Resource Allocation Optimization for Future Wireless Communication Systems

Binbin Su

School of Computing and Communications

Lancaster University

A thesis submitted for the degree of

Doctor of Philosophy

March 2020

To my family

Acknowledgements

In the first place, I would like to express my deepest appreciation and sincere thanks to my supervisor, Prof. Qiang Ni, for his continuous support, patient guidance, and kind assistance, throughout the whole process of my Ph.D. study. He impressed me a lot with the hard-working attitude, rigorous research spirit, and strong technical expertise, which will have a significant and long-term impact on my future research career. I could not imagine having a better supervisor for the Ph.D. journey.

I would like to thank Dr. Wenjuan Yu and Dr. Haris Pervaiz, who always provided sound technical guidance and helpful suggestions on my research works and directions. Besides, I would like to give my sincere thanks to all academic and administrative staff in the School of Computing and Communications (SCC), who offered regular interesting seminars, professional support, and valuable suggestions. Working with every one of you in SCC has been my great pleasure and will remain forever in my memory.

Last but not the least, my deepest gratitude goes to my deeply loved parents, who gave birth to me and always unconditionally supported every decision I made, especially pursuing the Ph.D. degree abroad. I also want to thank my wife Miss. Yu Huang, my newborn lovely son Xingcheng, and all my friends, for supporting and encouraging me all the time.

Abstract

To meet the ever-increasing requirements of high data rate, extremely low latency, and ubiquitous connectivity for the fifth generation (5G) and beyond 5G (B5G) wireless communications, there is imperious demands for advanced communication system design. Particularly, efficient resource allocation is regarded as the fundamental challenge whereas an effective way to improve system performance. The term "resource" refers to scare quantities such as limited bandwidth, power and time in wireless communications. Moreover, the development of wireless communication systems is accompanied by the innovation of applied technologies. Motivated by the above observations, efficient resource allocation strategies for several promising 5G and B5G technologies in terms of non-orthogonal multiple access (NOMA), mobile edge computing (MEC) and Long Range (LoRa) are addressed and investigated in this thesis.

Firstly, the strong user's data rate maximization problem for simultaneous wireless information and power transfer (SWIPT)-enabled cooperative NOMA system, considering the presence of channnel uncertainties, is proposed and investigated. Two major channel uncertainty design criteria in terms of the outage-based constraint design and the worst-case based optimization are adopted. In addition to the high-complexity optimal two-dimensional exhaustive search method, the low-complexity suboptimal solution is further proposed. The advantages of SWIPT-enabled cooperation in robust NOMA are confirmed with simulations.

Secondly, considering the application of NOMA and user cooperation (UC) in a wireless powered MEC under the non-linear energy harvesting model, a computation efficiency maximization problem subject to the quality of service (QoS) and power budget constraint, is studied and analyzed. The formulated problem is nonconvex, which is challenging to solve. The semidefinite relaxation (SDR) approach is first applied, then the sequential convex approximation (SCA)-based solution is further proposed to maximize the system computation efficiency.

Finally, taking into consideration the aspect of energy efficiency (EE), this thesis investigates the energy efficient resource allocation in LoRa networks to maximize the system EE (SEE) and the minimal EE (MEE) of LoRa users, respectively. The energy efficient resource allocation is formulated as NP-hard problems. A low-complexity user scheduling scheme based on matching theory is proposed to allocate users to channels, then the heuristic SF assignment solution is designed for LoRa users scheduled on the same channel. The optimal power allocation strategy is further proposed to maximize the corresponding EE.

List of Publications

Journal Papers

1. **B. Su**, Q. Ni, and W. Yu, "Robust Transmit Beamforming for SWIPT-Enabled Cooperative NOMA with Channel Uncertainties,", IEEE Transactions on Communications, vol. 67, no. 6, pp. 4381-4392, June 2019.
2. **B. Su**, Z. Qin, and Q. Ni, "Energy Efficient Uplink Transmissions in LoRa Networks", IEEE Transactions on Communications, vol. 68, no. 8, pp. 4960-4972, May 2020.
3. **B. Su**, Q. Ni, W. Yu, and H. Pervaiz, "Computation Efficiency Maximization for Wireless Powered NOMA-Assisted Mobile Edge Computing with User Cooperation", submitted to IEEE Transactions on Communications, 2020.
4. K. Anam, Ali. Syed, **B. Su**, H. Pervaiz, and Q. Ni, "Minimizing the Transaction Time Difference for NOMA-Based Mobile Edge Computing", IEEE Communication Letters, vol. 24, no. 4, pp. 853-857, Jan. 2020.

Conference Papers

1. **B. Su**, Q. Ni, and B. He, "Robust Transmit Designs for Secrecy Rate Constrained MISO NOMA System", in IEEE 29th PIMRC, Bologna, Italy, Sep. 2018, pp. 1-5.
2. **B. Su**, Z. Qin, and Q. Ni, "Energy Efficient Resource Allocation for Uplink LoRa Networks", in Proc. IEEE GLOBECOM, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1-7.

3. **B. Su**, Q. Ni, W. Yu, and H. Pervaiz, "Outage Constrained Robust Beamforming Design for SWIPT-Enabled Cooperative NOMA System", in IEEE Int. Conf. Commun. (ICC), Shanghai, China, May 2019, pp. 1-6.

4. N. Gao, X. Jing, Q. Ni, and **B. Su**, "Active Spoofing Attack Detection: An Eigenvalue Distribution and Forecasting Approach", in IEEE 29th PIMRC, Bologna, Italy, Sep. 2018, pp. 1-6.

Contents

List of Tables

List of Figures

List of Acronyms

1G	First generation
2ES	Two-sided exchange stable
2G	Second generation
3G	Third generation
3GPP	Third generation partnership project
4G	Fourth generation
5G	Fifth generation
B5G	Beyond 5G
ADR	Adaptive data rate
AF	Amplify and forward
AGM	Arithmetic geometric mean
AP	Access point
AWGN	Additive white Gaussian noise
BS	Base station
CDMA	Code division multiple access
CSI	Channel sate information
CSS	Chirp spread spectrum
DC	Difference of convex
DF	Decode and forward
EE	Energy efficiency
eMBB	Enhanced mobile broadband

ETSI	European Telecommunications Standards Institute
FDMA	Frequency division multiple access
ICT	Information and communication technology
ID	Information decoding
IoT	Internet of Things
IR	Infrared radiation
KKT	Karush-Kuhn-Tucker
LMI	Linear matrix inequality
LoRa	Long Range
LPWA	Low-power wide-area
LTE	Long-term evolution
MA	Multiple access
MEC	Mobile edge computing
MEE	Max-min energy efficiency
MIMO	Multi-input muti-output
MISO	Multi-input single-output
mMTC	Massive machine-type communications
MRC	Maximal ratio combination
NOMA	Non-orthogonal multiple access
OFDM	Orthogonal frequency division multiplexing
OFDMA	Orthogonal frequency division multiple access
OMA	Orthogonal multiple access
PS	Power splitting
QoS	Quality of service
RAN	Radio access network
RF	Radio frequency
SC	Superposition coding

SCA	Successive convex approximation
SDR	Semidefinite relaxation
SE	Spectral efficiency
SEE	System energy efficiency
SF	Spreading factor
SIC	Successive interference cancellation
SINR	Signal to interference plus noise ratio
SISO	Single-input single-output
SNR	Signal to noise ratio
SWIPT	Simultaneous wireless information and power transfer
TDMA	Time division multiple access
UC	User cooperation
URLLC	Ultra-reliable low-latency communications
WPCNs	Wireless powered communication networks
WPT	Wireless power transfer

List of Symbols and Mathematical Operators

Symbols

B	Channel bandwidth
β	Power splitting ratio
ϑ	Energy harvesting efficiency
$\boldsymbol{w}$	Beamforming vector
$\boldsymbol{W}$	Semidefinite positive matrix
P_{c}^{k}	Circuit power of the k-th user
$\boldsymbol{S}$	User clustering indicator matrix
S_m	The number of allocated LoRa users for the m-th channel
B_m	Bandwidth allocated to the m-th channel
$\boldsymbol{P}$	Power allocation coefficients
$\boldsymbol{\ell}$	Computation bits allocation coefficients
$\boldsymbol{t}$	Time allocation coefficients
η	System energy efficiency
$\tilde{\eta}$	System computation efficiency
$\eta_{m,k}$	Individual energy efficiency of the k-th user over channel m

Mathematical Operators

Pr(A)	Probability of an event A
$\exp, e$	Exponential function

$\mathbb{E}[\cdot]$	Expectation operation
$\mathrm{Tr}(\cdot)$	Trace function
I	Identity matrix
$\lvert\cdot\rvert$	Absolute value
$\lvert\lvert\cdot\rvert\rvert$	Norm function
$\mathcal{CN}(0, \sigma^2)$	Circularly symmetric complex Gaussian distribution with mean 0 and variance σ^2
argmax	Domain points at which the function values are maximized
s.t.	Subject to
$\ln(\cdot)$	Natural logarithm
$\log_i$	Logarithm base i
$\sum$	Summation operation
$\prod$	Product operation
max	Maximum function
min	Minimum function
$\in$	Set membership (Belongs to a set)
$\forall$	Universal quantifier (For all)
$\exists$	Existential quantifier (There exists)
$\cap$	Intersection
$\frac{\mathrm{d}f(x)}{\mathrm{d}x}, f'(x)$	First derivative of function f with respect to x
$\frac{\mathrm{d}^2f(x)}{\mathrm{d}x^2}, f''(x)$	Second derivative of function f with respect to x
$\frac{\partial f(x,y)}{\partial x}$	Partial derivative of function f with respect to x
$[x]^+$	Maximum between zero and some real number x
$(\cdot)^T$	Transpose of a matrix

Chapter 1

Introduction

1.1 Background and Motivation

The concept of wireless communication can date back to the 19th century when radio transmission was first introduced by Marconi, and wireless communication systems have made great progress since then. Wireless communication system refers to the system that enables information transmission from devices to devices through the air by utilizing electromagnetic waves such as radio frequency (RF), infrared radiation (IR) and satellite signal in the air. As a consequence, wireless communication system can be divided into the variety of mobile communication systems, infrared wireless communications, broadcast radio and satellite communication systems. Due to the increasing popularity of computers, mobile phones and tablets, mobile communication system becomes the key to establish the connections between mobile devices and transmitters like access points (AP) and base stations (BS). Therefore, mobile communication system is the main focus of this thesis.

During the past few decades, mobile communication system has evolved from the first generation (1G) to the fourth generation (4G), and the fifth generation (5G) and B5G communication network is on its way. Compared to 4G, with the goal of achieving at least 1,000-fold capacity increase, reducing energy consumption on the order of several magnitudes, and improving spectral efficiency (SE) by 10 times for 5G and beyond 5G (B5G) networks [2], it remains a huge challenge to realize efficient communication systems design. Towards these directions, a variety of methods such as efficient resource allocation strategies, better data compression algorithms and improved channel coding schemes can be exploited. Due to the complication and impossibility to address all the methods, in this thesis, we concentrate on the investigation of resource allocation optimization. Typically, resources represent bandwidth,

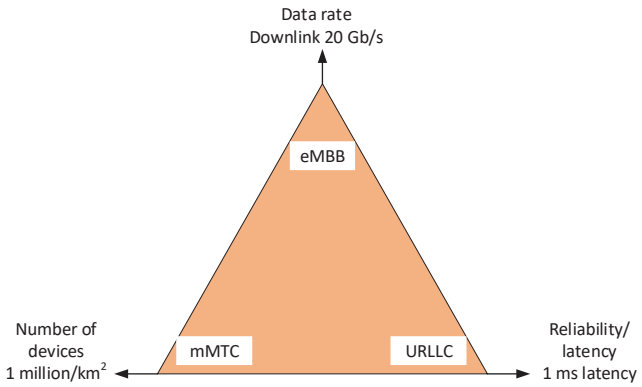


Figure 1.1: 5G and B5G services [1].

power, frequency, and time in wireless communication systems. In a wireless communication system, the amount of information that can be received at the receiver is constrained by the available resources as well as the resource allocation strategies implemented at the transmitter. Different resource allocation strategies result in diverse system performance. Efficient resource allocation strategy aims to allocate the limited resources to receivers in an effective way, which can help make the most of the scarce resources to achieve the best system performance. Henceforth, it is of vital significance to design efficient resource allocation schemes.

Furthermore, resource allocation needs to cope with the novel emerging communication technologies to better satisfy the corresponding requirements of 5G and B5G services. As can be seen from Fig. 1.1, 5G and B5G communication is categorized into three services, i.e., enhanced mobile broadband (eMBB) to provide high data rates (downlink 20 Gbit/s, uplink 10 Gbit/s), ultra-reliable low-latency communications (URLLC) to enable ultra-reliable and delay-critical tasks (1 ms end-to-end latency), and massive machine-type communications (mMTC) to accommodate massive number of devices for scenarios like Internet of Things (IoT) (connection density of 1 million devices per km^2) [3]. However, it is impossible to provide all the services at the same time with one technique, which motivates us to investigate non-orthogonal multiple access (NOMA) [4,5], mobile edge computing (MEC) [6,7], and Long Range (LoRa) [8,9]. In this thesis, we aim to optimize the corresponding resource allocation, and propose the optimal or suboptimal solutions to improve system performance for NOMA, MEC, and LoRa networks.

1.2 Related Works

The related works regarding resource allocation management for the technologies aforementioned are discussed in this section.

1.2.1 Non-orthogonal Multiple Access

Note that the appearance of the new communication system is accompanied by multiple access (MA) technique innovation. It is well known that frequency division MA (FDMA) for 1G, time division MA (TDMA) for the second generation (2G), code division MA (CDMA) for the third generation (3G), and orthogonal frequency division MA (OFDMA) for 4G can all be categorized into orthogonal MA (OMA) schemes. Recently, power-domain NOMA [1] has been demonstrated to possess the potential to significantly improve SE and accommodate massive connections [10–12]. Though NOMA has not been accepted for 5G, it is under consideration for B5G. NOMA has been shown to be more beneficial than conventional OMA schemes in many aspects [13, 14]. Thus far, extensive works have been conducted to address the resource allocation management for both downlink and uplink NOMA transmission scenarios. In regard to downlink NOMA transmission, Cui et al. [15] investigated the power allocation scheme to address the concerns of power consumption and user fairness in the NOMA system. An optimal power allocation algorithm was proposed in [16] to maximize the energy efficiency (EE) of a NOMA system. Considering the uplink NOMA transmission, a joint user clustering and power allocation strategy was proposed in [17] to maximize the sum throughput of the uplink NOMA system. A two-step resource allocation optimization strategy, which includes separate channel assignment and power allocation, has been proposed in [18] to maximize the sum rate for uplink NOMA transmission. In a downlink NOMA transmission, by utilizing the superposition coding (SC) technique, the BS sends the superimposed information containing all users' messages, then the users with strong channel conditions can obtain the prior information of the weak users[2], after applying successive interference cancellation (SIC) to remove the co-channel interference. The obtained prior information can be fully exploited with a cooperative relay transmission scheme, to improve the weak user's reception reliability [4]. The cooperative scheme can be designed based on amplify and forward (AF) relay protocol and decode and forward (DF) relay protocol. Regarding DF relaying, Liu et al. [19] applied simultaneous wireless information and

[1] Power-domain NOMA is simplified to NOMA in the rest of this thesis.
[2] Here, the weak user means the user that is far from the BS.

power transfer (SWIPT) to the cooperative NOMA system and proved the advantages of cooperative SWIPT NOMA from the perspective of outage probability and system throughput. In [20], considering a cooperative multiple-input single-output (MISO) SWIPT NOMA scheme, the power splitting (PS) ratio and the beamforming vectors were optimized to maximize the data rate of the strong user[3] while satisfying the quality of service (QoS) requirements of the weak user. For AF relaying, the authors in [21] investigated the outage probability for multiple-antenna relaying NOMA networks and demonstrated the advantages of NOMA over OMA. Take both AF and DF protocols into account, the system performance for NOMA-based user cooperation with SWIPT was characterized by pairwise error probability in [22] to show the superiority of NOMA.

1.2.1.1 Wireless Power Transfer Assisted NOMA Transmission

Note that energy efficient communications have drawn tremendous attention due to the fact that the ever-increasing energy consumption of the information and communication technologies (ICT) contributes more and more to the greenhouse gas emissions [23]. Therefore, EE becomes a key concern for 5G and B5G wireless communications [24]. To provide energy efficient communications, overcome the insufficient power supply and prolong the sustainable operation for mobile users, wireless power transfer (WPT) has emerged as an effective solution via energizing mobile devices remotely [25]. Specifically, WPT is used to charge the battery of energy harvesting devices by adopting the dedicated radio frequency (RF) energy transmitters. Wireless powered communication networks (WPCNs) [26] and SWIPT [27] are the main WPT applications to achieve sustainable communications. Extensive researches have been carried out to integrate WPCNs with NOMA. For instance, by considering two types of decoding orders, an efficient greedy algorithm was proposed in [28] to maximize the minimal rate for a wireless powered NOMA system. The authors in [29] investigated resource allocation optimization to maximize the rate region of the wireless powered NOMA communication. The optimal time switching and PS strategies were studied to maximize the achievable rate regions of the wireless powered NOMA systems in [30].

Moreover, SWIPT has drawn remarkable attention to realize more energy efficient communications [31]. Specifically, the application of SWIPT to NOMA has been studied by assuming that NOMA users can harvest energy and acquire information from the received RF signals at the same time [32]. For example, the SE performance

[3]Here, the strong user means the user that is near the BS.

comparison between NOMA and OMA for a two-user SWIPT system has been addressed in [33]. The cooperative SWIPT NOMA protocol was investigated in [5], in which near NOMA users act as energy harvesting relays to help far NOMA users without draining their batteries.

1.2.1.2 NOMA with Imperfect CSI

Channel state information (CSI) errors are universal present in wireless communication systems, and perfect CSI is quite difficult to obtain due to channel estimation errors, feedback delay and quantization errors [34, 35]. Therefore, it is more practical to consider imperfect CSI scenarios. To address CSI errors, various channel uncertainty models can be found in the existing literature. A common one is the worst-case SINR constrained problem [36], in which the CSI errors are assumed to lie in a bounded uncertainty set. The other is the outage-based constrained formulation [37] where the outage probability of the signal to interference plus noise ratio (SINR) must be less than a given value. Regarding worst-case robust model, by considering a more practical scenario that the BS only knows imperfect CSI, a robust beamforming design problem for MISO NOMA systems was investigated in [38] to maximize the achievable sum rate subject to the transmit power constraint. In [39], the beamformers were designed for a robust power minimization problem by incorporating the norm-bounded channel uncertainties to satisfy the required QoS at each user. In addition, to tackle the EE maximization problem, the robust beamforming design was proposed in [40] for a massive multiple-input multiple-output (MIMO) NOMA downlink system with imperfect CSI considered. For the outage-based model, the optimal power allocation strategy was studied in [41] to maximize system utility for the outage constrained MIMO-NOMA system.

1.2.2 Mobile Edge Computing

The boosting computation-intensive applications prevalent in the IoT networks as well as the growing number of mission-critical tasks in future-generation networks pose significant challenges in real-time communication system design [42]. To address the requirements of the increasing demand for massive computing and overcome the resource limitations (i.e., small size, low power, and limited computing capability) of mobile devices, MEC has been proposed as a promising solution to enhance mobile users computation capability and realize low-latency communications [43].

Different from conventional cloud computing, where cloud server is deployed far from mobile devices leading to high transmission cost and long latency, the cloud-like

server is integrated with the AP at the edge of MEC networks [44]. The leverage of MEC enables the resource-limited mobile users to offload tasks for remote execution at the more powerful MEC server in their proximity, which brings the benefit of improved computation capability and reduced latency. Besides, WPT is regarded as the potential solution to provide sustainable power supply for battery-limited devices in MEC networks, and NOMA can improve the connections between mobile devices and MEC servers. Recent studies show that MEC performance can be further enhanced by incorporating WPT and NOMA. Therefore, wireless powered MEC and MEC with NOMA will be introduced in the next two sections.

1.2.2.1 Wireless Powered MEC

To avoid consuming energy of the limited batteries for mobile devices, the application of WPT into MEC networks has drawn considerable attention. A new paradigm called wireless powered MEC is introduced to fuse MEC and WPT techniques, in which mobile devices can realize bits computing with the energy harvested from the RF signals and thus tackle the limitation of finite battery lifetime [45]. For example, the joint computation offloading and computing resource allocation has been investigated in [46] to minimize the system energy consumption for wireless powered multi-user MEC system. The authors in [47] maximized the sum computation rate for wireless powered MEC under binary offloading by jointly optimizing the computing mode selection and transmission time allocation. A wireless powered cooperative MEC system has been presented in [48] to maximize the computation rate, where nearby devices are exploited as MEC servers.

1.2.2.2 MEC with NOMA

The integration of NOMA and MEC is envisioned to significantly improve computation performance. NOMA-MEC design criterion is based on two modes, namely hybrid NOMA-MEC and pure NOMA-MEC [4], where multiple users can offload data simultaneously for NOMA-MEC within the same block slot duration, while for hybrid NOMA-MEC, a user can first offload part of the tasks by occupying the time allocated to another user and then offload the rest of tasks by using the extra time slot [49]. For hybrid NOMA-MEC design, Ding et al. [49] derived the closed-form solutions of the power and time allocation for NOMA-MEC and showed that the performance of MEC offloading with hybrid NOMA is superior to that of OMA and

[4]Pure NOMA-MEC is simplified as NOMA-MEC in the following sections.

NOMA scheme. In [50], a NOMA-MEC offloading taking into account the hybrid NOMA was considered and the latency was minimized by utilizing the reinforcement learning approach.

In regard to MEC offloading with NOMA, the authors in [51] minimized the overall delay of mobile users for the NOMA-assisted MEC system. By decomposing the formulated problem into sub-problems as computation offloading and time allocation, Wu et al. [51] minimized the overall delay for NOMA-assisted MEC system. It also showed that NOMA-assisted MEC outperforms conventional OMA-assisted MEC. Considering the multi-antenna NOMA-assisted MEC under both partial and binary offloading, where the BS was equipped with multiple antennas while the users were equipped with single antenna, Wang et al. [52] minimized the weighted system energy consumption.

1.2.3 Long Range (LoRa) Communication

Driven by the massive connectivity, low data rate, and low power consumption requirements in IoT networks, low-power wide-area (LPWA) networks have emerged as a potential solution to enable long-distance power efficient wireless communications [53, 54]. Compared with traditional technologies prevalent in IoT networks, such as Bluetooth, Wi-Fi, and Long-Term Evolution (LTE), LPWA techniques achieve better tradeoffs of coverage range, data rates, and power consumption. Among all the emerging LPWA technologies, LoRa [55, 56], which operates in the unlicensed bands, has attracted extensive attention.

LoRa network is composed of LoRa users, LoRa gateways, and the network server. It adopts typical star topology, in which the data and/or requests of LoRa end devices are collected by the LoRa gateway and then it forwards them to the LoRa server [9]. The core of LoRa lies in the adopted chirp spread spectrum (CSS) technique and multiple orthogonal spreading factors (SFs). The system throughput is enhanced as multiple LoRa end devices can transmit at the same time and frequency slot in one channel, but with different SFs. Different SFs result in diverse signal to noise ratio (SNR) sensitives, which leads to different transmission rates and coverage ranges.

So far, extensive research has been carried out to investigate the impact of perfect and imperfect SF orthogonality. For instance, the authors in [57] have adopted the stochastic geometry tool to analyze the co-channel interference caused by LoRa users using the same SF over the same channel. Besides, the influence of imperfect SF orthogonality on the system throughput has been analyzed in [58] to provide insights on the SF assignment design of uplink LoRa networks. The packet loss caused by

inter-SF collisions has been numerically analyzed and then validated with experiments based on commercial devices in [59]. The joint SF assignment and transmit power allocation algorithm has been investigated in [60] to improve throughput fairness by considering both co-SF and inter-SF interferences.

Typically, adaptive data rate (ADR) mechanism can be enabled in the LoRaWAN to adjust the SF assignment and transmit power, based on the messages obtained from previous uplink measurements [61]. The ADR mechanism has been presented in [62] from the perspective of the average coverage time. In addition, the near-far problem and fair data rate deployment ratios have been addressed in [63] to achieve the data rate fairness among LoRa nodes based on LoRaWAN. By taking the uplink throughput and data transmission times of a single end device as the performance metrics, the authors in [64] have analyzed the capacity and scalability performance for uplink LoRaWAN.

However, the centralized ADR scheme is inefficient as it requests a number of uplink and downlink information exchanges to update the transmit power and SF step by step. Moreover, ADR is unable to deal with user collisions, which boosts the need to design efficient user scheduling and power allocation schemes. Besides, by adopting an interference-based simulation model, the authors in [65] have analyzed the LoRa scalability, i.e., the number of end devices that can be served per gateway. The performance of a LoRa network has been investigated in [66] to guarantee the fairness among LoRa users, with particular focus on the effects of interference caused by LoRa users using the same SF.

1.3 Thesis Outline and Contributions

Motivated by the discussions aforementioned, note that NOMA is demonstrated to significantly improve data rate and also regarded as the advanced MA technique for mMTC, MEC is considered to be a key enabler to achieve URLLC, and LoRa is the leading technology of LPWA techniques to realize mMTC. Therefore, this thesis aims to optimize the corresponding resource allocation schemes for NOMA, MEC, and LoRa to better meet the requirements of 5G and B5G services.

The main contributions of each chapter are summarized as follows.

Chapter 1 provides the background on the development of wireless communication systems and stands out the necessity of resource allocation optimization in future wireless communication systems. Besides, a detailed literature review of the related research works is presented.

Chapter 2 presents the background theory related to the system design of this thesis. We first give a brief introduction to resource allocation management and convex optimization theory. Then, we provide the description of three potential 5G and B5G technologies, i.e., NOMA, MEC, and LoRa, and point out the feature of each technology. It is worth mentioning that the purpose of this chapter is to present a comprehensive background overview, which helps readers better understand the rest of this thesis.

In Chapter 3, the robust beamforming design for a SWIPT-enabled cooperative NOMA system is studied. Concerning the SWIPT-enabled cooperative NOMA system, the strong user acts a relay to improve the connection between the BS and the weak user. To avoid consuming the strong user's energy, the power utilized to transmit the information of the weak user is assumed to be obtained from SWIPT. By considering a more practical scenario, it is assumed that only imperfect CSI is known at the BS. Two major design criteria, namely outage-based constraint design and the worst-case based optimization, are adopted. With the considered two different channel uncertainty models, the strong user's data rate maximization problems are formulated by designing the robust beamforming vector and power splitting (PS) ratio.

In Chapter 4, considering the practical non-linear energy harvesting model, the application of user cooperation (UC) and NOMA in a wireless powered MEC system is investigated. The mobile users first harvest energy from a multi-antenna AP, then both users simultaneously offload tasks to the MEC server with the harvested energy, by performing NOMA protocol. UC scheme is further conducted, where the near user acts as a relay to help the far user offload tasks to the AP. To achieve energy efficient design, the computation efficiency measurement metric, defined as the ratio of the system computation bits to the consumed energy, is introduced and adopted. The objective is to maximize the computation efficiency by jointly optimizing beamforming vectors, time and power allocations.

In Chapter 5, the energy efficient resource allocation is investigated to maximize the system EE (SEE) and the minimal EE (MEE) of LoRa users in LoRa networks, respectively. To deal with the formulated nonconvex problems, the corresponding EE is maximized by separately exploiting user scheduling, SF assignment, and transmit power allocations. A suboptimal algorithm including the low-complexity user scheduling scheme based on matching theory and the heuristic SF assignment solution for LoRa users scheduled on the same channel is first proposed. To deal with the power allocation, an optimal algorithm is proposed for the SEE problem. Concerning

MEE, an iterative power allocation algorithm is proposed to maximize the achieved minimal EE achieved of LoRa users.

Finally, this thesis is summarized in Chapter 6, and several potential future research topics are further presented.

Chapter 2

Background Theory and Fundamental Concepts

2.1 Resource Allocation and Convex Optimization Theory

In the development of wireless communication systems, resource allocation management always remains the primary challenge, and its role becomes more and more important. Resource allocation management refers to the process of allocating resources to each user, based on the users' CSI and QoS requirements. The main resource in wireless communication systems is bandwidth, power, and time. The necessity of efficient resource allocation management is twofold: the dynamic nature of wireless networks caused by fading channels and mobile devices' mobility, and the scarceness of the limited wireless resources in practice. Moreover, the overall system performance enhancement does not only rely on a single wireless resource increase, but lies in the joint resource allocation method. By optimally allocating the available wireless resources, in addition to the significant system performance improvement, the system can also be more flexibly adapted to the channel feature and QoS requirements, hence further realizing a flexible communication service structure. Consequently, the optimal solutions to resource allocation continue to be a major concern.

Note that the convex optimization method is one of the most effective approaches to deal with resource allocation optimization problems in wireless communication systems. Therefore, a brief introduction to convex optimization theory is provided to help better understand the related mathematical steps for the investigated research topics.

Typically, a standard mathematical optimization problem [67, 68] can be charac-

terized as follows

$$\begin{aligned} \min \quad & f_0(x) \\ s.t. \quad & f_i(x) \leq 0,\ i = 1, \ldots, m, \\ & h_i(x) = 0,\ i = 1, \ldots, p, \end{aligned} \tag{2.1}$$

where $x \in \mathbb{R}^n$ is the optimization variable, f_0: $\mathbb{R}^n \to \mathbb{R}$ denotes the objective function or the cost function, $f_i : \mathbb{R}^n \to \mathbb{R}$ and $h_i : \mathbb{R}^n \to \mathbb{R}$ represent the inequality and equality functions, respectively. Equation (2.1) aims to find an x under the constraints $f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p$, that minimizes $f_0(x)$. The optimal value is denoted as $\mu^* = \inf\{f_0(x) | f_i(x) \leq 0,\ i = 1, \ldots, m, h_i(x) = 0,\ i = 1, \ldots, p\}$, which means that the optimal solution x^* provides the smallest objective value while satisfying the constraints.

Generally, standard optimization problems are difficult to solve, but convex optimization problems can be reliably tackled, which motivates us to transform general optimization problems into convex expressions. Before reformulating (2.6) as a convex problem, we first present the definition of convex functions. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if **dom** f is a convex set and

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta) f(y), \tag{2.2}$$

where the inequality holds for all x, $y \in$ **dom** f, and $0 \leq \theta \leq 1$. If a function $-f$ is convex, then f is concave. Moreover, f is strictly convex if strict inequality holds in (2.2) with $x \neq y$ and $0 < \theta < 1$. It is noted that for an affine function $f(x) = ax + b$ on $\mathbb{R}$, and a, $b \in \mathbb{R}$, it is both convex and concave.

Furthermore, a standard form convex optimization problem can be characterized as the same expression given in Eq. (2.1), on the condition that $f_0, \ldots, f_m$ are convex functions, and equality constraints are affine functions. Compared with the general form, the convexity is guaranteed with the following conditions satisfied: 1) the objective function $f_0(x)$ must be convex; 2) the inequality functions $f_i,\ i = 1, \ldots, m$, must be convex; 3) the equality functions $h_i = a_i^T x + b_i,\ i = 1, \ldots, p$, must be affine [67].

The domain is defined as

$$\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom} f_i \cap \bigcap_{i=1}^{p} \mathbf{dom} h_i, \tag{2.3}$$

where $x \in \mathcal{D}$ is the implicit constraint for the optimization problem.

It has been proved that any local optimal solution of a convex problem is also globally optimal. Consequently, a globally optimal solution $x^* \in \mathcal{D}$ to the convex

problem can be found by applying the convex optimization approaches such as interior point method, ellipsoid method and subgradient method [67]. Therefore, if a practical problem can be transformed into convex forms, then it can be efficiently solved.

Furthermore, successive convex approximation (SCA) is considered to an effective method to approximate nonconvex optimization problems. The key idea of SCA is to sequentially optimize the nonconvex function by establishing a convex trust region around the original nonconvex spatial points. Though the approximation results may heavily depend on the initial points, it has been verified that SCA often works well in practical applications [67]. Take (2.1) as an example, denote solution of the k-th iteration as $x^{(k)}$, form convex approximation $\hat{f}_i$ of f_i, affine estimate $\hat{g}_i$ of g_i, then the optimal point during the $(k+1)$-th iteration can be obtained by solving the following approximated convex problem:

$$\begin{aligned} \min \quad & \hat{f}_0(x) \\ s.t. \quad & \hat{f}_i(x) \leq 0,\ i = 1, \ldots, m, \\ & \hat{h}_i(x) = 0,\ i = 1, \ldots, p, \\ & x \in \Gamma^{(k)}, \end{aligned} \tag{2.4}$$

where $\Gamma^{(k)}$ is the convex trust region that can be denoted as the box around the current point, i.e., $\Gamma^{(k)} = \{x \mid\mid x_i - x_i^{(k)} \mid\leq \triangle_i\}$, $\triangle_i$ represents random positive values.

Besides, semedefinite relaxation (SDR) is a powerful technique to approximate nonconvex quadratically constrained quadratic program (QCQP). The general complex-value QCQP can be characterized as [69]

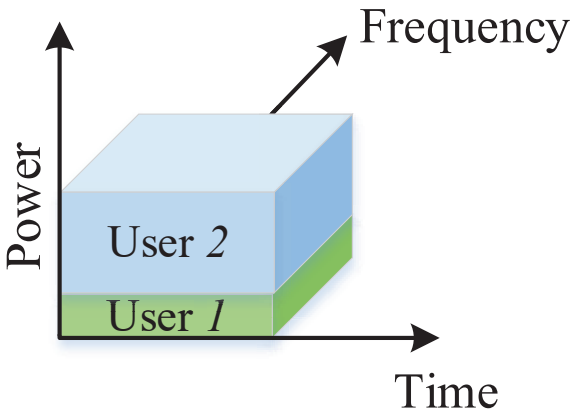
$$\begin{aligned} \min_{x \in \mathbb{C}^n} \quad & \boldsymbol{x}^{\boldsymbol{H}} \boldsymbol{C} \boldsymbol{x} \\ s.t. \quad & \boldsymbol{x}^{\boldsymbol{H}} \boldsymbol{A}_i \boldsymbol{x} \geq b_i,\ i = 1, \ldots, m, \end{aligned} \tag{2.5}$$

where $\boldsymbol{C}, \boldsymbol{A}_i \in \mathbb{H}^n$, with $\mathbb{H}^n$ being the $n \times n$ Hermitian matrices.

