Resource Allocation and Interference Management for Heterogeneous Networks Governed by Massive MIMO Macro Cell

?, 万明

https://doi.org/10.15017/1931936

Resource Allocation and Interference Management for Heterogeneous Networks Governed By Massive MIMO Macro Cell

Wanming Hao

Graduate School of Information Science and Electrical Engineering
Kyushu University

This dissertation is submitted for the degree of
Doctor of Engineering

Kyushu University March 2018

Acknowledgements

I would like to express my special appreciation and thanks to my supervisor Assoc. Prof. Muta Osamu for his support and guidance. I am very grateful for the pleasant research environment and the tremendous opportunity to pursue a Ph.D. he gave me. His prudence and perseverance for research became a model for me. He gave me many chance to attend national and international conference to improve myself. Besides this, he gave me lots of useful advice during our weekly meeting. He helped me improve my paper, presentation and encourage me to exploit my full potential in academic studies. I will cherish the friendship with him.

I would like to thank Prof. Furukawa Hiroshi. His expertise and wisdom in the research area are significant value to me. His exceptional motivation, vision and integrity will always be an inspiring role model for my future career. I also would like to thank Prof. Okamura Koji and Assoc. Prof. Jitsumatsu Yutaka for their comments and advice for improving my thesis. I also would like to thank Dr. Haris Gacanin for his help in modifying my paper, and gave me many useful advice for writing and presentation.

I would like to thank my colleague in our laboratory, namely Dr. Togashi Hiroaki, Mr. Kojima Yuki, Mr. Kageyama Tomoya, Mr. Matsuzaki Kouki and so on. I also would like to thank my roommates, Mr. Kai Wen and, Mr. Xiaochen Yang. I also would like to thank my friends, Miss Ting Cheng, Mr. Shiyan Feng, Mr. Liang Shang and so on. They gave me lots of help in my life and research.

Last, my special gratitude goes to my family for their unconditional love and infinite encouragement. Thank you all for supporting me in all my pursuits.

Wanming Hao
Fukuoka, Japan, March, 2018

Abstract

Given the 1000x capacity increase requirement for the next-generation cellular networks, massive multiple input multiple output MIMO (mMIMO), small cells (SCs) and cognitive radio (CR) have been proposed as important techniques. Therefore, the mMIMO coexists with CR or SCs to form an mMIMO heterogeneous network (HetNet), i.e., mMIMO-CR HetNet and mMIMO-SC HetNet, will be promising schemes. However, it brings more challenges due to the combination, especially for pilot contamination and interference management. The theme of this thesis is to propose advanced schemes for improving the achievable capacities (including per-user transmission rate and system sum rate) in mMIMO-HetNet by reducing the pilot contamination and coordinating the interference, which is divided to three parts.

For the first part (i.e., chapter 2), we study the pilot allocation problem in mMIMO homogeneous network for reducing the pilot contamination. To reduce the required complexity for finding the optimum pilot allocation, we propose a low-complexity pilot allocation algorithm. In addition, to improve users' fairness, we formulate a fairness aware pilot allocation problem and solve the formulated problem using a similar algorithm. Simulation results show that our proposed pilot allocation scheme can improve per-user transmission rate by about 17% in comparison with the conventional pilot allocation scheme.

For the second part (i.e., chapters 3 and 4), we study the pilot and power allocation problems in mMIMO-CR HetNet. We first propose a price-based iterative pilot allocation algorithm to obtain a win-win paradigm between primary network (PN) and cognitive network (CN) in chapter 3. The results show that the PN and CN can obtain positive revenue, which implies that pilot sharing concept between PN and CN is effective in improving the performance of both PN and CN. Next, to avoid producing serious interference from the CN to PN, we investigate the power allocation problem of the CN in mMIMO-CR HetNet with pilot contamination in chapter 4. We propose an orthogonal pilot sharing scheme at pilot transmission phase, where cognitive users are allowed to use pilots for channel estimation only when there are temporarily unused orthogonal pilots. Following this, we formulate the power allocation optimization problem of the CN to maximize the downlink sum rate of the CN subject to the total transmit power and primary users' signal to interference plus noise ratio (SINR) constraints. Then, we propose an iterative algorithm to solve the formulated

problem. The numerical results show that our proposed scheme can improve the sum rate of the CN by about 10% in comparison with the conventional scheme.

For the third part (i.e., chapters 5 and 6), we investigate the pilot allocation and interference management problems in mMIMO-SC HetNet. We first propose a pilot allocation scheme for maximizing ergodic downlink sum rate of the system in chapter 5, where the uplink pilot overhead and inter-tier interference are jointly considered. Then, we propose a low complexity one dimensional search algorithm to obtain the optimum pilot allocation. In addition, we propose two suboptimal pilot allocation algorithms to simplify the computational process and improve users' fairness, respectively. Simulation results show that our proposed scheme can improve the sum rate of the system by about 12% in comparison with the conventional scheme. Based on this, we investigate the dynamic SC clustering strategy and their precoding design problem for interference coordination in mMIMO-SC HetNet in chapter 6. An interference graph-based dynamic SC clustering scheme is proposed. Based on this, we formulate an optimization problem to design precoding weights at macro base station and clustered SCs for maximizing the downlink sum rate of SC users subject to the power constraint of each SC base station. A non-cooperative game-based distributed algorithm is proposed to solve the formulated problem. Simulation results show that our proposed scheme can improve the sum rate of SC users by about 40% in comparison with the conventional scheme.

In conclusion, through the above analysis and results, this thesis clarifies that the proposed schemes (pilot allocation, power allocation, SC clustering, and precoding design) are effective in increasing the achievable capacities in two types of NetNets (i.e., CR-type and SC-type) governed by mMIMO macro cell.

Table of contents

List of figures

List of tables

List of abbreviations

Acronyms / Abbreviations

AWGN	Additive White Gaussian Noise
BBU	Baseband Processing Unit
BS	Base Station
CBS	Cognitive Base Station
CC	Cognitive Cell
CCU	Central Control Unit
CDF	Cumulative Distribution Function
CN	Cognitive Network
CoMP	Coordinated Multiple Process
CR	Cognitive Radio
CSBD	Clustered Small Cell Block Diagonalization
CSI	Channel State Information
CU	Cognitive User
DoF	Degree Of Freedom
FCC	Federal Communications Commission
H-CRAN	Heterogeneous Centralized Radio Access Network
HetNet	Heterogeneous Networks

MBS	Macro Base Station
MF	Match Filter
MIMO	Multiple Input Multiple Output
MIP	Mixed Integer Programming
mMIMO	Massive Multiple Input Multiple Output
MMSE	Minimum Mean Squared Error
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MU	Macro User
NE	Nashi Equilibria
PBS	Primary Base Station
OFDMA	Orthogonal Frequency Division Multiple Access
PC	Primary Cell
PCS	Price Control Side
PN	Primary Network
PU	Primary User
QoS	Quality Of Service
RRH	Remote Radio Head
SBS	Small Cell Base Station
SC	Small Cell
SINR	Signal To Interference Plus Noise Ratio
SU	Small Cell User
SVD	Singular Value Decomposition
TDD	Time Division Duplex

WCDMA	Wideband Code Division Multiple
ZF	Zero Forcing

Chapter 1

Introduction

1.1 Background

Over the past a few decades, the wireless communications and networks have witnessed an unprecedented growth and steady evolution from the first to the fourth generation wireless networks. Meanwhile, some advanced techniques, like wideband code division multiple access (WCDMA), orthogonal frequency division multiple access (OFDMA) etc., have significantly contributed towards this gradual evolution. However, in recent years, the mobile data traffic (e.g., mobile video conference, streaming video and online game, etc.) and the advanced communication devices (e.g., smartphones, tablets and laptops, etc.) have been increasing rapidly. Fig. 1.1 and Fig. 1.2 show the demand for mobile data traffic and devices from 2016 to 2021, respectively [1]. The global mobile data traffic is expected to increase to 49 exabytes per month by 2021, and the number of mobile devices and connections are expected to grow to 11.3 billion by 2021. Although the increasing smartphones and multimedia services satisfy users' experiences and requirements, the recent mobile network has not achieved enough capacity to provide such huge increase of the video traffic in future [2]-[4]. As a result, how to satisfy the increasing traffic requirement has become one of challenges in future mobile networks. Two promising techniques to solve this issue are massive multiple input multiple output (mMIMO) and heterogeneous network (HetNet) concepts.

1.1.1 Massive MIMO and Heterogeneous Network

mMIMO is an effective technique to improve the capacity, where the base station (BS) is equipped with a large number of antennas to serve multiple users with the same time frequency resource, i.e., the number of user terminals is much less than the number of BS

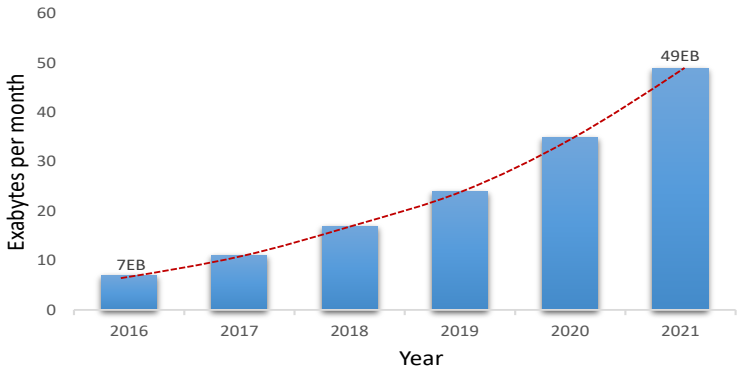


Fig. 1.1 Global mobile data traffic growth.

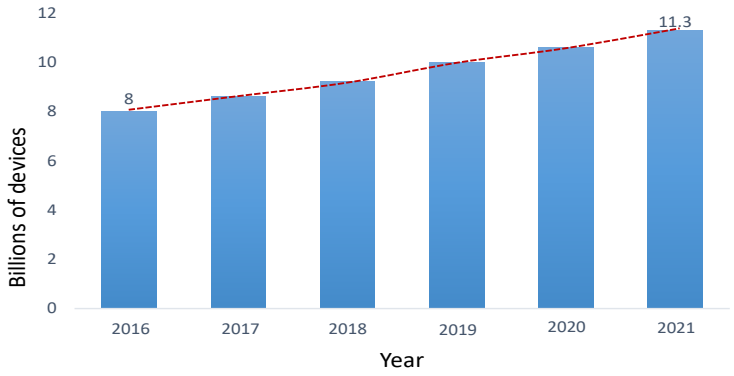


Fig. 1.2 Global mobile devices and connections growth.

antennas [5]. In this case, the huge throughput can be obtained because of the high degrees of freedom for mMIMO BS. Meanwhile, it has been verified that the full advantage of the mMIMO can be exploited by using simple linear-based approaches such as maximum ratio transmission (MRT), maximum ratio-combining (MRC) or zero forcing (ZF) [6]. The effects of fast fading, intra-cell interference and uncorrelated noise tend to disappear as the number of BS antennas grows enough large.

On the other hand, the HetNet is also a promising scheme to improve capacity. Different from the conventional homogeneous network which is governed by only one type of network, HetNet is defined as the combination of different types of networks. We can categorize HetNet into two types in a viewpoint of what system coexists, i.e., cognitive radio (CR) type and small cell (SC) type. In the former case, the primary network (PN) and cognitive network (CN) coexist to form a CR-type HetNet as shown in Fig. 1.3, where both networks have different priorities. In the later case, the macro cell (MC) network and SC network coexist to form a SC-type HetNet as shown in Fig. 1.4, where both networks have the same priority

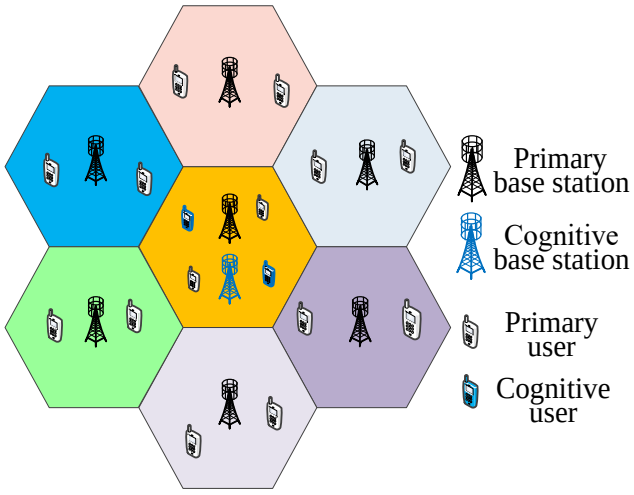


Fig. 1.3 The CR type HetNet model.

unlike the CR type. In this study, it is assumed that both types of HetNets are governed by MC with mMIMO to improve the system capacity. In the following sections, these two types of HetNets are explained, respectively.

1.1.2 Cognitive Radio Type Heterogeneous Network with Massive MIMO

The advantages of the CR-type HetNet can be summarized as follows: It is well known that more spectrum can bring higher throughput, but it seems that the wireless spectrum has crowded, and no more possible assignments for new users or services. However, the real problem of the spectrum scarcity has been shown by Federal Communications Commission (FCC) due to the fixed assignment of radio resource [7]. For example, Fig. 1.5 shows that the spectrum are not used at all time, and there always exists spectrum idle. Therefore, how to fully utilize those temporarily unused spectrum and improve the spectrum utilization is critical. Based on this, the CR technology has been proposed [8], which is defined as an intelligent wireless communication system that is aware of its surrounding environment in real-time. The CR can scan the frequency band of interest to assess the presence of active primary users (PUs) through a spectrum sensing process. For a given sensing result, CR needs to implement an adequate protocol for using the spectrum, namely spectrum access technique. There are three main spectrum access schemes: underlay, interweave and overlay [9]. Under the underlay design, cognitive users (CUs) are allowed to share the licensed spectrum with PUs as long as the interference to PUs is below a given threshold. By contrast, under the interweave design, CUs are requested to use the licensed spectrum

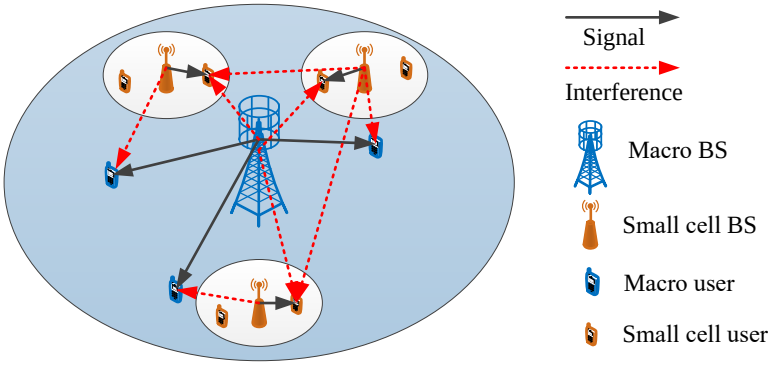


Fig. 1.4 The SC type HetNet model.

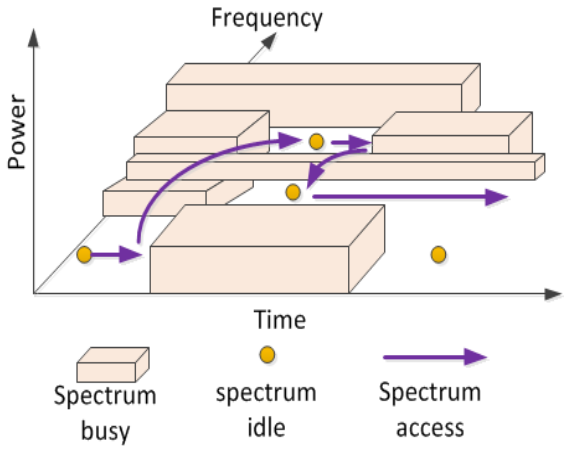


Fig. 1.5 The wireless spectrum utilization.

only when the spectrum are not used by PUs. Similarly with the first design, under the overlay design, CUs are allowed to share the licensed spectrum with PUs. However, CUs are requested to cooperate with PUs' communication by using some sophisticated signal processing and coding technology, while obtaining the chance for their own communication. From the above analysis, it is clear that the CR technology coexisted with primary MC is treated as the HetNet, and different priority should be considered, namely the PUs have the high priority to use the resource. In other words, transmission power and available resources for the CN are strictly restricted unlike the PN.

We have analyzed the advantages of mMIMO and HetNet in improving the capacity in previous section. Therefore, to future improve the performance of the CR-type HetNet, the

primary BS (PBS) and cognitive BS (CBS) can be equipped with a large number of antennas and form a CR-type HetNet with mMIMO. Hereafter, we denote it as mMIMO-CR HetNet in this thesis.

1.1.3 Small Cell Type Heterogeneous Network with Massive MIMO Macro Cell

For the SC-type HetNet, it is well known that the deployment of SCs can effectively improve the throughput of the system. For example, in a multi-user network, users in the coverage of a cell share the available bandwidth. Thus, reducing the cell size and deploying more cells also reduce the coverage of per cell, and in turn increases the bandwidth available to each user. Meanwhile, the deployment of SCs shortens the distance between terminals and BSs, and thus 1) lowering the transmit power, 2) improving the signal to noise ratio and 3) realizing the dense spectrum reuse, such as femtocells, picocells and microcells. In addition, different from the CR-type HetNet, all users own the same priority to use resource.

Similar to the mMIMO-CR HetNet, all BSs can be also equipped with a large number of antennas in the SC-type HetNet for improving the achievable capacity of the system. However, we know that there are two classes of BSs. One is the macro BS (MBS) that has the high transmit power and large coverage area. The other is the SC BS (SBS) that has the low transmit power and small coverage area, and its physical size is also small. Therefore, it is unnecessary and difficult to equip a large number of antennas at SBS. Based on the above discussion, in this study, it is assumed that a large number of antennas is equipped on only MBS coexisted with SCs to form a SC-type HetNet with mMIMO MC. Hereafter, we denote it as mMIMO-SC HetNet in this thesis.

1.2 Technical Challenges for Heterogeneous Network with Massive MIMO

1.2.1 Application of Massive MIMO to Heterogeneous Network

We have analyzed the advantages of the mMIMO-CR and mMIMO-SC HetNets in the previous subsection. However, there also exist technical challenges for the application of mMIMO techniques. In general, the precoding is used at mMIMO BS to cancel multi-user interference. In this case, the channel state information (CSI) should be obtained by channel estimation. Channel estimation is usually based on pilot, including uplink pilot (from users to the BS) and downlink pilot (from the BS to users). For the downlink pilot, the demand

of the orthogonal pilots (here, orthogonal pilots means that the different pilot codes are orthogonal) is huge when there are a large number of BS antennas. This is because that each antenna needs to be allocated one orthogonal pilot. As a result, the huge number of orthogonal pilots is required due to the pilot overhead and thus it results in decreasing the transmission efficiency. On the contrary, when the uplink pilot is used, namely each served users transmits one orthogonal pilot. In this case, the number of orthogonal plots is equal to the number of users. In general, the number of served users is much less than that of the BS antennas. Therefore, the pilot overhead is small for the uplink pilot transmission, which is usually used in mMIMO system [5]. According to this, the time division duplex (TDD) is adopted because the estimated uplink CSI can be used for downlink. Here, in TDD systems the pilot and data transmission occupy different time in each frame. Although the orthogonal pilots can be used at each cell, they have to be reused in different cells due to the limited coherence time. As a result, different users will use the same pilot, which causes pilot interference (i.e., pilot contamination) [10]. Therefore, when the mMIMO is applied in HetNet, the pilot contamination must be considered and solved.

We first analyze the basic pilot contamination problem in mMIMO homogeneous network. We assume that a total of L cells share the same set of K pilot signals. In each cell, the BS is equipped with a large number of antennas M to serve K user terminals. In this case, the received signal of the jth BS can be written as

$$\mathbf{r}_j = \sqrt{p_u} \sum_{l=1}^{L} \mathbf{G}_{jl} \mathbf{x}_l + \mathbf{n}_j, \tag{1.1}$$

where p_u is the average transmit power of each terminal, $\mathbf{G}_{jl}$ is the $M \times K$ channel matrix between the K terminals in the lth cell and the BS antennas in the jth cell, where$[\mathbf{G}_{jl}]_{mk} = g_{mjkl} = \sqrt{\beta_{jkl}} h_{mjkl}$, $\mathbf{x}_l$ denotes the transmit symbols from the lth cell, and $\mathbf{n}_j$ is vector of receiver noise. Let $\hat{\mathbf{G}}_{jj}$ denotes the estimate for the $M \times K$ propagation matrix between the M base station antennas of the jth cell, and the K terminals in the jth cell, which can be written as

$$\hat{\mathbf{G}}_{jj} = \sqrt{p_t} \sum_{l=1}^{L} \mathbf{G}_{jl} + \mathbf{v}_j, \tag{1.2}$$

where p_t is the pilot transmit power, and $\mathbf{v}_j$ denotes the received noise. The BS processes its received signal by MRC and yields

$$\hat{\mathbf{r}} = \hat{\mathbf{G}}_{jj}^H \mathbf{r}_j = \left[\sqrt{p_t} \sum_{l=1}^{L} \mathbf{G}_{jl} + \mathbf{v}_j \right]^H \left[\sqrt{p_u} \sum_{l=1}^{L} \mathbf{G}_{jl} \mathbf{x}_l + \mathbf{n}_j \right]. \tag{1.3}$$

As M grows infinity the L2-norm of these vector grows proportional to M, while the inner products of uncorrelated vectors grows at a lesser rate. For a large M, the products of identical quantities remain significant and we have

$$\frac{1}{M}\mathbf{G}_{jl_1}^H\mathbf{G}_{jl_2} = \mathbf{D}_{\bar{\beta}_{jl_1}}^{1/2}\left(\frac{\mathbf{H}_{jl_1}^H\mathbf{H}_{jl_2}}{M}\right)\mathbf{D}_{\bar{\beta}_{jl_2}}^{1/2}, \tag{1.4}$$

where $\mathbf{D}_{\bar{\beta}_{jl}}$ is a $K\times K$ diagonal matrix and $\left[\bar{\beta}_{jl}\right]_k = \beta_{jkl}$, $\mathbf{H}_{jl}$ is a $M\times K$ fast fading coefficients matrix and $\left[\mathbf{H}_{jl}\right]_{mk} = h_{mjkl}$. As M goes into infinity we have $\frac{1}{M}\mathbf{H}_{jl_1}^H\mathbf{H}_{jl_2} \to \mathbf{I}_K\delta_{l_1l_2}$, where $\mathbf{I}_K$ is the $K\times K$ identity matrix. Then, we have

$$\frac{1}{M\sqrt{p_tp_u}}\hat{\mathbf{r}}_j \to \sum_{l=1}^{L}\mathbf{D}_{\bar{\beta}_{jl}}\mathbf{x}_l. \tag{1.5}$$

The kth component of the processed signal becomes

$$\frac{1}{M\sqrt{p_tp_u}}\hat{r}_{kj} \to \beta_{jkj}x_{kj} + \sum_{l\neq j}\beta_{jkl}x_{kl}. \tag{1.6}$$

Therefore, the user's rate can be written as

$$R_{lk} = \log_2\left(1+\frac{\beta_{lkl}^2}{\sum_{l\neq j}\beta_{jkl}^2}\right). \tag{1.7}$$

From (1.7), it is clear that the user's rate is affected by pilot contamination from other cells. Therefore, how to reduce the pilot contamination is a key problem when the mMIMO technique is applied. In this thesis, we will first investigate the pilot contamination and propose effective pilot allocation schemes in a homogeneous mMIMO network to reduce pilot contamination, which will be the fundament for investigating the mMIMO-CR and mMIMO-SC HetNets.

1.2.2 Challenges for Cognitive Radio Type Heterogeneous Network with Massive MIMO

We have analyzed that pilot contamination should be considered in mMIMO. Since PBS and SBS are all equipped with a large number antennas, it is necessary to investigate the pilot allocation problem in mMIMO-CR HetNet. In addition, for CR HetNet, it is well known that the PN has the high priority to use the resource. In other words, although the CN is allowed to share the resource with PN, the serious interference produced by CN to

PN should be avoided. Otherwise, the communication of the PN will be affected. Based on this, when the CN shares the resource with PN, the CBS must control the transmit power so that the produced interference to PN below to a tolerated level. Therefore, except for the pilot allocation, power allocation to CUs must be considered, which is also a challenge in mMIMO-CR HetNet.

1.2.3 Challenges for Small Cell Type Heterogeneous Network with Massive MIMO Macro Cell

Since the MBS is equipped with a large number of antennas, the pilot contamination should be considered in mMIMO-SC HetNet. Different from the mMIMO-CR HetNet, the MUs and SUs have the same priority to use the resource. In this case, the MUs' and SUs' interference should all be effectively coordinated. In fact, from the Fig. 1.4, it is clear that the SU's interference from the MBS is serious due to the high transmit power of the MBS. In addition, when there are a lot of SCs covered with MC, the interference among SCs has to be considered. As a result, how to reduce the interference from the MBS to SUs and coordinate the interference among SCs are also critical and challenges in mMIMO-SC HetNet. Based on this, for the interference from the MBS to SUs, effective precoding design must be considered. For the interference coordination among SCs, SC clustering should be one of the solutions.

1.3 Motivations and Contributions of This Thesis

1.3.1 Motivations of This Thesis

Based on the above analysis, it is clear that there exists a common challenge to apply mMIMO technique to CR- and SC-HetNets, i.e., pilot contamination. Meanwhile, there are also different challenges for these two types of HetNet. Concretely, in mMIMO-CR HetNet, we need to consider how to control the transmit power of the CBS to avoid producing serious interference to PUs. In mMIMO-SC HetNet, we need to consider how to coordinate the interference among MUs and SUs, and the interference among SCs. Based on this, the motivation of this paper is how to solve the above challenges in these two types of HetNets. To this end, objective of this thesis is divided into the following three parts to investigate the above challenges.

- The first part is to study the pilot allocation problem in mMIMO homogeneous network and clarify that the pilot allocation is effective to improve the capacity for a mMIMO network (i.e., chapter 2).

- The second part is to study the pilot and power allocation problems in mMIMO-CR HetNet where CN and PN have different priorities (i.e., chapters 3 and 4).

- The third part is to study the pilot allocation and interference management problems in mMIMO-SC HetNet where SC and MC have the same priority (i.e., chapters 5 and 6).

1.3.2 Contributions of This Thesis

The key contributions of this thesis are summarized as follows:

- We propose a low complexity pilot allocation algorithm to maximize the uplink rate of the mMIMO homogeneous network. Meanwhile, to improve the users' fairness, we formulate a fairness aware pilot allocation as maximization problem of sum of user's logarithmic and apply the similar algorithm to obtain the solution (i.e., chapter 2).

- We propose a price-based iterative pilot allocation algorithm to obtain a win-win paradigm between PN and CN in mMIMO-CR HetNet (i.e., chapter 3). Next, we propose an iterative power allocation algorithm to maximize the downlink sum rate of the CN subject to the transmit power and the SINR constraints of the PUs in mMIMO-CR HetNet (i.e., chapter 4).

- We propose an optimum pilot allocation scheme to maximize the downlink sum rate of the mMIMO-SC HetNet. Meanwhile, two suboptimal algorithms are proposed to simplify the optimization process and improve the SUs' fairness, respectively (i.e., chapter 5). Next, to reduce the interference among SCs, an interference graph-based dynamic SC clustering scheme is proposed. Then, a non-cooperative game-based precoding design algorithm is proposed to maximize the downlink sum rate of the SUs in mMIMO-SC HetNet (i.e., chapter 6).

1.4 Organization of This Thesis

This thesis is organized in seven chapters, which are summarized as follows:

Chapter 1 provides a broad introduction on target system and related techniques, i.e., mMIMO, mMIMO-CR HetNet and mMIMO-SC HetNet. Next, we study the basic problem

in chapter 2, namely the pilot allocation problem in a mMIMO homogeneous network. Then, we study the pilot and power allocation problem in mMIMO-CR HetNet. We divide two chapters (i.e., chapters 3 and 4) to consider them. Specifically, in chapter 3, we study the pilot allocation problem based on the infinite number antenna at BS. After that, the limited number antenna and power allocation at BS is studied in chapter 4. To this end, we study the pilot allocation and interference management problems in mMIMO-SC HetNet. We still divide two chapters (i.e., chapters 5 and 6) to consider the above problem. Specifically, in chapter 5, we only investigate the pilot allocation problem in mMIMO-SC HetNet. Based on the chapter 5, we study the SC clustering and precoding design to solve the interference problem in chapter 6. Chapter 7 summarizes this thesis and gives the future research direction.

Chapter 2

Pilot Allocation for Massive MIMO Homogeneous Network

2.1 Introduction

In this chapter, we will investigate how to reduce pilot contamination by effective pilot allocation so as to improve the performance of the mMIMO system. In fact, pilot contamination problem has been studied widely in the literature [11]-[15]. In [11], a pilot assignment scheme is proposed to mitigate pilot contamination problem, where the allocation of pilot sequences is optimized to maximize the signal-to-interference power ratio on the uplink. The work in [12] proposes the users scheduling per cell in order to maximize the spectral efficiency, but for the given number of users in each cell, the approach does not take into consideration the pilot allocation strategy. In [13], a fractional pilot reuse scheme is proposed, where users in different cells are allowed to reuse the same pilot sequence if they are close to their BSs. Otherwise, if users are located far away from BS in different cells, the orthogonal pilot sequences must be used. Thus, the pilot allocation is not considered for users located closely to their BSs. In [14], a graph coloring based pilot allocation is proposed to reduce the pilot contamination. The authors first construct an interference graph according to the strength of potential pilot contamination between any two users in different cells with the same pilot. Then, they allocate pilots among users in order to minimize potential pilot contamination term in the graph. In [15], the authors assume that a subset of pilots is owned by each cell and then, cells may cooperate to utilize pilots from other cells and support more users. However, the pilot-to-user allocation is not considered. Although some pilot allocation schemes in works [11]-[15] are proposed to improve the capacity of the system, they are all

not global optimum pilot allocation schemes when considering the sum rate maximization of the system.

In this chapter, we assume that an uplink communication is established in two phases: (i) pilot and (ii) data signaling. Thus, by reducing the interference between utilized pilots from adjacent cells, the (data) uplink user sum rate may be improved. The optimum pilot allocation is decided by a central control unit (CCU) that acts as master BS. Then, we formulate the pilot allocation optimization problem of maximizing the uplink sum rate of the mMIMO systems. To decrease the complexity, we propose an iterative pilot allocation optimization algorithm, where the original problem is transformed into a number of subproblems which can be solved as one-to-one matching problem. The Hungarian algorithm [16] can be applied to find the optimum pilot allocation problem in each subproblem. In addition, to improve the users' fairness, we formulate a users' fairness aware pilot allocation as maximization problem of sum of user's logarithmic rate and use a similar algorithm to obtain the corresponding pilot allocation.

2.2 System Model

We consider an uplink multi-cell system composed of L hexagonal cells as shown in Fig. 2.1. The radius of each cell is r_c, and white area in each cell denotes the cell-hole (users are not located within the center disk of radius r_h). One of the BSs works as CCU, while each BS is equipped with M antennas and serves K ($M \gg K$) single-antenna users. We assume that there is time-frequency coherent block of S symbols in each frame. K orthogonal pilot signals $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \cdots, \boldsymbol{\psi}_K]^T \in \mathscr{C}^{K\times K} (\boldsymbol{\psi}_i = [\psi_{i1}, \cdots, \psi_{iK}]^T)$ are reused in adjacent cells due to the limited coherence time, while different users in each cell use orthogonal pilots to avoid severe interference, and we assume that $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{I}_K$. Here, $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and Hermitian transpose, respectively.

During the training phase, the received signal at the BS of the l-th cell can be expressed as:

$$\mathbf{Y}_l = \sqrt{p_p} \sum_{j=1}^{L} \sum_{k=1}^{K} \mathbf{h}_{ljk} \boldsymbol{\psi}_k^T + \mathbf{Z}_l, \tag{2.1}$$

where p_p denotes the pilot transmit power, $\mathbf{Z}_l \in \mathscr{C}^{M\times K}$ is an independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) defined as $\mathscr{CN}(0,\ \delta_z^2)$, $\mathbf{h}_{ljk} \in \mathscr{C}^{M\times 1}$ is the channel coefficient between BS in the l-th cell and the k-th user in the j-th cell. $\mathbf{h}_{ljk} = \sqrt{\beta_{ljk}}\mathbf{g}_{ljk}$, where β_{ljk} and $\mathbf{g}_{ljk} \sim \mathscr{CN}(\mathbf{0},\ \mathbf{I}_M)$ denote the large-scale fading coefficient and small-scale fading vector, respectively.

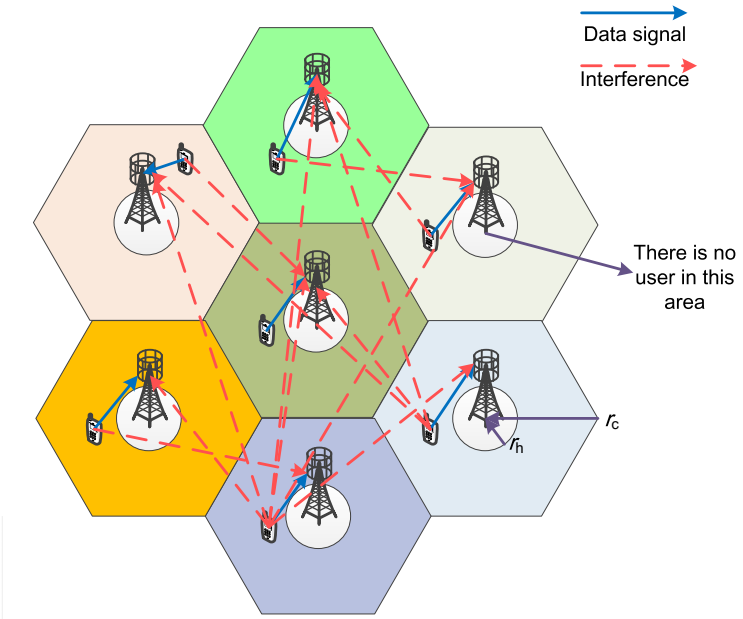


Fig. 2.1 Uplink interference model for the multi-cell mMIMO system.

The channel estimate of the k-th user in the l-th cell is obtained by correlating $\mathbf{Y}_l$ with $\boldsymbol{\psi}_k^*$ as follows:

$$\begin{aligned}\tilde{\mathbf{h}}_{llk} &= \mathbf{h}_{llk}\boldsymbol{\psi}_k^T\boldsymbol{\psi}_k^* + \sum_{j\neq l}^{L}\sum_{i=1}^{K}\mathbf{h}_{lji}\boldsymbol{\psi}_i^T\boldsymbol{\psi}_k^* + \frac{1}{\sqrt{p_p}}\mathbf{Z}_l\boldsymbol{\psi}_k^* \\ &= \mathbf{h}_{llk} + \sum_{j\neq l}^{L}\sum_{i=1}^{K} f[\theta(j,i),\theta(l,k)]\mathbf{h}_{lji} + \mathbf{w}_{lk},\end{aligned} \tag{2.2}$$

where $(\cdot)^*$ denotes the complex conjugate, $\mathbf{w}_{lk}$ denotes the equivalent noise, $\boldsymbol{\psi}_{\theta(j,i)}$ $(\theta(j,i) \in \{1,\cdots,K\})$ denotes that the $\theta(j,i)$-th pilot is used by the i-th user in the j-th cell with $\theta(j,k) \neq \theta(j,k')$ when $k \neq k'$. In the above expression, $f[\cdot] \in \{0,1\}$ represents the pilot reuse index, $f[\theta(j,i),\theta(l,k)] = 1$ when $\theta(j,i) = \theta(l,k)$, else $f[\theta(j,i),\theta(l,k)] = 0$.

During the data phase, the received signal at the BS of the l-th cell can be expressed as:

$$\mathbf{y}_l = \sqrt{p_t}\sum_{j=1}^{L}\sum_{k=1}^{K}\mathbf{h}_{ljk}x_{jk} + \mathbf{n}_l, \tag{2.3}$$

where p_t denotes the uplink data transmit power, x_{jk} denotes the data transmitted by the k-th user in the l-th cell with $E[|x_{lk}|^2] = 1$ and $\mathbf{n}_l \sim \mathscr{CN}(\mathbf{0},\ \sigma_l^2 \mathbf{I}_M)$ denotes the noise, where $E[\cdot]$ denotes the expectation operator.

Using the channel estimate of the k-th user in (2), the matched-filter (MF) detector is applied to obtain the decision variables of the k-th user as:

$$\begin{aligned}\tilde{x}_{lk} = \tilde{\mathbf{h}}_{llk}^H \mathbf{y}_l = &\underbrace{\sqrt{p_t}\mathbf{h}_{llk}^H \mathbf{h}_{llk} x_{lk}}_{\text{Desired signal}} + \underbrace{\sqrt{p_t}\sum_{n\neq k}^{K} \mathbf{h}_{llk}^H \mathbf{h}_{lln} x_{ln}}_{\text{intra-cell interference}} \\ &+ \underbrace{\sqrt{p_t}\sum_{j\neq l}^{L}\sum_{i=1}^{K}\sum_{m=1}^{L}\sum_{n=1}^{K} f[\theta(j,i),\theta(l,k)]\mathbf{h}_{lji}^H \mathbf{h}_{lmn} x_{mn}}_{\text{pilot contamination}} \\ &+ \underbrace{\sqrt{p_t}\sum_{m\neq l}^{L}\sum_{n=1}^{K} \mathbf{h}_{llk}^H \mathbf{h}_{lmn} x_{mn}}_{\text{inter-cell interference}} + \underbrace{\omega_{lk}}_{\text{uncorrelated noise}},\end{aligned} \tag{2.4}$$

where $\omega_{lk} = \mathbf{h}_{llk}^H \mathbf{n}_l + \sum_{j\neq l}^{L}\sum_{i=1}^{K} f[\theta(j,i),\theta(l,k)]\mathbf{h}_{lji}^H \mathbf{n}_l + {\mathbf{w}_{lk}}^H \mathbf{n}_l$. In (2.4), the first term denotes the desired signal component, the second term denotes the intra-cell interference, the third term denotes the pilot contamination, the fourth term denotes the inter-cell interference, and the last term denotes the uncorrelated noise after MF filtering. According to (2.4), the average uplink rate of the user can be expressed as

$$r_{lk} = \mathrm{E}\left\{\log_2\left(1 + \frac{\left|\mathbf{h}_{llk}^H \mathbf{h}_{llk}\right|^2}{\mathrm{IN}_{lk} + |\omega_{lk}|^2/p_t}\right)\right\} \tag{2.5}$$

where $\mathrm{IN}_{lk} = \sum_{n\neq k}^{K}\left|\mathbf{h}_{llk}^H \mathbf{h}_{lln}\right|^2 + \sum_{m\neq l}^{L}\sum_{n=1}^{K}\left|\mathbf{h}_{llk}^H \mathbf{h}_{lmn}\right|^2 + \sum_{j\neq l}^{L}\sum_{i=1}^{K}\sum_{m=1}^{L}\sum_{n=1}^{K} f[\theta(j,i),\theta(l,k)]\left|\mathbf{h}_{lji}^H \mathbf{h}_{lmn}\right|^2$.

