EFFICIENT DEVICE TO DEVICE COMMUNICATIONS UNDERLAYING HETEROGENEOUS NETWORKS

by

Xue Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

Approved:

Dr. Rose Qingyang Hu Major Professor	Dr. Jacob Gunther Committee Member
Dr. Bedri Cetiner Committee Member	Dr. Tam Chantem Committee Member
Dr. Anthony Chen Committee Member	Dr. Mark R. McLellan Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2016

Abstract

Efficient Device to Device Communications Underlaying Heterogeneous Networks

by

Xue Chen, Doctor of Philosophy

Utah State University, 2016

Major Professor: Dr. Rose Qingyang Hu
Department: Electrical and Computer Engineering

In this dissertation, we have investigated cross-layer optimization, radio resource allocation and interference management algorithms to significantly improve user experience, system spectral efficiency, and energy efficiency for D2D communications underlaying wireless heterogeneous networks. By exploiting frequency reuse and multi-user diversity, this research work aims to design wireless system level algorithms to utilize the spectrum and energy resources efficiently in the next generation wireless heterogeneous network.

First an analytical evaluation of coverage for the D2D communications underlaying cellular network is given, which is derived from stochastic geometry theory. The SINR distributions for both cellular users and D2D users in uplink and downlink resource sharing scenario are analyzed under various network environments to find out the critical parameters that influence the network performance. The conclusions drawn from the analysis provide us a guideline in design of D2D communication network when considering power control, interference management and resource allocation.

Second, we discuss the joint power and spectrum allocation for D2D communications and try to find an optimal algorithm to improve overall network efficiency. A sub-optimal distributed resource and power allocation scheme based on Stackelberg game framework is proposed and the problem is decomposed into sub-problems and solved in a two-step

approach. We also include the mode selection for D2D users and derive an optimal resource allocation scheme for the D2D communication in an OFDM based cellular system. We develop a dual optimization framework to transform the intractable problem into equivalent problem and solve it with reasonable computational complexity. In order to investigate aspects of network energy efficiency for the D2D communication networks, resource allocation between cellular users and D2D users are modeled as a non-cooperative game, where each user tries to determine which resource blocks to select and how much power they plan to transmit correspondingly so as to maximize a utility function. A unique Nash equivalence exists when the channel is assumed as flat fading.

We also study the tradeoff between energy efficiency and spectral efficiency in presence of statistical QoS requirements for delay constrained communication. To exploit the EE-SE relationship under different SNR regimes, we propose a generic close-form approximation with curve fitting. When the circuit power is incorporated in the energy model, it turns out that in the high SNR regime, QoS has a dominant impact on the EE-SE tradeoff, while circuit power impacts EE-SE tradeoff more in the low SNR regime. We also propose a joint uplink and downlink resource optimization scheme for mobile association in the heterogeneous network.

(145 pages)

Public Abstract

Efficient Device to Device Communications Underlaying Heterogeneous Networks

by

Xue Chen, Doctor of Philosophy

Utah State University, 2016

Major Professor: Dr. Rose Qingyang Hu
Department: Electrical and Computer Engineering

Device-to-Device communications have the great potential to bring significant performance boost to the conventional heterogeneous network by reusing cellular resources. In cellular networks, Device-to-Device communication is defined as two user equipments in a close range communicating directly with each other without going through the base station, thus offloading cellular traffic from cellular networks. In addition to improve network spectral efficiency, D2D communication can also improve energy efficiency and user experience.

However, the co-existence of D2D communication on the same spectrum with cellular users can cause severe interference to the primary cellular users. Thus the performance of cellular users must be assured when supporting underlay D2D users.

In this work, we have investigated cross-layer optimization, resource allocation and interference management schemes to improve user experience, system spectral efficiency and energy efficiency for D2D communication underlaying heterogeneous networks. By exploiting frequency reuse and multi-user diversity, this research work aims to design wireless system level algorithms to utilize the spectrum and energy resources efficiently in the next generation wireless heterogeneous network.

To my parents

Acknowledgments

First of all, I would like to thank my advisor Prof. Rose Qingyang Hu for her great support and guidance throughout these years. Her boundless energy, continual encouragement, and strong devotion to work have been truly inspirational. Prof. Rose Qingyang Hu not only advises me on how to do great research, but also teaches me how to be a rightful person. She has been and will be the role model for me in both my academic and person lives.

I am also grateful to Professors Jacob Gunther, Bedri Cetiner, Tam Chantem, Anthony Chen for serving as my committee members and teaching me in classes. During my Ph.D. study and proposal research, they provide great insights and suggestions on how to conduct a real research. I would also like to thank Professor Yi Qian, Dr. Geng Wu, Dr. Jeongho Jeon, Dr. Clara Qian Li, and Dr. Lili Wei, who give great helps on my research work. I especially thank Dr. Jeongho Jeon for his wonderful job. A lot of work in this thesis are results of our collaboration.

I am also deeply indebted to current and past members of Communications Network Innovation Lab, including Professor Xianfu Lei, Dr. Bei Xie, Dr. Tao He, Dr. Junlin Zhang, Yiran Xu, Zhengfei Rui, Zekun Zhang, Haijian Sun, Xuan Xie, David Neal, Dr. Zhouyuan Li. I also need to thank all my friends at Utah State University, for making the years so enjoyable.

Last but not least, I want to thank my parents, who always give me their endless love and strong mental and emotional supports. I would not succeed in my PhD study without their support and encouragement.

Xue Chen

Contents

List of Figures

Acronyms

AWGN	Additive White Gaussian Noise
BS	Base Station
CINR	Channel Gain-to-Interference-plus-Noise-Ratio
CRN	Cognitive Radio Networks
CSI	Channel State Information
D2D	Device to Device
EE	Energy Efficiency
eNB	Evolved NodeB
LTE	Long Term Evolution
M2M	Machine-to-Machine
MB	Mobile Staion
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MTC	Machine Type Communication
OFCDM	Orthogonal Frequency and Code Division Multiplexing
OFDM	Orthogonal Frequency Division Multiplexing
PPP	Poisson Point Processes
QoS	Quality of Service
SE	Spectral Efficiency
SNR	Signal-to-Noise Ratio
WLANs	Wireless Local Area Networks
WPAN	Wireless Person Area Network
UE	user equipment

Chapter 1

Introduction

1.1 Background

Wireless communication networks have witnessed a tremendous growth in the past decades, which is boosted by ubiquitous communication services such as video streaming, online gaming, social networking, and so on. And this trend will keep on growing exponentially in the next decade. However, the progress to improve wireless network infrastructure is far from satisfying the increasing demand for communication service, especially with the boom of local area services. The future success of wireless networks critically depends on the two factors: network spectral efficiency (SE) and energy efficiency (EE). As a non-renewable natural resource, spectrum must be efficiently used for supporting ever increasing wireless traffic growth and quality of services (QoS) demands from end users. Furthermore, system energy efficiency is becoming more and more important due to the green gas emission control and relatively slow progress on battery technologies.

In order to meet capacity demands from the quick expansion of data traffic growth, heterogeneous network with base stations (BSs) of diverse sizes and various transmission powers are expected to achieve a higher spectral efficiency and energy efficiency. A typical heterogeneous network model consists of Macro-Base Station (M-BS), Pico-Base Station (P-BS),Femto-Base Station (F-BS) and relay base-stations (R-BS). An M-BS transmits at a high power and hence serves a larger coverage area; other types of BSs transmit at a relatively lower power so that their coverage size is also smaller. M-BSs are normally deployed for blanket coverage while other low power BSs are deployed more or less for capacity expansion and coverage extension. Heterogeneous networks have a number of prominent advantages compared to the traditional homogeneous networks. First, a heterogeneous network can greatly improve the wireless link quality since the BSs are now much closer to

the mobiles. Second, due to the coexistence of BSs with different transmit powers, the heterogeneous network can be more energy and spectral efficient. Compared to the traditional homogeneous networks, issues such as mobile association, load balancing, interference management all need to be studied carefully in order to realize the performance gain in a wireless heterogeneous network.

On the other hand, researchers have been seeking for new paradigms to revolutionize the existing wireless networking technologies. Device-to-Device (D2D) communications in the wireless heterogeneous network have been lately used to facilitate proximity-aware services and data traffic offloading, especially with the boom of local area communication services in social networks. D2D communications in cellular networks provide a direct communication between two mobile users without going through a BS and can provide four types of performance gain. The first one is proximity gain as short range communication using a D2D link enables high bit rates, low delays, and lower power consumption. The second one is hop gain as D2D communications use one hop rather than two hops consisting of one uplink and one downlink. The third one is reuse gain as D2D communications can reuse cellular spectrum in an underlay mode. The last one is paring gain, which facilitates new types of wireless services. A UE with D2D capability has the flexibility to switch between cellular mode and D2D communication mode as needed. System spectral efficiency and energy efficiency can be significantly boosted from this new communication paradigm. Meanwhile, new challenges and issues are also arising. How to maximize system capacity while guaranteeing service quality for both cellular users and D2D users stays as a big challenge, especially when dense D2D users are supported in an underlay mode. In order to understand the problems and develop various mechanisms to support desirable D2D communications in cellular networks, we need to be empowered with effective analytical and simulation tools, among which stochastic geometry theory based analytical approaches have been widely used in cellular network study and considered as an effective tool for this purpose.

Furthermore, when evaluating the performance of a system design, QoS, SE, and EE

are usually considered among the most important performance metrics. In reality these three system performance metrics are not independent with each other. Improvement in one of them does not necessarily boost another, sometimes even has a negative impact on another one. For wireless communication in a point to point additive white Gaussian noise (AWGN) channel, SE and EE relationship has been investigated extensively [1–4]. It is either a cup shape curve without considering circuit power or a bell shape curve if circuit power is incorporated. More and more research works have been done to study the tradeoff between EE and SE in the presence of statistically QoS requirements in wireless systems. The concept of effective capacity was first proposed by Wu et al. [5], which is used to model the physical layer fading channel with link layer parameters, such as delay and data rate, provides an effective tool to measure SE and EE with respect to QoS requirements in wireless systems. Under this context, SE is defined as effective capacity per unit bandwidth and EE is defined as energy consumed per effective capacity bit. Hence, analysis of EE-SE relation under the QoS constraint is becoming much more direct for wireless communication in our study.

1.2 Literature Survey

In this section, we provide a survey on state-of-the-art techniques that support D2D communications in wireless heterogeneous network to improve SE and EE.

1.2.1 Related Technologies

The increasing demand for local area services and high data rates have triggered extensive research efforts on improving system capacity and achieving better user QoS. Current 4G cellular technologies have significantly improved physical and MAC layer performance, but they are still lagging behind mobile booming data demands. It is predicted that by 2020, there will be seven trillion wireless devices serving billions of people [6], which is mainly attributed to the advent of new devices such as wearable and machine type communication (MTC) devices. Given the limited availability of spectrum and marginal improvement on spectral efficiency, capacity provision for this enormous number of devices through the con-

ventional cellular communication connecting all of them to the base stations (BSs) may not be sustainable. Hence, researchers have been seeking for new paradigms to revolutionize the traditional communication wireless cellular network. D2D communication using a direct communication link between two mobile users without going through any BS has been considered as a promising technology to improve SE and EE and to provide better user experience in next generation cellular networks [7]. A comprehensive survey on these topics is provided in [8].

In research community, D2D communication was first proposed to provide multi-hop relays in cellular networks [9]. It was then used to support other services such as peer-to-peer communication, multicasting, content distribution, machine-to-machine (M2M) communications, cellular offloading, and so on. The first implementation of D2D communication in cellular network was made by Qualcomm's FlashLinQ [10]. By joint optimizing PHY and MAC layers, FlashLinQ creates an efficient method for timing synchronization, peer discovery, and link management based on OFDM/OFDMA technologies in D2D-enabled cellular networks. If categorized by spectrum reuse mode, the related technologies on D2D communication can be grouped into two types: inband D2D and outband D2D [8]. Inband D2D communication reuses cellular spectrum either orthogonally or non-orthogonally. In the orthogonal mode, part of the cellular resources are dedicated to D2D communication exclusively, while in the non-orthogonal mode D2D communication shares the same radio resources with cellular users. Non-orthogonal mode tends to provide a higher SE than the orthogonal mode. However, it creates interference between D2D communication and cellular communication, which inevitably leads to performance degradation for both. Hence, advanced interference management algorithms are required and they may increase the complexity and computational overhead of cellular and D2D users. For an outband D2D scheme, D2D users generally contain two radio interfaces: one can operate in the cellular spectrum just as normal and the other one can operate in an independent spectrum such as ISM spectrum. Outband D2D communication faces a few challenges in coordinating communications over two different bands.

Comparing D2D communications with other wireless technologies of similar architecture, e.g., wireless local-area network (WLAN) based on IEEE 802.11 standards, wireless person-area network (WPAN) such as Bluetooth and Ultra Wideband technologies, the main difference lies in a central entity in the cellular network such as evolved NodeB (eNB) that is involved in the D2D communication. A general session setup of D2D communication includes following steps [11]: 1) a D2D user initiates a communication request. 2) The BS checks if the communication source and destination are in the same subnet or not. 3) If a number of criteria are met, BS can set up a D2D link for communication. These criteria may include minimum data requirement, D2D capable devices, higher SE/EE with D2D communications, etc. Even if a D2D connection has been set up, UEs can still switch to cellular communication mode if needed. The availability of a supervising/ managing central entity in D2D communications resolves many challenges such as spectrum hole detection, collision avoidance, and synchronization, which may exist in a network without a supervising/managing central entity, such as Cognitive Radio Networks (CRN). Furthermore, D2D communication operating in a licensed band owned by a cellular network can provide a better interference-controlled environment. M2M communication also has a similar architecture as D2D, but M2M communication is between two devices with the help of infrastructure nodes. It is different from D2D communication in the sense that its communication is not constrained by any distance requirement, and it is application-oriented and technology-independent. D2D aims at proximity connectivity and it is technology dependent.

1.2.2 Spectral Efficiency

D2D communication can significantly increase cellular SE, benefited from frequency reuse and multi-user diversity. The main challenge is to deal with co-channel interference between D2D users and cellular users caused by spectrum resource reuse. Extensive research efforts have been spent on solving the problem through efficient interference management [12–16], mode selection, resource allocation and network coding. Paper [12] proposed a scheme to use cellular uplink resources for D2D communication. Since reusing

uplink resources generates interference to the received signals at BS, D2D users monitor the received power of downlink control signals and estimate the pathloss between D2D transmitters and the BS. In order to avoid excessive interference to cellular users, D2D users keep the transmit power below a threshold. Paper [13] proposes two mechanisms to tackle the interference between cellular users and D2D users on the cellular uplink. D2D users read the resource block allocation information from the control channel and avoid using resource blocks that are used by the cellular users in the proximity. Furthermore, D2D interference is broadcast among all D2D users so that D2D users can adjust their transmission power and resource block selection. Interference from D2D communication to uplink transmission is thus kept below a tolerable threshold.

Paper [14] proposes an interference control mechanism based on user locations. First, a dedicated control channel is allocated for D2D users. Cellular users listen to this channel and measure the received SINR. If the SINR is higher than a pre-defined threshold, a report is sent to the eNB. Accordingly, the eNB stop scheduling cellular users on the resource blocks currently occupied by D2D users. The eNB also sends broadcast information regarding the location of the cellular users and their allocated resource blocks. Hence, D2D users can avoid using resource blocks which interfere with cellular users. Paper [15] proposes a scheme to minimize the maximum D2D received power from cellular users. Very similar to the approach in [14], D2D users also measure the signal power levels of cellular users and feed them back to the BS, which then avoids allocating the same frequency-time slot to cellular and D2D users that have strong interference to each other. Another interference cancellation algorithm is proposed in [17] by using Han-Kobayashi rate splitting technique to improve throughput of D2D communications. In rate splitting, the message is divided into two parts, namely, private and public. The private part, as its name suggests, can be decoded only by the intended receiver while the public part can be decoded by any receiver. This technique helps D2D interfered victims to cancel the interference from the public part of the message by running a best-effort successive interference cancellation algorithm. Their simulation results show that throughput improvement is prominent when two D2D users

are far from the BS but close to each other.

A new interference management scheme is proposed in [16,18], where interference control is not achieved by limiting D2D transmission power as in other conventional D2D interference management schemes. The proposed scheme is based on the concept of interference limited area, in which cellular users and D2D users should not be allocated the same resources. Hence, the interference between D2D and cellular users is avoided. But this physical separation limits the scheduling alternatives for the BS and as a consequence multi-user diversity is not fully exploited. Nevertheless, numerical results show that the capacity loss due to multi-user diversity reduction is negligible compared to the gain achieved by their proposal. In [18], the authors propose an interference limited area according to the amount of tolerable interference and minimum SINR requirements for successful transmission, which consists of 1) defining interference limited areas where cellular and D2D users cannot use the same resources; and 2) allocating the resources in a manner that D2D and cellular users within the same interference area use different resources.

Doppler et al. also study several aspects of D2D communications in cellular networks to improve network spectrum efficiency in [19–24]. They discuss optimal mode selection strategies for D2D communication in [19,20] and propose a joint D2D communication and network coding scheme in [21]. In [19], some semi-analytical studies are performed to optimally select the mode of D2D communication in a single cell scenario with one cellular user and one D2D pair. By utilizing power optimization and optimal mode selection, the sum rate increases sevenfold for a D2D connection separated by 10% of the cell radius. The sum rate increase is threefold when supporting a rate guarantee to the cellular user. In [20], they first study the optimal selection of possible resource sharing modes with the cellular network in a single cell, based on which they propose a mode selection procedure for a multi-cell environment. The mode selection algorithm is not only based on the D2D link quality but also takes into account the quality of the cellular link and the interference level under each possible mode. Simulation results show that in the local area scenario, the proposed mode selection improves the sum rate in the network by 50% compared to

pure cellular communication and the ratio of successful D2D communications is more than doubled. The same research group also proposes a joint D2D communication and network coding scheme in [21], where D2D communication is used for uplink message exchange among cellular users before the messages are transmitted to the BS. Then each user sends the coded data containing the original data from both users to the BS. In their scheme, they also propose to group users with complementary characteristics to improve network coding performance rather than just randomly selecting cooperative users, which is much less efficient.

1.2.3 Energy Efficiency

Energy efficiency is another important research area for D2D enabled cellular networks. A common technique to achieve this is to adaptively select operation mode for D2D communication based on user's location and CSI. Usually, the resource allocation problem is formulated as linear or non-linear programming, which is a NP hard, and there is no direct way for the solution. Due to the complexity of these problems, a heuristic algorithm is proposed to solve the problem and only a sub-optimal solution is available. In [25], Xiao et al. propose a power optimization scheme for OFDMA-based cellular networks. They address the joint resource allocation and mode selection problem in a D2D communications, aiming at minimizing total downlink power consumption and propose a heuristic approach using existing subcarrier and bit allocation algorithms in [26, 27]. The heuristic first performs subcarrier and bit allocation for all users in cellular mode and then selects a proper transmission mode for each D2D pair between the direct links and cellular links. Simulation results show that their proposed heuristic algorithm can save the downlink power consumption of the network around 20% compared with the traditional OFDMA system without D2D.

Yu et al. consider resource allocation and power control for D2D communication in a single cell scenario where one cellular user and one D2D pair share the same radio resources [19, 24]. They analyze two power control cases. In the first case, cellular and D2D are treated as competing services without priority. The system is aiming for a greedy sum-rate

maximization under a maximum transmit power constraint. In the second case, cellular users are the prioritized users with guaranteed minimum transmission rate, under the same maximum transmit power constraint. They assume that the instantaneous Channel State Information (CSI) of all links is available at BS which controls the transmit power and resource allocation for D2D links. Optimality is discussed under practical constraints for different resource sharing modes, namely non-orthogonal sharing mode, orthogonal sharing mode, and cellular mode. Authors in paper [28] extend the scenario in [24] to multiple D2D pair multiple resource allocation and propose a maximum-weight bipartite algorithm for optimal power control. The scheme is divided into three steps. First, it performs admission control for D2D connection based on QoS requirement, then allocates powers for each admissible D2D pair and its potential cellular partner. Finally, a maximum-weight bipartite matching based scheme is proposed for resource allocation for cellular and D2D users to maximize overall system throughput. Simulation results show that their approach can significantly improve system performance in terms of D2D access rate and overall network throughput, which provides up to 70% throughput gain compared with other approaches in [12, 15, 29].

In [30], the authors aim to minimize the overall transmission power in a multi-cell OFDM cellular network. They first formulate the problem of joint mode selection, scheduling and power control as mixed integer linear programming, which is proven to be NP-hard in the strong sense and results in the solution of brute-force approach. To reduce the computational complexity, they propose the load control policy with distributed algorithm which performs mode selection and resource allocation cell by cell. The performance of proposed heuristic method is compared with other two schemes: 1) cellular mode in which transmission should go through the BS; and 2) D2D mode in which all D2D users can only communicate directly and passing through the BS is not allowed. Simulation and analysis show that the gain of power efficiency of the proposed method over conventional cellular networks is significant (up to 100%) when the distance between D2D users is less than 150m.

Different from other research work on power efficiency in D2D communication which

usually focus on minimizing transmission power under various constraints, system object in [31] is directly to optimize energy efficiency. In the work, authors of the paper define energy efficiency as a function of transmission rate and power consumption in different transmission mode (cellular and D2D). They propose a heuristic approach which performs power allocation and mode selection to maximize energy efficiency in two steps. First, they obtain the energy efficiency for all possible mode selections of each user through the suboptimal power allocation. Then based on this information, the algorithm selects a mode sequence which has the maximal energy efficiency among all possible mode combinations of users in the second step. Simulation results demonstrate that proposed algorithm can achieve up to 100% gain over other schemes.

1.2.4 Other Aspects

For most problems in D2D communication such as power control, mode selection, scheduling or resource allocation, no matter the proposed algorithm is BS centralized or distributed approach based, most of these algorithms critically rely on Channel State Information (CSI). However, CSI is usually obtained from channel estimation, which could be inaccurate in reality and also causes high signaling overhead in some scenarios. Moreover, if there are a large number of D2D communication users in the cellular network, to obtain CSI of different D2D communication links such as from D2D users to cellular users or from D2D users to BS, is both time consuming and bandwidth consuming, which causes system delay, and requires extra system resources to transmit the CSI.

Approaches based on stochastic geometry theory have been widely used to analyze complex wireless system design issues [32]. By modeling the spacial distribution of network nodes and mobile users as homogeneous spatial Poisson Point Processes (PPP), it is more convenient to study the large dimension wireless system problems through analytical approaches instead of seeking complex system-level simulations for cellular network in conventional way, which is usually modeled by a large number of parameters (e.g. grid model). One obvious advantage of analytical approach is that it only uses the link's statistical information and user distribution to evaluate system performance such as network spectral

efficiency, energy efficiency and coverage, and there is no need for user's instant CSI. Secondly, the simulation approach based on conventional grid model for cellular network is highly idealized and becoming less and less accurate as cell size shrinks to support a dense user capacity.

These research papers [32–35] are focusing on D2D communication in cellular network to evaluate network performance and find out the bottleneck of improving spectral efficiency and energy efficiency. In [32], authors develop a general model to evaluate downlink coverage /outage probability and rate for multi-cell heterogeneous network using stochastic geometry. The analytical results are compared to the grid model and actual base station deployment, which suggests to be more tractable and capable of capturing opportunistic and dense placement of base stations. Research work in [33, 34] gives out downlink SINR distribution for multi-tier heterogeneous cellular network, which consists of multiple tiers of transmitters (e.g., macro-, pico-cell and femto-cells). Authors in [35] give out coverage analysis for OFDMA-based cellular networks, where two types of interference management schemes: strict fractional frequency reuse and soft frequency reuse are discussed. Based on the analysis expressions, they propose a SINR-proportional resource allocation strategy which can increase sum-rate as well as coverage for cell-edge users.

1.3 Our Approach and Thesis Outline

The major goal of this research is to investigate resource allocation and interference management algorithms to improve user experience, system spectral efficiency, and energy efficiency for D2D communication underlaying heterogeneous networks. By exploiting multi-user diversity and CSI, this research work aims to design integrated algorithms to utilize the spectrum and energy resources efficiently for the heterogeneous wireless networks.

First, we provide an extensive review of background and current technology development on D2D communications in Chapter 1. Research work on D2D communications for improving system spectral efficiency, energy efficiency, users' QoS are discussed. This chapter provides new insights to current research works which lead to our own research topics about analytical evaluation of SINR distribution for D2D communications, power

and resource allocation in D2D underlaying cellular network, energy efficient resource allocation in D2D communications and tradeoff between EE and SE in delay constrained communications.

In Chapter 2, we present an approach to jointly optimize the downlink and uplink resources for mobile association in a heterogeneous network. The proposed scheme considers both capacity and uplink power consumption during mobile association. A gradient descent search algorithm is developed to search for the optimal mobile association that can maximize the system capacity and also minimize mobile uplink transmission power consumption. The simulation for the network model is based on 3GPP case 1, which demonstrates a good performance improvement on network spectral efficiency and energy efficiency in our proposed scheme.

In Chapter 3, we give out an analytical evaluation of SINR distribution in a D2D communications underlaying cellular network model, which is derived based on stochastic geometry theory. The users' 2-dimensional location is draw from a Poisson Point Process (PPP). Only statistical channel information and user distribution is needed for evaluation of system metrics such as network coverage, outage probability and throughput. The SINR distribution is analyzed for both cellular users and D2D users in the uplink and downlink resource sharing scenario, respectively. We also validate our analysis with the simulated network model. The conclusion draw from this chapter can provide a guideline in design of D2D communication network when considering power control, interference management and resource allocation.

We begin to discuss the power/resource allocation for D2D communications in Chapters 4 and 5. A sub-optimal distributed resource allocation and power scheme based on Stackelberg game framework is proposed for improving network capacity in Chapter 4. The system aims to maximize the number of supportable underlay D2D users while guaranteeing QoS of the prioritized cellular users. Thereafter, the problem and system objective are formulated with Stackelberg game theoretical model. Due to computational complexity, we decompose the problem into sub-problems and solve it in two steps, first grouping

DUEs that share the same radio resource of a CUE, and then allocating resources to them distributively through a price mechanism. Simulation results show that our proposed distributed algorithm converges fast and the system capacity of D2D communication network is significantly improved. In Chapter 5, we consider a joint resource/power allocation and mode selection for D2D communication in the OFDMA based cellular network. The optimization problem is formulated as a mixed integer nonlinear programming, which is proven NP-complete. Thus we develop a dual optimization framework to transform the intractable problem into equivalent problem and solve it with reasonable computational complexity. Analytical results show that our scheme can achieve a much higher system throughput compared with other schemes.

In Chapter 6, we investigate aspects of network energy efficiency when allocating radio resources for D2D communication networks. The resource allocation between CUEs and DUEs is modeled as a non-cooperative game, where cellular users or D2D users determine which resource blocks to allocate and how much power they plan to transmit correspondingly so as to maximize a utility function. The utility function in the work is defined as the achievable rate normalized by power consumption. In flat fading channel, we prove there exist a unique point of Nash equivalence for our proposed game model. We also propose a method for the game to converge to its Nash equivalence.

In Chapter 7, we study the fundamental tradeoff between energy efficiency and spectral efficiency in presence of statistical QoS requirements for the delay constrained communication. System QoS metric is incorporated and measured through effective capacity, based on which the spectral efficiency is defined as effective capacity per unit bandwidth and energy efficiency is defined as energy consumed per effective capacity bit. Total power consumption consists of both circuit power and transmission power. To exploit the EE-SE relation under different SNR regime, we propose a generic close-form approximation by using a curve fitting approach.

In Chapter 8, we make conclusions and summarize contributions for the dissertation.

Chapter 2

Joint Uplink and Downlink Optimal Mobile Association in a Wireless Heterogeneous Network

In this chapter, we discuss the architecture and system model of the heterogeneous networks, then we present the mathematical formulations for the proposed mobile association scheme and develop a gradient descent algorithm to search the sub-optimal solutions for two scenarios, full frequency reuse and partial frequency reuse respectively. In the end, simulation results and numerical analysis are provided.

2.1 Heterogeneous Network Structure

A typical heterogeneous network is illustrated in 2.1, which has a prominent advantage compared to the traditional homogeneous networks. A heterogeneous network can greatly help reduce the uplink transmission power since the BSs are much closer to the mobiles. Furthermore, due to the coexistence of base stations with different transmitting powers, high power base stations can offer blanket coverage while low power nodes can be capacity boosters. So the heterogeneous network can be more energy and spectral efficient. In our study, there exist two types of base stations. One type is a Macro-Base Station (M-BS) that transmits at a higher power and hence serves a larger coverage area; the other type is a Micro-Base Station (m-BS) that transmits at a lower transmitting power with a smaller coverage area. There will be one M-BS each sector while several m-BSs can be deployed each sector per capacity needs. In this paper, we focus on a specific type of m-BSs, called relay node (RN), due to its unique multi-hop feature that imposes extra complexity to the problem under investigation. A RN transmits at a low power and can help forward the information between MSs and M-BSs on both uplink and downlink. As shown in Figure 2.1, in a relay network, a MS can connect to the wireless network either through a direct link

(D-link) to a M-BS or through an indirect link (I-link) to a RN, which is further connected to its donor M-BS via a backhaul.

The following notations are used. N_c denotes the number of M-BSs in network. N_r denotes the total number of RNs per sector. The total number of MSs in the network is N_u. We use $h_{k,0,i}$ and $h_{k,j,i}$ to denote the channel gain on the D-link between k^{th} MS and M-BS in the i^{th} sector, and the channel gain on the I-link between k^{th} MS and j^{th} RN in the i^{th} sector, respectively. For simplicity but without loss of generality, in the work, we assume channels are reciprocal, i.e., an uplink channel and a downlink channel between the same communicating parties have the same channel gain. $C_{k,0,i}$ represents the D-link bandwidth needed to support MS k if it is associated with M-BS in the i^{th} sector while $c_{k,j,i}$ represents the I-link bandwidth needed to support MS if it is associated with j^{th} RN in the i^{th} sector. C denotes the total system bandwidth. Deploying multiple RNs in each sector will create cell splitting within that sector. Each RN may reuse total bandwidth C or part of C. $X_{k,j,i}$ indicates the association status between the k^{th} MS and the j^{th} node in the i^{th} sector. Here $j = 0$ represents the M-BS in that sector and $j > 0$ represents the RNs in that sector. $X_{k,j,i} = 1$ indicates that the k^{th} MS is associated with the defined node while $X_{k,j,i} = 0$ indicates otherwise. The transmit power of the M-BS i is P_i^b and the transmission power for the RN j is P_j^r.

During the uplink open loop power control, the target SNR at the receiving node is set to 10dB. All the BSs bear the same noise level at σ_W^2 per resource block (RB). $P_{k,j,i}$ represents the desired transmission power of MS k when it is associated with j node in the i^{th} sector. The goal of uplink power control is to achieve the same level of the designated SNR at the receiving node for all MSs. So we have $P_{k,j,i} = \frac{10\sigma_W^2}{\mathrm{PL}_{k,j,i}}$, where $\mathrm{PL}_{k,j,i}$ denotes the average channel gain or simply pathloss between k^{th} MS and j^{th} node in the i^{th} sector.

In the heterogeneous networks, due to the difference between the transmission powers of M-BSs and the RNs, conventional best power or best-quality based association schemes may lead to a highly uneven traffic distribution and thus a low resource utilization at RNs. The recently proposed range-expansion association scheme uses a bias to offset the power

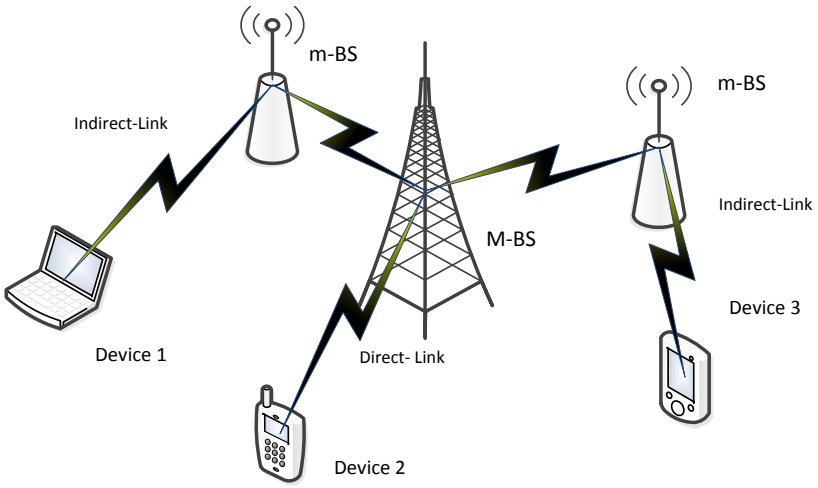

Fig. 2.1: A heterogeneous relay network

difference between M-BSs and RNs so that more MSs can be associated with RNs [36]. However, there has been little study on how to select the best bias or how to best save mobile power consumptions when doing mobile association. Furthermore, most of the existing mobile association schemes are based on either only downlink information or only uplink information but not both [37]. Little has been done to jointly consider downlink and uplink information during mobile association. We propose an association scheme that can optimize an objective function that jointly considers downlink capacity and the mobile uplink power consumption. The scheme will be formulated and evaluated in two scenarios: full frequency reuse and partial frequency reuse.

2.2 Full Frequency Re-use Scheme

In an OFDM based wireless network, usually different RBs are allocated to different MSs within a sector so that there is not much co-channel interference among MSs in the same sector. The co-channel interference mostly exists between MSs in the different sectors or different cells. In the full frequency scenario, all the M-BSs and RNs use the same frequency band, for both downlink and uplink transmissions. There is no inter-cell interference coordination between the M-BSs and RNs. It represents the most aggressive spectrum reuse

scenario and also causes the strongest co-channel interference between M-BSs and RNs. In the full frequency reuse case, the downlink received signal-to-interference-noise (SINR) at the k^{th} MS through direct link and indirect link, can be expressed as

$$\mathrm{SINR}_{k,0,i}^{D-DL} = \frac{|h_{k,0,i}|^2 P_i^b}{\sum_{i' \neq i}^{N_c} |h_{k,0,i'}|^2 P_{i'}^b + \sum_{i=1}^{N_c} \sum_{j=1}^{N_r} |h_{k,j,i}|^2 P_j^r + \sigma_w^2}. \tag{2.1}$$

$$\mathrm{SINR}_{k,j,i}^{D-DL} = \frac{|h_{k,j,i}|^2 P_j^r}{\sum_{i=1}^{N_c} |h_{k,0,i}|^2 P_{i'}^b + \sum_{i=1}^{N_c} \sum_{j' \neq j}^{N_r} |h_{k,j',i}|^2 P_{j'}^r + \sigma_w^2}. \tag{2.2}$$

The corresponding uplink SINRs for k^{th} MS are:

$$\mathrm{SINR}_{k,0,i}^{D-UL} = \frac{|h_{k,0,i}|^2 P_{k,0,i}}{\sum_{i' \neq i}^{N_c} \frac{1}{N_u^{i'}} \sum^{N_u^{i'}} |h_{k,0,i'}|^2 P_{k,0,i'} + \sum_{i=1}^{N_c} \sum_{j=1}^{N_r} \frac{1}{N_u^i} \sum^{N_u^i} |h_{k,0,i}|^2 P_{k,j,i} + \sigma_w^2}. \tag{2.3}$$

$$\mathrm{SINR}_{k,j,i}^{I-UL} = \frac{|h_{k,j,i}|^2 P_{k,j,i}}{\sum_{i=1}^{N_c} \frac{1}{N_u^i} \sum^{N_u^i} |h_{k,j,i}|^2 P_{k,0,i} + \sum_{l=1}^{N_c} \sum_{\substack{l \neq i \\ j' \neq j}}^{N_r} \frac{1}{N_u^l} \sum^{N_u^l} |h_{k,j,i}|^2 P_{k,j',l} + \sigma_w^2}. \tag{2.4}$$

On the uplink, the co-channel interference for a MS comes from the MSs that use the same RBs in the neighboring cells. In reality, the MSs that cause interference change from one scheduling cycle to another cycle due to the scheduling dynamics on RB allocations. However, from mobile association point of view, an average interference to a MS is more important than the instantaneous interference level at each scheduling cycle. Thus in our study the average interference to a MS is calculated. $\frac{1}{N_u^{i'}} \sum^{N_u^{i'}} |h_{k,0,i'}|^2 P_{k,0,i'}$ in 2.3 and 2.4 represents the average interference to MS k from all the MSs the neighboring M-BSs while $\sum_{l=1}^{N_c} \sum_{\substack{l \neq i \\ j' \neq j}}^{N_r} \frac{1}{N_u^l} \sum^{N_u^l} |h_{k,j,i}|^2 P_{k,j',l}$ represents the average interference from all the MSs in the neighboring RNs.

According to Shannon's theorem, the downlink resource allocated to MS k for the D-link and I-link can be determined through formulation.

$$c_{k,j,i} = \frac{\Phi_k}{\log_2(1 + \mathrm{SINR}_{k,j,i})}. \tag{2.5}$$

Φ_k is a parameter reflecting user's QoS requirement.