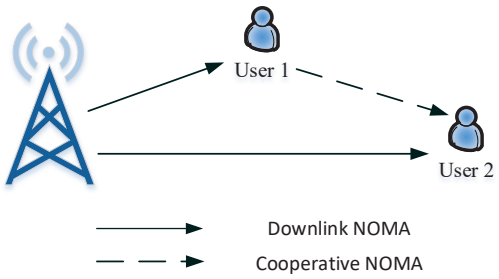
Based on the observation $\boldsymbol{x}^{\boldsymbol{H}} \boldsymbol{C} \boldsymbol{x} = \mathrm{Tr}(\boldsymbol{x}^{\boldsymbol{H}} \boldsymbol{C} \boldsymbol{x}) = \mathrm{Tr}(\boldsymbol{C} \boldsymbol{x}^{\boldsymbol{H}} \boldsymbol{x})$, by introducing $\boldsymbol{X} = \boldsymbol{x}^{\boldsymbol{H}} \boldsymbol{x}$ and dropping the rank-one constraint, (2.5) can be reformulated as

$$\begin{aligned} \min_{X \in \mathbb{C}^n} \quad & \mathrm{Tr}(\boldsymbol{C} \boldsymbol{X}) \\ s.t. \quad & \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \geq b_i,\ i = 1, \ldots, m. \end{aligned} \tag{2.6}$$

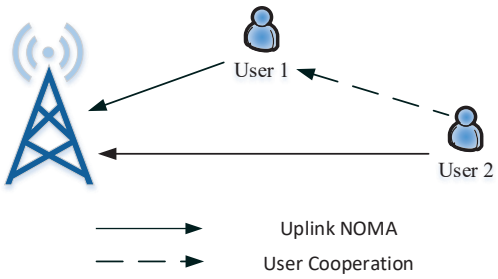
Problem (2.6) is known as a SDR of (2.5), which can be solved by convex solver. To establish the equivalence, the rank-one optimality needs to be further proved.



(a) Power multiplexing NOMA



(b) Downlink NOMA



(c) Uplink NOMA

Figure 2.1: NOMA transmission structure.

2.2 Non-orthogonal Multiple Access

2.2.1 Basic Principles of NOMA

NOMA is regarded as the potential MA technique for 5G and B5G networks. A power multiplexing NOMA scheme is illustrated in Fig. 2.1(a), where multiple users can share the same time/frequency/code domain by performing power multiplexing. The key enabling techniques in NOMA are SC [70, 71] and SIC technology [72, 73].

Specifically, for a downlink NOMA transmission, take a two-user scenario as an example (solid line in Fig. 2.1(b)), the BS transmits the superposition coded signals to all users. At the user side, the user with poor channel conditions (User 2) is allocated with more power to decode its signal by treating the other user's signal as noise. By invoking SIC technology, the user with better channel conditions (User 1) can first decode the message of User 2 and remove the interference from the signals, after

which User 1 decodes its own information. Regarding uplink NOMA transmission (solid line in Fig. 2.1(c)), User 1 and User 2 transmit their own signals to the BS. At the BS side, the signal of User 1 is first decoded by taking User 2's signal as noise. Then SIC is implemented at the BS to remove the signal of User 1, and the data of User 2 is further decoded.

Though it has been demonstrated that NOMA achieves evident gain improvement over existing MA schemes, some practical issues related to SIC application arise in NOMA communication systems. Particularly, due to the utilization of SIC technology, the user with better channel conditions needs to decode the information intended for weak users in the downlink transmission, while for the uplink scenario, the BS has to decode the information of all users with a given decoding order. As a consequence, the decoding complexity (for the user with better channel conditions in downlink transmission, and at the BS for uplink transmission) increases rapidly when there are a large number of users. To reduce the decoding complexity, user pairing is proposed to separate users into groups, where each group contains only a limited number of users that can be multiplexed on the same channel. It has been proved that a better tradeoff between decoding complexity and system performance can be achieved when two users allocated on the same channel [74, 75]. The reason is that, there is a high probability that two users have distinct channel differences to achieve the largest performance gain, while the receiver possesses low complexity. Therefore, similar to most previous works [20, 75], for the researches related to NOMA in this thesis, we assume that the users have already been grouped into pairs. Therefore, the case of two users is investigated in Chapter 3 and Chapter 4.

2.2.2 Cooperative NOMA

Consider a downlink NOMA transmission, the BS communicates with two users where User 1 is a strong user and User 2 is a weak user, which is shown in Fig. 2.1(b). Due to the application of SIC, the prior information of User 2 can be obtained at User 1, which can be further exploited with a cooperative transmission scheme, i.e., User 1 is regarded as a DF relay to help transmit signal of User 2. Two time slots are included in a cooperative NOMA transmission. During the first direct transmission slot (solid line in Fig. 2.1(b)), the BS sends the SC signals to two users. In the second cooperative transmission slot (dash line in Fig. 2.1(b)), User 1 helps transmit the decoded signal to User 2. The advantages of cooperative NOMA are twofold: i) the reception reliability of weak users in cooperative NOMA can be greatly enhanced,

Table 2.1: Comparison Between MCC and MEC.

Paradigm	MCC	MEC
Deployment	Centralized	Distributed
Distance from users to servers	High	Low
Latency	High	Low
Power at server	Ample	Limited
Storage capacity at server	Ample	Limited

thus improving user fairness. ii) multi-path fading can be better controlled since cooperative NOMA provides higher diversity gain.

Unlike downlink NOMA transmission, SIC is performed at the BS to decode the users' signals following a given decoding order for uplink NOMA transmission. Furthermore, UC can still be exploited at the strong user to improve the fairness of the weak user (dash line in Fig. 2.1(c)). To exploit the benefit of UC while maintaining the advantages of NOMA, three slots are included for UC-enabled uplink NOMA transmission. The offloaded information of both users is divided into two segments, where the two segments are transmitted to the BS directly in the first and the third slot for User 1. For User 2, the first segment is transmitted collaboratively to the BS in the first and second slots, and the second segment is transmitted directly to the BS in the third slot. Specifically, in the first slot, two users simultaneously transmit signals to the BS, while User 1 can receive the signal of User 2 at the same time. During the next slot, User 1 acts as a DF relay to help forward the signal of User 2 to the BS. The transmitted data of User 2 is constrained by the effect of UC. In the third slot, User 1 and User 2 send their own information to the BS. Note that Chapter 4 of this thesis is based on uplink NOMA transmission with UC, as it can overcome the "doubly near-far" effect in the wireless powered MEC system.

2.3 Mobile Edge Computing

2.3.1 Comparison Between MCC and MEC

The last decade has witnessed the data explosion in the IoT networks as well as the 5G and B5G wireless communications, which brings along computation-intensive applications and mission-critical tasks. Though new mobile devices are equipped with more powerful CPU, mobile devices are still most likely unable to meet these requirements due to the limited processing capability and restricted battery energy. To address

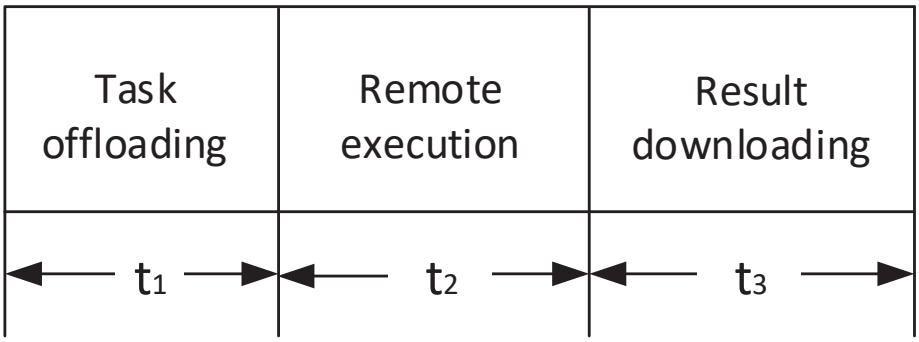


Figure 2.2: Time allocation of computation offloading.

the concerns mentioned aforementioned, the centralized mobile cloud computing (MCC) paradigm has been proposed to offer remote powerful computing by leveraging the vast resources at the distant cloud [76]. Though MCC greatly enhances system performance by exploiting the cloud's adequate computational energy and storage capacity, MCC introduces high latency due to the far distance of clouds.

To overcome the drawback of high latency for MCC while maintaining the higher computing capability, it is desired to deploy servers in close proximity to mobile users, which introduces the new paradigm MEC. The concept of MEC was firstly proposed and defined by the European Telecommunications Standards Institute (ETSI) in 2014 [77], and MEC is intended to provide cloud computing capabilities within the radio access network (RAN). The detailed comparison between MCC and MEC is illustrated in Table 2.1.

2.3.2 MEC Computation Offloading

Typically, the MEC system features a three-layer architecture, which consists of mobile devices, MEC servers, and the cloud, and the cloud-like MEC server is integrated with the BS or AP at the edge of networks. MEC design involves resource allocation of both computing and communication processes, as computation tasks can be computed locally at the devices and remotely at the MEC servers, while the connections between mobile devices and MEC servers are established by wireless communication techniques to realize task offloading and result downloading. As a consequence, three phases are embodied in the computation offloading, which is shown in Fig. 2.2. Mobile devices first offload computation tasks to MEC servers, and then MEC servers compute the offloaded tasks. After that, mobile devices download the result from the MEC servers.

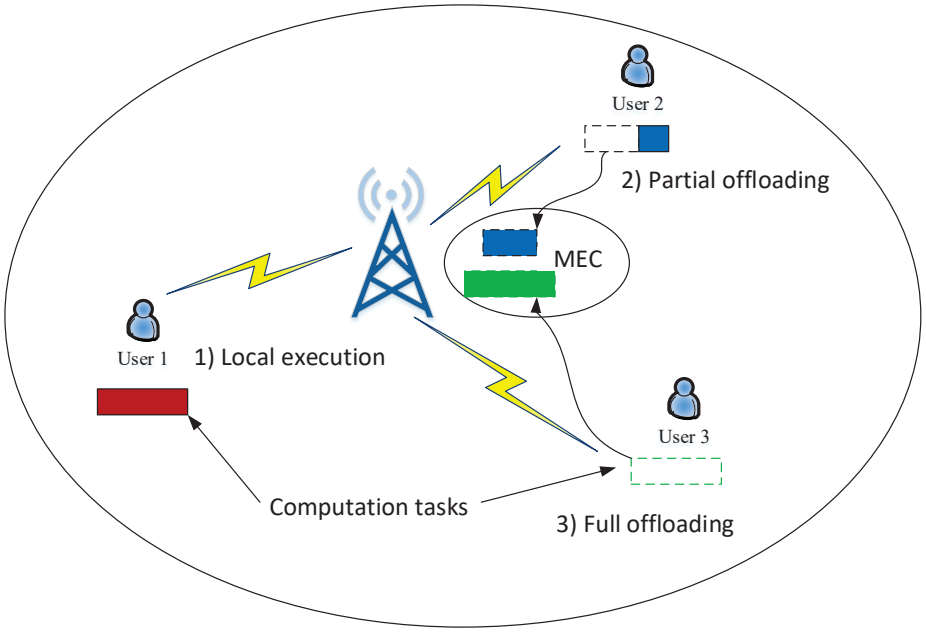


Figure 2.3: Model of computation offloading decisions.

The boosting computation-intensive latency-critical mobile applications benefit from MEC, by executing the offloaded computing tasks remotely at the MEC server. According to the feature of computation tasks, computation offloading decisions are divided into three categories, namely local execution, partial offloading and full offloading, which is shown in Fig. 2.3. Specifically, the bubbles represent the total required computed tasks for each user, where solid bubbles denote the part of tasks that can be executed by local computing, and hollow bubbles mean the tasks executed by MEC serves. Regarding local execution, the whole computation bits are executed locally at mobile users only. For partial offloading, it is assumed that the computation tasks can be artificially divided, where part of the computation bits is computed locally while the rest is executed at the MEC server. Full offloading denotes that the whole computation bits are processed remotely at the MEC server, which is suitable for the case that the computational capability of local computing is negligible compared to MEC servers. The major challenges for MEC design include computation offloading decisions, computing resources allocation, and communication resource management. In this thesis, the efficient resource allocation of the MEC system is considered in Chapter 4.

2.4 Long Range Communication

2.4.1 Basic Principles of LoRa

To complement the weakness of conventional cellular and short-range communications in IoT scenarios, LPWA networks are envisioned as effective ways to support the massive number of end devices as well as the long coverage and low power consumption feature in IoT networks. LWPA stands for a series of technologies such as LoRa [78], Sigfox [79], LTE-M [9], and narrow-band (NB)-IoT [80], which can be capable of providing long-distance power-efficient communications at the expense of low data rates. With the advantages of long-range capability, low power consumption, and massive network capacity, LoRa becomes the utmost promising LPWA technology for IoT networks. LoRa was first proposed by Semtech and further developed by the LoRa Alliance [55]. In addition, LoRa is on the way towards standardization and commercialization. LoRa operates with license-free ISM bands (Europe: 868 MHz and 433 MHz, US: 915 MHz) to enable flexible transmission distance range and data rates with different spreading factors (SFs). SF ranges from 7 to 12, and higher SFs provides long-distance communication at the cost of reduced data rate, whereas lower SFs provide high data rate but limited transmission range. Furthermore, to support the LoRa physical layer operation, the higher layers were defined by LoRa Alliance and called LoRaWAN. LoRa supports long-distance communication links while LoRaWAN defines the system architecture for the networks.

2.4.2 LoRaWAN Architecture

The LoRaWAN architecture is illustrated in Fig. 2.4. Specifically, the MAC protocol in LoRaWAN adopts pure ALOHA access (communication starts when end devices have data ready to transmit) with duty cycle limitations, which reduces energy consumption based on listening and sensing mode. Depending on different application scenarios and service requirements, three classes of end devices, namely Class A, Class B, and Class C, are defined in LoRaWAN. The most energy efficient Class A is intended for battery-powered sensors, which supports maximum battery lifetime but allows the biggest latency. All end devices are required to support the functionality of Class A. Class B is used by battery-powered latency-controlled sensors and actuators, which realizes synchronization by receiving a beacon from the LoRa gateway. Class C end devices are actuators with sufficient external power supply, which have the minimum latency and can stay in a continuous receive window during the transmission.

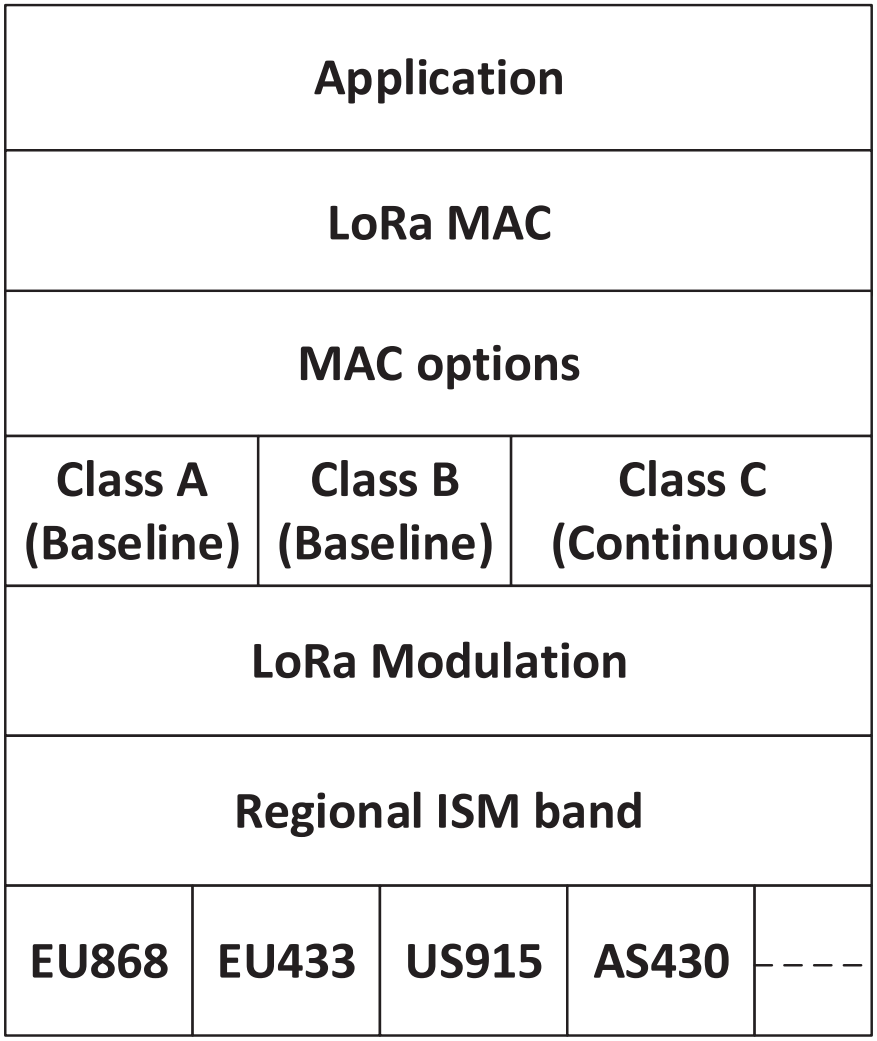


Figure 2.4: System architecture of LoRaWAN.

2.5 Energy Efficient Communications

Recently, energy efficient communications have drawn tremendous attention in both industry and academia, due to the fact that the ever-increasing energy consumption of the information and communication technologies (ICT) contributes more and more to the greenhouse gas emissions [23]. To reveal the system efficiency from the perspective of the computed bits per Joule, according to the feature of systems, two energy efficient definitions in terms of EE [81,82] and computation efficiency [83] are adopted for LoRa networks in Chapter 5 and MEC systems in Chapter 4, respectively.

2.5.1 Energy Efficient Resource Allocation

In wireless communication systems, EE is considered to be an essential measurement metric in system design from the perspective of both operators and mobile users [81]. The purpose of EE is to provide a tradeoff between the transmitted data rate and energy consumption. The main aim of energy efficient communications is to maximize the amount of transmitted data per Joule of consumed energy by allocating resources optimally.

EE is commonly defined as the ratio of transmitted bits to energy consumption, which is given as

$$\eta = \frac{RT}{(P + P_c)T} = \frac{R}{(P + P_c)} \text{ bits/Joule}, \tag{2.7}$$

where R represents the achievable data rate, T is a given time duration, P is the total transmit power, and P_c denotes the constant circuit power consumption, which is independent of the transmit power. Specifically, $R = B\log_2(1 + \frac{P|h|^2}{\sigma^2})$, where B denotes the bandwidth, $|h|^2$ is the channel gain and σ^2 is the additive white Gaussian noise (AWGN) variance. Moreover, P_c stands for the circuit power consumed by the mixer, the low-noise amplifier, frequency synthesizer, transmit filter, and digital-to-analog converter [84]. In this thesis, the energy efficient transmission for uplink LoRa networks is investigated in Chapter 5.

2.5.2 Computationally Efficient Resource Allocation

Regarding MEC, most previous resource allocation works on MEC systems focus on either maximizing the sum computation rates [47, 85], or minimizing the consumed energy [86, 87], which cannot achieve good tradeoff between the energy consumption and the computation bits (which is defined as the total number of computed bits at both the users and MEC server). To realize energy efficient design, it is desirable to show the MEC efficiency from the perspective of maximizing the achievable computation bits per Joule.

Different from conventional communication systems, MEC involves not only wireless communication but also computing processes. Moreover, the MEC system is required to handle computation-intensive yet latency-critical tasks, which poses strict delay budget demand. Therefore, it may become the second consideration to maximize the communication throughput, compared to the requirement of satisfying the latency constraint for offloading through wireless communications. For such a latency-constrained scenario, to capture the efficiency of the energy considering both the local computing and remote task execution phase, the computation efficiency measurement

metric [88], defined as the ratio of the computation bits to the energy consumption [1], is given as

$$\tilde{\eta} = \frac{\ell^{loc} + RT}{E^{loc} + PT} \text{ bits/Joule}, \tag{2.8}$$

where R is the achievable offloading data rate, T is the offloading time duration, P denotes the transmit power, ℓ^{loc} and E^{loc} represent the computation bits and consumed energy for local computing, respectively. The efficient resource allocation approach to maximize the computation efficiency is further discussed in Chapter 4.

2.6 Summary

In this chapter, we first give a brief description of background knowledge on resource allocation and introduce the convex optimization theory. Then we present the basic principles and technical foundation of several promising technologies for 5G and B5G communications. Finally, energy efficient communications are described to illustrate the research motivations.

[1]The time consumption for remote execution at the MEC server and computed results downloading is assumed to be negligible.

Chapter 3

Robust Transmit Beamforming for SWIPT-Enabled Cooperative NOMA with Channel Uncertainties

3.1 Introduction

In this chapter, we study the robust beamforming design for a SWIPT-enabled system, with cooperative NOMA protocol applied. A novel cooperative NOMA scheme is proposed, where the strong user with better channel conditions adopts power splitting (PS) scheme and acts as an energy-harvesting relay to transmit information to the weak user. The presence of channel uncertainties is considered and incorporated in our formulations to improve the design robustness and communication reliability. Specifically, only imperfect CSI is assumed to be available at the BS, due to the reason that the BS is far away from both users and suffers serious feedback delay. To comprehensively address the channel uncertainties, two major design criteria are adopted, which are the outage-based constraint design and the worst-case based optimization. Then, our aim is to maximize the strong user's data rate, by optimally designing the robust transmit beamforming and PS ratio, while guaranteeing the correct decoding of the weak user. With two different channel uncertainty models respectively incorporated, the proposed formulations yield to challenging nonconvex optimization

The works presented in this Chapter have been published at the IEEE Transactions on Communications, June 2019, and IEEE International Conference on Communications (ICC 2019), Shanghai, China, May 2019.

problems. For the outage-based constrained optimization, we first conservatively approximate the probabilistic constraints with the Bernstein-type inequalities, which are then globally solved by two-dimensional exhaustive search. To further reduce the complexity, an efficient low-complexity algorithm is then proposed with the aid of successive convex approximation (SCA). For the worst-case based scenario, we firstly apply semedefinite relaxation (SDR) method to relax the quadratic terms and prove the rank-one optimality. Then the nonconvex max-min optimization problem is readily transformed into convex approximations based on S-procedure and SCA. Simulation results show that for both channel uncertainty models, the proposed algorithms can converge within a few iterations, and the proposed SWIPT-enabled robust cooperative NOMA system achieves better system performance than existing protocols.

The remainder of this chapter is organized as follows. In Section 3.2, we give a brief introduction to the system model of the proposed SWIPT-enabled robust cooperative NOMA system. The probabilistic SINR constrained optimization problem is formulated and analyzed in Section 3.3. Further, In Section 3.4, the data rate maximization problem for the strong user is formulated and solved by adopting the worst-case based channel uncertainty model. Finally, numerical results are given in Section 3.5, followed by conclusions in Section 3.6.

3.2 System Model and Problem Formulation

3.2.1 System Model

Consider a downlink TDMA MISO transmission system, as shown in Fig. 3.1, wherein the BS is equipped with N_t antennas and all users are equipped with the single antenna. There are two users in each beam and the BS performs MISO transmission with N users through M beams, where $N = 2M$ [1]. Assume that the total time duration T is equally divided into M slots where the time duration of each slot is $t_m = \frac{T}{M}$. Then during the time slot t_m, only the m-th user-pair is allowed to transmit, while the other user-pairs remain silent. Therefore, the inner-cell interference between pairs of users does not exist. As a result, when formulating the rate-maximization problems and designing robust beamforming, we can just focus on one beam.

Let us take the first beam as an example. It is assumed that NOMA protocol is adopted for the two users. Without loss of generality, we assume that user 1 is a

[1]It is assumed that users have already been grouped into pairs, and we can refer to [89] for how to do user pairing.

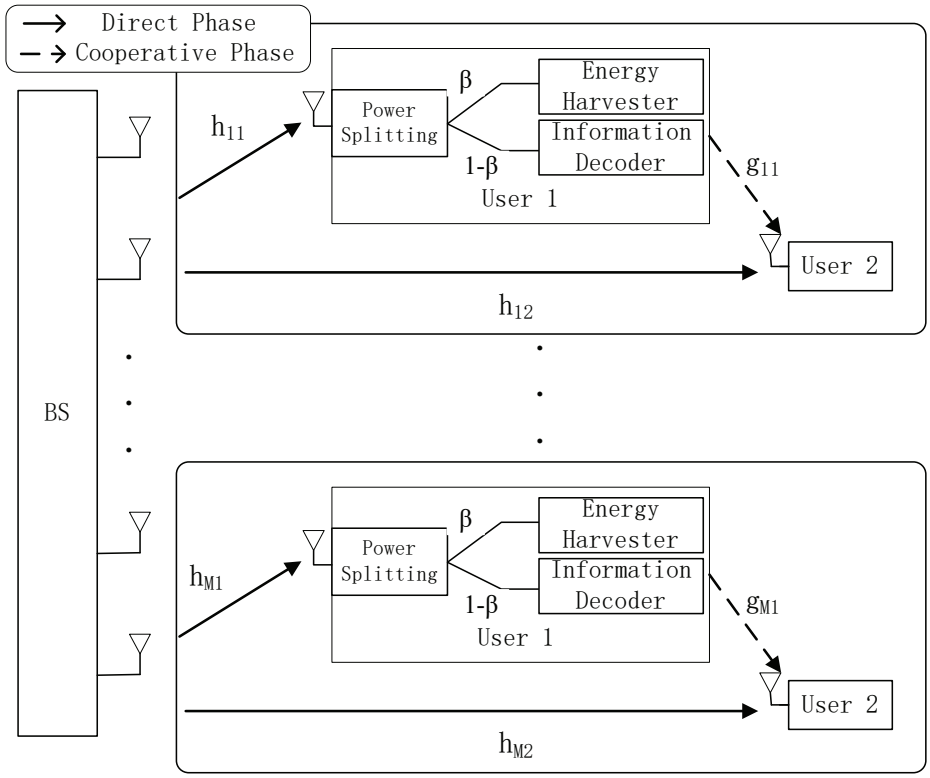


Figure 3.1: System model for the SWIPT-enabled cooperative NOMA.

strong user with better channel conditions and user 2 is a weak user. According to NOMA protocol, user 1 removes the interference of user 2 by applying SIC and then detect its own information, while the user 2 treats the user 1's message as noise. Since user 1 can obtain prior information of the messages for user 2 and thus can act as a relay to improve the connection between the BS and user 2. In order to help user 2 without draining user 1's battery, we assume that the power utilized to transmit the information of user 2 can be obtained from SWIPT.

Since the imperfect CSI case at the BS is considered, we first introduce the CSI error model. The actual channels between the BS and two users can be characterized as

$$\boldsymbol{h}_i = \tilde{\boldsymbol{h}}_i + \boldsymbol{e_i}, \quad i = 1, 2, \tag{3.1}$$

where $\boldsymbol{h}_i$ denotes the actual channel gain, $\tilde{\boldsymbol{h}}_i$ is the estimated channel gain at the BS, and $\boldsymbol{e_i}$ represents the channel errors of two users. The detailed expression of $\boldsymbol{e_i}$ for the two channel uncertainty models will be introduced in the next two sections.

Two phases are involved in the SWIPT-enabled robust cooperative NOMA transmission. At the first *robust direct transmission phase*, user 1 coordinates the process

of information decoding (ID) and energy harvesting from the received signal by adopting PS scheme. Specifically, as can be seen in Fig. 1, the received signal at user 1 is split into the information decoder and the energy harvester. As for user 2, it receives the direct transmission signal from the BS at this phase. Then in the second *cooperative transmission phase*, user 1 forwards the decoded user 2's message to user 2 with the harvested energy. The detailed process is summarized as follows.

3.2.2 Robust Direct Transmission Phase

During this phase, the signals for two users are superposition coded at the BS, i.e., $\boldsymbol{x} = \boldsymbol{w}_1 x_1 + \boldsymbol{w}_2 x_2$, where x_1 and x_2 are the messages for user 1 and user 2 respectively. The power of the transmitted symbol is normalized, i.e., $\mathbb{E}\|x_1\|^2 = \mathbb{E}\|x_2\|^2 = 1$, and $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are the corresponding precoding vector. Then, for the weak user, i.e., user 2, the observation is given by

$$y_2^{(1)} = \boldsymbol{h}_2^H(\boldsymbol{w}_1 x_1 + \boldsymbol{w}_2 x_2) + n_2, \tag{3.2}$$

where $\boldsymbol{h}_2^H$ denotes the Hermitian transpose of $\boldsymbol{h}_2 \in \mathbb{C}^{Nt\times 1}$, $n_2 \sim \mathcal{CN}(0,1)$ is the AWGN. In this chapter, we assume that all channels have the same noise value as σ^2=1. Then the SINR obtained by user 2 from the direct transmission can be expressed as

$$\mathrm{SINR}_2^{(1)} = \frac{|\boldsymbol{h}_2^H \boldsymbol{w}_2|^2}{1 + |\boldsymbol{h}_2^H \boldsymbol{w}_1|^2}. \tag{3.3}$$

Due to the assumption that there is not enough power to forward the signal of x_2 to user 2, user 1 needs to replenish the energy from the BS based on the 'harvest-then-transmit' protocol proposed in [90]. The PS scheme is employed at user 1 to perform SWIPT. Then, the information received at user 1 is given by

$$y_1 = \sqrt{1-\beta}\boldsymbol{h}_1^H(\boldsymbol{w}_1 x_1 + \boldsymbol{w}_2 x_2) + n_1, \tag{3.4}$$

where $\beta \in [0,1]$ is the PS ratio, $\boldsymbol{h}_1 \in \mathbb{C}^{Nt\times 1}$ is the channel coefficient between the BS and user 1, and $n_1 \sim \mathcal{CN}(0,1)$ is the AWGN. With SIC carried out at user 1, i.e., user 1 firstly decodes the message for user 2 and then removes the information of user 2 to decode its own information, the received SINR for user 1 to detect the message of user 2, is given by

$$\mathrm{SINR}_{1,2} = \frac{(1-\beta)|\boldsymbol{h}_1^H \boldsymbol{w}_2|^2}{1 + (1-\beta)|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2}. \tag{3.5}$$

After removing the message of user 2 from y_1, the corresponding SNR of user 1 can be expressed as

$$\mathrm{SNR}_1 = (1-\beta)|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2, \tag{3.6}$$

which will be our optimization objective in the next section.

Besides, to ensure the correct decoding capability in a given order, we have the following inequality requirements [91]:

$$|\boldsymbol{h}_1^H \boldsymbol{w}_2|^2 \geq |\boldsymbol{h}_1^H \boldsymbol{w}_1|^2, \tag{3.7a}$$

$$|\boldsymbol{h}_2^H \boldsymbol{w}_2|^2 \geq |\boldsymbol{h}_2^H \boldsymbol{w}_1|^2. \tag{3.7b}$$

Furthermore, with PS protocol applied at user 1 to harvest energy from the BS, the harvested energy can be given as [92]

$$E = \vartheta\beta(|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H \boldsymbol{w}_2|^2)\mathrm{T}, \tag{3.8}$$

where ϑ and T denote the energy harvesting efficiency and the transmission time fraction, respectively. Without loss of generality, we set $\mathrm{T} = \frac{1}{2}$ which means that equal time duration is assigned for direct and cooperative transmission stages. Hence, the available average power of user 1 can be expressed as

$$P_r = \frac{\vartheta\beta(|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H \boldsymbol{w}_2|^2)T}{1-T} = \vartheta\beta(|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H \boldsymbol{w}_2|^2). \tag{3.9}$$

It is worthwhile to point out that only when user 1 can successfully decode the signals of user 2, it can then use the harvested energy to forward the signals to user 2. This means that it is more important for user 1 to decode the signals than performing energy harvesting.

3.2.3 Cooperative Transmission Phase

In the cooperative transmission phase, if the received SINR for user 1 to detect the message of user 2 is larger than or equal to the target SINR of user 2, we can assume that user 1 can correctly decode the received symbols of user 2 [93]. Then user 1 forwards signal x_2 to user 2 using the harvested energy. The observation of user 2 at this phase can be characterized as

$$y_2^{(2)} = \sqrt{P_r} g x_2 + n_3, \tag{3.10}$$

where P_r is the available power of user 1, g is the perfectly known channel coefficient between user 1 and user 2, and $n_3 \sim \mathcal{CN}(0,1)$ is the normalized AWGN. The achievable SNR of user 2 at this phase can be written as

$$\mathrm{SNR}_2^{(2)} = \vartheta\beta|g|^2(|\boldsymbol{h}_1^H\boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H\boldsymbol{w}_2|^2). \tag{3.11}$$

Combining the observation from both phases and using maximal ratio combination (MRC), the equivalent SINR of user 2 can be finally obtained as

$$\mathrm{SINR}_2 = \mathrm{SINR}_2^{(1)} + \mathrm{SNR}_2^{(2)} \quad = \frac{|\boldsymbol{h}_2^H\boldsymbol{w}_2|^2}{1+|\boldsymbol{h}_2^H\boldsymbol{w}_1|^2} + \vartheta\beta|g|^2(|\boldsymbol{h}_1^H\boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H\boldsymbol{w}_2|^2). \tag{3.12}$$

In the next two sections, we aim to maximize the data rate of user 1, which is equivalent to maximize the SNR of user 1, subject to the outage-based and worst-case based constraints respectively.

3.3 Outage-based Constrained Optimization

In this section, the outage-based probabilistic constraints caused by imperfect CSI will be investigated, where the unsuccessful decoding of weak user falls into the scope of outage. The goal is to design beamforming vectors $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ to maximize the data rate of user 1, which is equivalent to maximize the SNR of user 1 while guaranteeing the outage requirements. Specifically, the outage for strong user happens when it is not able to decode the weaker user's information, while for the weak user, the outage means that it can not successfully decode its own information.

The study of outage-based constrained robust optimization is a meaningful design criterion as CSI errors are universal present in practical systems, and they may cause severe outage if not handled properly [94]. However, as the probability functions can not yield straightforward closed-form expressions, how to deal with probabilistic constraints is of vital importance. To tackle the problem, we will resort to Bernstein-type inequality approach to deal with the probability constraints.

We consider the following robust beamforming design problem [2]:

$$\max_{\beta,\boldsymbol{w}_1,\boldsymbol{w}_2} \quad (1-\beta)|\boldsymbol{h}_1^H\boldsymbol{w}_1|^2 \tag{3.13a}$$

$$s.t. \quad \Pr\left(\mathrm{SINR}_{1,2} \geq \gamma\right) \geq 1-\rho_1, \tag{3.13b}$$

$$\Pr\left(\mathrm{SINR}_2 \geq \gamma\right) \geq 1-\rho_2, \tag{3.13c}$$

$$||\boldsymbol{w_1}||_2^2 + ||\boldsymbol{w_2}||_2^2 \leq P_{\max}, \tag{3.13d}$$

$$\left|\boldsymbol{h}_1^H\boldsymbol{w}_2\right|^2 \geq \left|\boldsymbol{h}_1^H\boldsymbol{w}_1\right|^2, \tag{3.13e}$$

$$\left|\boldsymbol{h}_2^H\boldsymbol{w}_2\right|^2 \geq \left|\boldsymbol{h}_2^H\boldsymbol{w}_1\right|^2, \tag{3.13f}$$

$$0 \leq \beta \leq 1, \tag{3.13g}$$

where γ is the target SINR of user 2, $\rho_i \in [0,1)$, $i=1,2$, is the maximum tolerable outage probability for two users, and $P_{\max}$ is the maximum available power at the BS. Constraints (3.13e) and (3.13f) represent the given order decoding capability requirements [91].