2.3 Problem Formulation and Solution

In this section, we first formulate a pilot allocation optimization problem to maximize uplink sum rate of the system. Then, we propose a low-complexity algorithm to obtain the optimal solution. Next, considering users' rate fairness, we formulate a fairness aware pilot allocation as maximization problem of sum of user's logarithmic rate and use the similar method to solve the formulated problem.

2.3.1 Problem Formulation Based on Sum Rate Maximization

We formulate the pilot allocation optimization problem for maximizing uplink sum rate of the system as follows:

$$\begin{aligned} &\max_{\boldsymbol{\theta}} \; R(\boldsymbol{\theta}) = \sum_{l=1}^{L}\sum_{k=1}^{K}(1-\eta)r_{lk} \\ s.t \quad &\theta(l,k) \in \{1,2,\cdots,K\}, \forall l,k, \\ &\theta(l,k) \neq \theta(l,k'), k \neq k', \end{aligned} \tag{2.6}$$

where $\boldsymbol{\theta} = [\theta(l,k)]_{LxK}$ denotes the pilot allocation index for each user, $\eta = K/S$. Note that accurate CSI is needed to estimate user's rate r_{lk} for solving the optimization problem (2.6). However, based on the fact that CSI can not be obtained before determining pilot allocation, it seems that it is not possible to solve the problem (2.6).

According to [17], when the number of BS antennas M goes to infinity, the uplink rate can be approached using only large-scale fading coefficients as

$$r_{lk} \approx \log_2\left(1 + \frac{\beta_{llk}^2}{\sum_{j\neq l}^{L}\sum_{i=1}^{K} f[\theta(j,i),\theta(l,k)]\beta_{lji}^2}\right). \tag{2.7}$$

It can be observed from (2.7) that the uplink rate in the optimization problem can be approximated with only the large-scale fading coefficients, which can be easily tracked by the BSs. Here, we propose to use approximated rate in (2.7) for solving the problem (2.6). The details of the proposed algorithm to solve (2.6) is mentioned in next subsection.

2.3.2 Proposed Sum Rate Maximization Scheme

Problem (2.6) is known as mixed integer programming (MIP) problem. The challenge of this problem is the discrete nature of the pilot allocation index. Exhaustive search can be used to find the optimum pilot allocation, but it requires high computational complexity given as $O((K!)^L)$. Thus, exhaustive search is not feasible solution for a large number of users in multi-cell mMIMO system.

To decrease the computational complexity, we decouple (2.6) into L subproblems, where in each subproblem, we aim at optimizing the pilot allocation of K users in one particular cell and fix pilot allocation in other L-1 cells. Based on the above description, we can get one of subproblems as follows:

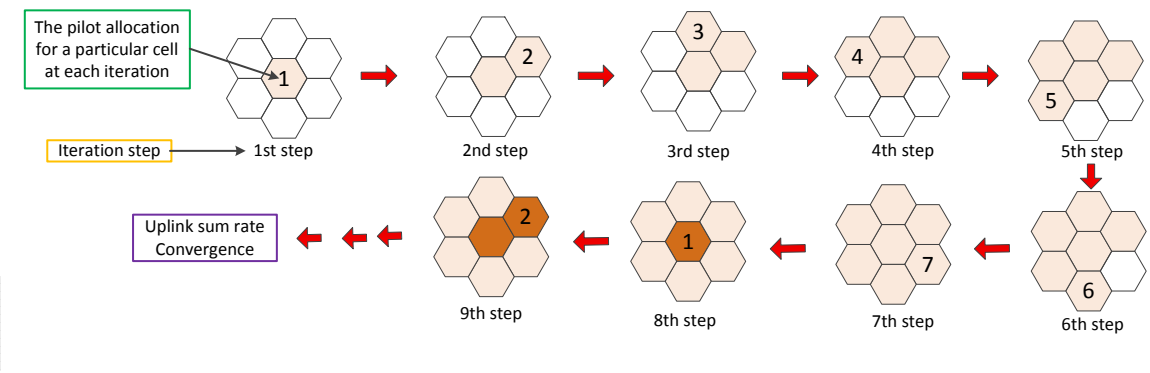


Fig. 2.2 Iteration diagram for pilot allocation in the proposed algorithm.

$$\begin{aligned} &\max_{\boldsymbol{\theta}_m} \quad R_m(\boldsymbol{\theta}_{-m}, \boldsymbol{\theta}_m) \\ s.t. \quad &\theta(m,k) = \{1,2,\ldots,K\}, \forall k \\ &\theta(m,k) \neq \theta(m,k'), k \neq k' \end{aligned} \tag{2.8}$$

where $R_m(\boldsymbol{\theta}_{-m}, \boldsymbol{\theta}_m) = \sum_{l=1}^{L}\sum_{k=1}^{K}(1-\eta)\log_2\left(1+\frac{\beta_{llk}^2}{\sum_{j\neq l}^{L}\sum_{i=1}^{K} f[\theta(j,i),\theta(l,k)]\beta_{lji}^2}\right)$, $\boldsymbol{\theta}_{-m}$ denotes the pilot allocation decision matrix except for the m-th cell, and $\boldsymbol{\theta}_m$ the pilot allocation matrix in the m-th cell. For (2.8), since pilot allocation in other cells have been decided in advance (at the beginning, we assume that the pilots are randomly allocated in these cells), we just need to allocate pilots to users in the m-th cell for maximizing the sum rate of the system. Exhaustive search is not feasible because the required complexity is given as $(O(K!))$ and significantly increased with a large K.

To reduce the required complexity for finding the optimum solution, we propose a low-complexity pilot allocation scheme. Since we have fixed pilot allocation in other $L-1$ cells, the problem (2.8) is reduced to a one-to-one matching problem, namely K users select K pilots. Next, we define the one-to-one matching problem as follows:

Definition: We assume that there are K users and K pilots, and we need to allocate the K pilots to K users. The allocation rule is that every user is assigned one pilot and each pilot is only assigned to one user. Each possible allocation between the i-th pilot and the k-th user is associated a utility U_{ik} (the U_{ik} can be regarded as the revenue of the k-th user when it uses the i-th pilot), which is given in Table 2.1.

Table 2.1 The utility of pilot allocation

Pilot \ User	1	2	3	$\cdots$	K
1	U_{11}	U_{12}	U_{13}	$\cdots$	U_{1K}
2	U_{21}	U_{22}	U_{23}	$\cdots$	U_{2K}
3	U_{32}	U_{32}	U_{33}	$\cdots$	U_{3K}
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\ddots$	$\vdots$
K	U_{K1}	U_{K2}	U_{K3}	$\cdots$	U_{KK}

Then, the matching problem can be presented by the following optimization problem:

$$\begin{aligned} \max_{c_{nm}} \quad & \sum_{n=1}^{K}\sum_{m=1}^{K} c_{nm}U_{nm} \\ s.t. \quad & \sum_{n=1}^{K} c_{nm} = 1, \ \forall n, \\ & \sum_{m=1}^{K} c_{nm} = 1, \ \forall m, \\ & c_{nm} \in \{0,1\}, \ \forall n,m, \end{aligned} \tag{2.9}$$

where c_{nm} denotes the binary assignment variable, and $c_{nm} = 1$ means that pilot n is allocated to user m, and $c_{nm} = 0$, otherwise. $\sum_{n=1}^{K} c_{nm} = 1$ denotes that each pilot is only allocated to one user, $\sum_{m=1}^{K} c_{nm} = 1$ denotes that each user is only allocated one pilot.

As for the problem (2.9), the optimal matching problem can be solved by applying the well-known Hungarian algorithm [18], which is a combinatorial optimization algorithm that solves the assignment problem in polynomial time. Therefore, the subproblem (2.8) can be solved by using the similar method. We rewrite the subproblem (2.8) as follows:

$$
\begin{aligned}
\max_{\boldsymbol{\theta}_m} \quad & \sum_{a=1}^{K}\sum_{p=1}^{K} c_{ap} R_m^{ap}(\boldsymbol{\theta}_{-m}, \boldsymbol{\theta}_m) \\
s.t. \quad & R_m^{ap}(\boldsymbol{\theta}_{-m}, \boldsymbol{\theta}_m) = \begin{cases} R_m(\boldsymbol{\theta}_{-m}, \boldsymbol{\theta}_m), \\ \theta(m,a) = p, \end{cases} \\
& \sum_{a=1}^{K} c_{ap} = 1, \ \forall a, \\
& \sum_{p=1}^{K} c_{ap} = 1, \ \forall p, \\
& c_{ap} \in \{0,1\}, \ \forall a, p.
\end{aligned} \tag{2.10}
$$

where a and b denote the pilots and users index in the m-th cell, respectively. We can find that the subproblem (2.10) is also an one-to-one matching problem and the optimum pilot allocation can be obtained by applying the Hungarian algorithm. Next, we move to the next cell and use the same method to optimize pilot allocation for next subproblem. After multiple iterations, the global optimum pilot allocation for problem (2.6) can be obtained according to the Proposition 1. To describe our proposed algorithm more clearly, we present the iterative diagram in Fig. 2.2. For example, at the first step, the $m = 1$ in problem (2.10), namely, we only optimize the pilot allocation at the 1st cell while fixing pilot allocation in other cells. After solving problem (2.10), we can obtain the uplink sum rate. Then, similar to the first step, we optimize the pilot allocation at the 2nd cell as the second step of Fig. 2.2. This process is continued until the uplink sum rate is converged. We also summarize the above method in Algorithm 1.

Proposition 1: For given L and K, global optimum pilot allocation converges after a finite number of iterations.

Proof: In solving each subproblem (iteration), the pilot allocation is obtained according to the Hungarian method, and the sum rate of the system is maximized in this optimization (iteration). Therefore, the objective of problem (2.6) increases over each iteration until converges.

2.3.3 Problem Formulation Based on Users' Fairness

When the pilot allocation is optimized for maximizing the sum rate of the system, the cell-edge user's rate (i.e., users' fairness) is not taken into account. If the same pilot is allocated to cell-edge users at different cells, the pilot contamination occurs and it intensively deteriorates

Algorithm 1: Proposed SR-M Algorithm

Initialize cell index l, pilot allocation $\boldsymbol{\theta}_{-l}$ (assume $l=1$), tolerance ε, iterative index $t=1$.
repeat
  Obtain the optimum pilot allocation $\boldsymbol{\theta}_l$ at the lth cell according to the Hungarian method.
  Get the pilot allocation results $\boldsymbol{\theta}^t$.
  Compute the uplink rum rate according to $R(\boldsymbol{\theta}^t)$.
  Update $t \leftarrow t+1$, $l \leftarrow l+1$.
  if $l > L$ **then**
    Update $l \leftarrow 1$.
  end if
until $R(\boldsymbol{\theta}^{t+1}) - R(\boldsymbol{\theta}^t) < \varepsilon$;
Obtain optimum pilot allocation $\boldsymbol{\theta}^t$.

the rates of these users. Thus, users' fairness aware pilot allocation should be considered. For this purpose, we formulate the pilot allocation optimization problem for maximizing the sum of user's logarithmic rate as follows:

$$\begin{aligned} & \max_{\boldsymbol{\theta}} \; R(\boldsymbol{\theta}) = \sum_{l=1}^{L}\sum_{k=1}^{K} \log\left((1-\eta) r_{lk}\right) \\ s.t \quad & \theta(l,k) \in \{1,2,\cdots,K\}, \forall l,k, \\ & \theta(l,k) \neq \theta(l,k'), k \neq k'. \end{aligned} \tag{2.11}$$

As for problem (2.11), we can use the similar algorithm to problem (2.6) to obtain the corresponding pilot allocation. The algorithm consists of the following four steps:

1. Divide problem (2.11) into L subproblems.
2. Optimize pilot allocation for users in one cell while fixing pilot allocation in others cell.
3. Move to the next cell and do the same optimization as step 2.
4. Repeat steps 2 and 3 until sum logarithmic rate $\log\left((1-\eta) r_{lk}\right)$ converges.

We call the above algorithm as user's fairness aware (UF-A) algorithm. Since the similar algorithm in (i.e., Algorithm 1) is applied, we omit the detailed explanations of the algorithm. The related results will be presented in simulation section directly.

Table 2.2 Simulation Parameters.

Parameters	Value
Radius of cell r_c	500 m
Radius of cell hole r_h	100 m
Number of users K	$2 \leq \mathrm{M} \leq 8$
Number of BS antennas M	$10 \leq \mathrm{M} \leq 500$
Number of cells L	7
Transmit power of users	0dBm
Time-frequency coherent block size S	100 symbols
Bandwidth	20 MHz
Noise Power	-174dBm/Hz

2.4 Numerical Results and Discussion

In this section, we evaluate the average uplink rate of the proposed pilot allocation schemes. We consider an $L = 7$ typical hexagonal cellular network where each BS is equipped with M antennas, and there are K users in each cell. Therefore, the proposed algorithm works to maximize total sum rate of 7 cells as defined in problem (2.6). We assume that cell radius is $r_c = 500$ meters, and cell-hole radius $r_h = 100$ meters. The large-scale fading coefficient captures the path-loss effect as follows $\beta_{ljk} = 1/d_{ljk}^{\alpha}$ [19], where d_{ljk} denotes the distance between the l-th BS and the k-th user in the j-th cell, and $\alpha = 3.8$ is the path-loss exponent. Users are distributed randomly within each cell, and Monter-Carlo method is applied with 10^4 simulation for single user having random location in each trail. Note that (2.5) is used to compute the uplink rate of each user, while the approximated user-rate in (2.7) is used to solve the problems (2.6). The system parameters are summarized in Table 2.2.

In fact, similarly to [20], the inter-cluster interference should be also considered. Fig. 2.3 shows system model where there are multiple clusters (different colors stand for different clusters). Since there is no any cooperation among clusters, the cluster cannot know necessary information of adjacent clusters such as user's location information and pilot allocation formation. Thus, the average interference power from outer-cluster cells should be estimated without the above information. For this purpose, we propose the following approximate scheme. We only consider the interference from adjacent outer-cluster cells due to the very slight interference for non-adjacent outer-cluster cells. When each cell estimates the average interference from outer-cluster cells, the BS'location is assumed as the user's location.

Fig. 2.4 plots the average uplink rate versus number of BS antennas with different algorithms when the number of users in each cell is 4. It can be clearly found that the average

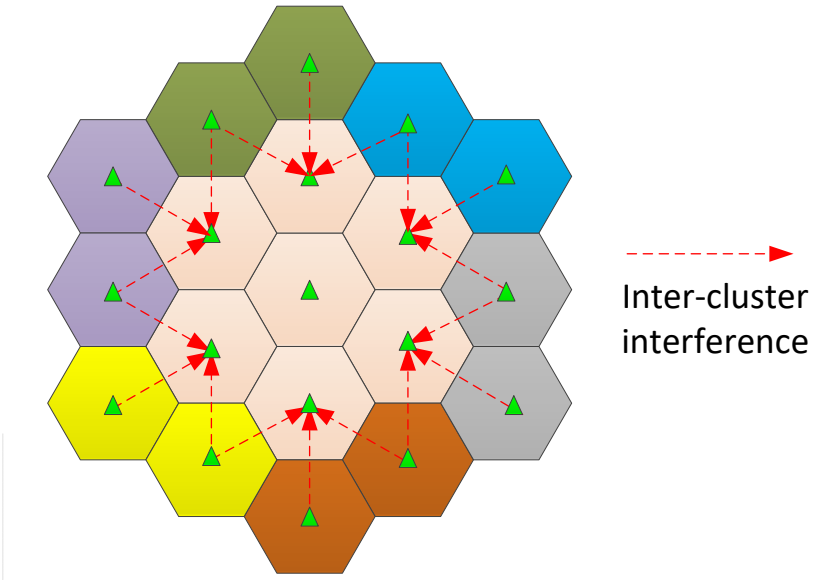


Fig. 2.3 The model of the inter-cluster interference and different colors denote the different clusters.

uplink rate increases with M under all algorithms, and the average uplink rate under the proposed SR-M algorithm is almost the same with that under the exhaustive search algorithm. In exhaustive search scheme, the best pilot allocation to maximize the average uplink sum rate is selected among all possible candidates. In random allocation scheme, pilot allocation is randomly determined regardless of the achievable uplink sum rate. We can find that the average uplink rate of the proposed UF-A algorithm is lower than that of the proposed SR-M algorithm and is higher than that of the random allocation algorithm. The reason is that the achievable sum rate has to be sacrificed for improving the users' fairness with the proposed UF-A algorithm. In addition, we can also find that per-user rate can be improved by about 17% by using the proposed SR-M algorithm in comparison with the random allocation algorithm.

Fig. 2.5 shows that the average uplink rate per user versus the number of users in each cell with different algorithms. We can find that the average uplink rate decreases with K increases. In fact, there are two reasons for this result. The first is that (1-η) decreases as K increases, which reduces the uplink rate per user. The second is that the degree of freedom (DoF) of the BS antennas decreases with the number of serviced users increases, which leads to the decline of the average rate. It is also easy to understand that more number of BS antennas leads to higher rate. Although the average uplink rate decreases with the number

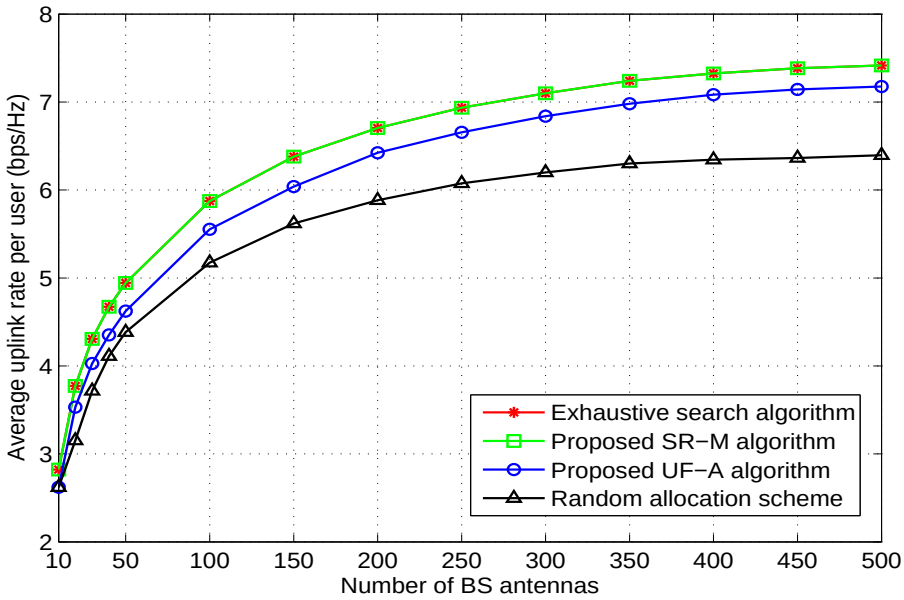


Fig. 2.4 The average uplink rate versus the number of BS antennas with different algorithms ($K = 4$).

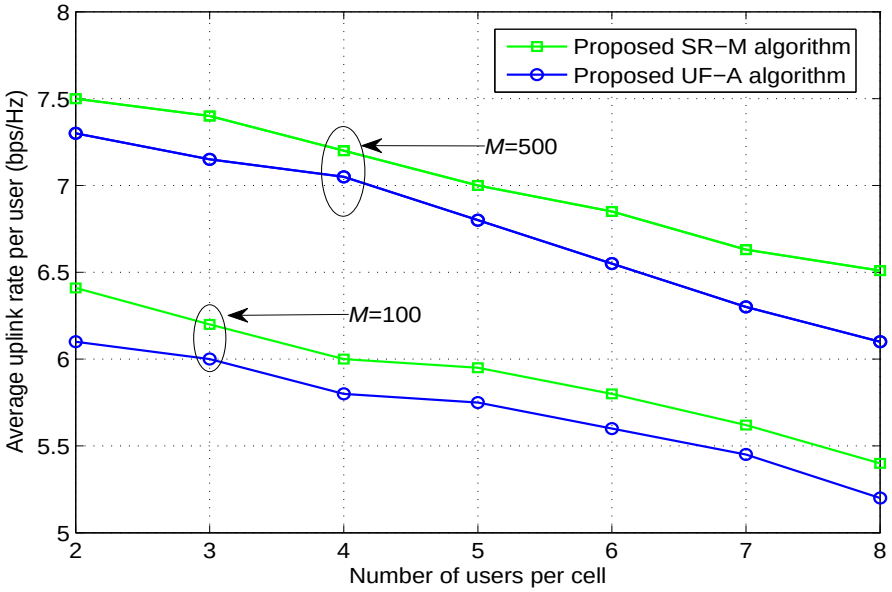


Fig. 2.5 The average uplink rate versus the number of users per cell.

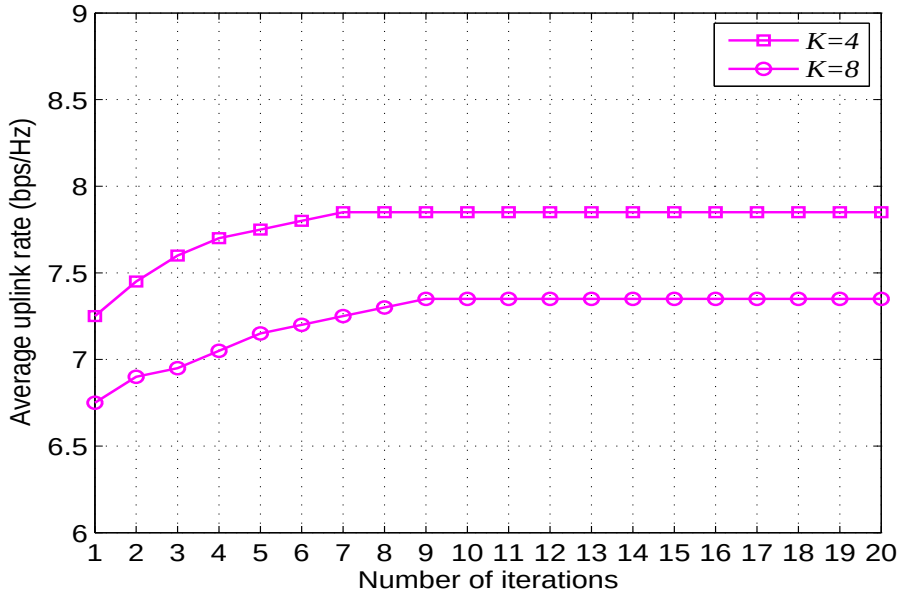


Fig. 2.6 The average uplink rate versus the number of iterations.

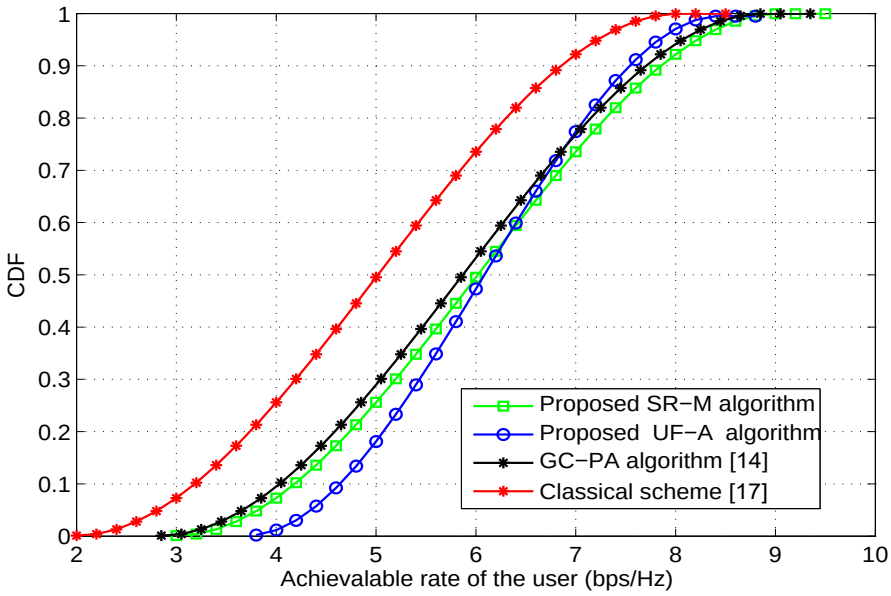


Fig. 2.7 CDF versus users' uplink achievable rate (bps/Hz) (K=4,M=100).

of users, the uplink sum rate increases, and we can get it by the proposed low complexity algorithms according to Fig. 2.5. On the other hand, we can get that the uplink sum rate of the system will increase when it services more users, but the average uplink rate per each user will decrease, which lowers each user's experience. Therefore, in practice, the tradeoff between number of serviced users and each user's experience needs to be considered .

Fig. 2.7 shows the cumulative distribution function (CDF) curve of users' uplink achievable rate with $K = 4$ and $M = 100$. The graph coloring based pilot allocation (GC-PA) [14] and classical random pilot allocation scheme [17] are compared with our proposed schemes. We can find that that the uplink rate with our proposed SR-M algorithm is higher than that with GC-PA algorithm. Meanwhile, it can be verified that the user's rate is more concentrated with UF-A algorithm than that with SR-M algorithm, which means that the UF-A improves the users' fairness. In addition, it is clear that the classical scheme has the worst performance compared with other algorithms.

2.5 Conclusions

In this chapter, we have proposed an optimum pilot allocation scheme to improve uplink sum rate in mMIMO systems. Firstly, we formulate the pilot allocation optimization problem for maximizing uplink sum rate of the system. Since the high complexity for solving the original problem, we transform the formulated problem into several subproblems. In each subproblem, we obtain the optimum pilot allocation by applying Hungarian method. Through multiple iterations, the optimum pilot allocation is found. For improving users' fairness, we formulate the maximization problem of sum of user' logarithmic rate and use the similar algorithm to obtain the corresponding pilot allocation. Simulation results show that per-user rate can be improved by about 17% by using the proposed SR-M algorithm in comparison with the conventional random allocation algorithm.

Chapter 3

Pilot Allocation for Cognitive Radio Type Heterogeneous Network with Massive MIMO

3.1 Introduction

Different from chapter 2, in this chapter, we will investigate the pilot allocation in mMIMO-CR HetNet. Although there are some related works, most of them focus on traditional MIMO-CR system with regular antennas. For example, [21] considers the achievable rate and power efficiency for mMIMO in spectrum-sharing networks. For maximizing the quality of channel estimation for the CN, a pilot decontamination algorithm is proposed [22]. In [23], a reciprocity-based CR beamforming scheme is proposed to reduce the interference from CUs to PUs. However, works in [21, 23] do not involve the pilot allocation problem between PN and CN.

In TDD-based mMIMO system, it is well known that the uplink pilot symbol is a significant and limited resource because the short channel coherence time limits the number of orthogonal pilots. Although the orthogonal pilots are used for channel estimation in each cell of the PN, the same pilots have to be reused in adjacent cells due to limited orthogonal pilots, which results in pilot contamination. Similarly to the PN, the CN is required to allocate the orthogonal pilots to CUs for their channel estimation. However, if CUs use the same pilots or non-orthogonal pilots with PUs, it causes the serious pilot contamination between PUs and CUs. Therefore, an effective pilot sharing scheme between the PN and CN is necessary.

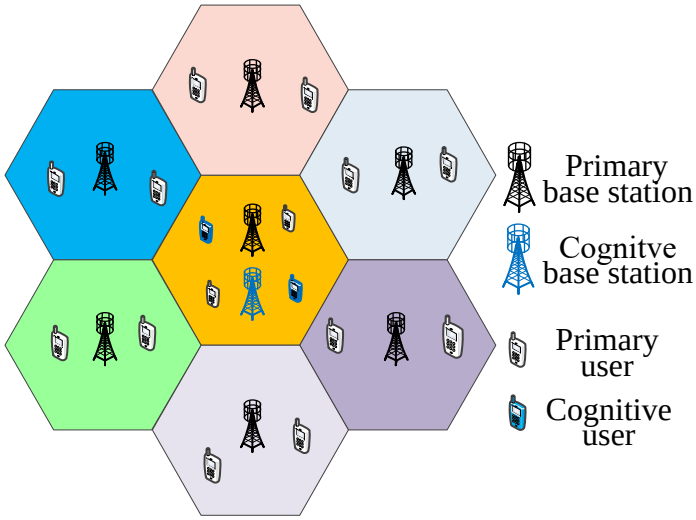


Fig. 3.1 System model for spectrum-sharing mMIMO networks.

In this chapter, we study pilot allocation problem mMIMO-CR HetNet. In our approach, PN and CN are regarded as the leaser and lessee, respectively. The CN is allowed to lease a part of available orthogonal pilots from PN. Consequently, PN can obtain profits by leasing pilots to CN. We assume that PN and CN are rational and selfish, and they aim at maximizing their revenue when pilots are traded. To guarantee success of pilot trade, we propose a three-side pilot trade platform, including price control side (PCS), PN side and CN side. Specifically, for given pilot lease price, PN side will lease the optimum pilots to CN for maximizing its revenue. Then, CN side allocates these pilots to some of CUs for maximizing its revenue. To realize the above, we propose a price-based iterative optimum pilot allocation algorithm to maximize the profits of PN and CN.

3.2 System Model and Proposed Scheme

3.2.1 System Model

We consider a communication system that consists of one L hexagonal primary-cells (PCs) PN and one single hexagonal cognitive-cell (CC) CN, as shown in Fig. 3.1. We assume that CC has the same coverage area with the central PC. For convenience, the central PC and CC are denoted as the 1st PC and CC (target cell), respectively. Each PC consists of a PBS equipped with M antennas and K ($\mathscr{K} = \{1,\ldots,K\}$) single-antenna PUs ($M \gg K$). We assume that the same K orthogonal pilots sequences $\mathbf{\Psi} = \{\boldsymbol{\psi}_1,\ldots,\boldsymbol{\psi}_K\} \in \mathbb{C}^{K\times K}$ are

assigned to K PUs uniquely in each PC, where $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{I}_M$, and no pilots are assigned to CUs. $\boldsymbol{\psi}_{\theta(l,k)}(\theta(l,k) \in \{1,\ldots,K\})$ denotes that the $\theta(l,k)$-th pilot is used by the k-th PU in the l-th PC, where $\theta(l,k) \neq \theta(l,k')$ when $k \neq k'$. Uplink rate of the k-th PU in the 1st PC can be expressed as [5]

$$r_{1,k}^p = \log_2(1 + \mathrm{SINR}_{1,k}^p), \tag{3.1}$$

where $\mathrm{SINR}_{1,k}^p \approx \frac{{\beta_{11k}^p}^2}{\sum_{l\neq 1}^{L}\sum_{i=1}^{K} f(\theta(l,i),\theta(1,k)){\beta_{1li}^p}^2}$ when $M \to \infty$, β_{1li}^p denotes the large-scale fading coefficient (LFC) between the PBS in the 1st PC and the i-th PU in the l-th PC. $\beta_{1li}^p = 1/{d_{1li}^p}^{\alpha}$, where d_{1li}^p denotes the distance between PBS in the 1st PC and the i-th PU in the l-th PC , and α is path-loss exponent. $\sum_{j\neq l}\sum_{i=1}^{K} f(\theta(j,i),\theta(1,k)){\beta_{1ji}^p}^2$ denotes pilot contamination caused by pilot reuse in adjacent cells, $f(\theta(j,i),\theta(l,k)) = 1$ when $\theta(j,i) = \theta(l,k)$, else $f(\theta(j,i),\theta(l,k)) = 0$.

In CC, there are a M-antenna CBS and K single-antenna CUs. Similarly, uplink rate of the k-th CU in the CC can be expressed as

$$r_{1,k}^s = \log_2(1 + \mathrm{SINR}_{1,k}^s), \tag{3.2}$$

where $\mathrm{SINR}_{1,k}^s \approx \frac{{\beta_{11k}^s}^2}{\sum_{l\neq 1}^{L}\sum_{i=1}^{K} f(\theta(l,i),\theta(1,k)){\beta_{1li}^s}^2}$ when $M \to \infty$, β_{11i}^s denotes the LFC between CBS and the i-th CU in the CC, and β_{1li}^s denotes the LFC between CBS in the 1st PC and the i-th PU in the l-th PC.

CN can lease pilots from PN for channel estimation. PN can get some revenue from CN (i.e., lease fee). We assume that PN and CN are rational and selfish. For given pilot lease price, PN always leases the optimum pilots to CN for maximizing its revenue, while CN always optimally allocates these pilots to CUs for maximizing its own revenue.

Since we assume that CC has the same coverage area with the 1st PC, CN leases pilots from the 1st PC. Therefore, the revenue of PN can be given as follows:

$$\max_{\mathbf{\Psi}_\mathrm{S},\mathscr{K}_\mathrm{P}} \quad \underbrace{m|\mathbf{\Psi}_\mathrm{S}|}_{\substack{\text{Price of leased pilots}\\ \text{from PN to CN}}} \quad \underbrace{-n \sum_{i\in\mathscr{K}_\mathrm{P},\, \psi_{\theta(1,i)}\in\mathbf{\Psi}_\mathrm{S}} r_{1,i}^p}_{\text{Lost utility(PUs' rate)}}, \tag{3.3a}$$

$$= \max_{\mathbf{\Psi}_\mathrm{S},m} m|\mathbf{\Psi}_\mathrm{S}| - \min_{\mathbf{\Psi}_\mathrm{S},\mathscr{K}_\mathrm{P}} \quad n \sum_{\substack{i\in\mathscr{K}_\mathrm{P}\\ \psi_{\theta(1,i)}\in\mathbf{\Psi}_\mathrm{S}}} r_{1,i}^p, \tag{3.3b}$$

where m is the lease price per pilot, n denotes the price per rate of PN which is determined by PN. $\mathbf{\Psi}_\mathrm{S}$ ($\mathbf{\Psi}_\mathrm{S} \subset \mathbf{\Psi}$) and $|\mathbf{\Psi}_\mathrm{S}|$ denote the leased pilots set to CN and the number in set $\mathbf{\Psi}_\mathrm{S}$,

respectively. $\mathscr{K}_{\mathrm{P}}(\mathscr{K}_{\mathrm{P}} \subset \mathscr{K})$ denotes the pilots allocated to PUs in the 1st PC before leasing pilots. In (3.3a), the first term denotes the obtained revenue because of leased pilots. The second term denotes the lost utility because those PUs will not connect the PBS. From (3.3b), we can find that the first term is a increasing linear function with $|\mathbf{\Psi}_{\mathrm{S}}|$ for given m, and the second term is a function of growing faster with $|\mathbf{\Psi}_{\mathrm{S}}|$. Therefore, for given pilot lease price, the 1st PC can find the optimum pilots to the CN for maximizing its revenue. Since there is no relationship between the secondary term and m, (3.3b) is a monotone non-decreasing function with m. This analysis assumes that users are randomly distributed in each cell.

At the CN side, for given pilot price m and pilots $\mathbf{\Psi}_{\mathrm{S}}$, the total revenue can be expressed as:

$$\max_{\mathscr{K}_{\mathrm{S}}} \quad \underbrace{c \sum_{\substack{i \in \mathscr{K}_{\mathrm{S}} \\ \psi_{\theta(1,i)} \in \mathbf{\Psi}_{\mathrm{S}}}} r^{\mathrm{s}}_{1,i}}_{\text{Obtained utility(CUs' rate)}} - \underbrace{m|\mathbf{\Psi}_{\mathrm{S}}|}_{\substack{\text{Fee for leased pilots} \\ \text{from PN to CN}}}, \tag{3.4a}$$

$$\stackrel{\Delta}{=} \max_{\mathscr{K}_{\mathrm{S}}} \; c \sum_{\substack{i \in \mathscr{K}_{\mathrm{S}} \\ \psi_{\theta(1,i)} \in \mathbf{\Psi}_{\mathrm{S}}}} \frac{{\beta^s_{11i}}^2}{\sum\limits_{j \neq 1} \sum\limits_{m=1}^{K} f(\theta(j,m),\theta(1,i)) {\beta^s_{1jm}}^2}, \tag{3.4b}$$

where c denotes the price per rate of CN which is determined by CN, and $\mathscr{K}_{\mathrm{S}}(\mathscr{K}_{\mathrm{S}} \subset \mathscr{K})$ is a set of CUs allocated pilots. In (3.4a), the first term denotes obtained utility because of leased pilots from the 1st PC. To obtain more revenue, CN will optimally allocate these pilots to some of CUs. Since the second term in (3.4a) is given constant, we only need to maximize (3.4b). (3.4b) is an optimal matching problem, and matching rule is that one pilot is assigned to one CU, but different users are not assigned the same pilot. To maximize CUs' total SINR, (3.4b) can be solved by applying the well-known Hungarian algorithm, as our description in Chapter 3.

To guarantee the 1st PC's service quality, number of available pilots in CN must be limited. We define provided primary user ratio (PUR) as the number of PUs under service normalized by the number of all PUs in the 1st PC. We also define the required minimum PUR as p, which is expressed as:

$$|\mathbf{\Psi}_{\mathrm{S}}| \leq (1-p)K. \tag{3.5}$$

where $(1-p)K$ denotes the number of maximum orthogonal pilots leased to the CN.

