfied, by assuming that $p_k^{DL,min} > p_k^{UL,min}$ and $p_k^{DL,max} > p_k^{UL,max}$, $\forall k \in \{1, 2, ..., K\}$, it can be concluded that $p_k^{DL,min} r_{k_b,k_u}^{-\alpha} > \mathfrak{C}\left(p_k^{UL,max}\right)$, $\forall k \in \{1, 2, ..., K\}$. Then it is easy to observe that the best response of player k for any fixed $\boldsymbol{p}_{-k}$ is to transmit with its maximum power, both in uplink and downlink (that is, once $\boldsymbol{p}_{-k}$ is fixed, the larger the uplink and downlink power, the higher will be the payoff). This means $\left(p_k^{UL,max}, p_k^{DL,max}\right)$ is a strictly dominant strategy for all players. All players will play their dominant strategy at any NE. Since there is one strictly dominant strategy for each player, the game has only one NE. The equilibrium strategy $\boldsymbol{p}_{k,NE}$ is given by $\left(p_k^{UL,max}, p_k^{DL,max}\right), \forall k \in \mathcal{K}$. □

The only NE for the above game is all nodes transmitting their maximum possible transmit power. Therefore, at the NE, the players also experience the maximum interference. Thus, the NE may not be the Pareto optimal power/payoff profile. However, in the next section, I show that by formulating a repeated game, a solution that Pareto dominates $\boldsymbol{p}_{k,NE}$ can be obtained.

5.4 Repeated Game and Analysis of Public Perfect Equilibrium

5.4.1 Repeated Game, $\mathcal{G}_r$

I define the game $\mathcal{G}_s$ be repeated infinitely many times and denote the corresponding repeated game by $\mathcal{G}_r$. The main components of $\mathcal{G}_r$ are described below.

Set of players, $\mathcal{K}$: Naturally, the set of players of $\mathcal{G}_r$ is the same as that of $\mathcal{G}_s$, i.e., $\mathcal{K} = \{0, 1, 2, ..., K\}$.

Action set of player k, $\mathcal{A}_k$: The action set of a generic player k at each step of $\mathcal{G}_r$ is same as the action set of that player in $\mathcal{G}_s$, previously denoted by $\mathcal{A}_k$. Moreover, the action of player k at t^{th} step of the game $\mathcal{G}_r$ is denoted by $\boldsymbol{a}_k(t) = \boldsymbol{p}_k(t)$.

Payoff of player k, v_k: The payoff of player k is defined as

$$v_k = (1-\delta) \lim_{T\to\infty} \sum_{t=0}^{T} (\delta)^t \pi_k(t), \tag{5.11}$$

where $0 \leq \delta < 1$ is the discount factor. The discount factor can also be seen as the probability of not ending the game at a certain time step. Also, $\boldsymbol{v} = [v_0\ v_1 \quad v_2 \quad ... \quad v_K]^T$ is the payoff profile (payoff vector) of all players.

History at step t, $\boldsymbol{h}(t)$: History of the game $\mathcal{G}_r$ at step t is composed of actions played by all the players until step $(t-1)$. The action profile of the system at step t is given by

$$\boldsymbol{a}(t) = [\boldsymbol{a}_0(t)\ \boldsymbol{a}_1(t)\ ...\ \boldsymbol{a}_K(t)]^T.$$

Therefore, at time step t, the history is given by

$$\boldsymbol{h}(t) = [\boldsymbol{a}(1)\ \boldsymbol{a}(2)\ ...\ \boldsymbol{a}(t-1)]. \tag{5.12}$$

Perfect monitoring requires every player to have knowledge of $\boldsymbol{h}(t)$. Thus, players need to exchange information about their actions after each step of the game. Since such scenario results in excessive overhead, it is not realistic for a small cell network with limited backhaul. In order to avoid this problem, I proceed as follows: I assume that the MBS measures the interference at its user. I denote the MBS's measurement on interference at its user at time t by $I\left(\boldsymbol{p}_{-0}(t)\right)$, which can be written as

$$I\left(\boldsymbol{p}_{-0}(t)\right) = \left(I_{0_u}^{UL} + I_{0_u}^{DL}\right) + \eta, \tag{5.13}$$

where $\eta \sim \mathcal{N}\left(\mu, \sigma^2\right)$ is the estimation error. Based on $I\left(\boldsymbol{p}_{-0}(t)\right)$, the MBS transmits a public message $m(t)$ which is drawn from a finite set of messages given by $\mathcal{M} = \{\mathrm{m}_1, \mathrm{m}_2, ..., \mathrm{m}_M\}$. The public history at step t can be then written as

$$\boldsymbol{h}_{pub}(t) = \left[m(1)\ m(2)\ ...\ m(t-1)\right]. \tag{5.14}$$

Due to the broadcast nature of wireless transmission, it is realistic to assume that this signal is heard by all network nodes. Thus, with very low overhead, the public history becomes available at all nodes.

Public strategy of player k, $\boldsymbol{s}_k$: The public strategy of player k is a mapping of any possible public history to an action. Since $\mathcal{G}_r$ is a repeated game with infinite time horizon, it has an infinite number of possible public histories. Thus the public strategy space ($\mathcal{S}_k$) of each player k is also infinite. Also, the public strategy profile of the system is given by $\boldsymbol{s} = [\boldsymbol{s}_0\ \boldsymbol{s}_1\ ...\ \boldsymbol{s}_K]$.

The equilibrium concept for a repeated game with public history and public strategies is called perfect public equilibrium (PPE). The formal definition of PPE is given below.

Definition (*Perfect Public Equilibrium [PPE]*): A public strategy profile $\bar{\boldsymbol{s}} = [\bar{\boldsymbol{s}}_0\ \bar{\boldsymbol{s}}_1\ ...\ \bar{\boldsymbol{s}}_K]$ is a perfect public equilibrium if for any public history $\boldsymbol{h}_{pub}(t)$, the continuation public strategy given by $\bar{\boldsymbol{s}}|h(t)$ is a Nash equilibrium of the continuing subgame for all players, i.e.,

$$v_k(\bar{\boldsymbol{s}}|h_{pub}(t)) \geq v_k(\hat{\boldsymbol{s}}_k, \bar{\boldsymbol{s}}_{-k}|h_{pub}(t)),\ \ \forall \hat{s}_k \in \mathcal{S}_k,\ \forall k. \tag{5.15}$$

5.4.2 Analysis of Perfect Public Equilibrium Set

Before characterizing the set of PPEs of game $\mathcal{G}_r$, I define the following terms.

Definition (*Minmax Payoff*): The minmax payoff is defined for the stage game. For any player k, the minmax payoff, $\pi_{k,minmax}$, is the maximum achievable payoff when every other player aims at minimizing the payoff of k. Formally,

$$\pi_{k,minmax} = \min_{\boldsymbol{a}_{-k} \in \mathcal{A}_{-k}} \max_{\boldsymbol{a}_k \in \mathcal{A}_k} \pi_k\left(\boldsymbol{a}_k, \boldsymbol{a}_{-k}\right). \tag{5.16}$$

Note that the NE of the stage game $\mathcal{G}_s$ results in the minmax payoff for all players.

Definition (*Enforceability*): A payoff profile $\boldsymbol{v}$ of the repeated game $\mathcal{G}_r$ is enforceable if $v_k \geq \pi_{k,minmax}$, $\forall k$. (Strictly enforceable if $v_k > \pi_{k,minmax}\ \forall k$.)

Definition (*Feasibility*): A payoff vector $\boldsymbol{v}$ of the repeated game $\mathcal{G}_r$ is feasible if there exist rational, non-negative values of $\zeta_{\boldsymbol{a}} \forall \boldsymbol{a} \in \mathcal{A}$, such that $v_k \forall k \in \mathcal{K}$ can be expressed as

$$\sum_{\boldsymbol{a} \in \mathcal{A}} \zeta_{\boldsymbol{a}} \pi_k\left(\boldsymbol{a}\right), \quad \text{where} \quad \sum_{\boldsymbol{a} \in \mathcal{A}} \zeta_{\boldsymbol{a}} = 1. \tag{5.17}$$

Definition (*Individual Full Rank*): Let $\Pi_k\left(\boldsymbol{a}_{-k}\right)$ be a matrix of size $|\mathcal{A}_k| \times |\mathcal{M}|$. For

an arbitrary player k and a fixed $\boldsymbol{a}_{-k}$, $\Pi_k(\boldsymbol{a}_{-k})$ as follows.

$$\Pi_k(\boldsymbol{a}_{-k}) = \begin{pmatrix} \Pr(m_1/\boldsymbol{a}_1, \boldsymbol{a}_{-k}) & \Pr(m_2/\boldsymbol{a}_1, \boldsymbol{a}_{-k}) & ... & \Pr(m_{|\mathcal{M}|}/\boldsymbol{a}_1, \boldsymbol{a}_{-k}) \\ \Pr(m_1/\boldsymbol{a}_2, \boldsymbol{a}_{-k}) & \Pr(m_2/\boldsymbol{a}_2, \boldsymbol{a}_{-k}) & ... & \Pr(m_{|\mathcal{M}|}/\boldsymbol{a}_2, \boldsymbol{a}_{-k}) \\ . & ... & ... & . \\ . & ... & ... & . \\ \Pr(m_1/\boldsymbol{a}_{|\mathcal{A}_k|}, \boldsymbol{a}_{-k}) & \Pr(m_2/\boldsymbol{a}_{|\mathcal{A}_k|}, \boldsymbol{a}_{-k}) & ... & \Pr(m_{|\mathcal{M}|}/\boldsymbol{a}_{|\mathcal{A}_k|}, \boldsymbol{a}_{-k}) \end{pmatrix}. \tag{5.18}$$

An action profile $\boldsymbol{a}$ has *individual full rank* for player k if $\Pi_k(\boldsymbol{a}_{-k})$ has rank $|\mathcal{A}_k|$.

Folk Theorem *for repeated games with imperfect public monitoring*:

Let $\mathcal{V}^*$ denote the set of feasible and enforceable payoffs of a repeated game $\mathcal{G}_r$. In [112], it is shown that for any smooth subset W in the interior of $\mathcal{V}^*$, there exists a $\hat{\delta}$ such that for all $\delta \in (\hat{\delta}, 1)$, each point in W is a PPE, if $\mathcal{G}_r$ satisfies the following conditions for all players: i) Any pure action profile has individual full rank, and ii) the number of public messages exceeds $|\mathcal{A}_k| - 1$.

This means that any point in $\mathcal{V}^*$ can be obtained as a PPE for patient players, i.e., when $\delta \longrightarrow 1$. Thus, there can be infinitely many PPEs. My objective is to find a desired operating point for the system which is also a PPE. The minmax action profile can then be used as a punishment to motivate players to stay at the desired point.

Even with the assistance of the Folk theorem, solving the formulated repeated game is challenging. First, all players need to find a PPE operating point. Second, the equilibrium strategy should ensure that all rational players would stay truthful, i.e., the players should not have any incentive to deviate from the PPE operating point. To find a PPE operating point and to implement an equilibrium strategy, I propose a distributed mechanism that consists of the following two phases: i) learning

phase, and ii) operating phase. Detailed analyses of these two phases are provided in the next two sections.