To solve the problem (3.13), we first relax it by applying SDR approach and drop the rank-one constraint. Specifically, we replace the beamforming vector $\boldsymbol{w}_i$ by semidefinite positive matrices $\boldsymbol{W}_i$, i.e,

$$\boldsymbol{W}_i = \boldsymbol{w}_i\boldsymbol{w}_i^H, i=1,2. \tag{3.14}$$

Then, since imperfect CSI is considered at the BS, the channel errors can be modeled as

$$\boldsymbol{e_i} = \boldsymbol{C}_i^{\frac{1}{2}}\boldsymbol{e}, \quad i=1,2, \tag{3.15}$$

where $\boldsymbol{C}_i \succeq \mathbf{0}$ denotes some known error covariance and $\boldsymbol{e} \sim \mathcal{CN}(\mathbf{0}, \mathrm{I}_{N_t})$.

By replacing $\boldsymbol{h}_i$ with $\tilde{\boldsymbol{h}}_i + \boldsymbol{C}_i^{\frac{1}{2}}\boldsymbol{e}$ and denoting that $\Gamma = (\tilde{\boldsymbol{h}}_1 + \boldsymbol{C}_1^{\frac{1}{2}}\boldsymbol{e})^H(\boldsymbol{W}_2 - \gamma\boldsymbol{W}_1)(\tilde{\boldsymbol{h}}_1 + \boldsymbol{C}_1^{\frac{1}{2}}\boldsymbol{e})$, the probabilistic SINR constraint (3.13b) can be recast as

$$\Pr\left(\Gamma \geq \frac{\gamma}{1-\beta}\right) \geq 1-\rho_1. \tag{3.16}$$

On the other hand, for the outage-based SINR constraint (3.13c), we transform it by introducing an auxiliary variable θ. Firstly, (3.13c) can be decomposed into the following two sub-problems:

$$\Pr\left\{\vartheta\beta|g|^2(|\boldsymbol{h}_1^H\boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H\boldsymbol{w}_2|^2) \geq \gamma - \theta\right\} \geq 1-\rho_2, \tag{3.17a}$$

[2]We do not pose any threshold requirement for user 1, since the primal purpose of cooperative NOMA is to guarantee fairness of the weak user, thus the QoS requirement of the weak user is more important. In this case, it is assumed that once user 1 can detect the signal of user 2, it can decode its own signal to help user 2.

$$\frac{|\boldsymbol{h}_2^H \boldsymbol{w}_2|^2}{|\boldsymbol{h}_2^H \boldsymbol{w}_1|^2 + 1} \geq \theta, \tag{3.17b}$$

where the optimality of the decomposition can be assured when (3.17b) holds with equality. It is worth noting that (3.17a) has the same form as (3.13b). Furthermore, with the application of SDR, (3.17b) can be further described as

$$\theta \mathrm{Tr}(\boldsymbol{H}_2 \boldsymbol{W}_1) \leq \mathrm{Tr}(\boldsymbol{H}_2 \boldsymbol{W}_2) - \theta, \tag{3.18}$$

where $\boldsymbol{H}_i \triangleq \boldsymbol{h}_i \boldsymbol{h}_i^H, i = 1, 2$.

Further, by introducing several auxiliary variables, i.e., Q_i, r_i and $s_i, i = 1, 2$, the following correspondence to (3.13b) and (3.17a) can be shown

$$\boldsymbol{Q}_1 = \boldsymbol{C}_{\mathbf{1}}^{\frac{1}{2}} (\boldsymbol{W}_2 - \gamma \boldsymbol{W}_1) \boldsymbol{C}_{\mathbf{1}}^{\frac{1}{2}}, \tag{3.19a}$$

$$\boldsymbol{r}_1 = \boldsymbol{C}_{\mathbf{1}}^{\frac{1}{2}} (\boldsymbol{W}_2 - \gamma \boldsymbol{W}_1) \tilde{\boldsymbol{h}}_1, \tag{3.19b}$$

$$s_1 = \tilde{\boldsymbol{h}}_1^H (\boldsymbol{W}_2 - \gamma \boldsymbol{W}_1) \tilde{\boldsymbol{h}}_1 - \frac{\gamma}{1 - \beta}, \tag{3.19c}$$

$$\boldsymbol{Q}_2 = \boldsymbol{C}_{\mathbf{1}}^{\frac{1}{2}} (\boldsymbol{W}_1 + \boldsymbol{W}_2) \boldsymbol{C}_{\mathbf{1}}^{\frac{1}{2}}, \tag{3.19d}$$

$$\boldsymbol{r}_2 = \boldsymbol{C}_{\mathbf{1}}^{\frac{1}{2}} (\boldsymbol{W}_1 + \boldsymbol{W}_2) \tilde{\boldsymbol{h}}_1, \tag{3.19e}$$

$$s_2 = \tilde{\boldsymbol{h}}_1^H (\boldsymbol{W}_1 + \boldsymbol{W}_2) \tilde{\boldsymbol{h}}_1 - \frac{\gamma - \theta}{\vartheta \beta |g|^2}. \tag{3.19f}$$

Finally, the original probabilistic outage constraint (3.13b) and the reformulated (3.17a) can be written as the following same structure

$$\Pr\left\{\boldsymbol{e}^H \boldsymbol{Q}_i \boldsymbol{e} + 2Re\{\boldsymbol{e}^H \boldsymbol{r}_i\} + s_i \geq 0\right\} \geq 1 - \rho_i, i = 1, 2. \tag{3.20}$$

3.3.1 Bernstein-type Inequality Method

To deal with a probabilistic constraint that has a form as (3.20), we adopt the Bernstein-type inequality to construct a convex approximation. Firstly, the following lemma is introduced which serves as a basis [95]:

Lemma 1. *Let $\boldsymbol{e} \in \mathcal{CN}(0, I_n)$, $\boldsymbol{Q} \in \mathbb{H}^n$ and $\boldsymbol{r} \in \mathbb{C}^n$. Then, for any $\varepsilon > 0$, we have that*

$$\Pr\{\boldsymbol{e}^H \boldsymbol{Q}_i \boldsymbol{e} + 2Re\{\boldsymbol{e}^H \boldsymbol{r}_i\} \geq \Gamma(\varepsilon)\} \geq 1 - e^{-\varepsilon}, \tag{3.21}$$

where the function Γ is defined by that:

$$\Gamma(\varepsilon) = \mathrm{Tr}(\boldsymbol{Q}_i) - \sqrt{2\varepsilon} \sqrt{||\boldsymbol{Q}_i||_F^2 + 2||\boldsymbol{r}_i||^2} - \varepsilon \lambda^+(\boldsymbol{Q}_i), \tag{3.22}$$

with $\lambda^+(\boldsymbol{Q}_i) = \max\{\lambda_{max}(-\boldsymbol{Q}_i), 0\}$, and λ_{max} denotes the corresponding maximum eigenvalue of $-\boldsymbol{Q}_i$.

The above inequality is the well-known Bernstein-type inequality, which can also be expressed by the inverse mapping Γ^{-1} as follows due to the monotonically decreasing characteristic of $\Gamma(\varepsilon)$:

$$\Pr\left\{\boldsymbol{e}^H\boldsymbol{Q}_i\boldsymbol{e}+2\mathrm{Re}\{\boldsymbol{e}^H\boldsymbol{r}_i\}+s_i\geq 0\right\}\geq 1-e^{-\Gamma^{-1}(-s_i)}. \tag{3.23}$$

Compare Eq.(3.23) with the reformulated inequality (3.20), it is easy to find that the inequality (3.20) can be satisfied if we replace $e^{-\Gamma^{-1}(-s_i)}$ with ρ_i on the condition that $e^{-\Gamma^{-1}(-s_i)}\leq\rho_i$ holds. By adopting the Bernstain-type inequality and using the monotonically decreasing characteristic of Γ, we can obtain that

$$\mathrm{Tr}(\boldsymbol{Q}_i)+\ln(\rho_i)\lambda^+(\boldsymbol{Q}_i)+s_i-\sqrt{-2\ln(\rho_i)}\sqrt{||\boldsymbol{Q}_i||_F^2+2||\boldsymbol{r}_i||^2}\geq 0. \tag{3.24}$$

Furthermore, by introducing two slack variables $\Lambda_1\in\mathbb{R}$ and $\Lambda_2\in\mathbb{R}$, (3.24) can be reformulated as the following convex conic inequalities:

$$\mathrm{Tr}(\boldsymbol{Q}_i)-\sqrt{-2\ln(\rho_i)}\Lambda_1+\ln(\rho_i)\Lambda_2+s_i\geq 0, \tag{3.25a}$$

$$\sqrt{||\boldsymbol{Q}||_F^2+2||\boldsymbol{r}_i||^2}\leq\Lambda_1, \tag{3.25b}$$

$$\Lambda_2\mathrm{I}_{N_t}+\boldsymbol{Q}_i\succeq\boldsymbol{0}, \tag{3.25c}$$

$$\Lambda_2\geq 0. \tag{3.25d}$$

Therefore, one can note that the probabilistic inequality (3.20) is transformed into efficiently computable convex restrictions as (3.25a)-(3.25d).

Finally, by applying the SDR approach and Bernstein-type method, (3.13) is

reformulated as

$$\max_{\beta,\boldsymbol{W}_1,\boldsymbol{W}_2} \quad (1-\beta)\mathrm{Tr}(\boldsymbol{H}_1\boldsymbol{W}_1) \tag{3.26a}$$

$$s.t. \quad \mathrm{Tr}(\boldsymbol{Q}_1) - \sqrt{-2\ln(\rho_1)}t_1 + \ln(\rho_1)t_2 + s_1 \geq 0, \tag{3.26b}$$

$$\sqrt{||\boldsymbol{Q}_1||_F^2 + 2||\boldsymbol{r}_1||^2} \leq \Lambda_1, \tag{3.26c}$$

$$\Lambda_2 \mathrm{I}_n + \boldsymbol{Q}_1 \succeq \boldsymbol{0}, \tag{3.26d}$$

$$\mathrm{Tr}(\boldsymbol{Q}_2) - \sqrt{-2\ln(\rho_2)}\Lambda_3 + \ln(\rho_2)\Lambda_4 + s_2 \geq 0, \tag{3.26e}$$

$$\sqrt{||\boldsymbol{Q}_2||_F^2 + 2||\boldsymbol{r}_2||^2} \leq \Lambda_3, \tag{3.26f}$$

$$\Lambda_4 \mathrm{I}_n + \boldsymbol{Q}_2 \succeq \boldsymbol{0}, \tag{3.26g}$$

$$\Lambda_2 \geq 0, \Lambda_4 \geq 0, \tag{3.26h}$$

$$\theta\mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W}_1) \leq \mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W}_2) - \theta, \tag{3.26i}$$

$$\mathrm{Tr}(\boldsymbol{H}_1\boldsymbol{W}_2) \geq \mathrm{Tr}(\boldsymbol{H}_1\boldsymbol{W}_1), \tag{3.26j}$$

$$\mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W}_2) \geq \mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W}_1), \tag{3.26k}$$

$$\mathrm{Tr}(\boldsymbol{W}_1) + \mathrm{Tr}(\boldsymbol{W}_2) \leq P_{\max}, \tag{3.26l}$$

$$0 \leq \beta \leq 1, \tag{3.26m}$$

where Q_i, r_i and $s_i, i = 1, 2$, are defined as (3.19a)-(3.19f).

Remark 1. *The optimal solution to problem (3.26) can be found through two-dimensional exhaustive search of variables β and θ.*

However, the complexity of two-dimensional exhaustive search is too high, which motivates us to find a low-complexity suboptimal solution based on SCA and arithmetic geometric mean (AGM) [96].

3.3.2 SCA-based Transformation

In this subsection, before we solve the formulated problem (3.26), we first transform it into a convex program. By applying epigraph reformulation and introducing two auxiliary variables μ an ν, the objective function (3.26a) can be recast as:

$$\max \quad \mathrm{u} \tag{3.27a}$$

$$s.t. \quad \nu^2 \geq \mu, \tag{3.27b}$$

$$\begin{bmatrix} 1-\beta & \nu \\ \nu & \mathrm{Tr}(\mathrm{H}_1\mathrm{W}_1) \end{bmatrix} \succeq 0. \tag{3.27c}$$

Hence, (3.26a) is converted into a linear objective function (3.27a), a convex linear matrix inequality (LMI) (3.27c) and a nonconvex quadratic inequality (3.27b) which can be then approximated by the SCA method. To approximate (3.27b), a convex lower bound for ν^2 needs to be obtained by applying first-order Taylor approximation as below:

$$\nu^2 \geq 2\nu^{(n)}\nu - (\nu^{(n)})^2, \tag{3.28}$$

where $\nu^{(n)}$ denotes the value of variable ν at the n-th iteration. By replacing ν^2 with the inequality (3.28), (3.27b) can be approximated by a stringent constraint given as

$$2\nu^{(n)}\nu - (\nu^{(n)})^2 \geq \mu. \tag{3.29}$$

In addition, by applying AGM method, the constraint (3.26i) can be approximated using the following convex function:

$$\left(a_1^{(n)}\theta\right)^2 + \left(\mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W}_1)/a_1^{(n)}\right)^2 \leq 2\mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W}_2) - 2\theta, \tag{3.30}$$

where the setting of $a_1^{(n)}$ can be given by

$$a_1^{(n)} = \sqrt{(\mathrm{Tr}(\boldsymbol{H}_2^H\boldsymbol{W}_1)^{(n-1)}/\theta^{(n-1)}}. \tag{3.31}$$

Now the remaining problem lies in (3.26e), as the formation of s_2 is nonconvex. By introducing a slack variable ξ, s_2 can be reformulated as

$$s_2 = \tilde{\boldsymbol{h}}_1^H(\boldsymbol{W}_1 + \boldsymbol{W}_2)\tilde{\boldsymbol{h}}_1 - \frac{\gamma}{\vartheta\beta|g|^2} + \xi, \tag{3.32a}$$

$$\left(a_2^{(n)}\beta\right)^2 + \left(\xi/a_2^{(n)}\right)^2 \leq \frac{2\theta}{\vartheta|g|^2}, \tag{3.32b}$$

where $a_2^{(n)} = \sqrt{\xi^{(n-1)}/\beta^{(n-1)}}$. Here, (3.32b) is obtained with the AGM-inequality method and the transformation process is omitted for simplicity.

As a result, after applying the proposed approximation methods, the original problem (3.26) can be transformed to a convex program. During the n-th iteration, the following convex optimization problem needs to be solved:

$$\max_{\mu,\nu,\beta,\boldsymbol{W}_1,\boldsymbol{W}_2} \quad \mu \tag{3.33a}$$

$$s.t. \quad 2\nu^{(n)}\nu - (\nu^{(n)})^2 \geq \mu, \tag{3.33b}$$

$$\begin{bmatrix} 1-\beta & \nu \\ \nu & \mathrm{Tr}(\boldsymbol{H}_1\boldsymbol{W}_1) \end{bmatrix} \succeq 0, \tag{3.33c}$$

$$(3.26b),\ (3.26c),\ (3.26d),\ (3.26e),\ (3.26f), \tag{3.33d}$$

$$(3.26g),\ (3.26h),\ (3.26j),\ (3.26k),\ (3.26l), \tag{3.33e}$$

$$(3.27c),\ (3.29),\ (3.30),\ (3.32a),\ (3.32b). \tag{3.33f}$$

Table 3.1: SCA-based Method to Solve Eq. (3.33)

Input $\mu_0 = 0.001$, $\nu_0 = 0.01$ $n = 0$, $\eta = 1$ and the tolerance $\epsilon = 10^{-3}$.
While $\Delta \geq \epsilon$
Set $\mu^n = \mu^{n-1}$;
Update $\Delta = \mu^n - \mu^{n-1}$;
Update $n = n + 1$;
End While
Output $\boldsymbol{W}_1^n$ and $\boldsymbol{W}_2^n$.

Finally, to solve the problem (3.33), we provide the SCA-based iterative algorithm, outlined in Table 3.1.

To prove the effectiveness of the proposed SCA-based iterative algorithm, we provide the following proposition.

Proposition 1. *The proposed SCA-based algorithm can converge to a Karush-Kuhn-Tucker (KKT) point of problem (3.26) whenever problem (3.33) is feasible.*

Proof: The convergence of SCA method will be proved in Proposition 3 in the next section.

Table 3.2: Ratio of Rank-one Solutions

$P_{\max}$	10 dB	20 dB	30 dB
Ratio	$\frac{297}{297}$	$\frac{859}{876}$	$\frac{954}{971}$

It is noted that both problem (3.26) and (3.33) are formulated by dropping the rank-one constraint. We verify the rank-one characteristic of problem (3.33) via simulations by setting $\gamma = 1$ Mbps, and $\rho = 0.1$. The solution is declared as rank-one if the following condition holds:

$$\frac{\lambda_{max}(\boldsymbol{W}_i)}{\mathrm{Tr}(\boldsymbol{W}_i)} = 1, i = 1, 2. \tag{3.34}$$

We iteratively solve the optimization problems for 1,000 times based on empirical knowledge, since we believe that 1,000 times are large enough to average the channel realizations and acquire fair number of feasible solutions. Consequently, the value of the ratio does not change significantly. As can be seen from Table 3.2, in the ratio column, the denominator represents the number of feasible points while the numerator denotes the amount of rank-one solutions. The probability of rank-one solutions is

higher than 98%, which means the rank-one solutions are *usually* obtained. Hence, we can conclude that the solutions of problem (3.33) guarantee that the rank-one constraints can be satisfied with a high probability, which provides a tight upper bound on the optimal values of the original problem formulation.

3.4 Worst-case Based Optimization

Apart from the outage-based channel uncertainty model, in this section the channel uncertainties are modeled based on the worst-case scenario. Firstly, let us discuss the channel mismatches. For worst-case based optimization, the channel mismatches are assumed to lie in the bounded sets $\{\tilde{\boldsymbol{e}}_{h_1} : ||\tilde{\boldsymbol{e}}_{h_1}||^2 \leq \epsilon_{h_1}^2\}$ and $\{\tilde{\boldsymbol{e}}_{h_2} : ||\tilde{\boldsymbol{e}}_{h_2}||^2 \leq \epsilon_{h_2}^2\}$, where $\epsilon_{h_1}^2$ and $\epsilon_{h_2}^2$ are known constants that model the channel errors.

Then, our objective is to maximize the SNR of strong user while guaranteeing the correct signal decoding of the weak user for the channel mismatches, i.e., $\tilde{\boldsymbol{e}}_{h_1}$, $\tilde{\boldsymbol{e}}_{h_2}$, bounded in the known sets. Hence, the following robust beamforming design problem can be formulated:

$$\max_{\beta,\boldsymbol{w}_1,\boldsymbol{w}_2} \min_{\tilde{\boldsymbol{e}}_h \in \epsilon_h} \quad (1-\beta)|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2 \tag{3.35a}$$

$$s.t. \qquad \frac{(1-\beta)|\boldsymbol{h}_1^H \boldsymbol{w}_2|^2}{(1-\beta)|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2 + 1} \geq \gamma, \tag{3.35b}$$

$$\frac{|\boldsymbol{h}_2^H \boldsymbol{w}_2|^2}{1+|\boldsymbol{h}_2^H \boldsymbol{w}_1|^2} + \vartheta\beta|g|^2(|\boldsymbol{h}_1^H \boldsymbol{w}_1|^2 + |\boldsymbol{h}_1^H \boldsymbol{w}_2|^2) \geq \gamma, \tag{3.35c}$$

$$|\boldsymbol{h}_1^H \boldsymbol{w}_2|^2 \geq |\boldsymbol{h}_1^H \boldsymbol{w}_1|^2, \tag{3.35d}$$

$$|\boldsymbol{h}_2^H \boldsymbol{w}_2|^2 \geq |\boldsymbol{h}_2^H \boldsymbol{w}_1|^2, \tag{3.35e}$$

$$||\boldsymbol{w}_1||_2^2 + ||\boldsymbol{w}_2||_2^2 \leq P_{\max}, \tag{3.35f}$$

$$0 \leq \beta \leq 1, \tag{3.35g}$$

where γ is the target SINR of user 2 and $P_{\max}$ is the maximum available power at the BS. It can be easily verified that problem (3.35) is nonconvex. This is not only due to the quadratic terms of the objective and constraints, but also for the hidden inner minimization constraint over $\tilde{\boldsymbol{e}}_h$.

Similarly, the first step is also to replace the beamforming vector $\boldsymbol{w}_i$ with semidefinite positive matrices $\boldsymbol{W}_i$, i.e,

$$\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H, i = 1, 2. \tag{3.36}$$

Then, we decompose (3.35a) and (3.35b) by introducing several auxiliary variables to explore its hidden convexity. By denoting that $\boldsymbol{W} = \boldsymbol{W}_1 + \boldsymbol{W}_2$, the original problem (3.35) can be reformulated as

$$\max_{\beta, \boldsymbol{W}_1, \boldsymbol{W}_2} \min_{\tilde{\boldsymbol{e}}_h \in \epsilon_h} \quad \mu \tag{3.37a}$$

$$s.t. \quad (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1})^H \boldsymbol{W}_1 (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1}) \geq \frac{\mu}{1-\beta}, \tag{3.37b}$$

$$(\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1})^H \boldsymbol{W}_2 (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1}) \geq \frac{\gamma(\mu+1)}{1-\beta}, \tag{3.37c}$$

$$(\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1})^H \boldsymbol{W} (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1}) \geq \frac{\gamma - t}{\vartheta\beta|g|^2}, \tag{3.37d}$$

$$\frac{(\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2})^H \boldsymbol{W}_2 (\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2})}{1 + (\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2})^H \boldsymbol{W}_1 (\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2})} \geq \theta, \tag{3.37e}$$

$$(\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1})^H (\boldsymbol{W}_2 - \boldsymbol{W}_1) (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1}) \geq 0, \tag{3.37f}$$

$$(\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2})^H (\boldsymbol{W}_2 - \boldsymbol{W}_1) (\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2}) \geq 0, \tag{3.37g}$$

$$\mathrm{Tr}(\boldsymbol{W}_1) + \mathrm{Tr}(\boldsymbol{W}_2) \leq P_{\max}, \tag{3.37h}$$

$$\boldsymbol{W}_1, \boldsymbol{W}_2 \succeq \boldsymbol{0} \text{ and } (3.35g), \tag{3.37i}$$

where μ and θ are two auxiliary variables that can be respectively interpreted as the SNR of user 1 and $\mathrm{SINR}_2^{(1)}$. Note that problem (3.37) is a relaxed version by dropping the nonconvex rank-one constraints, i.e., $\mathrm{rank}(\boldsymbol{W}_i) = 1$, $i = 1, 2$. The advantage of this relaxation lies in that the transformed inequalities are linear to $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$. Although the rank-one constraints are dropped, we provide the following proposition to show that the relaxed problem (3.37) can still achieve an optimal solution which satisfies the rank-one constraints.

Proposition 2. *There is always an optimal solution* $(\boldsymbol{W}_1{}^*, \boldsymbol{W}_2{}^*)$ *to problem (3.37) with* $\mathrm{rank}(\boldsymbol{W}_i^*) = 1$, $i = 1, 2$, *whenever it is feasible.*

Proof: The proof is provided in Appendix A.

Further, we note that in problem (3.37), only constraints (3.37b)-(3.37g) are nonconvex. To deal with the inner minimization problem for the bounded error, (3.37b) can be first expressed as follows by introducing the inner bounded sets of channel mismatches:

$$\tilde{\boldsymbol{e}}_{h_1}^H \boldsymbol{W}_1 \tilde{\boldsymbol{e}}_{h_1} + 2Re(\tilde{\boldsymbol{h}}_1^H \boldsymbol{W}_1 \tilde{\boldsymbol{h}}_1) + \tilde{\boldsymbol{h}}_1^H \boldsymbol{W}_1 \tilde{\boldsymbol{e}}_1 - \tau_1 \geq 0, \tag{3.38a}$$

$$-\tilde{\boldsymbol{e}}_{h_1}^H \tilde{\boldsymbol{e}}_{h_1} + \epsilon_{h_1}^2 \geq 0. \tag{3.38b}$$

$$\tau_1 \geq \frac{\mu}{1-\beta}. \tag{3.38c}$$

Then, in order to make the problem more tractable to solve, we introduce the S-procedure with Lemma 2 [67].

Lemma 2. *Let $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ be symmetric matrices, $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ be vectors, h_1 and h_2 be real numbers, then the following implication*

$$\boldsymbol{x}^T\boldsymbol{F}_1\boldsymbol{x} + 2\boldsymbol{x}^T g_1\boldsymbol{x} + h_1 \leq 0, \tag{3.39a}$$

$$\Longrightarrow \boldsymbol{x}^T\boldsymbol{F}_2\boldsymbol{x} + 2\boldsymbol{x}^T g_2\boldsymbol{x} + h_2 \leq 0, \tag{3.39b}$$

holds if and only if there exists a nonnegative number $\lambda \geq 0$ such that

$$\begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{g}_1 \\ \boldsymbol{g}_1^T & h_1 \end{bmatrix} \succeq \lambda \begin{bmatrix} \boldsymbol{F}_2 & \boldsymbol{g}_2 \\ \boldsymbol{g}_2^T & h_2 \end{bmatrix}, \tag{3.40}$$

provided that there exists a point $\hat{\boldsymbol{x}}$ with $\hat{\boldsymbol{x}}^T\boldsymbol{F}_1\hat{\boldsymbol{x}} + 2\hat{\boldsymbol{x}}^T g_1\boldsymbol{x} + h_1 \leq 0$.

According to Lemma 2, we note that both inequalities, i.e., (3.38a) and (3.38b), can be satisfied with a proper $\tilde{\boldsymbol{e}}_{h_1}$ if and only if there exists a $\nu_1 \geq 0$ such that

$$\begin{bmatrix} \nu_1\boldsymbol{I}_{N_t} + \boldsymbol{W}_1 & \boldsymbol{W}_1\tilde{\boldsymbol{h}}_1 \\ \tilde{\boldsymbol{h}}_1^H\boldsymbol{W}_1 & \tilde{\boldsymbol{h}}_1^H\boldsymbol{W}_1\tilde{\boldsymbol{h}}_1 - \nu_1{\epsilon_{h_1}}^2 - \tau_1 \end{bmatrix} \succeq 0. \tag{3.41}$$

Notice that the inequality (3.41) is a convex LMI and can be easily implemented with standard convex solvers such as CVX [97]. Therefore, one can note that constraint (3.37b) has been transformed into convex forms.

After applying similar steps and introducing two slack variables τ_2 and τ_3, (3.37c), (3.37d), (3.37f) and (3.37g) can be transformed into the following formulations:

$$\begin{bmatrix} \nu_2\boldsymbol{I}_{N_t} + \boldsymbol{W}_2 & \boldsymbol{W}_2\tilde{\boldsymbol{h}}_1 \\ \tilde{\boldsymbol{h}_1}^H\boldsymbol{W}_2 & \tilde{\boldsymbol{h}}_1^H\boldsymbol{W}_2\tilde{\boldsymbol{h}}_1 - \nu_2{\epsilon_{h_1}}^2 - \tau_2 \end{bmatrix} \succeq 0, \tag{3.42a}$$

$$\begin{bmatrix} \nu_3\boldsymbol{I}_{N_t} + \boldsymbol{W} & \boldsymbol{W}\tilde{\boldsymbol{h}}_1 \\ \tilde{\boldsymbol{h}}_1^H\boldsymbol{W} & \tilde{\boldsymbol{h}}_1^H\boldsymbol{W}\tilde{\boldsymbol{h}}_1 - \nu_3{\epsilon_{h_1}}^2 - \tau_3 \end{bmatrix} \succeq 0, \tag{3.42b}$$

$$\begin{bmatrix} \nu_4\boldsymbol{I}_{N_t} + \boldsymbol{W}_2 - \boldsymbol{W}_1 & (\boldsymbol{W}_2 - \boldsymbol{W}_1)\tilde{\boldsymbol{h}}_1 \\ \tilde{\boldsymbol{h}}_1^H(\boldsymbol{W}_2 - \boldsymbol{W}_1) & \tilde{\boldsymbol{h}}_1^H(\boldsymbol{W}_2 - \boldsymbol{W}_1)\tilde{\boldsymbol{h}}_1 - \nu_4{\epsilon_{h_1}}^2 \end{bmatrix} \succeq 0, \tag{3.42c}$$

$$\begin{bmatrix} \nu_5\boldsymbol{I}_{N_t} + \boldsymbol{W}_2 - \boldsymbol{W}_1 & (\boldsymbol{W}_2 - \boldsymbol{W}_1)\tilde{\boldsymbol{h}}_2 \\ \tilde{\boldsymbol{h}}_2^H(\boldsymbol{W}_2 - \boldsymbol{W}_1) & \tilde{\boldsymbol{h}}_2^H(\boldsymbol{W}_2 - \boldsymbol{W}_1)\tilde{\boldsymbol{h}}_2 - \nu_5{\epsilon_{h_1}}^2 \end{bmatrix} \succeq 0, \tag{3.42d}$$

$$\tau_2 \geq \frac{\gamma(\mu+1)}{1-\beta}, \tag{3.42e}$$

$$\tau_3 \geq \frac{\gamma - t}{\vartheta\beta|g|^2}, \tag{3.42f}$$

where ν_2, ν_3, ν_4 and ν_5 are nonnegative variables and the above inequalities are all convex LMIs.

To deal with (3.37e), we introduce some auxiliary variables and (3.37e) can be then written as

$$\left|\boldsymbol{h}_2^H \boldsymbol{w}_2\right|^2 \geq \tau_4, \tag{3.43a}$$

$$\frac{\tau_4 - \theta}{\theta} \geq \tau_5, \tag{3.43b}$$

$$\left|\boldsymbol{h}_2^H \boldsymbol{w}_1\right|^2 \leq \tau_5. \tag{3.43c}$$

For (3.43a) and (3.43c), S-procedure can be used to convert it into LMI formulation as follows

$$\begin{bmatrix} \nu_6 \boldsymbol{I}_{N_t} + \boldsymbol{W}_2 & \boldsymbol{W}_2 \tilde{\boldsymbol{h}}_2 \\ \tilde{\boldsymbol{h}}_2^H \boldsymbol{W}_2 & \tilde{\boldsymbol{h}}_2^H \boldsymbol{W} \tilde{\boldsymbol{h}}_2 - \nu_4 \epsilon_{h_2}^2 - \tau_4 \end{bmatrix} \succeq 0, \tag{3.44a}$$

$$\begin{bmatrix} \nu_7 \boldsymbol{I}_{N_t} - \boldsymbol{W}_1 & -\boldsymbol{W}_1 \tilde{\boldsymbol{h}}_2 \\ -\tilde{\boldsymbol{h}}_2^H \boldsymbol{W}_1 & -\tilde{\boldsymbol{h}}_2^H \boldsymbol{W}_1 \tilde{\boldsymbol{h}}_2 - \nu_5 \epsilon_{h_2}^2 + \tau_5 \end{bmatrix} \succeq 0. \tag{3.44b}$$

Remark 2. *Through the above transformation, the optimal solution can be found through two-dimensional exhaustive search of variables β and θ.*

To reduce the complexity introduced by the exhaustive search method, (3.38c), (3.42e) and (3.42f) can be reformulated as follows:

$$\begin{bmatrix} \tau_1 & \iota_1 \\ \iota_1 & 1 - \beta \end{bmatrix} \succeq 0, \tag{3.45a}$$

$$\iota_1^2 \geq \mu, \tag{3.45b}$$

$$\begin{bmatrix} \tau_2 & \iota_2 \\ \iota_2 & 1 - \beta \end{bmatrix} \succeq 0, \tag{3.45c}$$

$$\iota_2^2 \geq \gamma(\mu + 1), \tag{3.45d}$$

$$\begin{bmatrix} \tau_3 & \iota_3 \\ \iota_3 & \vartheta \beta |g|^2 \end{bmatrix} \succeq 0, \tag{3.45e}$$

$$\iota_3^2 \geq \gamma - \theta, \tag{3.45f}$$

which consist of three LMIs (3.45a), (3.45c), and (3.45e), and three nonconvex quadratic inequalities (3.45b), (3.45d) and (3.45f) that need to be further transformed.

Then, the SCA method can be applied to iteratively approximate (3.45b), (3.45d) and (3.45f) by performing the first-order Taylor approximation as below

$$2\iota_1^{(n)} \iota_1 - (\iota_1^{(n)})^2 \geq \mu, \tag{3.46a}$$

$$2\iota_2^{(n)} \iota_2 - (\iota_2^{(n)})^2 \geq \gamma(\mu + 1), \tag{3.46b}$$

Table 3.3: SCA-based Method to Solve Eq. (3.49)

Given randomly generated feasible solution $\Lambda^{(0)}$
n=0;
Repeat
Update $\Delta^{(n)}$ by solving problem (3.49);
Set $n = n + 1$;
Until convergence or required number of iterations.

$$2\iota_3^{(n)}\iota_3 - (\iota_3^{(n)})^2 \geq \gamma - \theta, \tag{3.46c}$$

where $\iota_i^{(n)}, i = 1, 2, 3$, denotes the variable value of ι_i at the n-th iteration. As a result, the original nonconvex constraints can be approximated by SCA-based expressions and LMIs. The equivalence is guaranteed since (3.46a), (3.46b) and (3.46c) must hold with equality at optimum.

With respect to (3.43b), we apply the AGM inequality-based method to get its convex approximation that can be represented as:

$$(a_3^{(n)}\tau_5)^2 + (\theta/(a_3^{(n)})^2 \leq 2\tau_4 - 2\theta, \tag{3.47}$$

where $a_3^{(n)}$ represents the value of a at the n-th iteration and can be calculated as

$$a_3^{(n)} = \sqrt{\frac{\theta^{(n-1)}}{\tau_5^{n-1}}}. \tag{3.48}$$

As a result, after applying the proposed approximation methods, the original problem (3.35) can be transformed into convex program. At the n-th iteration, the following optimization problem needs to be solved:

$$\max_{\mu,\theta,\beta,\boldsymbol{w_1},\boldsymbol{w_2}} \quad \mu \tag{3.49a}$$

$$s.t. \quad \nu_1, \nu_2, \nu_3, \nu_4, \nu_5, \nu_6, \nu_7 \geq 0, \tag{3.49b}$$

$$(3.41),\ (3.42a),\ (3.42b),\ (3.42c), \tag{3.49c}$$

$$(3.42d)\ ,(3.44a),\ (3.44b),\ (3.45a),\ (3.45c), \tag{3.49d}$$

$$(3.45e),\ (3.46a),\ (3.46b),\ (3.46c), \tag{3.49e}$$

$$(3.47),\ (3.37h) \text{ and } (3.37i). \tag{3.49f}$$

Accordingly, to solve the problem (3.49), the SCA-based iterative algorithm is provided in Table 3.3.

To prove that the above proposed algorithm converges, we provide the following propositions.

Proposition 3. *A non-decreasing sequence of the objective values can be obtained from the proposed beamforming design, i.e.,* $\mu^{n+1} \geq \mu^n$. *Thus, The algorithm provided in Table 3.3 can continuously converge to a stationary point.*

Proof: The proof is provided in Appendix B.