In this chapter we assume that the wireless backhaul quality $\mathrm{SINR}_{j,i}^{BH}$ is the same for all j, i. The radio resource needed for MS k on the backhaul is

$$c_{k,j,i}^{b} = \frac{\Phi_k}{\log_2(1 + \mathrm{SINR}_{j,i}^{BH})}. \tag{2.6}$$

The optimal mobile association proposed in this paper will maximize the downlink system capacity as well as minimize the MS uplink power consumption. The optimization formulation is expressed as:

$$\begin{aligned}
\max \quad G(x) \;=\; & \sum_i \sum_j \sum_k x_{k,j,i} - \rho_1 \Big(\sum_i \sum_k x_{k,0,i} c_{k,0,i} \\
& + \sum_i \sum_j \sum_k x_{k,j,i} (W c_{k,j,i} + c_{k,j,i}^{b}) \Big) - \rho_2 \sum_i \sum_j \sum_k x_{k,j,i} P_{k,j,i} \qquad (2.7) \\
\text{s.t.} \quad & \\
& \sum_k x_{k,0,i} c_{k,0,i} + \sum_j \sum_k x_{k,j,i} c_{k,j,i}^{b} < C_i^{\mathrm{BS}}, \qquad (2.8) \\
& \sum_k x_{k,j,i} (c_{k,j,i} + c_{k,j,i}^{b}) < C_j^{\mathrm{RN}}, \qquad (2.9) \\
& \sum_i \sum_j x_{k,j,i} = 1 \;\; \text{or} \;\; 0. \qquad (2.10)
\end{aligned}$$

The first term in (2.7) is the total number of accepted mobiles. The second term in (2.7) is the total consumed hypothetical resources while the third item is the total sum of the mobile power consumption on the uplink. We choose to minimize the sum of power density instead of total actual mobile power consumptions so that all the mobiles will have equal chance to save unit power consumption regardless of the sizes of their allocated uplink

bandwidth. Coefficients ρ_1, ρ_2 are the relative weights of the total consumed resources, and of the sum of uplink power density with respect to the total system capacity in terms of the number of granted mobiles. Coefficient W in (2.7) specifies the relative weight of the RN resources with respect to the M-BS resources in the objective function. We can set $W < 1$ so that RNs resources are less weighted than the M-BS resources. Through proper choice of W , there would be more traffic offloaded from M-BSs to RNs, so that RN resources can be more effectively utilized, which actually enlarges a RNs coverage area [38].

The above optimization problem is a 0-1 knapsack problem, and it is NP-hard. An optimal solution is difficult to derive, especially given a large number of MSs and BSs in the network. In the following, we propose a pseudo-optimal solution based on a gradient descent method. To apply the gradient descent method, we relax integer variables $x_{k,j,i}$ into real variables in the range of $[0, 1]$. The value for each $x_{k,j,i}$ indicates the probability of mobile k being associated with node j in sector i . When the gradient search completes, all the $x_{k,j,i}$ are ranked in a descending order, so that the nodes with higher probability values will be selected to serve the MSs. The whole procedure stops when all the MSs are accepted in the network (under-loaded case) or the boundaries of the constraints are reached (over-loaded case). The gradient value of the objective function with respect to each association probability is evaluated as follows

$$\frac{\partial G(x)}{\partial x_{k,0,i}} = 1 - \rho_1 c_{k,0,i} - \rho_2 P_{k,j,i}, \tag{2.11}$$

$$\frac{\partial G(x)}{\partial x_{k,j,i}} = 1 - \rho_2 (W c_{k,j,i} + c^b_{k,j,i}) - \rho_2 P_{k,j,i}. \tag{2.12}$$

The association probability $x_{k,j,i}$ is updated along the direction $\Delta_{k,0,i} = \partial G(x)/\partial x_{k,0,i}$ and $\Delta_{k,j,i} = \partial G(x)/\partial x_{k,j,i}$ according to equations (2.13) and (2.14). The weight coefficient W is crucial in the optimization of the network capacity and resource utilization. A low value of W reduces the weight of the RN resources in the objective function (2.7), which

helps to increase the RN utility.

$$x_{k,0,i}(t+1) = x_{k,0,i}(t) + \delta\Delta_{k,0,i}, \tag{2.13}$$

$$x_{k,j,i}(t+1) = x_{k,j,i}(t) + \delta\Delta_{k,j,i}. \tag{2.14}$$

Coefficient δ indicates the step size for each update. After each update, total resources consumed at each node (M-BS or RN) are calculated. The update for $x_{k,j,i}$ continues till constraints (2.23) and (2.24) are reached with equality.

2.3 Partial Frequency Re-use for Better Interference Management

Due to the co-existence of high power and low power nodes, MSs served by the low power nodes, especially the ones located at the cell edge, may subject to strong interference from the neighboring high power nodes in the full frequency reuse scenario. Inter-cell interference control techniques such as partial frequency reuse have been widely used to improve the cell edge user performance in the traditional homogenous networks [39]. In this chapter, we extend the idea of partial frequency reuse to the heterogeneous network by using different frequency sub-bands for the M-BSs and RNs in the same sector. More specifically, the M-BS will vacate a portion of its bandwidth C and allow all the RNs in its sector to use the vacated bandwidth, thus effectively creating cell-splitting within each sector with no inter-layer interference between high power M-BS and low power RNs. The bandwidth division between M-BS and RNs in each sector can be adjusted in order to maximize the system capacity. Compared with the full frequency reuse scheme, a better SINR is achieved in the partial frequency reuse scheme due to the suppressed interference from the adjacent cells. As illustrated in Figure 2.2, RNs operate in a half-duplex TDD mode to avoid self-interference. In time slot T_1 , M-BS transmits to its MSs on band F_{11} and simultaneously transmits to its connected RNs on band F_{12} . In time slot T_2 , M-BS transmits to its MSs on band F_{21} while all the RNs transfer the data received from M-BS to the MSs on band F_{22}. A similar reuse scheme applies to the uplink. The total radio bandwidth C for each sector is the same. The resource partition between M-BS and RN

resource can be expressed as:

$$C_1 = (\frac{T_1}{T_1+T_2} + \frac{T_2}{T_1+T_2} \cdot \frac{F_{21}}{F_{21}+F_{22}})C, \quad (2.15)$$

$$C_2 = \frac{T_2}{T_1+T_2} \cdot \frac{F_{22}}{F_{21}+F_{22}}C. \quad (2.16)$$

C_1 is the equivalent total bandwidth for a M-BS and C_2 is the equivalent total bandwidth for a RN in the partial frequency re-use case. On the downlink transmission, the received SINR for MS k can be expressed as

$$\mathrm{SINR}_{k,0,i}^{D-DL} = \frac{|h_{k,0,i}|^2 P_i^b}{\sum_{i' \neq i}^{N_c} |h_{k,0,i'}|^2 P_{i'}^b + \sigma_w^2}, \quad (2.17)$$

$$\mathrm{SINR}_{k,j,i}^{I-DL} = \frac{|h_{k,j,i}|^2 P_j^r}{\sum_{i=1}^{N_c} \sum_{j' \neq j}^{N_r} |h_{k,j',i}|^2 P_{j'}^r + \sigma_w^2}. \quad (2.18)$$

And for uplink, the expressions are similarly defined:

$$\mathrm{SINR}_{k,0,i}^{D-UL} = \frac{|h_{k,0,i}|^2 P_{k,0,i}}{\sum_{i' \neq i}^{N_c} \frac{1}{N_u^{i'}} \sum^{N_u^{i'}} |h_{k,0,i'}|^2 P_{k,0,i'} + \sigma_w^2}, \quad (2.19)$$

$$\mathrm{SINR}_{k,j,i}^{I-UL} = \frac{|h_{k,j,i}|^2 P_{k,j,i}}{\sum_{l=1}^{N_c} \sum_{\substack{l \neq i \\ j' \neq j}}^{N_r} \frac{1}{N_u^l} \sum^{N_u^l} |h_{k,j',l}|^2 P_{k,j',l} + \sigma_w^2}. \quad (2.20)$$

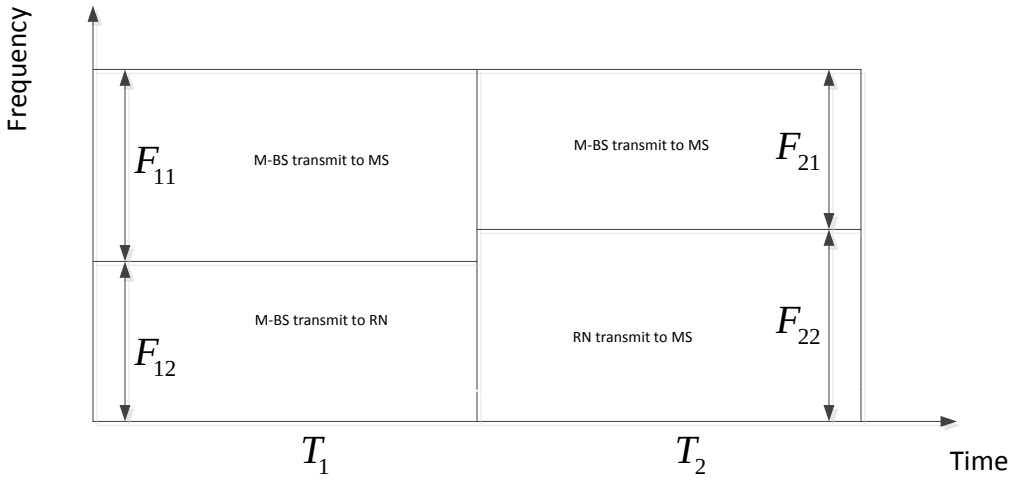

Fig. 2.2: Illustration of partial frequency re-use scheme

Let $r = \frac{T_1}{T_1+T_2} + \frac{T_2}{T_1+T_2} \cdot \frac{F_{21}}{F_{21}+F_{22}}$. The optimal mobile association can be formulated as the following:

$$\max \quad G(x) = \sum_i \sum_j \sum_k x_{k,j,i} - \rho_1 (\sum_i \sum_k x_{k,0,i} c_{k,0,i} \tag{2.21}$$

$$+ \sum_i \sum_j \sum_k x_{k,j,i} (W c_{k,j,i} + c^b_{k,j,i})) - \rho_2 \sum_i \sum_j \sum_k x_{k,j,i} P_{k,j,i} \tag{2.22}$$

s.t.

$$\sum_k x_{k,0,i} c_{k,0,i} + \sum_j \sum_k x_{k,j,i} c^b_{k,j,i} < C_1, \tag{2.23}$$

$$\sum_k x_{k,j,i} (c_{k,j,i} + c^b_{k,j,i}) < C_2, \tag{2.24}$$

$$\sum_i \sum_j x_{k,j,i} = 1 \quad \text{or} \quad 0. \tag{2.25}$$

Where $C_1 = r \times C$, $C_2 = (1 - r) \times C$. Coefficient W is the weight of RN resources with respect to M-BS resources. We can apply gradient descent algorithm in the similar way to search the pseudo-optimal solution.

2.4 Numerical Simulation and Analysis

To verify the performance of the proposed mobile association scheme, a simulation platform has been built in Matlab by following 3GPP case 1 simulation methodology specified in [40]. The system model is based on a 19-cell 3-sector three-ring hexagonal cell structure. In each sector 4 RNs are uniformly deployed around the M-BS. All the M-BSs transmit at 46dBm and all the RNs transmit at 30dBm. Maximum MS transmission power is set to 23dBm. We use Log-normal shadowing and Okumura-Hata model as the propagation model for both direct and indirect links on downlink and uplink. In the simulation, we set $W = 1/6, \rho_1 = 1, \rho_2 = 0.015$. As shown in Fig. 2.3, radio resource utilizations of M-BSs and RNs are compared between our proposed scheme and the traditional best power mobile association scheme in the full frequency reuse scenario. MSs are uniformly distributed in the network. The first 57 nodes represent M-BSs while the rest represent the RNs. From the results, we observe that both mobile association schemes achieve 100% utilization at M-BSs. The performance difference lies in the utilization rates for RNs. Most of the RNs in the joint optimal scheme have a utilization rate of more than 90%, while in best power scheme, RNs can only achieve an average of 48% utilization rate. Through the proposed optimal mobile association, more MSs are offloaded from M-BS to the RNs and traffic load is better balanced. So the overall network capacity increases. And the same improvement for load balancing and network capacity can be observed in the partial frequency re-use scenario.

In Figures 2.4 and 2.5, we simulate the MS received downlink SINR distribution for both full frequency and partial frequency reuse scenarios. The results from range expansion and best power association schemes are also plotted in the same figures for comparison. SINR distributions are plotted separately for MSs associated with M-BSs and with RNs. As expected, the SINR performance of the proposed scheme lies in between the range expansion and best power schemes. With joint consideration of downlink and uplink during the mobile association, more MSs would be offloaded from M-BSs to RNs, although some of them may receive a stronger signal from the M-BSs. Due to the coverage shrinkage

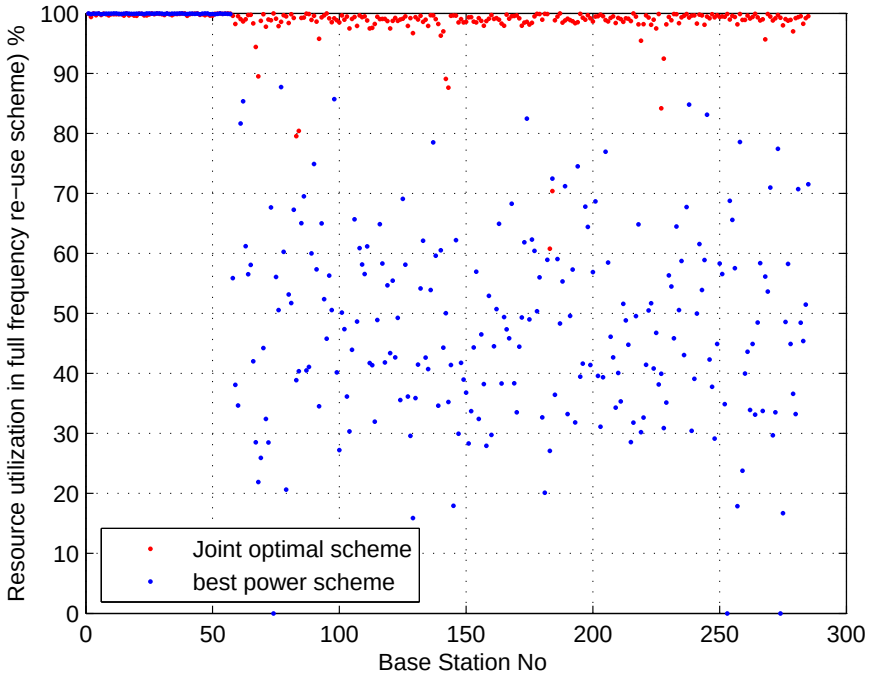

Fig. 2.3: Resource utilization rate comparison between best power and joint optimization

of M-BSs and expansion of RNs by using the optimal mobile association scheme, M-BS associated MSs will get a better SINR distribution while RN associated MSs will have a worse SINR distribution, compared with the best power association scheme. This has been demonstrated in Figure 2.4. From the same figure, we can also observe the proposed scheme can achieve a much better SINR distribution for the MSs associated with RNs than the range expansion scheme, which leaves 35% RN associated MSs below -10dB in the full frequency reuse case. However, the SINR difference among all three mobile association schemes becomes smaller in the partial frequency reuse case, in which high power M-BSs do not cause any interference to the low power RNs.

Simulation results in Figure 2.6 show the uplink transmission power and uplink SINR distributions for the full frequency reuse scenario. The mobile power consumption for the joint optimization scheme lies in between the best power and range expansion schemes. As the target received SNR at base stations is fixed to 10dB, the less the transmission power of a MS, the less interference the MS will cause to others. With the joint optimization

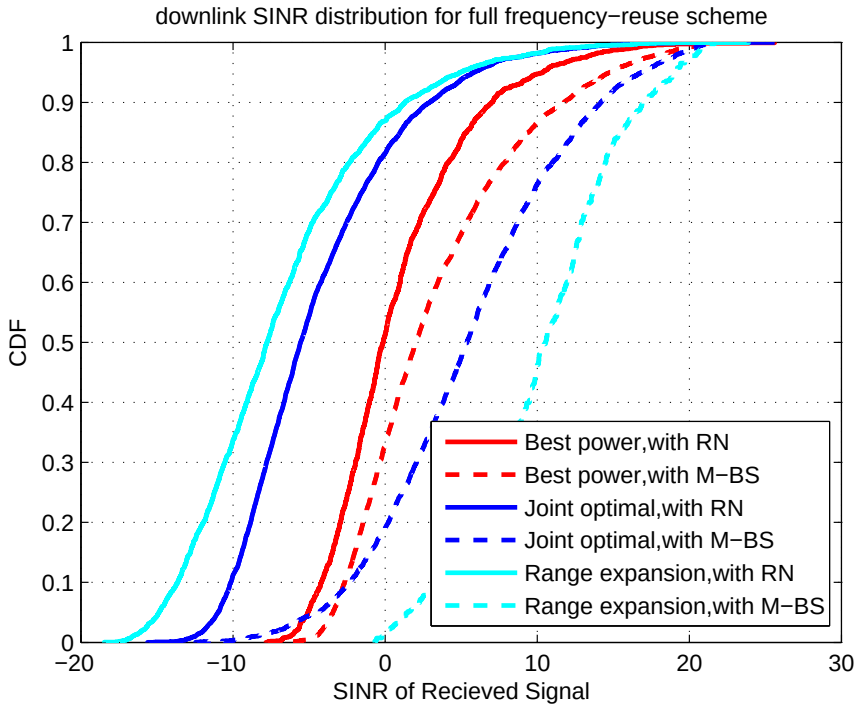

Fig. 2.4: MS received SINR in full frequency re-use scenario

scheme, uplink SINR distribution is better than that of the best power association scheme but slightly worse than that of the range expansion scheme. In the partial frequency reuse scenario as shown in Figure 2.7, inter-layer interference is eliminated, so the SINR distributions for all the schemes are improved. Without uplink inter-layer interference, best power association scheme actually achieves a slightly better uplink SINR performance than the joint optimization scheme while the range expansion has the poorest uplink SINR distribution. When considering uplink and downlink jointly in mobile association, the optimal scheme will provide the best compromise among downlink SINR, uplink SINR and mobile power consumptions. More importantly, it also gives the highest system capacity among all schemes as will be shown later. Worth of mentioning, by adjusting ρ_1 and ρ_2 , the optimal scheme can achieve different compromise between uplink and downlink.

2.5 Summary

This chapter presents a new approach to jointly optimize downlink and uplink based

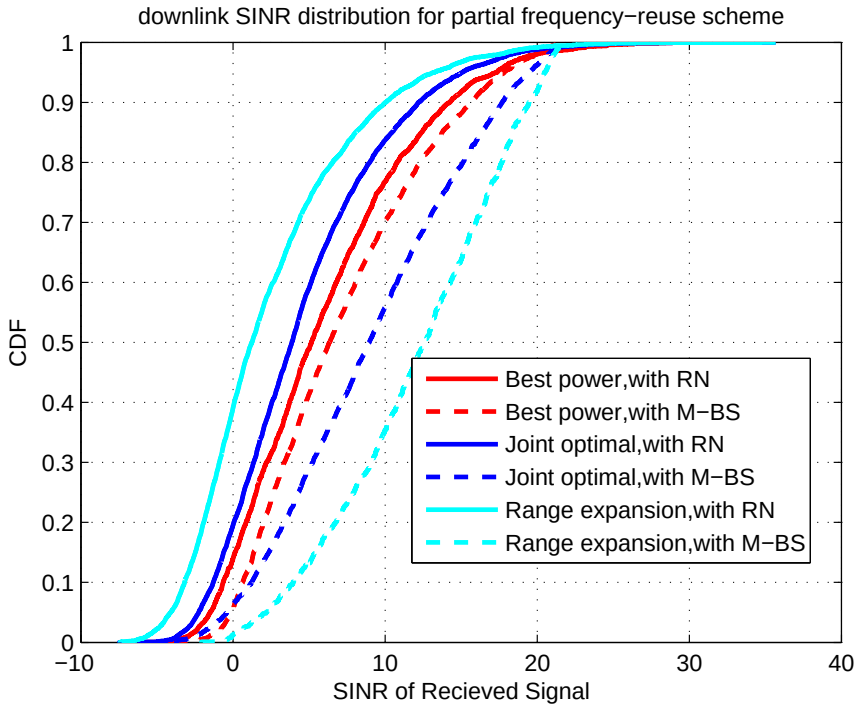

Fig. 2.5: MS received SINR in partial frequency re-use scenario

mobile association in a wireless heterogeneous network. The proposed scheme considers both capacity and uplink power consumption during mobile association. A gradient descent search algorithm is developed to search for the optimal mobile association that can maximize the system capacity and also minimize mobile uplink transmission power consumption. The simulation results demonstrate that the proposed scheme can effectively achieve both high spectrum efficiency and energy efficiency. The mathematical model proposed in this paper can also be used for a more generic mobile association study by considering different design objectives.

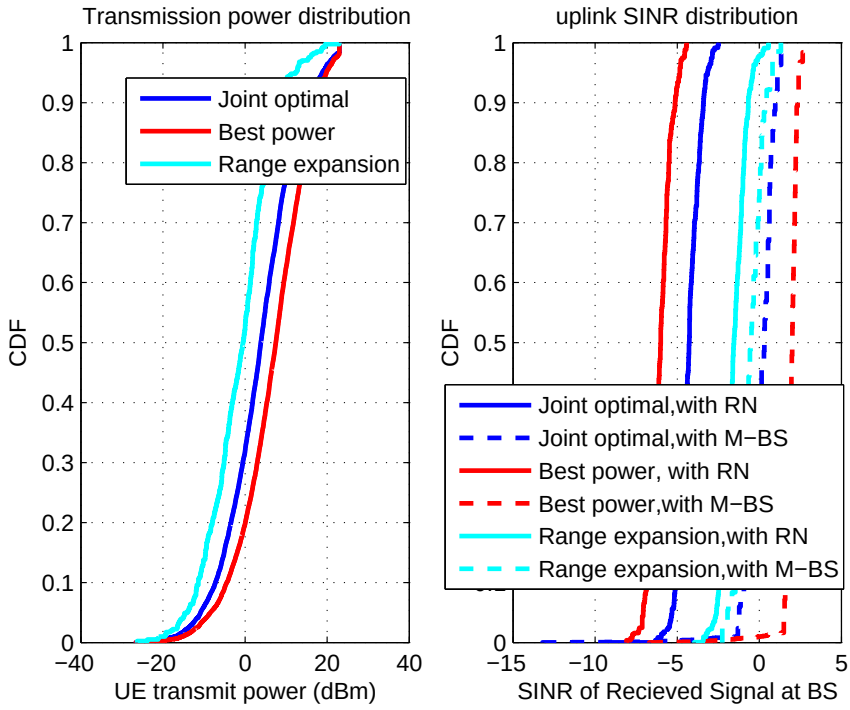

Fig. 2.6: MS transmit signal and SINR for full frequency re-use scheme

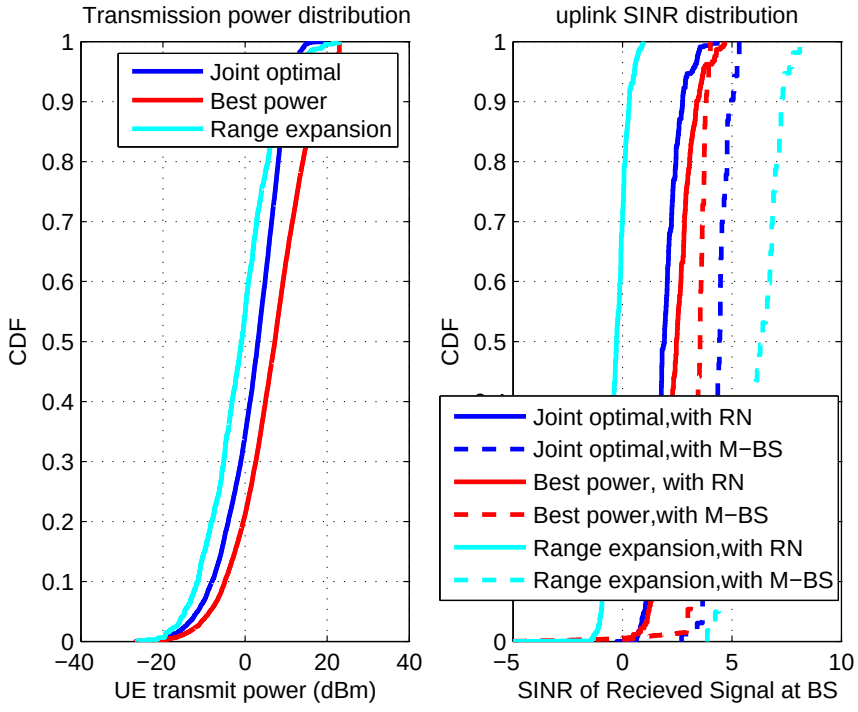

Fig. 2.7: MS transmit signal and SINR for partial frequency re-use scheme

Chapter 3

Downlink and Uplink Coverage of Device-to-Device Communications

In this chapter, we consider the D2D communication in a single-cell scenario and provide an analytical framework to evaluate the SINR distribution and coverage in a D2D underlaid cellular network based on stochastic geometry theory. The spacial distribution of D2D pairs are modeled with the homogeneous Poisson Point Process (PPP). The instantaneous channel state information (CSI) of cellular links and D2D links is not available at the base station. we analyze and derive the uplink and downlink SINR distribution for both cellular users and D2D users given only statistical channel information. Simulation results validate our theoretical analysis based on the proposed system model. This analytical tool can be conveniently used to evaluate other aspects of D2D underlaying cellular network such as outage probability, network throughput and provide critical insights for network design guidelines on power control, interference management, resource allocation and so on.

3.1 Problem Formulation

3.1.1 System Model

The system consists of one BS located in the cell center and user equipments (UEs) scattered around the cell. There are two types of UEs in the network, namely Cellular UE (CUE) and D2D UE (DUE). A CUE communicates directly to the BS for both its downlink and uplink communications. CUEs are considered as the primary service users to allocate cellular radio resources and thus their performance is of primary interests when designing such a D2D underlay network. D2D communication provides complementary services when the direct communication of a D2D pair becomes desirable. Although the communication

between a D2D pair can go through the cellular networks by using two hops, i.e., uplink and then downlink, the performance and spectrum efficiency may go down, compared with the D2D direct communication. In our study, we assume that a D2D pair is randomly dropped in a virtual circle, the center of which is derived from the PPP distribution. The radius of circle R_d is set below certain threshold to keep the link quality of D2D communication. Spatial distribution of CUEs and DUEs both follow an independent homogeneous PPP model with an intensity of λ_c and λ_d, respectively.

We consider a co-channel deployment between CUEs and DUEs. Be more specific, a cellular channel (either a frequency band or a time slot) can be allocated to one CUE and multiple D2D pairs. So there is no co-channel interference among CUEs within a cell. But the co-channel interference can exist between a CUE and multiple D2D pairs as well as between two D2D pairs. The BS has in total N channels. In the underlay communication mode, it is desirable for the DUEs not to cause harmful interference to the CUEs. The channel path loss exponent between BS and a CUE is α_c, and between a D2D pair is α_d. We assume the fast fading between BS and a CUE and the fast fading between any two DUEs are independent and identically distributed Rayleigh fading process. In the work, we number CUEs and DUEs separately. h_i^{bc} denotes the fading coefficient between the BS and i_{th} CUE, h_j^{bd} is the fading coefficient between the BS and j_{th} DUE, h_{ij}^{cd} denotes the fading coefficient between the ith CUE and j_{th} DUE, h_{jk}^{dd} denotes the fading coefficient between the j_{th} DUE and the k_{th} DUE, h_{jj}^{dd} denotes the fading coefficient between the transmitter DUE and receiver DUE in the jth D2D pair. Downlink and uplink channels are assumed to be symmetric. N_0 is average power level for the AWGN noise. The following notations are also used in the paper. A illustration of our system model is shown as in 3.1.

- $\Phi_{d,i}$: the set of DUEs that share the same channel with CUE i in a cell
- K: Number of D2D pairs
- P_b: BS transmission power
- P_d: D2D user transmission power

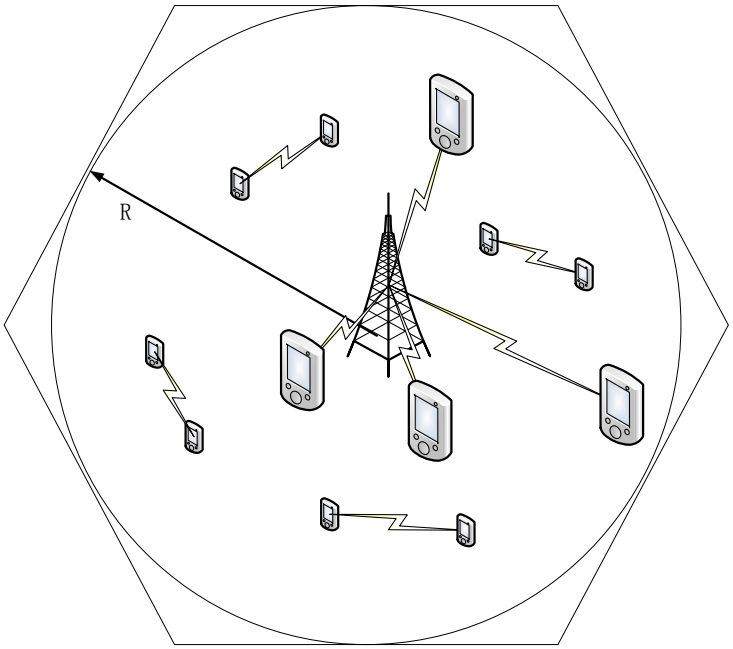

Fig. 3.1: Illustration of Device-to-Device underlaying cellular network

- P_c: Cellular user uplink transmission power
- R: Radius of the cellular cell

3.1.2 D2D Pair Location Model

We assume that all the D2D pairs in the cell have been pre-formed and are randomly distributed within the cell. The instantaneous D2D channel state information is not available at the BS. The 2-D geometry location distribution of D2D pairs is modeled based on the homogeneous Poisson point process with an intensity λ_d. In our system model, a channel can be shared among a number of D2D pairs and one CUE that can be in any location within the cell. Since there are N channels in the system, each D2D pair has the probability $1/N$ to be allocated to a particular channel. According to Coloring Theorem [41], the location model for co-channel D2D users can be still considered as a new PPP with intensity λ_d/N. In the work, for simplicity but without loss of generality, we still use λ_d to represent the effective intensity for the co-channel D2D pairs in the cell. The PPP model has the following properties [42].

- The average number of DUE pairs in an area with size $\mathcal{A}_\mathcal{S}$ is denoted as $\lambda\mathcal{A}_\mathcal{S}$.
- The probability that m DUE pairs locate in any considered area with a radius R_a is given by

$$\mathbb{P}(M=m)=\frac{(\lambda\pi R_a^2)^m}{m!}\exp(-\lambda\pi R_a^2),\quad k\geq 0 \tag{3.1}$$

- The m DUE pairs in the cell are i.i.d. uniformly distributed in the area. The distance from a D2D pair to the BS follows the following distribution:

$$f_{M=m}(r)=\begin{cases}\frac{2r}{{R_a}^2} & 0\leq r\leq R_a,\\ 0 & \text{otherwise.}\end{cases} \tag{3.2}$$

3.1.3 Distance Distribution of Two DUEs Forming a D2D Pair

The probability density function (PDF) of the distance between the two DUEs forming a pair can be expressed as:

$$f_{\Delta_R}(r)=\begin{cases}\frac{2r({R_d}^2+r^2)}{{R_d}^4}-\frac{4r^2}{\pi {R_d}^3}G(\arccos(\frac{r}{2R_d})), & 0<r<R_d,\\ \frac{4\pi-3\sqrt{3}}{3\pi R_d}, & r=R_d,\\ -\frac{4r^2}{\pi {R_d}^3}G(\arccos(\frac{r}{2R_d})), & R_d<r<2R_d,\\ 0, & \text{otherwise.}\end{cases} \tag{3.3}$$

Proof. See Appendix A. □

3.1.4 Channel Model and Interference Distribution

The wireless channel considered in this paper can be characterized as:

$$P_r=\frac{P_t*h^2}{r^\alpha}, \tag{3.4}$$

where P_t is the transmission power, P_r is the power at receiver end. h denotes the channel coefficient accounting for multipath fading and shadowing. α represents the pass loss

exponent, which usually ranges from 1.6 (indoor environment) to 5 (dense urban area) depending on the spacial environment between the receiver and the transmitter [43]. We assume Rayleigh fading channel model, i.e., $h^2 \sim \exp(\mu)$. The D2D communication can use either the cellular downlink or uplink resources [44]. In the downlink case, I_c^{DL} represents interference to a CUE which comes from all co-channel D2D transmission pairs while I_d^{DL} denotes the interference to a D2D link which contains the one from BS and other co-channel D2D links. Similar for downlink case, I_c^{UL} denotes interference to a CUE from all co-channel D2D links and I_d^{UL} is the interference to a D2D transmission which comes from co-channel CUE uplink transmission as well as other co-channel D2D transmission pairs. We have the following interference model for cellular link and D2D link.

When the D2D communication uses the cellular downlink resources:

$$I_{c_i}^{DL} = \sum_{j \in \Phi_{d,i}} P_d (h_{ji}^{cd})^2 (r_{ji}^{cd})^{-\alpha_d}, \tag{3.5}$$

$$I_{d_j}^{DL} = \sum_{\substack{k \in \Phi_{d.i} \\ k \neq j}} P_d (h_{jk}^{dd})^2 (r_{jk}^{dd})^{-\alpha_d} + P_b (h_j^{bd})^2 (r_j^{bd})^{-\alpha_c}. \tag{3.6}$$

When D2D communication uses the cellular uplink resources:

$$I_{c_i}^{UL} = \sum_{j \in \Phi_{d,i}} P_d (h_j^{bd})^2 (r_j^{bd})^{-\alpha_c}, \tag{3.7}$$

$$I_{d_j}^{UL} = \sum_{\substack{k \in \Phi_{d,i} \\ k \neq j}} P_d (h_{jk}^{dd})^2 (r_{jk}^{dd})^{-\alpha_d} + P_{c,i} (h_{ij}^{cd})^2 (r_{ij}^{cd})^{-\alpha_d}. \tag{3.8}$$

r_{ji}^{cd} denotes the distance from the ith CUE to the jth D2D pair. r_{jk}^{dd} denotes the distance from the jth D2D pair to kth D2D pair. r_j^{bd} denotes the distance from the BS to the jth D2D pair. P_b, P_d, $P_{c,i}$ represent the transmission powers of BS, DUE and CUE i, respectively. In this work, we assume all the channels are reciprocal, i.e., the channel gain from the ith user to jth user is the same to that from jth user to ith user.

Based on properties of the PPP model and the assumption of i.i.d. Rayleigh fading channel on all the wireless links, we can calculate the Laplace Transform of interference variable by the definitions in [41] and [45].

For a CUE, the interference only comes from the co-channel D2D pairs, on both downlink and uplink.

$$\begin{aligned}\mathcal{L}_{I_r}(s) &= \mathbb{E}_{I_r}[\exp(-sI_r)] \\ &= \exp(-\lambda \int_{\mathcal{S}} \{1 - \mathbb{E}_{h^2}[\exp(-sP_d h^2 r^{-\alpha})]\} dr).\end{aligned} \tag{3.9}$$

For a D2D communication, the interference consists of two parts. One part comes from all co-channel D2D pairs and another part comes from the co-channel cellular communications. These two parts are independent to each other and denoted as as I_{r1} and I_{r2} respectively. The joint Laplace transform of I_{r1} and I_{r2} is given by:

$$\begin{aligned}\mathcal{L}_{I_r}(s_1, s_2) &= \mathbb{E}_{I_r}[\exp(-s_1 I_{r1} - s_2 I_{r2})] \\ &= \exp(-\lambda \int_{\mathcal{S}} \{1 - \mathbb{E}_{h_1^2}[\exp(-s_1 P_1 h_1^2 r_1^{-\alpha_1})]\} dr_1) \\ &\quad \times \mathbb{E}_{h_2^2}[\exp(-s_2 P_2 h_2^2 r_2^{-\alpha_2})].\end{aligned} \tag{3.10}$$

3.2 Coverage Analysis when DUEs Using Downlink Cellular Channels

This section analyzes the CUE and DUE SINR characteristics when D2D communications share the cellular downlink resources. For a CUE which is r_i^{bc} from the BS, the downlink SINR can be expressed as:

$$\text{SINR}_c^{DL}|_{r=r_i^{bc}} = \frac{P_b (h_i^{bc})^2 (r_i^{bc})^{-\alpha}}{\sum\limits_{j \in \Phi_{d,i}} P_d (h_{ji}^{cd})^2 (r_{ji}^{cd})^{-\alpha_d} + N_0}. \tag{3.11}$$

r_{ij}^{cd} is the distance from j_{th} D2D pair center to the considered CUE. r_j^{bd} denotes the distance

between j_{th} D2D pair to the BS. The distance between the D2D transmitter and receiver in j_{th} D2D pair is denoted as r_{jj}^{dd}. The SINR for the D2D communication can be expressed as: $\text{SINR}_d^{DL}|_{r=r_j^{bd}}$

$$= \frac{P_d(h_{jj}^{dd})^2 {r_{jj}^{dd}}^{-\alpha_d}}{\sum\limits_{\substack{k\neq j \\ k\in\Phi_{d,i}}} P_d(h_{jk}^{dd})^2(r_{jk}^{dd})^{-\alpha_d} + P_b(h_j^{bd})^2(r_j^{bd})^{-\alpha} + N_0}. \tag{3.12}$$

The channel coefficient of the link in the considered j_{th} D2D pair is h_{jj}^{dd} and channel coefficient of the link between j_{th} and k_{th} D2D pairs is h_{jk}^{dd}.