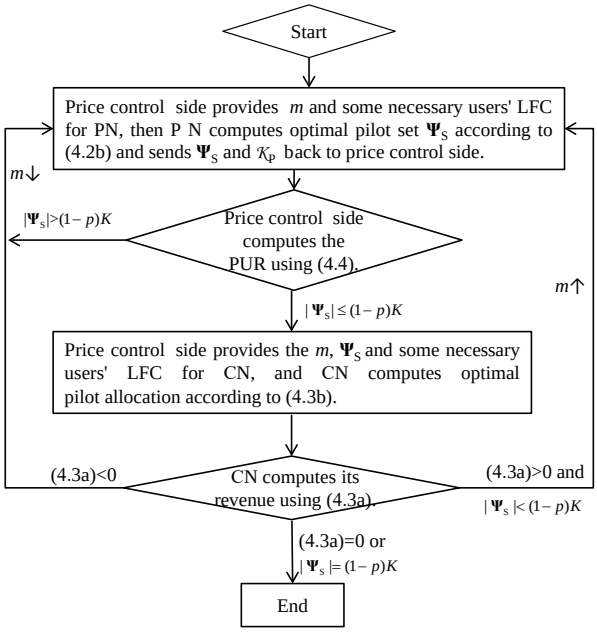


Fig. 3.2 Proposed pilot allocation algorithm flow diagram.

3.2.2 Proposed Scheme

To ensure the success of pilot trade, we propose a three-side pilot trade platform, including PCS, PN side and CN side, where PCS is in charge of the pilot trade between PN and CN, while PN selects the set $\mathbf{\Psi}_S$ only according to m given by PCS. The mechanism to encourage the pilot trade is explained below. According to (3.3) and (3.4), if m is decreased, CN's revenue tends to be decreased, while CN's revenue tends to be increased. Consequently, the pilot trade from PN to CN is encouraged by decreasing m. PN and CN are able to know their own users' LFC, respectively. Here, the revenue of CN can not be negative, and the pilot trade will be finished when one of the following conditions is satisfied: i) (3.5) becomes tight constraint, ii) revenue of CN becomes 0. Based on the above analysis, a price-based iteration optimal pilot allocation algorithm for maximizing the profits of the PN and SN can be described as follows:

1) PCS provides a pilot lease price m and users' CSI information for PN, then PN computes optimal pilot set $\mathbf{\Psi}_S$ according to (3.3b) and sends $\mathbf{\Psi}_S$ and $\mathscr{K}_S$ back to PCS.

2) If PCS finds that the SUR, namely, (3.5) is not satisfied, it will decrease the pilot lease price and repeat 1) until the SUR is satisfied. Then PCS provides the pilot lease price, $\mathbf{\Psi}_S$ and users' CSI information for CN, and CN computes optimal pilot allocation according to (3.3b).

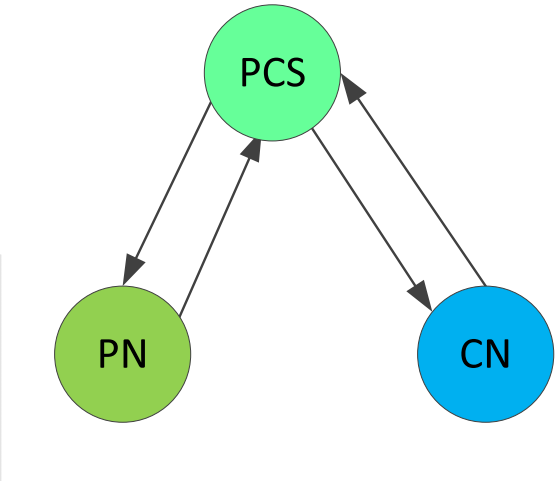


Fig. 3.3 The relation among PCS, PN and CN.

3) If CN finds that its revenue is negative, CN does not rent these pilots and feedback this information to PCS. Then PCS will decrease price and repeat 1) and 2) until CN's revenue nonnegative. If CN finds that its profits is 0, pilot trade is finished. Else, CN will send the pilot allocation results to PCS.

4) When PCS finds that the constraint (3.5) is slack, it will increase price and repeat 1)∼4) until pilot lease price converges.

The flow of the proposed algorithm is illustrated in Fig. 3.2, and the relation among PCS, CN and CN is shown in Fig. 3.3

3.3 Simulation Results and Discussions

We consider a typical hexagonal cellular network with 7 PCs and 1 CC (as shown in Fig.3.1), and each BS is equipped with M antennas (M goes infinite) [22]. We assume there are 100 PUs in each PC and 100 CUs in CC, and all users are randomly distributed in each cell. The number of orthogonal pilots is 100. The cell radius is $r_c = 500$ meters, and the cell-hole radius is $r_h = 100$ meters (the terminals do not figure in this scenario). $n = 1$, $c = 1$ and α=3.8 are assumed. m, n, c and revenue are regarded as the actual currency.

Fig. 3.4 shows the revenue of PN and CN. We can find that PN's revenue increases with (1-p) and gets maximum at higher (1-p) region. On the other hand, CN's revenue first increases and then decreases with (1-p). This is because the first term in (3.4a) is a logarithmic function with respect to (w.r.t.) (1-p), and the second term in (3.4a) is a linear

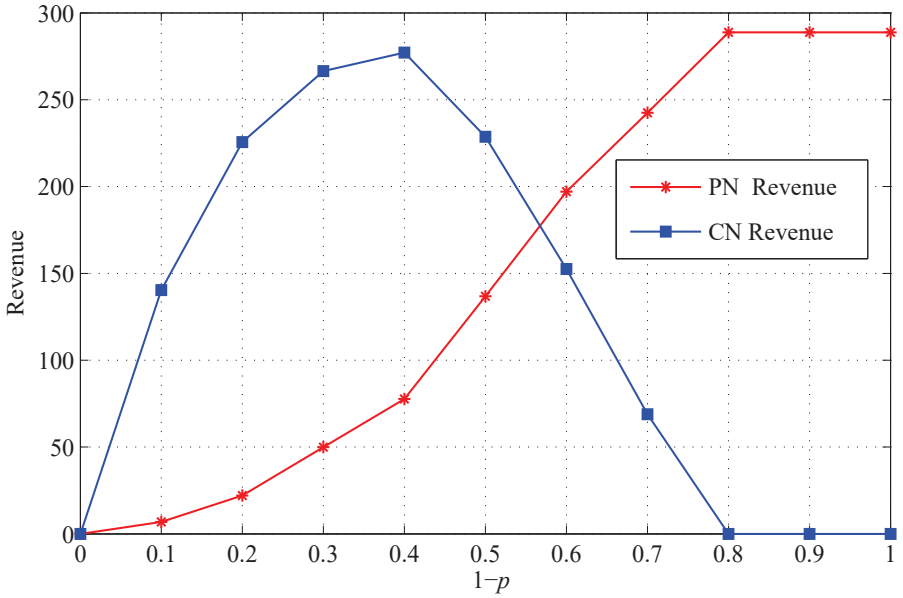


Fig. 3.4 Revenue versus (1-p).

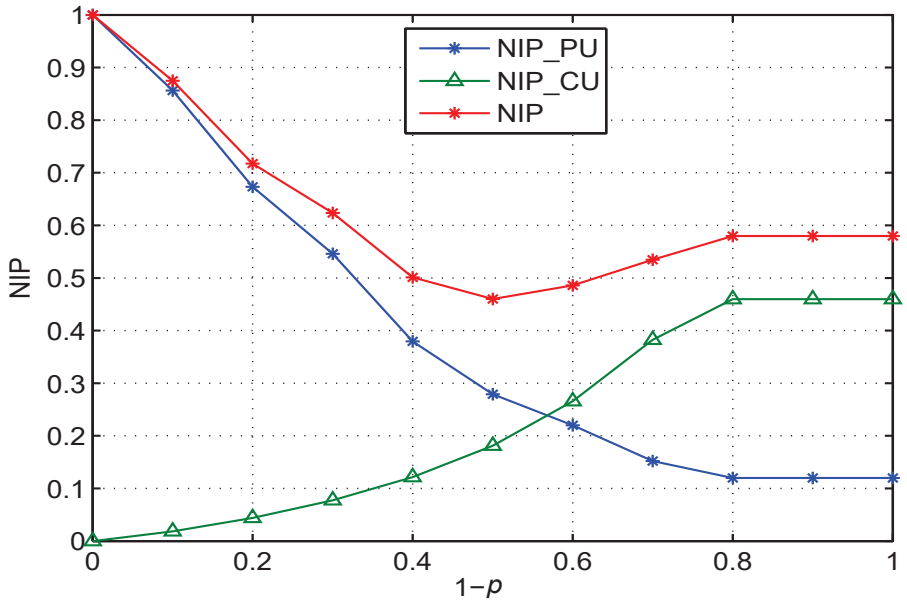


Fig. 3.5 NIP versus (1-p).

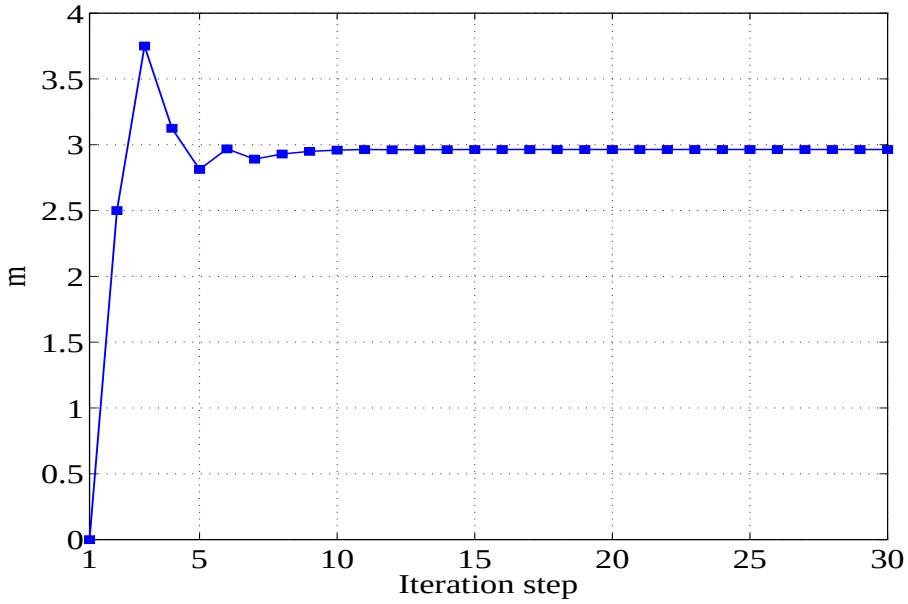


Fig. 3.6 Pilot lease price versus iteration step when p=0.8.

function w.r.t. (1-p). Therefore, as (1-p) increases, according to the basic math theory, CN's revenue first increases and then decreases. In Fig.. 3.5, we define the normalized interference power (NIP) as $\mathrm{I_{after}}/\mathrm{I_{before}}$:

$$\mathrm{NIP} = \frac{\mathrm{I_{after}}}{\mathrm{I_{before}}} = \frac{\mathrm{I_{after,PU}} + \mathrm{I_{after,CU}}}{\mathrm{I_{before}}} = \underbrace{\frac{\mathrm{I_{after,PU}}}{\mathrm{I_{before}}}}_{\mathrm{NIP\_PU}} + \underbrace{\frac{\mathrm{I_{after,CU}}}{\mathrm{I_{before}}}}_{\mathrm{NIP\_CU}} \tag{3.6}$$

where $\mathrm{I_{after}}$ and $\mathrm{I_{before}}$ denote the total interference power from both PUs and CUs to adjacent PBSs after pilot leasing and before leasing, respectively. $\mathrm{I_{after,PU}}$ and $\mathrm{I_{after,CU}}$ denote interference power to adjacent PBSs from PUs and interference power to adjacent PBSs from CUs, respectively. As the number of leased pilots to CN increases, NIP_PU decreases while NIP_CU increases. Thus, NIP first decreases and then increases, and an optimum value of NIP can be observed for a given $(1-p) = 0.5$. As (1-p) increases, more pilots will be leased to CN and less PUs can connect to PBS, which results in the decrease of NIP_PU. On the contrary, as (1-p) increases, more CUs will be allocated pilots and connected to CBS, which leads to the increase of NIP_CU. Note that, NIP_PU, NIP_CU and NIP are kept constant when $(1-p) > 0.8$. This is because there is no pilot trade for $(1-p) > 0.8$. In addition, the total interference from central cell to adjacent cells is lower compared after pilots leasing. This is because PN will lease those pilots having serious pilot interference to CN, and CN will allocate those pilots to CUs having the smallest pilot interference, which

results in the decrease of total interference after pilot trade. Fig. 3.6 presents the pilot lease price convergence behavior at each iteration when p=0.8.

3.4 Conclusions

In this chapter, we have studied the pilot allocation problem in mMIMO-CR HetNet. The CN is allowed to rent the orthogonal pilots from the PN for their channel estimation, while the number of available pilots in the PN is reduced. The proposed system model and algorithm achieve a win-win paradigm between PN and CN, The results show that the PN and CN can obtain positive revenue, which implies that pilot sharing concept between PN and CN is effective in improving the performance of both PN and CN. In other words, although the PN must sacrifice some pilots, it can get more profit in addition to decrease the total interference to adjacent cells, while the CUs are allowed to share the orthogonal pilots with PUs.

Chapter 4

Power Allocation for Cognitive Radio Type Heterogeneous Network with Massive MIMO

4.1 Introduction

In chapter 3, we have proposed a pilot allocation scheme for the mMIMO-CR HetNet, but the infinite number antennas at BS is assumed and the power allocation is also not considered for simplicity. Therefore, in this chapter, we will consider the limited number antennas at BS and power allocation problem. For traditional MIMO-CR networks [24]-[27], PUs' interference is caused by CUs' data transmission. However, for mMIMO-CR networks, PUs' interference is also impacted by pilot transmission. The authors in [28] guarantee PUs' QoSs in mMIMO-CR networks by setting the peak interference level and study the impact of large-scale PBS antennas on mMIMO CN. However, the pilot contamination is not considered. A pilot allocation scheme for mMIMO-CR networks in [29] has been proposed to maximize the channel estimation quality of CUs while minimizing a negative impact on PN's channel estimation. In [30], a reciprocity-based mMIMO-CR beamforming scheme has been proposed to reduce the interference from CUs to PUs. To reduce pilot contamination and training overhead, a full-space spatial spectrum-sharing for mMIMO-CR networks with reduced training overhead is proposed in [31] with an efficient 2-dimensional-discrete Fourier transform aided direction of arrival and angular spread estimation. However, in [29]-[31], the power allocation of the CN is not considered. Efficient power allocation scheme may eliminate (or significantly reduce) harmful interference to the PN while maximizing the

performance of the CN. In [19], the authors investigate the pilot and power allocation problem to maximize EE of multi-cell mMIMO networks. However, the CR network is not considered.

In this chapter, we study the power allocation problem in mMIMO-CR HetNet with pilot contamination. Unlike the conventional approach, where PUs' tolerated interference levels are imposed to guarantee their QoSs, we introduce the required signal-to-interference-plus-noise-ratio (SINR) for PUs further improving the performance of the CN. The main contributions are summarized as follows:

- We propose an orthogonal pilot sharing scheme in TDD-based mMIMO-CR HetNet, where CUs always share the overall spectrum with PUs if they are allowed to access the primary spectrum. Since the orthogonal pilots are preferentially allocated to PUs, CUs are only allowed to access the primary spectrum when there are temporarily unused orthogonal pilots. Then, these CUs use obtained orthogonal pilots for channel estimation in pilot transmission phase.
- We derive the CU's ergodic downlink rate and formulate the power allocation optimization problem to maximize the downlink sum rate of the CN subject to the total transmit power constraint. To guarantee PUs' QoSs, the required SINR for each PU is considered. Since the formulated problem is nonconvex, we transform it into a convex one by using convex approximation techniques. Then, an iterative algorithm is proposed to obtain the solution. Meanwhile, we prove that the obtained solution satisfies the necessary Karush-Kuhn-Tucker (KKT) conditions of the original problem.
- We theoretically analyze and discuss the impact of the number of CBS and PBS antennas on the downlink rate of the PN and CN. Our findings illustrate that: *i*) for the fixed number of CBS antennas the downlink rate of each CU is close to zero when the number of PBS antennas approaches to infinity; *ii*) for the fixed number of PBS antennas the downlink rate of some PUs is close to zero when the number of CBS antennas approaches to infinity; *iii*) when the number of PBS and CBS antennas grows simultaneously the downlink rate of PUs or CUs is affected by the transmit power and pilot contamination from adjacent cells.

4.2 System Model

We consider a downlink communication system that consists of a multi-cell multi-user mMIMO-PN and a single cell multi-user mMIMO-CN as shown in Fig. 4.1. We assume that there are L cells in PN, where each cell consists of a M_{P} antennas PBS and K_{P} single-antenna

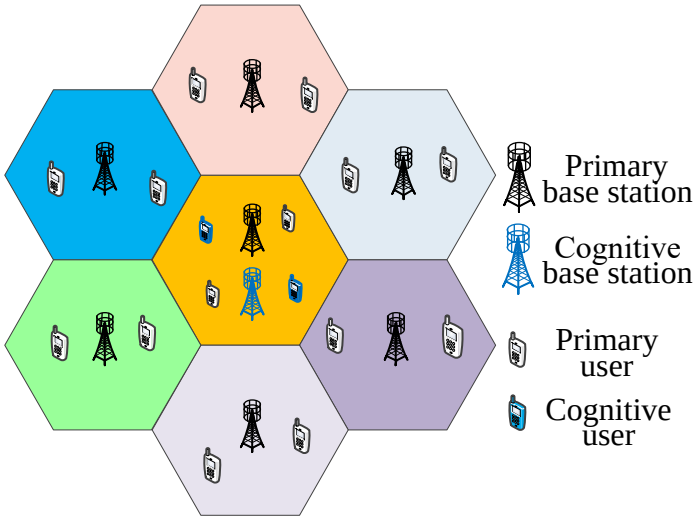


Fig. 4.1 System model for mMIMO-CR networks.

PUs. In CN, there are a M_{S} antennas CBS and K_{S} single-antenna CUs. All PBSs or CBS serves their own users with the same time-frequency resource. The CUs located in the coverage region of the central PC share resource (i.e., pilots and spectrum) with PUs. It means that, theoretically, the single cell CN can have similar coverage area with that of the central PC (i.e., the CBS has the similar location with the central PBS). The central cell in PN is labeled as the 1-st cell. To avoid serious pilot interference, all PUs in each cell use orthogonal pilots $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \cdots, \boldsymbol{\psi}_{K_{\mathrm{P}}}]^T \in \mathscr{C}^{K_{\mathrm{P}} \times K_{\mathrm{P}}}$, where $\boldsymbol{\psi}_i^{\dagger} \boldsymbol{\psi}_i = \rho$ and ρ is the pilot signal power. The i-th pilot is allocated to the i-th PU for channel estimation in each cell. The same orthogonal pilot sequences are reused in adjacent cells. When there are K_{T} ($K_{\mathrm{S}} \geq K_{\mathrm{T}}$) inactive PUs in the 1-st cell, the pilots $\{1, \cdots, K_{\mathrm{P}} - K_{\mathrm{T}}\}$ are used by PUs in the 1-st cell and the pilots $\{K_{\mathrm{P}} - K_{\mathrm{T}} + 1, K_{\mathrm{P}} - K_{\mathrm{T}} + 2, \cdots, K_{\mathrm{P}}\}$ can be used by CUs. Similarly, we assume that the $(K_{\mathrm{P}} - K_{\mathrm{T}} + n)$-th pilot in the 1-st cell is allocated to the $(K_{\mathrm{P}} - K_{\mathrm{T}} + n)$-th CU ($n = \{1, 2, \cdots, K_{\mathrm{T}}\}$).

The following assumptions are adopted in our study:

- A high speed backhaul link is available between PBS and CBS for feedback of the channel state information and users' locations information [32]-[34].
- Ideal synchronization between PN and CN is realized by a backhaul link [32]-[34].
- The PBS periodically shares the available primary pilot information with the CBS by Operation, Administration and Management system via backhaul link [32]-[34].
- The deployment of CBSs is sparser in comparison with PBSs. Thus, the interference between CBSs can be neglected.

4.2.1 Uplink Training Transmission

During the uplink training phase, the CBS receives the signal matrix $\mathbf{Y}_1^{\mathrm{S}} \in \mathscr{C}^{M_{\mathrm{S}} \times K_{\mathrm{P}}}$, which is expressed as

$$\mathbf{Y}_1^{\mathrm{S}} = \sum_{j=2}^{L} \sum_{k=1}^{K_{\mathrm{P}}} \sqrt{\beta_{1jk}^{\mathrm{PS}}} \mathbf{h}_{1jk}^{\mathrm{PS}} \boldsymbol{\psi}_k^T + \sum_{m=1}^{K_{\mathrm{P}}-K_{\mathrm{T}}} \sqrt{\beta_{11m}^{\mathrm{PS}}} \mathbf{h}_{11m}^{\mathrm{PS}} \boldsymbol{\psi}_m^T + \sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}} \sqrt{\beta_{11n}^{\mathrm{SS}}} \mathbf{h}_{11n}^{\mathrm{SS}} \boldsymbol{\psi}_n^T + \mathbf{V}_1. \quad (4.1)$$

In the above expression, $\beta_{1jk}^{\mathrm{PS}}$ and $\mathbf{h}_{1jk}^{\mathrm{PS}}$ denote the large-scale fading coefficient and the $M_{\mathrm{S}} \times 1$ small-scale fading vector for the channel between the CBS in the 1-st cell and the k-th PU in the j-th cell, respectively. $\beta_{11k}^{\mathrm{SS}}$ and $\mathbf{h}_{11k}^{\mathrm{SS}}$, respectively, denote the large-scale channel coefficient and the $M_{\mathrm{S}} \times 1$ small-scale fading vector for the channel between the CBS and the k-th CU in the 1-st cell. $\mathbf{V}_1 \in \mathscr{C}^{M_{\mathrm{S}} \times K_{\mathrm{P}}}$ is AWGN at CBS whose elements are distributed according to $\mathscr{CN}(0, \sigma^2)$. We assume each fading vector $\mathbf{h}_{ijk}^{\star} \sim \mathscr{CN}(\mathbf{0}, \mathrm{I}_{M_{\mathrm{S}}})$ [19] where $\star \in \{\mathrm{SS}, \mathrm{PS}\}$. The large-scale fading coefficient is assumed capturing the path-loss effect as $\beta_{1jk}^{\star} = 1/d_{1jk}^{\alpha}$ [19]: d_{1jk} denotes the distance between CBS and the k-th PU (CU) in the j-th cell and α is the path-loss exponent.

The minimum mean-squared error (MMSE) estimate of the channel $\mathbf{h}_{11n}^{\mathrm{SS}}$ can be expressed as [35]

$$\hat{\mathbf{h}}_{11n}^{\mathrm{SS}} = \sqrt{\beta_{11n}^{\mathrm{SS}}} \mathbf{Q}_{1n} \mathbf{Y}_1^{\mathrm{S}} \boldsymbol{\psi}_n, \quad (4.2)$$

where $\mathbf{Q}_{1n} = \left(\sigma^2 \mathbf{I}_{M_{\mathrm{S}}} + \rho \mathbf{I}_{M_{\mathrm{S}}} \left(\sum_{j=2}^{L} \beta_{1jn}^{\mathrm{PS}} + \beta_{11n}^{\mathrm{SS}} \right) \right)^{-1}$. The estimated channel can be written as

$$\hat{\mathbf{h}}_{11n}^{\mathrm{SS}} = \mathbf{h}_{11n}^{\mathrm{SS}} - \tilde{\mathbf{h}}_{11n}^{\mathrm{SS}}, \quad (4.3)$$

where $\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}}$ is the error term. From (2), we easily observe that $\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}$ is $\mathscr{CN}(\mathbf{0}, \boldsymbol{\Theta}_{11n})$. Since $\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}}$ is independent of $\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}$, $\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}}$ is also $\mathscr{CN}(\mathbf{0}, \mathrm{I}_{M_{\mathrm{S}}} - \boldsymbol{\Theta}_{11n})$ with

$$\boldsymbol{\Theta}_{11n} = \rho \beta_{11n}^{\mathrm{SS}} \mathbf{Q}_{1n} = \frac{\rho \beta_{11n}^{\mathrm{SS}}}{\left(\sigma^2 \mathbf{I}_{M_{\mathrm{S}}} + \rho \mathbf{I}_{M_{\mathrm{S}}} \left(\sum_{j=2}^{L} \beta_{1jn}^{\mathrm{PS}} + \beta_{11n}^{\mathrm{SS}} \right) \right)}. \quad (4.4)$$

4.2.2 Downlink Data Transmission

During the downlink data transmission phase, the received signal at the n-th CU in the 1-st cell can be written as

$$\begin{aligned} y_{1n}^{\mathrm{S}} = & \sum_{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}} \sqrt{\beta_{11n}^{\mathrm{SS}} P_{1i}} \mathbf{h}_{11n}^{\mathrm{SS}\dagger} \mathbf{w}_{11i}^{\mathrm{SS}} x'_{1i} + \sum_{m=1}^{K_{\mathrm{P}}-K_{\mathrm{T}}} \sqrt{\beta_{11n}^{\mathrm{SP}} p_{\mathrm{t}}} \mathbf{h}_{11n}^{\mathrm{SP}\dagger} \mathbf{w}_{11m}^{\mathrm{PP}} x_{1m} \\ & + \sum_{j=2}^{L} \sum_{k=1}^{K_{\mathrm{P}}} \sqrt{\beta_{j1n}^{\mathrm{SP}} p_{\mathrm{t}}} \mathbf{h}_{j1n}^{\mathrm{SP}\dagger} \mathbf{w}_{jjk}^{\mathrm{PP}} x_{jk} + v_{1n}, \end{aligned} \tag{4.5}$$

where $\beta_{j1i}^{\mathrm{SP}}$ and $\mathbf{h}_{j1i}^{\mathrm{SP}}$, respectively, denote the large-scale fading coefficient and the $M_{\mathrm{P}} \times 1$ small-scale fading vector for the channel between the PBS in the j-th cell and the i-th CU in the 1-st cell. $\mathbf{w}_{11i}^{\mathrm{SS}}$ and $\mathbf{w}_{11m}^{\mathrm{SS}}$, respectively, represent the precoding for the CU and PU. Each PU's transmit power is assumed the same and denoted by p_{t}, where P_{1i} is the transmit power of the i-th CU in the 1-st cell. x_{jk} and x'_{1i} denote the data symbols of the PU and CU, respectively, with unit average power, i.e., $\mathbb{E}\{|x_{jk}|^2\} = 1$ and $\mathbb{E}\{|x'_{1i}|^2\} = 1$. v_{1i} is AWGN received by the i-th CU in the 1-st cell.

When the number of BS antennas approaches to infinity, the performance of MF or ZF detector optimally converges. Due to the high complexity of ZF detector (high dimension channel matrix inversion [36]) and mathematical trackability of MF detector, we only consider MF detector. Following this, we have $\mathbf{w}_{11i}^{\mathrm{SS}} = \hat{\mathbf{h}}_{11i}^{\mathrm{SS}} / \|\hat{\mathbf{h}}_{11i}^{\mathrm{SS}}\|$ and $\mathbf{w}_{jjk}^{\mathrm{PP}} = \hat{\mathbf{h}}_{jjk}^{\mathrm{PP}} / \|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\|$, where $\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}$ is the MMSE estimate of $\mathbf{h}_{jjk}^{\mathrm{PP}}$.

We rewrite (4.5) as follows

$$\begin{aligned} y_{1n}^{\mathrm{S}} = & \underbrace{\sqrt{\beta_{11n}^{\mathrm{SS}} P_{1n}} \mathbf{h}_{11n}^{\mathrm{SS}\dagger} \mathbf{w}_{11n}^{\mathrm{SS}} x'_{1n}}_{\text{Desired signal}} + \underbrace{\sum_{\substack{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1 \\ i \neq n}}^{K_{\mathrm{P}}} \sqrt{\beta_{11n}^{\mathrm{SS}} P_{1i}} \mathbf{h}_{11n}^{\mathrm{SS}\dagger} \mathbf{w}_{11i}^{\mathrm{SS}} x'_{1i}}_{\text{Intra-cell interference from CN}} + \underbrace{\sum_{m=1}^{K_{\mathrm{P}}-K_{\mathrm{T}}} \sqrt{\beta_{11n}^{\mathrm{SP}} p_{\mathrm{t}}} \mathbf{h}_{11n}^{\mathrm{SP}\dagger} \mathbf{w}_{11m}^{\mathrm{PP}} x_{1m}}_{\text{Intra-cell interference from PN}} \\ & + \underbrace{\sum_{j=2}^{L} \sum_{k=1}^{K_{\mathrm{P}}} \sqrt{\beta_{j1n}^{\mathrm{SP}} p_{\mathrm{t}}} \mathbf{h}_{j1n}^{\mathrm{SP}\dagger} \mathbf{w}_{jjk}^{\mathrm{PP}} x_{jk}}_{\text{Inter-cell interference}} + \underbrace{v_{1n}}_{\text{Noise}}. \end{aligned} \tag{4.6}$$

From (4.6), we find that the interference of the CU includes three components: *i*) the first component is the interference among CUs, *ii*) the second component comes from PBS of the same cell with CBS and *iii*) the third component comes from PBSs of adjacent cells.

The ergodic downlink rate of the n-th CU in the 1-st cell is obtained as follows:

$$R_{1n} = \log_2\left(1 + \frac{P_{1n}\beta_{11n}^{\mathrm{SS}}\left|\mathbb{E}\left\{\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\}\right|^2}{\mathbf{IN}}\right) \tag{4.7}$$

with the interference term **IN** given by

$$\begin{aligned}\mathbf{IN} =& P_{1n}\beta_{11n}^{\mathrm{SS}}\mathrm{var}\left\{\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\} + \sum_{j=2}^{L}\sum_{k=1}^{K_{\mathrm{P}}}\beta_{j1n}^{\mathrm{SP}}p_{\mathrm{t}}\mathbb{E}\left\{\left|\mathbf{h}_{j1n}^{\mathrm{SP}\dagger}\mathbf{w}_{jjk}^{\mathrm{PP}}\right|^2\right\} \\ &+ \sum_{m=1}^{K_{\mathrm{P}}-K_{\mathrm{T}}}\beta_{11n}^{\mathrm{SP}}p_{\mathrm{t}}\mathbb{E}\left\{\left|\mathbf{h}_{11n}^{\mathrm{SP}\dagger}\mathbf{w}_{11m}^{\mathrm{PP}}\right|^2\right\} + \sum_{\substack{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ i\neq n}}^{K_{\mathrm{P}}}\beta_{11n}^{\mathrm{SS}}P_{1i}\mathbb{E}\left\{\left|\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11i}^{\mathrm{SS}}\right|^2\right\} + \sigma^2.\end{aligned}$$

For the given $\mathbf{P} = \left[P_{1\{K_{\mathrm{P}}-K_{\mathrm{T}}+1\}}, \cdots, P_{1K_{\mathrm{P}}}\right]$, the rate can be represented by

$$R_{1n}(\mathbf{P}) = \log_2\left(1 + \frac{(1/\tau_{1n}^{\mathrm{S}})P_{1n}{\beta_{11n}^{\mathrm{SS}}}^2\mathbb{E}^2\{\vartheta\}}{I_1 + I_2 + I_3 + I_4 + \sigma^2}\right), \tag{4.8}$$

where

$$\begin{cases} I_1 = P_{1n}\beta_{11n}^{\mathrm{SS}}\left(\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}\mathrm{var}\{\vartheta\} + 1 - \frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}\right) + \sum\limits_{\substack{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ i\neq n}}^{K_{\mathrm{P}}}\beta_{11n}^{\mathrm{SS}}P_{1i}, \\ I_2 = \sum\limits_{j=2}^{L}\beta_{j1n}^{\mathrm{SP}}p_{\mathrm{t}}\left(\frac{\beta_{j1n}^{\mathrm{SP}}}{\tau_{jn}^{\mathrm{P}}}\mathbb{E}\{\varepsilon^2\} + 1 - \frac{\beta_{j1n}^{\mathrm{SP}}}{\tau_{jn}^{\mathrm{P}}}\right), \\ I_3 = \sum\limits_{j=2}^{L}\sum\limits_{k\neq n}^{K_{\mathrm{P}}}\beta_{j1n}^{\mathrm{SP}}p_{\mathrm{t}}, \\ I_4 = \sum\limits_{m=1}^{K_{\mathrm{P}}-K_{\mathrm{T}}}\beta_{11n}^{\mathrm{SP}}p_{\mathrm{t}}. \end{cases} \tag{4.9}$$

In the above expressions, $\vartheta = \sqrt{\sum_{m=1}^{M_{\mathrm{S}}}|u_m|^2}$ and $\varepsilon = \sqrt{\sum_{m=1}^{M_{\mathrm{P}}}|u_m|^2}$, $\{u_m\}$ is i.i.d. $\mathscr{CN}(0,1)$. τ_{1n}^{S}, τ_{jn}^{P}, and the detailed proof can be found in Appendix A.

4.3 Problem Formulation

In CR networks, the CN must control its transmit power to avoid harmful interference to PUs. Therefore, we usually impose restriction to interference power of the CN and guarantee PUs' QoSs (e.g., an average interference power constraint [37]-[38] and a peak interference power constraint [39]-[40]). In this chapter, we consider the minimum required SINR of each PU as

the constraint as $\mathrm{SINR}_{jk} \geq \eta_{jk}$, where η_{jk} denotes the minimum required SINR of the k-th PU in the j-th cell. Therefore, the power optimization problem is formulated as

$$\max_{\mathbf{P}} \quad \mathscr{R}(\mathbf{P}) = \sum_{n=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}} R_{1n}(\mathbf{P}) \tag{4.10a}$$

$$\text{s.t.} \sum_{n=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}} P_{1n} \leq P_{\max}, \tag{4.10b}$$

$$\mathrm{SINR}_{jk} \geq \eta_{jk} \begin{cases} j \in \{1\}, \quad k \in \{1,\cdots,K_\mathrm{P}-K_\mathrm{T}\}, \\ j \in \{2,\cdots,L\}, k \in \{1,\cdots,K_\mathrm{P}\}, \end{cases} \tag{4.10c}$$

$$P_{1n} \geq 0 \qquad n \in \{K_\mathrm{P}-K_\mathrm{T}+1,\cdots,K_\mathrm{P}\}, \tag{4.10d}$$

where(4.10b) is the total power constraint of the CBS, and (4.10c) is used to guarantee the each PU's QoS. In fact, (4.10c) is function of $\mathbf{P}$ and thus problem (4.10) can be rewritten as

$$\max_{\mathbf{P}} \quad \mathscr{R}(\mathbf{P}) = \sum_{n=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}} R_{1n}(\mathbf{P}) \tag{4.11a}$$

$$\text{s.t.} \sum_{n=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}} P_{1n} \leq \min\{I_\mathrm{P}, P_{\max}\}, \tag{4.11b}$$

$$\sum_{\substack{n=K_\mathrm{P}-K_\mathrm{T}+1 \\ n\neq k}}^{K_\mathrm{P}} P_{1n} + \xi_{1k} P_{1k} \leq \mathscr{I}_{jk} \begin{cases} k \in \{K_\mathrm{P}-K_\mathrm{T}+1,\cdots,K_\mathrm{P}\}, \\ j \in \{2,\cdots,L\}, \end{cases} \tag{4.11c}$$

$$P_{1n} \geq 0 \qquad n \in \{K_\mathrm{P}-K_\mathrm{T}+1,\cdots,K_\mathrm{P}\}, \tag{4.11d}$$

with

$$\begin{cases} I_\mathrm{P} = \min\left\{[\mathscr{I}_{jk}]_{L\times(K_\mathrm{P}-K_\mathrm{T})}\right\}, \\ \mathscr{I}_{jk} = \frac{1}{\beta_{1jk}^{\mathrm{PS}}}\left(\frac{(1/\tau_{jk}^{\mathrm{P}})p_\mathrm{t}{\beta_{jjk}^{\mathrm{PP}}}^2\mathbb{E}^2\{\varepsilon\}}{\eta_{jk}} - I_1' - I_2' - I_3'\right), \\ \xi_{1k} = \left(\frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}}\mathbb{E}\{\vartheta^2\} + 1 - \frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}}\right). \end{cases}$$

I_1', I_2', I_3' and detailed proof can be found in Appendix B.

From the proof in Appendix B, we observe that the PUs' SINR increases with the number of PBS antennas and decreases with the transmit power for CUs. Therefore, there is an advantage to use (4.10c) as constraint. For example, when the number of PBS antennas increases, the CBS can transmit a higher power, which improves the downlink sum rate of the CN. However, this constraint brings more difficulties to design the power allocation strategy.

Since the power, $\mathbf{P}$, appears in the denominator of each CU's SINR, the objective function (4.11a) is not concave and (4.11) is a non-convex optimization problem, which is difficult to directly solve. Next, we propose an iterative algorithm to address the above problem.

4.4 The Solution of The Optimization Problem

In this section, we first transform the original non-convex optimization problem into a convex one by approximation. Then, we iteratively solve the approximated convex optimization problem. Finally, we prove that the obtained solution satisfies the necessary KKT conditions of the original problem (4.11).

