Energy Efficient Resource Allocation for Non-Orthogonal Multiple Access (NOMA) Systems

by

Fang Fang

M.A.Sc., Lanzhou University, 2013
B.Sc., Lanzhou University, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE COLLEGE OF GRADUATE STUDIES

(Electrical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

December 2017

The following individuals certify that they have read, and recommend to the College of Graduate Studies for acceptance, a thesis/dissertation entitled: Energy Efficient Resource Allocation for Non-Orthogonal Multiple Access (NOMA) Systems submitted by Fang Fang in partial fulfilment of the requirements of the degree of Doctor of Philosophy

Dr. Julian Cheng, School of Engineering

Supervisor

Dr. Md. Hossain Jahangir, School of Engineering

Supervisory Committee Member

Dr. Feng Chen, School of Engineering

Supervisory Committee Member

Dr. Joshua Brinkerhoff, School of Engineering

University Examiner

Dr. Lin Cai, Electrical & Computer Engineering, University of Victoria

External Examiner

Abstract

Non-orthogonal multiple access (NOMA) is a promising technique for the fifth generation (5G) mobile communication due to its capability of achieving high spectral efficiency and high data rate. A popular NOMA scheme uses power domain to achieve multiple access. By applying successive interference cancellation (SIC) technique at the receivers, multiple users with different power levels can be multiplexed on the same frequency band, providing higher sum rate than that of conventional orthogonal multiple access (OMA) schemes. The energy consumption has increased rapidly in recent years. To save energy and meet the requirement of green communications in 5G, we focus on the energy efficient resource optimization for NOMA systems. Our research aim is to maximize the system energy efficiency in NOMA systems by considering perfect channel state information (CSI) and imperfect CSI via resource management.

We first study the energy efficient resource allocation for a downlink single cell NOMA network with perfect CSI. The energy efficient resource allocation is formulated as a non-convex problem. A low-complexity suboptimal algorithm based on matching theory is proposed to allocate users to subchannels. A novel power allocation is designed to further maximize the system energy efficiency. However, the perfect CSI is challenging to obtain in practice. We subsequently investigate energy efficiency improvement for a downlink NOMA single cell network by considering imperfect CSI. To balance the system performance and computational complexity, we propose a new suboptimal user scheduling scheme, which closely attains the optimal performance. By utilizing Lagrangian approach, an iterative power allocation algorithm is proposed to maximize the system energy efficiency.

Implementing NOMA in Heterogeneous networks (HetNets) can alleviate the cross-

tier interference and highly improve the system throughput via resource optimization. By considering the cochannel interference and cross-tier interference, an iterative algorithm is proposed to maximize the macro cell and small cells energy efficiency. Simulations results show that the proposed algorithm can converge within ten iterations and can achieve higher system energy efficiency than that of OMA schemes.

Preface

A list of my publications at The University of British Columbia is provided in the following.

Refereed Journal Publications

J1. Haijun Zhang, **Fang Fang**, Julian Cheng, Wei Wang, Keping Long, and Victor C. M. Leung,"Energy-efficient resource allocation in 5G NOMA heterogeneous networks,"*IEEE Wireless Communications Magazine*, Feb. 2017. (Submitted)

J2. **Fang Fang**, Haijun Zhang, Julian Cheng, Sébastien Roy, and Victor C. M. Leung, "Joint user scheduling and power allocation optimization for energy efficient NOMA systems with imperfect CSI," Accepted for publication in *IEEE Journals on Selected Areas on Communications*, 2017.

J3. **Fand Fang**, Haijun Zhang, Julian Cheng, and Victor C. M. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access (NOMA) network," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3722–3732, July 2016.

Refereed Conference Publications

C1. **Fang Fang**, Haijun Zhang, Julian Cheng, and Victor C. M. Leung, "Energy-efficient resource scheduling for NOMA systems with imperfect channel state information," *Proceedings of IEEE International Conference on Communications (ICC 2017)*, Paris, France, May 21-25, 2017.

C2. **Fang Fang**, Haijun Zhang, Julian Cheng, and Victor C. M. Leung, "Energy efficiency of resource scheduling for non-orthogonal multiple access (NOMA) wireless network," *Proceedings of IEEE International Conference on Communications (ICC 2016)*, Kuala Lampur, Malaysia, May 23-27, 2016.

Table of Contents

List of Figures

List of Acronyms

Acronyms	Definitions
1G	First Generation
2G	Second Generation
3G	Third Generation
4G	Fourth Generation
5G	Fifth Generation
AMPS	Advanced Mobile Phone System
ATSC	Advanced Television Systems Committee
AWGN	Additive White Gaussian Noise
BS	Base Station
CRNN	Channel Response Normalized by Noise
CSI	Channel State Information
EE	Energy Efficiency
FDE	Frequency Domain Equalization
FTPA	Fractional Transmit Power Allocation
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications

HetNet	Heterogeneous Network
ICT	Information and Communication Technology (ICT)
IMT	International Mobile Telecommunications
IoT	Internet of Things
IP	Internet Protocol
LTE	Long Term Evolution
LTE-Advanced	Long Term Evolution-Advanced
ITU	International Telecommunication Union
ITU-R	International Telecommunications Union-Radio
KKT	Karush-Kuhn-Tucker
LDM	Layer Division Multiplexing
MBS	Macro Base Station
MIMO	Multiple-Input Multiple-Output
MUE	Macro User Equipments
MUST	Multiple-User Superposition Transmission
NOMA	Non-orthogonal Multiple Access
OFDMA	Orthogonal Frequency Division Multiple Access
OMA	Orthogonal Multiple Access
Q1	First Quarter
QoS	Quality of Service
SBS	Small Base Station
SC	Subchannel

SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
SOMSA	Suboptimal Matching Scheme for Subchannel Assignment
SUE	Small User Equipment
UE	User Equipment
UT	User Terminals

List of Symbols

Symbols	Definitions
$\arg\max\{\cdot\}$	Points of the domain of the function at which the function values are maximized
$\log_2(\cdot)$	The log function with base 2
$\max\{\cdot\}$	The maximum value of the function
$\min\{\cdot\}$	The minimum value of the function
s.t.	Subject to
$E[\cdot]$	The statistical expectation operation
$Q_1(\cdot)$	The first-order Marcum Q-function
$\log(\cdot)$	The log function with base 10
$\Pr[\cdot]$	The probability of an event
$\lvert\cdot\rvert$	The absolute value of the argument

Acknowledgements

I am deeply grateful to my supervisor Dr. Julian Cheng for his constant guidance, advice, encouragement and support for my PhD study. He granted me a great flexibility and freedom in my research work. He taught me academic knowledge, research skills and writing skills. I will continue to be deeply influenced by his work enthusiasm, rigorous scholarship, clarity in thinking and professional integrity. It is my honor to study and do research under his supervision.

I would like to express my thanks to Dr. Lin Cai from University of Victoria to serve as my external examiner. I am honored to have her on my committee. I would also like to thank Dr. Joshua Brinkerhoff for being my University Examiner and Dr. Md. Jahangir Hossian and Dr. Chen Feng for being my committee members. I really appreciate their valuable time. Besides, I would like to give special thanks to Dr. Huijun Zhang for insightful discussions and valuable suggestions on my research work.

I want to thank my love Tianlong Liu for his continuing encouragement and accompanying on my PhD adventure. I would also like to thank my dear landlord family (Taiji family) in Kelowna, Bruce Taiji, Jane Taiji, Sian Taiji, grandma and grandpa, for their continuing support and encouragement when I was in depression during my PhD journey.

I would like to thank my dear colleagues Bingcheng Zhu, Md. Zoheb Hassan, Hui Ma, Guanshan Ye and Fan Yang for sharing their academic experiences and constructive viewpoints during my PhD study at The University of British Columbia. I also would like to thank all my dear friends who helped me a lot when I was in Kelowna and I miss the fun we had together.

Finally, I would like to thank my parents and my younger sister for their patience,

understanding, unconditional love and support over all these years. All my achievements would not have been possible without their constant encouragement and support.

Chapter 1

Introduction

1.1 Background and Motivation

The evolution of communication systems began with use of drums, smoke signal and semaphore in the early human history [1]. Since the invention of the telephone by Alexander Graham Bell in 1876, the instant communication across long distance has ignited the revolution of communication systems. The discovery of radio waves demonstrated communication systems using radio signals. The time of wireless communication had begun since Marconi first demonstrated radio transmission in 1895 from the Isle of Wight to a tugboat that was 18 miles away. Radio technology was developed rapidly to enable transmission over longer distances with better quality, less power, and smaller, cheaper devices, thereby enabling public and private radio communications, television, and wireless networking [2]. The researchers at AT&T Bell Laboratories developed the cellular concept to solve the capacity problem emerged during the 1950s and 1960s [3].

During the 1980s, the first generation (1G) mobile telecommunication systems were invented for commercial use. The first analogue cellular system, advanced mobile phone system (AMPS), was widely deployed in North America. With the increasing demand of capacity and high quality of communications, the development of digital cellular technology became significant. The second generation (2G) mobile telecommunication networks were commercially launched in Finland by Radiolinja in 1991. This network used the global system for mobile communications (GSM) standard [2]. The 2G systems had higher spectrum efficiency and offered more mobile data services than the 1G systems. The introduction of the general packet radio service (GPRS) became the first major step in the evolution

of GSM networks towards the third generation (3G) telecommunication technology. In the early 1980s, the third generation telecommunication technology was developed by the International Telecommunication Union (ITU). Compered with the 2G networks, the 3G networks offered higher data rate and greater security. By using the bandwidth and location information of 3G devices, global positioning system (GPS), location-based services, mobile Internet access, video calls and mobile TV were developed into applications. A new generation of cellular standards has appeared approximately every ten years since the 1G systems were introduced. In March 2008, requirements of the fourth generation mobile telecommunication (4G) technology standards were specified by the International Telecommunications Union-Radio communications sector (ITU-R) [4]. A major step from 3G to 4G is that the 4G systems can support all-Internet Protocol (IP) based communication, such as IP telephony, instead of traditional circuit-switched telephony service. The spread spectrum radio technology used in the 3G networks was abandoned in the 4G systems. The main technology in 4G systems were the orthogonal frequency division multiple access (OFDMA) multi-carrier transmission and other frequency domain equalization (FDE) schemes.

In the fourth generation mobile communication systems such as long-term evolution (LTE) and LTE-Advanced [5], OFDMA has been widely adopted to achieve higher data rate. The demand for mobile traffic data volume is expected to be 500-1,000 times larger in 2020 than that in 2010 [6].

Figure 1.1 shows total global monthly data (ExaBytes per month) and voice traffic from Q1 2012 to Q1 2017 [7]. It depicts a continued strong growth in data traffic. This growth is driven by increased smart phone subscriptions and average data volume per subscription, which has been fueled primarily by more viewing of video content. Data traffic grew around 12% quarter-on-quarter and around 70% year-on-year.

To further meet the overwhelming requirement of data rates, various new techniques have been proposed in recent years, and these techniques include massive multiple-input multiple output (MIMO) [8], millimeter wave communications [9], LTE-U[10], C-RAN [11],

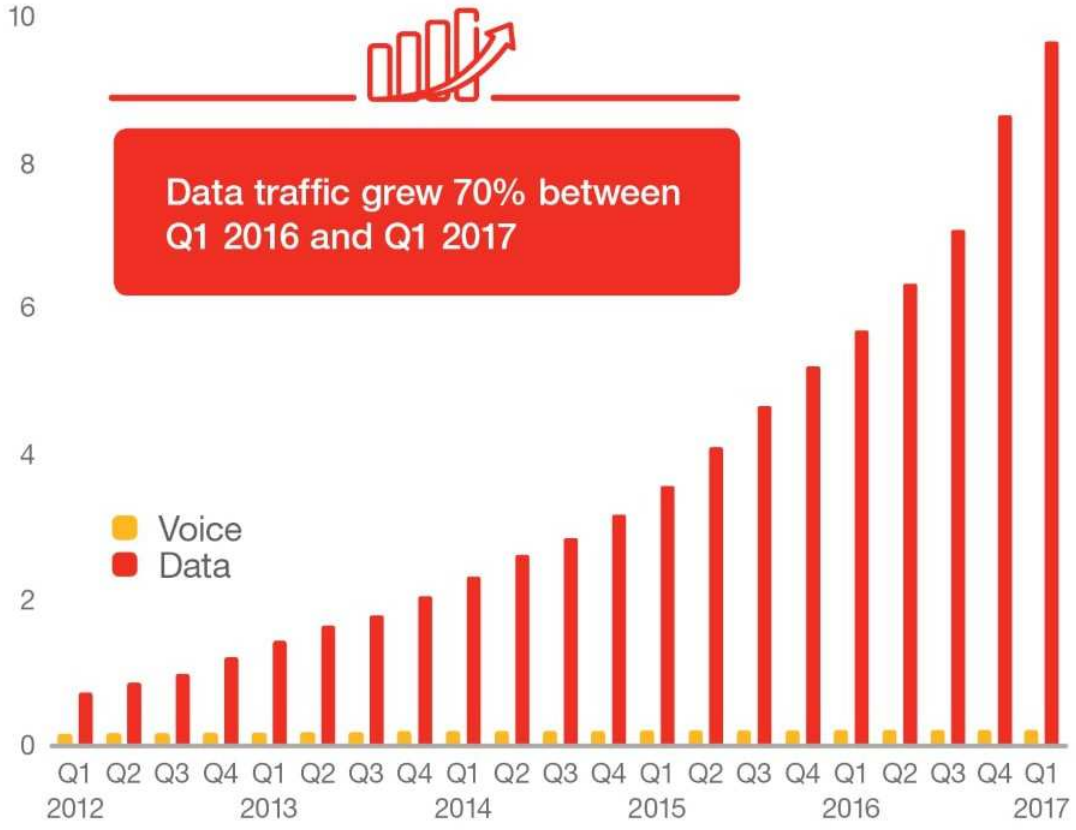

Figure 1.1: Mobile data traffic growth predicted by Ericsson [7].

SON [12] and non-orthogonal multiple access (NOMA)[13]. Among them, NOMA takes advantage of spectrum efficiency by allowing multiple users to occupy the same subchannel, which is different from the resource allocation in OFDM systems [14–16]. By applying successive interference cancellation (SIC) in the NOMA systems, superposition coded signal can be correctly decoded and demodulated at the receiver [17–20]. Therefore, NOMA has been well considered as a promising candidate for the next generation mobile communication systems.

With the overwhelming increase of the traffic data and mobile devices, the energy cost has rapidly increased and become an important issue in the green cellular network because of the increasing amount of CO_2 emission levels caused by energy consumption [22–25]. Thus, fast growing energy consumption and limited global energy resources are the important motivations for the research on energy efficient communications in NOMA networks. In this thesis, we mainly focus on the energy efficient resource allocation for NOMA networks. We aim to design different resource allocation schemes to maximize the energy efficiency of different NOMA systems.

1.2 Literature Review

Since the basic concept of NOMA was introduced and the cell-edge user throughput performance improvement was demonstrated in [26], NOMA has attracted much research attention. The NOMA system has also been envisioned as a key technology in the fifth generation mobile communication systems [27]. A popular NOMA scheme uses power domain to achieve multiple access. More specifically, by utilizing the SIC technology, multiple users can be multiplexed on the same frequency band with different power levels [28, 29]. This is different from the conventional orthogonal multiple access (OMA) scheme. In the downlink NOMA system, the user who has the higher channel state information (CSI) can remove the interference signal from the other users who has less CSI [30, 31]. This protocol enables NOMA to achieve higher spectrum efficiency than the OMA scheme. In addi-

tion to its spectral efficiency gain, academic and industrial research has also demonstrated that NOMA can effectively support massive connectivity, which is important for ensuring that the forthcoming 5G network can support the Internet of Things (IoT) functionalities [32–35].

The multiuser superposition transmission scheme has been proposed in the third generation partnership project long-term evolution advanced (3GPP-LTE-A) networks [36], where NOMA is referred to as multi-user superposition transmission (MUST). A variation of NOMA, termed layer division multiplexing (LDM) was proposed for next generation digital TV standard advanced television systems committee (ATSC) 3.0. Extensive studies for NOMA system have been conducted. By considering a single frequency band in the NOMA system with uniformly deployed mobile users, the outage probability was first derived with perfect CSI, and the numerical results confirmed the derived outage probability expressions [37]. Furthermore, the outage performance of a NOMA system with two types of imperfect CSI was analyzed for the downlink single cell NOMA system, and numerical results showed that the average sum rate matches well with the Monte Carlo simulations [38]. By exploiting the outage probability from [38], the problem of energy efficient user scheduling and power allocation was investigated for NOMA systems by considering only two users multiplexed on one subchannel and imperfect CSI [39].

Besides the performance analysis, the resource allocation was investigated in NOMA systems. By using fractional transmit power allocation (FTPA) among users and equal power allocation across subchannels, the authors in [40, 41] compared system-level performance of the NOMA system with the OMA system, and showed that the overall cell throughput, cell-edge user throughput, and the degree of proportional fairness of NOMA are all superior to those of the OMA scheme. Though FTPA is simple to implement, it fails to optimally allocate power among multiplexed users on each subchannel. Thus, a new power allocation scheme based on water filling was proposed to achieve high spectral efficiency [42]. A cooperative relay system based on NOMA was proposed in [43], where the improvement of the spectral efficiency was presented by numerical results. A greedy

subchannel and power allocation algorithm was proposed for the NOMA system [44], and a cooperative NOMA transmission scheme, where some users have prior information of the other users' message, was proposed in [45] to improve spectrum efficiency. Driven by the rapidly increasing of the energy cost, the energy-efficient power allocation was investigated for NOMA systems [46]. By using statistical channel state information at the transmitter, a near optimal power allocation scheme was proposed to maximize the system energy efficiency [46].

The problem of joint subcarrier allocation and power allocation was investigated in a full-duplex NOMA system [47], where the proposed suboptimal iterative scheme with low computation complexity can achieve close-to-optimal performance. The application of combining NOMA with MIMO technologies has attracted recent research attention. The MIMO NOMA design for small packet transmission and the multi-user detection for uplink NOMA systems were investigated in [48] and [49], respectively. The fairness clustering problem was solved in MIMO NOMA scenarios [50], where an algorithm was proposed to achieve a good tradeoff between the complexity and the throughput.

Driven by the rapid increase of wireless terminal equipments and wide usage of mobile Internet, HetNets have emerged as one of the most promising network infrastructures to provide high system throughput and large coverage of indoor and cell edge scenarios in the 5G wireless communication systems. In such an architecture, a macro cell is overlaid by several small cells, e.g., microcell, picocell and femtocell, to significantly improve the system throughput and the spectral efficiency. For HetNets, frequency band sharing between macro cell and small cells is viable, and it is also more efficient to reuse the frequency bands within a macro cell. However, the cross-tier interference can severely degrade the quality of the wireless transmission. The advantage of the HetNet also comes with fundamental challenges such as cross-tier interference mitigation and resource scheduling. In previous research, the cross-tier interference control and cancellation in HetNets have been investigated, e.g. precoding technique and resource management [15]. In the traditional OFDMA HetNets, the frequency band can be divided into several sub frequency bands, and users in the same

small cell are assigned to different sub frequency bands in order to avoid the inter-cell interference [15]. Spectrum sharing between macro cell and small cells is applicable in HetNets; however, the cross-tier interference and the co-channel interference can severely degrade the communication quality in HetNet. Therefore, implementing NOMA in HetNets can alleviate the cross-tier interference and significantly improve the system throughput via resource optimization [51]. To maximize the sum rate of small cells, an iterative subchannel allocation and power allocation algorithm is proposed to closely approach the optimal solution within limited iterations [51].

To this end, most research works on NOMA systems have focused on the case that the base station (BS) knows the perfect knowledge of the CSI [40, 41, 52]. Since the perfect CSI is challenging to obtain in practice, power-efficient resource allocation scheme was designed to minimize the total transmit power [53], where the proposed scheme can achieve close-to-optimal performance. We assume that a channel estimation error model where the BS only knows the estimated channel gain and *a priori* knowledge of the variance of the estimation error [54, 55]. In this situation, the scheduled user data rate may exceed the maximum achievable data rate due to the estimated channel gain. Therefore, an outage probability requirement should be considered for the resource allocation to maximize system energy efficiency.

1.3 Thesis Outline and Contributions

In this thesis, we present the research work conducted on the following three topics:

- Energy efficient resource allocation for downlink NOMA networks with perfect CSI
- Energy efficient resource allocation for downlink NOMA networks with imperfect CSI
- Energy efficient resource allocation for downlink NOMA HetNets.

The summary and contributions of each chapter are as follows.

Chapter 1 presents background knowledge on the history and development of cellular communication systems. In addition, this chapter provides a detailed literature review related to the rest of the thesis.

Chapter 2 presents the required technical background for the entire thesis. We first introduce the SIC technology and discuss practical issues associated with SIC implementation. After that, we introduce the NOMA system and analyze its performance gain compared to the conventional OMA scheme. As the energy efficient design has an important role in modern communications, a detailed introduction of energy efficiency in communication is provided. Finally, we discuss the resource allocation types and optimization tools in wireless communication systems.

Chapter 3 studies the energy efficient resource allocation for NOMA system with the perfect channel state information. Unlike most previous works focusing on resource allocation to maximize the throughput, we aim to optimize subchannel assignment and power allocation to maximize the energy efficiency for the downlink NOMA network. Assuming perfect knowledge of the channel state information at the base station, we propose a low-complexity suboptimal algorithm which includes energy-efficient subchannel assignment and power proportional factors determination for subchannel multiplexed users. We also propose a novel power allocation across subchannels to further maximize energy efficiency. Since both optimization problems are non-convex, difference of convex programming is used to transform and approximate the original non-convex problems to convex optimization problems. Solutions to the resulting optimization problems can be obtained by iteratively solving the convex subproblems.

In Chapter 4, we investigate energy efficiency improvement for a downlink NOMA single cell network by considering imperfect CSI. The energy-efficient resource optimization problem is formulated as a non-convex optimization problem with the constraints of outage probability limit, the maximum power of the system, the minimum user data rate and the maximum number of multiplexed users sharing the same subchannel. To efficiently solve this problem, the probabilistic mixed problem is first transformed into a non-probabilistic

problem. An iterative algorithm for user scheduling and power allocation is proposed to maximize the system energy efficiency. The optimal user scheduling based on exhaustive search serves as a system performance benchmark, but it has high computational complexity. To balance the system performance and the computational complexity, a new suboptimal user scheduling scheme is proposed to schedule users on different subchannels. Based on the user scheduling scheme, the optimal power allocation expression is derived by the Lagrange approach. By transforming the fractional-form problem into an equivalent subtractive-form optimization problem, an iterative power allocation algorithm is proposed to maximize the system energy efficiency. Simulation results demonstrate that the proposed user scheduling algorithm closely attains the optimal performance. In this work, multiple users can occupy the same subchannel, which is different from [39, 56, 57] where only two users can be supported on the same subchannel.

In Chapter 5, NOMA is extended and applied to HetNets. In this chapter, we design a joint resource allocation scheme to maximize the system energy efficiency via the Lagrangian approach. We maximize not only small cells energy efficiency but also macro cell energy efficiency. Simulation results show that our proposed resource allocation scheme for NOMA HetNets can achieve higher energy efficiency than that of the conventional OMA scheme.

Chapter 6 summarizes the entire thesis and lists our contributions. In addition, some future works related to our current research are also suggested.

Chapter 2

Background on Energy Efficient Non-Orthogonal Multiple Access (NOMA) Systems

In this chapter, we provide a brief description of NOMA system and SIC, and SIC is the main technology applied at the receiver in NOMA systems. Then, the resource allocation background knowledge in wireless network is presented. Finally, we introduce the energy efficiency concept in wireless communications and present our research motivations.

2.1 Non-Orthogonal Multiple Access (NOMA) Systems

2.1.1 Successive Interference Cancelation (SIC) Technology in NOMA Systems

Non-orthogonal multiple access has been recognized as a promising multiple access technique for the fifth generation networks due to its superior spectral efficiency. In NOMA systems, the main technology is SIC technology. Considering a downlink transmission scenario where a single base station transmits superposition coded information to two users. The channel gains from the base station to these two users are h_1 and h_2, respectively. Assume that $|h_1| < |h_2|$, and both h_1 and h_2 are perfectly known to both the transmitter

and receivers. The transmit signal or the sum of two signals, can be expressed as

$$x = x_1 + x_2 \tag{2.1}$$

where x_k is the signal intended for user $k, k = 1, 2$. Therefore, the received signal at user k can be written by

$$y_k = h_k x + w_k, \quad k = 1, 2 \tag{2.2}$$

where $w_k \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ is independent and identically distributed (i.i.d.) complex additive white Gaussian noise (AWGN) with mean zero and variance σ_n^2. The transmit signal x has an average power constraint of $P = P_1 + P_2$ where P_1 and P_2 are the transmit powers for User 1 and User 2, respectively. The key idea of SIC technique is that users who have higher channel gains can decode the data that has been successfully decoded by the users who have less channel gains [58]. In this case, since User 2 has a higher channel gain than User 1, User 2 can decode the data that has been successfully decoded by User 1. Thus, the superposition coding scheme can be implemented by the following steps [59]:

1. The transmit signal is superposition coded signals of the two users.
2. At the receiver, User 1 treats User 2's signal as noise and decodes its data from y_1.
3. User 2 with better channel performs SIC, i.e., it decodes the data of User 1 and subtracts User 1's signal from y_2. After that, User 2 can decode its data.

According to the Shannon's capacity formula, the data rate for User 1 and User 2 with bandwidth B can be achieved by

$$R_1 = B \log\left(1 + \frac{P_1|h_1|^2}{P_2|h_1|^2 + \sigma_n^2}\right) \quad \text{bits/s} \tag{2.3}$$

$$R_2 = B \log\left(1 + \frac{P_2|h_2|^2}{\sigma_n^2}\right) \quad \text{bits/s.} \tag{2.4}$$

We consider K users are distributed in a downlink network with SIC at the receivers

and $|h_K| \geq |h_{K-1}| \geq \cdots \geq |h_1|$. The boundary of the capacity region is given by

$$R_k = B \log \left(1 + \frac{P_k |h_k|^2}{\sigma_n^2 + |h_k|^2 \sum\limits_{j=k+1}^{K} P_j} \right), k = 1, 2, \cdots, K \tag{2.5}$$

for all possible splits $P = \sum\limits_{k=1}^{K} P_k$ of the total power at the base station. The optimal points are achieved by superposition coding at the transmitter and SIC at each of the receivers. The cancellation order at every receiver is always to decode the weaker users before decoding its own data.

We have discussed the advantages of SIC technology in the downlink network. SIC has a significant performance gain over the conventional orthogonal multiple access (OMA) techniques. It takes advantage of the strong channel of the nearby user to give it high rate while providing the weak user with the best possible performance. Here we discuss several potential practical issues in applying SIC in a wireless system [59].

- Complexity will increase when the number of the users increases: In the downlink, applying SIC at the mobile receivers means that the user needs to decode information intended for some of the other users, which would not happen in the conventional system. Thus the decoding complexity at each mobile user will increase when the number of users multiplexed on the same frequency band increases. However, we have seen that superposition coding in conjunction with SIC has the largest performance gain when the users have large disparate channels from the base station. To avoid the high complexity of SIC, user grouping can be a solution in practice. In order to reduce the complexity of decoding, it is suggested to separate users in the cell into groups containing small number of users. Each group users can be multiplexed on the same subchannel and superposition coding based SIC is performed. Therefore, the SIC can achieve the performance gain with low complexity.

- Imperfect channel state information estimation: The interfering signal from the other

users must be reconstructed before it is removed from the received signal. This contribution depends on the estimation of channel state information. The imperfect estimation of CSI will lead to residual cancellation errors. One concern is that if the difference of the users' received powers is large, the residual error from cancelling the stronger user can still swamp the weaker users signal. On the other hand, it is also easier to get an accurate channel estimate when the user has high CSI. It turns out that these two effects compensate each other and the effect of residual errors does not grow with the power disparity.

– Analog-to-digital quantization error: When the difference of received powers of the users is large, a large dynamic range of the analog-to-digital (A/D) converter is required. For example, if the power disparity is 20 dB, even 1-bit accuracy for the weak signal would require an 8-bit A/D converter. This may well pose an implementation constraint on how much gain SIC can offer.

2.1.2 Performance Analysis for NOMA Systems

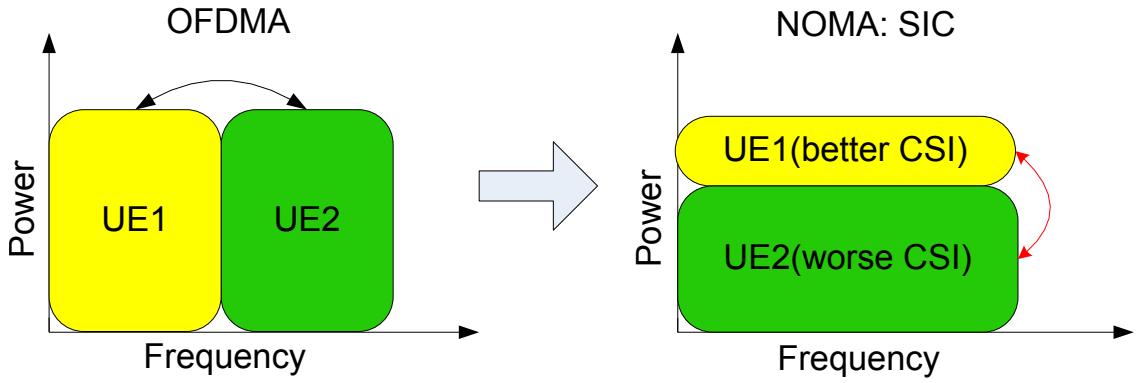

Figure 2.1: OFDMA versus NOMA systems.

In this section, we discuss basic NOMA with SIC and analyze its performance gain over OFDMA schemes. A popular NOMA scheme uses power domain to achieve multiple access. Figure 2.1 presents the power domain comparison over frequency of the NOMA system and the OFDMA system. By applying SIC technique at the receivers, multiple

users with different power levels can be multiplexed on the same frequency band, providing higher sum rate than that of conventional orthogonal multiple access (OMA) schemes.

In NOMA systems, SIC is applied at the receivers. Let us focus on the two-user case for a downlink NOMA system. Consider that two users are multiplexed on the same subchannel with the channel gains $|h_1|^2 \geq |h_2|^2$ shown in Fig. 2.1, where $h_m = g_m \cdot PL^{-1}(d), m = 1, 2$, and where g_m is assumed to be the Rayleigh fading channel gain and $PL^{-1}(d)$ is the path loss function between the BS and UT_m at distance d. Denote the assigned power on SC_n by p_n. The bandwidth of the subchannel is B_{sc} and the power proportional factors for User 1 and User 2 are β_1 and β_2, respectively.