3.2.1 CUE Downlink Coverage

We define the UE coverage probability as the probability that the UE's SINR is above a certain threshold Γ. So when we talk about coverage in this work, it is always associated with a threshold value. The coverage probability of CUE i located r_i^{bc} away from BS is given by

$$\begin{aligned}
\mathbb{P}(\text{SINR}_c^{DL} > \Gamma) &= \mathbb{P}(\frac{P_b(h_i^{bc})^2(r_i^{bc})^{-\alpha_c}}{I_r + N_0} > \Gamma) \\
&= \mathbb{E}_{I_r}[\mathbb{P}((h_i^{bc})^2 > \frac{\Gamma(r_i^{bc})^{\alpha_c}}{P_b}(I_r + N_0)|I_r)] \\
&= \mathbb{E}_{I_r}[\exp(-\frac{\mu\Gamma}{P_b}{r_i}^{\alpha_c}(I_r + N_0))] \\
&= \exp(-\frac{\mu\Gamma {r_i}^{\alpha_c} N_0}{P_b})\mathcal{L}_{I_r}(\frac{\mu\Gamma {r_i}^{\alpha_c}}{P_b}),
\end{aligned} \tag{3.13}$$

where $\mathcal{L}_{I_r}(s)$ is the Laplace transform of random variable I_r evaluated at s conditioned on the distance from considered UE to the BS.

From property of PPP [41] and $I_r = \sum_{j \in \Phi_d} P_d (h_{ji}^{cd})^2 (r_{ji}^{cd})^{-\alpha_d}$, we have $\mathcal{L}_{I_r}(s)$

$$
\begin{aligned}
&= \mathbb{E}_{I_r}[\exp(-sI_r)] \\
&= \mathbb{E}_{\Phi_{d,i}}[\prod_{j \in \Phi_{d.i}} \mathbb{E}_{(h_{ji}^{cd})^2}[\exp(-sP_d h^2 {r_{ji}}^{-\alpha_d})]] \\
&= \exp(-\lambda \int_{\mathcal{S}} \{1 - \mathbb{E}_{(h_{ji}^{cd})^2}[\exp(-sP_d h^2 r^{-\alpha_d})]\} dr) \\
&= \exp(-\lambda \{ \int_0^{R - r_i^{bc}} f(s,r) 2\pi r dr + \\
&\quad \int_{R - r_i^{bc}}^{\sqrt{R^2 - (r_i^{bc})^2}} f(s,r)[2\pi - 2\arctan(\frac{y_0}{x_0 - r_i^{bc}})] r dr + \\
&\quad \int_{\sqrt{R^2 - (r_i^{bc})^2}}^{R + r_i^{bc}} f(s,r) 2\arctan(\frac{y_0}{r_i^{bc} - x_0}) r dr \}) \\
&= \exp(-\lambda \{ \int_0^{\sqrt{R^2 - (r_i^{bc})^2}} f(s,r) 2\pi r dr + \\
&\quad \int_{R - r_i^{bc}}^{R + r_i^{bc}} f(s,r) 2\arctan(\frac{y_0}{r_i^{bc} - x_0}) r dr \}) \\
&= \exp\{-\lambda [\pi (R^2 - (r_i^{bc})^2) \cdot \\
&\quad {}_2F_1(1; \frac{2}{\alpha_d}; \frac{2 + \alpha_d}{\alpha_d}; -(R^2 - (r_i^{bc})^2)^{\frac{\alpha_d}{2}} \frac{\mu_d}{sP_d}) + \\
&\quad \int_0^{\pi} \frac{-2Rr_i^{bc} \sin(x) \arctan[\frac{\sin(x)}{\cos(x) - \frac{r_i^{bc}}{R}}]}{1 + \frac{\mu_d}{sP_d}[R^2 + (r_i^{bc})^2 - 2Rr_i^{bc}\cos(x)]^{\frac{\alpha_d}{2}}} dx]\},
\end{aligned}
\tag{3.14}
$$

where $f(s,r) = 1 - \mathbb{E}_{(h_{ji}^{cd})^2}[\exp(-sP_d h^2 r^{-\alpha_d})] = \frac{sP_d r^{-\alpha_d}}{sP_d r^{-\alpha_d} + \mu_d}$, and ${}_2F_1(a;b;c;z)$ is hypergeometric function from [46]. For the last step in the above expression, we follow the same approach used in proof of Appendix A, and x_0, y_0 also have the same meaning as defined in Appendix A. Although there is no closed-form expression derived, numerical evaluation can be easily applied to calculate the integration.

3.2.2 DUE Downlink Coverage

Assume the distance between the two DUEs forming a pair is r_d and the distance between the D2D pair j and the BS is r_j^{bd}. The coverage probability for D2D pair j is given by

$$
\begin{aligned}
&\mathbb{P}(\mathrm{SINR}_d^{DL} > \Gamma) \\
&= \mathbb{P}(\frac{P_d (h_{ii}^{dd})^2 {r_d}^{-\alpha_d}}{I_r + N_0} > \Gamma) \\
&= \int_0^{2R_d} \mathbb{P}((h_{ii}^{dd})^2 > \frac{\Gamma {r_d}^{\alpha_d}}{P_d}(I_r + N_0) | I_r, (h_{ii}^{dd})^2, r_d) f_{\Delta_R}(r_d) dr_d \\
&= \int_0^{2R_d} \exp(-\frac{\mu_d \Gamma N_0 {r_d}^{\alpha_d}}{P_d}) \mathcal{L}_{I_r|r_d}(s_1, s_2) f_{\Delta_R}(r_d) dr_d,
\end{aligned}
\tag{3.15}
$$

where $s_1 = s_2 = \frac{\mu_d \Gamma {r_d}^{\alpha_d}}{P_d}$ and $\mathcal{L}_{I_r|r_d}(s_1, s_2)$ is equal to:

$$
\begin{aligned}
&= \exp(-\lambda \int_{\mathcal{S}} \{1 - \mathbb{E}_{(h_{jk}^{dd})^2}[\exp(-s_1 P_d h^2 r^{-\alpha_d})]\} dr) \times \\
&\quad \mathbb{E}_{(h_j^{bd})^2}[\exp(-s_2 P_b (h_j^{bd})^2 (r_j^{bd})^{-\alpha_c})] \\
&= \exp(-\lambda \{ \int_0^{\sqrt{R^2 - (r_j^{bd})^2}} f(s_1, r) 2\pi r dr + \\
&\quad \int_{R - r_j^{bd}}^{R + r_j^{bd}} f(s_1, r) 2 \arctan(\frac{y_0}{r_j^{bd} - x_0}) r dr \}) \times \frac{\mu}{\mu + s_2 P_b (r_j^{bd})^{-\alpha_c}}.
\end{aligned}
\tag{3.16}
$$

$f(s_1, r)$ and the integral part have the same form as in equation (3.14).

3.3 Coverage Analysis when DUEs Using Uplink Cellular Channels

This section gives the CUE and DUE coverage analysis when the D2D communications share the cellular uplink resources.

CUE j has a distance r_i^{bc} from the BS. Its uplink SINR is expressed as

$$
\mathrm{SINR}_c^{UL}|_{r = r_i^{bc}} = \frac{P_c (h_i^{bc})^2 (r_i^{bc})^{-\alpha_c}}{\sum\limits_{j \in \Phi_{d,i}} P_d (h_j^{bd})^2 (r_j^{bd})^{-\alpha_c} + N_0}.
\tag{3.17}
$$

Assume the distance between the two DUEs forming D2D pair j is r_{jj}^{dd}. The distance

between the D2D pair j and its co-channel CUE i is r_{ij}^{dc}. D2D pair j is located at $r = r_j^{bd}$ from the BS. The SINR distribution for D2D pair can be expressed as

$$\mathrm{SINR}_d^{UL}|_{r=r_j^{bd}} = \frac{P_d (h_{jj}^{dd})^2 (r_j^{bd})^{-\alpha_d}}{\sum\limits_{\substack{k\neq j \\ k\in\Phi_{d,i}}} P_d (h_{kj}^{dd})^2 r_{kj}^{dd\,-\alpha_d} + P_c (h_{ij}^{dc})^2 (r_{ij}^{dc})^{-\alpha_d} + N_0}. \tag{3.18}$$

3.3.1 CUE Uplink Coverage

The CUE uplink coverage probability can be expressed as

$$\begin{aligned} &\mathbb{P}(\mathrm{SINR}_c^{UL} > \Gamma) \\ &= \exp[-\frac{\mu\Gamma (r_i^{bc})^{\alpha_c} N_0}{P_c} - \lambda\pi R^2 \cdot \\ &\quad {}_2F_1(1; \frac{2}{\alpha_c}; \frac{2+\alpha_c}{\alpha_c}; -R^{\alpha_c} \frac{P_c}{\Gamma (r_i^{bc})^{\alpha_c} P_d})]. \end{aligned} \tag{3.19}$$

Proof. $\mathbb{P}(\mathrm{SINR}_c^{UL} > \Gamma)$

$$\begin{aligned} &= \mathbb{P}(\frac{P_c (h_i^{bc})^2 (r_i^{bc})^{-\alpha_c}}{I_r + N_0} > \Gamma) \\ &= \mathbb{E}_{I_r}[\mathbb{P}((h_i^{bc})^2 > \frac{\Gamma (r_i^{bc})^{\alpha_c}}{P_c}(I_r + N_0)|I_r)] \\ &= \mathbb{E}_{I_r}[\exp(-\frac{\mu\Gamma (r_i^{bc})^{\alpha_c}(I_r + N_0)}{P_c})] \\ &= \exp(-\frac{\mu\Gamma (r_i^{bc})^{\alpha_c} N_0}{P_c}) \mathcal{L}_{I_r}(\frac{\mu\Gamma (r_i^{bc})^{\alpha_c}}{P_c}). \end{aligned} \tag{3.20}$$

$\mathcal{L}_{I_r}(s)$ is the Laplace transform of random variable I_r evaluated at s conditioned on the distance from considered cellular user to BS. Similarly,

$$\begin{aligned}
\mathcal{L}_{I_r}(s) &= \mathbb{E}_{\Phi_{d,i},(h_j^{bd})^2}[\exp(-s\sum_{j\in\Phi_{d,i}} P_d(h_j^{bd})^2(r_j^{bd})^{-\alpha_c}) \\
&= \exp(-\lambda\int_{\mathcal{S}}\{1-\mathbb{E}_{h_j}[\exp(-sP_d(h_j^{bd})^2 r^{-\alpha_c})]\}dr) \\
&= \exp(-\lambda\int_0^R \frac{sP_d r^{-\alpha_c}}{sP_d r^{-\alpha_c}+\mu}2\pi r dr) \\
&= \exp[-\lambda\pi R^2\cdot {}_2F_1(1;\frac{2}{\alpha_c};\frac{2+\alpha_c}{\alpha_c};-R^{\alpha_c}\frac{\mu}{sP_d}]. \qquad (3.21)
\end{aligned}$$

□

3.3.2 DUE Uplink Coverage

D2D pair j is located at $r = r_j^{bd}$ from the BS. The distance between the two DUEs forming the pair is r_{jj}^{dd}. The DUE coverage probability can be expressed as $\mathbb{P}(\text{SINR}_d^{UL} > \Gamma)|_{r=r_j^{bd}}$

$$\begin{aligned}
&= \mathbb{P}(\frac{P_d(h_{jj}^{dd})^2(r_{jj}^{dd})^{-\alpha_d}}{I_r+N_0} > \Gamma) \\
&= \int_0^{2R_d} \mathbb{P}((h_{jj}^{dd})^2 > \frac{\Gamma {r_d}^{\alpha_d}}{P_d}(I_r+N_0)|I_r, r_d) f_{\Delta_R}(r_d)dr_d \\
&= \int_0^{2R_d} \exp(-\frac{\mu_d\Gamma N_0 {r_d}^{\alpha_d}}{P_d})\mathcal{L}_{I_r|r_d}(s_1,s_2) f_{\Delta_R}(r_d)dr_d, \qquad (3.22)
\end{aligned}$$

Where $s_1 = s_2 = \frac{\mu_d \Gamma {r_d}^{\alpha_d}}{P_d}$, and $\mathcal{L}_{I_r|r_d}(s_1, s_2)$ is

$$\begin{aligned} &= \exp(-\lambda \int_{\mathcal{S}} \{1 - \mathbb{E}_{(h_{kj}^{dd})^2}[\exp(-s_1 P_d h^2 r^{-\alpha_d})]\} dr) \times \\ &\quad \mathbb{E}_{(h_{ij}^{dc})^2, r_{ij}^{dc}}[\exp(-s_2 P_c h^2 r^{-\alpha_d})] \\ &\approx \exp(-\lambda \{ \int_0^{\sqrt{R^2 - (r_j^{bd})^2}} f(s_1, r) 2\pi r dr + \\ &\quad \int_{R - r_j^{bd}}^{R + r_j^{bd}} f(s_1, r) 2 \arctan(\frac{y_0}{r_j^{bd} - x_0}) r dr \}) \times \frac{\mu_d}{\mu_d + s_2 P_c \bar{r}_{ij}^{-\alpha_d}}. \end{aligned} \tag{3.23}$$

$f(s_1, r)$ has the same form as in equation (3.14), and $\mathbb{E}_{{h^2}_{ji}, r_{ij}}[\exp(-s_2 P_c h^2 {r_{ij}}^{-\alpha_d})] \approx \frac{\mu_d}{\mu_d + s_2 P_c \bar{r}_{ij}^{-\alpha_d}}$. $\bar{r}_{li}$ denotes average distance between cellular user and the considered D2D user. Numerical approach can be applied to calculate the integration.

3.4 Numerical Results

In this section, we evaluate the performance of the given D2D network. We consider a normalized circular cell with a unity radius and set the BS transmit power P_b to ensure that the cell edge user Signal-to-Noise Ratio (SNR) is at least 1dB. All other transmit powers and distances are given with respect to these two values in the simulation. The path loss model used in the work is based on [47] $PL(d) = PL(d_0) + 10\alpha \log(d)(dB)$, where $PL(d)$ is the receiver's path loss at a distance d away from the transmitter. $PL(d_0)$ represents the reference path loss at distance d_0 and α is the path loss exponent. The path loss exponent for all the channels is assumed to be 4.

3.4.1 System Validation

In our system model, all the D2D users appear as communication pairs and their 2-D spacial locations are distributed according to homogeneous PPP with intensity λ_d. DUEs and CUEs can share radio resources. First, we verify the developed system model and derived formulas in Section III and Section IV. Simulation based on Monte Carlo approach is applied to estimate the statistical SINR distribution of CUEs and DUEs for both downlink

and uplink, which are compared to the analytical results given by the corresponding formula expression. It can be observed that the simulation curves matches very well with the analytical results.

3.4.2 Downlink and Uplink Coverage

Figure 3.2 presents the coverage probability for the CUEs when D2D communication uses cellular downlink resources. The curves with a legend "S" represent the results from simulation while the curves with a legend "A" are the results from the analysis. For a downlink CUE, there is no interference from other CUEs. All the interferences are from the co-channel D2D users. We evaluate the impact of parameters, such as CUE distance to BS, D2D user density λ_d, DUE transmission power P_d, on the performance of CUE downlink coverage. As expected, a CUE closer to BS, i.e., $r_i = 0.3R$ has a better coverage than the one further away, i.e., $r_i = 0.5R$, from the BS. Also a higher DUE transmission power and a higher DUE density will cause a larger interference to the CUEs, resulting in a lower CUE coverage probability and thus a higher outage for the CUEs. It is worth noting that simulation results in the plot match very well with our analytical formula from (3.13), which confirms the accuracy of our theoretical study.

Figure 3.3 shows the coverage probability of DUEs in the downlink resource sharing scenario. We again use different system parameter settings to study the network performance and verify our analysis. The curves with the same color share the same parameter settings for DUE density λ_d, DUE transmission power P_d, and D2D pair circle radius R_d. The curve marked with circle represents the scenario that the DUE under observation is $0.5R$ away from the BS, while the curve marked with square represents the scenario that the DUE under observation is $0.3R$ away from the BS. The closer the DUE is to the BS, the stronger the interference it will receive from the BS downlink transmission and thus the smaller its coverage probability will be, i.e., the coverage becomes worse. It can be also observed from the plots that a higher DUE density, a lower DUE transmission power, and a larger circle radius for D2D pairs will degrade the DUE coverage performance. There is a tradeoff between the selection of maximum D2D transmission power and D2D commu-

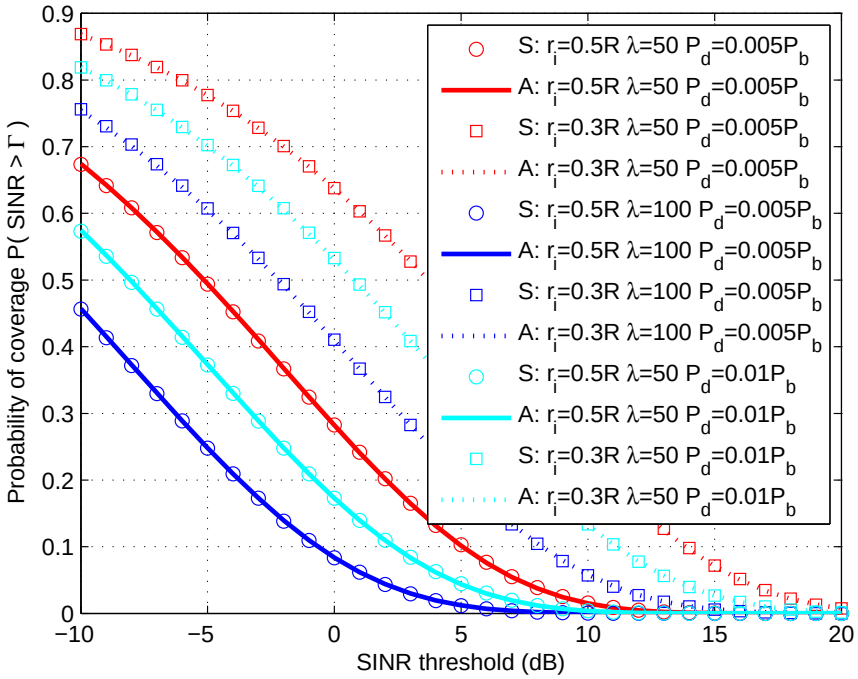

Fig. 3.2: Probability of coverage for downlink cellular user

nication distance. The optimal selection becomes even more complicated when the system needs to manage the DUE interference to CUEs on the downlink.

On the uplink, the CUEs will need power control so that the received SNR at BS for all CUEs can reach the target value 5dB. Since the CUEs are considered as the primary users, we consider a simple approach to reducing the interference to the CUEs on the uplink transmission by restricting the operation area of DUEs. When operating on the uplink resource sharing mode, the DUEs are not allowed to use the cellular channels when its distance to the BS is less than R_g. Thus We simulate two different scenarios, i.e., with and without guard area. In these two scenarios, the total number of D2D pairs in the cell are set to be the same. Without a guard area, the DUEs can locate anywhere in the cell while with a guard area, the DUEs can only locate outside the circular area around the BS with a radius of R_g. Although the area density of DUEs with a guard area case is higher than that without a guard area, the interference to the CUEs is still reduced significantly when guard area is applied, which can be seen from the CUEs SINR distributions in Figure 3.4.

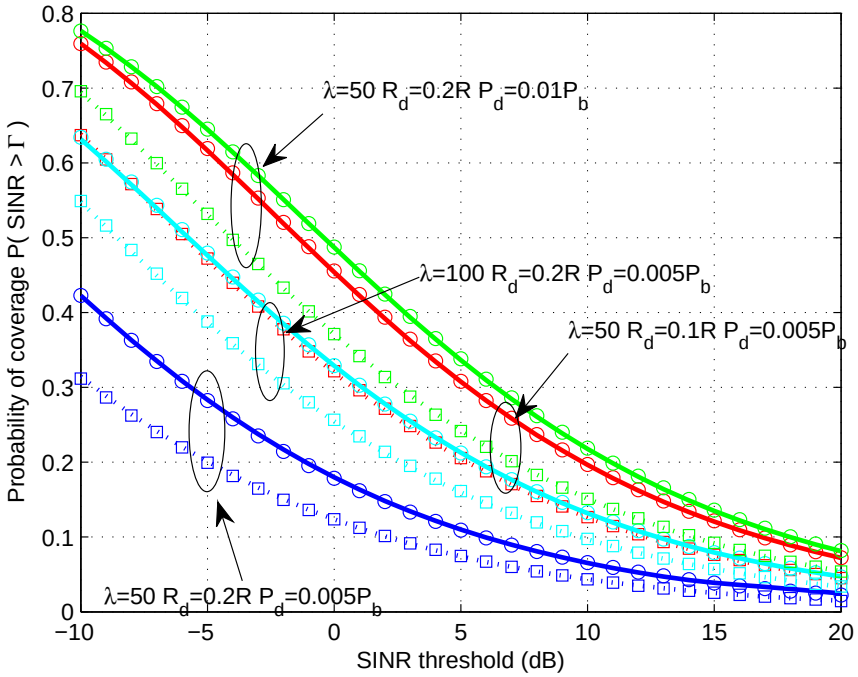

Fig. 3.3: Probability of coverage for downlink D2D user

From the same figure we can see that a lower co-channel DUE density and a lower DUE transmission power can always help improve the CUE SINR performance. Thus in reality, when serving D2D users, the system needs to be careful with what power level to use for the DUEs and which cellular channel to use for the DUEs. This can be done though intelligent interference coordination and power control between CUEs and DUEs.

Figure 3.5 illustrates the coverage probability for DUEs in the uplink scenario. We study the impact of system parameters such as D2D user density, D2D circle radius and DUE transmission power on the performance of DUEs in uplink resources sharing mode. The analytical curve is plotted based on the approximation formula in (3.22), (3.23) by using an average distance $\bar{R}_{cd}$ between CUEs and DUEs, which still matches the simulation results very well. It is observed that when D2D communication uses cellular uplink resources, the transmit power of DUEs has a very limited influence on the DUE SINR performance, although its influence on the CUEs is significant (as can be seen from 3.4). it can be explained in the following. When increasing the DUE transmit power, the increase of

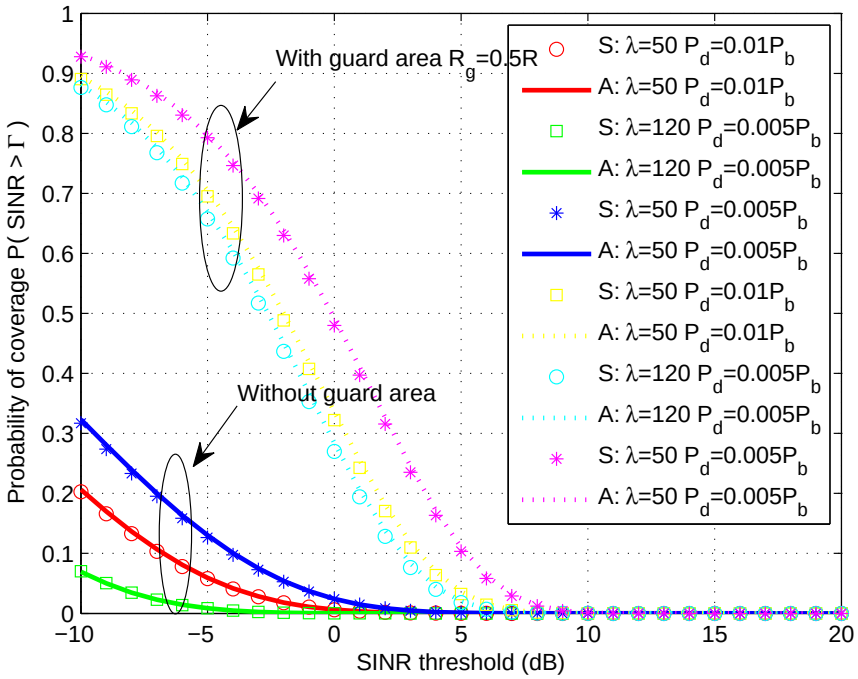

Fig. 3.4: Probability of coverage for uplink cellular user

interference from other co-channel DUEs offsets the improvement of received transmit signal power, leading to limited SINR improvement for the DUEs. For this reason, it is better to limit the transmit power of DUEs to reduce the interference to CUEs in the uplink resource sharing operation mode. Similarly, other parameters such as DUE density, distance constraint impact the DUE performance in the similar way to the downlink case.

3.5 Summary

In the work, an analytical evaluation of downlink and uplink SINR performance for the D2D communication underlaying a cellular network is derived. We model the 2-D spacial location of D2D pairs with a homogeneous PPP. The D2D users are pre-paired with a maximum communication distance constraint to guarantee the link budget and QoS of D2D user. The CUEs and DUEs can share the same channels, either on the uplink or on the downlink. The coverage probability for both DUEs and CUEs is derived based on user distribution and channel statistics without any knowledge of instantaneous channel information. The

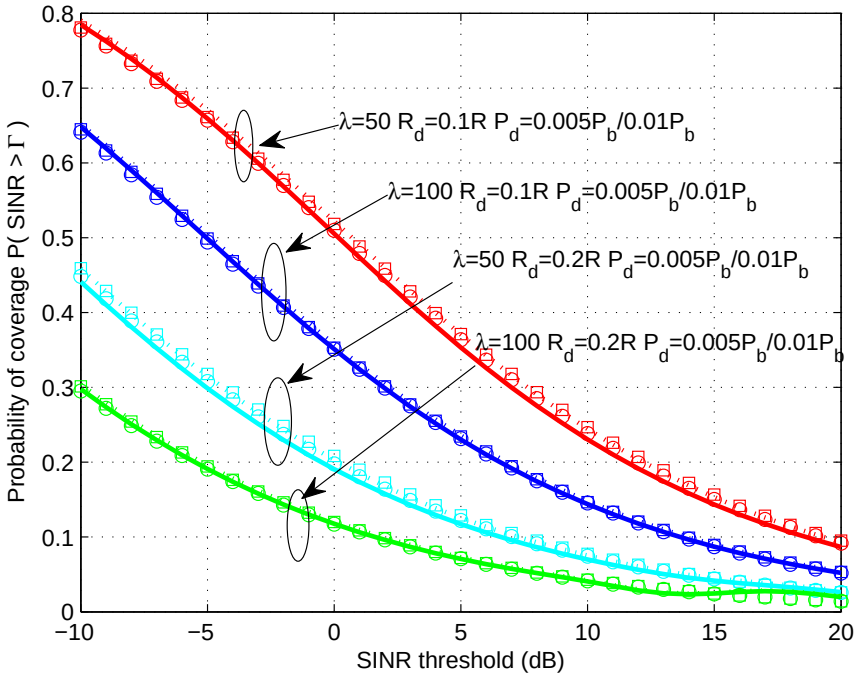

Fig. 3.5: Probability of coverage for uplink D2D user

analytical results match very well with the simulation work. We investigate the impact of system parameters such as DUE density, BS transmission power, DUE transmission power on the DUE and CUE coverage performance. When considering CUEs as the prioritized users in such a D2D cellular network, intelligent interference coordination through resource allocation and power control is very critical in achieving satisfactory coverage performance for both primary CUEs and secondary DUEs. The proposed analytical tools provide convenient and accurate studies on the co-channel deployed D2D communications in a cellular network.

Chapter 4

Distributed Resource Allocation for D2D Communication - A Stackelberg Game Model

In this chapter, we discuss a distributed resource allocation and power control scheme based on Stackelberg game framework to improve network efficiency in D2D communication networks. The system aims to maximize the number of supportable underlay D2D users while guaranteeing Quality of Service (QoS) of the prioritized cellular users. D2D users are supported in an underlay mode by sharing radio resources with the cellular downlink communications. Network throughput and spectral efficiency can be further improved by exploiting multiuser diversity in a D2D underlaying cellular communication, in which one cellular user's resource can be shared by multiple D2D pairs and one D2D pair can reuse resources from different cellular users. We formulate the joint optimization on D2D power control and resource allocation with a distributed Stackelberg game theoretical model and decompose it into two steps to approach the game equilibrium. In the work, the Stackelberg game framework [48,49] is used to model the interaction between Base Station (BS), cellular users and D2D pairs for power allocation on each resource block. Given the cost of resource block, BS/cellular user can share the resource based on the D2D communication channel condition and D2D interference to other users and the game equilibrium is reached at the maximum utility.

The rest of this chapter is organized as following. First, we discuss the system model of D2D communication underlaying cellular network. Then we propose the resource allocation and power control scheme based on the Stackelberg game model. Next, we present simulation results and numerical analysis. Finally, conclusions are drawn in the last section.

4.1 System Model

In this work, we consider underlay D2D communications that share cellular downlink resources with the prioritized cellular users (CUEs). D2D users (DUEs) come in as pairs. So each DUE consists of two UEs that form a pair. The system has N CUEs and M DUEs. All CUEs directly communicate with the BS and each CUE is assigned one downlink cellular resource. Thus there are in total N downlink cellular resources. We allow each CUE resource to be shared among multiple DUEs in order to maximize the spectrum reuse gain and each DUE can be assigned multiple resources from different CUEs. CUEs have a higher priority than DUEs when they are served as co-channel users so that CUE QoS will not be compromised by serving underlay DUEs. In order to achieve that, the resource sharing between CUEs and DUEs should be well coordinated. As part of that coordination, the maximum transmission power of DUEs is P^D and transmission distance from the transmitter to the receiver of a DUE is constrained to R_d to maintain the link quality of D2D communication as well as to throttle the DUE interference to CUEs.

Figure 4.1 depicts the system model for D2D communications underlaying the cellular network. We denote f_i as the resource allocated to CUE i. As an illustrative example of resource allocation in Figure 4.1, CUE1, CUE2 ,CUE3 are allocated orthogonal resources f_1, f_2, and f_3 , respectively. DUE1, DUE2, DUE3, DUE4 are competing to share these resources in an underlay mode. There is no intra-cell co-channel interference among CUEs. One possible D2D underlay resource sharing is as follows. DUE1 shares f_3, DUE2 shares f_1, DUE3 shares f_1 and f_2, DUE4 shares f_2 and f_3. The following notations are used:

- P_i^B: Transmission power allocated to CUE i
- $P_j^{d_i}$: Power allocated to DUE j for transmitting on resource i owned by CUE i
- g_i^c: channel gain from BS to CUE i
- $g_j^{d_i}$: channel gain from BS to DUE j
- $g_{jj}^{d_i}$: channel gain between the two UEs in DUE j on resource i
- $g_{jj'}^{d_i}$: channel gain from the DUE j to DUE j' D2D pair on resource i

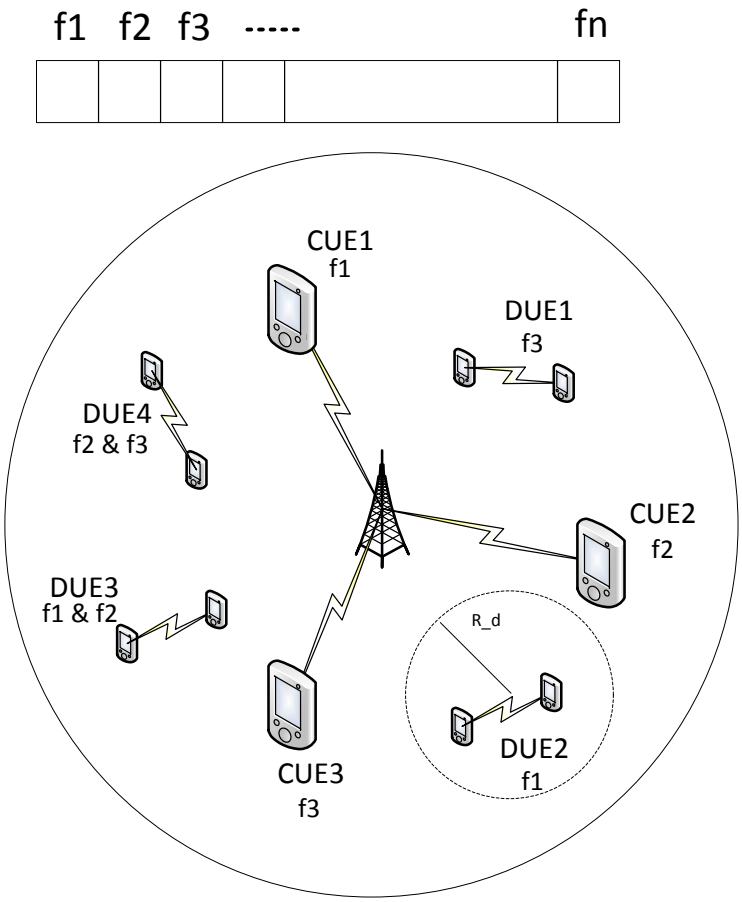

Fig. 4.1: Illustration of Device-to-Device underlaying cellular network

- g_{ji}^{dc}: channel gain from the DUE j to CUE i
- N_0: AWGN noise power density

We consider a Rayleigh fast fading, which follows an exponential distribution denoted as $h \sim \exp(\mu)$. The pathloss model is given as $PL(d) = PL(d_0)d^{-\alpha}$. The channel gain is expressed as $g = PL(d_0)h \cdot d^{-\alpha}$. Based on the proposed underlay resource sharing mode, the SINR for CUE i and DUE j can be expressed as following.

$$\text{SINR}_i^c = \frac{P_i^B g_i^c}{\sum_j x_{ji} P_{ji}^{dd} g_{ji}^{dc} + N_0}, \tag{4.1}$$

$$\mathrm{SINR}_{ji}^{d} = \frac{P_j^{d_i} g_{jj}^{d_i}}{\sum_{j' \neq j} x_{j'i} P_{j'}^{d_i} g_{j'j}^{d_i} + P_i^B g_j^{d_i} + N_0}, \tag{4.2}$$

where x_{ji} represents the allocation status of resource f_i to DUE j. If $x_{ji} = 1$, f_i is allocated to DUE j. $x_{ji} = 0$ otherwise. There is no co-channel interference among CUEs within the cell. The interference to a CUE comes from all the underlay co-channel DUEs while the interference to a DUE includes the co-channel CUE transmission and other co-channel DUEs, as illustrated in formula (4.1) and (4.2).

Our objective is to maximize the sum-rate of DUEs while guaranteeing QoS of prioritized CUEs. The optimization problem is formulated as:

$$\max \quad \sum_j \sum_i x_{ji} \log_2(1 + \mathrm{SINR}_{ji}^{d}) \tag{4.3}$$

s.t.

$$x_{ji} = \{0 \text{ or } 1\}, \tag{4.4}$$

$$\frac{P_i^B g_i^c}{\sum_j x_{ji} P_j^{d_i} g_{ji}^{dc} + N_0} \geq \gamma_i, \qquad i = 1 \cdots N; \tag{4.5}$$

$$\sum_i x_{ji} P_j^{d_i} \leq P^D, \qquad j = 1 \cdots M. \tag{4.6}$$

Constraint (4.5) denotes the minimum channel quality requirement for each CUE. Condition in (4.6) is the maximum DUE transmission power constraint. To achieve the maximum transmission rate, each DUE adapts its transmission power allocation among multiple shared resources based on the channel quality on each resource. The power allocation problem is closely tied to interference management, which makes the mixed integer nonlinear resource and power allocation optimization problem mathematically intractable.

4.2 Resource Allocation and Power Control for D2D User

Although jointly considering DUE resource allocation and power control is highly desirable from network spectral efficiency's perspective, the joint optimization is usually com-

putationally complex and mathematically intractable. Furthermore, the power and resource allocation occurs in a fast manner, e,g, every 1 ms, to exploit diversity gain. Therefore, we need to pursue a computationally efficient and mathematically tractable scheme. Hence, an algorithm with a lower complexity is preferred, even though it is suboptimal. We resolve the problem into a two-step approach, with each step separately addressing resource allocation and power allocation. In this section, we first discuss a grouping mechanism, in which all the DUEs that share the same resource of a CUE are deemed as a group. Furthermore, by using this grouping scheme, power allocation based on Stackelberg game model is developed to optimally distribute power among resources for each DUE. To make our assumptions more clear, each CUE takes one resource while each DUE can use multiple resources.

4.2.1 A Sub-optimal CUE-DUE Grouping Scheme

We discuss a sub-optimal CUE-DUE grouping scheme in this section. In observation of co-channel interference among DUEs sharing the same resource, we group together DUEs in the way that can boost network capacity most while limit the interference to each other. Full transmission power is assumed during grouping. To measure cellular network capacity improvement by sharing resources with DUEs, C_{ji}^{gain} denotes the throughput gain when DUE j shares the resource with CUE i.

$$\begin{aligned} C_{ji}^{gain} &= \log_2(1+\frac{P_B g_i^c}{P_j^{d_i} g_{ji}^{dc}+N_0}) \\ &+ \log_2(1+\frac{P_j^{d_i} g_{jj}^{d_i}}{P_B g_j^d+N_0}) - \log_2(1+\frac{P_B g_i^c}{N_0}). \end{aligned} \tag{4.7}$$

We also denote the group of DUEs sharing the same resource with CUE i as group Ω_i, $i = 1, ..., N$. The grouping algorithm aims to maximize the total throughput gain by efficiently reusing the cellular resources. A DUE is allowed to share a resource if and only if the throughout gain is positive. On the other hand, the grouping and resource sharing also need to consider the co-channel interference among different DUEs. Detailed CUE-DUE grouping algorithm is described in Table 4.1.