5.5 Learning Phase: Finding a PPE Operating Point

5.5.1 Learning Algorithm

During the learning phase, the players are expected to learn a PPE operating point distributively. In [113], the authors propose a distributed algorithm that converges to a social optimal action profile (i.e., an action profile which maximizes the sum of the payoffs of all players) for a system with a finite number of players with finite action set. The algorithm is developed based on the theory of perturbed Markov chains. By modifying the algorithms presented in [113] and [114], I develop a learning model that finds a PPE operating point for the formulated repeated game $\mathcal{G}_r$. The proposed learning model is entirely distributed and does not require any information exchange among the network nodes. Each node only has access to its own historical payoffs and actions, and decisions are made solely based on individual payoff as described below.

Each player k has a state given by $\mathfrak{m}_k(t)$, also called the *mood*, which can take two values, namely, *Content* (C) or *Discontent* (D). The state or mood of a player reflects its inclination towards experimenting new actions. More precisely, if a certain player is *Content*, it will occasionally experiment new actions; in contrast, if a player is *Discontent*, it will experiment new actions more often. Once the current state $\mathfrak{m}_k(t)$, current action $\boldsymbol{a}_k(t)$, and the individual instantaneous payoff $\pi_k(t)$ at step t are known, the next action is selected according to the rules given in **Algorithm 2**. In **Algorithm 2**, $\nu \in (0, 1)$ and $c > |K|$ are constants. From **Algorithm 2**, it can be

observed that if a certain player is in mood C, then it is more likely to stay with the same action in the next step. A *Discontent* player, on the contrary, is more likely to experiment other actions in the next step by randomizing over all actions belonging to its action set.

Algorithm 2 Action updating algorithm

1: **Initialization:** Select constants $\nu \in (0,1)$ and $c > |\mathcal{K}|$.
2: **if** $(\mathfrak{m}_k(t) == C)$ **then**
3: Select $\boldsymbol{a}_k(t+1)$ according to following rule:

$$\Pr(d) = \begin{cases} 1-\nu^c, & \text{if } d == \boldsymbol{a}_k(t) \\ \frac{1-\nu^c}{|\mathcal{A}_k|-1}, & \text{otherwise} \end{cases}$$

4: **else**
5: Select $\boldsymbol{a}_k(t+1)$ according to following rule:

$$\Pr(d) = \frac{1}{|\mathcal{A}_k|}, \quad \forall d \in \mathcal{A}_k$$

6: **end if**

The state of a player is updated based on its individual instantaneous payoff $\pi_k(t)$ and current action $\boldsymbol{a}_k(t)$. Once the next action $\boldsymbol{a}_k(t)$ is selected by executing **Algorithm 2**, the next state is updated by following the rules given in **Algorithm 3**.

From **Algorithm 3**, it can be observed that, the larger the instantaneous payoff, the higher is the probability that the players are *Content*. In essence, if $\pi_k(t) = 1$ (i.e., maximum possible $\pi_k(t)$), then $\mathfrak{m}_k(t+1) = C$ with probability 1 as $\nu^{(1-\pi_k(t))} = 1$. The *Content* players are reluctant to experiment new actions according to **Algorithm 2**. Therefore, all players are likely to stay with the actions which give higher payoffs.

All network nodes keep updating their moods and actions until the maximum number of steps ($Iter_{max}$) for learning phase is reached or until the instantaneous payoff does not change for a significant number of iterations (which can be predefined).

Algorithm 3 Mood updating algorithm

if $(\mathfrak{m}_k(t) == C)$ **then**
  if $(\boldsymbol{a}_k(t+1) == \boldsymbol{a}_k(t))$ **then**
    $\mathfrak{m}_k(t+1) = C$
  else

$$\mathfrak{m}_k(t+1) = \begin{cases} C, & \text{with prob.} \quad \nu^{(1-\pi_k(t))} \\ D, & \text{with prob.} \quad 1-\nu^{(1-\pi_k(t))} \end{cases}$$

  end if
else

$$\mathfrak{m}_k(t+1) = \begin{cases} C, & \text{with prob.} \ \nu^{(1-\pi_k(t))} \\ D, & \text{with prob.} \ 1-\nu^{(1-\pi_k(t))} \end{cases}$$

end if

I present the complete learning method in **Algorithm 4**.

Algorithm 4 Complete learning algorithm

Initialization: $Iter_{max}$= maximum number of iterations, select $\boldsymbol{a}_k$ randomly, $\mathfrak{m}_k = D$, $t = 1$, $c = |\mathcal{K}| + 1$
repeat
  $\nu = \frac{1}{\sqrt{t}}$
  Step 1: Select $\boldsymbol{a}_k(t)$ by following the rules given in **Algorithm 2**
  Step 2: Update mood $\mathfrak{m}_k(t)$ by following the rules given in **Algorithm 3**
  $t = t + 1$
until $t \geq Iter_{max}$
$\boldsymbol{p}_k^{con} = \boldsymbol{a}_k$
Inform $\boldsymbol{p}_k^{con}$ to the MBS.

5.5.2 Theoretical Analysis of the Learning Phase

Before proceeding to theoretical analysis, in the following, I define the social optimal action profile ($\boldsymbol{a}^{SO}$) and interdependence property of one-shot games.

Definition (*Social Optimal Action Profile*): A social optimal action profile, $\boldsymbol{a}^{SO}$, is referred to as a pure strategy action profile of the stage game $\mathcal{G}_s$ that maximizes the

sum of the payoffs of all players. Formally,

$$\boldsymbol{a}^{SO} = \arg\max_{\boldsymbol{a}\in\mathcal{A}} \sum_{k\in\mathcal{K}} \pi_k(\boldsymbol{a}). \tag{5.19}$$

Definition (*Interdependence*): Consider a finite game $\mathcal{G}$ with a set of players $\mathcal{J}$. Any arbitrary player j's set of actions is denoted by $\mathcal{A}_j$. The set of pure strategy action profiles of the system is then given by $\boldsymbol{a} \in \prod_{j\in\mathcal{J}} \mathcal{A}_j$. Let $\bar{\mathcal{J}}$ denote a non-empty subset of $\mathcal{J}$. The pure strategy action profile of players in subset $\bar{\mathcal{J}}$ is shown by $\boldsymbol{a}_{\bar{J}}$, and the pure strategy action profile of players which do not belong to the subset $\bar{\mathcal{J}}$ is shown by $\boldsymbol{a}_{-\bar{J}}$. Then, the game $\mathcal{G}$ is interdependent if the following condition is satisfied:

- For any arbitrary $\bar{\mathcal{J}}$ and for any arbitrary action profile $\boldsymbol{a} = \boldsymbol{a}_{\bar{J}} \cup \boldsymbol{a}_{-\bar{J}} \in \mathcal{A}$, there is at least one player $i \notin \bar{\mathcal{J}}$ and an action profile $\hat{\boldsymbol{a}}_{\bar{J}} \in \{\prod_{k\in\bar{\mathcal{J}}} \mathcal{A}_k\}$ such that $\pi_i(\hat{\boldsymbol{a}}_{\bar{J}}, \boldsymbol{a}_{-\bar{J}}) \neq \pi_i(\boldsymbol{a}_{\bar{J}}, \boldsymbol{a}_{-\bar{J}})$.

In words, interdependency means that any subset of players has the ability to change the payoff of at least one player outside the subset, by changing the actions of the players inside the subset. In other words, if a game is interdependent, it is not possible to split up players into two different sets which do not mutually affect each other. Now the following results on the proposed learning model can be stated.

Lemma 5.5.1. *Game $\mathcal{G}_s$ is interdependent.*

Proof. See **Appendix B.1**. □

Theorem 5.5.1. *If all players update their actions and moods according to* ***Algorithm 4*** *at every stage of repeated game $\mathcal{G}_r$, then the system's asymptotic convergence point has the following properties:*

1. *The action profile at the convergence point is social optimal, i.e.,* $\boldsymbol{a}(t)_{t\longrightarrow\infty} = \boldsymbol{a}^{SO}$.

2. *All players are* Content *after convergence, i.e.,* $\mathfrak{m}_k(t)_{t\longrightarrow\infty} = C, \ \forall k \in \mathcal{K}$.

Proof. See **Appendix B.2**. □

Corollary 5.5.1. *If* ***Algorithm 4*** *is applied to the repeated game* $\mathcal{G}_r$*, the instantaneous payoff profile and the action profile at the convergence point of the algorithm are respectively equivalent to the Pareto optimal payoff profile and the corresponding action profile of the stage game* $\mathcal{G}_s$.

Proof. An action profile $\boldsymbol{a}^{PO}$ is said to be Pareto optimal if the payoff profile achieved by playing $\boldsymbol{a}^{PO}$ is not Pareto dominated by any other any other payoff profile. Let $\boldsymbol{a}^{con}$ be the action profile at the convergence point of **Algorithm 4** for game $\mathcal{G}_r$. I prove that $\boldsymbol{a}^{con}$ is Pareto optimal by following the method of proof by contradiction. First, assume

$$\boldsymbol{a}^{con} \neq \boldsymbol{a}^{PO}. \tag{5.20}$$

Then, $\boldsymbol{a}^{PO}$ should Pareto dominate $\boldsymbol{a}^{con}$ and hence the following can be deducted:

$$\sum_{k\in\mathcal{K}} \pi_k\left(\boldsymbol{a}^{con}\right) < \sum_{k\in\mathcal{K}} \pi_k\left(\boldsymbol{a}^{PO}\right). \tag{5.21}$$

However, from *Theorem 5.5.1* we have

$$\boldsymbol{a}^{con} = \boldsymbol{a}^{SO} = \arg\max_{\boldsymbol{a}\in\mathcal{A}} \sum_{k\in\mathcal{K}} \pi_k\left(\boldsymbol{a}\right). \tag{5.22}$$

Therefore, (5.21) and (5.22) contradict with each other. Consequently, the assumption (5.20) is false. Hence, the convergence point of the algorithm is a Pareto optimal point, i.e., $\boldsymbol{a}^{con} = \boldsymbol{a}^{PO}$. □

Theorem 5.5.2. *Any feasible payoff vector obtained by following **Algorithm 4** in $\mathcal{G}_r$ (denoted by $\boldsymbol{v}^A$) is enforceable if the players are patient.*

Proof. According to *Corollary 5.5.1*, if all players follow **Algorithm 4**, after convergence, the instantaneous payoff profile is Pareto optimal. Let the system converge after $\bar{T}$ time slots. Then, the payoff obtained by any player k in game $\mathcal{G}_r$ can be written as

$$v_k^A = (1-\delta)\left(\sum_{t=0}^{\bar{T}}(\delta)^t\pi_k(t) + \sum_{t=\bar{T}+1}^{\infty}(\delta)^t\pi_k\left(\boldsymbol{a}^{PO}\right)\right). \tag{5.23}$$

For smaller values of δ (i.e., for patient players) $\sum_{t=\bar{T}+1}^{\infty}(\delta)^t\pi_k\left(\boldsymbol{a}^{PO}\right) >> \sum_{t=0}^{\bar{T}}(\delta)^t\pi_k(t)$. Hence we have

$$v_k^A = (1-\delta)\left(\sum_{t=\bar{T}+1}^{\infty}(\delta)^t\pi_k\left(\boldsymbol{a}^{PO}\right)\right) = (1-\delta)\frac{\pi_k\left(\boldsymbol{a}^{PO}\right)}{(1-\delta)} = \pi_k\left(\boldsymbol{a}^{PO}\right) \geq \pi_{k,minmax}. \tag{5.24}$$

Since $v_k^A \geq \pi_{k,minmax}\ \forall k \in \mathcal{K}$, payoff profile $\boldsymbol{v}^A$ is enforceable. □

From *Theorem 5.5.2*, it can be observed that action profile $\boldsymbol{a}^{con}$ (i.e., $\boldsymbol{a}^{PO}$) and payoff profile v_k^A are together a PPE operating point if the game $\mathcal{G}_r$ satisfies the conditions required by Folk theorem, as stated in Section 5.4.2.