4.4.1 Problem Transformation

First, we present the following lower bound [41]

$$a\ln x + b \leq \ln(1+x), \tag{4.12}$$

that is tight at a given $x = x_0$. The coefficients a and b are selected as

$$a = \frac{x_0}{1+x_0}, \qquad b = \ln(1+x_0) - \frac{x_0}{1+x_0}\ln x_0. \tag{4.13}$$

After some manipulations by using (4.12) and (4.13), we define the lower bound on the achievable downlink rate for the n-th CU as

$$\begin{aligned} &\tilde{R}_{1n}(\mathbf{P}, a_{1n}, b_{1n}) \\ &= \left(a_{1n} \ln \left(\frac{(1/\tau_{1n}^{\mathrm{S}}) P_{1n} {\beta_{11n}^{\mathrm{SS}}}^2 \mathbb{E}^2\{\vartheta\}}{I_1 + I_2 + I_3 + I_4 + \sigma^2} \right) + b_{1n} \right) \log_2 e, \end{aligned} \tag{4.14}$$

where a_{1n} and b_{1n} denote the coefficients given in (4.13). The coefficients are updated by replacing x_0 with P_{1n} in each iteration. Since (4.14) is still non-concave, we define $P_{1n} \triangleq e^{\tilde{P}_{1n}}$ and the lower bound of achievable ergodic downlink rate (4.14) can be transformed into (4.15). According to [41], ln-sum-e is convex and (4.15) is concave function of $\tilde{\mathbf{P}}$.

$$\begin{aligned}\tilde{R}_{1n}(e^{\tilde{\mathbf{P}}},a_{1n},b_{1n}) &= \left(a_{1n}\left(\ln\left(C_{1n}e^{\tilde{P}_{1n}}\right)-\ln\left(D_{1n}e^{\tilde{P}_{1n}}\beta_{11n}^{\mathrm{SS}}+\sum_{\substack{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ i\neq n}}^{K_{\mathrm{P}}}\beta_{11n}^{\mathrm{SS}}e^{\tilde{P}_{1i}}+\mathscr{N}_0\right)\right)+b_{1n}\right)\log_2 e\\ &= \left(a_{1n}\left(\ln(C_{1n})+\tilde{P}_{1n}-\ln\left(D_{1n}e^{\tilde{P}_{1n}}\beta_{11n}^{\mathrm{SS}}+\sum_{\substack{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ i\neq n}}^{K_{\mathrm{P}}}\beta_{11n}^{\mathrm{SS}}e^{\tilde{P}_{1i}}+\mathscr{N}_0\right)\right)+b_{1n}\right)\log_2 e.\end{aligned} \tag{4.15}$$

$$C_{1n}=\frac{{\beta_{11n}^{\mathrm{SS}}}^2\mathbb{E}^2\{\vartheta\}}{\tau_{1n}^{\mathrm{S}}}, D_{1n}=\left(\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}\mathrm{var}\{\vartheta\}+1-\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}\right), \mathscr{N}_0=I_2+I_3+I_4.$$

Next, we transform the original problem into the convex optimization problem as follows:

$$\max_{\tilde{\mathbf{P}}} \quad \tilde{\mathscr{R}}(e^{\tilde{\mathbf{P}}},\mathbf{a},\mathbf{b})=\sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}}\tilde{R}_{1n}(e^{\tilde{\mathbf{P}}},a_{1n},b_{1n}) \tag{4.15a}$$

$$\text{s.t.} \sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}}e^{\tilde{P}_{1n}}\leq \min\{I_{\mathrm{P}},P_{\max}\}, \tag{4.15b}$$

$$\sum_{\substack{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ n\neq k}}^{K_{\mathrm{P}}}e^{\tilde{P}_{1n}}+\xi_{1k}e^{\tilde{P}_{1k}}\leq \mathscr{I}_{jk}\begin{cases}k\in\{K_{\mathrm{P}}-K_{\mathrm{T}}+1,\cdots,K_{\mathrm{P}}\},\\ j\in\{2,\cdots,L\},\end{cases} \tag{4.15c}$$

$$e^{\tilde{P}_{1n}}\geq 0 \qquad n\in\{K_{\mathrm{P}}-K_{\mathrm{T}}+1,\cdots,K_{\mathrm{P}}\}, \tag{4.15d}$$

where $\mathbf{a}=[a_{1\{K_{\mathrm{P}}-K_{\mathrm{T}}+1\}},\cdots,a_{1K_{\mathrm{P}}}]$ and $\mathbf{b}=[b_{1\{K_{\mathrm{P}}-K_{\mathrm{T}}+1\}},\cdots,b_{1K_{\mathrm{P}}}]$.

Since (4.15) is a convex optimization problem, the duality gap is zero and solving its dual problem is equivalent to solve the original problem [42]. Therefore, we address its dual problem to obtain the solution of (4.15). Then the obtained power is transformed to the P-space by $P_{1n}\triangleq e^{\tilde{P}_{1n}}$. We note here that the optimal solution of the above problem is a lower bound of the original problem given by (4.11). Based on the following theorem, the efficient solution of problem (4.11) can be obtained by iteratively solving the above defined problem in (4.15).

Theorem 4.4.1 *The value of the objective function given by (4.11) will be either improved at the $(t+1)$-th iteration, or remained at the same value as the previous iteration, namely $\mathscr{R}\left(\mathbf{P}^{(t)}\right)\leq\mathscr{R}\left(\mathbf{P}^{(t+1)}\right)$.*

Proof We assume that $\mathbf{P}^{(t)}$ is optimal solution of problem (4.15) at the t-th iteration. Then, we have

$$\begin{aligned}\tilde{\mathscr{R}}\left(e^{\tilde{\mathbf{P}}^{(t)}},\,\mathbf{a}^{(t)},\mathbf{b}^{(t)}\right) &\overset{(a)}{\leq} \mathscr{R}\left(\mathbf{P}^{(t)}\right) \overset{(b)}{=} \tilde{\mathscr{R}}\left(e^{\tilde{\mathbf{P}}^{(t)}},\,(\mathbf{a})^{(t+1)},\mathbf{b}^{(t+1)}\right)\\ &\overset{(c)}{\leq} \tilde{\mathscr{R}}\left(e^{\tilde{\mathbf{P}}^{(t+1)}},\,(\mathbf{a})^{(t+1)},\mathbf{b}^{(t+1)}\right)\\ &\overset{(d)}{\leq} \mathscr{R}\left(\mathbf{P}^{(t+1)}\right).\end{aligned} \tag{4.16}$$

From (4.16), we observe that the optimal value of the objective function in (4.11) at the $(t+1)$-th iteration is larger than that at the t-th iteration. In fact, inequality (a) follows from (4.12); equality (b) holds because of the tight approximation at the current power $\{e^{\tilde{\mathbf{P}}^{(t)}},\,\mathbf{a}^{(t+1)},\mathbf{b}^{(t+1)}\}$; inequality (c) holds because $e^{\tilde{\mathbf{P}}^{(t+1)}}$ and $e^{\tilde{\mathbf{P}}^{(t)}}$, respectively, are the optimal and feasible solution for problem (4.15) at the $(t+1)$-th iteration; the last inequality directly comes from (4.12). ∎

4.4.2 Problem Solution

Next, we focus on solving the dual problem of (4.15). First, we define the following Lagrange dual function

$$g(\boldsymbol{\lambda},\boldsymbol{\mu}) = \max_{\tilde{\mathbf{P}}\in\boldsymbol{\Omega}} L(\tilde{\mathbf{P}},\boldsymbol{\lambda},\boldsymbol{\mu}), \tag{4.17}$$

where

$$\begin{aligned}L(\tilde{\mathbf{P}},\boldsymbol{\lambda},\boldsymbol{\mu}) = &\sum_{n=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}}\left(a_{1n}\left(\ln(C_{1n})+\tilde{P}_{1n}-\ln\left(D_{1n}e^{\tilde{P}_{1n}}\beta_{11n}^{\mathrm{SS}}+\sum_{\substack{i=K_\mathrm{P}-K_\mathrm{T}+1\\ i\neq n}}^{K_\mathrm{P}}\beta_{11n}^{\mathrm{SS}}e^{\tilde{P}_{1i}}+\mathscr{N}_0\right)\right)+b_{1n}\right)\log_2 e\\ &+\boldsymbol{\lambda}\left(\min\{I_\mathrm{P},P_{\max}\}-\sum_{n=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}}e^{\tilde{P}_{1n}}\right)+\sum_{j=2}^{L}\sum_{k=K_\mathrm{P}-K_\mathrm{T}+1}^{K_\mathrm{P}}\mu_{jk}\left(\mathscr{I}_{jk}-\left(\sum_{\substack{n=K_\mathrm{P}-K_\mathrm{T}+1\\ n\neq k}}^{K_\mathrm{P}}e^{\tilde{P}_{1n}}+\xi_{1k}e^{\tilde{P}_{1k}}\right)\right).\end{aligned} \tag{4.18}$$

$\boldsymbol{\Omega}$ denotes the feasible domain defined by (4.15b)–(4.15d), $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}=\{\mu_{jk}\}(j\in\{2,\ldots,L\},k\in\{K_\mathrm{P}-K_\mathrm{T}+1,\ldots,K_\mathrm{P}\})$ denote the value and vector of the dual variables associated with constraint conditions (4.15b) and (4.15c), respectively. Based on this, the dual optimization

problem is given as

$$\begin{aligned} &\min_{\lambda,\boldsymbol{\mu}} \quad g(\lambda,\boldsymbol{\mu}) \\ &\text{s.t.} \quad \lambda,\boldsymbol{\mu} \geq 0. \end{aligned} \tag{4.19}$$

Since a dual function is always convex [43], we minimize $g(\lambda,\boldsymbol{\mu})$ through subgradient method. The dual variables can be updated as follows:

$$\begin{aligned} \lambda^{(s+1)} &= \left[\lambda^{(s)} + \varsigma^{(s)}\left(\sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}} P_{1n} - \min\{I_{\mathrm{P}}, P_{\max}\}\right)\right]^{+}, \\ \mu_{jk}^{(s+1)} &= \left[\mu_{jk}^{(s)} + \varsigma_{jk}^{(s)}\left(\sum_{\substack{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1 \\ n\neq k}}^{K_{\mathrm{P}}} P_{1n} + \xi_{1k}P_{1k} - \mathscr{I}_{jk}\right)\right]^{+}, \end{aligned} \tag{4.20}$$

where s represents iteration index. $\varsigma^{(s)}$ and $\varsigma_{jk}^{(s)}$, respectively, represent the step sizes in the s-th iteration. The step size of dual variables is chosen based on the diminishing step size rule to guarantee convergence. Note that in (4.20), we have backed to the P-space.

However, solving the dual problem in (4.19) involves the optimal $\tilde{\mathbf{P}}$ for given dual variables λ and $\boldsymbol{\mu}$. We apply the KKT condition [43] and obtain the power as

$$P_{1n} = \frac{\sqrt{(\mathscr{P}_{1n}+\mathscr{N}_0)^2 + \frac{4D_{1n}\beta_{11n}^{\mathrm{SS}}a_{1n}(\mathscr{P}_{1n}+\mathscr{N}_0)\log_2 e}{\lambda+\mathscr{U}_{1n}}} - \mathscr{P}_{1n} - \mathscr{N}_0}{2D_{1n}\beta_{11n}^{\mathrm{SS}}}, \tag{4.21}$$

where $\mathscr{P}_{1n} = \sum\limits_{\substack{i=K_{\mathrm{P}}-K_{\mathrm{T}}+1 \\ i\neq n}}^{K_{\mathrm{P}}} \beta_{11n}^{\mathrm{SS}}P_{1i}$, $\mathscr{U}_{1n} = \sum_{j=2}^{L}\mu_{jn}\xi_{1n} + \sum_{j=2}^{L}\sum\limits_{\substack{k=K_{\mathrm{P}}-K_{\mathrm{T}}+1 \\ k\neq n}}^{K_{\mathrm{P}}}\mu_{jk}$.

Here, we find that (4.21) is a fixed point equation, namely $\mathbf{P}$ also appears in the right side of equation (4.21). According to [41], we can obtain the value of P_{1n} based on the fixed-point power update. Summarily, to solve our formulated original problem (4.11), we first initialize the parameters $\mathbf{a}$ and $\mathbf{b}$. Then, (4.11) can be transformed into a convex optimization problem (4.15) by approximation, which can be solved by classical Lagrange dual and subgradient methods. Next, we update $\mathbf{a}$ and $\mathbf{b}$ with obtained power according to (4.13) and resolve problem (4.15). The above process is repeated until the sum rate converges. We summarize the above method as Algorithm 2.

Next, we have the following theorem.

Theorem 4.4.2 *The solution obtained by iteratively solving the approximated problem in (*4.15*) satisfies the necessary KKT conditions of the original problem given by (*4.11*).*

Algorithm 2: Proposed Power Allocation Algorithm

Initialize $\mathbf{a}^{(t)} = \mathbf{1}, \mathbf{b}^{(t)} = \mathbf{0}$, the maximum tolerate ε, iteration index $t = 0$.
repeat
  Initialize dual variables λ and $\boldsymbol{\mu}$.
  repeat
    Obtain the power allocation P_{1n} via (4.21) and fixed-point power update [41].
    Update dual variables λ and $\boldsymbol{\mu}$ via (4.20).
  until Dual variables converge;
  Compute the sum rate $\tilde{\mathscr{R}}(\mathbf{P}, \mathbf{a}^{(t)}, \mathbf{b}^{(t)})$.
  Update $t = t + 1$.
  Update $\mathbf{a}^{(t)}$ and $\mathbf{b}^{(t)}$ via (4.13).
until $|\tilde{\mathscr{R}}(\mathbf{P}, \mathbf{a}^{(t+1)}, \mathbf{b}^{(t+1)}) - \tilde{\mathscr{R}}(\mathbf{P}, \mathbf{a}^{(t)}, \mathbf{b}^{(t)})| \leq \varepsilon$;

Proof First, we consider the following optimization problem:

$$\max_{\mathbf{X}} \ f_0(\mathbf{X}) \tag{4.22a}$$

$$\text{s.t.} \ f_i(\mathbf{X}) \leq 0, i = 1, 2, \cdots, M, \tag{4.22b}$$

where the objective function $f_0(\mathbf{X})$ and constraints $f_i(\mathbf{X})$ are assumed as nonconvex. Next, we select a convex function $f_i(\mathbf{X})$ so that $\tilde{f}_i(\mathbf{X}) \approx f_i(\mathbf{X})$ $(i = 0, 1, 2, \cdots, M)$. The approximate problem is a convex optimization problem, which can be solved by standard convex method. According to [44], the solution of approximate problem can converge to a point that satisfies the KKT conditions of the original problem, if the approximations satisfy the following conditions:

(1) $\tilde{f}_i(\mathbf{X}) \leq f_i(\mathbf{X})$ for any i.

(2) $\tilde{f}_i(\mathbf{X}_0) = f_i(\mathbf{X}_0)$, where $\mathbf{X}_0$ is the optimal solution of the approximate problem in the previous iteration.

(3) $\nabla \tilde{f}_i(\mathbf{X}_0) = \nabla f_i(\mathbf{X}_0)$ for any i, where ∇ means the derivation operation.

According to the defined parameters $\mathbf{a}$ and $\mathbf{b}$ in (4.12), (4.13) and (4.14), it is easy to verify that our proposed convex approximation satisfies $(1) - (3)$ simultaneously. Therefore, the proposed iterative algorithm would converge to the solution that satisfies the KKT conditions of the original problem. We finish the proof. Accordingly, the obtained solution at least reaches a local optimum of the original problem, which indicates that it has the potential to reach the global optimality.

4.5 Performance Analysis of the PN and CN with mMIMO

In this section, we assume that the PBS or CBS is equipped with very large-scale antenna arrays and then, we analyze the downlink rate of the PN and CN.

4.5.1 $M_{\mathrm{P}} \to \infty$ and M_{S} is fixed

The downlink rate of the CN

We find that all other terms in (4.8) are limited constants except for the interference term I_2. I_2 can be expressed as follows

$$\begin{aligned} I_2 &= \sum_{j=2}^{L} \beta_{j1n}^{\mathrm{SP}} p_{\mathrm{t}} \left(\frac{\beta_{j1n}^{\mathrm{SP}}}{\tau_{jn}^{\mathrm{P}}} \mathbb{E}\{\varepsilon^2\} + 1 - \frac{\beta_{j1n}^{\mathrm{SP}}}{\tau_{jn}^{\mathrm{P}}} \right) \\ &= \sum_{j=2}^{L} \beta_{j1n}^{\mathrm{SP}} p_{\mathrm{t}} \left(\frac{\beta_{j1n}^{\mathrm{SP}}}{\tau_{jn}^{\mathrm{P}}} M_{\mathrm{P}} + 1 - \frac{\beta_{j1n}^{\mathrm{SP}}}{\tau_{jn}^{\mathrm{P}}} \right). \end{aligned} \tag{4.23}$$

From (4.23), it can be observed that the interference term I_2 increases with M_{P}, and $I_2 \to \infty$ when $M_{\mathrm{P}} \to \infty$. Therefore, according to (4.8), the downlink rate of each CU will be close to zero when $M_{\mathrm{P}} \to \infty$.

The downlink rate of the PN

From Appendix B, we can get the SINR of the k-th PU in the j-th cell as follows

$$\mathrm{SINR}_{jk} = \frac{(1/\tau_{jk}^{\mathrm{P}}) p_{\mathrm{t}} {\beta_{jjk}^{\mathrm{PP}}}^2 \mathbb{E}^2\{\varepsilon\}}{I_1' + I_2' + I_3' + I_4' + \sigma^2}. \tag{4.24}$$

Before analyzing, two formulas are brought: duplication and Stirling's formulas.

$$\begin{aligned} &\Gamma(m)\Gamma(m+\frac{1}{2}) = 2^{(1-2m)}\sqrt{\pi}\Gamma(2m), \\ &\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n} n^n e^{(-n)}} = 1. \end{aligned} \tag{4.25}$$

According to (4.25), we have

$$\begin{aligned}
&\lim_{M_{\mathrm{P}}\to\infty}\frac{1}{\sqrt{M_{\mathrm{P}}}}\frac{\Gamma(M_{\mathrm{P}}+\frac{1}{2})}{\Gamma(M_{\mathrm{P}})}=\lim_{M_{\mathrm{P}}\to\infty}\sqrt{\frac{\pi}{M_{\mathrm{P}}}}2^{(1-2M_{\mathrm{P}})}\frac{(2M_{\mathrm{P}}-1)!}{(M_{\mathrm{P}}-1)!(M_{\mathrm{P}}-1)!}\\
&=\lim_{M_{\mathrm{P}}\to\infty}\sqrt{\frac{\pi}{M_{\mathrm{P}}}}2^{(1-2M_{\mathrm{P}})}\frac{\sqrt{2\pi(2M_{\mathrm{P}}-1)}(2M_{\mathrm{P}}-1)^{((2M_{\mathrm{P}}-1))}}{2\pi(M_{\mathrm{P}}-1)(M_{\mathrm{P}}-1)^{2(M_{\mathrm{P}}-1)}e}.\\
&=\lim_{M_{\mathrm{P}}\to\infty}\sqrt{\frac{2M_{\mathrm{P}}-1}{2M_{\mathrm{P}}}}\left(1+\frac{1}{2(M_{\mathrm{P}}-1)}\right)^{2M_{\mathrm{P}}-1}e^{-1}\\
&=ee^{-1}=1.
\end{aligned}\tag{4.26}$$

Thus, $\lim_{M_{\mathrm{P}}\to\infty}\frac{\mathbb{E}^2\{\varepsilon\}}{M_{\mathrm{P}}}=1$ and $\lim_{M_{\mathrm{P}}\to\infty}\frac{\mathrm{var}\{\varepsilon\}}{M_{\mathrm{P}}}=0$.

When $k\le K_{\mathrm{P}}-K_{\mathrm{T}}$, (4.24) can be approximated as follows

$$\begin{aligned}
\lim_{M_{\mathrm{P}}\to\infty}\mathrm{SINR}_{jk}&=\lim_{M_{\mathrm{P}}\to\infty}\frac{(1/\tau_{jk}^{\mathrm{P}})p_{\mathrm{t}}{\beta_{jjk}^{\mathrm{PP}}}^2\mathbb{E}^2\{\varepsilon\}}{I_1'+I_2'+I_3'+I_4'+\sigma^2}\\
&=\lim_{M_{\mathrm{P}}\to\infty}\frac{(1/\tau_{jk}^{\mathrm{P}})p_{\mathrm{t}}{\beta_{jjk}^{\mathrm{PP}}}^2}{\frac{I_1'+I_2'+I_3'+I_4'+\sigma^2}{M_{\mathrm{P}}}}=\frac{{\beta_{jjk}^{\mathrm{PP}}}^2}{\sum_{l\neq j}^{L}{\beta_{ljk}^{\mathrm{PP}}}^2}.
\end{aligned}\tag{4.27}$$

Accordingly, the downlink rate of each PU is only affected by pilot contamination from adjacent cells.

When $K_{\mathrm{P}}-K_{\mathrm{T}}+1\le k\le K_{\mathrm{P}}$, (4.24) can be approximated as follows

$$\lim_{M_{\mathrm{P}}\to\infty}\mathrm{SINR}_{jk}=\frac{{\beta_{jjk}^{\mathrm{PP}}}^2}{\sum_{\substack{l\neq j\\ l\neq 1}}^{L}{\beta_{ljk}^{\mathrm{PP}}}^2}.\tag{4.28}$$

In fact, (4.28) is similar to (4.27), and only difference is that the 1-st cell does not produce interference to PUs located in other cells .

4.5.2 $M_{\mathrm{S}} \to \infty$ and M_{P} is fixed

The downlink rate of the CN

We rewrite the SINR of the n-th CU as follows

$$\mathrm{SINR}_{1n} = \frac{(1/\tau_{1n}^{\mathrm{S}}) P_{1n} {\beta_{11n}^{\mathrm{SS}}}^2 \mathbb{E}^2\{\vartheta\}}{I_1 + I_2 + I_3 + I_4 + \sigma^2}. \tag{4.29}$$

From (4.29), we find that the denominator is a limited constant when M_{P} is fixed, so $\lim\limits_{M_{\mathrm{S}}\to\infty} \frac{I_1+I_2+I_3+I_4+\sigma^2}{M_{\mathrm{S}}} = 0$. According to (4.25) and (4.26), we have $\lim\limits_{M_{\mathrm{S}}\to\infty} \frac{\mathbb{E}^2\{\vartheta\}}{M_{\mathrm{S}}} = 1$. Therefore, the downlink rate of each CU goes to infinite when $M_{\mathrm{S}} \to \infty$.

The downlink rate of the PN

When $k \leq K_{\mathrm{P}} - K_{\mathrm{T}}$, we rewrite the SINR of the k-th PU in the j-th cell as follows

$$\mathrm{SINR}_{jk} = \frac{(1/\tau_{jk}^{\mathrm{P}}) p_{\mathrm{t}} {\beta_{jjk}^{\mathrm{PP}}}^2 \mathbb{E}^2\{\varepsilon\}}{I_1' + I_2' + I_3' + I_4' + \sigma^2}. \tag{4.30}$$

From Appendix B, we find that there is no relationship between (4.30) and M_{S}. Thus, the SINR of the PU is affected by M_{P}, transmit power for PUs and CUs, and pilot contamination. When $K_{\mathrm{P}} - K_{\mathrm{T}} + 1 \leq k \leq K_{\mathrm{P}}$, it is clear that all other terms in (4.24) are limited constants except for the interference I_4' that is represented as

$$\begin{aligned} I_4' &= \sum_{\substack{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1 \\ n\neq k}}^{K_{\mathrm{P}}} \beta_{1jk}^{PS} P_{1n} + \beta_{1jk}^{\mathrm{PS}} P_{1k} \left(\frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}} \mathbb{E}\{\vartheta^2\} + 1 - \frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}} \right) \\ &= \sum_{\substack{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1 \\ n\neq k}}^{K_{\mathrm{P}}} \beta_{1jk}^{PS} P_{1n} + \beta_{1jk}^{\mathrm{PS}} P_{1k} \left(\frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}} M_{\mathrm{S}} + 1 - \frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}} \right). \end{aligned} \tag{4.31}$$

It can be observed from (4.31) that the interference term I_4' increases with M_{S}, and $I_4' \to \infty$ when $M_{\mathrm{S}} \to \infty$. Therefore, the downlink rate of each PU will be close to zero when $M_{\mathrm{S}} \to \infty$.

4.5.3 $M_{\mathrm{S}} \to \infty$ and $M_{\mathrm{P}} \to \infty$

The downlink rate of the CN

According to (4.8), we have

$$\begin{aligned}\lim_{\substack{M_{\mathrm{S}}\to\infty\\ M_{\mathrm{P}}\to\infty}} \mathrm{SINR}_{1n} &= \lim_{\substack{M_{\mathrm{S}}\to\infty\\ M_{\mathrm{P}}\to\infty}} \frac{(1/\tau_{1n}^{\mathrm{S}})P_{1n}{\beta_{11n}^{\mathrm{SS}}}^2\mathbb{E}^2\{\vartheta\}}{I_1+I_2+I_3+I_4+\sigma^2} \\ &= \lim_{\substack{M_{\mathrm{S}}\to\infty\\ M_{\mathrm{P}}\to\infty}} \frac{(1/\tau_{1n}^{\mathrm{S}})P_{1n}{\beta_{11n}^{\mathrm{SS}}}^2}{\frac{I_1+I_2+I_3+I_4+\sigma^2}{M_{\mathrm{P}}}} = \frac{P_{1n}{\beta_{11n}^{\mathrm{SS}}}^2}{p_{\mathrm{t}}\sum_{j=2}^{L}{\beta_{j1n}^{\mathrm{PS}}}^2}.\end{aligned} \tag{4.32}$$

It is observed in (4.32) that the downlink rate of each CU is affected by two factors: transmit power for PUs and CUs, and pilot contamination.

The downlink rate of the PN

When $k \le K_{\mathrm{P}} - K_{\mathrm{T}}$, we have

$$\begin{aligned}\lim_{\substack{M_{\mathrm{P}}\to\infty\\ M_{\mathrm{S}}\to\infty}} \mathrm{SINR}_{jk} &= \lim_{\substack{M_{\mathrm{P}}\to\infty\\ M_{\mathrm{S}}\to\infty}} \frac{(1/\tau_{jk}^{\mathrm{P}})p_{\mathrm{t}}{\beta_{jjk}^{\mathrm{PP}}}^2\mathbb{E}^2\{\varepsilon\}}{I_1'+I_2'+I_3'+I_4'+\sigma^2} \\ &= \lim_{\substack{M_{\mathrm{P}}\to\infty\\ M_{\mathrm{S}}\to\infty}} \frac{(1/\tau_{jk}^{\mathrm{P}})p_{\mathrm{t}}{\beta_{jjk}^{\mathrm{PP}}}^2}{\frac{I_1'+I_2'+I_3'+I_4'+\sigma^2}{M_{\mathrm{P}}}} = \frac{{\beta_{jjk}^{\mathrm{PP}}}^2}{\sum_{l\neq j}^{L}{\beta_{ljk}^{\mathrm{PP}}}^2}.\end{aligned} \tag{4.33}$$

It can be found that (4.33) is the same with (4.24), and there is no interference from the CN.

When $K_{\mathrm{P}} - K_{\mathrm{T}} + 1 \le k \le K_{\mathrm{P}}$,

$$\lim_{\substack{M_{\mathrm{P}}\to\infty\\ M_{\mathrm{S}}\to\infty}} \mathrm{SINR}_{jk} = \frac{p_{\mathrm{t}}{\beta_{jjk}^{\mathrm{PP}}}^2}{p_{\mathrm{t}}\sum_{\substack{l\neq j\\ l\neq 1}}^{L}{\beta_{ljk}^{\mathrm{PP}}}^2 + P_{1k}{\beta_{1jk}^{\mathrm{PS}}}^2}. \tag{4.34}$$

It is clear that the downlink rate of the PU is affected by two factors: transmit power for PUs and CUs, and pilot contamination.

According to the above analysis, for a given transmit power for PUs and CUs, we observe that:

1) The downlink rate of all CUs will be affected seriously when the number of PBS antennas approaches to infinity.
2) The downlink rate of some PUs (PUs located in adjacent cells use the same pilots with CUs in the 1-st cell) is seriously affected when the number of CBS antennas approaches to infinity.
3) When the number of CBS and PBS antennas approaches to infinity simultaneously, the downlink rate of CUs and PUs will be affected by transmit power and pilot contamination.

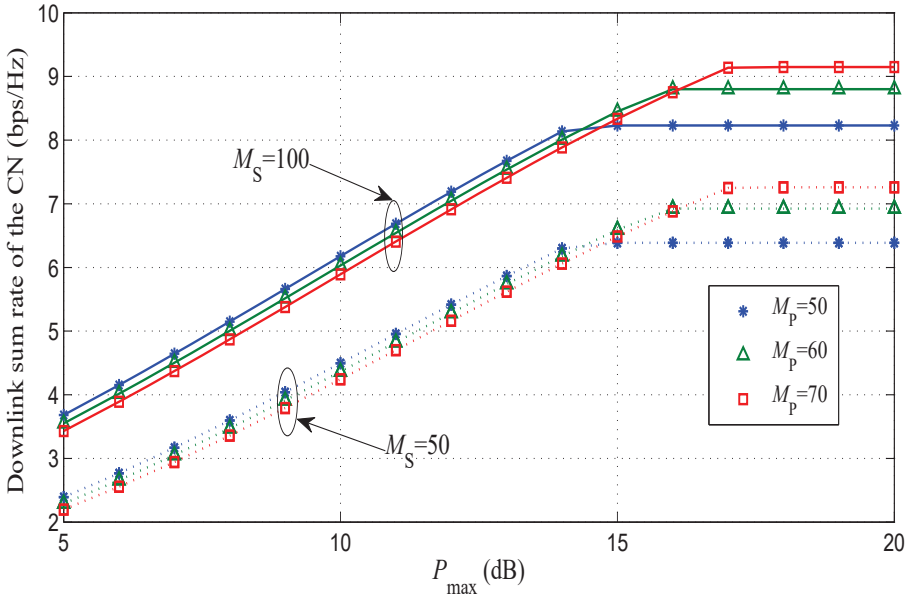
Thus, for fixed transmit power, the number of CBS antennas should not be great larger than that of PBS antennas. Otherwise, some PUs' QoSs may not be guaranteed according to (4.31). On the other hand, when the number of antennas at PBS is larger than that of CBS, the CN should not be allowed to access spectrum because the downlink rate of CUs will be very low according to (4.23). Therefore, for the PN and CN, to fully utilize the limited time-frequency resource, the number of PBS and CBS antennas should be close each other.

Note that the power control of the CN is not considered in above analysis. For a given transmit power of the PN and the number of PBS antennas, the CN has to decrease its transmit power to guarantee PUs' QoSs as the number of CBS antennas increases. Therefore, the increase of CBS antennas may not result in higher rate of the CN. On the other hand, for a given number of CBS antennas, the increase of PBS antennas degrades the rate of the CN, but this allows the CN to transmit higher power for CUs so that the rate of the CN may not decrease. The above discussion are presented in following section.

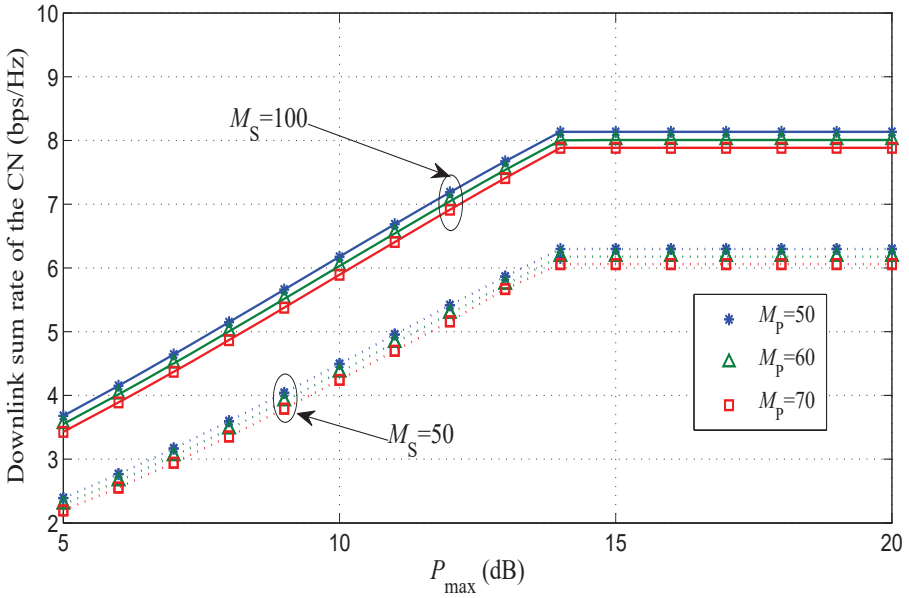
4.6 Numerical Results and Discussions

We consider a mMIMO-CR HetNet that consists of a PN ($L = 3$ PCs) and a CN (single CC). There are $K_{\mathrm{P}} = 4$ PUs in each PN cell and $K_{\mathrm{S}} = 4$ CUs in the SC. The total number of orthogonal pilots is 4 and PUs located in the same PN cell use orthogonal pilots and different PN cells reuse the same pilots. The same transmit power p_{t}=10 dB is assumed for all PUs. The cell radius (from center to vertex) is $r_c = 1000$ meters and the cell-hole radius is $r_h = 100$ meters (the users are not located in this area). We assume that one of the PBSs has the same geographic location with the CBS. The noise power is -174 dBm/Hz and the total bandwidth is 10 MHz. The pilot SNR is 5 dB. We assume that there are 2 inactive PUs in the PN cell which has the same geographic region with the CN cell and all PUs in other PN cells are active. The path-loss exponent is $\alpha = 3.8$.

Fig. 4.2(a) plots the downlink sum rate of the CN versus the maximum transmit power $P_{\max}$ for different number of antennas at PBS and CBS when $\eta = 8$ dB for all PUs. It can be

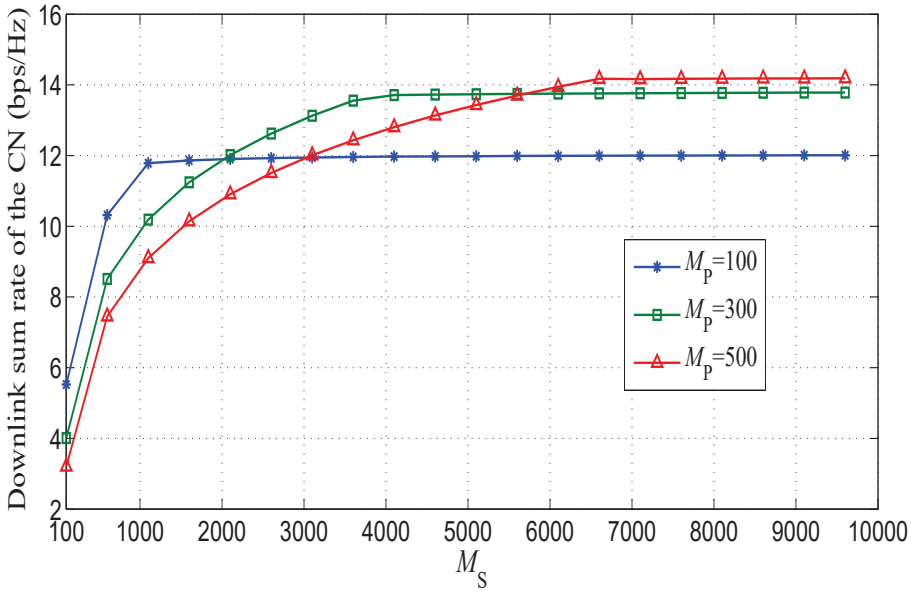


(a) SINR constraint

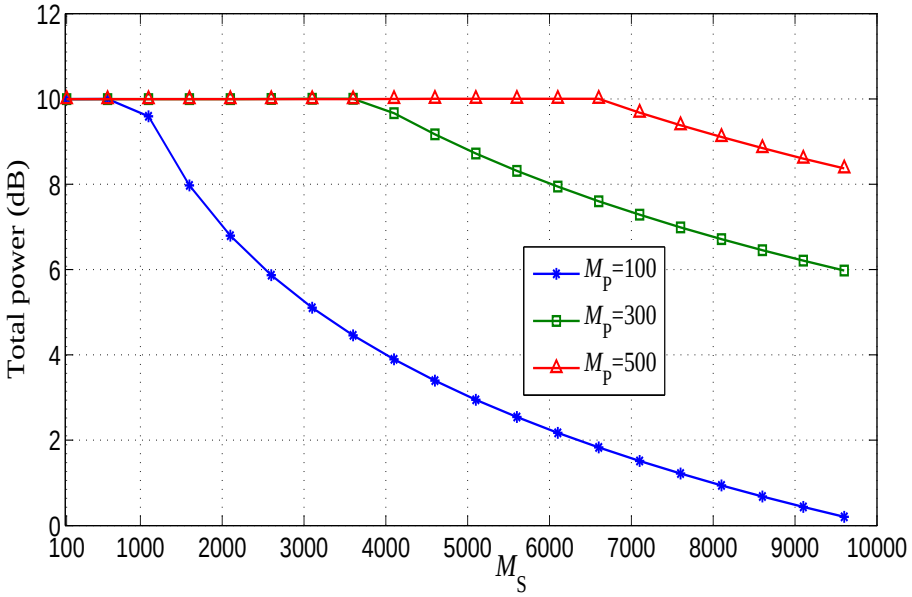


(b) Peak interference power constraint

Fig. 4.2 Downlink sum rate of the CN versus $P_{\max}$ with η=8 dB.



(a) Downlink sum rate



(b) Total transmit power

Fig. 4.3 Downlink sum rate and corresponding transmit power of the CN versus M_S with η=8 dB, $P_{max} = 10$ dB.

observed that the downlink sum rate of the CN first increases and then stabilizes with $P_{\max}$. This is because the CN is not allowed to transmit higher power due to the constraint of η. Before reaching stabilization, for the same M_{S}, the downlink sum rate of the CN is slightly higher when the PBS is equipped with less number of antennas. The reason is that all power is efficient utilized by the CN. On the contrary, after reaching stable state, for given M_{S}, the downlink sum rate of the CN is slightly lower when PBS is equipped with less number of antennas. This is because the PU's SINR increases with M_{P}, so more power is consumed by the CN for a given η, resulting in a higher downlink sum rate. We can also find that the downlink sum rate of the CN is higher when $M_{\mathrm{S}} = 100$ in comparison with $M_{\mathrm{S}} = 50$. It is clear that more antennas at CBS result in a higher rate.

We compare the downlink sum rate of the CN under different constraint conditions, i.e., we used SINR constraint and traditional peak interference power constraint [39]-[40]. We consider the following equivalent preprocessing: According to $\mathrm{SINR}_{jk} \geq \eta_{jk}$, we obtain the interference term caused by the CN as $I'_4 \leq \frac{1}{\eta}(1/\tau^{\mathrm{P}}_{jk}) p_{\mathrm{t}} {\beta^{\mathrm{PP}}_{jjk}}^2 \mathbb{E}^2\{\varepsilon\} - (I'_1 + I'_2 + I'_3 + \sigma^2)$. Then, we set $M_P = 50$ and define the interference power constraint as $I_{th} = \frac{1}{\eta}(1/\tau^{\mathrm{P}}_{jk}) p_{\mathrm{t}} {\beta^{\mathrm{PP}}_{jjk}}^2 \mathbb{E}^2\{\varepsilon\} - (I'_1 + I'_2 + I'_3 + \sigma^2)$ when $M_P = 50$, namely using $I'_4 \leq I_{th}$ replaces the constraint condition (4.11c). We plot the downlink sum rate of the CN versus the maximum transmit power $P_{\max}$ for different number of antennas at PBS and CBS when $\eta = 8$ dB for all PUs in Fig. 4.2(b). We find that the downlink sum rate of the CN is the same under these two schemes when $P_{\max} \leq 14$. However, the downlink sum rate of the CN with the SINR constraint is higher than that with traditional peak interference power constraint when $P_{\max} \geq 14$. In fact, for traditional peak interference power constraint, the number of PBS antennas does not influence the transmit power of the CN. On the contrary, the PU's SINR is related to the number of PBS antennas and more PBS antennas will lead to higher SINR. Thus, for a given η, the CN will be allowed to transmit higher power for a large number of PBS antennas, which improves the CU's rate. Thus, by using the SINR constraint, a higher rate of the CN is achieved for a high $P_{\max}$, which can be improved by about 10% in comparison with the conventional peak interference power constraint.

