RESOURCE ALLOCATION FOR COGNITIVE RADIO NETWORKS WITH COOPERATIVE RELAYING

by

Zheng Wang
A Dissertation
Submitted to the Graduate Faculty of George Mason University
In Fulfillment of The Requirements for the Degree of Doctor of Philosophy
Electrical and Computer Engineering

Committee:

Dr. Brian L. Mark, Dissertation Director
Dr. Yariv Ephraim, Committee Member
Dr. Kai Zeng, Committee Member
Dr. Song Min Kim, Committee Member
Dr. Monson H. Hayes, Department Chair
Dr. Kenneth S. Ball, Dean, Volgenau School of Engineering

Date: Fall Semester 2020
George Mason University
Fairfax, VA
Resource Allocation for Cognitive Radio Networks with Cooperative Relaying

A dissertation submitted in fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Zheng Wang
Master of Science
Tsinghua University, 2012
Bachelor of Science
Tsinghua University, 2008

Director: Dr. Brian L. Mark, Professor
Department of Electrical and Computer Engineering

Fall Semester 2020
George Mason University
Fairfax, VA
I dedicate this dissertation to my father Fuling and my mother Guifen for their love and support during my PhD study, for which I will always cherish.
I wish to express my sincerest gratitude to my advisor Dr. Brian Mark for all his mentoring in the research work during my PhD study, and his guidance in my paper writing and most of all, allowing me to think and work in my own way. Without his help, I would not get every solid step to the completion of my dissertation.

I would like to thank Dr. Bernd-Peter Paris, Dr. Bijan Jabbari and Dr. Tarek Lahlou for their collaborations on my teaching work. I would like to thank Dr. Zhu Han from University of Houston for his help on my research. I would also like to thank the committee members, Dr. Yariv Ephraim, Dr. Song Min Kim and Dr. Kai Zeng for their suggestions over my research work. I really benefited a lot from their expertise and insightful feedback.

I would like to thank Dr. Daniel Sun and Dr. Joseph Bruno for their help and discussions in the early years of my PhD program. I would also thank our group members Hanke Cheng and Massieh Kordi Boroujeny for valuable discussions in research and the collaborative atmosphere in the lab.

In the end, I would like to thank my family who always stand with me during my PhD study. I would also like to thank all my friends and supporters who have made this happen.

My dissertation work was supported in part by the U.S. National Science Foundation under Grants CNS-1205453 and CNS-1421869.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>Abstract</td>
<td>x</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Problem Statement</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1 Exclusion Zone Approximation</td>
<td>2</td>
</tr>
<tr>
<td>1.2.2 Spectrum Resource Allocation Problem Under Cooperative Relaying</td>
<td>4</td>
</tr>
<tr>
<td>1.2.3 Hierarchical Game for Power Control in Cooperative Relaying</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Contributions of the Dissertation</td>
<td>9</td>
</tr>
<tr>
<td>2 Background and Literature Review</td>
<td>12</td>
</tr>
<tr>
<td>2.1 Gaussian Random Field Model (GRFM)</td>
<td>12</td>
</tr>
<tr>
<td>2.1.1 Overview of GRFM</td>
<td>13</td>
</tr>
<tr>
<td>2.1.2 Theoretical Analysis to GRFM</td>
<td>14</td>
</tr>
<tr>
<td>2.2 Stackelberg Game</td>
<td>16</td>
</tr>
<tr>
<td>2.2.1 Stackelberg Game for Spectrum Resource Allocation</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2 Extension of Stackelberg Game</td>
<td>18</td>
</tr>
<tr>
<td>2.3 Mean Field Game</td>
<td>19</td>
</tr>
<tr>
<td>2.3.1 Introduction to Differential Game</td>
<td>20</td>
</tr>
<tr>
<td>2.3.2 Mean Field Game for Power Control in Cooperative Relaying</td>
<td>21</td>
</tr>
<tr>
<td>3 Gaussian Random Field Model for Exclusion Zone Approximation</td>
<td>24</td>
</tr>
<tr>
<td>3.1 System Model</td>
<td>25</td>
</tr>
<tr>
<td>3.1.1 Network Model</td>
<td>25</td>
</tr>
<tr>
<td>3.1.2 Aggregate Interference</td>
<td>28</td>
</tr>
<tr>
<td>3.2 GRFM Framework</td>
<td>29</td>
</tr>
<tr>
<td>3.2.1 Overview of GRFM</td>
<td>30</td>
</tr>
<tr>
<td>3.2.2 GRFM in Cognitive Network Models</td>
<td>32</td>
</tr>
<tr>
<td>3.3 Estimation of Exclusion Zone</td>
<td>34</td>
</tr>
<tr>
<td>3.4 Experimental Results</td>
<td>38</td>
</tr>
</tbody>
</table>
3.4.1 Received Signal Power Approximation .................................. 39
3.4.2 Aggregate Interference Power Approximation .......................... 41
3.4.3 Exclusion Zone Approximation ........................................... 45
3.5 Conclusion ............................................................................. 46

4 Stackelberg Game Theoretic Approach to Resource Allocation with Cooperative Relaying .......................................................... 48
4.1 System Model ........................................................................ 50
4.1.1 Network Model ............................................................... 50
4.1.2 Decode-and-Forward (DF) Relay Transmission Scheme ........ 53
4.1.3 Direct Transmission Scheme ............................................. 56
4.2 Stackelberg Game Analysis ..................................................... 57
4.2.1 Extended Stackelberg Game Framework ............................. 57
4.2.2 Analysis to Case 1 ............................................................ 63
4.2.3 Analysis for Case 2 ........................................................... 65
4.2.4 Optimal DF Transmission Strategy ................................... 71
4.2.5 Optimal Direct Transmission (DT) Strategy ...................... 72
4.3 Joint Allocation and Scheduling Algorithm ............................... 72
4.4 Experimental Results and Analysis ......................................... 77
4.4.1 Parameter Settings .......................................................... 77
4.4.2 Performance Study of DF Transmission ............................. 77
4.4.3 Experimental Study of Hybrid Scheduling Algorithm .......... 80
4.5 Conclusions .......................................................................... 85

5 Hierarchical Stackelberg / Mean Field Game for Distributed Power Control in Dense Cooperative Relaying Network .......................... 87
5.1 System Model ....................................................................... 88
5.1.1 Network Model ............................................................... 88
5.1.2 State Space .................................................................... 93
5.1.3 Cost Function ............................................................... 94
5.2 Stackelberg Game Analysis ..................................................... 95
5.2.1 Stackelberg Game Framework for D2D Coordination .......... 95
5.2.2 Follower Analysis ........................................................... 98
5.2.3 Leader Analysis ............................................................. 101
5.2.4 Stackelberg Game Nash Equilibrium ............................... 103
5.3 Mean Field Game for Power Control ...................................... 104
5.3.1 Optimal Power Control Problem ..................................... 105
5.3.2 Mean Field Approximation ........................................... 106
5.3.3 Mean Field Game Analysis .......................................... 109
5.4 Distributed Mean-Field Game Power Control Policy .............. 110
  5.4.1 FPK Equation Formulation ........................................ 111
  5.4.2 Lagrange Relaxation to HJB ...................................... 112
6 Conclusions and Future Work .......................................... 115
  6.1 GRFM for Exclusion Zone Approximation ......................... 115
  6.2 Stackelberg Game Theoretic Approach to Resource Allocation with Cooperative Relaying ......................................................... 116
  6.3 Hierarchical Stackelberg / Mean Field Game for Distributed Power Control in Dense Cooperative Relaying Networks .................. 117
Bibliography ......................................................................... 119
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Simulation Parameters</td>
<td>40</td>
</tr>
<tr>
<td>4.1 Simulation Parameters</td>
<td>79</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 The Stackelberg game interactions between the leader and the follower for optimal transmission strategy.</td>
<td>18</td>
</tr>
<tr>
<td>2.2 The mean field game framework for power control in cooperative relaying.</td>
<td>23</td>
</tr>
<tr>
<td>3.1 Geometry of cognitive network.</td>
<td>27</td>
</tr>
<tr>
<td>3.2 SUs selected for GRFM approximation.</td>
<td>34</td>
</tr>
<tr>
<td>3.3 GRFM approximation for received signal power.</td>
<td>41</td>
</tr>
<tr>
<td>3.4 GRFM approximation for aggregate interference power.</td>
<td>43</td>
</tr>
<tr>
<td>3.5 Relationship between threshold $I_{th}$ and exclusion zone radius $R_0$.</td>
<td>45</td>
</tr>
<tr>
<td>4.1 Cognitive Radio Network Model with Downlink Channel Sharing.</td>
<td>52</td>
</tr>
<tr>
<td>4.2 Illustration of the four cases in DF cooperative relaying.</td>
<td>62</td>
</tr>
<tr>
<td>4.3 Device Deployment in the Proposed CRN Model.</td>
<td>78</td>
</tr>
<tr>
<td>4.4 CDF for logarithm of capacity ratio under different values of $\beta$.</td>
<td>80</td>
</tr>
<tr>
<td>4.5 SU link rate distribution under different $\beta$, $D_{sr,1} = D_{\text{max}}/2$.</td>
<td>81</td>
</tr>
<tr>
<td>4.6 Accumulative sum rate versus the number of scheduled SU pairs.</td>
<td>82</td>
</tr>
<tr>
<td>4.7 Average sum rate versus TTIs, $\delta_{df} = \delta_{dt} = 0$.</td>
<td>83</td>
</tr>
<tr>
<td>4.8 Average sum rate versus TTIs regarding fairness ($\beta = 1$).</td>
<td>85</td>
</tr>
<tr>
<td>4.9 Average sum rate versus TTIs regarding fairness ($\beta = 1, M = 6$).</td>
<td>86</td>
</tr>
<tr>
<td>5.1 Dense D2D cooperative relaying network model.</td>
<td>90</td>
</tr>
<tr>
<td>5.2 The interactions between the coupled HJB and FPK equations.</td>
<td>110</td>
</tr>
</tbody>
</table>
Abstract

RESOURCE ALLOCATION FOR COGNITIVE RADIO NETWORKS WITH COOPERATIVE RELAYING
Zheng Wang, PhD
George Mason University, 2020
Dissertation Director: Dr. Brian L. Mark

Resource allocation is a key issue to efficiently utilize the spectrum in wireless communication systems. A promising technology for efficient spectrum usage, cognitive radio (CR), allows the secondary users (SUs) to make use of the spectrum that has been assigned to the licensed users, also known as primary users (PUs). The spectrum can be accessed by SUs as long as the PUs are idle or the interference caused by the SUs is below some threshold. Therefore, the management of such spectrum resources needs to be carefully investigated. Allocation is a problem in resource management which aims to find a way for the SUs to access proper licensed channels based on certain criteria. Meanwhile, the geographical distribution of the system users leads to spatial reuse and primary system protection, e.g., exclusion zone approximation problem. As the number of devices significantly grows, the coordination among different devices in such networks needs to be carefully considered as well.

In this dissertation, we investigate the following issues in spectrum resource allocation for cognitive radio networks (CRNs). For the exclusion zone approximation problem, theoretical analysis has been proposed in prior work to find the range of the zone. For the spectrum allocation problem, game theoretic approaches have been proposed so that the
optimal channels could be allocated to the SUs. All the above approaches could be further extended to more realistic and complicated scenarios such as relay transmission. In our work, we first adopt a Gaussian random field model (GRFM) to approximate the radius of an exclusion zone to protect PUs from SU interference. This approach allows us to find the coverage of exclusion zone using limited observations of signal power and interference received by the SUs. Second, we propose a novel Stackelberg-game based framework to model the behavior of SUs and PUs in cooperative relaying, and design hybrid scheduling algorithms to solve the resource allocation problem in CRNs. Numerical results show that significant improvement in terms of CRN system capacity can be achieved under cooperative relaying. Third, we propose a hierarchical Stackelberg game/mean field game framework to coordinate the behavior of the cognitive devices and control transmit power under cooperative relaying, and develop a finite difference method for determining the optimal power control scheme in this setting.
1.1 Motivation

The spectrum resource is playing a key role in wireless communication networks due to its scarcity. As the demand for spectrum resource driven by wireless applications grows, efficient utilization and management of the spectrum are required. One major branch of enhancing the efficiency of spectrum utilization is cognitive radio (CR), which allows secondary users (SUs) to access the spectrum that has been licensed to the primary users (PUs) based on the detection of licensed spectrum usage. In practical scenarios, many licensed bands are underutilized. Meanwhile, the impact that the SUs bring to the primary system should be considered as well. Therefore, further improvement of spectrum utilization will rely on carefully designed resource allocation and management schemes.

We first consider the spatial reuse of spectrum resource in cognitive radio networks. The SUs operating in a cognitive system can be regarded as signal and interference detectors. As long as the SUs are far away enough from the primary system, they can share the licensed spectrum of the primary system, which is referred to as spatial reuse. The PUs could use the licensed spectrum simultaneously if the interference from the SUs is tolerable. The geographical relationship among the users is critical in spatial reuse, as they may interfere each other when transmitting. Hence, the concept of exclusion zone has been proposed in the literature, and a theoretical analysis to characterize an exclusion zone was provided in [84]. Yet this analysis is impractical, as the network model it adopted is not realistic. To the best of our knowledge, very few works have addressed the issue of estimating the exclusion zone coverage in practical scenarios.

Secondly, the temporal reuse of spectrum resource in cognitive radio networks is considered, which leads to a spectrum resource allocation problem. For spectrum allocation,
the SUs could sense the channel first and obtain channel parameters. In existing work, game-theoretic approaches have been proposed to tackle resource allocation problem based on sensing results. A meaningful direction of research is to study the spectrum allocation problem under cooperative relaying, as cooperative relaying among SUs could extend transmission range and promote network flexibility compared to direct transmission. To study this, a game-theoretic framework for cooperative relaying in cognitive radio network needs to be investigated.

Thirdly, we consider spectrum resource reuse in a dense device-to-device (D2D) network which is assisted by cognitive radio techniques. The D2D users share the spectrum of macrocell user equipments and communicate directly with each other by consuming battery energy. As the energy is always limited, a power control scheme needs to be carefully considered. In a pioneering work [94], a game-theoretic power control scheme for dense D2D networks was investigated. To extend transmission range and increase system capacity, cooperative relaying could be adopted in D2D networks. Yet power control schemes in dense D2D networks with cooperative relaying have not been well investigated. Power control is challenging as both interactions within a D2D link and the interactions among different D2D links should be taken into account simultaneously. In this dissertation, we aim to design a power control scheme for D2D cooperative relaying network.

1.2 Problem Statement

1.2.1 Exclusion Zone Approximation

A cognitive radio network consists of primary users which have priority to access the licensed spectrum, and secondary users which may utilize the same spectrum under some predefined access rules [84] [87] [80]. The primary users, including primary transmitters and receivers, can be regarded as a primary system which requires a certain level of communication quality guarantee. The SUs, mostly mobile devices, often exist in large numbers within the cognitive network. Accordingly, a high demand of available spectrum resources
emerges in order to meet the communication and interaction requirements of the SUs. Contemporary measurements show that spectrum holes are common in the area covered by the same spectrum band [84]. As a consequence, underutilization of spectrum resource is widely observed in network activities and dynamic spectrum sharing by detecting temporarily idle spectrum is thus adopted to improve spectrum utilization. Since the activities of the SUs may cause harmful interference to the primary system, the access rules are needed to allow the SUs to dynamically access the spectrum while maintaining the degradation of the primary node transmission below some threshold [84] [87] [80].

As long as the SUs are far away enough from the primary devices, the interference the primary system received can be maintained at an acceptable level. Thus, the SUs may reuse the spectrum in those remote areas, which are also known as spatial spectrum holes. From the perspective of the primary system, it is protected by deactivating the SUs when they enter its neighboring area. In this case, the major source of the interference introduced can be suppressed. Such area defined above is called an exclusion zone, and is widely adopted in scenarios such as ad-hoc [82] [38] and heterogeneous networks [45] [60]. To react to the exclusion zone, it is crucial to find out if the SUs are inside the exclusion zone or not, and the primary system should be aware of the exclusion zone coverage as well. Therefore, the boundary of the exclusion zone needs to be carefully identified.

Existing approaches attempt to perform theoretical analysis by fixing the parameter settings of the primary and SUs [84] [87] [80] [82]. Assuming omnidirectional antenna for the transmitters, the basic assumption becomes that the exclusion zone has the geometry of a disk [84] [87] [82]. By applying performance constraints (e.g., aggregated interference constraint or signal-to-interference-plus-noise constraint), the radius of the disk can be found. However, the theoretical results are derived with relatively simple network geometry under a static environment, in which all the SUs and primary users are assumed to be stationary. A more complicated network geometry or the lack of system knowledge (e.g. the propagation environment [83]) may prevent us from obtaining the exact analytical coverage of the exclusion zone. Furthermore, the stochastic property and the mobility of
the SUs also make such an attempt more difficult. In practical scenarios, the cognitive network is working under time-varying characteristics and parameters, and the exclusion zone defined above is thus dynamic as well. In other words, a given area may be covered by the exclusion zone at some time instant, while the SUs inside it may transmit during other time periods. Thus, some alternative approach to effectively capture the exclusion zone is required.

In this work, a more realistic model for characterizing exclusion zone coverage is considered instead of an idealistic model. Theoretical analysis has been performed in the idealistic model to get an estimate of exclusion zone [84], yet it can not be adopted in practical scenarios. From the perspective of the SUs, the SUs may exchange data with their neighbors when PUs are idle. As each SU is often aware of the information whether its vicinity is within the exclusion zone or not, an approach for estimating the exclusion zone with limited data exchange is desired. Given the geographical locations of the primary system and the cognitive band, the existing analytical approach calculates the power and the interference received by the primary system. However, the analytical solutions of such analytical results are often hard to derive. To make approximations, the boundary conditions need to be relaxed, which brings negative impacts to approximation accuracy. Furthermore, the cognitive band has been assumed to be a continuous medium, which is impractical as well. In practical scenarios where the SUs are discrete and mutually independent, SU observations consist of the signal and interference they receive from the PUs through the wireless propagation environment. Therefore, our goal in this problem is to design an exclusion zone approximation approach which makes use of the SU observations above.

1.2.2 Spectrum Resource Allocation Problem Under Cooperative Relaying

With the rapid growth of wireless communications, spectrum scarcity has become an important issue. Cognitive radio is a promising technology that aims to improve spectrum utilization by allowing unlicensed SUs (SUs) to access and share licensed spectrum that
is not temporarily being used by the primary users (PUs). Cognitive radio networks (CRNs) require suitable power control and resource allocation schemes to avoid harmful interference to the PUs and promote system capacity. With proper interference coordination, the SUs can efficiently access the available spectrum, while higher system capacity can be achieved without compromising the PUs.

The SUs form a CRN either via a base station or via direct communications, whereby the SUs establish direct links or adopt other SUs as relays to transmit information. Communications using direct links among SUs could be an effective alternative to an infrastructured network, where the SUs form an underlay ad-hoc network with multiple link requirements. Two SUs with certain traffic demand between them is known as a SU transmitter-receiver pair, and they access the spectrum provided by the PUs when transmitting signals. Meanwhile, as the SU pairs might be far apart from each other, relay nodes are often necessary to support transmission. Among the relaying schemes, amplify-and-forward (AF) and decode-and-forward (DF) are the two main schemes that could be adopted by relay nodes to forward the signal. In AF, the noisy signal received by the relay is directly amplified and then forwarded to the destination SU. The destination SU can merge two or more independently transmitted noisy signals to make a decision and derive the ultimate signal. On the other hand, the noisy signal received by the relay node in DF transmission is decoded before it is forwarded to the destination. The noise in the received signal is removed in this step and the message is re-encoded before it is relayed to the next hop.

For a SU underlay network with multiple SU pairs and multiple relay nodes, we assume a number of channels have already been licensed to a group of PUs, while the SUs are trying to access these channels for signal transmission. The following problems needs to be answered: Which relay nodes should be selected to support SU transmission? What strategy should be used in relay transmission? What spectrum needs to be accessed? What is the appropriate transmit power in SU relay transmission? All the above questions can be summarized as spectrum resource allocation in SU cooperative relaying. Clearly, the topology of the
network, the deployment of devices and the available spectrum resource all need to be taken into account. In addition, spectrum resource allocation becomes increasingly complex as the communication demand between SU pairs grows. Hence, spectrum resource allocation in such SU underlay networks emerges as a major research challenge. To answer the proceeding problems, two approaches need to be considered:

1. The first approach is to formulate the behaviors given the SUs and the PUs which provide the spectrum resources. Game theory has been proposed as an effective mathematical tool to model the interactions among wireless network devices and predict their future actions \[21, 67, 75\]. In CRNs, the SUs interact with the PUs to compete for access to spectrum resources. By assuming that the PUs and SUs are independent rational players, their own transmission strategies can be derived. Such transmission strategies are known as the Nash equilibrium (NE) of the game, in which no player gains more revenue by making unilateral change of its own strategy \[63, 66\]. Among a variety of game models, Stackelberg games have been intensively studied \[22, 30, 86\]. A Stackelberg game is a hierarchical game with a leader and a follower. The leader announces the price for the spectrum resource it would like to sell, and the follower decides the amount of resource it would like to buy subsequently. Both the leader and the follower aim to maximize their utility functions respectively by selecting their own transmission strategies. Yet existing Stackelberg frameworks model only the interactions between PUs and SUs in a direct transmission scheme, whereas the potential advantages of cooperative relaying are ignored in this context \[86, 105\]. The Stackelberg game has also been applied to model the interactions between different hops in cooperative relaying in \[6, 54\], but spectrum allocation is not considered in these models. To address the above concerns, a game framework is required to characterize the interactions between SUs under cooperative relaying and the PUs.

2. The second approach aims to formulate an optimization problem in terms of some certain network performance (i.e., maximize the sum rate) and find the optimal or sub-optimal solution by designing a scheduling algorithm \[64\]. In CRNs, the SU pairs could
either adopt relay nodes for transmission or use direct transmission based on the geographical location of devices and channel parameters, and each SU pair needs to be mapped to appropriate PUs when accessing the spectrum resource. The mapping problem needs to be solved in a centralized manner with all the related system information, as a decentralized approach often requires impractical data exchange among all the devices [97]. When the CRN scale grows, it is often hard to find the global optimal mapping scheme for all the PUs and the SU pairs. Therefore, a scheduling algorithm taking into account both direct transmission and cooperative relaying needs to be designed. Meanwhile, time-domain scheduling should be taken into account to maintain fairness among different devices. Yet very few works have investigated hybrid scheduling for SU cooperative relaying.

By means of two approaches above, our goal is to design a complete scheme which allocates channel resource based on the formulated behavior of the SUs in cooperative relaying. The new scenario of cooperative relaying is a meaningful extension, as the cognitive radio network could have higher system capacity and better transmission flexibility compared to SU direct transmission. Apparently, the allocation process becomes more complicated, as the relay nodes would also share the spectrum. This will lead to the increase of algorithm complexity, and needs to be carefully considered as well.

1.2.3 Hierarchical Game for Power Control in Cooperative Relaying

In recent years, device-to-device (D2D) communications have been widely used as the underlay to cellular primary systems to enhance spectrum efficiency. Due to the proximity between D2D devices, direct communication is available by reusing spectrum licensed to the primary devices, with the assistance of cognitive radio technologies. Yet, the transmission range of a D2D pair is usually small under co-channel interference from other D2D devices. Basically, D2D transmission qualities degrade significantly as D2D transmitter and receiver are located further apart from each other. Therefore, the performance requirements are often not met, which is usually essential in practical scenarios. In previous work [6], cooperative relaying communication has been shown to support long range transmission and
promote higher system capacity by utilizing relays to forward signals. For D2D networks, the idea of cooperative relaying transmission can be naturally borrowed to extend D2D transmission range and improve system performance. The relays nodes are D2D devices as well and they can work as both transmitter and receiver when receiving signals and forwarding them to the desired destinations. Hence, D2D links over longer transmission ranges could be established in this way.

In relay transmission, two major schemes, namely amplify-and-forward (AF) and decode-and-forward (DF), are frequently used. In AF scheme, the signal received by any relay node is directly amplified without removing the noise or interference inside it, and then forwarded to the next hop. Apparently, the noise will be amplified and forwarded as well so that the interference will accumulate at the D2D receiver. Though the receiver can receive and combine signals sent from both transmitter and relay nodes, the devices that are far from the receiver contributes very little to the link performance. Besides, the relay in a certain D2D link is more likely to introduce higher interference to other D2D links if the noise and interference are not removed from the signal. Yet in contrast, DF scheme aims to decode the signal received by relay node and remove the interference inside it, and re-encode the signal before transmitting to the next hop. In this way, the receiver will not receive accumulative interference from the relay node, while the other D2D users will suffer less interference from the current D2D link. Therefore, we consider DF scheme in this problem for D2D cooperative communications.

Mutual interference that exists among different D2D devices is another key issue that worthwhile to consider. When the number of D2D links grows significantly, channel reuse is adopted by various D2D links due to the scarcity of spectrum resource. Hence, the interference brought by dense deployed devices also exacerbates network performance degradation when the number of D2D links is large. Since D2D devices are usually powered by batteries with finite energy supply, power control is essential to extending device lifetime and improving overall system performance. Hence, in this work, we propose a power control scheme for D2D pairs under cooperative relaying scheme so as to mitigate their mutual interference.
and promote system performance. Power control is a dynamic process that determines the energy consumption for each D2D user over time slots by taking the interactions of different D2D users into account. As signal transmission takes the major portion of energy consumption in D2D devices, the power control of both the D2D transmitter and the relay node need to be considered, which makes the problem more complicated.

To effectively characterize the dynamics in power control process, the interactions between different devices must be carefully studied and modeled. Basically, two kinds of interactions exist among different D2D devices. On the one hand, for a certain D2D pair that uses relay transmission, intra-tier interactions exist between different hops. In this work, we assume that all D2D pairs use two hops in relay transmission. Thus, the interactions between the first hop and the second hop in the same D2D transmission pair needs to be carefully coordinated (intra-tier coordination), i.e., the relay node needs to select an appropriate transmit strategy based on the transmit strategy adopted in the first hop. On the other hand, inter-tier interaction exist among devices in different D2D links, in which the D2D links would reuse the same channel and cause interference to each other. To tackle this problem, the behaviors of different D2D links need to be coordinated as well, i.e., inter-tier coordination.

To the best of our knowledge, very few works have considered power control for dense D2D cooperative relaying networks. Therefore, we design a novel framework which incorporates both intra-tier coordination and inter-tier coordination for power control regarding all the concerns above. As game-theoretic approaches have been successfully adopted in prior work for D2D device coordination \[70, 92\], some of the key ideas from these works can naturally be adopted in our work.