5.6 Operation Phase: Implementing an equilibrium strategy

5.6.1 Steps in the Operation Phase

Phase 2 of the learning mechanism is called the *Operation Phase*. The goal of this phase is to implement an equilibrium strategy which enforces the potential PPE operating point learned by executing Phase 1 (the learning phase) and provide the

incentive to all players to stay at the learned operating point. This phase is composed of the following two steps:

1. **Detection:** This step is performed by the MBS in order to detect if any player is deviating from the desired operating point learned at Phase 1.

2. **Punishment:** In case a deviation from the desired operating point (i.e., cheating) is detected, a punishment should be given to ensure that the cheating players do not benefit in the long run.

In next two subsections, these two steps are explained in details.

5.6.2 Step 1: Detection of the Point of Change of Interference

As mentioned in Section 5.4, at each step of $\mathcal{G}_r$, the MBS obtains a noisy measurement on the interference at the macro user. The MBS then broadcasts a public message $m(t) \in \mathcal{M} = \{\mathrm{m}_1, \mathrm{m}_2\}$ based on these interference measurements. The measured interference at step t, $I(t)$, can be written as

$$I(t) = \left(I_{0_u}^{UL}(t) + I_{0_u}^{DL}(t)\right) + \eta, \tag{5.25}$$

where $\eta \sim \mathcal{N}(\mu, \sigma^2)$. All small cell nodes are supposed to transmit at the power levels learned in Phase 1. If at least one small cell node (the SBS or the small cell user) cheats by increasing its transmit power at a certain time and continues transmitting with the increased power, the statistics of the interference measurement I also changes from that time onwards. Detection of statistical changes of a data sequence is a fundamental problem in many applications [115]. I propose an algorithm to identify cheating by detecting the statistical changes of the interference measurements, $I(t)$. To this end, I take advantage of the *Page-Hinkley* test [116], which is a simple

approach to detect any change point in the statistical mean of a sequence of random variables. The proposed cheating detection rules are given in **Algorithm 5**.

Algorithm 5 Detection of cheating

Initialization: Select a threshold value th, $\Lambda(0) = 0$
repeat At each step $t \in \{1, 2, ...\}$
  $m(t-1) = \mathrm{m}_1$
  $\bar{I} = \frac{1}{t}\sum_{\tau=1}^{t} I(\tau)$
  $\Lambda(t) = \Lambda(t-1) + I(t) - \bar{I}(t)$
  $\Lambda_{min} = \min\{\Lambda(\tau) | \tau \in \{1, 2, 3, ..., t\}\}$
until $(\Lambda(t) - \Lambda_{min}) > th$
$m(t) = \mathrm{m}_2$
$T_c = \arg\min\{\Lambda(\tau) | \tau \in \{1, 2, ..., t\}\}$

According to **Algorithm 5**, if no increase is detected in the statistical mean of I, the public message sent by the MBS to the small cells equals m_1. In case of cheating, the public signal equals m_2. If small cells receive m_2 as the public message, they will trigger the punishment strategy, which will be explained in next subsection. This algorithm also estimates the step where the player(s) started deviating from the learned solution, $\boldsymbol{a}^{PO}$. This change point is denoted by T_c. Note that in **Algorithm 5**, th is a pre-defined threshold, based on which a change in the mean value of the sequence of measured interferences is alarmed. Thus, the value of th affects the algorithm's performance measures such as the false alarm rate and the detection delay (i.e., time taken to detect cheating). More precisely, smaller th results in faster detection, but the probability of false alarm increases as an undesired side effect. I will analyze this trade-off numerically in Section 5.8.

Now, I consider a simplified version of the repeated game $\mathcal{G}_r$, denoted by $\bar{\mathcal{G}}_r$. In $\bar{\mathcal{G}}_r$, every player can select among only two actions, i.e., $\mathcal{A}_k = \{a_{1,k}, a_{2,k}\}, \ \forall k \in \mathcal{K}$. I state the following result.

Theorem 5.6.1. *In $\bar{\mathcal{G}}_r$, the payoff profile obtained by executing **Algorithm 4** (i.e.,*

$\bar{v}_k^A$) is a PPE payoff profile if public messages are selected according to ***Algorithm 5****.*

Proof. At any round t, for any arbitrary action $\boldsymbol{a}_{i,k}$ of any arbitrary player k, for any fixed $\boldsymbol{a}_{-k}$ and fixed sequence of interference measurements until time $(t-1)$, it can be written

$$\begin{aligned} & \Pr\left(m_1/\boldsymbol{a}_{i,k}, \boldsymbol{a}_{-k}\right) = \Pr\left(\Lambda(t) - \Lambda_{min} - T > 0\right) \qquad (5.26) \\ = \; & \Pr\left(\Lambda(t-1) + I(t) - \frac{1}{t}\sum_{\tau=1}^{t} I(\tau) - \min\{\Lambda(0), \Lambda(1), \Lambda(2), ..., \Lambda(t)\} - T > 0\right). \end{aligned}$$

Also, from (5.25), $I(t) = I_{actual}(t) + \eta$, where $I_{actual}(t)$ is the actual interference the macro user experiences at time t. Let $\Lambda(t-1) + \left(1 - \frac{1}{t}\right) I_{actual}(t) - \frac{1}{t}\sum_{\tau=1}^{t-1} I(\tau) = K_1$ and $\Lambda_{min}^{(t-1)} = \min\{\Lambda(0), \Lambda(1), \Lambda(2), ..., \Lambda(t-1)\}$, since $I_{actual}(t)$ and the sequence of interference measurements until time $(t-1)$ are not random. By substituting in the above equation I have

$$\Pr\left(m_1/\boldsymbol{a}_{i,k}, \boldsymbol{a}_{-k}\right) = \Pr\left(K_1 + \left(1 - \frac{1}{t}\right)\eta - \min\{\Lambda_{min}^{(t-1)}, K_1 + \left(1 - \frac{1}{t}\right)\eta\} - T > 0\right). \qquad (5.27)$$

Let Ω be the sample space of the values that can be taken by η. Consider the following two events: i) $A = \left\{\eta \in \Omega |\;\; \eta < \frac{\Lambda_{min}^{(t-1)} - K_1}{\left(1 - \frac{1}{t}\right)}\right\}$, and ii) $A^c = \left\{\eta \in \Omega |\;\; \eta > \frac{\Lambda_{min}^{(t-1)} - K_1}{\left(1 - \frac{1}{t}\right)}\right\}$.

By applying the law of total probability,

$$\begin{aligned}
& \Pr\left(m_1/\boldsymbol{a}_{i,k}, \boldsymbol{a}_{-k}\right) \\
= \; & \Pr\left(A\right) \Pr\left(\left(K_1 + \left(1-\frac{1}{t}\right)\eta - \min\{\Lambda_{min}^{(t-1)}, K_1 + \left(1-\frac{1}{t}\right)\eta\} - T > 0\right) \mid A\right) \\
+ \; & \Pr\left(A^c\right) \Pr\left(\left(K_1 + \left(1-\frac{1}{t}\right)\eta - \min\{\Lambda_{min}^{(t-1)}, K_1 + \left(1-\frac{1}{t}\right)\eta\} - T > 0\right) \mid A^c\right) \\
= \; & \Pr\left(A\right) \Pr\left(-T > 0\right) + \Pr\left(A^c\right) \Pr\left(K_1 + \left(1-\frac{1}{t}\right)\eta - \Lambda_{min}^{(t-1)} - T > 0\right) \\
= \; & \Pr\left(\eta < \frac{K_1 - \Lambda_{min}^{(t-1)}}{\left(1-\frac{1}{t}\right)}\right) \Pr\left(\eta < \frac{K_1 - \Lambda_{min}^{(t-1)} - T}{\left(1-\frac{1}{t}\right)}\right) \\
= \; & \frac{1}{4}\left[1 + \operatorname{erf}\left(\frac{\frac{K_1 - \Lambda_{min}^{(t-1)}}{\left(1-\frac{1}{t}\right)}}{\sqrt{2}\sigma}\right)\right]\left[1 + \operatorname{erf}\left(\frac{\frac{K_1 - \Lambda_{min}^{(t-1)} - T}{\left(1-\frac{1}{t}\right)}}{\sqrt{2}\sigma}\right)\right].
\end{aligned} \tag{5.28}$$

$\Pi_k\left(\boldsymbol{a}_{-k}\right)$ for player k of $\bar{\mathcal{G}}$ can be written as $\Pi_k\left(\boldsymbol{a}_{-k}\right) = \begin{pmatrix} \Pr\left(m_1/\boldsymbol{a}_{1,k}, \boldsymbol{a}_{-k}\right) & 1 - \Pr\left(m_1/\boldsymbol{a}_{1,k}, \boldsymbol{a}_{-k}\right) \\ \Pr\left(m_1/\boldsymbol{a}_{2,k}, \boldsymbol{a}_{-k}\right) & 1 - \Pr\left(m_1/\boldsymbol{a}_{2,k}, \boldsymbol{a}_{-k}\right) \end{pmatrix}$. We know that $\boldsymbol{a}_{-k}$ is fixed and K_1 is an increasing function of a_k. Hence, from (5.28), it is clear that $\Pr\left(m_1/\boldsymbol{a}_{1,k}, \boldsymbol{a}_{-k}\right) \neq \Pr\left(m_1/\boldsymbol{a}_{2,k}, \boldsymbol{a}_{-k}\right)$. Thus the determinant of Π_k is non-zero and Π_k has individual full rank. By *Theorem 5.5.2*, $\bar{v}_k^A$ is enforceable. Hence, according to the *Folk theorem* (see Section 5.4.2), the payoff profile obtained by executing **Algorithm 4** is that for a PPE, if public messages are selected according to **Algorithm 5**. □

5.6.3 Step 2: Punishment

Step 1 is followed by Step 2 only when cheating is detected, that is, when m_2 is received as the public message. A punishment is performed to prevent players from

evaluating cheating as worthy. Generally, the two following punishment techniques are applied in repeated games [117]: i) Grim trigger strategy and ii) punish and forgive strategy. I will describe these two punishment rules in the following subsections, and for the latter strategy, I will propose a novel approach to efficiently tune the *duration of punishment*.

Grim trigger strategy

Once cheating is detected, all players start transmitting at maximum power both in uplink and downlink forever, which is also the NE power profile. Formally,

$$\boldsymbol{p}_k(t+1) = \begin{cases} \boldsymbol{p}_k^{con}, & \text{if} \quad h(t) = [\mathrm{m}_1, \mathrm{m}_1, ..., \mathrm{m}_1] \\ \boldsymbol{p}_{k,NE}, & \text{otherwise.} \end{cases}$$

If for all k, $\pi_k(\boldsymbol{p}^{con}) > \pi_k(\boldsymbol{p}_{NE})$, the grim trigger strategy causes a reduction of long-term payoff for any cheating player. Therefore, if all players are rational, they have no incentive to deviate from $\boldsymbol{p}^{con}$ and grim trigger strategy. Consequently, the grim trigger strategy is stable.