Fig. 4.3 shows the downlink sum rate and the corresponding total transmit power of the CN versus M_{S}, respectively. We set $\eta = 8$ dB for all PUs and $P_{\max} = 10$ dB. From Fig. 4.3(a), we observe that the downlink sum rate of the CN first increases until the stabilization is reached. The similar reasons have been explained in Fig. 4.2(b). On one hand, more antennas at CBS lead to higher downlink rate. On the contrary, more antennas at CBS will cause more interference to the PN. To guarantee PUs' QoSs, the CN has to decrease the transmit power as shown in Fig. 4.3(a). Therefore, the total transmit power of the CN decreases with M_{S}, but the downlink sum rate of the CN does not decease thanks to the increase of CBS antennas.

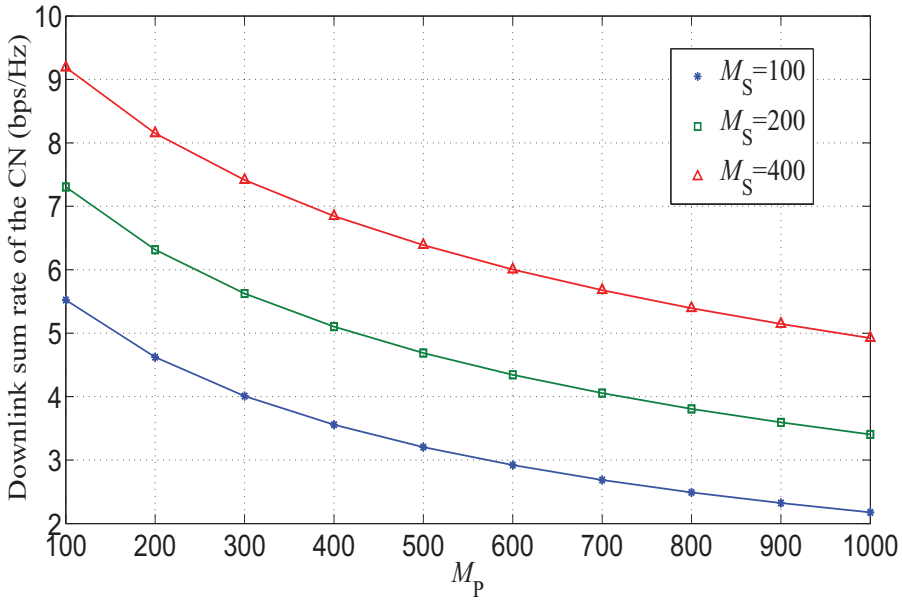


Fig. 4.4 Downlink sum rate of the CN versus M_P with η=8 dB, $P_{max} = 10$ dB.

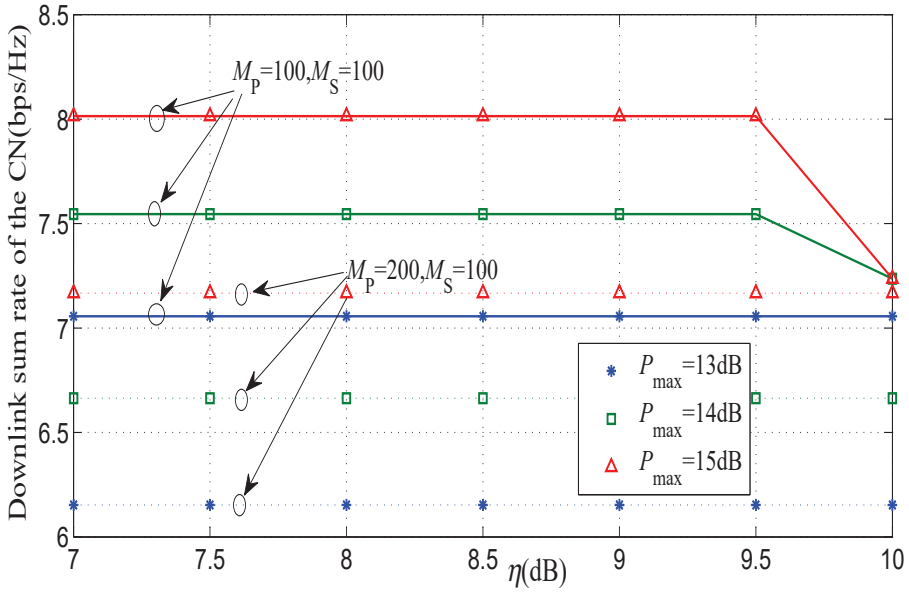


Fig. 4.5 Downlink sum rate of the CN versus η under different P_{max}.

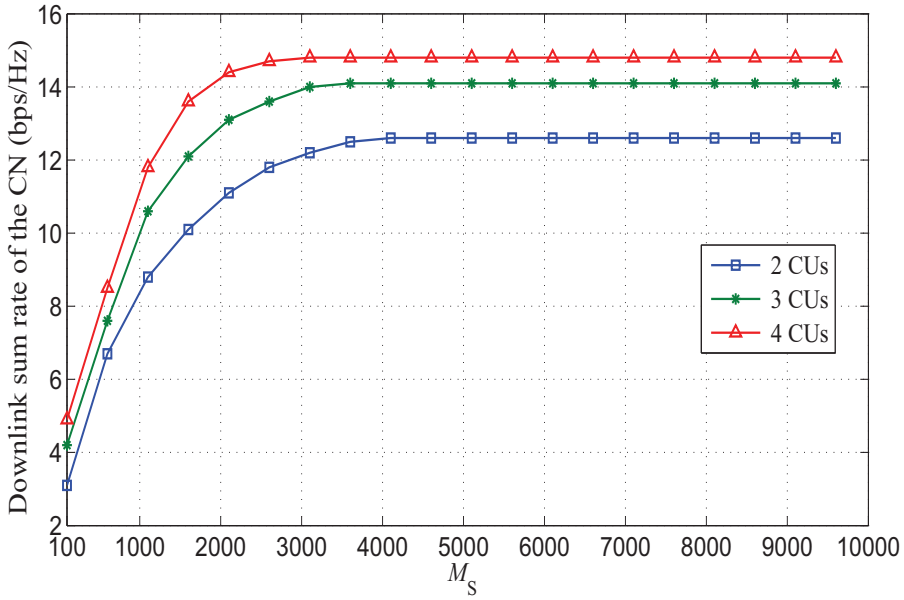


Fig. 4.6 Downlink sum rate of the CN versus M_S with η=8 dB, $P_{max} = 10$ dB under different number of active CUs.

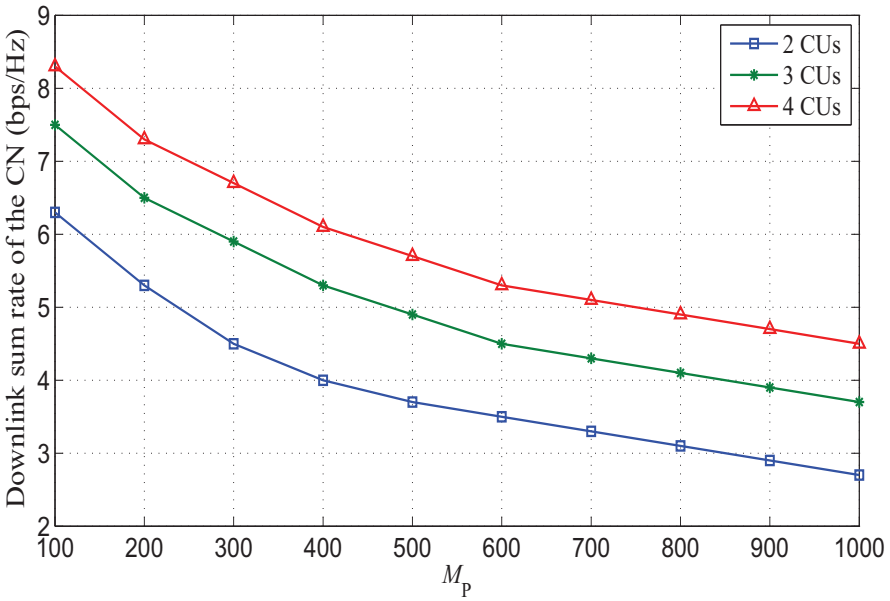


Fig. 4.7 Downlink sum rate of the CN versus M_P with η=8 dB, $P_{max} = 10$ dB under different number of active CUs.

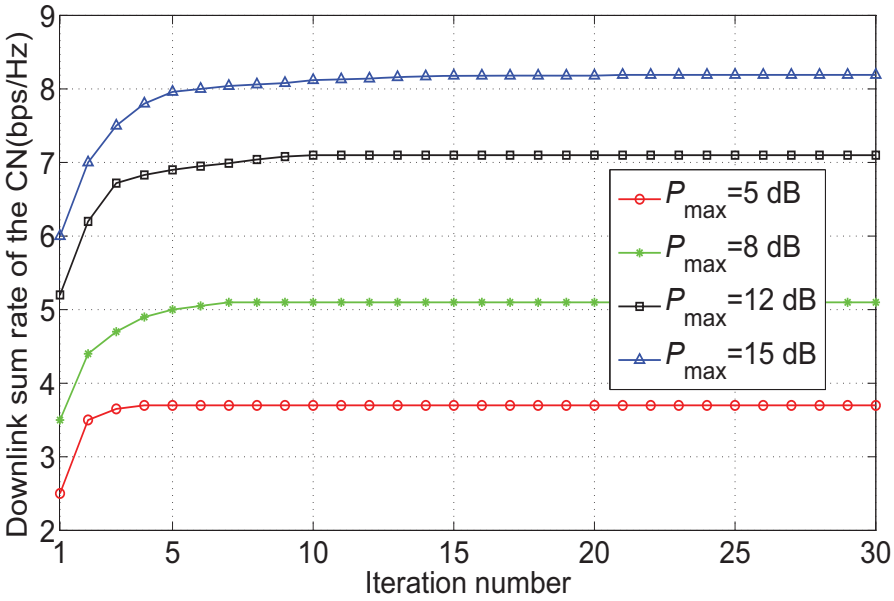


Fig. 4.8 Downlink sum rate of the CN versus iteration number.

Fig. 4.4 shows the downlink sum rate of the CN versus M_{P} for different M_{S} with η=8 dB for all PUs. For a given M_{S}, we find that the downlink sum rate of the CN decreases with M_{P}. The impact of η on the downlink sum rate of the CN under different M_{P} when M_{S}=100 is shown in Fig. 4.5. For a larger M_{P} (e.g., M_{P}=200), the PU's SINR is higher and (4.11c) is a loose constraint for given η (7 dB $\sim$ 10 dB) when $P_{\max}$=13 dB, 14 dB and 15 dB. Therefore, the downlink sum rate of the CN is a constant for given total transmit power. However, for a smaller M_{P} (e.g. M_{P}=100), the PU's SINR is lower. To guarantee PUs' QoSs, the CN is not allowed to transmit higher power for a larger η. Therefore, the downlink sum rate of the CN will decrease.

In Figs. 4.6 and 4.7, we assume 8 PUs in each PC and different number of active CUs in SC. Fig. 4.6 shows that the downlink sum rate of the CN versus M_{S} when $M_{\mathrm{P}} = 300$. The legend "3 CUs" denotes that 3 CUs share the primary pilots. Similarly to Fig. 4.3(a), the sum rate first increases and then stabilizes with M_{S}. Meanwhile, we find that the sum rate increases when more CUs are allowed to share the primary pilots, but the increased ratio decreases due to the fixed total transmit power. In Fig. 4.7, $M_{\mathrm{S}} = 200$ and the SINR constraint is considered. Similarly to Fig. 4.4, the sum rate decreases with M_{P} and increases with the number of active CUs.

Fig. 4.8 shows the downlink sum rate of the CN versus iteration number. We set $M_{\mathrm{S}} = 100, M_{\mathrm{P}} = 50, \eta = 8$ dB. Here, we compare the convergence speed under different transmit power $P_{\max}$. For a low $P_{\max}$, we find that the sum rate speedily converges. For example,

4 iterations are needed to guarantee convergence for $P_{\max} = 5$ dB. As $P_{\max}$ increases, the convergence becomes slower. For example, about 15 iterations are needed when $P_{\max} = 15$ dB. Thus, $P_{\max}$ affects the performance of the convergence.

4.7 Conclusions

In this chapter, we have studied the power allocation problem for the mMIMO-CR HetNet with pilot contamination. Following this, we formulated the power allocation optimization problem to maximize the downlink sum rate of the CN. To effectively protect PUs from harmful interference, we adopted PUs' SINR constraint or not traditional interference power constraint. Then, a convex approximation-based iterative method was proposed to solve the formulated problem, where the obtained solution can be guaranteed to satisfy KKT points of the original problem. We analyzed the performance of the PN and CN when the number of PBS or CBS antennas was assumed huge. The results show that the sum rate of the CN can be improved by about 10% by using the proposed SINR constraint in comparison with the conventional peak interference power constraint.

Chapter 5

Pilot Allocation for Small Cell Type Heterogeneous Network with Massive MIMO Marco Cell

5.1 Introduction

In chapters 3 and 4, we consider that the mMIMO and CR coexist to form a mMIMO-CR HetNet, including PN and CN. In this chapter, we will consider the pilot allocation problem in mMIMO-SC HetNet. It is well known that TDD is considered in mMIMO-SC HetNet [17] [45, 46]. However, due to short coherence time, the number of available orthogonal pilots for channel estimation is limited. In the case that orthogonal pilots are provided for all SUs and MUs, an excessive pilot overhead is generated and data transmission efficiency degrades. When the coverage regions of any two SCs are non-overlapping and their distance is relatively large the same pilot may be reused by users in adjacent SCs [46], [17]. However, even in such cases, the inter-tier interference from MBS to SUs in downlink still occurs and seriously affects the achievable rate in SCs. Thus, to mitigate the inter-tier interference, it is necessary to estimate the CSI of the links between MBS and SUs while minimizing the uplink pilot overhead.

In this chapter, we propose a new pilot allocation scheme for the two-tier TDD mMIMO-SC HetNet, where a part of SCs is allowed to use the orthogonal pilots. We consider the uplink pilot overhead and inter-tier interference coordination while maximizing the ergodic downlink sum rate of MUs and SUs. Unlike the previous works in [17] and [46], the proposed scheme enables the MBS to estimate not only CSIs of MBS-MUs links but also those of

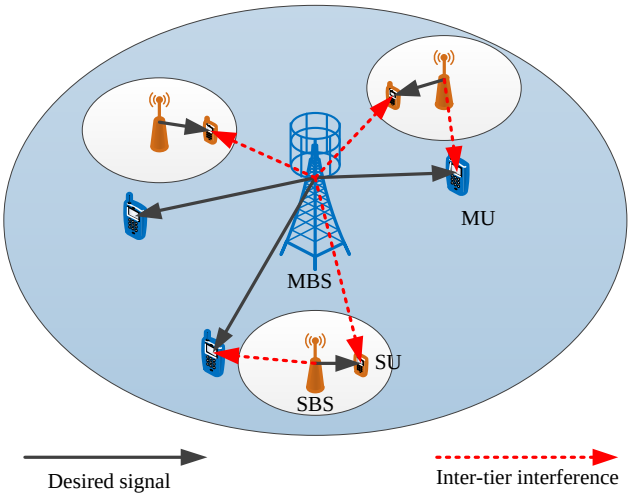


Fig. 5.1 The mMIMO-SC HetNet model.

MBS-SUs links. Consequently, the inter-tier interference from MBS to SUs can be mitigated by a downlink beam-forming such as ZF technique.

5.2 System Model

We consider a downlink two-tier mMIMO-SC HetNet system as shown in Fig. 5.1, where K SCs share the same time-frequency resource with an MC. In each SC, SBS equipped with N_S antenna serves one single-antenna SU one time. As for MC, there is a N_{M}-antenna MBS and $M(N_{\mathrm{M}} \gg M)$ single-antenna MUs are served by the MBS simultaneously. We assume that the two-tier mMIMO-SC HetNet system operates in TDD mode, and the uplink pilot and downlink data transmission are completely synchronous in both tiers. For simplicity, we denote the SU in SC k by the SU k.

The received signal of the MU m can be expressed as:

$$
\begin{aligned}
y_m^{\mathrm{M}} &= \sum_{i=1}^{M} \sqrt{P_{\mathrm{M}} \beta_{0,m}^{\mathrm{M,M}}} \mathbf{h}_{0,m}^{\mathrm{M,M}} \mathbf{w}_i x_i + \sum_{k=1}^{K} \sqrt{P_{\mathrm{S}} \beta_{k,m}^{\mathrm{S,M}}} \mathbf{h}_{k,m}^{\mathrm{S,M}} \mathbf{v}_k s_k + n_m^{\mathrm{M}} \\
&= \underbrace{\sqrt{P_{\mathrm{M}} \beta_{0,m}^{\mathrm{M,M}}} \mathbf{h}_{0,m}^{\mathrm{M,M}} \mathbf{w}_m x_m}_{\text{Desired signal}} + \underbrace{\sum_{i=1, i \neq m}^{M} \sqrt{P_{\mathrm{M}} \beta_{0,m}^{\mathrm{M,M}}} \mathbf{h}_{0,m}^{\mathrm{M,M}} \mathbf{w}_i x_i}_{\text{Intra-tier interference}} \\
&\quad + \underbrace{\sum_{k=1}^{K} \sqrt{P_{\mathrm{S}} \beta_{k,m}^{\mathrm{S,M}}} \mathbf{h}_{k,m}^{\mathrm{S,M}} \mathbf{v}_k s_k}_{\text{Inter-tier interference}} + \underbrace{n_m^{\mathrm{M}}}_{\text{Noise}},
\end{aligned}
\tag{5.1}
$$

where P_{M} and $\mathbf{w}_i$ denote the transmit power and $N_{\mathrm{M}} \times 1$ precoding vector for MU i at the MBS, respectively. We assume the same transmit power for each MU. P_{S} and $\mathbf{v}_k$, respectively, represent the transmit power and $N_{\mathrm{S}} \times 1$ precoding vector at SBS k and P_{S} is assumed the same for each SBS. The x_i and s_i are the transmit data for MU i and SU i, respectively, while n_m^{M} denotes the i.i.d. AWGN defined as $CN(0, \delta^2)$. $\beta_{a,b}^{\mathrm{A,B}}$ and $\mathbf{h}_{a,b}^{\mathrm{A,B}}$ ($a, b \in \{0, 1, \ldots, \max\{M, K\}\}$, A, B $\in \{\mathrm{M}, \mathrm{S}\}$) are the large-scale fading coefficient and small-scale fading vector between MBS ($a = 0, \mathrm{A} = \mathrm{M}$) or SBS a ($a \neq 0, \mathrm{A} = \mathrm{S}$) and MU b (B=M) or SU b (B=S), where $\mathbf{h}_{0,b}^{\mathrm{M,B}} \in C^{1 \times N_{\mathrm{M}}}$, $\mathbf{h}_{a,b}^{\mathrm{S,B}} \in C^{1 \times N_{\mathrm{S}}}$, and $\mathbf{h}_{0,b}^{\mathrm{M,B}} \sim CN(\mathbf{0}_{N_{\mathrm{M}}}, \mathrm{I}_{N_{\mathrm{M}}})$, $\mathbf{h}_{a,b}^{\mathrm{S,B}} \sim CN(\mathbf{0}_{N_{\mathrm{S}}}, \mathrm{I}_{N_{\mathrm{S}}})$.

Similarly, the received signal at SU k can be written as:

$$
\begin{aligned}
y_k^{\mathrm{S}} &= \sum_{i=1}^{K} \sqrt{P_{\mathrm{S}} \beta_{i,k}^{\mathrm{S,S}}} \mathbf{h}_{i,k}^{\mathrm{S,S}} \mathbf{v}_i s_i + \sum_{m=1}^{M} \sqrt{P_{\mathrm{M}} \beta_{0,k}^{\mathrm{M,S}}} \mathbf{h}_{0,k}^{\mathrm{M,S}} \mathbf{w}_m x_m + n_k^{\mathrm{S}} \\
&= \underbrace{\sqrt{P_{\mathrm{S}} \beta_{k,k}^{\mathrm{S,S}}} \mathbf{h}_{k,k}^{\mathrm{S,S}} \mathbf{v}_k s_k}_{\text{Desired signal}} + \underbrace{\sum_{i=1, i \neq k}^{K} \sqrt{P_{\mathrm{S}} \beta_{i,k}^{\mathrm{S,S}}} \mathbf{h}_{i,k}^{\mathrm{S,S}} \mathbf{v}_i s_i}_{\text{Intra-tier interference}} \\
&\quad + \underbrace{\sum_{m=1}^{M} \sqrt{P_{\mathrm{M}} \beta_{0,k}^{\mathrm{M,S}}} \mathbf{h}_{0,k}^{\mathrm{M,S}} \mathbf{w}_m x_m}_{\text{Inter-tier interference}} + \underbrace{n_k^{\mathrm{S}}}_{\text{Noise}},
\end{aligned}
\tag{5.2}
$$

where n_k^{S} denotes the AWGN.

5.3 Problem Formulation and Solution

In this section, we first derive the lower bound on ergodic downlink rate of the MU and SU. Then, we formulate the pilot allocation problem for maximizing the ergodic downlink sum rate. Finally, pilot allocation algorithms are proposed.

5.3.1 Ergodic Downlink Rate of The MU and SU

Firstly, we assume that the coverage regions of any two SCs are non-overlapping and the distance between different SCs is large enough (to neglect their pilot interference), so that the same pilot is reused in all SCs. At the same time, the MC uses another set of pilots that are mutually orthogonal to SC pilot. Then, we can obtain the ergodic downlink rate of MU m and SU k as follows:

$$R_m^{\mathrm{M}} = \left(1 - \frac{1+M}{S}\right) \mathbb{E}\left\{\log_2(1+\mathrm{SINR}_m^{\mathrm{M}})\right\}, \tag{5.3}$$

$$R_k^{\mathrm{S}} = \left(1 - \frac{1+M}{S}\right) \mathbb{E}\left\{\log_2(1+\mathrm{SINR}_k^{\mathrm{S}})\right\}, \tag{5.4}$$

where “1" and M denote that one orthogonal pilot symbol is shared by all SCs and M orthogonal pilots symbols are uniquely allocated to M MUs, respectively. S denotes the total transmission symbols per frame with symbol duration T and time duration per frame ST.

$$\mathrm{SINR}_m^{\mathrm{M}} = \frac{P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\mathbf{w}_m\right|^2}{\sum\limits_{i=1,i\neq m}^{M} P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\mathbf{w}_i\right|^2 + \sum\limits_{k=1}^{K} P_{\mathrm{S}}\beta_{k,m}^{\mathrm{S,M}}\left|\mathbf{h}_{k,m}^{\mathrm{S,M}}\mathbf{v}_k\right|^2 + \delta^2}, \tag{5.5}$$

$$\mathrm{SINR}_k^{\mathrm{S}} = \frac{P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}\left|\mathbf{h}_{k,k}^{\mathrm{S,S}}\mathbf{v}_k\right|^2}{\sum\limits_{i=1,i\neq k}^{K} P_{\mathrm{S}}\beta_{i,k}^{\mathrm{S,S}}\left|\mathbf{h}_{i,k}^{\mathrm{S,S}}\mathbf{v}_i\right|^2 + \sum\limits_{m=1}^{M} P_{\mathrm{M}}\beta_{0,k}^{\mathrm{M,S}}\left|\mathbf{h}_{0,k}^{\mathrm{M,S}}\mathbf{w}_m\right|^2 + \delta^2}. \tag{5.6}$$

For MUs, we apply ZF precoding scheme at MBS. Firstly, we define $\mathbf{H} = [(\mathbf{h}_{0,1}^{\mathrm{M,M}})^T, (\mathbf{h}_{0,2}^{\mathrm{M,M}})^T, \ldots, (\mathbf{h}_{0,M}^{\mathrm{M,M}})^T]^T$, and then we have:

$$\overline{\mathbf{W}} = \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}. \tag{5.7}$$

Therefore, the precoding vector $\mathbf{w}_m$ can be defined as $\mathbf{w}_m = \overline{\mathbf{w}}_m/\|\overline{\mathbf{w}}_m\|$, where $\overline{\mathbf{w}}_m$ is the m-th column vector of $\overline{\mathbf{W}}$. For SUs, since only one SU is served one time in each SC, we apply MF precoding scheme at SBS as follows: $\mathbf{v}_k = \mathbf{h}_{k,k}^{\mathrm{S,S}H}/\|\mathbf{h}_{k,k}^{\mathrm{S,S}H}\|$.

Then, we have the following proposition:

Proposition 5.3.1 *With perfect CSI, the downlink achievable rate for MU m and SU k can be lower bounded as follows:*

$$\underline{R}_m^{\mathrm{M}} = \left(1 - \frac{1+M}{S}\right) \log_2 \left(1 + \frac{P_{\mathrm{M}} \beta_{0,m}^{\mathrm{M,M}} (N_{\mathrm{M}} - M)}{\sum_{k=1}^{K} P_{\mathrm{S}} \beta_{k,m}^{\mathrm{S,M}} + \delta^2}\right). \tag{5.8}$$

$$\underline{R}_k^{\mathrm{S}} = \left(1 - \frac{1+M}{S}\right) \log_2 \left(1 + \frac{P_{\mathrm{S}} \beta_{k,k}^{\mathrm{S,S}} (N_{\mathrm{S}} - 1)}{\sum_{m=1}^{M} P_{\mathrm{M}} \beta_{0,k}^{\mathrm{M,S}} + \sum_{i=1, i\neq k}^{K} P_{\mathrm{S}} \beta_{i,k}^{\mathrm{S,S}} + \delta^2}\right). \tag{5.9}$$

The detailed proof can be found in Appendix C.

Based on the above analysis, it is clear that SUs seriously suffer from inter-tier interference from MBS. To suppress inter-tier interference from MBS to SUs in downlink, MBS needs to know CSIs of MBS-SUs links as well as CSIs with respect to MUs in own cell. However, it is not practical solution to allocate orthogonal pilots to all MU and all SUs. To obtain the optimal performance of the system, we propose a new pilot allocation scheme that allows a part of SUs to use orthogonal pilots. The detailed description is presented below.

5.3.2 Problem Formulation

We assume that there are K_{O} SUs using the orthogonal pilots, where K_{O} is a parameter needed to be optimized ($0 \leq K_{\mathrm{O}} \leq K$) for maximizing the ergodic downlink sum rate. We denote K SUs as $\mathscr{K} = \{1, 2, \cdots, K\}$, K_{O} SUs as $\mathscr{K}_{\mathrm{O}} = \{\theta_1, \theta_2, \cdots, \theta_{K_{\mathrm{O}}}\}$ and other $K - K_{\mathrm{O}}$ SUs as $\mathscr{K}_{\mathrm{S}} = \{\vartheta_1, \vartheta_2, \cdots, \vartheta_{K_{\mathrm{S}}}\}$, where $\mathscr{K}_{\mathrm{O}} \bigcup \mathscr{K}_{\mathrm{S}} = \mathscr{K}$, $\mathscr{K}_{\mathrm{O}} \bigcap \mathscr{K}_{\mathrm{S}} =$, $\mathscr{K}_{\mathrm{O}} =$ for $K_{\mathrm{O}} = 0$ and $\mathscr{K}_{\mathrm{S}} =$ for $K_{\mathrm{S}} = 0$. Then, the MBS can obtain $\mathscr{K}_{\mathrm{O}}$ SUs' CSIs, and the ZF precoding can be used at MBS to cancel the interference from MBS to these $\mathscr{K}_{\mathrm{O}}$ SUs as follows:

$$\overline{\mathbf{W}}_{\mathrm{IC}} = \widetilde{\mathbf{H}}^H \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H\right)^{-1}, \tag{5.10}$$

where $\widetilde{\mathbf{H}} = \left[(\mathbf{h}_{0,1}^{\mathrm{M,M}})^T, \ldots, (\mathbf{h}_{0,M}^{\mathrm{M,M}})^T, (\mathbf{h}_{0,\theta_1}^{\mathrm{M,S}})^T, \ldots, (\mathbf{h}_{0,\theta_{K_{\mathrm{O}}}}^{\mathrm{M,S}})^T\right]^T$, and $\mathbf{h}_{0,\theta_k}^{\mathrm{M,S}}$ denotes the small-scale fading vector between SU θ_k and MBS.

Proposition 5.3.2 *Ergodic downlink rate of MU m and SU θ_k can be obtained as follows:*

$$\underline{\hat{R}}_m^{\mathrm{M}} = \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right) \log_2 \left(1 + \frac{P_{\mathrm{M}} \beta_{0,m}^{\mathrm{M,M}} (N_{\mathrm{M}} - M - K_{\mathrm{O}})}{\sum\limits_{k=1}^{K} P_{\mathrm{S}} \beta_{k,m}^{\mathrm{S,M}} + \delta^2}\right), \tag{5.11}$$

$$\underline{R}_{\theta_k}^{\mathrm{S}} = \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right) \log_2 \left(1 + \frac{P_{\mathrm{S}} \beta_{\theta_k,\theta_k}^{\mathrm{S,S}} (N_{\mathrm{S}} - 1)}{\sum\limits_{i=1, i\neq\theta_k}^{K} P_{\mathrm{S}} \beta_{i,\theta_k}^{\mathrm{S,S}} + \delta^2}\right), \tag{5.12}$$

where $\xi = \mathrm{sgn}(K - K_{\mathrm{O}})$ and $\mathrm{sgn}(x) = 1$ for $x > 0$ and $\mathrm{sgn}(x) = 0$ for $x = 0$. Similary, the ergodic downlink rate of SU ϑ_k can be expressed by

$$\underline{R}_{\vartheta_k}^{\mathrm{S}} = \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right) \log_2 \left(1 + \frac{P_{\mathrm{S}} \beta_{\vartheta_k,\vartheta_k}^{\mathrm{S,S}} (N_{\mathrm{S}} - 1)}{\sum\limits_{m=1}^{M} P_{\mathrm{M}} \beta_{0,\vartheta_k}^{\mathrm{M,S}} + \sum\limits_{\substack{i=1 \\ i\neq\vartheta_k}}^{K} P_{\mathrm{S}} \beta_{i,\vartheta_k}^{\mathrm{S,S}} + \delta^2}\right). \tag{5.13}$$

The proof of (5.11)-(5.13) is similar to the proof for (5.8) and (5.9) (see Appendix C) and it is omitted. Then, we formulate the following optimization pilot allocation (i.e., $\mathscr{K}_{\mathrm{O}}$) problem

$$\begin{aligned} \max_{\mathscr{K}_{\mathrm{O}}} \quad & \sum_{m=1}^{M} \underline{\hat{R}}_m^{\mathrm{M}} + \sum_{\theta_k \in \mathscr{K}_{\mathrm{O}}} \underline{R}_{\theta_k}^{\mathrm{S}} + \sum_{\vartheta_k \in \mathscr{K}_{\mathrm{S}}} \underline{R}_{\vartheta_k}^{\mathrm{S}} \\ s.t. \quad & K_{\mathrm{O}} \leq K, \\ & \mathscr{K}_{\mathrm{O}} \cup \mathscr{K}_{\mathrm{S}} = \mathscr{K}, \\ & \mathscr{K}_{\mathrm{O}} \cap \mathscr{K}_{\mathrm{S}} = O. \end{aligned} \tag{5.14}$$

Here, we assume $K + M \leq N_{\mathrm{M}}$, and we need to find the optimal $\mathscr{K}_{\mathrm{O}}$ for maximizing ergodic downlink sum rate. The exhaustive search can be used to find the optimal $\mathscr{K}_{\mathrm{O}}$, but the computational complexity is high as $\sum_{i=0}^{K} C_K^i = \sum_{i=0}^{K} \frac{K!}{i!(K-i)!}$ and for a large number of SUs it becomes practically infeasible.

Algorithm 3: Optimal SR-M algorithm

Initialize $\mathscr{R}_1 = \left\{r_1^1, r_2^1, \cdots, r_K^1\right\}$, $\mathscr{R}_2 = \left\{r_1^2, r_2^2, \cdots, r_K^2\right\}$ according to (5.12) and (5.13), $\Delta\mathscr{R} = \left\{\Delta r_{\gamma_1}, \Delta r_{\gamma_2}, \cdots, \Delta r_{\gamma_K}\right\}$
for $K_{\mathrm{O}} = 0 : K$ **do**
  $\mathscr{K}_{\mathrm{O}}(K_{\mathrm{O}}) = \{\gamma_0, \gamma_1, \gamma_2, \cdots, \gamma_{K_{\mathrm{O}}}\}$, where $\gamma_0 = \{\emptyset\}$
  Computing ergodic downlink sum rate $F(K_{\mathrm{O}})$ according to (5.16)
end for
$K_{\mathrm{O}}^* = \operatorname{argmax} F(K_{\mathrm{O}})$
$\mathscr{K}_{\mathrm{O}}^{\mathrm{optimal}}(K_{\mathrm{O}}^*) = \{\gamma_0, \gamma_1, \gamma_2, \cdots, \gamma_{K_{\mathrm{O}}^*}\}$

5.3.3 Problem Solution

Optimal Sum-Rate Maximization (SR-M) Algorithm

To reduce the computational complexity, we first transform the original problem as:

$$\begin{aligned} F(K_{\mathrm{O}}) &\triangleq \max_{\mathscr{K}_{\mathrm{O}}} \sum_{m=1}^{M} \underline{\hat{R}}_m^{\mathrm{M}} + \sum_{\theta_k \in \mathscr{K}_{\mathrm{O}}} \underline{R}_{\theta_k}^{\mathrm{S}} + \sum_{\vartheta_k \in \mathscr{K}_{\mathrm{S}}} \underline{R}_{\vartheta_k}^{\mathrm{S}} \\ &= \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right) \left(\sum_{m=1}^{M} \underline{\hat{r}}_m^{\mathrm{M}} + \max_{\mathscr{K}_{\mathrm{O}}} \left(\sum_{\theta_k \in \mathscr{K}_{\mathrm{O}}} \underline{r}_{\theta_k}^{\mathrm{S}} + \sum_{\vartheta_k \in \mathscr{K}_{\mathrm{S}}} \underline{r}_{\vartheta_k}^{\mathrm{S}}\right)\right), \end{aligned} \tag{5.15}$$

where

$$\begin{cases} \underline{\hat{r}}_m^{\mathrm{M}} = \log_2\left(1 + \dfrac{P_{\mathrm{M}} \beta_{0,m}^{\mathrm{M,M}} (N_{\mathrm{M}} - M - K_{\mathrm{O}})}{\sum\limits_{k=1}^{K} P_{\mathrm{S}} \beta_{k,m}^{\mathrm{S,M}} + \delta^2}\right), \\ \underline{r}_{\theta_k}^{\mathrm{S}} = \log_2\left(1 + \dfrac{P_{\mathrm{S}} \beta_{\theta_k,\theta_k}^{\mathrm{S,S}} (N_{\mathrm{S}} - 1)}{\sum\limits_{i=1, i \neq \theta_k}^{K} P_{\mathrm{S}} \beta_{i,\theta_k}^{\mathrm{S,S}} + \delta^2}\right), \\ \underline{r}_{\vartheta_k}^{\mathrm{S}} = \log_2\left(1 + \dfrac{P_{\mathrm{S}} \beta_{\vartheta_k,\vartheta_k}^{\mathrm{S,S}} (N_{\mathrm{S}} - 1)}{\sum\limits_{m=1}^{M} P_{\mathrm{M}} \beta_{0,\vartheta_k}^{\mathrm{M,S}} + \sum\limits_{i=1, i \neq \vartheta_k}^{K} P_{\mathrm{S}} \beta_{i,\vartheta_k}^{\mathrm{S,S}} + \delta^2}\right). \end{cases}$$

For any SU k, we use (5.12) and (5.13) to obtain $\underline{r}_{\theta_k}^{\mathrm{S}}$ and $\underline{r}_{\vartheta_k}^{\mathrm{S}}$, and denote as r_k^1 and r_k^2, respectively. Then, $\underline{\hat{r}}_m^{\mathrm{M}}$ can be obtained using (5.11). We define $\mathscr{R}_1 = \left\{r_1^1, r_2^1, \cdots, r_K^1\right\}$, $\mathscr{R}_2 = \left\{r_1^2, r_2^2, \cdots, r_K^2\right\}$ and $\Delta\mathscr{R} = \left\{\Delta r_{\gamma_1}, \Delta r_{\gamma_2}, \cdots, \Delta r_{\gamma_K}\right\}$, where $\Delta r_{\gamma_i} \geq \Delta r_{\gamma_j}$ when $\gamma_i \leq \gamma_j$, and

Algorithm 4: Sub-optimal SR-M algorithm

1 **Initialize** $\mathscr{R}_1 = \left\{r_1^1, r_2^1, \cdots, r_K^1\right\}$, $\mathscr{R}_2 = \left\{r_1^2, r_2^2, \cdots, r_K^2\right\}$ according to (5.18) and (5.19), $\Delta\mathscr{R} = \left\{\Delta r_{\gamma_1}, \Delta r_{\gamma_2}, \cdots, \Delta r_{\gamma_K}\right\}$
2 **for** $K_{\mathrm{O}} = 0 : K$ **do**
3 $\quad \mathscr{K}_{\mathrm{O}}(K_{\mathrm{O}}) = \{\gamma_0, \gamma_1, \gamma_2, \cdots, \gamma_{K_{\mathrm{O}}}\}$, where $\gamma_0 = \{\emptyset\}$
4 $\quad$ Computing ergodic downlink sum rate $F(K_{\mathrm{O}})$ according to (5.16)
5 **end for**
6 $K_{\mathrm{O}}^* = \operatorname{argmax} F(K_{\mathrm{O}})$
7 $\mathscr{K}_{\mathrm{O}}^{\mathrm{optimal}}(K_{\mathrm{O}}^*) = \{\gamma_0, \gamma_1, \gamma_2, \cdots, \gamma_{K_{\mathrm{O}}^*}\}$

$\Delta r_{\gamma_k} = r_{\gamma_k}^1 - r_{\gamma_k}^2 (\gamma_k \in \mathscr{K})$. According to (5.15), we have:

$$\begin{aligned} F(K_{\mathrm{O}}) &\triangleq \max_{\mathscr{K}_{\mathrm{O}}} \sum_{\theta_k \in \mathscr{K}_{\mathrm{O}}} \underline{R}_{\theta_k}^{\mathrm{S}} + \sum_{\vartheta_k \in \mathscr{K}_{\mathrm{S}}} \underline{R}_{\vartheta_k}^{\mathrm{S}} + \sum_{m=1}^{M} \hat{\underline{R}}_m^{\mathrm{M}} \\ &= \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right) \left(\sum_{m=1}^{M} \hat{\underline{r}}_m^{\mathrm{M}} + \sum_{k=1}^{K} r_k^1 + \max_{\mathscr{K}_{\mathrm{O}}} \sum_{k \in \mathscr{K}_{\mathrm{O}}} \Delta r_{\gamma_k}\right) \\ &= \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right) \left(\sum_{m=1}^{M} \hat{\underline{r}}_m^{\mathrm{M}} + \sum_{k=1}^{K} r_k^1 + \sum_{i=1}^{K_{\mathrm{O}}} \Delta r_{\gamma_i}\right), \end{aligned} \tag{5.16}$$

where $\sum_{i=1}^{K_{\mathrm{O}}} \Delta r_{\gamma_i} = 0$ when $K_{\mathrm{O}} = 0$. According to (5.16), since $K_{\mathrm{O}} \in [0,\ K]$, we obtain the optimal $\mathscr{K}_{\mathrm{O}}$ and K_{O} using one-dimension search and summarize the above description in Algorithm 3 (Optimal SR-M algorithm).