Table 4.1: Sub-optimal CUE-DUE grouping scheme

Initialize $\Omega_i^{'} = \emptyset$ $\Omega_i = \emptyset$ for $i = 1, 2, \cdots, N$, and $\mathcal{D}_j$ for $\{j = 1, 2, \cdots, M\}$,
Calculate C_{ji}^{gain} for $i = 1, 2, \cdots, N$, $j = 1, 2, \cdots, M$
Step 2: For $i = 1$ to N,
  a) sort out C_{ji}^{gain} in descending order,
  b) update $\Omega_i^{'} = \Omega_i^{'} \bigcup j$ with $C_{ji}^{gain} > 0$ for all $j = 1, 2, \cdots, M$

Step 3: For $i = 1$ to N,

For $j \in \Omega_i$, $j = \arg\max_{j \in \Omega_i^{'}} C_{ji}^{gain}$, $\Omega_i^{'} = \Omega_i^{'} \setminus j$,
evaluate $p1 = \frac{g_{jj}^{d_i}}{g_{kj}^{d_i}}$, $p2 = \frac{g_{kk}^{d_i}}{g_{jk}^{d_i}}$ for all $k \in \Omega_i$
if $P1 > \gamma$ and $P2 > \gamma$
$\Omega_i = \Omega_i \bigcup j$, $\mathcal{D}_j = \mathcal{D}_j \bigcup i$

We allow each resource to be shared among multiple DUEs as long as the sharing can improve the network capacity. DUEs that are too close to each other may not want to use the same resource due to the strong interference to each other even if each of them can boost network efficiency separately. In Step 3 of Table 4.1, $g_{jk}^{d_i}$ denotes the channel gain from DUE j to DUE k when sharing resource with CUE i. Threshold γ prevents those proximity DUEs from sharing resources. Based on the proposed CUE-DUE grouping algorithm, we can assume that dominant interference to DUEs comes from BS. As each DUE can use multiple resources, how D2D allocate power among different resources is discussed in the next section.

4.2.2 Stackelberg Game Model Based Power Allocation

We use Stackelberg game to group DUEs and allocate power distributively. The game is divided into two levels: CUEs and BS are modeled as buyers, aiming to achieve the maximum throughput gain by buying powers from the DUEs; DUEs are modeled as sellers and aim to gain most payment from selling their powers to contribute to the network throughput. On the buyer side, given a certain price offer, CUEs calculate the amount of power bought from each DUE. If a seller asks for an overly high price, CUEs would bid for less power or even give up their bids. On the seller side, each DUE sells its power

to the CUE who can provide a higher price to achieve the maximum utility. However, the competition among seller DUEs will make sure each DUE asks for a reasonable price based on its location and channel condition. From system's perspective, the price depends on the nexus between CUEs and DUEs, which influences power allocation. On the other hand, the power allocation for each DUE in turn determines the price offer given by sellers. Ultimately, there is a final price concurred by both CUEs and DUEs and this final price and power allocation are the equilibrium of the game.

Buyer Level Utility

The goal of the power control and resource allocation algorithm is to maximize the sum rate D2D throughput while guaranteeing CUE QoS. Therefore, the utility function of buyers is defined as:

$$U_{ji} = \beta \sum_{i}^{N} \sum_{j \in \Omega_i} \log_2(1 + \frac{P_j^{d_i} g_{jj}^{d_i}}{P_B g_j^d + + N_0}) - C, \tag{4.8}$$

where $C = \sum_i^N \sum_{j \in \Omega_i} r_{ji} P_j^{d_i}$ denotes the total cost paid by CUEs to buy power from DUEs. r_{ji} is the unit price of power sold by DUE j to contribute throughput on CUE i's resource and $P_j^{d_i}$ denotes how much power CUE will buy from DUE j given the price. It is worth noting that β in (4.8) is used to balance the sum rate capacity and total cost. A larger value of β puts more weight on capacity improvement, thus rendering more power bought by BS given the same price offer. At the buyer level, the goal of the game is to maximize the buy utility function under the constraint of D2D transmission power while guaranteeing QoS of prioritized CUEs. The buyer optimization problem is formulated as:

$$\max \quad U_b = \sum_{i}^{N} \sum_{j \in \Omega_i} [\beta \log_2(1 + \frac{P_j^{d_i} g_{jj}^{d_i}}{P_B g_j^d + N_0}) - r_{ji} P_j^{d_i}] \tag{4.9}$$

s.t.

$$\frac{P_B g_i^c}{\sum_{j \in \Omega_i} P_j^{d_i} g_{ji}^{dc} + N_0} \geq \Gamma_i, \quad \text{for} \;\; i = 1, \cdots, N; \tag{4.10}$$

$$\sum_{i=1}^{N} P_j^{d_i} \leq P_D, \quad \text{for} \;\; j = 1, \cdots, M \;\; j \in \Omega_i. \tag{4.11}$$

The first constraint represents SINR requirement for CUEs and the second constraint states the maximum transmit power for each DUE. For the given price $r_{ji}, i = 1, \cdots, N; j = 1, \cdots, M$ offered by DUEs, buyer level objective function U_b is a concave function of power allocation $P_j^{d_i}$.

Remark 4.2.1. *CUE i can only share its resource to DUE j when the price offered by DUE j is no greater than* $\frac{\beta g_{jj}^{d_i}}{\ln 2(P_B g_j^d + N_0)}$.

Let's consider the objective function U_b.

$$\frac{\partial U_b}{\partial P_j^{d_i}} = \frac{\alpha \log_2 e g_{jj}^{d_i}}{P_B g_j^d + N_0 + g_{jj}^{d_i} P_j^{d_i}} - r_{ji}. \tag{4.12}$$

The first derivative of U_b *is a strictly decreasing function of* $P_j^{d_i}$. *Without considering the two constraints (4.10) and (4.11), the maximum objective function value is achieved when* $P_{ji}^{dd} = \frac{\alpha}{\ln 2 r_{ji}} - \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}}$. *However, when* $r_{ji} \geq \frac{\alpha g_{jj}^{d_i}}{\ln 2(P_B g_j^d + N_0)}$, *we have* $\frac{\partial U_b}{\partial P_j^{d_i}} < 0$. *Thus the objective function* U_b *achieves its maximum value of zero at* $P_j^{d_i} = 0$. *Therefore, if DUE j provides a higher price than that, CUE i would gain no benefit by buying power from DUE j.*

The optimal solution for the buyer level game is not straightforward to solve. We can transform the original problem into a dual problem by introducing additional variables. The Lagrange function $L(p_j^{d_i}, \lambda_i, \mu_j)$ is defined as :

$L(p_j^{d_i}, \lambda_i, \mu_j)$

$$
\begin{aligned}
&= \sum_i \sum_{j\in\Omega_i} [r_{ji} p_j^{d_i} - \beta \log_2(1 + \frac{p_j^{d_i} g_{jj}^{d_i}}{P_B g_j^d + N_0})] \\
&+ \sum_i \lambda_i [\sum_{j\in\Omega_i} P_j^{d_i} g_{ji}^{dc} - \frac{P_B g_i^c}{\Gamma_i} + N_0] + \sum_j \mu_j [\sum_{\substack{i \\ j\in\Omega_i}} P_j^{d_i} - P_D] \\
&= \sum_i \sum_{j\in\Omega_i} [r_{ji} P_j^{d_i} - \beta \log_2(1 + \frac{P_j^{d_i} g_{jj}^{d_i}}{P_B g_j^d + N_0}) + \lambda_i P_j^{d_i} g_{ji}^{dc} \\
&+ \mu_j P_j^{d_i}] + \sum_i \lambda_i (N_0 - \frac{P_B g_i^c}{\Gamma_i}) - \sum_j \mu_j P_D, \qquad (4.13)
\end{aligned}
$$

where the λ_i and μ_j are the Lagrange multipliers associated with two inequality constraints (4.10) and (4.10). By introducing the Lagrange function, the original problem can be solved through solving the following dual optimization problem. $g(\lambda_i, \mu_j)$

$$
\begin{aligned}
&= \sum_i \sum_{j\in\Omega_i} \inf_{P_j^{d_i}} [P_j^{d_i} (r_{ji} + \lambda_i g_{ji}^{dc} + \mu_j) - \beta \log_2(1 + P_j^{d_i} \\
&\times \frac{g_{jj}^{d_i}}{P_B g_j^d + N_0})] + \sum_i \lambda_i (N_0 - \frac{P_B g_i^c}{\Gamma_i}) - \sum_j \mu_j P_D \\
&= - \sum_i \sum_{j\in\Omega_i} \sup_{P_j^{d_i}} [P_{ji}^d (-r_{ji} - \lambda_i g_{ji}^{dc} - \mu_j) + \beta \log_2(1 + P_j^{d_i} \\
&\times \frac{g_{jj}^{d_i}}{P_B g_j^d + N_0})] + \sum_i \lambda_i (N_0 - \frac{P_B g_i^c}{\Gamma_i}) - \sum_j \mu_j P_D \\
&= \sum_i \sum_{j\in\Omega_i} [\beta \log_2(r_{ji} + \lambda_i g_{ji}^{dc} + \mu_j) - \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}} (r_{ji} \\
&+ \lambda_i g_{ji}^{dc} + \mu_j)] + \sum_i \lambda_i (N_0 - \frac{P_B g_i^c}{\Gamma_i}) - \sum_j \mu_j P_D + D.
\end{aligned}
\qquad (4.14)
$$

$D = \sum_i \sum_{j\in\Omega_i} [\beta \log_2(\frac{e \ln 2 (P_B g_j^d + N_0)}{\beta g_{jj}^{d_i}})]$ is a constant. And $F^*(y)$ is the conjugate function of $F(P_j^{d_i}) = -\beta \log_2(1 + P_j^{d_i} \frac{g_{jj}^{d_i}}{P_B g_j^d + N_0})$. $F^*(y) = \sup_x (x * y - F(x))$ as defined in [50]. The conjugate function of $F^*(y) = \sup_{P_j^{d_i}} (P_j^{d_i} * y + \beta \log_2(1 + P_j^{d_i} \frac{g_{jj}^{d_i}}{P_B g_j^d + N_0}))$ can be evaluated as follows.

- When $y \geq 0$, $F^*(P_j^{d_i}) = +\infty$.
- When $y < 0$, first order derivation of $F^*(P_j^{d_i}, y) = P_j^{d_i} * y + \beta \log_2(1 + P_j^{d_i} \frac{g_{jj}^{d_i}}{P_B g_j^d + N_0})$ is evaluated as

$$\frac{\partial F^*(P_j^{d_i}, y)}{\partial P_j^{d_i}} = y + \frac{\beta \log_2 e g_{jj}^{d_i}}{P_B g_j^d + N_0 + g_{jj}^{d_i} P_j^{d_i}}. \tag{4.15}$$

By letting $\frac{\partial F^*(P_j^{d_i}, y)}{\partial P_j^{d_i}} = 0$, we get $P_j^{d_i} = -\frac{\beta \log_2 e}{y} - \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}}$. Replacing this $P_j^{d_i}$ into $F^*(y)$, we have

$$\begin{aligned} F^*(y) &= \beta \log_2 \frac{\beta \log_2 e g_{jj}^{d_i}}{-y(P_B g_j^d + N_0)} - \beta \log_2 e \\ &- \frac{y(P_B g_j^d + N_0)}{g_{jj}^{d_i}}, y < 0. \end{aligned} \tag{4.16}$$

Thus the dual optimization problem can be formulated as

$$\max \qquad g(\lambda_i, \mu_j) \tag{4.17}$$

s.t.

$$\lambda_i \geq 0, \quad i = 1 \cdots N; \tag{4.18}$$

$$\mu_j \geq 0, \quad j = 1 \cdots M. \tag{4.19}$$

By checking Slater's condition [50], the duality gap between the Lagrange dual problem and the original problem is zero and optimal solution to dual optimization problem is equal to the solution to the original problem. It should be noted that the above dual problem is a function of λ_i and μ_j and it is always concave according to the property of dual function. For the above convex optimization problem, the following gradient based search algorithm

can be applied to find the optimal solution.

$$\begin{aligned}\nabla_{\lambda_i} g(\lambda_i, \mu_j) &= \sum_{j\in\Omega_i} [\frac{\beta \log_2 e g_{ji}^{dc}}{r_{ji} + \lambda_i g_{ji}^{dc} + \mu_j} - \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}} g_{ji}^{dc}] \\ &+ N_0 - \frac{P_B g_i^c}{\Gamma_i}, \end{aligned} \tag{4.20}$$

$$\begin{aligned}\nabla_{\mu_j} g(\lambda_i, \mu_j) &= \sum_{\substack{i=1 \\ (j\in\Omega_i)}}^{N} [\frac{\beta \log_2 e}{r_{ji} + \lambda_i g_{ji}^{dc} + \mu_j} - \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}}]. \\ &- P_D \end{aligned} \tag{4.21}$$

Update for dual variable along with gradient direction is provided as:

$$\lambda_i(t+1) = \lambda_i(t) + \delta \nabla_{\lambda_i} g(\lambda_i, \mu_j), \quad i = 1 \cdots N. \tag{4.22}$$

$$\mu_j(t+1) = \mu_j(t) + \delta \nabla_{\mu_j} g(\lambda_i, \mu_j), \quad j = 1 \cdots M. \tag{4.23}$$

δ is the update step size.

KKT Conditions for Optimality

For the original constrained convex optimization problem, the KKT conditions are sufficient and necessary for both primary and dual optimality. Let $\widetilde{p_j^{d_i}}$, and $\widetilde{\lambda_i}$, $\widetilde{\mu_j}$ be the optimal solutions to the original and dual problems respectively.

$$\begin{aligned}\nabla_{p_j^{d_i}} L(\widetilde{p_j^{d_i}}, \widetilde{\lambda_i}, \widetilde{\mu_j}) &= r_{ji} - \frac{\beta g_{jj}^{d_i}}{\ln 2 (\widetilde{p_j^{d_i}} g_{jj}^{d_i} + P_B g_j^d + N_0)} \\ &+ \widetilde{\lambda_i} g_{ji}^{dc} + \widetilde{\mu_j} = 0, \end{aligned} \tag{4.24}$$

$$\widetilde{\lambda_i} [\sum_{j\in\Omega_i} (\widetilde{p_j^{d_i}} g_{ji}^{dc}) - \frac{P_B g_i^c}{\Gamma_i} + N_0] = 0, \tag{4.25}$$

$$\widetilde{\mu_j}[\sum_{\substack{i=1\\ j\in\Omega_i}}^{N}(\widetilde{p_j^{d_i}}) - P_D] = 0, \tag{4.26}$$

$$\sum_{j\in\Omega_i}\widetilde{p_j^{d_i}} g_{ji}^{dc} - \frac{P_B g_i^c}{\Gamma_i} + N_0 \leq 0, \tag{4.27}$$

$$\sum_{\substack{i\\ j\in\Omega_i}}\widetilde{p_j^{d_i}} - P_D \leq 0, \tag{4.28}$$

$$\widetilde{\lambda_i} \geq 0, \quad i = 1\cdots N, \tag{4.29}$$

$$\widetilde{\mu_j} \geq 0, \quad j = 1\cdots M. \tag{4.30}$$

By evaluating (4.24), the optimal power allocation $\widetilde{p_j^{d_i}}$ is

$$\widetilde{p_j^{d_i}} = \frac{\beta}{\ln 2(r_{ji} + \widetilde{\mu_j} + \widetilde{\lambda_i} g_{ji}^{dc})} - \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}}. \tag{4.31}$$

Till now, given the price offer r_{ji}, the optimal power allocation $\widetilde{p_j^{d_i}}$ can be indirectly gained by first evaluating dual optimality $\widetilde{\lambda_i}$ and $\widetilde{\mu_j}$, and then substituting $\widetilde{\lambda_i}$ and $\widetilde{\mu_j}$ into equation (4.31).

Seller Level Utility

At the seller level, DUE j aims to maximize the utility by selling power to CUE i given price r_{ji}, $\forall i, j$. The utility function for DUE j can be defined as follows:

$$\max \quad U_b = \sum_{\substack{i\\ j\in\Omega_i}}^{N}(r_{ji} - c_{ji})(p_j^{d_i} - K), \qquad j = 1, \cdots, M, \tag{4.32}$$

where c_{ji} denotes the cost of DUE j by sharing resource with CUE i. K is the power allocation threshold. The tradeoff between buyers and sellers can be understood in the

following. On each resource, a DUE sells power to the CUE at a proper price to make the maximum profit. Based on the given price, the CUE determines the DUE's power allocation on the resource. If the price offered by the DUE for a given resource is low, the DUE power allocated to the resource is low. Therefore, the total utility gained by DUE from the trade is low. With the increase of price, the total utility gained by DUEs will increase at the beginning. But if DUEs offer an overly high price for the power on the allocated resource block, CUEs will buy much less power, the total utility gained by DUEs will decrease. Therefore, there exists a maximum utility that can be achieved for each DUE.

By substituting the optimal power allocation at the buyer level into seller problem, we can solve the optimal price r_{ji} as the function of $\widetilde{\mu_j}$ and $\widetilde{\lambda}_i$. The proposed iterative algorithm to find the optimal power allocation is detailed in table 7.1.

$$r_{ji} = \sqrt{\frac{\beta \log_2 e(c_{ji} + \widetilde{\mu_j} + \widetilde{\lambda}_i g_{ji}^{dc})}{K + \frac{P_B g_j^d + N_0}{g_{jj}^{d_i}}}} - (\widetilde{\mu_j} + \widetilde{\lambda}_i g_{ji}^{dc}). \tag{4.33}$$

4.3 Performance Evaluation

We consider D2D communications underlaying downlink cellular network. CUEs are uniformly distributed in a single cell. DUEs are uniformly distributed. Since DUEs come in as pairs, the distance between two D2D users forming a pair has to be less than R_d. We set the BS and maximum DUE transmit power to be 46 dBm and 24 dBm. The channel model parameters include unit mean for the Rayleigh fading process, a pathloss coefficient of 4, and a standard deviation of 8 dBm for the log-normal shadowing. Number of CUEs N is fixed at 100, and the number of DUEs varies from 20 to 150 to simulate different D2D density. The value of $\Gamma_i, \forall i$ follows a uniform distribution between 7 to 13 dB with a mean of 10 dB. Threshold γ is set to be 30 dB.

We first show the performance of the proposed distributed power allocation algorithm. Fig. 4.2 presents the convergence of the buyer (CUE) utility function. Three utility curves are presented, each corresponding to a different D2D density. At each iteration, the buyer level updates its power allocation based on the price offered from each DUE. Based on the

Table 4.2: Distributed-resource-and-power-allocation-ALGORITHM FLOW

Step 1: proceed the CUE-DUE grouping scheme from table4.1,
Step 2: Initialize $\widetilde{\mu_j}$, $\widetilde{\lambda_i}$ and r_{ji} for $i = 1, 2, \cdots, N$, $\{j = 1, 2, \cdots, M\}$,
Step 3: Initialize the step size parameter δ
repeat
    Update the $\widetilde{\mu_j}$, $\widetilde{\lambda_i}$ in (4.22) and (4.22)
    calculate power allocation $p_j^{d_i}$ in (4.31)
    if $p_j^{d_i} < K$ **then**
        Delete j, i from set Ω_i and $\mathcal{D}_j$ and return to Step 2.
    else
        Update r_{ji} in (4.33)
    end if
until solution for r_{ji} converge

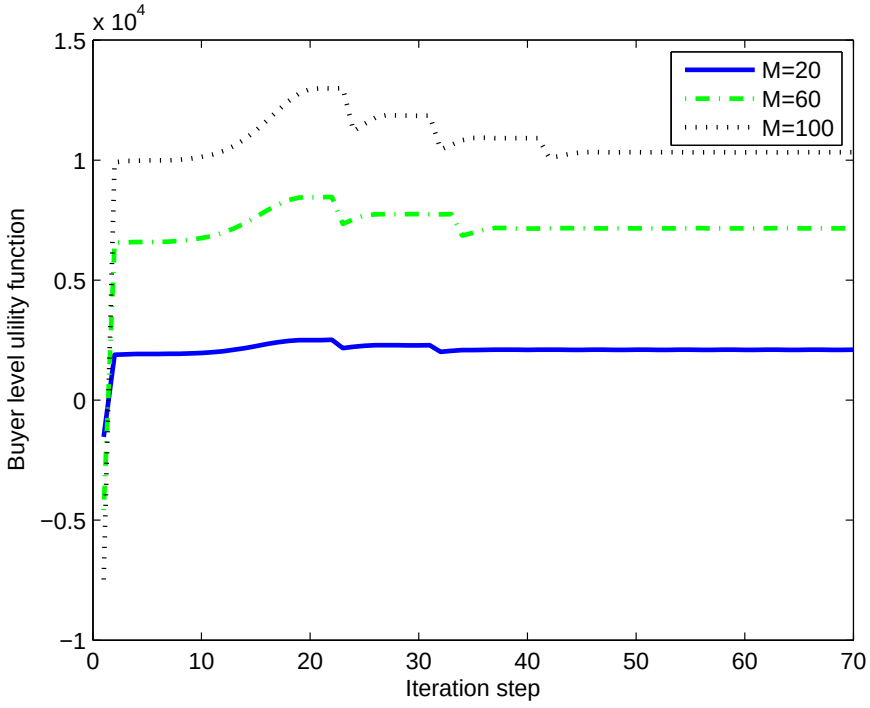

Fig. 4.2: Convergence of power allocation algorithm for different number of D2D users, where D2D cluster radius $r = 30m$, cell radius $R = 1km$

new power allocation result, the seller level calculates the new price that can achieve the maximum seller (D2D) utility. The iteration continues till both buyers and sellers achieve the equilibrium. The algorithm converge very fast at 40 to 50 iteration steps. It is obvious that a higher capacity gain can be achieved at a higher D2D density. However, the increase

rate is becoming less when more and more D2D users join the network, which indicates that the network can only support up to a certain D2D density level in order to guarantee the QoS requirement of CUEs.

Fig. 4.3 shows the average accessible ratio of DUEs to the network with respect to the D2D cluster radius. The accessible ratio is defined as the percentage of DUEs that are selected for underlay resource sharing. Since each cellular resource can be shared by multiple DUEs, the number of supported underlay DUEs is higher than that of the primary CUEs. The DUE accessible ratio is significantly influenced by D2D communication distance, which is the distance between the transmitter and the receiver of a DUE. The higher the transmission distance, the lower the accessible ratio and more DUEs will be excluded from underlay sharing due to their lower link quality and higher interference to CUEs. Thus from the system's perspective, it encourages short range D2D communications. Furthermore, when the D2D communication distance is fixed, the number of DUEs barely influences the accessible ratio.

Fig. 4.4 shows the net system capacity gain. With the increase of DUE density, the system capacity gain also improves, especially when the D2D communication distance is short. However, the capacity improvement becomes less effective when the D2D communication distance becomes large. This has been already explained in Fig. 4.3, which shows the DUE access ratio is limited significantly due to poor link quality. A large number of DUEs can be accepted to network as long as the co-channel interference to the prioritized cellular user is under control and its QoS is guaranteed.

4.4 Summary

This chapter studied the resource and power allocation problem in a cellular network with underlaid D2D communications. A distributed power and resource allocation algorithm based on Stackelberg game model was proposed. To maximally improve system capacity while not comprising cellular user's QoS, we allow each cellular downlink resource to be shared by multiple D2D users and every D2D user can use multiple resources from different cellular users. We formulate an optimization problem on joint power control and resource

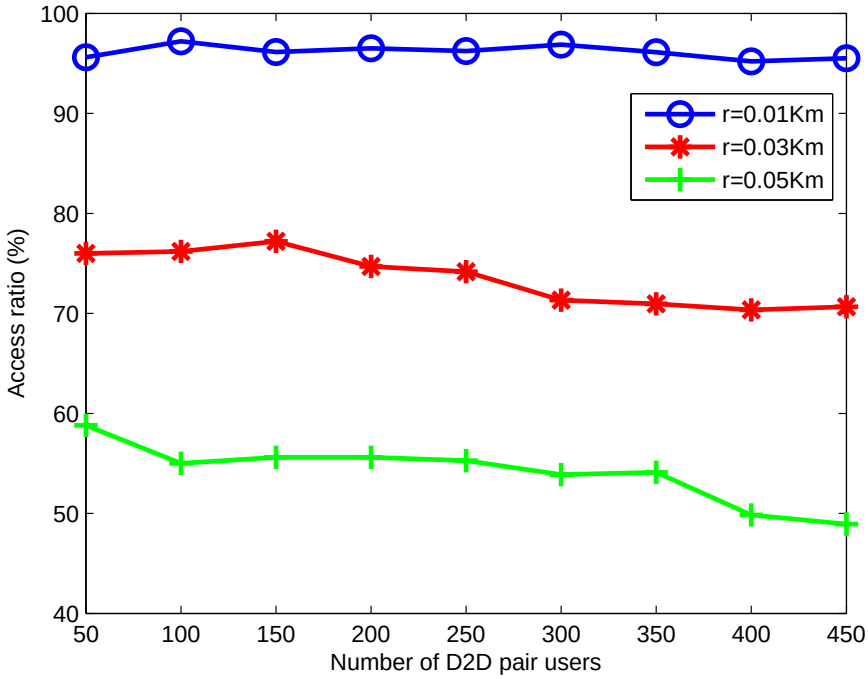

Fig. 4.3: Illustration of D2D user access ratio

allocation and further decompose the problem into a two-step approach to obtain a sub-optimal solution to reduce computational complexity. The simulation results show that our proposed algorithm can converge very fast and the system capacity is improved significantly by supporting underlay D2D users.

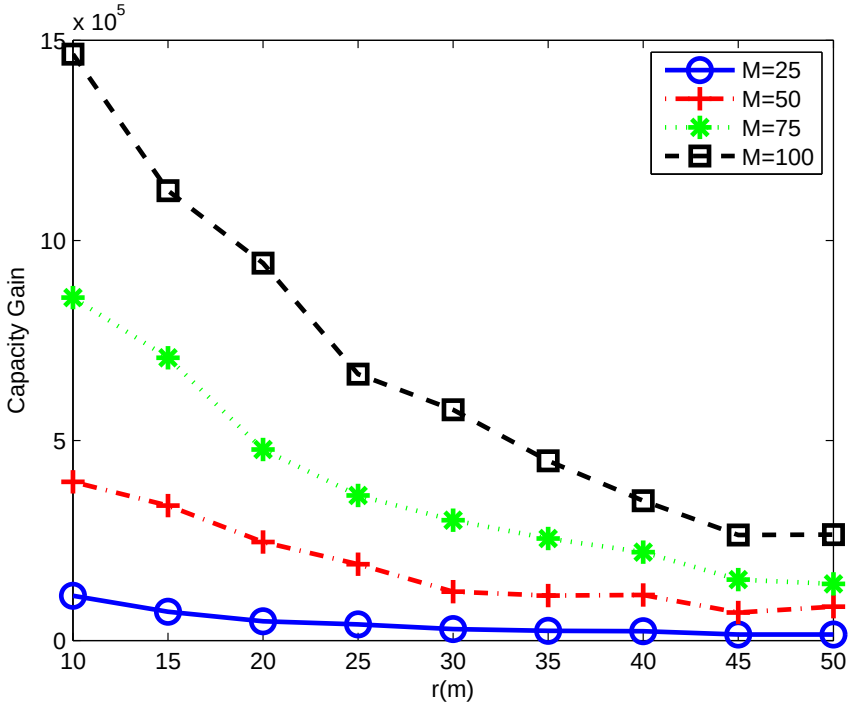

Fig. 4.4: Illustration of system capacity gain

Chapter 5

Joint Resource Allocation and Mode Selection Scheme for D2D Communication

In this chapter, we consider the device-to-device (D2D) communications underlaying cellular networks to support local communication needs. In particular, we focus our attention on the design of an optimal resource allocation and mode selection algorithm for both cellular and D2D users. In the design, communication mode selection for D2D users is taken into account so that a D2D source-destination pair has an option to either directly communicate or indirectly communicate through the base station (BS). On the other hand, it is necessary and important to provide a certain level of Quality of Service (QoS) to users. The users can be further differentiated by assigning different weighting factors and QoS requirements. To this end, we formulate a problem of maximizing the weighed sum rate of all users constrained by their power and QoS requirements. The tool that we use to handle this problem is the primal-dual technique which transforms the original problem into the equivalent problem showing lower computational complexity. We present our simulation results to show our proposed scheme outperforms the previous heuristic scheme by jointly optimizing the power and resource allocation along with mode selection for D2D communications.

5.1 System Model and Problem Formulation

5.1.1 System Model

This chapter studies an optimal power control and resource allocation problem in a cellular network with underlaying D2D communications. More particularly, the study focuses on a downlink cellular network so the D2D communication reuses cellular downlink

resources. There are two types of user equipments (UE) defined in the system, namely cellular UE (CUE) and D2D UE (DUE). CUE is defined as a UE that communicates with another UE located in a different cell while DUE is defined as a UE that communicates with another UE located in the same cell. The two communicating UEs in the same cell together are called a DUE in order to facilitate the mathematical expressions and analysis in following. The system in total has K resources blocks (RBs), indexed by the set RB $\mathcal{K} = \{1, ..., k, ...K\}$. RB is the scheduling and channel feedback granularity in this study. There are in total L UEs, indexed by the set $\mathcal{L} = \{1, ..., M, M+1, ...M+N\}$. Among L UEs, there are M CUEs, indexed by $\mathcal{M} = \{1, ..., M\}$, and N DUEs, indexed by $\mathcal{N} = \{1, ..., N\}$. $\mathcal{L} = \mathcal{M} \bigcup \mathcal{N}$. Furthermore, the system supports three communication modes, namely CUE cellular mode or C-C mode, D2D cellular mode or D-C mode, and D2D mode or D-D mode. In the D-C mode, the communication consists of two hops, first uplink from one D2D UE to BS and then downlink from BS to another D2D UE. In the D-D mode, communication between two D2D UEs (together as one DUE) in a close proximity can be directly carried out between each other in one hop. Consequently, a DUE needs to do mode selection before the real communication starts and mode selection is an essential part of the problem formulation in this study. We further use τ to numerically represent different communication mode, i.e., $\tau = 0$ represents the C-C mode, $\tau = 1$ represents the D-C mode, and $\tau = 2$ represents the D-D mode. Illustrations for three communication modes are shown in Fig.5.1.

Extensive studies have shown that inter-cell interference can be effectively mitigated with inter-cell interference control mechanisms such as power control or radio resource management. Joint resource and power allocation can be performed to achieve the most efficient spectral usage and power saving in a cellular network by fully exploiting the knowledge of channel state information (CSI). In a network with underlay DUEs, radio resources can be shared among CUEs and DUEs in a non-orthogonal mode or DUE and CUE are allocated resources in an orthogonal mode. In this work, our study assumes an orthogonal resource sharing mode. The channel model between any two communication parties. i.e. between

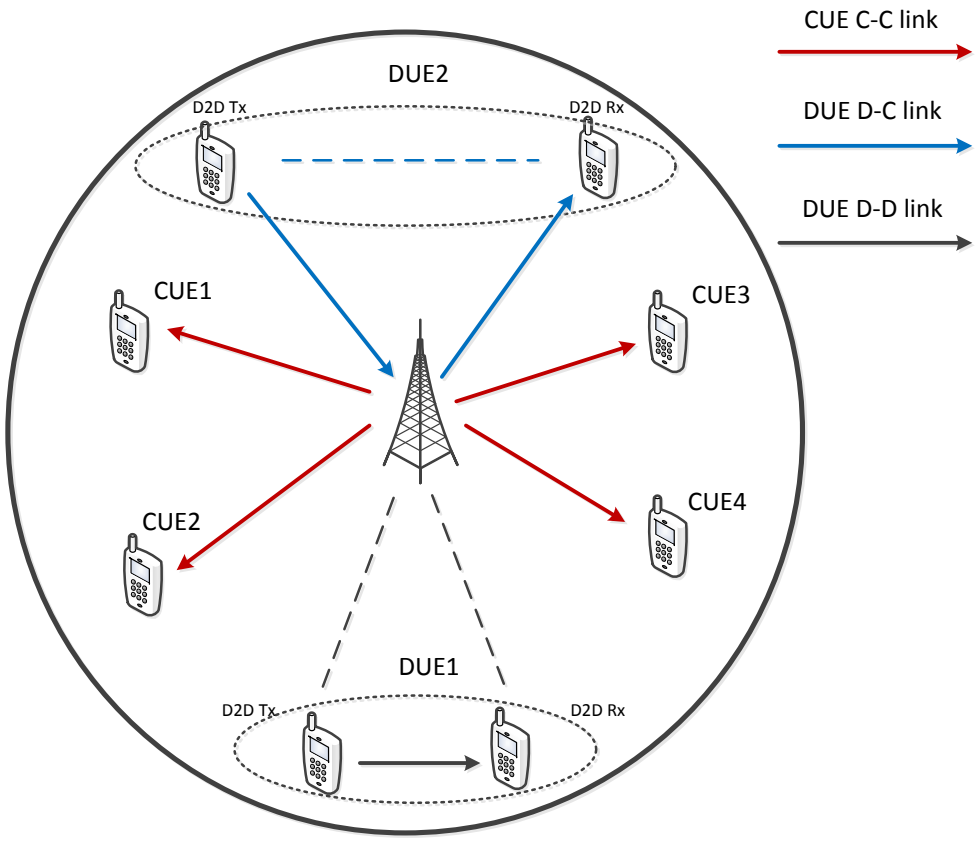

Fig. 5.1: Cellular network with D2D communication in OFDMA downlink system

BS and UE or between two UEs) consists of path loss, shadowing and fast fading. The channel gain for CUE i on k_{th} RB can be expressed as

$$g_{i,k}^{\tau=0} = g_0 \mu_{i,B}^k \phi_{i,B}^k d_{i,B}^{-\alpha}, \tag{5.1}$$

where g_0 is a system dependent constant, $\mu_{i,B}^k$ is the fast fading channel gain with an exponential distribution, $\phi_{i,B}^k$ is the slow fading gain with a log-normal distribution, α is the path loss exponent, and $d_{i,B}$ is the distance between CUE i and the BS. Similarly, the channel gain for DUE j can be expressed as $g_{j,k}^{\tau}, \tau = 1$ or 2.

5.1.2 Problem Formulation

In this work we study the mode selection, resource allocation and power control altogether in one problem to realize the maximum system spectral efficiency. As primary

users, CUEs are guaranteed with a minimum QoS. Let $p_{l,k}$ and $r_{l,k}$ denote the transmission power and rate of user l on RB k, respectively. We have the following expressions for the achievable rate at each RB for different UEs supported in different communication modes. In our study, the size of one RB is normalized into 1, which makes the rate of each RB the same as spectral efficiency.