Punish and forgive strategy

Once a deviation from $\boldsymbol{p}^{con}$ is detected, pulling the grim trigger is equivalent to playing the stage game forever. This will substantially decrease the efficiency of the entire system. As a solution, a punish and forgive strategy can be implemented. The idea is that, when a deviation is detected, all players will transmit at the maximum possible power for some rounds T_p, and afterward the players play $\boldsymbol{p}^{con}$ again. The duration of punishment, T_p, should be long enough to diminish the benefit obtained by cheating.

Therefore, choosing T_p is critical when implementing the punish and forgive strategy. To make any cheating worthless for players, the duration of punishment T_p needs to satisfy the following condition: The maximum possible benefit obtained by cheating (denoted by $\tilde{b}^{max}$) should be less than the minimum possible payoff reduction (denoted by $\tilde{r}^{min}$) during the punishment.

Although the proposed change detection algorithm can detect any deviation from the desired operating point, it is incapable of identifying the dishonest player(s). In this case, $\tilde{b}^{max}$ and $\tilde{r}^{min}$ should be estimated. Since all players update their actions at the end of the learning phase (see **Algorithm 4**), $\boldsymbol{p}^{con}$ is known to the MBS. Moreover, the initial location information of the small cell nodes is known to the MBS. Using this information, the MBS estimates the instantaneous payoff of player k as

$$\begin{aligned}&\tilde{\pi}_k\left(\boldsymbol{p}_k, \boldsymbol{p}_{-k}\right)\\ =\ & \theta \frac{\log\left(1+\frac{p_k^{DL} r_{k_b,k_u}^{-\alpha}}{N_0+I_{k_u}^{UL}(\boldsymbol{p}_{-k})+I_{k_u}^{DL}(\boldsymbol{p}_{-k})}\right)}{\log\left(1+\mathrm{SINR}_k^{DL,max}\right)} + (1-\theta)\frac{\log\left(1+\frac{p_k^{UL} r_{k_u,k_b}^{-\alpha}}{N_0+I_{k_b}^{UL}(\boldsymbol{p}_{-k})+I_{k_b}^{DL}(\boldsymbol{p}_{-k})}\right)}{\log\left(1+\mathrm{SINR}_k^{UL,max}\right)}.\end{aligned} \tag{5.29}$$

Now, I define

$$\tilde{b}^{max} = \max\{k \in \{1,2,...,|\mathcal{K}|\} \mid \left(\tilde{\pi}_k\left(\boldsymbol{p}_k^{NE}, \boldsymbol{p}_{-k}^{con}\right) - \tilde{\pi}_k\left(\boldsymbol{p}_k^{con}, \boldsymbol{p}_{-k}^{con}\right)\right)\}, \quad \text{and} \tag{5.30}$$

$$\tilde{r}^{min} = \min\{k \in \{1,2,...,|\mathcal{K}|\} \mid \left(\tilde{\pi}_k\left(\boldsymbol{p}_k^{con}, \boldsymbol{p}_{-k}^{con}\right) - \tilde{\pi}_k\left(\boldsymbol{p}_k^{NE}, \boldsymbol{p}_{-k}^{NE}\right)\right)\}. \tag{5.31}$$

Assume that **Algorithm 5** detects a deviation at time T_d, and let T_c be the change point as estimated by **Algorithm 4**. The punishment duration T_p is then calculated

as

$$T_p = \left\lceil \frac{(T_d - T_c)\,\tilde{b}^{max}}{\tilde{r}^{min}} \right\rceil. \tag{5.32}$$

The proposed punish and forgive strategy is summarized in **Algorithm 6**. When

Algorithm 6 Punish and forgive strategy

1: **Initialize:** Set $pd = 0$.
2: For every step t
3: **if** $((m(t) == \mathrm{m}_1)$ && $(pd == 0))$ **then**
4: $\quad \boldsymbol{p}_k(t+1) = \boldsymbol{p}_k^{con}$
5: **else**
6: $\quad \boldsymbol{p}_k(t+1) = \boldsymbol{p}_{k,NE}$ and $pd = pd + 1$
7: **end if**
8: **if** $(pd == T_p)$ **then**
9: $\quad pd = 0$
10: **end if**

T_p is selected according to (5.32) and the actions are selected to $\boldsymbol{p}_k(t+1)$ as given in **Algorithm 6**, cheating players gather no benefit in the long run, since the payoff reduction during the punishment duration is larger than the increase in payoff obtained by cheating. As an immediate result, if the punish and forgive strategy is implemented, the rational players would not deviate from the social optimal action profile.

5.7 Distributed Power Control Algorithm

In this section, I present the complete distributed power control algorithm based on the two phases explained in the previous section. During Phase 1, the players (i.e., small cells) learn Pareto optimal solution (which is also the social optimal solution in this case) by acting according to **Algorithm 4**. In practice, **Algorithm 4** can actually be executed at the SBSs on behalf of both base station and user. Moreover, in Phase 1, the MBS measures the interference at its user and perform steps 4, 5, and

6 of **Algorithm 5**. Note that during the learning phase, the MBS does not alarm any change in the mean interference experienced by the macro user. At the end of Phase 1, all small cells report their current action (i.e., $\boldsymbol{p}_k^{con}$) to the MBS. I can identify two types of dishonest behavior by small cells, after they find out the social optimal operating point.

- *Type 1*: In this case, the small cells try to mislead the MBS by reporting transmission power (by SBS and/or the small cell user) which is higher than the true $\boldsymbol{p}_k^{con}$. The dishonest small cell nodes start transmitting with the reported power from the first step of Phase 2. Since the MBS has collected interference measurements at the macro user during the learning phase (by performing steps 4, 5, and 6 of **Algorithm 5**), this type of cheating is also detected as an increase in the interference. However, the MBS needs $\boldsymbol{p}^{con}$ to calculate the appropriate duration of punishment, which is unknown. Thus, in this case, the grim trigger strategy is used as a punishment.

- *Type 2*: In this case, the small cells report the true values of $\boldsymbol{p}_k^{con}$ to the MBS, but increase their transmission power at some later time during Phase 2. In this case, the punish and forgive strategy can be used as explained in the previous section.

The complete distributed power control algorithm is shown in Fig. 5.1. For each step of the learning algorithm, every small cell preforms only a few simple mathematical operations including generating random numbers. **Algorithm 4** has a constant time-complexity of $\mathcal{O}\left(Iter_{max}\right)$. The Page-Hinckley test based detection algorithm (**Algorithm 4**) has a time-complexity of $\mathcal{O}\left(\bar{T}^2\right)$, where $\bar{T}$ is the average run length of the detection algorithm [118].

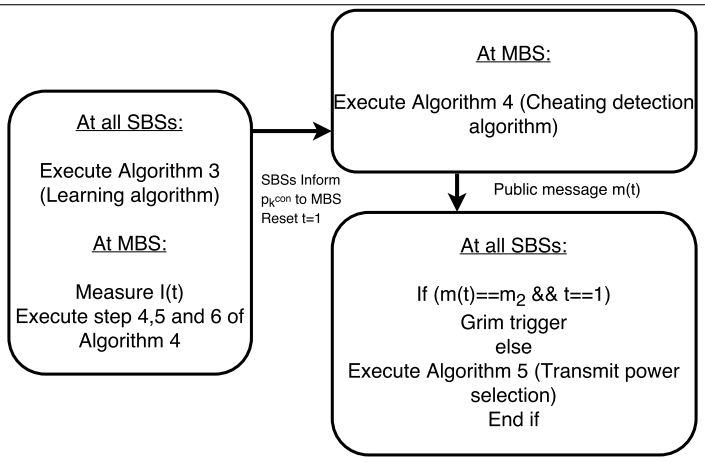


Figure 5.1: Distributed power control mechanism.

5.8 Numerical Results and Discussion

In this section, I evaluate the performance of the proposed algorithm numerically. In the first step, I consider a toy system model consisting of two small cells underlaying the macrocell. Simulation parameters are summarized in Table 5.2. I consider the case where small cells do not cheat on the MBS. Therefore, the MBS does not take part in the game. The MBS transmits in the downlink direction with some fixed transmission power. The small cell users transmit in the uplink direction with some fixed transmit power. For the downlink transmission, SBSs select a transmission power from the set $\{15,\ 20\}$. The game matrix is given in Table 5.3. From this matrix, it can be observed that the Pareto optimal solution corresponds to both the SBSs transmitting the lowest possible transmit power.

5.8.1 Phase 1: Learning Pareto Optimal Solution

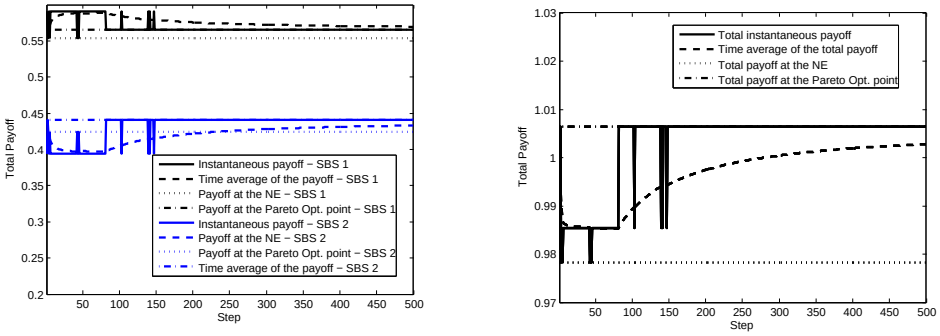
Fig. 5.2 shows the convergence of the system during Phase 1 (Learning Phase). It can be seen from Fig. 5.2(a) that, for both players, the learning phase converges to an operating point that results in some payoff higher than the NE. Moreover, from Fig.

Table 5.2: Chapter 5: Simulation parameters

Parameter	Value
(x, y) coordinates of SBS 1	$(-2.8010, 5.6566)$
(x, y) coordinates of user of SBS 1	$(-2.8233, 7.7278)$
(x, y) coordinates of SBS 2	$(-5.7635, -2.8713)$
(x, y) coordinates of user of SBS 2	$(-2.9953, -4.1508)$
N_0	0.001 W
α	4
p_1^{UL},p_2^{UL}	2 W
$\mathcal{P}_1^{DL}, \mathcal{P}_2^{DL}$	$\{15,\ 20\}$ W
γ	0.001
δ	1
θ	0.7

Table 5.3: Game matrix

$p_1^{DL}, p_2^{DL}(W)$	π_1, π_2
$(15, 15)$	$(0.5659, 0.4406)$
$(15, 20)$	$(0.5288, 0.4698)$
$(20, 15)$	$(0.5909, 0.3946)$
$(20, 20)$	$(0.5540, 0.4241)$



(a) Payoff of each player. (b) Aggregate payoff of the system.

Figure 5.2: Phase 1: Payoff for the toy model.

5.2(b), the learning phase eventually converges to the social optimal solution, which yields the highest aggregate payoff according to the game matrix shown in Table 5.3. This convergence point is also Pareto optimal since it is not Pareto dominated by any other payoff profile.