Though proposition 3 demonstrates that the proposed algorithm converges to a stationary point, the global optimality of the problem still cannot be guaranteed due to the nonconvex characteristic of problem (3.37). However, we can verify that the calculated solutions converge to a KKT point under some specific conditions, as summarized in the following proposition.

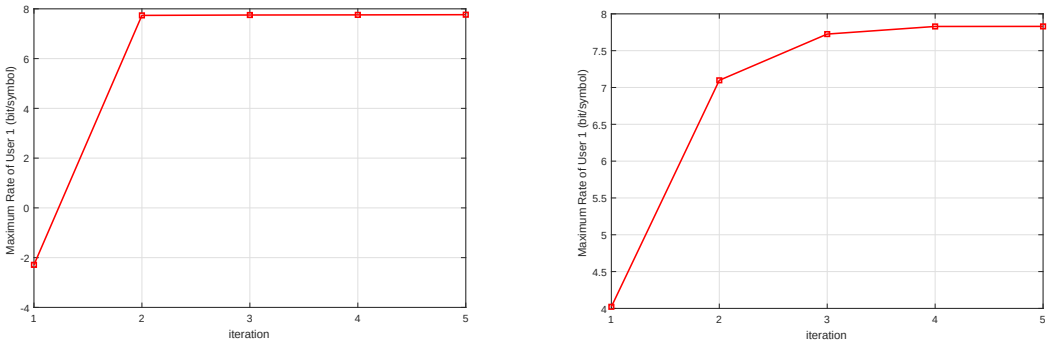
Proposition 4. *The calculated solutions by using SCA-based algorithm continuously converge to a KKT point of problem (3.37) when the iteration number tends to infinity.*

Proof: The proof is provided in Appendix C.

3.5 Numerical Results

In this section, numerical results are provided to evaluate the performance of the proposed algorithms for a cellular SWIPT-enabled robust cooperative NOMA system through Monte Carlo simulations. Firstly the outage-based constrained optimization problem will be examined, followed by the worst-case based optimization problem. For both cases, we iteratively solve the robust optimization problems for 2,000 times. In the following simulations, it is assumed that the BS has two antennas, i.e. $N_t = 2$, while user 1 and user 2 each has one. The estimated channel coefficient can be modeled as $\tilde{\boldsymbol{h}}_k = \boldsymbol{g}_k d_k^{-\frac{\alpha}{2}}$, $k = \{1, 2\}$, where d_k is the distance from the BS to the k-th user, α is the path loss exponent. Here, we assume $\alpha = 2.5$ and $\boldsymbol{g}_k$ follows Rayleigh fading distribution with zero mean and unit variance. We set the energy harvesting efficiency ϑ=0.7, unless otherwise stated. Without loss of generality, the bandwidth is set to be 1 MHz. All the background noise power is assumed to be 1 Watt, and the transmit power is defined in dB relative to the noise power. In addition, for the purpose of system performance comparison, robust noncooperative NOMA, cooperative/noncooperative NOMA with perfect CSI, non-robust cooperative NOMA, and robust TDMA schemes are introduced as follows, which will be then compared with the proposed robust cooperative NOMA model:

- For robust noncooperative NOMA scheme, the BS serves two users simultaneously by performing NOMA and there is no cooperative transmission between



(a) For algorithm of Table 3.1 (b) For algorithm of Table 3.3

Figure 3.2: The convergence procedure of two algorithms.

user 1 and user 2. In addition, the BS only has imperfect CSI of two users and robust beamforming design is applied.

- Regarding cooperative/noncooperative NOMA with perfect CSI scheme, the algorithm proposed in [20] is adopted as the benchmark, where beamforming vectors and PS ratio are acquired for cooperative/noncooperative NOMA with perfect CSI.

- For the non-robust cooperative NOMA scheme, the beamforming vectors and PS ratio for the cooperative NOMA system with perfect CSI is first obtained by using the beamforming design algorithm proposed in [20]. Since we want to check the performance of the non-robust design in a system with channel uncertainties, hence after obtaining $\boldsymbol{W}_i$ and β, if the constraints of problem P4 are not satisfied, the achievable rate of user 1 is 0. Otherwise, the rate of user 1 is computed by introducing the channel mismatches.

- For robust TDMA scheme, the system operates with TDMA mode and the time resource is equally allocated to two users. Furthermore, channel uncertainties exist in the connections between the BS and two users and robust design scheme is applied.

Before we examine the performance of the proposed SWIPT-enabled robust cooperative NOMA system, we first provide insight on the convergence property of the proposed algorithms. It can be observed from Fig. 3.2 that both algorithms converge to the maximum values within about 5 iterations, which proves the effectiveness of the proposed algorithms.

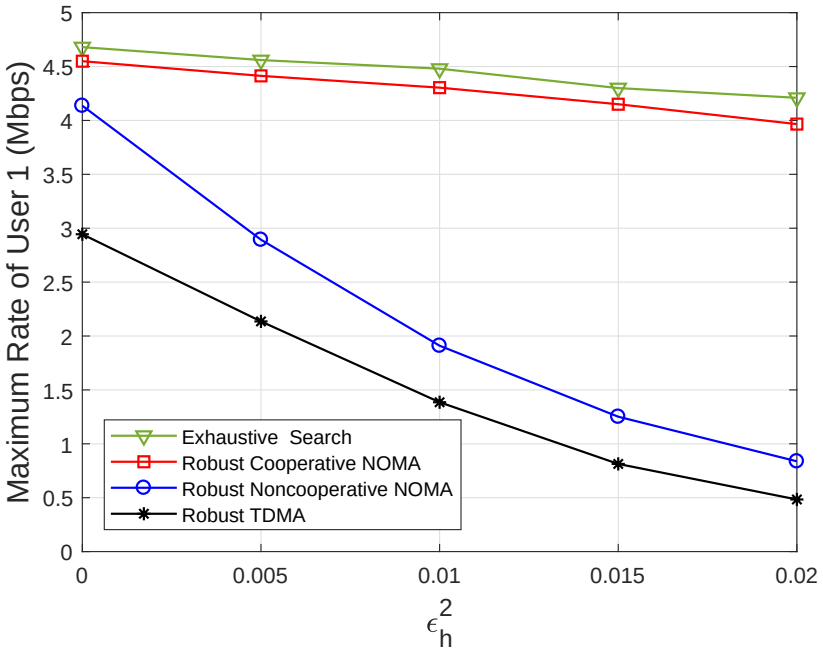


Figure 3.3: Achievable rate of user 1 versus error variance with γ=1, for the outage-based constrained optimization.

3.5.1 Outage-based Constrained Optimization Simulation

In Fig. 3.3, the impact of the error variance is shown for the outage-based constrained optimization problem. Specifically, we set error covariances $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ be the same value as ϵ_h^2, the desired data rate of user 2 as 1 Mbps, the available maximum power at the BS be 20 dB, and the outage is set to be 0.1 which means that the system has a chance of 90% or higher probability to satisfy the SINR requirements. The figure illustrates that the proposed SCA-based algorithm achieves similar system performance as exhaustive search method, but has significantly reduced computational complexity. Furthermore, we can observe that although the maximum achievable data rate of user 1 decreases for all of the schemes when the error variance becomes larger, the benefit of using the proposed SWIPT-enabled robust cooperative NOMA scheme becomes more significant since the gap between the proposed model and the other two schemes becomes larger. Moreover, it can be seen that the two NOMA schemes illustrated in this figure yield better performance than TDMA which shows the advantage of applying NOMA in the outage-based constrained optimization problem.

To investigate the performance of the proposed system model, Fig. 3.4 illustrates the maximum achievable data rate of user 1 versus the available transmission power

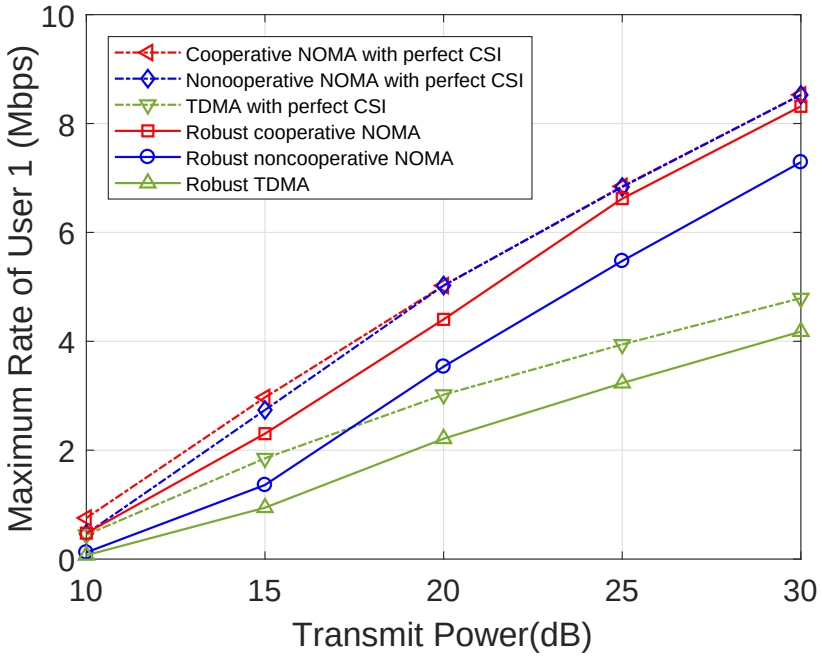


Figure 3.4: Achievable rate of user 1 versus transmit power with γ=1, for the outage-based constrained optimization.

at the BS for the following schemes: the proposed robust cooperative NOMA, cooperative NOMA with perfect CSI, robust noncooperative NOMA, noncooperative NOMA with perfect CSI, classical robust TDMA and TDMA with perfect CSI. This figure is plotted for the outage-based constrained optimization problem. To provide fair comparison results, we set ϑ=1 here. First, it demonstrates that when perfect CSI is available at the BS, cooperative NOMA outperforms noncooperative NOMA in the low power region and achieves the same data rate in the high power regime. Moreover, Fig. 3.4 indicates that the proposed robust cooperative NOMA system always achieves better performance than the robust noncooperative NOMA and TDMA, which means that it is beneficial to adopt the cooperative transmission design for the situations with only imperfect CSI available.

Furthermore, in order to study the relationship between the achievable rate of user 1 and the target rate of user 2, we plot Fig. 3.5 to investigate the rate tradeoff between the two users for robust cooperative NOMA, robust noncooperative NOMA and robust TDMA schemes. This figure is plotted for the outage-based constrained optimization problem. Firstly, we can find that the robust cooperative NOMA yields the largest achievable data rate for user 1 among all three schemes. For example, when the target data rate of user 2 is 1.5 Mbps, the maximum achievable rate of user

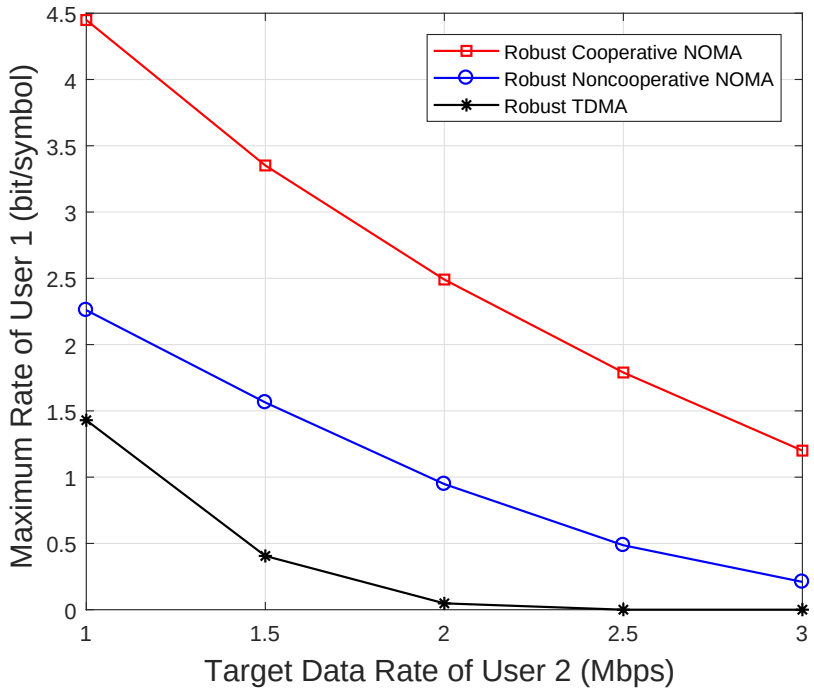


Figure 3.5: Rates tradeoff for the outage-based constrained optimization with ϵ_h^2=0.01 and $P_{\max}$=20 dB .

1 for robust cooperative NOMA is 3.4 Mbps, while for the robust noncooperative NOMA and TDMA schemes, the maximum achievable rate of user 1 are 1.6 Mbps and 0.4 Mbps respectively. Furthermore, from Fig. 3.5, we can also notice that when the target data rate of user 2 increases, the achievable data rate of user 1 decreases for all three schemes as more power is allocated to user 2 in order to satisfy its rate requirements.

3.5.2 Worst-case Based Optimization Simulation

In Fig. 3.6, the impact of the channel mismatch for the worst-case based optimization problem is presented. Here the desired QoS rate of user 2 is set to be 1 Mbps, and the maximum available power is 35 dB. Further, we set $\epsilon_{h_1}^2$ and $\epsilon_{h_2}^2$ to be the same value, denoted as ϵ_h^2. Similar to Fig. 3.3, Fig. 3.6 also shows that for the worst-case based optimization problem, the performance of the proposed SCA-based robust cooperative design is close to that of the exhaustive search method. Furthermore, when perfect CSI is available, i.e., $\epsilon_h^2 = 0$, user 1 achieves almost the same rate for the proposed robust cooperative NOMA, non-robust cooperative NOMA and

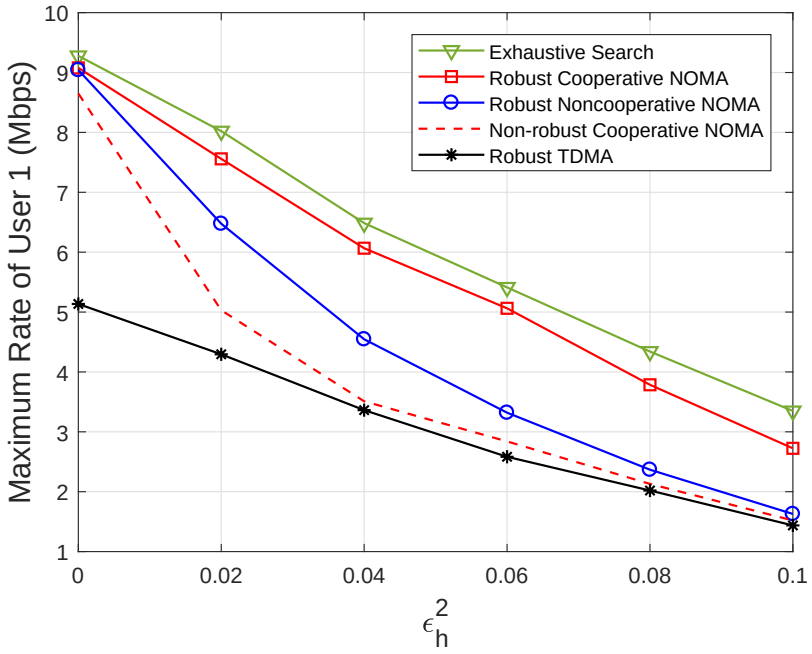


Figure 3.6: Achievable rate of user 1 versus error variance with γ=1 for the worst-case optimization.

robust noncooperative NOMA. However, when there exists channel mismatch, the proposed robust cooperative scheme is more beneficial than the non-robust design. In addition, Fig. 3.6 shows that the robust cooperative NOMA always outperforms the robust noncooperative NOMA scheme. The reason is that, for the robust cooperative scheme, the cooperative phase with perfect CSI can be utilized to improve the weak user's reception reliability under the condition of limited available power at the BS. Furthermore, though the gap between the robust noncooperative NOMA and robust TDMA scheme decreases with the error variance, it can still be observed that NOMA scheme always performs better than TDMA scheme which demonstrates the superiority of NOMA. Specifically, the advantage of NOMA is more significant when the error variance ϵ_h^2 is relatively small.

To study the performance of the proposed robust cooperative NOMA scheme, Fig. 3.7 is plotted to compare different schemes: robust cooperative NOMA, robust noncooperative NOMA, non-robust cooperative NOMA and traditional robust TDMA, for the worst-case optimization. The channel mismatch is set as $\epsilon_{h_1}^2$=$\epsilon_{h_2}^2$=0.05. Firstly, we can notice that the proposed SWIPT-enabled robust cooperative NOMA produces the best performance among all schemes. Especially, the robust design improves the data rate of user 1 greatly, compared to its non-robust counterparts. Furthermore,

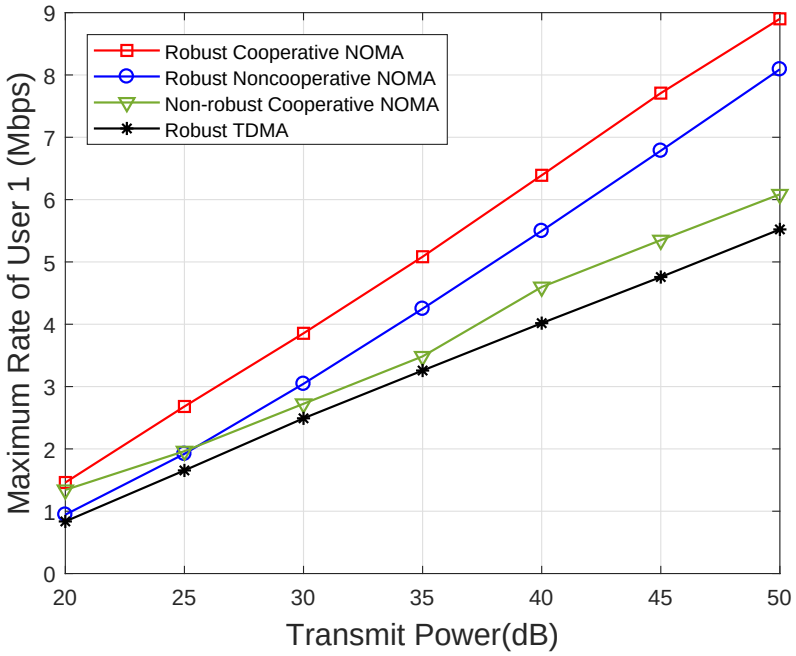


Figure 3.7: Achievable rate of user 1 versus power with γ=1 for the worst-case optimization.

both the proposed scheme and the robust noncooperative NOMA outperforms the traditional robust TDMA scheme which indicates the advantage of NOMA in improving system spectral efficiency when channel uncertainty exists.

Fig. 3.8 shows the influence of user 2's target data rate on the achievable data rate performance of user 1 for the robust cooperative NOMA, robust noncooperative NOMA and robust TDMA schemes. This figure is illustrated based on the worst-case optimization. Firstly, Fig. 3.8 demonstrates that the proposed SWIPT-enabled robust cooperative NOMA achieves higher maximum achievable rate for user 1, compared to the robust noncooperative NOMA and TDMA schemes. In addition, Fig. 3.8 shows that the achievable data rate of user 1 decreases with the increase of target data rate of user 2 for all schemes. This is because when user 2 has a higher target data rate, more power is allocated to satisfy its requirement and as a result, the power available to user 1 becomes less.

3.6 Summary

In this chapter, we have investigated the robust beamforming and PS design to maximize the strong user's data rate for a SWIPT-enabled robust cooperative NOMA

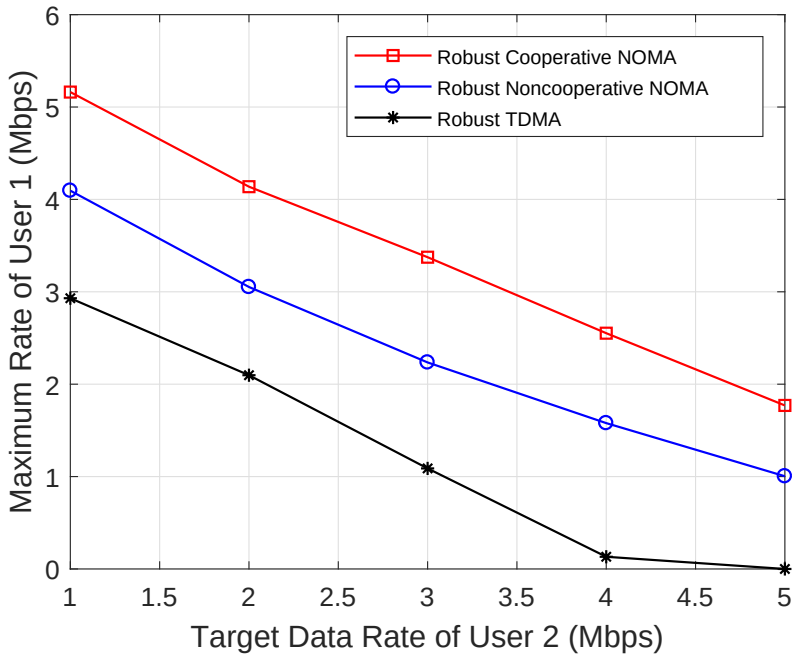


Figure 3.8: Rates tradeoff for the worst-case optimization with ϵ_h^2=0.05 and $P_{\max}$=35 dB .

system. Two kinds of channel uncertainties are considered, which respectively lead to an outage-based SINR constrained optimization and a worst-case based optimization problem. For both cases, the original problem was first transformed into a more tractable form by using SDR technique. Specifically, as to the outage-based SINR constrained optimization problem, the Bernstein-type inequality was applied to convert the probabilistic constraints into manageable and computable approximations that can be globally solved by two-dimensional exhaustive search. An iterative method was further developed to reduce the high complexity. On the other hand, to solve the worst-case based optimization problem, the rank-one optimality of the SDR approach was first proved. Then, by applying the S-procedure, the nonconvex problem was reformulated as convex ones which can be finally solved using the proposed SCA-based algorithm. Simulation results demonstrated the superiority of the proposed SWIPT-enabled cooperation in robust NOMA design over other schemes.

Chapter 4

Computation Efficiency Maximization for Wireless Powered NOMA-Assisted Mobile Edge Computing with User Cooperation

4.1 Introduction

We investigate the application of user cooperation (UC) and NOMA for wireless powered MEC system, in which two single-antenna mobile users first harvest energy from a multi-antenna AP integrated with the MEC server. Considering the non-linear feature of energy harvesting circuits, a more practical non-linear energy harvesting model is adopted. Then, during the computation offloading phase, both users simultaneously offload tasks to the MEC server with the harvested energy, by performing NOMA protocol. To further improve the system performance, UC scheme is carried out, where the near user acts as a DF relay to help the far user offload computation tasks to the AP. To obtain energy efficient communications, our objective in this chapter is to maximize the computation efficiency (i.e., the total computation bits divided by the energy consumption) by jointly designing the transmit beamforming, time and power allocations, which yields to a challenging nonconvex optimization problem. To

The works presented in this Chapter have been submitted to the IEEE Transactions on Communications.

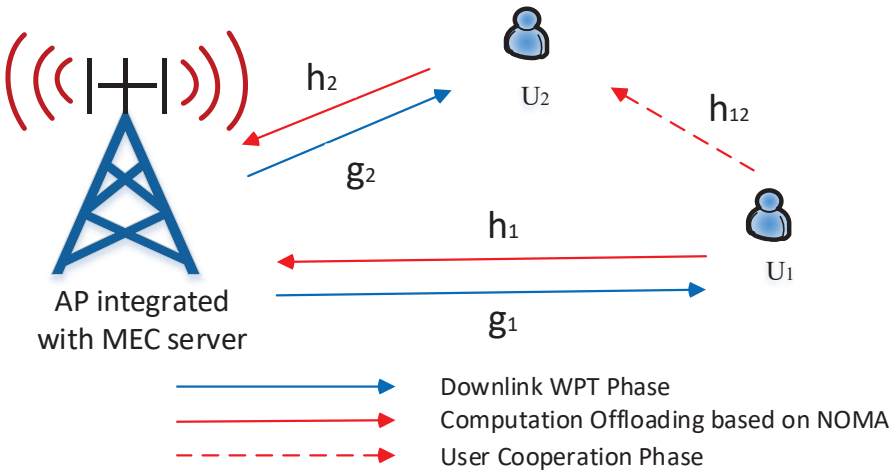


Figure 4.1: NOMA-assisted wireless powered MEC with user cooperation.

deal with it, the original problem is first transformed into a more tractable formulation by applying the SDR technique and then solved by utilizing the SCA approach. Numerical results demonstrate that UC makes a great impact when two users are relatively close, while NOMA makes effect when two users are relatively far. Combining both NOMA and UC, the proposed scheme, named NOMA-UC MEC, yields better system performance than the benchmark schemes.

The remainder of this chapter is organized as follows. In Section 4.2, we present the system model of the wireless powered NOMA-assisted MEC with user cooperation and formulate the computation efficiency optimization problem. Then, A solution approach based on SDR and SCA is developed in Section 4.3 to maximize computation efficiency. Finally, simulation results are given in Section 4.4, followed by conclusions in Section 4.5.

4.2 System Model and Problem Formulation

4.2.1 System Model

Consider a wireless powered MEC system, as shown in Fig. 4.1, which is composed of an N_t-antenna AP integrated with an MEC server and two single-antenna users. Without loss of generality, U_1 is assumed to be far from the AP and U_2 is close to the AP. Let d_1, d_2, and d_{12} denote the distance between U_1 and the AP, U_2 and the AP, and that from U_1 to U_2, respectively. Particularly, $d_1 \geq d_{12}$ is assumed to guarantee that U_2 has an advantage in decoding U_1's message than the AP.

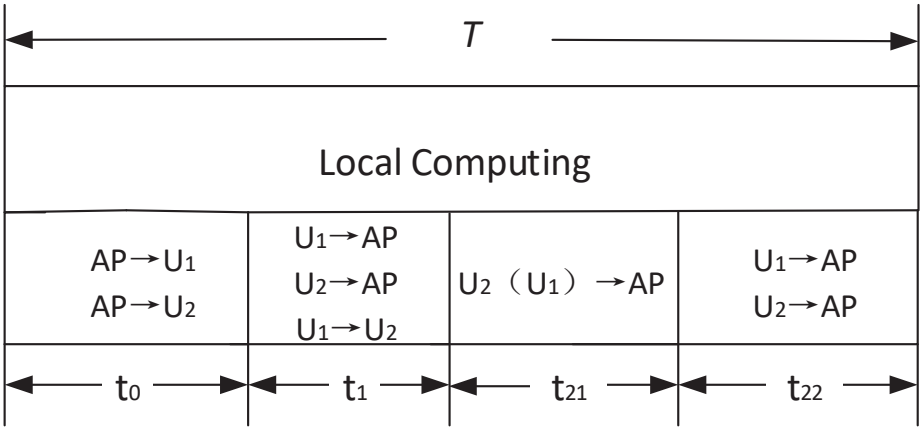


Figure 4.2: Time allocation structure for the wireless powered NOMA-assisted MEC with user cooperation.

The system is assumed to be divided into block slots, and the duration of each block is T seconds. T is chosen to be no more than the user latency requirement and the channel coherence time, hence the channels remain unchanged during one block. It is assumed that perfect CSI is available at the AP [1]. For a given block, two processes, namely the WPT phase and the computation offloading phase, will be performed. The AP first charges the users via employing the RF signal. Then, based on the harvested energy, part of the tasks can be executed by local computing, while the remaining computation tasks can be offloaded to the MEC server for remote execution.

4.2.2 WPT Phase

During the WPT phase, the AP broadcasts wireless energy via downlink transmission and the received signals at both users can be expressed as

$$y_i = \boldsymbol{g}_i^H \boldsymbol{w} x + n_i,\ i = \{1, 2\}, \tag{4.1}$$

where $\boldsymbol{g}_i \in \mathbb{C}^{N_t \times 1}$ is the channel gain from AP to U_i, $i = \{1, 2\}$, $\boldsymbol{w} \in \mathbb{C}^{N_t \times 1}$ denotes the RF energy beamforming vector, x is the RF energy signal with normalized transmit power, i.e., $\mathbb{E}[\|x\|^2] = 1$, and n_i is the AWGN following $n_i \sim \mathcal{CN}(0, \sigma^2)$.

The received RF power at the receiver can be denoted as

$$P_i(\boldsymbol{w}) = |\boldsymbol{g}_i^H \boldsymbol{w}|^2, i = \{1, 2\}. \tag{4.2}$$

[1]It is noted that the system performance may degrade when only imperfect CSI is known at the AP. Therefore, the perfect CSI scenario serves as the upper bound. For imperfect CSI case, robust optimization methods proposed in Chapter 3 may be applied, which is left for our future work.

For the considered non-linear energy harvesting model, according to [98, 99], the harvested energy at the users during the WPT phase occupying the time period t_0 can be expressed as

$$E_i = t_0 \left[\frac{\Psi_i}{X_i} - Y_i \right], \ i = \{1, 2\}, \tag{4.3}$$

with

$$\Psi_i = \frac{M_i}{1 + \exp(-a_i(P_i(\boldsymbol{w}) - b_i))}, \tag{4.4a}$$

$$X_i = \frac{\exp(a_i b_i)}{1 + a_i b_i}, \tag{4.4b}$$

$$Y_i = \frac{M_i}{\exp(a_i b_i)}, \tag{4.4c}$$

where M_i, a_i and b_i are constants capturing the non-linear properties of the energy harvesting system. Specifically, M_i denotes the maximum output power of the energy harvesting circuits, while a_i and b_i reflect the hardware phenomena, i.e., the capacitance, the resistance and the circuit sensitivity.

4.2.3 Computation Offloading Phase

The partial offloading case is considered, where the computation task of each user is divided into two parts for remote execution at the AP and local computing, respectively. The time allocation structure of the computation offloading phase is illustrated in Fig. 4.2. To exploit the benefit of UC while maintaining the advantages of NOMA, three slots are included for UC-enabled uplink NOMA transmission. The offloaded information of both users is divided into two segments, where the two segments are transmitted to the AP directly in the first and the third slot for User 2. For User 1, the first segment is transmitted collaboratively to the AP in the first and second slots, and the second segment is transmitted directly to the AP in the third slot. Specifically, during the subsequent period t_1, due to the application of NOMA protocol, U_1 and U_2 offload some input-bits simultaneously with power p_{11} and p_{20}. Then, both the AP and U_2 can decode the signal of U_1, while the AP also needs to decode U_2's information. For information decoding at the AP, the user with the better channel gain is firstly decoded for uplink NOMA, i.e., the AP first detects U_2's message by treating the message of U_1 as noise, and then removes it with SIC to further decode U_1's information. The remaining time is divided into two parts, given as t_{21} and t_{22}. UC is applied during the second period t_{21}, i.e., U_2 acts a DF relay to forward the signal of U_1 to the AP with power p_{21}. In the third slot t_{22}, U_1 and U_2 offload their own input-bits to the AP with power p_{12} and p_{22}.

Combing the observation from both t_1 and t_{22}, the offloaded data size of U_2 can be characterized as

$$\ell_2^{off} \leq t_1 B\log_2(1+\frac{p_{20}|\boldsymbol{h}_2|^2}{I_1+\sigma^2}) + t_{22}B\log_2(1+\frac{p_{22}|\boldsymbol{h}_2|^2}{I_2+\sigma^2}), \tag{4.5}$$

where $\boldsymbol{h}_i, i = \{1,2\}$, denotes the uplink channel gain from the users to the MEC server. Then, $I_1 = p_{11}|\boldsymbol{h}_1|^2$ and $I_2 = p_{12}|\boldsymbol{h}_1|^2$ represent the interference caused by U_1 during t_1 and t_{22}.

For uplink NOMA transmission, to guarantee the correct SIC decoding in a given order and allocate non-trivial data rate to U_2, the following inequality should be satisfied [17]:

$$p_{20}|\boldsymbol{h}_2|^2 \geq p_{11}|\boldsymbol{h}_1|^2, \tag{4.6a}$$

$$p_{22}|\boldsymbol{h}_2|^2 \geq p_{12}|\boldsymbol{h}_1|^2. \tag{4.6b}$$

After removing the interference signal of U_2, the offloaded data size of U_1 can be given as follows

$$\ell_1^{off} = \ell_{1,1} + \ell_{1,2}, \tag{4.7}$$

where $\ell_{1,1}$ represents the offloaded data size of U_1 via UC. Based on [100], $\ell_{1,1}$ is expressed as $\ell_{1,1} \leq \min\{\ell_{1,direct}, \ell_{1,relay}\}$, where $\ell_{1,direct}$ and $\ell_{1,relay}$ are the offloaded data size of U_1 at the AP and U_2, which are given as $\ell_{1,direct} = t_1 B\log_2(1+\frac{p_{11}|\boldsymbol{h}_1|^2}{\sigma^2}) + t_{21}B\log_2(1+\frac{p_{21}|\boldsymbol{h}_2|^2}{\sigma^2})$, $\ell_{1,relay} = t_1 B\log_2(1+\frac{p_{11}|h_{12}|^2}{\sigma^2})$, respectively. Moreover, $\ell_{1,2}$ denotes the offloaded data size during the period t_{22}, which can be expressed as $\ell_{1,2} \leq t_{22}B\log_2(1+\frac{p_{12}|\boldsymbol{h}_1|^2}{\sigma^2})$.

Similar to [101], we assume that the time consumption of two processes, i.e., task execution at the MEC server and MEC server transmits computed results back to users, are negligible. The reason is that, sufficient CPU-capability and energy are assumed to be available at the MEC-integrated AP, and the output data sizes are much smaller compared with that of the input data sizes. Furthermore, U_1's information decoding time at U_2 is also ignored, as it is much smaller compared with the computation offloading time. Therefore, the system latency constraint including the WPT and computation offloading can be given as

$$t_0 + t_1 + t_{21} + t_{22} \leq T. \tag{4.8}$$

During this phase, the consumed energy of U_1 and U_2 can be respectively denoted as

$$E_1^{off} = p_{11}t_1 + p_{12}t_{22}, \tag{4.9a}$$

$$E_2^{off} = p_{20}t_1 + p_{21}t_{21} + p_{22}t_{22}. \tag{4.9b}$$

4.2.4 Local Computing

Furthermore, during the whole block duration T, $\ell_i^{loc}, i = \{1, 2\}$, input-bits are executed by local computing at the users. Similar to [6, 102], identical CPU frequency $f_i = \frac{C_i \ell_i^{loc}}{T}$ is adopted for CPU cycle, where $C_i, i = \{1, 2\}$, denotes the number of required CPU cycles to compute one input-bit locally. f_i is constrained by a maximum CPU frequency $f_{\max}$, which can be equivalently expressed as

$$C_i \ell_i^{loc} \leq T f_{\max}. \tag{4.10}$$

Accordingly, the consumed energy for local computing can be given as

$$E_i^{loc} = \frac{\kappa_i C_i^3 (\ell_i^{loc})^3}{T^2}, \ i = \{1, 2\}, \tag{4.11}$$

where κ_i is a constant denoting the effective capacitance coefficient and the value is dependent on the chirp architecture [7].