Rate of UE l in the C-C mode:

$$r_{l,k}^{\tau=0} = \log_2(1 + \frac{p_{l,k}^{\tau=0} g_{l,k}^{\tau=0}}{\Gamma\sigma^2}), \quad \forall l \in \mathcal{M}. \tag{5.2}$$

Rate of UE l in the D-C mode:

$$r_{l,k}^{\tau=1} = \frac{1}{2}\log_2(1 + \frac{p_{l,k}^{\tau=1} g_{l,k}^{\tau=1}}{\Gamma\sigma^2}), \quad \forall l \in \mathcal{N}. \tag{5.3}$$

Rate of UE l operating in the D-D mode:

$$r_{l,k}^{\tau=2} = \log_2(1 + \frac{p_{l,k}^{\tau=2} g_{l,k}^{\tau=2}}{\Gamma\sigma^2}), \quad \forall l \in \mathcal{N}. \tag{5.4}$$

Obviously the combination of τ and UE l is only valid for certain values. For examples, $\tau = 1$ and $l \in \mathcal{M}$ or $\tau = 0$ and $l \in \mathcal{N}$ can not be valid combinations. These constraints will be considered in the problem formulation. For the D-D communication mode, we assume its uplink channel and downlink channel are symmetric. It takes 2 RBs of the serving cell, one for uplink and one for downlink, to deliver one-round communication. Thus in one RB the supported rate is cut by half, as shown in Equation (5.3). As a comparison, communication modes 0 and 2 take only one RB in the serving cell to deliver one-round communication. σ^2 is the additive white Gaussian noise on each RB, Γ defines the SNR gap between the ideal Shannon channel rate and a more practical rate based on a selected modulation and coding scheme. For example, if the M-ary quadrature amplitude modulation (QAM) applied in the system, $\Gamma = [Q^{-1}(\frac{BER}{4})]^2/3$, where $Q^{-1}(x)$ is the inverse Q-function. The overall system

optimization problem ($\mathbb{P}_1$) can be formulated as follows.

$$\mathbb{P}_1 \ : \quad \max_{p^{\tau}_{l,k}} \quad \left\{ \sum_{l \in \mathcal{L}} \omega_l \sum_{k \in \mathcal{K}} \sum_{\tau \in \{0,1,2\}} r^{\tau}_{l,k} x^{\tau}_{l,k} \right\}, \tag{5.5}$$

subject to

$$\sum_{k \in \mathcal{K}} \sum_{\tau \in \{0,1,2\}} r^{\tau}_{l,k} x^{\tau}_{l,k} \geq R_c, \quad \forall l \in \mathcal{M}, \tag{5.6}$$

$$\sum_{k \in \mathcal{K}} \sum_{\tau \in \{0,1,2\}} r^{\tau}_{l,k} x^{\tau}_{l,k} \geq R_d, \quad \forall l \in \mathcal{N}, \tag{5.7}$$

$$\sum_{k \in \mathcal{K}} \sum_{\tau \in \{1,2\}} p^{\tau}_{l,k} x^{\tau}_{l,k} \leq P^{l}_{D}, \quad \forall l \in \mathcal{N}, \tag{5.8}$$

$$\sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \sum_{\tau \in \{0,1\}} p^{\tau}_{l,k} x^{\tau}_{l,k} \leq P_B, \tag{5.9}$$

$$\sum_{l \in \mathcal{L}} \sum_{\tau \in \{0,1,2\}} x^{\tau}_{l,k} \leq 1, \quad \forall k \in \mathcal{K}, \tag{5.10}$$

$$x^{\tau}_{l,k} = \begin{cases} 0 \text{ or } 1 & \tau = 0, \ \ \forall l \in \mathcal{M}, \ \forall k \in \mathcal{K}, \\ 0 \text{ or } 1 & \tau = 1 \text{ or } 2, \ \forall l \in \mathcal{N}, \ \forall k \in \mathcal{K}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.11}$$

The above problem aims to maximize the sum of a weighted system spectral efficiency by assigning a set of weighting factors $\omega_l, \{l \in \mathcal{L}\}$, to each UE. $\omega_m > \omega_n$ gives CUEs a higher weight thus a higher priority service than DUEs. Constraints (5.6) and (5.7) enforce the minimum QoS requirements for CUEs and DUEs, respectively. The total DUE power consumption, including modes 1 and 2, is constrained to a maximum amount P^n_D in (5.8). In constraint (5.9), the total base station power consumption, including transmission power to CUEs in mode 0 and transmission power to DUEs in mode 1, is limited to P_B. P^n_D and P_B are defined as overall energy consumption caps. It should be noted that we assume a symmetric channels for uplink and downlink in mode 1 and thus the same rate is transmitted on the uplink and downlink channels. Constraint (5.10) denotes an exclusive RB assignment

rule, i.e., a single RB can be assigned to only one UE, either CUE or DUE. Constraint (5.11) ensures the right combination of mode τ and UE l.

5.2 An Optimal Power Allocation and Mode Selection Algorithm

A multi-user resource allocation problem is normally formulated as a mixed-integer nonlinear programming (MINLP) problem, which is NP-complete [51]. It implies that there is no known polynomial-time algorithm to find the optimal solution, which renders it very difficult and complex to solve. Consequently, a heuristic method or sub-optimal solution is normally pursued. ($\mathbb{P}_1$) is a typical MINLP, for which we propose a dual optimization framework in this work to solve it.

5.2.1 Dual Optimization Framework

we first get rid of all binary variables $x_{l,k}^{\tau}$ in the original problem of ($\mathbb{P}_1$). The converted problem will be incorporated into the Lagrangian dual function to form a new dual problem. The Lagrangian function is defined over a domain $\mathcal{D}$ as:

$$\begin{aligned}
&L\left(\bar{\lambda}, \{p_{l,k}^{\tau}\}, \{r_{l,k}^{\tau}\}\right) \\
&= \sum_{l\in\mathcal{L}} \sum_{k\in\mathcal{K}} \sum_{\tau\in\{0,1,2\}} \omega_l r_{l,k}^{\tau} + \sum_{l\in\mathcal{L}} \lambda_{A,l} \left(\sum_{k\in\mathcal{K}} \sum_{\tau\in\{0,1,2\}} r_{l,k}^{\tau} - R_{\{d,c\}} \right) \\
&+ \sum_{l\in\mathcal{N}} \lambda_{B1,l} \left(P_D^l - \sum_{k\in\mathcal{K}} \sum_{\tau\in\{1,2\}} p_{l,k}^{\tau} \right) \\
&+ \lambda_{B2} \left(P_B - \sum_{l\in\mathcal{L}} \sum_{k\in\mathcal{K}} \sum_{\tau\in\{0,1\}} p_{l,k}^{\tau} \right),
\end{aligned} \tag{5.12}$$

where domain $\mathcal{D}$ is defined as the set of all non-negative $p_{l,k}^{\tau}$'s $\forall l \in \mathcal{L}$, $k \in \mathcal{K}$ and $\tau \in \{0,1,2\}$ such that for each k, only one $p_{l,k}^{\tau}$ is positive $\forall l \in \mathcal{L}$ and $\tau \in \{0,1,2\}$, from constraint (5.10). $\bar{\lambda} = [\overline{\lambda_A} \quad \overline{\lambda_{B1}} \quad \lambda_{B2}]$ is the vector consisting of all Lagrangian multipliers derived

from constraints (5.6)-(5.9). Then, the Lagrangian dual function is formed as:

$$g(\bar{\lambda}) = \max_{\{p^{\tau}_{l,k}\},\{r^{\tau}_{l,k}\}\in\mathcal{D}} L\Big(\bar{\lambda}, \{p^{\tau}_{l,k}\}, \{r^{\tau}_{l,k}\}\Big). \tag{5.13}$$

It is observed that the dual function $g(\bar{\lambda})$ is a pointwise maximum of a family of affine functions of $\bar{\lambda}$. Thus $g(\bar{\lambda})$ is a convex function of $\bar{\lambda}$. Hence the original problem can be solved from the dual problem ($\mathbb{P}_2$).

$$\mathbb{P}_2 \ : \quad \min_{\bar{\lambda}\succeq 0} \quad g(\bar{\lambda}), \tag{5.14}$$

where $\succeq$ denotes the element-wise greater than or equal to symbol in above constraint. By letting $g^{'}(\bar{\lambda})$ defined in (5.15) be the function of $\bar{\lambda}$, we have the following expression:

$$\begin{aligned} g^{'}(\bar{\lambda}) =& \max_{\{p^{\tau}_{l,k}\},\{r^{\tau}_{l,k}\}\in\mathcal{D}} \Bigg\{ \sum_{l\in\mathcal{L}} \sum_{k\in\mathcal{K}} \sum_{\tau\in\{0,1,2\}} \Big(\omega_l r^{\tau}_{l,k} + \lambda_{A,l} r^{\tau}_{l,k} \Big) \\ &- \sum_{l\in\mathcal{N}} \sum_{k\in\mathcal{K}} \sum_{\tau\in\{1,2\}} \lambda_{B1,l} p^{\tau}_{l,k} - \sum_{l\in\mathcal{L}} \sum_{k\in\mathcal{K}} \sum_{\tau\in\{0,1\}} \lambda_{B2} p^{\tau}_{l,k} \Bigg\}, \qquad (5.15) \\ =& \sum_{k\in\mathcal{K}} \max_{\{p^{\tau}_{l,k}\},\{r^{\tau}_{l,k}\}\in\mathcal{D}} \Bigg\{ \sum_{l\in\mathcal{L}} \sum_{\tau\in\{0,1,2\}} \Big(\omega_l r^{\tau}_{l,k} + \lambda_{A,l} r^{\tau}_{l,k} \Big) \\ &- \sum_{l\in\mathcal{N}} \sum_{\tau\in\{1,2\}} \lambda_{B1,l} p^{\tau}_{l,k} - \sum_{l\in\mathcal{L}} \sum_{\tau\in\{0,1\}} \lambda_{B2} p^{\tau}_{l,k} \Bigg\}, \qquad (5.16) \\ =& \sum_{k\in\mathcal{K}} \max_{\substack{\tau\in\{0,1,2\} \\ l\in\mathcal{L}}} \left\{ \max_{\substack{\{p^{\tau}_{l,k}\}, \\ \{r^{\tau}_{l,k}\}\in\mathcal{D}}} \big\{ f_1(p^{\tau=0}_{l,k}), f_2(p^{\tau=1}_{l,k}), f_3(p^{\tau=2}_{l,k}) \big\} \right\}. \qquad (5.17) \end{aligned}$$

Functions f_1, f_2, and f_3 are respectively defined as:

$$\begin{aligned} f_1(p^{\tau=0}_{l,k}) \ &= \ (\omega_l + \lambda_{A,l}) r^{\tau=0}_{l,k} - \lambda_{B2} p^{\tau=0}_{l,k}, \\ &\qquad \forall k \in \mathcal{K}, \ \ \forall l \in \mathcal{M}. \end{aligned} \tag{5.18}$$

$$f_2(p_{l,k}^{\tau=1}) = (\omega_l + \lambda_{A,l}) r_{l,k}^{\tau=1} - (\lambda_{B1,l} + \lambda_{B2}) p_{l,k}^{\tau=1}, \quad \forall k \in \mathcal{K}, \quad \forall l \in \mathcal{N}. \tag{5.19}$$

$$f_3(p_{l,k}^{\tau=2}) = (\omega_l + \lambda_{A,l}) r_{l,k}^{\tau=2} - \lambda_{B1,l} p_{n,k}^{\tau=2}, \quad \forall k \in \mathcal{K}, \quad \forall l \in \mathcal{N}. \tag{5.20}$$

Equation (5.16) is derived since power and rate variables can be separable across different RBs and Equation (5.17) is satisfied because an RB is exclusively assigned to a UE with the selected mode. Thus the original problem can be decomposed into the K independent optimization problems, each of which is a per-RB optimization problem and computation complexity is prominently decreased.

Also we have

$$g(\bar{\lambda}) = g^{'}(\bar{\lambda}) - \sum_{l \in \mathcal{L}} \lambda_{A,l} R_{\{d,c\}} + \sum_{l \in \mathcal{N}} \lambda_{B1,l} P_D^l + \lambda_{B2} P_B. \tag{5.21}$$

It is not difficult to verify that $f_1(p_{l,k}^{\tau=0})$, $f_2(p_{l,k}^{\tau=1})$,$f_3(p_{l,k}^{\tau=2})$ are all concave functions of $p_{l,k}^{\tau=0}$, $p_{l,k}^{\tau=1}$, $p_{l,k}^{\tau=2}$, respectively. By evaluating the first derivative of each of them, we can attain the optimal power allocation for each user with the best mode selection.

$$p_{l,k}^{\tau=0*} = \left[\frac{1}{\ln 2} \frac{\omega_l + \lambda_{A,l}^*}{\lambda_{B2}^*} - \frac{\Gamma \sigma^2}{g_{l,k}^{\tau=0}} \right]^+, \quad \forall k \in \mathcal{K}, \quad \forall l \in \mathcal{M}. \tag{5.22}$$

$$p_{l,k}^{\tau=1*} = \left[\frac{1}{2 \ln 2} \frac{\omega_l + \lambda_{A,l}^*}{\lambda_{B1,l}^* + \lambda_{B2}^*} - \frac{\Gamma \sigma^2}{g_{l,k}^{\tau=1}} \right]^+, \quad \forall k \in \mathcal{K}, \quad \forall l \in \mathcal{N}. \tag{5.23}$$

$$p_{l,k}^{\tau=2*} = \left[\frac{1}{\ln 2} \frac{\omega_l + \lambda_{A,l}^*}{\lambda_{B1,l}^*} - \frac{\Gamma \sigma^2}{g_{l,k}^{\tau=2}} \right]^+, \quad \forall k \in \mathcal{K}, \quad \forall l \in \mathcal{N}. \tag{5.24}$$

$[\cdot]^+$ denotes the nonnegative value function. The optimal mode selection for UE l on RB k is:

$$[l,\tau]_k^* = \underset{\substack{m\in\mathcal{M},n\in\mathcal{N}\\ \tau\in\{0,1,2\}}}{\arg\max} \left\{ f_1^*(p_{m,k}^{\tau=0}), f_2^*(p_{n,k}^{\tau=1}), f_3^*(p_{n,k}^{\tau=2}) \right\}, \tag{5.25}$$

The optimal values for binary variables can be determined in the following.

$$(x_{l,k}^{\tau})^* = \begin{cases} 1 & \{l,\tau\} = [l,\tau]_k^*, \quad \forall k \in \mathcal{K}; \\ 0 & \text{otherwise.} \end{cases} \tag{5.26}$$

Corresponding optimal rate allocations $r_{l,k}^{\tau=0^*}$, $r_{l,k}^{\tau=1^*}$, and $r_{l,k}^{\tau=2^*}$ can be evaluated by inserting above optimal power expressions and best mode selection to Equations (5.2)-(5.4).

Proposition 5.2.1. *For the dual optimization problem* $(\mathbb{P}_2)$ *with a Lagrangian dual function* $g(\bar{\lambda})$ *defined in (5.13), the following vector* S *is a subgradient for* $g(\bar{\lambda})$*:*

$$S = \left[\overline{S_1} \quad \overline{S_2} \quad P_B - \sum_{l\in\mathcal{L}} \sum_{k\in\mathcal{K}} \sum_{\tau\in\{0,1\}} p_{l,k}^{\tau *} \right], \tag{5.27}$$

where

$$\overline{S_{1,l}} = \sum_{k\in\mathcal{K}} \sum_{\tau=0} r_{l,k}^{\tau *} - R_c, \qquad \text{if } \ l \in \mathcal{M}, \tag{5.28}$$

$$\overline{S_{1,l}} = \sum_{k\in\mathcal{K}} \sum_{\tau\in\{1,2\}} r_{l,k}^{\tau *} - R_d, \qquad \text{if } \ l \in \mathcal{N}, \tag{5.29}$$

$$\overline{S_{2,l}} = P_D^n - \sum_{k\in\mathcal{K}} \sum_{\tau\in\{1,2\}} p_{l,k}^{\tau *}, \qquad \forall l \in \mathcal{N}. \tag{5.30}$$

$p_{l,k}^{\tau *}$ *and* $r_{l,k}^{\tau *}$ *optimize the maximization.*

Proof. For any $\bar{\nu} = [\overline{\nu_A} \;\; \overline{\nu_{B1}} \;\; \nu_{B2}] \succeq 0$, we have the following.

$$\begin{aligned} g(\bar{\nu}) &= \max_{\{p_{l,k}^{\tau}\},\{r_{l,k}^{\tau}\}\in\mathcal{D}} L\left(\bar{\nu},\{p_{l,k}^{\tau}\},\{r_{l,k}^{\tau}\}\right) \geq L\left(\bar{\nu},\{p_{l,k}^{\tau*}\},\{r_{l,k}^{\tau*}\}\right) \\ &= g(\bar{\lambda}) + \sum_{l\in\mathcal{L}}(\nu_{A,l}-\lambda_{A,l})\cdot\left(\sum_{k\in\mathcal{K}}\sum_{\tau\in\{0,1,2\}} r_{l,k}^{\tau*} - R_l\right) \\ &+ \sum_{l\in\mathcal{N}}(\nu_{B1,l}-\lambda_{B1,l})\cdot\left(P_D^l - \sum_{k\in\mathcal{K}}\sum_{\tau\in\{1,2\}} p_{l,k}^{\tau*}\right) \\ &+ (\nu_{B2}-\lambda_{B2})\cdot\left(P_B - \sum_{l\in\mathcal{L}}\sum_{k\in\mathcal{K}}\sum_{\tau\in\{0,1\}} p_{l,k}^{\tau*}\right) \\ &= g(\bar{\lambda}) + S\cdot(\bar{\nu}-\bar{\lambda})^T. \end{aligned} \tag{5.31}$$

□

5.2.2 Ellipsoid Method Based Optimal Search

Theoretically, the solution for the dual problem $\mathbb{P}_2$ just provides an upper bound for the original primary problem if the primary problem is not a convex problem, which is the case for the primary problem $\mathbb{P}_1$. However, it has been proved in [52] that when the total number of users is becoming large enough (> 8) as in our case, the duality gap is approaching zero. Thus we can solve the dual problem and find the optimal dual variables $\overline{\lambda}^*$. Then by substituting them back into the primary problem, the optimal values for the primary variables ${p_{l,k}^{\tau}}^*$ and ${r_{l,k}^{\tau}}^*$ can be found. Through the above analysis, a joint optimal resource allocation and mode selection based on a dual framework is developed and the detailed algorithm flow is illustrated in Table 5.1.

The search for dual variables $\overline{\lambda}$ is done by using the ellipsoid method, which has a better performance compared with the subgradient method [50]. Ellipsoid method converges in $\mathcal{O}(n^2)$ steps where n is the number of variables. In our problem, overall optimization needs $\mathcal{O}((L+N+1)^2)$ runs of an optimization problem with a complexity of $\mathcal{O}((L+N)K)$. Hence, $\mathcal{O}(K(L+N)^3)$ executions are required to find the optimal solutions by using the proposed algorithm. As an example of one constraint parameter setting, convergence of

Table 5.1: Optimal Resource Allocation and Mode Selection Algorithm

Input: R_m, R_n, P_D^n, P_B and ω_l, $\forall m \in \mathcal{M},\ \forall n \in \mathcal{N}$
Output: ${p_{l,k}^{\tau}}^*$, ${r_{l,k}^{\tau}}^*$, ${x_{l,k}^{\tau}}^*$, $\forall l \in \mathcal{L},\ \forall k \in \mathcal{K},\ \forall \tau \in \{0,1,2\}$
1: Initialize $\lambda_{A,l}$, $\lambda_{B1,n}$ and λ_{B2}, $\forall l \in \mathcal{L},\ \forall n \in \mathcal{N}$
2: Initialize parameter for Ellipsoid search
3: **while** $g(\bar{\lambda})$ has not converge **do**
4: **for all** $l \in \mathcal{L},\ k \in \mathcal{K},\ \tau \in \{0,1,2\}$ **do**
5: calculate optimal power $p_{m,k}^{\tau=0^*}$,$p_{n,k}^{\tau=1^*}$,$p_{n,k}^{\tau=2^*}$ from (6.19)–(5.24), and $r_{m,k}^{\tau=0^*}$,$r_{n,k}^{\tau=1^*}$,$r_{n,k}^{\tau=2^*}$ correspondingly.
6: **end for**
7: **for all** $k \in \mathcal{K}$ **do**
8: Calculate function f_1, f_2,f_3 in (5.18) – (5.20), respectively.
9: Update RB allocation index ${x_{l,k}^{\tau}}^*$ by (5.26).
10: **end for**
11: Evaluate subgradient S in (5.27).
12: Update $\lambda_{A,l}$, $\lambda_{B1,n}$ and λ_{B2} using ellipsoid method.
13: Evaluate Lagrangian function $g(\bar{\lambda})$ in (5.21).
14: **end while**

some of the Lagrangian multipliers and the dual objective function is shown in Fig. 5.2. Here the convergence of λ_{B2} and $\lambda_{A,m}$ in the figure implies the rate constraint (5.6) and power constraint (5.9) are all satisfied in our algorithm.

5.3 Numerical Analysis

In this section, performance of the proposed algorithm is evaluated in simulations by considering a single cell network with a radius $R = 1$ Km, where regular CUEs are uniformly distributed in the cell. The two UEs of a DUE are randomly located in a circle with a radius of $R_d = 100$ m and DUEs are uniformly distributed in the cell. In total 64 RBs are considered in the system. The wireless channels are modeled by using path loss, shadow fading and Rayleigh fading. A distance-based path loss of each wireless link is modeled as $PL = 128 + 40\log_{10}(d)$, where d is the distance of the link in kilometer. The variance for log-normal shadow fading is 8 dB. The noise power spectral density is set to be -174 dBm/Hz. Transmission power limit for DUEs is $P_D^n = 24$ dBm and the peak transmission power for a BS is set as $P_B = 46$ dBm.

In Fig. 5.2, we first show the convergence of the proposed optimization algorithm. 5

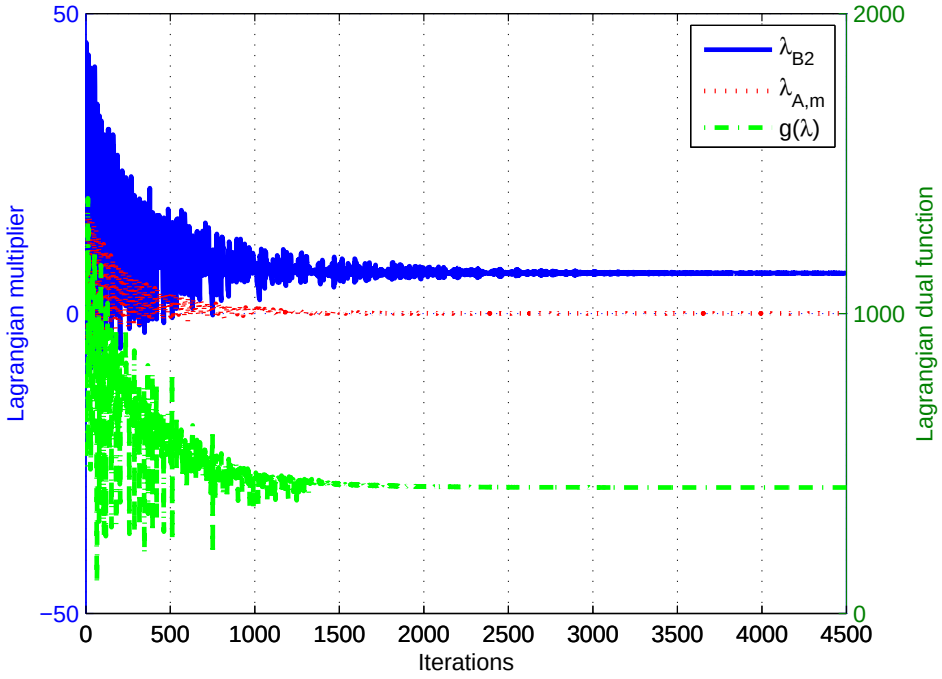

Fig. 5.2: Convergence of dual objective function

CUEs and 5 DUEs are uniformly distributed around the cell. QoS requirements for CUEs and DUEs are set to be $R_m = 10$ bits/s/Hz and $R_n = 5$ bits/s/Hz, respectively. The convergence of the Lagrangian dual function is shown in Fig. 5.2. To further verify the performance of the algorithm, we compare our dual optimization approach with a heuristic scheme based on the greedy water-filling approach proposed in [25] in a similar system environment. We set $\omega = [1\ 1 \cdots\ 1]$, so that weighting factors for all CUEs and DUEs are the same. The system objective is to maximize the total system spectral efficiency. In the approach defined in [25], exclusive RB allocation is done by greedy water-filling for each user, assuming all DUEs operate in the cellular mode. The ultimate operation mode for a DUE is fixed according to transmission power increment at each given RB and mode selection. Although the system objective in work [25] is different from the one defined in this work, it is not hard to derive a very similar heuristic approach in our system for comparison purpose. In the simulation, 5 CUEs and 15 DUEs are dropped uniformly in the cell. QoS requirements for CUEs and DUEs are set to be $R_c = 20$ bits/s/Hz and $R_d = 10$ bits/s/Hz,

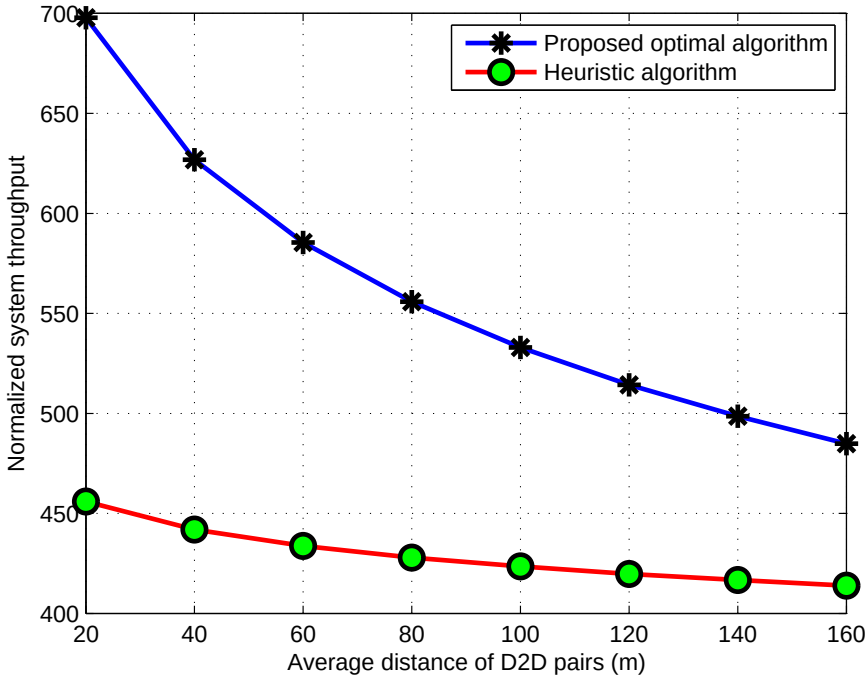

Fig. 5.3: Total downlink system throughput vs average distance between D2D pair

respectively. From Fig. 5.3, it can be seen that our approach achieves significantly higher system throughput than the chosen heuristic method which performs RB, power allocation and mode selection in a heuristic way. This is because more RBs are allocated to the D-D mode in our case, which provides a higher spectral efficiency than the D-C mode. But the rate allocated to DUEs decrease when increasing the average distance between the two UEs within a DUE, while the rate for CUEs almost keeps unchanged. Thus, the total system rate decreases when the average distance between D2D users in a pair increases.

In Fig. 5.4, we investigate the average rates for CUEs and DUEs under different weighting factors. The average distance between D2D UEs is set to 100 m. QoS requirements for CUEs and DUEs are set to be $R_c = 20$ bits/s/Hz and $R_d = 10$ bits/s/Hz, respectively. The weighting factor for CUE m is ω_m and it is ω_n for DUE n. All the CUEs have the same weighting factors. All The DUEs have the same weighting factor as well. The ratio r_o between a CUE weighting factor and a DUE weighting factor, i.e.,, $r_o = \omega_m : \omega_n$, changes from $1 : 8$ to $8 : 1$. When the value of the weighting factor for CUEs increases, more RBs

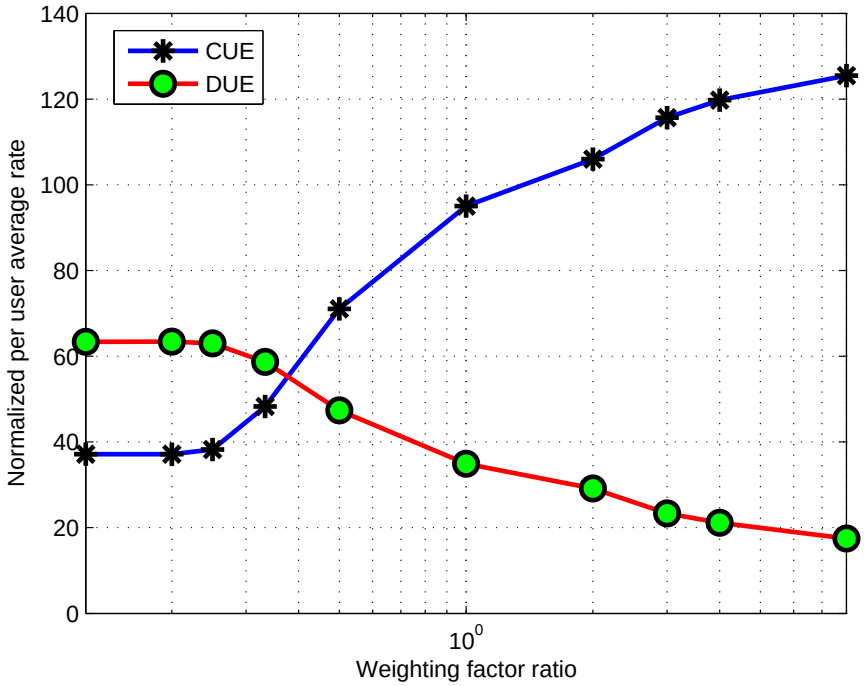

Fig. 5.4: Average cellular and D2D user data rate under different weighting factor

will be assigned to CUEs and less RBs to DUEs, so that the total weighted system throughput increases. Similarly, when the weighting factor for DUEs increases, more RBs would be assigned to DUEs and less to CUEs. It should be noted that when the weighting factor become favorable to DUE or CUE, their rate increment or decrement is diminishing. This is because both cellular user and D2D users are constrained by BS or their own transmission power and QoS requirement.

5.4 Summary

In this chapter, a joint optimal resource/power allocation and mode selection scheme is proposed in a cellular network with underlay D2D commutation. The problem is a NP-complete mixed integer nonlinear programming problem. We develop a dual optimization framework to solve the problem with a reasonable computational complexity of $\mathcal{O}(K(L+N)^3)$. Simulation results show that the D2D users can intelligently select the transmission mode, either through D2D direct transmission or through two-top transmission via base

station depending on the channel condition and resource constraint. The comparison with other schemes show that the proposed scheme can achieve a much higher total system throughput.

Chapter 6

Energy Efficient Resource Allocation for D2D Communication

In this chapter, we discuss the topic of energy efficient resource allocation for D2D communication underlaying cellular network. Energy efficiency of each user is defined as the achievable rate normalized by the power consumption, which is in [bits/sec/Joule]. We formulate the problem a non-cooperative game, where mobile users, either legacy cellular users or D2D users, decide their respective transmission power over available resource blocks (RBs) with the goal of maximizing their own utility function. Such a utility function well reflects the users' satisfaction in reality when users are mobile and subject to the availability of energy due to the finite battery capacity and limited recharging facility. The scenario where the cellular and the D2D connections share the same resources is considered, in which the interference management between cellular and D2D communications is of great importance to guarantee the performance of high-priority cellular users. The most energy efficient strategy turns out that a CUE allocates optimal power on its assigned channel for maximum energy efficiency, while a DUE should allocate the least amount of power on the channel which has the best channel gain-to-interference-noise-ratio (CINR). Yet the DUE still achieves the highest data rate on that channel.

6.1 System Model

We consider a network consisting of a single cell with one BS in the center and a plurality of UEs distributed in the cell. In the system, two types of UEs are supported, namely CUE and DUE. CUEs directly communicate with the BS, whereas DUEs communicate with each other over direct D2D links while remain under the control of the BS. Since CUEs are generally considered as high-priority primary users, the performance guarantee of CUEs

is of the main concern in designing cellular networks with underlay D2D communication. D2D communication can be established when two UEs are in a close proximity.

To further enhance efficient utilization of radio resources, D2D communication is allowed to share the resources used by the cellular communications in an underlay mode. This is possible since D2D communications are usually short-range communications so that interference to CUEs is rather low. Extensive research have shown that the inter-cell interference can be controlled at an acceptable level via proper radio resource management in the overlay cellular networks. Thus in this work, we confine the interference problem in the intra-cell level and focus on the intra-cell interference coordination between cellular and the D2D communications.

An illustration of the network model is shown in Figure 6.1, where D2D communication reuses cellular radio spectrum in an underlay mode, i.e., D2D communication shall control the interference to the cellular users by sharing the same resources. This is in line with the current discussion on the D2D communication in Long Term Evolution (LTE) Release-12 [7]. A example of D2D underlay resource sharing is as follows: DUE1 shares f_3, DUE2 shares f_1, DUE3 shares f_1 and f_2, DUE4 shares f_2 and f_3 when CUE1, CUE2, CUE3 take up orthogonal resources f_1, f_2, and f_3, respectively, as shown in Figure 6.1. In a later section, we will elaborate on how power allocation and rate distribution are dependent on the channel CINR if the objective is to achieve the maximum energy efficiency.

The notations used in this work are as follows. Total frequency resources are divided into K orthogonal channels, or interchangeably resource blocks (RBs) in LTE terminology. There are M active CUEs, each of which would be assigned one RB, assuming $K \geq M$. N denotes the number of D2D communication pairs, each of which can be assigned multiple RBs as long as the interference to the cellular communication is kept below an acceptable level and the total power constraint of a device, which is denoted by P_{dd}, is not violated. The channel with pathloss, Rayleigh fading and log-normal shadowing is considered. The channel gain from CUE m to the BS over channel k is $g_m^k = |h_m^k|^2 = g_0^k * \varsigma_m^k * \mu_m^k * d_m^{-\alpha}$, where g_0^k is the path-loss at the reference distance, ς_m^k is the shadowing, μ_m^k is the squared

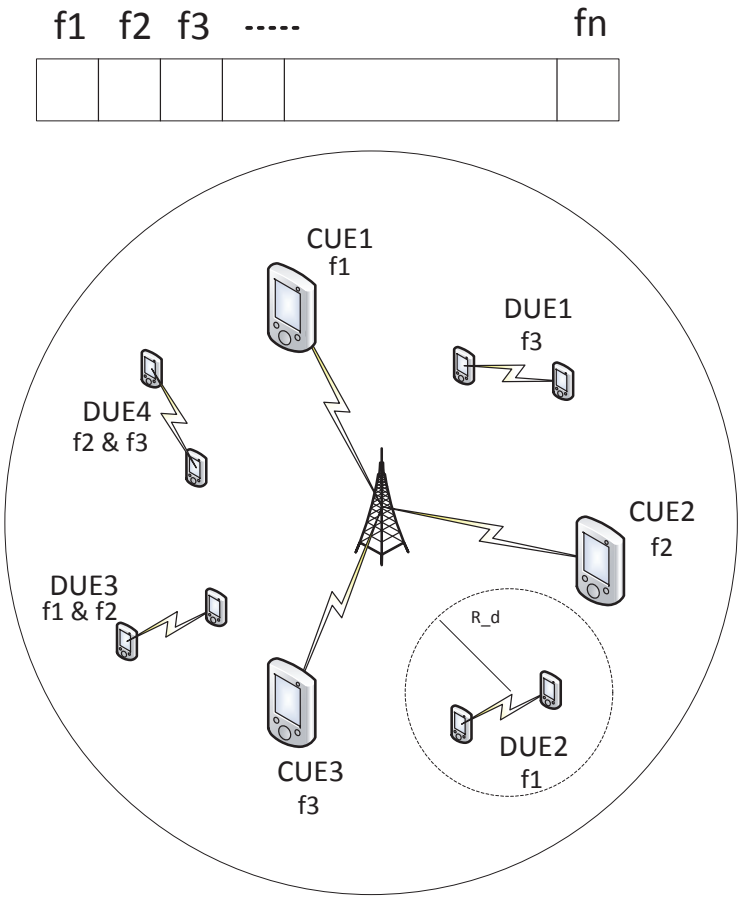

Fig. 6.1: Illustration of Device-to-Device underlaying cellular network.

magnitude of the Rayleigh fading coefficient that follows an exponential distribution with unit mean, d_m is the distance between CUE m and BS, and α is the path-loss exponent. A DUE denotes the pair of UEs that form a D2D connection. The channel gain from DUE i to DUE j over channel k is $g_{ij}^k = |h_{ij}^k|^2 = g_0^k * \varsigma_{ij}^k * \mu_{ij}^k * d_{ij}^{-\alpha}$. If DUEs operate in the uplink underlay mode, i.e., DUEs reuse cellular uplink resources, the received signal-to-interference-plus-noise-ratio (SINR) of CUE m at the BS is given by[1]

$$\text{SINR}_m^k = \frac{p_m^k g_m^k}{\sum_{n=1}^{N} p_n^k g_n^k + \sigma^2}, \tag{6.1}$$

where p_m^k is the transmit power of CUE m over channel k, p_n^k is the transmit power of the

[1]For simplicity of exposition, the subscript indicating the receiver is omitted when the receiver is the BS.

source of the DUE n over channel k, g_n^k is the channel gain between DUE n and BS, and σ^2 is the noise power. Due to the exclusive channel assignment among CUEs, the interference to the cellular link only comes from DUEs using the same uplink channel. However, the interference at DUE is from both the CUE operating on that channel, and other co-channel DUEs. Therefore, the received SINR of DUE n is given by

$$\mathrm{SINR}_n^k = \frac{p_n^k g_{nn}^k}{\sum_{i=1 i\neq n}^{N} p_i^k g_{in}^k + p_m^k g_{mn}^k + \sigma^2}, \tag{6.2}$$

where g_{nn}^k is the channel gain on channel k between the two UEs that form DUE n, g_{in}^k is the channel gain between DUE i and DUE n; g_{mn}^k is the channel gain between CUE m and DUE n. Here the corresponding CINR is

$$\Gamma_n^k = \mathrm{CINR}_n^k = \frac{g_{nn}^k}{\sum_{i=1 i\neq n}^{N} p_i^k g_{in}^k + p_m^k g_{mn}^k + \sigma^2}. \tag{6.3}$$

Based on Shannon capacity, the achievable rate of UE, either a CUE m or a DUE n, over channel k is given by

$$r_{(n \text{ or } m)}^k = w \log_2 (1 + p_{(n \text{ or } m)}^k \Gamma_{(n \text{ or } m)}^k). \tag{6.4}$$

w is the bandwidth of each channel. We will show later $\Gamma_{(n \text{ or } m)}^k$ is a critical decision parameter in the Nash equilibrium. In the downlink underlay mode, all the expressions are similar and will not be elaborated here. In the following, we will use index n to denote DUE, m to denote CUE. l to denote either a CUE or a DUE. Thus we have $n = 1, \cdots, N$, $m = 1, \cdots, M$, $l = 1, \cdots, N + M$.