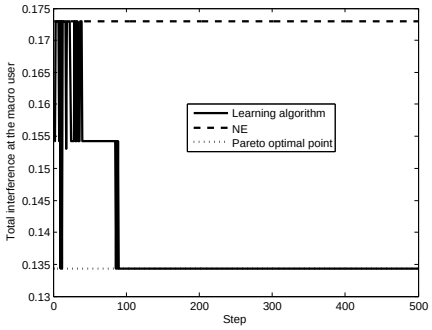


Figure 5.3: Phase 1: Aggregate interference at the macro user for the toy model.

Fig. 5.3 plots the interference measured at the macro user during Phase 1. The macro user experiences the highest interference if the system is at the NE, due to the fact that all small cell nodes transmit at the maximum transmit power at the NE. It can be seen from Fig. 5.3 that the interference at the macro user is reduced at the operating point learned during the learning phase. For the presented toy model, this convergence point also causes the minimum interference at the macro user since

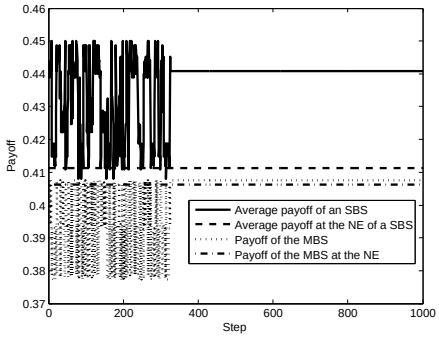


Figure 5.4: Phase 1: Payoff for a system with two small cells and one macro cell.

both the SBSs transmit at their minimum possible transmit power. In the next step I observe the learning phase when the macro cell also participates in the game with the two small cells. Fig. 5.4 plots the average payoff of a small cell player and the payoff of the macro player during the learning phase. Similar to the previous case, the learning algorithm converges to a point which yields a payoff better than that of the NE for all players.

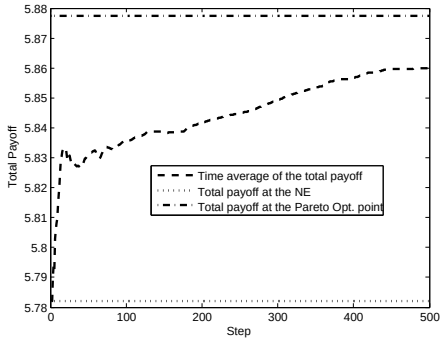


Figure 5.5: Phase 1: Payoff for a system with 8 small cells.

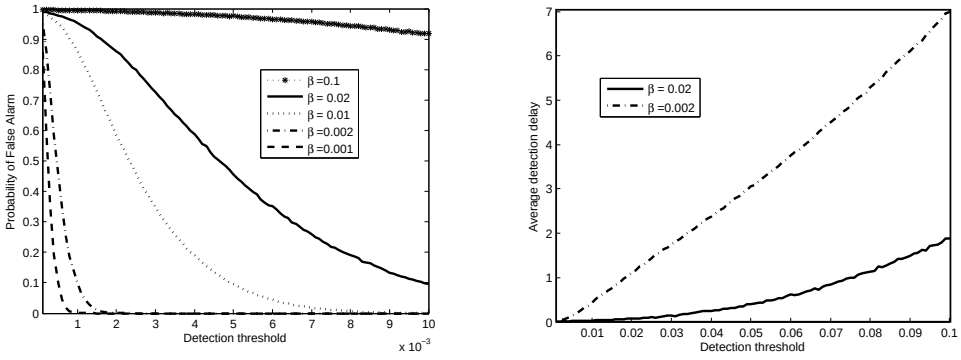
Next, in Fig. 5.5, I observe the behavior of the learning phase for a larger network with 8 small cells. The set of actions for each player is similar to that in the toy model. In this case also, the learning algorithm results in a social optimal payoff better than that for the NE; nevertheless, this system shows a slower convergence rate compared to the toy model, due to a larger number of players. Moreover, a

smaller difference between the payoff values leads to slower convergence, due to the following reason. In **Algorithm 3**, the probability of being *Content* (or *Discontent*) at a certain step t is given by $\nu^{(1-\pi_k(t))}$; thus, similar payoff values result in similar probabilities. Consequently, the algorithm needs more time to distinguish those action profiles which deliver similar payoffs. Due to the same reason, the learning mechanism might also converge to some operating point in which the aggregate payoff is slightly less than the social optimal payoff, but still better than that of NE.

5.8.2 Phase 2: Operation Phase

In this section, I discuss the performance of Phase 2 of the proposed power control scheme. For all simulations, I consider the toy model described in the previous section. Moreover, player 1 cheats by transmitting its maximum transmit power (i.e., 20 W), starting 10 rounds after the beginning of Phase 2. For simulations, I set the interference measurement error (η) as follows:

$\eta = \beta \times \bar{n} \times$ (maximum average interference at the macro user), where $\bar{n} \sim \mathcal{N}(0, 1)$.



(a) Detection rule: Probability of false alarm.

(b) Detection rule: Detection delay.

Figure 5.6: Effect of change-threshold on the performance of Page-Hinkly test.

I first show the performance of the deviation detection rule based on the Page-Hinckley test described in **Algorithm 5**. In Fig. 5.6(a) I illustrate the variation of the false alarm probability as a function of the detection threshold th. For any value of β, smaller values of detection threshold yield higher probability of false alarm. Moreover, it can be observed that the false alarm probability increases with the value of β, which can be explained as follows. For larger measurement errors, there is a higher chance for the Page-Hinckley test to mistakenly alarm a change in the mean interference even if there is no actual change. Therefore, in order to avoid a high false alarm probability, one has to increase the detection threshold. However, increasing the detection threshold th increases the delay of detecting any deviation. This is evident from Fig. 5.6(b), which shows the variation of the detection delay[5] with detection threshold. It can also be observed that detection delay is smaller for larger values of β. The reason is as follows. In case of cheating, for higher values of β, the deviation from the mean interference increases, and thus it is easily detectable.

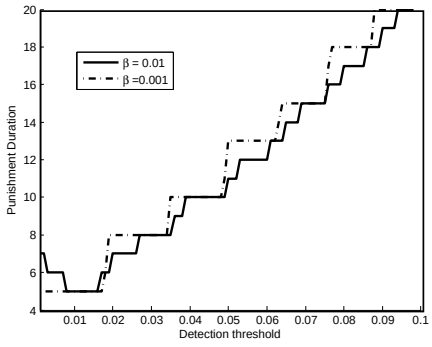


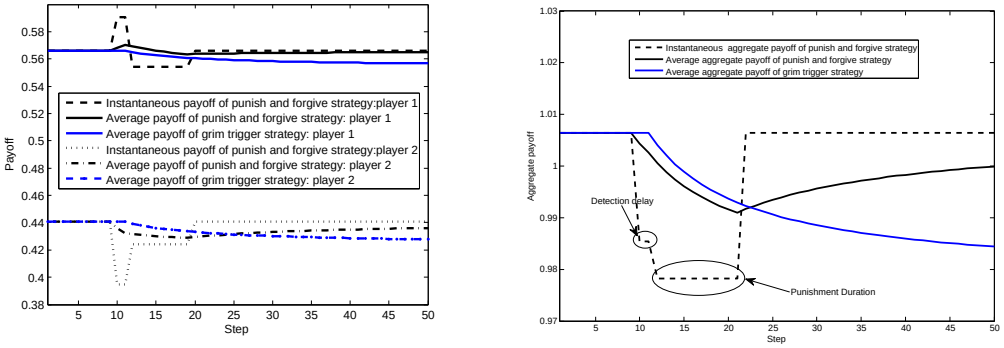
Figure 5.7: Punishment rule: Average punishment duration.

To analyze the performance of the punishment rule, I first compute the average punishment duration, T_p. Fig. 5.7 shows the variation of the average punishment duration as a function of the detection threshold, for different values of β. For $\beta = 0.001$

[5]Detection delay is defined as the number of steps required to detect a deviation *after* the deviation actually occurs.

and for $\beta = 0.01$, if the detection threshold is larger than 0.01, the punishment duration increases with the detection threshold. This can be justified by the following argument. As discussed before, the detection delay increases when the threshold increases. Moreover, the dishonest users gain benefits only before the cheating is detected. As a result, the average punishment duration also increases with increasing detection threshold, to void the effect of a larger detection delay. When the threshold is less than 0.01, variations of the punishment duration show a different and interesting behavior. we can observe that, for $\beta = 0.01$, the average punishment duration decreases until the threshold value reaches 0.01. Moreover, when the detection threshold is less then 0.01, the average punishment duration for $\beta = 0.01$ is larger than that for $\beta = 0.001$. This is in fact the inverse effect of that in the previously-discussed $th \geq 0.01$ scenario. This behavior can be explained based on the false alarm probability of the detection algorithm. Clearly, a punishment can also be triggered due to a false alarm. On one hand, from Fig. 5.6(a), when $\beta = 0.01$ and $th < 0.01$, the false alarm probability is non-zero. On the other hand, for $\beta = 0.001$ and $th < 0.01$, the false alarm probability is almost zero. Moreover, it is clear that the false alarm probability also decreases with th. Therefore, for $\beta = 0.01$, the average number of punishments (hence the average punishment duration) due to false alarm decreases with th. However, when $th > 0.01$, the average punishment duration increases with th, as the false alarm probability is almost zero and larger threshold values result in higher detection delays.

Fig. 5.8 plots the payoff of the toy model once a deviation from the optimal solution has taken place. As previously explained, I assume that player 1 cheats at 10^{th} round of Phase 2. I also select $\beta = 0.002$ and $th = 0.05$. For this particular case, the deviation is detected two steps later and hence the value of punishment



(a) Payoff of each player.

(b) Aggregate payoff of the system.

Figure 5.8: Punishment rule: Payoff for the toy model.

duration T_p is 8. Fig. 5.8(a) shows the individual payoffs of each player. In can be observed that player 1 gains an instantaneous payoff before cheating is detected. However, player 1's instantaneous payoff decreases during punishment duration, and the decrement is sufficient to reduce the average payoff to a value smaller than the average payoff obtained before cheating. Grim trigger, in comparison, decreases the payoffs of both the players forever. Both grim trigger and punish and forgive strategies can incentivize rational players to cooperate. However, from Fig. 5.8(b), it can be easily concluded that the punish and forgive strategy is more productive than the grim trigger strategy. The objective of any rational player is to maximize its average payoff over time. Thus, the proposed power control mechanism is cheat-proof if all participating players are rational.

5.9 Chapter Summary

I have studied the distributed power control problem for a co-channel deployed two-tier network with full-duplex small cells underlaying a macro cell. I have formulated the power control problem as a repeated game with imperfect public monitoring. The existence and uniqueness of Nash equilibrium for the stage game is established

and the perfect public equilibrium payoff set is characterized. I have proposed a distributed learning mechanism to achieve the Pareto optimal solution which, unlike the Nash equilibrium of the stage game, is efficient in the sense of maximizing the social welfare. I also have developed a deviation detection technique based on Page-Hinckley test to detect any cheating and presented a punishment policy to motivate cheating players to cooperate. It is also proven that the proposed algorithm results in a perfect public equilibrium operating point of the formulated repeated game. To this end, I have proposed a distributed cheat-proof power control mechanism that not only is efficient for small cell nodes, but also satisfies the SINR requirement of the macro user. Intensive numerical analysis also verified the algorithm's capability to prevent cheating if all small cell nodes (i.e., SBSs and small cell users) are rational and protect the macro user from harmful interference caused by the small cells.