### 1.3 Contributions of the Dissertation

The contributions of this dissertation are outlined as follows:

In Chapter 2 we provide a review of the relevant literature and study the background of the given problems. Theoretical models and key techniques for exclusion zones in cognitive
radio networks, spectrum resource allocation, Stackelberg game and mean-field game for power control are also discussed in Chapter 2.

In Chapter 3, a simulation based approach incorporating Gaussian random field model (GRFM) for approximating the coverage of exclusion zone is proposed. Since the existing approach is analytical with impractical assumptions, the results are less accurate and hard to derive in practical scenarios. Furthermore, the boundaries of the exclusion zone are relaxed when theoretical analysis is adopted for approximation. In our approach, we use a more realistic discrete model in which a finite number of SUs are mutually independent. The transmit and interference power of the primary system are regarded as spatial distributed Gaussian random fields, and the SUs are regarded as spatial samplers. By using the samples taken from the SUs as input to GRFM, the radius of the exclusion zone can be derived. Detailed analysis as well as the simulations are proposed to demonstrate the performance of the GRFM approach.

In Chapter 4, a Stackelberg game theoretic approach for spectrum resource allocation is investigated. In prior work [86], a Stackelberg game framework are successfully adopted in spectrum resource allocation problems in cognitive radio networks. For more complicated scenarios such as relay cooperative relaying, an extended Stackelberg game framework is needed. In this chapter, we propose a novel Stackelberg game model for cooperative relaying. Based on Stackelberg game outcomes, a hybrid scheduling algorithm regarding fairness for spectrum resource allocation is also proposed. The simulation results demonstrate the performance of the Stackelberg game framework and the scheduling algorithm.

In Chapter 5, we propose a hierarchical Stackelberg / mean field game framework for power control in cognitive D2D network. In prior work [70, 92], mean field game has been successfully applied to power control problems in D2D network. In this chapter, we consider a dense D2D network under cooperative relaying and use a hierarchical game framework to characterize the interactions among D2D devices. In the hierarchical game framework, Stackelberg game aims to find out the optimal transmission strategy by modeling the interactions within a D2D pair, while a mean field game is used to model the interactions
between a certain D2D pair and all other D2D pairs. A distributed power control scheme is later formulated based on the theoretical analysis of the proposed hierarchical game.

In Chapter 6, our research on GRFM for exclusion zone approximation, Stackelberg game for spectrum resource allocation and hierarchical Stackelberg/mean field game for power control in D2D cooperative relaying, together with the corresponding results, is summarized. Future directions of research are discussed as well.
Chapter 2: Background and Literature Review

In this chapter, we discuss the background and key techniques related to exclusion zone approximation, spectrum resource allocation in cooperative relaying, and power control in dense D2D cooperative relaying network. The exclusion zone problem has been theoretically studied and the approximation for exclusion zone radius is presented in pioneer research. Although the results appears to be promising, the scenario is impractical due to the unrealistic assumptions of the network infrastructure. Therefore, we adopt a practical scenario instead and Gaussian random field model (GRFM) as well to approximate the scenario. The background of GRFM is discussed in details in this chapter. For spectrum resource allocation problem in cooperative relaying, game theoretic approaches have been successfully adopted in existing work. In this chapter, we briefly review the representative game frameworks for network behavior modeling, with an emphasize on Stackelberg game framework as well as its extension that would be adopted in our research. For power control in D2D cooperative relaying, we first study the background of the power control problem in D2D networks. Mean field game in existing work shows its good performance in characterizing the power control problem as a dynamic process, while Stackelberg game could be adopted to model the intra-tier interactions. Hence, we also study the properties of mean field game and Stackelberg game in this power control problem.

2.1 Gaussian Random Field Model (GRFM)

In exclusion zone approximation problems, we consider a learning framework called Gaussian Random Field (GRF), which makes use of data samples and provides an effective way to construct local metamodels for functions with unknown landscapes. These metamodels are called Gaussian Random Field Model (GRFM) [1]. As an efficient learning tool
for function approximation, GRFM is selected based on the following considerations: (1) The exclusion zone is an area such that the SUs inside it are silenced to avoid harmful interference when the primary system is operating. This implies that the interference the active SUs cause to the primary system should not exceed some threshold defined by SNR or outage probability requirements. The transmit signal and interference power can be considered as 2-dimensional spatial functions in the network area; (2) The SUs can make measurements and derive the corresponding information (e.g., received power and interference) by sensing the propagation environment. Thus, these SUs could act as spatial samplers in the network model. With all the predictions above, the exclusion zone coverage can be derived consequently. Yet to the best of our knowledge, very few work has been done using this approach. The detailed introduction to GRFM is given below.

2.1.1 Overview of GRFM

Prior to introducing GRFM, we briefly review the approach adopted in [84] in approximating the exclusion zone radius. A primary transmitter is assumed to be at the center, while the cognitive band consists of secondary users are assumed to be a continuous medium. The primary receiver is receiving the lowest level of tolerable SINR when it is located on the border of the exclusion zone. Therefore, we can calculate the interference the primary receiver receives given the position of the cognitive band and the primary receiver. However, this scenario is impractical and it is hard to derive a closed-form of the integral in most cases. Thus, an alternative approach is required.

Clearly, the power transmitted by the PUs and the SUs could be detected at any point as long as it is located in the network area. Hence, the spatial distribution of the transmit power provides multi-domain information in the real-world network environment, i.e., radio environment map (REM) [24, 40, 68, 96, 103], which could help identify the coverage of the exclusion zone. Existing work such as neighborhood interpolation [65, 72] and Kriging-based approaches [40, 68, 95, 99] have been adopted to derive the power distribution in the network area. Yet these approaches have not been adopted in exclusion zone approximation.
In this work, we consider GRFM, which can better utilize the characteristics of signal power distribution. GRFM can be viewed as a function predictor that applies a Gaussian random field to determine values of the function at some desired points in the parameter space where the actual function values are unknown or hard to derive. The inputs of a GRFM are the coordinates of a set of sample points as well as their corresponding function values. GRFM can be interpreted as a stochastic learning process using a group of training data from the sampling points. The predictive distribution of the function value is Gaussian, and the quality of the model and its computational cost are determined by the training data as well as the distribution of the sample points.

2.1.2 Theoretical Analysis to GRFM

Let the function to be predicted be \( f(x) \) within a \( d \)-dimensional parameter space. We would like to predict the function value \( f(x^*) \) at a desired point \( x^* \). Under the GRFM, the predicted value is modeled as a Gaussian random variable with mean \( \mu \) and variance \( \sigma^2 \).

Calculation of the predicted value requires the key concept of correlation. Assume we have two arbitrary sample points \( x_i \) and \( x_j \) in the parameter space. Intuitively, \( x_i, x_j \) are highly correlated when they are sufficiently close to each other. Therefore, a Gaussian correlation function between \( x_i \) and \( x_j \) is defined as follows:

\[
\text{Corr}(x_i, x_j) = \exp \left( -\sum_{l=1}^{d} \theta_l |x_{il} - x_{jl}|^2 \right),
\]

where \( \theta_l \) is the correlation parameter in the \( l \)th dimension which determines the rate of change in the correlation as \( x_i \) approaches \( x_j \) in dimension \( l \). If the function parameter space is isotropic, \( \theta_l \) will be equal to some constant for \( l = 1, \ldots, d \); otherwise, the \( \theta_l \) values will be mutually different. The approach to determine \( \theta_l \) can be found in \[58\].

More generally, given a set of \( n \) sample points \( x_1, \cdots, x_n \) and the corresponding output
function values $y_1, \cdots, y_n$, we define the correlation matrix $R$ by:

$$R_{ij} = \text{Corr}(x_i, x_j), \quad i, j = 1, \cdots, n,$$  \hspace{1cm} (2.2)

where $R_{ij}$ is the $(i, j)$th entry of matrix $R$. The likelihood function $L$ of the observed data is given by

$$L = \frac{1}{(2\pi\sigma^2)^n|R|^{1/2}} \exp\left(-\frac{(y - \mu 1)^T R^{-1} (y - \mu 1)}{2\sigma^2}\right),$$  \hspace{1cm} (2.3)

where $1$ is a column vector of ones, $y = \text{col}(y_1, \cdots, y_n)$ is the column vector of all outputs, and the superscript $^T$ denotes matrix transpose. By setting its first order derivative to 0, i.e., $\frac{\partial L}{\partial \mu} = 0$, in maximizing the likelihood function (2.3), optimal estimates of the output mean and variance can be derived as follows:

$$\hat{\mu} = \frac{1^T R^{-1}y}{1^T R^{-1} 1}, \quad \hat{\sigma}^2 = \frac{1}{n} (y - 1\hat{\mu})^T R^{-1} (y - 1\hat{\mu}).$$  \hspace{1cm} (2.4)

By applying the GRFM, the prediction of the function value $\hat{y}(x^*)$ at the desired point $x^*$ yields

$$\hat{y}(x^*) = \hat{\mu} + r^T R^{-1} (y - 1\hat{\mu}),$$  \hspace{1cm} (2.5)

where $r = \text{col} [\text{Corr}(x^*, x_1), \cdots, \text{Corr}(x^*, x_n)]$ is a column vector containing Gaussian correlation coefficients between the desired point $x^*$ and all sample points $x_1, \cdots, x_n$. Furthermore, the uncertainty of the prediction, denoted by the predicted variance of the function value at $x^*$, is given by

$$\hat{\sigma}^2(x^*) = \sigma^2 \left[ 1 - r^T R^{-1} r + \frac{(1 - r^T R^{-1} r)^2}{1^T R^{-1} 1} \right].$$  \hspace{1cm} (2.6)

which is the best linear unbiased predictor [46, 58]. With the GRFM prediction of the power
distribution, the exclusion zone could be approximated.

2.2 Stackelberg Game

The approaches to spectrum allocation problem in cognitive communication systems depends on a well designed mapping scheme between SUs and PUs by observing their interactions. Game theory has been proposed to offer an effective mathematical tool to model such interactions among network devices and give predictions to their future actions. In cognitive radio networks, the SUs interact with the PUs to compete for necessary spectrum resource. By assuming that the PUs and SUs are independent rational players, their own strategies can be derived under some reasonable game frameworks. Game theoretic approaches such auction-based game [8, 41, 79, 88, 90], coalition game [4, 31, 55, 73] and Stackelberg game [7, 85, 86, 89, 98] have been already successfully applied in resource allocation in recent years. Among all these game models, Stackelberg game is the most widely adopted game theoretic approaches for tackling spectrum resource allocation problems and has proved its efficiency in pioneer research. Yet, existing Stackelberg game frameworks often model the interactions in SU direct transmission, or the interactions among different hops in relay transmission, while the interactions between multiple PUs and multiple SUs are rarely considered. Hence, the Stackelberg game needs to be extended to fit cooperative relaying. We first discuss the Stackelberg game model for SU direct transmission, then extend the framework to cooperative relaying scenario.

2.2.1 Stackelberg Game for Spectrum Resource Allocation

Stackelberg game has been introduced to model the behavioral relationship between the leader who aims to make gains from its follower, and the follower seeks to buy resource from the leader. In a leader-follower pair, the PU plays the role of leader as it provides the channel resource, and the SU plays the role of follower which seeks to access the channel. Consider downlink transmission, we have a PU P and a SU pair (including a transmitter $S_T$ and a receiver $S_R$) in a leader-follower pair, in which $S_T$ aims to directly transmit to $S_R$. 
We assume that the original utilities of PU P and the SU pair are denoted by $u_p$ and $u_s$, respectively. The SU transmit power is denoted by $p_s$. The leader has the right to charge the follower and make profits, while the follower needs to pay the price for resources. After bargaining, the profit will be added to the original leader utility, i.e.,

$$u_l(\alpha, p_s) = u_p + \alpha \beta p_s g_{sp} \quad (2.7)$$

where $u_l(\alpha, p_s)$ is the leader utility, $\alpha > 0$ is the charging price, $\beta$ is a scale factor that determines the revenue gained by the leader, and $g_{sp}$ is the channel coefficient between . The objective is to maximize the leader’s utility by asserting the optimal price, i.e.,

$$\max u_l(\alpha, p_s), \quad \text{s.t. } \alpha > 0. \quad (2.8)$$

Accordingly, by subtracting the cost it has to pay for buying spectrum resource from the original follower utility, the follower utility yields:

$$u_f(\alpha, p_s) = u_s - \alpha p_s g_{sp} \quad (2.9)$$

The optimization objective for follower utility is to find its maximum value by finding a proper value of $p_s$, i.e.,

$$\max u_f(\alpha, p_s), \quad \text{s.t. } p_{\min} \leq p_s \leq p_{\max}. \quad (2.10)$$

In the proposed Stackelberg game framework, the leader moves first and the follower moves based on its observations. The leader knows ex ante that the follower is observing its action and sets the optimal price. The follower selects an appropriate SU transmit power. A pair of strategies $(\alpha, p_s)$ reaches Nash equilibrium if no unilateral deviation in such strategies leads to higher utility for either the leader or the follower. At this time, the Nash equilibrium is a stable outcome of the Stackelberg game.

The interactions between the leader and the follower in the proposed Stackelberg game
Figure 2.1: The Stackelberg game interactions between the leader and the follower for optimal transmission strategy.

2.2.2 Extension of Stackelberg Game

Stackelberg game approach can well coordinate the behavior of the PU and the SU pair in direct transmission so that the leader-follower pair do not have an insufficient outcome. However, if cooperative relaying is adopted in SU transmission, the interactions between the PUs, the SUs and the relay node become more complicated. We assume that the SU relay transmission allows up to two hops (i.e., at most one relay node is adopted by each SU pair). An intuitive way to tackle this problem is to extend the Stackelberg game framework in direct transmission to fit the cooperative relaying scenario. As the SU transmitter and the relay node would each access a channel when transmitting signals, we assume that two independent PUs as a group play the role of the leader, while a SU pair and the
corresponding relay node as a group play the role of the follower.

As described in the original Stackelberg game framework, the leader charges the follower some fee for accessing its spectrum resources. The leader decides the optimal price and the follower choose its transmit power based on the price to maximize its utility. Such operations will lead to the Nash equilibrium under Stackelberg game framework. In the leader-follower pair which consists multiple PUs and multiple SUs, the leader aims to sell its spectrum resources and the follower pays the price when accessing the channels in relay transmission. Without loss of generality, we assume that the follower pays different prices for accessing the two channels. Hence, the optimal price for accessing the two channels needs to be searched in a two dimensional space, which can be denoted by a strategy pair \((\alpha_1, \alpha_2)\). The follower’s transmission strategy is given by the transmit power of the SU transmitter and the relay node, and denoted by \((p_s, p_r)\). The optimal values for \((p_s, p_r)\) could be derived under the leader’s strategy \((\alpha_1, \alpha_2)\).

### 2.3 Mean Field Game

Mean field game is a game framework which characterizes the crowd dynamics in a decision making process, where the ‘small’ players exist in large quantities \([3, 19, 25, 51]\). By ‘small’ players, we mean that each player has very little influence on the overall system performance. During this process, each player reacts to the statistical behavior of all other players and performs the same type of strategy in its optimal decisions. The decision making process refers to the power control strategy in D2D cooperative relaying in our work. In D2D power control problems, the decision making process is often modeled as a differential game which involves the state space of the devices such as energy and cumulative interference \([2, 70, 94, 101]\). The key concept of the mean field is the statistical distribution of the device state space. By incorporating the concept of mean field into the differential game for power control, the power control problem of the devices can be formulated as a mean field game. In the following subsections, we first give fundamental introduction to differential game, then discuss the mean field game in power control for D2D cooperative relaying.
2.3.1 Introduction to Differential Game

Differential game belong to a subclass of dynamic games called state space game. By adopting the tools and models in optimal control theory, differential game extends static non-cooperative continuous-kernel game into dynamic environment, which is appropriate for power control problems in D2D networks [37]. In a differential game, a set of state variables are introduced to describe the state space of the system at any particular time instant during play. Meanwhile, each player interacts with the system and the payoff-relevant influence of past events is adequately summarized into the state variables. The state variables subsequently determine the optimal control problem in the differential game.

Consider a simple scenario with only one player, the objective for the optimal control problem is to maximize its payoff (or utility) function over a certain period of time, say, $[0, T]$. The standard model in the optimal control theory can be formulated as a partial differential equation, which is given by:

$$\dot{s}(t) = f(s(t), u(t)),$$  \hfill (2.11)

where $s(t)$ denotes the state of the system, $u(t)$ is the control function at time instant $t$. $\dot{s}(t)$ is the first order derivative of the state $s(t)$, and $f(s(t), u(t))$ formulates the dynamics of the state space. The payoff function $J[u(.)]$ of the player in the differential game consists of two parts and is defined as:

$$J[u(.)] = \int_{t=0}^{T} g(s(t), u(t))dt + h(s(T)),$$  \hfill (2.12)

where $g(s(t), u(t))$ is known as the running payoff, while $h(s(t))$ is the terminal payoff. $J[u(.)]$ is the ultimate payoff of the player which needs to be optimized. Hence, we define a value function $v(s, t)$ as:

$$v(s, t) = \arg \max_{u(.)} J[u(.)],$$  \hfill (2.13)
which is the value at the final state at time instant $T$. According to the Bellman’s optimality principle \cite{37}, the Hamilton-Jacobi-Bellman (HJB) equation for the value function $v(s, t)$ is given by:

$$
\partial_t v(s, t) + \max_{u(.)} \{ f(s, u) \cdot \nabla_s v(s, t) + g(s, u) \} = 0,
$$

By solving the HJB function, the optimal $v^*(s, t)$ that maximizes the value function could be derived. This is also the Nash equilibrium of the proposed differential game associated with the player. If the system contains multiple players, each player would share a state that obeys the control function from all the players, and solve its own optimal control problem by optimizing the corresponding payoff function. Assume that there are $N$ players, each player has a control function $u_i(t)$ and a terminal payoff $h_i(s(T))$. The payoff function for the $i$th player is given by:

$$
J_i[u(.)] = \int_{t=0}^{T} g(s(t), u_1(t), ..., u_N(t)) dt + h_i(s(T)),
$$

which can be optimized using the approach above.

### 2.3.2 Mean Field Game for Power Control in Cooperative Relaying

Power control for D2D cooperative relaying can be viewed as an non-ccooperative optimal control problem. Thus, the state space in the game is defined as the residual energy of the D2D devices, as the devices are supported by batteries. As time evolves continuously, the evolution of the state space over time can be modeled as a set of differential equations under differential game framework. By defining an appropriate payoff function for each
player, the value function in (2.13) could be optimized as the residual energy runs out during the time period. In pioneer work, as D2D pairs exist in large quantities in the network, each D2D device contributes very little interference to other devices, and the impact of all D2D pairs to a certain D2D node could be modeled as a mean field \[69,92\]. Yet very little work has been done to incorporate power control in D2D cooperative relaying scenario.

Naturally, each D2D pair would decide its own transmission strategy under cooperative relaying and interact with other D2D pairs during power control process. For a certain D2D pair, the transmitter seeks for the aid of the relay node to forward its signal to the receiver. Hence, the interactions between the transmitter and the relay node could be modeled as a Stackelberg game, in which the leader (i.e., relay node) sells its transmit power in forwarding the signal, while the follower (i.e., transmitter) pays the price. The Nash equilibrium of the Stackelberg game denotes the optimal transmission strategy. The D2D pair then interacts with the statistical impact of all other D2D pairs (i.e., the mass) by observing the mean field and performing its own strategy to optimize the payoff in (2.12). Therefore, the mean field game framework for cooperative relaying is hierarchical, which is depicted in 2.2.

As all D2D devices are supported by batteries with limited energy supply, the power control for the transmitter and the relay node needs to be considered simultaneously. Hence, the state of the system is defined as the residual energy of D2D transmitters and relay nodes. The mean field \(m(s,t)\) denotes the distribution of the device residual energy at time instant \(t\) and satisfies:

\[
\int m(s,t) \, ds = 1, \quad \forall t \in [0, T].
\] (2.17)

Meanwhile, the dynamics of the mean field \(m(s,t)\) according to the player’s action is modeled by a Fokker-Planck-Kolmogorov (FPK) equation, which is:

\[
\partial_t m(s,t) + \nabla (m(s,t) \cdot \partial_t s(t)) = 0.
\] (2.18)
HJB and FPK equations are coupled in the mean field game. Therefore, by jointly solving the equations (2.15) (2.17) (2.18), the power control strategy for cooperative relaying can be derived under mean field game framework.
Chapter 3: Gaussian Random Field Model for Exclusion Zone Approximation

In a cognitive radio network, the primary transmit system, including primary transmitters and primary receivers, can communicate using predefined spectrum over a certain area while no secondary users inside it may transmit. Such an area is called an exclusion zone (EZ). The primary system is protected by banning the secondary users (SUs) inside the exclusion zone so that the interference caused by the secondary users outside the zone is under a predefined tolerance level.

In this chapter, we focus on the problem of exclusion zone approximation in a more practical scenario. A novel approach called Gaussian random field model (GRFM) is adopted to implement such approximation in simulations. We assume a cognitive radio network in which all the devices are deployed in the same plane. The primary transmitter is located in the center of the network and the secondary nodes are uniformly distributed outside the exclusion zone. As the primary transmit power and the aggregated interference can both be regarded as a 2-dimensional spatial field, GRFM is applied to approximate the equivalent transmit power as well as the aggregated interference at desired points using measurement data from the secondary users. Consider the worst case in which the primary receiver is at the edge of the exclusion zone. The spatial samples of received power and interference at the secondary nodes are used to construct GRFMs. Compared to the theoretical analysis on the coverage of exclusion zone, the GRFM network model consists of a finite number of independent secondary users rather than a continuous cognitive band. The observations taken by the SUs are used as the inputs to GRFM, while the outputs are the approximations of signal and interference power at certain locations. The exclusion zone coverage is thus derived using the GRFM approximations given above. The results allow us to further
study the impact of network parameters and trade-off relations among them on the exclusion zone coverage based on GRFM. We further derive the radius of the exclusion zone with other system parameters as well. Furthermore, GRFM may act as an effective framework to detect exclusion zones in practical scenarios even if no analytical definitions of exclusion zone are available.

The remainder of the chapter is given as follows: Section 3.1 gives a brief introduction to exclusion zone and discusses the system model adopted in GRFM exclusion zone approximation. In Section 3.2, we discuss the background of GRFM and describe the detailed framework in exclusion zone approximation. The simulation results for exclusion zone and related discussions are proposed in Section 3.4. The conclusions are given in Section 3.5.

3.1 System Model

In a cognitive radio system, primary users (PUs) licensed with certain spectrum resource have the highest priority to access to it, while the secondary users (SUs) need to access the same spectrum under some predetermined rules [80, 84, 87]. The PUs form a primary system that requires a certain level of communication quality guarantee. The SUs, mostly mobile devices, may constitute a large population within the cognitive radio network. Accordingly, the secondary system generates a high demand for available spectrum resources. Spectrum measurement studies show that spectrum resources are often highly underutilized by the primary system [84]. However, the activities of SUs will definitely cause interference to the primary system, which leads to the problem of protecting the primary system.

To tackle the problem given above, the concept of exclusion zone is briefly introduced first. The system model description for GRFM is given in the following subsections.

3.1.1 Network Model

Dynamic spectrum sharing is regarded as an approach to improving spectrum utilization by allowing SUs to sense the spectrum and make use of spectrum holes. Since the activities of SUs may cause interference to the primary system, access rules are needed to allow SUs to
dynamically access the spectrum while maintaining the interference to the primary system below a certain threshold [80, 84, 87]. If the SUs are sufficiently far from the PUs, the interference to the primary system can be maintained at an acceptable level. The areas in which SUs may reuse the spectrum are referred to as spatial holes. From the perspective of the primary system, protection from interference by SUs can be achieved by deactivating SUs when they enter a neighboring region. Such a region is referred to as an exclusion zone. By applying performance constraints, for example, an aggregate interference constraint or SINR constraint, the radius of the disk can be derived. The most widely adopted criterion is that the performance in terms of SINR is above a predefined threshold. The concept of exclusion zone has been widely adopted in scenarios such as ad hoc networks [38, 82] and heterogeneous networks [33, 45, 60]. In order to be effective, the boundary of the exclusion zone needs to be carefully identified. A number of approaches to determining exclusion zone boundaries have involved theoretical analysis by fixing the parameter settings of the primary and secondary systems [80, 82, 84, 87]. Under the assumption of omnidirectional antennas for the primary and secondary transmitters, the exclusion zone has the geometry of a disk [11, 26, 27, 82, 84, 87].

In practical scenarios, the characteristics of the cognitive radio network are often time-varying, such that the exclusion zone is dynamic over time. Thus, alternative measurement-based approaches are needed to effectively characterize the exclusion zone. Besides, unlike the abstract system model in which the SUs are assumed to be distributed spatially as a continuous medium [84], the SUs should be mutually independent from each other, while the signal and interference are received by the SUs individually.

We consider a system model of a cognitive radio network, similar to the model adopted in [84], consisting of a single primary transmitter, at least one primary receiver, and multiple SUs equipped with cognitive radios. The primary transmitter is assumed to be located in the center, while the primary receivers are located within a disk of radius $R_0$ meters, which defines the exclusion zone. PUs and SUs are all equipped with omnidirectional antennas. Furthermore, the SUs should be at least $\epsilon$ meters away from any of the primary receivers.
to ensure that the aggregate interference to these receivers does not become harmful. This *guard band* isolates the exclusion zone from the SUs. The SUs are distributed within an annular area, referred to as the *cognitive band*, with inner radius $R_0 + \epsilon$ and outer radius $R$. The system model is depicted in Figure 3.1.

The focus of this work is to determine the coverage of the exclusion zone subject to certain performance constraints. If the SUs are mobile or dynamically joining/leaving the cognitive band, the randomness of the interference will result in large deviations in the primary system performance. Hence, it is necessary to ensure that the outage probability is above some threshold and the primary receiver is able to operate normally. In our work, we assume a scenario in which the cognitive users are stationary and analyze the system performance. In this scenario, the area of the exclusion zone is such that the outage probability in terms of signal-to-interference ratio (SIR) of the primary receiver inside it, is
below some threshold. The extreme case is that the primary receiver on the boundary of
the exclusion zone satisfies such SIR constraint.