6.2 Non-cooperative Resource Allocation Game in D2D Network

The objective of our game is to achieve an energy efficient resource allocation for D2D communication, which is different from the one maximizing the spectral efficiency [53]. As explained earlier, each CUE is assigned one channel, or equivalently one RB. RBs are assigned exclusively among CUEs. Each DUE, on the other hand, can be assigned multiple

RBs. Furthermore, one RB can be shared among one CUE and multiple DUEs. Fig. 6.1 exemplifies such an allocation. All the UEs that share the same channel will interfere with each other. The study here assumes a perfect knowledge on the instantaneous channel conditions at the transmitter side.

6.2.1 Energy Efficiency Based Utility Function

The following power consumption model is considered, which captures not only the transmission power at the RF front-end but also the power consumption along the entire circuitry including mixer, analog-to-digital converter (ADC), digital-to-analog converter (DAC), filters, and digital signal processing (DSP) blocks [54]. The circuit power consumption can be divided into the static background power consumption and the dynamic one that is dependent on the communication data rate $\epsilon R + P_s$, where ϵ is the power consumption per unit rate, R is the data rate, and P_s is the static background power consumption. By taking into account the transmission power P at the RF front-end, the total power consumption can be expressed as

$$P_{\text{total}} = \zeta P + \epsilon R + P_s, \tag{6.5}$$

where ζ is a factor reflecting the compensation for the power amplifier loss. We now define the utility of UE l, either CUE or DUE, in terms of the rate per unit power as

$$U_l(\mathbf{P}_l, \mathbf{P}_{-l}) = \frac{R_l}{\zeta P_l + \epsilon R_l + P_s}, \tag{6.6}$$

where $\mathbf{P}_l = \left\{p_l^1, \cdots, p_l^K\right\}$ is the transmission power vector of UE l over K RBs,

$$\mathbf{P}_{-l} = \{\mathbf{P}_1, \cdots, \mathbf{P}_{l-1}, \mathbf{P}_{l+1}, \cdots, \mathbf{P}_{N+M}\} \tag{6.7}$$

is the collection of the transmission power vector of all the UEs including both CUEs and DUEs except UE l. $P_l = \sum_{k=1}^{K} p_l^k$, and $R_l = \sum_{k=1}^{K} r_l^k$, representing the total consumed power and total transmission rate for UE l, respectively. It should be noted that the power

allocation vector $\mathbf{P}_{-l}, \{l = N+1, \cdots, N+M\}$ for a CUE can only have no more than one non-zero element as each CUE is only assigned one channel.

Given the knowledge on the power allocation of all other UEs, $\mathbf{P}_{-l}$, the optimal power allocation by UE l is to maximize its own utility, i.e.,

$$\mathbf{P}_l^*(\mathbf{P}_{-l}) = \arg\max_{\mathbf{P}_l} U_l(\mathbf{P}_l, \mathbf{P}_{-l}). \tag{6.8}$$

Proposition 6.2.1. *The utility function $U_l(\mathbf{P}_l, \mathbf{P}_{-l})$ is a strict quasi-concave function of transmission rate r_l^k for all $k \in 1, \cdots, K$.*

Proof. From the definition of quasi-concavity [50], it suffices to show that the sub-level set $\left\{r_l^k | U_l(\mathbf{P}_l, \mathbf{P}_{-l}) \geq \alpha\right\}$ is a convex set for all possible values of α. When $\alpha \leq 0$, the sub-level set is the set of real numbers, which is obviously a convex set. When $\alpha > 0$, it is straightforward to see that the sub-level set is written as $\left\{r_l^k | \alpha\zeta \sum_{k=1}^K p_l^k + (\alpha\epsilon - 1) \sum_{k=1}^K r_l^k + \alpha P_s \leq 0\right\}$. Note that from Eq. 6.4 the transmission power of UE l over channel k can be expressed as

$$p_l^k = \frac{2^{r_l^k/w} - 1}{\Gamma_l^k}, \tag{6.9}$$

which is a strictly convex function of r_l^k. Since a summation of convex functions and linear functions is still convex, the corresponding sub-level set is also a convex set, which proves the proposition. □

A Nash equilibrium of non-cooperative games is achieved if no UE can increase its utility by unilaterally altering its power allocation strategy. In our game model in Eq. (6.8), each UE selects its transmission power vector to maximize its own utility. Thus, at Nash equilibrium (if exists), all the UEs achieve their own maximum utility by non-cooperatively choosing the transmission power vector.

Proposition 6.2.2. *In the game model in* (6.8)*, there exists at least one Nash equilibrium. When the non-cooperative game reaches Nash equilibrium, power allocation and transmission rate for the UE l on channel k must satisfy one of the following two conditions: 1) if*

$\zeta \sum_{i=1 i\neq k}^{K} p_l^i + P_s - (\sum_{i=1, i\neq k}^{K} r_l^i)\zeta \frac{\ln 2}{\Gamma_l^k w} \leq 0$, *then the optimal power and rate allocation must be* $p_l^k = r_l^k = 0$ *2) if* $\zeta \sum_{i=1, i\neq k}^{K} p_l^i + P_s - (\sum_{i=1 i\neq k}^{K} r_l^i)\zeta \frac{\ln 2}{\Gamma_l^k w} > 0$, *then the power allocation must satisfy* $\zeta P_l + P_s - R_l \zeta \frac{\partial P_l}{\partial r_l^k} = 0$.

Proof. As proved before, 1) the utility function U_l is a continuous and strictly quasi-concave function on $\mathbf{R}_l = [r_l^1, r_l^2 \cdots, r_l^K]$, $\forall l$, and 2) the set $\mathbf{R}_l$ is a nonempty, convex subset of the real number. These two properties guarantee the existence of a Nash equilibrium [55]. Next, let's consider the first-order partial derivatives of the utility function and express it as

$$U_l{}' = \frac{\partial U_l}{\partial r_l^k}, \ \forall k \in 1, \ldots, K. \tag{6.10}$$

Given other UE's transmit power and channel gain, transmit rate r_l^k of UE l over RB k is only dependent on the power allocation p_l^k on this RB, thus we have

$$\frac{\partial U_l}{\partial r_l^k} = \frac{\zeta P_l + P_s - R_l \zeta \frac{\partial P_l}{\partial r_l^k}}{(\zeta P_l + \epsilon R_l + P_s)^2}.$$

Define $F(r_l^k) \triangleq \zeta P_l + P_s - R_l \zeta \frac{\partial P_l}{\partial r_l^k}$. We then find that $F'(r_l^k) = -R_l \zeta \frac{\partial^2 P_l}{\partial {r_l^k}^2}$.

We have $\frac{\partial P_l}{\partial r_l^k} = \frac{2^{r_l^k/w} \ln 2}{\Gamma_l^k w}$ and $\frac{\partial^2 P_l}{\partial {r_l^k}^2} = \frac{2^{r_l^k/w} (\ln 2)^2}{\Gamma_l^k w^2}$, both of which are positive. Thus, $F'(r_l^k)$ is always non-positive and $F(r_l^k)$ is a non-increasing function of r_l^k, for all $k \in 1, \ \ldots, \ K$. We further evaluate the function $F(r_l^k)$ at $r_l^k = 0$, which is given by $F(r_l^k)|_{r_l^k=0} = \zeta \sum_{i=1, i\neq k}^{K} p_l^i + P_s - (\sum_{i=1, i\neq k}^{K} r_l^i)\zeta \frac{\ln 2}{\Gamma_k^k w}$.

1. If $F(r_l^k)|_{r_l^k=0} \leq 0$, the optimal transmit power and the corresponding rate over channel k is zero. Thus, the utility over that channel is also zero.

2. If $F(r_l^k)|_{r_l^k=0} > 0$, the optimal transmit power on the channel satisfies the equation $\zeta P + P_s - R\zeta \frac{\partial P}{\partial r_l^k} = 0$, which is equivalent to

$$\begin{aligned} \frac{\zeta P + P_s}{R} &= \zeta \frac{\partial P}{\partial r_l^k}|_{r_l^{*k}}, \\ &= \frac{\zeta 2^{r_l^{*k}/w} \ln 2}{\Gamma_l^k w}, \\ &= \zeta \frac{\ln 2(1 + p_l^{*k}\Gamma_l^k)}{\Gamma_l^k w}. \end{aligned} \tag{6.11}$$

From the second condition, we can observe that the right side of equation (6.11) is the power increase rate while the left side is not related to channel index k. So if multiple channels are allocated to user l, they must have the same power increase rate on each channel at Nash equilibrium. Next, we consider the system operating in flat fading channel scenario, in which we assume the UE experiences the same channel fading over all RBs. Thus the channel gain is the same on all RBs for UE. We will study how the proposed game in the flat channel environment works. □

6.2.2 Power Allocation in Flat-fading Channel

In a flat fading system, a UE experiences the same statistical channel gain on all RBs. Based on the proposed non-cooperative game model, CUE can only be assigned a single RB. DUEs can use multiple RBs and they will need to allocate their power on different RBs.

Proposition 6.2.3. *With flat fading channels, there exists a unique equilibrium for the non-cooperative game. The best action taken by CUE m is $\mathbf{P}_m^*(\mathbf{P}_{-m}) = \arg\max_{\mathbf{P}_m} U_n(\mathbf{P}_m, \mathbf{P}_{-m})$, which is a standard function [56] and has the following properties.*

1. *Positivity: $\mathbf{P}_m^*(\mathbf{P}_{-m}) > 0$.*

2. *Concavity: $\mathbf{P}_m^*(\mathbf{P}_{-m})$ is strictly concave of $\mathbf{P}_{-m}$.*

3. *Monotonicity: if* $\mathbf{P}_{-m} \succ \mathbf{Q}_{-m}$*, then* $\mathbf{P}^*_m(\mathbf{P}_{-m}) > \mathbf{P}^*_m(\mathbf{Q}_{-m})$*, where* $\succ$ *denotes the vector inequality.*

4. *Scalability: If* $\beta > 1$*, then* $\beta\mathbf{P}^*_m(\mathbf{P}_{-m}) > \mathbf{P}^*_m(\beta\mathbf{P}_{-m})$.

Proof. We prove the following properties for CUEs. The conclusions can be extended to DUEs as well.

From $\mathbf{P}^*_m(\mathbf{P}_{-m}) = \arg\max_{\mathbf{P}_m} \quad U_m(\mathbf{P}_m, \mathbf{P}_{-m})$, one can see that $U_m(\mathbf{P}_m, \mathbf{P}_{-m}) \geq 0$ and $U_m(\mathbf{P}_m, \mathbf{P}_{-m}) = 0$ only when $\mathbf{P}_m = 0$. Thus $\mathbf{P}^*_m(\mathbf{P}_{-m}) > 0$.

We further prove monotonicity and concavity. Each CUE can transmit on any single RB. Thus, the sufficient and necessary condition for equilibrium from equation (6.11) must be satisfied:

$$\frac{\partial U_m}{\partial r^k_m} = 0. \tag{6.12}$$

By applying chain rule $\frac{\partial U_m}{\partial p^k_m} = \frac{\partial U_m}{\partial r^k_m} \cdot \frac{\partial r^k_m}{\partial p^k_m}$, above condition is equivalent to $\frac{\partial U_m}{\partial p^k_m} = 0$, as $\frac{\partial r^k_m}{\partial p^k_m} = \frac{w\Gamma^k_m}{\ln 2(1+p^k_m\Gamma^k_m)} \neq 0$. Then Nash equilibrium condition becomes:

$$R^{'}(P^*_m) = \frac{\zeta R}{\zeta P + P_s}. \tag{6.13}$$

$R(p^k_m) = r^k_m$ and $P = p^k_m$ for CUE m. Denote $\Gamma^k_m = \frac{g^k_{mm}}{I+\sigma^2}$. Here $I = \sum_{i=1}^{N} p^k_i g^k_{im}$ denotes the sum interference, which is a linear combination of $\mathbf{P}_{-m}$. By substituting $R^{'}(p^k_m) = \frac{w\Gamma^k_m}{\ln 2(1+p^k_m\Gamma^k_m)}$, we can get

$$\begin{aligned} g^k_{mm}(P^*_m + \frac{P_s}{\zeta}) = & (I + \sigma^2 + g^k_{mm}P^*_m) \times \\ & \ln(1 + g^k_{mm}\frac{P^*_m}{I + \sigma^2}) \end{aligned} \tag{6.14}$$

By evaluating the derivative of P_m^* as a function of I on both sides, the first order derivative is expressed as

$$\frac{\partial P_m^*}{\partial I} = \frac{p_m^k \Gamma_m^k - \ln(1 + g_{mm}^k \Gamma_m^k)}{g_{mm}^k \ln(1 + g_{mm}^k \Gamma_m^k)}. \tag{6.15}$$

We can find that $\frac{\partial P_m^*}{\partial I} > 0$ for all $p_m^k \Gamma_m^k > 0$, as $x - \ln(1 + x) > 0$ for all $x > 0$. Thus P_m^* is a strictly increasing function of I. monotonicity is proved.

Furthermore, we can evaluate the second order derivative of ∂P_m^* as

$$\frac{\partial^2 P_m^*}{\partial I^2} = -\frac{[(1 + p_m^k \Gamma_m^k) \ln(1 + p_m^k \Gamma_m^k) - p_m^k \Gamma_m^k]^2}{{g_{mm}^k}^2 (1 + p_m^k \Gamma_m^k)[\ln(1 + p_m^k \Gamma_m^k)]^3}. \tag{6.16}$$

$\frac{\partial^2 P_m^*}{\partial I^2} < 0$, thus P_m^* is a strictly concave function of I. As interference I is a linear combination of $\mathbf{P}_{-m}$, then P_m^* is strictly concave in $\mathbf{P}_{-m}$. Concavity is proved.

Last we prove scalability. We define a function $G(\beta) = \beta \mathbf{P}_m^*(\mathbf{P}_{-m}) - \mathbf{P}_m^*(\beta \mathbf{P}_{-m})$. Then $G(1) = 0$ and $G^{''}(\beta) > 0$. It has been proved that the $\mathbf{P}_m^*(\mathbf{P}_{-m})$ is always positive and concave. We have $G^{'}(\beta)|_{\beta=1} = \mathbf{P}_m^*(\mathbf{P}_{-m}) - \mathbf{P}_{-m} \mathbf{P}_m^{'*}(\mathbf{P}_{-m}) > 0$. Thus $G(\beta) > 0$ for all $\beta > 1$. Scalability is proved.

The best strategy function $\mathbf{P}_m^*$ is a standard function, thus guarantees the uniqueness of Nash equilibrium [56].

For the DUE case, multiple RBs may be employed for power allocation and transmission, a similar approach can be applied for proof. We will not detail the proof here. □

Next, we investigate the impact of channel CINR on power and rate distribution when this energy efficient game arrives at Nash equilibrium. Above proposition suggests that the power allocation on each RB either has the same power increase rate $\frac{\partial P}{\partial r_n^k}|_{r_n^{*k}}$ for $k = 1 \cdots K$ or has value zero. Thus for DUE n, we have

$$U_n^{\max} = \frac{\zeta(p_n^1 + p_n^2 + \cdots + p_n^K) + P_s}{(r_n^1 + r_n^2 + \cdots + r_n^K)\zeta}$$
$$= \frac{2^{r_n^1/w} \ln 2}{\Gamma_n^1 w} = \frac{2^{r_n^2/w} \ln 2}{\Gamma_n^2 w} = \cdots \frac{2^{r_n^K/w} \ln 2}{\Gamma_n^K w}. \tag{6.17}$$

We rank the channels $1, \cdots, K$ in the increasing order of CINR Γ_n^k, and assume $p_n^1 > 0$, $r_n^1 > 0$. By incorporating eq. (6.4) and eq. (6.9), power and rate distribution on any two channels k_1 and k_2 is related as

$$r_n^{k_1} = r_n^{k_2} - w(\log_2 \Gamma_n^{k_2} - \log_2 \Gamma_n^{k_1}), \tag{6.18}$$

$$p_n^{k_1} = p_n^{k_2} - (\frac{1}{\Gamma_n^{k_1}} - \frac{1}{\Gamma_n^{k_2}}). \tag{6.19}$$

From the above expression, for DUEs, we can observe that the optimal power and rate allocations on each RB are directly related to the channel CINR. For example, when $\Gamma_n^{k_2} > \Gamma_n^{k_1}$, we will have $r_n^{k_2} > r_n^{k_1}$ and $p_n^{k_2} < p_n^{k_1}$ at the Nash equilibrium.

6.2.3 A Special Case: D2D Power Allocation in Downlink Underlay Mode

To further study the properties of the game model, we investigate DUE downlink underlay mode. We continue to assume the system operate in flat-fading scenario. In this case, DUEs still work under the non-cooperative game framework to compete for RBs. BSs transmit at a fixed power on each RB. This is aligned with the understanding that LTE uses rate control instead of power control on the downlink. This study aims to discover the relationship among D2D energy efficiency, BS transmit power, and CINR. We can prove in the same way as in the previous section that there exists a unique Nash equilibrium for DUEs in this case.

With flat fading channels on the downlink, CUEs experience the same channel gain and also the same co-channel interference on all RBs. According to the analysis in the previous section, each DUE would choose to allocate its power uniformly on all RBs to achieve the

maximum utility. Power allocation and transmission rate will be the same on all K RBs. Thus we have $P = K * p_n^k$, $R = K * r_n^k$, $\Gamma_n^1 = \Gamma_n^2 = \cdots \Gamma_n^k = \cdots \Gamma_n^K$ and substitute this into equation (6.11), we have

$$\frac{\zeta p_n^{k*} + P_s}{w \log_2(1 + p_n^{k*}\Gamma_n^k)} = \frac{\zeta(1 + p_n^{k*}\Gamma_n^k)\ln 2}{\Gamma_n^k w}. \tag{6.20}$$

Denote the receiving SINR as $\gamma_n^* = p_n^{k*}\Gamma_n^k$,

$$\zeta\gamma_n^* + \frac{P_s\Gamma_n^k}{K} = \zeta(1 + \gamma_n^*)\ln(1 + \gamma_n^*). \tag{6.21}$$

Remark: as a function of γ_n^*, the left side of the above equation is a linear function with a slope of ζ, which intersects the vertical axis at $\frac{P_s\Gamma_n^k}{K} > 0$, while the right side is a convex function intersecting the vertical axis at 0. Hence there exists only one solution of $\gamma_n^* > 0$. Obviously, $\gamma_n^*(\Gamma_n^k)$ is a strictly increasing function of Γ_n^k. Expressing Γ_n^k in γ_n^*, we have

$$\Gamma_n^k = \frac{K\zeta[(1 + \gamma_n^*)\ln(1 + \gamma_n^*) - \gamma_n^*]}{P_s}, \tag{6.22}$$

and

$$\frac{\zeta P + P_s}{R\zeta} = \frac{\ln 2 P_s}{Kw} \cdot \frac{1}{\ln(1 + \gamma_n^*) + \frac{1}{1+\gamma_n^*} - 1}. \tag{6.23}$$

From the above equations, we can derive the relationship between user energy efficiency U_n and CINR Γ_n^k. First, we consider the properties of function $g(\gamma^*) = \ln(1 + \gamma^*) + \frac{1}{1+\gamma^*}$, whose first order derivative $g'(\gamma^*) = \frac{\gamma^*}{(1+\gamma^*)^2} > 0$. Left side of the above equation $\frac{\zeta P + P_s}{R\zeta}$ is strictly decreasing with the increase of γ_n^*. From equation (6.21), we can see that energy efficiency function $U_n(\gamma^*) = \frac{R}{\zeta P + \epsilon R + P_s} = \frac{1}{\frac{\zeta P + P_s}{R} + \epsilon}$ is a strictly increasing function of $\gamma_n^*(\Gamma_n^k)$ and also of Γ_n^k.

Similarly, we can express optimal power allocation p_n^{k*} as a function of γ_n^*,

$$p_n^{k*}(\gamma_n^*) = \frac{\gamma_n^* P_s}{K\zeta[(1+\gamma_n^*)\ln(1+\gamma_n^*) - \gamma_n^*]}. \tag{6.24}$$

We find that the optimal power allocation $p_n^{k*}(\gamma_n^*)$ is a strictly decreasing function of γ_n^*, since the first order derivation of $p_n^{k*}(\gamma_n^*)$ is $\frac{dp_n^{k*}(\gamma_n^*)}{d\gamma_n^*} = \frac{P_s[\ln(1+\gamma_n^*)-\gamma_n^*]}{K\zeta[(1+\gamma_n^*)\ln(1+\gamma_n^*)-\gamma_n^*]^2} < 0$ for all $\gamma_n^* > 0$. Using $\bar{\gamma_n} = \ln(1+\gamma_n^*)$ and with some algebraic operations, we can re-rewrite equation (6.21) into

$$\exp(\bar{\gamma_n} - 1)(\bar{\gamma_n} - 1) = \frac{P_s\Gamma_n^k - K\zeta}{eK\zeta}. \tag{6.25}$$

Observing that this is the form of Lambert-W function $W(x)$ [57], which is the solution of $W(x)\exp(W(x)) = x$, we can write equation (6.25) into

$$W(\frac{P_s\Gamma_n^k - K\zeta}{eK\zeta}) = \bar{\gamma_n} - 1, \tag{6.26}$$

which can be solved and gives $\gamma_n^* = \exp[W(\frac{P_s\Gamma_n^k - K\zeta}{eK\zeta}) + 1] - 1$. The optimal power allocation on each channel $\min(P_{dd}/K, \frac{\gamma_n^*}{\Gamma_n^k})$.

From the above analysis, we can draw the following conclusions. In Nash equilibrium, each DUE achieves its own maximum energy efficiency by allocating its power uniformly on each RBs. The UE with a higher channel gain would have a higher energy efficiency and a higher transmission rate.

6.3 Performance Study

In this section, we simulate the proposed power and resource allocation approach. UEs are uniformly deployed in a circular cell of radius 500 meters. DUEs are dropped in a pair with a maximum pair distance R_d, which is set to 40 meters. The adopted path loss model is $PL(d) = PL(d_0) + 10\alpha\log(d/d_0)$ [dB], where $PL(d_0)$ is the path loss at the reference distance d_0 and α is the path loss exponent, which is set to 4 in our simulation. Each RB has a 15 kHz bandwidth. We assume the number of RBs is equal to the number of CUEs

in the simulation. The compensation factor for power amplifier loss ζ is set to be 30%, the power consumption per unit rate is set to 2mW/Kbps, and the static background power consumption P_s is set to 10dBm. The total number of DUEs is 40 and the number of CUEs varies from 10 to 50.

With flat-fading channels, there exists a unique Nash equilibrium for both uplink and downlink underlay modes. Fig. 6.2 shows the optimal transmission power and energy efficiency with respect to CINR. Obviously, as CINR increases, the optimal transmit power decreases and the corresponding energy efficiency increases. The figure also shows how the number of RBs can affect the optimal transmit power and energy efficiency, which is consistent with our previous analysis in section III. With more RBs, the users tend to allocate less power on each RB while keeping the energy efficiency unchanged. The reason for this is because UE always has a higher energy efficiency at a lower transmit power and rate when the static circuit power consumption is fixed. The Nash equilibrium condition (2) can exactly explain this point. Actually, increasing the number of available RBs has the same effect as decreasing the static circuit power P_s, as shown in equation (6.25).

We next investigate the optimal transmit power and the achievable rate at Nash equilibrium. To study how interference from CUE impacts DUEs, we assume a DUE experiences the same interference on all its RBs, which is the case in the downlink underlay mode, where BS transmits to all the CUEs using the same power, thus generating the same interference to a DUE on different RBs. We denote P_{inf} as the interference to DUE at a reference distance of half cell radius. Fig. 6.3 and Fig. 6.4 show the data rate and the transmit power distribution of CUEs at the Nash equilibrium. Obviously, with the increase of interference power, DUEs have to increase transmit power at Nash equilibrium, but the transmit rate still decreases. This is because higher interference from CUE offsets the DUE power increase and thus decreases DUE's CINR. By comparing Fig. 6.3 and Fig. 6.4, we can find that DUEs with higher data rates are also the ones with lower transmit powers in equilibrium, which means the higher data rate users have better channels (or higher CINRs).

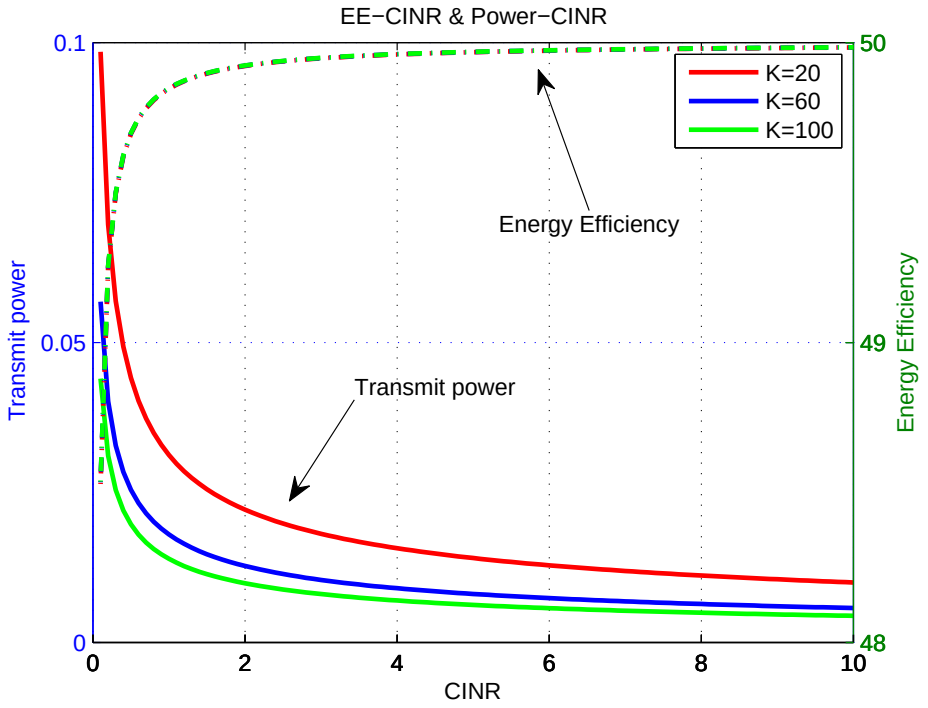

Fig. 6.2: Optimal power allocation and energy efficiency with respect to CINR

6.4 Summary

In this work, we have studied an game theory based energy efficient resource allocation for D2D communication underlaying cellular networks. In our non-cooperative game model, each UE acts selfishly to maximize its own utility function, which is defined as the achievable rate per unit power. The existence and uniqueness of Nash equilibrium were shown for both uplink and downlink underlay modes. We also make simulation of power allocation for proposed game model to find Nash equilibrium. Both theoretical analysis and simulation results show that UEs always allocate less power to RBs with higher CINR to achieve the maximum energy efficiency.

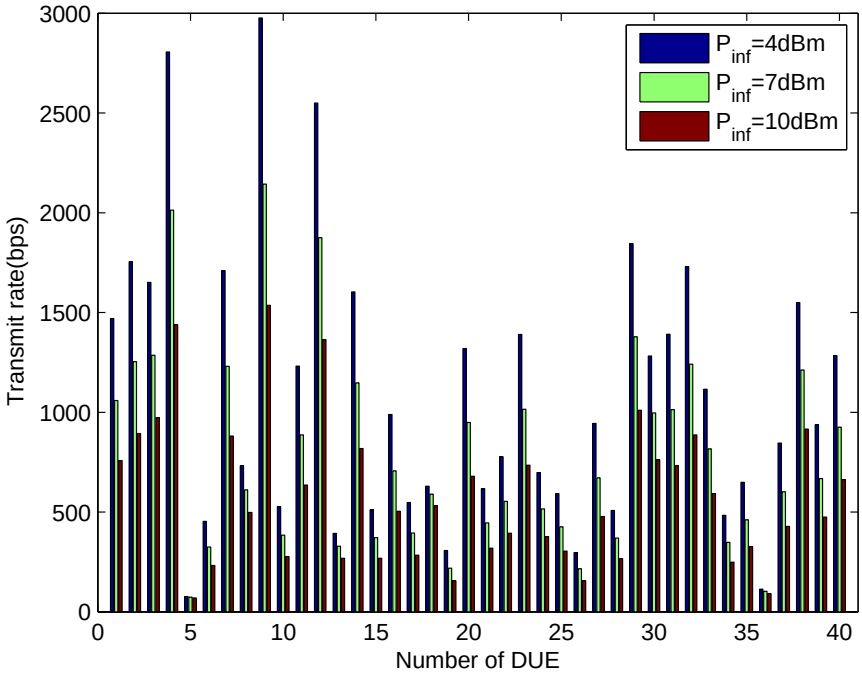

Fig. 6.3: Transmit rate distribution of DUE

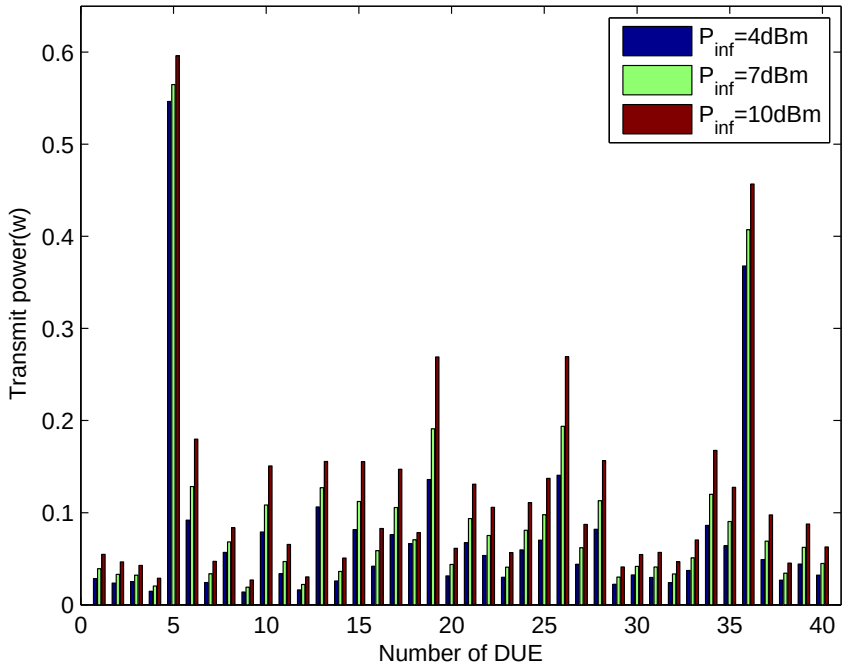

Fig. 6.4: Power allocation distribution of DUEs

Chapter 7

Tradeoff Between Energy Efficiency and Spectral Efficiency in a Delay Constrained Wireless System

In this chapter, we study the fundamental tradeoff between Energy Efficiency (EE) and Spectral Efficiency (SE) in presence of statistical QoS requirement in the wireless transmission system. Performance with QoS requirement for wireless transmission can be measured through effective capacity, which is introduced to model the physical layer fading channel with link layer parameters, such as delay and data rate [5]. Thus SE is defined as effective capacity per unit bandwidth under QoS requirement, and EE is energy per effective capacity bit. Both the circuit power and transmission power are considered in the energy model, based on which we derive the quasi-convex generalized EE formulation. To further exploit the tradeoff between EE and SE, we propose a generic close-form approximation for EE-SE formulation by employing a curve fitting approach. The impacts of QoS and circuit power consumption on EE-SE tradeoff are analyzed. QoS requirement and circuit power consumption affect the EE-SE tradeoff differently. In the low SNR regime, circuit power shows more impact on the EE-SE tradeoff while QoS impacts EE-SE tradeoff more in the high SNR regime.

7.1 Effective Capacity in a Wireless Channel

A transmission in a wireless channel can usually be formulated as

$$y = \sqrt{p}hx + n, \tag{7.1}$$

where p is the transmission power, n denotes the additive white Gaussian noise (AWGN), and h is the channel gain, including path loss, shadowing and short term fading. For exam-

ple, Rayleigh fading model is utilized to capture the fading in a dense urban environment when there is no line of sight signal between the receiver and the transmitter [58]. In this work, we characterize the channel statistics by the square amplitude of coefficient $z = |h|^2$ with probability density function $p(z)$.

7.1.1 Shannon Capacity

Shannon capacity provides the theoretic upper bound of the information rate that can be supported in a wireless channel at a given SNR.

$$C = W\mathrm{E}\{\log_2(1 + \rho|h|^2)\}, \tag{7.2}$$

where W denotes the channel bandwidth, $\rho = \frac{P_t}{N_0 W}$ is the transmission SNR, and N_0 is the power spectral density of noise.

7.1.2 Effective Capacity

Shannon capacity provides insights into the maximum capacity of a wireless channel subject to noise and interference. But due to the statistical fluctuations of the wireless channel, the actual data rate served will be smaller than the Shannon capacity when link layer queuing delay is considered. Effective capacity incorporates the link layer QoS requirements and captures the decay rate of buffer occupancy violation at the large buffer size region. It can be characterized by triple parameters: source data rate r_s, delay bound D_{max}, delay-violation probability ϵ. D_{max} and ϵ satisfy:

$$P_r\{D(\infty) \geq D_{max}\} \leq \epsilon. \tag{7.3}$$

$D(\infty)$ is the steady-state queuing delay experienced by the traffic flow at the link layer. In large deviation theory, the decay rate of queue occupancy probability with a queue of large enough buffer size which is infused with a constant source data rate can be approximated as

$$P_r(Q \geq q_{max}) \approx e^{-\theta q_{max}}. \tag{7.4}$$

Q is the steady state queue length at the transmitter and $q_{max} = r_s D_{max}$. θ is the QoS exponent characterizing the delay constraints. A larger value of θ denotes a more stringent QoS requirement and delay constraint. When $\theta = 0$, there is no delay constraint and the source can bear unlimited delay. According to the theorem proposed in [59], relationship between queue size and delay violation probability is established through $P_r\{D(\infty) \geq D_{max}\} \leq m\sqrt{P_r\{Q \geq q_{max}\}}$ under large queue length assumption, where m is a positive constant. In a block fading channel, where the channel fading process $h(t)$ is constant during time T, effective capacity $R(\theta)$ is defined as a log-moment generation function [60]:

$$R(\theta) = -\frac{1}{\theta T}\log_e\{\mathrm{E}\{e^{-\theta T C(t)}\}\}, \tag{7.5}$$

where $C(t)$ is the instantaneous wireless channel Shannon capacity. Here by re-defining the system SE based on effective capacity that incorporates statistical QoS requirement, we substitute $C(t) = W\log_2(1+\rho|h(t)|^2)$ into the above formula. The SE formulation based on effective capacity is expressed as:

$$\eta_{\mathrm{SE}}(\theta,\rho) = \frac{R(\theta)}{W} = -\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h(t)|^2)^{-A}\})\mathrm{bit/s/Hz}, \tag{7.6}$$

$A = \theta TW/\ln 2$. According to this definition, ergodic Shannon capacity C can be considered as the special case of $R(\theta)$ with QoS exponent $\theta = 0$, which can be verified in the following. Applying L'Hosptital Rule to equation (7.6), we have

$$\begin{aligned}
\mathrm{R}(\theta = 0) &= \lim_{A\to 0} -\frac{W}{A}\log_2(\mathrm{E}\{(1+\rho|h(t)|^2)^{-A}\}) \\
&= \lim_{A\to 0} -\log_2 e \frac{W\mathrm{E}\{\ln(1+\rho|h(t)|^2)^{-1}(1+\rho|h|^2)^{-A}\}}{\mathrm{E}(1+\rho|h(t)|^2)^{-A}} \\
&= W\mathrm{E}\{\log_2(1+\rho|h(t)|^2)\} \\
&= C.
\end{aligned} \tag{7.7}$$

7.1.3 Special Case: Effective Capacity Based SE in a Rayleigh Fading Channel

For a Rayleigh fading channel, square of the channel amplitude $x = |h|^2$ obeys an

exponential distribution with a probability density function $p(x) = \frac{1}{\Omega}e^{-\frac{x}{\Omega}}$, with parameter Ω denoting the average channel gain. Substitute $p(x)$ into formula (7.6), the Rayleigh channel's SE can be calculated as

$$\begin{aligned}
\eta_{\mathrm{SE}}(\theta,\rho) &= -\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\}) \\
&= -\frac{1}{A}\log_2\int_0^{\infty}(1+\rho x)^{-A}\frac{1}{\Omega}e^{-\frac{x}{\Omega}}dx \\
&= -\frac{1}{A}\log_2\{(\rho\Omega)^{-A}e^{\frac{1}{\rho\Omega}}\int_{\frac{1}{\rho\Omega}}^{\infty}\frac{1}{x^A e^x}dx\} \\
&= \log_2(\rho\Omega) - \frac{\log_2 e}{A\rho\Omega} - \frac{1}{A}\log_2\Gamma(1-A,\frac{1}{\rho\Omega}).
\end{aligned} \tag{7.8}$$

$\Gamma(p,q) = \int_q^{\infty}\frac{1}{x^{1-p}e^x}dx$ is the upper incomplete gamma function. When considering the special case with no QoS requirement, i.e., $\theta = 0$, SE can be written as

$$\begin{aligned}
\eta_{SE}(0,\rho) &= \mathrm{E}\{\log_2(1+\rho|h|^2)\} \\
&= \int_0^{\infty}\log_2(1+\rho x)\frac{1}{\Omega}e^{-\frac{x}{\Omega}}dx \\
&= \log_2\mathrm{e}\cdot(e^{\frac{1}{\rho\Omega}})\int_{\frac{1}{\rho\Omega}}^{\infty}\frac{1}{xe^x}dx \\
&= \log_2\mathrm{e}\cdot(e^{\frac{1}{\rho\Omega}})\Gamma(0,\frac{1}{\rho\Omega}).
\end{aligned} \tag{7.9}$$

SE at High SNR

The impact of QoS constraints on performance in the high SNR region can be captured by two measurements, high-SNR slope S_∞ and power offset L_∞, which are defined in [61] as

$$S_\infty = \lim_{\rho\to\infty}\frac{\eta_{\mathrm{SE}}(\theta,\rho)}{\log_2\rho}, \quad L_\infty = \lim_{\rho\to\infty}(\log_2\rho - \frac{\eta_{\mathrm{SE}}(\theta,\rho)}{S_\infty}). \tag{7.10}$$

Thus at high SNR, system SE can be approximated as

$$\eta_{\mathrm{SE}}(\theta,\rho) = S_\infty(\log_2\rho - L_\infty) + o(1). \tag{7.11}$$

where $o(1)$ denotes a finite constant value when SNR approaches infinity. Substituting the SE expression (7.8) into formula (7.10),we derive that

$$S_\infty = \begin{cases} 1 & 0 < A \leq 1, \\ \frac{1}{A} & A > 1. \end{cases} \tag{7.12}$$

and

$$L_\infty = \begin{cases} \log_2 \frac{\Gamma(1-A)^{\frac{1}{A}}}{\Omega}, & 0 < A \leq 1, \\ \log_2 \frac{\Gamma(A-1)}{\Gamma(A)\Omega}, & A > 1. \end{cases} \tag{7.13}$$

The approximated SE at high-SNR can be expressed in the work as:

$$\eta_{\mathrm{SE}}(\theta, \rho) \approx \begin{cases} \log_2 \frac{\rho\Omega}{\Gamma(1-A)^{\frac{1}{A}}} + o(1), & 0 < A \leq 1, \\ \frac{1}{A}\log_2 \frac{\rho\Omega\Gamma(A)}{\Gamma(A-1)} + o(1), & A > 1. \end{cases} \tag{7.14}$$

Proof. See Appendix. □

For the case when $\theta = 0$ or $A = 0$, SE approximation in the high-SNR regime can be expressed as $\eta_{\mathrm{SE}}(0, \rho) = \log_2(\rho\Omega) - \gamma\log_2 e + o(1)$. The high-SNR slope S_∞ is calculated as

$$\begin{aligned} S_\infty|_{\theta=0} &= \lim_{\rho\to\infty} \frac{\log_2 \mathrm{e}(e^{\frac{1}{\rho\Omega}})\Gamma(0, \frac{1}{\rho\Omega})}{\log_2 \rho} \\ &= -\log_2 e \lim_{z\to 0} \frac{e^z \int_z^\infty \frac{1}{xe^x} dx}{\log_2 z + \log_2 \Omega} \\ &= -\log_2 e \lim_{z\to 0} \frac{\frac{-1}{ze^z}}{\frac{1}{z}\log_2 e} = 1. \end{aligned} \tag{7.15}$$

In the above expression, we replace $\frac{1}{\rho\Omega}$ with z in the second equation. The power offset L_∞ is equal to

$$\begin{aligned} L_\infty|_{\theta=0} &= \frac{1}{A}\log_2 \mathrm{E}\{(|h|^2)^{-A}\} \\ &= \lim_{A\to 0} \log_2 \frac{[\int_0^\infty \frac{1}{x^A e^x} dx]^{\frac{1}{A}}}{\Omega} \\ &= \lim_{A\to 0} \log_2 \frac{e^{\frac{\ln \int_0^\infty \frac{1}{x^A e^x} dx}{A}}}{\Omega} \\ &= \log_2 \frac{e^\gamma}{\Omega}, \end{aligned} \tag{7.16}$$

where $\gamma = -\int_0^\infty \ln x e^{-x} dx \approx 0.5772156649$ is Euler-Mascheroni constant [62].