Sub-optimal SR-M Algorithm

To simplify the Algorithm 3, we propose a suboptimal algorithm that approximates the SU's location with the SBS's location. Specifically, since coverage of each SC is small compared with to MC with short distance between a SU and its associated SBS, the SU's location can be approximated as SBS's location. Let us assume two SCs located in the two-tier mMIMO-SC HetNet system as shown in Fig. 5.2. In this figure, the large-scale fading coefficient can be approximated as

$$\beta_{0,j}^{\mathrm{M,S}} \approx \widehat{\beta}_{0,j}^{\mathrm{M,S}} \text{and } \beta_{0,\mathrm{j}}^{\mathrm{M,S}} \approx \widehat{\beta}_{0,\mathrm{j}}^{\mathrm{M,S}}. \tag{5.17}$$

Therefore, the interference terms $\sum_{i \neq \theta_k}^{K} P_{\mathrm{S}} \beta_{i,\theta_k}^{\mathrm{S,S}} + \delta^2$ and $\sum_{m=1}^{M} P_{\mathrm{M}} \beta_{0,\vartheta_k}^{\mathrm{M,S}} + \sum_{i \neq \vartheta_k}^{K} P_{\mathrm{S}} \beta_{i,\vartheta_k}^{\mathrm{S,S}} + \delta^2$ in (5.12) and (5.13) can be approximated as $\sum_{i \neq \theta_k}^{K} P_{\mathrm{S}} \widehat{\beta}_{i,\theta_k}^{\mathrm{S,S}} + \delta^2$ and $\sum_{m=1}^{M} P_{\mathrm{M}} \widehat{\beta}_{0,\vartheta_k}^{\mathrm{M,S}} + \sum_{i \neq \vartheta_k}^{K} P_{\mathrm{S}} \widehat{\beta}_{i,\vartheta_k}^{\mathrm{S,S}} + \delta^2$, respectively. Hence, these two terms can be regarded as constants even when

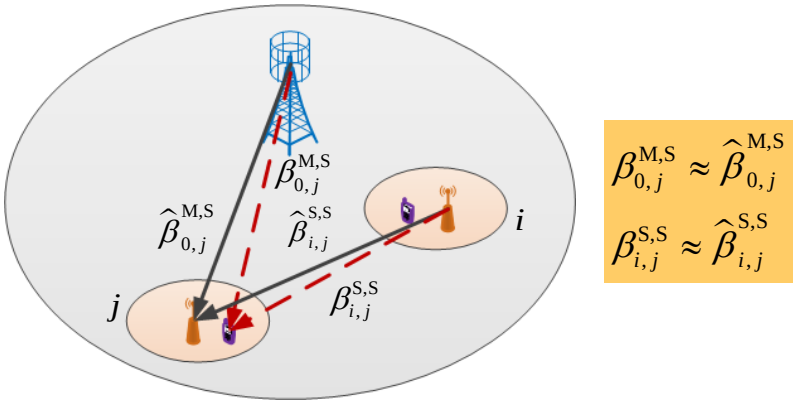


Fig. 5.2 Illustration of approximation in Algorithm 4.

SU's location changes. We can see that the MBS does not need to update $\sum_{i\neq\theta_k}^{K} P_{\mathrm{S}}\beta_{i,\theta_k}^{\mathrm{S,S}}+\delta^2$ and $\sum_{m=1}^{M} P_{\mathrm{M}}\beta_{0,\vartheta_k}^{\mathrm{M,S}}+\sum_{i\neq\vartheta_k}^{K} P_{\mathrm{S}}\beta_{i,\vartheta_k}^{\mathrm{S,S}}+\delta^2$ regularly, which can simplify the pilot allocation algorithm. According to the above approximation, we have:

$$\underline{R}_{\theta_k}^{\mathrm{S}} \approx \left(1-\frac{K_{\mathrm{O}}+M+\xi}{S}\right)\log_2\left(1+\frac{P_{\mathrm{S}}\beta_{\theta_k,\theta_k}^{\mathrm{S,S}}(N_{\mathrm{S}}-1)}{\sum\limits_{i=1,i\neq\theta_k}^{K} P_{\mathrm{S}}\hat{\beta}_{i,\theta_k}^{\mathrm{S,S}}+\delta^2}\right), \tag{5.18}$$

$$\underline{R}_{\vartheta_k}^{\mathrm{S}} \approx \left(1-\frac{K_{\mathrm{O}}+M+\xi}{S}\right)\log_2\left(1+\frac{P_{\mathrm{S}}\beta_{\vartheta_k,\vartheta_k}^{\mathrm{S,S}}(N_{\mathrm{S}}-1)}{\sum\limits_{m=1}^{M} P_{\mathrm{M}}\hat{\beta}_{0,\vartheta_k}^{\mathrm{M,S}}+\sum\limits_{\substack{i=1\\ i\neq\vartheta_k}}^{K} P_{\mathrm{S}}\hat{\beta}_{i,\vartheta_k}^{\mathrm{S,S}}+\delta^2}\right). \tag{5.19}$$

Then, we can get $\mathscr{R}_1$, $\mathscr{R}_2$ and $\Delta\mathscr{R}$, which have the similar computing process with Algorithm 4. We call this suboptimal algorithm as Algorithm 4 (Sub-optimal SR-M algorithm).

SUs' Fairness-Aware Algorithm

To consider SUs' fairness, based on the original problem (5.14) we propose a SUs' fairness-aware algorithm. Here, the orthogonal pilots are preferentially allocated to relatively low-rate SUs and then to maximize the ergodic downlink sum rate. According to the above analysis, we sort $\mathscr{R}_1$ with ascending order and denote as $\mathscr{R}_1^* = \left\{r_{\lambda_1}^1, r_{\lambda_2}^1, \cdots, r_{\lambda_K}^1\right\}$ $(\lambda_k \in$

Algorithm 5: SUs' fairness-aware algorithm

Initialize $\mathscr{R}_1^* = \left\{r_{\lambda_1}^1, r_{\lambda_2}^1, \cdots, r_{\lambda_K}^1\right\}$, $\mathscr{R}_2^* = \left\{r_{\lambda_1}^2, r_{\lambda_2}^2, \cdots, r_{\lambda_K}^2\right\}$ according to (5.12) and (5.13), $\Delta\mathscr{R}^* = \left\{\Delta r_{\lambda_1}, \Delta r_{\lambda_2}, \cdots, \Delta r_{\lambda_K}\right\}$
for $K_{\mathrm{O}} = 0 : K$ **do**
  $\mathscr{K}_{\mathrm{O}}(K_{\mathrm{O}}) = \{\lambda_0, \lambda_1, \lambda_2, \cdots, \lambda_{K_{\mathrm{O}}}\}$, where $\lambda_0 = \{\emptyset\}$
  Computing ergodic downlink sum rate $F(K_{\mathrm{O}})$ according to (5.20)
end for
$K_{\mathrm{O}}^* = \operatorname{argmax} F(K_{\mathrm{O}})$
$\mathscr{K}_{\mathrm{O}}^{\mathrm{optimal}}(K_{\mathrm{O}}^*) = \{\lambda_0, \lambda_1, \lambda_2, \cdots, \lambda_{K_{\mathrm{O}}^*}\}$

Table 5.1 Simulation parameters.

Parameters	Value
Radius of macro cell	1000 m
Radius of micro cell	30 m
Number of MUs	50
Number of SBS antennas	4
Transmit power of MBS	46dBm
Transmit power of SBS	23dBm
Frame duration	$200T$
Pathloss between MBS and MU or SU	27.3+39.1log10(d)
Pathloss between SBS and MU or SU	36.8+36.7log10(d)
Downlink Bandwidth	10 MHz
Noise Power	-174dBm/Hz

$\mathscr{K}$), where $r_{\lambda_i}^1 \leq r_{\lambda_j}^1$ when $\lambda_j \geq \lambda_i$, then we can get $\mathscr{R}_2^* = \left\{r_{\lambda_1}^2, r_{\lambda_2}^2, \cdots, r_{\lambda_K}^2\right\}$ and $\Delta\mathscr{R}^* = \left\{\Delta r_{\lambda_1}, \Delta r_{\lambda_2}, \cdots, \Delta r_{\lambda_K}\right\}$ ($\Delta r_{\lambda_k} = r_{\lambda_k}^1 - r_{\lambda_k}^2$). Therefore, the $F(K_{\mathrm{O}})$ can be written as follows:

$$F(K_{\mathrm{O}}) = \left(1 - \frac{K_{\mathrm{O}} + M + \xi}{S}\right)\left(\sum_{m=1}^{M} \hat{\underline{r}}_m^{\mathrm{M}} + \sum_{k=1}^{K} r_k^1 + \sum_{i=1}^{K_{\mathrm{O}}} \Delta r_{\lambda_i}\right). \tag{5.20}$$

We summarize the above algorithm as Algorithm 5 (SUs' fairness-aware algorithm).

5.4 Simulation Results and Discussions

The ergodic downlink sum rate of MUs and SUs for the proposed algorithms is compared in this section. We consider a single MC with a radius of 1000 meters, where the MBS is

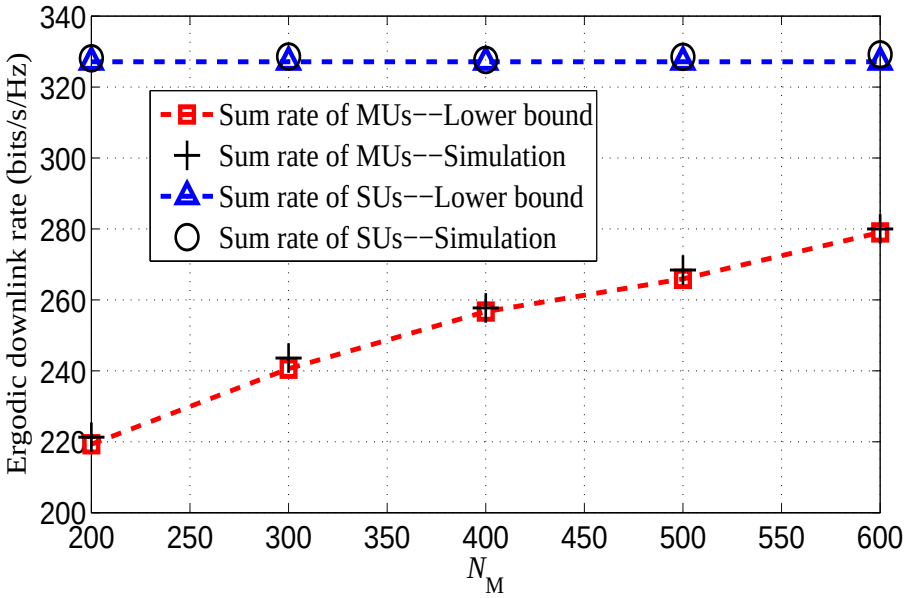


Fig. 5.3 The ergodic downlink rate versus N_{M} with $K = 50$ and $K_{\mathrm{O}} = 20$.

located at the center of the MC and MUs are uniformly distributed in the MC. We assume the MC-hole radius is 100 meters (all MUs or SUs do not figure in this scenario), and the radius of each SC is 30 meters. All SCs are randomly located within the MC, and SUs are randomly located in each SC. As a typical example, we assume that the minimum distance between the SU and SBS is 5 meters, and then the distance between any two SBSs is longer than 120 meters. The results are averaged over 10^3 trials. Other related simulation parameters are listed in Table 5.1.

Fig. 5.3 plots the ergodic downlink sum rate of MUs and SUs versus N_{M} with $K = 50$, respectively. In this figure, the lower bound of ergodic downlink rate is compared with the simulation results, where $K_{\mathrm{O}} = 20$ is used in the Algorithm 3. From Fig. 5.3, we can find that the gap between lower bound of ergodic downlink rate and the simulation results is small enough. Although the downlink sum rate of MUs increases with the number of MBS antennas N_M, the downlink sum rate of SUs is kept constant, because SUs' rate is given regardless of N_M. In the following simulation, we just consider the lower bound of ergodic downlink sum rate of MUs and SUs.

Fig. 5.4 plots the ergodic downlink sum rate versus K_{O} with $N_{\mathrm{M}} = 500$ and $K = 50$. We find that the ergodic downlink sum rate is a convex function of K_{O}. The performance gap of the Algorithms 3 and 4 is very small, and the maximum ergodic downlink sum rate is nearly the same. Since Algorithm 3 considers the SUs' fairness (i.e., orthogonal pilots are

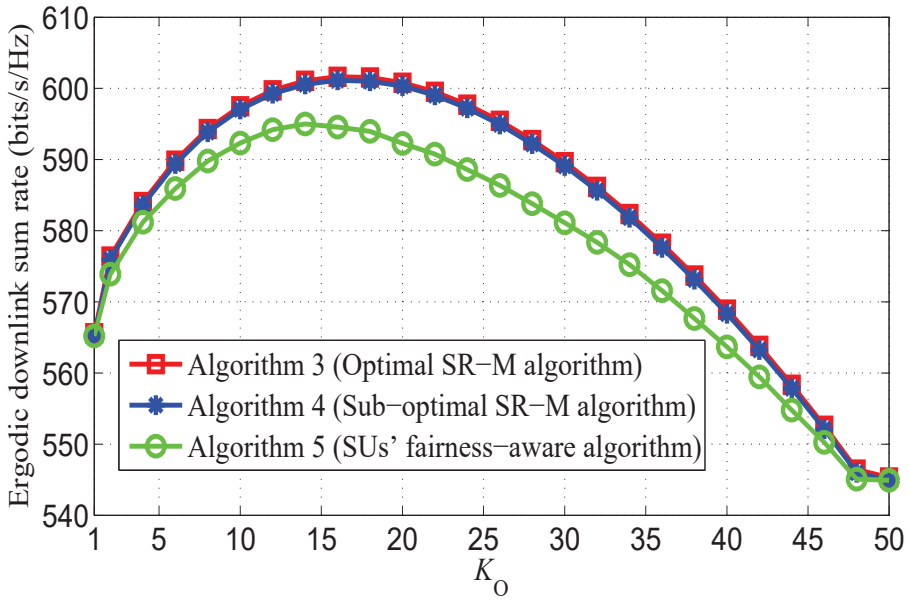


Fig. 5.4 The ergodic downlink sum rate versus K_O with $N_M = 500$ and $K = 50$.

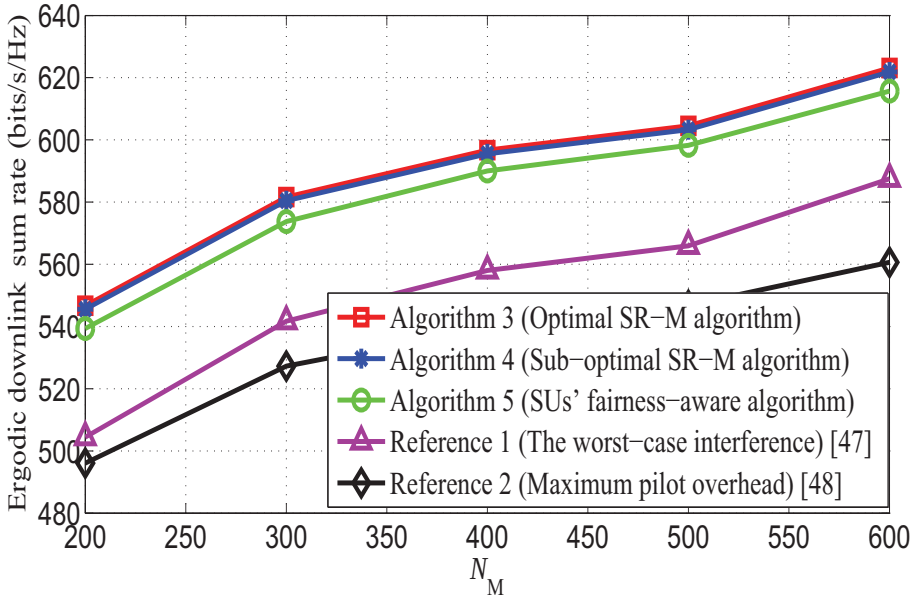


Fig. 5.5 The ergodic downlink sum rate versus N_M with $K = 50$.

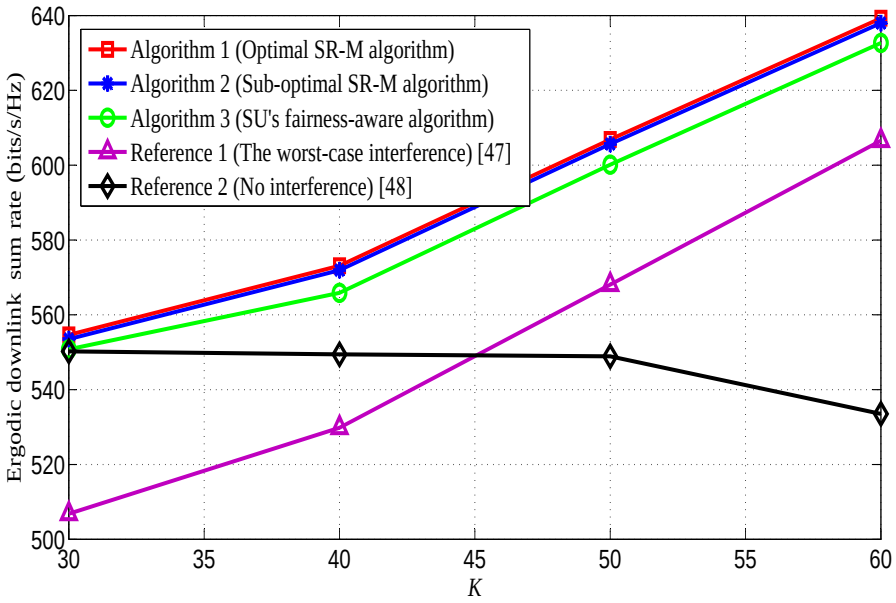


Fig. 5.6 The ergodic downlink sum rate versus K with $N_{\mathrm{M}} = 500$.

preferentially allocated to relatively low-rate SUs to improve their achievable rate), the sum rate is lower than those of Algorithms 3 and 4.

Fig. 5.5 shows that the ergodic downlink sum rate versus N_{M} with $K = 50$. To compare with the performance of our proposed algorithms, we also plot the ergodic downlink sum rate of traditional pilot allocation schemes. For example, reference 1 considers the worst-case interference between MBS and SUs, where all SCs use the same pilot [47], and reference 2 considers no interference between MBS and SUs, where all SCs use orthogonal pilots [48]. Obviously, our proposed algorithms effectively improve the sum rate of the system (about 12%) in comparison with the references 1 and 2. In addition, we also find that the performance of reference 2 is the worst. Although the inter-tier interference from MBS to SUs can be canceled, the large number of pilots results in decreasing data transmission in one coherent time block. Fig. 5.5 also shows that ergodic downlink sum rate in all schemes is improved as the number of antennas N_M increases, because of increased antenna gain. Similarly to the results in Fig. 5.4, the sum rates of Algorithms 3 and 4 are nearly the same. To preferentially improve achievable rate of lower-rate SUs, Algorithm 5 will lose some overall performance gains, and this point is also shown in Fig. 5.5.

Fig. 5.6 plots the ergodic downlink sum rate versus K with $N_{\mathrm{M}} = 500$, we can get that the ergodic downlink sum rate increases with K, and the performance of the first two algorithms are still better than that of the third algorithm. In addition, it can be found that the performance of our proposed algorithms is always better than that of references 1 and 2.

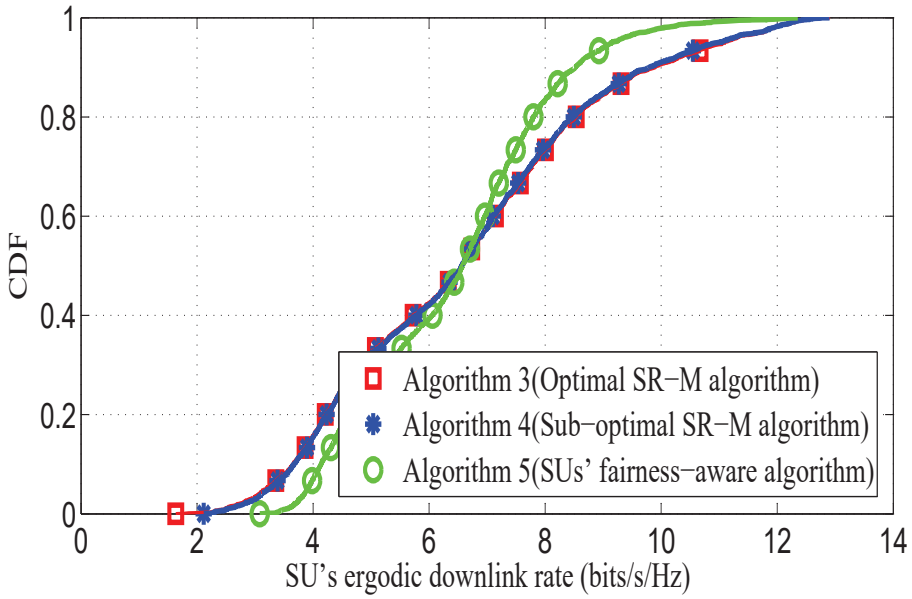


Fig. 5.7 The CDF of SU's ergodic downlink rate with $N_{\mathrm{M}} = 500$ and $K = 50$.

Meanwhile, we find that the sum rate of reference 2 decreases with the number of SCs. The reason is that more SCs lead to more orthogonal pilots, which results in less time for data transmission.

Fig. 5.7 shows that the cumulative distribution function (CDF) curve of the SU's ergodic downlink rate with $N_{\mathrm{M}} = 500$ and $K = 50$. Compared with the first two algorithms, we find that there is lower proportion of SUs with the smaller downlink rate in Algorithm 5. Meanwhile, for the higher downlink rate, the proportion of SUs is higher for the first two algorithms. The result suggests that Algorithm 5 preferentially improves achievable rate of relatively low-rate SUs by sacrificing the sum-rate performance of the system.

5.5 Conclusions

In this chapter, we have investigated the uplink pilot allocation problem for the two-tier TDD mMIMO-SC HetNet. An optimal pilot allocation algorithm has been proposed to maximize the ergodic downlink sum rate of MUs and SUs. In addition, a suboptimal algorithm has been proposed, in which the SU's location is approximated as SBS's location based on the assumption that the coverage area of SCs is small enough. Then, for guaranteeing the SUs' fairness, we have presented another algorithm to preferentially improve achievable rate of relatively low-rate SUs. Simulation results have demonstrated that our proposed scheme

can improve the sum rate of the system by about 12% in comparison with the conventional schemes.

Chapter 6

Small Cell Clustering and Precoding Design for Small Cell Type Heterogeneous Network with Massive MIMO Marco Cell

6.1 Introduction

In previous chapters, we have proposed some pilot and power allocation schemes to reduce the pilot contamination and signal interference. In this chapter, we will consider how to coordinate the interference by SC clustering and precoding design in a mMIMO-SC HetNet. In literature [49]-[52], there have been some studies about SC clustering for interference coordination. *Zhou et al.* in [49] propose a graph-based approach combining SC clustering and user clustering for mitigating interference among SCs. According to the downlink SINR user receives, SCs are grouped into multiple clusters. A rate loss-based low-complexity algorithm for SC clustering is proposed by *Seno et al.* in [50]. Each SBS's transmit power must be decided in advance for the above clustering schemes, which are inapplicable in our study. *Hong et al.* in [51] investigate the BSs clustering and precoding design for partially coordinated transmission to maximize the utility of the system. Since one user can be served by different BS clusters in their scheme, the intra-cluster interference can not be cancelled completely. A dynamic greedy algorithm for cooperative BSs clustering is proposed in [52], but the number of BSs in each cluster must be known in advance. *Fan et al.* in [53] propose a distance-based SBS clustering scheme. Although the proposed algorithm is simple, it is not appropriate for time-varying environment. In fact, [49]-[52] consider the clustering at user

level, namely the cluster is formed by considering the real-time interference between users and BS, which is superior than clustering at SC level (e.g., [53]).

Several precoding schemes for interference coordination in MIMO system have also been investigated in literature [54]-[57]. For example, *Zhang et al.* in [54] study the precoding design optimization problem for maximizing the weighted sum rate of all users in multi-cell system. Since the multi-user interference can be eliminated by using the proposed block diagonalization (BD) precoding technique, the original problem can be transformed into convex optimization problem. Similarly, in [55] and [56], the authors design effective precoder to eliminate the multi-user interference so that the original problem can be transformed into convex optimization problem. Then, the optimal precoder can be obtained by using interior-point method or convex optimization toolbox directly. *Niu et al.* in [57] propose a joint interference alignment and power allocation problem for reducing the intra- and inter-tier interference, which is solved by some simply linear algorithms. Although the linear algorithm has low complexity, it cannot be used in non-convex problems such as ours due to per-SBS power constraint and interference among clusters.

Unlike previous works, in this chapter we investigate a new SC clustering strategy and their precoding designs for maximizing downlink sum rate of SUs in two-tier mMIMO-SC HetNet. The main contributions are summarized as follows:

- To reduce the interference among SCs, an interference graph-based dynamic SC clustering scheme is proposed, where SCs are grouped into multiple SC clusters according to their interference channel strength. On this basis, the SUs' signals are jointly designed in each cluster.
- To achieve joint interference coordination mentioned above, we formulate an optimization problem to design precoding weights at MBS and clustered SCs for maximizing the downlink sum rate of SUs subject to per-SBS power constraint. Precoding weights at MBS are designed to eliminate the multi-MU and inter-tier interference, while precoding weights at clustered SCs are designed to cancel the intra-cluster interference and mitigate inter-cluster interference.
- To eliminate multi-MU and inter-tier interference simultaneously, we propose a clustered SC BD (CSBD) precoding scheme for MBS. Specifically, we use the singular value decomposition (SVD) to find the null space of the inter-tier interference channels. Following this, the ZF downlink precoding weight matrix for MUs is projected onto the above null space to simultaneously cancel the multi-MU and inter-tier interference.
- To cancel the intra-cluster interference and coordinate inter-cluster interference, the precoding vector of each SU at clustered SCs is designed as the product of the following

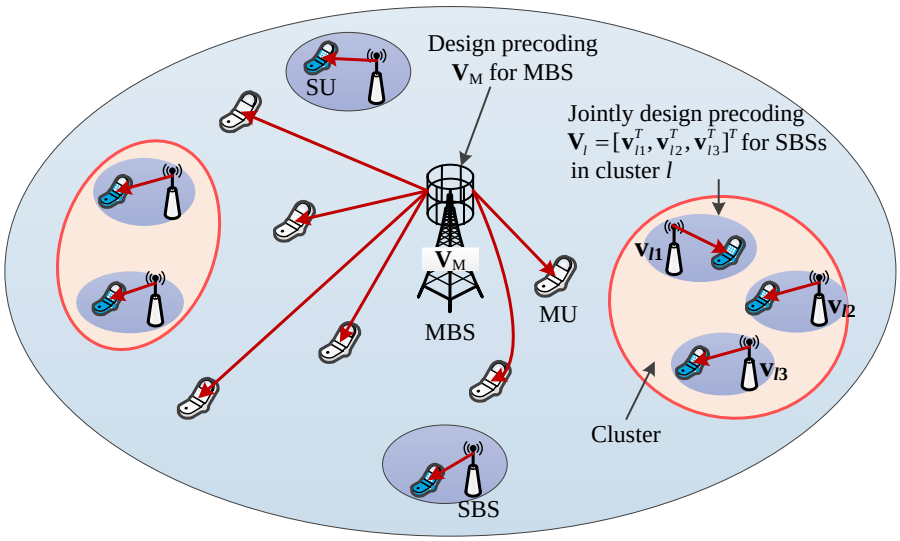


Fig. 6.1 System model for small cluster-based two-tier downlink mMIMO-SC HetNet.

two parts. The first part is designed with SVD to remove the intra-cluster interference. The second part is designed to coordinate the inter-cluster interference for maximizing the downlink sum rate of SUs. This is a non-convex optimization problem that is difficult to directly solve. We propose a cluster-based non-cooperative game and develop a distribute algorithm to obtain a suboptimal solution. Finally, we prove the existence and uniqueness of the Nash equilibria (NE) for the formed game.

6.2 System Model and Problem Formulation

6.2.1 System Model

A two-tier downlink mMIMO-SC HetNet system is considered as shown in Fig. 6.1, which is composed of a MC and J overlaid SCs ($\mathbf{J} = \{1, 2, \ldots, J\}$). We assume that K_{M} single-antenna MUs are served by the central MBS equipped with M antennas ($M \gg K_{\mathrm{M}}$) , and each SBS associated with N antennas serves K_{S} single-antenna SUs ($N \geq K_{\mathrm{S}}$) . The MC and all SCs share the overall spectrum while users (MUs and SUs) are served with the same time-frequency resource. In this case, there exist interference among SCs. Cooperative downlink transmission among SCs eliminates the inter-SC interference. However, transmission data must be exchanged and shared among SBSs, which needs huge backhaul overhead. SC clustering approach is effective in decreasing the required overhead, where the transmission

data are only shared within each cluster. To harvest the benefits of the mMIMO antennas, we assume that TDD protocol is applied with perfect CSI at MBS and SBSs [45], [46].

We assume that all SCs are grouped into C clusters, where each SU is served by all SBSs belonging to the same cluster. Let C_l denote the number of SCs in the lth cluster with the total number of SUs and antennas across all SBSs in the lth cluster denoted as $K_l = C_l K_{\mathrm{S}}$ and $N_l = C_l N$, respectively. Meanwhile, we assume that $((O_l-1)N+1)$th to $(O_l N)$th antennas are taken as the N SBS antennas in the O_lth SBS in the lth cluster with $O_l = 1,2,\ldots,C_l$. Similarly, the indices of SUs in the O_lth SBS in the lth cluster can be denoted as the $((O_l-1)K_{\mathrm{S}}+1)$th to $(O_l K_{\mathrm{S}})$th.

The received signal by the kth SU in the lth cluster can be expressed as:

$$\begin{aligned} y_{lk} &= \sum_{i=1}^{C}\sum_{j=1}^{K_i}\mathbf{h}_{ilk}\mathbf{v}_{ij}x_{ij} + \sum_{m=1}^{K_{\mathrm{M}}}\mathbf{h}_{0lk}\mathbf{v}_{0m}x_{0m} + n_{lk} \\ &= \underbrace{\mathbf{h}_{llk}\mathbf{v}_{lk}x_{lk}}_{\text{Desired signal}} + \underbrace{\sum_{j\neq k}^{K_l}\mathbf{h}_{llk}\mathbf{v}_{lj}x_{lj}}_{\text{Intra-cluster interference}} + \underbrace{\sum_{i\neq l}^{C}\sum_{j=1}^{K_i}\mathbf{h}_{ilk}\mathbf{v}_{ij}x_{ij}}_{\text{Inter-cluster interference}} + \underbrace{\sum_{m=1}^{K_{\mathrm{M}}}\mathbf{h}_{0lk}\mathbf{v}_{0m}x_{0m}}_{\text{Inter-tier interference}} + \underbrace{n_{lk}}_{\text{Noise}}, \end{aligned} \tag{6.1}$$

where $\mathbf{h}_{ilk} \in \mathbb{C}^{1\times N_i}$ and $\mathbf{h}_{0lk} \in \mathbb{C}^{1\times M}$, respectively, denote the downlink channel from all C_i SBSs and the MBS to the kth SU in the lth cluster. $\mathbf{v}_{ij} \in \mathbb{C}^{N_i\times 1}$ and $\mathbf{v}_{0m} \in \mathbb{C}^{M\times 1}$, respectively, denote the precoding vector for the jth SU in the ith cluster and the mth MU. x_{ij} and x_{0m} denote the transmit signals of the jth SU in the ith cluster and the mth MU, respectively. We assume $\mathbb{E}[|x|^2] = 1$ and n_{lk} is an i.i.d. AWGN defined as $CN(0,\ \delta^2)$.

Similarly, the received signal at the kth MU can be expressed as follows:

$$\begin{aligned} y_{0k} &= \sum_{m=1}^{K_{\mathrm{M}}}\mathbf{h}_{00k}\mathbf{v}_{0m}x_{0m} + \sum_{i=1}^{C}\sum_{j=1}^{K_i}\mathbf{h}_{i0k}\mathbf{v}_{ij}x_{ij} + n_{0k} \\ &= \underbrace{\mathbf{h}_{00k}\mathbf{v}_{0k}x_{0k}}_{\text{Desired signal}} + \underbrace{\sum_{m\neq k}^{K_{\mathrm{M}}}\mathbf{h}_{00k}\mathbf{v}_{0m}x_{0m}}_{\text{intra-tier interference}} + \underbrace{\sum_{i=1}^{C}\sum_{j=1}^{K_i}\mathbf{h}_{i0k}\mathbf{v}_{ij}x_{ij}}_{\text{Inter-tier interference}} + \underbrace{n_{0k}}_{\text{Noise}}, \end{aligned} \tag{6.2}$$

where $\mathbf{h}_{00k} \in \mathbb{C}^{1\times M}$ and $\mathbf{h}_{i0k} \in \mathbb{C}^{1\times N_i}$ denote the downlink channel from the MBS and all C_i SBSs in the ith cluster to the kth MU, respectively.

Following this, the constraints for eliminating inter-tier interference from MBS to SUs and intra-cluster interference are given as:

$$\mathbf{h}_{0lk}\mathbf{v}_{0m} = 0, \forall l,k,m = \{1,\ldots,K_{\mathrm{M}}\}, \tag{6.3a}$$

$$\mathbf{h}_{llk}\mathbf{v}_{lj} = 0, \forall k \neq j, l = \{1,\ldots,C\}, \tag{6.3b}$$

Therefore, the received signal by the kth SU in the lth cluster can be rewritten as:

$$y_{lk} = \mathbf{h}_{llk}\mathbf{v}_{lk}x_{lk} + \sum_{i\neq l}^{C}\sum_{j=1}^{K_i}\mathbf{h}_{ilk}\mathbf{v}_{ij}x_{ij} + n_{lk}, \tag{6.4}$$

and its rate can be expressed as follows:

$$R_{lk} = \log_2\left(1 + \frac{\mathbf{h}_{llk}\mathbf{v}_{lk}\mathbf{v}_{lk}^H\mathbf{h}_{llk}^H}{\sum\limits_{i\neq l}^{C}\sum\limits_{j=1}^{K_i}\mathbf{h}_{ilk}\mathbf{v}_{ij}\mathbf{v}_{ij}^H\mathbf{h}_{ilk}^H + \delta^2}\right). \tag{6.5}$$

6.2.2 Problem Formulation

Since total transmit antennas in each cluster comes from more than one SBS, the per-SBS power constraint is expressed as follows:

$$\sum_{k=1}^{K_l}\mathrm{Tr}(\mathbf{B}_{O_l}\mathbf{v}_{lk}\mathbf{v}_{lk}^H) \leq P, \forall l, O_l = \{1,\ldots,C_l\}, \tag{6.6}$$

with

$$\mathbf{B}_{O_l} \triangleq \mathrm{Diag}(\underbrace{0,\ldots,0}_{(O_l-1)N},\underbrace{1,\ldots,1}_{N},\underbrace{0,\ldots,0}_{(C_l-O_l)N}). \tag{6.7}$$

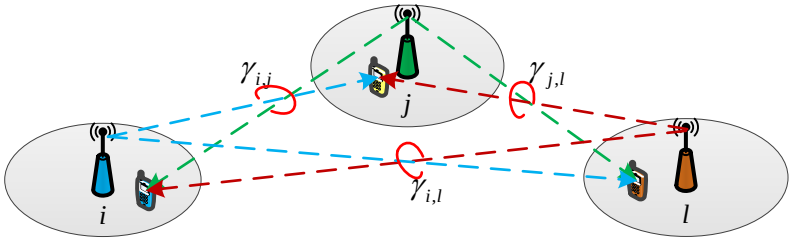


Fig. 6.2 An example for interference graph.