Chapter 6

Conclusion and Future Direction

6.1 Conclusion

In order to support the constantly rising demand for wireless traffic, future wireless networks will be composed of thousands of different types of small cells. Due to the increased density of base stations, their random deployment and heterogeneity, traditional centralized network control and manual intervention will not be realistic for future wireless networks. Consequently, self-organization has been proposed as a significant feature for future heterogeneous wireless network, i.e., network nodes are expected to take individual decisions. Distributed resource allocation is one of the desirable feature for self-organizing networks. In this thesis, I proposed a framework for distributed resource allocation in self organizing small cell networks underlaying macro-cellular networks.

Game theory is a strong mathematical tool that can model and analyze distributed interactions among entities with conflicting interests. Specifically, in this thesis, I have used evolutionary games, mean field games and repeated games to model three different scenarios of distributed resource allocation in self-organizing small cell networks.

Moreover, stochastic geometry is used to model the spatial distribution of the network nodes and analyze the performance metrics of the network in order to derive payoff functions for the formulated evolutionary game and mean field game.

In summary, I have developed three distributed resource allocation techniques for self-organizing small cell networks underlaying macro-cellular networks. The proposed three resource allocation schemes are intended to be applied in three different scenarios (i.e., small cell networks with three different requirements). Fig. 6.1 provides a comparison and a summary on the three resource allocation paradigms that has been proposed in this thesis.

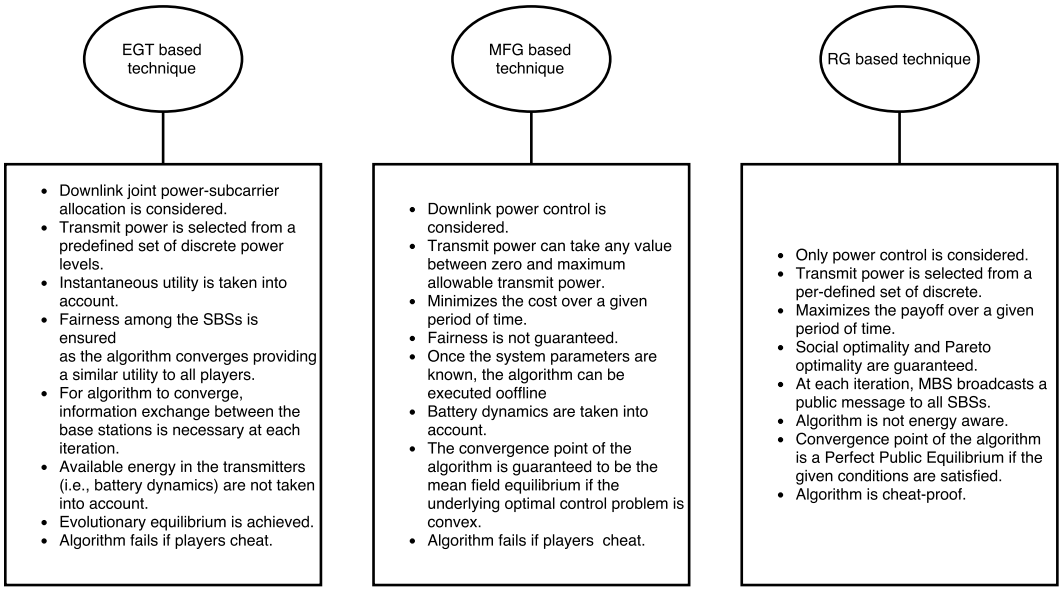


Figure 6.1: Summary of the resource allocation techniques proposed in this thesis.

In next section, a brief discussion on practical implementation of the proposed algorithms is given.

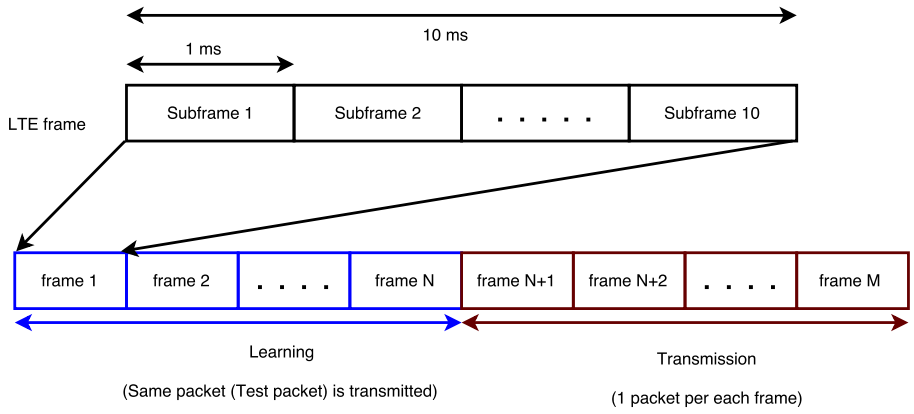


Figure 6.2: Implementation of Algorithms.

6.1.1 Implementation of Proposed Algorithms in Practice

Frame length of a general LTE system is 10ms; each frame is composed of 1ms sub-frames as shown in fig. 6.2. Each step of the proposed algorithms can be considered to be equivalent to one LTE frame. During learning phases of each algorithm, network nodes may transmit a test packet or same data packet as they do not achieve expected performances until the learning duration is over; Ex: If algorithm takes 50 iterations to converge a network node may transmit a test packet for 500 LTE frames, i.e., 0.5 s. After learning phase, actual transmission can take place, i.e., one packet per LTE frame as shown in fig. 6.2. However, algorithm has to be reset after a pre-defined time; Ex: 10 s (1000 LTE frames) as system parameters (such as channel gains, distance between transmitter and receiver, etc).

6.2 Future Research Directions

Some of the potential research directions in designing distributed resource allocation techniques for self-organizing small cell networks are outlined below.

1. *Multi radio access technology (Multi-RAT):*

In future networks, different radio access technologies (e.g., Wi-Fi, different types of small cells) are expected to be integrated in order to provide seamless service to the user. Access control between different technologies should be done in a self-organizing way to achieve the optimal performance. Access control is closely related with resource allocation, since access control algorithms generally allow a new user only if there are sufficient resources to serve that user. Therefore, distributed coordinated access control and resource allocation for Multi-RAT to provide seamless connection to the user is an important research direction.

2. *Signaling overhead-optimal performance trade-off:*

   There is always a trade-off between the signaling overhead and optimal performance of a network. A network may deliver optimal performance with complete information but the signaling cost for implementing such algorithms would be very high. On the other hand, the algorithms that rely on less information or incomplete information may deliver slightly degraded performance. Addressing this issue and quantifying the trade-off is significant in order to achieve near-optimal or optimal performance in self-organizing networks.

3. *Context-awareness:*

   Context awareness, which is a powerful feature in many intelligent systems, has recently been applied for enhancing the self-organizing features in small cell networks. The idea is to utilize the context information, i.e., information from the users' environment, behavior, and social media, to enhance the provision of services and applications. Context information can also be obtained by recently proposed mobile crowdsensing techniques. The algorithms should be devised considering the efficient exploitation of context aware information taken from

different sources. The reliability of the different information sources would also be an important issue.

Bibliography

[1] A. Goldsmith, *Wireless Communications*. Cambridge university press, 2005.

[2] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Communications Magazine, IEEE*, vol. 46, pp. 59–67, 2008.

[3] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: Key challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, vol. 8, 2013.

[4] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution towards 5g multi-tier cellular wireless networks: An interference management perspective," *Wireless Communications, IEEE*, vol. 8, p. 118127, June 2014.

[5] O. Aliu, A. Imran, M. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys and Tutorials*, vol. 15, pp. 336–361, 2012.

[6] H. Claussen, L. T. Ho, and L. G. Samuel, "An overview of the femtocell concept," *Bell Labs Technical Journal*, vol. 13, pp. 221–245, 2008.

[7] S. Haykin, "Cognitive dynamic systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007.*, vol. 4, 2007, pp. IV–1369–IV–1372.

[8] Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201–220, 2005.

[9] A. Spilling, A. Nix, M. Beach, and T. Harrold, "Self-organisation in future mobile communications," *Electronics and Communication Engineering Journal*, vol. 12, pp. 133–147, 2000.

[10] C. Prehofer and C. Bettstetter, "Self-organization in communication networks: principles and design paradigms," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 78–85, 2005.

[11] S. Hämäläinen, H. Sanneck, C. Sartori *et al.*, *LTE Self-Organising Networks (SON): Network Management Automation for Operational Efficiency.* John Wiley & Sons, 2012.

[12] I. Viering, M. Dottling, and A. Lobinger, "A mathematical perspective of self-optimizing wireless networks," in *IEEE International Conference on Communications*, 2009, pp. 1–6.

[13] M. Dirani and Z. Altman, "Self-organizing networks in next generation radio access networks: Application to fractional power control," *Computer Networks*, vol. 55, no. 2, pp. 431–438, 2011.

[14] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game theory in wireless and communication networks.* Cambridge University Press, 2012.

[15] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 13, pp. 372–395, 2011.

[16] M. Bennis, S. M. Perlaza, and M. Debbah, "Learning coarse correlated equilibria in two-tier wireless networks," in *IEEE International Conference on Communications (ICC)*, 2012, pp. 1592–1596.

[17] M. K. H. Yeung and Y.-K. Kwok, "A game theoretic approach to power aware wireless data access," *IEEE Transactions on Mobile Computing*, vol. 5, pp. 1057–1073, 2006.

[18] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1074–1084, 2006.

[19] H. Li, Y. Gai, Z. He, K. Niu, and W. Wu, "Optimal power control game algorithm for cognitive radio networks with multiple interference temperature limits," in *Vehicular Technology Conference, Spring 2008.* IEEE, 2008, pp. 1554–1558.

[20] P. Reichl, B. Tuffin, and R. Schatz, "Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience," *Telecommunication Systems*, pp. 1–14, 2011.

[21] L. Huang, Y. Zhou, X. Han, Y. Wang, M. Qian, and J. Shi, "Distributed coverage optimization for small cell clusters using game theory," in *IEEE Wireless Communications and Networking Conference (WCNC).* IEEE, 2013, pp. 2289–2293.

[22] S. Subramani, T. Basar, S. Armour, D. Kaleshi, and Z. Fan, "Noncooperative equilibrium solutions for spectrum access in distributed cognitive radio networks," in *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks,DySPAN 2008.* IEEE, 2008, pp. 1–5.

[23] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks*, 2012.

[24] S. Cheng and S. Lien, "On exploiting cognitive radio to mitigate interference in macro/femto heterogeneous networks," *IEEE Wireless Communications*, pp. 40–47, 2011.

[25] X. Kang, Y.-C. Liang, and H. K. Garg, "Distributed power control for spectrum-sharing femtocell networks using Stackelberg game," in *IEEE International Conference on Communications (ICC'11)*, 2011.

[26] S. Guruacharya and D. Niyato, "Hierarchical competition in femtocell-based cellular networks," in *IEEE Global Telecommunications Conference (GLOBECOM'10)*, 2010.