For NOMA systems, SIC is applied at User 1 with a higher channel gain than User 2. According to the SIC protocol, User 1 can cancel the interference signal from User 2. The data rates of User 1 and User 2 in NOMA systems can be respectively represented as

$$R_1 = B_{sc}\log_2\left(1 + \frac{|h_1|^2\beta_1 p_n}{\sigma_n^2}\right) \tag{2.6}$$

and

$$R_2 = B_{sc}\log_2\left(1 + \frac{|h_2|^2\beta_2 p_n}{|h_2|^2\beta_1 p_n + \sigma_n^2}\right). \tag{2.7}$$

The sum rate can be written by

$$R_{NOMA} = B_{sc}\log_2\left(1 + \frac{|h_1|^2\beta_1 p_n}{\sigma_n^2}\right) + B_{sc}\log_2\left(1 + \frac{|h_2|^2\beta_2 p_n}{|h_2|^2\beta_1 p_n + \sigma_n^2}\right). \tag{2.8}$$

In OFDMA systems, we assume OFDMA with orthogonal user multiplexing. The total bandwidth B_{sc} is occupied by these two users (assume that each user has half bandwidth). The data rates of User 1 and User 2 in OFDMA systems can be respectively represented as

$$R_1 = \frac{1}{2}B_{sc}\log_2\left(1 + \frac{|h_1|^2 p_n}{\sigma_n^2}\right) \tag{2.9}$$

and

$$R_2 = \frac{1}{2} B_{sc} \log_2 \left(1 + \frac{|h_2|^2 p_n}{\sigma_n^2} \right). \tag{2.10}$$

Therefore, the sum rate of these two users in OFDM system can be expressed as

$$R_{OFDM} = \frac{1}{2} B_{sc} \log_2 \left(1 + \frac{|h_1|^2 p_n}{\sigma_n^2} \right) + \frac{1}{2} B_{sc} \log_2 \left(1 + \frac{|h_2|^2 p_n}{\sigma_n^2} \right). \tag{2.11}$$

Using the parameters values in [60], we set $\beta_1 = 1/5$, $\beta_2 = 4/5$, $B = 1$ Hz, $\frac{|h_1|^2 p_n}{\sigma_n^2} = 20$ dB and $\frac{|h_2|^2 p_n}{\sigma_n^2} = 0$ dB . Assume each user has the same weighted bandwidth ($B_{sc} = 1$ Hz). Therefore, $R_{OFDM} = 3.33 + 0.50 = 3.83$ bits/sec and $R_{NOMA} = 4.39 + 0.74 = 4.53$ bits/sec. The gain of the NOMA system is 34% more than that of the OFDMA scheme.

Now consider better channel conditions, which means we improve the channel gains by setting $\frac{|h_1|^2 p_n}{\sigma_n^2} = 30$ dB and $\frac{|h_2|^2 p_n}{\sigma_n^2} = 10$ dB. Assume each user has the same weighted bandwidth ($B_{sc} = 1$ Hz), $\beta_1 = 1/5$, $\beta_2 = 4/5$. Therefore, $R_{OFDM} = 4.99 + 1.73 = 6.72$ bits/sec and $R_{NOMA} = 7.65 + 1.87 = 9.52$ bits/sec. The gain of the NOMA system is 42% more than that of the OFDMA scheme.

Based on the above numerical examples, it can be concluded that the sum data rate of the NOMA system will be improved when the channel gains increase, and the performance gain of NOMA over OFDMA increases when the channel conditions improve.

The difference of two systems's sum rate in high signal-to-noise ratio (SNR) region can be calculated. Consider that two users are multiplexed on the same subchannel with the

channel gains $|h_1|^2 \geq |h_2|^2$. Therefore, the difference of sum rate can be written as [61]

$$\begin{aligned}
&R_{NOMA} - R_{OFDM} \\
=&B_{sc}\log_2\left(1+\frac{|h_1|^2\beta_1 p_n}{\sigma_n^2}\right) + B_{sc}\log_2\left(1+\frac{|h_2|^2\beta_2 p_n}{|h_2|^2\beta_1 p_n+\sigma_n^2}\right) \\
&-\frac{1}{2}B_{sc}\log_2\left(1+\frac{|h_1|^2 p_n}{\sigma_n^2}\right) - \frac{1}{2}B_{sc}\log_2\left(1+\frac{|h_2|^2 p_n}{\sigma_n^2}\right) \\
&\underset{\frac{p_n}{\sigma_n^2}\to\infty}{\to} \log_2\left(\frac{p_n}{\sigma_n^2}|h_1|^2\beta_1\right) + \log_2\left(\frac{1}{\beta_1}\right) - \log_2\left(\frac{p_n}{\sigma_n^2}\,|h_1|\,|h_2|\right) \\
=&\log_2\left(|h_1|\right) - \log_2\left(|h_2|\right)
\end{aligned} \tag{2.12}$$

which is not a function of SNR. From (2.12), it can be concluded that the sum rate gap between NOMA and OFDMA will be increased when the channel gain difference of the two users is enlarged. Therefore, NOMA systems can outperform OFDMA systems only if the channel gain difference exists.

2.2 Energy Efficiency in Communication Networks

2.2.1 Research Motivation of Energy Efficiency

From an operation point of view, approximately 600 TWh of the world wide electrical energy is consumed by the information and communication technology (ICT). By the end of 2030, this number is expected to grow to 1700 TWh [21]. This increasing energy consumption becomes an important issue in the green cellular network because of the increasing amount of CO_2 emission levels caused by energy consumption [22–25]. Therefore, the fast growing energy consumption and limited global energy resources are the important motivations for the research on energy efficient wireless communication systems.

2.2.2 Energy Efficiency Definition in Wireless Communications Networks

During the past decades, many research works have been conducted to improve system throughput. However, with the exponential growth of wireless data traffic, energy consumption of wireless networks has been rapidly increasing. Therefore, finding the trade-off of high data rate and energy saving is an urgent task in the next generation wireless communication systems.

Energy efficiency is commonly defined as the ratio of data rate to the power consumption. Bits per Joule is commonly used to measure energy efficiency performance in wireless networks [62, 63]. For energy-efficient communication, it is desirable to send the maximum amount of data with a given amount of energy. With bandwidth B, the achievable data rate $R = B\log\left(1+\frac{P|h|^2}{\sigma_n^2}\right)$, where P is the transmit power, σ_n^2 is the AWGN power and $|h|^2$ is the channel power gain between transmitter and receiver. Given an amount of energy ΔE that is consumed in a duration ΔT, we have $\Delta E = P\Delta T$. Therefore, the energy efficiency (EE) is defined as

$$EE = \frac{R\Delta T}{\Delta E} = \frac{R}{P} \quad \text{bits/Joule.} \tag{2.13}$$

The power consumption includes transmit power and circuit power consumption. The circuit power consumption is the additional device power consumption, which includes signal processing and active circuit blocks such as analog-to-digital converter, digital-to-analog converter, synthesizer, and mixer during the transmission [64]. Denote the additional device power consumption, circuit power, as P_c. Thus, the overall power assumption is $P + P_c$. Energy efficiency needs to be redefined as data rate bits/s per unit energy, where an additional circuit power factor, P_c, needs to be taken into consideration. Therefore, the energy efficiency is defined as

$$EE = \frac{R\Delta T}{\Delta E} = \frac{R}{P+P_c} \quad \text{bits/Joule.} \tag{2.14}$$

Note that the circuit power consumption P_c is independent of the transmit power.

In this thesis, we consider two definitions of system energy efficiency. In Chapter 3, the system energy efficiency is formulated as a summation of each subchannel energy efficiency. In Chapter 4, the system energy efficiency is formulated as the ratio of system sum rate to the total power consumption. Different resource allocation schemes are proposed to improve the system energy efficiency. Regardless different definitions of system energy efficiency, our research conclusion will be the same. By resource allocation, the energy efficiency of a NOMA system can be made higher than that of an OFDMA system.

2.2.3 Resource Optimization and Convex Optimization Approaches

Resource management plays an important role to improve the energy efficiency in wireless communication systems. The main resource management in wireless communications is frequency, time and power optimization. There are different mechanisms for resource management in wireless networks. The most important ones include congestion control, routing, subchannel allocation and power control [65]. In this thesis, we mainly focus on subchannel allocation (user scheduling) and power allocation to maximize the system energy efficiency in NOMA networks. Subchannel allocation means that the schedular needs to assign different users to different subchannels. Since different subchannels have different gains, different subchannel allocation schemes can achieve different performance. Power allocation means that the schedular needs to allocate different powers to the users, which can also achieve different performance gains. One of the most common and effective mathematical tools to solve the resource allocation problem in wireless communication networks is the convex optimization method.

Let us consider a standard form convex problem:

$$\begin{aligned} \min_{x} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, i = 1, 2, ..., m \\ & h_i(x) = 0, i = 1, 2, ..., p \\ & x \in \mathcal{C}. \end{aligned} \tag{2.15}$$

Equation (2.15) describes the problem of finding an x that minimizes $f_0(x)$ among all x values that satisfy the conditions $f_i(x) \leq 0,\ i = 1, 2, ..., m,\ h_i(x) = 0,\ i = 1, 2, ..., p$ and $x \in \mathcal{C}$. We call $x \in \mathcal{C}$ the optimization variable and f_0 the objective function or cost function. $f_i(x)$ and $h_i(x)$ are the inequality and equality constraint functions, respectively, and $\mathcal{C}$ is the constraint set. The domain of the objective and constraint functions are defined as

$$\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\ f_i \cap \bigcap_{i=0}^{m} \mathbf{dom}\ h_i \cap \mathcal{C}. \tag{2.16}$$

The convexity of this problem can be proved by the following conditions [66]. First the objective function $f_0(x)$ should be convex. Second the inequality constraint functions f_i $(i = 1, 2, ..., m)$ and equality constraint functions h_i $(i = 1, 2, ..., p)$ should be convex. Thus, it is proved that the problem (2.15) is convex, and we can find a global optimal solution $x^* \in D$ to this problem by using standard algorithms from convex optimization theory [66], e.g, interior point method and sequential quadratic programming.

2.3 Summary

In this chapter, we presented the essential technical background knowledge on the NOMA system. Compared with the OFDMA system, the performance gain of the NOMA system is presented. Moreover, energy efficiency aspect in wireless communications was provided and the convex optimization approach was briefly introduced.

Chapter 3

Energy Efficient Resource Allocation for Downlink NOMA Network with Perfect CSI

In this chapter, we aim to optimize subchannel assignment and power allocation to maximize the energy efficiency for the downlink NOMA network. Assuming perfect knowledge of the channel state information at base station, we propose a low-complexity suboptimal algorithm that includes energy-efficient subchannel assignment and power proportional factors determination for subchannel multiplexed users. We also propose a novel power allocation across subchannels to further maximize energy efficiency. Since both optimization problems are non-convex, difference of convex programming is used to transform and approximate the original non-convex problems to convex optimization problems. Solutions to the resulting optimization problems can be obtained by iteratively solving the convex subproblems. Simulation results show that the NOMA system equipped with the proposed algorithms yields much better sum rate and energy efficiency performance than the conventional orthogonal frequency division multiple access scheme.

3.1 System Model

Figure 3.1 shows a downlink NOMA network. A BS transmits its signals to M user terminals (UTs) through N subchannels, and SIC is employed at the receiver of UTs. We denote n as index for the nth subchannel where $n \in \{1, 2, \cdots, N\}$ and denote m

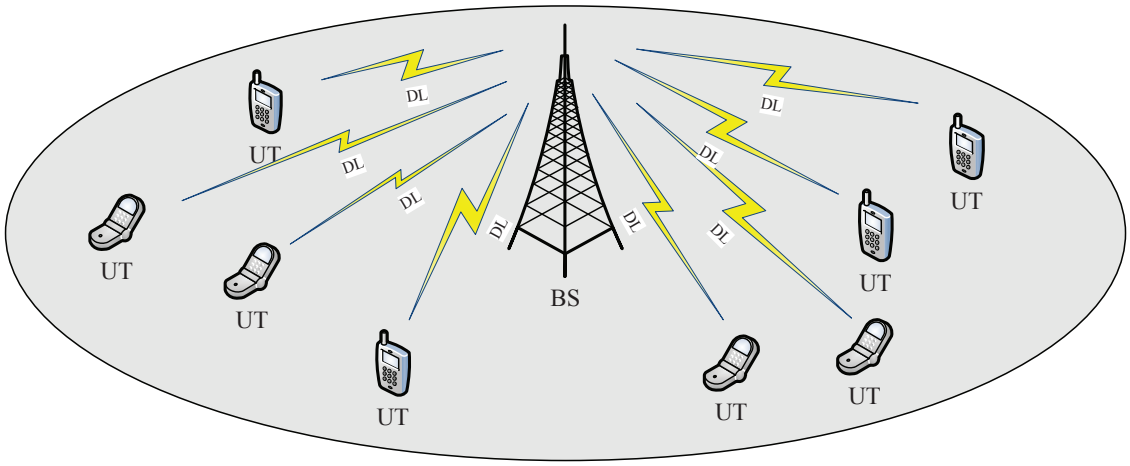

Figure 3.1: System model of a downlink NOMA single-cell network.

as index for the mth mobile user where $m \in \{1, 2, \cdots, M\}$. In the cell, M users are uniformly distributed in a circular region with radius R. The total bandwidth of the system, BW, is equally divided into N subchannels where the bandwidth of each subchannel is $B_{sc} = BW/N$. Let $M_n \in \{M_1, M_2, \cdots, M_N\}$ be the number of users allocated on the subchannel n (SC_n) and the power allocated to the lth user on SC_n is denoted by $p_{l,n}$. Then, the subchannel and BS power constraints are given by $\sum_{l=1}^{M_n} p_{l,n} = p_n$ and $\sum_{n=1}^{N} p_n = P_s$, where p_n and P_s are, respectively, the allocated power on SC_n and the total transmitted power of the BS. In NOMA systems, we assume that the BS has full knowledge of the channel state information. According to the NOMA protocol [26], multiple users can be allocated to the same subchannel with SIC technique. A block fading channel is considered in the system model, where the channel fading of each subchannel remains the same, but it varies independently across different subchannels. Based on the parameters and constraints of the system, the BS needs to assign multiple users (with different power levels) to different subchannels and allocate different powers across subchannels. Considering M_n users are allocated on SC_n, the symbol transmitted by the BS on each subchannel SC_n can be expressed as

$$x_n = \sum_{i=1}^{M_n} \sqrt{p_{i,n}} s_i \tag{3.1}$$

where s_i is the modulated symbol of the ith user on SC_n, which is denoted by $UT_{i,n}$[1]. The received signal at the lth user on SC_n is

$$y_{l,n} = h_{l,n}x_n + z_{l,n} = \sqrt{p_{l,n}}h_{l,n}s_l + \sum_{i=1,i\neq l}^{M_n} \sqrt{p_{i,n}}h_{l,n}s_i + z_{l,n} \tag{3.2}$$

where $h_{l,n} = g_{l,n} \cdot PL^{-1}(d)$ is the coefficient of SC_n from the BS to $UT_{l,n}$, and where $g_{l,n}$ is assumed to have Rayleigh fading channel gain, and $PL^{-1}(d)$ is the path loss function between the BS and $UT_{l,n}$ at distance d. The impact of users' channel conditions on the performance gain of NOMA over OFDMA was studied in [61]. In this work, the authors presented that the performance gain of the NOMA over OFDMA will increase when the difference of channel gain of users become larger. The authors in [37] showed that the distances between BS and UTs will affect the performance of the NOMA system. In this paper, we assume these distances of different users are known by BS. Let $z_{l,n} \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ be the additive white Gaussian noise with mean zero and variance σ_n^2. In a downlink NOMA network, each subchannel can be shared by multiple users. Each user on SC_n receives its signals as well as interference signals from the other users on the same subchannel. Therefore, without SIC at receiver, the received signal-to-interference-plus-noise ratio (SINR) of the lth user on the SC_n is written by

$$SINR_{l,n} = \frac{p_{l,n}|h_{l,n}|^2}{\sigma_n^2 + \sum\limits_{i=1,i\neq l}^{M_n} p_{i,n}|h_{l,n}|^2} = \frac{p_{l,n}H_{l,n}}{1 + \sum\limits_{i=1,i\neq l}^{M_n} p_{i,n}H_{l,n}} \tag{3.3}$$

where $\sigma_n^2 = E[|z_{l,n}|^2]$ is the noise power on SC_n and $H_{l,n} \triangleq |h_{l,n}|^2/\sigma_n^2$ represents the channel response normalized by noise (CRNN) of the lth user on SC_n. Based on the Shannon's capacity formula, the achievable sum rate of SC_n is written by

$$R_n = B_{sc}\sum_{l=1}^{M_n} \log_2\left(1 + SINR_{l,n}\right) = B_{sc}\sum_{l=1}^{M_n} \log_2\left(1 + \frac{p_{l,n}H_{l,n}}{1 + I_{l,n}}\right) \tag{3.4}$$

[1]Without causing notational confusion, $UT_{i,n}$ denotes the ith user on SC_n, while UT_m denotes the mth user in the cell, where $m \in \{1, 2, \cdots, M\}$.

where $I_{l,n}$ is the interference that $UT_{l,n}$ receives from the other users on the SC_n, which can be expressed as

$$I_{l,n} = \sum_{i=1, i \neq l}^{M_n} p_{i,n} H_{l,n}. \tag{3.5}$$

In NOMA systems, the SIC process is implemented at UT receiver to reduce the interference from the other users on the same subchannel. The optimal decoding order for SIC is the increasing order of CRNNs. Based on this order, any user can successfully and correctly decode the signals of the other users with smaller CRNN values. Thus, the interference from the users having poorer channel condition can be cancelled and removed by the user who has better channel condition. In order to maximize the sum rate of SC_n, the NOMA protocol allocates higher power to the users with lower CRNN [26], i.e., for two users $UT_{i,n}$ and $UT_{j,n}$ sharing the same SC_n with CRNNs $|H_{i,n}| \geq |H_{j,n}|$ [2], we always set $p_{i,n} \leq p_{j,n}$ to guarantee the weak user's communication quality. This assumption is widely used in the NOMA scheme [40, 41]. Consider that M_n users are allocated on SC_n with CRNNs order

$$|H_{1,n}| \geq |H_{2,n}| \geq \cdots \geq |H_{l,n}| \geq |H_{l+1,n}| \geq \cdots \geq |H_{M_n,n}|. \tag{3.6}$$

According to the optimal SIC decoding order, User l can successfully decode and remove the interference symbols from users $i > l$. However, the interference symbol from User i ($i < l$) cannot be removed and will be treated as noise by User l. Therefore, the SINR of User l with SIC at receiver can be written as

$$\widetilde{SINR}_{l,n} = \frac{p_{l,n} H_{l,n}}{1 + \sum_{i=1}^{l-1} p_{i,n} H_{l,n}}. \tag{3.7}$$

[2]Without causing notational confusion, $h_{l,n}$ is the channel gain of $UT_{l,n}$, while $H_{l,n} \triangleq |h_{l,n}|^2/\sigma_n^2$ denotes the channel response normalized by noise of $UT_{l,n}$

Then the data rate of the lth user on SC_n can be expressed as

$$R_{l,n}\left(p_{l,n}\right) = B_{sc}\log_2\left(1 + \frac{p_{l,n}H_{l,n}}{1 + \sum\limits_{i=1}^{l-1} p_{i,n}H_{l,n}}\right). \tag{3.8}$$

Therefore, the overall sum rate of the NOMA system can be written as

$$R = \sum_{n=1}^{N}\sum_{l=1}^{M_n} R_{l,n}\left(p_{l,n}\right) = \sum_{n=1}^{N} R_n\left(p_n\right). \tag{3.9}$$

3.2 Problem Formulation

In this section, we formulate the energy-efficient subchannel assignment and power allocation as an optimization problem. For energy-efficient communication, it is desirable to maximize the amount of transmitted data bits with a unit energy, which can be measured by energy efficiency. For each subchannel in the NOMA system, given assigned power p_n on SC_n and additional circuit power consumption P_c, the energy efficiency over SC_n is defined as

$$EE_n = \frac{R_n}{P_c + p_n}. \tag{3.10}$$

Then the overall energy efficiency of the system can be given by

$$EE = \sum_{n=1}^{N} EE_n. \tag{3.11}$$

For the downlink NOMA network, SIC technique is well investigated in [17, 20]. The implementation complexity of SIC at the receiver increases with the maximum number of the users allocated on the same subchannel. In order to keep the receiver complexity comparatively low, we consider a simple case where only two users are allocated on the same subchannel. This assumption is important because it also restricts the error propagation.

In this case, given that the two users sharing SC_n with CRNNs $|H_{1,n}| \geq |H_{2,n}|$, the sum rate of SC_n can be expressed as

$$R_n\left(p_n\right) = B_{sc}\log_2\left(1+\beta_n p_n H_{1,n}\right) + B_{sc}\log_2\left(1+\frac{\left(1-\beta_n\right)p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right) \tag{3.12}$$

where β_n is the power proportional factor for the two users on SC_n. Generally, β_n is used for the user who performs SIC on SC_n and $\beta_n \in (0,1)$. The optimal power proportional factor can be decided within our proposed subchannel assignment scheme. To obtain an energy-efficient resource allocation scheme for this system, we formulate the energy efficiency optimization problem as

$$\max_{p_n>0} \sum_{n=1}^{N} \frac{R_n\left(p_n\right)}{P_c + p_n} \tag{3.13}$$

$$\begin{aligned} \text{subject to } C1 &: R_{l,n}(p_n) \geq R_{\min} \\ C2 &: \sum_{n=1}^{N} p_n = P_s \end{aligned} \tag{3.14}$$

where $C1$ guarantees user minimum data rate constraint and $R_{\min}$ is denoted as minimum data rate determined by quality of service (QoS) requirement. The constraint $C2$ ensures the maximum BS power constraint. Since this optimization problem is non-convex and NP-hard, it is challenging to find the global optimal solution within polynomial time. To solve this problem efficiently, we will treat subchannel assignment and subchannel power allocation separately.

3.3 Subchannel Allocation

In this section, we investigate the energy-efficient matching algorithm for subchannel assignment in the NOMA network. For the optimization problem (3.13), it can be shown that the subchannel assignment and power allocation for subchannels are coupled with each other in terms of energy efficiency. Due to the considerable complexity of global optimum solution, we decouple subchannel assignment and power allocation to obtain a suboptimal

solution. We first propose a greedy subchannel-user matching algorithm by assuming equal power is allocated on each subchannel, in which each power proportional factor β_n is also determined to allocate different powers to the multiplexed users on the same subchannel. We define the parameter β_n as the proportional factor of assigned power to the user who performs SIC on SC_n. By decomposing the objective function into difference of convex functions, the suboptimal matching scheme for subchannel assignment is decided by a DC programming approach.

3.3.1 Subchannel Matching Problem Formulation

To describe the dynamic matching between the users and the subchannels, we consider subchannel assignment as a two-sided matching process between the set of M users and the set of N subchannels. Considering only two users can be multiplexed on the same subchannel due to the complexity of decoding, we assume $M = 2N$ [26]. We say UT_m and SC_n are matched with each other if UT_m is allocated on SC_n. Based on the perfect channel state information, the preference lists of the users and subchannels can be denoted by

$$\begin{aligned} PF\_UT &= [PF\_UT(1), \cdots, PF\_UT(m), \cdots, PF\_UT(M)]^T \\ PF\_SC &= [PF\_SC(1), \cdots, PF\_SC(n), \cdots, PF\_SC(N)]^T \end{aligned} \tag{3.15}$$

where $PF\_UT(m)$ and $PF\_SC(n)$ are the preference lists of UT_m and SC_n, respectively. We say UT_m prefers SC_i to SC_j if UT_m has higher channel gain on SC_i than that on SC_j, and it can be expressed as

$$SC_i\,(m) \succ SC_j\,(m)\,. \tag{3.16}$$

As an example, we consider four users and two subchannels with the following channel gain matrix

$$H = \left[\ 0.197,\ 0.778;\ 0.437,\ 0.143;\ 0.322,\ 0.545;\ 0.272,\ 0.478\ \right]$$

where row index denotes the users and column index denotes the subchannels. Therefore, we have the preference list of the users as

$$PF\_UT(1) = \begin{bmatrix} 2 & 1 \end{bmatrix}^T, PF\_UT(2) = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$$
$$PF\_UT(3) = \begin{bmatrix} 2 & 1 \end{bmatrix}^T, PF\_UT(4) = \begin{bmatrix} 2 & 1 \end{bmatrix}^T$$

and the preference list of the subchannels as

$$PF\_SC(1) = \begin{bmatrix} 2 & 3 & 4 & 1 \end{bmatrix}^T$$
$$PF\_SC(2) = \begin{bmatrix} 1 & 3 & 4 & 2 \end{bmatrix}^T.$$

We say SC_n prefers user set q_m to user set q_n (q_n, q_m is denoted as subsets of $\{1, 2, \cdots, M\}$) if the users in set q_m can provide higher energy efficiency than users in set q_n on SC_n, and we represent this scenario as

$$EE_n(q_m) > EE_n(q_n), q_m, q_n \subset \{UT_1, UT_2, \cdots, UT_{M_n}\}. \tag{3.17}$$

Matching theory has been studied in [67, 68], where various properties and types of preferences have been discussed. Based on the preference lists of users and subchannels, the subchannel assignment problem is formulated as a two-sided matching problem [67, 68].

Definition 1: *(Two-sided Matching)* Consider users and subchannels as two disjoint sets, $\boldsymbol{M} = \{1, 2, \cdots, M\}$ and $\boldsymbol{N} = \{1, 2, \cdots, N\}$. A two-to-one, two-sided matching $\mathcal{M}$ is a mapping from all the subsets of users $\boldsymbol{M}$ into the subchannel set $\boldsymbol{N}$ satisfying $UT_m \in \boldsymbol{M}$ and $SC_n \in \boldsymbol{N}$

1) $\mathcal{M}(UT_m) \in \boldsymbol{N}$.

2) $\mathcal{M}^{-1}(SC_n) \subseteq \boldsymbol{M}$.

3) $|\mathcal{M}(UT_m)| = 1$, $|\mathcal{M}^{-1}(SC_n)| = 2$.

4) $SC_n \in \mathcal{M}(UT_m) \Leftrightarrow UT_m \in \mathcal{M}^{-1}(SC_n)$.

Condition 1) states that each user matches with one subchannel, and Condition 2) represents each subchannel can be matched with a subset of users. Condition 3) states that the number of users can be allocated on each subchannel is limited to two. Condition 4) means that UT_m and SC_n are matched with each other.

Definition 2: *(Preferred Matched Pair)* Given a matching $\mathcal{M}$ that $UT_m \notin \mathcal{M}^{-1}(SC_n)$ and $SC_n \notin \mathcal{M}(UT_m)$. If $EE_n\left(S_{new}\right) > EE_n\left(\mathcal{M}^{-1}(SC_n)\right)$ where $S_{new} \subseteq \{UT_m\} \cup S$ and $S = \mathcal{M}^{-1}(SC_n)$, where S is the user set that has been assigned to SC_n, S_{new} becomes the preferred users set for subchannel n and $(UT_m,\ SC_n)$ is a preferred matched pair. Based on the above definition, we will describe in Section 3.3.2 the matching action between the users and the subchannels. If each subchannel has to select the best subset of users to allocate, it will cause considerable complexity especially when the number of users is large. Because the optimal solution requires to search all the possible combinations of the users to maximize energy efficiency. To reduce the complexity, a suboptimal matching algorithm is proposed for subchannel assignment as follows.

3.3.2 Suboptimal Matching for Subchannel Assignment Algorithm in NOMA

In this subsection, we propose a suboptimal matching algorithm for subchannel assignment. The main idea of this matching model is that each user sends the matching request to its most preferred subchannel according to its preference list. This preferred subchannel has the right to accept or reject the user according to energy efficiency that the all users can provide on this subchannel. Based on the equal power allocation across subchannels, the user selection algorithm is a process of finding the preferred matching pair for each user and subchannel.

Algorithm 1 describes the proposed low-complexity suboptimal matching scheme for a subchannel assignment (SOMSA) scheme to maximize the system energy efficiency. This algorithm includes initialization and matching procedures. In the initialization step, preferences lists of subchannels and users are decided according to the channel state information,

Algorithm 1: Suboptimal Matching for Subchannel Assignment

Initialize the matched list $S_{Match}(n)$ to record users matched on SC_n for all the subchannels $\forall n \in \{1, 2, \cdots, N\}$.
Initialize preference lists $PF\_UT\,(m)$ and $PF\_SC\,(n)$ for all the users $\forall m \in \{1, 2, \cdots, M\}$ and all the subchannels $\forall n \in \{1, 2, \cdots, N\}$ according to CRNNs.
Initialize the set of unmatched users $S_{UnMatch}$ to record users who has not been allocated to any subchannel.
while $\{S_{UnMatch}\}$ is not empty **do**
  for $m = 1$ to M **do**
    Each user sends matching request to its most preferred subchannel $\hat{n}$ according to $PL\_UT\,(m)$.
    if $|S_{Match}\,(\hat{n})| < 2$ **then**
      Subchannel $\hat{n}$ adds user m to $S_{Match}\,(\hat{n})$, and removes user m from $\{S_{UnMatch}\}$
    end if
    if $|S_{Match}\,(\hat{n})| = 2$ **then**
      a) Find power proportional factor β_n for every two users in S_{q_m}, $S_{q_m} \subset \{S_{match}(\hat{n}), m\}$ by using (3.18), or exhaustive search method or DC programming algorithm in Section 3.4.1.
      b) Subchannel $\hat{n}$ selects a set of 2 users S_{q_m} satisfying maximum energy efficiency $E_{\hat{n}}(q_m) \geq E_{\hat{n}}(q_n), q_m, q_n \subset \{S_{match}(\hat{n}), m\}$.
      c) Subchannel $\hat{n}$ sets $S_{match}(\hat{n}) = q_m$, and rejects other users. Remove the allocated users from $\{S_{UnMatch}\}$, add the unallocated user to $\{S_{UnMatch}\}$.
      d) The rejected user removes subchannel from their preference lists.
    end if
  end for
end while

and $S_{Match}(n)$, $\forall n \in \{1, 2, \cdots, N\}$ and $S_{UnMatch}$ are initialized to record the allocated users on SC_n and unallocated users of the system, respectively. In the matching procedure, at each round, each user sends the matching request to its most preferred subchannel. According to the preferred list of each user ($PF\_UT(m)$, $\forall m \in \{1, 2, \cdots, M\}$) which is a list of subchannels ordered by decreasing channel gains, the mth user will find the first non-zero entry in $PF\_UT(m)$ and send matching request to the corresponding subchannel. The subchannel accepts the user directly if the number of allocated users on this subchannel is less than two. When the number of the allocated users equals to two, only the subset of users that can provide higher energy efficiency will be accepted or it will be rejected. This matching process will terminate when there is no user left to be matched. After that, the allocated user and the corresponding subchannels in the preference list are set to zero. The proposed SOMSA converges to a stable matching after a limited number of iterations [68].

3.3.3 Power Ratio Factor Determination

In Algorithm 1, it is required to determine the power proportional factor β_n for every two subchannel users. In this section, we will first review the existing fractional transmit power allocation scheme and the exhaustive searching method. Then we will introduce a new energy-efficient power allocation algorithm based on DC programming in Section 3.4.1. It will be shown in the simulation results that the new algorithm can result in improved energy efficiency.

Fractional transmit power allocation

According to the SINR expression in (3.7), the transmit power allocation to one user affects the achievable sum rate as well as the energy efficiency on each subchannel. Due to its low computational complexity, FTPA is widely adopted in OFDMA systems and NOMA systems [20, 26]. In the FTPA scheme, the transmit power of UT_m on SC_n is allocated

according to the channel gains of all the multiplexed users on SC_n, which is given as

$$p_{l,n} = p_n \frac{{H_{l,n}}^{-\alpha}}{\sum_{i=1}^{M_n} {H_{i,n}}^{-\alpha}} \tag{3.18}$$

where $\alpha\,(0 \leq \alpha \leq 1)$ is a decay factor. In the case $\alpha = 0$, it corresponds to equal power allocation among the allocated users. From (3.18), it is clear that when α increases, more power is allocated to the user with poorer CRNN. Note that the same decay factor should be applied to all subchannels and transmission times.

Exhaustive searching method

In finding power proportional factor β_n, the method of exhaustion can also be exploited for $\beta_n \in (0, 1)$. The optimal value can be found through searching all β_n values in $(0, 1)$ using a sufficiently small step size. Therefore, the optimal power proportional factors for the multiplexed users can be obtained. However, the computational complexity of the exhaustion method is much higher than FTPA. Therefore, in the following, we consider a suboptimal but efficient DC programming to allocate power among multiple users to maximize the energy efficiency.