### 3.1.2 Aggregate Interference

Considering a certain primary receiver, the outage constraint must hold for the worst
case in which the receiver is located on the border of the exclusion zone and the \( \epsilon_p \) guard
band. Thus, the primary receiver is closest to the cognitive band and receives the maximum
amount of aggregate interference power from the SUs. For simplicity, we assume that the
receivers do not adopt multiuser detection. Thus, the aggregate interference to a given
primary receiver, denoted as \( I_0 \), comes from the transmit power of the SUs. We adopt the
simple path-loss channel model used in [84] given by:

\[
h(d) = \frac{A}{d^{\alpha/2}},
\]

where \( A \) is a frequency-dependent amplitude, \( d \) is the distance between the transmitter and
receiver, and \( \alpha \) is the path-loss exponent. For simplicity, we shall normalize \( A \) to 1. In
addition to path-loss, our system model accounts for the presence of shadowing.

We assume that the SUs are uniformly distributed over the geographical area covered by
the cognitive radio network, with density \( \lambda \), while all the SUs may operate concurrently. The
network model proposed in [84] assumes that the cognitive band around the exclusion zone
is a continuous medium that may transmit signals simultaneously. Thus, the aggregated
interference in such continuous case can be written as an integral of transmit power within
the cognitive band. The transmit power from either transmitter all satisfies the pathloss
property. Following the analysis in [84], the expected value of aggregated interference from
continuous cognitive band, \( E[I_0] \), has the following form:

\[
E[I_0] = \int_{R_0}^{R} \int_{0}^{2\pi} \frac{\lambda r P_{cn} dr d\theta}{(r^2 + R_0^2 - 2rR_0 \cos \theta)^{\alpha/2}}
\]
where $P_{cn}$ is the transmit power per unit area of the cognitive band, and $\alpha$ is the pathloss exponent. If $\alpha$ is not an even integer, the aggregate interference does not have a closed-form expression, and the value of the integral needs to be approximated numerically. This can be regarded as a baseline while comparing results with the GRFM approach. The definition only works within the $\epsilon_p$ guard band and the exclusion zone. The expected interference grows to infinity as the primary receiver approaches the cognitive band. However, the cognitive band in a practical network model consists of a finite number of SUs. Hence, the aggregate interference $I_0$ received by the primary receiver has the following form:

$$I_0 = \sum_{i=1}^{N} P_{su} \cdot |h(d_{i,P})|^2,$$

(3.3)

where $N$ is the total number of independent SUs, $P_{su}$ is the transmit power of each SU, and $d_{i,P}$ is the geographical distance from the $i$th SU to the primary receiver.

### 3.2 GRFM Framework

Gaussian Random Field (GRF) theory provides a powerful and efficient framework that allows the constructions of metamodels (a.k.a. Gaussian Random Field Model, GRFM) for predicting the values of an unknown objective function through a set of observations or measurement data. Under the basic assumption that the objective function is a sample of a Gaussian stochastic process, the value of the function at some certain point can be approximated through the observations in its vicinity. Apart from the predicted value, the GRFM also provides a meaningful measure of prediction uncertainty, i.e., variance, which directly indicates the quality of the GRFM. In continuous spatial functions, the function values between any pair of neighboring sample points are considered to be highly correlated. Thus, GRFM is appropriate to serve as a local model and can replace the precise objective function at certain desired points when sufficient knowledge of their vicinity is known. Intuitively, the closer a desired point is to the space where the function is sampled, the
more accurate the metamodel is considered to be. The diversity and the number of sample points also have an impact on the quality of GRFM, such that a large number of dense sample points would increase its accuracy.

In this section, we first give an overview of GRFM, which includes its theoretical background, then we present our proposed approach of exclusion zone estimation using GRFM approximations.

### 3.2.1 Overview of GRFM

In this section, we provide a detailed introduction to GRFM and key relevant formulas. A GRFM can be viewed as a function predictor that applies a Gaussian random field to determine values of the function at some desired points in the parameter space where the actual function values are unknown or otherwise hard to derive. The inputs of a GRFM are the coordinates of a set of sample points as well as their corresponding function values. Let the function to be predicted be \( f(x) \) within a \( d \)-dimensional parameter space. We would like to predict the function value \( f(x^*) \) at some desired point \( x^* \). Under the GRFM framework, the predicted value is modeled as a Gaussian random variable with mean \( \mu \) and variance \( \sigma^2 \).

Calculation of the predicted value requires the key concept of correlation. Assume we have two arbitrary sample points \( x_i \) and \( x_j \) in the parameter space. Intuitively, \( x_i, x_j \) are highly correlated when they are close enough to each other. Therefore, a Gaussian correlation function \( \text{Corr}(x_i, x_j) \) between \( x_i \) and \( x_j \) is defined as follows:

\[
\text{Corr}(x_i, x_j) = \exp \left( -\sum_{l=1}^{d} \theta_l |x_{il} - x_{jl}|^2 \right),
\]

where \( \theta_l \) is the correlation parameter in the \( l \)th dimension which determines the rate of change in the correlation as \( x_i \) approaches \( x_j \) in dimension \( l \). If the function parameter space is isotropic, \( \theta_l \) will be equal to some constant for \( l = 1, \ldots, d \); otherwise, the \( \theta_l \) values will be mutually different. Formulas to determine \( \theta_l \) can be found in [58], in which the values
for $\theta_l, l = 1, \ldots, d$ are optimized by maximizing the likelihood functions. More generally, given a set of $n$ sample points $x_1, \cdots, x_n$ and the corresponding output function values $y_1, \cdots, y_n$, we define the correlation matrix $R$ by

$$R_{ij} = \text{Corr}(x_i, x_j), \quad i, j = 1, \cdots, n,$$  

(3.5)

where $R_{ij}$ is the $(i, j)$th entry of matrix $R$. Additionally, the likelihood function $L$ of the observed data is given by

$$L = \frac{1}{(2\pi\sigma^2)^n|R|^{1/2}} \exp\left(-\frac{(y - \mu 1)^T R^{-1}(y - \mu 1)}{2\sigma^2}\right),$$  

(3.6)

where $1$ is a column vector of ones, $y = \text{col}(y_1, \cdots, y_n)$ is the column vector of all outputs, and the superscript $T$ denotes matrix transpose. By setting its first order derivative to 0, i.e., $\frac{\partial L}{\partial \mu} = 0$, in maximizing the likelihood function (3.6), optimal estimates of the output mean and variance can be derived as follows:

$$\hat{\mu} = \frac{1^T R^{-1} y}{1^T R^{-1} 1},$$  

(3.7)

$$\hat{\sigma}^2 = \frac{1}{n} (y - 1\hat{\mu})^T R^{-1} (y - 1\hat{\mu}).$$  

(3.8)

By applying the GRFM framework, the prediction of the function value $\hat{y}(x^*)$ at the desired point $x^*$ yields:

$$\hat{y}(x^*) = \hat{\mu} + r^T R^{-1}(y - 1\hat{\mu}),$$  

(3.9)

where $r = [\text{Corr}(x^*, x_1), \cdots, \text{Corr}(x^*, x_n)]^T$ is a column vector containing Gaussian correlation coefficients between the desired point $x^*$ and all sample points $x_1, \cdots, x_n$. Furthermore, the uncertainty of the prediction, denoted by the predicted variance of the function
value at $x^*$, is given by

$$
\hat{\sigma}^2(x^*) = \hat{\sigma}^2 \left[ 1 - r^T R^{-1} r + \frac{(1 - r^T R^{-1} r)^2}{1^T R^{-1} 1} \right].
$$

(3.10)

This predictor has been proved to be the best linear unbiased predictor \[58\]. GRFM estimation can be regarded as a stochastic learning process using a group of training data from the set of sampling points, and the predictive distribution of the function value satisfies $f(x) \sim N(\hat{y}(x), \hat{\sigma}^2(x))$, which is a Gaussian distribution. The quality of the model and the computational cost are both determined by the training data as well as the distribution of the sample points. Intuitively, if the number of sampling points is large, the model is considered to be more accurate. However, this requires more computational resources and may conflict with real-time requirements. On the other hand, using fewer sample points accelerates the computation process, yet sacrifices the model accuracy. Hence, it is necessary to use a proper number of observations to construct GRFMs regarding practical problems. Note that GRFM is also a local modeling technique, the function cannot be well captured if the observation samples are relatively sparse among fast changing landscapes, or they are too far away from the desired area. In other words, the sampled observations should have a reasonable distribution as well. In this chapter, the DACE toolbox \[58\] is adopted to implement GRFM.

### 3.2.2 GRFM in Cognitive Network Models

In the network model proposed in Section \[5.1\] the cognitive band is assumed to be an ideal continuous medium and thus can transmit power from any part of it. However, the SUs are discrete in practical cases. To adopt GRFM in approximations, the cognitive users are deployed within the cognitive band and assumed to be located apart from each other. Thus, the interference power received by each cognitive user can be derived if the channel model is given. This is essential to the construction of the GRFM, since it implies that each SU can capture the related information of the network at its own position. If all the SUs
can be well synchronized, a rough graph of interference distribution within the cognitive band can be obtained at the sample instant.

Next we discuss GRFM applied in a cognitive network simulation model. As described in Section 3.2.1 a GRF can be applied to construct metamodels for predicting function values at the desired points, when the function is continuous and smooth in the search space. Apparently, the interference power can be regarded as a continuous 2-dimensional spatial function and thus GRFM can be applied to estimate the distribution of interference power. For each point in the practical cognitive band that is devoid of SUs, the equivalent interference power can be estimated by using GRFM and the measurements from the neighboring SUs in its vicinity. Unlike the continuous medium model, the discrete interference power model is continuous among the entire cognitive network plane and has a smooth edge at the border of the exclusion zone. Even if the SUs are some distance away from the selected desired point, the interference could still be approximated using GRFM. Thus, the interference power at a desired point in $\epsilon_p$ band or exclusion zone can be obtained in the same way. Generally speaking, GRFM well fits low dimensional practical function approximations, and the efficiency makes GRFM a desirable choice for approximating the interference power distribution. For the worst case in which the primary receiver is located on the border of the exclusion zone, a sufficient large number of cognitive users should be used. However, too many data samples will lead to high computational overhead, which decreases the efficiency of GRFM in real-time scenarios. Hence, we often need to properly select a moderate amount of SUs and make sure that they are dense enough to sample the interference power distribution reasonably. In addition, the selected SUs should be as close as possible to the desired point, since the correlation among them will become stronger and thus fit GRFM requirements.

The selection criterion of the SUs is illustrated in Fig. 3.2 in which the closest cognitive users to the primary receiver are deployed in the intersection between the cognitive band and the circle centered at the primary receiver (i.e., dark grey area in Fig. 3.2). Samples from these nodes will be used to construct GRFMs. In the following sections, we will consider two
scenarios in which the transmit power and the interference power are considered as GRFM objective function respectively, experimentally study the performance of GRFM approach, and then make comparisons to the analytical results.

### 3.3 Estimation of Exclusion Zone

In this section, we investigate the coverage of exclusion zone as well as the relationships between the bound and other parameters of the network. According to the descriptions in Section 3.2, the performance of the primary receiver should satisfy some constraints in the presence of the exclusion zone. The exclusion zone is defined such that the outage
probability of the primary receiver inside it is below some predetermined threshold. Assume that the rate of the primary transmitter is $T_0$, the outage constraints thus yields:

$$\Pr[T_0 \leq C_0] \leq \beta$$

(3.11)

where $C_0$ and $\beta$ are the thresholds for actual transmission rate and outage probability, respectively. This equation ensures that the transmission rate of the primary receiver is at least $C_0$ for $1 - \beta$ fraction of time. Considering the worst case in which the primary receiver is on the edge of the exclusion zone and assuming the Shannon capacity for the desired transmission rate, we have $T_0 = \log(1 + \frac{P_T}{R_0(I_0 + \sigma^2)})$, where $I_0$ is the aggregated interference power, $P_T$ is the primary transmit power, and $\sigma^2$ is the additive white Gaussian noise. By substituting $T_0$ into the outage constraint, an equivalent constraint in terms of SINR yields:

$$\Pr[I_0 \leq I_{th}] \leq \beta$$

(3.12)

where $I_{th} = \frac{P_T}{2^{\frac{C_0}{\alpha_0}} - 1} - \sigma^2$ is the threshold for SINR. In the model where additive channel noise is not taken into consideration, the threshold $I_{th}$ becomes $\frac{P_T}{2^{\frac{C_0}{\alpha_0}} - 1}$, while the SINR is actually the signal-to-interference ratio (SIR). For different values of transmission rate threshold $C_0$, the relationship between the constraint and the bound of exclusion zone (i.e., $R_0$) needs to be investigated. Thus, the key issue is to find out the SIR received by the primary user. In theoretical analysis for the continuous cognitive band, the transmit power and aggregated interference power distribution can be directly derived from their corresponding closed form expressions, and the outage probability of the primary receiver can be calculated as well. According to the basic assumptions in GRFM, the output is assumed to be a Gaussian random variable. Thus, the equivalent SIR is a random variable as well, and the probability that the SIR is above threshold $\beta$ can be calculated if its distribution is known. As described in Section 3.2.1, GRFM not only provides a mean value estimation at the desired point, but also a measure of uncertainty in terms of variance. Since the transmit power and aggregated
interference power are two distinct variables and independent from each other, equivalent SIR in dB scale equals received power minus aggregated interference. The probability that equivalent SIR is greater than some threshold thus becomes the following:

\[ \text{Pr}[\text{SIR} > I_{th}] = \int_{-\infty}^{\infty} \text{Pr}(P_r - P_i > I_{th} | P_i > \theta) d\theta \]  

(3.13)

where SIR is the actual SIR at the primary receiver, \( P_r \) and \( P_i \) are the received signal power and aggregated interference power at the primary receiver in dB scale, respectively. In our model where the SUs are discrete, a straightforward approach to estimate the coverage of the exclusion zone is to investigate its radius under some predefined outage constraint. Therefore, the outage probability constraint of the primary receiver is given by:

\[ \text{Pr}(\text{SIR} \leq I_{th}) = \text{P}(P_r - P_i \leq I_{th}) \leq \beta, \]  

(3.14)

Note that \( P_r \) and \( P_i \) in dB scale are both assumed to be Gaussian distributed. According to the basic assumption in GRFM, \( SIR = P_r - P_i \) has a Gaussian distribution as well. The mean and variance estimate for SIR are \( \hat{P}_r - \hat{P}_i \) and \( \hat{\sigma}_i^2 + \hat{\sigma}_r^2 \), respectively, where \( \hat{P}_r \), \( \hat{P}_i \), \( \hat{\sigma}_i^2 \), \( \hat{\sigma}_r^2 \) are the mean and variance estimates for \( P_r \) and \( P_i \), respectively. Thus, the probability that the SIR is above the threshold \( I_{th} \) can be expressed in terms of a Q-function, which can be rewritten as:

\[ \text{Pr}(\text{SIR} \leq I_{th}) = Q \left( \frac{(\hat{P}_r - \hat{P}_i) - I_{th}}{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_r^2}} \right), \]  

(3.15)

For any fixed point inside the inner circle of the cognitive band, this probability can be calculated from (3.15). Intuitively, as the primary receiver gets closer to the center of the network, the SIR it receives increases and eventually exceeds the threshold \( \beta \). We select a number of points with various distances from the primary transmitter and approximate their SIR. By fixing the outage probability constraint \( \beta \), the relationship between the exclusion
zone radius \( R_0 \) and the SIR threshold \( I_{th} \) can be derived as

\[
I_{th} = \hat{P}_r(R_0) - \hat{P}_i(R_0) - Q^{-1}(\beta)\sqrt{\hat{\sigma}_i^2(R_0) + \hat{\sigma}_r^2(R_0)}.
\] (3.16)

In our experiments, we test a number of different values of outage threshold \( \beta \) and derive the corresponding results and compare them to the theoretical analysis in which the cognitive band is continuous (continuous case). Based on the analysis of the continuous case, the received signal power and aggregate interference power can be directly derived from their corresponding closed-form expressions, and the outage probability of the primary receiver can be consequently calculated given \( I_{th} \). Under a certain level of transmission rate \( C_0 \), the value of \( R_0 \) that satisfies Eq. (3.15) defines the edge of the exclusion zone. We assume that the Shannon capacity of the primary system is \( C \), and the maximum capacity that the primary system can achieve becomes

\[
C = \log_2(1 + \frac{P_T}{E[I_0]}),
\] (3.17)

where the aggregated interference \( E[I_0] \) is defined by Eq. (3.2). Note that \( E[I_0] \) is determined by the distance to the primary transmitter, the capacity \( C \) is thus a function of the distance as well. When \( \alpha = 4 \) is an even number, Eq. (3.2) has a closed-form expression. In this work, we assume that the \( \alpha = 4 \), and Eq. (3.2) yields:

\[
E[I_0]_{\alpha=4} = \int_{R_0+\epsilon}^{R} \int_0^{2\pi} \frac{\lambda r P_c n dr d\theta}{(r^2 + R_0^2 - 2rR_0 \cos \theta)^{\alpha/2}} \bigg|_{\alpha=4} \bigg( R^2 \bigg) \bigg( R_0 + \epsilon \bigg)^2
\] (3.18)

More details about the derivation of \( E[I_0] \) can be found in [84]. By substituting Eq.
\( (3.2) \) into the expression of \( C \) \( (3.17) \), we have the following:

\[
C\big|_{\alpha=4} = \log_2(1 + \frac{P_T}{E[I_0]}) = \log_2(1 + \frac{P_T}{\lambda \pi P_{cn}\left[\frac{R^2}{R^2-R_0^2} + \frac{(R_0+\epsilon)^2}{\epsilon^2(2R_0+\epsilon)^2}\right]})
\] (3.19)

At any certain sample point within the inner bound of cognitive band, the practical threshold of transmission rate \( C_0 \) should not exceed the value of \( C \). Accordingly, the threshold for aggregated interference \( \frac{P_T}{R_0^2-1} \) could be theoretically calculated as well. Here \( C_0 \) is the capacity of the primary system under the existence of aggregate SU interference power. With the SIR distribution derived from GRFM, the probability that the SIR will exceed the threshold is derived from the Q-function expression in \( (3.15) \). This probability is compared to the given threshold \( \beta \) to decide whether the current point is inside the exclusion zone or not. The value of \( R_0 \) such that the probability equals the threshold directly yields the radius of the exclusion zone.

3.4 Experimental Results

In this section, we study the performance of the proposed GRFM through simulation. To effectively characterize the exclusion zone, the equivalent SIR performance at different locations in the network needs to be determined. For the primary receiver, the signal power comes from the primary transmitter, and the aggregate interference comes from all the SUs in the cognitive band. Therefore, we apply GRFM in both scenarios to derive the estimate of primary transmit power and aggregate interference. To adopt GRFM for approximation, we make use of the MATLAB toolbox DACE [58]. The inputs are the coordinates of the SUs as well as the observation samples, i.e., signal and aggregate interference power. The outputs are the mean and variance estimates of the function values at the desired points.
3.4.1 Received Signal Power Approximation

The first experiment is to apply GRFM to approximate the distribution of transmit power from the primary transmitter. Generally, we can derive both the mean and the variance of the function value simultaneously in GRFM approximation, even if the objective function does not have some simple closed expression, i.e., the objective could be a black-box function with a very complex landscape. Clearly, in this scenario, the expression for the transmit power has a relatively simple form, and the propagation medium is assumed to be homogeneous. According to the path loss model in Section 3.1, assume that $d$ is the propagation distance from the primary transmitter and $W_P$ is the lognormal shadowing, the signal power received by SU $i$ from the primary transmitter, in dB scale, is given by

$$P_{R,dB}(i) = P_{T,dB} - 10\alpha \log_{10}(d_i) + W_P,$$

(3.20)

where $P_{T,dB}$ is the power from the primary transmitter in dB, $d_i$ is the distance between primary transmitter and SU $i$, and the shadowing noise $W_P$ is Gaussian distributed with zero mean and variance $\sigma^2_S$. The effect of shadowing noise can be suppressed by measurement averaging, though it cannot be entirely removed from the received signal. By averaging using $M$ observation samples from a given SU, the variance of the shadowing noise can effectively be reduced by a factor of $M$, at the expense of additional computational overhead. Our objective is to approximate the received power at the primary receiver using a set of secondary nodes. The reason to use dB scale for power level rather than linear is that the fluctuation of power distribution might vary significantly in linear scale, which is likely to result in an ill-conditioned matrix in GRFM computation and large estimation error. If we transform the power scale into log scale, the difference between different entries of the matrix will become much smaller and thus reduce the estimation error.

In our simulation model, the 2-dimensional propagation plane is first divided into small square grids of equal size $\frac{1}{\sqrt{\lambda}} \times \frac{1}{\sqrt{\lambda}}$ and then SUs are generated randomly within each grid according to a uniform distribution. This approach balances the need to represent both
Table 3.1: Simulation Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density of cognitive users</td>
<td>$\lambda = 16/m^2$</td>
</tr>
<tr>
<td>Primary transmit power</td>
<td>$P_T = 40$ dBm</td>
</tr>
<tr>
<td>Secondary transmit power</td>
<td>$P_{su} = 0$ dBm</td>
</tr>
<tr>
<td>Pathloss exponent</td>
<td>$\alpha = 4$</td>
</tr>
<tr>
<td>Amplitude</td>
<td>$A = 1$</td>
</tr>
<tr>
<td>Number of selected SUs</td>
<td>$T = 80$</td>
</tr>
<tr>
<td>Cognitive band radii</td>
<td>$R_i = 1.1$ m; $R = 5$ m</td>
</tr>
<tr>
<td>Number of measurements</td>
<td>$M = 100$</td>
</tr>
<tr>
<td>Standard deviation of shadowing</td>
<td>$\sigma_s = 2$ dB</td>
</tr>
</tbody>
</table>

...randomness and diversity in the locations of the SUs. If the density of SUs is $\lambda$, the width of the grid thus becomes $\frac{1}{\sqrt{\lambda}}$. Once the SUs are generated, the inner and outer bound of the cognitive band are fixed for the simulation. The simulation parameter settings for the exclusion zone are summarized in Table 3.1.

As discussed above, $R_i = R_0 + \epsilon$. The mean number of SUs in the cognitive band is given by $\lambda \pi (R^2 - R_i^2) = 1,196$. Thus, the selected SUs used to obtain the GRFM comprise about $80/1196 \approx 6.7\%$ of the SUs in the cognitive band. The wireless propagation environment is considered to be homogeneous. To test the performance of the GRFM, we deploy a set of test points in the network with mutually different distances to the primary transmitter, and use the corresponding SUs illustrated in Fig. 3.2 to approximate the received signal power at these points. GRFM can well describe the landscape in the vicinity of the sampler nodes in the network plane. The further the desired point is from the data samplers (i.e., SUs), the less accurate the approximation is expected to be. In our simulation model, the deployment of the SUs is fixed once they are generated. This implies that the distance from the cognitive band to the primary transmitter is fixed as well and equals $R_0 + \epsilon_p$. To ensure the accuracy of GRFM approximation, the desired points should be within a reasonable range to the cognitive band.

Fig. 3.3 compares the theoretical result from [84] to the GRFM approximation. Along the horizontal axis, $d_t$ [m] denotes the distance between a given desired point and the...
primary transmitter, i.e., . Fig. 3.3 shows the actual power a primary receiver can receive when it is a distance $d_t$ away from the transmitter and its approximated received power if only the power measurements of the SUs are given. When $d_t < 1.1$, it is equivalent to the case that $R_0 = d_t$ and $\epsilon = 1.1 - d_t$. As $d_t$ becomes smaller, the discrepancy between the two curves becomes larger. When $d_t$ is sufficiently small, no SUs are deployed in the vicinity of the desired points, so the GRFM provides a less accurate approximation to the primary transmitter power.

3.4.2 Aggregate Interference Power Approximation

In this experiment, we study the performance of GRFM in approximating the aggregated interference power to the primary receiver. Such an interference can be regarded as a
2-dimensional spatial function as well, since the equivalent aggregated interference can be
derived at any point within the propagation space as it can be regarded as a continuous spa-
tial function. Similar to the analysis in Section 3.4.1, the SUs may sample the interference
power instead of transmit power when the primary transmitter is idle. Assuming that all
SUs can transmit signals during the same time frame with the same power, the statistical
aggregated interference one SU receives is thus the sum of transmit powers from all other
SUs regardless of all other types of noise. Besides, the number of SUs is often sufficiently
large so that the aggregated interference at each node is approximately a Gaussian variable.
Regarding the basic assumptions in GRFM, it is appropriate to apply it in this scenario.

Using the same simulation model as in Section 3.4.1, a certain SU could only receive
the aggregated interference power from all other SUs. The sum of all SU transmit powers
is a continuous function at any point in the network space. However, the samples taken
by each SU are not exactly the equivalent aggregated interference in the interference power
field, as the SUs are discrete in the network model. For the primary receiver, the aggregated
interference sample is the sum of transmit powers from all SUs. Since an SU cannot actually
cause interference to itself, it is necessary to add an equivalent term $P_{su}$ to the aggregate
interference samples taken by SUs before applying GRFM. Let $N$ denote the total number
of SUs, the equivalent aggregate interference of the $i$th SU can be written as follows:

$$P_{I,S}(i) = P_{su} + \sum_{j=1, j\neq i}^{N} P_{I}(i,j),$$

(3.21)

where $P_{I,S}(i)$ is the total equivalent aggregate interference power received by SU $i$, and
$P_{I}(i,j)$ is the interference power from SU $j$, all in linear scale. The term $\sum_{j=1, j\neq i}^{N} P_{I}(i,j)$
indicates the aggregated interference directly sampled by $i$th SU, while $P_{I,S}(i)$ will serve
as the input of GRFM. The assumption for the equivalent term $P_{su}$ is that it equals the
secondary transmit power, as it is constant for all SUs and no attenuation occurs at exactly
the position of each SU. Therefore, the basic assumptions in GRFM still hold, which makes
GRFM appropriate for aggregate interference power approximation. Similar to (3.20), the
interference power from SU $j$ in the presence of shadowing in dB is given by

$$P_{I,dB}(i,j) = P_{su} - 10\alpha \log_{10}(d_{ij}) + W_{S,ij}, \quad (3.22)$$

where $d_{ij}$ is the distance between SUs $i$ and $j$, and $W_{S,ij}$ represents the lognormal shadowing noise.

Same as in Section 3.4.1 the region in which we take samples is a ring which is neither too close nor too far away from the cognitive band. The inner and outer radius of the ring are set to $0.6R_0$ and $R_0$, respectively. The inner radius of the cognitive band is always fixed once the SUs are generated, and set to 1.1. The sample points are thus taken inside the region defined above, with their distances to the primary transmitter given as 0.6m, 0.7m,
0.8m, 0.9m and 1.0m, respectively. The SU transmit power is set to 1 mW (0 dBm), and the rest of the GRFM parameter settings for this experiment are the same as in the received signal power approximation in Section 3.4.1.