Next, we formulate the optimization problem to maximize the downlink sum rate of SUs as follows:

$$\max_{\{\mathbf{V}_1,\ldots,\mathbf{V}_C\},\mathbf{V}_M} \quad \sum_{l=1}^{C}\sum_{k=1}^{K_l} R_{lk} \tag{6.8a}$$

$$\text{s.t.} \quad \mathbf{h}_{0lk}\mathbf{v}_{0m} = 0, \forall l,k,m = \{1,\ldots,K_M\}, \tag{6.8b}$$

$$\mathbf{h}_{llk}\mathbf{v}_{lj} = 0, \forall k \neq j, l = \{1,\ldots,C\}, \tag{6.8c}$$

$$\sum_{k=1}^{K_l} \mathrm{Tr}(\mathbf{B}_{O_l}\mathbf{v}_{lk}\mathbf{v}_{lk}^H) \leq P, \forall l, O_l, \tag{6.8d}$$

$$\mathbf{v}_{lk} \succeq \mathbf{0}, \mathbf{v}_{0m} \succeq \mathbf{0}, \forall l,k,m, \tag{6.8e}$$

where $\mathbf{V}_l = [\mathbf{v}_{l1}^T, \ldots, \mathbf{v}_{lK_l}^T]^T$ and $\mathbf{V}_M = [\mathbf{v}_{01}^T, \ldots, \mathbf{v}_{0K_M}^T]^T$.

To solve the above problem (6.8), we consider the following three steps. The first step is to design a SC clustering scheme so that all SCs form multiple clusters as shown. The second step designs precoding at MBS to eliminate inter-tier interference, namely (6.8b). The last step designs precoding at each cluster to maximize the downlink sum rate of SUs, namely (6.8a), (6.8c) and (6.8d).

6.3 SC Clustering Scheme for Interference Coordination

We define the average interference channel strength between two SCs i, j as follows:

$$\gamma_{i,j} = \frac{1}{NK_T}\sum_{k=1}^{K_S}\left(\left\|\bar{\mathbf{h}}_{ijk}\right\| + \left\|\bar{\mathbf{h}}_{jik}\right\|\right), i,j = \{1,\ldots,C\}, \tag{6.9}$$

where K_T denotes the total number of SUs in SCs i and j ($K_T = 2K_S$). $\bar{\mathbf{h}}_{ijk} \in \mathbb{C}^{1\times N}$ denotes the downlink interference channel from the ith SC to the kth SU in the jth SC. $\gamma_{i,j}$ in (6.9) represents the potential average interference strength level between SCs i and j. For

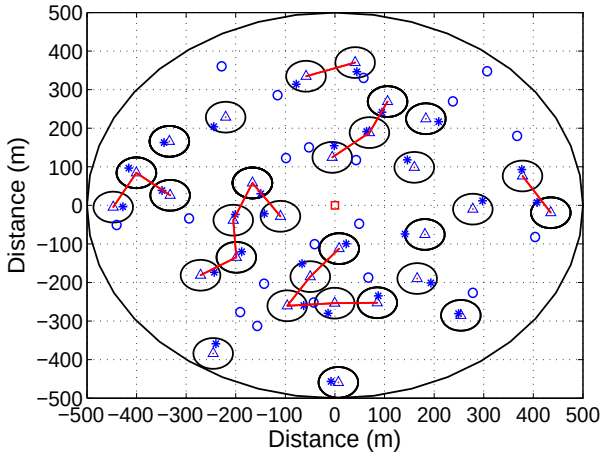


Fig. 6.3 An example for SC clustering with γ_{th}=-100dB.

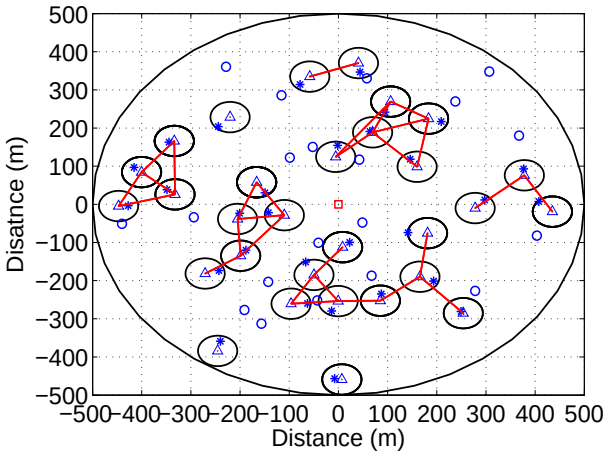


Fig. 6.4 An example for SC clustering with γ_{th}=-105dB.

Algorithm 6: Interference graph-based dynamic SC clustering algorithm

Input : J, γ_{th}, $n = 1$.
Output : $\mathbf{C}$.

```
for i = 1 : J do
    C{n} = {i}.
    for j = 1 : J (i ≠ j) do
        Compute γ_{i,j} according to (6.9).
        if γ_{i,j} ≥ γ_th then
            The SCs i and j form a new cluster, namely C{n} = C{n} ∪ {j}.
            if The SC i or j has belonged to any other cluster n' then
                The SCs i, j and their all cluster members form a new cluster, namely
                C{n} = C{n} ∪ C{n'}, n = n − 1.
            end if
        end if
    end for
    n = n + 1.
end for
```

Note: Here, $\mathbf{C}$ should be a set consisting of several subsets. Each subset indicates a cluster and its elements represent SCs' index. $\mathbf{C}\{n\}$ denotes the nth subset in set $\mathbf{C}$.

example, a larger $\gamma_{i,j}$ denotes higher interference and vice versa. Therefore, two SCs form a cluster when $\gamma_{i,j}$ is high.

The potential interference relationship among all SCs can be constructed as the interference graph. First, we set an interference threshold γ_{th} that is used to determine whether two SCs should form a cluster. We assume three SCs in Fig. 6.2 denoted as i, j, l. If we have $\gamma_{i,j} \geq \gamma_{th}$, $\gamma_{j,l} \geq \gamma_{th}$ and $\gamma_{i,l} < \gamma_{th}$, according to the above definition, the SC j will belong to two different clusters. For simplifying the problem and coordinating more interference among SCs, the SCs i, j and l will form one cluster under the above situation.

Based on the above analysis, an interference graph can be constructed as an undirected graph $G(\mathbf{J},\mathbf{E})$ in SBSs, where $\mathbf{J}$ vertices denote all SCs and $E(u,v)$ edges stand the potential interference between SCs u and v, $\forall u,v \in \mathbf{J}$. Figs. 6.3 and 6.4 illustrate the SC clustering results with different γ_{th} having the same randomly located SCs. It can be clearly observed from the figure that higher interference threshold leads to smaller-size clusters with fewer SCs. This reduces the information exchange within each cluster, which decreases the backhaul overhead and system latency. On the contrary, lower interference threshold leads to larger-size clusters with more SCs. This increases information exchange within each cluster, which increases the backhaul overhead and system latency. Therefore, in practice, the interference threshold can be determined according to the required criteria, e.g., the system sum rate maximization, delay and/or backhaul overhead minimization. In this section, the interference

threshold is empirically selected to maximize the system sum rate for simplicity. We note here that information exchange is only done between SBSs within each cluster as indicated by $\mathbf{h}_{llk}$ and x_{lk} in (6.4). The information exchange among users is not needed.

The SC clustering scheme is presented in Algorithm 6. We first compute $\gamma_{i,j}$ between SCs i and j (in line 4). Next, two SCs are decided whether to form a cluster or not based on γ_{th} (in lines 5-6). When SC i and j form a cluster according to the above scheme, if any one of them has belonged to other cluster, SCs i, j and that cluster members will reform one cluster (in lines 7-9). The above procedure is repeated until all SCs are clustered.

6.4 CSBD Precoding Design for MBS

To eliminate the inter-tier interference from MBS to SUs and multi-MU interference simultaneously, we propose a CSBD precoding scheme for MBS.

First, we define the inter-tier interference channels from MBS to SUs as follows:

$$\mathbf{H}_{\text{in}} = [\mathbf{h}_{011}^T, \ldots, \mathbf{h}_{01K_1}^T, \ldots, \mathbf{h}_{0C1}^T, \ldots, \mathbf{h}_{0CK_C}^T]^T. \tag{6.10}$$

where $\mathbf{H}_{\text{in}} \in \mathbb{C}^{JK_{\text{S}} \times M}$. To obtain the null space of the inter-tier interference channels Null ($\mathbf{H}_{\text{in}}$), we apply the classical SVD to the matrix $\mathbf{H}_{\text{in}}$, yielding

$$\mathbf{H}_{\text{in}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H, \tag{6.11}$$

where $\mathbf{U} \in \mathbb{C}^{JK_{\text{S}} \times JK_{\text{S}}}$ denotes the left-singular-vector matrix, $\mathbf{V} \in \mathbb{C}^{M \times M}$ denotes the right-singular-vector matrix, and $\boldsymbol{\Sigma} \in \mathbb{C}^{JK_{\text{S}} \times M}$ denotes the singular values as follows:

$$\Sigma = \begin{bmatrix} \widehat{\boldsymbol{\Sigma}}_r & \mathbf{0}_{r \times (M-r)} \\ \mathbf{0}_{(JK_{\text{S}}-r) \times r} & \mathbf{0}_{(JK_{\text{S}}-r) \times (M-r)} \end{bmatrix}. \tag{6.12}$$

where $r = \text{rank}(\mathbf{H}_{\text{in}})$ is the rank of $\mathbf{H}_{\text{in}}$, and $\widehat{\boldsymbol{\Sigma}}_r = \text{Diag}\{\sigma_1, \ldots, \sigma_{\text{r}}\}$.

Therefore, the null space of $\mathbf{H}_{\text{in}}$ can be found by spanning the columns of $\mathbf{V}$ as follows:

$$\widehat{\mathbf{V}} = [\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_M]. \tag{6.13}$$

Note that there is a constraint condition for the existence of $\widehat{\mathbf{V}} \in \mathbb{C}^{M \times (M-r)}$, namely the number of SUs must be lower than the number of MBS antennas ($JK_S \leq M$). Then, we have the following:

$$\mathbf{H}_{\text{in}}\widehat{\mathbf{V}} = \mathbf{0}. \tag{6.14}$$

According to (6.14), we can find that the inter-tier interference can be cancelled completely if a column vector for each MU is randomly chosen from $\widehat{\mathbf{V}}$, but it may cause serious multi-MU interference. To eliminate the inter-tier and multi-MU interference simultaneously, we first define the projection matrix $\widetilde{\mathbf{V}}_{\text{in}}$ based on the null space Null($\mathbf{H}_{\text{in}}$) as follows:

$$\widetilde{\mathbf{V}}_{\text{in}} = \widehat{\mathbf{V}}\widehat{\mathbf{V}}^H. \tag{6.15}$$

Accordingly, we can project ZF precoding matrix for MUs onto the null space Null($\mathbf{H}_{\text{in}}$) and obtain the finall precoding matrix as:

$$\begin{aligned}
\mathbf{W}_{\text{CSBD}} &= \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H\right)^{-1} \\
&= \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\widetilde{\mathbf{V}}_{\text{in}}^H\mathbf{H}_{\text{M}}^H\right)^{-1} \\
&= \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H \left(\mathbf{H}_{\text{M}}\widehat{\mathbf{V}}\widehat{\mathbf{V}}^H\widehat{\mathbf{V}}\widehat{\mathbf{V}}^H\mathbf{H}_{\text{M}}^H\right)^{-1} \\
&= \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H \left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H\right)^{-1},
\end{aligned} \tag{6.16}$$

where $\mathbf{H}_{\text{M}} = [\mathbf{h}_{001}^T, \mathbf{h}_{002}^T, \ldots, \mathbf{h}_{00K_{\text{M}}}^T]^T$ denotes the multi-MU downlink channel matrix. Meanwhile, the necessary condition for the existence of $\mathbf{W}_{\text{CSBD}}$ is $JK_S + K_{\text{M}} \leq M$, namely the number of SUs and MUs should be lower than the number of MBS antennas. Next, we provide the proof for the above necessary condition. For the i.i.d. Rayleigh fading channel, rank($\mathbf{H}_{\text{in}}$) should be JK_{S}. Thus, $\widehat{\mathbf{V}}$ is a $M \times (M - JK_{\text{S}})$ matrix, where rank($\widehat{\mathbf{V}}$) $= M - JK_{\text{S}}$ and rank($\widetilde{\mathbf{V}}_{\text{in}}$) $\leq M - JK_{\text{S}}$. In addition, since $\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H$ is a $K_{\text{M}} \times K_{\text{M}}$ matrix, the condition for existence of $(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H)^{-1}$ is rank($\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H$) $= K_{\text{M}}$. Due to rank($\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H$) $\leq$ min$\{M - JK_{\text{S}}, K_{\text{M}}\}$, we have $M - JK_{\text{S}} \geq K_{\text{M}}$, namely, $JK_{\text{S}} + K_{\text{M}} \leq M$.

According to the obtained precoding $\mathbf{W}_{\text{CSBD}}$, we have the following:

$$\mathbf{H}_{\text{M}}\mathbf{W}_{\text{CSBD}} = \mathbf{H}_{\text{M}}\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H\right)^{-1} = \mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H\right)^{-1} = \mathbf{I}, \tag{6.17a}$$

$$\mathbf{H}_{\text{in}}\mathbf{W}_{\text{CSBD}} = \mathbf{H}_{\text{in}}\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\right)^H\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H\right)^{-1} = \mathbf{H}_{\text{in}}\widehat{\mathbf{V}}\widehat{\mathbf{V}}^H\mathbf{H}_{\text{M}}^H\left(\mathbf{H}_{\text{M}}\widetilde{\mathbf{V}}_{\text{in}}\mathbf{H}_{\text{M}}^H\right)^{-1} = \mathbf{0}, \tag{6.17b}$$

where (6.17a) and (6.17b), respectively, illustrate that the multi-MU and inter-tier interference can be cancelled.

The precoding vector for the kth MU can be written as:

$$\mathbf{v}_{0k} = \frac{\sqrt{P_{0k}}\mathbf{W}^k_{\mathrm{CSBD}}}{\left\|\mathbf{W}^k_{\mathrm{CSBD}}\right\|}, \tag{6.18}$$

where $\mathbf{W}^k_{\mathrm{CSBD}}$ is the kth column vector of $\mathbf{W}_{\mathrm{CSBD}}$, and P_{0k} denotes the transmit power for the kth MU.

In addition, when $JK_S+K_{\mathrm{M}} \leq M$, the ZF precoding can be also used to cancelled the multi-MU and inter-tier interference directly. For example, we define the following channel matrix:

$$\mathbf{H}_{\mathrm{ZF}} = [\mathbf{h}^T_{001}, \mathbf{h}^T_{002}, \ldots, \mathbf{h}^T_{00K_{\mathrm{M}}}, \mathbf{h}^T_{011}, \ldots, \mathbf{h}^T_{01K_1}, \ldots, \mathbf{h}^T_{0C1}, \ldots, \mathbf{h}^T_{0CK_C}]^T. \tag{6.19}$$

On this basis, we obtain the ZF precoding as follows:

$$\mathbf{W}_{\mathrm{ZF}} = \mathbf{H}^H_{\mathrm{ZF}} \left(\mathbf{H}_{\mathrm{ZF}}\mathbf{H}^H_{\mathrm{ZF}}\right)^{-1}. \tag{6.20}$$

The precoding vector for the kth MU can be written as:

$$\mathbf{v}_{0k} = \frac{\sqrt{P_{0k}}\mathbf{W}^k_{\mathrm{ZF}}}{\left\|\mathbf{W}^k_{\mathrm{ZF}}\right\|}, \tag{6.21}$$

where $\mathbf{W}^k_{\mathrm{ZF}}$ is the kth column vector of $\mathbf{W}_{\mathrm{ZF}}$, and P_{0k} denotes the transmit power for the kth MU.

However, for ZF precoding, we find that $\mathbf{H}_{\mathrm{ZF}}\mathbf{H}^H_{\mathrm{ZF}}$ is a $(JK_{\mathrm{S}}+K_{\mathrm{M}}) \times (JK_{\mathrm{S}}+K_{\mathrm{M}})$ matrix. Accordingly, we need to inverse a high dimension matrix ($(JK_{\mathrm{S}}+K_{\mathrm{M}}) \times (JK_{\mathrm{S}}+K_{\mathrm{M}})$) in (6.20) with high complexity, especially for a lager JK_{S} (i.e., number of SUs). In contrast, for our proposed CSBD precoding, we only need to inverse a low dimension matrix $\mathbf{H}_{\mathrm{M}}\widetilde{\mathbf{V}}_{\mathrm{in}}\mathbf{H}^H_{\mathrm{M}}$ ($K_{\mathrm{M}} \times K_{\mathrm{M}}$), reducing the computational complexity.

6.5 Non-Cooperative Game-Based Precoding Design for Clustered SCs

The precoding vector of the kth SU in the lth cluster is designed as the product as follows:

$$\mathbf{v}_{lk} = \mathbf{T}_{lk}\mathbf{s}_{lk}, \tag{6.22}$$

where $\mathbf{T}_{lk}$ is used to remove the intra-cluster interference, and $\mathbf{s}_{lk}$ is designed to coordinate the inter-cluster interference for maximizing the downlink sum rate of SUs.

We first define the intra-cluster interference channels of the kth SU in the lth cluster as follows:

$$\mathbf{H}_{lk} = [\mathbf{h}_{ll1}^T, \ldots, \mathbf{h}_{ll(k-1)}^T, \mathbf{h}_{ll(k+1)}^T, \ldots, \mathbf{h}_{llK_l}^T]^T, \tag{6.23}$$

where $\mathbf{H}_{lk} \in \mathbb{C}^{(K_l-1)\times N_l}$. Then, we can obtain the null space of the interference channel matrix $\mathbf{H}_{lk}$ by SVD of $\mathbf{H}_{lk}$. The $\mathbf{h}_{llk}$ is mutually independent for any k and we obtain:

$$\mathbf{H}_{lk} = \mathbf{U}_{lk}\Sigma_{lk}\left[\mathbf{V}_{lk}\widetilde{\mathbf{V}}_{lk}\right]^H, \tag{6.24}$$

where $\widetilde{\mathbf{V}}_{lk} \in \mathbb{C}^{N_l\times(N_l-K_l+1)}$ denotes the orthogonal basis of the null space of $\mathbf{H}_{lk}$, namely $\mathbf{H}_{lk}\widetilde{\mathbf{V}}_{lk} = 0$ and $\widetilde{\mathbf{V}}_{lk}^H\widetilde{\mathbf{V}}_{lk} = \mathbf{I}$. Thus, we have $\mathbf{T}_{lk} = \widetilde{\mathbf{V}}_{lk}$. Following this, the original problem (6.8) can be transformed as:

$$\max_{\{\Phi_{lk}\}} \quad \sum_{l=1}^{C}\sum_{k=1}^{K_l}\log_2\left(1+\frac{\mathbf{h}_{llk}\mathbf{T}_{lk}\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H\mathbf{h}_{llk}^H}{\sum\limits_{i\neq l}^{C}\sum\limits_{j=1}^{K_i}\mathbf{h}_{ilk}\mathbf{T}_{ij}\mathbf{\Phi}_{ij}\mathbf{T}_{ij}^H\mathbf{h}_{ilk}^H+\delta^2}\right) \tag{6.25a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K_l}\mathrm{Tr}(\mathbf{B}_{O_l}\mathbf{T}_{lk}\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H) \leq P, \forall l, O_l, \tag{6.25b}$$

$$\mathrm{rank}\,(\mathbf{\Phi}_{lk}) = 1, \tag{6.25c}$$

where $\mathbf{\Phi}_{lk} = \mathbf{s}_{lk}\mathbf{s}_{lk}^H \in \mathbb{C}^{(N_l-K_l+1)\times(N_l-K_l+1)}$. Here, we consider the single-antenna SU and $\mathbf{s}_{lk}$ is a $(N_l-K_l+1)\times 1$ vector, so we have (6.25c).

Next, we define $\widetilde{\mathbf{h}}_{llk} = \mathbf{h}_{llk}\mathbf{T}_{lk}$, $\widetilde{\mathbf{h}}_{ilj} = \mathbf{h}_{ilk}\mathbf{T}_{ij}$ and the final optimization problem can be written as:

$$\max_{\{\mathbf{\Phi}_{lk}\}} \quad \sum_{l=1}^{C}\sum_{k=1}^{K_l}\widetilde{R}_{lk} \tag{6.26a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K_l}\mathrm{Tr}(\mathbf{B}_{O_l}\mathbf{T}_{lk}\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H) \leq P, \forall l, O_l, \tag{6.26b}$$

$$\mathrm{rank}\,(\mathbf{\Phi}_{lk}) = 1, \tag{6.26c}$$

where $\widetilde{R}_{lk} = \log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\mathbf{\Phi}_{lk}\widetilde{\mathbf{h}}_{llk}^H}{\sum_{i\neq l}^{C}\sum_{j=1}^{K_i}\widetilde{\mathbf{h}}_{ilj}\mathbf{\Phi}_{ij}\widetilde{\mathbf{h}}_{ilj}^H+\delta^2}\right)$. We note here that (6.26) is a non-convex optimization problem due to the non-concave objective function (6.26a) and the rank-one constraint (6.26c). Problem in (6.26) is very difficult to be directly solved even through a centralized algorithm. On this basis, we devise a distributed scheme based on the non-cooperative game. Finally, we prove the existence and uniqueness of the NE for the formulated non-cooperative game and propose an iterative algorithm to obtain NE solution.

6.5.1 The Formulated Non-Cooperative Game Model

With the certain price, the non-cooperative game for the cluster player is defined as:

$$\mathtt{G} = \left\{\mathtt{C}, \{\mathbf{\Phi}_l\}_{l\in\mathtt{C}}, \{U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})\}\right\}, \tag{6.27}$$

where $\mathtt{C} = \{1,\ldots,C\}$ is the set of all clusters; $\mathbf{\Phi}_l = [\mathbf{\Phi}_{l1}^T,\ldots,\mathbf{\Phi}_{lK_l}^T]^T (l \in \mathscr{C})$ denotes the precoding matrix of the lth cluster; $U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})$ is the utility function of the lth cluster; $\mathbf{m} = [m_1,\ldots,m_{\mathrm{C}}]$ denotes the interference price for clusters; $\mathbf{\Phi}_{-l} = [\mathbf{\Phi}_1^T,\ldots,\mathbf{\Phi}_{l-1}^T,\ldots,\mathbf{\Phi}_{l+1}^T,\ldots,\mathbf{\Phi}_{\mathrm{C}}^T]^T$ is precoding matrix of other $(C-1)$ clusters. We define the utility function as follows:

$$\begin{aligned} U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l}) &= \sum_{k=1}^{K_l}\widetilde{R}_{lk} - \sum_{k=1}^{K_l} L_{lk}(\mathbf{\Phi}_{lk}) \\ &= \sum_{k=1}^{K_l}\widetilde{R}_{lk} - \sum_{k=1}^{K_l}\sum_{i\neq l}^{C}\sum_{j=1}^{K_i} m_i \mathbf{h}_{lij}\mathbf{T}_{lk}\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H\mathbf{h}_{lij}^H, \end{aligned} \tag{6.28}$$

where L_{lk} denotes the interference imposed by the precoding vector of the kth SU in the lth cluster to all SUs in other $(C-1)$ clusters.

From (6.28), we find that the second term of the utility function accounts for the cost due to generated interference to other clusters, which discourages the lth cluster from maximizing its own sum rate selfishly. However, when the price vector $\mathbf{m} = \mathbf{0}$, the cluster will maximize its own sum rate uniquely.

Therefore, for the cluster player l $(l \in \mathtt{C})$, we solve the following problem:

$$\max_{\{\mathbf{\Phi}_l\}} \quad U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l}) \tag{6.29a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K_l}\mathrm{Tr}(\mathbf{B}_{O_l}\mathbf{T}_{lk}\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H) \leq P, \forall O_l, \tag{6.29b}$$

$$\mathrm{rank}(\mathbf{\Phi}_{lk}) = 1, \forall k. \tag{6.29c}$$

Problem (6.29) includes the non-convex constraint (6.29c). To overcome the obstacle of non-convexity, we first reformulate the problem without the rank-one constraint and then, the obtained closed-form solution can be guaranteed to be rank-one. Following this, we formulate the optimization problem as follows:

$$\max_{\{\mathbf{\Phi}_l\}} \quad U_l\left(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l}\right) \tag{6.30a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K_l} \text{Tr}(\mathbf{B}_{O_l} \mathbf{T}_{lk} \mathbf{\Phi}_{lk} \mathbf{T}_{lk}^H) \leq P, \forall O_l. \tag{6.30b}$$

Definition of NE: A strategy profile $\mathbf{\Phi} = [\mathbf{\Phi}_1^T, \ldots, \mathbf{\Phi}_C^T]^T$ is a NE, if $\mathbf{\Phi}_l$ is the best response to $\mathbf{\Phi}_{-l}$ for every player l. Formally, the strategy profile $\mathbf{\Phi}_l$ is an NE if

$$U_l\left(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l}\right) \geq U_l\left(\mathbf{m}, \mathbf{\Phi}_l', \mathbf{\Phi}_{-l}\right), l \in \{1, \ldots, C\}, \tag{6.31}$$

where $\mathbf{\Phi}_l'$ is an arbitrary profile of player l in strategy space.

Theorem 6.5.1 *There exists a NE for the non-cooperative game* $\mathtt{G}$ *in (6.27).*

Proof Please refer to Appendix D for proof.

6.5.2 The Solution of the Non-Cooperative Game

Before proving the uniqueness of NE, we first solve the optimization problem (6.30), where a cluster obtains its best response for given other clusters' action. Since we have proved that the objective function in (6.30) is concave and the constraint is a convex set w.r.t. $\mathbf{\Phi}_l$, (6.30) is a convex optimization problem and can be solved using standard convex optimization techniques, e.g., the interior point method [43] and standard determinant maximization (MAXDET) software [58]. However, our interest is to design an algorithm for solving (6.30), which is based on dual method due to the zero gap between problem (6.30) and its dual [43].

The Lagrange dual function of (6.30) is defined as:

$$g(\boldsymbol{\mu}_l) = \max_{\mathbf{\Phi}_l \succeq \mathbf{0}} \quad L\left(\mathbf{\Phi}_l, \boldsymbol{\mu}_l\right), \tag{6.32}$$

where

$$\begin{aligned} L(\boldsymbol{\Phi}_l, \boldsymbol{\mu}_l) &= \sum_{k=1}^{K_l} \left(\log_2 \left(1 + \frac{\widetilde{\mathbf{h}}_{llk} \boldsymbol{\Phi}_{lk} \widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk} + \delta^2} \right) - \sum_{i \neq l}^{C} \sum_{j=1}^{K_i} m_i \mathbf{h}_{lij} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H \mathbf{h}_{lij}^H \right) \\ &\quad + \sum_{O_l=1}^{C_l} \mu_{O_l} \left(P - \sum_{k=1}^{K_l} \mathrm{Tr}(\mathbf{B}_{O_l} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H) \right) \\ &= \sum_{k=1}^{K_l} \left(\log_2 \left(1 + \frac{\widetilde{\mathbf{h}}_{llk} \boldsymbol{\Phi}_{lk} \widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk} + \delta^2} \right) - \sum_{i \neq l}^{C} \sum_{j=1}^{K_i} m_i \mathbf{h}_{lij} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H \mathbf{h}_{lij}^H \right. \\ &\quad \left. - \sum_{O_l=1}^{C_l} \mu_{O_l} \mathrm{Tr}(\mathbf{B}_{O_l} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H) \right) + \sum_{O_l=1}^{C_l} u_{O_l} P, \end{aligned} \tag{6.33}$$

and $\boldsymbol{\mu}_l = [\mu_1, \ldots, \mu_{C_l}]$ a vector of dual variables each associated with one corresponding power constraint given in (6.30b). Accordingly, the dual optimization problem is as follows:

$$\min_{\boldsymbol{\mu}_l \succeq \mathbf{0}} \quad g(\boldsymbol{\mu}_l). \tag{6.34}$$

It is obvious that (6.34) is convex and satisfies the Slater's condition [43], so the dual gap between the optimal objective value of (6.30) and that of (6.34) is zero. Therefore, we can solve (6.34) to obtain the optimal value of (6.30). The subgradient method [43] can be used to minimize $g(\boldsymbol{\mu}_l)$, and the dual variables $\boldsymbol{\mu}_l$ are updated as follows:

$$\mu_{O_l}(n+1) = \left[\mu_{O_l}(n) + \zeta(n) \left(P - \sum_{k=1}^{K_l} \mathrm{Tr}(\mathbf{B}_{O_l} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H) \right) \right]^+, \tag{6.35}$$

where $\zeta(n)$ is the diminishing step size, and n is the iterative index.

In addition, solving dual problem (6.34) involves determining the optimal $\boldsymbol{\Phi}_l$ at given dual variables $\boldsymbol{\mu}_l$. Next, we focus on solving $\boldsymbol{\Phi}_l$ for fixed $\boldsymbol{\mu}_l$. We find that problem (6.32) can be divided into K_l independent subproblems and each only involves $\boldsymbol{\Phi}_{lk}$. Since $\sum_{O_l=1}^{C_l} u_{O_l} P$ is a constant for fixed $\boldsymbol{\mu}_l$, for the kth SU in the lth cluster, the corresponding subproblem can be expressed as:

$$\begin{aligned} \max_{\boldsymbol{\Phi}_{lk} \succeq \mathbf{0}} \quad & \log_2 \left(1 + \frac{\widetilde{\mathbf{h}}_{llk} \boldsymbol{\Phi}_{lk} \widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk} + \delta^2} \right) - \sum_{i \neq l}^{C} \sum_{j=1}^{K_i} m_i \mathbf{h}_{lij} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H \mathbf{h}_{lij}^H \\ & - \mathrm{Tr}(\mathbf{B}_{\mu} \mathbf{T}_{lk} \boldsymbol{\Phi}_{lk} \mathbf{T}_{lk}^H), \end{aligned} \tag{6.36}$$

where $\mathbf{B}_{\mu} = \sum_{O_l=1}^{C_l} \mu_{O_l} \mathbf{B}_{O_l}$.

Next, we define the objective function of (6.36) as $\mathscr{L}(\mathbf{\Phi}_{lk})$ and then, we obatin:

$$\begin{aligned}\mathscr{L}(\mathbf{\Phi}_{lk}) &= \log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\mathbf{\Phi}_{lk}\widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk}+\delta^2}\right)-\sum_{i\neq l}^{C}\sum_{j=1}^{K_i} m_i \mathrm{Tr}(\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H\mathbf{h}_{lij}^H\mathbf{h}_{lij}\mathbf{T}_{lk})-\mathrm{Tr}(\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H\mathbf{B}_{\mu}\mathbf{T}_{lk}) \\ &= \log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\mathbf{\Phi}_{lk}\widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk}+\delta^2}\right)-\mathrm{Tr}\left(\mathbf{\Phi}_{lk}\left(\sum_{i\neq l}^{C}\sum_{j=1}^{K_i} m_i\mathbf{T}_{lk}^H\mathbf{h}_{lij}^H\mathbf{h}_{lij}\mathbf{T}_{lk}\right)\right)-\mathrm{Tr}(\mathbf{\Phi}_{lk}\mathbf{T}_{lk}^H\mathbf{B}_{\mu}\mathbf{T}_{lk}) \\ &= \log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\mathbf{\Phi}_{lk}\widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk}+\delta^2}\right)-\mathrm{Tr}(\mathbf{Z}_{lk}\mathbf{\Phi}_{lk}),\end{aligned} \tag{6.37}$$

where $\mathbf{Z}_{lk}=\sum\limits_{i\neq l}^{C}\sum\limits_{j=1}^{K_i} m_i\mathbf{T}_{lk}^H\mathbf{h}_{lij}^H\mathbf{h}_{lij}\mathbf{T}_{lk}+\mathbf{T}_{lk}^H\mathbf{B}_{\mu}\mathbf{T}_{lk}$ and $\mathbf{Z}_{lk}\in\mathbb{C}^{(N_l-K_l+1)\times(N_l-K_l+1)}$. Meanwhile, for the above derivation, some equations are used, such as $\mathrm{Tr}(\mathbf{XY})=\mathrm{Tr}(\mathbf{YX})$ and $a\mathrm{Tr}(\mathbf{X})+b\mathrm{Tr}(\mathbf{Y})=\mathrm{Tr}(a\mathbf{X}+b\mathbf{Y})$. We then have the following theorem.

Theorem 6.5.2 *For the problem in (6.36) to have a bounded objective value, matrix* $\mathbf{Z}_{lk}$ *should be positive definite.*

Proof Please refer to Appendix E for proof.

We rewrite $\mathbf{\Phi}_{lk}$ in its original form, i.e., $\mathbf{\Phi}_{lk}=\mathbf{s}_{lk}\mathbf{s}_{lk}^H$. Accordingly, the problem (6.36) can be transformed as follows:

$$\max_{\mathbf{s}_{lk}\succeq\mathbf{0}}\log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\mathbf{s}_{lk}\mathbf{s}_{lk}^H\widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk}+\delta^2}\right)-\mathbf{s}_{lk}^H\mathbf{Z}_{lk}\mathbf{s}_{lk}. \tag{6.38}$$

According to Cholesky decomposition [59], we have $\mathbf{Z}_{lk}=\hat{\mathbf{Z}}_{lk}\hat{\mathbf{Z}}_{lk}^H$ and $\hat{\mathbf{Z}}_{lk}$ is reversible due to $\mathbf{Z}_{lk}$ is positive definite. We define $\hat{\mathbf{s}}_{lk}=\mathbf{s}_{lk}^H\hat{\mathbf{Z}}_{lk}$ and then, (6.38) can be rewritten as follows:

$$\max_{\hat{\mathbf{s}}_{lk}\succeq\mathbf{0}}\log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\hat{\mathbf{Z}}_{lk}^{-H}\hat{\mathbf{s}}_{lk}^H\hat{\mathbf{s}}_{lk}\hat{\mathbf{Z}}_{lk}^{-1}\widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk}+\delta^2}\right)-\hat{\mathbf{s}}_{lk}\hat{\mathbf{s}}_{lk}^H. \tag{6.39}$$

Next, we define $\mathbf{a}_{lk}=\widetilde{\mathbf{h}}_{llk}\hat{\mathbf{Z}}_{lk}^{-H}/||\widetilde{\mathbf{h}}_{llk}\hat{\mathbf{Z}}_{lk}^{-H}||$. It is clear that the optimal precoding $\hat{\mathbf{s}}_{lk}$ has the same direction with $\mathbf{a}_{lk}$, i.e., $\hat{\mathbf{s}}_{lk}=\sqrt{p_{lk}}\mathbf{a}_{lk}$, where p_{lk} needs to be optimized to maximize (6.39). Based on this, substituting $\hat{\mathbf{s}}_{lk}=\sqrt{p_{lk}}\mathbf{a}_{lk}$ into (6.39), we get the following optimization problem:

$$\max_{p_{lk}\succeq\mathbf{0}}\log_2\left(1+\frac{p_{lk}\alpha_{lk}}{\Xi_{lk}+\delta^2}\right)-p_{lk}. \tag{6.40}$$

Algorithm 7: Non-Cooperative Game-Based Precoding Design For Each Cluster

Form multiple SBS clusters according to Algorithm 6.
Given the price vector $\mathbf{m}$, initialize the feasible precoding $\mathbf{\Phi}^{(0)} = [\mathbf{\Phi}_1^{(0)}, \ldots, \mathbf{\Phi}_C^{(0)}]$, set a counter $n = 1$.
repeat
    Design precoding for each cluster $l \in \{1, \ldots, C\}$.
    Initialize $\boldsymbol{\mu}_l$.
    repeat
        Compute $\mathbf{\Phi}_{lk}(k = 1, \ldots, K_l)$ according to (6.42).
        Update dual varies according to (6.35).
    until $\boldsymbol{\mu}_l$ converges;
    Obtain $\mathbf{\Phi}^{(n)}$.
    Update $n \leftarrow n + 1$.
until $\|\mathbf{\Phi}^{(n)} - \mathbf{\Phi}^{(n-1)}\| \leq \xi$, for some prescribed ξ;
Obtain the optimal precoding $\mathbf{\Phi}^{(n)}$.

where $\alpha_{lk} = ||\widetilde{\mathbf{h}}_{llk}\hat{\mathbf{Z}}_{lk}^{-H}||^2$, and optimal p_{lk} can be obtained by the standard water-filling algorithm [43]:

$$p_{lk} = \left(\frac{1}{\ln 2} - \frac{\Xi_{lk} + \delta^2}{\alpha_{lk}}\right)^+. \tag{6.41}$$

Finally, we obtain the precoding as follows:

$$\mathbf{s}_{lk} = \sqrt{p_{lk}}(\mathbf{a}_{lk}\hat{\mathbf{Z}}_{lk}^{-1})^H, \text{and } \mathbf{\Phi}_{lk} = p_{lk}(\mathbf{a}_{lk}\hat{\mathbf{Z}}_{lk}^{-1})^H\mathbf{a}_{lk}\hat{\mathbf{Z}}_{lk}^{-1}. \tag{6.42}$$

From (6.42), it is clear that $\mathbf{\Phi}_{lk}$ is a rank-one solution. Therefore, the solution of relaxed problem (6.30) is also the solution of the original problem (6.29) via our proposed algorithm. Based on the above results, we define the following theorem:

Theorem 6.5.3 *There exists an unique NE for the non-cooperative game* G *in (6.27).*

Proof Please refer to Appendix F for proof.

6.5.3 NE Searching Algorithm

For SC clustering, the SBSs send SUs' CSI to MBS through high speed backhaul links. Next, the MBS executes Algorithm 6 according to a predefined interference threshold and then, communicates the final clustering information with an interference price of each cluster

Table 6.1 Simulation Parameters.

Parameters	Value
Radius of MC	500 m
Radius of SC	40 m
Number of MUs	20
Number of SCs	20
Number of SUs each SC	2
Number of SBS antennas	2
Number of MBS antennas	500
Transmit power of MBS	46 dBm
Maximum transmit power of SBS	30 dBm
Pathloss between MBS and MU or SU	$27.3+39.1\log 10(d)$
Pathloss between SBS and MU or SU	$36.8+36.7\log 10(d)$
Downlink bandwidth	10 MHz
Noise power	-174 dBm/Hz

to SBSs through backhaul links. After forming the clusters, each cluster sets the initial feasible precoding (we assume that one of SBSs takes charge of the precoding design and denoted as SBS header). Since the SBS header has obtained all CSI through sharing among SBSs belonging to the same cluster, the precoding can be computed when SUs send back the received interference to SBS header. The SBSH will update precoding when SUs send back the updated interference, and the process is executed until convergence. We assume that the update of the interference price and precoding strategy among clusters is ideal synchronous. According to the above analysis, we find that information exchange is not needed among clusters. We summarize the distributed precoding design scheme for each cluster as Algorithm 7.