Due to the fact that the consumed energy at the users cannot exceed the harvested energy obtained from WPT, we have that

$$E_i^{loc} + E_i^{off} \leq E_i, \ i = \{1, 2\}. \tag{4.12}$$

The computation efficiency is defined as a ratio of the total calculated data bits to the system energy consumption, which can be given as

$$\tilde{\eta} = \frac{\sum_{i=1}^{2} \ell_i^{loc} + \ell_i^{off}}{t_0 |\boldsymbol{w}|^2}. \tag{4.13}$$

Finally, with the aim of obtaining an energy efficient design, we formulate the computation efficiency maximization problem as follows

$$\max_{\boldsymbol{w}, \boldsymbol{t}, \boldsymbol{p}, \boldsymbol{\ell}} \quad \tilde{\eta}, \tag{4.14a}$$

$$s.t. \quad t_0 + t_1 + t_{21} + t_{22} \leq T, \tag{4.14b}$$

$$E_i^{loc} + E_i^{off} \leq E_i, i = 1, 2, \tag{4.14c}$$

$$\ell_i^{loc} + \ell_i^{off} \geq L_i, i = 1, 2, \tag{4.14d}$$

$$|\boldsymbol{w}|^2 \leq P_{\max}, \tag{4.14e}$$

$$p_{20} |\boldsymbol{h}_2|^2 \geq p_{11} |\boldsymbol{h}_1|^2, p_{22} |\boldsymbol{h}_2|^2 \geq p_{12} |\boldsymbol{h}_1|^2, \tag{4.14f}$$

$$\boldsymbol{t} \geq 0, \boldsymbol{p} \geq 0, \boldsymbol{\ell} \geq 0, \tag{4.14g}$$

where $\mathbf{t} = [t_0, t_1, t_{21}, t_{22}]$, $\mathbf{p} = [p_{11}, p_{12}, p_{20}, p_{21}, p_{22}]$, and $\boldsymbol{\ell} = [\ell_1^{off}, \ell_2^{off}, \ell_1^{loc}, \ell_2^{loc}]$ denote the time allocation vector, the power allocation vector and the calculated data size sets for computation offloading and local computing, respectively. Further, constraint (4.14d) denotes the minimum required computing data bits for user i, where $i = \{1, 2\}$. The maximum available power at the AP is limited by (4.14e).

4.3 Solution Approach

It is noted that (4.14) is a nonconvex problem, which cannot be solved directly. The challenge is twofold, i.e., 1) the objective is a fractional function involving transmit beamfoming vector, 2) the expressions of U_2's offloading data size and the adopted non-linear energy harvesting model are complicated. In this section, we first provide the optimal time allocation conditions. Then, we relax the problem by leveraging the SDR approach. For the relaxed problem, the objective function, the energy-limited constraints and $U_2's$ offloading bits ℓ_2^{off} are further converted into convex approximations with SCA.

4.3.1 SCA-based Approach

Firstly, to solve (4.14), the optimal time utilization is obtained with the following Lemma 3.

Lemma 3. *The maximum computation efficiency of (4.14) can be achieved with* $t_0 + t_1 + t_{21} + t_{22} = T$.

Proof: The proof is provided in Appendix D.

To deal with the beamforming vector $\boldsymbol{w}$, SDR technique is applied to transform (4.14) into a more tractable form. Specifically, $\boldsymbol{w}$ is replaced by the semidefinite positive matrix, i.e., $\boldsymbol{W} = \boldsymbol{w}\boldsymbol{w}^H$. The constraint (4.14e) can be then reformulated as follows

$$\mathrm{Tr}(\boldsymbol{W}) \leq P_{\max}, \tag{4.15a}$$

$$\boldsymbol{W} \succeq 0, \tag{4.15b}$$

$$\mathrm{rank}(\boldsymbol{W}) \leq 1. \tag{4.15c}$$

Then, by introducing some slack variables $\boldsymbol{\zeta} = [\zeta_1, \zeta_2]$, $\boldsymbol{\tau} = [\tau_1, \tau_2]$, and several substitution variables, i.e., $\boldsymbol{E} = [E_{11}, E_{12}, E_{20}, E_{21}, E_{22}]$, where $E_{11} \triangleq t_1 p_{11}$, $E_{12} \triangleq t_{22} p_{12}$, $E_{20} \triangleq t_1 p_{20}$, $E_{21} \triangleq t_{21} p_{21}$, and $E_{22} \triangleq t_{22} p_{22}$, (4.14c) can be decoupled into the following constraints

$$E_{11} + E_{12} + \frac{\kappa_1 C_1^3 (\ell_1^{loc})^3}{T^2} \leq \zeta_1, \tag{4.16a}$$

$$(\frac{\zeta_1}{t_0} + Y_1) X_1 \leq \frac{M_1}{1 + \exp(-a_1(\tau_1 - b_1))}, \tag{4.16b}$$

$$\tau_1 \leq \mathrm{Tr}(\boldsymbol{G}_1 \boldsymbol{W}), \tag{4.16c}$$

$$E_{20} + E_{21} + E_{22} + \frac{\kappa_2 C_2^3 (\ell_2^{loc})^3}{T^2} \leq \zeta_2, \tag{4.16d}$$

$$(\frac{\zeta_2}{t_0} + Y_2)X_2 \leq \frac{M_2}{1+\exp(-a_2(\tau_2 - b_2))}, \tag{4.16e}$$

$$\tau_2 \leq \mathrm{Tr}(\boldsymbol{G}_2\boldsymbol{W}), \tag{4.16f}$$

where $\boldsymbol{G}_i \triangleq \boldsymbol{g}_i\boldsymbol{g}_i^H$, $i = \{1, 2\}$. Note that after the reformulations, (4.16b) and (4.16e) are still nonconvex constraints.

Further, (4.6a) and (4.6b) can be reformulated as

$$E_{20}|\boldsymbol{h}_2|^2 \geq E_{11}|\boldsymbol{h}_1|^2, \tag{4.17a}$$

$$E_{22}|\boldsymbol{h}_2|^2 \geq E_{12}|\boldsymbol{h}_1|^2. \tag{4.17b}$$

By further introducing two slack variables and applying the epigraph reformulation, (4.16b) can be reformulated as

$$(\upsilon_1 + Y_1)X_1 \leq \frac{M_1}{1+\exp(-a_1(\tau_1 - b_1))}, \tag{4.18a}$$

$$\begin{bmatrix} \upsilon_1 & \omega_1 \\ \omega_1 & t_0 \end{bmatrix} \succeq 0, \tag{4.18b}$$

$$\omega_1^2 \geq \zeta_1, \tag{4.18c}$$

where (4.18a) is a convex function, (4.18b) is a convex LMI, and the nonconvex function is (4.18c).

Moreover, SCA can be adopted to obtain the convex approximation of (4.18c). Due to the convex feature of ω_1^2, the lower bound approximation can be derived by performing the first-order Taylor approximation:

$$\omega_1^2 \geq 2\omega_1^{(n)}\omega_1 - (\omega_1^{(n)})^2, \tag{4.19}$$

where $\omega_1^{(n)}$ denotes the value of ω_1 during the n-th iteration.

Hence, (4.18c) is transformed into the following inequality:

$$2\omega_1^{(n)}\omega_1 - (\omega_1^{(n)})^2 \geq \zeta_1, \tag{4.20}$$

Similarly, (4.16e) can be approximated as

$$(\upsilon_2 + Y_2)X_2 \leq \frac{M_2}{1+\exp(-a_2(\tau_2 - b_2))}, \tag{4.21a}$$

$$\begin{bmatrix} \upsilon_2 & \omega_2 \\ \omega_2 & t_0 \end{bmatrix} \succeq 0, \tag{4.21b}$$

$$2\omega_2^{(n)}\omega_2 - (\omega_2^{(n)})^2 \geq \zeta_2. \tag{4.21c}$$

Then, by introducing auxiliary variables μ and Υ, the objective function can be approximated as follows

$$\max_{\boldsymbol{w},\boldsymbol{t},\boldsymbol{p},\boldsymbol{\ell}} \quad \mu, \tag{4.22a}$$

$$s.t. \quad \sum_{i=1}^{2} \ell_i^{loc} + \ell_i^{off} \geq \sqrt{\mu\Upsilon}, \tag{4.22b}$$

$$t_0 \mathrm{Tr}(\boldsymbol{W}) \leq \sqrt{\Upsilon}, \tag{4.22c}$$

where the equivalence is guaranteed when (4.22b) and (4.22c) hold with equality at optimum.

It is noted that $\sqrt{\mu\Upsilon}$ is a joint concave function with respect to μ and Υ, which can be approximated by its upper bound as below

$$\sqrt{\mu\Upsilon} \triangleq g(\mu, \Upsilon) \leq g^{'}(\mu, \Upsilon, \mu^{(n)}, \Upsilon^{(n)}), \tag{4.23a}$$

$$g^{'}(\mu, \Upsilon, \mu^{(n)}, \Upsilon^{(n)}) = \sqrt{\mu^{(n)}\Upsilon^{(n)}} + 0.5\sqrt{\frac{\mu^{(n)}}{\Upsilon^{(n)}}}(\Upsilon - \Upsilon^{(n)}) + 0.5\sqrt{\frac{\Upsilon^{(n)}}{\mu^{(n)}}}(\mu - \mu^{(n)}), \tag{4.23b}$$

where $\mu^{(n)}$ and $\Upsilon^{(n)}$ denote the value of variables μ and Υ at the n-th iteration respectively, and $g^{'}(\mu, \Upsilon, \mu^{(n)}, \Upsilon^{(n)})$ represents the first-order Taylor approximation around $(\mu^{(n)}, \Upsilon^{(n)})$.

Accordingly, (4.22b) can be reformulated as

$$\sum_{i=1}^{2} \ell_i^{loc} + \ell_i^{off} \geq g^{'}(\mu, \Upsilon, \mu^{(n)}, \Upsilon^{(n)}). \tag{4.24}$$

For (4.22c), AGM method [103] can be applied to transform it into convex approximation as

$$(\nu^{(n)} t_0)^2 + (\mathrm{Tr}(\boldsymbol{W})/\nu^{(n)})^2 \leq 2\sqrt{\Upsilon}, \tag{4.25}$$

where $\nu^{(n)}$ can be updated as below during the n-th iteration

$$\nu^{(n)} = \sqrt{\mathrm{Tr}(\boldsymbol{W})^{(n-1)}/t_0^{(n-1)}}. \tag{4.26}$$

Moreover, by substituting $\boldsymbol{E}$ into (4.5), it can be reformulated as below

$$\begin{cases} \ell_1^{off} = \ell_{1,1} + \ell_{1,2}, \\ \ell_{1,1} \leq t_1 B \log_2(1 + \frac{E_1|\boldsymbol{h}_1|^2}{t_1\sigma^2}) + t_{21} B \log_2(1 + \frac{E_{21}|\boldsymbol{h}_2|^2}{t_{21}\sigma^2}), \\ \ell_{1,1} \leq t_1 B \log_2(1 + \frac{E_1|h_{12}|^2}{t_1\sigma^2}), \\ \ell_{1,2} \leq t_{22} B \log_2(1 + \frac{E_{12}|h_1|^2}{t_{22}\sigma^2}). \end{cases} \tag{4.27}$$

As a function $f^{'}(x) = \log_2(1 + \frac{x}{\sigma^2})$ is concave, its perspective function $tf^{'}(\frac{x}{t}) = \log_2(1 + \frac{x}{t\sigma^2})$ is also concave. This indicates that the constraints in (4.27) are all convex.

Then, by introducing two slack variables $\ell_{2,1}^{off}$ and $\ell_{2,2}^{off}$, the offloaded data size of U_2 can be recast as

$$\ell_2^{off} \leq \ell_{2,1}^{off} + \ell_{2,2}^{off}, \tag{4.28a}$$

$$\ell_{2,1}^{off} \leq t_1 B \log_2(1 + \frac{E_{20}|\boldsymbol{h}_2|^2}{E_{11}|\boldsymbol{h}_1| + \sigma^2 t_1}), \tag{4.28b}$$

$$\ell_{2,2}^{off} \leq t_{22} B \log_2(1 + \frac{E_{22}|\boldsymbol{h}_2|^2}{E_{12}|\boldsymbol{h}_1| + \sigma^2 t_{22}}), \tag{4.28c}$$

where the optimality can be guaranteed when (4.28b) and (4.28c) hold with equality.

To deal with the nonconvex functions (4.28b) and (4.28c), we introduce the difference of convex (DC) programming [104] as follows:

Lemma 4. *DC programming can be expressed as a difference of convex functions, which is shown as*

$$\min_{x \in \chi} \quad f_0(x) - g_0(x) \tag{4.29a}$$

$$s.t. \quad f_i(x) - g_i(x) \leq 0, \ i = 1, \ldots, m, \tag{4.29b}$$

where $f_i(x)$ and $g_i(x)$ represent continuous convex or quasi-convex functions, χ is a convex set.

Though (4.29) is nonconvex, it has been demonstrated DC programming can be efficiently solved [105, 106].

Therefore, to further transform (4.28b), it can be firstly rewritten as $\ell_{2,1}^{off} \leq m_1(\boldsymbol{E}, t_1) - z_1(\boldsymbol{E}, t_1)$, where $m_1(\boldsymbol{E}, t_1)$ and $z_1(\boldsymbol{E}, t_1)$ are defined as

$$m_1(\boldsymbol{E}, t_1) = t_1 B \log_2(1 + \frac{E_{11}|\boldsymbol{h}_1|^2 + E_{20}|\boldsymbol{h}_2|^2}{\sigma^2 t_1}), \tag{4.30a}$$

$$z_1(\boldsymbol{E}, t_1) = t_1 B \log_2(1 + \frac{E_{11}|\boldsymbol{h}_1|^2}{\sigma^2 t_1}). \tag{4.30b}$$

It is worth noting that both $m_1(\boldsymbol{E}, t_1)$ and $z_1(\boldsymbol{E}, t_1)$ are joint concave functions with respect to $\boldsymbol{E}$ and t_1. Therefore, we can see that $m_1(\boldsymbol{E}, t_1) - z_1(\boldsymbol{E}, t_1)$ is a DC programming function, which can be converted into convex expression with the aid of SCA. As $z_1(\boldsymbol{E}, t_1)$ is a concave function, an upper bound can be given by using its first-Taylor expansion as below:

$$z_1(\boldsymbol{E}, t_1) \leq z_1(\boldsymbol{E}^{(n)}, t_1^{(n)}) + \nabla z_1(\boldsymbol{E}^{(n)})(E_{11} - E_{11}^{(n)}) \quad + \nabla z_1(t_1^{(n)})(t_1 - t_1^{(n)}), \tag{4.31}$$

where $E_{11}^{(n)}$ and $t_1^{(n)}$ represent the values of E_{11} and t_1 at the n-th iteration. $\nabla z_1(\boldsymbol{E}^{(n)})$ and $\nabla z_1(t_1^{(n)})$ denote the gradients of $z_1(\boldsymbol{E}, t_1)$ over E_{11} and t_1, which are expressed as

$$\begin{cases} \nabla z_1(\boldsymbol{E}^{(n)}) = \frac{B t_1^{(n)} |\boldsymbol{h}_1|^2}{(E_{11}^{(n)} |\boldsymbol{h}_1|^2 + \sigma^2 t_1^{(n)}) \ln 2}, \\ \nabla z_1(t_1^{(n)}) = \mathrm{Blog}_2(1 + \frac{E_{11}^{(n)} |\boldsymbol{h}_1|^2}{\sigma^2 t_1^{(n)}}) - \frac{B E_{11}^{(n)} |\boldsymbol{h}_1|^2}{(E_{11}^{(n)} |\boldsymbol{h}_1|^2 + \sigma^2 t_1^{(n)}) \ln 2}. \end{cases} \tag{4.32}$$

As a result, (4.28b) can be reformulated as

$$\ell_{2,1}^{off} \leq m_1(\boldsymbol{E}, t_1) - z_1(\boldsymbol{E}^{(n)}, t_1^{(n)}) - \nabla z_1(\boldsymbol{E}^{(n)})(E_{11} - E_{11}^{(n)}) - \nabla z_1(t_1^{(n)})(t_1 - t_1^{(n)}). \tag{4.33}$$

Furthermore, following a similar procedure, (4.28c) can be then recast as

$$\ell_{2,2}^{off} \leq m_2(\boldsymbol{E}, t_{22}) - z_2(\boldsymbol{E}^{(n)}, t_{22}^{(n)}) - \nabla z_2(\boldsymbol{E}^{(n)})(E_{12} - E_{12}^{(n)}) - \nabla z_2(t_{22}^{(n)})(t_{22} - t_{22}^{(n)}), \tag{4.34}$$

where $m_2(\boldsymbol{E}, t_{22})$, $z_2(\boldsymbol{E}, t_{22})$, $\nabla z_2(\boldsymbol{E}^{(n)})$, and $\nabla z_2(t_{22}^{(n)})$ are defined as

$$\begin{cases} m_2(\boldsymbol{E}, t_{22}) = t_{22} \mathrm{Blog}_2(1 + \frac{E_{12} |\boldsymbol{h}_1|^2 + E_{22} |\boldsymbol{h}_2|^2}{\sigma^2 t_{22}}), \\ z_2(\boldsymbol{E}, t_{22}) = t_{22} \mathrm{Blog}_2(1 + \frac{E_{12} |\boldsymbol{h}_1|^2}{\sigma^2 t_{22}}), \\ \nabla z_2(\boldsymbol{E}^{(n)}) = \frac{B t_{22}^{(n)} |\boldsymbol{h}_1|^2}{(E_{12}^{(n)} |\boldsymbol{h}_1|^2 + \sigma^2 t_{22}^{(n)}) \ln 2}, \\ \nabla z_2(t_{22}^{(n)}) = B \mathrm{log}_2(1 + \frac{E_{12}^{(n)} |\boldsymbol{h}_1|^2}{\sigma^2 t_{22}^{(n)}}) - \frac{B E_{12}^{(n)} |\boldsymbol{h}_1|^2}{(E_{12}^{(n)} |\boldsymbol{h}_1|^2 + \sigma^2 t_{22}^{(n)}) \ln 2}. \end{cases} \tag{4.35}$$

Finally, the original problem (4.14) can be transformed into a convex formulation by dropping the rank-one constraint. During the n-th iteration, the following convex

Table 4.1: Computation Efficiency Maximization Algorithm

Initialize $\boldsymbol{E}^{(0)}$ and $\boldsymbol{t}^{(0)}$, set $n=0$, $\mu^0=0$, $\mu^1=1$, and the tolerance $\epsilon=10^{-3}$.
While $\lvert\mu^{n+1}-\mu^n\rvert \geq \epsilon$
Update $\boldsymbol{E}^{(n)}$ and $\boldsymbol{t}^{(n)}$ by solving (4.36);
Update $\mu^n=\mu^{n-1}$;
Update $n=n+1$;
End While
Output $\boldsymbol{W}^{(n)}$, $\boldsymbol{E}^{(n)}$ and $\boldsymbol{t}^{(n)}$.

optimization problem needs to be solved:

$$\max_{\boldsymbol{W},\boldsymbol{\omega},\boldsymbol{\upsilon},\boldsymbol{t},\boldsymbol{E},\mu} \quad \mu, \tag{4.36a}$$

$$s.t. \quad (4.16a),\ (4.16c),\ (4.16d),\ (4.16f),\ (4.17a), \tag{4.36b}$$

$$(4.17b),\ (4.18a),\ (4.18b),\ (4.20),\ (4.21a), \tag{4.36c}$$

$$(4.21b),\ (4.21c),\ (4.24),\ (4.25),\ (4.27), \tag{4.36d}$$

$$(4.28a),\ (4.33),\ (4.34), \tag{4.36e}$$

$$\mathrm{Tr}(\boldsymbol{W}) \leq P_{\max}, \tag{4.36f}$$

$$t_0+t_1+t_{21}+t_{22}=T, \tag{4.36g}$$

$$\boldsymbol{t} \succeq 0, \boldsymbol{E} \succeq 0, \boldsymbol{W} \succeq 0, \boldsymbol{\ell} \succeq 0, \tag{4.36h}$$

where $\boldsymbol{\omega}=[\omega_1,\omega_2], \boldsymbol{\upsilon}=[\upsilon_1,\upsilon_2]$.

Therefore, the proposed computation efficiency maximization algorithm is provided in Table 4.1 to outline the detailed process to solve problem (4.36).

Note that the nonconvex rank-one constraint, i.e., $\mathrm{rank}(\boldsymbol{W}) \leq 1$ is dropped for (4.36). To demonstrate the equivalence between (4.36) and (4.14), we provide the following theorem.

Proposition 5. *There is always an optimal solution* $\boldsymbol{W}^*$ *to* (4.36)*, whenever the problem is feasible.*

Proof: The proof is provided in Appendix E.

Moreover, to prove the proposed algorithm converges, we have the following theorem.

Proposition 6. *Algorithm provided in Table 4.1 produces a non-decreasing sequence of the objective values, i.e.,* $\mu^{(n+1)} \geq \mu^{(n)}$*, which indicates the convergence of the proposed algorithm.*

Proof: The proof is provided in Appendix F.

Proposition 7. *The proposed algorithm continuously converges to a KKT point of problem (4.14) whenever problem (4.36) is feasible.*

Proof: The proof is provided in Appendix G.

4.3.2 Complexity Analysis

Note that the computational complexity of the proposed computation efficiency maximization algorithm consists of two loops: the outer iteration loop and the inner loop to solve problem (4.36). Specifically, denote the maximum iteration number of the algorithm provided in Table 4.1 as L_{max}, while the complexity of the interior point method to solve (4.36) is proportional to $O(r^{3.5}\delta)$ [107], where r denotes the number of total optimization variables, and δ is the number of bits needed to represent the entries in the optimization problem. In summary, the whole complexity of the proposed algorithm is $O(L_{max}r^{3.5}\delta)$, where r is the total number of variables $(\boldsymbol{W}, \boldsymbol{t}, \boldsymbol{E}, \boldsymbol{\zeta}, \boldsymbol{\tau}, \boldsymbol{\omega}, \boldsymbol{\upsilon}, \mu, \Upsilon)$ to solve (4.36).

4.4 Simulation Results

Numerical results are provided to evaluate the performance of the proposed scheme. The parameters are set as follows, unless otherwise stated. It is assumed that the AP is situated at the edge of the network with a coordinate of (0, 5 m). The two users are randomly distributed in a 8 m×10 m coverage region. Furthermore, we set the bandwidth $B = 1$ MHz, the capacitance coefficient $\kappa_i = 10^{-28}$, the maximum CPU frequency $f_{\max} = 2$ GHz, and the noise power $\sigma^2 = 10^{-9}$ W. The required CPU cycles to locally compute one bit for two users are given as 1000 [108]. Without loss of generality, the channel reciprocity is assumed to hold for the downlink and uplink, i.e., $\boldsymbol{h}_i = \boldsymbol{g}_i, \{i = 1, 2\}$, and the channel coefficient is modeled as $\boldsymbol{h}_i = 10^{-1.5}\tilde{\boldsymbol{h}}_i d_i^{-\frac{\alpha}{2}}, i = \{1, 2\}$, where $\alpha = 3$ denotes the path loss exponent, and $\tilde{\boldsymbol{h}}$ follows the Rayleigh fading distribution. Without loss of generality, we set $L_1 = L_2 = L$, which indicates that two users have the same computation rate requirement. For the non-linear energy harvesting model, the parameters are set as $M_1 = M_2 = 24$ mW, $a_1 = a_2 = 150$ and $b_1 = b_2 = 0.024$.

For simplicity, the proposed scheme is referred to as "NOMA-UC MEC" in the following figures. To provide a comprehensive study, we also simulate the baseline schemes, which are described as follows:

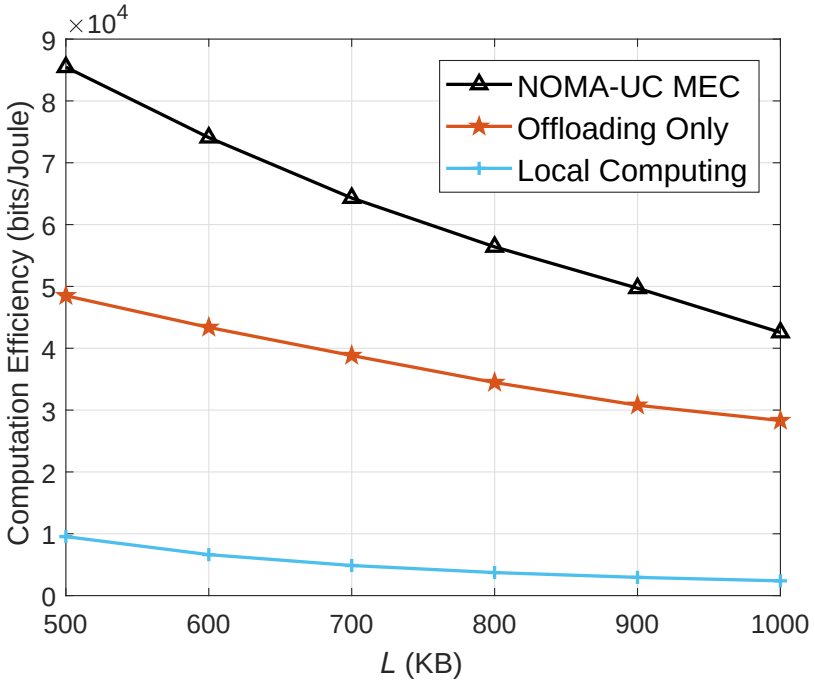


Figure 4.3: Maximum computation efficiency vs. L.

- UC-MEC represents the wireless powered UC-enabled MEC strategy, and TDMA is adopted as the MA scheme.
- For NOMA-MEC, it denotes the wireless powered MEC scheme, with NOMA protocol applied.
- MEC denotes the conventional MEC scheme based on TDMA protocol without the application of UC.
- With regards to the local computing scheme, the users execute the computation task by local computing only, which corresponds to the condition of $\ell_1^{off} = 0$ and $\ell_2^{off} = 0$.
- For offloading only scheme, the computation tasks are fully executed by the MEC integrated with the AP by setting $\ell_1^{loc} = 0$ and $\ell_2^{loc} = 0$.

In Fig. 4.3, we present the relationship between the maximum computation efficiency and the required computation data size. To demonstrate the effectiveness of the partial offloading, the results of offloading only and local computing schemes are provided for comparison. As can be observed from the figure, the computation efficiency decreases with larger required data bits for all three schemes, which implies

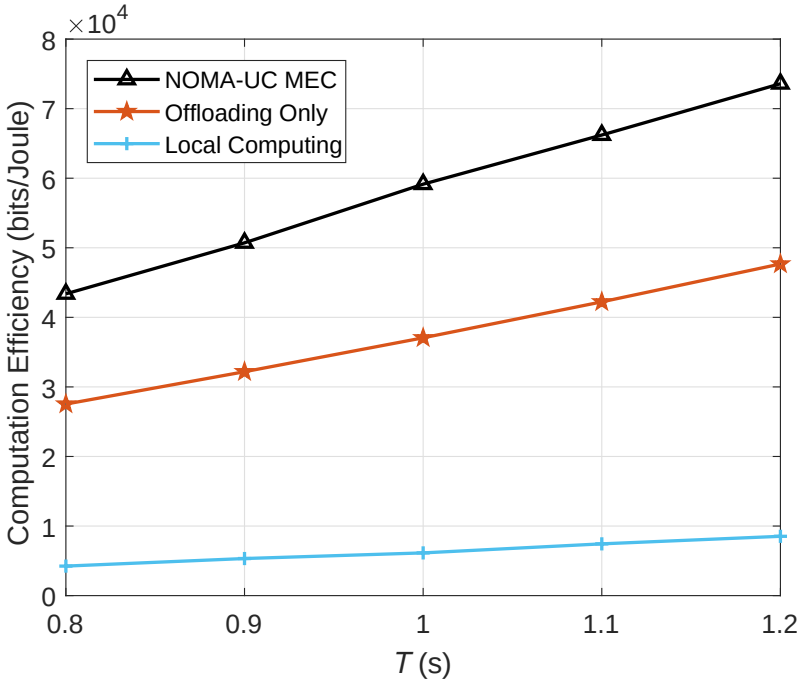


Figure 4.4: Maximum computation efficiency vs. T.

that the required energy to compute grows faster than the growth of data size. In addition, it is obvious that the proposed NOMA-UC MEC scheme is superior to the baseline schemes. Specifically, the local computing scheme yields the worst performance, indicating that the application of MEC greatly contributes to performance improvement.

Fig. 4.4 shows the influence of the block slot duration on the computation efficiency. It is noted that the computation efficiency increases monotonically with the block slot duration for all the schemes, and the proposed NOMA-UC MEC scheme outperforms the benchmark schemes. For example, when the time duration is 1s, the achievable computation efficiency for NOMA-UC MEC is about 6×10^4 bits/Joule, while for offloading only and local computing are 3.8×10^4 bits/Joule and 0.7×10^4 bits/Joule, respectively.

To show the effect of NOMA and UC application in MEC design, four schemes, namely the proposed NOMA-UC MEC, NOMA-MEC, UC-MEC, and MEC schemes are presented in Fig. 4.5. The computation efficiency performs decreasing trends with the increase of required data bits for all the schemes, while the proposed NOMA-UC MEC scheme produces the best performance. In addition, compared with the MEC scheme, both NOMA-UC and UC-MEC achieve higher computation efficiency,

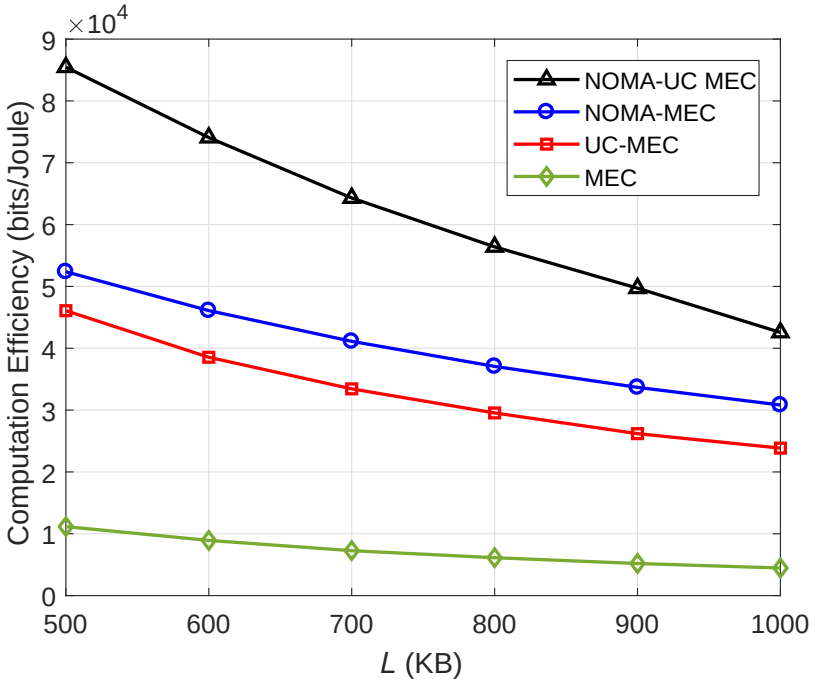


Figure 4.5: Maximum computation efficiency vs. L.

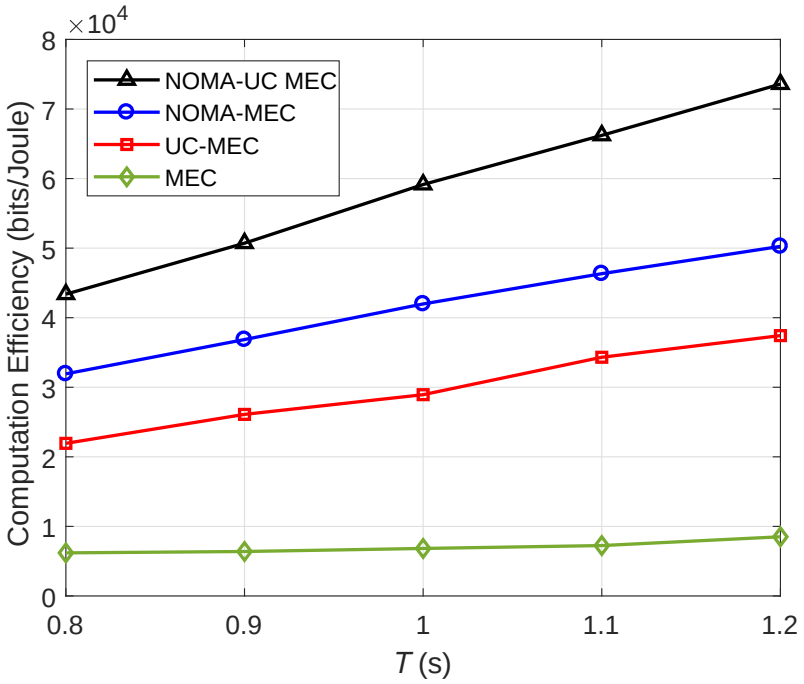


Figure 4.6: Maximum computation efficiency vs. T.

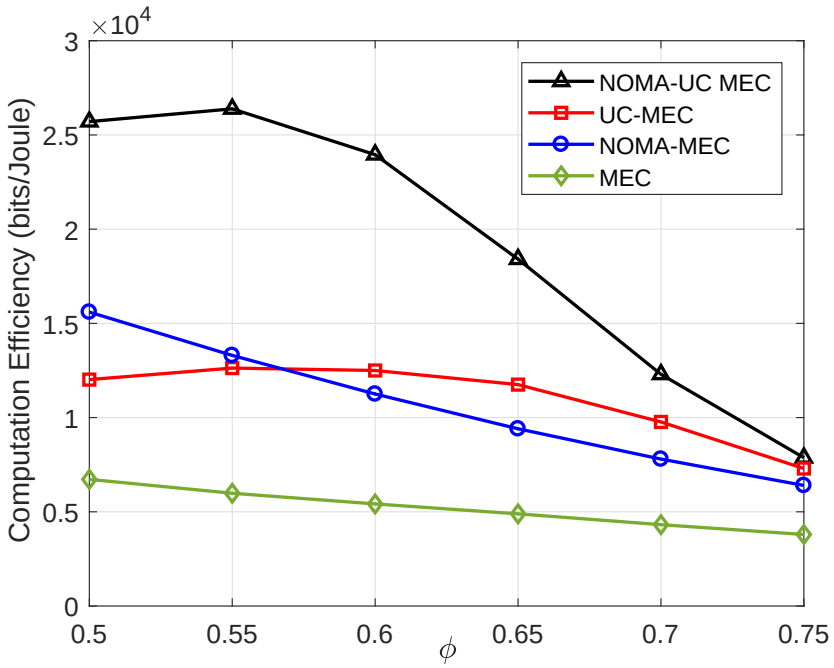


Figure 4.7: Maximum computation efficiency vs. ϕ.

proving the benefit of applying NOMA and UC into MEC design in dealing with the doubly near-far effect. Moreover, compared with Fig. 4.3, the computation efficiency of 'local computing' is marginally lower than that of 'MEC', which is determined by the feature of the adopted wireless powered MEC system. Due to the serious double near-far effect and limited computational capability of users, more resources will be allocated to the user with bad channel conditions to satisfy the corresponding computation requirements for MEC scheme, and thus the benefit of the user with good channel conditions is sacrificed. The system performance becomes even worse when the connection between mobile users and MEC servers is relatively poor. As a result, the performance improvement of MEC is limited compared to 'local computing' scheme.