3.3.4 Complexity Analysis

The optimal subchannel assignment scheme can only be obtained by searching over all possible combinations of the users and selecting the one that maximizes the system energy efficiency. If we have M users and N subchannels ($M = 2N$). The scheduler needs to search $\frac{(2N)!}{2^N}$ combinations. The time complexity of exhaustive searching is $O(\frac{(2N)!}{2^N})$. In order to compare the complexity of different algorithms, we take natural logarithm of the complexity. The logarithm complexity is $O(\ln((2N)!) - N) = O(\ln((2N)!))$. By using the Stirling's formula, $\ln(n!) = n \ln n - n + O(\ln(n))$, the logarithm complexity of the exhaustive searching can be written as $O(N \ln N)$. In the SOMSA algorithm, the complexity of the worst case is $O(N^2)$. Taking natural logarithm of the complexity, the logarithm complexity

is $O(\ln N)$. Since $O(\ln N) < O(N \ln N)$ and actual complexity of SOMSA is much less than the complexity of the worst case, the complexity of SOMSA algorithm is much less than the optimal subchannel assignment scheme. It can be shown that for a small number of users ($M = 4$), the SOMSA will yield the identical results from the exhaustive search.

3.4 Power Allocation

As mentioned in Section 3.3, equal power allocation is assumed across subchannels in SOMSA. In order to further improve the energy efficiency of the NOMA system, we consider to obtain the energy-efficient subchannel power allocation instead of equal power allocation. In this section, we introduce the DC programming approach and discuss its application in finding power proportional factors as well as power allocation across subchannels.

3.4.1 DC Programming

DC programming approach has been studied recently to solve non-convex optimization problems [69]. It is shown that DC programming can be applied if the objective function can be written as a minimization of a difference of two convex functions, which is represented as

$$\min_{\boldsymbol{x} \in \chi} q(\boldsymbol{x}) = f(\boldsymbol{x}) - g(\boldsymbol{x}) \tag{3.19}$$

where $\boldsymbol{x} = [x_1, x_2, \cdots x_L]^T$ and χ is a convex set; $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are continuous, convex or quasi-convex [69]. In general, the problem defined in (3.19) is non-convex. However, it can be solved suboptimally by using Algorithm 2. The key idea of Algorithm 2 is to convert a non-convex problem to convex subproblems by using successive convex approximations. In this algorithm, ε is the difference tolerance and the term $-g(\boldsymbol{x})$ in the objective function (3.19) is replaced by $-g\left(\boldsymbol{x}^{(k)}\right) - \nabla g^T\left(\boldsymbol{x}^{(k)}\right)\left(\boldsymbol{x} - \boldsymbol{x}^{(k)}\right)$ in (3.20). The convex optimization problem in (3.20) can be solved by using standard algorithms from convex optimization theory [66, 70, 71], i.e., interior point method and sequential quadratic programming.

Algorithm 2: Iterative, Suboptimal Solution for DC Problems [70]

Initialize $\boldsymbol{x}^{(0)}$, set iteration number $k = 0$.
while $\left|q\left(\boldsymbol{x}^{(k+1)}\right) - q\left(\boldsymbol{x}^{(k)}\right)\right| > \varepsilon$ **do**
Define convex approximation of $q^{(k)}(\boldsymbol{x})$ as

$$\hat{q}^{(k)}(\boldsymbol{x}) = f(\boldsymbol{x}) - g\left(\boldsymbol{x}^{(k)}\right) - \nabla g^T\left(\boldsymbol{x}^{(k)}\right)\left(\boldsymbol{x} - \boldsymbol{x}^{(k)}\right) \tag{3.20}$$

Solve the convex problem

$$\boldsymbol{x}^{(k+1)} = \arg\min_{\boldsymbol{x}\in\chi} \hat{q}^{(k)}(\boldsymbol{x}) \tag{3.21}$$

$k \leftarrow k + 1$
end while

The convergence of Algorithm 2 can be easily proved by

$$q\left(\boldsymbol{x}^{(k)}\right) = \hat{q}^{(k)}\left(\boldsymbol{x}^{(k)}\right) \geq \hat{q}^{(k)}\left(\boldsymbol{x}^{(k+1)}\right) \geq q\left(\boldsymbol{x}^{(k+1)}\right) \tag{3.22}$$

where $q\left(\boldsymbol{x}^{(k)}\right) = \hat{q}^{(k)}\left(\boldsymbol{x}^{(k)}\right)$ is the kth iteration step, and $\hat{q}^{(k)}\left(\boldsymbol{x}^{(k)}\right) \geq \hat{q}^{(k)}\left(\boldsymbol{x}^{(k+1)}\right)$ can be obtained by (3.21). Therefore, $q\left(\boldsymbol{x}^{(k)}\right)$ monotonically decreases when k increases. Under an additional assumption that $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are continuous and differentiable on the constraint set. In this case, Algorithm 2 always returns a point of $q(\boldsymbol{x})$, which may not be the global optimal solution [69].

3.4.2 Power Proportional Factor

Considering two users UT_1 and UT_2 who are to be multiplexed over SC_n with CRNNs $H_{1,n} \geq H_{2,n}$. According to the principle of SIC decoding sequences, UT_1 can cancel the interfering power term of UT_2, whereas UT_2 treats the symbol power UT_1 as noise. The problem of finding β_n to maximize energy efficiency of SC_n can be formulated as

$$\max_{\beta_n\in(0,1)} \frac{B_{sc}\log_2\left(1 + \beta_n p_n H_{1,n}\right)}{P_c + p_n} + \frac{B_{sc}\log_2\left(1 + \frac{(1-\beta_n)p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right)}{P_c + p_n} \tag{3.23}$$

which can be rewritten as

$$\max_{\beta_n \in (0,1)} \frac{B_{sc}\log_2\left(1+\beta_n p_n H_{1,n}\right) + B_{sc}\log_2\left(\frac{1+p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right)}{P_c + p_n}. \tag{3.24}$$

In order to use the DC programming approach, we can convert (3.24) to DC representation

$$\min_{\beta_n \in (0,1)} -\frac{B_{sc}\log_2\left(1+\beta_n p_n H_{1,n}\right)}{P_c + p_n} - \frac{B_{sc}\log_2\left(\frac{1+p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right)}{P_c + p_n} \tag{3.25}$$

or

$$\min_{\beta_n \in (0,1)} \left(f\left(\beta_n\right) - g\left(\beta_n\right)\right) \tag{3.26}$$

where $f\left(\beta_n\right) = -\frac{B_{sc}\log_2(1+\beta_n p_n H_{1,n})}{P_c+p_n}$ and $g\left(\beta_n\right) = \frac{B_{sc}\log_2\left(\frac{1+p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right)}{P_c+p_n}$, and both terms are convex functions with respect to β_n because $\nabla^2 f\left(\beta_n\right) > 0$ and $\nabla^2 g\left(\beta_n\right) > 0$. Therefore, the DC programming approach can be used to find β_n by replacing $\boldsymbol{x}$ with β_n in Algorithm 2.

3.4.3 Subchannel Power Allocation by DC Programming

Given the subchannel-user matching scheme and power proportional factors on different subchannels by Algorithm 1, the optimization problem in (3.13) can rewritten as

$$\max_{p_n \geq 0} \sum_{n=1}^{N} \left\{ \frac{B_{sc}\log_2\left(1+\beta_n p_n H_{1,n}\right)}{P_c + p_n} + \frac{B_{sc}\log_2\left(\frac{1+p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right)}{P_c + p_n} \right\} \tag{3.27}$$

$$\text{subject to } C1: R_{l,n}(p_n) \geq R_{\min};\ C2: \sum_{n=1}^{N} p_n = P_s \tag{3.28}$$

where $R_{l,n}(p_{l,n})$ is defined in (3.8). Since $R_{l,n}(p_{l,n})$ is a linear function respect to the assigned power p_n on SC_n. The constraint $C1$ can be converted to $p_n > p_{n,\min}$, where $p_{n,\min}$ is the minimum assigned power on SC_n and it is determined by $R_{\min}$. Condition $C2$ in (3.28) guarantees BS power constraint. Note that the optimization problem in (3.27) is

non-convex with respect to p_n. However, the representation of (3.27) is similar to the DC problem representation. Thus (3.27) can be rewritten as

$$\begin{aligned} &\min_{p_n \geq 0} -\sum_{n=1}^{N} \left\{ \frac{B_{sc}\log_2\left(1+\beta_n p_n H_{1,n}\right)}{P_c+p_n} + \frac{B_{sc}\log_2\left(\frac{1+p_n H_{2,n}}{1+\beta_n p_n H_{2,n}}\right)}{P_c+p_n} \right\} \\ &\min_{p_n \geq 0} F\left(\mathbf{P}\right) - G\left(\mathbf{P}\right) \end{aligned} \tag{3.29}$$

where $\mathbf{P} = [p_1, p_2, \cdots, p_n, \cdots, p_N]^T$ represents the allocated powers on the subchannels, and

$$\begin{aligned} F\left(\mathbf{P}\right) &= -\sum_{n=1}^{N} \frac{B_{sc}\log_2\left(1+\beta_n p_n H_{1,n}\right)}{P_c+p_n} - \sum_{n=1}^{N} \frac{B_{sc}\log_2\left(1+p_n H_{2,n}\right)}{P_c+p_n}; \\ G\left(\mathbf{P}\right) &= -\sum_{n=1}^{N} \left(\frac{B_{sc}\log_2\left(1+\beta_n p_n H_{2,n}\right)}{P_c+p_n} \right). \end{aligned}$$

Problem (3.27) can be written as

$$\min_{\mathbf{P} \succ \mathbf{0}} Q\left(\mathbf{P}\right) = \min_{\mathbf{P} \succ \mathbf{0}} F\left(\mathbf{P}\right) - G\left(\mathbf{P}\right)$$

$$\text{subject to } C1 : \mathbf{P} \succ \mathbf{P}_{\min};\ C2 : \|\mathbf{P}\|_1 = P_s \tag{3.30}$$

where $\mathbf{P}_{\min} = [p_{1,\min}, p_{2,\min}, \cdots, p_{n,\min}, \cdots, p_{N,\min}]^T$ and $\mathbf{P} \succ \mathbf{P}_{\min}$ means all the elements in $\mathbf{P}$ are larger than the corresponding elements in $\mathbf{P}_{\min}$, $p_n > p_{n,\min}$. *Proposition* 1 proves convexity of $F\left(\mathbf{P}\right)$ and $G\left(\mathbf{P}\right)$, Therefore, the DC programming approach can be applied to realize energy-efficient power allocation using Algorithm 3. Once the power allocation over subchannels is obtained, we replace the equal power allocation with our new power allocation scheme to achieve higher energy efficiency of the system.

In Algorithm 3, $\nabla G\left(\mathbf{P}^{(k)}\right)$ is the gradient of $G(\mathbf{P})$ at the point $\mathbf{P}^{(k)}$ and it is calculated

Algorithm 3: DC Programming Algorithm for Power Allocation across Subchannels

Initialize $\mathbf{P}^{(0)}$, set iteration number $k = 0$. The Objective function $Q(\mathbf{P})$, convex functions $F(\mathbf{P})$ and $G(\mathbf{P})$.

while $\left|Q\left(\mathbf{P}^{(K+1)}\right) - Q\left(\mathbf{P}^{(k)}\right)\right| > \varepsilon$ **do**

Define convex approximation of $G^{(k)}(\mathbf{P})$ at $\mathbf{P}^{(k)}$ as

$$Q^{(k)}(\mathbf{P}) = F(\mathbf{P}) - G\left(\mathbf{P}^{(k)}\right) - \nabla G^T\left(\mathbf{P}^{(k)}\right)\left(\mathbf{P} - \mathbf{P}^{(k)}\right) \tag{3.31}$$

Solve the convex problem

$$\mathbf{P}^{(k)} = \underset{p_n \geq p_{n,\min},\ \|\mathbf{P}\|_1 = \mathrm{P}_s}{\arg\min} Q^{(k)}(\mathbf{P}) \tag{3.32}$$

$k \leftarrow k + 1$

end while

by

$$\nabla G\left(\mathbf{P}^{(k)}\right) = \sum_{n=1}^{N} \frac{B_{sc}\log_2(1 + \beta_n p_n H_{2,n}) - (P_c + p_n)\frac{\beta_n H_{2,n}}{(1+\beta_n p_n H_{2,n})\ln 2}}{(P_c + p_n)^2}. \tag{3.33}$$

Since (3.32) and the power domain are convex, problem (3.33) can be solved by either the interior point method or the sequential quadratic programming. In order to use the DC programming approach, the quasi-convexity of $F(\mathbf{P})$ and $G(\mathbf{P})$ needs to be established. It is easy to show that

$$f(\mathbf{P}) = -\sum_{n=1}^{N} B_{sc}\log_2(1 + \beta_n p_n H_{1,n}) - \sum_{n=1}^{N} B_{sc}\log_2(1 + p_n H_{2,n})$$

and

$$g(\mathbf{P}) = -\sum_{n=1}^{N} B_{sc}\log_2(1 + \beta_n p_n H_{2,n})$$

are convex since $\nabla^2 f(\mathbf{P})$ and $\nabla^2 g(\mathbf{P})$ are positive semi-definite matrices.

Proposition 1: If

$$-f_1(p_n) = B_{sc}\log_2(1 + \beta_n p_n H_{1,n}) + B_{sc}\log_2(1 + p_n H_{2,n})$$

and

$$-g_1(p_n) = B_{sc}\log_2(1 + \beta_n p_n H_{2,n})$$

are strictly concave in p_n, $-F_1(p_n) = \frac{-f_1(p_n)}{P_c+p_n}$ and $-G_1(p_n) = \frac{-g_1(p_n)}{P_c+p_n}$ are quasi-concave with constant P_c. Inspired by [72], we can prove *Proposition* 1 as follows.

Proof: Denote the α-sublevel sets of function $-F_1(p_n)$ as

$$S_\alpha = \{p_n > 0 | -F_1(p_n) \geq \alpha\}. \tag{3.34}$$

$-F_1(p_n)$ is strictly quasi-concave if and only if S_α is strictly convex for any α. In this case, when $\alpha < 0$, there are no points satisfying $-F_1(p_n) = \alpha$. Therefore, S_α is strictly convex when $\alpha \leq 0$. When $\alpha > 0$, we can rewrite S_α as $S_\alpha = \{p_n > 0 | \alpha(P_c + p_n) + f_1(p_n) \leq 0\}$. Since $f(p_n)$ is strictly convex in p_n, S_α is therefore also strictly convex. Hence, $-F_1(p_n)$ and $-G_1(p_n)$ are strictly quasi-concave. Therefore, $F_1(p_n)$ and $G_1(p_n)$ are quasi-convex. As a result, $F(\mathbf{P})$ and $G(\mathbf{P})$ are quasi-convex. ■

3.5 Simulation Results

In this section, Monte Carlo simulation results are presented to evaluate the performance of the proposed resource allocation algorithms for NOMA systems. In the simulations, we consider one base station located in the cell center and all the user terminals are uniformly distributed in a circular range with radius of 500 m. We set the minimum distance among all the users to be 40 m, and the minimum distance from users to BS is 50 m. The bandwidth is 5 MHz. Let M users be randomly distributed in the cell. In NOMA systems, to reduce demodulating complexity of the SIC receiver, we consider each subchannel is only allocated with two users. In OFDMA schemes, each user can only be assigned to one subchannel. In the simulations, we compare the performance of NOMA systems with OFDMA systems, both with resource allocation algorithms. For the subchannel power

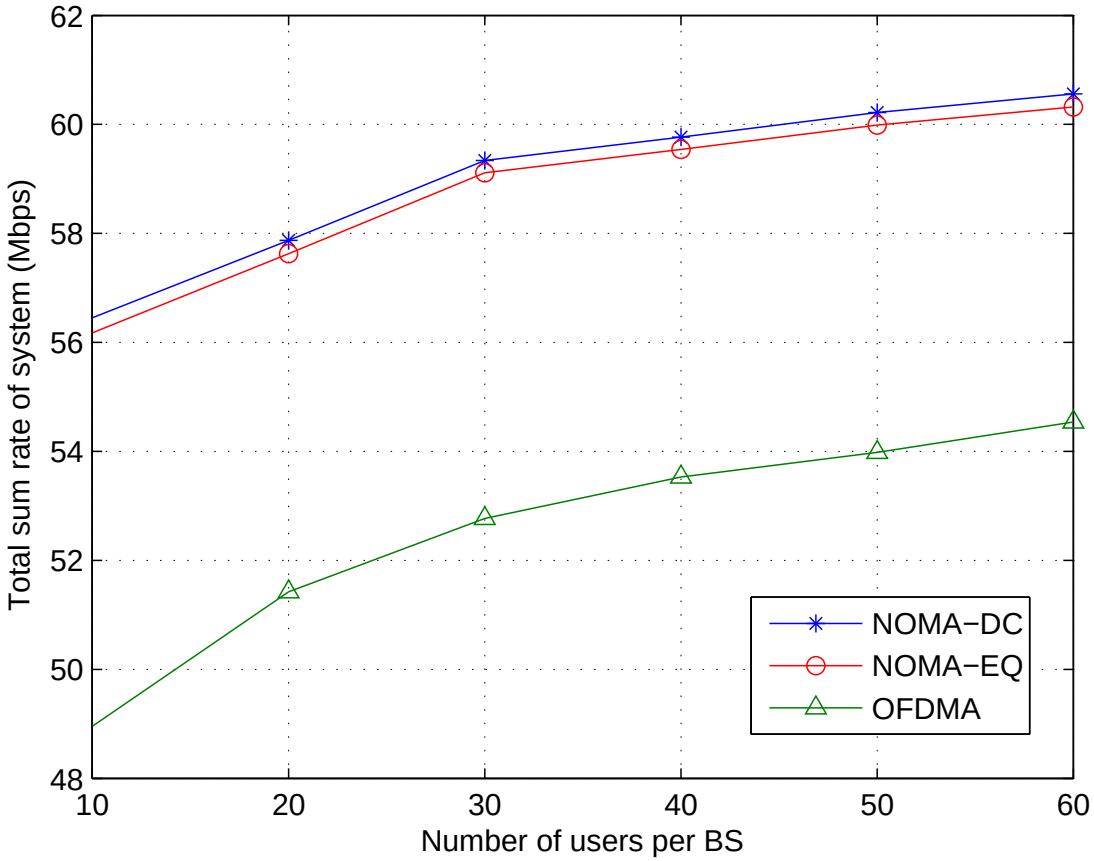

Figure 3.2: Sum rate of the system versus different number of users.

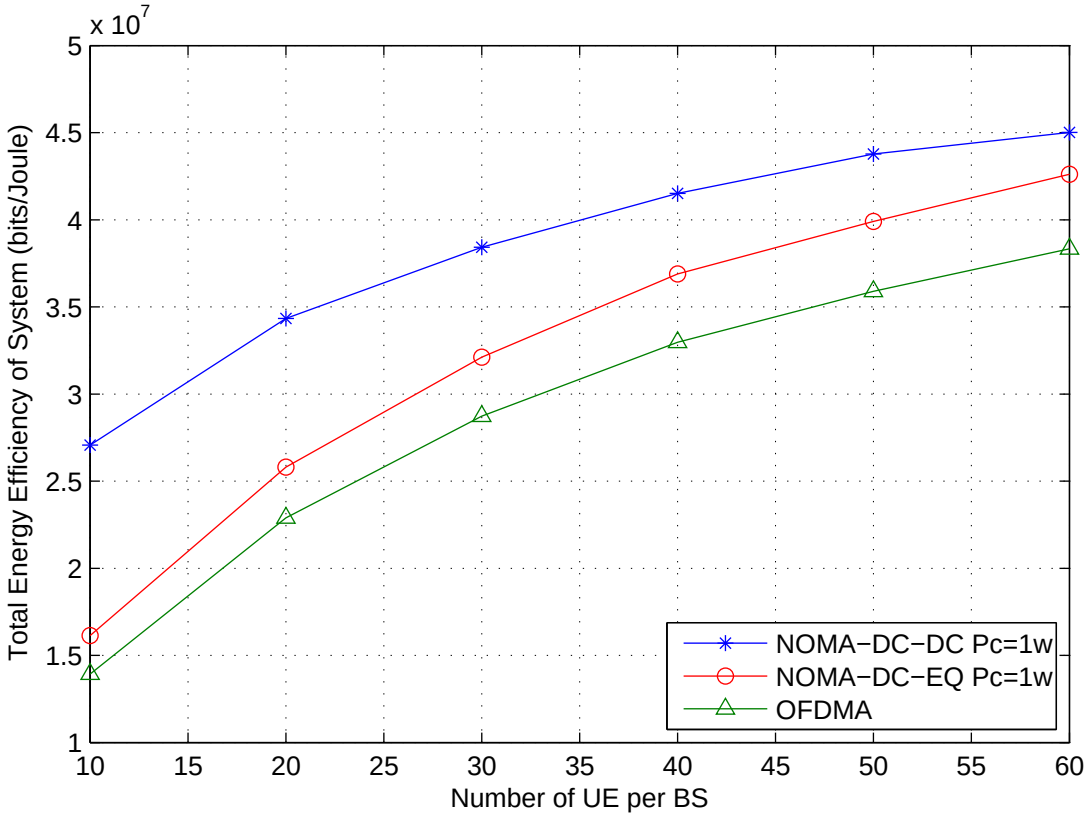

Figure 3.3: Energy efficiency of the system versus different number of users.

allocation, we compare our proposed suboptimal algorithm with equal power allocation scheme based on our proposed subchannel assignment scheme. FTPA for multiplexed users on subchannel is also compared with our proposed algorithms. In our simulations, we set BS peak power, P_s, to be 41 dBm and circuit power consumption $P_c = 1$ W [73]. The maximum number of users is 60 and $\sigma_n^2 = \frac{BW}{N} N_0$, where $N_0 = -174$ dBm/Hz is the AWGN power spectral density. In the simulations, we set the value of α as 0.4 [40].

In Fig. 3.2, the performance of the total sum rate is evaluated with the number of users M (M varies from 10 to 60). We set difference tolerance $\varepsilon = 0.01$, and the bandwidth is limited to 5 MHz. It is shown that the total sum rate increases when the number of the users grows. As the number of users grows larger, the sum rate continues to increase, but the rate of growth becomes slower, as expected from the Shannon's formula in calculating

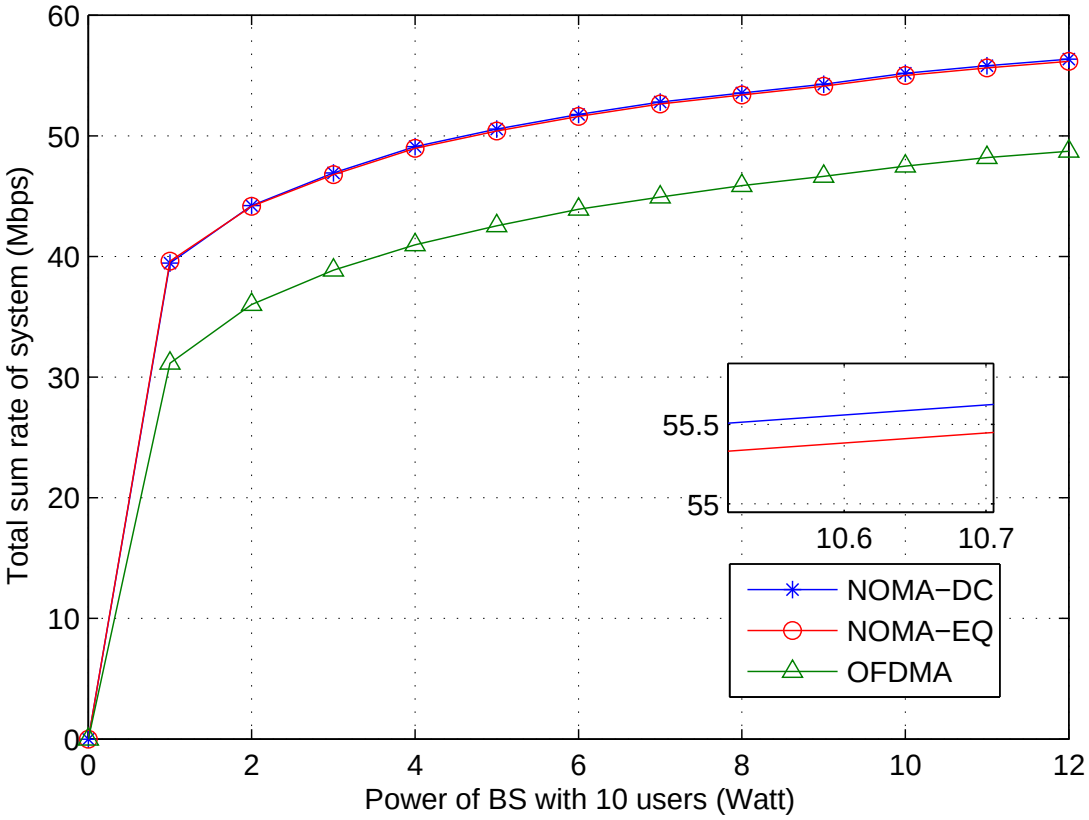

Figure 3.4: Sum rate of the system versus BS power.

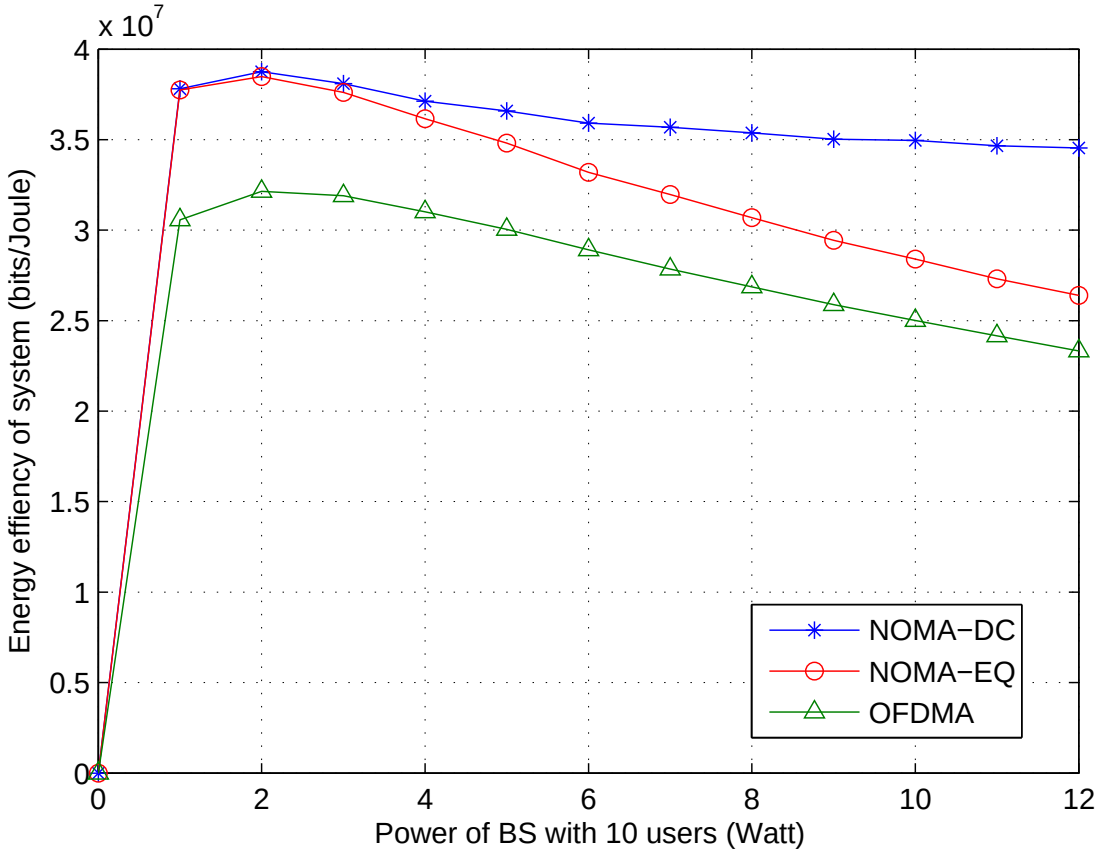

Figure 3.5: Energy efficiency of the system versus the maximum BS power P_s.

the sum rate. From Fig. 3.2, we observe that the performance of NOMA system with the proposed resource allocation algorithms, including subchannel assignment and power allocation, is much better than the OFDMA scheme. For example, when the number of users is 30, the sum rate of the proposed algorithm (NOMA-DC) [3] is 12.5% more than that of the OFDMA scheme, and the sum rate of equal power allocation (NOMA-EQ) is 11.9% more than that of the OFDMA scheme. That is because in the OFDMA scheme, each subchannel can only be used by one user. As a result, BS cannot fully use the spectrum resources. For different subchannel power allocation schemes, the sum rate of NOMA-DC is higher than that of NOMA-EQ.

Figure 3.3 shows the energy efficiency versus the number of users with the same constraints of Fig. 3.2. It can be observed that the energy efficiency also increases when the number of users grows. The trend of curve is similar to the sum rate curves due to the energy efficiency expression. From this figure, the performance of our proposed subchannels and power allocation is much more energy-efficient than the OFDMA scheme. Our proposed subchannel power allocation through the DC programming achieves better performance than the equal power allocation. When the number of users is 30, the energy efficiency of NOMA-DC is 33% more than that of the OFDMA scheme and 19% more than NOMA-EQ.

In Fig. 3.4, the performance of the total sum rate versus BS power with a fixed circuit power of $P_c = 1$ W, the total number of users $M = 10$, and the BS power is from 1 W to 12 W. In Fig. 3.4, the sum rate of the system increases as the BS power grows. In NOMA systems, our proposed algorithm using DC for subchannel power allocation performs better than equal power allocation. Both algorithms outperform the OFDMA system.

Figure 3.5 illustrates the total energy efficiency versus BS power with the fixed circuit power of $P_c = 1$ W, and the number of users is $M = 10$, and the BS power ranges from 1 W to 12 W. It shows that the total energy efficiency first increases from 0 when BS transmit power increases. After the power reaches a certain level, the total energy efficiency

[3] NOMA-DC uses DC programming to allocate power across subchannels.

begins to decrease. That is because there is a tradeoff between transmission capacity and power consumption for the energy-efficient power allocation. From Fig. 3.5, it is seen that NOMA-DC can achieve better performance than NOMA-EQ and the OFDMA scheme. For larger BS power levels, NOMA-DC achieves much better performance than NOMA-EQ and OFDMA.

Figure 3.6 shows the total energy efficiency versus circuit power to BS power ratio P_c/P_s. The system energy efficiency decreases when the ratio P_c/P_s increases. With the fixed BS power of 12 W, the system achieves less energy efficiency when the circuit power increases. According to the definition of energy efficiency, its value will become smaller when P_c increases. However, the NOMA system equipped with the proposed resource allocation algorithms still outperforms the OFDMA system.

In Fig. 3.7, FTPA among multiplexed users with equal subchannel power scheme is compared with the NOMA-DC, NOMA-EQ and OFDMA schemes. Fig. 3.7 shows that the energy efficiency increases as the number of users grows. NOMA-DC-DC[4] performs the best among those schemes. When user number is 20, the energy efficiency of NOMA-DC-DC is 35% more than that of the OFDMA scheme and 30% more than NOMA-FTPA-EQ[5]. NOMA-DC-EQ[6]achieves 12% more than the OFDMA scheme in terms of energy efficiency.

3.6 Summary

By assigning only two users to the same subchannel, we proposed energy-efficient resource allocation algorithms for a downlink NOMA wireless network. These algorithms include subchannel assignment, power proportional factors determination for multiplexed users and power allocation across subchannels. By formulating the subchannel assignment problem as a two-sided matching problem, we proposed the SOMSA algorithm to maximize

[4]NOMA-DC-DC uses DC programming approach to determine the power proportional factors and allocate different powers across the subchannels.