Under the above parameter settings, we repeat the GRFM approximation 1,000 times for each case. The relationship between aggregate interference power approximation at each sample point and the distance to the primary transmitter is shown in Fig. 3.4. The actual aggregate interference in both discrete case and continuous case are shown as well. The aggregate interference in the continuous case was obtained using the theoretical analysis, in which the expected interference to the primary receiver is expressed as an integral over the cognitive band in Eq. (3.18). The actual aggregate interference (discrete case) was obtained by summing up the transmission power contributions from the finite set of SUs in the cognitive band. It can be observed that GRFM approximation curve approximates the actual aggregate interference curve quite well. When the test point moves to the primary transmitter, both the GRFM approximation and actual aggregated interference slowly decrease, yet still have an acceptable discrepancy with the theoretical result in the continuous case. This is reasonable because all of the parameters adopted in the continuous and discrete cases are the same. This result demonstrates the effectiveness of the GRFM in estimating the actual aggregate interference distribution.

It can be found that when a sample point gets closer to the target, the theoretical aggregated interference and the upper/lower bounds all go to infinity. As a contrast, GRFM approximation results are finite and fall slightly below the theoretical lower bound when $R_0 = 1$, since the simulation model adopted naturally prevents ill-state matrix computation in GRFM. When the sample point moves closer to the center of the network, the GRFM approximation results achieves comparable or even better quality compared to the theoretical lower bound. The result clearly demonstrates the effectiveness of GRFM on estimating the aggregated interference distribution.
3.4.3 Exclusion Zone Approximation

According to the definition of exclusion zone in Section 3.3, the outage probability is calculated through the distribution of received signal power and interference power in the expression of SIR. In the channel model, the approximated aggregated interference $\hat{P}_i$ defined in (3.15) not only contains the part that comes directly from the aggregate interference power approximation in Section 3.4.2, but also contains the equivalent shadowing that has not been completely removed from the approximated received signal power. Assume that the approximated interference power is $\hat{P}_{I,S}$ with variance $\hat{\sigma}^2_{I,S}$, the mean value and variance for $\hat{P}_i$ become $\hat{P}_{I,S}$ and $\hat{\sigma}^2_{I,S} + \sigma^2_M$, respectively, from properties of Gaussian random variables.

To evaluate the exclusion zone approximation based on the GRFM, we use the same
parameters as adopted in Section 3.4.2. As the width of the guard band $\epsilon$ is not an independent parameter since the inner radius of the cognitive band (i.e., $R_0 + \epsilon$) is constant after the cognitive band is generated. Under different transmission rate requirements, the relationship between SIR threshold $I_{th}$ and exclusion zone radius $R_0$ can be derived. Here, three different values of outage probability constraint $\beta$ are selected for evaluating the exclusion zone radius $R_0$, which are 0.05, 0.1 and 0.15, respectively. The relationship between $R_0$ and the SIR threshold under different outage requirements is shown in Fig. 3.5.

It can be observed that the SINR threshold decreases to a relative low level (around 13 dB) when the exclusion zone radius extends to $R_0 = 1$, as the harmful interference the to the primary receiver significantly increases near the cognitive band. By decreasing the radius of the exclusion zone, the guard band width grows as well and this implies that a higher SINR threshold is reasonable. Meanwhile, under the same exclusion zone radius, the growth of outage probability leads to the increase of SINR threshold. However, the exclusion zone size has a more significant impact on the SINR constraint, since it determines the maximum transmit power and the minimum interference power received by the primary user in the worst case.

In this experiment, the GRFM approximation is shown to be effective in determining the exclusion zone radius and the relationship between $R_0$ and other parameters in the proposed cognitive network model.

### 3.5 Conclusion

In this chapter, we proposed a novel modeling framework, GRFM, to approximate the coverage of the exclusion zone. This is an approach of interest from a new perspective in evaluating the coverage of the exclusion zone. Instead of using theoretical analysis, the distribution of two spatial functions, the primary transmit power and the aggregated interference from all cognitive users are both approximated directly using the observation samples from the SUs. A discrete network model is adopted as well to replace the continuous model in theoretical analysis. The approximation results show that the performance of
GRFM is rather accurate as long as the desired point for prediction is within a reasonable range to the SUs. Based on these results and the assumptions in GRFM, we studied the coverage of the exclusion zone and the relationship between the network parameters and the area it covers. The simulation results show that good approximation quality of power distribution as well as exclusion zone coverage can be achieved under the proposed GRFM framework.
Chapter 4: Stackelberg Game Theoretic Approach to Resource Allocation with Cooperative Relaying

Cooperative relaying in cognitive radio networks (CRN) has been intensively studied in past decades. In CRNs, direct link communications between secondary user (SU) pairs can bring more potential benefit in terms of system capacity. The direct SU communication scheme aims to set up a pattern in which the spectrum usage and SU transmit power are well designed while maximizing the system capacity. Yet, the scheme still faces the following problems: (1) Transmission between SUs will cause significant interference to the primary system; (2) Long-distance transmission between SUs may not be reliable enough; (3) As the direct communication demand grows, sophisticated coordination for spectrum utilization is required between the SUs and the primary system. (4) The fairness over multiple devices needs to be taken into account. The coordination scheme generally include aspects such as power control and resource allocation, which have been studied in prior works. For power control issues, a reasonable transmit power for SUs should be selected to satisfy the communication requirements while keeping the harmful interference to the primary system under some threshold.

Game theory has been proposed to offer an effective mathematical tool to model the interactions among the network devices and give predictions of their future actions. In CRNs, the SUs interact with the primary users (PUs) to compete for necessary spectrum resources. By assuming that the PUs and SUs are independent rational players, their own strategies can be derived under a reasonable game framework. Among existing game-theoretic approaches, Stackelberg game exhibits good performance in modeling the interactions among non-cooperative devices with their own goals. Under the proposed Stackelberg game framework, the PUs are willing to sell their channel resources at an optimal price to maximize
their utilities. Subsequently, the SUs try to select the optimal relay node for cooperative relaying and determine their optimal transmit powers when buying channel resources from the PUs. The Nash equilibrium (NE) of the proposed game is derived with full knowledge of channel gains and transmit power, and the optimal transmission strategy of the PUs and SUs can be derived. However, the above challenges are still not well tackled.

In this chapter, we assume that certain transmission demands exist among a number of SU pairs, and we present a novel extended Stackelberg game-theoretic approach for downlink resource allocation in an SU cooperative relay communication scenario, which jointly manage spectrum resources and coordinate PUs and SUs with decode-and-forward (DF) cooperative relaying capability to extend network coverage. To the best of our knowledge, little prior work has been done on the proposed problem. In our framework, the extended Stackelberg game is first adopted to model the interaction between given SUs and PUs that are the potential spectrum resource providers, in which the SUs attempt to decide whether or not to use relay nodes and determine their optimal transmit power. Since the outcome given by the Stackelberg game is in a generalized form, it is necessary to map the SU and PU into reasonable groups to maximize the overall system rate. Thus, a mapping scheme is proposed at this stage in which the PUs compete for the SUs until no available spectrum resource left. In practical CRNs, multiple PUs and SUs are deployed in the network area and a scheduling approach is required to map the PUs and SUs to optimize the network performance. Our main contributions are as follows:

- We propose a novel framework for downlink transmission with SU cooperative relaying under a decode-and-forward (DF) scheme, in which the interactions between the devices in cooperative relaying (i.e., the SUs) and the potential spectrum providers (i.e., the PUs) are modeled by a Stackelberg game.

- Theoretical analysis of the Nash equilibrium (NE) of the proposed Stackelberg game is carried out to derive the optimal transmission strategies of the devices.

- Based on Stackelberg game outcomes, we design a hybrid priority-based scheduling
scheme to jointly select appropriate relay nodes for SU pairs with traffic demand and allocate spectrum resource for both direct transmission (DT) and DF cooperative relaying. Time-domain scheduling regarding fairness is also incorporated in the proposed algorithm among consecutive time slots.

- We analyze the performance of the proposed scheduling algorithm by running extensive simulations. Our numerical results show that significant performance improvement can be achieved by employing DF cooperative relaying among the SUs compared to using DT only.

The remainder of the chapter is organized as follows. The system model is described in Section 4.1. The extended Stackelberg game as well as the theoretical analysis of the above Stackelberg game are given in Section 4.2. Based on Stackelberg game outcomes, a novel hybrid scheduling algorithm incorporating both direct transmission and DF cooperative relay transmission is proposed to tackle the problem in Section 4.3. The experimental results are discussed and analyzed in Section 4.4. Finally the conclusions are given in Section 4.5.

4.1 System Model

4.1.1 Network Model

The adopted cognitive radio network (CRN) model consists of one base station (BS), multiple primary users (PUs) and multiple secondary users (SUs), which are deployed over a transmission plane. We assume that the BS, PUs and SUs are each equipped with an omni-directional antenna. Each PU is preallocated with a fixed licensed channel so that it can communicate with the BS. Meanwhile, the SUs can access these channels, transmit and receive signals as long as the PUs are detected as being. Basically, the SU nodes can be classified into two categories, namely SU pairs and independent SUs. In the first category, two SUs act as transmitter and receiver respectively in an SU pair, while certain communication demands exist between them. The independent SUs do not have such
communication requirements. They could either communicate with the BS or with other nodes depending on their needs. In both categories, each licensed channel used by the PUs is only allowed to be accessed by one SU transmission link at one time. Thus, the secondary system can be regarded as an overlay to the primary system to improve spatial efficiency and system capacity of the entire network.

To meet the communication demands given above, links between each SU pair need to be established. If the deployment of SUs is dense enough to guarantee a sufficiently small transmission distance for each SU pair, a direct link can be set up between the SU transmitter and receiver. However, if the transmitter and receiver in the same SU pair are geographically separated from each other, a direct link might not be available to maintain a satisfactory quality of SU service. To tackle this problem, multi-hop communication can be adopted to avoid the drawbacks of using direct links between certain SU transmitters and receivers. By adopting relay nodes that forward the signal from SU transmitter to SU receiver, better system performance (i.e., overall throughput) can be achieved. Apparently, relay nodes are cognitive users as well that can sense the channels individually. Therefore, the independent SUs could naturally play the role of relay nodes.

We focus on downlink channel access and consider the effect of co-channel interference between PU and SU nodes. In the downlink transmission process, the PUs receives signals from the BS and the interference from the SUs that access the channels, while the SU receivers often suffer severe interference from the BS. As the number of PUs grows, the BS needs to properly allocate its total transmit power to all the PUs over different channels, which leads to smaller interference to each SU device that access the channels. The SUs could also coordinate with the corresponding PUs and select appropriate transmit powers. In this case, downlink period transmission could be investigated. A scenario illustrating the proposed spectrum sharing scheme in a cognitive radio network is presented in Fig. 4.1. For simplicity, we assume that at most one relay node can be adopted by each SU pair. As depicted in Fig. 4.1, the SU pair \((S_1, S_2)\) uses DT, which shares the channel that has been licensed to PU \(P_1\), and \(S_2\) suffers interference from the
BS. Meanwhile, PU $P_1$ receives interference from $S_1$. As a contrast, the SU pair $(S_3, S_4)$ uses DF transmission using SU $S_5$ as the relay. We assume half-duplex transmission and that the two hops do not use the same channel and hence do not cause interference to each other. The PUs $P_2$ and $P_3$ are selected to provide the channel resources to the first and the second hop, respectively. Apparently, $S_4$ receives only the signal forwarded from $S_5$. The PU channel selection can be done via wideband spectrum sensing \cite{16}. 

Figure 4.1: Cognitive Radio Network Model with Downlink Channel Sharing.
4.1.2 Decode-and-Forward (DF) Relay Transmission Scheme

We consider decode-and-forward in cooperative relaying in this chapter. As the signal in amplify-and-forward (AF) scheme is simply scaled and forwarded by the relay node, different hops need to access the same channel to implement the AF scheme, and the PU that provides the channel would suffer from higher interference from all the hops. In contrast, the relay node in DF cooperative relaying fully decodes the signal received from the previous hop and then forwards the signal to the next hop [21]. Hence, the relay node could remodulate the signal and select an appropriate channel to transmit, which is more flexible compared to AF scheme. Besides, the DF scheme can better utilize the spectrum resource provided by different PUs. The investigation of DF cooperative relaying is given as follows.

In DF scheme, time-division multiplexing (TDM) is assumed such that either the SU transmitter or the relay node can transmit during each transmission time interval (TTI) [86, 100]. In a complete transmission process, the transmitter in an SU pair transmits the signal in the first TTI. Once the signal is received by the relay node, it selects a channel and transmits the signal to the next hop using an appropriate power in the following TTI. Such forwarding process terminates once the signal is received by the corresponding SU receiver. We assume that at most one relay node could be used in each SU end-to-end link. Since the relay node needs extra processing effort to decode the signal, the DF cooperative relaying scheme takes two consecutive TTIs to complete.

In our model, we assume a total number of $M$ PUs in the CRN coverage area, which implies a maximum of $M$ orthogonal channels. The number of SU pairs with certain traffic demand is assumed to be $N$. Since the independent SUs are often randomly deployed in the network area, each independent SU could serve an SU pair as a relay node as long as it is close enough to both transmitter and receiver. Clearly, a relay node could serve multiple SU pairs as long as the above constraints are satisfied. The total number of relay nodes that can serve at least one SU pair is assumed to be $R$. Hence, the total number of active SUs in the CRN is $2N + R$. In each TTI, the SU transmitters and relay nodes that receive a
signal from the first hop need to transmit a signal, hence they may access different channels considering frequency-division multiplexing. We further assume that the BS uses equal and constant transmit power to each PU, and we focus on one SU pair under DF cooperative relaying. In TDM transmission, since no interference occurs between two consecutive TTIs in DF cooperative relaying, the two hops could either access the same channel or different channels. If the SU transmitter and the relay node access the channel provided by PU $i$ and PU $j$, respectively, the signals received by PUs $i, j$ are given as follows. Without loss of generality, we assume that the first hop and the second hop access the channel licensed to PU $i$ and PU $j$ are, respectively,

\[ y_p^i = \sqrt{p_b g_{bp,i}} s_{b,i} + \sqrt{p_s g_{sp}} s_s + n_{p,i} \tag{4.1} \]
\[ y_p^j = \sqrt{p_b g_{bp,j}} s_{b,j} + \sqrt{p_r g_{rp}} s_r + n_{p,j} \tag{4.2} \]

where $s_{b,i}, s_{b,j}$ are the symbols transmitted from BS to PU $i, j$, and $s_s, s_r$ are the symbols transmitted by SU transmitter and relay node, respectively. The transmit powers of the BS to each PU, the SU transmitter and the relay node per symbol are denoted by $p_b, p_s, p_r$, respectively. The noise received by PU $i, j$ are $n_{p,i}, n_{p,j}$, respectively. The channel coefficients from the BS to PUs $i$ and $j$, from the SU transmitter to PU $i$, and from the relay node to PU $j$ are denoted by $g_{bp,i}, g_{bp,j}, g_{sp}, g_{rp}$, respectively. Similarly, the signals received by relay node and SU receiver are given by:

\[ y_r = \sqrt{p_s g_{sr}} s_s + \sqrt{p_b g_{br}} s_{b,i} + n_r \tag{4.3} \]
\[ y_s = \sqrt{p_r g_{rs}} s_s + \sqrt{p_b g_{bs}} s_{b,j} + n_s \tag{4.4} \]

where $g_{sr}, g_{br}, g_{rs}, g_{bs}$ denote the channel coefficients from SU transmitter to relay, from BS to relay, from relay to SU receiver and from BS to SU receiver, respectively. The noise received by relay node are $n_{p,i}, n_{p,j}$ respectively. Assume that all the licensed channels have the same bandwidth $W_0$ and the power spectral density (PSD) of the additive white
Gaussian noise is $n_0$, and the noise power received by each device is given by $N_0 = n_0 \cdot W_0$. Therefore, the signal-to-noise-plus-interference ratio (SINR) received by the PUs are given by:

\[
\gamma_{p,i} = \frac{p_b g_{bp,i}}{p_s g_{sp} + N_0}, \quad \gamma_{p,j} = \frac{p_b g_{bp,j}}{p_r g_{rp} + N_0},
\]

(4.5)

Similarly, the SINR received by the relay node and the SU transmit power are given by:

\[
\gamma_r = \frac{p_s g_{sr}}{p_b g_{br} + N_0}, \quad \gamma_s = \frac{p_r g_{rs}}{p_b g_{bs} + N_0},
\]

(4.6)

(4.7)

According to Shannon’s capacity formula, the channel rate of the two PUs, the first and the second hop in SU transmission are given by:

\[
r_{p,i} = \log_2(1 + \gamma_{p,i}), \quad r_{p,j} = \log_2(1 + \gamma_{p,j}),
\]

(4.8)

\[
r_1 = \log_2(1 + \gamma_r),
\]

(4.9)

\[
r_2 = \log_2(1 + \gamma_s).
\]

(4.10)

A Rayleigh fading channel model is adopted in this chapter. Hence, for any devices $i, j$ in the proposed CRN, the channel coefficient is given by $g_{ij} = |h_{ij}|/d_{ij}^2$, where $d_{ij}$ is the distance between device $i$ and $j$, and $h_{ij}$ follows a complex Gaussian distribution (cf. [86]). Each SU pair aims to select a relay node and then decide the appropriate transmit power to optimize its objective. Meanwhile, the PUs providing channels to the SU pair need to maximize their revenues as well. Therefore, a non-cooperative game between a SU pair and its corresponding PUs is established, and a Stackelberg game-theoretic model is proposed to investigate the optimal transmission strategy.
4.1.3 Direct Transmission Scheme

In the proposed DF scheme, each SU pair must adopt one relay node for transmission if it does not satisfy the geographical distance constraint. However, for a SU pair that meets the constraint, it could either adopt relay nodes or use direct transmission (DT). Therefore, we have the following analysis:

For a certain SU pair \((S_T, S_R)\) without a relay node, it shares a licensed channel with a PU, say, \(P\). During downlink period, the BS transmits signal \(s_b\) to the PU and the SU transmitter \(S_T\) transmits signal \(s_s\) to its receiver \(S_R\). The received signal at \(P\) can be written as follows:

\[
y_p = \sqrt{p_b g_{bp}} s_b + \sqrt{p_s g_{sp}} s_s + n_p \tag{4.11}
\]

And the signal received by \(S_R\) yields:

\[
y_s = \sqrt{p_s g_{ss}} s_s + \sqrt{p_b g_{bs}} s_b + n_s \tag{4.12}
\]

where \(p_b\) is the BS transmit power using channel assigned to \(P\); \(p_s\) is the SU transmit power. \(y_p\) and \(y_s\) denote the total power received by PU \(P\) and SU receiver \(S_R\), respectively. Here \(g_{bp}\) is the channel gain between BS and \(S_T\), \(g_{sp}\) is the channel gain between \(S_T\) and \(P\), \(g_{ss}\) is the channel gain between \(S_T\) and \(S_R\), and \(g_{ss}\) is the channel gain between \(S_T\) and \(S_R\). \(n_p\) and \(n_s\) are the additive white Gaussian noise (AWGN) at PU/SU receivers. For all channels, the AWGN is assumed to have the same PSD \(N_0\). The subscripts \(b, p, s\) in channel parameters denote BS, PU and SU, respectively. Hence, the SINR received by PU \(P\) can be written as:

\[
\gamma_p = \frac{p_b g_{bp}}{p_s g_{sp} + N_0} \tag{4.13}
\]
The SINR received by SU receiver $S_R$ can be written as:

$$\gamma_s = \frac{p_s g_{ss}}{p_b g_{bs} + N_0}$$

(4.14)

And the channel rate of the PU and the SU pair are given by:

$$r_p = \log_2(1 + \frac{p_b g_{bp}}{p_s g_{sp} + N_0})$$

(4.15)

and

$$r_s = \log_2(1 + \frac{p_b g_{bp}}{p_s g_{sp} + N_0})$$

(4.16)

respectively. As in the DF scheme, the distribution of the channel coefficients follows a Rayleigh fading channel model.

4.2 Stackelberg Game Analysis

In this section, we present a detailed analysis of the extended Stackelberg game framework adopted in this chapter. The extension of the original Stackelberg game enables us to derive the outcome of the equilibria under the DF scheme.

4.2.1 Extended Stackelberg Game Framework

In the proposed CRN system model, the SUs act as an overlay to the primary system, and we focus on power control and channel scheduling for the SUs pairs and relay nodes \[20\,59\]. To optimize the overall system performance while causing as little harmful interference to the primary system as possible, a power control scheme is required, which consists of at least two basic steps. In the first step, the SU pairs attempt to find an appropriate relay node as well as a transmission strategy. After this, the SU pairs and the corresponding relay nodes are mapped to the PUs that provide spectrum resources to form PU-SU groups. In
order to achieve optimal system performance, a carefully designed PU-SU mapping scheme is presented. In this section, we focus on the first step in which the SU pairs aim to find an optimal pricing strategy, and the power control phase is analyzed under a Stackelberg game-theoretic model.

Stackelberg game is a powerful mathematical tool to describe the interactions between two rational players by modeling them as a leader-follower pair \[6, 86\], in which both the leader and the follower seek to maximize their own utilities. In recent works, Stackelberg game has been successfully applied to the interactions between two independent devices in DT schemes, where the leader charges some price from the follower for utilizing its channel resources. However, unlike DT, SU cooperative relaying utilizes multiple SU devices to accomplish signal transmission, while multiple channels need to be accessed in the scheduling process. Hence, we propose an extended Stackelberg game-theoretic framework for SU DF cooperative relaying in which the leader consists of the two PUs, while the follower consists of the SU pair and the relay node. We assume that the relay node has the same transmit power range as the SU transmitter, i.e., \([p_{\text{min}}, p_{\text{max}}]\). Clearly, the hop with a smaller channel rate in SU link is the bottleneck. We further assume that the charging prices of SU transmit power (per unit power) are the same for accessing both channels, which can be denoted by a positive value \(\alpha_d\). The leader gains revenue by selling its channels to the follower at the optimal value of \(\alpha_d\). For simplicity, the two PUs in the SU pair are denoted by PU 1 and PU 2, respectively. Hence, the leader utility function can be written as:

\[
u_{l,r}(\alpha_d, p_s, p_r) = r_{p,1} + r_{p,2} + \alpha_d \beta p_s g_{sp} + \alpha_d \beta p_r g_{rp}
= \log_2(1 + \frac{p_{\text{g}_{bp,1}}}{p_{s}g_{sp} + N_0}) + \log_2(1 + \frac{p_{\text{g}_{bp,2}}}{p_{r}g_{rp} + N_0})
+ \alpha_d \beta p_s g_{sp} + \alpha_d \beta p_r g_{rp}
\] (4.17)

where \(\beta\) is the ratio between the leader’s revenue and the follower’s payment \(\alpha_d p_s g_{sp} + \alpha_d p_r g_{rp}\). Clearly, the hop with the smaller channel rate determines the overall rate in
DF transmission, and it takes two TTIs to transmit under DF cooperative relaying. The follower also aims to maximize its utility, which can be written as:

$$u_{f,r}(\alpha_d, p_s, p_r) = \frac{1}{2} \min\{r_1, r_2\} - \alpha_d p_s g_{sp} - \alpha_d p_r g_{rp}$$

$$= \frac{1}{2} \min\left\{ \log_2(1 + \frac{p_s g_{sr}}{p_b g_{sr} + N_0}), \log_2(1 + \frac{p_r g_{rs}}{p_b g_{rs} + N_0}) \right\}$$

$$- \alpha_d p_s g_{sp} - \alpha_d p_r g_{rp}$$  \hspace{1cm} (4.18)

From the above equations, it can be observed that the SU link capacity depends on the channel parameters as well as BS transmit power. Since the SU receiver will not receive interference from the first hop, either of the two hops could be the bottleneck in relay transmission, and we select the minimum channel rate among the two hops. The leader first decides the charging price $\alpha_d$ for accessing its spectrum, and the follower decides the transmit powers $(p_s, p_r)$ based on the price $\alpha_d$ for the SU transmitter and relay node, respectively. Stackelberg game can be interpreted as a two-stage game which consists the game at the leader’s side and then the game at the follower’s side. The full strategy of the proposed Stackelberg game is denoted by \{\alpha_d, p_s, p_r\}, while the follower’s strategy is determined by the leader’s strategy. Hence, the follower strategy can be denoted by \{\alpha_d, p_s(\alpha_d), p_r(\alpha_d)\}. The Stackelberg game Nash equilibrium (SNE) is achieved at an optimal strategy $\{\alpha_d^*, p_s^*, p_r^*\}$ when no unilateral deviation of $\alpha_d^*$ and $\{p_s^*, p_r^*\}$ leads to higher leader utility or follower utility, i.e.,

$$u_{f,r}(\alpha_d^*, (p_s^*, p_r^*)) \geq u_{f,r}(\alpha_d^*, (p_s, p_r)) \hspace{1cm} (4.19)$$

$$u_{l,r}(\alpha_d^*, (p_s(\alpha_d^*), p_r(\alpha_d^*))) \geq u_{l,r}(\alpha_d, (p_s(\alpha_d), p_r(\alpha_d))) \hspace{1cm} (4.20)$$

From (4.18), one sees that calculating the link rate is a key issue in follower utility.
Define:

\[ C := \frac{\gamma_r}{\gamma_s} \]  \hspace{1cm} (4.21)

\[ \rho := \frac{p_{\min}}{p_{\min} + p_{\max}}. \]  \hspace{1cm} (4.22)

Since both \( p_s \) and \( p_r \) are between \( p_{\min} \) and \( p_{\max} \), we know that \( \rho \leq \frac{p_s}{p_r} \leq \rho^{-1} \), and the bottleneck rate \( r^* \) of the two hops can be determined by comparing \( p_s/p_r \) and \( C \). Depending on the channel parameters, there are four exhaustive and mutually exclusive cases:

- **Case 1**, \( C \geq \rho^{-1} \): Clearly \( r_1 \leq r_2 \), so \( r^* = r_1 \).
- **Case 2**, \( 1 \leq C \leq \rho^{-1} \): Either \( r_1 \) or \( r_2 \) could be the bottleneck;
- **Case 3**, \( \rho^{-1} \leq C \leq 1 \): Either \( r_1 \) or \( r_2 \) could be the bottleneck;
- **Case 4**, \( C \leq \rho \): In this case, \( r_1 \geq r_2 \), so \( r^* = r_2 \).