Fig. 4.6 is plotted to compare the computation efficiency performance with the block slot duration T. It shows that the trend for all the four curves are similar, and the proposed NOMA-UC MEC yields the best performance. This indicates NOMA-UC MEC can improve the computation efficiency of the system. Moreover, the performance of both NOMA-MEC and UC-MEC is superior to that of MEC, proving the advantage of applying NOMA and UC into the MEC design.

To evaluate the impact of users' locations, it is assumed that the AP and two

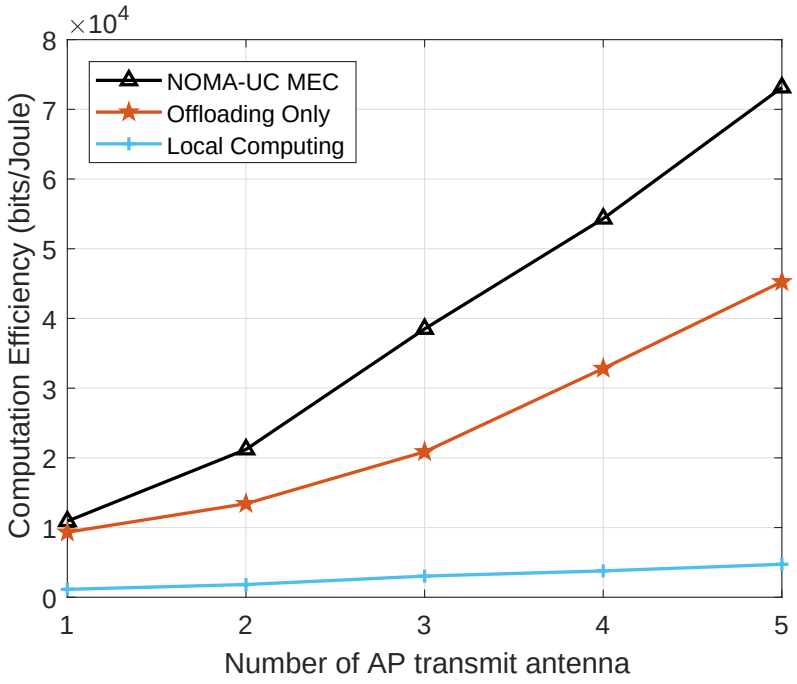


Figure 4.8: Maximum computation efficiency vs. number of AP transmit antenna.

users are placed in a line, i.e., $d_1 = 8$ m, $d_2 = \phi d_1$, and $d_{12} = (1-\phi)d_1$, where $\phi \in [0.5, 0.75]$. As can be observed from Fig. 4.7, the plain MEC scheme still yields the worst performance. Moreover, both NOMA-UC MEC and UC-MEC show a similar trend, where the computation efficiency first increases and then decreases when ϕ becomes larger. The reason is that, with UC applied to NOMA-UC MEC and UC-MEC schemes, it can make a great impact when ϕ is relatively small. When ϕ becomes larger, the far user's channel gain degradation dominates the effect of UC, and thus the computation efficiency decreases. Moreover, NOMA-MEC outperforms UC-MEC in the small ϕ regime, while UC-MEC gains better performance with a larger ϕ. This is because the channel gain of U_2 decreases when ϕ increases, the performance of NOMA-MEC decreases monotonically. However, the performance of UC-MEC degrades with a smaller ϕ, since two users are relatively far from each other and UC is inefficient. When two users get closer to each other with the increase of ϕ, the channel degradation can be better compensated for UC-MEC scheme.

Fig. 4.8 shows the achieved computation efficiency versus the number of transmit antennas N_t equipped at the AP. The AP and two users are placed in a line and the distance is set as $d_1 = 8$ m, $d_2 = 4.8$ m, and $d_{12} = 3.2$ m. Due to the additional degrees of freedom introduced by the increasing number of transmit antennas, the

computation efficiency can be improved for all the schemes. Particularly, compared with NOMA-UC MEC and offloading only scheme, the performance improvement of local computing is limited. The reason is that, the channel gain can be improved in both the downlink WPT transmission and uplink communications for NOMA-UC MEC and offloading only scheme, whereas only downlink WPT transmission is affected for local computing.

4.5 Summary

We investigated the application of NOMA and UC in a wireless powered MEC system under the non-linear energy harvesting model, in which the joint optimization of transmit beamforming, time and power allocations was proposed to maximize the system computation efficiency. To solve the formulated nonconvex problem, SDR technique was first applied to transform the original problem into a more tractable expression. Then, the transformed problem was reformulated with variables substitution, which can be finally solved by applying the SCA method. Numerical results demonstrated the superiority of applying NOMA and UC in wireless powered MEC design.

Chapter 5

Energy Efficient Uplink Transmissions in LoRa Networks

5.1 Introduction

LoRa has been recognized as one of the most promising LPWA techniques. Since LoRa devices are usually powered by batteries, EE is an essential consideration. In this chapter, we investigate the energy efficient resource allocation in LoRa networks to maximize the system EE (SEE) and the minimal EE (MEE) of LoRa users, respectively. Specifically, our objective is to maximize the corresponding EE by jointly exploiting user scheduling, SF assignment, and transmit power allocations. To solve them efficiently, we first propose a suboptimal algorithm, which includes the low-complexity user scheduling scheme based on matching theory and the heuristic SF assignment solution for LoRa users scheduled on the same channel. Then, to deal with the power allocation, for the case considering SEE, an optimal algorithm is proposed to maximize the SEE. With regards to MEE, an iterative power allocation algorithm based on the generalized fractional programming and sequential convex programming is proposed to maximize the minimal EE achieved by LoRa users accessing the same channel. Numerical results show that the proposed user scheduling algorithm achieves near-optimal EE performance, and the proposed power allocation algorithms outperform the benchmarks.

The works presented in this Chapter have been published at the IEEE Transactions on Communications, May 2020, and IEEE Global Communications Conference (GLOBECOM 2018), Abu Dhabi, UAE, Dec. 2018.

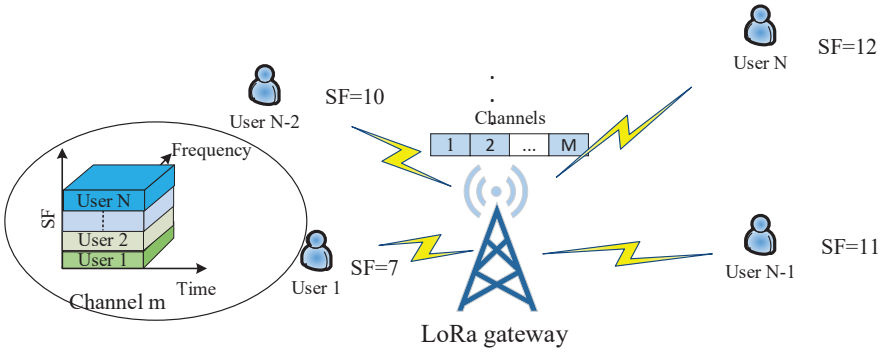


Figure 5.1: System model of resource allocation in LoRa networks.

The remainder of this chapter is organized as follows. In Section 5.2, we present the system model of the uplink LoRa networks and formulate two EE optimization problems. A low-complexity energy efficient user scheduling and heuristic SF assignment scheme are developed in Section 5.3. In Section 5.4, the power allocation algorithms for both problems are provided. Simulation results are presented in Section 5.5 and finally the chapter is concluded in Section 5.6.

5.2 System Model

5.2.1 System Model

Considering the uplink transmission in LoRa networks, N active LoRa users communicate with one LoRa gateway through M channels, which is shown in Fig. 5.1. LoRa users located within the same channel share the same time and frequency slots by adopting different SFs. Both the LoRa gateway and LoRa users are equipped with the single antenna. Denote $\boldsymbol{M} = \{1, ...M\}$ and $\boldsymbol{N} = \{1, ...N\}$ to be the channels set and users set, respectively. B_m Hz is the bandwidth of the m-th channel, SC_m. The number of LoRa users scheduled on SC_m is labelled as S_m, i.e., $S_m = \sum_{l=1}^{N} s_{m,l}$, where $s_{m,l} \in \{0, 1\}$ is used to indicate whether an arbitrary user, U_l, is allocated to SC_m. If $s_{m,l} = 1$, it indicates that U_l occupies SC_m, and $s_{m,l} = 0$ if otherwise. Let $\boldsymbol{S} = \{s_{m,l} | m \in \boldsymbol{M}, l \in \boldsymbol{N}\}$ denote the set of user clustering. It is noted that S_m should be no more than 6 as the available SFs range from 7 to 12 [109], which limits the maximum number of active LoRa users that can be served simultaneously in one channel. Let $p_{m,l}$ be the power allocated to U_l using SC_m and $\boldsymbol{P} = \{p_{m,l} | m \in \boldsymbol{M}, l \in \boldsymbol{N}\}$ represent the set of power allocation coefficients. It is assumed that perfect CSI is available at the LoRa gateway.

Assuming that S_m users are allocated within SC_m, the signal received at the LoRa gateway through SC_m can be expressed as

$$y_m = \sum_{l=1}^{N} \sqrt{p_{m,l}} h_{m,l} x_l + n_m, \tag{5.1}$$

where $h_{m,l} = g_{m,l}\varsigma_m d_l^{-a}$ indicates the channel coefficient between U_l and the LoRa gateway on SC_m, $g_{m,l}$ represents the Rayleigh fading channel gain, d_l denotes the distance from U_l to the LoRa gateway, a is the channel path loss exponent, and ς_m is a constant depending on path loss of SC_m. Denote $n_m \in \boldsymbol{CN}(0, \sigma_m^2)$ as the AWGN with noise variance σ_m^2.

The SINR of U_l, received at the gateway over SC_m is characterized as

$$\mathrm{SINR}_{m,l}(\boldsymbol{S}, \boldsymbol{P}) = \frac{s_{m,l} p_{m,l} |h_{m,l}|^2}{I_{SF} + \sigma_m^2}, \tag{5.2}$$

where $I_{SF} = \sum_{k=1, k\neq l}^{S_m} p_{m,k} |h_{m,k}|^2 \psi(l,k)$ represents the interference caused by LoRa users adopting different SFs in the same channel, and $\psi \in [0,1]$ represents the cross correlation factors between the coded LoRa waveforms with different SFs. Note that the interference is introduced due to imperfect SF orthogonality.

Similar to [110], we assume that a transmission rate of Shannon's upper bound can be achieved by a perfect coding. Therefore, Shannon rate is adopted to model the LoRa-specific rates for mathematical tractability. The achievable data rate and the overall power consumption for the l-th LoRa user over SC_m can be denoted as follows

$$\begin{cases} R_{m,l}(\boldsymbol{S}, \boldsymbol{P}) = B_m \log_2 (1 + \mathrm{SINR}_{m,l}), \\ P_{m,l}(\boldsymbol{S}, \boldsymbol{P}) = \zeta_{m,l} p_{m,l} + P_c^l, \end{cases} \tag{5.3}$$

where $\zeta_{m,l} \geq 1$ is a constant denoting the power inefficiency, P_c^l represents the additional circuit power consumption of U_l owing to inevitable electronic operations [23].

Therefore, the achievable sum rate and the total power consumption of the system can be expressed as

$$\begin{cases} R(\boldsymbol{S}, \boldsymbol{P}) = \sum_{m=1}^{M} \sum_{l=1}^{S_m} R_{m,l}(\boldsymbol{S}, \boldsymbol{P}), \\ P(\boldsymbol{S}, \boldsymbol{P}) = \sum_{m=1}^{M} \sum_{l=1}^{S_m} P_{m,l}(\boldsymbol{S}, \boldsymbol{P}). \end{cases} \tag{5.4}$$

For energy efficient uplink transmissions in LoRa networks, the goal is to maximize SEE, which is defined as information bits within a unit energy. Therefore, SEE is

formulated as a ratio of the system sum rate to the total power consumption, which can be characterized as

$$\eta = \frac{R\left(\boldsymbol{S}, \boldsymbol{P}\right)}{P\left(\boldsymbol{S}, \boldsymbol{P}\right)}. \tag{5.5}$$

Moreover, MEE is defined as the ratio of individual LoRa user rate to the corresponding consumed power, which can be given as

$$\eta_{m,l} = \frac{R_{m,l}\left(\boldsymbol{S}, \boldsymbol{P}\right)}{P_{m,l}\left(\boldsymbol{S}, \boldsymbol{P}\right)}. \tag{5.6}$$

5.2.2 Problem Formulations

To acquire an energy efficient resource allocation design for the considered LoRa networks, we formulate the EE optimization problems based on two major performance measurement criteria in terms of SEE and MEE, respectively.

5.2.2.1 System Energy Efficiency

The SEE maximization problem is formulated as follows

$$\max_{\boldsymbol{S},\boldsymbol{P},\tau_{SF}} \quad \eta, \tag{5.7a}$$

$$s.t. \quad 0 \leq p_{m,l} \leq p_{\max}, \tag{5.7b}$$

$$s_{m,l} \in \{0,1\}\,,\ \forall\ m,\ l, \tag{5.7c}$$

$$\sum_{m=1}^{M} s_{m,l} \leq 1,\ \forall\ l, \tag{5.7d}$$

$$\sum_{l=1}^{N} s_{m,l} \leq \Lambda_{\max},\ \forall\ m, \tag{5.7e}$$

$$\frac{p_{m,l}|h_{m,l}|^2}{\sigma_m^2} \geq \tau_{SF},\ \forall\ m,\ l. \tag{5.7f}$$

In problem (5.7), (5.7a) represents the formulated problem to maximize SEE. Constraint (5.7b) limits the transmit power of each LoRa user. In (5.7c), the value of the user cluster indicator $s_{m,l}$ is either 0 or 1. Constraint (5.7d) indicates that each LoRa user can access at most one channel. Due to the maximum number of available SF and interference control, we assume that at most $\Lambda_{\max}$ users can be assigned to the same channel, which is guaranteed by (5.7e). It is worth noting that the desired LoRa user adopting a given SF can be detected successfully only if the received SNR is no less than the threshold τ_{SF}, which is guaranteed by constraint (5.7f). Table 5.1 [57] shows the relationship between the required SNR and SF.

Table 5.1: Relationship Between Distance Range and Spreading Factors.

Spreading factor (SF)	7	8	9	10	11	12
Distance Range (km)	2	4	6	8	10	12
Required SNR (dB)	-7.5	-10	-12.5	-15	-17.5	-20

5.2.2.2 Max-min Energy Efficiency

The MEE optimization problem is given by

$$\max_{\boldsymbol{S},\boldsymbol{P},\tau_{SF}} \min_{m,l} \quad \eta_{m,l}, \tag{5.8a}$$

$$s.t. \qquad (5.6b) - (5.6f), \tag{5.8b}$$

where (5.8a) denotes the objective to maximize MEE.

Theorem 1. *The formulated problems of both (5.7) and (5.8) are NP-hard.*

Proof: The proof is provided in Appendix H.

Since the formulated problems are nonconvex and NP-hard, it is challenging and intractable to solve (5.7) and (5.8) within polynomial time.

Furthermore, the current adaptive data rate (ADR) mechanism adopted in LoRaWAN fails to perform the channel selection effectively, a more efficient distributed user scheduling scheme needs to be designed. Besides, the ADR scheme optimizes the transmit power and SF based on some previous uplink messages, which achieves low resource efficiency. Therefore, the optimal transmit power allocation scheme that can be easily implemented at the LoRa gateway is required. Moreover, as can be observed from the expression of the objective functions (5.7a) and (5.8a), the channel and power allocations are coupled with each other for both MEE and SEE. As the formulated problem is NP-hard, to avoid the considerable complexity of the global optimum solution, we will exploit user scheduling, SF assignment, and power allocation schemes separately. Specifically, LoRa users first perform self-matching to match with the corresponding channels. Due to the linear SF inequality constraint and the inner relationship between distance range and SF, the SF can be determined by the distance. Then, a low-complexity distance-based SF assignment algorithm is proposed to operate at the LoRa gateway. Finally, the LoRa gateway assigns SFs and allocates power for LoRa users sharing the same channel, on the basis of the proposed low-complexity distance-based SF assignment scheme and optimal power allocation algorithms. The details are described in next two sections.

5.3 Energy-efficient User Scheduling

In this section, for the two formulated problems, energy efficient user scheduling scheme is firstly proposed to maximize the corresponding EE. For LoRa users scheduled on the same channel, a heuristic distance-based SF assignment scheme is then proposed.

5.3.1 User Scheduling

In this section, we provide a matching theory based user scheduling scheme with low complexity. Firstly, by assuming that each user is allocated with the maximum power and a given SF, (5.7) and (5.8) can be reformulated as

$$\max_{\boldsymbol{S}} \quad \eta, \tag{5.9a}$$

$$\text{or} \quad \max_{\boldsymbol{S}} \min_{m,l} \quad \eta_{m,l}, \tag{5.9a'}$$

$$s.t. \quad s_{m,l} \in \{0,1\}, \tag{5.9b}$$

$$\sum_{m=1}^{M} s_{m,l} \leq 1, \ \forall\, l, \tag{5.9c}$$

$$\sum_{l=1}^{N} s_{m,l} \leq \Lambda_{\max}, \ \forall\, m. \tag{5.9d}$$

It is noted that (5.9) is a many-to-one matching problem for both SEE and MEE, as at most one channel can be allocated to a LoRa user while a subset of LoRa users can be assigned into the same channel. Moreover, due to the interference term in (5.2), each user's preference on the channel is not only influenced by the channel conditions, but also the other LoRa users sharing the same channel. Similarly, each channel not only cares which LoRa users to match with, but also the co-channel interference introduced by the other subset of LoRa users with different SFs. Hence, this is a many-to-one matching game with peer effects [111].

To better illustrate the matching model with peer effects, we firstly introduce a preference ordering for LoRa users, in which for any given user $U_l \in \boldsymbol{N}$, any two channels SC_m, $SC_{m'} \in \boldsymbol{M}$, any two matchings φ and φ' are defined as

$$(SC_m, \varphi) \succ_{U_l} (SC_{m'}, \varphi') \Leftrightarrow R_{m,l}(\varphi) > R_{m,l}(\varphi'), \tag{5.10}$$

which means that LoRa user U_l prefers channel SC_m in φ rather than $SC_{m'}$ in φ' only if U_l achieves higher rate over channel SC_m than over $SC_{m'}$. It is worth mentioning that the preference order is based on achievable rates. The reason is that, with

given transmit power allocation, the power consumption in the denominator of the objective is a fixed value, the EE objectives are equivalent to the corresponding rates optimization problems.

By defining channel SC_m's preference value on the user set as $R_m^{SEE} = \sum_{l=1}^{N} R_{m,l}$ and $R_m^{MEE} = \min R_{m,l}, \forall l \in SC_m$ for SEE and MEE, respectively, the preference ordering for channel SC_m can be obtained similarly.

Due to the existence of peer effects, stable matching is not straightforward guaranteed. Therefore, the two-sided exchange stability has been introduced to depict the impact of peer effects on the matching game [111]. Firstly, a swap matching and swap blocking pair are defined as follows:

Definition 1. *A swap matching behaviour* $\varphi_l^j = \varphi \backslash \{(U_l, SC_m), (U_j, SC_n)\} \cup \{(U_j, SC_m), (U_l, SC_n)\}$ *is defined as* $\varphi(U_l) = SC_m$ *and* $\varphi(U_j) = SC_n$.

Note that a swap-matching is realized by performing a swap operation, which motivates us to introduce swap-blocking pair.

Definition 2. *Given a matching* φ *with a pair* (U_l, U_j)*, if there exists* $\varphi(U_l) = SC_m$ *and* $\varphi(U_j) = SC_n$ *such that*
1) $\forall q \in \{U_l, U_j, SC_m, SC_n\}, \varphi_l^j(q) \succeq_q \varphi(q)$*;*
2) $\exists q \in \{U_l, U_j, SC_m, SC_n\}, \varphi_l^j(q) \succ_q \varphi(q)$,
which means the swap matching φ_l^j *is approved, and we call* (U_l, U_j) *as a swap-blocking pair in* φ.

The definition demonstrates that the achievable data rates of any player, i.e., U_l and SC_m, will not decrease by employing a swap matching, and the data rates of at least one player will increase. Then a stable matching status can be achieved through a set of swap matching operations, known as a two-sided exchange stable matching that is described in Definition 3.

Definition 3. *A two-sided exchange stable (2ES) matching* φ *can be achieved if it is not blocked by any swap-blocking pair.*

Based on the above definition, the proposed user scheduling algorithm is described in Table 5.2, which consists of initialization step and swap matching step. The initialization step is a deferred acceptance algorithm [112], which aims to generate the initial matching. Specifically, the CSI-based preference list is constructed for each LoRa user, i.e., is the most preferred channel for LoRa user U_i. For the LoRa gateway, it constructs the distance-based preference list, i.e., $j^* = \underset{j \in \boldsymbol{M}}{\text{argmax}}$, the highest preference

Table 5.2: User Scheduling Algorithm for LoRa Networks Based on Matching Theory

Initialization step
While there exists unmatched users and channels
1) $j^* = \underset{j \in \boldsymbol{M}}{\text{argmax}} \quad |h_{j,i}|^2$.
2) LoRa user U_i matches with its most preferred channels that it has not been rejected before.
3) Remove U_i from $\boldsymbol{N}$.
Swap matching step
Repeat
1) For user $U_i \in \boldsymbol{N}$, it searches $\boldsymbol{N} \backslash U_j$.
2) **if** U_i and U_j is swap-blocking pair, **then**
U_i exchanges its matching with U_j and set $\varphi = \varphi_i^j$.
3) **else**
4) keep the current matching.
5) **end if**
Until no swap blocking can be formed for all users.
Return the stable matching φ

is the closest LoRa user, due to the fact that LoRa provides long-range communications. Each LoRa user proposes to the highest preference channel, and each channel picks at most $\Lambda_{\max}$ users based on its preference list. Then the remaining LoRa users propose to their second preference, and the process stops until no unmatched users exist. In the swap matching step, each LoRa user keeps searching for swap-blocking pairs to perform swap matching operation if approved. The searching terminates until no swap-blocking pairs can be formed, and the final stable matching is returned.

Theorem 2. *The proposed user scheduling algorithm converges to a $2ES$ matching φ^* within a finite number of swap operations.*

Proof: The proof is provided in Appendix I.

Theorem 3. *The computational complexity of the proposed user scheduling algorithm is $O(MN + \frac{1}{2}I\Lambda_{\max}N(M-1))$ at worst, where I denotes the number of iterations for swap-matching step.*

Proof: The proof is provided in Appendix J.

5.3.2 SF Assignment

As can be seen from the structure of the formulated problems, i.e., (5.7) and (5.8), the SF is only related to the linear inequality constraint, i.e., the SNR threshold

Table 5.3: Distance-based SF Assignment Algorithm for LoRa Networks

1. **Initialization**: Distance-based SFs are assigned to LoRa users scheduled within one channel according to Table 5.1.
2. Initialize the set of $\mathbf{SF}_k$ to record users who has been allocated with SF of k
3. **while** $\{\boldsymbol{SF}_k \geq 2\}$ **do**
4. For LoRa users j_1,...j_n sharing the same SF, sort the distance as $d_{j_1} < ... < d_{j_n}$.
5. Assign SF of k to LoRa user j_1.
6. Increase SF by one in sequence for LoRa user j_2,...j_n.
7. **end while**
8. **Until** all users in one channel occupy different SFs.

requirement. With given power allocation and user scheduling scheme, we need to check whether the SF constraint is satisfied. By applying the proposed user scheduling scheme, LoRa users are allocated into the corresponding channels. For a given channel, the distance of each scheduled LoRa user is easily obtained at the LoRa gateway. Since LoRa network aims to realize long-range communications up to 40 km, the large-scale fading becomes the main effect for the channel gain. Therefore, a heuristic distance-based SF assignment scheme is proposed, which is summarized as Table 5.3. Specifically, in each channel, a predefined SF is assigned to each LoRa user according to the relationship between the distance range and corresponding SF based on Table 5.1. Then the LoRa gateway keeps searching for the SF assigned to more than one LoRa user. For LoRa users sharing the same SF, a higher available SF in the network is reassigned to the one with longer distance to the gateway. The searching iteration stops until all LoRa users accessing the same channel occupy a unique SF.

5.4 Energy-efficient Power Allocation Algorithms

In this section, we focus on the optimal power allocations to maximize SEE and MEE, respectively, with the obtained user scheduling and SF assignment schemes given in the last section. It is worth mentioning that though the proposed power allocation algorithms are centralized approach, it is easy to be implemented at the LoRa gateway. The reason is that, the LoRa gateway can easily obtain the CSI of LoRa users assigned within each channel, due to the fact that the number of LoRa users sharing the same channel is limited after the user scheduling. Moreover, as the LoRa gateway is more

powerful than LoRa users, the centralized power allocation algorithms implemented at the LoRa gateway help LoRa users avoid energy consumption for power allocation.

5.4.1 Energy-Efficient Power Allocation for SEE

With given user scheduling and SF assignment schemes, we consider the power allocation problem (5.7) with constraints (5.7b) and (5.7f). The difficulty lies in the nonconvex objective function as all the constraints are linear inequalities. To transform the objective into a more tractable form, we first approximate it with the lower bound by using the following inequality [113]

$$\ln(1+\gamma) \geq \alpha \ln\gamma + \Upsilon \tag{5.11}$$

where

$$\alpha = \frac{\tilde{\gamma}}{1+\tilde{\gamma}}, \tag{5.12a}$$

$$\Upsilon = \ln(1+\tilde{\gamma}) - \frac{\tilde{\gamma}}{1+\tilde{\gamma}}\ln(\tilde{\gamma}). \tag{5.12b}$$

The approximation is tight when the constants α and Υ are chosen with $\gamma = \tilde{\gamma}$. The proof can be easily acquired by substituting $\gamma = \tilde{\gamma}$ into the right side of inequality (5.11).

With inequality (5.11), we get a lower bound for the achievable data rate of U_l accessing SC_m as follows

$$\begin{aligned} R_{m,l} \geq \tilde{R}_{m,l} &= \frac{B_m}{\ln 2}\left(\alpha_{m,l}\ln(\mathrm{SINR}_{m,l}) + \Upsilon_{m,l}\right) \\ &= \frac{B_m}{\ln 2}\left(\alpha_{m,l}\ln\left(\frac{p_{m,l}|h_{m,l}|^2}{\sum_{k=1,k\neq l}^{S_m} p_{m,k}|h_{m,k}|^2 + \sigma_m^2}\right) + \Upsilon_{m,l}\right). \end{aligned} \tag{5.13}$$

However, $\tilde{R}_{m,l}$ is still nonconvex. To convert it into a concave expression, we introduce a variable transformation as $x_{m,l} = \ln(p_{m,l})$. Then we have

$$\ln(\mathrm{SINR}_{m,l}) = \ln(|h_{m,l}|^2) + x_{m,l} - \ln(\sum_{k=1,k\neq l}^{S_m} e^{x_{m,k}}|h_{m,k}|^2 + \sigma_m^2), \tag{5.14}$$

which is concave over $\boldsymbol{x} = \{x_{m,l} | m \in \boldsymbol{M}, l \in \boldsymbol{N}\}$ since the log-sum-exp function is convex. Through the above transformation, the original objective can then be approximated by its lower bound function, i.e., $\eta \geq \frac{\sum_{m=1}^{M}\sum_{l=1}^{S_m}\frac{B_m}{\ln 2}\left(\alpha_{m,l}\ln(\mathrm{SINR}_{m,l})+\Upsilon_{m,l}\right)}{\sum_{m=1}^{M}\sum_{l=1}^{S_m}(\zeta_{m,l}p_{m,l}+P_c^l)}$.

Table 5.4: Optimal Power Allocation Algorithm for Solving SEE

1. Initialize feasible power allocation variables $\boldsymbol{P^1}$.
2. Set $n = 1$. The value of (5.16a) is calculated with $\boldsymbol{P^1}$, denoted as η^0.
3. **while** $\frac{\eta^n - \eta^{n-1}}{\eta^{n-1}} > \epsilon$, where ϵ is a given constant.
4. Set $n = n + 1$.
5. Update $\tilde{x}_{m,l}^n$ and ϕ^n by solving (5.16).
6. Update power allocation variables $\boldsymbol{P}^n$ by $p_{m,l}^n = e^{\frac{\tilde{x}_{m,l}^n}{\phi^n}}$.
7. Update the objective value η^n.
8. **end while**
9. Output the optimal $\boldsymbol{P}^*$.

The lower bound approximation is a concave-convex fractional function as it consists of a concave numerator $f(p_{m,l}) = \sum_{m=1}^{M}\sum_{l=1}^{S_m} \frac{B_m}{\ln 2}\left(\alpha_{m,l}\ln(\mathrm{SINR}_{m,l}) + \Upsilon_{m,l}\right)$, and an affine denominator $g(p_{m,l}) = \sum_{m=1}^{M}\sum_{l=1}^{S_m}(\zeta_{m,l}p_{m,l} + P_c^l)$. The concave-convex fractional problem can be efficiently solved with the Charnes-Cooper transformation [114], which is given as follows:

Lemma 5. *A concave-convex fractional problem, max* $\frac{f(x)}{g(x)}$*, where f is concave and g is convex, can be reformulated as a concave problem*

$$\max \qquad \phi f(\frac{y}{\phi}) \tag{5.15a}$$

$$s.t. \qquad \phi g(\frac{y}{\phi}) \leq 1, \tag{5.15b}$$

with the Charnes-Cooper transformation $y = \frac{x}{g(x)}$, $\phi = \frac{1}{g(x)}$ *and* $\phi > 0$.

As a result, the original problem is recast as the following equivalent convex optimization problem

$$\max_{\tilde{x}_{m,l}} \qquad \phi\sum_{m=1}^{M}\sum_{l=1}^{S_m}\tilde{R}_{m,l}\left(\frac{\tilde{x}_{m,l}}{\phi}\right) \tag{5.16a}$$

$$s.t. \qquad \phi(\sum_{m=1}^{M}\sum_{l=1}^{S_m} e^{\frac{\tilde{x}_{m,l}}{\phi}} + P_c) \leq 1, \tag{5.16b}$$

$$\phi e^{\frac{\tilde{x}_{m,l}}{\phi}} \geq 0, \tag{5.16c}$$

$$\phi(e^{\frac{\tilde{x}_{m,l}}{\phi}} - p_{\max}) \leq 0,\ \forall\ m,\ l, \tag{5.16d}$$

$$\phi(\sigma_m^2\tau_{SF} - e^{\frac{\tilde{x}_{m,l}}{\phi}}|h_{m,l}|^2) \leq 0,\ \forall\ m,\ l, \tag{5.16e}$$

$$\phi > 0, \tag{5.16f}$$

where

$$\tilde{R}_{m,l}(\frac{\tilde{x}_{m,l}}{\phi}) = \frac{B_m}{\ln 2}\left(\alpha_{m,l}\ln(\tilde{\text{SINR}}_{m,l}) + \Upsilon_{m,l}\right), \tag{5.17a}$$

$$\ln(\tilde{\text{SINR}}_{m,l}) = \ln\left(\frac{e^{\frac{\tilde{x}_{m,l}}{\phi}}|h_{m,l}|^2}{\sum_{k=1,k\neq l}^{S_m} e^{\frac{\tilde{x}_{m,k}}{\phi}}|h_{m,k}|^2 + \sigma_m^2}\right). \tag{5.17b}$$

The equivalence between (5.14) and (5.17b) is guaranteed with the aid of Charnes-Cooper transformation introduced by Lemma 5. It is noted that (5.16) is a convex problem, which can be efficiently solved with standard convex solvers [97]. The corresponding procedure is outlined in Table 5.4.

To prove that the proposed power allocation algorithm converges, we provide the following proposition.

Proposition 8. *The value of η is improved continuously in each iteration, and finally the proposed power allocation algorithm converges to a KKT point of the original problem.*

Proof: The proof is provided in Appendix K.

5.4.2 Energy-Efficient Power Allocation for MEE

For the case of MEE, given user scheduling and SF assignment, denote η_{MEE}^{opt} and $\boldsymbol{P}^{opt}$ as the optimal solution to MEE and power allocation coefficients, the following results hold

$$\eta_{MEE}^{opt} = \max_{\boldsymbol{P}} \min_{m,l} \frac{R_{m,l}(\boldsymbol{P})}{P_{m,l}(\boldsymbol{P})} = \min_{m,l} \frac{R_{m,l}(\boldsymbol{P}^{opt})}{P_{m,l}(\boldsymbol{P}^{opt})}. \tag{5.18}$$

Then we have Theorem 4 as below:

Theorem 4. *The optimal solution η_{MEE}^{opt} is obtained on condition that*

$$\begin{aligned} &\max_{\boldsymbol{P}} \min_{m,l} [R_{m,l}(\boldsymbol{P}) - \eta_{MEE}^{opt} P_{m,l}(\boldsymbol{P})] \\ &= \min_{m,l} [R_{m,l}(\boldsymbol{P}^{opt}) - \eta_{MEE}^{opt} P_{m,l}(\boldsymbol{P}^{opt})] = 0. \end{aligned} \tag{5.19}$$

Proof: The proof is provided in Appendix L.

Theorem 4 indicates that the optimal solutions of the original problem can be obtained by equivalently solving (5.19). As the value of η_{MEE}^{opt} is unknown in advance, the properties of (5.19) need to be further revealed.

Denote that $\pi(\eta_{m,l}) = \max_{\boldsymbol{P}} \min_{m,l} [R_{m,l}(\boldsymbol{P}) - \eta_{m,l} P_{m,l}(\boldsymbol{P})]$, we have the following theorem:

Table 5.5: The Bisection Method to Solve MEE

1. Given $\eta_{m,l}^{min}$ and $\eta_{m,l}^{max}$, the tolerance $\epsilon = 10^{-3}$, and iteration index $i = 0$.
2. **while (1) do**
3. $\eta_{m,l}^{i} = \frac{\eta_{m,l}^{min} + \eta_{m,l}^{max}}{2}$.
4. update $\boldsymbol{P}^i$ by solving problem (5.21);
5. **if** $\lvert\Delta\rvert = \lvert \min_{m,l} R_{m,l}(\boldsymbol{P}^i) - \eta_{m,l}^{i} P_{m,l}(\boldsymbol{P}^i)\rvert < \epsilon$
6. $\quad \boldsymbol{P}^{opt} = \boldsymbol{P}^i$,
7. $\quad$ break.
8. **else if** $\Delta < 0$ $\quad \eta_{m,l}^{max} = \eta_{m,l}^{i}$.
9. **else if** $\Delta > 0$ $\quad \eta_{m,l}^{min} = \eta_{m,l}^{i}$.
10. **end if**
11. $i = i + 1$.
12. **end while**

Theorem 5. *i)* $\pi(\eta_{m,l})$ *strictly decreases with* $\eta_{m,l}$.

ii) With $\eta_{m,l} > 0$*, we obtain that*

$$\pi(\eta_{m,l}) = \begin{cases} > 0 & if \quad \eta_{m,l} < \eta_{MEE}^{opt}, \\ = 0 & if \quad \eta_{m,l} = \eta_{MEE}^{opt}, \\ < 0 & if \quad \eta_{m,l} > \eta_{MEE}^{opt}, \end{cases} \tag{5.20}$$

Proof: The proof is provided in Appendix M.