[27] C. Ko and H. Wei, "On-demand resource-sharing mechanism design in two-tier OFDMA femtocell networks," *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 1059–1071, 2011.

[28] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A Stackelberg game approach," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 538–549, Apr. 2012.

[29] M. Bennis and M. Debbah, "On spectrum sharing with underlaid femtocell networks," in *IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*, Sep. 2010.

[30] M. Bennis and S. Perlaza, "Decentralized cross-tier interference mitigation in cognitive femtocell networks," in *IEEE International Conference on Communications (ICC)*, 2011.

[31] L. Giupponi and C. Ibars, "Distributed interference control in OFDMA-based femtocells," *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1201–1206, Sep. 2010.

[32] L. Rose and S. Lasaulce, "Learning equilibria with partial information in decentralized wireless networks," *IEEE Communications Magazine*, pp. 136–142, Aug. 2011.

[33] S. Perlaza and H. Tembine, "On the fictitious play and channel selection games," in *IEEE Latin-American Conference on Communications (LATINCOM)*, 2010.

[34] E. J. Hong, S. Y. Yun, and D.-H. Cho, "Decentralized power control scheme in femtocell networks: A game theoretic approach," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2009, pp. 415–419.

[35] W. Ma, H. Zhang, W. Zheng, and X. Wen, "Differentiated-pricing based power allocation in dense femtocell networks," in *15th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. IEEE, 2012, pp. 599–603.

[36] V. Chandrasekhar, J. G. Andrews, T. Muharemovic, Z. Shen, and A. Gatherer, "Power control in two-tier femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 4316–4328, 2009.

[37] I. W. Mustika, K. Yamamoto, H. Murata, and S. Yoshida, "Potential game approach for self-organized interference management in closed access femtocell networks," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*. IEEE, 2011, pp. 1–5.

[38] L. Giupponi and C. Ibars, "Distributed interference control in ofdma-based femtocells," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2010, pp. 1201–1206.

[39] J.-H. Yun and K. G. Shin, "Adaptive interference management of OFDMA femtocells for co-channel deployment," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1225–1241, 2011.

[40] F. Pantisano, M. Bennis, W. Saad, M. Latva-aho, and R. Verdone, "Enabling macrocell-femtocell coexistence through interference draining," in *Wireless Communications and Networking Conference Workshops (WCNCW)*, 2012, pp. 81–86.

[41] M. Bennis, S. M. Perlaza, P. Blasco, Z. Han, and H. V. Poor, "Self-organization in small cell networks: A reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 3202–3212, 2013.

[42] S. Guruacharya, D. Niyato, M. Bennis, and D. Kim, "Dynamic coalition formation for network mimo in small cell networks," *IEEE Transactions on Wireless Communications*, 2013.

[43] J. W. Huang and V. Krishnamurthy, "Cognitive base stations in LTE/3GPP femtocells: a correlated equilibrium game-theoretic approach," *IEEE Transactions on Communications*, vol. 59, pp. 3485–3493, 2011.

[44] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 15, no. 3, pp. 996–1019, Third 2013.

[45] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[46] J. Weibull, *Evolutionary Game Theory*. M.I.T. Press, 1997.

[47] M. Khan, H. Tembine, and A. Vasilakos, "Evolutionary coalitional games: design and challenges in wireless networks," *IEEE Wireless Communications*, vol. 19, no. 2, pp. 50–56, 2012.

[48] M. Khan, H. Tembine, and V. Vasilakos, "Game dynamics and cost of learning in heterogeneous 4g networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 198–213, 2012.

[49] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 2008–2017, May 2009.

[50] H. Tembine, E. Altman, R. El-Azouzi, and Y. Hayel, "Evolutionary games in wireless networks," *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics*, vol. 40, pp. 634–46, Jun. 2010.

[51] K. Zhu, E. Hossain, and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Transactions on Mobile Computing*, pp. 1–14, 2013.

[52] M. Bennis and L. Giupponi, "Interference management in self-organized femtocell networks: the BeFEMTO approach," in *2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology (Wireless VITAE)*, Chennai, 2011, pp. 5–10.

[53] M. Simsek, A. Czylwik, and M. Bennis, "On interference management techniques in LTE heterogeneous networks," in *IEEE 21st International Conference on Computer Communications and Networks (ICCCN'12)*, 2012, pp. 1–5.

[54] S. Sirinivasa and M. Haenggi, "Modeling interference in finite uniformly random networks," in *International Workshop on Information Theory for Sensor Networks (WITS'07)*, 2007, pp. 1–12.

[55] S. Srinivasa and M. Haenggi, "Distance distributions in finite uniformly random networks: Theory and applications," vol. 59, pp. 940–949, 2010.

[56] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks, Volume I: Theory.* NOW Publishers, 2009.

[57] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modelling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, pp. 996–1019, Jan. 2013.

[58] M. Haenggi and J. G. Andrews, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, pp. 1029–1046, 2009.

[59] Y. Kuznetsov, *Elements of Applied Bifurcation Theory.* Berlin, Germany: Springer-verlag, 2004.

[60] D. Poole, *Linear Algebra: A Modern Introduction.* CengageBrain. com, 2010.

[61] F. Mazenc and S. Niculescu, "Lyapunov stability analysis for nonlinear delay systems," *Systems & Control Letters*, 2001.

[62] Y.-Y. Cao and P. Frank, "Stability analysis and synthesis of nonlinear time-delay systems via linear Takagi-Sugeno fuzzy models," *Fuzzy Sets and Systems*, vol. 124, pp. 213–229, Dec. 2001.

[63] F. Aurenhammer, "Voronoi diagrams - A survey of a fundamental geometric data structure," *ACM Computing Surveys*, vol. 23, pp. 345–405, Sep. 1991.

[64] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, 2007.

[65] M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled lqg problems with nonuniform agents: Individual-mass behavior and decentralized ε-nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, pp. 1560–1571, 2007.

[66] M. Huang, R. P. Malhame, and P. E. Caines, "Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, pp. 221–252, 2006.

[67] O. Guéant, J.-M. Lasry, and P.-L. Lions, "Mean field games and applications," in *Paris-Princeton Lectures on Mathematical Finance 2010*. Springer, 2011, pp. 205–266.

[68] H. Tembine, R. Tempone, and P. Vilanova, "Mean field games for cognitive radio networks," in *American Control Conference (ACC)*. IEEE, 2012, pp. 6388–6393.

[69] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, "Electrical vehicles in the smart grid: A mean field game analysis," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 1086–1096, 2012.

[70] F. Mériaux, V. Varma, and S. Lasaulce, "Mean field energy games in wireless networks," in *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2012, pp. 671–675.

[71] D. Bauso, B. M. Dia, B. Djehiche, H. Tembine, and R. Tempone, "Mean-field games for marriage," *PloS one*, vol. 9, p. e94933, 2014.

[72] M. Burger and J. M. Schulte, "Adjoint methods for hamilton-jacobi-bellman equations," 2010.

[73] F. Mériaux, S. Lasaulce, and H. Tembine, "Stochastic differential games and energy-efficient power control," *Dynamic Games and Applications*, vol. 3, pp. 3–23, 2013.

[74] F. Mériaux and S. Lasaulce, "Mean-field games and green power control," in *5th International Conference on Network Games, Control and Optimization (NetGCooP)*. IEEE, 2011.

[75] H. Tembine, S. Lasaulce, and M. Jungers, "Joint power control-allocation for green cognitive wireless networks using mean field theory," in *Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM), 2010 Proceedings of the Fifth International Conference on*. IEEE, 2010.

[76] J. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. Reed, "Femtocells: Past, present, and future," *IEEE Journal on Selected Areas in Communications,*, vol. 30, no. 3, pp. 497–508, April 2012.

[77] S. Srinivasa and M. Haenggi, "Modeling interference in finite uniformly random networks," in *International Workshop on Information Theory for Sensor Networks (WITS07)*, 2007, pp. 1–12.

[78] ——, "Distance distributions in finite uniformly random networks: Theory and applications," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 940–949, 2010.

[79] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, 2009, vol. 501.

[80] M. Huang, P. E. Caines, and R. P. Malhamé, "Uplink power adjustment in wireless communication systems: A stochastic control analysis," *IEEE Transactions on Automatic Control,*, vol. 49, pp. 1693–1708, 2004.

[81] M. Huang, P. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions," in *42nd IEEE Conference on Decision and Control, 2003. Proceedings.*, vol. 1. IEEE, 2003, pp. 98–103.

[82] T. Basar, G. J. Olsder, G. Clsder, T. Basar, T. Baser, and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1995, vol. 200.

[83] R. Bellman, "Dynamic programming and lagrange multipliers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 42, no. 10, p. 767, 1956.

[84] ——, *Dynamic Programming*. Princeton University Press, 1957.

[85] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer, 2008.

[86] H. Soner, "Controlled markov processes, viscosity solutions and applications to mathematical finance," *Viscosity Solutions and Applications*, pp. 134–185, 1997.

[87] L. C. Evans, "Partial differential equations," 1998.

[88] O. Guéant, "A reference case for mean field games models," *Journal de mathématiques pures et appliquées*, vol. 92, pp. 276–294, 2009.

[89] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume 1: THEORY*. Now Publishers Inc, 2009, vol. 1.

[90] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications,*, vol. 27, no. 7, pp. 1029–1046, 2009.

[91] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *Communications, IEEE Transactions on*, vol. 59, no. 11, pp. 3122–3134, 2011.

[92] R. Ahmed, "Numerical schemes applied to the burgers and buckley-leverett equations," Ph.D. dissertation, Department of Mathematics, University of Reading, 2004.

[93] M. Carter, "The fredholm alternative and the lagrange multiplier theorem," Supplementary note for foundations of mathematical economics, 2001.

[94] S. Guruacharya, D. Niyato, D. I. Kim, and E. Hossain, "Hierarchical competition for downlink power allocation in ofdma femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 1543–1553, April 2013.

[95] V. N. Ha and L. B. Le, "Fair resource allocation for ofdma femtocell networks with macrocell protection," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 3, pp. 1388–1401, March 2014.

[96] A. Abdelnasser, E. Hossain, and D. I. Kim, "Clustering and resource allocation for dense femtocells in a two-tier cellular ofdma network," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1628–1641, March 2014.

[97] D. T. Ngo, L. B. Le, T. Le-Ngoc, E. Hossain, and D. I. Kim, "Distributed interference management in two-tier cdma femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 979–989, March 2012.

[98] K. Zhu, E. Hossain, and A. Anpalagan, "Downlink power control in two-tier cellular ofdma networks under uncertainties: A robust stackelberg game," *Communications, IEEE Transactions on*, vol. 63, no. 2, pp. 520–535, Feb 2015.

[99] P. Semasinghe, E. Hossain, and K. Zhu, "An evolutionary game for distributed resource allocation in self-organizing small cells," *Mobile Computing, IEEE Transactions on*, vol. 14, no. 2, pp. 274–287, 2015.

[100] K. M. Thilina, H. Tabassum, E. Hossain, and D. I. Kim, "Medium access control design for full duplex wireless systems: Challenges and approaches," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 112–120, 2015.

[101] B. Di, S. Bayat, L. Song, and Y. Li, "Radio resource allocation for full-duplex ofdma networks using matching theory," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2014, pp. 197–198.