Since the \((p_r, p_s)\) parameter space is symmetric, we need only analyze Cases 1 and 2. The analyses for Cases 3 and 4 can be derived in the same way as Cases 2 and 1, respectively. Based on the definition of the four cases, the proof of existence of the Stackelberg game Nash equilibrium is given below. According to the analysis in [63], the existence of SNE is guaranteed by the following theorem:

**Theorem 3.1:** The proposed Stackelberg game admits at least one SNE if \( \forall i \in \{1, \ldots, N\} \), the follower strategy set \((p_s, p_r)\) is a nonempty compact subset of the Euclidean space and the utility function \( u_{f,r}(\alpha_d, p_s, p_r) \) is continuous and quasi-concave on \((p_s, p_r)\).

Apparently, the follower strategies \( p_s, p_r \in [p_{\min}, p_{\max}] \) guarantee a convex set in 2-dimensional Euclidean space. For Case 1 and 4, the follower utility \( u_{f,r} \) is continuous. For Case 2 and 3, since we take the minimum rate as the bottleneck in \( u_{f,r} \), it yields the same value when \( r_1 = r_2 \). If hop 1 is adopted, it is easy to verify that \( u_{f,r} = \frac{1}{2} r_1 - \alpha_d p_s g_{sp} \) is continuous. Similarly, if hop 2 is adopted, \( u_{f,r} = \frac{1}{2} r_2 - \alpha_d p_s g_{sp} \) is continuous as well.
Therefore, \( u_{f,r} \) is continuous in the entire parameter space \((p_s, p_r)\). As for quasi-concavity, we take the second order derivative of the follower utility for each case. If hop 1 is adopted, we have the following:

\[
\frac{\partial^2 u_{f,r}}{\partial p_s^2} = -\frac{1}{2 \ln 2} \left( \frac{g_{sr}}{p_s g_{sr} + p_b g_{br} + N_0} \right)^2 < 0,
\]

\[
\frac{\partial^2 u_{f,r}}{\partial p_r^2} = 0, \quad \frac{\partial^2 u_{f,r}}{\partial p_s \partial p_r} = 0.
\] (4.23)

Otherwise if hop 2 is adopted, the second order derivative yields:

\[
\frac{\partial^2 u_{f,r}}{\partial p_r^2} = -\frac{1}{2 \ln 2} \left( \frac{g_{rs}}{p_r g_{rs} + p_b g_{bs} + N_0} \right)^2 < 0,
\]

\[
\frac{\partial^2 u_{f,r}}{\partial p_s^2} = 0, \quad \frac{\partial^2 u_{f,r}}{\partial p_r \partial p_s} = 0.
\] (4.24)

Hence, the joint quasi-concavity of the follower utility is proved, and the proposed Stackelberg game has at least one SNE.

In the Stackelberg game, the leader knows that the follower will react to its behavior. Thus, it analyzes all the possible follower strategies and asserts the optimal price. The follower subsequently decides the optimal SU and relay transmit power based on the optimal price. Therefore, the proposed game can be solved by backward induction, i.e., the follower’s strategy with respect to price is derived first and the leader decides the price by observing the follower’s strategy.

The four cases in the two-dimensional \((p_r, p_s)\) parameter space are intuitively depicted in Fig. 4.2, where the slope of each dashed line denotes the value of \(C\) in the corresponding case. The follower strategy tuple \(\{p_s^*, p_r^*\}\) is chosen from the square area \([p_{\min}, p_{\max}] \times [p_{\min}, p_{\max}]\). For Cases 1 and 4, no matter what the strategy tuple is, the bottleneck of the SU link will always be the first hop and the second hop, respectively. For Case 2, the follower strategies in two subregions (Region I and Region II,III,IV) within the square area
yield different utility functions. Hence, the strategies in each subregion will be investigated and the optimal follower strategy is derived by comparing them. The same investigation process could be applied to Case 3, in which the two subregions are Region I,II,III and Region IV, respectively.

According to all discussions above, the analysis for Case 1 and Case 2 are given in the following sections.
4.2.2 Analysis to Case 1

In Case 1, the first hop is apparently the bottleneck. Therefore, the follower utility functions yield:

\[ u_{f,r,1}(\alpha_d, p_s, p_r) = \frac{1}{2} r_1 - \alpha_d p_s g_{sp} - \alpha_d p_r g_{rp} \]

\[ = \frac{1}{2} \log_2 (1 + \frac{p_s g_{sr}}{p_b g_{br} + N_0}) - \alpha_d p_s g_{sp} - \alpha_d p_r g_{rp} \]  \hspace{1cm} (4.27)

while the leader’s utility is given by Eq. (4.17). To maximize the follower utility, its first order derivatives with respect to SU transmit power and relay transmit power are obtained as:

\[ \frac{\partial u_{f,r,1}}{\partial p_s} = \frac{1}{2} \ln 2 \frac{g_{sr}}{p_s g_{sr} + p_b g_{br} + N_0} - \alpha_d g_{sp}, \]  \hspace{1cm} (4.28)

\[ \frac{\partial u_{f,r,1}}{\partial p_r} = -\alpha_d g_{rp}. \]  \hspace{1cm} (4.29)

Apparently, Eq. (4.28) and (4.29) are mutually independent. From Eqn (4.29), as the first order derivative is always negative, it is easy to conclude that the follower utility is minimized when the relay transmit power yields \( p_{min} \). For the SU transmit power, we set \( \frac{\partial u_{f,r,1}}{\partial p_s} = 0 \) and the optimal estimate of SU transmit power is:

\[ \hat{p}_s = \frac{1}{2\alpha_d g_{sp} \ln 2} - \frac{p_b g_{br} + N_0}{g_{sr}} \]  \hspace{1cm} (4.30)

Naturally, the SU transmit power estimate should satisfy \( p_{min} \leq \hat{p}_s \leq p_{max} \). By substituting \( \hat{p}_s \) into the inequality, we have:
\begin{equation}
\alpha_{d_{\text{min},1}} = \frac{g_{sr}}{2(p_{\text{max}} g_{sr} + p_{bg} + N_0) g_{sp} \ln 2}, \hspace{1cm} (4.31)
\end{equation}

\begin{equation}
\alpha_{d_{\text{max},1}} = \frac{g_{sr}}{2(p_{\text{min}} g_{sr} + p_{bg} + N_0) g_{sp} \ln 2}, \hspace{1cm} (4.32)
\end{equation}

From the leader’s perspective, it aims to maximize its own utility as well. Therefore, by substituting the optimal SU and relay transmit power estimate into leader’s utility (4.17), we have:

\begin{equation}
u_{l,1} (\alpha_d) = \frac{\beta}{2 \ln 2} - \alpha_d \beta g_{sp} p_{bg} g_{sr} + N_0 g_{sp} + \alpha_d \beta p_{\text{min}} g_{rp}
+ \log_2 \left[ 1 + p_{bgp,1} \left( \frac{1}{2 \alpha_d \ln 2} - g_{sp} \frac{p_{bg} + N_0}{g_{sr}} + N_0 \right)^{-1} \right]
+ \log_2 \left( 1 + \frac{p_{bgp,2}}{p_{\text{min}} g_{rp} + N_0} \right), \hspace{1cm} (4.33)
\end{equation}

Define:

\begin{equation}
A_1 := p_{bgp,1}, B_1 := \frac{1}{2 \ln 2},
\end{equation}

\begin{equation}
C_1 := -g_{sp} \frac{p_{bg} + N_0}{g_{sr}} + N_0, D_1 := p_{\text{min}} g_{rp}, \hspace{1cm} (4.34)
\end{equation}

By taking the first order derivative of the leader’s utility (4.33) and set \( \frac{\partial u_{l,1}}{\partial \alpha_d} = 0 \), we have:

\begin{equation}
\frac{\partial u_{l,1}}{\partial \alpha_d} = \frac{2 A_1 B_1^2}{(C_1 \alpha_d + B_1)(A_1 + C_1) \alpha_d + B_1}
+ (C_1 - N_0 + D_1) \beta = 0, \hspace{1cm} (4.35)
\end{equation}

Similar to the analysis in [86], the optimal charging price, denoted by \( \alpha_{d,1}^* \), is selected
from the set \( \{ \phi_1, \phi_2, \phi_3, \alpha_{dmin,1}, \alpha_{dmax,1} \} \), where:

\[
\phi_1 = \frac{2B_1}{\beta(N_0 - D_1)} - \frac{B_1}{A_1},
\]

\[
\phi_2 = \frac{B_1}{A_1} - \frac{2B_1}{\beta(N_0 + A_1 - D_1)},
\]

\[
\phi_3 = -B_1(A_1 + 2C_1) + \sqrt{A_1^2B_1^2 - \frac{8A_1B_1^2C_1(A_1+C_1)}{\beta(C_1-N_0+D_1)}} / 2C_1(A_1 + C_1)
\]

Based on the optimal price, the optimal SU and relay transmit power are given by

\[
p_{s,1}^* = \frac{1}{2\alpha_{d,1}g_{sp} \ln 2} - \frac{p_b g_{hr} + N_0}{g_{sr}}, \quad p_{r,1}^* = p_{\min},
\]

respectively. The leader and follower utilities in Case 1 are given by \( u_{l,r,1} = u_{l,r}(\alpha_{d,1}^*, p_{s,1}^*, p_{r,1}^*) \) and \( u_{f,r,1} = u_{f,r}(\alpha_{d,1}^*, p_{s,1}^*, p_{r,1}^*) \), respectively.

### 4.2.3 Analysis for Case 2

In Case 2, where \( 1 \leq C \leq \rho^{-1} \), either hop could be the bottleneck, depending on the channel coefficients and the threshold \( C \). Hence, to derive the Stackelberg game equilibrium in Case 2, we need to find the optimal transmit power that maximizes the utility functions. It is easy to verify that in the search space, the straight line \( p_s = Cp_r \) intersects \( p_r = p_{\min} \) at \( p_s = Cp_{\min} \). Therefore, the entire search space can be divided into two non-overlapping rectangular sub-areas, i.e., \( p_s \in [p_{\min}, Cp_{\min}] \), \( p_r \in [p_{\min}, p_{\max}] \) and \( p_s \in [Cp_{\min}, p_{\max}] \), \( p_r \in [p_{\min}, p_{\max}] \). To calculate the optimal transmit power and utility functions in Case 2, it is necessary to apply Stackelberg game analysis within both rectangular sub-areas, i.e.,

1) If \( p_s \in [p_{\min}, Cp_{\min}] \), the first hop becomes the bottleneck.

2) If \( p_s \in [Cp_{\min}, p_{\max}] \), either hop could be the bottleneck.

In each sub-area, an optimal price that maximizes leader utility uniquely exists and the
SU transmit strategy can be uniquely decided. We compare Stackelberg game outcomes in both sub-areas and select the optimal transmit strategy \((p_s, p_r)\) that maximizes the utility functions as the final outcome of Case 2. The analysis for the above two scenarios is given as follows.

- **Scenario 1**

  In sub-area 1, the Stackelberg game model is the same as in Section 4.2.2, except that the upper bound for \(p_s\) is \(C_{p_{\text{min}}}\) instead of \(p_{\text{max}}\). Therefore, the optimal SU transmit power is searched within \([p_{\text{min}}, C_{p_{\text{min}}}]\), which is denoted by \(p_{s,2,1}^*\) in Scenario 1. The optimal relay transmit power \(p_{r,2,1}^*\) in Scenario 1 takes the minimum value \(p_{\text{min}}\), and the optimal charging price is denoted by \(\alpha_{d,2,1}^*\). The corresponding leader and follower utility functions in Scenario 1 are denoted by \(u_{l,r,1} = u_{l,r}(\alpha_{d,2,1}^*, p_{s,2,1}^*, p_{r,2,1}^*)\) and \(u_{f,r,1} = u_{f,r}(\alpha_{d,2,1}^*, p_{s,2,1}^*, p_{r,2,1}^*)\), respectively.

- **Scenario 2**

  When \(p_s \geq C_{p_{\text{min}}}\), either of the two hops could be the bottleneck. If the optimal transmit strategy satisfies \(p_s \leq C_{p_r}\), the first hop becomes the bottleneck and the follower utility becomes:

  \[
  u_{f,r,2}(\alpha_d, p_s, p_r) = \frac{1}{2} \log_2(1 + \frac{p_sw_{sr}}{p_bw_{br} + N_0}) - \alpha_d p_s w_{sp} - \alpha_d p_r w_{rp} \tag{4.38}
  \]

  Since \(p_r \geq \frac{p_s}{C}\), we know that \(-\alpha_d p_r w_{rp} \leq -\frac{\alpha_d p_s w_{sp}}{C}\), and we have \(u_{f,r,2} \leq r_1/2 - \alpha_d p_s \gamma\), where \(p_s \in [C_{p_{\text{min}}}, p_{\text{max}}]\) and

  \[
  \gamma := \frac{w_{sp} + w_{rp}}{C} \tag{4.39}
  \]

  Clearly, the follower utility is maximized when the relay transmit strategy \((p_s, p_r)\) is searched on the straight line \(p_s = C_{p_r}\), i.e., maximizing follower utility is equivalent to
maximizing the following:

$$\max \quad \frac{1}{2} \log_2(1 + \frac{p_s g_{sr}}{p_b g_{br} + N_0}) - \alpha d p_s \gamma$$  \hspace{1cm} (4.40)$$

On the other hand, if $p_s \geq C p_r$, the second hop is the bottleneck. Hence, the follower utility becomes:

$$u_{f,r,2}(\alpha_d, p_s, p_r) = \frac{1}{2} \log_2(1 + \frac{p_r g_{rs}}{p_b g_{bs} + N_0})$$

$$- \alpha d p_s g_{sp} - \alpha d p_r g_{rp}.$$  \hspace{1cm} (4.41)$$

Similarly, we know $-\alpha d p_s g_{sp} \leq -\alpha d p_r g_{rp} C$, and thus we have $u_{f,r,2} \leq r_2/2 - \alpha d p_r C \gamma$. The relay transmit strategy $(p_s, p_r)$ is searched on the straight line $p_s = C p_r$ as well, which is equivalent to maximizing the following:

$$\max \quad \frac{1}{2} \log_2(1 + \frac{p_r g_{rs}}{p_b g_{bs} + N_0}) - \alpha d p_r C \gamma$$  \hspace{1cm} (4.42)$$

It is easy to verify that Eq. (4.38) and (4.41) yield the same expression when $p_s = C p_r$. Therefore, the follower utility in Scenario 2 is maximized when $p_s = C p_r$. For simplicity, we substitute this relation into (4.38) and set its first order derivative with respect to $p_s$ as 0, i.e.,

$$\frac{\partial u_{f,r,2}}{\partial p_s} \Bigg|_{p_s = C p_r} = \frac{1}{2 \ln 2} \frac{g_{sr}}{p_s g_{sv} + p_b g_{br} + N_0} - \alpha d \gamma = 0,$$  \hspace{1cm} (4.43)$$

The optimal SU transmit power in Scenario 2 can be derived as:

$$p_{s,2,2} = \frac{1}{2 \alpha_d \gamma \ln 2} - \frac{p_b g_{br} + N_0}{g_{sr}}.$$  \hspace{1cm} (4.44)$$
and the corresponding optimal relay transmit power is:

\[ p_{r,2,2} = \frac{1}{2C \gamma \ln 2} - \frac{p_b g_{br} + N_0}{C g_{sr}} \]  
(4.45)

From the leader’s perspective, the charging price \( \alpha_d \) should be set neither too high nor too low to prevent insufficient outcome. Similar to the analysis in \[86\], \( \alpha_d \) is set so that 
\[ C p_{\min} \leq p_{s,2,2} \leq p_{\max} \]
holds, and the lower and upper bound for \( \alpha_d \) in Scenario 2 are denoted by \( \alpha_{d,\min,2,2} \) and \( \alpha_{d,\max,2,2} \), respectively. To derive the optimal \( \alpha_d \) in Scenario 2, we substitute \( p_{s,2,2} \) and \( p_{r,2,2} \) into the leader utility function (4.17), i.e.,

\[
u_{l,r,2}(\alpha_d) = \frac{\beta}{2 \ln 2} - \alpha_d \beta \gfrac{p_b g_{br} + N_0}{g_{sr}}
+ \log_2 \left[ 1 + p_b g_{bp,1} \left( \frac{g_{sp}}{2C \gamma \ln 2} - \frac{p_b g_{br} + N_0}{g_{sr}} + N_0 \right)^{-1} \right]
+ \log_2 \left[ 1 + p_b g_{bp,2} \left( \frac{g_{rp}}{2C \gamma \ln 2} - \frac{g_{rp} p_b g_{br} + N_0}{C g_{sr}} + N_0 \right)^{-1} \right],
\] 

(4.46)

To maximize the leader utility (4.46), we take its first order derivative and set it to zero, i.e.,

\[
\left. \frac{\partial \nu_{l,r,2}}{\partial \alpha_d} \right|_{p_s=C p_r} = \frac{2 A_1 B_1^2 E_1^2}{(C_1 \alpha_d + B_1 E_1) [ (A_1 + C_1) \alpha_d + B_1 E_1 ]}
+ \frac{2 A_2 B_1^2 E_2^2}{(C_2 \alpha_d + B_1 E_2) [ (A_2 + C_2) \alpha_d + B_1 E_2 ]}
+ \beta (C_1 + C_2 - 2N_0) = 0,
\] 

(4.47)
where $C_1$ is as defined in (4.34). The other parameters are defined as follows:

\[ A_2 := p_b g_{2p,2}, \]
\[ C_2 := -g_r p_b g_{br} + N_0 + N_0, \]
\[ E_1 := \frac{g_s p}{\gamma}, E_2 := \frac{g_r p}{C \gamma} \]  

(4.48)

It is easy to verify that $\partial u_{l,r,2}(\alpha_d)/\partial \alpha_d = 0$ yields a quartic polynomial equation, i.e.,

\[ F_4 \alpha_d^4 + F_3 \alpha_d^3 + F_2 \alpha_d^2 + F_1 \alpha_d + F_0 = 0, \]  

(4.49)

where:

\[ F_4 := \beta(C_1 + C_2 - 2N_0)C_1 C_2 (A_1 + C_1)(A_2 + C_2), \]
\[ F_3 := \beta(C_1 + C_2 - 2N_0)[C_1 (A_1 + C_1)B_1 E_1 (A_2 + 2C_2) \]
\[ + C_2 (A_2 + C_2)B_1 E_1 (A_1 + 2C_1)], \]
\[ F_2 := \beta(C_1 + C_2 - 2N_0)[B_1^2 E_1 E_2 (A_1 + 2C_1)(A_2 + 2C_2) \]
\[ + C_2 B_1^2 E_1^2 (A_2 + C_2) + C_1 B_1^2 E_2^2 (A_1 + C_1)] \]
\[ + 2A_1 B_1^2 E_1^2 C_2 (A_2 + C_2) + 2A_2 B_1^2 E_2^2 C_1 (A_1 + C_1), \]
\[ F_1 := \beta(C_1 + C_2 - 2N_0)[B_1^3 E_1 E_2^2 (A_1 + 2C_1) \]
\[ + B_1^3 E_2^2 E_2 (A_2 + 2C_2)] + 2A_1 B_1^3 E_1^2 E_2 (A_2 + 2C_2) \]
\[ + 2A_2 B_1^3 E_1 E_2^2 (A_1 + 2C_1), \]
\[ F_0 := [\beta(C_1 + C_2 - 2N_0) + 2(A_1 + A_2)]B_1^4 E_1^2 E_2^2 \]  

(4.50)
It is known that an analytical expression for the roots exist in the quartic equation. For more details about analytical root expressions, please refer to [62]. According to quartic function properties, there are at most two real roots that indicate local maxima in (4.49), and we denote them by \( \hat{\alpha}_{d,2,2,1}, \hat{\alpha}_{d,2,2,2} \). On the other hand, as the charging price is always positive, the positive real roots of (4.49) need to be searched within the support \([\alpha_{d_{\text{min}},2,2}, \alpha_{d_{\text{max}},2,2}]\). Hence, the optimal price \( \alpha_{d,2,2}^* \) that maximizes the leader’s utility in Scenario 2 is searched from the set \( \{\hat{\alpha}_{d,2,2,1}, \hat{\alpha}_{d,2,2,2}, \alpha_{d_{\text{min}},2,2}, \alpha_{d_{\text{max}},2,2}\} \), \( i = 1, 2 \) if \( \alpha_{d_{\text{min}},2,2} \leq \hat{\alpha}_{d,2,2,2} \leq \alpha_{d_{\text{max}},2,2}, i = 1, 2 \). Otherwise, the optimal price in Scenario 2 is searched from \( \{\alpha_{d_{\text{min}},2,2}, \alpha_{d_{\text{max}},2,2}\} \). The optimal SU transmit power in Scenario 2 is given by:

\[
p_{s,2,2}^* = \frac{1}{2 \ln 2} \cdot \frac{1}{\alpha_{d,2,2}^* \gamma} - \frac{p_{b} g_{br} + N_0}{g_{sr}}, \tag{4.51}
\]

and the corresponding optimal relay node transmit power is \( p_{r,2,2}^* = p_{s,2,2}^*/C \). The optimal leader and follower utilities in Scenario 2 are given by \( u_{l,r,2} = u_{l,r}(\alpha_{d,2,2}^*, p_{s,2,2}^*, p_{r,2,2}^*) \) and \( u_{f,r,2} = u_{f,r}(\alpha_{d,2,2}^*, p_{s,2,2}^*, p_{r,2,2}^*) \), respectively. In this work, \( \hat{\alpha}_{d,2,2,2} \), \( i = 1, 2 \) are derived using a polynomial equation solver in Matlab.

**Optimal Strategy in Case 2**

The optimal strategy in Case 2 can be derived by summarizing Scenarios 1 and 2. Since the leader attempts to profit as much as possible by sharing its licensed channels to the follower, we select the Stackelberg outcomes from Scenarios 1,2 such that the leader utility in Case 2 is maximized, i.e.,

\[
i^* = \arg \max_{i \in \{1, 2\}} u_{l,r}(\alpha_{d,2,i}^*, p_{s,2,i}^*, p_{r,2,i}^*) \tag{4.52}
\]

and the optimal charging price, SU transmit power and relay node transmit power in Case 2 are denoted by \( \alpha_{d,2}^*, p_{s,2}^*, p_{r,2}^* \), respectively. The leader and follower utilities in Case 2 are
given by \( u_{l,r,2} = u_{l,r}(\alpha_{d,2}^*, p_{s,2}^*, p_{r,2}^*) \) and \( u_{f,r,2} = u_{f,r}(\alpha_{d,2}^*, p_{s,2}^*, p_{r,2}^*) \), respectively.

### 4.2.4 Optimal DF Transmission Strategy

As discussed in Section 4.2.1, the only difference between Case 1 and 4, Case 2 and 3 is the rate bottleneck in DF cooperative relaying. Thus, by swapping the corresponding channel parameters of the two hops in the follower in pairwise, i.e.,

\[
g_{bp,1} \leftrightarrow g_{bp,2}, g_{sr} \leftrightarrow g_{rs}, g_{sp} \leftrightarrow g_{rp}, g_{br} \leftrightarrow g_{bs}
\]

the Stackelberg game Nash equilibrium of Case 3 and 4 can be derived in the same way as Case 2 and 1, respectively. Using the approach proposed in Section 4.2.3 and 4.2.2, Stackelberg game outcomes can be derived for Cases 3 and 4, which are denoted by tuples \( (\alpha_{d,3}^*, p_{s,3}^*, p_{r,3}^*) \) and \( (\alpha_{d,4}^*, p_{s,4}^*, p_{r,4}^*) \), respectively. The follower and leader utilities in Cases 3 and 4 are denoted by

\[
\begin{align*}
    u_{f,r,3} &= u_{f,r}(\alpha_{d,3}^*, p_{s,3}^*, p_{r,3}^*), \\
    u_{l,r,3} &= u_{l,r}(\alpha_{d,3}^*, p_{s,3}^*, p_{r,3}^*), \\
    u_{f,r,4} &= u_{f,r}(\alpha_{d,4}^*, p_{s,4}^*, p_{r,4}^*), \\
    u_{l,r,4} &= u_{l,r}(\alpha_{d,4}^*, p_{s,4}^*, p_{r,4}^*).
\end{align*}
\]

respectively. The optimal DF transmission strategy is selected from the tuples \( (\alpha_{d,i}^*, p_{s,i}^*, p_{r,i}^*) \) while the follower and leader utilities are selected from \( \{u_{f,r,i}, u_{l,r,i}\} \) for \( i \in \{1, 2, 3, 4\} \), depending on the channel parameters. For simplicity, the optimal utility of the follower and the corresponding SU rate in DF scheme are denoted by \( u^{df} \) and \( C^{df} \), respectively.
4.2.5 Optimal Direct Transmission (DT) Strategy

As discussed earlier, the SU pairs could choose DT scheme as long as they satisfy distance constraints. Therefore, the same Stackelberg game which employs a single PU and a single SU pair as the leader-follower pair is still adopted to derive the optimal strategy in DT transmission. The leader utility and follower utility are given by:

\[ u_{l,r}(\alpha_d, p_s) = \log_2(1 + \frac{p_b g_{bp}}{p_s g_{sp} + N_0}) + \alpha_d \beta p_s g_{sp} \tag{4.55} \]

and

\[ u_{f,r}(\alpha_d, p_s) = \log_2(1 + \frac{p_s g_{ss}}{p_b g_{bs} + N_0}) - \alpha_d p_s g_{sp} \tag{4.56} \]

where the optimal SU transmit power \( p_s \) is subsequently determined by the charging price \( \alpha_d \). The Stackelberg game Nash equilibrium \( \{\alpha^*_d, p^*_s\} \) is achieved when either the leader or the follower gains no more revenue by changing their strategies unilaterally, i.e.,

\[ u_{f,r}(\alpha^*_d, p^*_s) \geq u_{f,r}(\alpha_d, p_s) \tag{4.57} \]

\[ u_{l,r}(\alpha^*_d, (p^*_s(\alpha_d))) \geq u_{l,r}(\alpha_d, p^*_s(\alpha_d)) \tag{4.58} \]

More details can be found in [86]. The optimal utility of the follower and the corresponding SU rate in DT scheme are denoted by \( u^{dt} \) and \( C^{dt} \), respectively.