[5]NOMA-FTPA-EQ uses FTPA to determine the power proportional factors, and equal power allocation across the subchannels.

[6]NOMA-DC-DC uses DC programming approach to determine the power proportional factors and equal power allocation across the subchannels.

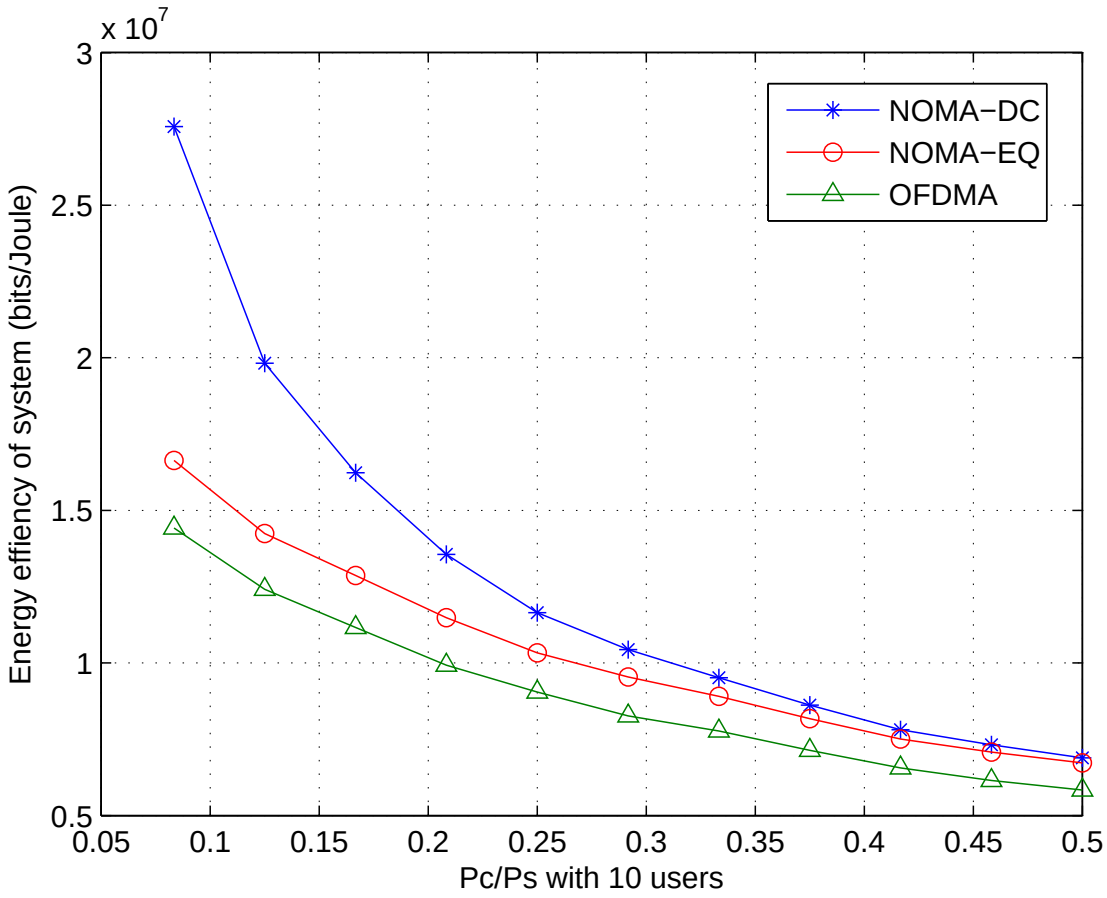

Figure 3.6: Energy efficiency of the system versus P_c/P_s.

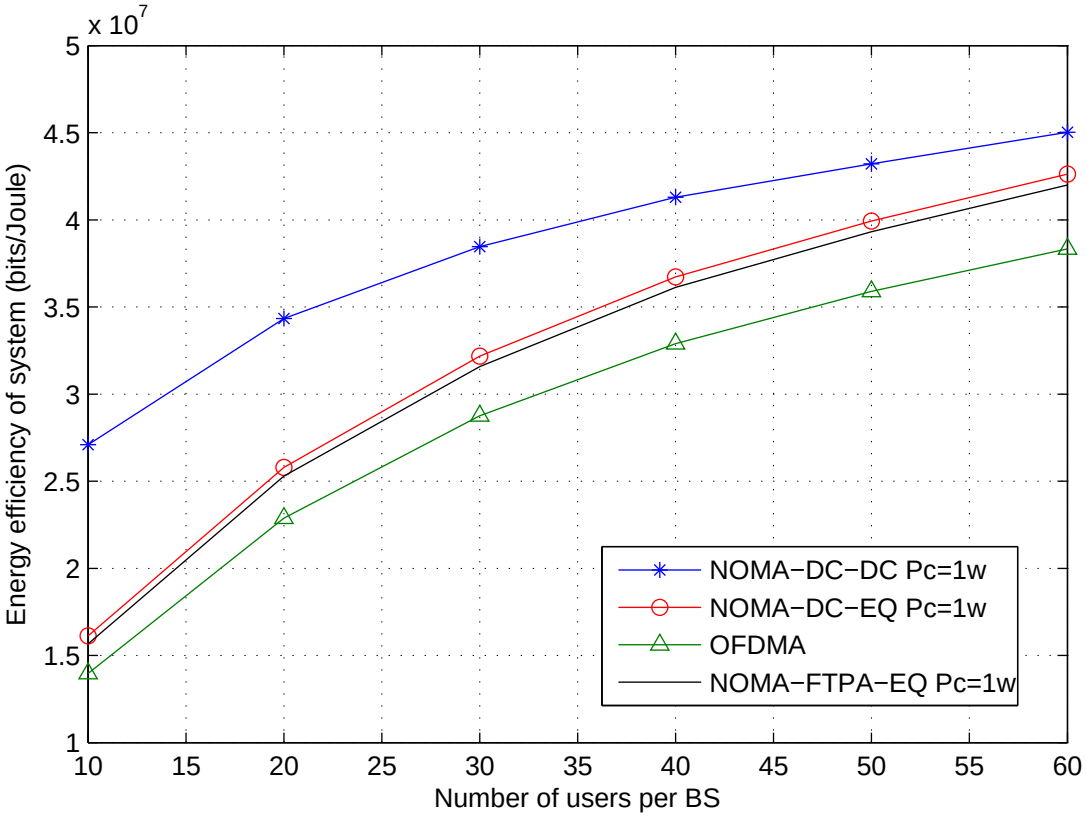

Figure 3.7: Energy efficiency of the system versus different number of users.

the system energy efficiency. Power proportional factors for the multiplexed users on each subchannel are determined by SOMSA. In the power allocation across subchannels scheme, since the objective function is non-convex, DC programming was utilized to approximate the non-convex optimization problem into the convex subproblem. Therefore, a suboptimal power allocation across subchannels was obtained by solving the convex subproblems iteratively. Based on the resource scheduling from the proposed SOMSA algorithm, further improvement in the system energy efficiency was achieved by the proposed subchannel power allocation scheme. Through extensive simulations, the performance of the proposed algorithms for resource allocation was compared with the OFDMA system. It was shown that the total sum rate and energy efficiency of NOMA system are both higher than the OFDMA scheme. The proposed power allocation for subchannel users outperforms the FTPA scheme. Moreover, the proposed subchannel power allocation achieves better performance than the equal power allocation scheme.

Chapter 4

Energy Efficient Resource Allocation for Downlink NOMA Network with Imperfect CSI

In this chapter, we investigate energy efficiency improvement for a downlink NOMA single-cell network by considering imperfect CSI. The energy efficient resource scheduling problem is formulated as a non-convex optimization problem with the constraints of outage probability limit, the maximum power of the system, the minimum user data rate and the maximum number of multiplexed users sharing the same subchannel. Different from previous works, the maximum number of multiplexed users can be greater than two, and the imperfect CSI is first studied for resource allocation in NOMA. The probabilistic mixed problem is first transformed into a non-probabilistic problem. To balance the system performance and the computational complexity, a new suboptimal user scheduling scheme is proposed to schedule users on different subchannels. Based on the user scheduling scheme, the optimal power allocation expression is derived by the Lagrangian approach. By transforming the fractional-form problem into an equivalent subtractive-form optimization problem, an iterative power allocation algorithm is proposed to maximize the system energy efficiency. Finally, we present some selected numerical results to demonstrate that the proposed user scheduling algorithm closely attains the optimal performance.

4.1 System Model

We consider a downlink single cell NOMA network where a BS is located in the cell center and M users are uniformly distributed within the cell. We define M and N as the numbers of users and subchannels, respectively. The BS transmits its signals to M users through N subchannels. We denote n as the index for the subchannel where $n \in \{1, 2, \cdots, N\}$. The total bandwidth B is shared by N subchannels and each subchannel occupies a bandwidth of $B_{sc} = B/N$. According to the NOMA protocol, multiple users can share the same subchannel. We denote the set of users on subchannel n (SC_n) as U_n and denote the number of users on SC_n as $M_n \triangleq |U_n|, n \in \{1, 2, \cdots, N\}$ where $M = M_1 + M_2 + \cdots + M_N$. We denote m as the index of the mth user multiplexed on each subchannel. P_T is denoted as the total transmitted power of the system and let $\sum_{n}^{N} \sum_{m}^{M_n} p_{m,n} \leq P_T$ where $p_{m,n}$ is the allocated power to the mth user (UE_m) on SC_n.

We assume that M_n users are multiplexed on the subchannel n. The signal transmitted by the BS through SC_n can be expressed as

$$x_n = \sum_{m=1}^{M_n} \sqrt{p_{m,n}} s_m \tag{4.1}$$

where s_m is the modulated symbol, which represents the transmitted symbol of UE_m on SC_n. We define $h_{m,n}$ as the channel gain from the BS to the mth user on SC_n. Without loss of generality, we assume channel gains of the M_n users on SC_n are perfectly known and are sorted as $|h_{M_n,n}| \geq |h_{M_n-1,n}| \geq \cdots \geq |h_{m,n}| \geq \cdots \geq |h_{1,n}|$. In the NOMA system, by exploiting the SIC technique at the receivers, the signal received by UE_m on SC_n can be formulated as

$$\begin{aligned} y_{m,n} &= h_{m,n} x_n + z_{m,n} \\ &= h_{m,n} \sqrt{p_{m,n}} s_m + h_{m,n} \sum_{i=m+1}^{M_n} \sqrt{p_{i,n}} s_i + z_{m,n} \end{aligned} \tag{4.2}$$

where $h_{m,n} = D_{m,n} g_{m,n}$, where $D_{m,n} = d_{m,n}^{-\frac{1}{2}}$ and $g_{m,n} \sim \mathcal{CN}(0, 1)$ respectively account for

path loss coefficient and the Rayleigh fading channel gain between the BS to the mth user on SC_n, and where $d_{m,n}$ is the distance from the BS to the mth user on SC_n. The noise term $z_{m,n}$ is a zero-mean complex AWGN random variable with variance σ_z^2.

4.1.1 Imperfect Channel Model

Most previous works assumed that the BS has the whole knowledge of CSI. However, the perfect CSI is challenging to obtain in practice. In this paper, we investigate the energy efficiency optimization by assuming that the small scale fading channel is estimated at the BS. Since the path loss and shadowing are large scale fading factors and are slowly varying, we assume that the path loss and shadowing coefficients $D_{m,n}$ can be estimated perfectly by the BS [74, 75]. By using the minimum mean square error channel estimation error model [73, 76, 77], we can model the Rayleigh fading coefficient between the BS and the mth user on SC_n as

$$g_{m,n} = \hat{g}_{m,n} + e_{m,n} \tag{4.3}$$

where $g_{m,n}$ is the real Rayleigh fading channel coefficient between the BS to the mth user on SC_n; $\hat{g}_{m,n} \sim \mathcal{CN}(0, 1-\sigma_e^2)$ denotes the estimated channel gain and $e_{m,n}$ is the estimated error which follows a complex Gaussian distribution with mean zero and the variance σ_e^2. We assume that random variables $\hat{g}_{m,n}$ and $e_{m,n}$ are uncorrelated.

4.1.2 Scheduled Channel Capacity and Outage Capacity

If the BS has the perfect CSI $h_{m,n} = D_{m,n} g_{m,n}$, according to the Shannon's capacity formula, the maximum achievable data rate of the mth user on SC_n can be written as

$$C_{m,n} = B_{sc} \log_2 (1 + \Gamma_{m,n}) \tag{4.4}$$

where

$$\Gamma_{m,n} = \frac{p_{m,n}|h_{m,n}|^2}{|h_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n} + \sigma_z^2}. \tag{4.5}$$

In (5), $\Gamma_{m,n}$ is the signal-to-interference-plus-noise ratio for the mth user on SC_n; $|h_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n}$ is the interference from the users who have higher channel gains on SC_n. However, in practice, the BS only has the estimated fading channel coefficient $\hat{g}_{m,n}$. The data rate may exceed the maximum achievable data rate. Therefore, in this paper, we adopt outage probability as a metric to measure the performance of the case when the scheduled data rate exceeds the achievable data rate with imperfect CSI. The scheduled data rate of the mth user on SC_n can be written as

$$r_{m,n} = B_{sc}\log_2\left(1 + \Phi_{m,n}\right) \tag{4.6}$$

where

$$\Phi_{m,n} = \frac{p_{m,n}|\hat{h}_{m,n}|^2}{|\hat{h}_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n} + \sigma_z^2} \tag{4.7}$$

and where $\hat{h}_{m,n} = D_{m,n}\hat{g}_{m,n}$. Therefore, the system average outage sum rate is defined as [73]

$$R(\mathcal{U}, \mathcal{P}) = \sum_{n=1}^{N}\sum_{m=1}^{M_n} r_{m,n} \Pr\left[r_{m,n} \leq C_{m,n}|\hat{g}_{m,n}\right] \tag{4.8}$$

where $\Pr\left[r_{m,n} \leq C_{m,n}|\hat{g}_{m,n}\right]$ denotes the probability of the event when the scheduled data rate $r_{m,n}$, which is based on the estimated channel gain $\hat{g}_{m,n}$, is less than or equal to the maximum data rate $C_{m,n}$, which is based on the real channel gain $g_{m,n}$. The resource allocation includes user scheduling policy ($\mathcal{U}$) and power allocation policy ($\mathcal{P} = \{p_{m,n}, \forall m, n\}$ for $p_{m,n}$).

4.2 Problem Formulation

For energy efficient communication, we aim to maximize the total communication data rate with the unit power cost. Therefore, we formulate the system energy efficiency as a ratio of the system sum rate to the total power consumption. The energy efficiency of the system is formulated as

$$EE(\mathcal{U},\mathcal{P}) = \frac{R(\mathcal{U},\mathcal{P})}{P_s(\mathcal{U},\mathcal{P})} \tag{4.9}$$

where $P_s(\mathcal{U},\mathcal{P}) = P_c + P_T$ is the total power consumption of the system; P_c and $P_T = \sum_{n=1}^{N}\sum_{m=1}^{M_n} p_{m,n}$ are the circuit power consumption and transmission power, respectively. The energy efficiency maximization problem can be formulated as

$$\max_{\mathcal{U},\mathcal{P}} EE(\mathcal{U},\mathcal{P}) \tag{4.10}$$

$$\begin{aligned}
\text{s.t. } & C1: \Pr[C_{m,n} < r_{m,n}|\hat{g}_{m,n}] \leq \varepsilon_{out}, \forall m,n \\
& C2: \sum_{n=1}^{N}\sum_{m=1}^{M_n} p_{m,n} \leq P_{\max}, \\
& C3: r_{m,n} \geq R_{\min}, \forall m,n \\
& C4: p_{m,n} \geq 0, \forall m,n \\
& C5: |U_n| \leq U_{\max}, \forall n
\end{aligned} \tag{4.11}$$

where $C1$ specifies the channel outage probability requirement ε_{out}; $C2$ is the transmission power limit constraint for the BS in the downlink single cell network and $P_{\max}$ is the maximum transmitted power of the BS; $C3$ describes that each user data rate must be larger than the minimum user data rate $R_{\min}$, which is determined by the QoS requirement; $C4$ ensures that the power allocated to each user is positive, and $C5$ specifies the user number allocated on each subchannel must be less than the maximum number of the users $U_{\max}$, which can be greater than two.

4.3 Methodology and Problem Solution

The objective function in (4.10) is a non-convex function with non-convex probabilistic constraints. The global optimal resource allocation cannot easily be obtained in practice since this problem cannot be optimally solved in polynomial time. In order to solve this optimization problem efficiently, we first deal with the non-convex probabilistic constraint $C1$ in (4.11) and transform the probabilistic mixed problem into a non-probabilistic problem. An iterative resource allocation algorithm for user scheduling and power allocation is shown as Algorithm 4.

4.3.1 Optimization Problem Transformation

The outage probability requirement in $C1$ is a complicated non-convex function of scheduled data rate and powers. Therefore, the exact closed form expression for resource allocation cannot be obtained. However, we can address this issue by the following approximations. We first rewrite the scheduled data rate as

$$\begin{aligned} r_{m,n} &= B_{sc}\log_2\left(1+\Phi_{m,n}\right) \\ &= B_{sc}\log_2\left(1+\frac{a_{m,n}}{b_{m,n}}\right) \end{aligned} \tag{4.12}$$

where $\Phi_{m,n} = \frac{a_{m,n}}{b_{m,n}} = 2^{\frac{r_{m,n}}{B_{sc}}} - 1$ and $a_{m,n} = b_{m,n}(2^{\frac{r_{m,n}}{B_{sc}}} - 1)$, and the achievable data rate is

$$\begin{aligned} C_{m,n} &= B_{sc}\log_2\left(1+\Gamma_{m,n}\right) \\ &= B_{sc}\log_2\left(1+\frac{c^U_{m,n}}{c^D_{m,n}}\right) \end{aligned} \tag{4.13}$$

where $c^U_{m,n} = p_{m,n} D^2_{m,n} |g_{m,n}|^2, c^D_{m,n} = D^2_{m,n} |g_{m,n}|^2 \sum_{i=m+1}^{M_n} p_{i,n} + \sigma_z^2$. The outage probability condition can be bounded by

$$\begin{aligned} &\Pr[C_{m,n} < r_{m,n} | \hat{g}_{m,n}] \leq \varepsilon_{out} \\ =&\Pr[\Gamma_{m,n} < \Phi_{m,n} | \hat{g}_{m,n}] \leq \varepsilon_{out}. \end{aligned} \tag{4.14}$$

According to the total probability theorem, the outage probability can be written as

$$\begin{aligned} &\Pr[\Gamma_{m,n} < \Phi_{m,n} | \hat{g}_{m,n}] \\ =&\Pr\left[\frac{c^U_{m,n}}{c^D_{m,n}} < 2^{\frac{r_{m,n}}{B_{sc}}} - 1 | \hat{g}_{m,n}\right] \\ =&\Pr[E1] \cdot \Pr\left[c^U_{m,n} \leq a_{m,n} | \hat{g}_{m,n}\right] \\ &+ \Pr[E2] \cdot \Pr\left[c^U_{m,n} > a_{m,n} | \hat{g}_{m,n}\right] \end{aligned} \tag{4.15}$$

where

$$\begin{aligned} \Pr[E1] =& \Pr\left[\frac{c^U_{m,n}}{c^D_{m,n}} < 2^{\frac{r_{m,n}}{B_{sc}}} - 1 | c^U_{m,n} \leq a_{m,n}, \hat{g}_{m,n}\right] \\ \Pr[E2] =& \Pr\left[\frac{c^U_{m,n}}{c^D_{m,n}} < 2^{\frac{r_{m,n}}{B_{sc}}} - 1 | c^U_{m,n} > a_{m,n}, \hat{g}_{m,n}\right]. \end{aligned} \tag{4.16}$$

Following [73, 76, 77], we can show that the outage probability constraint $C1$ can be derived from

$$\Pr\left[c^D_{m,n} \geq b_{m,n} | \hat{g}_{m,n}\right] \leq \varepsilon_{out}/2 \tag{4.17}$$

and

$$\Pr\left[c^U_{m,n} \leq a_{m,n} | \hat{g}_{m,n}\right] = \varepsilon_{out}/2. \tag{4.18}$$

To show this, we note that since $b_{m,n} = a_{m,n}/(2^{\frac{r_{m,n}}{B_{sc}}} - 1)$, we have

$$\begin{aligned} &\Pr\left[c^D_{m,n} \geq b_{m,n}|\hat{g}_{m,n}\right] \\ =&\Pr\left[c^D_{m,n} \geq a_{m,n}/(2^{\frac{r_{m,n}}{B_{sc}}} - 1)|\hat{g}_{m,n}\right] \\ =&\Pr\left[\frac{a_{m,n}}{c^D_{m,n}} \leq 2^{\frac{r_{m,n}}{B_{sc}}} - 1|\hat{g}_{m,n}\right]. \end{aligned} \tag{4.19}$$

Therefore, eq. (4.17) can be rewritten as $\Pr\left[\frac{a_{m,n}}{c^D_{m,n}} \leq 2^{\frac{r_{m,n}}{B_{sc}}} - 1|\hat{g}_{m,n}\right] \leq \frac{\varepsilon_{out}}{2}$. When $c^U_{m,n} > a_{m,n}$, we can always have

$$\Pr[E_2] = \Pr\left[\frac{c^U_{m,n}}{c^D_{m,n}} \leq 2^{\frac{r_{m,n}}{B_{sc}}} - 1|\hat{g}_{m,n}\right] \leq \varepsilon_{out}/2. \tag{4.20}$$

From (4.18), we have $\Pr\left[c^U_{m,n} > a_{m,n}|\hat{g}_{m,n}\right] = 1 - \varepsilon_{out}/2$. Since $\Pr[E_1] \leq 1$, we have from (4.15), (4.17) and (4.18)

$$\begin{aligned} \Pr[\Gamma_{m,n} < \Phi_{m,n}|\hat{g}_{m,n}] &\leq \varepsilon_{out}/2 + (\varepsilon_{out}/2)(1 - \varepsilon_{out}/2) \\ &= \varepsilon_{out} - \varepsilon^2_{out}/4 \approx \varepsilon_{out}, \quad \text{for} \quad \varepsilon_{\text{out}} \ll 1. \end{aligned} \tag{4.21}$$

Therefore, the constraint $C1$ can be rewritten as (4.17) and (4.18).

We now integrate the probabilistic constraints in (4.17) and (4.18) into (4.10) and transform to a revised optimization problem as follows. By using the Markov inequality, we have [73, 77, 78]

$$\begin{aligned} &\Pr\left[c^D_{m,n} \geq b_{m,n}|\hat{g}_{m,n}\right] \\ =&\Pr\left[D^2_{m,n}|g_{m,n}|^2 \sum_{i=m+1}^{M_n} p_{i,n} \geq b_{m,n} - \sigma^2_z|\hat{g}_{m,n}\right] \\ \leq&\frac{E\left[D^2_{m,n}|g_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n}\right]}{b_{m,n} - \sigma^2_z} \\ =&\frac{D^2_{m,n}|g_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n}}{b_{m,n} - \sigma^2_z}. \end{aligned} \tag{4.22}$$

According to (4.17), let the right side of (4.22) equals to $\varepsilon_{out}/2$, we have

$$\frac{D_{m,n}^2|g_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n}}{b_{m,n}-\sigma_z^2} = \frac{\varepsilon_{out}}{2}. \tag{4.23}$$

Since $|g_{m,n}|^2 \sim \mathcal{CN}(\hat{g}_{m,n}, \sigma_e^2)$ is a non-central chi-squared distributed random variable with two degrees of freedom, the left side of (4.18) can be rewritten as

$$\begin{aligned}
&\Pr\left[c_{m,n}^U \le a_{m,n}|\hat{g}_{m,n}\right] \\
=&\Pr\left[p_{m,n}D_{m,n}^2|g_{m,n}|^2 \le a_{m,n}|\hat{g}_{m,n}\right] \\
=&\Pr\left[|g_{m,n}|^2 \le \frac{a_{m,n}}{D_{m,n}^2 p_{m,n}}|\hat{g}_{m,n}\right] \\
=&F_{|g_{m,n}|^2}\left(\frac{a_{m,n}}{D_{m,n}^2 p_{m,n}}\right) \\
=&1-Q_1\left(\sqrt{\frac{2|\hat{g}_{m,n}|^2}{\sigma_e^2}}, \sqrt{\frac{2}{\sigma_e^2}\frac{a_{m,n}}{D_{m,n}^2 p_{m,n}}}\right)
\end{aligned} \tag{4.24}$$

where

$$Q_1(a,b) = \exp\left(-\frac{a^2+b^2}{2}\right)\sum_{k=0}^{\infty}\left(\frac{a}{b}\right)^k I_k(ab)$$

is the first-order Marcum Q-function and $I_k(\cdot)$ is the kth order modified Bessel function of the first kind. Based on (4.19), let (4.24) equal to $\varepsilon_{out}/2$, then we have

$$a_{m,n} = F_{|g_{m,n}|^2}^{-1}(\varepsilon_{out}/2)\cdot D_{m,n}^2 p_{m,n} \tag{4.25}$$

where $F_{|g_{m,n}|^2}^{-1}(\cdot)$ is the inverse function of $F_{|g_{m,n}|^2}(\cdot)$. Based on $|g_{m,n}|^2 = |\hat{g}_{m,n}|^2 + \sigma_e^2$, $b_{m,n} = a_{m,n}/(2^{\frac{r_{m,n}}{B_{sc}}}-1)$, (4.23) and (4.24), we have

$$\begin{aligned}
\frac{D_{m,n}^2|g_{m,n}|^2 \sum\limits_{i=m+1}^{M_n} p_{i,n}}{a_{m,n}/(2^{\frac{r_{m,n}}{B_{sc}}}-1)-\sigma_z^2} &= \frac{\sum\limits_{i=m+1}^{M_n} p_{i,n}D_{m,n}^2\left(|\hat{g}_{m,n}|^2+\sigma_e^2\right)}{\frac{F_{|g_{m,n}|^2}^{-1}(\varepsilon_{out}/2)\cdot D_{m,n}^2 p_{m,n}}{2^{\frac{r_{m,n}}{B_{sc}}}-1}-\sigma_z^2} \\
&= \frac{\varepsilon_{out}}{2}.
\end{aligned} \tag{4.26}$$

Therefore, the data rate for the mth user on SC_n ($r_{m,n}$) with the outage probability constraint can be written by

$$\tilde{R}_{m,n} = B_{sc}\log_2\left(1+\tilde{\Phi}_{m,n}\right) \tag{4.27}$$

where

$$\tilde{\Phi}_{m,n} = \frac{\varepsilon_{out}F^{-1}_{|g_{m,n}|^2}\left(\varepsilon_{out}/2\right)\cdot D^2_{m,n}p_{m,n}}{\varepsilon_{out}\sigma_z^2 + 2D^2_{m,n}\left(|\hat{g}_{m,n}|^2+\sigma_e^2\right)\sum\limits_{i=m+1}^{M_n}p_{i,n}}. \tag{4.28}$$

Now, the transformed average sum rate for the entire system can be written by [59]

$$\tilde{R}(\mathcal{U},\mathcal{P}) = \sum_{n=1}^{N}\sum_{m=1}^{M_n}\left(1-\varepsilon_{out}\right)\tilde{R}_{m,n}. \tag{4.29}$$

The energy efficient optimization problem can be reformulated as

$$\max_{\mathcal{U},\mathcal{P}} \widetilde{EE}(\mathcal{U},\mathcal{P}) = \frac{\tilde{R}(\mathcal{U},\mathcal{P})}{P_s(\mathcal{U},\mathcal{P})} \tag{4.30}$$

$$\begin{aligned} \text{s.t. } C1: &\sum_{n=1}^{N}\sum_{m=1}^{M_n}p_{m,n} \le P_{\max}, \\ C2: &\tilde{R}_{m,n} \ge R_{\min}, \forall m,n \\ C3: &p_{m,n} \ge 0, \forall m,n \\ C4: &|U_n| \le U_{\max}, \forall n. \end{aligned} \tag{4.31}$$

The transformed non-probabilistic optimization problem (4.30) is still non-convex. The optimal solution to this non-convex optimization problem (4.30) is challenging to obtain in practice. In order to solve (4.30) efficiently, we will design an iterative algorithm to maximize the system energy efficiency.

4.4 Energy Efficient Resource Allocation Scheme

In this section, we design an iterative algorithm for the energy efficient joint user scheduling and power allocation, as shown in Algorithm 4. This algorithm includes user scheduling

subproblem and power allocation subproblem. In each iteration, user scheduling and power allocation are updated iteratively until convergence. We first assume equal power allocation for all the users. The user scheduling is optimized by the proposed Algorithm 5 which provides better performance than the corresponding algorithm in [57]. Based on the obtained user scheduling scheme, the power allocation scheme can be updated by Algorithm 6, where the optimal power allocation policy can be obtained by the Lagrangian approach. An iterative power allocation scheme is proposed and the optimal closed form power expression is derived for each user. Due to the non-convexity of the problem in (4.30), the optimal solution cannot be obtained in polynomial time. However, in this iterative algorithm, the system energy efficiency slightly improves at each iteration and it converges at the end of the procedure. It is observed that the constraints in (4.31) satisfy the necessary Karush-Kuhn-Tucker (KKT) conditions. Algorithm 4 finds at least one locally optimal solution with the potential of being a global optimum.

Algorithm 4: Energy Efficient Resource Allocation Algorithm

1: Initialize the power allocation for each user $p_{m,n} = P_{\max}/M$.
2: Initialize the maximum iterations $L_{\max}$, the index $l = 1$ and maximum tolerance ε.
3: **while** $\left|EE^{(l)} - EE^{(l-1)}\right| > \varepsilon$ or $l \leq L_{\max}$ **do**
4: 1. Given $P^{(l)}$, obtain user scheduling scheme U^l by the proposed suboptimal algorithm (Algorithm 5).
5: 2. Update the power allocation scheme $P^{(l+1)}$ according to the propose power allocation policy shown in Algorithm 6 of Section 4.4.2.
6: 3. Set $l = l + 1$ and compute $EE^{(l+1)}$.
7: **end while**

4.4.1 User Scheduling Scheme Design

In this section, we design a user scheduling scheme to assign users to different subchannels in order to maximize the system energy efficiency. Since the maximum number of users that can be multiplexed on the same subchannel is less than $U_{\max}$, the global optimal solution can only be obtained by the exhaustive search method, which has exponential complexity with respect to the number of subchannels. Thus we propose a novel suboptimal

user scheduling scheme to reduce the complexity. The user scheduling can be expressed as

$$\mathcal{U}_{optimal} = \arg\max_{\mathcal{U}} EE(\mathcal{U}) = \frac{\tilde{R}(\mathcal{U})}{P_s(\mathcal{U})}. \tag{4.32}$$

Algorithm 5: A Novel Suboptimal User Scheduling Algorithm

Initialize the power allocation for each user $P_{m,n}$.
Construct the estimate channel gain $\mathbb{H} \triangleq [|\hat{h}_{m,n}|]_{M\times N}$.
Initialize the sets $\mathbb{U}_{un}$ to record the unallocated user in the system.
Initialize the lists for all the subchannels $EE(n)$ to record the energy efficiency of SC_n.
while $\mathbb{U}_{unun} \neq \emptyset$ **do**
  Find the maximum value $|\hat{h}_{m,n}|$ in $\mathbb{H}$ using

$$|\hat{h}_{m,n}| = \underset{m\in\mathbb{U}_{un}, n\in\mathbb{H}_{un}}{\arg\max} (\mathbb{H}).$$

  $EE_{n,possible}$=$\emptyset$.
  $EE_{n,i}$=$\emptyset$.
  if the number of multiplexed users on this subchannel is less than $U_{\max}$ **then**
    a) Schedule the user m onto the subchannel n.
    b) $\mathbb{U}_{un}$=$\mathbb{U}_{un}\backslash UE_m$.
  end if
  if the number of multiplexed users on this subchannel equals $U_{\max}$ **then**
    a) Assume UE_m is allocated on SC_n and the user set is now $U_{n,possible}$.
    b) Calculate the energy efficiency $EE_{n\_possible}$ of the $U_{\max}$ users from $U_{n,possible}$ on SC_n.
    c) $U_n = \underset{U\in U_{n,possible}}{\arg\max} (EE_{n\_possible})$ and $UE_i \notin U_n$.
    d) $\mathbb{U}_{un}$=$\mathbb{U}_{un}\backslash U_n$.
    Let the ith and nth row's elements in $\mathbb{H}$ be zeros.
    Let the nth column's elements in $\mathbb{H}$ be zeros.
  end if
end while

Algorithm 5 describes the proposed suboptimal user scheduling process. We first initialize the power allocation scheme. When Algorithm 5 is used for the first time, we assume equal power allocation for each user $p_{m,n} = \frac{P_{\max}}{M}$. $\mathbb{U}_{un}$ is initialized to record the users who have not been allocated to any subchannel. In the scheduling procedure, we need to find the user who has the maximum channel gain and allocate it to the corresponding

subchannel if the number of users multiplexed on this subchannel is less than $U_{\max}$. If the user number is equal to $U_{\max}$, the users should be selected from the user set $\mathbb{U}_{n,possible}$. The users who can provide the maximum energy efficiency on this subchannel will be allocated on this subchannel. The user who has been selected will be returned to $\mathbb{U}_{un}$. This process terminates if there is no user left to be allocated.