[102] M. Al-Imari, M. Ghoraishi, P. Xiao, and R. Tafazolli, "Game theory based radio resource allocation for full-duplex systems," in *IEEE 81st Vehicular Technology Conference (VTC Spring), 2015*, May 2015, pp. 1–5.

[103] G. J. Mailath and L. Samuelson, *Repeated Games and Reputations*. Oxford University Press, Oxford, 2006, vol. 2.

[104] Y. Xiao, J. Park, and M. van der Schaar, "Repeated games with intervention: Theory and applications in communications," *IEEE Transactions on Communications*, vol. 60, pp. 3123–3132, October 2012.

[105] M. Le Treust and S. Lasaulce, "A repeated game formulation of energy-efficient decentralized power control," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 2860–2869, Sept. 2010.

[106] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 517–528, April 2007.

[107] F. Shen and E. Jorswieck, "Universal non-linear cheat-proof pricing framework for wireless multiple access channels," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 1436–1448, March 2014.

[108] Y. Xiao and M. van der Schaar, "Dynamic spectrum sharing among repeatedly interacting selfish users with imperfect monitoring," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 1890–1899, November 2012.

[109] P. Semasinghe, K. Zhu, and E. Hossain, "Distributed resource allocation for self-organizing small cell networks: An evolutionary game approach," in *IEEE Globecom Workshops (GC Wkshps)*, Dec. 2013, pp. 702–707.

[110] P. Semasinghe and E. Hossain, "Downlink power control in self-organizing dense small cells underlaying macrocells: A mean field game," *IEEE Transactions on Mobile Computing*, vol. 15, no. 2, pp. 350–363, doi: 10.1109/TMC.2015.2 417 880, Feb. 2016.

[111] J. Nash, "Non-cooperative games," *Annals of Mathematics*, pp. 286–295, 1951.

[112] D. Fudenberg, D. Levine, and E. Maskin, "The folk theorem with imperfect public information," *Econometrica*, vol. 62, pp. 997–1039, 1994.

[113] J. R. Marden, H. P. Young, and L. Y. Pao, "Achieving pareto optimality through distributed learning," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 2753–2770, 2014.

[114] A. Menon and J. S. Baras, "Convergence guarantees for a decentralized algorithm achieving pareto optimality," in *American Control Conference (ACC)*. IEEE, 2013, pp. 1932–1937.

[115] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice Hall Englewood Cliffs, 1993, vol. 104.

[116] D. V. Hinkley, "Inference about the change-point in a sequence of random variables," *Biometrika*, vol. 57, pp. 1–17, 1970.

[117] D. T. Hoang, X. Lu, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Applications of repeated games in wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, pp. 2102–2135, 2015.

[118] G. Andrienko, N. Andrienko, M. Mladenov, M. Mock, and C. Poelitz, "Extracting events from spatial time series," in *Information Visualisation (IV), 2010 14th International Conference*. IEEE, 2010, pp. 48–53.

[119] C. Xu, M. Sheng, X. Wang, C. X. Wang, and J. Li, "Distributed subchannel allocation for interference mitigation in OFDMA femtocells: A utility-based learning approach," *IEEE Transactions on Vehicular Technology*, vol. 64, pp. 2463–2475, June 2015.

Appendix A

A.1 Laplace transform of the aggregate interference

Let

$$I_1 = \int_{x=0}^{\infty} \mathbf{E}_p \mathbf{E}_h [1 - e^{-vphx^{-\alpha}}] \; x \; \mathrm{d}x. \tag{A.1}$$

Since the integration can be interchanged with the expectation,

$$I_1 = \mathbf{E}_p \mathbf{E}_h \int_{x=0}^{\infty} \left(1 - e^{-vphx^{-\alpha}}\right) \; x \; \mathrm{d}x. \tag{A.2}$$

Substituting $y = vphx^{-\alpha}$, I_1 can be written as

$$\begin{aligned} I_1 &= \mathbf{E}_p \mathbf{E}_h \int_{y=0}^{\infty} (1 - e^{-y}) \left(\frac{1}{\alpha}\right) y^{\frac{-2}{\alpha}-1} (vph)^{\frac{2}{\alpha}} \; \mathrm{d}y \\ &= \mathbf{E}_p \left(p^{\frac{2}{\alpha}}\right) \mathbf{E}_h \left(h^{\frac{2}{\alpha}}\right) \frac{v^{\frac{2}{\alpha}}}{\alpha} \int_{y=0}^{\infty} \frac{1 - e^{-y}}{y^{\frac{2}{\alpha}+1}} \; \mathrm{d}y. \end{aligned} \tag{A.3}$$

Let $I_2 = \int_{y=0}^{\infty} \frac{1-e^{-y}}{y^{\frac{2}{\alpha}+1}} \; \mathrm{d}y$. By performing integration by parts on I_2, we have

$$I_2 = \frac{\alpha}{2} \Gamma \left(1 - \frac{2}{\alpha}\right). \tag{A.4}$$

Hence,

$$I_1 = \mathbf{E}_p\left(p^{\frac{2}{\alpha}}\right)\mathbf{E}_h\left(h^{\frac{2}{\alpha}}\right)\frac{v^{\frac{2}{\alpha}}}{2}\Gamma\left(1-\frac{2}{\alpha}\right). \tag{A.5}$$

A.2 Average SINR of a Generic User

For an interference-limited network (i.e., $N_0 = 0$) and assuming path-loss exponent $\alpha = 4$, we can derive $\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right]$ as in (A.6).

$$\begin{aligned}&\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \\ &\int_{t=0}^{\infty}\exp\left\{-\pi\lambda_m\sqrt{p_m}\mathbf{E}[\sqrt{h_{m,k}^{(n)}}]\sqrt{v}\Gamma\left(0.5\right)\right\} \\ &\exp\left\{-\pi\lambda_s^{(n)}\mathbf{E}[\sqrt{p_s}]E[\sqrt{h_{i,k}^{(n)}}]\sqrt{v}\Gamma\left(0.5\right)\right\}\,\mathrm{d}t.\end{aligned} \tag{A.6}$$

Since $h_{m,k}^{(n)}$ and $h_{i,k}^{(n)}$ are i.i.d. and they are exponentially distributed with mean μ,

$$\mathbf{E}\left[\sqrt{h}\right] = \int_{h=0}^{\infty}\sqrt{h}\mu e^{-\mu h}\mathrm{d}h. \tag{A.7}$$

By substituting $x = \mu h$, we have

$$\begin{aligned}\mathbf{E}\left[\sqrt{h}\right] &= \frac{1}{\sqrt{\mu}}\int_{x=0}^{\infty}\left(\sqrt{x}e^{-x}\right)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{\mu}}\Gamma\left(1.5\right) = \frac{1}{2}\sqrt{\frac{\pi}{\mu}}.\end{aligned} \tag{A.8}$$

Also, after substituting the value of $\Gamma(0.5) = \sqrt{\pi}$ and $v = \frac{\mu t r_s^{\alpha}}{p_l}$ in (A.6), we have ,

$$\begin{aligned}
&\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \\
&\int_{t=0}^{\infty} \exp\left\{-\frac{\pi^2 \lambda_m \sqrt{p_m} \sqrt{t} r_s^2}{2\sqrt{p_l}}\right\} \exp\left\{-\frac{\pi^2 \lambda_s^{(n)} \mathbf{E}[\sqrt{p_s}] \sqrt{t} r_s^2}{2\sqrt{p_l}}\right\} \mathrm{d}t \\
&\int_{t=0}^{\infty} \exp\left\{-\frac{A}{2\sqrt{p_l}} \left(\lambda_m \sqrt{p_m} + \lambda_s^{(n)} \mathbf{E}[\sqrt{p_s}]\right) \sqrt{t}\right\} \mathrm{d}t, \qquad \text{(A.9)}
\end{aligned}$$

where $A = \pi^2 r_s^2$.

The expression in (A.9) has the following form:

$$\begin{aligned}
I_3 &= \int_0^{\infty} e^{-k\sqrt{t}} \mathrm{d}t \\
&= \frac{2}{k^2} \int_0^{\infty} x e^{-x} \mathrm{d}x = \frac{2}{k^2} \Gamma(2) = \frac{2}{k^2}, \qquad \text{(A.10)}
\end{aligned}$$

where the second step follows from the change of variables by substituting $k\sqrt{t} = x$. Hence, (A.9) can be re-written as follows:

$$\mathbf{E}\left[\mathrm{SINR}_l^{(n)}\right] = \frac{8 p_l}{A^2 \left(\lambda_m \sqrt{p_m} + \lambda_s^{(n)} \mathbf{E}[\sqrt{p_s}]\right)^2}. \qquad \text{(A.11)}$$

Appendix B

B.1 Proof of Lemma 5.5.1

Let $\mathcal{J}$ be any arbitrary subset of base stations, i.e., $\mathcal{J} \subset \mathcal{K}$. The action profiles of $\mathcal{J}$ and $\mathcal{J}^c$ are given by $\boldsymbol{a}_J$ and $\boldsymbol{a}_{-J}$, respectively. Consider any base station $i \notin \mathcal{J}$. The payoff of i is a function of the corresponding received SINRs, i.e., $\pi_i = \mathcal{F}\left(\text{SINR}_i^{UL}, \text{SINR}_i^{DL}\right)$. Also, from (5.1) and (5.2), it is evident that both uplink and downlink SINRs are functions of uplink and downlink transmit powers of all network nodes, i.e., $\pi_i = \mathcal{F}\left(\boldsymbol{a}_J \cup \boldsymbol{a}_{-J}\right)$. Therefore, any change in $\boldsymbol{a}_J$ will change π_i. Therefore, game $\mathcal{G}_s$ is interdependent.

B.2 Proof of Theorem 5.5.1

When all players play a finite game repeatedly and follow the learning rules described in **Algorithm 4**, the dynamics of the system induces a perturbed Markov process with a finite state space $\mathcal{X} = \mathcal{A} \times \mathcal{MO}$, with $\mathcal{MO} = \left(\{C, D\}^{|\mathcal{K}|}\right)$ being the set of mood profiles of the system. In [113] and [119], it is shown that when any interdependent finite game is repeated following **Algorithm 4** and $\nu \longrightarrow 0$, a certain state $\bar{x} = [\bar{\boldsymbol{a}} \quad \bar{\mathfrak{m}}] \in \mathcal{X}$ is stochastically stable if and only if: i) $\bar{\boldsymbol{a}}$ is the social optimal action profile and ii) $\bar{\mathfrak{m}}_k = C, \quad \forall k \in \mathcal{K}$. The game $\mathcal{G}_s$ is finite. It is also interdependent

according to *Lemma 5.5.1*. In the game $\mathcal{G}_r$, the players repeat the stage game $\mathcal{G}_s$ infinitely many times. Therefore, if all players in $\mathcal{G}_r$ update their actions and moods according to **Algorithm 4** at every stage, it's asymptotic convergence point is the social optimal action profile, $\boldsymbol{a}^{SO}$. Moreover, the convergence of **Algorithm 4** is guaranteed when the experimenting rate (ν) is time-varying and goes to zero with the iterations [114]. Therefore, to guarantee the asymptotic convergence of the learning model, we select $\nu = \frac{1}{\sqrt{t}}$ in **Algorithm 4**.