4.3 Joint Allocation and Scheduling Algorithm

In the proposed Stackelberg game framework, for any given devices and configurations of the system, the optimal transmission of the follower can be uniquely decided. The PUs, SUs and relay nodes in the entire CRN are then scheduled to transmit signals in each TTI. Therefore, a scheduling algorithm needs to be conducted to properly map the devices into Stackelberg game leader-follower pairs in a number of consecutive TTIs. In
our work, a hybrid priority-based scheduling algorithm that jointly assigns relay nodes and allocates channels to each SU pair regarding DF cooperative relaying is proposed. Consider the deployment of the devices in the CRN, the SU pairs may either select DT or DF transmission depending on device geographical locations and the channel parameters. As utility functions indicate the satisfaction of the devices in Stackelberg game, the proposed algorithm is designed as a hybrid priority-based approach, in which the priorities of the devices all depend on follower utilities. The follower utility definition in DT is directly inherited from [86] as no relay node is used in such a scheme, while the follower utility definition in DF cooperative relaying is described in Section 4.2.1. Since the follower utilities in DT and DF transmission scheme cannot be compared directly, we maintain two separate queues for DT and DF transmission priorities denoted as queue 1 and queue 2, respectively. Each queue contains all the possible leader-follower combinations in DT and DF schemes respectively, and these combinations are sorted by their priorities in descending order.

We assume that the channel parameters among a small number of consecutive TTIs keep constant throughout the scheduling process. The fairness among different SU pairs is also taken into account to balance their transmission demand. If a SU pair has been scheduled in previous TTIs, then it should be assigned a lower priority so that it is less likely to be selected in current TTI. In both DT and DF queues, if a SU pair has been scheduled in previous TTIs, some penalties will be added to the corresponding follower priorities. As described in [86], the penalties can be decided by the cummulative follower utilities. This type of definition could be applied to the priorities of both queues. Besides, if a certain SU pair has been scheduled in previous TTIs, the priorities of the leader-follower pairs in both queues that contain the above SU pair will be affected as well. Therefore, we define the priorities as follows:
We first define two sets of binary indicator variables which will be recorded in each TTI:

\( \{ x_{i,k}(t) \}, i \in \mathcal{M}, k \in \mathcal{N} \)

\( \{ y_{i,j,k,l}(t) \}, i, j \in \mathcal{M}, k \in \mathcal{N}, l \in \mathcal{R} \)

where \( x_{i,k}(t) = 1 \) indicates that \( k \)th SU pair is sharing the channel provided by \( i \)th PU at the \( t \)th TTI and \( x_{i,k}(t) = 0 \) otherwise. Similarly, \( y_{i,j,k,l}(t) = 1 \) indicates that the \( k \)th SU uses the \( l \)th relay node to support DF transmission and access the channels provided by \( i \)th and \( j \)th PU at the \( t \)th TTI, and \( y_{i,j,k,l}(t) = 0 \) otherwise. As mentioned in Section 4.1, \( M, N, R \) denote the number of PUs, SU pairs and all available relay nodes, respectively. We use \( \mathcal{M}, \mathcal{N}, \mathcal{R} \) to denote the sets of PUs, SU pairs and relay nodes for simplicity. The priorities in queue 1 and 2 at the \( t \)th TTI are defined as follows:

\[
P_{i,k}^{\text{dt}}(t) = u_{i,k}^{\text{dt}}(t) - q_{k}^{\text{dt}}(t), t \geq 0; \quad (4.59)
\]

\[
P_{i,j,k,l}^{\text{df}}(t) = u_{i,j,k,l}^{\text{df}}(t) - q_{k,l}^{\text{df}}(t), t \geq 0. \quad (4.60)
\]

where \( u_{i,k}^{\text{dt}}(t), u_{i,j,k,l}^{\text{df}}(t) \) are the follower utilities in DT and DF transmission scheme, respectively. \( q_{k}^{\text{dt}}(t) \) and \( q_{k,l}^{\text{df}}(t) \) are the penalties in the two queues, respectively. They are accumulative with time and need to be updated per TTI, as all previous schedules of corresponding followers lead to lower priorities. The corresponding link capacities in DT and DF transmission are denoted by \( C_{i,k}^{\text{dt}} \) and \( C_{i,j,k,l}^{\text{df}} \). For the ease of representation, we denote the leader-follower pairs in (4.59) as \((i, k)_{\text{dt}}\) and \((i, j, k, l)_{\text{df}}\), respectively. Therefore, the penalty received by \((i, k)_{\text{dt}}\) at the \( t \)th TTI in DT scheme is defined as:

\[
q_{k}^{\text{dt}}(t) = \sum_{\tau=0}^{t-1} \sum_{i=1}^{M} \delta_{\text{dt}} \cdot x_{i,k}(\tau) \cdot u_{i,k}(\tau) \quad (4.61)
\]

where \( \delta_{\text{dt}} \) is a coefficient that decides the influence of previous channel assignments on
priorities in current TTI. On the other hand, if a leader-follower pair in DF scheme has been scheduled in previous TTIs, the followers that involves corresponding devices will all suffer from such penalties. We assume a leader-follower pair that consists PU $i,j$, SU pair $k$ and relay node $l$ at the $t$th TTI. If PU $i$ or PU $j$ has been scheduled in previous TTIs, some penalties will be added to the priority of $(i,j,k,l)_{df}$. Furthermore, if both PU $i,j$ has been scheduled in previous TTIs, all leader-follower pairs containing both PU$i,j$ will receive a higher penalty in current TTI. Therefore, we define the penalty received by leader-follower pair $(i,j,k,l)_{df}$ as:

$$q_{i,j,k,l}^{df}(t) = t - 1 \sum_{\tau=0}^{t-1} \delta_{df} \left( \sum_{i'=1}^{M} y_{i',j,k,l}(\tau) u_{i',j,k,l}^{df}(\tau) + \sum_{j'=1}^{M} y_{i,j',k,l}(\tau) u_{i,j',k,l}^{df}(\tau) \right)$$ (4.62)

where $\delta_{df}$ is similar to $\delta_{dt}$ and denotes the importance of previous scheduling results in DF scheme. According to the analysis above, we initialize two separate queues by calculating the priorities for all possible combinations of leader-follower pairs and sort them in descending order for both DT and DF queues at $t = 0$. During the hybrid scheduling process, we take the head of the two queues and compare their link capacities in each TTI. The leader-follower pair with a larger link capacity is scheduled and all the devices scheduled are recorded. Naturally, if a leader-follower pair in DT scheme is scheduled, the corresponding devices could not be used in DF scheme and vice versa. The two indicator variables $x_{i,k}(t), y_{i,j,k,l}(t)$ need to be updated in each TTI, and the algorithm is summarized in Algorithm 4.3.

It is also worth noting that if $(i,j,k,l)_{df}$ is scheduled, all the leader-follower pairs containing either PU $i$ or PU $j$ will be penalized. Therefore, we also set $y_{j,i,k,l} = 1$ in Algorithm 4.3 when updating. Once the optimal leader-follower pair among the two queues is selected, we delete all the leader-follower pairs that use any of the recorded devices in each queue. The scheduling is executed until all PUs or SU pairs are scheduled in each TTI. After this, the priorities of all the leader-follower pairs in DT and DF transmission are updated, and both queues are sorted in descending order based on the new priorities for the coming TTI. The entire scheduling process terminates when the predetermined number of TTIs, $N_{TTI}$,
is met. Hence, the full scheduling algorithm is summarized in Algorithm 2. If fairness is not taken into account in the scheduling process, we can simply set $\delta_{df} = \delta_{dt} = 0$.

---

**Algorithm 1** Update $x_{i,k}(t)$ and $y_{i,j,k,l}(t)$

1: Given the scheduled leader-follower pair $(i, k)_{dt}$ or $(i, j, k, l)_{df}$

2: **if** $(i, k)_{dt}$ is scheduled **then**

3: Set $x_{i,k} = 1$

4: **for** $k \in M$ **do**

5: **for** $l \in R$ **do**

6: Set $y_{i,i,k,l} = 1$

7: **else**

8: Set $x_{i,k} = 1, x_{j,k} = 1$

9: **for** $k \in M$ **do**

10: **for** $l \in R$ **do**

11: Set $y_{i,j,k,l} = 1, y_{j,i,k,l} = 1$

---

**Algorithm 2** Hybrid Scheduling Algorithm Regarding Fairness

1: Initialize $P_{dt}^{i,k}(0)$ and calculate $C_{dt}^{i,k}, \forall i, k$

2: Initialize $P_{df}^{i,j,k,l}(0)$ and calculate $C_{df}^{i,j,k,l}, \forall i, j, k, l$

3: Initialize $t = 0$

4: **while** $t < N_{TTI}$ **do**

5: **repeat**

6: Select queue heads $(i_1, k_1)_{dt}$ and $(i_2, j_2, k_2, l_2)_{df}$

7: **if** $C_{dt}^{i_1,k_1} > C_{df}^{i_2,j_2,k_2,l_2}$ **then**

8: Schedule (PU $i_1$, SU $k_1$)

9: Delete all pairs containing PU $i_1$ or SU $k_1$ from queues 1 and 2

10: Update $x_{i,k}(t), y_{i,j,k,l}(t)$ using Algorithm 4.3

12: **else**

13: Schedule [(PU $i_2$, PU $j_2$), (SU $k_2$, Relay $l_2$)]

14: Delete all pairs containing PU $i_2$, PU $j_2$, SU $k_2$, or relay $l_2$ from queues 1 and 2

16: Update $x_{i,k}(t), y_{i,j,k,l}(t)$ using Algorithm 4.3

17: **until** all PUs or SUs pairs have been scheduled

18: $t = t + 1$

19: Update $P_{dt}^{i,k}(t), P_{df}^{i,j,k,l}(t), \forall i, j, k, l$

20: Sort queue 1 and 2 in descending order
4.4 Experimental Results and Analysis

In this section, we present our analysis to the outcomes of the proposed Stackelberg game framework in this chapter. The results and analysis to the hybrid scheduling algorithm are presented as well.

4.4.1 Parameter Settings

We assume that the BS is located at the center of the cell, while all the PUs, SU pairs and relay nodes are deployed according to a uniform distribution in the cell. We assume the threshold $D_{\text{max}}$ is the maximum distance for DT between any pair of nodes. Naturally, the two hops in DF transmission must satisfy the maximum distance constraint simultaneously. Given a transmitter and a receiver, the received signal power is $P_R = P_T g_{TR} = P_T |h_{TR}|^2/d_{TR}^2$, where $P_T$ and $P_R$ are the transmit and received power of the devices respectively, and $d_{TR}$ is the distance. The coefficient $h_{TR}$ follows a complex Gaussian distribution $\mathcal{CN}(0, 1)$. The parameters used in the simulation are summarized in Table 4.1.

We also set the number of SUs large enough so that it could support all the SU pairs are deployed such that each of them can use direct transmission, while the relay nodes are randomly deployed in the network coverage area. Therefore, only the relay nodes that supports DF transmission for at least one SU pair are available in the network model. The actual number of available relay nodes is 130 out of a total of 200 in Table 4.1. The deployment of the devices in the proposed CRN is shown in 4.3, where the red dot, the black dots and the blue dots denote the BS, PUs and relay nodes, respectively. The black circles connected by red dashed lines denote the SU pairs with certain traffic demand.

4.4.2 Performance Study of DF Transmission

We first focus on a single SU pair under DF cooperative relaying and make comparisons to DT scheme. The distance between SU transmitter and receiver is assumed to be the maximum value allowed (i.e., $D_{\text{max}}$), while they have equal distance to the BS. We set the coordinates of the SU transmitter, the SU receiver and the BS as (-25 m,0), (25 m,0) and
Figure 4.3: Device Deployment in the Proposed CRN Model.

(0,50 m) on the same plane, respectively. The distance between SU transmitter and receiver is $D_{\text{max}} = 50$ m and the distance between BS and SU devices is 55.90 m. The relay node is located on the straight line between SU transmitter and receiver, and its coordinate can be denoted by $(R_0, 0)$, $-D_{\text{max}}/2 < R_0 < D_{\text{max}}/2$. We further assume the same group of channel coefficients is used in the simulation. The SU link capacities are derived from Stackelberg game equilibria and then compared to the link capacity under DT, which is a constant under the above assumptions. However, the link capacity will change along with the position of the relay node. Thus, we first find out the maximum link capacity by setting the relay node on a set of equally spaced locations. We further define the capacity
Table 4.1: Simulation Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Radius $R$</td>
<td>100 m</td>
</tr>
<tr>
<td>BS Transmit Power</td>
<td>23 dBm</td>
</tr>
<tr>
<td>SU Transmit Power Range</td>
<td>0 – 23 dBm</td>
</tr>
<tr>
<td>Relay Node Transmit Power Range</td>
<td>0 – 23 dBm</td>
</tr>
<tr>
<td>Maximum Direct Transmission Distance $D_{\text{max}}$</td>
<td>50 m</td>
</tr>
<tr>
<td>Number of PUs</td>
<td>16</td>
</tr>
<tr>
<td>Number of SU pairs</td>
<td>8</td>
</tr>
<tr>
<td>Number of Relay Nodes</td>
<td>200</td>
</tr>
<tr>
<td>Noise Power Density</td>
<td>-174 dBm/Hz</td>
</tr>
<tr>
<td>Bandwidth BW</td>
<td>180 kHz</td>
</tr>
<tr>
<td>Transmission Time Interval (TTI)</td>
<td>1 ms</td>
</tr>
</tbody>
</table>

ratio as the ratio of maximum DF link capacity to DT link capacity. The parameter $\beta$ in utility functions is selected from \{1, 2, 5, 10\} and we run 1,000 independent simulations for each $\beta$. Since some capacity ratios could be large, we consider the logarithm of the capacity ratios, i.e., $\log\left(\frac{\max\{r_1(R_0), r_2(R_0)\}}{r_{\text{dt}}(R_0)}\right)$. $r_1(R_0), r_2(R_0)$ are the rate of the first hop and the second hop in DF transmission, respectively. $r_{\text{dt}}$ is the capacity of direct transmission. The cumulative distribution function (CDF) for each $\beta$ is shown in Fig. 4.4. When $\beta$ is relatively small, i.e. $\beta = 1$, the probability to have a lower maximum capacity compared to DT (i.e., the logarithm is below zero) is around 20%. Such a probability will reach 30% and 60% as $\beta = 5, 10$, respectively, meaning that DF transmission is more likely to show worse performance than DT when $\beta$ becomes large. Yet when $\beta$ grows, DF transmission is more likely to achieve higher transmission rate and result in better overall performance in some cases.

Next we study the rate distributions of SU pairs and PUs under different values of $\beta$ by fixing the relay node at the middle, i.e. the coordinates of the relay yield (0, 0). For $\beta = 1, 2, 5, 10$, we run the simulation 2,000 times under different groups of channel coefficients. The SU link rate distributions are given in Fig. 4.5. We can see that the follower’s relay transmission rate increases as $\beta$ grows, since the follower is prone to select
Figure 4.4: CDF for logarithm of capacity ratio under different values of $\beta$.

4.4.3 Experimental Study of Hybrid Scheduling Algorithm

In this section, we study the performance of the proposed hybrid scheduling algorithm. We first consider the case that the number of PUs $M$ is large enough to support all transmission schemes (i.e., $M \geq 2N$), and the number of PUs, SU pairs and relay nodes are given in Table 4.1. We assume that each SU pair satisfies a maximum distance constraint to adopt DT. Based on the above assumptions, some of the relay nodes could be adopted by multiple SU pairs for DF relay transmission. Once such a relay node is assigned to a certain SU pair, it cannot be adopted by any other SU pairs for DF transmission in the same TTI.

We first study the cumulative SU sum rate by executing the hybrid scheduling algorithm in a single TTI. Therefore, fairness is not taken into account in this scenario. The DT
scheduling algorithm in [86] is also performed to make a comparison. The results for the proposed hybrid scheduling and DT scheduling are given in Fig. 4.6. It can be seen that DF transmission always achieves higher sum rate compared to DT under the same value of $\beta$. As $\beta$ grows, the follower is likely to buy larger SU and relay transmit power due to their relatively low cost, which results in a higher SU sum rate. Yet when $\beta$ is sufficiently large, very little sum rate improvement can be achieved using DF transmission compared to DT scheme. This is due to the fact that when the follower uses higher transmit power, the probability of achieving higher rate in DF transmission significantly decreases, which can be observed in Fig. 4.4. Since only a small number of relay nodes could be adopted by each SU pair, it is very likely that none of the relay nodes would be scheduled during the process. In this case, the performance of hybrid scheduling is almost the same as that of
Next we study the average sum rate over a number of consecutive TTIs, where the scheduling algorithm is conducted in each TTI. If two PUs $i, j$, a SU pair $k$ and a relay node $l$ are scheduled under DF transmission in some TTI $t$, then PU $j$ is preserved and cannot be scheduled in the following TTI $t + 1$ as it provides channel for the second hop. Therefore, we can skip all the leader-follower pairs that either contain PU $j$ or relay node $l$ in the two queues in TTI $t + 1$. Yet SU transmitter $k$ could still be scheduled under either DT or DF scheme as long as it accesses a channel other than channel $i$. Based on the considerations above, we design the following experiments.

In the first experiment, we do not consider the fairness among different devices, i.e., $\delta_{df} = \delta_{df} = 0$. Assume that the maximum number of TTIs is $N = 20$ and we use pipelining
as described above in the whole scheduling process. The relationship between the average sum rate and the number of consecutive TTI is given in Fig. 4.7. Clearly, DT scheduling outcomes is always the same under the same $\beta$ in each TTI, as DT scheduling does not preserve any PUs for the following TTI, while the SU pairs receive no penalty from previous scheduling outcomes. As the number of TTIs increases, the average sum rate tends to converge under all values of $\beta$ and significant improvement is made when $\beta$ is relatively small. To protect the PUs from harmful SU interference, a small to moderate value of $\beta$ should be selected.
In the second experiment, we take fairness into account so that the previous scheduling outcomes will affect the scheduling in current TTI. The value of $\beta$ is set to 1 and we test different values for the fairness coefficients $\delta_{dt}, \delta_{df}$. We also set $\delta_{dt} = \delta_{df} = \delta, \delta \in \{0.01, 0.02, 0.1, 0.2\}$. The total number of TTIs is set to 50 when running the scheduling algorithm. The average sum rate under different values of the coefficients are depicted in Fig. 4.8. It can be observed that the average rates in both DT and DF schemes are all significantly affected by the penalty and begin decreasing after some TTIs for all values of $\delta$. Clearly, the larger $\delta$ is, the higher the penalty becomes, and the earlier the average sum rate begin to decrease. This is due to the fact that the leader-follower pairs with higher priorities suffer from higher penalty when $\delta$ is large, and those leader-follower pairs with lower initial priorities are scheduled after a number of TTIs. It can also be observed that the proposed hybrid scheduling approach maintains advantages comparing to DT scheduling for all values of $\theta$.

In the third experiment, we consider the case in which the available spectrum resource is scarce, i.e., the number of PUs $M$ is smaller than the number of SU pairs $N$. The value $M$ is set to 6 so that not all the SU pairs can be scheduled in the same TTI, while the other parameter settings are the same in experiment 2. The average sum rate under different values of the coefficients are depicted in Fig. 4.9. Apparently, the average sum rate is smaller compared to experiment 2 as much less SU links could be supported. As the value of $\delta$ grows, hybrid scheduling achieves more significant improvement in contrast to DT scheme in terms of sum rate. The decrease of average sum rate is slower compared to DT as well. For $\delta = 0.2$, the average sum rate even rises slightly after 20 TTIs. The reason for the above observations is that the leader-follower pairs with relatively high priorities will not be scheduled in a same TTI, and the penalties received by each SU pair is likely to be smaller. Therefore, hybrid scheduling algorithm shows eminent performance under scarce channel resources.
Figure 4.8: Average sum rate versus TTIs regarding fairness ($\beta = 1$).

4.5 Conclusions

In this chapter, we proposed a novel Stackelberg game framework for SU DF cooperative relaying in a cognitive radio network. Under the proposed framework, we analyzed the SNE of the Stackelberg game and derived the optimal transmission strategies for SU relay pairs. The simulation result demonstrates that a significant increase in SU link capacity can be achieved under the proposed DF cooperative relaying scheme compared to direct transmission. It also shows that the scale factor $\beta$ has a significant impact on the achievable link capacity. Based on the Stackelberg game outcomes, we proposed a hybrid scheduling algorithm combining SU direct transmission and DF relay transmission regarding fairness among different devices. The experimental results show that good sum rate improvement
can be achieved under small to moderate values of $\beta$. Considering small $\beta$ value and fairness, experimental results show the clear advantages of the proposed scheduling algorithm in terms of sum rate, especially under scarce channel resources.
Chapter 5: Hierarchical Stackelberg / Mean Field Game for Distributed Power Control in Dense Cooperative Relaying Networks

Device-to-device (D2D) communications are widely adopted as the underlay to cellular primary networks [28, 29]. Assisted by cognitive radio and spectrum sensing technologies, D2D users can significantly enhance system spectrum efficiency by reusing the licensed spectrum and directly transmitting signals to the proximity devices [76, 107]. However, D2D communications usually face the restrictions of transmission range and limited energy supply. Besides, mutual interference among different devices will further increase energy consumption and thus exacerbate communication QoS (quality-of-service) degradation [77, 106].

In this chapter, we consider a scenario in which a large number of dense deployed D2D links attempt to transmit under cooperative relaying simultaneously. A hierarchical game framework which consists of both a Stackelberg game and a mean field game (MFG) for distributed power control is proposed. The main contribution of this work is a hierarchical Stackelberg/mean-field game framework which models the interactions among D2D pairs or devices and gives an optimal distributed power control policy taking into account both D2D transmitters and relay nodes. In this framework, the number of D2D pairs is assumed to be a large number. A Stackelberg game is first adopted to model the relationship between a D2D transmitter and the corresponding relay node under cooperative relaying to decide their transmission strategies, while the mean field game in the higher hierarchy gives the power control policy based on the Stackelberg game outcomes. In the Stackelberg game, the D2D transmitter and the relay node act as a leader-follower pair and they both seek to maximize their own utilities. After this, the MFG is adopted to model the interactions
between each D2D pair and the statistical interference coming from other D2D transmitters and relay nodes \cite{3,37,41,51}. In the proposed game framework, the distributed power control problem can be interpreted as a cost function minimization issue, while the constraints of available energy in D2D devices needs to be taken into consideration. The Nash equilibrium of the proposed game framework can be derived using a finite-step differential approach. A detailed theoretical analysis of the above hierarchical game framework is presented.

The remainder of the chapter is organized as follows. In Section \ref{sec:5.1} we propose the system model and formulate the power control problem in a dense D2D cooperative relaying network. In Section \ref{sec:5.2} we analyze the interactions within a D2D pair under a Stackelberg game framework. In Section \ref{sec:5.3} we develop a mean-field game analysis of the D2D pair under cooperative relaying in a dense D2D network. The distributed power control scheme based on the hierarchical game framework is analyzed in Section \ref{sec:5.4}.

\section{5.1 System Model}

\subsection{5.1.1 Network Model}

We consider a dense D2D network macrocell in which a large number of D2D pairs are deployed and share uplink spectrum resources of macro user equipments (MUEs) \cite{9,39,47,53,78,94}. Each D2D pair is assumed to adopt cooperative relaying to increase its transmission range. A macrocell eNodeB is assumed to be located in the network, which can communicate directly with the MUEs. In each D2D pair, a D2D transmitter aims to transmit signal to its corresponding receiver. The MUEs use predetermined channels with equal bandwidth to transmit signals. The number of D2D pairs is denoted by \( N \), and each D2D pair adopts only one relay node when transmitting. The relay nodes are also densely deployed D2D devices that can transmit directly to other D2D users, and we assume that the number of relay nodes is also \( N \). We assume full frequency reuse for all D2D transmitters and relay nodes while they do not interfere with each other, i.e., all D2D transmitters reuse the same channel and all relay nodes reuse another common channel. To satisfy such
demand, two MUEs provide their channels to allow D2D transmitters and relay nodes to transmit. In this case, D2D transmitters would only cause interference to relay nodes in other D2D pairs, while relay nodes will only cause interference to D2D receivers in other D2D pairs. In D2D cooperative relaying, a decode and forward (DF) scheme is assumed such that the signal received by the relay will be fully decoded before it is forwarded to the receiver. Thus, a D2D receiver will not receive any signal directly from its corresponding transmitter. Here we use $D_{T,i}$, $D_{R,i}$ and $R_i$ to denote the transmitter, the receiver and the relay node respectively in the $i$th D2D pair under cooperative relaying, and the proposed D2D cooperative relaying network is depicted in Fig. 5.1 in which the solid lines with arrows denote the signals and the dashed lines with arrows denote the interference. We also assume that all D2D transmitters access the channel of MUE 1 and all relay nodes access the channel of MUE 2 in the MUE uplink transmission phase.

Among the D2D network devices, the interference of $i$th D2D transmitter/relay node to all other D2D pairs and the interference received by $i$th D2D transmitter/relay node from all other D2D devices are termed as intra-tier and inter-tier interference, respectively. At some time instant $t$, the intra-tier interference introduced by $i$th D2D transmitter and $i$th relay node to all other D2D devices are given by:

$$I_{i,D}(t) = p_{d,i}(t) \sum_{j=1, j \neq i}^{N} g_{dr,i,j}(t), \quad (5.1)$$

and

$$I_{i,R}(t) = p_{r,i}(t) \sum_{j=1, j \neq i}^{N} g_{rd,i,j}(t), \quad (5.2)$$

respectively, where $p_{d,i}(t), p_{r,i}(t)$ are the transmit power of the $i$th D2D transmitter $D_{T,i}$ and the $i$th relay node $R_i$, respectively. Here, $g_{dr,i,j}(t), g_{rd,i,j}(t)$ denote the channel gain from $i$th D2D transmitter to $j$th relay node, and the channel gain from $i$th relay node to
Figure 5.1: Dense D2D cooperative relaying network model.

\[ I_{R,i}(t) = \sum_{j=1, j \neq i}^{N} p_{d,j}(t) g_{dr,j,i}(t), \]  (5.3)

where \( p_{d,j}(t) \) is the transmit power of the \( j \)th D2D transmitter \( D_j^T \), and \( g_{dr,j,i}(t) \) denotes the channel gain from the \( j \)th D2D transmitter \( D_j^T \) to \( i \)th relay node \( R_i \). Similarly, the
inter-tier interference received by D2D receiver $D_{R,i}$, denoted by $I_D(t)$, can be formulated as:

$$I_{D,i}(t) = \sum_{j=1, j\neq i}^{N} p_{r,j}(t)g_{rd,j,i}(t), \quad (5.4)$$

where $p_{r,j}(t)$ is the transmit power of the $j$th relay node $R_i$, and $g_{rd,j,i}(t)$ denotes the channel gain from $j$th relay node $R_j$ to $i$th D2D receiver $D_{R,i}$.