The complexity of the proposed algorithm is less than the optimal exhaustive search. Since the user scheduling subproblem is non-convex, the optimal user scheduling algorithm can only be obtained through exhaustive search. For a given power allocation scheme, the best user scheduling scheme can be obtained to maximize the system energy efficiency by searching all the combinations of users and subchannels. The exhaustive search depends on the maximum number of users multiplexed on each subchannel $U_{\max}$. For the optimal solution, the BS needs to search $N!\left(\binom{M}{1}+\binom{M}{2}+\cdots+\binom{M}{U_{\max}}\right)$ combinations. The time complexity of exhaustive search is $O(N!2^M)$. The performance of the proposed user scheduling algorithm is compared with the optimal solution and the existing scheme in [57]. Simulation comparison is shown in Section 4.5. With the same complexity, the proposed algorithm can achieve improved energy efficiency performance.

4.4.2 Energy Efficient Power Allocation Algorithm

After user scheduling, the power allocation for all the users is still equal power allocation scheme. In order to further improve the system energy efficiency, a closed form optimal power allocation expression is derived. Given the user scheduling scheme, the optimization problem in (4.30) is still a non-convex optimization problem with respect to $p_{m,n}$. The objective function in (4.30) has a non-linear fractional form which makes the problem challenging to solve. In order to reduce the complexity in solving the power allocation problem, successive convex approximation and parameter transformation are exploited to solve this problem. We design an iterative algorithm for power allocation in Algorithm 6.

The lower bound of $\log_2(1+q)$ can be expressed as $\alpha\log_2 q+\beta \leq \log_2(1+q)$ for any $q \geq 0$ [79]. The bound will become tight when $q=\bar{q}$, $\alpha=\frac{\bar{q}}{1+\bar{q}}$ and $\beta=\log_2(1+\bar{q})-\frac{\bar{q}}{1+\bar{q}}\log_2\bar{q}$. By

utilizing the lower bound, the lower bound of data rate for the mth user on the subchannel n can be written as

$$R^*_{m,n} = B_{sc}\alpha_{m,n}\log_2\left(\tilde{\Phi}_{m,n}\right) + \beta_{m,n}. \tag{4.33}$$

Therefore, the energy efficient optimization problem can be rewritten by

$$\max_{\mathbf{P}\succ 0} EE(\mathbf{P}) = \frac{R^*(\mathbf{P})}{P_s(\mathbf{P})} \tag{4.34}$$

$$\begin{aligned} \text{s.t. } C1 &: \sum_{n=1}^{N}\sum_{m=1}^{M_n} p_{m,n} \le P_{\max}, \\ C2 &: R^*_{m,n} \ge R_{\min}, \forall m, n \end{aligned} \tag{4.35}$$

where $R^* = \sum_{n=1}^{N}\sum_{m=1}^{M_n}(1-\varepsilon_{out})R^*_{m,n}$, $\mathbf{P} \triangleq [p_{m,n}]_{M_n\times N}$ and $\mathbf{P}\succ 0$ means all elements of $\mathbf{P}$ are positive. The objective function in (4.34) is non-convex. In order to avoid high complexity of the solution to this problem, we introduce the following parameter transformation. Without loss of generality, we define the maximum energy efficiency of the system as

$$t^* = \max_{\mathbf{P}\succ 0}\frac{R^*(\mathbf{P})}{P_s(\mathbf{P})} = \frac{R^*(\mathbf{P}^*)}{P^*_s(\mathbf{P}^*)}. \tag{4.36}$$

Note that the problem in (4.34) is a concave-convex fractional optimization problem which is a nonlinear fractional program, and it can be transformed to an equivalent parameterized non-fractional form as follows [66]

$$\max_{\mathbf{P}\succ 0} R^*(\mathbf{P}) - tP_s(\mathbf{P}) \tag{4.37}$$

$$\begin{aligned} \text{s.t. } C1 &: \sum_{n=1}^{N}\sum_{m=1}^{M_n} p_{m,n} \le P_{\max}, \\ C2 &: R_{m,n} \ge R_{\min}, \forall m, n \end{aligned} \tag{4.38}$$

where t is a parameter introduced to scale the weight of P_s. For a given value t, the solution to problem (4.37) is denoted by $\mathbf{P}$ and the optimal solution to (4.37) is defined as $\mathbf{P}^*$. Let

Algorithm 6: An Iterative Power Allocation Algorithm

Initialize the maximum number of the iterations $L_{\max}$ and the maximum tolerance ε.
Initialize the energy efficiency t and the iteration index $l = 0$.
while $\left|R^*\left(\mathbf{P}^l\right) - t^* P_s\left(\mathbf{P}^l\right)\right| > \varepsilon$ or $l \leq L_{\max}$ **do**
1. Given the energy efficiency t, solve the power allocation using (4.45).
2. Set $l = l + 1$ and let $t^l = \frac{R^*\left(\mathbf{P}^{l-1}\right)}{P_s^*\left(\mathbf{P}^{l-1}\right)}$.

end while

us define the function

$$f(t) \triangleq \max_{\mathbf{P} \succ 0} \left\{R^*\left(\mathbf{P}\right) - tP_s\left(\mathbf{P}\right)\right\}. \tag{4.39}$$

From (4.37), $f(t)$ is negative when t approaches infinity while $f(t)$ is positive when t approaches minus infinity. Obviously, $f(t)$ is convex with respect to t. Hence, a proposed algorithm can be exploited to determine the maximum energy efficiency. Therefore, solving (4.32) is equivalent to finding the maximum energy efficiency t^*, which can be achieved if and only if

$$\begin{aligned} f\left(t^*\right) &= \max_{\mathbf{P}} \left\{R^*\left(\mathbf{P}\right) - t^* P_s\left(\mathbf{P}\right)\right\} \\ &= R^*\left(\mathbf{P}^*\right) - t^* P_s\left(\mathbf{P}^*\right) \\ &= 0 \end{aligned} \tag{4.40}$$

where $\mathbf{P}^*$ is the optimal power allocation policy. Therefore, an iterative resource allocation algorithm for power allocation can be proposed in Algorithm 6. Given the maximum iteration number and maximum tolerance, the energy efficiency improves for each iteration until the algorithm converges. In each iteration, the Lagrange multiplier approach is used to solve the problem in (4.37). Appendix A proves the convergence of Algorithm 6. Given the energy efficiency t, the optimal power allocation for each user will be developed in the following subsection.

4.4.3 Power Allocation Expression Derivation

In Algorithm 6, given the energy efficiency t, the optimization in (4.37) is a concave maximization problem with respect to power allocation policy $\mathbf{P}$. This problem can be solved by its dual problem and the difference between the primal and dual solution is zero when strong duality holds [66, 80]. In this section, we solve the primal problem of (4.37) by solving its associated dual problem. The Lagrangian function can be written by

$$\begin{aligned} L(\mathbf{P}, \lambda, \boldsymbol{\upsilon}) = & \sum_{n=1}^{N} \sum_{m=1}^{M_n} (1 - \varepsilon_{out}) R_{m,n}^{*}(\mathbf{P}) - t P_s(\mathbf{P}) \\ & + \lambda \left(P_{\max} - \sum_{n=1}^{N} \sum_{m=1}^{M_n} p_{m,n} \right) \\ & + \sum_{n=1}^{N} \sum_{m=1}^{M_n} \upsilon_{m,n} \left(R_{m,n}^{*}(\mathbf{P}) - R_{\min} \right) \end{aligned} \tag{4.41}$$

where λ and $\boldsymbol{\upsilon} = [\upsilon_{m,n}]_{M_n \times N}$ are the Lagrange multipliers corresponding to the constraints $C1$ and $C2$ in (4.36). The constraints are KKT conditions [66] for optimizing power allocation, thus the dual problem of (4.37) is

$$\min_{\lambda, \boldsymbol{\upsilon}} \quad g(\lambda, \boldsymbol{\upsilon}) \tag{4.42}$$

$$\text{s.t.} \quad \lambda \geq 0, \boldsymbol{\upsilon} \succeq 0 \tag{4.43}$$

where

$$g(\lambda, \boldsymbol{\upsilon}) = \max_{\mathbf{P} \succ 0} \quad L(\mathbf{P}, \lambda, \boldsymbol{\upsilon}). \tag{4.44}$$

To solve (4.37), we decompose it into two layers: inner layer and outer layer. We first solve the inner layer problem to obtain the power allocation policy, and then use the outer layer to compute the dual variables λ and $\boldsymbol{\upsilon_{m,n}}$ iteratively.

For a fixed Lagrange multiplier and a given energy efficiency t, the problem is a standard optimization problem with the KKT conditions. The optimal power allocation policy for

the mth user on SC_n can be derived in Appendix B as

$$p_{m,n} = \frac{\alpha_{m,n}\frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{m,n}\right)}{\ln 2\left(\lambda+t\right)+\sum\limits_{l=1}^{m-1} A(p_{l,n})} \tag{4.45}$$

where

$$A(p_{l,n}) = \left(1-\varepsilon_{out}+\upsilon_{l,n}\right)\alpha_{l,n}\frac{2\left(|\hat{g}_{l,n}|^2+\sigma_e^2\right)D_{l,n}^2\tilde{\Phi}_{l,n}}{p_{l,n}\varepsilon_{out}F^{-1}_{|g_{l,n}|^2}\left(\varepsilon_{out}/2\right)}. \tag{4.46}$$

Given the power allocation scheme in (4.43), the outer layer primal problem can be solved by the gradient method since the objective function is differentiable. Therefore, the dual variables can be updated with gradient descent as

$$\lambda\left(l^{'}+1\right) = \left[\lambda\left(l^{'}\right)-\xi_1\left(l^{'}\right)\left(P_{\max}-\sum_{n=1}^{N}\sum_{m=1}^{M_n}p_{m,n}\right)\right]^{+} \tag{4.47}$$

$$\upsilon_{m,n}\left(l^{'}+1\right) = \left[\upsilon_{m,n}\left(l^{'}\right)-\xi_2\left(l^{'}\right)\left(R^{*}_{m,n}-R_{\min}\right)\right]^{+}, \quad \forall m,n \tag{4.48}$$

where $l^{'}$ is the iteration index. $\xi_1\left(l^{'}\right)$ and $\xi_2\left(l^{'}\right)$ are positive step sizes at iteration $l^{'}$. Since the transformed problem (4.37) is concave, this guarantees that the iteration converges to an optimal solution to problem (4.37) based on appropriate step sizes.

4.5 Simulation Results

In this section, we evaluate the performance of our proposed resource allocation algorithms for NOMA system through Monte Carlo simulations. The system parameters used in our simulations are given as follows. We consider one BS located in the cell center and M users are uniformly distributed on the circular range with radius of 500 m. In the simulations, we set the minimum distance among users as 40 m and the minimum distance between base station and users as 50 m. The total bandwidth in this system is 5 MHz that is divided into N subchannels. In the NOMA and OFDMA systems, we assume the

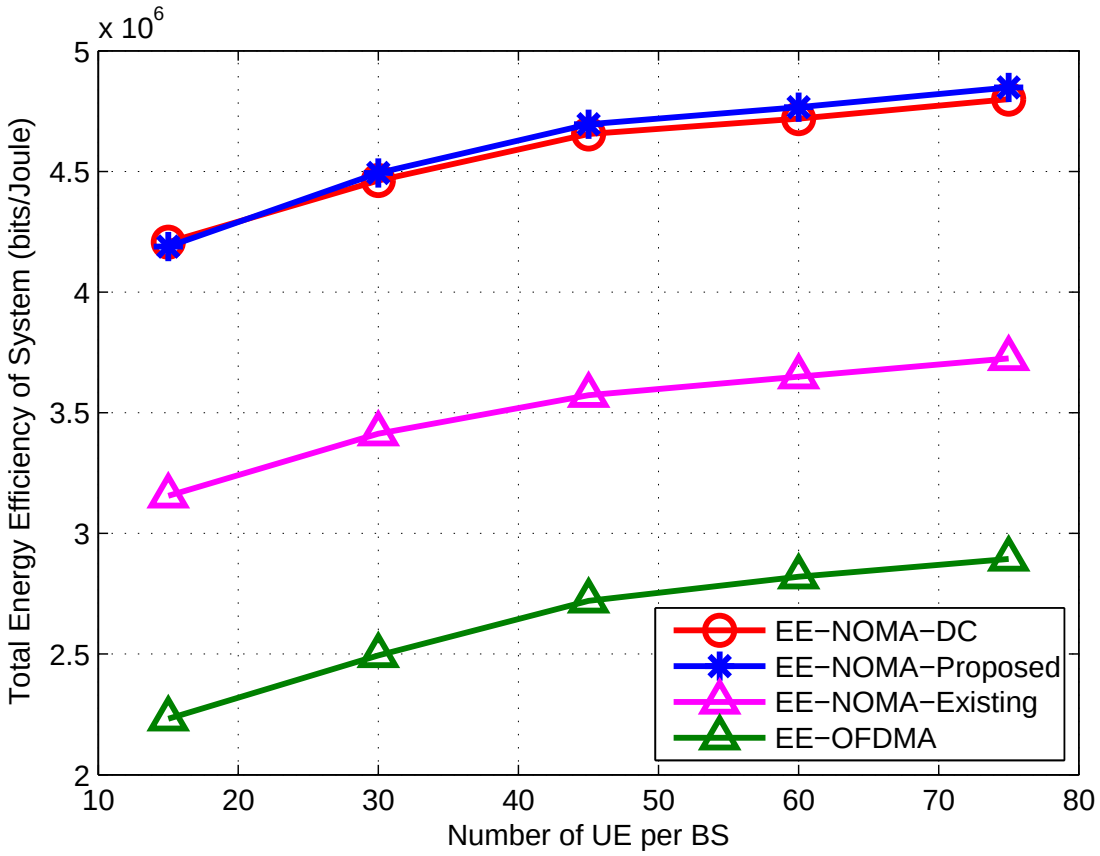

Figure 4.1: Energy efficiency performance versus the number of users

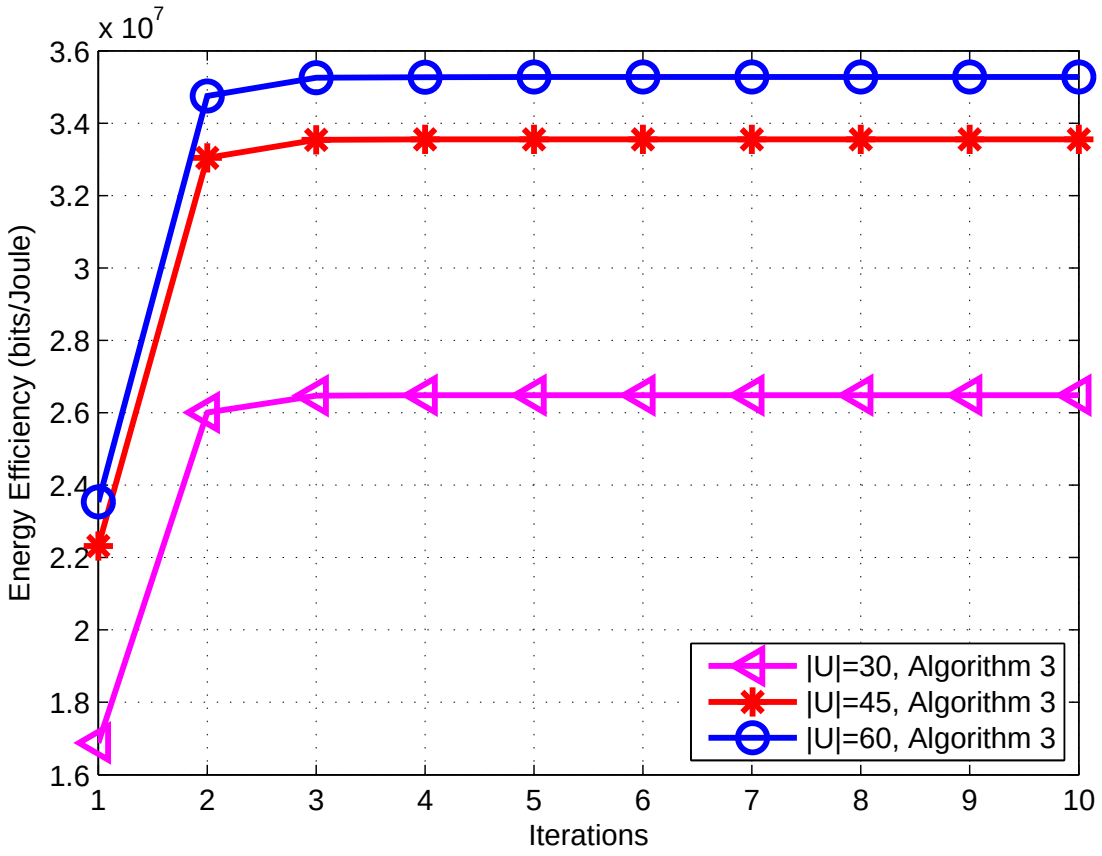

Figure 4.2: Energy efficiency performance versus the number of iterations for Algorithm 6.

small-scale Rayleigh fading channels between the base station and users, and the 3GPP urban path loss model with a path loss factor of 3.76 [81] [7]. In OFDMA systems, each user can only be assigned to one subchannel. In NOMA systems, the maximum users can be multiplexed on the same subchannel is three. In the simulations, we compare our proposed resource allocation algorithms for NOMA systems with a conventional OFDMA system. We set circuit power consumption $P_c = 30$ dBm and the BS peak power $P_{\max}$ is from 10 dBm and 40 dBm. The maximum number of users is 75 and $\sigma_z^2 = \frac{B}{N} N_0$, where $N_0 = -174$ dBm/Hz is the AWGN power spectral density. The maximum error tolerance for the algorithms is set to $\epsilon = 0.01$.

Figure 4.1 compares the proposed user scheduling scheme with the suboptimal subchannel allocation in [57]. In this figure, we set $P_{\max} = 41$ dBm. The outage probability is 0.05 and the variance of estimated error for channel gain is 0.1. Based on the equal power allocation for users, the new proposed user scheme can achieve higher energy efficiency than the existing schemes in [44] and [57]. For example, when the number of users is 60, the energy efficiency of our proposed resource allocation scheme for the NOMA system is 2% more than the subchannel algorithm in [57] and 30% more than that of the existing scheme in [44] with equal power allocation.

Figure 4.2 evaluates the energy efficiency performance versus the number of iterations of Algorithm 6. We set the minimum normalized data rate normalized by bandwidth for each user as 7 bits/s/Hz and the maximum transmit power as $P_{\max} = 40$ dBm. The outage probability is 0.1 and the variance of estimated error for channel gain is 0.1. Based on the user scheduling scheme shown in Algorithm 4, the convergence of the proposed iterative power allocation algorithm (Algorithm 6) is presented with different number of users. As observed from Fig. 4.2, the system energy efficiency converges after 4 iterations. The system with 60 users can achieve higher energy efficiency than the system with 45 and 30 users.

[7] Let D be the distance from the base station to the different users. The path loss model from the base station to its users is $PL\_dB = 15.3 + 37.6\log_{10} D$.

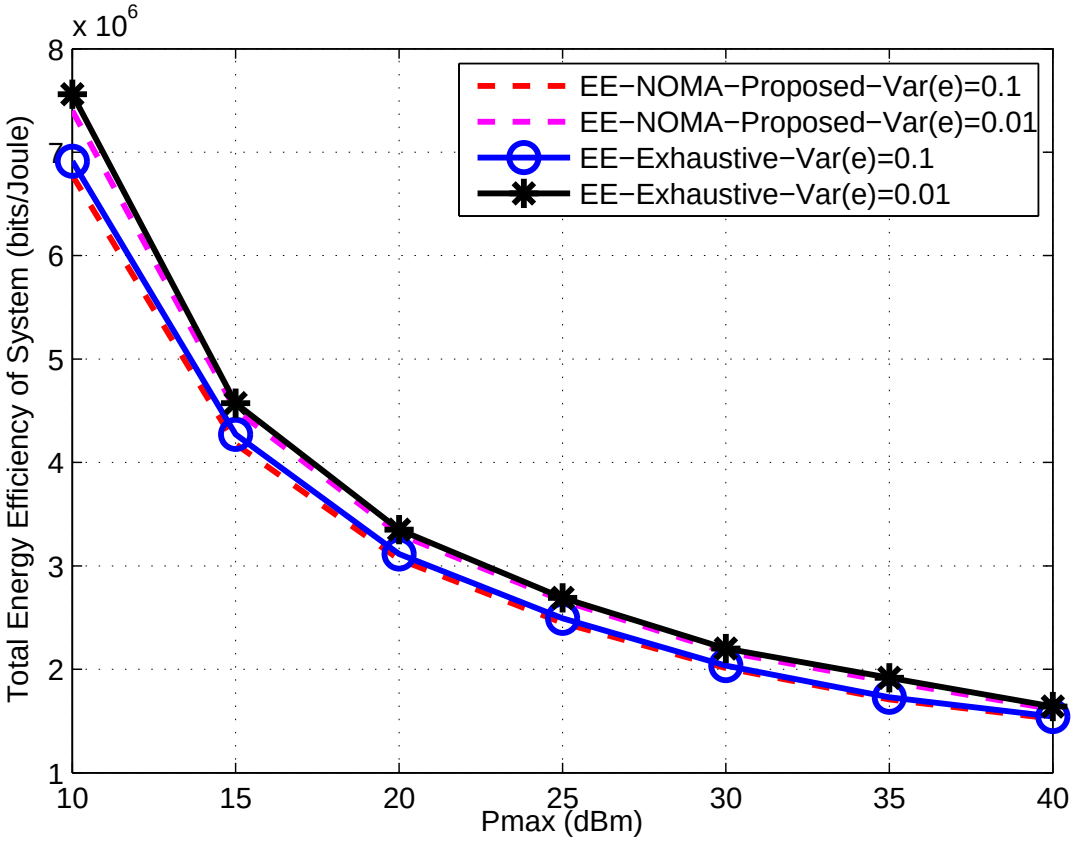

Figure 4.3: Energy efficiency performance versus the number of users.

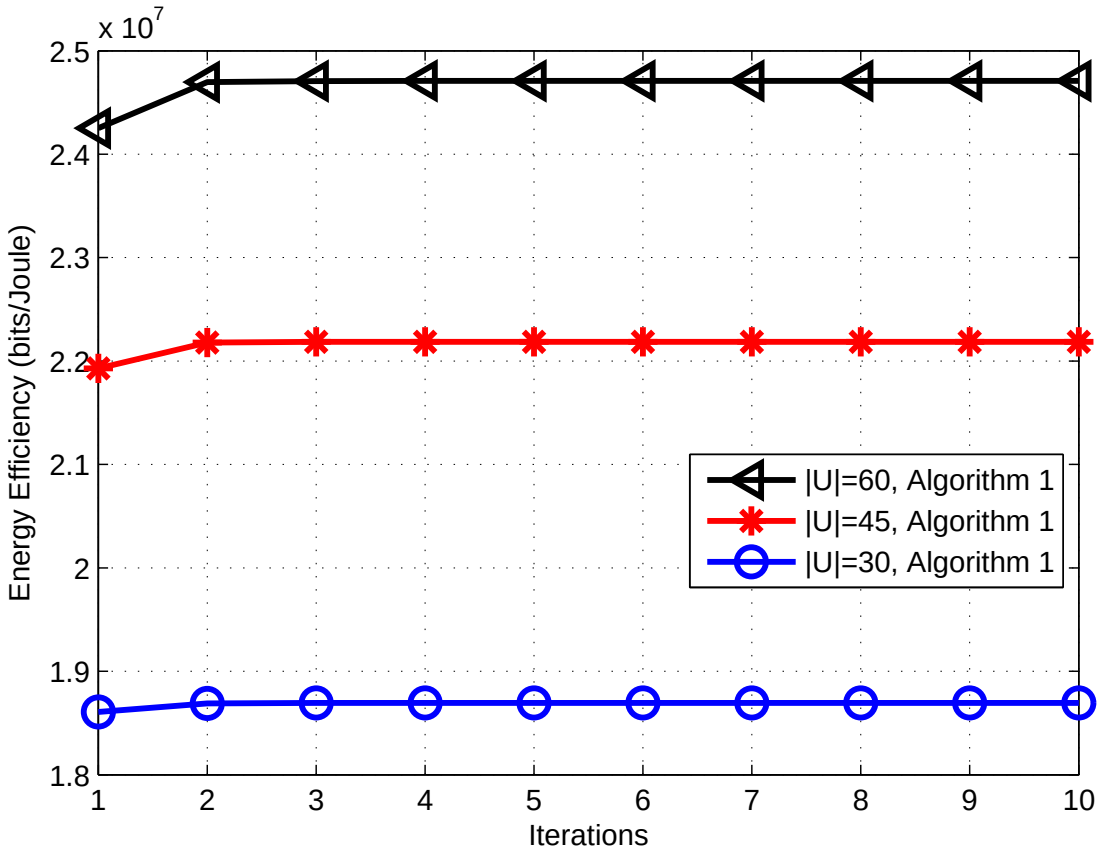

Figure 4.4: Energy efficiency performance versus the number of iterations for Algorithm 4.

Figure 4.3 shows the proposed user scheduling comparison with the optimal user scheduling through exhaustive search. To evaluate the performance of the proposed user scheduling algorithm, we adopt the equal power allocation scheme for each user and the total transmit power for this system is set from 10 dBm to 40 dBm. Since the number of users is fixed, the energy efficiency decreases when the total transmit power consumption increases. This is because the transmit power consumption grows faster than the system sum rate. As shown in this figure, the energy efficiency of the NOMA system with our proposed user scheduling scheme is close to the exhaustive search solution especially when the transmit power is large. The complexity of the proposed user scheduling scheme is lower than the optimal solution especially when the user number is large. When $P_{\max} = 25$ dBm, the optimal algorithm can only achieve 1.1% more than the proposed user scheduling scheme.

Figure 4.4 shows the convergence of the joint user scheduling and power allocation algorithm for different number of users with $P_{\max} = 40$ dBm. It is observed that the overall iterative algorithm converges after 5 iterations. This result demonstrates that the energy efficiency increases when the user number increases.

In Fig. 4.5, the performance of energy efficiency is evaluated versus the number of users. In our simulation, we set the channel estimation error variance as $\sigma_e^2 = 0.05$ and the outage probability ε_{out} as 0.1. It is observed that the system energy efficiency increases when the number of users increases. As the number of users grows, the energy efficiency continues to increase, but the rate of growth becomes slower. The performance of our proposed resource allocation scheme for the NOMA system achieves higher energy efficiency than that of OFDMA scheme as well as NOMA system with equal power allocation scheme. For example, when the number of users is 30, the energy efficiency of the proposed resource allocation scheme for the NOMA system is 38% more than that of the OFDMA scheme with equal power allocation.

Figure 4.6 demonstrates the energy efficiency performance of NOMA systems with different estimation error variances versus the number of users M. As observed in Fig. 6, the energy efficiency of the system deteriorates when the error variances increases. When

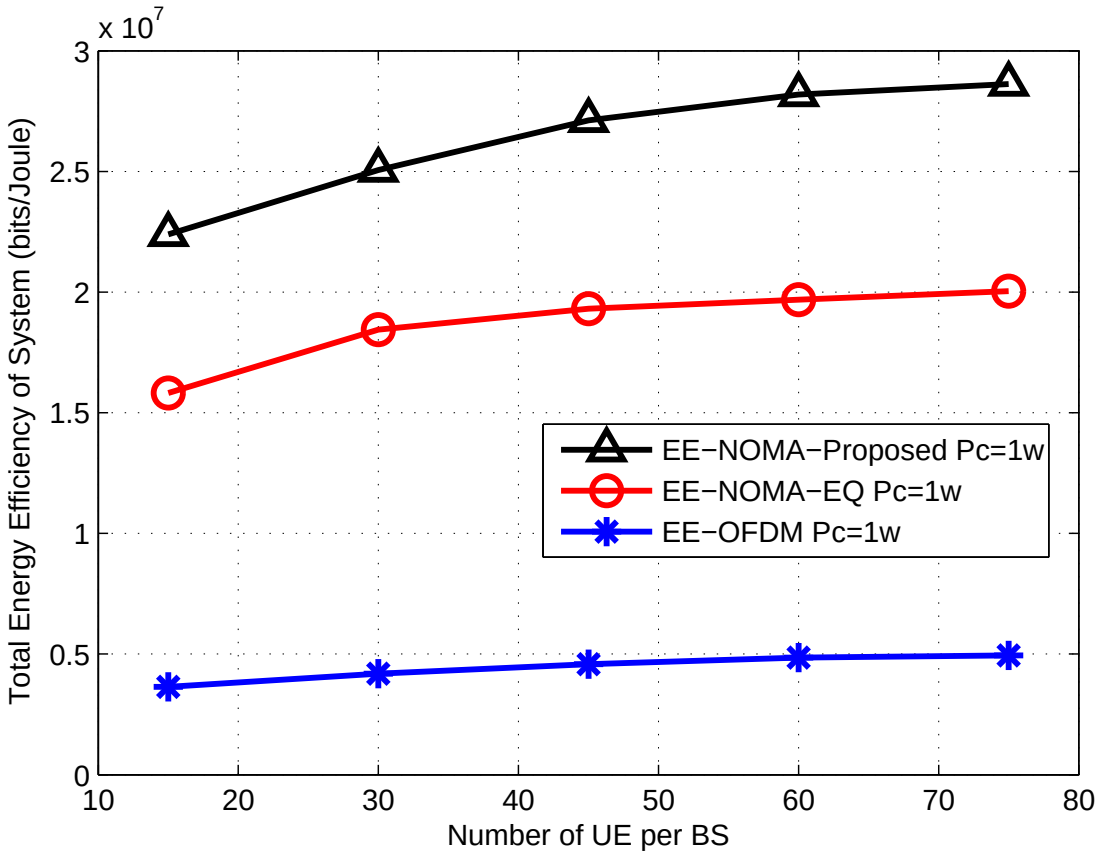

Figure 4.5: Energy efficiency performance versus BS maximum power.