Therefore, based on the properties of $\pi(\eta_{m,l})$, we can apply the bisection method to solve it. The initial lower and upper bounds can be set as $\eta_{m,l}^{min} = 0$ and $\eta_{m,l}^{max}$, where $\eta_{m,l}^{max}$ is a relatively large constant. Given $\eta_{m,l}^{min}$ and $\eta_{m,l}^{max}$, the algorithm based on the bisection method is summarized in Table 5.5.

For the i-th iteration with a given $\eta_{m,l}^{i}$ at line 4 of Table 5.5, to update the power allocation coefficients, we need to deal with the following optimization problem:

$$\max_{\boldsymbol{P}} \min_{m,l} R_{m,l}(\boldsymbol{P}) - \eta_{m,l}^{i} P_{m,l}(\boldsymbol{P}), \tag{5.21a}$$

$$s.t. \qquad 0 \le p_{m,l} \le p_{\max}, \tag{5.21b}$$

$$\frac{p_{m,l}|h_{m,l}|^2}{\sigma_m^2} \ge \tau_{SF}, \ \forall\ m,\ l. \tag{5.21c}$$

All the constraints in (5.21) are convex, the difficulty lies in the nonconvex objective. To deal with it, an auxiliary variable μ is introduced to denote that $\min_{m,l} R_{m,l}(\boldsymbol{P}) - \eta_{m,l}^{i} P_{m,l}(\boldsymbol{P}) \ge \mu$. Due to the minimization operator, $R_{m,l}(\boldsymbol{P}) - \eta_{m,l}^{i} P_{m,l}(\boldsymbol{P}) \ge \mu$ can

be satisfied for all LoRa users. For a given $\eta_{m,l}^{i}$, the problem is reformulated as

$$\max_{\boldsymbol{P}} \quad \mu \tag{5.22a}$$

$$s.t. \quad R_{m,l}\left(\boldsymbol{P}\right) - \eta_{m,l}^{i} P_{m,l}\left(\boldsymbol{P}\right) \geq \mu, \tag{5.22b}$$

$$0 \leq p_{m,l} \leq p_{\max}, \tag{5.22c}$$

$$\frac{p_{m,l}|h_{m,l}|^2}{\sigma_m^2} \geq \tau_{SF}, \ \forall \ m, \ l. \tag{5.22d}$$

where the difficulty only lies in the nonconvex constraint (5.22b).

The left side of constraint (5.22b) can be further denoted as $f_i(\boldsymbol{P}) - z_i(\boldsymbol{P})$, where $f_i(\boldsymbol{P})$ and $z_i(\boldsymbol{P})$ are defined as

$$f_i(\boldsymbol{P}) = B_m \log_2 \left(\sum_{k=1}^{S_m} p_{m,k}|h_{m,k}|^2 + \sigma_m^2 \right) - \eta_{m,l}^{i} P_{m,l}\left(\boldsymbol{P}\right), \tag{5.23a}$$

$$z_i(\boldsymbol{P}) = B_m \log_2 (\sum_{k=1,k\neq l}^{S_m} p_{m,k}|h_{m,k}|^2 + \sigma_m^2). \tag{5.23b}$$

Constraint (5.22b) is equivalent to that

$$f_i(\boldsymbol{P}) - z_i(\boldsymbol{P}) \geq \mu. \tag{5.24}$$

Moreover, we can find that $f_i(\boldsymbol{P})$ and $z_i(\boldsymbol{P})$ are both concave functions with respect to $\boldsymbol{P}$, thus inequality (5.24) is a DC programming function [105]. Due to the concave feature of $z_i(\boldsymbol{P})$, we can further approximate it by its upper bound with the first-Taylor expansion as follows:

$$z_i(\boldsymbol{P}) \leq z_i(\boldsymbol{P}^n) + \nabla {z_i}^T(\boldsymbol{P}^n)(\boldsymbol{P} - \boldsymbol{P}^n), \tag{5.25}$$

where $\boldsymbol{P}^n$ is the value of $\boldsymbol{P}$ at the n-th iteration, $\nabla {z_i}^T(\boldsymbol{P})$ represents the gradient of $z_i(\boldsymbol{P})$ that can be denoted as

$$\nabla z_i(\boldsymbol{P}) = \frac{\nu_{m,k}}{\sum\limits_{k=1,k\neq l}^{S_m} p_{m,k}|h_{m,k}|^2 + \sigma_m^2}. \tag{5.26}$$

Specifically, $\nu_{m,k}$ is a S_m-dimensional vector that can be given as

$$\nu_{m,k} = \begin{cases} 0 & \text{if} \quad k = l, \\ \frac{B_m |h_{m,k}|^2}{\ln 2} & \text{if} \quad k \neq l. \end{cases} \tag{5.27}$$

Table 5.6: DC-based Power Allocation Algorithm for MEE

1.Initialize $\boldsymbol{P}^0$, set $n=0$, $q^0=0$, $q^1=1$, and the tolerance $\epsilon=10^{-3}$.
2. **while** $\lvert q^{n+1}-q^n\rvert \geq \epsilon$
3. update $\boldsymbol{P}^n$ by solving problem (5.28).
4. update $q^n=\min_{m,l}[f_i(\boldsymbol{P}^n)-z_i(\boldsymbol{P}^n)]$.
5. update $n=n+1$.
6. **end while**
7. Output the optimal power allocation coefficients $\boldsymbol{P}^{opt}=\boldsymbol{P}^n$.

As a result, the original problem has been converted into the convex form, and during the n-th iteration, we need to tackle the following convex problem

$$\max_{\boldsymbol{P}} \quad \mu \tag{5.28a}$$

$$s.t. \quad f_i(\boldsymbol{P})-\left(z_i(\boldsymbol{P}^n)+\nabla {z_i}^T(\boldsymbol{P}^n)(\boldsymbol{P}-\boldsymbol{P}^n)\right)\geq \mu, \tag{5.28b}$$

$$0\leq p_{m,l}\leq p_{\max}, \tag{5.28c}$$

$$\frac{p_{m,l}\lvert h_{m,l}\rvert^2}{\sigma_m^2}\geq \tau_{SF},\ \forall\ m,\ l. \tag{5.28d}$$

Therefore, the detailed process of the DC programming approach to solve power allocations for MEE is outlined in Table 5.6.

To prove that the proposed DC-based power allocation algorithm converges, we have the following proposition.

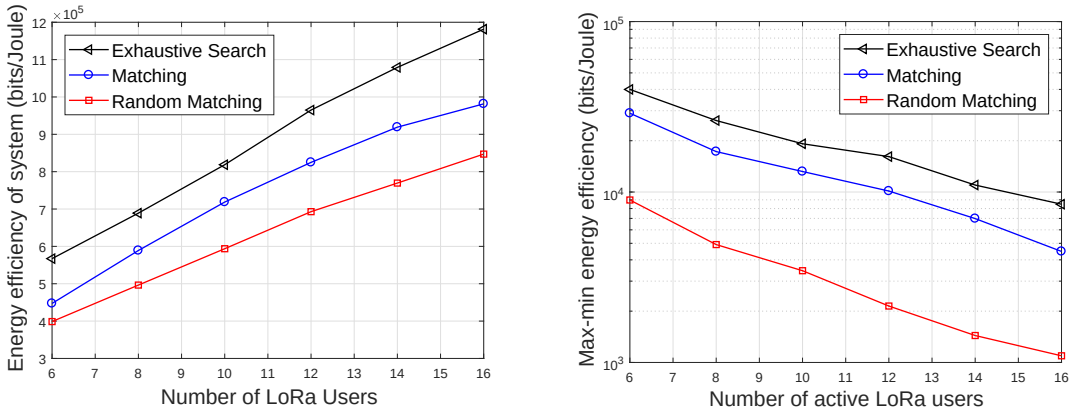
Proposition 9. *The proposed DC-based power allocation algorithm continuously converges to a stationary point of* (5.21) *with given* $\eta_{m,l}^i$.

Proof: The proof is provided in Appendix N.

5.4.3 Complexity Analysis

With regards to the energy efficient power allocation for SEE, denote $L_{max}^{(1)}$ as the maximum iteration number of the proposed power allocation algorithm given in Table 5.4, whereas the computational time to solve (5.16) by interior point method is proportional to $O(r^{3.5}\delta)$ [107], where r denotes the number of variables, and δ is the number of bits needed to represent the entries in the optimization problem. Therefore, the whole complexity to solve SEE is $O(L_{max}^{(1)}(N+1)^{3.5}\delta^{(1)})$.

As for MEE, the computational complexity comes from the algorithm provided in Table 5.5 and Table 5.6. Note that by appropriately setting the initial values



(a) SEE (b) MEE

Figure 5.2: Energy efficiency versus number of active LoRa users, N.

as $\eta_{m,l}^{min} \leq \eta_{MEE}^{opt} \leq \eta_{m,l}^{max}$, with a given accuracy ϵ, the complexity of the bisection method to solve (5.21) is $\log_2(\epsilon^{-1}(\eta_{m,l}^{max} - \eta_{m,l}^{min}))$. For the DC-based power allocation algorithm, denote the maximum iteration number as $L_{max}^{(2)}$, while the complexity of the interior point method to solve (5.28) is proportional to $O(N^{3.5}\delta^{(2)})$. In conclusion, the computational complexity of MEE is $O\left(\log_2(\epsilon^{-1}(\eta_{m,l}^{max} - \eta_{m,l}^{min}))L_{max}^{(2)}(N^{3.5}\delta^{(2)})\right)$.

5.5 Numerical Results

In this section, numerical results are provided to evaluate the performance of the proposed algorithms. In the simulations, the simulation parameters are set following LoRa specifications [55]. It is assumed that the LoRa gateway located in the cell center and all the LoRa users are uniformly distributed in a circular range with the radius of 12 km, which is consistent with LoRa characteristics to enable long-range transmission. The number of channels is set to be $M = 3$ working at 868 MHz. The bandwidth of each channel is set to be $B_m = 125$ kHz. We set the path loss factor to be $\alpha = 3.5$. Moreover, the duty cycle is set as 1% by following the LoRaWAN specification. The noise is defined as $\sigma^2 = -174 + 10\log_{10}B_m$ dBm. Without loss of generality, we assume that $P_c^1 = \cdots = P_c^N = P_c$, which indicates the same circuit power consumption is adopted for all LoRa users. Besides, the cross correlation factor ψ is a random variable between 0 and 1, which keeps the same for different SFs within a given channel realization, and ψ varies for different channel realizations.

Fig. 5.2 illustrates the effectiveness of the proposed user scheduling scheme versus the number of LoRa users with $p_{\max}$=20 dBm for both SEE and MEE. The results of the exhaustive search approach and random matching method are provided for

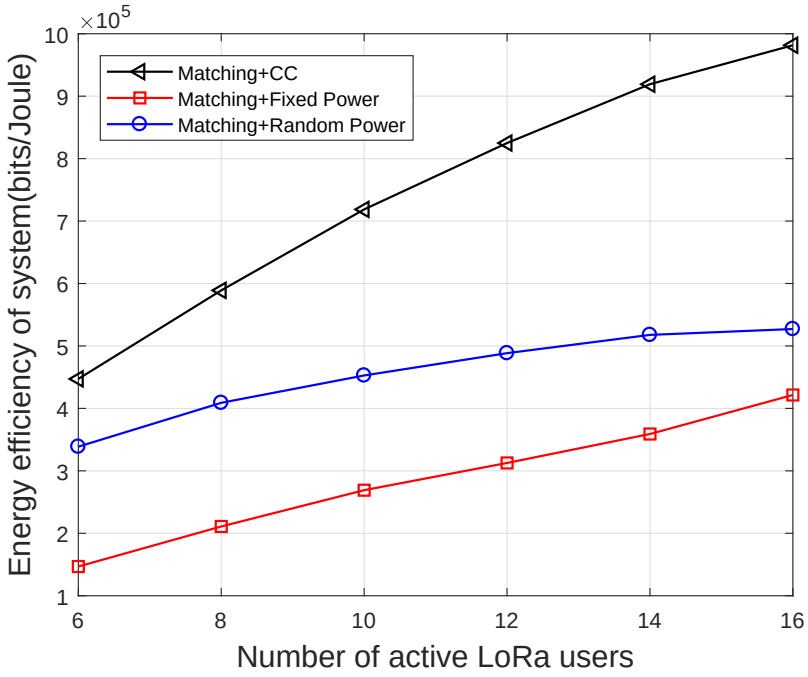


Figure 5.3: System energy efficiency comparison of the proposed power allocation schemes versus number of active LoRa users, N.

comparison. In the "random matching" scheme, the LoRa user randomly chooses a channel among $\boldsymbol{M}$, whereas adopting the proposed SF assignment and power allocation schemes. We observe that the system EE increases monotonically with the number of LoRa users for all the presented methods in the figure, while the max-min EE shows the inverse trend. It is noted that the performance of the proposed low-complexity user scheduling algorithm is very close to that of the exhaustive search method for both cases. Furthermore, the proposed matching algorithm yields much better performance than the random matching scheme. In addition, the gap between the proposed matching algorithm and random matching increases with the larger number of active LoRa users. The reason is that, when the number of active LoRa users increases, the intra-channel interference caused by LoRa users with different SFs can be well controlled by the proposed user scheduling scheme, whereas it cannot be suppressed by random user scheduling. Besides, random user scheduling scheme will schedule channels with poor conditions, which decreases the EE in terms of both SEE and MEE. Furthermore, the proposed matching algorithm plays a more important role in MEE design due to the fact that the max-min EE can be more easily affected by the interference.

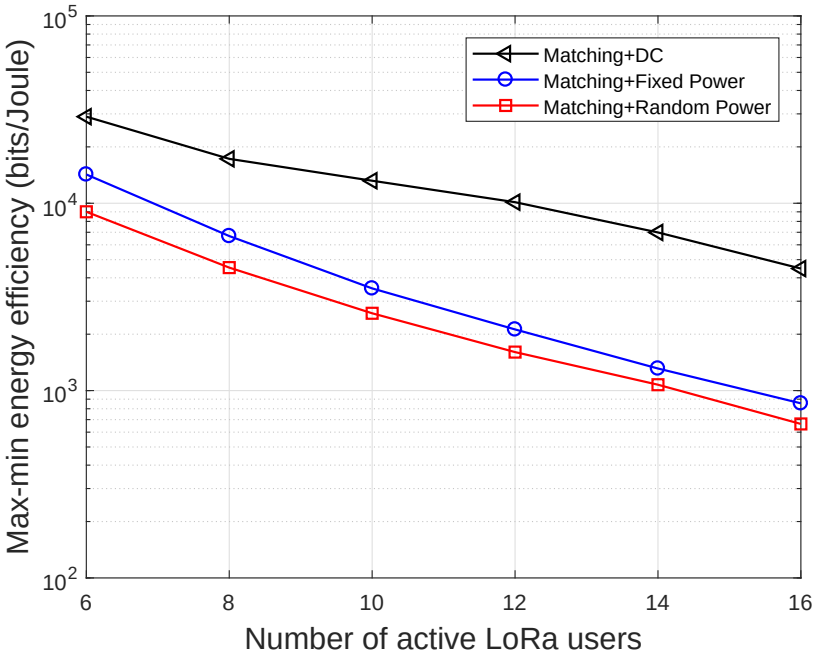


Figure 5.4: Max-min energy efficiency comparison of the proposed power allocation schemes versus number of active LoRa users, N.

In Fig. 5.3, the performance of the energy efficient power allocation schemes for SEE are evaluated with the active LoRa users ranges from 6 to 16 for $p_{\max}$=20 dBm. For simplification, the proposed resource allocation algorithm for SEE is denoted as "Matching+CC" where CC stands for Charnes-Cooper transformation. To provide fair comparison, the proposed user scheduling and SF assignment algorithms are adopted for all the three methods presented in the figure. As can be observed from Fig. 5.3, for SEE, the system EE performs increasing trends with the number of active LoRa users for all three schemes, and especially, the proposed 'matching+CC' algorithm produces the best performance among all three schemes. For instance, when the number of LoRa users is 12, the available system EE for matching+CC is 8.1×10^5 bits/Joule, while for matching+fixed power and matching+random power are 4.9×10^5 bits/Joule and 3.1×10^5 bits/Joule, respectively. Moreover, the advantage of the proposed scheme is more obvious when there are more LoRa users since the gap becomes larger. The reason is that, the intra-channel interference for LoRa users with different SFs cannot be effectively suppressed for random and fixed power allocation.

In Fig. 5.4, we provide the performance of the proposed power allocation scheme named as "Matching+DC" with the different number of LoRa users for MEE. It shows that the trend for three curves is similar, and the proposed "Matching+DC" scheme

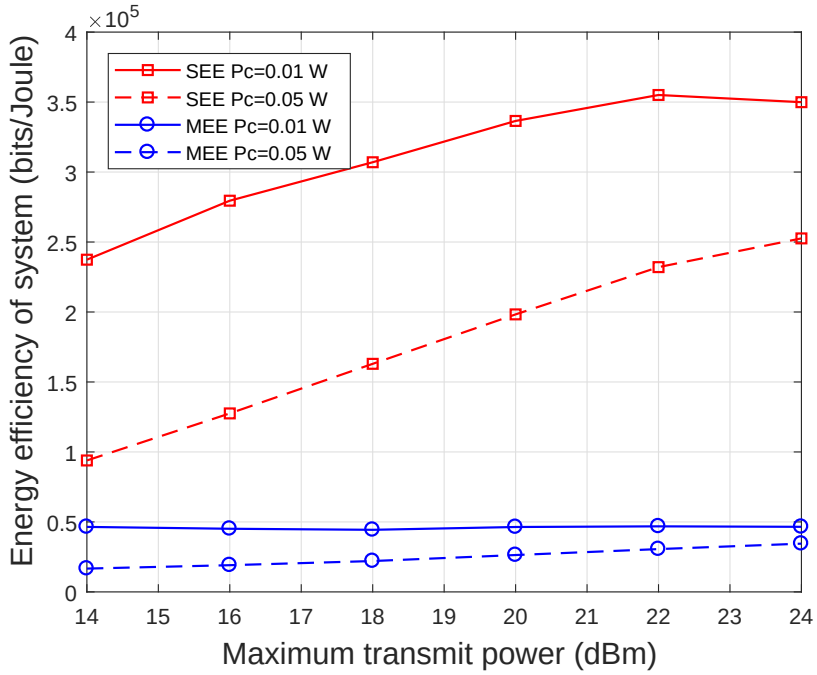


Figure 5.5: Energy efficiency of system comparison between SEE and MEE methods versus $p_{\max}$.

yields the best system performance among all three schemes. It is noted that the max-min EE descends with the increasing number of available users, due to the fact that increasing the number of active LoRa users results in more users scheduled in one channel, which increases interference caused by LoRa users allocated in one channel. As a result, the max-min EE declines. In addition, the benefit of the proposed power allocation scheme becomes more obvious with the increasing number of active LoRa users as the gap becomes larger, which is the same as that of Fig. 5.3.

The system EE performance comparison between SEE and MEE is illustrated in Fig. 5.5, based on the same random channel realization. We plot the relationship between system EE and the maximum transmit power for both SEE and MEE. We can see that the SEE design achieves significantly higher system EE compared with MEE design. The system EE gap between $P_c = 0.01$ W and $P_c = 0.05$ W decreases with $p_{\max}$ for both SEE and MEE design. This is because the proportion of circuit power consumption decreases with the increasing $p_{\max}$, and the effect of circuit power consumption is more obvious in the low power regime. Moreover, when $P_c = 0.01$ W, the system EE firstly increases with $p_{\max}$ and reaches the peak at 22 dBm, and then it decreases.

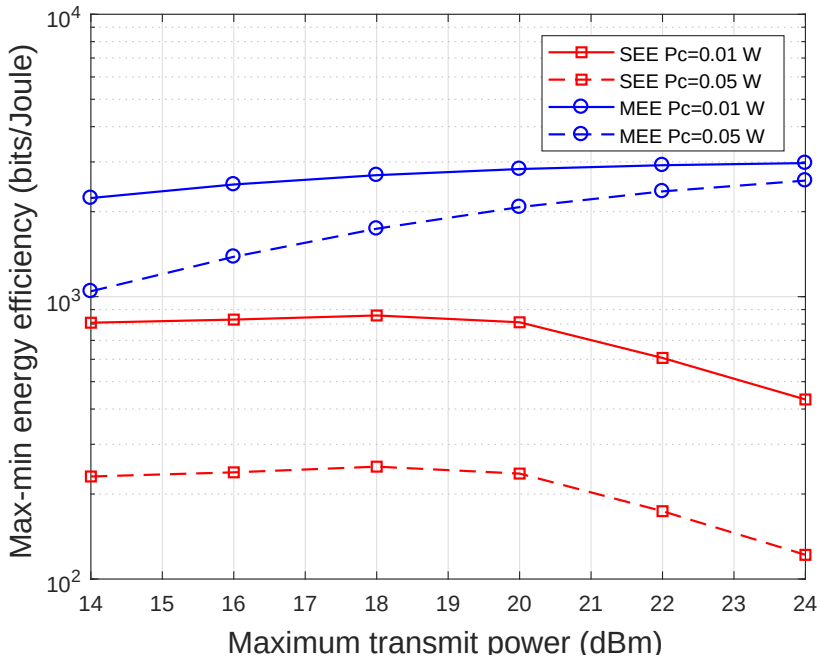


Figure 5.6: Max-min energy efficiency comparison between SEE and MEE methods versus $p_{\max}$.

Fig. 5.6 is plotted to compare the max-min EE performance for SEE and MEE with the given random channel, which indicates that the max-min EE for MEE slightly improves with the increasing $p_{\max}$. However, for SEE, the max-min EE firstly keeps constant, and then decreases. The reason is that, when $p_{\max}$ becomes larger, more resources will be allocated to users with bad channel conditions, which guarantees fairness among users for MEE design. On the other hand, to improve system performance, SEE will schedule more power to users in good channel conditions and sacrifice the benefit of bad users, and the corresponding max-min EE decreases. Moreover, for MEE, the gap between $P_c = 0.01$ W and $P_c = 0.05$ W becomes smaller with $p_{\max}$, due to the reason that $p_{\max}$ dominates P_c in the high power regime. Combining Fig. 5.5 and Fig. 5.6, it can be seen that SEE and MEE have completely different preferences in EE design, and the network can adopt the corresponding strategies according to the system requirement.

5.6 Summary

In this chapter, we have investigated the uplink LoRa networks to maximize the EE of the whole network and the minimal EE of LoRa users, named as SEE and MEE,

respectively. Particularly, we decompose the formulated problems into three subproblems, including user scheduling, SF assignment, and power allocation. Moreover, we have proposed a low-complexity user scheduling scheme to solve the channel assignment problem by formulating it as a many-to-one two-sided matching problem with peer effects. The SF is assigned to LoRa users scheduled on the same channel based on the distance between LoRa user and gateway, which is obtained from the stable matching. Moreover, for energy efficient power allocation to maximize SEE, we approximate the fractional nonconvex function by its lower bound, which can be further transformed into convex approximations with Charnes-Cooper transformation. While to maximize MEE, an iterative method based on generalized fractional programming and DC programming has been proposed. Numerical results have shown that the proposed matching algorithms and power allocation schemes outperform the existing schemes in terms of both SEE and MEE.

Chapter 6

Conclusions and Future Work

In this chapter, the contributions of this thesis are first summarized. Furthermore, several potential research extensions and promising future directions related to the works in this thesis are presented.

6.1 Summary

This thesis mainly focuses on efficient resource allocation optimization design for several potential future wireless communication technologies, i.e., NOMA, MEC, and LoRa. We first investigated the strong user's data rate optimization problem for the SWIPT-enabled cooperative NOMA system considering imperfect CSI. Then we studied the application of NOMA and UC in wireless powered MEC systems. Moreover, the energy efficient transmission for uplink LoRa networks was addressed to maximize SEE and MEE.

The main contributions and insights of this thesis can be summarized as follows:

In Chapter 3, a novel robust cooperative NOMA scheme was proposed, where the strong user acts as DF relay to transmit information to the weak user, and only imperfect CSI was assumed to be available at the BS. The robust beamforming design for SWIPT-enabled cooperative NOMA with channel uncertainties was investigated to maximize the strong user's data rate. To present a comprehensive study, two major channel uncertainties design criteria in terms of outage-based design and worst-case based optimization were adopted. For the outage-based constraint design, the formulated probabilistic nonconvex problem was first transformed into a non-probabilistic problem with Bernstein-type inequality, which can be globally solved by a two-dimensional exhaustive search method. To reduce the high computational complexity introduced by the exhaustive approach, the low-complexity suboptimal solution based on SCA was further proposed. With regards to the worst-case based

optimization problem, an iterative algorithm based on SDR and SCA was proposed. It was demonstrated that the proposed algorithms converged within a few iterations and achieved near-optimal performance. Moreover, numerical results showed that the proposed robust design for the SWIPT-enabled cooperative system outperforms the benchmark schemes.

In Chapter 4, the application of NOMA and UC in wireless powered MEC system under the non-linear energy harvesting model was exploited. A computation efficiency maximization problem was formulated, subject to the QoS requirement as well as the harvested energy limitations at the users. SDR technique was first utilized to transform the nonconvex problem into a more tractable form, and the rank-one optimality was proved to establish the equivalence of the transformation. The reformulated problem was converted into the convex expression with the aid of SCA, which can be further solved with the proposed iterative algorithm. The convergence and complexity analysis was provided. The performance of the proposed algorithm was verified by computer simulations. Numerical results showed that the partial offloading design outperforms offloading only and local computing schemes. Furthermore, it was demonstrated that NOMA plays a great role when the distance of two users is relatively far, while UC contributes to performance improvement when two users are closer. Combing both NOMA and UC, it was shown that the proposed NOM-UC MEC scheme yields the best system performance

In Chapter 5, the energy efficient resource allocation problem was formulated for uplink LoRa networks to maximize SEE and MEE, respectively. To avoid the considerable complexity of the global optimum solution due to the NP-hardness of the original problems, we decomposed it into three sub-problems as user scheduling, S-F assignment, and power allocations. A low-complexity suboptimal solution based on matching theory was first proposed to enable the self-matching of LoRa users with proper channels. For LoRa users scheduled on the same channel, a distance-based heuristic algorithm was provided to realize efficient SF assignment. To allocate power across channels, centralized optimal power allocation algorithms that can be implemented at the LoRa gateway were proposed. Specifically, by deriving the lower bound approximation of MEE, it was further transformed into convex expression by applying Charnes-Cooper transformation. Moreover, by utilizing generalized fractional programming and DC programming, an iterative power allocation algorithm was proposed to maximize MEE. Simulation results revealed the unique design preference feature of SEE and MEE. Also, it was shown that the proposed user scheduling and

power allocation scheme significantly improved the EE performance compared with the existing schemes.

6.2 Future Work

Based on the current outcome of this thesis, the extensions of present works and some promising topic directions can be further conducted in future work, which is summarized as below:

1. Spectral-energy efficiency tradeoff for NOMA system with imperfect CSI: The present researches of NOMA focus on either both SE and EE maximization with only perfect CSI or SE maximization with imperfect CSI (In Chapter 3 of this thesis). On one hand, to avoid the additional system overhead and SE reduction introduced by obtaining perfect CSI, imperfect CSI scenario is more practical from the implementation perspective. On the other hand, EE is a significant concern for state-of-art communication design. Based on the above observations, it motivates us to investigate the spectral-energy efficiency tradeoff of the NOMA system with imperfect CSI.

2. MEC system with machine learning: The application of machine learning in the wireless communication system is envisioned to greatly enhance system performance [115,116]. The advantages of combing machine learning with MEC are threefold: i) Intelligent AI-model achieves better performance in channel control and offloading selection. ii) By deploying MEC servers in proximity to mobile devices, real-time data process can be realized at edge learning. iii) Due to the powerful processing capability, edge learning can support much more complex AI models than on-device learning. Despite the above benefits, how to integrate machine learning with the MEC system is still an open challenge.

3. UAV-assisted MEC network with NOMA: Different from traditional MEC networks where the location of MEC servers is fixed, UAV-assisted architecture provides adjustable deployment of servers. The line-of-sight (LoS) transmit links further improve communication performance. Note that NOMA can improve SE, it is natural to investigate the application of NOMA with UAV-assisted MEC network to increase the computation capacity.

4. Energy efficient LoRa design considering both co-SF (collisions of devices with the same SFs) and inter-SF (collisions of devices with different SFs) interferences: Note that only inter-SF interference was investigated for the energy efficient uplink transmissions in LoRa networks in Chapter 5. However, due to the large number of LoRa devices, it is shown that both inter-SF and co-SF interferences are prevalent in LoRa networks. Combing both interferences, the system model and problem formulation will be extremely complicated. Therefore, how to make energy efficient design is a challenging but important future research topic.

Appendix A

Proof of Proposition 2

Proof. For any given β and θ, problem (3.37) can be degraded into the following problem:

$$\max_{\beta, \boldsymbol{W}_1, \boldsymbol{W}_2} \min_{\tilde{\boldsymbol{e}}_h \in \epsilon_h} \quad \mu \tag{A.1a}$$

$$s.t. \quad (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1})^H \boldsymbol{W}_1 (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1}) \geq \frac{\mu}{1-\beta}, \tag{A.1b}$$

$$(\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1})^H \boldsymbol{W}_2 (\tilde{\boldsymbol{h}}_1 + \tilde{\boldsymbol{e}}_{h_1}) \geq \frac{\gamma(\mu+1)}{1-\beta}, \tag{A.1c}$$

$$\frac{\mu}{1-\beta} + \frac{\gamma(\mu+1)}{1-\beta} \leq \frac{\gamma-\theta}{\beta|g|^2}, \tag{A.1d}$$

$$(\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2})^H (\boldsymbol{W}_2 - \theta \boldsymbol{W}_1)(\tilde{\boldsymbol{h}}_2 + \tilde{\boldsymbol{e}}_{h_2}) \geq \theta, \tag{A.1e}$$

$$\gamma(\mu+1) \geq \mu, \tag{A.1f}$$

$$Tr(\boldsymbol{W}_1) + \mathrm{Tr}(\boldsymbol{W}_2) \leq P_{\max}, \tag{A.1g}$$

$$\boldsymbol{W}_1, \boldsymbol{W}_2 \succeq \boldsymbol{0}. \tag{A.1h}$$

Particularly, (A.1d) is acquired by substituting the constraints of (3.37b) and (3.37c) into (3.37d) and the inequality can be satisfied based on the fact that the summation of two individual lower bound values is always smaller or equal to the global lower bound. Constraint (A.1f) is obtained by replacing constraint (3.37f) with (3.7b) and (3.37c), and constraint (3.37g) is omitted as it can be satisfied if (A.1e) holds since $\theta > 0$. Assume that problem (A.1) is feasible and it is also dual feasible. As can be seen from problem (A.1), there are four linear constraints (A.1b, A.1c, A.1e and A.1g) related to the optimal solution $(\boldsymbol{W}_1{}^*, \boldsymbol{W}_2{}^*)$ and according to [117, Theorem 3.2], we have that

$$\mathrm{rank}^2(\boldsymbol{W}_1{}^*) + \mathrm{rank}^2(\boldsymbol{W}_2{}^*) \leq 4. \tag{A.2}$$

If problem (A.1) is feasible, from (A.1b), we can find that $\boldsymbol{W}_1{}^* \succeq \boldsymbol{0}$ and $\boldsymbol{W}_1{}^* \neq \boldsymbol{0}$; from (A.1c), we have that $\boldsymbol{W}_2{}^* \succeq \boldsymbol{0}$ and $\boldsymbol{W}_2{}^* \neq \boldsymbol{0}$. Further, with constraint (A.2)

considered, we can conclude that only when $\text{rank}(\boldsymbol{W}_i) = 1, i = 1, 2$, the inequality (A.2) can be satisfied. Hence, we can conclude that problem (3.37) always has an optimal solution $\boldsymbol{W}_1{}^*$ and $\boldsymbol{W}_2{}^*$. Then, the optimal beamforming vector $\boldsymbol{w}_1{}^*$ and $\boldsymbol{w}_2{}^*$ can be respectively obtained from $\boldsymbol{W}_1{}^*$ and $\boldsymbol{W}_2{}^*$ by using eigen-decomposition. Otherwise, a suboptimal solution can be attained by Gaussian randomization. The proof is completed.

□

Appendix B

Proof of Proposition 3

Proof. In order to prove that the acquired objective value is non-decreasing for each iteration, we need to first demonstrate that the solution to problem (3.49) at the n-th iteration is also a feasible point for the iteration $(n+1)$.

Let us assume that the optimal solutions to problem (3.49) at the n-th iteration are $\boldsymbol{W}_1^*$, $\boldsymbol{W}_2^*$, ι_1^*, ι_2^*, ι_3^*, and a^*. The constraints which use the SCA-based method to get convex approximation are constraints (3.46a), (3.46b) and (3.46c). Here, we take the constraint (3.46a) as an example.

$$2\iota_1^{(n)}\iota_1^* - (\iota_1^{(n)})^2 \geq \mu. \tag{B.1}$$

We then replace the variables at the iteration (n+1) with the optimal solutions obtained in iteration n, e.g., $\iota_1^{(n+1)} = \iota_1^*$. It is obvious that the constraints (3.46a) and (3.45a) can be satisfied. In addition, during the iteration of $(n+1)$ for (3.45b) with the updated parameter, the following result can be obtained:

$$\begin{aligned} 2\iota_1^*\iota_1^* - (\iota_1^*)^2 &= \iota_1^{*2} && \text{(B.2a)} \\ &\geq 2\iota_1^{(n)}\iota_1^* - (\iota_1^{(n)})^2 && \text{(B.2b)} \\ &\geq \mu, && \text{(B.2c)} \end{aligned}$$

where (B.2a) is derived by substituting the solutions of iteration n. The inequality (B.2b) is gained by performing the first-order Taylor approximation for ι_1^{*2} around ι_1^* which is a lower bound of the original function. Finally, we can get (B.2c) with the application of (B.1). Similarly, the optimal solutions obtained at the n-th iteration also satisfy the constraints (3.46b) and (3.46c) for the iteration $n+1$. The detailed analysis for the constraints (3.46b) and (3.46c) at iteration $(n+1)$ is omitted here, but can be provided following similar steps.