SE at low SNR

At low-SNR, we can approximate system SE with the second order Taylor expansion as

$$\eta_{\mathrm{SE}}(\theta,\rho) = \dot{\eta_{\mathrm{SE}}}(\theta,\rho)\rho + \ddot{\eta_{\mathrm{SE}}}(\theta,\rho)\frac{\rho^2}{2} + o(\rho^2), \tag{7.17}$$

where

$$\begin{aligned} \dot{\eta_{\mathrm{SE}}}(\theta,\rho) &= \log_2 e \frac{\mathrm{E}\{(1+\rho|h|^2)^{(-A-1)}|h|^2\}}{\mathrm{E}\{(1+\rho|h|^2)^{-A}\}}|_{\rho=0} \\ &= \log_2 e \mathrm{E}\{|h|^2\} \\ &= \log_2 e \Omega, \end{aligned} \tag{7.18}$$

and

$$\begin{aligned} \ddot{\eta_{\mathrm{SE}}}(\theta,\rho) &= \log_2 e \{A(\frac{\mathrm{E}\{(1+\rho|h|^2)^{-A-1}|h|^2\}}{\mathrm{E}^2\{(1+\rho|h|^2)^{-A}})^2 \\ &- (A+1)\frac{\mathrm{E}\{(1+\rho|h|^2)^{-A-2}|h|^4\}}{\mathrm{E}\{(1+\rho|h|^2)^{-A}\}}\}|_{\rho=0} \\ &= \log_2 e (A\mathrm{E}^2\{|h|^2\} - (A+1)\mathrm{E}\{|h|^4\}) \\ &= -\log_2 e (A+2)\Omega^2. \end{aligned} \tag{7.19}$$

7.1.4 Numerical Analysis and Discussion

Throughout the work, we set channel block length to $T = 1$ms and system bandwidth to $W = 1$MHz in the numerical study. In Figures 7.1 and 7.2, we have plotted the approximation curves for the SE with different QoS requirements in a Rayleigh fading channel in the low-SNR and high-SNR regimes. For the comparison purpose, we also plot the accurate expressions correspondingly. As illustrated in the graphs, in the low-SNR region below -10dB and high-SNR region above 20dB, approximate expressions (7.11) and (7.17) accurately represent the system EE under statistical QoS requirement.

Note that when $A = 0$ or equivalently $\theta = 0$, there is no delay constraint and the corresponding effective capacity is equal to the ergodic Shannon capacity. As expected, QoS constraint decreases the SE, which means a lower data rate can be supported at the physical link layer when data link layer queuing delay is considered. As observed in Figure 7.1, in the low-SNR regime, SE is better approximated by equation (7.17) when A is smaller or QoS is less demanding. But in the high-SNR regime, a better approximation can be achieved using equation (7.11) with a more stringent QoS, as shown in Figure 7.2. Besides, in the low-SNR regime, all the curves have the same slope at zero SNR $\dot{\eta_{\text{SE}}}(\theta, 0) = \log_2 e \text{E}\{|h|^2\}$ regardless of the QoS parameter. This could be explained in the following way. Because in low SNR, low transmission rate causes diminishing arrival rate supported by the channel, hence buffer violation probability is also decreased no matter of the QoS. However, a strict QoS always decreases the increase rate of system SE especially in the low SNR regime as illustrated in the figure and formula (7.19). In the High-SNR regime, curve slope S_∞ inversely increases with QoS parameter A. But QoS parameter has less impact on the increase rate of the curve. Obviously, system transmission rate can always be satisfied when the transmission SNR is high enough.

7.2 Optimal Energy Efficiency with QoS

In this section we study the EE of a wireless channel with QoS consideration. Here we define EE or η_{EE} as power consumption per information bit transmission under a link layer delay constraint. In our model, energy consumption includes both transmission power P_t

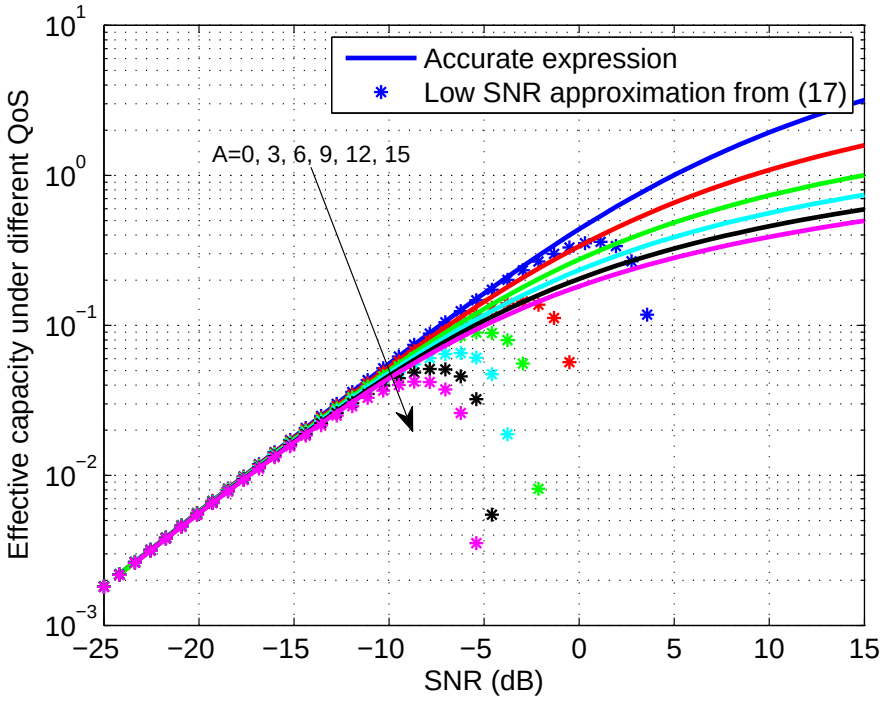

Fig. 7.1: SE approximation in low SNR

and the circuit power consumption P_c. The transmission power can be adapted in order to achieve a desirable SNR and thus a target transmission rate. Without loss of generality, we normalize E_b with N_0 for simplification. Hence, the system EE function based on effective capacity can be expressed as

$$\eta_{EE}(\theta,\rho)=\frac{P_t+P_c}{\eta_{\mathrm{SE}}(\theta,\rho)N_0W}=\frac{\rho+\frac{P_c}{N_0W}}{-\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\})}, \tag{7.20}$$

where A has been defined earlier as $A=\theta TW/\ln 2$.

Proposition 7.2.1. *Energy efficiency function defined in (7.20) is a strictly quasi-convex function with respect to transmission SNR ρ.*

Proof. We first prove the denominator in (7.20), which denotes the system SE, is a concave function of ρ. Define function $g(\rho)=(1+\rho z)^{-A}$.

$$\dot{g}=-A(1+\rho z)^{-A-1}z,\quad \ddot{g}=A(A+1)(1+\rho z)^{-A-2}z^2>0.$$

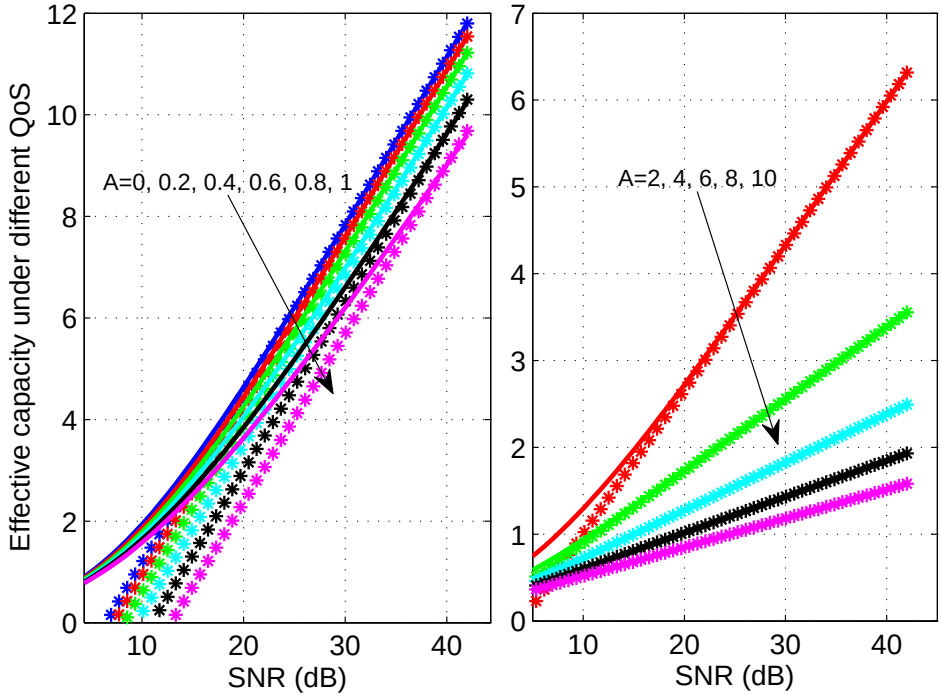

Fig. 7.2: SE approximation in high SNR

$\dot{g}$ and $\ddot{g}$ denote the first and second order derivation of $g(\rho)$ with respect to ρ. Since $z = |h|^2 > 0$, then $\ddot{g} > 0$. So $g(\rho)$ is a strictly convex function of ρ. $\mathrm{E}\{(1+\rho z)^{-A}\} = \int (1+\rho z)^{-A} p(z) dz$ is also a convex function of ρ because $\int f(z,v)dv$ is convex if $f(z,v)$ is convex with respect to z for each v in the domain [50]. As function $-\frac{1}{A}\log_2(z)$ is a strictly decreasing function of z, $\eta_{SE}(\theta,\rho) = -\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\})$ is a concave function of ρ.

We further prove in the following η_{EE} is a strictly quasi-convex function. If sub-level set $\eta_{EE} = \{\rho | \frac{\rho + \frac{P_c}{N_0 W}}{-\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\})} \leq \alpha, \rho \geq 0\}$ is a convex set for any α, then $\eta_{EE}(\rho)$ is a quasi-convex function with respect to ρ.

- when $\alpha \leq 0$, there is no feasible ρ.
- when $\alpha > 0$, according to the conclusion derived above, $\frac{\alpha}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\}) + \rho + \frac{P_c}{N_0 W}$ is a convex function of ρ, so set $\{\rho | \frac{\alpha}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\}) + \rho + \frac{P_c}{N_0 W} \leq 0, \rho \geq 0\}$ is also convex. Therefore, $\{\rho | \frac{\rho + \frac{P_c}{N_0 W}}{-\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\})} \leq \alpha, \rho \geq 0\}$ is a convex set.

So the energy efficiency function is a strictly quasi-convex function. □

Proposition 7.2.2. *There exists a unique minimum value for the energy efficiency function. The optimal SNR is achieved at $\rho_0 > 0$, with $\dot{\eta}_{EE}|_{\rho=\rho_0} = 0$. $\dot{\eta}_{EE}$ denotes the first order derivation of η_{EE} with respect to ρ.*

Proof. As a first step, we check the asymptotic bounds for η_{EE}.

- When $\rho \to 0$, $\eta_{\mathrm{SE}}(\theta, \rho) \to 0$, so $\eta_{EE}|_{\rho=0} \to +\infty$;
- When $\rho \to +\infty$, as

$$\begin{aligned} \eta_{EE} &= \frac{\rho + \frac{P_c}{N_0 W}}{-\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\})} \\ &\geq \frac{\rho + \frac{P_c}{N_0 W}}{\frac{1}{A}\mathrm{E}\{-\log_2((1+\rho|h|^2)^{-A})\}} \\ &= \frac{\rho + \frac{P_c}{N_0 W}}{\mathrm{E}\{\log_2((1+\rho|h|^2))\}} \\ &\geq \frac{\rho + \frac{P_c}{N_0 W}}{\log_2(1+\rho\mathrm{E}\{|h|^2\})}. \end{aligned} \tag{7.21}$$

The two inequalities in the above formula are derived from Jensen's inequality by applying it to the concave function $\log_2(\cdot)$ twice. Obviously, the last expression in (7.21) approaches $+\infty$ when $\rho \to +\infty$, thus energy efficiency function η_{EE} also approaches $+\infty$. Therefore, the optimal SNR for minimum energy efficiency function must be achieved between 0 and $+\infty$ if it exists.

Next, we prove that the optimal SNR exists and is uniquely achieved at $\dot{\eta}_{EE}|_{\rho=\rho_0} = 0$. The first order derivation for energy efficiency function can be expressed as

$$\dot{\eta}_{EE} = \frac{\eta_{\mathrm{SE}}(\theta,\rho) - (\rho + \frac{P_c}{N_0 W})\dot{\eta}_{\mathrm{SE}}(\theta,\rho)}{\eta^2_{\mathrm{SE}}(\theta,\rho)} = \frac{f(\rho)}{\eta^2_{\mathrm{SE}}(\theta,\rho)}, \tag{7.22}$$

where function $f(\rho)$ is defined as $f(\rho) = \eta_{\mathrm{SE}}(\theta) - (\rho + \frac{P_c}{N_0 W})\dot{\eta_{\mathrm{SE}}}(\theta,\rho)$ and

$$\dot{\eta}_{\mathrm{SE}}(\theta) = \log_2 e \frac{\mathrm{E}\{(1+\rho|h|^2)^{(-A-1)}|h|^2\}}{\mathrm{E}\{(1+\rho|h|^2)^{-A}\}}. \tag{7.23}$$

Further calculation shows

- when $\rho = 0$, $f(\rho) = -\log_2 e \frac{P_c}{N_0 W} \mathrm{E}\{|h|^2\} < 0$;
- when $\rho = +\infty$, $f(\rho) = -\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\}) - \log_2 e \to +\infty$.

The derivation of function $f(\rho)$ can be expressed as

$$\dot{f}(\rho) = -(\rho + \frac{P_c}{N_0 W})\ddot{\eta}_{\mathrm{SE}}(\theta, \rho) \geq 0. \tag{7.24}$$

$\dot{f}(\rho)$ is always positive since we have proved that $\eta_{\mathrm{SE}}(\theta, \rho)$ is a concave function of ρ and thus $\ddot{\eta}_{\mathrm{SE}}(\theta, \rho) \leq 0$. $f(\rho)$ is an increasing function of ρ. Therefore, there exists an $\rho_0 > 0$, for which $f(\rho) < 0$ when $\rho < \rho_0$ and $f(\rho) > 0$ when $\rho > \rho_0$, and the zero-crossing point is uniquely achieved at ρ_0. From (7.22), we can conclude that $\dot{\eta}_{EE} < 0$ when $\rho < \rho_0$ and $\dot{\eta}_{EE} > 0$ when $\rho > \rho_0$. That means η_{EE} first monotonically decreases when $\rho < \rho_0$ and then monotonically increases when $\rho > \rho_0$. Therefore, minimum value of energy efficiency function is unique and is only achieved at ρ_o, where $\dot{\eta}_{EE}|\rho_0 = 0$. □

7.3 Binary Search for Optimal ρ_0

By setting $\dot{\eta}_{EE}|_{\rho=\rho_0} = 0$, we get:

$$-\frac{1}{A}\ln(\mathrm{E}\{(1+\rho|h|^2)^{-A}\}) = (\rho + \frac{P_c}{N_0 W})\frac{\mathrm{E}\{(1+\rho|h|^2)^{(-A-1)}|h|^2\}}{\mathrm{E}\{(1+\rho|h|^2)^{-A}\}}. \tag{7.25}$$

The close-form expression for optimal ρ_0 is very difficult to derive from above formulation. We propose a two-step binary search algorithm to find the optimal solution for the quasi-convex function η_{EE}. For the optimal energy efficiency problem, we introduce a new variable t and convert the quasi-convex optimization problem into a convex problem, whose optimal solution can be searched using the two step binary search algorithm. The primary quasi-convex optimization problem is

$$\min \quad \eta_{EE} = \frac{\rho + \frac{P_c}{N_0 W}}{-\frac{1}{A}\log_2(\mathrm{E}\{(1+\rho\| \ h \ \|^2)^{-A}\})} \tag{7.26}$$

subject to

$$\rho \geq 0.$$

By letting $t = \eta_{EE}$, the primary optimization problem can be reformulated as

$$\min \quad t$$

subject to

$$\rho + \frac{t}{A}\log_2(\mathrm{E}\{(1+\rho\| \ h \ \|^2)^{-A}\}) + \frac{P_c}{N_0 W} \leq 0,$$
$$\rho \geq 0. \tag{7.27}$$

The binary search algorithm is detailed in Table 7.1.

7.3.1 Numerical Results and Discussion

In the conventional analytical results, the best EE is achieved at zero SNR when no delay or circuit power is considered [63]. In this section, we numerically given out the optimal operation point for EE and SNR when both circuit power consumption and link layer delay constraint are taken into considerations.

Figure 7.3 plots the EE function with respect to ρ under different delay requirements. EE with and without circuit power consumptions are compared. Curves with QoS parameter $A = 9, 7, 5, 0$ are plotted from top to bottom in the two separate cases. As expected, η_{EE} is a quasi-convex function, which first decreases and then increases with ρ when circuit power is considered. In the case with no circuit power, η_{EE} monotonically increases with ρ and the best η_{EE} is always achieved at $\rho = 0$, i.e., min $\eta_{EE} = \frac{1}{\log_2 e \mathrm{E}\{|h|^2\}}$, irrespective of the QoS requirement. With the increase of A value, EE gets worse in both cases, which means more power is required to achieve a better QoS with the same data rate delivery.

In Figure 7.4, we illustrate the impact of circuit power on EE. Obviously, adding circuit power will lower down system EE. The higher the circuit power, the less the EE will be. Figure 7.5 illustrates the optimal EE under various QoS constraints and circuit power consumptions. Figure 7.6 presents the transmission SNR ρ that achieves the optimal EE under different QoS constraints and circuit power consumptions. Again, as shown in Figure 7.5, both QoS parameter and circuit power will increase the minimum energy needed for per bit transmission. In Figure 7.6, circuit power increases the SNR value that can achieve the best EE while a more strict QoS constraint actually decreases the SNR value in order to achieve the best EE.

7.4 EE-SE Tradeoff with QoS Consideration in Rayleigh Fading Channel

SE, EE and QoS have been consistently considered as the most important metrics for the wireless system performance evaluation. However, the close-form expression between EE with respect to SE or EE-SE tradeoff with statistical QoS provision has not yet been derived due to its complexity [64]. In this section, we would exploit the relationship between EE and SE under delay constraint through a curve fitting method.

Table 7.1: TWO-STEP-BINARY-SEARCH-ALGORITHM FLOW

Step 1: Initialize $t = t1 = t2 = t0 > 0$,
if the solution for constraint (7.27) is feasible **then**
    repeat
        $t = t1 = t1/2$;
        Using gradient descent method to search the minimum value of (7.27)
    until solution for ρ becomes infeasible
    t2=t1*2;
else
    repeat
        $t = t2 = t2 * 2$;
        Using gradient descent method to search the minimum value of (7.27)
    until solution for ρ becomes feasible
    t1=t2/2;
end if
Step2: Initialize $\epsilon = 1e - 4$;
while $|t1 - t2| \geq \epsilon$ **do**
    $t = (t1 + t2)/2$;
    gradient descent search the minimum value of (7.27)
    if solution for ρ is feasible **then**
        t1=t;
    else
        t2=t;
    end if
end while

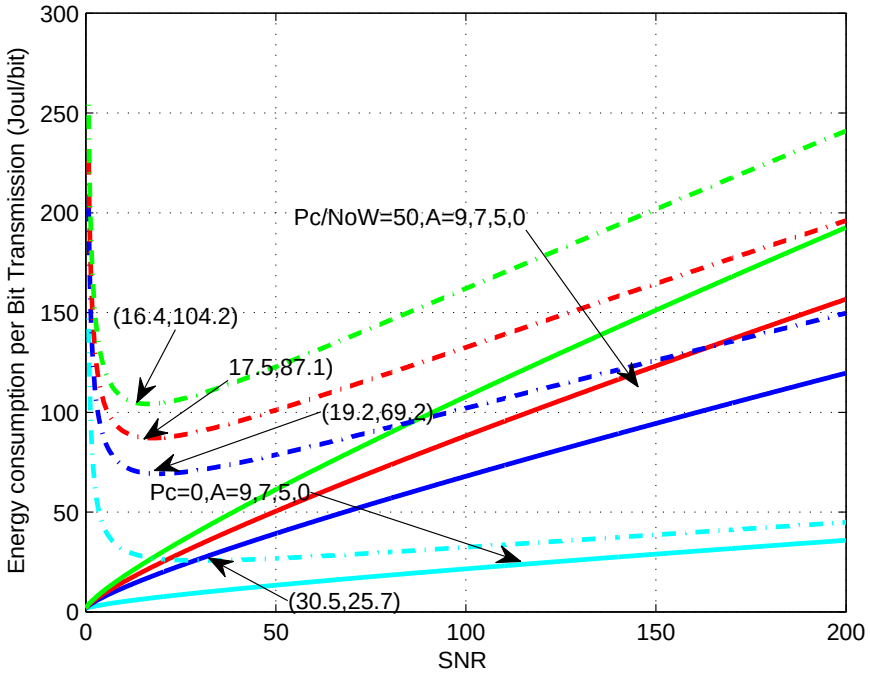

Fig. 7.3: Impact of QoS parameter on energy efficiency

7.4.1 Close-form Approximation for EE-SE Tradeoff

We derive the close-form expression for EE-SE tradeoff over the Rayleigh fading channel through a heuristic curve fitting method. In observation of EE expression in (7.20), if we can express SNR ρ in terms of SE, the EE-SE relationship can then be defined. Inspired from approximation curve in Figures 7.1 and 7.2, we find that inverse function for the item containing ρ in SE function (7.8) can be tightly approximated by using a group of logarithmic barrier functions $z_i(x) = A\log_2\frac{1}{1-x^{m_i}}$. Thus we use the logarithmic barrier function $z_i(x) = A\log_2\frac{1}{1-x^{m_i}}$ as the basis function and express ρ with respect to η_{SE} in a Rayleigh fading channel in the following form.

$$\rho = \frac{-1}{A \cdot \Omega \sum_{i=1} \alpha_i \log_2[1 - (\frac{1}{2})^{m_i A \eta_{SE}}]}, \tag{7.28}$$

where α_i's are the parameters solved from curve fitting. Parameter m_i determines the curvature of the logarithmic barrier function. By inserting the above heuristic ρ expression

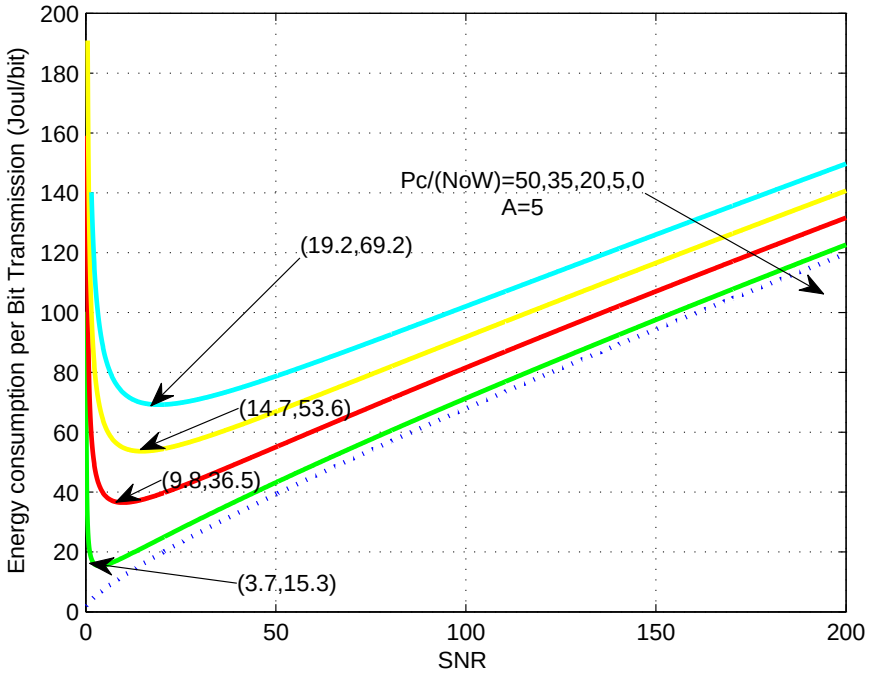

Fig. 7.4: Impact of circuit power on energy efficiency

into equation (7.20), a close-form expression of the EE-SE tradeoff with QoS requirement over Rayleigh fading channel can be formulated as

$$\eta_{EE} = \frac{\frac{P_c}{N_0 W} A \cdot \Omega \sum_i \alpha_i \log_2[1 - (\frac{1}{2})^{A \cdot \eta_{SE} \cdot m_i}] - 1}{\Omega \sum_i \alpha_i \log_2[1 - (\frac{1}{2})^{A \cdot \eta_{SE} \cdot m_i}]^{A \cdot \eta_{SE}}}. \tag{7.29}$$

The basis function $z_i(x)$ used for approximation has the following property: $Dom_x = (0, 1] \rightarrow (0, +\infty)$ and is monotonically increasing over D_x, which just resembles the required property. To determine the parameters α_i, m_i and solve the curve fitting problem, a desirable fitting criterion needs to be established. Usually, curve fitting problems are formulated as minimizing a least square error over a given data set. A linear combination of a group of basis functions are used to approximate the inverse function. And a larger number of basis functions will have a higher accuracy for approximation. A direct way is to solve the problem

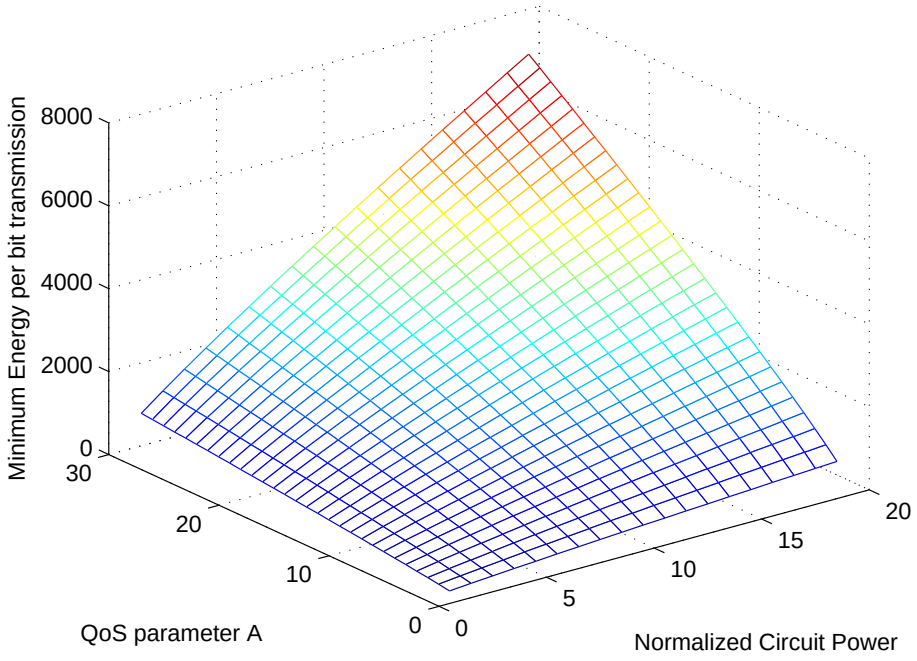

Fig. 7.5: Best energy efficiency under different QoS parameter and circuit power combination

$$\min \quad \sum_{i=1}^{N} (Z(u_i) - \frac{1}{\rho_i})^2, \tag{7.30}$$

where $Z(x) = \sum \alpha_i z_i(x)$ and ρ_i is known from the sample set. N in the sum expression denotes number of samples from the original function. Another way is to find sparse descriptions [50], where fewer basis functions are used to approximate the objective function. The curve fitting problem is decomposed into and completed in two steps. In the first step, an optimal subset of basis functions are selected from the entire set and then the ratio for this optimal group is determined from the least square problem. This curve fitting problem can be formulated as

$$\min \quad \sum_{i=1}^{N} (Z(u_i) - \frac{1}{\rho_i})^2 + \lambda ||\alpha||_1, \tag{7.31}$$

where $||\alpha||_1$ denotes the norm-1 of coefficient vector $[\alpha_1 \alpha_2 \cdots \alpha_K]$, K is the number of basis

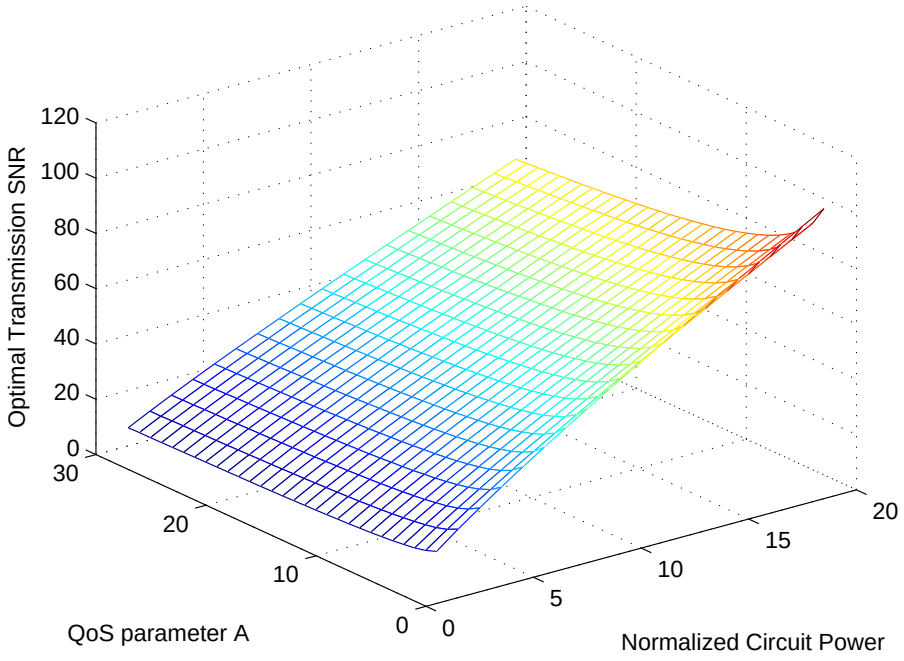

Fig. 7.6: Optimal SNR under different QoS parameter and circuit power combination

functions. $\lambda > 0$ is a parameter used to trade off the quality of the fitting and the sparsity of the coefficient vector. The solution to the above problem is denoted as set Ψ, which includes all the $\alpha_i's$, $\alpha_i \neq 0$. The second step solves the least-squares problem min $\sum_{i=1}^{N} (Z(u_i) - \frac{1}{\rho_i})^2$ with variables $\alpha_i \in \Psi$. It is noticed that the curve fitting problem can be easily solved from some well-developed convex optimization algorithm or toolbox.

7.4.2 Numerical Results and Discussions

To verify that expression (7.29) accurately establishes the relationship between EE and SE in a Rayleigh fading channel under statistical QoS requirement, Figure 7.7 and Figure 7.8 are plotted for EE approximation error against each SE under different QoS parameter A and $\frac{P_c}{N_0 W}$. In a wide range of QoS and $\frac{P_c}{N_0 W}$, the approximated EE well reflect the real value, just as shown in Figure 7.9 and Figure 7.10 that the approximated EE tightly fit with the nearly-exact EE-SE curve.

In Figures 7.9 and 7.10, we plot the EE-SE tradeoff in the Rayleigh fading channel

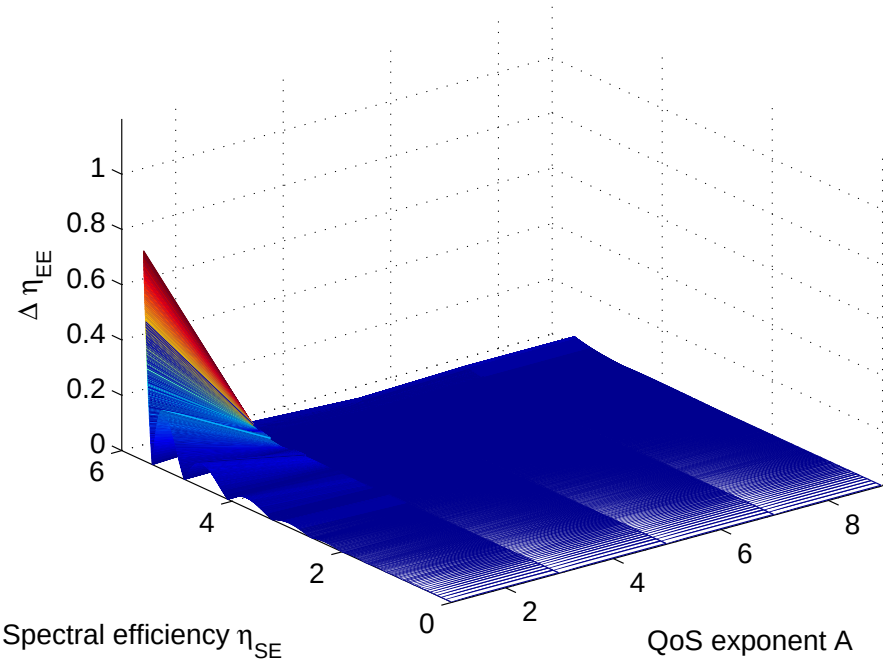

Fig. 7.7: Approximation error in term of η_{EE} between proposed CFA and nearly-exact result as a function of the SE and QoS

under different QoS and circuit power consumption values. In the two figures, we compare our curve fitting results from equations (7.29) and (7.28), with the accurate expressions of η_{EE} and η_{SE} that can be derived from equations (7.20) and (7.8). As what has been proven before, η_{EE} is a non-monotonic and quasi-convex function of SNR ρ, two possible SNR values might be derived from the inverse function $\rho = f^{-1}(\eta_{EE})$, which correspond to two separate spectral efficiency values. However, one can easily attain the SNR $\rho = f^{-1}(\eta_{SE})$ from spectral efficiency function through simple line search algorithm due to its monotonicity. Then, inserting this value in (7.20), we can plot the relationship of η_{EE} with respect to η_{SE}. The derived close-form approximation for the EE-SE tradeoff has been plotted for comparison to verify the accuracy in Figure 7.9 and Figure 7.10.