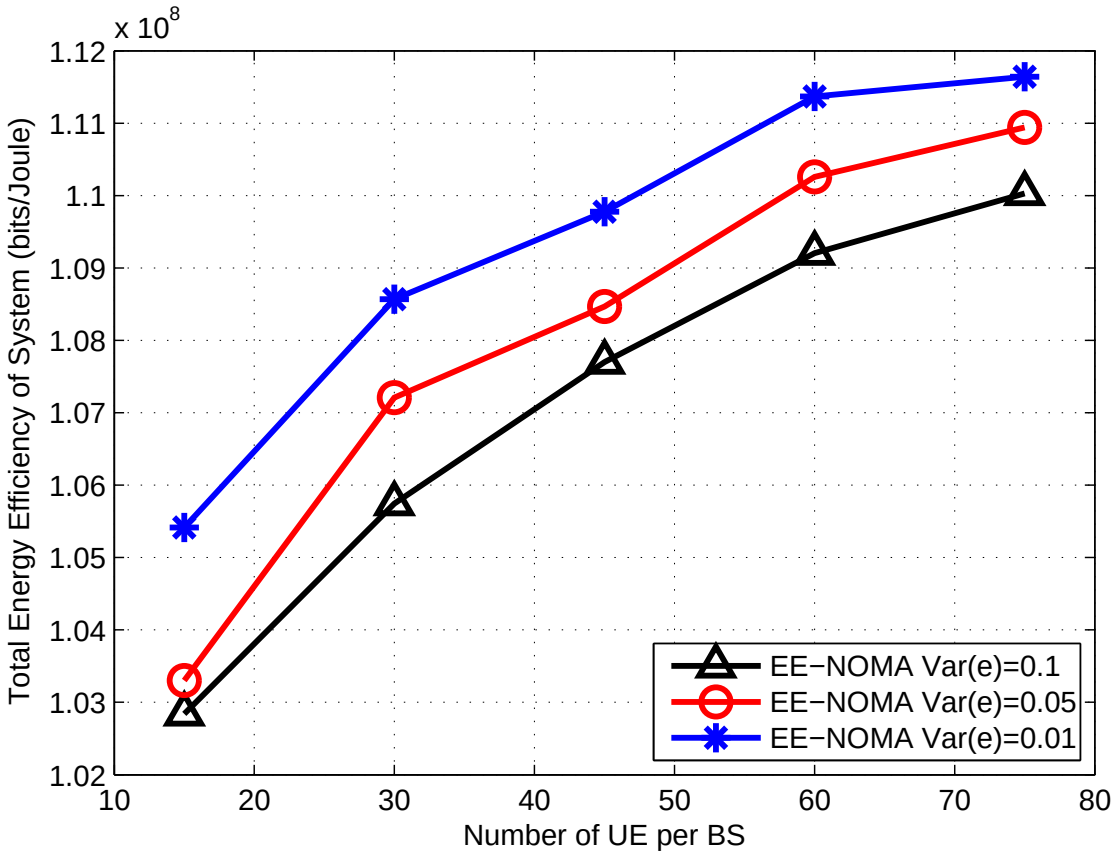

Figure 4.6: Energy efficiency performance versus users.

the number of users is 60, the energy efficiency of the proposed resource allocation scheme with $\sigma_e^2 = 0.01$ is 1.2% more than that with $\sigma_e^2 = 0.05$ and is 2.2% more than that with $\sigma_e^2 = 0.1$. Thus, as expected, the channel estimation error can degrade the energy efficiency performance.

4.6 Summary

For the NOMA system with imperfect CSI, we solved the resource allocation optimization problem by transforming a probabilistic mixed non-convex optimization problem to a non-probabilistic problem. A novel low-complexity suboptimal user scheduling algorithm was proposed to maximize the system energy efficiency. Given the user scheduling scheme, we proposed an optimal power allocation scheme and derived a closed form power allocation expression for users on each subchannel where the maximum user number can be greater than two. The effectiveness of the proposed scheme was verified by computer simulations and compared to the existing scheme in terms of energy efficiency. It was shown that the energy efficiency of the NOMA system with the proposed resource allocation scheme is higher than the considered referential scheme as well as the OFDMA scheme.

Chapter 5

Energy Efficient Resource Allocation for NOMA Heterogeneous Networks (HetNets)

Implementing NOMA in HetNets can alleviate the cross-tier interference and improve the system throughput via resource optimization. In this chapter, we aim to maximize the whole system energy efficiency in NOMA HetNet via subchannel allocation and power allocation. The task of energy efficient subchannel and power allocation in both macro cell and small cells is formulated as an integer mixed nonconvex problem. An iterative algorithm is proposed to maximize the macro cell and small cells energy efficiency. By considering the cochannel interference and cross-tier interference, an iterative algorithm is proposed to solve this optimization problem. In the proposed algorithm, convex relaxation and dual decomposition approaches are exploited to find the closed form optimal power allocation expression in each iteration. Finally, the complexity analysis and simulation results are provided to evaluate the system energy efficiency performance.

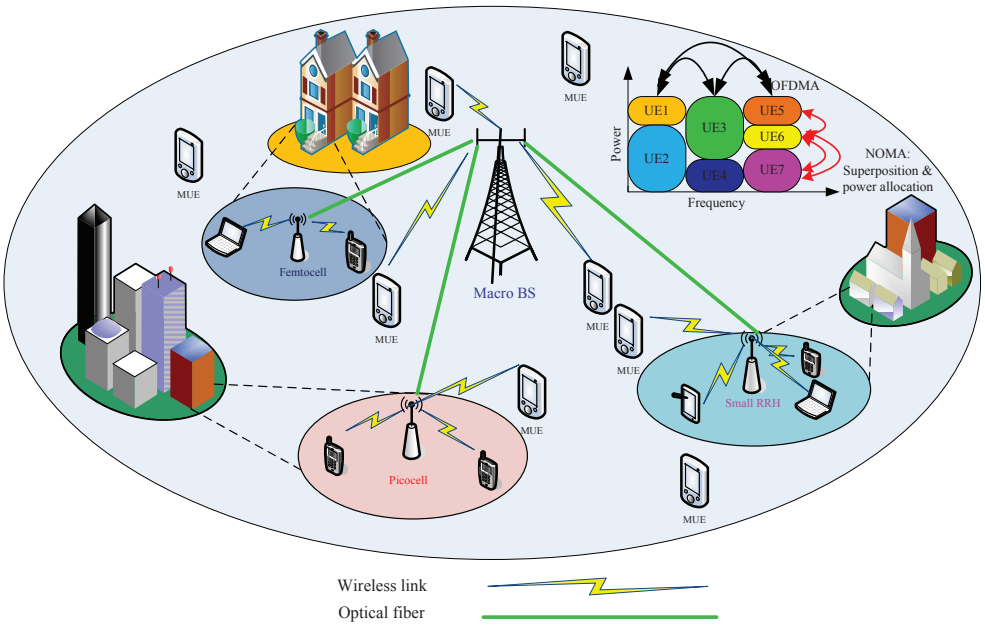

Figure 5.1: NOMA based heterogeneous networks.

5.1 System Model and Problem Formulation

5.1.1 NOMA HetNet System Model

In this system, we consider a downlink NOMA heterogeneous small cell network shown as Fig. 1, where one macro base station (MBS) is located in the center of macro cell. We assume that M macro user equipments (MUEs) are uniformly distributed within the macro cell overlaid by S small cells, e.g., picocells and femtocells. The index for the MUE and small cells are defined as $m \in \{1, 2, \cdots, M\}$ and $s \in \{1, 2, \cdots, S\}$, respectively. In each small cell, one small base station (SBS) is located in the small cell center and U user equipments (SUEs) are uniformly distributed within the small cell. We define $u \in \{1, 2, \cdots, U\}$ as the index of SUEs in each small cell. The MBS transmits signals to M MUEs through N subchannels and each small cell occupies one SC. The total bandwidth is B, which is divided into N subchannels and each subchannel occupies bandwidth $B_{sc} = B/N$.The maximum transmit power for MBS is $P^M_{\max}$ and the maximum transmit power of all small cells is $P^S_{\max}$.

A block fading channel is adopted in this system model, where the channel fading of

each subcarrier is assumed to be the same within a subchannel, but it varies independently across different subchannels. In this system, we allow MUEs and SUEs to reuse these N subchannels in order to improve the system spectrum efficiency. The cross-tier interference caused by the macro cell will be effectively mitigated by applying the NOMA technique. Note that the interference between small cells is neglected due to the following two reasons: first, because different small cells will occupy different subchannels and the subchannels are independent with each other; second, the inter-cell interference can be ignored due to the severe wall penetration loss [16]. We assume each UE (MUE or SUE) is equipped with one single antenna and is connected to one BS (MBS or SBS). We assume user association to the MBS and SBSs are completed before the resource allocation, which means that the users in the small cell are only connected to their corresponding SBS, and MUEs are connected to the MBS. Both the MBS and SBSs have the full knowledge of the channel state information obtained by the backhaul between the MBS and SBSs.

5.1.2 Channel Description

According to the NOMA protocol, superposition coding and successive interference cancelation are implemented in the BSs and UEs, respectively. In NOMA HetNets, each small cell applies NOMA, which means that the users in the same cell can be multiplexed on the same subchannel. The MUEs share the same subchannels with the SUEs. On each subchannel, by applying the NOMA protocol, the users who have larger channel gain will decode and remove the message from the users who have smaller channel gain. Denote $p^s_{u,n}$ and $p_{m,n}$ as the assigned power to the uth UE in the small cell s on the nth subchannel and transmit power from MBS to MUE m on SC n. The SBS s sends messages to SUE u through subchannel n. Define $g^s_{u,n}$ and $h^{MS}_{s,u,n}$, respectively, as the channel gain on subchannel n from SBS s to SUE u and the channel gain on the link n from MBS to SUE u in SBS s. Denote $\boldsymbol{\alpha} = [\alpha_{s,n}]_{S\times N}$ as the subchannel indicator matrix for the small cells, where $\alpha_{s,n} = 1$ denotes that subchannel n is assigned to the small cell s. Denote $\boldsymbol{\beta} = [\beta_{m,n}]_{M\times N}$ as subchannel indicator matrix for the MUEs, where $\beta_{m,n} = 1$ denotes that subchannel n is assigned to

the MUE m. We assume that each small cell has U SUEs. Generally, the channel gains of SUEs are sorted as $|g^s_{U,n}| \geq |g^s_{U-1,n}| \geq \cdots \geq |g^s_{u,n}| \geq \cdots \geq |g^s_{1,n}|$. The received signal by SUE u by SBS s on subchannel n is

$$y^s_{u,n} = g^s_{u,n}\sqrt{p^s_{u,n}}s^s_{u,n} + g^s_{u,n}\sum_{l=u+1}^{U}\sqrt{p^s_{l,n}}s^s_{l,n} + \sum_{m=1}^{M}\beta_{m,n}h^{MS}_{s,u,n}\sqrt{p_{m,n}}s_{m,n} + z^s_{u,n}. \tag{5.1}$$

The first term is the expected received signal from SBS s to SUE u; the second term is the interference from the users in the same small cell; the third term is the cross-tier interference; $z^s_{u,n} \sim \mathcal{CN}(0, \sigma_z^2)$ is the zero-mean complex additive white Gaussian noise random variable with variance σ_z^2. We assume that the MBS and SBS know the perfect CSI and are connected by wired links [83]. According to the Shannon's capacity formula, the data rate from SBS s to SUE u on SC n is

$$R^s_{u,n} = B_{sc}\log_2(1 + \gamma^s_{u,n}) \tag{5.2}$$

where

$$\gamma^s_{u,n} = \frac{\left|g^s_{u,n}\right|^2 p^s_{u,n}}{\left|g^s_{u,n}\right|^2 \sum\limits_{l=u+1}^{U} p^s_{l,n} + \sum\limits_{m=1}^{M}\beta_{m,n}\left|h^{MS}_{s,u,n}\right|^2 p_{m,n} + \sigma_z^2} \tag{5.3}$$

is the SINR of SUE u in small cell s. The total sum rate of small cells is

$$R^S = \sum_{s=1}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N}\alpha_{s,n}R^s_{u,n}. \tag{5.4}$$

The energy efficiency of small cell tier is defined as a ratio of the total sum rate of small cells to the total power consumption of small cells

$$EE^S = \frac{R^S}{P^S_T + P^S_c} \tag{5.5}$$

where $P^S_T = \sum\limits_{s=1}^{S}\sum\limits_{n=1}^{N}\sum\limits_{u=1}^{U}\alpha_{s,n}p^s_{u,n}$ is the total transmit power consumption and P^S_c is the total circuit power consumption of the small cells.

In the macro cell, we define $\boldsymbol{H} = [h^M_{m,n}]_{M\times N}$ as the channel gain on the link from MBS to MUE m on subchannel n. The interference from the small cell is also considered in this work. Denote $f^{SM}_{s,m}$ as the channel gain between SBS s to MUE m on the same subchannel. Without loss of generality, we assume M_n users are multiplexed on the subchannel n and channel gains are sorted as $|h^M_{M_n,n}| \geq \cdots \geq |h^M_{m,n}| \geq \cdots \geq |h^M_{1,n}|$. According to the Shannon's capacity formula, the data rate of MUE m can be written as

$$R_{m,n} = B_{sc}\log_2(1+\gamma_{m,n}) \tag{5.6}$$

where

$$\gamma_{m,n} = \frac{\left|h^M_{m,n}\right|^2 p_{m,n}}{\sum\limits_{i=m+1}^{M_n}\beta_{m,n}\left|h^M_{i,n}\right|^2 p_{i,n} + \sum\limits_{n=1}^{N}\sum\limits_{u=1}^{U}\alpha_{s,n}p^s_{u,n}\left|f^{SM}_{s,m}\right|^2 + \sigma_z^2}. \tag{5.7}$$

Therefore, the sum rate of macro cell is

$$R^M = \sum_{m=1}^{M}\sum_{n=1}^{N}\beta_{m,n}R_{m,n}. \tag{5.8}$$

The energy efficiency of the macro cell tier is defined as a ratio of the total sum rate to the total power consumption in the macro cell

$$EE^M = \frac{R^M}{P^M_T + P^M_c} \tag{5.9}$$

where $P^M_T = \sum\limits_{m=1}^{M}\sum\limits_{n=1}^{N}\beta_{m,n}p_{m,n}$ is the total transmit power consumption of the macro cell and P^M_c is the total circuit power consumption of the macro cell.

5.1.3 Problem Formulation

Our goal is to maximize the entire system energy efficiency including macro cell energy efficiency and small cell energy efficiency. For each tier, network's energy efficiency is formulated as a ratio of system sum rate to the total power consumption. The objective

is to maximize summation of network tier energy efficiency. Assume the MBS and SBSs have the perfect CSI. The resource allocation is performed by the entire system under the following definitions and constraints:

- The total power constraint:

$$P_T^S = \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U}\alpha_{s,n}p_{u,n}^s \leq P_{\max}^S \tag{5.10}$$

$$P_T^M = \sum_{m=1}^{M}\sum_{n=1}^{N}\beta_{m,n}p_{m,n} \leq P_{\max}^M. \tag{5.11}$$

- Quality of service requirement: The SUE data rate should be guaranteed for their basic communication, which requires the following constraint:

$$R_{u,n}^s \geq R_{\min}, \forall s, u, n. \tag{5.12}$$

  The MUE data rate should also be guaranteed for their basic communication, which requires

$$R_{m,n} \geq R_{\min}, \forall m, n. \tag{5.13}$$

- Cross-tier interference constraints: The interference from SBS s to MUEs who are also multiplexed on subchannel n. The cross-tier interference limit is constrained by a threshold $I_{n,th}^s$

$$I_n^s = \sum_{n=1}^{N}\alpha_{s,n}\sum_{u=1}^{U}p_{u,n}^s\left|f_{s,m}^{SM}\right|^2 \leq I_{n,th}^s, \quad \forall n. \tag{5.14}$$

  The interference from MBS to SUE u in small cell s is also limited by a threshold $I_{n,th}^M$

$$I_n^M = \sum_{m=1}^{M}\beta_{m,n}p_{m,n}|h_{s,u,n}^{MS}|^2 \leq I_{n,th}^M, \forall n. \tag{5.15}$$

The energy efficiency of NOMA HetNets can be defined as

$$EE = EE^S + EE^M. \tag{5.16}$$

Therefore, the energy efficient resource allocation for a downlink NOMA HetNet system can be formulated as

$$\max_{\{\alpha_{s,n}\},\{\beta_{m,n}\},\{p^s_{u,n}\},\{p_{m,n}\}} EE \tag{5.17}$$

$$\begin{aligned}
\text{s.t. } & C1: P_T^S \leq P_{\max}^S; P_T^M \leq P_{\max}^M, \\
& C2: p^s_{u,n} \geq 0, \forall s, u, n; p_{m,n} \geq 0, \forall m, n \\
& C3: R^s_{u,n} \geq R_{\min}, \forall s, u, n; R_{m,n} \geq R_{\min}, \forall m, n \\
& C4: I^s_n \leq I^s_{n,th}, \forall s; I^M_n \leq I^M_{n,th}, \forall n \\
& C5: \alpha_{s,n} \in \{0,1\} \forall s, n; \beta_{m,n} \in \{0,1\}, \forall m, n \\
& C6: \sum_{s=1}^{S} \alpha_{s,n} \leq 1, \forall n \\
& C7: \sum_{s=1}^{S} \sum_{u=1}^{U} \alpha_{s,n} + \sum_{m=1}^{M} \beta_{m,n} \leq U_{\max}, \forall n
\end{aligned} \tag{5.18}$$

where constraint $C1$ is the transmitted power limitation for all SBSs and MBS; Constraint $C2$ demonstrates that the transmitted power of BS should be no less than zero; Constraint $C3$ describes the heterogeneous QoS requirement that the data rata of each UE should be no less than the minimum user data rate $R_{\min}$. In constraint $C4$, the cross-tier interferences from small cell and macro cell are limited by $I^s_{n,th}$ and $I^M_{n,th}$, respectively; Constraints $C5$ and $C6$ are imposed to guarantee that each subchannel can only be assigned to at most one small cell according to $C5$; Constraint $C7$ limits the user number on the same subchannel.

5.2 Energy Efficient Resource Allocation for NOMA HetNets

It is challenging to find the global optimal solution to the problem (5.17) within polynomial time. To solve this problem efficiently, we first deal with macro cell subchannel allocation to MUEs with equal power allocation policy. Based on the value of $\beta_{m,n}$, the energy-efficient subchannel allocation and power allocation for SUEs can be iteratively solved by Algorithm 8 where a closed form optimal power allocation expression is derived by the Lagrangian approach. Finally, we update the power allocation for MUEs to further improve the system energy efficiency shown in Algorithm 9.

5.2.1 Energy Efficiency Optimization for the Entire System Algorithm Design

Assume the MBS knows the entire knowledge of channel statement information $\boldsymbol{H} = [h^M_{m,n}]_{M\times N}$. To reduce the complexity of the global optimal solution, we first decouple the problem into macro cell energy efficiency maximization and small cell energy efficiency maximization subproblems. A low-complexity suboptimal algorithm is designed to maximize the system energy efficiency, as shown in Algorithm 7.

In this algorithm, we first determine subchannel allocation for MUEs in macro cell. Equal power is allocated for all MUEs. We define a set as $\mathbb{U}^M_{un}$ to record the unallocated MUEs and let it equal to $1, 2, \cdots, M$. For each user, we will find subchannel n^* who has the maximum channel gain among N subchannels. We need to check the number of MUEs on subchannel n^*. Allocate this MUE on subchannel n^* if the user number is less than $U^M_{\max}$. However, if the MUE number is equal to $U^M_{\max}$, this subchannel n^* will choose the user set who can provide the maximum energy efficiency [57]. The MUE who has not been chosen will be put back into the set $\mathbb{U}^M_{un}$. For this unallocated MUE, we will repeat the allocation progress until it has been allocated on one subchannel. This subchannel allocation for MUEs procedure will terminate until all the MUEs has been allocated on subchannels.

Algorithm 7: An Iterative Energy Efficient Resource Allocation Algorithm for NOMA HetNets

Initialize the power allocation for MUEs $p_{m,n} = P_{\max}/M$.
Initialize the sets $\mathbb{U}_{un}^{M} = 1, 2, \cdots, M$ to record the unallocated user in the system.
while $\mathbb{U}_{un}^{M}$ is not empty **do**
  for $m = 1$ to M **do**
    1. Find the subchannel n^* who has the maximum channel gain in $\boldsymbol{H}$

$$n^* = \max_{n} \boldsymbol{H}$$

    if the number of MUEs on this subchannel n^* is less than $U_{\max}^{M}$ **then**
      a) Schedule the MUE m onto the subchannel n^*.
      b) $\mathbb{U}_{un}^{M}$=$\mathbb{U}_{un}^{M} \backslash m$.
    end if
    if the number of multiplexed users on this subchannel equals $U_{\max}$ **then**
      a) Subchannel n^* selects a set of $U_{\max}^{M}$ users who can provide maximum energy efficiency.
      b) $\mathbb{U}_{un}^{M}$=$\mathbb{U}_{un}^{M} \backslash m$.
      c) The unchosen MUE will go back to Step 1 and find the maximum channel gain among $\{1, 2, \cdots, N\} \backslash n^*$ repeat this step until it have been allocated on one subchannel.
    end if
  end for
end while
After all MUEs are allocated on subchannels, $\alpha_{s,n}$ and $P_{u,n}^{s}$ can be optimized for small cells by Algorithm 8.
Power allocation for MUEs can be updated by Algorithm 9.

After all the MUEs have been assigned to different subchannels, we focus on subchannel allocation for small cells and power allocation for SUEs. An iterative algorithm is proposed to solve this problem, as shown in Algorithm 8. To further improve the entire system energy efficiency, another iterative algorithm is proposed to update the power allocation for MUEs, as shown in Algorithm 9 of Section 5.2.4. In both Algorithm 8 and Algorithm 9, closed form power allocation expressions for SUEs and MUEs are derived by the Lagrangian approach. Details of the derivation can be found in Appendix C.

5.2.2 Energy Efficiency Optimization for Small Cells

In this subsection, we design an iterative algorithm to allocate subchannels to small cell and power allocation for SUEs in order to maximize the small cells energy efficiency. Based on the value of $\beta_{m,n}$, the energy efficient resource allocation for small cells can be formulated as

$$\max_{\{\alpha_{s,n}\},\{p^s_{u,n}\}} \frac{\sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N}\alpha_{s,n}R^s_{u,n}}{\sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N}\alpha_{s,n}p^s_{u,n}+P^S_c} \tag{5.19}$$

$$\begin{aligned}
\text{s.t. } C1: &\sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U}\alpha_{s,n}p^s_{u,n} \le P^S_{\max},\\
C2: &\, p^s_{u,n} \ge 0, \forall s,u,n\\
C3: &\, \alpha_{s,n}R^s_{u,n} \ge R_{\min}, \forall s,u,n\\
C4: &\sum_{s=1}^{S}\sum_{u=1}^{U}\alpha_{s,n}p^s_{u,n}\left|f^{SM}_{s,m}\right|^2 \le I^s_{n,th}, \forall n\\
C5: &\, \alpha_{s,n} \in \{0,1\}, \forall s,n\\
C6: &\sum_{s=1}^{S}\alpha_{s,n} \le 1, \forall n.
\end{aligned} \tag{5.20}$$

This problem is mixed-integer programming due to constraint $C5$. The optimal solution to this non-convex problem has an extremely high complexity. To efficiently solve this

problem, by using convex relaxation, we first relax the subchannel indication variable $\alpha_{s,n}$ to be a continuous real variable in [0,1]. Since the range of $\alpha_{s,n}$ is between zero and one, we could consider it as a time-sharing factor for subchannel n. It denotes the fraction of time that small cell s occupies subchannel n during one block transmission. This relaxation was first proposed to modify the integer mixed problem with the relaxed constraints for subcarrier allocation in OFDM [84]. The duality gap of the relaxed problem is proved to be zero [85].

Since the problem is fractional nonlinear programming, and it can be transformed to an equivalent parameterized non-fractional form [66]. The equivalent subtractive problem with relaxation can be formulated as

$$\max_{\{\alpha_{s,n}\},\{p^s_{u,n}\}} \sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N} \alpha_{s,n} R^s_{u,n} - t\left(\sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N} \alpha_{s,n} p^s_{u,n} + P^S_c\right) \tag{5.21}$$

$$\begin{aligned}
\text{s.t. } C1 &: \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U} \alpha_{s,n} p^s_{u,n} \le P^S_{\max}, \\
C2 &: p^s_{u,n} \ge 0, \forall s,u,n \\
C3 &: \alpha_{s,n} R^s_{u,n} \ge R_{\min}, \forall s,u,n \\
C4 &: \sum_{s=1}^{S}\sum_{u=1}^{U} \alpha_{s,n} p^s_{u,n} \left|f^{SM}_{s,m}\right|^2 \le I^s_{n,th}, \forall n \\
C5 &: \alpha_{s,n} \in [0,1], \forall s,n \\
C6 &: \sum_{s=1}^{S} \alpha_{s,n} \le 1, \forall n
\end{aligned} \tag{5.22}$$

where t is a parameter introduced to scale the weight of the total power consumption of small cells. For a given value t, the solution to the problem can be denoted as $\{\alpha_{s,n}\}$ and $\{p^s_{u,n}\}$. We define

$$f(t) \triangleq \max_{\{\alpha_{s,n}\},\{p^s_{u,n}\}} \sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N} \alpha_{s,n} R^s_{u,n} - t\left(\sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N} \alpha_{s,n} p^s_{u,n} + P^S_c\right). \tag{5.23}$$

It is observed that $f(t)$ is negative when t approaches infinity, while $f(t)$ is positive when t approaches minus infinity. Therefore, $f(t)$ is convex with respect to t. Define $\{\alpha_{s,n}^*\}$ and $\{p_{u,n}^{s,*}\}$ are the optimal subchannel allocation policy and power allocation policy for problem (5.20). Therefore, the maximum energy efficiency t^* can be achieved if and only if

$$f(t^*) = R^S\left(\{\alpha_{s,n}^*\},\{p_{u,n}^{s,*}\}\right) - t^*\left(P_T^S\left(\{\alpha_{s,n}^*\},\{p_{u,n}^{s,*}\}\right) + P_c^S\right) = 0 \tag{5.24}$$

where the maximum energy efficiency of the small cells can be defined as

$$t^* = \frac{R^S\left(\{\alpha_{s,n}^*\},\{p_{u,n}^{s,*}\}\right)}{P_T^S\left(\{\alpha_{s,n}^*\},\{p_{u,n}^{s,*}\}\right) + P_c^S}. \tag{5.25}$$

For notational simplicity, we denote the actual power allocation to SUE u in small cell s on subchannel n as $\tilde{p}_{u,n}^s = \alpha_{s,n} p_{u,n}^s$. Thus, the data rate of SUE u in small cell s on subchannel n can be written by

$$\tilde{R}_{u,n}^s = B_{sc}\log_2\left(1 + \frac{\left|g_{u,n}^s\right|^2 \tilde{p}_{u,n}^s}{\left|g_{u,n}^s\right|^2 \sum\limits_{l=u+1}^{U} \tilde{p}_{l,n}^s + \alpha_{s,n} I_n^M + \alpha_{s,n}\sigma_z^2}\right). \tag{5.26}$$

Then the problem can be rewritten as

$$\max_{\{\alpha_{s,n}\},\{p_{u,n}^s\}} \sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N} \alpha_{s,n}\tilde{R}_{u,n}^s - t\left(\sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N} \tilde{p}_{u,n}^s + P_c^S\right) \tag{5.27}$$

$$
\begin{aligned}
\text{s.t. } C1: & \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s} \leq P_{\max}^{s}, \\
C2: & \tilde{p}_{u,n}^{s} \geq 0, \forall s,u,n \\
C3: & \alpha_{s,n}\tilde{R}_{u,n}^{s} \geq R_{\min}, \forall s,u,n \\
C4: & \sum_{s=1}^{S}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}\left|f_{s,m}^{SM}\right|^{2} \leq I_{n,th}^{s}, \forall n \\
C5: & \alpha_{s,n} \in [0,1], \forall s,n \\
C6: & \sum_{s=1}^{S}\alpha_{s,n} \leq 1, \forall n.
\end{aligned}
\tag{5.28}
$$

For a given value t, the Hessian matrix of the objective function in (5.27) with respect to $\tilde{p}_{u,n}^{s}$ and $\alpha_{s,n}$ is negative semi-definite. The objective function (5.27) is concave [66]. As the inequality constraints in (5.28) are convex, the feasible set of objective function is convex. Being a convex optimization problem, the transformed optimization problem in (5.27) has a unique optimal solution, i.e., the local solution is the optimal solution, and it can be obtained in polynomial time. Therefore, the dual decomposition method can be used to solve this problem. The Lagrangian function of the problem (5.27) can be written by

$$
\begin{aligned}
& L\left(\{\alpha_{s,n}\},\{p_{u,n}^{s}\},t,\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right) \\
=& \sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N}\alpha_{s,n}\tilde{R}_{u,n}^{s} - t\left(\sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N}\tilde{p}_{u,n}^{s} + P_{c}^{S}\right) \\
& + \lambda^{s}\left(P_{\max}^{s} - \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}\right) \\
& + \sum_{s}^{S}\sum_{u=1}^{U}\sum_{n=1}^{N}\nu_{u,n}^{s}\left(\alpha_{s,n}\tilde{R}_{u,n}^{s} - R_{\min}\right) \\
& + \sum_{n=1}^{N}\mu_{n}^{s}\left(I_{n,th}^{s} - \sum_{s=1}^{S}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}\left|f_{s,m}^{SM}\right|^{2}\right) \\
& + \sum_{n=1}^{N}\eta_{n}^{s}\left(1 - \sum_{s=1}^{S}\alpha_{s,n}\right)
\end{aligned}
\tag{5.29}
$$

where $\lambda^{s} \geq 0$, $\boldsymbol{\nu}^{s} \succeq 0$, $\boldsymbol{\mu}^{s} \succeq 0$, and $\boldsymbol{\eta}^{s} \succeq 0$ are the Lagrange multipliers corresponding to

the power constraints. The Lagrangian function can be rewritten as

$$\begin{aligned}
&L\left(\{\alpha_{s,n}\},\{p_{u,n}^{s}\},t,\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right)\\
=&\sum_{s=1}^{S}\sum_{n=1}^{N}L_{s,n}\left(\{\alpha_{s,n}\},\{p_{u,n}^{s}\},\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right)\\
&-tP_{c}^{S}+\lambda^{s}\left(P_{\max}^{S}\right)-\sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U}\nu_{u,n}^{s}R_{\min}+\sum_{n=1}^{N}\mu_{n}^{s}I_{n,th}^{s}+\sum_{n=1}^{N}\eta_{n}
\end{aligned}\tag{5.30}$$

where

$$\begin{aligned}
&L_{s,n}\left(\{\alpha_{s,n}\},\{p_{u,n}^{s}\},t,\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right)\\
=&\sum_{u=1}^{U}\alpha_{s,n}\tilde{R}_{u,n}^{s}-t\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}-\lambda^{s}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}+\sum_{u=1}^{U}\nu_{u,n}^{s}\alpha_{s,n}\tilde{R}_{u,n}^{s}\\
&-\mu_{n}^{s}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}\left|f_{s,m}^{SM}\right|^{2}-\eta_{n}^{s}\alpha_{s,n}.
\end{aligned}\tag{5.31}$$

Given by t, the dual problem of (5.27) is

$$\min_{\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}} g\left(\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right)\tag{5.32}$$

$$\text{s.t.}\quad \lambda^{s}\geq 0,\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\succeq 0\tag{5.33}$$

where

$$g\left(\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right)=\max_{\{\alpha_{s,n}\},\{p_{u,n}^{s}\}} L\left(\{\alpha_{s,n}\},\{p_{u,n}^{s}\},t,\lambda^{s},\boldsymbol{\nu}^{s},\boldsymbol{\mu}^{s},\boldsymbol{\eta}^{s}\right).\tag{5.34}$$

We decompose the dual problem into two layers: inner layer and outer layer. We first solve the inner layer problem to obtain the subchannel allocation power allocation policy for small cells, and then outer layer to compute the dual variables iteratively. To reduce the complexity of the optimal solutions, the lower bound is applied to achieve the optimal solution iteratively. With lower bound [79], the data rate of SUE u in small cell s on subchannel n can be rewritten by

$$\hat{R}_{u,n}^{s}=B_{sc}a_{u,n}^{s}\log_{2}(\gamma_{u,n}^{s})+b_{u,n}^{s}.\tag{5.35}$$

According to (5.30), the Lagrangian dual function can be decomposed into $S \times N$ subproblems. We define the actual optimal power allocation to SUE u in small cell s on subchannel n as $\tilde{p}_{u,n}^{s,*} = \alpha_{s,n} p_{u,n}^{s,*}$. According to the KKT conditions, assume we have the channel gains satisfying $|g_{U,n}^s| \geq |g_{U-1,n}^s| \geq \cdots \geq |g_{1,n}^s|$, the optimal power allocation for SUEs in small cell s can be derived as

$$p_{u,n}^{s,*} = \frac{\tilde{p}_{u,n}^{s,*}}{\alpha_{s,n}} = \frac{B_{sc} a_{u,n}^s \left(1 + \nu_{u,n}^s\right)}{\sum\limits_{l=1}^{u-1} B_{sc} a_{l,n}^s \left(1 + \nu_{l,n}^s\right) \left(\frac{\gamma_{l,n}^s}{p_{l,n}^{s,*}}\right) + \ln 2 \left(\eta^s + \lambda^s + \mu_n^s \left|f_{s,m}^{SM}\right|^2\right)} \tag{5.36}$$

where

$$\gamma_{u,n}^s = \frac{\left|g_{u,n}^s\right|^2 p_{u,n}^{s,*}}{\left|g_{1,n}^s\right|^2 \sum\limits_{l=u+1}^{U} p_{l,n}^{s,*} + I_n^M + \sigma_z^2}. \tag{5.37}$$

It can be observed from (5.36) and (5.37) that the power allocation policy is a fixed point equation. Denote $\boldsymbol{\alpha} = \{p_{u,n}^s\}$ and $\boldsymbol{P}^S = \{p_{u,n}^s\}$ where $s = 1, 2, \cdots, S$, $u = 1, 2, \cdots, U$ and $n = 1, 2, \cdots, N$. According to the positivity, monotonicity and scalability of variable $\boldsymbol{P}^S$, the power allocation can be updated by each iteration with (5.36) and (5.37).