Note that D2D transmitters and relay nodes access different MUE channels when transmitting, the signal in the first hop will not be directly received by the corresponding D2D receiver. Under DF transmission, the noise and interference received by the relay node will be removed from the signal before next hop. Therefore, the signal power received by the D2D receiver comes from the relay node. We assume that $p_{d,i}(t), p_{r,i}(t)$ are the transmit power of the $i$th D2D transmitter $D_{T,i}$ and the $i$th relay node $R_i$, respectively. The signal to interference-plus-noise ratio (SINR) received by relay node $R_i$ can be written as:

$$\gamma_{R,i}(t) = \frac{p_{d,i}(t)g_{dr,i,i}(t)}{I_{R,i}(t) + \sigma^2} = \frac{p_{d,i}(t)g_{dr,i,i}(t)}{\sum_{j=1, j\neq i}^{N} p_{d,j}(t)g_{dr,j,i}(t) + \sigma^2} \quad (5.5)$$

where $g_{dr,i,i}(t)$ is the channel gain between D2D transmitter $D_{T,i}$ and relay node $R_i$, and $\sigma^2$ is the channel noise power. The SINR at the $i$th D2D receiver at time $t$ yields:

$$\gamma_{D,i}(t) = \frac{p_{r,i}(t)g_{rd,i,i}(t)}{I_{D,i}(t) + \sigma^2} = \frac{p_{r,i}(t)g_{rd,i,i}(t)}{\sum_{j=1, j\neq i}^{N} p_{r,j}(t)g_{rd,j,i}(t) + \sigma^2} \quad (5.6)$$
where $g_{rd,i,i}(t)$ is the channel gain between relay node $R_i$ and D2D receiver $D_{R_i}$. 

Based on the discussions above, we summarize our power control policy for D2D pairs in dense D2D relay networks as follows: In each D2D pair, we need to manage the energy consumptions of all possible transmitters, i.e., the D2D transmitter and the relay node. Since the D2D transmitter and relay node are both powered by batteries, we assume that they have the same maximum available energy $E_{max}$, and the power control policy for each D2D pair is conducted over a time period $[0,T]$. The system is assumed to be using time division multiplexing, and the power control policy is defined as the sequence of transmitter powers for both D2D transmitter and the relay node in each time slot over $[0,T]$. Hence, at each time instant $t$, the power control policy for the $i$th D2D pair can be written as a tuple $(P_{i,d}^*(t), P_{i,r}^*(t))$ in which the two entries denote D2D transmit power and relay transmit power, respectively. During a given time period $[0,T]$, the power control problem of each D2D pair could be formulated in terms of D2D transmitter and relay node energy dynamics under a differential game framework [36, 61, 81]. Some basic definitions in such a differential game are given as follows:

- **Player Set $\mathcal{N}$**: In a differential game, we assume that each D2D pair (which contains the relay node) is a rational player which can decide its transmission policy. The transmission policy for each pair consists of both the D2D transmit power and the relay node transmit power, which are correlated due to intra-tier coordination. In this study, $\mathcal{N} = \{1, ..., N\}$.

- **Player Action Tuples $\{p_{d,i}(t), p_{r,i}(t)\}_{i \in \mathcal{N}}$**: In each D2D pair, the D2D transmitter selects its transmit power $p_{d,i}(t)$ in the first hop, and the relay node select its transmit power $p_{r,i}(t)$ when forwarding the signal based on the parameters and choice in the first hop. Clearly, D2D transmitter needs to interact with the relay node, and their relationship could be derived when modeled under Stackelberg framework.

- **State Space $\{S_i\}_{i \in \mathcal{N}}$**: We consider energy control for both the D2D transmitter and the relay node in every D2D pair. Thus, the state space that describes the $i$th player
is two dimensional, in which each dimension indicates the remaining power of the two devices.

- **Cost Function** $C_i(t)_{i \in S}$: A cost function is often an indicator of a D2D pair which numerically compares its performance (e.g., SINR) to some predefined threshold. In this paper, we consider both D2D pair performance and transmit power simultaneously when designing the cost function.

### 5.1.2 State Space

The state variable in the differential game is defined as the remaining available energy in the D2D transmitter and the relay node. At some time instant $t$, the state space of the $i$th D2D pair is denoted by $(E_{d,i}, E_{r,i})$. We assume that all D2D transmitters and relay nodes have the same maximum energy $E_{max}$ at time $t = 0$, and the devices cannot transmit once their energy is used up. Therefore, we have $0 \leq E_{d,i}(t), E_{r,i}(t) \leq E_{max}$. We focus on the $i$th D2D pair, and the energy dynamics of D2D transmitter $D_{T,i}$, can be written as:

$$dE_{d,i}(t) = -p_{d,i}(t)dt,$$

which indicates the power consumption process of $D_{T,i}$. The energy dynamics of relay node $R_i$ yields:

$$dE_{r,i}(t) = -p_{r,i}(t)dt,$$

Thus, the state space of the $i$th D2D pair is defined as follows:

$$s_i(t) = [E_{d,i}(t), E_{r,i}(t)],$$

which consists of the remaining energy of both the D2D transmitter and the relay node at time instant $t$. For a certain D2D pair, the interference received from all other D2D pairs would affect the performance of both hops and thus affect strategy selection of the D2D
transmitter and the relay node. Note that during each time slot, the signal is first transmitted to the relay node from the D2D transmitter, and then forwarded to the D2D receiver in the second hop. It can be assumed that the channel conditions do not change in two consecutive time slots. Therefore, once a D2D transmitter decides its transmission power in a certain time slot, the relay node can decide its own transmission strategy subsequently. The interactions within a D2D pair can be modeled in which the D2D transmitter tries to find an optimal price and the relay node tries to find its optimal transmit power. To decide the proper transmission power of the D2D transmitter, the D2D pair as a whole can be modeled as a rational player that interacts with all other D2D pairs. A cost function is thus defined as as the minimization objective in a mean field game taking both D2D transmission performance and transmit power into account. The cost function is developed in the following subsection.

5.1.3 Cost Function

As described above, each D2D pair interacts with other D2D pairs while deciding its own power control policy to minimize its cost. In such interactions, the only thing that each D2D pair needs to decide is the D2D transmit power, and the optimal price and the relay transmit power can then be derived by modeling the interactions within a D2D pair under Stackelberg game framework. The performance of a D2D pair can be characterized by the SINR received by D2D receiver, and we aim to find an optimal control policy for the D2D transmit power and the relay node transmit power. For simplicity, we focus on the signal quality at the D2D receiver. As the two hops in a D2D pair access different channels, only the relay nodes in other D2D pairs will cause interference to the current D2D receiver. We assume a common SINR threshold $\gamma_{D,th}$ for all D2D pairs, and the cost function of the $i$ th D2D pair is given by:

$$c_i(t) = [\gamma_{D,i}(t) - \gamma_{D,th}]^2 + \lambda(p_{d,i}(t) + p_{r,i}(t)), \quad (5.10)$$

which needs to be optimized over the time period $[0, T]$. The coefficient $\lambda$ is introduced to
balance the received SINR performance and power consumption. Basically, optimizing $c_i(t)$ will attempt to satisfy the QoS constraint, and similar cost functions are used in [42,43]. In this scenario, we take both D2D transmit power and relay transmit power into consideration.

The proposed differential game aims to coordinate inter-tier interactions among different D2D pairs, while the Stackelberg game coordinates intra-tier interactions between the two hops within each D2D pair. In a distributed power control scheme, each D2D pair senses the channels and makes its own transmission strategy decision based on channel parameters. The differential game equilibrium can then be derived subsequently based on Stackelberg game outcomes of each D2D pair. In the following section, Stackelberg game analysis applied to D2D pair coordination is first carried out to derive the optimal transmission strategy.

5.2 Stackelberg Game Analysis

Stackelberg game and mean field game are the two key game-theoretic approaches for handling interactions among various D2D devices in dense D2D relay networks. Stackelberg game can be adopted to tackle intra-tier coordination within a D2D pair, and mean field game can be adopted to tackle inter-tier coordination among different D2D pairs, i.e., the rational players in a differential game. In this section, Stackelberg game analysis is carried out for intra-tier D2D coordination.

5.2.1 Stackelberg Game Framework for D2D Coordination

Stackelberg game is a two-stage game framework which models two rational players as a leader-follower pair and the interactions between them as a trade relationship [23, 85, 100, 104]. For intra-tier coordination in DF transmission, a relay node will select an appropriate transmit strategy to forward the signal it receives from the D2D transmitter to the destination. In other words, the D2D link cannot be set up without the help of the relay node. Therefore, the D2D transmitter is willing to pay some price for adopting relay node in DF transmission, and the relay node can gain some revenue from D2D transmitter by helping establish the link. In each D2D pair, we model the relay node as the leader and D2D
transmitter as the follower \[6,18,23,56\]. The leader moves first and announces the charging price for relay transmit power. The follower reacts subsequently by deciding the amount of relay transmit power it would like to buy based on the price and makes a payment. Both the leader and the follower attempt to maximize their own utilities simultaneously. The Stackelberg game Nash equilibrium (SGNE) is achieved if neither the leader nor the follower can gain higher revenues by making unilateral changes to its own transmission strategy.

We assume that transmit power of \(D_{T,i}, p_{d,i}(t)\), is known a priori in the Stackelberg game analysis. The follower aims to gain as much benefit as possible at the least possible expenses. In DF cooperative relaying, a data packet takes two time slots to arrive at its destination, hence the capacity bottleneck of overall link rate is determined by the smaller one among the two hops. The highest rate \(r_{i,\text{max}}(t)\) that the link can achieve is thus given by:

\[
r_{i,\text{max}}(t) = \frac{1}{2} \min \{ \log_2(1 + \gamma_{R,i}(t)), \log_2(1 + \gamma_{D,i}(t)) \}
\]

\[
= \frac{1}{2} \min \left\{ \log_2(1 + \frac{p_{d,i}(t)g_{r,i,i}(t)}{N \sum_{j=1,j\neq i} p_{d,j}(t)g_{r,j,i}(t) + \sigma^2}), \log_2(1 + \frac{p_{r,i}(t)g_{r,d,i}(t)}{N \sum_{j=1,j\neq i} p_{r,j}(t)g_{r,d,j,i}(t) + \sigma^2}) \right\}
\]

\[
(5.11)
\]

Naturally, from the perspective of the follower, it hopes to achieve as high an end-to-end link capacity as possible. Therefore, we choose the following function \(U_{F,i}(t)\) as the
definition for $i$th follower utility:

$$U_{F,i}(t) = \frac{1}{2} \min \{ \log_2(1 + \gamma_{R,i}(t)), \log_2(1 + \gamma_{D,i}(t)) \} - \alpha_i p_{r,i}(t)$$

$$= \frac{1}{2} \min \left\{ \log_2(1 + \frac{p_{d,i}(t)g_{d,r,i}(t)}{N \sum_{j=1,j\neq i}^N p_{d,j}(t)g_{d,r,j,i}(t) + \sigma^2}), \log_2(1 + \frac{p_{r,i}(t)g_{r,d,i,i}(t)}{N \sum_{j=1,j\neq i}^N p_{r,j}(t)g_{r,d,j,i}(t) + \sigma^2}) \right\}$$

$$- \alpha_i p_{r,i}(t) \quad (5.12)$$

where $\alpha_i$ is the charging price for relay transmit power charged by the leader, and the follower needs to purchase the relay transmit power. The definition of $U_{F,i}(t)$, is related to a meaningful quantity in DF transmission, i.e., the end-to-end link capacity of the D2D pair. The term $-\alpha_i p_{r,i}(t)$ represents the expenses spent on relay transmit power. It is clear that the follower utility is zero when the relay node is not using any transmit power. The larger the price $\alpha_i$ is, the more eager the follower is willing to buy less relay transmit power from the leader, which results in a decrease of the D2D link capacity. Therefore, the follower seeks to maximize its utility by selecting the amount of relay transmit power it would like to buy, i.e.,

$$\max_{\{p_{r,i}(t)\}} U_{F,i}(t) \quad s.t. \{p_{r,i}(t) \geq 0\} \quad (5.13)$$

From the leader’s perspective, the relay node can make some revenue by selling its transmit power. Hence, we simply define the $i$th leader utility function $U_{L,i}(t)$ as follows:

$$U_{L,i}(t) = (\alpha_i - \alpha_0)p_{r,i}(t), \quad (5.14)$$

where $\alpha_0$ is a predetermined constant which denotes the unit cost of relay transmit power. The leader can make profits if the price is higher than the cost and the relay transmit
power is greater than zero. Hence, the optimization problem for the leader (relay node) is to maximize its utility by setting a proper charging price for relay transmit power, i.e.,

$$\max_{\alpha_i} U_{L,i}(t) \quad \text{s.t.} \{\alpha_i \geq \alpha_0\} \quad (5.15)$$

The price $\alpha_i$ charged by the leader and the relay transmit power $p_{r,i}(t)$ the follower would like to purchase constitute a strategy pair $(\alpha_i, p_{r,i}(t))$. Based on the definition of Stackelberg game, the proposed strategy pair is also the SGNE if neither the leader nor the follower can gain more profits by making unilateral changes to its own strategy, i.e.,

$$U_{F,i}(\alpha_i, p_{r,i}(t), t) \geq U_{F,i}(\alpha_i, p'_{r,i}(t), t) \quad (5.16)$$

$$U_{L,i}(\alpha_i, p_{r,i}(t), t) \geq U_{F,i}(\alpha'_i, p_{r,i}(t), t) \quad (5.17)$$

where both $\alpha'_i, p'_{r,i}(t)$ can take any values in their own searching spaces allowed. In the Stackelberg game, the follower chooses its transmission strategy based on the leader’s action. Naturally, the equilibrium of the proposed Stackelberg game can be solved by backward induction. The analysis to follower transmit power is carried out first, then the analysis of the leader is developed to derive the optimal charging price.

### 5.2.2 Follower Analysis

Given the unit charging price $\alpha_i$, the follower can maximize its own utility by selecting a proper relay transmit power. If the follower buys too much transmit power, the high expenditure will cause a significant decrease in its utility function. Yet if the follower buys a small relay transmit power, the D2D link capacity and D2D receiver SINR performance will degrade. Therefore, the best response for relay transmit power needs to be searched within its range. As $U_{F,i}(t)$ could have different analytical forms, we need to simplify Eq. (5.12) based on the relationship between $p_{r,i}(t)$ and other parameters. For simplicity, we omit the time variable $t$ in all variable notations. All possible cases are given as follows:
• Case 1: $\gamma_{R,i} \leq \gamma_{D,i}$

If $\gamma_{R,i} \leq \gamma_{D,i}$ is satisfied, the first hop (i.e., D2D transmitter to relay node) has a smaller capacity and thus becomes the bottleneck. We substitute Eq. (5.5) and (5.6) into $\gamma_{R,i} \leq \gamma_{D,i}$ and have:

$$p_{r,i} \geq \left( \frac{\sum_{j=1, j \neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{\sum_{j=1, j \neq i}^{N} p_{d,j} g_{dr,j,i} + \sigma^2} \right) g_{rd,i,i}$$

(5.18)

For simplicity, we rewrite the right side of Eq. (5.18) as $C_{p_{d,i}}$, in which $C$ is defined as:

$$C := \left( \frac{\sum_{j=1, j \neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{\sum_{j=1, j \neq i}^{N} p_{d,j} g_{dr,j,i} + \sigma^2} \right) g_{rd,i,i}$$

(5.19)

Therefore, we know $p_{r,i} \geq C_{p_{d,i}}$, and the follower utility in Case 1 becomes:

$$U_{F,i,1} = \frac{1}{2} \log_2 \left( 1 + \frac{p_{d,i} g_{dr,i,i}}{\sum_{j=1, j \neq i}^{N} p_{d,j} g_{dr,j,i} + \sigma^2} \right) - \alpha_i p_{r,i},$$

(5.20)

Apparently, $U_{F,i,1}$ is monotonically decreasing with $p_{r,i}$ for any given $\alpha_i$. We use $\hat{p}_{r,i,1}$ to denote the optimal relay transmit power in Case 1 and we have $\hat{p}_{r,i,1} = C_{p_{d,i}}$.

• Case 2: $\gamma_{R,i} > \gamma_{D,i}$

If condition $\gamma_{R,i} > \gamma_{D,i}$ stands, the second hop in D2D link becomes the bottleneck, i.e., $p_{r,i} < C_{p_{d,i}}$. The follower utility function Eq. (5.12) can be simplified as:

$$U_{F,i,2} = \frac{1}{2} \log_2 \left( 1 + \frac{p_{r,i} g_{rd,i,i}}{\sum_{j=1, j \neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2} \right) - \alpha_i p_{r,i},$$

(5.21)

where $U_{F,i,2}$ is the follower utility in Case 2. To find out the optimal $p_{r,i}$, we take the
first order derivative of \((5.21)\) and set it to zero, i.e.,

\[
\frac{\partial U_{F,i,2}}{\partial p_{r,i}} = \frac{1}{2 \ln 2} \cdot \frac{g_{rd,i,i}}{\sum_{j=1, j \neq i}^{N} p_{r,j} g_{rd,j,i} + p_{r,i} g_{rd,i,i} + \sigma^2} - \alpha_i = 0 \quad (5.22)
\]

The optimal relay transmit power \(\hat{p}_{r,i,2}\) in Case 2 yields:

\[
\hat{p}_{r,i} = \frac{1}{2 \alpha_i \ln 2} - \frac{\sum_{j=1, j \neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{g_{rd,i,i}} \quad (5.23)
\]

It is easy to verify that the second-order derivative of \(U_{F,i}\) in \((5.21)\) satisfies \(\frac{\partial^2 U_{F,i}}{\partial p_{r,i}^2} < 0\), and the optimal \(\hat{p}_{r,i}\) maximizes follower utility function. Considering the practical constraints, \(p_{r,i} < C_{p_{d,i}}\) must be satisfied simultaneously in Case 2.

Since Case 1 and Case 2 are mutually exclusive, we summarize the above analysis as follows. As relay transmit power is searched in \([0, +\infty)\), we further divide it into two non-overlapping subintervals, i.e., \([0, C_{p_{d,i}}]\) and \((C_{p_{d,i}}, +\infty)\). Clearly, Case 2 is satisfied in subinterval \([0, C_{p_{d,i}}]\), and the optimal relay transmit power is searched in \(\{0, \hat{p}_{r,i}, C_{p_{d,i}}\}\). The optimal value for follower utility function is thus searched within \(\{0, U_{F,i,2}(\hat{p}_{r,i,2}), U_{F,i,2}(C_{p_{d,i}})\}\).

Otherwise, if relay transmit power is searched in the subinterval \([C_{p_{d,i}}, +\infty)\), only Case 1 is satisfied and the optimal relay transmit power is \(C_{p_{d,i}}\). Hence, the corresponding follower utility becomes \(U_{F,i,1}(C_{p_{d,i}})\).

It is easy to verify that when the relay transmit power is exactly \(C_{p_{d,i}}\), the follower utility yields the same value in both subintervals, i.e., \(U_{F,i,1}(C_{p_{d,i}}) = U_{F,i,2}(C_{p_{d,i}})\). Hence, the optimal value for follower utility is searched in \(\{0, U_{F,i,2}(\hat{p}_{r,i,2}), U_{F,i,2}(C_{p_{d,i}})\}\), i.e.,

\[
\{0, \frac{1}{2} \log_2(1 + \frac{\hat{p}_{r,i} g_{rd,i,i}}{\sum_{j=1, j \neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}) - \alpha_i \hat{p}_{r,i}, \quad (5.24)
\]

\[
\frac{1}{2} \log_2(1 + \frac{p_{d,i} g_{dr,i,i}}{\sum_{j=1, j \neq i}^{N} p_{d,j} g_{dr,j,i} + \sigma^2}) - \alpha_i C_{p_{d,i}} \}
\]
5.2.3 Leader Analysis

From the leader’s perspective, the goal is to maximize relay node utility in D2D transmission by setting an appropriate price for the relay transmit power. The revenue (i.e., leader utility) in (5.14) has the following properties: (1) The revenue yields 0 if either relay transmit power is 0 or the follower would only like to pay $\alpha_0$. (2) If the price is set too large, the follower will not buy sufficient transmit power from the relay node and this leads to insufficient Stackelberg game outcomes. The leader utility will drop to zero as well. Therefore, the relay node would set its price such that the estimated relay transmit power satisfies $\hat{p}_{r,i} \geq 0$, i.e.,

$$\hat{p}_{r,i} = \frac{1}{2\alpha_i \ln 2} - \frac{\sum_{j=1, j\neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{g_{rd,i,i}} \geq 0$$  \hspace{1cm} (5.25)

and the charging price $\alpha_i$ satisfies $\alpha_i \leq \alpha_{i,max}$, where the upper bound $\alpha_{i,max}$ is:

$$\alpha_{i,max} = \frac{1}{2 \ln 2} \cdot \frac{g_{rd,i,i}}{\sum_{j=1, j\neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}$$  \hspace{1cm} (5.26)

By substituting the estimated optimal relay transmit power (5.23) into the leader utility (5.14), we have:

$$U_{L,i} = (\alpha_i - \alpha_0)\left( \frac{1}{2\alpha_i \ln 2} - \frac{\sum_{j=1, j\neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{g_{rd,i,i}} \right)$$

$$= -\alpha_i \cdot \frac{\sum_{j=1, j\neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{g_{rd,i,i}} - \frac{\alpha_0}{2\alpha_i \ln 2}$$

$$+ \frac{1}{2 \ln 2} + \alpha_0 \frac{\sum_{j=1, j\neq i}^{N} p_{r,j} g_{rd,j,i} + \sigma^2}{g_{rd,i,i}}$$  \hspace{1cm} (5.27)
Since $\alpha_0$ is a constant, the optimal price can be derived by setting the first order derivative $\partial U_{L,i}/\partial \alpha_i = 0$, i.e.,

$$\frac{\partial U_{L,i}}{\partial \alpha_i} = -\sum_{j=1,j\neq i}^{N} p_{r,j}g_{rd,j,i} + \sigma^2 \frac{\alpha_0}{g_{rd,i,i}} + \alpha_0 \frac{\alpha_0^2}{2\alpha_i^2 \ln 2} = 0,$$

(5.28)

and we derive the optimal estimated price $\hat{\alpha}_i$ as follows:

$$\hat{\alpha}_i = \sqrt{\frac{\alpha_0 g_{rd,i,i}}{2 \ln 2 (\sum_{j=1,j\neq i}^{N} p_{r,j}g_{rd,j,i} + \sigma^2)}},$$

(5.29)

It is easy to verify that $\partial^2 U_{L,i}/\partial \alpha_i^2 \leq 0$, as $\alpha_i$ is positive. Hence, the leader utility is maximized when the charging price is set to $\hat{\alpha}_i$. For simplicity, we define:

$$D := \frac{g_{rd,i,i}}{\sum_{j=1,j\neq i}^{N} p_{r,j}g_{rd,j,i} + \sigma^2},$$

(5.30)

and thus we can rewrite $\hat{\alpha}_i = \sqrt{\frac{\alpha_0 D}{2 \ln 2}}$ and $\alpha_{i,max} = \frac{D}{2 \ln 2}$. Note that the bounds for $\alpha_i$ are derived from relay transmit power bounds, we make the following remarks based on $\alpha_0$:

- $\alpha_0 < \alpha_{i,max}$: In this scenario, the charging price $\alpha_i$ is searched in the interval $[\alpha_0, \alpha_{i,max}]$; hence a unique $\alpha_i$ that maximizes leader utility must exist. It can be noticed that the optimal estimated charging price $\hat{\alpha}_i$ is the geometric mean of the lower bound and the upper bound, and naturally $\hat{\alpha}_i$ satisfies $\alpha_0 < \hat{\alpha}_i < \alpha_{i,max}$. Thus, the leader utility function is maximized when the optimal charging price is selected to be $\sqrt{\frac{\alpha_0 D}{2 \ln 2}}$.

- $\alpha_{i,max} \leq \alpha_0$: In this scenario, the upper bound $\alpha_{i,max}$ is below the threshold $\alpha_0$ which allows the leader to make profit from the follower. When the actual charging price $\alpha_i$ is greater than $\alpha_{i,max}$, the corresponding optimal transmit power $\hat{p}_{r,i}$ falls below zero, which is impractical. From the follower’s point of view, data cannot be transmitted
in such a D2D pair when the relay transmit power is 0. On the other hand, the leader cannot make any profit from the follower as long as the relay transmit power is 0. Hence, the optimal charging price for the leader is $\alpha_0$.

### 5.2.4 Stackelberg Game Nash Equilibrium

Based on the discussions above, the Stackelberg game Nash equilibrium of a given D2D pair can be achieved in the proposed Stackelberg game. Since different leader strategies need to be adopted under different system parameters, we make the following remarks:

(1) $\alpha_0 < \alpha_{i,\text{max}}$: As described in Section 5.2.3, the optimal charging price is $\hat{\alpha}_i = \sqrt{\frac{\alpha_0 D}{2\ln 2}}$. In this case, the corresponding optimal transmit power is always positive. By substituting $\hat{\alpha}_i$ into the optimal relay transmit power (5.23), we have:

$$\hat{p}_{r,i} = \frac{1}{\sqrt{2\alpha_0 D \ln 2}} - \frac{1}{D}. \quad (5.31)$$

Based on the relationship between $\hat{p}_{r,i}$ and $C_{p_{d,i}}$, we have the discussions for the following scenarios:

- $\hat{p}_{r,i} < C_{p_{d,i}}$: Clearly $\hat{p}_{r,i}$ satisfies the condition of Case 2 in 5.2.2. According to the proposed analysis, the follower will choose $\hat{p}_{r,i}$ as the optimal relay transmit power. Therefore, the SGNE is $(\sqrt{\frac{\alpha_0 D}{2\ln 2}}, \frac{1}{\sqrt{2\alpha_0 D \ln 2}} - \frac{1}{D})$.

- $\hat{p}_{r,i} \geq C_{p_{d,i}}$: Since $\hat{p}_{r,i}$ is beyond the interval $[0, C_{p_{d,i}}]$, we know that the optimal relay transmit power is searched within $\{0, C_{p_{d,i}}\}$ according to the analysis in 5.2.2. It is also easy to verify that $U_{F,i}$ is monotonically increasing over $[0, C_{p_{d,i}}]$. Hence, the overall optimal relay transmit power is selected as $C_{p_{d,i}}$. No matter what the charging price is, the follower cannot gain higher revenue by changing the transmit power unilaterally. On the other hand, the leader utility (5.14) becomes $U_{L,i} = (\alpha_i - \alpha_0)C_{p_{d,i}}$. Clearly, the higher the charging price is, the higher revenue the leader
can get. Hence, the charging price is set as $\alpha_{i,\text{max}}$, which is \( \frac{D^2}{2 \ln 2} \). The SGNE in this scenario can be written as \( (\frac{D^2}{2 \ln 2}, C_{p_{d,i}}) \).