In conclusion, it can be proved that the optimal solution of the n-th iteration obtained from Table 3.3 is a feasible point for problem (3.49) at the $(n+1)$-th iteration. As problem (3.49) is a concave problem, the objective value at the $(n+1)$-th iteration is larger or equal to that achieved from the n-th iteration. Hence, the proof is completed and the proposition is proved. □

Appendix C

Proof of Proposition 4

Proof. Let $\Lambda^n = \{\boldsymbol{W}_1^n, \boldsymbol{W}_2^n, \iota_1^n, \iota_2^n, \iota_3^n, a^n\}$ be the solution derived from Table 3.3 during the n-th iteration. According to Proposition 3, we have that $\Lambda^n \to \Lambda^*$ as $n \to \infty$ where Λ^* represents the optimal solution to (3.49). Besides, with the application of the SCA method, the introduced lower bound of (3.46a) has the same value and gradient value around the point Λ^n for any iterations (which still holds as $n \to \infty$). Therefore, we can conclude that the proposed algorithm provided in Table 3.3 can continuously coverage to a KKT point of problem (3.37) when the iteration number tends to infinity based on the above property. □

Appendix D

Proof of Lemma 3

Proof. Lemma 3 can be proved by contradiction approach. Suppose that $\{\boldsymbol{w}^*, \boldsymbol{t}^*, \boldsymbol{p}^*, \boldsymbol{\ell}^*\}$ is the optimal solution to problem (4.14) corresponding to the maximum objective $\tilde{\eta}^*$, and the time allocation satisfies $t_0^* + t_1^* + t_{21}^* + t_{22}^* < T$. Based on the expression of (4.14a), with fixed t_0^*, $\tilde{\eta}$ can be further improved as increasing $\{t_1 + t_{21} + t_{22}\}$ results in larger computation bits in the numerator, contradicting that the solution is optimal. Therefore, the maximum computation efficiency can be achieved with $t_0 + t_1 + t_{21} + t_{22} = T$. □

Appendix E

Proof of Proposition 5

Proof. Assume that problem (4.14) is feasible and it is also dual feasible. As can be seen from problem (4.36), there are three linear constraints (4.16c, 4.16f and 4.36f) related to the $\boldsymbol{W}^*$. According to [117, Theorem 3.2], we have that

$$\operatorname{rank}^2(\boldsymbol{W}^*) \leq 3. \tag{E.1}$$

If problem (4.36) is feasible, we can infer that $\boldsymbol{W}^* > \mathbf{0}$, according to (4.16f) and (4.36f). Moreover, with constraint (E.1) considered, we can conclude that only when $\operatorname{rank}(\boldsymbol{W}^*) = 1$, the inequality (E.1) can be satisfied. Hence, the relaxation is tight, and one can conclude that problem (4.36) always has an optimal solution $\boldsymbol{W}^*$.

Furthermore, it is worth noting that problem (4.36) is a convex optimization problem, hence the global optimal solution $(\boldsymbol{W}^*, \boldsymbol{t}^*, \boldsymbol{E}^*)$ can be obtained with the interior point method. If $\operatorname{rank}(\boldsymbol{W}^*) = 1$, we can get that $\boldsymbol{W}^* = \frac{\boldsymbol{W}^*}{t_0}$, and the optimal beamforming vector $\boldsymbol{w}$ can be computed from $\boldsymbol{W}^*$ by using eigen-decomposition. Otherwise, a suboptimal solution can be attained by Gaussian randomization [117]. □

Appendix F

Proof of Proposition 6

Proof. To reveal the convergence of the proposed algorithm in Table 4.1, we need to demonstrate that the sequence of the objective values obtained from Table 4.1 is non-decreasing for each iteration, i.e., $\mu^{(n+1)} \geq \mu^{(n)}$.

Denote $\boldsymbol{W}^*, \boldsymbol{t}^*, \boldsymbol{E}^*, \mu^*, \Upsilon^*$ as the optimal solution to problem (4.36) during the n-th iteration. Note that during the problem transformation, constraints (4.19), (4.21c), (4.24), (4.31), and (4.34) are approximated with SCA. Take (4.24) as an example, denoting that $\ell = \sum_{i=1}^{2} \ell_i^{loc} + \ell_i^{off}$, we have that

$$\ell \geq \sqrt{\mu^{(n)}\Upsilon^{(n)}} + 0.5\sqrt{\frac{\mu^{(n)}}{\Upsilon^{(n)}}}(\Upsilon^* - \Upsilon^{(n)}) + 0.5\sqrt{\frac{\Upsilon^{(n)}}{\mu^{(n)}}}(\mu^* - \mu^{(n)}). \tag{F.1}$$

The variables in iteration $n+1$ are updated accordingly, i.e., $\mu^{(n+1)} = \mu^{(*)}$, $\Upsilon^{(n+1)} = \Upsilon^{(*)}$, while (4.22b) can still be satisfied. By substituting the updated parameters into (4.24) during iteration $n+1$, we have

$$\begin{aligned}
&\sqrt{\mu^{(*)}\Upsilon^{(*)}} + 0.5\sqrt{\frac{\mu^{(*)}}{\Upsilon^{(*)}}}(\Upsilon^* - \Upsilon^{(*)}) + 0.5\sqrt{\frac{\Upsilon^{(*)}}{\mu^{(*)}}}(\mu^* - \mu^{(*)}) \\
&= \sqrt{\mu^{(*)}\Upsilon^{(*)}} && \text{(F.2a)} \\
&\leq \sqrt{\mu^{(n)}\Upsilon^{(n)}} + 0.5\sqrt{\frac{\mu^{(n)}}{\Upsilon^{(n)}}}(\Upsilon^* - \Upsilon^{(n)}) + 0.5\sqrt{\frac{\Upsilon^{(n)}}{\mu^{(n)}}}(\mu^* - \mu^{(n)}) && \text{(F.2b)} \\
&\leq l, && \text{(F.2c)}
\end{aligned}$$

where (F.2a) is derived by replacing $(\mu^{n+1}, \Upsilon^{n+1})$ with the obtained optimal solution (μ^*, Υ^*), (F.2b) is the upper-bounded approximation of (F.2a), and inequality (F.2c) is deduced with the aid of (F.1). Similar steps can be applied to prove the convergence of (4.19), (4.21c), (4.31) and (4.34), where the detailed process is omitted here.

In summary, it is proved that the solution of the n-th iteration is a feasible point of iteration $n+1$ for problem (4.36). Based on the above analysis, and due to the fact that the objective of problem (4.36) is a concave function, $\mu^{(n+1)} \geq \mu^{(n)}$ is proved. The proof is completed. □

Appendix G

Proof of Proposition 7

Proof. Moreover, denote χ^n as the optimal solutions to problem (4.36) during iteration n of Table 4.1, due to the convergence feature of the proposed algorithm introduced by Proposition 6, $\chi^n \to \chi^*$ holds when $n \to \infty$, where χ^* denotes the optimal solution to problem (4.36). In addition, we note that (4.36) is obtained from problem (4.14) by performing SCA, while the bound approximation introduced by the SCA has the same function value and gradient value around the original spatial point for any iterations. Therefore, we can conclude that the algorithm derived from Table 4.1 can continuously coverage to a KKT point. The proof is completed. □

Appendix H

Proof of Theorem 1

Proof. Take SEE as an example, Theorem 1 can be proved by considering two cases, i.e, $\Lambda_{max} = 1$ and $\Lambda_{\max} > 1$.

- When $\Lambda_{\max} = 1$, the original SEE reduces to the joint channel and power allocations for system EE maximization problem in an OFDMA system, the NP-harness has already been proved in [104].

- For the case of $\Lambda_{\max} > 1$, we prove that SEE is NP-hard even without considering power allocations. We construct a case of SEE with given power allocation coefficients and the NP-hardness can be proved by establishing the equivalence between the constructed instance and 3-dimensional matching problem, which is known to be NP-hard. The instance with N LoRa users, M channels, and $\Lambda_{\max} = 2$ is considered. Let $\mathfrak{X}$ and $\mathfrak{Y}$ be two different sets with $|\mathfrak{X}| = |\mathfrak{Y}| = \frac{N}{2}$ and V be a subset of $M \times \mathfrak{X} \times \mathfrak{Y}$. Assuming that any tripe $V_i = (m_i, x_i, y_i) \in V$, which means LoRa users $x_i \in \mathfrak{X}$, $y_i \in \mathfrak{Y}$ are selected on channel $m_i \in M$. With given power allocation coefficients, denote the maximized sum rate with any given V_i as $\mathfrak{R}_{V_i}$. Hence, we just need to verify the 3-dimensional problem if there exists $V^{'} \subset V$, satisfing that 1) $m_1 \neq m_2$, $x_1 \neq x_2$, and $y_1 \neq y_2$ for any two triples $(m_1, x_1, y_1) \in V^{'}$ and $(m_2, x_2, y_2) \in V^{'}$. 2)$V^{'}$=min$\{M, \frac{N}{2}\}$.

According to the definition, if the feasibility problem is proved to be NP-hard, then the original problem is also NP-hard [118]. Therefore, denote the sum rate for any triple as $R_{V_i^{'}}$, given the power allocations, the sum rate feasibility problem can be expressed as $\sum_{i=1}^{i=V^{'}} R_{V_i^{'}} \leq \epsilon$, where ϵ is a given constant. When ϵ becomes positive infinity, an instance of the feasibility problem corresponds to a 3-dimensional matching problem, then a special case of the original SEE is NP-harness, which proves

the original SEE is NP-hard. The proof for MEE is similar and omitted here. The proof is completed. □

Appendix I

Proof of Theorem 2

Proof. Theorem 2 is proved from two aspects. Firstly, the number of possible swap matching operations is finite since only a limited number of LoRa users can occupy the same channel. In addition, Due to the feature of swap matching given by Definition 2, if a swap-matching is approved, the achievable data rates of any player, i.e., U_l and SC_m, will not decrease by employing a swap matching, and the data rates of at least one player will increase. Therefore, the corresponding objectives, i.e., (5.7a) and (5.8a), will increase after each swap matching operation. The spectrum resources are limited, which restricts the upper bound of energy efficiency. Hence, there is a swap matching after which no further energy efficiency is improved and the algorithm of Table 5.2 converges to a $2ES$ matching. The proof is completed. □

Appendix J

Proof of Theorem 3

Proof. The initialization step is a deferred acceptance algorithm, the complexity depends on the process of user proposing, which is up to $O(MN)$ in the worst case. Besides, the computational complexity of the swap matching step lies in the number of iterations and swap operations. In each iteration, for any channel SC_m, the maximum assigned users is $\Lambda_{\max}$. For user U_j, there exists $(M-1)$ possible swap-blocking pairs in φ_l^j. The potential combinations for φ_l^j with j fixed is $\Lambda_{\max}(M-1)$. Since there are N LoRa users, we can conclude that the number of swap matchings is $\Lambda_{\max}N(M-1)$ during each iteration. Considering the number of iterations, the total complexity of swap matching is $O(\frac{1}{2}I\Lambda_{\max}N(M-1))$. Combining the above two phases, the complexity of the algorithm in Table 5.2 is $O(MN+\frac{1}{2}I\Lambda_{\max}N(M-1))$ in the worst case. □

Appendix K

Proof of Proposition 8

Proof. With inequality (5.11) and equation (5.14), the original SEE has been transformed into a concave-convex problem, which implies that KKT conditions are sufficient and any local maximum is the global maximum [119]. Besides, a feasible point set $\{\tilde{x}_{m,l}^n\}$ and $\{\phi^n\}$ at iteration n can be obtained by solving problem (5.16), the power allocated to each user is calculated as ${p_{m,l}}^n = e^{\frac{\tilde{x}_{m,l}^n}{\phi^n}}$. The system EE η^n can be further obtained according to (5.7a). Moreover, problem (5.16) is concave in $(\tilde{x}, \phi)$, with the interior point method, we can derive that $\eta^n \leq \eta^{n+1}$ which means the algorithm provided in Table 5.4 converges. The proof is completed. □

Appendix L

Proof of Theorem 4

Proof. The proof can be proved from two aspects. We first prove the necessity. From equation (5.19), we can deduce that

$$\eta_{MEE}^{opt} = \min_{m,l} \frac{R_{m,l}\left(\boldsymbol{P}^{opt}\right)}{P_{m,l}\left(\boldsymbol{P}^{opt}\right)} \geq \min_{m,l} \frac{R_{m,l}\left(\boldsymbol{P}\right)}{P_{m,l}\left(\boldsymbol{P}\right)}, \tag{L.1}$$

which is equivalent to the following

$$\min_{m,l}[R_{m,l}\left(\boldsymbol{P}\right) - \eta_{MEE}^{opt} P_{m,l}\left(\boldsymbol{P}\right)] \leq 0, \tag{L.2a}$$

$$\min_{m,l}[R_{m,l}\left(\boldsymbol{P}^{opt}\right) - \eta_{MEE}^{opt} P_{m,l}\left(\boldsymbol{P}^{opt}\right)] = 0. \tag{L.2b}$$

Therefore, the maximum value for the left side of (L.2a) can be achieved if and only if $\boldsymbol{P} = \boldsymbol{P}^{opt}$, which completes the necessity proof.

The we prove the sufficiency of Theorem 4. Assuming $\boldsymbol{P}^{'}$ as the optimal power allocation to (5.19), then for any feasible power allocation coefficient $\boldsymbol{P}$, we have the following function

$$\min_{m,l}[R_{m,l}\left(\boldsymbol{P}\right) - \eta_{MEE}^{opt} P_{m,l}\left(\boldsymbol{P}\right)] \leq \min_{m,l}[R_{m,l}\left(\boldsymbol{P}^{'}\right) - \eta_{MEE}^{opt} P_{m,l}\left(\boldsymbol{P}^{'}\right)] = 0. \tag{L.3}$$

As a result, we can obtain that

$$\eta_{MEE}^{opt} = \min_{m,l} \frac{R_{m,l}\left(\boldsymbol{P}^{'}\right)}{P_{m,l}\left(\boldsymbol{P}^{'}\right)}, \tag{L.4}$$

which means that $\boldsymbol{P}^{'}$ is the optimal power allocation of the original problem. The proof is completed. □

Appendix M

Proof of Theorem 5

Proof. Let us assume that $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$ be the optimal power solution corresponding to given max-min EE $\eta^1_{m,l}$ and $\eta^2_{m,l}$, with the condition $\eta^1_{m,l} > \eta^2_{m,l}$, then we can deduce the following

$$\pi(\eta^1_{m,l}) = \max_{\boldsymbol{P}} \min_{m,l} [R_{m,l}(\boldsymbol{P}) - \eta^1_{m,l} P_{m,l}(\boldsymbol{P})] \tag{M.1a}$$

$$= \min_{m,l} [R_{m,l}(\boldsymbol{P}^1) - \eta^1_{m,l} P_{m,l}(\boldsymbol{P}^1)] \tag{M.1b}$$

$$< \min_{m,l} [R_{m,l}(\boldsymbol{P}^1) - \eta^2_{m,l} P_{m,l}(\boldsymbol{P}^1)] \tag{M.1c}$$

$$\leq \min_{m,l} [R_{m,l}(\boldsymbol{P}^2) - \eta^2_{m,l} P_{m,l}(\boldsymbol{P}^2)] = \pi(\eta^2_{m,l}), \tag{M.1d}$$

where inequality (M.1c) is derived from the condition that $\eta^1_{m,l} > \eta^2_{m,l}$, (M.1d) is obtained due to the fact $\boldsymbol{P}^2$ is the optimal solution to $\eta^2_{m,l}$. Hence, $\pi(\eta_{m,l})$ is monotonically decreasing with $\eta_{m,l}$. In addition, incorporating the conclusion gained from Theorem 4, we can easily derive the property *ii*) of Theorem 5. □

Appendix N

Proof of Proposition 9

Proof. In order to prove the convergence of algorithm provided in Table 5.6, we need to confirm that the optimal solution to problem (5.28) at n-th iteration is also a feasible point of the iteration $n+1$. From the characteristic of inequality (5.28b), denote $\boldsymbol{P}^*$ as the optimal solution to problem (5.28) at iteration n, the following result can be obtained:

$$
\begin{aligned}
u &\leq q^n && \text{(N.1a)}\\
&= \min_{m,l} f_i(\boldsymbol{P}^*) - \big(z_i(\boldsymbol{P}^n) + \nabla {z_i}^T(\boldsymbol{P}^n)(\boldsymbol{P}^* - \boldsymbol{P}^n)\big) && \text{(N.1b)}\\
&= \min_{m,l} f_i(\boldsymbol{P}^{n+1}) - \big(z_i(\boldsymbol{P}^n) + \nabla {z_i}^T(\boldsymbol{P}^n)(\boldsymbol{P}^{n+1} - \boldsymbol{P}^n)\big) && \text{(N.1c)}\\
&\leq \min_{m,l} [f_i(\boldsymbol{P}^{n+1}) - z_i(\boldsymbol{P}^{n+1})] = q^{n+1}, && \text{(N.1d)}
\end{aligned}
$$

where the equality (N.1c) holds as $\boldsymbol{P}^{n+1}$ is the optimal solution of n-th iteration. Besides, inequality (N.1d) can be derived with inequality (5.25), since the first-Taylor approximation is an upper bound of $z_i(\boldsymbol{P})$.

In conclusion, we can obtain that the objective value at the iteration $n+1$ is larger or equal to that achieved from the n-th iteration, which proves the convergence of the proposed algorithm.

Moreover, denote constraint (5.28b) as $\varsigma(\boldsymbol{P}) = f_i(\boldsymbol{P}) - \big(z_i(\boldsymbol{P}^n) + \nabla {z_i}^T(\boldsymbol{P}^n)(\boldsymbol{P} - \boldsymbol{P}^n)\big)$. Since the proposed algorithm converges, then $\boldsymbol{P}^n = \boldsymbol{P}^{n+1}$ when $n \to \infty$. The first-order optimality condition [67] can be written as

$$
\begin{aligned}
\nabla {\varsigma}^T(\boldsymbol{P}^n)(\boldsymbol{P} - \boldsymbol{P}^n) &= (\nabla f_i^T(\boldsymbol{P}^n) - \nabla z_i^T(\boldsymbol{P}^n))(\boldsymbol{P} - \boldsymbol{P}^n) && \text{(N.2a)}\\
&= \nabla {\varsigma}^T(\boldsymbol{P}^{n+1})(\boldsymbol{P} - \boldsymbol{P}^{n+1}) && \text{(N.2b)}\\
&\leq 0. && \text{(N.2c)}
\end{aligned}
$$

Consequently, the first-order optimality condition of problem (5.21) is confirmed, which means the result obtained from Table 5.6 satisfies KKT conditions and is a stationary point of (5.21). The proof is completed. □

Bibliography

[1] ITU-R Rec. M.2083, "IMT vision framework and overall objectives of the future development of IMT for 2020 and beyond," Tech. Rep. 4, Oct. 2015.

[2] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, June 2017.

[3] D. Jiang and G. Liu, "An overview of 5G requirements," in *5G Mobile Commun.* Springer, 2017, pp. 3–26.

[4] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.

[5] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.

[6] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, Aug. 2017.

[7] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, Mar. 2017.

[8] M. Saari, A. M. bin Baharudin, P. Sillberg, S. Hyrynsalmi, and W. Yan, "LoRa-A survey of recent research trends," in *MIPRO 2018, Opatija, Croatia*, May 2018, pp. 0872–0877.

[9] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, Jan. 2017.

[10] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[11] L. Lv, J. Chen *et al.*, "Cooperative non-orthogonal multiple access in cognitive radio," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2059–2062, Oct. 2016.

[12] 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. Proc. Let.*, vol. 21, no. 12, pp. 1501–1505, July 2014.

[13] J. Ye, H. Lei, Y. Liu, G. Pan *et al.*, "Cooperative communications with wireless energy harvesting over Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5149–5164, Dec. 2017.

[14] W. Yu, L. Musavian, and Q. Ni, "Link-layer capacity of NOMA under statistical delay QoS guarantees," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4907–4922, Oct. 2018.

[15] J. Cui, Z. Ding, and P. Fan, "A novel power allocation scheme under outage constraints in NOMA systems," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1226–1230, Sep. 2016.

[16] Y. Zhang, H.-M. Wang, T.-X. Zheng, and Q. Yang, "Energy-efficient transmission design in non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2852–2857, Mar. 2016.

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

[18] R. Ruby, S. Zhong, H. Yang, and K. Wu, "Enhanced uplink resource allocation in non-orthogonal multiple access systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1432–1444, Dec. 2017.

[19] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.

[20] Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, Sep. 2017.

[21] J. Men and J. Ge, "Non-orthogonal multiple access for multiple-antenna relaying networks," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1686–1689, Oct. 2015.

[22] S. Li, L. Bariah, S. Muhaidat, P. Sofotasios, J. Liang, and A. Wang, "Error analysis of NOMA-based user cooperation with SWIPT," in *2019 15th Int. Conf. DCOSS.*, May 2019, pp. 507–513.

[23] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1–14, Jan. 2015.

[24] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 36–43, May 2014.

[25] K. Huang and V. K. Lau, "Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 902–912, Feb. 2014.

[26] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2013.

[27] L. R. Varshney, "Transporting information and energy simultaneously," in *2008 IEEE Int. Symp. Inf. Theory (ISIT)*. IEEE, Toronto, Oct. 2008, pp. 1612–1616.

[28] P. D. Diamantoulakis, K. N. Pappi, Z. Ding, and G. K. Karagiannidis, "Wireless-powered communications with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8422–8436, Dec. 2016.

[29] H. Chingoska, Z. Hadzi-Velkov, I. Nikoloska, and N. Zlatanov, "Resource allocation in wireless powered communication networks with non-orthogonal multiple access," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 684–687, Dec. 2016.

[30] J. Gong and X. Chen, "Achievable rate region of non-orthogonal multiple access systems with wireless powered decoder," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2846–2859, Dec. 2017.

[31] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[32] G. Pan, H. Lei, Y. Yuan, and Z. Ding, "Performance analysis and optimization for SWIPT wireless sensor networks," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2291–2302, May 2017.

[33] M. Hedayati and I.-M. Kim, "On the performance of OMA and NOMA in the two-user SWIPT system," *IEEE Trans. Veh. Technol.*, Nov. 2018.

[34] A. J. Goldsmith, *Wireless Commun.* Cambridge: Cambridge Univ. Press, 2005.

[35] C. C. Zarakovitis, Q. Ni, D. Skordoulis, and M. G. Hadjinicolaou, "Power-efficient cross-layer design for OFDMA systems with heterogeneous QoS, imperfect CSI, and outage considerations." *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 781–798, Feb. 2012.

[36] Y. Hao, Q. Ni, H. Li, and S. Hou, "Robust multi-objective optimization for EE-SE tradeoff in D2D communications underlaying heterogeneous networks," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4936–4949, Oct. 2018.

[37] Z. Zhu, Z. Chu, Z. Wang, and I. Lee, "Outage constrained robust beamforming for secure broadcasting systems with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7610–7620, Sep. 2016.

[38] Q. Zhang, Q. Li, and J. Qin, "Robust beamforming for nonorthogonal multiple-access systems in MISO channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 231–10 236, Dec. 2016.

[39] F. Alavi, K. Cumanan, Z. Ding, and A. G. Burr, "Robust beamforming techniques for non-orthogonal multiple access systems with bounded channel uncertainties," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2033–2036, Sep. 2017.

[40] S. Chinnadurai, P. Selvaprabhu, Y. Jeong, X. Jiang, and M. H. Lee, "Worst-case energy efficiency maximization in a 5G Massive MIMO-NOMA system," *Sensors*, vol. 17, no. 9, p. 2139, 2017.

[41] J. Cui, Z. Ding, and P. Fan, "Outage probability constrained MIMO-NOMA designs under imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8239–8255, Dec. 2018.

[42] S. Li *et al.*, "Energy-efficient resource allocation for industrial cyber-physical IoT systems in 5G era," *IEEE Trans. Ind. Inform.*, vol. 14, no. 6, pp. 2618–2628, June 2018.

[43] Y. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing:A key technology towards 5G," *White Paper, ETSI, Sophia Antipolis, France*, vol. 11, no. 11, pp. 1–16, 2015.

[44] C. Wang, C. Liang, F. R. Yu, Q. Chen, and L. Tang, "Computation offloading and resource allocation in wireless cellular networks with mobile edge computing," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 4924–4938, May 2017.

[45] C. Psomas and I. Krikidis, "Wireless powered mobile edge computing: Offloading or local computation?" *IEEE Commun. Lett.*, early access, 2020.

[46] F. Wang, J. Xu, X. Wang, and S. Cui, "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1784–1797, Mar. 2018.

[47] S. Bi and Y. J. Zhang, "Computation rate maximization for wireless powered mobile-edge computing with binary computation offloading," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4177–4190, June 2018.

[48] D. Wu, F. Wang, X. Cao, and J. Xu, "Wireless powered user cooperative computation in mobile edge computing systems," in *Proc. IEEE Globecom Workshops*, Dec, 2018, pp. 1–7.

[49] Z. Ding, J. Xu, O. A. Dobre, and V. Poor, "Joint power and time allocation for NOMA-MEC offloading," *IEEE Trans. Veh. Technol.*, June 2019.

[50] P. Yang, L. Li, W. Liang, H. Zhang, and Z. Ding, "Latency optimization for multi-user NOMA-MEC offloading using reinforcement learning," in *28th Wireless Optical Commun. Conf. (WOCC)*. IEEE, Beijing, China, July 2019, pp. 1–5.

[51] Y. Wu, K. Ni, C. Zhang, L. P. Qian, and D. H. Tsang, "NOMA-assisted multi-access mobile edge computing: A joint optimization of computation offloading and time allocation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12 244–12 258, Dec. 2018.

[52] F. Wang, J. Xu, and Z. Ding, "Multi-antenna NOMA for computation offloading in multiuser mobile edge computing systems," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2450–2463, Mar. 2018.

[53] Q. M. Qadir, T. A. Rashid, N. K. Al-Salihi, B. Ismael, A. A. Kist, and Z. Zhang, "Low power wide area networks: A survey of enabling technologies, applications and interoperability needs," *IEEE Access*, vol. 6, pp. 77 454–77 473, Nov. 2018.

[54] Z. Qin, F. Y. Li, G. Y. Li, J. A. McCann, and Q. Ni, "Low-power wide-area networks for sustainable IoT," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 140–145, June 2019.

[55] N. Sornin, M. Luis, T. Eirich, T. Kramp, and O. Hersent, "LoRaWAN specification, 1–82," 2015.

[56] Z. Qin, Y. Liu, G. Y. Li, and J. A. McCann, "Performance analysis of clustered LoRa networks," *IEEE Trans. Veh. Technol.*, 2019, accept to appear.

[57] O. Georgiou and U. Raza, "Low power wide area network analysis: Can LoRa scale?" *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 162–165, Jan. 2017.

[58] A. Waret, M. Kaneko, A. Guitton, and N. E. Rachkidy, "LoRa throughput analysis with imperfect spreading factor orthogonality," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 408–411, Oct. 2018.

[59] D. Croce, M. Gucciardo, S. Mangione, G. Santaromita, and I. Tinnirello, "Impact of LoRa imperfect orthogonality: Analysis of link-level performance," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 796–799, Apr. 2018.

[60] L. Amichi, M. Kaneko, E. H. Fukuda, N. E. Rachkidy, and A. Guitton, "Joint allocation strategies of power and spreading factors with imperfect orthogonality in LoRa networks," *arXiv preprint arXiv:1904.11303*, Apr. 2019.

[61] V. Hauser and T. Hegr, "Proposal of adaptive data rate algorithm for LoRaWAN-based infrastructure," in *IEEE 5th Int. Conf. FiCloud*, Aug. 2017, pp. 85–90.

[62] S. Li, U. Raza, and A. Khan, "How agile is the adaptive data rate mechanism of LoRaWAN?" in *Proc. IEEE GLOBECOM*, Abu Dhabi, UAE, Dec. 2018, pp. 206–212.

[63] K. Q. Abdelfadeel, V. Cionca, and D. Pesch, "Fair adaptive data rate allocation and power control in LoRaWAN," *arXiv preprint arXiv:1802.10338*, Feb. 2018.

[64] K. Mikhaylov, J. Petaejaejaervi, and T. Haenninen, "Analysis of capacity and scalability of the LoRa low power wide area network technology," in *Proc. 22nd Eur. Wireless Conf. , Oulu, Finland, May*, 2016, pp. 1–6.

[65] J. Haxhibeqiri, F. Van den Abeele, I. Moerman, and J. Hoebeke, "LoRa scalability: A simulation model based on interference measurements," *Sensors*, vol. 17, no. 6, p. 1193, May 2017.

[66] Z. Qin and J. A. McCann, "Resource efficiency in low-power wide-area networks for IoT applications," in *Proc. IEEE GLOBECOM*, Singapore, Dec. 2017, pp. 1–7.

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

[68] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer-Verlag New York, 2006.

[69] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[70] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[71] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[72] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Commun.*, vol. 12, no. 2, pp. 19–29, Apr. 2005.

[73] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 5, pp. 796–807, June 1994.

[74] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Sep. 2015.

[75] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3722–3732, July 2016.

[76] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches," *Wireless Commun. Mobile Comput.*, vol. 13, no. 18, pp. 1587–1611, 2013.

[77] M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, A. Neal *et al.*, "Mobile edge computing-Introductory technical white paper," *White paper,ETSI*, pp. 1089–7801, Sophia Antipolis, France, Sep. 2014.

[78] N. Sornin, M. Luis, T. Eirich, T. Kramp, and O.Hersent, "LoRaWAN specification," Jan. 2015. [Online]. Available: https://www.lora-alliance.org/portals/0/specs/LoRaWAN%20Specification%201R0.pdf

[79] "Sigfox", accessed Oct. 2016. [Online]. Available: https://www.sigfox.com/

[80] 3GPP, "Narrowband IoT," 3rd Generation Partnership Project, TS 151621, Sep. 2016.

[81] Y. Hao, Q. Ni, H. Li, and S. Hou, "On the energy and spectral efficiency tradeoff in massive MIMO-enabled hetnets with capacity-constrained backhaul links," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4720–4733, Nov. 2017.

[82] C. C. Zarakovitis, Q. Ni, and J. Spiliotis, "Energy-efficient green wireless communication systems with imperfect CSI and data outage," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3108–3126, Dec. 2016.

[83] F. Zhou, Y. Wu, R. Q. Hu, and Y. Qian, "Computation efficiency in a wireless-powered mobile edge computing network with NOMA," in *IEEE Int. Conf. Commun. (ICC), Shanghai, May, 2019*, pp. 1–7.

[84] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3767–3778, Jan. 2012.

[85] F. Zhou, Y. Wu, R. Q. Hu, and Y. Qian, "Computation rate maximization in UAV-enabled wireless-powered mobile-edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1927–1941, Sep. 2018.

[86] X. Hu, K. Wong, and K. Yang, "Wireless powered cooperation-assisted mobile edge computing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2375–2388, Apr. 2018.

[87] C. You, K. Huang, H. Chae, and B.-H. Kim, "Energy-efficient resource allocation for mobile-edge computation offloading," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1397–1411, Mar. 2017.

[88] H. Sun, F. Zhou, and R. Q. Hu, "Joint offloading and computation energy efficiency maximization in a mobile edge computing system," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3052–3056, Jan. 2019.

[89] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[90] L. Liu, R. Zhang, and K.-C. Chua, "Multi-antenna wireless powered communication with energy beamforming," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4349–4361, Dec. 2014.

[91] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Sep. 2016.

[92] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3269–3280, June 2014.

[93] S. Mallick, M. M. Rashid, and V. K. Bhargava, "Joint relay selection and power allocation for decode-and-forward cellular relay network with channel uncertainty," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3496–3508, Oct. 2012.

[94] M. B. Shenouda and T. Davidson, "Probabilistically-constrained approaches to the design of the multiple antenna downlink," in *42nd IEEE Asilomar Conf. SSC*, Oct. 2008, pp. 1120–1124.

[95] I. Bechar, "A Bernstein-type inequality for stochastic processes of quadratic forms of Gaussian variables," *arXiv preprint arXiv:0909.3595*, 2009.

[96] M. K. Vamanamurthy and M. Vuorinen, "Inequalities for means," *J. Math. Anal. Appl.*, vol. 183, no. 1, pp. 155–166, 1994.

[97] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008.

[98] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2082–2085, Dec. 2015.

[99] Y. Lu, K. Xiong, P. Fan, Z. Zhong, and K. B. Letaief, "Robust transmit beamforming with artificial redundant signals for secure SWIPT system under non-linear EH model," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2218–2232, Apr. 2018.

[100] Y. Liang and V. V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3284–3289, Sep. 2005.

[101] Z. Ding, D. W. K. Ng, R. Schober, and H. V. Poor, "Delay minimization for NOMA-MEC offloading," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1875–1879, Dec. 2018.

[102] X. Cao, F. Wang, J. Xu, R. Zhang, and S. Cui, "Joint computation and communication cooperation for energy-efficient mobile edge computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4188–4200, June 2019.

[103] H. Kosaki, "Arithmetic–geometric mean and related inequalities for operators," *J. Funct. Anal.*, vol. 156, no. 2, pp. 429–451, 1998.

[104] Y. Liu and Y. Dai, "On the complexity of joint subcarrier and power allocation for multi-user OFDMA systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 583–596, Nov. 2014.

[105] N. Vucic, S. Shi, and M. Schubert, "DC programming approach for resource allocation in wireless networks," in *Proc. Int. Symp. Modeling Optim. Ad Hoc Wireless Netw. (WiOpt)*, June 2010, pp. 380–386.

[106] P. Parida and S. S. Das, "Power allocation in OFDM based NOMA systems: A DC programming approach," in *IEEE Globecom Workshops*. IEEE, Dec. 2014, pp. 1026–1031.

[107] I. J. Lustig, R. E. Marsten, and D. F. Shanno, "Interior point methods for linear programming: Computational state of the art," *ORSA J. Comput.*, vol. 6, no. 1, pp. 1–14, 1994.

[108] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE Trans. Nets.*, vol. 24, no. 5, pp. 2795–2808, Oct. 2015.

[109] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do LoRa low-power wide-area networks scale?" in *Proc. of ACM Int. Conf. on Modeling, Analysis Simul. Wireless Mobile Systems*, Nov. 2016, pp. 59–67.

[110] M. Dianati, X. Shen, and K. Naik, "Cooperative fair scheduling for the downlink of CDMA cellular networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1749–1760, July 2007.

[111] E. Bodine-Baron, C. Lee, A. Chong, B. Hassibi, and A. Wierman, "Peer effects and stability in matching markets," in *Proc. 4th Symp. Algorithmic Game Theory (SAGT)*, Oct. 2011, pp. 117–129.

[112] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[113] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of Int. Commun. Conf. (ICC), Istanbul, Turkey*, vol. 7, 2006, pp. 3270–3275.

[114] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Research Logistics Quarterly*, vol. 10, no. 1, pp. 273–274, 1963.

[115] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, Third quatar, 2019.

[116] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Trans. Cogn.Commun. Netw.*, vol. 4, no. 4, pp. 648–664, Dec. 2018.

[117] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[118] M. R. Garey, "A guide to the theory of NP-Completeness," *Computers and intractability*, 1979.

[119] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.