To obtain $\alpha_{s,n}$, the partial derivation of the Lagrangian can be expressed as

$$\frac{\partial L_{s,n}(\cdots)}{\partial \alpha_{s,n}} = \Delta_{s,n} - \eta_n^s \begin{cases} < 0, & \alpha_{s,n} = 0 \\ = 0, & 0 < \alpha_{s,n} < 1 \\ > 0, & \alpha_{s,n} = 1 \end{cases} \quad \forall s, n \tag{5.38}$$

where

$$\begin{aligned}\Delta_{s,n} =& \sum_{u=1}^{U}\left(1+\nu_{u,n}^{s}\right)B_{sc}a_{u,n}^{s}\log_2\left(\frac{\left|g_{u,n}^{s}\right|^2 p_{u,n}^{s,*}}{\left|g_{u,n}^{s}\right|^2\sum\limits_{l=u+1}^{U}p_{l,n}^{s,*}+I_n^M+\sigma_z^2}\right)\\ &-\sum_{u=1}^{U}\left(1+\nu_{u,n}^{s}\right)\frac{B_{sc}a_{u,n}^{s}}{\ln 2}\left(\frac{I_n^M+\sigma_z^2}{\left|g_{u,n}^{s}\right|^2\sum\limits_{l=u+1}^{U}p_{l,n}^{s,*}+I_n^M+\sigma_z^2}\right)\\ &-\left(t+\lambda^s\right)\sum_{u=1}^{U}p_{u,n}^{s,*}-\mu_n^s\sum_{u=1}^{U}p_{u,n}^{s,*}\left|f_{s,m}^{SM}\right|^2.\end{aligned} \tag{5.39}$$

Subchannel n^* is assigned to small cell s with the largest $\Delta_{s,n}$, that is

$$\alpha_{s,n}^{*} = 1|_{n^*=\max\limits_{n}\Delta_{s,n}}, \forall s. \tag{5.40}$$

The outer layer primal problem can be solved by the gradient method since the objective function is differentiable. Therefore, the dual variables can be updated with gradient descent as

$$\lambda^s\left(i+1\right) = \left[\lambda^s\left(i\right)-\xi_1\left(i\right)\times\left(P_{\max}^{s}-\sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}\right)\right]^{+}, \tag{5.41}$$

$$\nu_{u,n}^{s}\left(i+1\right) = \left[\nu_{u,n}^{s}\left(i\right)-\xi_2\left(i\right)\times\left(\alpha_{s,n}\tilde{R}_{u,n}^{s}-R_{\min}\right)\right]^{+}, \forall s,u,n \tag{5.42}$$

$$\mu_n^s\left(i+1\right) = \left[\mu_n^s\left(i\right)-\xi_3\left(i\right)\times\left(I_{n,th}^{s}-\sum_{s=1}^{S}\sum_{u=1}^{U}\tilde{p}_{u,n}^{s}\left|f_{s,m}^{SM}\right|^2\right)\right]^{+}, \forall n \tag{5.43}$$

$$\eta_n^s\left(i+1\right) = \left[\eta_n^s\left(i\right)-\xi_4\left(i\right)\times\left(1-\sum_{s=1}^{S}\alpha_{s,n}\right)\right]^{+}, \forall n \tag{5.44}$$

where i is the iteration index. $\xi_1\left(i\right)$, $\xi_3\left(i\right)$, $\xi_3\left(i\right)$ and $\xi_4\left(i\right)$ are positive step sizes at iteration i. Since the transformed problem (5.27) is concave, this guarantees that the iteration process converges to an optimal solution to problem (5.27) based on appropriate step sizes.

5.2.3 Algorithm Design

In this subsection, we design an iterative algorithm to obtain the energy efficient subchannel allocation and power allocation in small cells, as shown in Algorithm 8. Note that the wire links between MBS and SBSs are assumed to help SBSs coordinate with MBS.

In Algorithm 8, given the maximum iteration number and maximum tolerance, the energy efficiency improves for each iteration until it converges. In each iteration, the Lagrange multiplier approach is used to solve the problem (5.27). Given the energy efficiency t, the optimal power allocation for each user and subchannel indication determination will be developed iteratively until convergence.

Algorithm 8: An Iterative Energy Efficient Resource Allocation Algorithm

Initialize the maximum number of the iterations $I_{\max}$ and the maximum tolerance ε.
Initialize the energy efficiency t and the iteration index $i = 0$.
Initialize $p_{u,n}^{s}$ with a uniform power distribution among all subchannels
Initialize subchannel allocation for small cells $\alpha_{s,n}$ method in
while $|R^S\left(\boldsymbol{\alpha}(i), \boldsymbol{P}^S(i)\right) - t(i-1)\left(P_T^S\left(\boldsymbol{\alpha}(i), \boldsymbol{P}^S(i)\right) + P_c^S\right)| > \varepsilon$ or $i \leq I_{\max}$ **do**
  for $n = 1$ to N **do**
    for $s = 1$ to S **do**
      for $u = 1$ to U **do**
        1. Given the energy efficiency $t(i)$, update $p_{u,n}^{s,*}$ according to the optimal power allocation policy eq.
        2. Calculate $\Delta_{s,n}$ according to eq. (5.36).
        3. Update $\alpha_{s,n}^{*}$ according to eq. (5.40).
        4. Update $\lambda^s, \nu^s, \mu^s, \eta^s$ by (5.41)-(5.44).
      end for
    end for
  end for
  Set $i = i + 1$ and $t(i) = \frac{R^S\left(\boldsymbol{\alpha}(i-1), \boldsymbol{P}^S(i-1)\right)}{P_T^S\left(\boldsymbol{\alpha}(i-1), \boldsymbol{P}^S(i-1)\right) + P_c^S}$.
end while

5.2.4 Macro Cell Energy Efficiency Maximization

Based on Algorithm 8, the power allocation and subchannel allocation have been determined for all the small cells. Now each small cell occupies one subchannel and has the interference to the MUEs who are multiplexed on the same subchannel. However, the MUEs

on the same subchannel are allocated with equal power. To improve the energy efficiency of the macro cell, we now update the power allocation of MUEs. We define the power allocation policy of MUEs as $\boldsymbol{P}^M = \{p_{m,n}\}$. The macro cell energy efficiency optimization problem can be formulated as

$$\max_{\boldsymbol{P}^M} \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} R_{m,n}}{\sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n} + P_c^M} \tag{5.45}$$

$$\begin{aligned} \text{s.t. } C1 &: \sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n} \leq P_{\max}^M \\ C2 &: p_{m,n} \geq 0, \forall m, n \\ C3 &: R_{m,n} \geq R_{\min}, \forall m, n \\ C4 &: \sum_{m=1}^{M} \beta_{m,n} p_{m,n} |h_{s,u,n}^{MS}|^2 \leq I_{n,th}^M. \end{aligned} \tag{5.46}$$

For the power allocation, this optimization problem can be transformed into an equivalent subtractive problem with a parameter η^M [66], which is defined as the system energy efficiency based on given power allocation following by the similar steps (5.23)-(5.25). We define a function of η^M function as

$$f\left(\eta^M\right) \triangleq \max_{\{p_{m,n}\}} \sum_{m=1}^{M}\sum_{n=1}^{N} R_{m,n} - \eta^M \left(\sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n} + P_c^M\right) \tag{5.47}$$

which is convex with respect to η^M. Therefore, the optimal energy efficiency for macro cell $\eta^{M,*}$ can be achieved when

$$f\left(\eta^{M,*}\right) = R^M\left(\{p_{m,n}^*\}\right) - \eta^{M,*}\left(P_T^M\left(p_{m,n}^*\right) + P_c^M\right) = 0 \tag{5.48}$$

with the optimal power allocation $\boldsymbol{P}^{M,*} = \{p_{m,n}^*\}$ for MUEs. Therefore, the transformed

subtractive form problem can be written by

$$\max_{\mathbf{p^M}} \sum_{m=1}^{M}\sum_{n=1}^{N} R_{m,n} - \eta^M \left(\sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n} + P_c^M\right) \tag{5.49}$$

$$\begin{aligned} \text{s.t. } C1 &: \sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n} \le P_{\max}^M, \\ C2 &: p_{m,n} \ge 0, \forall m,n, \\ C3 &: R_{m,n} \ge R_{\min}, \forall m,n \\ C4 &: \sum_{m=1}^{M} p_{m,n}|h_{s,u,n}^{MS}|^2 \le I_{u,n,th}^M, \forall u,n. \end{aligned} \tag{5.50}$$

To solve (5.49), an iterative algorithm is proposed to find the optimal power allocation for MUEs by iteratively solving the convex subproblems, as shown in Algorithm 9. By utilizing the dual decomposition method, the Lagrangian function of problem (5.49) can be written by

$$\begin{aligned} & L\left(\{p_{m,n}\}, \eta^M, \lambda^M, \boldsymbol{\nu}^M, \boldsymbol{\mu}^M\right) \\ &= \sum_{m=1}^{M_n}\sum_{n=1}^{N} R_{m,n} - \eta^M \left(\sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n} + P_c^M\right) + \lambda^M \left(P_{\max}^M - \sum_{m=1}^{M}\sum_{n=1}^{N} p_{m,n}\right) \\ &\quad + \sum_{m=1}^{M}\sum_{n=1}^{N} \nu_{m,n}^M R_{m,n} - R_{\min} + \sum_{n=1}^{N}\sum_{u=1}^{U} \mu_{u,n}^M \left(I_{u,n,th}^M - \sum_{m=1}^{M} p_{m,n}|h_{u,n}^{MS}|^2\right) \\ &= \sum_{m=1}^{M_n}\sum_{n=1}^{N} L_{m,n}\left(\{p^M\}, \eta^M, \lambda^M, \boldsymbol{\nu}^M, \boldsymbol{\mu}^M\right) - \eta^M P_c^M + \lambda^M \left(P_{\max}^s\right) \\ &\quad - \sum_{m=1}^{M}\sum_{n=1}^{N} \nu_{m,n}^M R_{m,n} + \sum_{n=1}^{N} \mu_n^M I_{u,n,th}^M \end{aligned} \tag{5.51}$$

where

$$\begin{aligned} & L_{m,n}\left(\{p_{m,n}\}, \eta^M, \lambda^M, \boldsymbol{\nu}^M, \boldsymbol{\mu}^M\right) \\ &= \left(1 + \nu_{m,n}^M\right) R_{m,n} - \left(\eta^M + \lambda^M\right) p_{m,n} - \sum_{u=1}^{U} \mu_{u,n}^M p_{m,n}|h_{u,n}^{MS}|^2 \end{aligned} \tag{5.52}$$

λ^M, $\boldsymbol{\nu}^M$ and $\boldsymbol{\mu}^M$ are the Lagrange multipliers corresponding to the power constraints.

Given η^M, the corresponding dual problem of (5.49) is

$$\min_{\lambda^M \geq 0, \boldsymbol{\nu}^M, \boldsymbol{\mu}^M \succ 0} \max_{\{p_{m,n}\}} L\left(\{p_{m,n}\}, \eta^M, \lambda^M, \nu^M, \mu^M\right). \tag{5.53}$$

The dual decomposition approach is exploited to solve the dual problem (5.53). Given by the parameter η^M and fixed Lagrangian multipliers λ^M, $\boldsymbol{\nu}^M$ and $\boldsymbol{\mu}^M$, the inner subproblem is a convex problem. Therefore, the optimal power allocation for MUEs on subchannel n can be written by

$$p_{m,n}^{M,*} = \frac{B_{sc} a_{2,n}^M \left(1 + \nu_{1,n}^M\right)}{\sum\limits_{l=1}^{m-1} B_{sc} a_{l,n}^M \left(1 + \nu_{l,n}^M\right) \left(\frac{\gamma_{l,n}^M}{p_{l,n}^{M,*}}\right) + \ln 2 \left(\eta^M + \lambda^M + \sum\limits_{u=1}^{U} \mu_{u,n}^M \left|h_{u,n}^{MS}\right|^2\right)} \tag{5.54}$$

where

$$\gamma_{m,n}^M = \frac{\left|h_{m,n}^M\right|^2 p_{m,n}}{\sum\limits_{i=m+1}^{M_n} \left|h_{m,n}^M\right|^2 p_{i,n} + \sum\limits_{n=1}^{N} \sum\limits_{u=1}^{U} \alpha_{s,n} p_{u,n}^s \left|f_{s,m}^{SM}\right|^2 + \sigma_z^2}. \tag{5.55}$$

Given the power allocation scheme in (5.54), the outer layer primal problem can be solved by the gradient method since the objective function is differentiable. Therefore, the dual variables can be updated with gradient descent as

$$\lambda^M(i+1) = \left[\lambda^M(i) - \zeta_1(i) \times \left(P_{\max}^M - \sum_{m=1}^{M} \sum_{n=1}^{N} p_{m,n}\right)\right]^+, \tag{5.56}$$

$$\nu_{m,n}^M(i+1) = \left[\nu_{u,n}^s(i) - \zeta_2(i) \times (R_{m,n} - R_{\min})\right]^+, \forall m, n, \tag{5.57}$$

$$\mu_{u,n}^M(i+1) = \left[\mu_{u,n}^M(i) - \zeta_3(i) \times \left(I_{u,n,th}^M - \sum_{m=1}^{M} p_{m,n} |h_{u,n}^{MS}|^2\right)\right]^+, \forall u, n \tag{5.58}$$

where i is the iteration index; $\zeta_1(i)$, $\zeta_2(i)$ and $\zeta_3(i)$ are positive step sizes at iteration i.

Based on (5.42)-(5.55), Algorithm 9 is proposed to update the power allocation for MUEs in macro cell. Since the transformed problem (5.49) is concave, this guarantees that the iteration process converges to an optimal solution to problem (5.49) based on appropriate step sizes.

Algorithm 9: An Iterative Energy Efficient Resource Allocation Algorithm

1: Initialize the maximum number of the iterations $I_{\max}$ and the maximum tolerance ε.
2: Initialize the energy efficiency η^M and the iteration index $i = 0$.
3: Initialize $p_{m,n}^M$ with a uniform power distribution among all subchannels.
4: **while** $|R^M\left(\boldsymbol{P}^M(i)\right) - \eta^M(i-1)\left(P_T^M\left(\boldsymbol{P}^M(i)\right) + P_c^M\right)| > \varepsilon$ or $i \leq I_{\max}$ **do**
5: 1. Given the energy efficiency η^M, update $p_{m,n}^*$ according to the optimal power allocation policy (5.54) and (5.55).
6: 2. Update η^M, λ^M, $\boldsymbol{\nu}^M$ and $\boldsymbol{\mu}^M$ by (5.55)-(5.58).
7: 3. Set $i = i + 1$ and $\eta^M(i) = \frac{R^M\left(\boldsymbol{P}^M(i-1)\right)}{P_T^M\left(\boldsymbol{P}^M(i-1)\right)+P_c^M}$.
8: **end while**

5.3 Simulation Results

Simulation results are presented to demonstrate the effectiveness of the proposed algorithms. In our simulations, we assume that all users are uniformly distributed in each small cell coverage area, and the small cells are uniformly distributed in the macro cell coverage area. The radius of the macro cell is 300 m. The radius of each small cell is 10 m. Small cell has a minimum distance of 50 m from the macro base station. The minimum distance between small cell base stations is 40 m. The path loss[8] model is based on [81]. We assume that the shadowing standard deviation between base station and the users is 10 dB. The channel fading is composed of shadowing fading, path loss, and Rayleigh fading. The additive white Gaussian noise power is set as σ^2=3.9811 $\times$ 10^{-14} W. We assume that the maximum transmit power is 40 dBm at the macro cell base station, the maximum transmit power is 17 dBm in each small cell, and circuit power of each user is 20 dBm. The minimum data rate of each user with unit bandwidth ($R_{\min}$) is 7 bits/s.

Figure 5.1 evaluates the energy efficiency performance versus the number of iterations of Algorithm 8. We set the minimum data rate with unit bandwidth for each user as 7 bits/s and the maximum transmit power as $P_{\max} = 40$ dBm. From Fig. 5.1, we can observe

[8]Let D be the distance from the corresponding base station to the different users, let R_{SBS} be the radius of each small cell, and the loss of wall L is 10 dB. The path loss models we use are listed as: 1. from small base station to its the small cell user, $PL\_dB = 38.46 + 20\log_{10}D + 0.7D$; 2. from small base station to macro cell user, $PL\_dB = \max((15.3 + 37.6\log_{10}(D - R_{SBS})), (38.46 + 20\log_{10}(D - R_{SBS}))) + 0.7R_{SBS} + L$; 3. from macro cell base station to small base station, $PL\_dB = 15.3 + 37.6\log_{10}D$; 4. from macro cell base station to small cell users, $PL\_dB = 15.3 + 37.6\log_{10}D + 2L$.

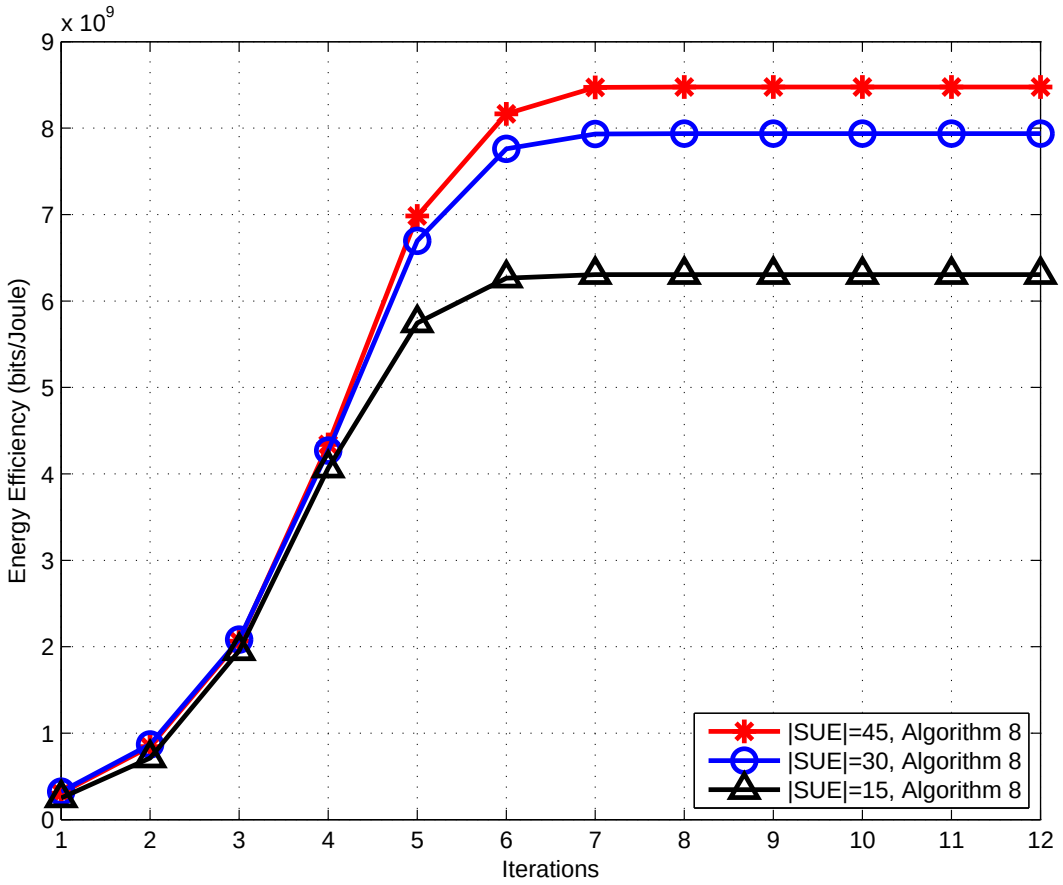

Figure 5.2: Energy efficiency versus the number of iterations for Algorithm 8.

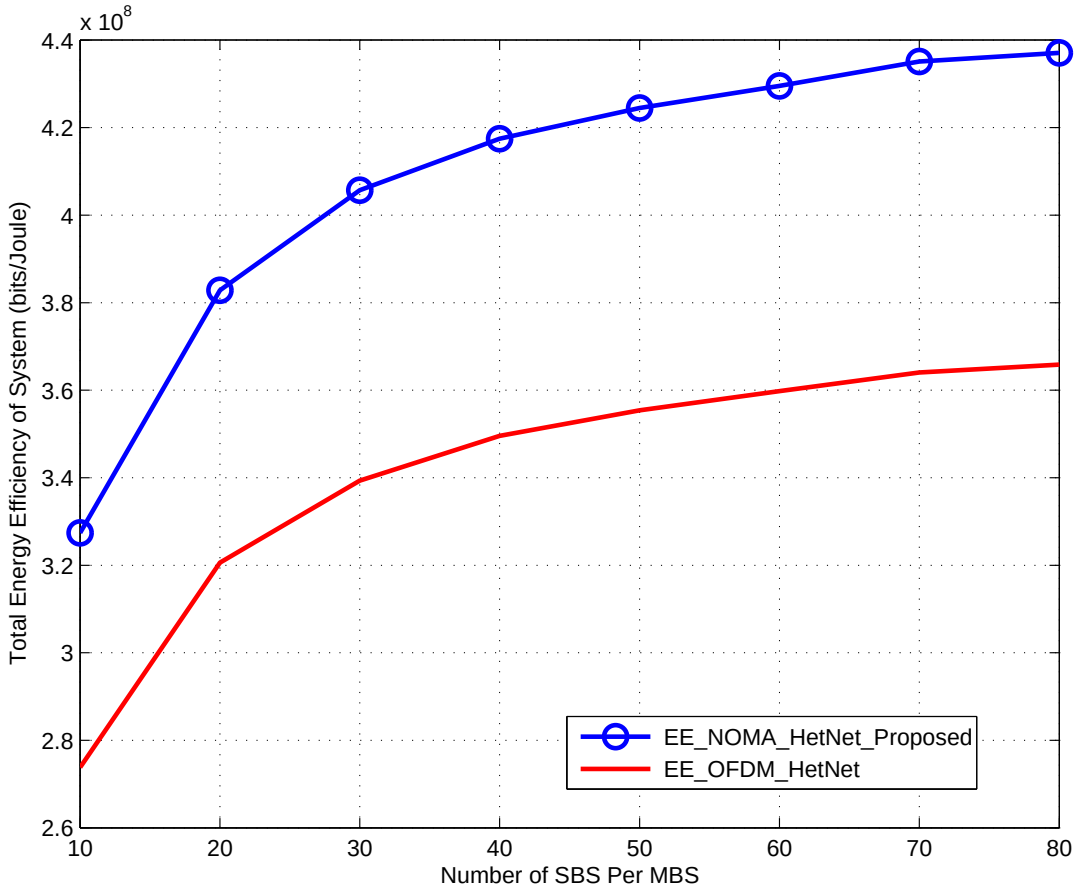

Figure 5.3: Energy efficiency versus number of small cells with perfect CSI.

that the convergence of the proposed iterative resource allocation algorithm (Algorithm 8) converges within 10 iterations, which suggests that the proposed scheme is practical. The system with 60 SUEs can achieve higher energy efficiency than the system with 45 and 15 users.

Figure 5.2 shows the performance of the system energy efficiency versus the number of the small cells. The bandwidth is limited to 10 MHz, and we set the peak power of the entire system to be 20 W and circuit power to be 0.1 W on each subchannel. Figure 5.2 indicates that the total energy efficiency increases when the number of the small cells increases. From Fig. 5.2, we can observe that the NOMA HetNet outperforms the OFDM heterogeneous network (OFDM HetNet) because OFDM HetNet cannot fully utilize the spectrum resources.

5.4 Imperfect CSI Discussion

In this subsection, we study the energy efficient resource allocation for the downlink NOMA heterogeneous network with imperfect CSI. In practice, perfect CSI at the transmitter is difficult to achieve due to channel estimation errors, feedback and quantization errors. To maximize the energy efficiency, we can formulate the energy efficient resource allocation as a probabilistic mixed non-convex optimization problem under the constraints of outage probability limitation, the maximum transmitted power and the number of multiplexed users on one subchannel.

The resource allocation scheme is optimized based on the imperfect CSI. To solve this problem, we decouple the problem into subchannel allocation and power allocation subproblems separately to maximize the system energy efficiency. We assume that the BSs have the estimated value of the CSI. In this situation, the user data rate may not meet the minimum data requirement determined by QoS. Therefore, an outage probability requirement is considered for the resource scheduling to maximize the system energy efficiency. The energy efficient optimization problem is formulated as a probabilistic mixed non-convex

optimization problem. In order to efficiently solve the optimization problem, we can first transform the probabilistic mixed non-convex optimization problem to a non-probabilistic optimization problem. The outage probability constraint can be transformed to minimum power constraints for UEs sharing the same subchannel. In this transformed problem, we can treat subchannel allocation and power allocation separately. The subchannel allocation starts with assigning equal power allocation on subchannels. The optimal solution to the user subchannel allocation subproblem is challenging to obtain in practice as it is required to search all the possible combinations. A suboptimal subchannel allocation algorithm can be obtained by using the estimated CSI. In order to cancel the interference from the other small cells, we let each small cell only occupy one subchannel. MUEs will multiplex on these subchannels, and hence the SUEs will not interfere with each other. The MUE with large channel gain will be chosen to multiplex on its corresponding subchannel. For the each subchannel, we set the maximum number of multiplexed users to reduce the complexity of decoding at the receivers. The subchannel allocation will terminate if there is no MUE left to be allocated. During the subchannel allocation, the power allocation for the multiplexed users on one subchannel can be determined by FTPA. The complexity of the proposed suboptimal algorithm is less than the optimal solution that can only be obtained by the exhaustive search.

Given the user subchannel allocation scheme, the underlying optimization problem can be shown to be a convex function respect to the power variable under the constraint of the maximum power of the system. Therefore, a unique optimal solution can be found by a gradient assisted binary search algorithm. Therefore, the new power allocation scheme can replace the equal power allocation to further improve the system energy efficiency.

Figure 5.4 shows the energy efficiency versus the number of the small cells with the channel gain estimation error variance 0.05. As shown, the energy efficiency increases when the number of the small cells grows. From Fig. 5.2, the NOMA HetNet outperforms OFDM HetNet in terms of energy efficiency.

Fig. 5.5 shows the energy efficiency of the NOMA HetNet performance with different

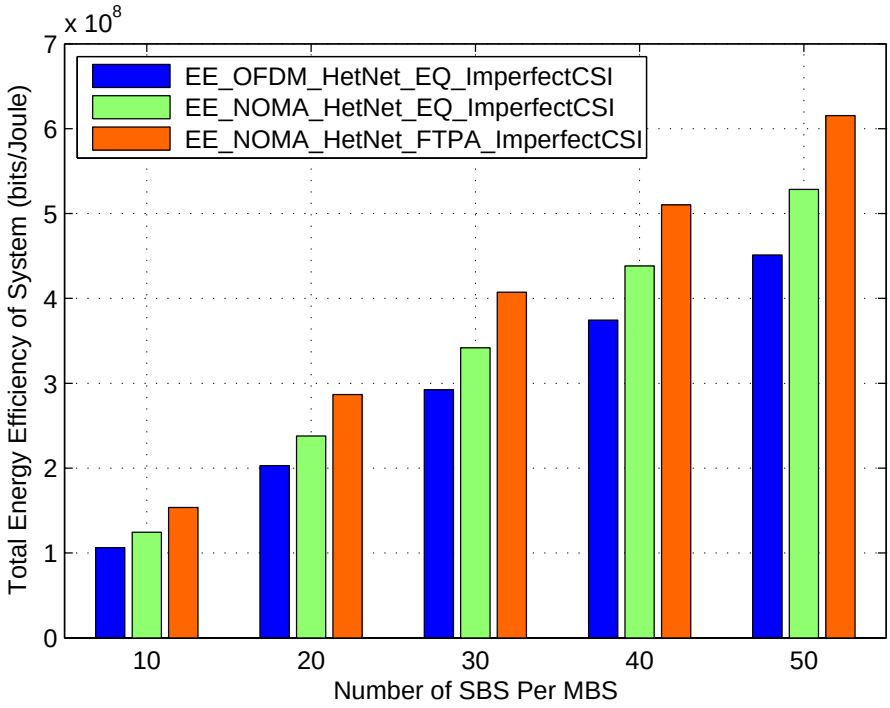

Figure 5.4: Energy efficiency of the system versus the number of the small cells with the estimation error variance 0.05.