(2) $\alpha_0 \geq \alpha_{i,\text{max}}$: In this scenario, the optimal charging price is $\alpha_0$, and the relay transmit power is set as zero. It can be noticed that when the leader increases its charging price, the profit it makes is always 0. Otherwise if the follower buys some relay transmit power, its revenue will decrease as $\alpha_0$ is positive. Thus, the strategy pair $(\alpha_0, 0)$ is the SGNE of the proposed game.

It can be observed from the above discussions that for a given D2D transmit power $p_{d,i}$, the optimal relay transmit power derived from the Stackelberg game is always bounded. Hence, mean field game analysis could be adopted to find the power control policy in dense D2D cooperative relaying network. In the following section, we will discuss the distributed power control policy under a mean-field game framework.

5.3 Mean Field Game for Power Control

Mean field game (MFG) is a promising game theoretic approach to model the interactions within a wireless network which consists of a huge number of D2D pairs \[3, 37, 51\]. It can be interpreted as a special form of differential game, in which each player interacts with the statistical behavior of all other players, i.e., the mean field. By solving the coupled Hamilton-Jacobi-Bellman (HJB) and Fokker-Plank-Kolmogorov (FPK) partial differential equations, the MFG equilibrium is achieved and the optimal power control path can be derived subsequently. The FPK equation evolves forward in time to update the distribution of the agents, while the HJB equation evolves backward in time to update the optimal control path. More details about mean field game can be referred to \[37\].

In dense D2D cooperative relaying network, each D2D pair can be modeled as an independent rational player that receives interference from all other players, as each pair can decide its own transmission strategy locally. This naturally fits the idea of MFG. In the rest of this section, we first formulate the power control problem for each pair, then develop
an MFG analysis to derive its optimal power control policy.

5.3.1 Optimal Power Control Problem

In a D2D cooperative relaying network, each D2D pair will determine its own optimal power control profile to minimize its cost function defined in (5.10) over a certain time period $[0, T]$ [12]. For the $i$th player, we denote the power control profile by $Q_i^*(t)$. According to Stackelberg game analysis, we can see that relay transmit power $p_{r,i}$ can be uniquely determined if a D2D transmit power $p_{d,i}$ and channel parameters are known. Thus, the optimal power control profile can be interpreted as a function of $p_{d,i}$, and the optimal power control problem can be defined as:

$$Q_i^*(t) = \arg\min_{p_{d,i}(t)} E\left[\int_0^T c_i(t)dt + c_i(T)\right]$$

(5.32)

where $E[\cdot]$ denotes the expected value of the given function. Clearly, the objective is to find out the optimal power control path over time period $[0, T]$ which minimizes the cost function. Meanwhile, we define a value function $u_i$ as follows,

$$u_i(t, s_i(t)) = \min_{p_{d,i}(t)} E\left[\int_t^T c_i(t)dt + u_i(T, s_i(T))\right]$$

(5.33)

where $u_i(t, s_i(t))$ is the value at time $t$ and achieves $u_i(T, s_i(T))$ when $t = T$. Naturally, the value function is monotonically increasing as $0 \leq t \leq T$. According to Bellman's principle of optimality [13], an optimal control policy should satisfy that, no matter what the initial decisions are, the remaining decisions must form an optimal policy with regard to the state resulting from the initial decisions [14, 17]. Hence, under the mean field game framework, Nash equilibrium is achieved if and only if the power control profile $Q_i^*(t)$ satisfies the
following:

\[ Q_i^*(t) = \arg \min_{p_{d,i}(t)} E \left[ \int_0^T c_i(p_{d,i}(t), p_{d,-i}^*) dt + c_i(T) \right] \]  

(5.34)

and subject to the state space dynamic constraints:

\[
\begin{cases}
  s_i(t) = [E_{d,i}(t), E_{r,i}(t)], \\
  dE_{d,i}(t) = -p_{d,i}(t) dt, \\
  dE_{r,i}(t) = -p_{r,i}(t) dt,
\end{cases}
\]  

(5.35)

where \( p_{d,-i}^* \) denotes the transmit power of all other D2D pairs. The power control game lies on the upper hierarchy of the entire game framework, which makes use of the outcomes of Stackelberg game on the lower hierarchy. None of the players in the mean-field game can have a smaller cost by unilaterally changing its own power control policy. Hence, the Nash equilibrium of the differential power control game can be derived by solving the coupled HJB-FPK equations. More details could be found in [37, 70, 94].

### 5.3.2 Mean Field Approximation

Mean field is defined as the statistical distribution of the proposed state space when the number of players goes to infinity [34, 35]. Given the two dimensional state space \( s_i(t) = [E_{d,i}(t), E_{r,i}(t)] \) in (5.9), the corresponding mean field \( m(s, t) \) is the probability that a certain state \( s_i(t) \) is likely to appear at time instant \( t \), i.e.,

\[
m(s, t) = \lim_{N \to \infty} \frac{1}{N} f_i(s_i(t), s)
\]  

(5.36)

where \( f_i(s_i(t), s) \) equals 1 if \( s_i(t) = s \) and 0 otherwise. To derive the mean field given above, mean field approximation is first applied to the statistical interference received by the relay node and D2D receiver. We assume that the number of players \( N \) is large enough such
that each player contributes very little interference to others when they transmit, and the
summation of all these tiny interferences makes up the whole interference signal in (5.3)
and (5.4). Hence, the interference received by the relay node (5.3) and the D2D
receiver (5.4) can be approximated by:

\[ I_{R,i}(t) \approx (N - 1)\bar{p}_d(t)\bar{g}_{dr,j,i}(t), \quad (5.37) \]

and

\[ I_{D,i}(t) \approx (N - 1)\bar{p}_r(t)\bar{g}_{rd,j,i}(t), \quad (5.38) \]

respectively. \( \bar{p}_d(t), \bar{p}_r(t) \) are known test powers which are identical to all D2D transmitters
and relay nodes, respectively, while \( \bar{g}_{dr,j,i}(t), \bar{g}_{rd,j,i}(t) \) are defined as the mean channel gains
among corresponding devices. The mean channel gains are unknown initially, thus we can
use the following approach to make an estimate of these parameters. We select \( \bar{p}_d(t), \bar{p}_r(t) \)
as the transmit power in \( i \)th D2D pair. The powers received by the relay node and D2D
receiver, denoted by \( p_{R,i}^R(t), p_{d,i}^R(t) \), are approximated by:

\[ p_{R,i}^R(t) \approx \bar{p}_d(t)\bar{g}_{dr,i,i}(t) + (N - 1)\bar{p}_d(t)\bar{g}_{dr,j,i}(t), \quad (5.39) \]

and

\[ p_{d,i}^R(t) \approx \bar{p}_r(t)\bar{g}_{rd,i,i}(t) + (N - 1)\bar{p}_r(t)\bar{g}_{rd,j,i}(t), \quad (5.40) \]

respectively. Thus, the mean field approximation of the channel gains can be written as:

\[ \bar{g}_{dr,j,i}(t) \approx \frac{p_{R,i}^R(t) - \bar{p}_d(t)\bar{g}_{dr,i,i}(t)}{(N - 1)\bar{p}_d(t)} \quad (5.41) \]

\[ \bar{g}_{rd,j,i}(t) \approx \frac{p_{d,i}^R(t) - \bar{p}_r(t)\bar{g}_{rd,i,i}(t)}{(N - 1)\bar{p}_r(t)} \quad (5.42) \]
In distributed power control, each D2D pair aims to decide the appropriate transmission strategy by itself, and the Stackelberg game outcomes are derived using mean field approximation as well. By substituting (5.39) (5.40) into (5.19), the ratio \( C \) is approximated by:

\[
\bar{C} = \frac{(I_{D,i}(t) + \sigma^2)g_{dr,i,i}(t)}{(I_{R,i}(t) + \sigma^2)g_{rd,i,i}(t)}
\]

\[
\approx \frac{(p_{R,i}(t) - \bar{p}_r(t)g_{rd,i,i}(t) + \sigma^2)g_{dr,i,i}(t)}{(p_{R,i}(t) - \bar{p}_d(t)g_{dr,i,i}(t) + \sigma^2)g_{rd,i,i}(t)}
\] (5.43)

Similarly, the ratio \( D \) in (5.30) is approximated by:

\[
\bar{D} \approx \frac{g_{rd,i,i}(t)}{p_{d,i}(t) - \bar{p}_r(t)g_{rd,i,i}(t) + \sigma^2}
\] (5.44)

and the \( i \)th D2D pair will decide the optimal transmission strategy based on \( \bar{C} \) and \( \bar{D} \).

Referring to the analysis in 5.2, we know that the SGNE tuple \((\hat{\alpha}_i(t), \hat{p}_{r,i}(t))\) is chosen from the following estimates:

\[
\left\{ \left( \sqrt{\frac{\alpha_0 D}{2 \ln 2}}, \frac{1}{\sqrt{2 \alpha_0 D \ln 2}}, \frac{1}{\bar{D}}, \left( \frac{\bar{D}}{2 \ln 2}, C p_{d,i}, (\alpha_0, 0) \right) \right) \right\}
\] (5.45)

where \( \hat{\alpha}_i(t), \hat{p}_{r,i}(t) \) are the optimal price and relay transmit power under mean field approximation, respectively. Hence, the SINR received by D2D receiver can be approximated by substituting \( \hat{p}_{r,i}(t) \) and (5.42) into (5.6), i.e.,

\[
\bar{\gamma}_{D,i}(t) \approx \frac{\hat{p}_{r,i}(t)g_{rd,i,i}(t)}{p_{d,i}(t) - \bar{p}_r(t)g_{rd,i,i}(t) + \sigma^2}.
\] (5.46)

We also substitute (5.46) and \( \hat{p}_{r,i} \) into cost function (5.10) and get:

\[
\bar{c}_i(t) = \left[ \bar{\gamma}_{D,i}(t) - \gamma_{D,th} \right]^2 + \lambda (p_{d,i}(t) + \hat{p}_{r,i}(t)),
\] (5.47)
We will use $\bar{c}_i(t)$ instead of $c_i(t)$ in the power control profile (5.32) to derive the mean-field game Nash equilibrium. In the following subsection, we will use the mean field cost function in analyzing the coupled HJB and FPK differential equations.

### 5.3.3 Mean Field Game Analysis

As described earlier, MFG can be formulated as a dynamic coupled system consisting of two differential equations, HJB and FPK. Nash equilibrium of the proposed MFG can be derived by solving these differential equations [10, 32, 93]. The FPK equation is defined as:

$$
\partial_t m(t, s) + \nabla (m(t, s) \cdot \partial_t s(t)) = 0,
$$

(5.48)

where $s$ is the state. The HJB equation is defined as:

$$
-\partial_t u(t, s) = \min_{p(t)} [c(t) + \partial_t s(t) \cdot \nabla u(t, s)],
$$

(5.49)

where the right side of (5.49) is also known as the Hamiltonian. The HJB and FPK equations govern the evolution of the optimal control path and mean field functions, respectively. The HJB is solved backwards in time domain from $T$ to 0, while the FPK is solved forwards from 0 to $T$. The value $u(t, s)$ and the mean field $m(t, s)$ are the solutions for the HJB and FPK equations, respectively. The mean field Nash equilibrium (MFNE) can be denoted by a stable tuple $(u^*(t, s), m^*(t, s))$ at any time $t$ and any state $s$, in which $u(t, s)$ and $m(t, s)$ interact with each other when HJB and FPK are jointly solved. The interactions between HJB and FPK are depicted in Fig. 5.2. The above mean-field Nash equilibrium can be obtained by solving the HJB and FPK equations associated with each player $i$, and a total of $N$ Nash equilibria needs to be solved for all the $N$ D2D pairs in the proposed network model. Yet by adopting mean field approximation in Section 5.3.2 only one Nash equilibrium needs to be derived. More details on MFNE can be found in [69, 92].
5.4 Distributed Mean-Field Game Power Control Policy

In this section, we propose a distributed mean-field power control policy to obtain the Nash equilibrium for each D2D pair. We discretize the partial differential equations FPK and HJB of a certain D2D pair and resort to a finite difference method, i.e., Lax-Friedrics scheme, to jointly solve them [15, 49, 71]. Firstly, the three dimensional search space is discretized into grids, where the three axes are the time $T$, D2D transmitter energy $E_d$ and relay node energy $E_r$. We assume that $T \in [0, T_{max}]$, $E_d \in [0, E_{max}]$, $E_r \in [0, E_{max}]$, and the power control policy aims to minimize the cost function over time interval $[0, T_{max}]$. We also assume that the number of steps in time, D2D transmitter energy and relay node energy are $N_t, N_d, N_r$, respectively. Hence, the iteration step size on these three dimensions are given by:

$$\delta_t = \frac{T_{max}}{N_t}, \delta_{E_d} = \frac{E_{max}}{N_d}, \delta_{E_r} = \frac{E_{max}}{N_r}$$ (5.50)

In this work, we simply set $N_t = N_d = N_r$. The solution to the mean-field game evolves in the space $[0, T_{max}] \times [0, E_{max}] \times [0, E_{max}]$ on those discretized grids, while the
optimal control path can be derived by solving the coupled HJB and FPK equations. A finite difference iterative approach is applied to solve the proposed discretized mean-field game equations using Lax-Friedrichs scheme in [5]. For more details about the discretized parameter space, please refer to [92].

5.4.1 FPK Equation Formulation

We rewrite the FPK equation by substituting parameter space \( s(t) \) in (5.9) into FPK (5.48):

\[
\partial_t m(t, s) + \nabla_E d m(t, s) E_d(t) + \nabla_E r m(t, s) E_r(t) = 0 \quad (5.51)
\]

where \( E_d(t), E_r(t) \) denote the remaining energy in the D2D transmitter and the relay node at time \( t \), respectively. As FPK governs the evolution of mean field, this equation can be iteratively solved in time domain by applying the Lax-Friedrichs scheme, and \( m(t, s) \) is updated forwards from 0 to \( T \). In the discretized approach, we define \( M(i, j, k), P_d(i, j, k), P_r(i, j, k) \) as the mean field, D2D transmitter energy level and relay node energy level on the \((i, j, k)\)th grid (i.e., the values at time \( i \) with D2D transmitter energy level \( j \) and relay node energy level \( k \)) in the discretized search space, respectively. Based on Stackelberg game outcomes in (5.45), we know that for any D2D transmit power \( P_d(i, j, k) \), the corresponding relay transmit power \( P_r(i, j, k) \) can be derived uniquely. Therefore, the update mechanism for the mean field under Lax-Friedrichs scheme is given by:

\[
M(i + 1, j, k) = \frac{1}{2} [M(i, j + 1, k) + M(i, j - 1, k) + M(i, j, k + 1) + M(i, j, k - 1)]
\]

\[
+ \frac{\delta_t}{2\delta E_d} [M(i, j + 1, k)P_d(i, j + 1, k) - M(i, j - 1, k)P_d(i, j - 1, k)]
\]

\[
+ \frac{\delta_t}{2\delta E_r} [M(i, j, k + 1)P_r(i, j, k + 1) - M(i, j, k - 1)P_r(i, j, k - 1)] \quad (5.52)
\]
5.4.2 Lagrange Relaxation to HJB

It is clear that the HJB equation itself does not have an explicit solution and thus cannot be analytically solved using the discretized approach. However, if we consider the dynamics of the value function \( u(t, s) \) in HJB and use the optimal path of the mean field \( m(t, s) \) in FPK as an equality constraint, we can reformulate the problem as follows:

\[
\min_{p_d,(t)} \mathbb{E} \left[ \int_0^T c_i(t) \, dt + c_i(T) \right]
\]

s.t. \( \partial_t m(t, s) + \nabla_{E_d} m(t, s) E_d(t) + \nabla_{E_r} m(t, s) E_r(t) = 0 \) (5.53)

Since the FPK function is considered as an equality constraint and the HJB is backwards in time, the proposed optimization problem can be solved by introducing a Lagrange multiplier \( \lambda(t, s) \) to the HJB function. We assume \( c_i(T) = 0 \), and the Lagrangian of the proposed problem is:

\[
L(m(t, s), P_d(t, s), \lambda(t, s)) = \mathbb{E} \left[ \int_0^T c_i(t) \, dt + c_i(T) \right]
\]

\[
+ \int_{t=0}^{T_{max}} \int_{E_d=0}^{E_{max}} \int_{E_r=0}^{E_{max}} \lambda(t, s) \left[ \partial_t m(t, s) + \nabla_{E_d} m(t, s) E_d(t) + \nabla_{E_r} m(t, s) E_r(t) \right] \, dt \, dE_d \, dE_r
\]

\[
= \int_{t=0}^{T_{max}} \int_{E_d=0}^{E_{max}} \int_{E_r=0}^{E_{max}} \left[ \lambda(t, s) \left( \partial_t m(t, s) + \nabla_{E_d} m(t, s) E_d(t) + \nabla_{E_r} m(t, s) E_r(t) \right) \right]
\]

\[
+ c_i(t, s) m(t, s) \right] \, dt \, dE_d \, dE_r
\]

(5.54)

Clearly, \( \mathbb{E}[\int_0^T c_i(t) \, dt + c_i(T)] \) in Eq. 5.54 can be calculated using the mean field \( m(t, s) \) as it also denotes the distribution of the cost function \( c_i(t) \). Assume \( \lambda(i, j, k) \) is the discretized Lagrange multiplier at the \((i, j, k)\)th grid, the discretized Lagrangian \( L(i, j, k) \) for (5.54) is
given by:

\[
L = \delta t \delta E_d \delta E_r \sum_{i=1}^{N_t+1} \sum_{j=1}^{N_d+1} \sum_{k=1}^{N_r+1} [C_o(i, j, k)M(i, j, k) + \lambda(i, j, k)(E + F + G)],
\]  

(5.55)

where \(C_o(i, j, k)\) denotes the cost function value at the \((i, j, k)\)th grid. The values of \(E, F, G\) in Eq. 5.55 are:

\[
E = \frac{1}{\delta t} \left[ M(i + 1, j, k) - \frac{1}{2} \left[ M(i, j + 1, k) + M(i, j - 1, k) + M(i, j, k + 1) + M(i, j, k - 1) \right] \right];
\]  

(5.56)

\[
F = \frac{1}{2\delta E_d} \left[ P_d(i, j + 1, k)M(i, j + 1, k) - P_d(i, j - 1, k)M(i, j - 1, k) \right];
\]  

(5.57)

\[
G = \frac{1}{2\delta E_r} \left[ P_r(i, j, k + 1)M(i, j, k + 1) - P_r(i, j, k - 1)M(i, j, k - 1) \right];
\]  

(5.58)

The optimal solution of the Lagrangian is given by the tuple \((M^*, P^*_d, \lambda^*)\) at any grid \((i, j, k)\). Thus, the optimal D2D transmit power can be derived by setting \(\frac{\partial L}{\partial P_d(i, j, k)} = 0\).

The discretized equation for \(\frac{\partial L}{\partial P_d(i, j, k)}\) is given by:

\[
\frac{\partial L}{\partial P_d(i, j, k)} = \sum_{j=1}^{N_d+1} \sum_{k=1}^{N_r+1} M(i, j, k) \frac{\partial C_o(i, j, k)}{\partial P_d(i, j, k)} \\
+ M(i, j, k) \left[ \frac{1}{2\delta E_d} + \frac{1}{2\delta E_r} \frac{\partial P_r(i, j, k)}{\partial P_d(i, j, k)} \right] \left[ \lambda(i, j + 1, k) - \lambda(i, j - 1, k) \right].
\]  

(5.59)

During the iterations, the Lagrangian multiplier \(\lambda(i, j, k)\) is updated by solving \(\frac{\partial L}{\partial M(i, j, k)} = 0\). As the HJB function needs to be solved backwards in time domain, we can derive \(\lambda(i - 1, j, k)\) in the iterations as:
\[ \lambda(i - 1, j, k) = \frac{1}{2} [\lambda(i, j + 1, k) + \lambda(i, j - 1, k) + \lambda(i, j, k + 1) - \lambda(i, j, k - 1)] - \delta t C_o(i, j, k) \]

\[ - \frac{\delta t}{2} \left[ \frac{P_d(i, j, k)}{\delta E_d} + \frac{P_r(i, j, k)}{\delta E_r} \right] \left[ \lambda(i, j + 1, k) - \lambda(i, j - 1, k) \right]. \]  

According to the discussions above, we adopt the Lax-Friedrichs based joint finite difference method to solve the Lagrangian proposed in Eq. 5.54. For more details about this method, please see [92].
Chapter 6: Conclusions and Future Work

In this dissertation, we first propose a GRFM based framework for approximating coverage of exclusion zone. Secondly, a novel Stackelberg game framework for channel resource allocation in a cognitive network under cooperative relaying is proposed. We also design a hybrid scheduling algorithm for cooperative relaying using the Stackelberg game outcomes. Thirdly, we propose a novel hierarchical game framework for power control in dense D2D cooperative relaying network. The contributions of the dissertation and future research work based on current results are summarized as follows.

6.1 GRFM for Exclusion Zone Approximation

In Chapter 3, we proposed a GRFM based approach for exclusion zone estimation, in which GRFM provides an efficient machine learning tool in approximating the coverage of the exclusion zone. Instead of using a continuous cognitive band as in the theoretical model of [84], we adopt a network model with discrete SUs uniformly distributed in the cognitive band for GRFM simulations. GRFM uses the signal and interference power samples received by the SUs as its inputs, and the exclusion zone radius can be calculated from GRFM approximations. Besides, the impact of shadowing on the received SIR is also taken into account, which is a an advancement compared to existing work. Simulation results are derived and compared to the theoretical analysis in [84], and the proposed GRFM based approach shows its effectiveness in obtaining the exclusion zone coverage under different SIR threshold values. It could be further extended to accomodate more complicated scenarios under different propagation models.

For future research, more complicated network models in GRFM approximation could be investigated. For example, the primary transmitters may be equipped with directional
antennas, and the SUs may no longer be uniformly distributed in the coverage area. In addition, fast fading and additive noise power can be considered in signal transmission. The suggested extensions above are made towards accommodating more realistic scenarios. Accordingly, parameter fine-tuning and the sample selection scheme for GRFM could be investigated further, and the approximation results could be compared to the state-of-the-art approaches.

6.2 Stackelberg Game Theoretic Approach to Resource Allocation with Cooperative Relaying

In Chapter 4, we proposed a novel Stackelberg game framework to model the downlink interactions among the SU pairs and PUs in a cognitive radio network under cooperative relaying. In the cognitive network model, each SU pair is allowed to use up to one relay node during transmission. Compared to existing Stackelberg game approaches, the novel Stackelberg game framework is proposed for cooperative relaying by assuming that the leader consists of two PUs instead of only one, while the follower consists of one SU pair and one relay node. By analyzing the proposed Stackelberg game framework, we find the optimal cooperative transmission strategies for all the possible leader-follower pairs in the cognitive network. Based on the optimal Stackelberg game outcomes, we proposed a hybrid scheduling algorithm which incorporates both direct transmission and cooperative relaying for the cognitive radio network. The SU pairs could select either direct transmission or cooperative relaying by observing the related channel parameters. The SU pairs and PUs are mapped into leader-follower pairs under both direct transmission and cooperative relaying in the proposed algorithm. The optimal leader-follower with highest priority is scheduled in each TTI until all the devices in the network are scheduled. The experimental results on the proposed Stackelberg game framework demonstrate that a significant SU rate improvement is achieved under cooperative relaying compared to direct transmission. As for the hybrid scheduling algorithm, our experimental results demonstrated good performance.
improvement in terms of sum rate using proper system parameter settings. As the available spectrum resource becomes scarce, the hybrid scheduling algorithm shows clear advantages in comparison to a direct transmission scheduling scheme in [86].

In future, interference to PUs as a constraint could be taken into account as well, i.e., the interference should not exceed some threshold determined by PU performance requirements. Therefore, we aim to design a scheduling scheme which incorporates interference constraints for cooperative relaying and make comparisons to existing scheduling schemes under different parameter settings.

### 6.3 Hierarchical Stackelberg / Mean Field Game for Distributed Power Control in Dense Cooperative Relaying Networks

In Chapter 5 we investigated the power control problem in a dense D2D cooperative relaying network. In system model, D2D pairs with cooperative sensing are densely deployed within the network in large numbers, while each D2D pair aims to transmit signals under a cooperative relaying scheme. The D2D devices face the constraint of limited energy supply, while the coordination within the same D2D pair and the coordination among different D2D pairs need to be considered simultaneously. Therefore, we designed a novel game framework called hierarchical Stackelberg/mean field game to formulate the interactions among the D2D devices. The Stackelberg game lies in the lower level of the hierarchical framework, while the mean field game lies in the upper level of the hierarchical framework. In the Stackelberg game, we model the relay as the leader which charges price for forwarding the signal, and the D2D transmitter as the follower which decides the relay transmit power it would like to buy. By analyzing the Nash equilibrium of the Stackelberg game, we obtained the optimal transmit strategy of the D2D pair, which is later used as the input to the upper level of the hierarchical game. In the mean field game, we aim to find a control path in the state space by adopting the concept of mean field. An approach which consists of
two coupled HJB and FPK differential equations is then adopted to search for the optimal power control strategy for the D2D pairs. Based on the analysis of the differential equations, we developed a distributed power control policy that can be implemented by using finite difference method and Lagrangian relaxation.

Our further research is to propose a full iterative approach that calculates the mean field and the optimal control path until convergence is achieved. The performance of the proposed power control policy in terms of mean field distribution, power consumption and cost function distribution can be further investigated. The spectrum and energy efficiency can also be studied and compared to the performance of the mean field game power control for dense D2D network under direct transmission scheme.
Bibliography
Bibliography


Curriculum Vitae

Zheng Wang grew up in China. He received his B.S. and M.S. in Electrical Engineering and Automation in 2008 and 2012, respectively, both in Tsinghua University, Beijing, China. Since 2013, he has been a PhD student at George Mason University in Electrical and Computer Engineering program. Under the supervision of Dr. Brian Mark, he has completed extensive research in cognitive radio network exclusion zone approximation, game-theoretic resource allocation in cooperative relaying, and hierarchical game-framework for power control in dense cooperative relaying networks.