In Figure 7.9, the EE-SE tradeoff has been plotted for different QoS requirements, i.e., $A = 9, 7, 5, 3, 1$, over Rayleigh fading channel. In this scenario, circuit power consumption is assumed to be fixed at $\frac{P_c}{N_0 W} = 17dB$. Note that the curve for $A = 0$ corresponds to

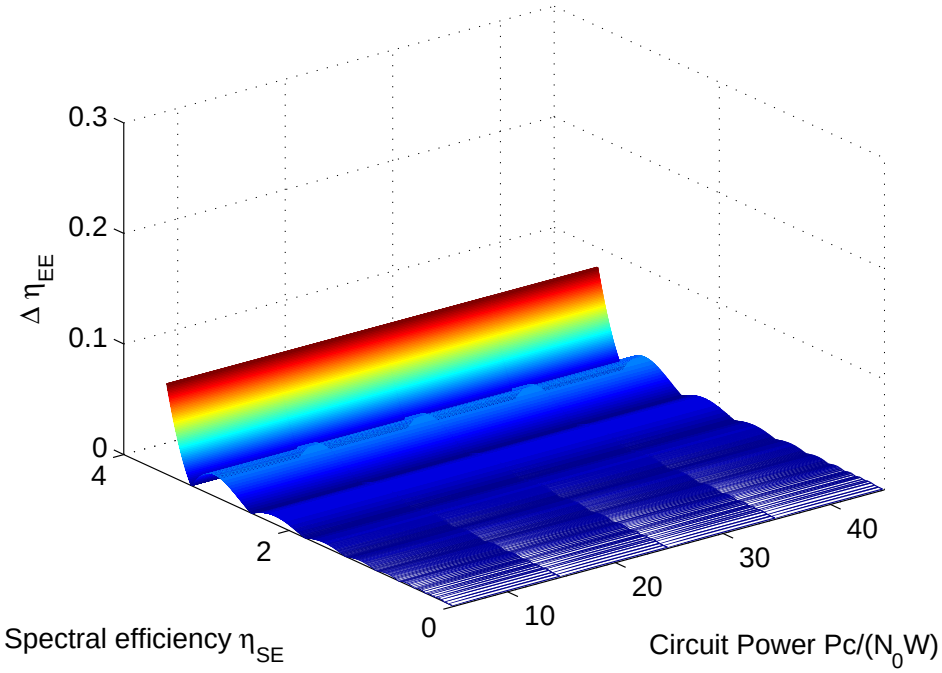

Fig. 7.8: Approximation error in term of η_{EE} between proposed CFA and nearly-exact result as a function of the SE and $\frac{P_c}{N_0 W}$

the Shannon capacity. Obviously, a more stringent QoS will lower down both EE and SE, especially at high SNR.

In Figure 7.10, the impact of circuit power consumption on EE-SE tradeoff has been studied in a Rayleigh fading channel. In the scenario, QoS parameter is fixed at $A = 5$ and we adjust the circuit power in the range of $\frac{P_c}{N_0 W} = 45, 35, 25, 15, 5$. A higher circuit power consumption always lowers down EE, as shown in Figure 7.10. However, opposite to the impact of QoS requirement, circuit power has a greater impact in the low SNR regime than the high SNR regime. This is because circuit power is dominant and contributes more to the EE in the low SNR regime. In the high SNR regime, transmission power is dominant, making EE and SE less relevant to the circuit power consumption.

7.5 Summary

In this study, we investigate the energy efficiency and spectrum efficiency performance

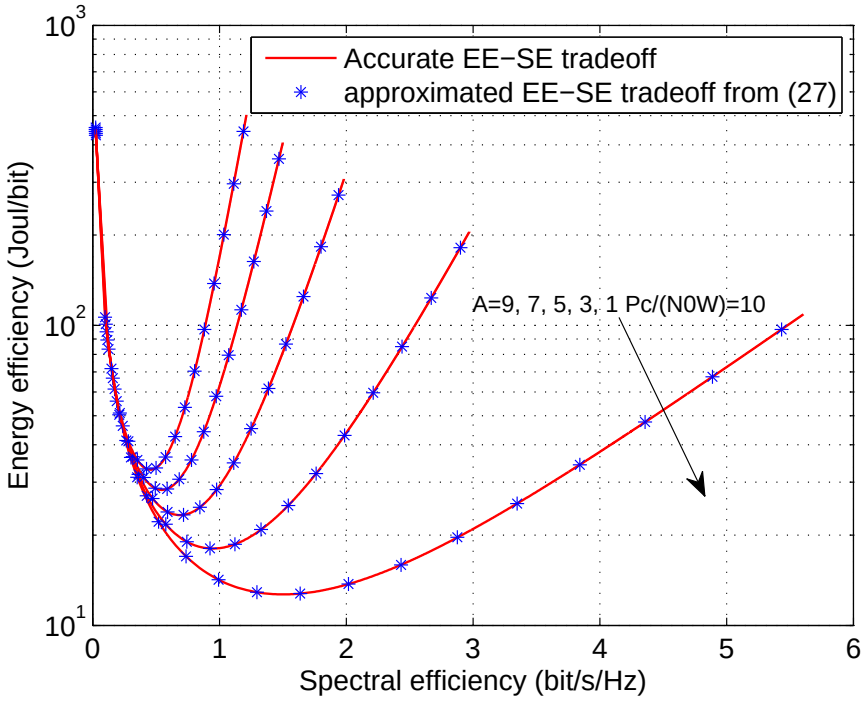

Fig. 7.9: EE-SE tradeoff under different QoS

in a wireless fading channel under the delay constraint. Effective capacity is used to measure the wireless system throughput under a delay constraint. We define the energy efficiency function as the total energy consumption, including both circuit power and transmission power consumptions, per information bit transmission. We have proved that the energy efficiency function is a quasi-convex function of transmission SNR. In order to find the best energy efficiency operation and optimal transmission SNR, we have proposed a two-step binary search algorithm to solve the quasi-convex problem. We further discuss the effective capacity in a Rayleigh fading channel case, and derive the approximate expressions for the effective capacity in the high-SNR and low-SNR regimes. To further exploit the tradeoff between EE and SE tradeoff under QoS and circuit power constraints, we develop a generic close-form approximation for EE-SE by employing a curve fitting approach. Numerical results clearly illustrate the impact of QoS and circuit power consumption on EE-SE tradeoff.

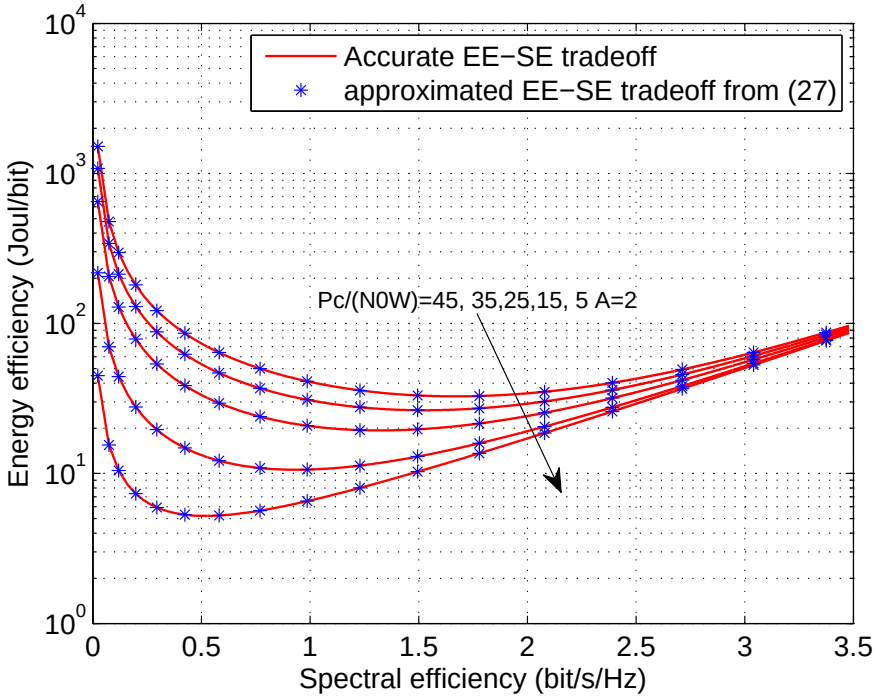

Fig. 7.10: EE-SE tradeoff under different circuit power consumption

Chapter 8

Conclusions

In this dissertation, we investigate various resource allocation and interference management algorithms to improve user experience, system spectral efficiency, and energy efficiency for D2D communication underlaying heterogeneous networks. A multiple range of mathematical tools, from stochastic geometry, optimization, to game theory, are employed to design and analyze the algorithms. The main contributions are summarized as follows.

First, we study the heterogeneous network by introducing a mobile association scheme which jointly optimizes the downlink and uplink resource to improve total network spectral and energy efficiency. The optimization problem is discussed in two scenarios: full frequency reuse and partial frequency reuse. A gradient descent search algorithm is developed to search for the optimal mobile association that can maximize the system capacity and also minimize mobile uplink transmission power consumption. Simulation results show that with joint consideration of downlink and uplink during the mobile association, more MSs would be offloaded from M-BSs to RNs due to shrinking coverage of M-BSs and expansion of RNs, which enhances the resource utilization rate on RNs. That also means improvement on total system spectral efficiency.

Second, we study the uplink and downlink coverage for both cellular users and D2D users in D2D communications by using stochastic geometry. Statistical channel information and user distribution are utilized to evaluated network performance such as coverage, outage probability, throughput. The analytical results match very well with the simulated SINR distribution curve in our work. In the downlink resource sharing scenario, selection of maximum D2D transmission power and D2D communication distance is very delicate, as higher D2D user density, a lower transmission power and larger distance always degrade D2D SINR distribution. In the uplink resource scenario, it is obvious to see that a guard

area around BS can effectively protect cellular users from interference caused by D2D transmission and link quality of cellular users can be guaranteed. Actually, high transmission does not necessary bring significant performance gain for D2D user due to increased co-channel interference and it also causes more interference to cellular users, thus degrades cellular user SINR distribution, especially for the case without a guard area.

Third, we propose a sub-optimal distributed algorithm based on game model for D2D communication underlaying cellular network, aiming to improve system spectral efficiency and energy efficiency. In the distributed resource allocation and power control scheme, a Stackelberg game framework is used to formulate the resource allocation process between two parties, where BS and CUEs are modeled as buyers, aiming to achieve the maximum throughout gain; DUEs are modeled as sellers and aim to gain payment from selling resources. The price mechanism adjusts the system to reach system equilibrium. A two-step approach is proposed to obtain a sub-optimal solution to reduce computational complexity, in which the first step is to group DUEs that share the same radio resource of a CUE, and then allocate resources to them distributively through the pricing mechanism. The amount of power allocated to each user mostly depends on the channel gain of that resource block. Besides, we consider a joint resource/power allocation and mode selection for D2D communications in OFDMA based cellular network and develop a dual optimization framework to solve the mixed integer nonlinear programming problem with a reasonable computational complexity. Analytical results show that our scheme can achieve a much higher system throughput compared with other schemes.

Fourth, we develop an energy efficient resource allocation scheme for D2D communications. Resource allocations between CUEs and DUEs are modeled as a non-cooperative game, where cellular users or D2D users determine which resource blocks to select and how much power they will transmit correspondingly so as to maximize the utility function. The utility function in the work is defined as the achievable rate normalized by power consumption. When using a flat fading channel, the most energy efficient strategy turns out that a CUE allocates the optimal power on its assigned channel for the maximum energy efficiency,

while a DUE should allocate the least amount of power on the channel that has the best channel gain-to-interference-noise-ratio (CINR). Yet the DUE still achieves the highest data rate on that channel.

Last but not least, we address the general tradeoff between energy efficiency and spectral efficiency in a delay constrained wireless system. System QoS requirement is incorporated and measured through the effective capacity. The spectral efficiency is defined as effective capacity per unit bandwidth and energy efficiency is defined as energy consumed per effective capacity bit. Through our analysis, we find that QoS requirements and circuit power consumption affect the EE-SE tradeoff differently. In the low SNR regime, circuit power shows more impact on the EE-SE tradeoff while QoS impacts EE-SE tradeoff more in the high SNR regime. Other aspects are discussed to improve both spectral and energy efficiency.

References

[1] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *Wireless Communications, IEEE Transactions on*, vol. 2, no. 4, pp. 630–643, 2003.

[2] M. C. Gursoy, "Mimo wireless communications under statistical queueing constraints," *Information Theory, IEEE Transactions on*, vol. 57, no. 9, pp. 5897–5917, 2011.

[3] M. Matthaiou, G. C. Alexandropoulos, H. Q. Ngo, and E. G. Larsson, "Analytic framework for the effective rate of miso fading channels," *Communications, IEEE Transactions on*, vol. 60, no. 6, pp. 1741–1751, 2012.

[4] C. Zhong, T. Ratnarajah, K.-K. Wong, and M.-S. Alouini, "Effective capacity of correlated miso channels," in *Communications (ICC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–5.

[5] D. Wu and R. Negi, "Effective capacity-based quality of service measures for wireless networks," in *Broadband Networks, 2004. BroadNets 2004. Proceedings. First International Conference on*. IEEE, 2004, pp. 527–536.

[6] K. Doppler, C. B. Ribeiro, and J. Kneckt, "Advances in d2d communications: Energy efficient service and device discovery radio," in *2011 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE)*.

[7] T. G. P. P. (3GPP), "Physical layer procedures (release 10) for evolved universal terrestrial radio access (e-utra)," *Technical Specification, 3GPP TS*.

[8] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *Communications Surveys & Tutorials, IEEE*, vol. 16, no. 4, pp. 1801–1819, 2014.

[9] Y.-D. Lin and Y.-C. Hsu, "Multihop cellular: A new architecture for wireless communications," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 2000, pp. 1273–1282.

[10] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 4, pp. 1215–1228, 2013.

[11] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to lte-advanced networks," *Communications Magazine, IEEE*, vol. 47, no. 12, pp. 42–49, 2009.

[12] B. Kaufman and B. Aazhang, "Cellular networks with an overlaid device to device network," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*. IEEE, 2008, pp. 1537–1541.

[13] T. Peng, Q. Lu, H. Wang, S. Xu, and W. Wang, "Interference avoidance mechanisms in the hybrid cellular and device-to-device systems," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE, 2009, pp. 617–621.

[14] S. Xu, H. Wang, T. Chen, Q. Huang, and T. Peng, "Effective interference cancellation scheme for device-to-device communication underlaying cellular networks," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*. IEEE, 2010, pp. 1–5.

[15] P. Jänis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-aware resource allocation for device-to-device radio underlaying cellular networks," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*. IEEE, 2009, pp. 1–5.

[16] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 12, pp. 3995–4000, 2011.

[17] C.-H. Yu and O. Tirkkonen, "Device-to-device underlay cellular network based on rate splitting," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. IEEE, 2012, pp. 262–266.

[18] X. Chen, L. Chen, M. Zeng, X. Zhang, and D. Yang, "Downlink resource allocation for device-to-device communication underlaying cellular networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*. IEEE, 2012, pp. 232–237.

[19] C.-H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," in *Communications, 2009. ICC'09. IEEE International Conference on*. IEEE, 2009, pp. 1–5.

[20] K. Doppler, C.-H. Yu, C. B. Ribeiro, and P. Jänis, "Mode selection for device-to-device communication underlaying an lte-advanced network," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*. IEEE, 2010, pp. 1–6.

[21] A. Osseiran, K. Doppler, C. Ribeiro, M. Xiao, M. Skoglund, and J. Manssour, "Advances in device-to-device communications and network coding for imt-advanced," *ICT Mobile Summit*, 2009.

[22] K. Doppler, M. P. Rinne, P. Jänis, C. Ribeiro, and K. Hugl, "Device-to-device communications; functional prospects for lte-advanced networks," in *Communications Workshops, 2009. ICC Workshops 2009. IEEE International Conference on*. IEEE, 2009, pp. 1–6.

[23] C.-H. Yu, K. Doppler, C. Ribeiro, and O. Tirkkonen, "Performance impact of fading interference to device-to-device communication underlaying cellular networks," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE, 2009, pp. 858–862.

[24] C.-H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 8, pp. 2752–2763, 2011.

[25] X. Xiao, X. Tao, and J. Lu, "A qos-aware power optimization scheme in ofdma systems with integrated device-to-device (d2d) communications," in *Vehicular Technology Conference (VTC Fall), 2011 IEEE*. IEEE, 2011, pp. 1–5.

[26] G. Zhang, "Subcarrier and bit allocation for real-time services in multiuser ofdm systems," in *Communications, 2004 IEEE International Conference on*, vol. 5. IEEE, 2004, pp. 2985–2989.

[27] G. Youjun, X. HaiBo, T. Hui, and Z. Ping, "A qos-guaranteed adaptive resource allocation algorithm with low complexity in ofdma system," in *Wireless Communications, Networking and Mobile Computing, 2006. WiCOM 2006. International Conference on*. IEEE, 2006, pp. 1–4.

[28] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlaying cellular networks," *Communications, IEEE Transactions on*, vol. 61, no. 8, pp. 3541–3551, 2013.

[29] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying lte network," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2010 IEEE 6th International Conference on*. IEEE, 2010, pp. 368–375.

[30] M. Belleschi, G. Fodor, and A. Abrardo, "Performance analysis of a distributed resource allocation scheme for d2d communications," in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*. IEEE, 2011, pp. 358–362.

[31] M. Jung, K. Hwang, and S. Choi, "Joint mode selection and power allocation scheme for power-efficient device-to-device (d2d) communication," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*. IEEE, 2012, pp. 1–5.

[32] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *Communications, IEEE Transactions on*, vol. 59, no. 11, pp. 3122–3134, 2011.

[33] S. Mukherjee, "Distribution of downlink sinr in heterogeneous cellular networks," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 575–585, 2012.

[34] ——, "Downlink sinr distribution in a heterogeneous cellular wireless network with biased cell association," in *Communications (ICC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 6780–6786.

[35] T. D. Novlan, R. K. Ganti, A. Ghosh, and J. G. Andrews, "Analytical evaluation of fractional frequency reuse for ofdma cellular networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 12, pp. 4294–4305, 2011.

[36] K. Okino, T. Nakayama, C. Yamazaki, H. Sato, and Y. Kusano, "Pico cell range expansion with interference mitigation toward lte-advanced heterogeneous networks," in *Communications Workshops (ICC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–5.

[37] Y. Yu, R. Q. Hu, C. S. Bontu, and Z. Cai, "Mobile association and load balancing in a cooperative relay cellular network," *Communications Magazine, IEEE*, vol. 49, no. 5, pp. 83–89, 2011.

[38] Q. Li, R. Q. Hu, G. Wu, and Y. Qian, "On the optimal mobile association in heterogeneous wireless relay networks," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 1359–1367.

[39] L. Liu, J. C. Zhang, J.-C. Yu, and J. Lee, "Intercell interference coordination through limited feedback," *International Journal of Digital Multimedia Broadcasting*, vol. 2010, 2010.

[40] E. U. T. R. Access, "Further advancements for e-utra physical layer aspects," 3GPP TR 36.814, Tech. Rep., 2010.

[41] J. F. C. Kingman, *Poisson processes*. Oxford university press, 1992, vol. 3.

[42] M. Z. Win, P. C. Pinto, L. Shepp *et al.*, "A mathematical theory of network interference and its applications," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 205–230, 2009.

[43] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[44] C.-H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "On the performance of device-to-device underlay communication with simple power control," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*. IEEE, 2009, pp. 1–5.

[45] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A new tractable model for cellular coverage," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 1204–1211.

[46] A. Jeffrey and D. Zwillinger, *Table of integrals, series, and products*. Academic Press, 2007.

[47] T. Rappaport, "Wireless communications: Principles and practice," 2001.

[48] R. Gibbons, *Game theory for applied economists*. Princeton University Press, 1992.

[49] G. Owen, "Game theory. 1995."

[50] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[51] M. R. Garey and D. S. Johnson, "Computers and intractability: a guide to the theory of np-completeness. 1979," *San Francisco, LA: Freeman*, 1979.

[52] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *Communications, IEEE Transactions on*, vol. 54, no. 7, pp. 1310–1322, 2006.

[53] X. Chen, R. Q. Hu, and Y. Qian, "Distributed resource and power allocation for device-to-device communications underlaying cellular network," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 4947–4952.

[54] C. Isheden and G. P. Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–6.

[55] T. S. Ferguson, *Mathematical statistics: A decision theoretic approach*. Academic press, 2014, vol. 1.

[56] R. D. Yates, "A framework for uplink power control in cellular radio systems," *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1341–1347, 1995.

[57] F. W. Olver, *NIST handbook of mathematical functions*. Cambridge University Press, 2010.

[58] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. prentice hall PTR New Jersey, 1996, vol. 2.

[59] L. Liu and J.-F. Chamberland, "On the effective capacities of multiple-antenna gaussian channels," in *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*. IEEE, 2008, pp. 2583–2587.

[60] D. Wu and R. Negi, "Effective capacity-based quality of service measures for wireless networks," *Mobile Networks and Applications*, vol. 11, no. 1, pp. 91–99, 2006.

[61] A. Lozano, A. M. Tulino, and S. Verdú, "High-snr power offset in multiantenna communication," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4134–4151, 2005.

[62] E. Whittaker and G. Watson, "A course of modern analysis. cambridge mathematical library," 1996.

[63] S. Verdú, "Spectral efficiency in the wideband regime," *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1319–1343, 2002.

[64] R. S. Chen, R. Hu, Q. Li, and G. Wu, "Energy efficiency in a delay constrained wireless network," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. IEEE, 2013, pp. 836–841.

Appendices

Appendix A

PDF of the Distance between Two DUEs Forming a Pair

As illustrated in Figure A.1, one D2D user is located r_i away from the circle center. Another D2D user randomly lies within the circle and is r_j away from the circle center. Thus $0 < r_i, r_j < R_d$. Let $\Delta_R = |\vec{r_i} - \vec{r_j}|$ denote the distance of D2D pair. The two D2D users are uniformly distributed in the circle. Both of them have the same distribution:

$$f_{r_i}(r) = f_{r_j}(r) = \begin{cases} \frac{2r}{{R_d}^2} & 0 \leq r \leq R_d, \\ 0 & \text{otherwise.} \end{cases} \tag{A.1}$$

The conditional cumulative distribution function of Δ_R is calculated by $\mathbb{P}(\Delta_R < r | R_i = r_i) =$

$$\begin{cases} P_1(r, r_i), & 0 < r < R_d - r_i, \\ P_2(r, r_i), & R_d - r_i < r < \sqrt{{R_d}^2 - {r_i}^2}, \\ P_3(r, r_i), & \sqrt{{R_d}^2 - {r_i}^2} < r < \sqrt{{R_d}^2 + {r_i}^2}, \\ P_4(r, r_i), & \sqrt{{R_d}^2 + {r_i}^2} < r < R_d + r_i, \\ P_5(r, r_i), & R_d + r_i < r, \end{cases} \tag{A.2}$$

where $P_1(r, r_i)$, $P_2(r, r_i)$, $P_3(r, r_i)$, $P_4(r, r_i)$, $P_5(r, r_i)$ are evaluated as following:

$$P_1(r, r_i) = \frac{\pi r^2}{\pi {R_d}^2}, \tag{A.3}$$

$$P_2(r, r_i) = \frac{r^2[\pi - \arctan(\frac{y_0}{x_0 - r_i})] + {R_d}^2 \arctan(\frac{y_0}{x_0}) - y_0 \cdot r_i}{\pi {R_d}^2}, \tag{A.4}$$

$$P_3(r, r_i) = \frac{r^2 \arctan(\frac{y_0}{r_i - x_0}) + {R_d}^2 \arctan(\frac{y_0}{x_0}) - y_0 \cdot r_i}{\pi {R_d}^2}, \tag{A.5}$$

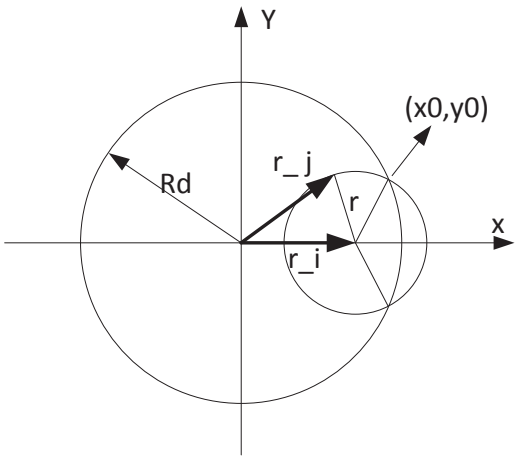

Fig. A.1: Two D2D user randomly located in a circle with a constraint radius of R_d

$$P_4(r, r_i)=\frac{r^2 \arctan(\frac{y_0}{r_i - x_0}) + {R_d}^2[\pi + \arctan(\frac{y_0}{x_0})] - y_0 \cdot r_i}{\pi {R_d}^2}, \tag{A.6}$$

$$P_5(r, r_i) = 1. \tag{A.7}$$

In the above probability expression, (x_0, y_0) is the intersection co-ordinate of the circle centered around the i_{th} user with radius r. So we have $x_0 = \frac{{r_i}^2 + {R_d}^2 - r^2}{2r_i}$, $y_0 = \frac{\sqrt{[(r+R_d)^2 - {r_i}^2][{r_i}^2 - (R_d - r)^2]}}{2r_i}$ from

$$\begin{cases} x^2 + y^2 = {R_d}^2, \\ (x - r_i)^2 + y^2 = r^2. \end{cases}$$

By conditional probability function of D2D distance, we can evaluate the distance CDF of D2D pair according to

$$\mathbb{P}(\Delta_R < r) =$$

$$\begin{cases} \int_0^{R_d - r} P_1(r, r_i) f_{R_i}(r_i) dr_i + \int_{R_d - r}^{\sqrt{{R_d}^2 - r^2}} P_2(r, r_i) f_{R_i}(r_i) dr_i \\ + \int_{\sqrt{{R_d}^2 - r^2}}^{R_d} P_3(r, r_i) f_{R_i}(r_i) dr_i, & 0 < r < R_d, \\ \int_0^{r - R_d} P_5(r, r_i) f_{R_i}(r_i) dr_i + \int_{r - R_d}^{\sqrt{r^2 - {R_d}^2}} P_4(r, r_i) f_{R_i}(r_i) dr_i \\ + \int_{\sqrt{r^2 - {R_d}^2}}^{R_d} P_3(r, r_i) f_{R_i}(r_i) dr_i, & R_d < r < \sqrt{2} R_d, \\ \int_0^{r - R_d} P_5(r, r_i) f_{R_i}(r_i) dr_i + \int_{r - R_d}^{R_d} P_4(r, r_i) f_{R_i}(r_i) dr_i, \\ & \sqrt{2} R_d < r < 2R_d. \end{cases} \tag{A.8}$$

By evaluating the derivation of CDF with respect to r, we can obtain the PDF for Δ_R. For clarity, we only provide derivation for the case when $\sqrt{2}R_d < r < R_d$. Similar derivations can be applied for the other two cases. When $\sqrt{2}R_d < r < 2R_d$, $f_{\Delta_R | \sqrt{2}R_d < r < 2R_d}(r)$ is equal to

$$\begin{aligned} &= \frac{dP(\Delta_R < r)}{dr} \\ &= \frac{d}{dr}[\int_{r - R_d}^{R_d} P_4(r, r_i) f_{R_i}(r_i) dr_i + \int_0^{r - R_d} P_5(r, r_i) f_{R_i}(r_i) dr_i] \\ &= \overbrace{\int_{r - R_d}^{R_d} \frac{\partial P_4(r, r_i)}{\partial r} f_{R_i}(r_i) dr_i}^{\text{Part 1}} - \overbrace{P_4(r, r - R_d) f_{R_i}(r - R_d)}^{\text{Part 2}} \\ &\quad + \frac{2(r - R_d)}{{R_d}^2} \\ &= -\frac{4r^2}{\pi {R_d}^3} G(\arccos(\frac{r}{2R_d})). \end{aligned} \tag{A.9}$$

And part 1 and part 2 in the above equation are further elaborated in the following:

Part 1

$$
\begin{aligned}
&= \int_{r-R_d}^{R_d} \frac{2r\arctan(\frac{y_0}{r_i-x_0})}{\pi {R_d}^2} \cdot \frac{2\pi r_i}{\pi {R_d}^2} dr_i \\
&= C \int_{r-R_d}^{R_d} \\
&\quad \arctan(\frac{\sqrt{[(r+R_d)^2-{r_i}^2][{r_i}^2-(R_d-r)^2]}}{{r_i}^2+r^2-{R_d}^2}) 2r_i dr_i \\
&\overset{(a)}{=} C \int_{2r(r-R_d)}^{r^2} \arctan(\frac{\sqrt{(2rR_d)^2-(2r^2-t)^2}}{t}) dt \\
&= C \int_{2r(r-R_d)}^{r^2} \arctan(\frac{2rR_d}{t}\sqrt{1-(\frac{r}{R_d}-\frac{t}{2rR_d})^2}) dt \\
&\overset{(b)}{=} C \int_0^{\arccos(\frac{r}{2R_d})} \arctan(\frac{\sin(y)}{\frac{r}{R_d}-\cos(y)}) 2rR_d \sin(y) dy \\
&= -\frac{4r^2}{\pi {R_d}^3} \int_0^{\arccos(\frac{r}{2R_d})} \arctan(\frac{\sin(y)}{\cos(y)-\frac{r}{R_d}}) \sin(y) dy \\
&= -\frac{4r^2}{\pi {R_d}^3} G(\arccos(\frac{r}{2R_d})),
\end{aligned} \tag{A.10}
$$

where $C = \frac{2\pi r}{(\pi {R_d}^2)^2}$. We use $t = {r_i}^2 + r^2 - {R_d}^2$ in (a) and $\cos(y) = \frac{r}{R_d} - \frac{t}{2rR_d}$ in (b) in the above expressions. The $G(\cdot)$ function defined here can be calculated as

$G(y)$

$$
\begin{aligned}
&= \int \arctan(\frac{\sin(y)}{\cos(y)-\frac{r}{R_d}}) \sin(y) dy \\
&= \frac{[(\frac{r}{R_d})^2-1]y - 2[(\frac{r}{R_d})^2+1]\arctan(\frac{(r+R_d)\tan(\frac{y}{2})}{r-R_d})}{4\frac{r}{R_d}} \\
&\quad + \frac{2\frac{r}{R_d}\sin(y) + 4\frac{r}{R_d}\cos(y)\arctan(\frac{\sin(y)}{\frac{r}{R_d}-\cos(y)})}{4\frac{r}{R_d}}.
\end{aligned} \tag{A.11}
$$

Part 2

$$
\begin{aligned}
&= \Big[\frac{\overbrace{r^2 \arctan(\frac{y_0}{r-R_d-x_0})}^{\equiv 0} + \pi {R_d}^2 + \overbrace{{R_d}^2 \arctan(\frac{y_0}{x_0})}^{\equiv 0}}{\pi {R_d}^2} \\
&\quad - \frac{\overbrace{y_0 \cdot (r-R_d)}^{\equiv 0}}{\pi {R_d}^2}\Big] \cdot \frac{2\pi(r-R_d)}{\pi {R_d}^2} \\
&= \frac{2(r-R_d)}{{R_d}^2}. \qquad \text{(A.12)}
\end{aligned}
$$

Appendix B

Proof of Approximated SE at High-SNR

First, we denote $z = \frac{1}{\rho\Omega}$. As [57]

$$\Gamma(a,z) = \frac{z^a e^{-z}}{\Gamma(1-a)} \int_0^\infty \frac{t^{-a}e^{-t}}{z+t} dt, \qquad |arg(z)| < \pi, Re(a) < 1, \tag{B.1}$$

where $\Gamma(x)$ is the gamma function .

Thus when $0 < A \leq 1$,

$$\begin{aligned} S_\infty &= \lim_{\rho\to\infty} \frac{\eta_{\mathrm{SE}}(\theta,\rho)}{\log_2\rho} \\ &= -\frac{1}{A} \lim_{\rho\to\infty} \frac{\ln\{(\frac{1}{\rho\Omega})^A e^{\frac{1}{\rho\Omega}} \int_{\frac{1}{\rho\Omega}}^\infty \frac{1}{x^A x^x} dx\}}{\ln\rho} \\ &= \frac{1}{A} \lim_{z\to 0} \frac{A\ln z + \ln\{e^z \Gamma(1-A,z)\}}{\ln z + \ln\Omega} \\ &= 1, \end{aligned} \tag{B.2}$$

and

$$\begin{aligned} L_\infty &= \lim_{\rho\to\infty} [\log_2\rho - \eta_{\mathrm{SE}}(\theta,\rho)] \\ &= \frac{1}{A} \lim_{\rho\to\infty} \mathrm{E}\{\frac{\rho^A}{(1+\rho|h|^2)^A}\} \\ &= \frac{1}{A}\log_2\{(|h|^2)^{-A}\} \\ &= \log_2 \frac{\Gamma(1-A)^{\frac{1}{A}}}{\Omega}. \end{aligned} \tag{B.3}$$

Therefore, when the delay requirement is not stringent, the results coincide with the conclusion made in the theorem 6 in [2].

When $A > 1$,

$$\begin{aligned} S_\infty &= \lim_{\rho\to\infty} \frac{\eta_{\mathrm{SE}}(\theta,\rho)}{\log_2\rho} \\ &= \frac{1}{A}\lim_{z\to 0} \frac{\ln\{z^A e^z \Gamma(1-A,z)\}}{\ln z + \ln\Omega} \\ &= \frac{1}{A}\lim_{z\to 0} \frac{\ln z + \ln\int_0^\infty \frac{t^{A-1}e^{-t}}{z+t}dt - \ln\Gamma(A)}{\ln z + \ln\Omega} \\ &= \frac{1}{A}, \end{aligned} \tag{B.4}$$

and

$$\begin{aligned} L_\infty &= \lim_{\rho\to\infty}[\log_2\rho - A\eta_{\mathrm{SE}}(\theta,\rho)] \\ &= \lim_{\rho\to\infty}[\log_2\rho + \log_2(\mathrm{E}\{(1+\rho|h|^2)^{-A}\})] \\ &= \lim_{\rho\to\infty}\{\log_2\rho + \log_2\frac{ze^{-z}}{\Gamma(A)}\int_0^\infty \frac{t^{A-1}e^{-t}}{z+t}dt\} \\ &= \lim_{z\to 0}\{\log_2\frac{1}{\Gamma(A)\Omega}\int_0^\infty \frac{t^{A-1}e^{-t}}{z+t}dt\} \\ &= \log_2\frac{\Gamma(A-1)}{\Gamma(A)\Omega}. \end{aligned} \tag{B.5}$$

Vita

Xue Chen

Education

- Ph.D., Electrical Engineering, Utah State University, 2016 (to be expected)
  Dissertation Topic: Efficient Device to Device Communications underlaying Heterogeneous Networks
- M.S., Electrical Engineering, Chinese Academy of Sciences, October 2009
- B.S., Electrical Engineering, Huazhong University of Science and Technology, June 2006

Published Journal Articles

- Tradeoff between energy efficiency and spectral efficiency in a delay constrained wireless system, X. Chen, R. Q. Hu, G. Wu and Q. C. Li, *Wireless Communications and Mobile Computing*, Mar, 2014.

Published Conference Papers

- Joint uplink and downlink optimal mobile association in a wireless heterogeneous network, X. Chen and R. Q. Hu, in *Proc. IEEE Global Communications Conference (GLOBECOM)*, pp.4131-4137, Dec. 2012.
- Energy efficiency in a delay constrained wireless network, R. X. Chen, R. Q. Hu, Q. Li and G. Wu, in *Proc. IEEE Wireless Communications and Networking Conference(WCNC)*,Apr. 2013.

- Mobile association for heterogeneous wireless relay networks with statistical QoS guarantees, X. Lei, X. Chen, R. Q. Hu and G. Wu, in *Proc. IEEE Global Communications Conference(GLOBECOM)*, Dec. 2013.

- Distributed resource and power allocation for device-to-device communications underlaying cellular network, X. Chen, R. Q. Hu and Y. Qian, in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2014.

- Coverage study of dense device-to-device communications underlaying cellular networks, X. Chen, Q. Hu, Z. Zhang and Y. Qian, in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2014.

- Energy efficient resource allocation for D2D communication underlaying cellular networks, X. Chen, R. Q. Hu, J. Jeon and G. Wu, in *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2015.

- Optimal resource allocation and mode selection for D2D communication underlaying cellular networks, X. Chen, R. Q. Hu, J. Jeon and G. Wu, in *Proc. IEEE Global Communications Conference (GLOBECOM)*, *Accepted*.