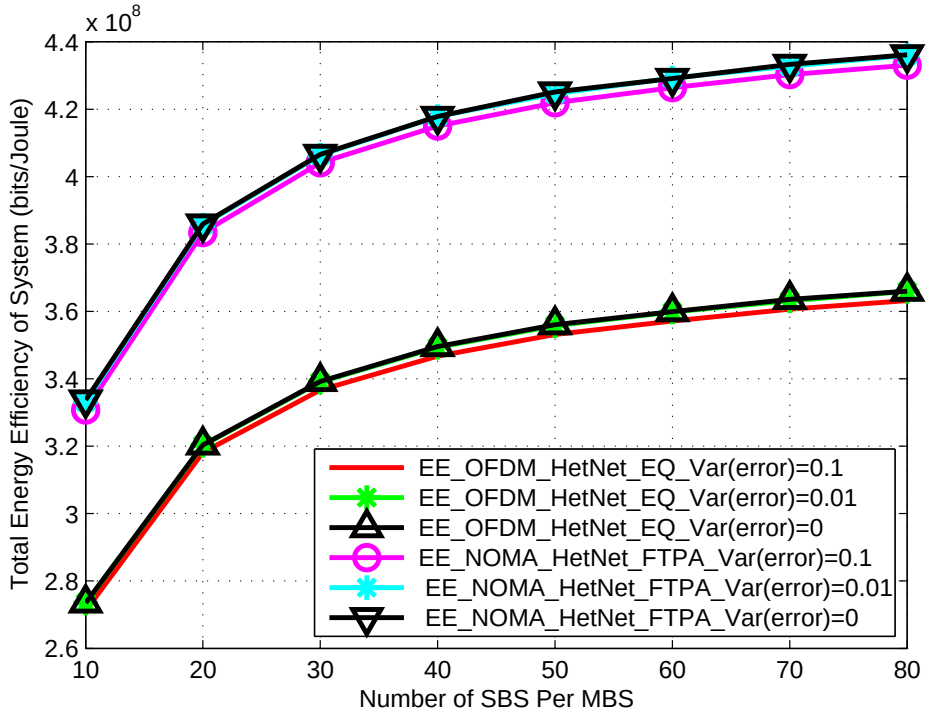

Figure 5.5: Energy efficiency of the system versus the number of the small cell with different estimation error variances.

estimation error variances. When the number of small cells is 20, the energy efficiency of NOMA HetNet with FTPA power allocation scheme is 27% more than that of the OFDM HetNet with the equal power (EQ) allocation scheme under the imperfect channel CSI.

5.5 Summary

In this chapter, we introduced the concept of NOMA HetNets and proposed an energy efficient subchannel allocation and power allocation scheme for a downlink NOMA HetNet with perfect CSI. We solved the resource optimization problem with the help of convex optimization. It is envisioned that a joint subchannel and power allocation approach can further enhance the energy efficiency of the overall system. We also discussed the energy efficiency maximization problem in the NOMA HetNet with imperfect CSI. The effectiveness of the proposed schemes was compared to the existing scheme and verified computer simulations to in terms of energy efficiency.

Chapter 6

Conclusions

In this chapter, we summarize the contributions of this thesis. Also, we present several potential future research topics that are related to our accomplished work.

6.1 Summary of Contributions

In this thesis, we investigated the energy efficient resource allocation design for downlink NOMA networks. We first studied the energy efficient resource allocation for a single cell NOMA network with perfect CSI. Then we extended the energy efficient resource allocation to the imperfect CSI case. Furthermore, the entire system energy efficiency maximization problem was studied in a NOMA HetNet. In the end, we discussed the related future work. Here we summarize the results obtained in each chapter.

Contribution 1: the energy-efficient resource allocation problem was formulated for a downlink NOMA wireless network by assigning only two users to the same subchannel. To efficiently solve this problem, we first decoupled it into subchannel allocation and power allocation subproblems. Matching theory was exploited by subchannel assignment subproblem formulation and a low-complexity suboptimal algorithm was proposed to maximize the system energy efficiency. To allocate powers across subchannels, DC programming was utilized to approximate the non-convex optimization problem into the convex subproblem. Therefore, a suboptimal power allocation across subchannels was obtained by solving the convex subproblems iteratively. The system energy efficiency was further improved by the proposed subchannel power allocation scheme. Numerical results showed that both sum rate and energy efficiency performance of the proposed resource allocation scheme outper-

formed the OFDMA system.

Contribution 2: For the NOMA system with imperfect CSI, the formulated probabilistic mixed non-convex resource optimization problem was first transformed into to a non-probabilistic problem. Based on the proposed low-complexity suboptimal user scheduling scheme, we proposed an iterative allocation algorithm for power allocation. In power allocation, we derived a closed form power allocation expression for users on each subchannel where the maximum user number can be greater than two. The performance of the proposed scheme was verified by computer simulations and compared to the existing scheme in terms of energy efficiency. It was shown that the energy efficiency of the NOMA system with our proposed resource allocation scheme is higher than the existing scheme in [57] as well as the OFDMA scheme.

Contribution 3: the single-cell NOMA system model was extended into HetNets. To maximize the entire NOMA HetNet, an energy efficient subchannel allocation and power allocation scheme was proposed for a downlink NOMA HetNet with perfect CSI. We solved the resource optimization problem with the help of convex optimization. The energy efficiency maximization was considered on both small cells and marco cell. In small cell, a joint subchannel and power allocation approach can enhance the system energy efficiency. The optimal power allocation scheme in macro cell was proposed to further improve the overall system energy efficiency. Moreover, imperfect CSI case was also discussed. Simulation results show that our proposed resource allocation scheme for NOMA HetNets can achieve higher energy efficiency than that of conventional OMA systems. Finally, the energy efficiency for NOMA HetNets with imperfect CSI was discussed, and the performance was evaluated by the simulation results.

6.2 Future Works

In Chapter 5, we presented an energy efficient resource scheme for a downlink NOMA HetNet. Resource allocation is one of the methods to improve the energy efficiency of 5G

NOMA networks. Small cell in HetNet can also be energy efficient because low power is typically required due to the short distance between transmitter and receiver. However, there are still many challenges and open issues in achieving energy-efficient 5G NOMA networks. 1) NOMA with MIMO: Not only user scheduling and power allocation, but also beamforming can be considered in the energy efficient resource allocation. Furthermore, different antenna configurations for macro cell and small cell in HetNet will make the problem more complex. Finding effective ways to combine NOMA and massive MIMO is a challenging but important topic in future research works. 2) NOMA with energy harvesting: To save more energy, 5G NOMA HetNet is expected to harvest energy from solar, wind, thermoelectric, and so on. One open problem is how to manage those renewable energy source because they are intermittent over time and space. 3) NOMA with game theory: User pairing is a key research direction in energy-efficient 5G NOMA networks. Matching game theory can be an effective tool to optimize the user pairing of energy-efficient 5G NOMA networks. 4) NOMA with open/closed access small cells: There are three access modes in small cell networks: open access, closed access, and hybrid access. Different access modes of NOMA small cells require different energy saving methods. Therefore, the energy efficient resource allocation in 5G NOMA networks can be both challenging and promising.

Bibliography

[1] A. A. Huurdeman, *The Worldwide History of Telecommunications.* New Jersey: John Wiley & Sons Press, 2003. → pages 1

[2] A. Goldsmith, *Wireless Communications.* Cambridge: Cambridge University Press, 2005. → pages 1

[3] F. A. Tobagi, "Modeling and performance analysis of multihop packet radio networks," *Proceedings of the IEEE,* pp. 135-155, Jan. 1987. → pages 1

[4] "Report ITU-R M.2134: Requirements related to technical performance for IMT-advanced radio interface(s)," ITU-R, 2008. → pages 2

[5] 3GPP TR 36.913 (V8.0.0), "Requirements for further advancements for E-UTRA (LTE-Advanced)," June 2008. → pages 2

[6] Qualcomm, "3GPP RAN Rel-12 and beyond," *Workshop on release 12*, June 2012. → pages 2

[7] "Ericsson mobility report June 2017," Ericsson, 2017. → pages xi, 2, 3

[8] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015. → pages 2

[9] X. Gao, L. Dai, S. Han, C.-L.I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE*

Journal on Selected Areas in Communications, vol. 34, no. 4, pp. 998–1009, Apr. 2016. → pages 2

[10] H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of Wi-Fi and heterogeneous small cell networks sharing unlicensed spectrum," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 158–164, Mar. 2015. → pages 2

[11] H. Zhang, C. Jiang, J. Cheng, and V. C. M. Leung, "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 92–99, June 2015.

[12] H. Zhang, C. Jiang, Q. Hu, and Y. Qian, "Self-organization in disaster resilient heterogeneous small cell networks," *IEEE Network*, vol. 30, no. 2, pp. 116–121, Mar. 2016. → pages 2

[13] J. Thompson, X. Ge, H. Wu, R. Irmer, H. Jiang, G. Fettweis, and S. Alamouti, "5G wireless communcation systems: Prospects and challenges," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 62–64, Feb. 2014. → pages 4

[14] H. Zhang, C. Jiang, X. Mao, and H. Chen, "Interference-limit resource optimization in cognitive femtocells with fairness and imperfect spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1761–1771, Mar. 2016. → pages 4

[15] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wang, and T. Quek, "Resource allocation for cognitive small cell networks: A cooperative bargaining game theoretic approach," *IEEE Transactions Wireless Communications*, vol. 14, no. 6, pp. 3481–3493, June 2015. → pages 4

[16] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2366–2377, July 2014. → pages 6, 7

[17] K. Higuchi and Y. Kishiyama, "Non-orthogonal access with random beamforming and intra-beam SIC for cellular MIMO downlink," in *Proceedings of IEEE 78th Vehicular Technology Conference (VTC2013-Fall)*, pp. 1–5, Sept. 2013. → pages 4, 75
→ pages 4, 24

[18] Y. Endo, Y. Kishiyama, and K. Higuchi, "Uplink non-orthogonal access with MMSE-SIC in the presence of inter-cell interference," in *Proceedings of IEEE International Symposium on Wireless Communication Systems (ISWCS 2012)*, Paris, France, Aug. 2012. → pages

[19] J. Umehara, Y. Kishiyama, and K. Higuchi, "Enhancing user fairness in non-orthogonal access with successive interference cancellation for cellular downlink," *Proceedings of IEEE International Conference on Complex Systems (ICCS)*, Nov. 2012. → pages

[20] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resouce allocation," *Proceedings of IEEE International Symposium on Wireless Communication Systems (ISWCS 2012)*, pp. 476–480, Aug. 2012. → pages 4, 24, 30

[21] I. Humar, X. Ge, L. Xiang, M. Jo, M. Chen, and J. Zhang, "Rethinking energy efficiency models of cellular networks with embodied energy," *IEEE Network*, vol. 25, no. 2, pp. 40-49, Mar./Apr. 2011. → pages 16

[22] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 55–62, Aug. 2011. → pages 4, 16

[23] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental tradeoffs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, June 2011. → pages

[24] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M.A. Imran, D. Sabella, M.J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run

a wireless network?" *IEEE Wireless Communications*, vol. 18, no. 5, pp. 40–49, Oct. 2011. → pages

[25] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 4, pp. 524–540, Nov. 2011. → pages 4, 16

[26] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," *Proceedings of IEEE 77th Vehicular Technology Conference (VTC2013-Spring)*, vol. 62, no. 7, pp. 1–5, Dresden, Germany, June 2013. → pages 4, 21, 23, 26, 30

[27] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Nonorthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, Sept. 2015. → pages 4

[28] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, Feb. 2017. → pages 4

[29] Z. Wei, Y. Jinhong, D. W. K. Ng, M. Elkashlan, and Z. Ding, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks," *ZTE Communications*, vol. 14, no. 4, pp. 17-25, Oct. 2016. → pages 4

[30] B. Wang, L. Dai, T. Mir, and Z. Wang, "Joint user activity and data detection based on structured compressive sensing for NOMA," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1473–1476, July 2016. → pages 4

[31] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016. → pages 4

[32] "5G radio access: Requirements, concepts and technologies," NTT DoCoMo, Inc., Tokyo, Japan, 5G White Paper, July 2014. → pages 5

[33] "5G, a technology vision," Huawei, Inc., Shenzhen, China, 5G White Paper, Mar. 2015. → pages

[34] "NGMN 5G white paper," NGMN Alliance, Frankfurt, Germany, White Paper, Feb. 2015. → pages

[35] "5G innovation opportunities- a discussion paper," techUK, London, U.K., 5G White Paper, Aug. 2015. → pages 5

[36] "Study on downlink multiuser supersition transmission (MUST) for LTE (Release 13)," 3GPP TR 36.859, Tech. Rep., Dec. 2015. → pages 5

[37] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014. → pages 5, 22

[38] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 654–667, Feb. 2016. → pages 5

[39] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-efficient resource scheduling for NOMA systems with imperfect channel state information," *Proceedings of IEEE International Conference on Communications (ICC 2017)*, Paris, France, 21–25 May 2017. → pages 5, 9

[40] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," *Proceedings of IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 611–615, Sept. 2013. → pages 5, 7, 23, 39

[41] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE enhancements," in *IEEE Globecom Workshops*, pp. 66–70, Dec. 2013. → pages 5, 7, 23

[42] M. R. Hojeij, J. Farah, C. A. Nour, and C. Douillard, "Resource allocation in downlink non-orthogonal multiple access (NOMA) for future radio access," *IEEE 81st Vehicular Technology Conference (VTC2015-Spring)*, pp. 1–6, Glasgow, U.K., May 2015. → pages 5

[43] J.-B. Kim and I.-H. Lee, "Capacity analysis of cooperative relaying systems using non-orthogonal multiple access," *IEEE Communications Letters*, vol. 19, no. 11, pp. 1949–1952, Nov. 2015. → pages 5

[44] P. Parida and S. Das, "Power allocation in OFDM based NOMA system: A DC programming approach," in *IEEE Globecom Workshops*, pp. 1026–1031, Dec. 2014. → pages 6, 66

[45] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1462–1465, June 2015. → pages 6

[46] Q. Sun, S. Han, C.-L. I, and Z. Pan, "Energy efficiency optimization for fading MIMO non-orthogonal multiple access systems," *Proceedings of IEEE International Conference on Communications (ICC 2015)*, pp. 2668–2673, London, U.K., June 2015. → pages 6

[47] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1077-1091, Mar. 2017. → pages 6

[48] Z. Ding, L. Dai, and H. V. Poor, "MIMO-NOMA design for small packet transmission in the internet of things," *IEEE Access*, vol. 4, pp. 1393–1405, Apr. 2016. → pages 6

[49] B. Wang, L. Dai, T. Mir, and Z. Wang, "Joint user activity and data detection based on structured compressive sensing for NOMA," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1473–1476, July 2016. → pages 6

[50] Y. Liu, M. Elkashlan, Z. Ding, and G. K. Karagiannidis, "Fairness of user clustering in MIMO non-orthogonal multiple access syste," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1465-1468, July 2016. → pages 6

[51] J. Zhao, Y. Liu, K. K. Chai, A. Nallanathan, Y. Chen, and Z. Han, "Spectrum allocation and power control for non-orthogonal multiple access in HetNets," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5825–5837, Sept. 2017. → pages 7

[52] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for MC-NOMA systems," *Proceedings of IEEE Globecom*, Washington, DC, USA, Dec. 2016. → pages 7

[53] Z. Wei, D. W. K. Ng, J. Yuan, and H.-M. Wang, "Optimal resource allocation for power-efficient MC-NOMA with imperfect channel state information," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3944-3961, Sept. 2017. → pages 7

[54] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203-2214, May 2006. → pages 7

[55] S. Han, S. Ahn, E. Oh, and D. Hong, "Effect of channel-estimation error on BER performance in cooperative transmission," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2083-2088, May 2009. → pages 7

[56] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy efficiency of resource scheduling for non-orthogonal multiple access (NOMA) wireless network," *Proceedings of IEEE International Conference on Communications (ICC 2016)*, Kuala Lampur, Malaysia, May 2016. → pages 9

[57] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access (NOMA) network," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3722–3732, July 2016. → pages 9, 57, 59, 66, 80, 102

[58] N. I. Miridakis and D. D. Vergados, "A survey on the successive interference cancellation performance for single-antenna and multipleantenna OFDM systems," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 312-335, Jan. 2013. → pages 11

[59] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge, U.K.: Cambridge Univ. Press, 2005. → pages 11, 12, 56

[60] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, andT. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," *Proceedings of International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) 2013*, Naha, Japan, Nov. 2013. → pages 15

[61] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple access downlink transmissions," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, Sept. 2016. → pages 16, 22

[62] T. Mao, G. Feng, L. Liang, X. Chu, S. Qin, and B. Wu, "Distributed energyefficient power control for macrofemto networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 718-731, Feb. 2016. → pages 17

[63] R. Xie, F. R. Yu, H. Ji, and Y. Li, "Energy-efficient resource allocation for heterogeneous cognitive radio networks with femtocells," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 3910-3920, Nov. 2012. → pages 17

[64] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089-1098, Aug. 2004. → pages 17

[65] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained Ad-Hoc wireless networks," *IEEE Communications Magazine*, vol. 9, no. 4, pp. 1536–1284, Aug. 2002. → pages 18

[66] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, New Yourk, USA, 2004. → pages 19, 32, 60, 62, 83, 85, 90

[67] A. Roth and M. Sotomayor, *Two-Sided Matching: A Study in Game Theoretic Modeling and Analysis*. Cambridge University Press, 1992. → pages 27

[68] S. Bayat, R. Louie, Z. Han, B. Vucetic, and Y. Li, "Distributed user association and femtocell allocation in heterogeneous wireless networks," *IEEE Transactions on Wireless Communications*, vol. 62, no. 8, pp. 3027–3043, Aug. 2014. → pages 27, 30

[69] N. Vucic, S. Shi, and M. Schubert, "DC programming approach for resource allocation in wireless networks," *Proceedings of International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, pp. 380–386, June 2010. → pages 32, 33

[70] D. P. Beretsekas, *Nonlinear Programming*. Athena Scientific, Belmont, MA, USA, 1999. → pages 32, 33

[71] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization, Philadelphia, PA, USA, 2001 → pages 32

[72] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 545–554, Feb. 2010. → pages 37

[73] D. W. K. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3292-3304, Sept. 2012. → pages 39, 49, 50, 53, 54

[74] S. Ikki and S. Aissa, "Two-way amplify-and-forward relaying with Gaussian imperfect channel estimations," *IEEE Communications Letters*, vol. 16, no. 7, pp. 956–959, July 2012. → pages 49

[75] C. Wang, T. K. Liu, and X. Dong, "Impact of channel estimation error on the performance of amplify-and-forward two-way relaying," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 3, pp. 1197–1207, Mar. 2012. → pages 49

[76] X. Wang, F.-C. Zheng, P. Zhu, and X. You, "Energy-efficient resource allocation in coordinated downlink multicell OFDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1395-1408, Mar. 2016. → pages 49, 53

[77] D. W. K. Ng and R. Schober, "Cross-layer scheduling for OFDMA amplify-and-forward relay networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1443-1458, Mar. 2010. → pages 49, 53, 54

[78] R. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks" *IEEE Transactions on Information Theory*, no. 9, pp. 4067–4086, Aug. 2009. → pages 54

[79] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009. → pages 59, 86

[80] J. Brehmer, *Utility Maximization in Nonconvex Wireless Systems.* Springer, New York, NY, 2012. → pages 62

[81] "Evolved universal terrestrial radio access: Further advancements for E-UTRA physical layer aspects," 3GPP TR 36.814, Tech. Rep., 2010. → pages 66, 93

[82] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, pp. 492–498, Mar. 1967. → pages 116

[83] V. Chandrasekhar and J. G. Andrews, "Femtocell networks: A survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, Sept. 2008. → pages 76

[84] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999. → pages 83

[85] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, July 2006. → pages 83

Appendix

Appendix A

Proof of Convergence of Algorithm 6

We will now prove the convergence of Algorithm 6 [82]. Based on (4.37) and (4.38), we first prove that $f(t)$ is a strictly monotonic decreasing function of t, i.e., if $t^{(1)} < t^{(2)}$, then $f(t^{(1)}) > f(t^{(2)})$.

Proof: Let $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ be the optimal power allocation schemes for $f(t^{(1)})$ and $f(t^{(2)})$, respectively. We have

$$\begin{aligned} f\left(t^{(2)}\right) &= \max_{\mathbf{P}\succ 0}\left\{R^*\left(\mathbf{P}\right) - t^{(2)}P_s\left(\mathbf{P}\right)\right\} \\ &= R^*\left(\mathbf{P}^{(2)}\right) - t^{(2)}P_s\left(\mathbf{P}^{(2)}\right) \\ &> R^*\left(\mathbf{P}^{(1)}\right) - t^{(2)}P_s\left(\mathbf{P}^{(1)}\right) \\ &> R^*\left(\mathbf{P}^{(1)}\right) - t^{(1)}P_s\left(\mathbf{P}^{(1)}\right) = f\left(t^{(1)}\right). \end{aligned} \tag{A.1}$$

Let $t^{(1)} = \frac{R^*\left(\mathbf{P}^{(1)}\right)}{P_s\left(\mathbf{P}^{(1)}\right)}$ and $\mathbf{P}$ be an arbitrary power allocation policy. We have $f\left(t^{(1)}\right) \geq 0$ because $f\left(t^{(1)}\right) = \max_{\mathbf{P}\succ 0}\left\{R^*\left(\mathbf{P}\right) - t^{(1)}P_s\left(\mathbf{P}\right)\right\} \geq R^*\left(\mathbf{P}^{(1)}\right) - t^{(1)}P_s\left(\mathbf{P}^{(1)}\right) = 0$. Now we are ready to prove the convergence of Algorithm 3. During each iteration, the energy efficiency t increases. We assume t^* is the maximum energy efficiency and its corresponding power allocation scheme is $\mathbf{P}^*$. Now we want to prove that the energy efficiency t is increased to the optimal t^* when $f(t^*) = 0$. First, let $\mathbf{P}^{(l)}$ be the optimal power allocation schemes in the lth iteration. We assume that $t^{(l)} \neq t^*$ and $t^{(l+1)} \neq t^*$. Based on (38), we have

$f\left(t^{(l)}\right) > 0$ and $f\left(t^{(l+1)}\right) > 0$. In Algorithm 3, $t^{(l+1)} = \frac{R^*\left(\mathbf{P}^{(l)}\right)}{P_s\left(\mathbf{P}^{(l)}\right)}$, then we have

$$\begin{aligned} f\left(t^{(l)}\right) &= R^*\left(\mathbf{P}^{(l)}\right) - t^{(l)} P_s\left(\mathbf{P}^{(l)}\right) \\ &= t^{(l+1)} P_s\left(\mathbf{P}^{(l)}\right) - t^{(l)} P_s\left(\mathbf{P}^{(l)}\right) \\ &= P_s\left(\mathbf{P}^{(l)}\right)\left(t^{(l+1)} - t^{(l)}\right) > 0. \end{aligned} \tag{A.2}$$

Since $P_s\left(\mathbf{P}^{(l)}\right) > 0$, we have $t^{(l+1)} - t^{(l)}$. When the number of iterations is sufficiently large, we will have $f\left(t^{(l)}\right) = 0$ and find t^*.

Appendix B

Derivation of the Optimal Power Allocation Policy in Chapter 4

For each subchannel, we assume that channel estimation has error and the users are assorted according to the estimated channel gains $|\hat{h}_{1,n}| \leq \cdots \leq |\hat{h}_{m,n}| \leq \cdots \leq |\hat{h}_{M_n,n}|$.

When $m = 1$, decode the first user and let

$$\frac{\partial L\left(P,\lambda\right)}{\partial p_{1,n}} = \frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{1,n}\right)\alpha_{1,n}\frac{1}{\ln 2}\frac{1}{p_{1,n}} - t - \lambda = 0. \tag{B.1}$$

$$p_{1,n} = \frac{\left(\frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{1,n}\right)\right)\alpha_{1,n}}{\ln 2\left(t+\lambda\right)}. \tag{B.2}$$

When $m = 2$, decode the second user and let

$$\begin{aligned}
\frac{\partial L\left(\mathbf{P},\lambda\right)}{\partial p_{2,n}} &= \frac{B}{N}\left(1-\varepsilon_{out}\right)\alpha_{2,n}\frac{1}{\ln 2}\frac{1}{p_{2,n}} \\
&+ \frac{B}{N}\left(1-\varepsilon_{out}\right)\left(\alpha_{1,n}\frac{1}{\ln 2}\tilde{\Phi}_{1,n}\frac{2\left(|\hat{g}_{1,n}|^2+\sigma_e^2\right)D_{1,n}^2}{p_{1,n}\varepsilon_{out}F^{-1}_{|g_{1,n}|^2}\left(\varepsilon_{out}/2\right)}\right) \\
&+ \frac{B}{N}\upsilon_{2,n}\alpha_{2,n}\frac{1}{\ln 2}\frac{1}{p_{2,n}} \\
&+ \frac{B}{N}\upsilon_{2,n}\left(\alpha_{1,n}\frac{1}{\ln 2}\tilde{\Phi}_{1,n}\frac{2\left(|\hat{g}_{1,n}|^2+\sigma_e^2\right)D_{1,n}^2}{p_{1,n}\varepsilon_{out}F^{-1}_{|g_{1,n}|^2}\left(\varepsilon_{out}/2\right)}\right) \\
&- t - \lambda = 0.
\end{aligned} \tag{B.3}$$

Then we can have

$$p_{2,n} = \frac{\alpha_{2,n}\frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{2,n}\right)}{\ln 2\left(t+\lambda\right)+A(p_{1,n})} \tag{B.4}$$

where

$$A\left(p_{1,n}\right) = \frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{1,n}\right)\alpha_{1,n} \left(\frac{2\left(\left|\hat{g}_{1,n}\right|^2+\sigma_e^2\right)D_{1,n}^2\tilde{\Phi}_{1,n}}{p_{1,n}\varepsilon_{out}F_{\left|g_{1,n}\right|^2}^{-1}\left(\varepsilon_{out}/2\right)}\right) \tag{B.5}$$

and

$$\tilde{\Phi}_{1,n} = \frac{\varepsilon_{out}F_{\left|g_{1,n}\right|^2}^{-1}\left(\varepsilon_{out}/2\right)\cdot D_{1,n}^2 p_{1,n}}{\varepsilon_{out}\sigma_z^2+2D_{1,n}^2\left(\left|\hat{g}_{1,n}\right|^2+\sigma_e^2\right)\sum\limits_{i=2}^{M_n}p_{i,n}}. \tag{B.6}$$

When $m=3$, decode the second user and let

$$\begin{aligned}\frac{\partial L\left(\mathbf{P},\lambda,\boldsymbol{\upsilon}\right)}{\partial p_{3,n}} &= \frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{3,n}\right)\alpha_{3,n}\frac{1}{\ln 2}\frac{1}{p_{3,n}} \\ &+\frac{\partial}{\partial p_{3,n}}\left(\frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{2,n}\right)\alpha_{2,n}\log_2\left(\tilde{\Phi}_{2,n}\right)\right) \\ &+\frac{\partial}{\partial p_{3,n}}\left(\frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{1,n}\right)\alpha_{1,n}\log_2\left(\tilde{\Phi}_{1,n}\right)\right) \\ &-t-\lambda=0.\end{aligned} \tag{B.7}$$

We can obtain

$$p_{3,n} = \frac{\alpha_{3,n}\left(\frac{B}{N}\left(1-\varepsilon_{out}+\upsilon_{3,n}\right)\right)}{\ln 2\left(t+\lambda\right)+A\left(p_{1,n}\right)+A\left(p_{2,n}\right)}. \tag{B.8}$$

Therefore, by deduction, the power of the mth users on SC_n can be derived as (4.43).

Appendix C

Derivation of the Optimal Power Allocation $\{p_{u,n}^{s,*}\}$ for SUEs in Chapter 5

For any subchannel n, we assume that the users are assorted according to the channel gains $|g_{1,n}^s| \leqslant |g_{2,n}^s| \leqslant \cdots \leqslant |g_{U,n}^s|$.

To obtain $\tilde{p}_{1,n}^s$, we let

$$\begin{aligned}\frac{\partial L_{s,n}(\cdots)}{\partial \tilde{p}_{1,n}^s} =& \alpha_{s,n} B_n a_{1,n}^s \left(1+\nu_{1,n}^s\right) \frac{1}{\ln 2} \frac{1}{\tilde{p}_{1,n}^s} - t - \lambda^s - \mu_n^s \left|f_{s,m}^{SM}\right|^2 \\ =& 0.\end{aligned} \tag{C.1}$$

Then we have

$$p_{1,n}^{s,*} = \frac{\tilde{p}_{1,n}^{s,*}}{\alpha_{s,n}} = \frac{B_n a_{1,n}^s \left(1+\nu_{1,n}^s\right)}{\ln 2 \left(t + \lambda^s + \mu_n^s \left|f_{s,m}^{SM}\right|^2\right)}. \tag{C.2}$$

To obtain $p_{2,n}^{s,*}$, we let

$$\begin{aligned}&\frac{\partial L_{s,n}(\cdots)}{\partial \tilde{p}_{2,n}^s} \\ =& \alpha_{s,n} B_n a_{1,n}^s \left(1+\nu_{1,n}^s\right) \frac{1}{\ln 2} \left(-\frac{\gamma_{1,n}^s}{p_{1,n}^s}\right) \\ &+ \alpha_{s,n} B_n a_{2,n}^s \left(1+\nu_{2,n}^s\right) \frac{1}{\ln 2} \left(\frac{1}{\tilde{p}_{2,n}^s}\right) - t - \lambda^s - \mu_n \left|f_{s,m}^{SM}\right|^2 \\ =& 0.\end{aligned} \tag{C.3}$$

Then we can have

$$p_{2,n}^{s,*} = \frac{\tilde{p}_{2,n}^{s,*}}{\alpha_{s,n}} = \frac{B_n a_{2,n}^s \left(1+\nu_{2,n}^s\right)}{B_n a_{1,n}^s \left(1+\nu_{1,n}^s\right)\left(\frac{\gamma_{1,n}^s}{p_{1,n}^{s,*}}\right) + \ln 2\left(t+\lambda^s+\mu_n^s\left|f_{s,m}^{SM}\right|^2\right)} \tag{C.4}$$

where

$$\gamma_{1,n}^s = \frac{|g_{1,n}^s|^2 p_{1,n}^{s,*}}{|g_{1,n}^s|^2 \sum\limits_{l=u+1}^{U} p_{l,n}^{s,*} + I_n^M + \sigma_z^2}. \tag{C.5}$$

Similarly, to obtain $p_{3,n}^{s,*}$, we let

$$\begin{aligned}
&\frac{\partial L_{s,n}(\cdots)}{\partial \tilde{p}_{3,n}^s} \\
=&\alpha_{s,n} B_n a_{1,n}^s \left(1+\nu_{1,n}^s\right) \frac{1}{\ln 2}\left(-\frac{\gamma_{1,n}^s}{p_{3,n}^s}\right) \\
&+ \alpha_{s,n} B_n a_{2,n}^s \left(1+\nu_{2,n}^s\right) \frac{1}{\ln 2}\left(-\frac{\gamma_{2,n}^s}{\tilde{p}_{3,n}^s}\right) \\
&+ \alpha_{s,n} B_n a_{3,n}^s \left(1+\nu_{3,n}^s\right) \frac{1}{\ln 2}\left(\frac{1}{\tilde{p}_{3,n}^s}\right) - t - \lambda^s - \mu_n^s\left|f_{s,m}^{SM}\right|^2 \\
=&0.
\end{aligned} \tag{C.6}$$

We can obtain

$$p_{3,n}^{s,*} = \frac{\tilde{p}_{3,n}^{s,*}}{\alpha_{s,n}} = \frac{B_n a_{3,n}^s \left(1+\nu_{3,n}^s\right)}{B_n a_{1,n}^s \left(1+\nu_{1,n}^s\right)\left(\frac{\gamma_{1,n}^s}{p_{1,n}^{s,*}}\right) + B_n a_{2,n}^s \left(1+\nu_{2,n}^s\right)\left(\frac{\gamma_{2,n}^s}{p_{2,n}^{s,*}}\right) + C} \tag{C.7}$$

where

$$C = \ln 2\left(t+\lambda^s+\mu_n^s\left|f_{s,m}^{SM}\right|^2\right). \tag{C.8}$$

Therefore, by deduction, the optimal power allocated on SUE u in small cell s on subchannel n can be derived as (5.35).