6.6 Numerical Results and Discussions

In this section, we provide numerical results to evaluate the performance of our proposed schemes. We consider a single MC with a radius of 500 meters, where the MBS is located at the center of the MC and MUs are randomly distributed in the MC. We assume that the MC-hole radius is 100 meters (all MUs and SUs do not locate in this area). We assume that the radius of each SC is 40 meters, where all SCs are randomly located within the MC but their coverage are not overlapped each other. The minimum distance between SUs and SBS

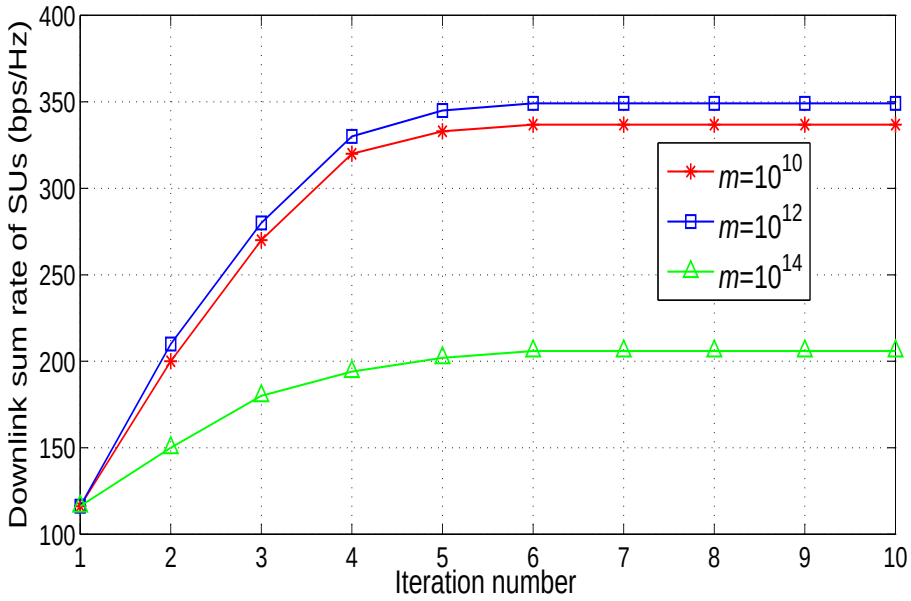


Fig. 6.5 Downlink sum rate of SUs versus iteration number.

is 5 meters. We assume that all prices for different clusters are the same for simplicity. Other related simulation parameters are listed in Table 6.1.

Fig. 6.5 shows the convergent speed under different interference prices, where $\gamma_{th} = -110$ dB. It is clearly found that the downlink sum rate of SUs is maximized after 6 iterations. The sum rate under different interference prices is different. To clearly analyze the impact of the interference price on the performance of the system, we plot Fig. 6.6 to shown the downlink sum rate of SUs with the interference price m. It can be observed that the sum rate first increases and then decreases with m. In other words, there exits an optimal m for maximizing the sum rate under a certain γ_{th}. Therefore, we can use some simple methods, e.g., one-dimension search, to find the optimal m for obtaining the maximum sum rate. Meanwhile, we can find that the sum rate increases with γ_{th} decreases. This is because more larger-size clusters are formed so that more interference are cancelled.

Fig. 6.7 shows that the downlink sum rate versus interference threshold γ_{th}. Here, the one-dimension search is used to obtain the maximum sum rate under each interference threshold. For a low γ_{th}, i.e., $\gamma_{th} = -120$ dB, all SCs form one cluster. In this case, the interference among SUs will be cancelled completely, so that the sum rate is maximized. Number of clusters increases as γ_{th} increases, while the sum rate decreases. This is because the interference increases among SUs due to the increase of the number of clusters. When γ_{th} is between -92 dB and -90 dB, there are no pairs of SCs to form the same cluster, resulting in the lowest sum rate.

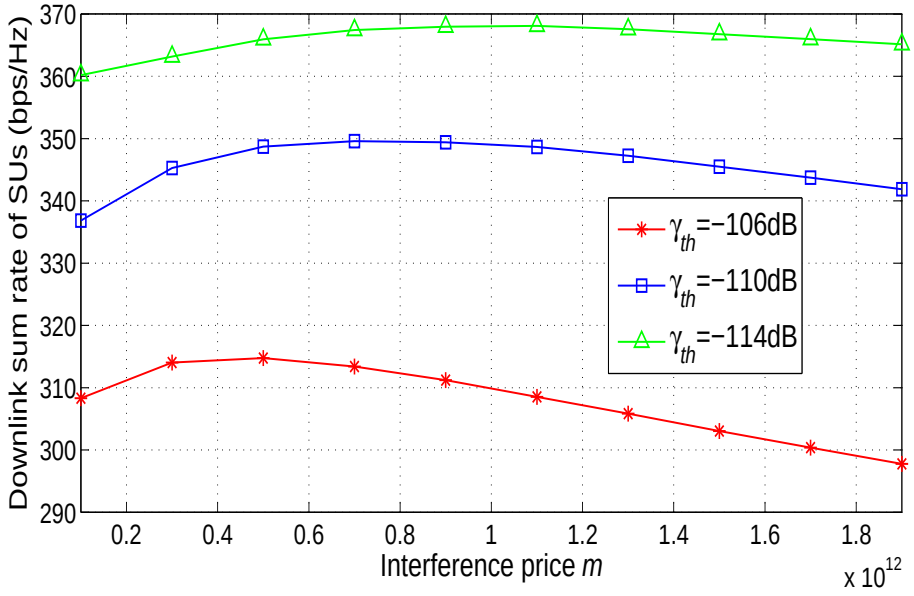


Fig. 6.6 Downlink sum rate of SUs versus interference price.

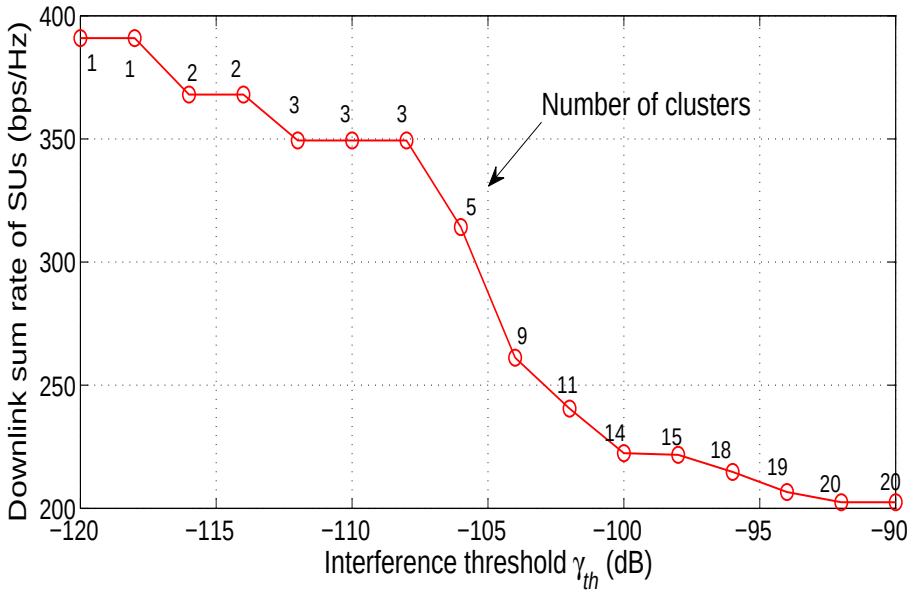


Fig. 6.7 Downlink sum rate of SUs versus interference threshold.

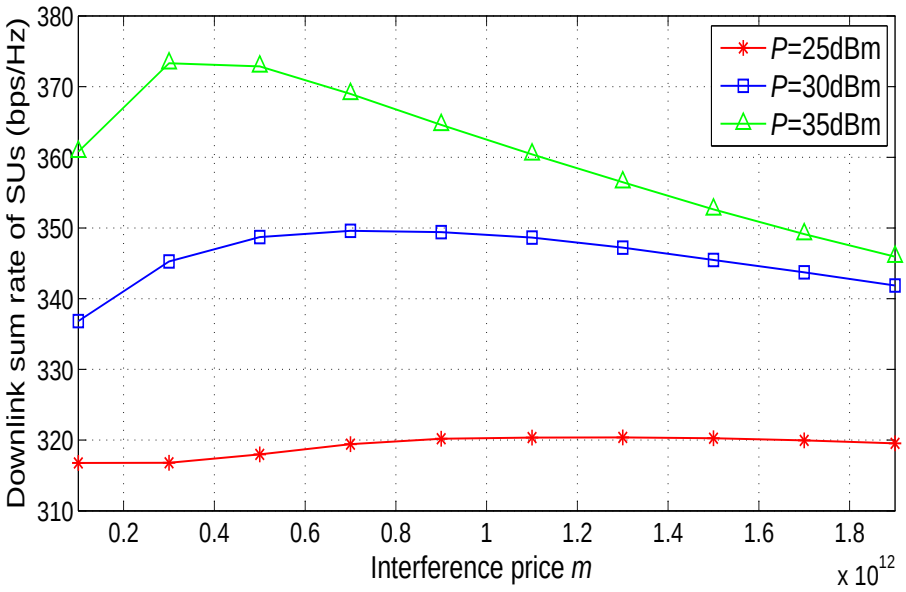


Fig. 6.8 Downlink sum rate of SUs versus interference price.

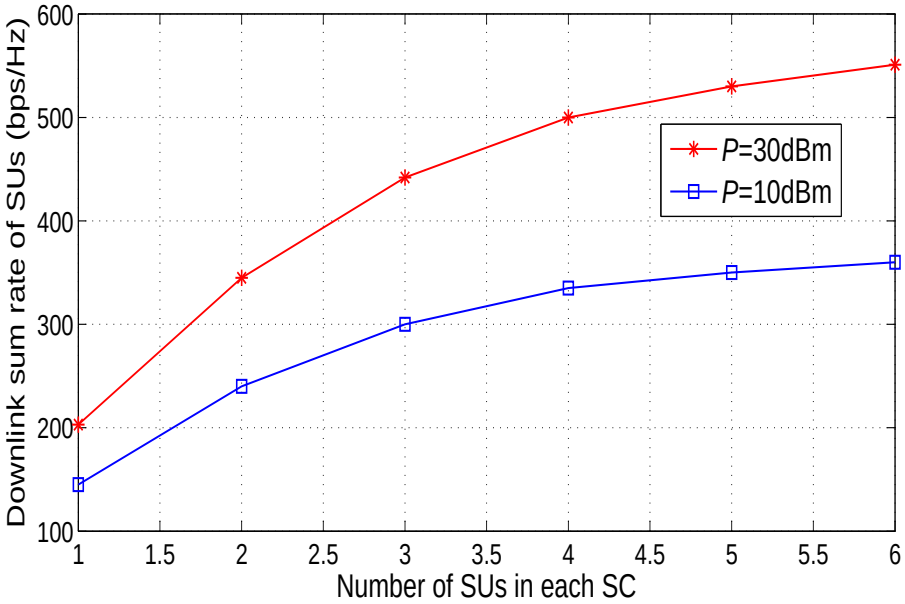


Fig. 6.9 Downlink sum rate of SUs versus number of SUs.

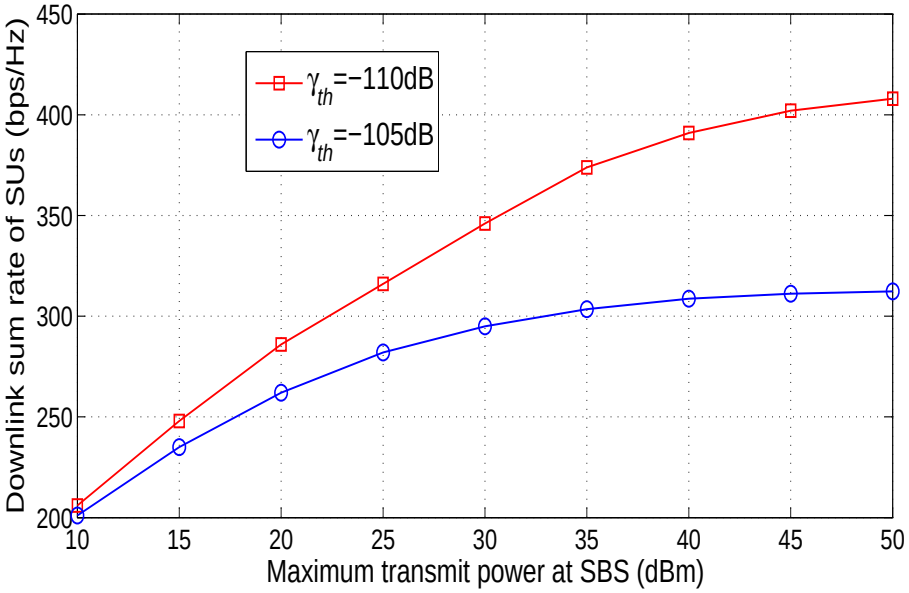


Fig. 6.10 Downlink sum rate of SUs versus transmit power.

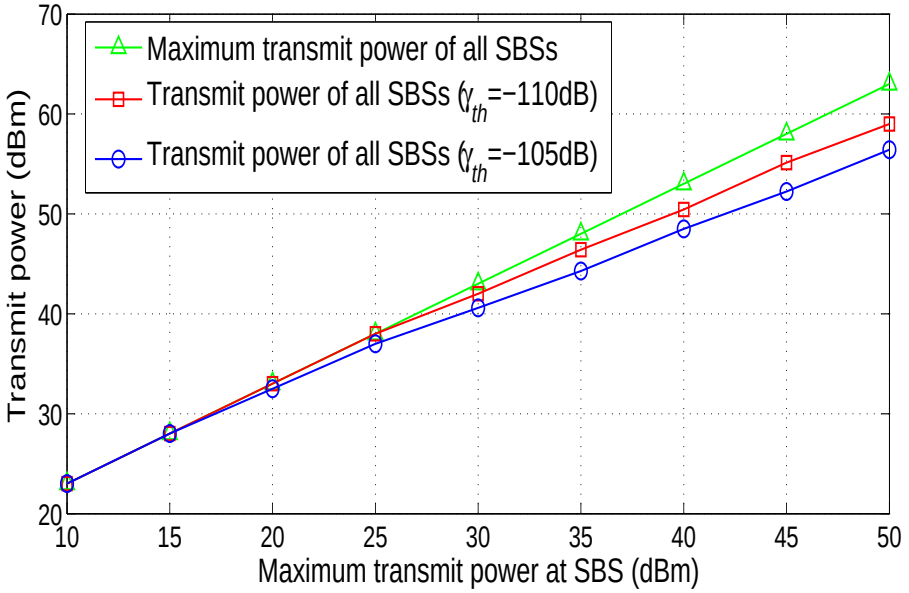


Fig. 6.11 Transmit power versus available transmit power at each SBS.

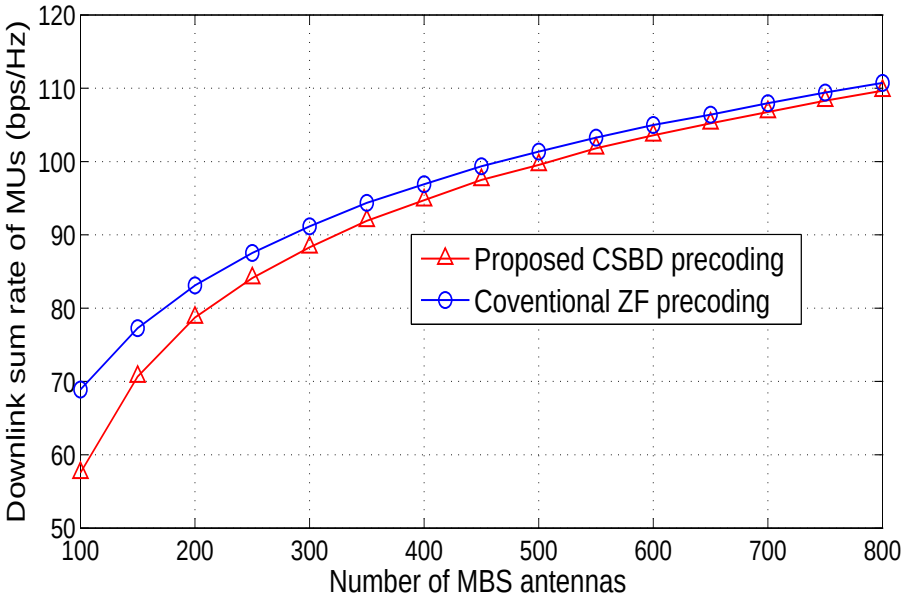


Fig. 6.12 Downlink sum rate of MUs versus number of MBS antenas.

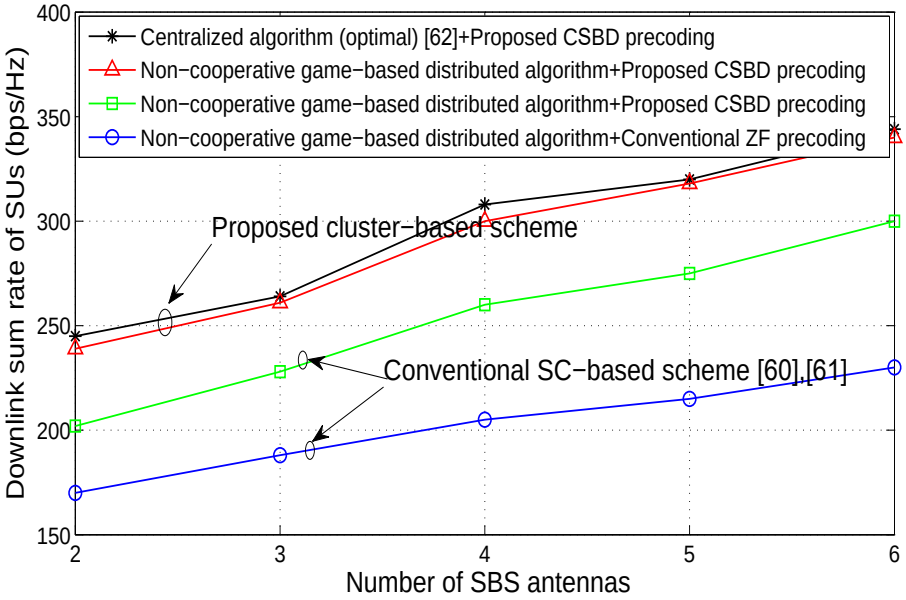


Fig. 6.13 Downlink sum rate of MUs versus number of SBS antenas.

Fig. 6.8 plots that the downlink sum rate versus m under different per-SBS transmit power, where $\gamma_{th} = -110$ dB. When the maximum transmit power of each SBS is high, as shown in Fig. 6.6, the sum rate first increases and then decreases with m. However, when the maximum transmit power of each SBS is low, the impact of m on the sum rate is slow. This is because for the former, each SBS is not allowed to fully transmit such large power for avoiding serious interference among clusters. In the contrary, for the later, each SBS can almost transmit overall power for improving sum rate.

Fig. 6.9 illustrates the downlink sum rate of SUs versus number of SUs in each SC, where $\gamma_{th} = -104$ dB and $N = 6$. It can be observed that the sum rate increases with the number of SUs, while the increased ratio is reduced. It is easy to understand that more SUs lead to higher gain gaps and improve the sum rate. However, the increased sum rate is limited because the total transmit power is a constant.

Figs. 6.10 and 6.11 show the downlink sum rate and transmit power versus the maximum transmit power of each SBS, respectively. It is easy to understand that the sum rate increases with the maximum transmit power. However, the increase of the sum rate is limited due to interference among clusters, especially for a lower γ_{th}. The detailed reason can be found in Fig. 6.11. When the maximum transmit power of each SBS is low, the power will be fully transmitted due to weak interference among clusters. However, as the maximum transmit power increases, to avoid serious interference, the transmit power at each SBS is not allowed such high.

Fig. 6.12 shows that the downlink sum rate of MUs versus number of MBS antennas. Here, “Conventional ZF precoding" denotes that only the ZF precoding is applied at MBS to eliminate the multi-MU interference, while the proposed CSBD precoding scheme works to eliminate the inter-tier and multi-MU interference simultaneously. From this figure, we can find that the downlink sum rate of MUs with proposed CSBD precoding is always lower than that with ZF precoding, but the sum-rate gap is slight when the number of MBS antennas is larger. The reason is that when the proposed CSBD precoding is applied at MBS, the MBS needs to sacrifice some DoFs to eliminate SUs’ interference, which results in the decrease in the sum rate of MUs. However, for a large antennas at MBS, it is suitable to apply the proposed CSBD precoding for eliminating SUs’ interference, which can effectively improve the downlink sum rate of SUs (as shown in Fig. 6.13). Meanwhile, the impact on the downlink sum rate of MUs is slight.

In Fig 6.13, we plot that the downlink sum rate of SUs versus number of SBS antennas, where $\gamma_{th} = -104$ dB. We compare the performance gain with the conventional SC-based non-cooperative game scheme, e.g., [60, 61]. It is clear that sum rate of SUs with our proposed cluster-based scheme is higher than that with conventional SC-based scheme,

which can be improved by about 40%. It means that it is necessary to form multiple clusters for interference cancellation, especially for ultra-dense SCs. In addition, to compare with the performance loss with optimal algorithm, we apply the centralized algorithm proposed in [62] to obtain the optimal solution. We can find the rate gap is slight between optimal centralized algorithm and our proposed distributed algorithm. Meanwhile, it can be observed that the proposed CSBD effectively reduces the inter-tier interference and improves the sum rate of the SUs.

In addition, we note here that a theoretical bound of the price-of-anarchy (i.e., the optimal solution) is very difficult if not impossible to be derived. This is because for the optimal algorithm [62] or our proposed suboptimal algorithm, the closed-form expression is a function of iteratively updated parameters such as Lagrange dual variables until convergence to reach a stable and optimal solution. Furthermore, the closed-form expression for optimal or suboptimal algorithm includes Lagrange dual variables with uncertain range, which additionally makes the bound derivation of price-of-anarchy very difficult if not impossible at all.

6.7 Conclusions

We have investigated the SC clustering and precoding design problems for the mMIMO-SC HetNet. An interference graph-based dynamic SC clustering scheme has been proposed in order to cooperative transmission among SBSs belonging to the same cluster. We designed precoding schemes for MBS to eliminate the inter-tier and multi-MU interference. Then, we presented the precoding design at clustered SCs as an optimization problem to maximize downlink sum rate of SUs under per-SBS power constraint. A non-cooperative game-based distributed algorithm was proposed to obtain a suboptimal solution. Simulation results show that our proposed cluster-based scheme can improve the sum rate of SUs by about 40% in comparison with the conventional SC-based scheme.

Chapter 7

Conclusions and Future Works

7.1 Conclusions

In this thesis, we introduced the recent development and some key techniques of the wireless communications, e.g., mMIMO, HetNet. We analyzed the potential problems and main challenges in mMIMO, mMIMO-CR and mMIMO-CR HetNet, e.g., pilot contamination and interference management. Then, we proposed effective pilot allocation and power allocation schemes to relieve the pilot contamination and coordinate the interference.

We divided our research into three parts. The first part (i.e., chapter 2) presented the basic pilot contamination problem in a mMIMO homogeneous network. Based on this, we proposed a low-complexity pilot allocation scheme to maximize the uplink rate of the system. Meanwhile, to improve the users' fairness, a fairness-based pilot allocation scheme was proposed. Simulation results showed that the our proposed scheme can improve per-user rate by about 17% in comparison with the conventional scheme.

The second part (i.e., chapters 3 and 4) solved the pilot and power allocation problems in mMIMO-CR HetNet. In chapter 3, we proposed a pilot allocation scheme to obtain a win-win paradigm between PN and CN. The results showed that the PN and CN can obtain positive revenue, which implies that pilot sharing concept between PN and CN is effective in improving the performance of both PN and CN. Then, we investigated the power allocation problem in chapter 4. Based on this, we formulated the power allocation optimization problem of the CN to maximize the downlink sum rate of the CN subject to the total transmit power and PUs' SINR constraints. An iterative power allocation algorithm was proposed. The numerical results presented that our proposed scheme can improve the sum rate of the CN by about 10% in comparison with the conventional scheme.

The third part (i.e., chapters 5 and 6) mainly focused on pilot allocation and interference coordination problems in mMIMO-SC HetNet. In chapter 5, we investigated pilot allocation

problem. We proposed a new pilot allocation scheme for maximizing ergodic downlink sum rate of the system. In addition, we proposed two suboptimal pilot allocation algorithms to simplify the computational process and improve SUs' fairness, respectively. Simulation results showed that our proposed scheme can improve the sum rate of the system by about 12% in comparison with the conventional scheme. To consider the interference coordination, we investigated the dynamic SC clustering strategy and their precoding design problem in chapter 6. An interference graph-based dynamic SC clustering scheme was proposed. Based on this, we formulated an optimization problem as design precoding weights at MBS and clustered SCs for maximizing the downlink sum rate of SUs subject to the power constraint of each SBS, while mitigating inter-cluster, eliminating inter-tier, intra-cluster and multi-MU interference. A non-cooperative game-based distributed algorithm was proposed. Simulation results showed that our proposed scheme can improve the sum rate of SUs by about 40% in comparison with the conventional SC-based scheme.

Summarily, in this thesis, we studied the pilot allocation and interference coordination problems in two types of mMIMO-HetNets, including mMIMO-CR HetNet and mMIMO-SC HetNet. According to the simulation results, our proposed schemes effectively improve the capacities of the system.

7.2 Future Works

Just as our analysis in this thesis, mMIMO-SC HetNet has been considered as a promising technique to meet the requirements of explosive data capacity in future wireless communications. However, the interference coordination to deal with inter- and intra-tier interference is still an important technical challenge. One of the effective interference coordination techniques is cooperative transmission and reception called as coordinated multipoint process (CoMP). To realize cooperative transmission among SBSs, the centralized network architectures have been investigated, under the assumption that CSI and the signaling information are shared among SCs through a high-speed backhaul link. However, since wireless data rate is increased especially when massive antennas are used, it increases the required amount of data to be shared which may need higher backhaul capacity.

The heterogeneous centralized radio access network (H-CRAN) is an emerging network architecture to realize the SC concept, where remote-radio heads (RRHs) are distributed to provide high rates for SUs while the MBS provides the seamless coverage. In this architecture, the centralized baseband processing unit (BBU) performs the baseband signal processing and resource allocation optimization while remote radio units called as RRH are distributed in service area. However, the full-scale coordination in a large-scale H-CRAN

requires very large channel matrices to perform cooperative processing, which leads to high computational complexity. Therefore, RRH cluster -based partial-scale coordination is an effective scheme to decrease the required complexity. In addition, fronthaul links are needed to share the required information for cooperative processing in H-CRAN. In this case, limited capacity provided by fronthaul links has to be considered. Therefore, we will investigate the resource allocation and interference management in a RRH cluster-based H-CRAN with limited fronthaul capacity.

References

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper", Mar. 2017.

[2] Cisco, "Visual Networking Index," White paper, Feb. 2015 [Online]. Available: www.Cisco.com.

[3] T. S. Rappaport, W. Roh and K. Cheun, "Wireless engineers long considered high frequencies worthless for cellular systems. They couldn't be more wrong," *IEEE Spectr.*, vol. 51, no. 9, pp. 34–58, Sep. 2014.

[4] B. Lars T., S. Andreas, P. Pascal and S. Daniel M, MIMO Power Line Communications: Narrow and Broadband Standards, EMC, and Advanced Processing. Devices, Circuits, and Systems. CRC Press, Feb. 2014.

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Top. Signal. Process.*, vol. 8, no. 5, pp. 742-758, Oct. 2014.

[6] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[7] F. C. Commission, "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," NPRM & Order ET Docket No. 03-108, FCC 03-322, Tech. Rep., Dec. 30, 2003.

[8] Y. Tachwali, B. F. Lo, I. F. Akyildiz and R. Agusti, "Multiuser resource allocation optimization using bandwidth-power product in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 451-463, Mar. 2013.

[9] Y. H. Zhang and C. Leung, "Resource allocation in an OFDM-based cognitive radio system," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1928-1931, Jul. 2009.

[10] F. Fernandes, A. Ashikhmin and T. L. Marzetta, "Inter-Cell Interference in Noncooperative TDD Large Scale Antenna Systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192-201, Feb. 2013.

[11] X. Zhu, Z. Wang, L. Dai and C. Qian, "Smart pilot assignment for massive MIMO," *IEEE Commun. Letters*, vol.19, no.9, pp.1644-1647, Sep. 2015.

[12] E. Bjornson, E. Larsson and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293-1308, Feb. 2016.

[13] I. Atzeni, J. Arnau and M. Debbah, "Fractional pilot reuse in massive MIMO systems," *in Proc. IEEE Conf. Commun. Workshop. (ICCW)*, pp.1030-1035, Jun. 2015.

[14] X. Zhu, L. Dai and Z. Wang, "Graph coloring based pilot allocation to mitigate pilot contamination for multi-cell massive MIMO systems," *IEEE Commun. Letters*, vol.19, no.10, pp.1842-1845, Oct. 2015.

[15] R. Mochaourab, E. Bjornson and M. Bengtsson, "Pilot clustering in asymmetric massive MIMO networks," *in Proc. IEEE Int. Workshop Signal Process. Advances in Wireless Commun. (SPAWC)*, Jun. 28-Jul. 1, 2015.

[16] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval research logistics quarterly*, pp.83-97, 1955.

[17] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[18] P. Hahn, T. Grant and N. Hall, "A branch-and-bound algorithm for the quadratic assignment problem based on the Hungarian method," *European Journal of Operational Research*, vol.108, no.3, pp.629-640, Aug. 1998.

[19] T. Nguyen, V. Ha and L. Le, "Resource allocation optimization in multi-user multi-cell massive MIMO networks considering pilot contamination," *IEEE Access*, vol. 3, pp.1272-1287, 2015.

[20] T. Seyama, D. Jitsukawa, T. Kobayashi, et al., "Study of coordinated radio resource scheduling algorithm for 5G ultra high-density distributed antenna systems," *IEICE Technical Report of RCS*, vol. 115, no, 472, pp. 181-186, 2015.

[21] L. Sboui, Z. Rezki and M.-S. Alouini, "A unified framework for the ergodic capacity of spectrum sharing cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 877-887, Feb. 2013.

[22] M. Filippou, D. Gesbert and H. Yin, "Decontaminating pilots in cognitive massive MIMO networks," *in Proc. Int. Symp. Wireless Commun. Syst.*, pp. 816-820, Aug. 2012.

[23] B. Kouassi, I. Ghauri and L. Deneire, "Reciprocity-based cognitive transmissions using a MU massive MIMO approach," *in Proc. IEEE ICC*, pp. 2738-2742, 2013.

[24] G. Scutari and D. P. Palomar, "MIMO cognitive Radio: A game theoretical approach," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 761-780, Feb. 2010.

[25] F. Moghimi, R. K. Mallik, and R. Schober, "Sensing time and power optimization in MIMO cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3398-3408, Sep. 2012.

[26] L. Fu, Y. J. A. Zhang and J. Huang, "Energy efficient transmissions in MIMO cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2420-2431, Nov. 2013.

[27] L. Zhang, Y. Xin and Y.-C. Liang, "Weighted sum rate optimization for cognitive radio MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2950-2959, Jun. 2009.

[28] L. Wang, H. Q. Ngo, M. Elkashlan, T. Q. Duong and K. Wong, "Massive MIMO in spectrum sharing networks: achievable rate and power efficiency," *IEEE Systems Journal*, vol. 11, no. 1, pp. 20-31, Mar. 2017.

[29] M. Filippou, D. Gesbert and H. Yin, "Decontaminating pilots in cognitive massive MIMO networks," *in Proc. Int. Symp. Wireless Commun. Syst.*, pp. 816-820, Aug. 2012.

[30] B. Kouassi, I. Ghauri and L. Deneire, "Reciprocity-based cognitive transmissions using a MU massive MIMO approach," *in Proc. IEEE ICC*, pp. 2738-2742, 2013.

[31] H. Xie, B. Wang, F. Gao and S. Jin, "A full-space spectrum-sharing strategy for massive MIMO cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol.34, no.10, pp.2537-2549, Oct. 2016.

[32] W. Wang, G. Yu and A. Huang, "Cognitive radio enhanced interference coordination for femtocell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 37-43, Jun. 2013.

[33] F. A. Khan, C. Masouros and T. Ratnarajah, "Interference-driven linear precoding in multiuser MISO downlink cognitive radio network," *IEEE Trans. Veh. Techol.*, vol. 61, no. 6, pp. 2531-2543, Jul. 2012.

[34] J. H. Noh and S. J. Oh, "Beamforming in a multi-user cognitive radio system with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 616-625, Feb. 2013.

[35] H. F Yin, D. Gesbert, M. Filippou and Y. Z Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264-273, Feb. 2013.

[36] Z. Mokhtari, M. Sabbaghian and R. Dinis, "Massive MIMO downlink based on single carrier frequency domain processing," *IEEE Trans. Commun.*, vol. PP, no.99, pp.1-1, 2017.

[37] W. B. Dang, M. X. Tao, H. Mu and J. W. Huang, "Subcarrier-pair based resource allocation for cooperative multi-relay OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1640-1649, May 2010.

[38] R. F. Fan and H. Jiang, "Optimal multi-channel cooperative sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1128-1138, Mar. 2010.

[39] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8 no. 4, pp. 2112-2120, Apr. 2009.

[40] S. Cui, A. J. Goldsmith and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 2349-2360, Apr. 2005.

[41] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory.*, vol. 55, no. 8, pp. 3711-3724, Aug. 2009.

[42] X. Kang, "Optimal power allocation for Bi-directional cognitive radio networks with fading channels," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 567-570, Oct. 2013.

[43] S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.

[44] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for non-convex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681-683, Jul. 1978.

[45] D. Bethanabhotla, O. Y. Bursalioglu, H. C. Papadopoulos and G. Caire, "Optimal user-cell association for massive MIMO wireless networks," *IEEE Trans. Wireless Commun.,* vol. 15, no. 3, pp. 1835-1850, Mar. 2016.

[46] Y. Xu and S. Mao, "User association in massive MIMO HetNets," *IEEE Systems Journal*, vol. 11, no. 1, pp. 7-19, Mar. 2017.

[47] G. Xu, A. Liu, W. Jiang, H. Xiang and W. Luo, "Energy-efficient beamforming for two-tier massive MIMO downlink," *China Commun.*, vol. 12, no. 10, pp. 64-75, Oct. 2015.

[48] Y. Liu, L. Lu, G. Li and Q. Cui, "Performance analysis and interference cancellation for heterogeneous network with massive MIMO," *IEEE GlobalSIP*, Orlando, FL, pp. 888-892, 2015.

[49] L. Zhou et al., "A dynamic graph-based scheduling and interference coordination approach in heterogeneous cellular networks," *IEEE Trans. Vehic. Tech.*, vol. 65, no. 5, pp. 3735-3748, May 2016.

[50] R. Seno, T. Ohtsuki, W. Jiang and Y. Takatori, "Complexity reduction of pico cell clustering for interference alignment in heterogeneous networks" *IEEE APCC*, pp. 267-271, 2015.

[51] M. Hong, R. Sun, H. Baligh and Z. Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 226-240, Feb. 2013.

[52] A. Papadogiannis, D. Gesbert and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," *IEEE ICC,* Beijing, China, pp. 4033-4037, 2008.

[53] S. Fan, J. Zheng and J. Xiao, "A clustering-based downlink resource allocation algorithm for small cell networks," 2015 *International Conference on Wireless Communications and Signal Processing (WCSP)*, Nanjing, China, pp. 1-5, 2015.

[54] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1435-1445, Dec. 2010.

[55] A. Wiesel, Y. C. Eldar and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409-4418, Sep. 2008.

[56] L. N. Tran, M. Juntti, M. Bengtsson and B. Ottersten, "Weighted sum rate maximization for MIMO broadcast channels using dirty paper coding and zero-forcing methods," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2362-2373, Jun. 2013.

[57] Q. Niu, Z. Zeng, T. Zhang, Q. Gao and S. Sun, "Joint interference alignment and power allocation in heterogeneous networks," *IEEE PIMRC*, pp.733-737, 2014.

[58] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," *IEEE Int. Symp. Computer-Aided Control Systems Design (CACSD)*, Taipei, Taiwan, R.O.C., pp. 284-289, Sep. 2004.

[59] J. H. Wilkinson, "The Algebraic Eigenvalue Problem." Oxford, U.K.: Clarendon Press, 1965.

[60] S. Fu, B. Wu, H. Wen, P. H. Ho and G. Feng, "Transmission scheduling and game theoretical power allocation for interference coordination in CoMP," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 112-123, Jan. 2014.

[61] A. Y. Al-Zahrani and F. R. Yu, "A game theory approach for inter-cell interference management in OFDM networks," *IEEE ICC*, Kyoto, pp. 1-5, 2011.

[62] S. He, Y. Huang, S. Jin, and L. Yang, "Coordinated beamforming for energy efficient transmission in multicell multiuser systems," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4961-4971, Dec. 2013.

[63] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

[64] D. Fudenberg and J. Tirole, "Efficient power control via pricing in wireless data networks," Game Theory, MIT Press, 1991.

[65] R. Yates, "A framework for uplink power control in cellular radio system," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341-1347, Sep. 1995.

Appendix A

Firstly, we have

$$\begin{aligned}\mathbb{E}\left\{\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\}&=\mathbb{E}\left\{\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\frac{\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}}{\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\|}\right\}=\mathbb{E}\left\{\left(\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}+\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}}\right)^{\dagger}\frac{\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}}{\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\|}\right\}\\&=\mathbb{E}\left\{\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\|\right\}+\mathbb{E}\left\{\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}\dagger}\frac{\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}}{\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\|}\right\}=\sqrt{\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}}\mathbb{E}\{\vartheta\},\end{aligned}\tag{A.1}$$

where $\tau_{1n}^{\mathrm{S}}=\left(1/\gamma_{\mathrm{p}}+\sum_{j=2}^{L}\beta_{1jn}^{\mathrm{PS}}+\beta_{11n}^{\mathrm{SS}}\right)$ and γ_{p} ($\gamma_{\mathrm{p}}=\frac{\rho}{\sigma^2}$) denotes the SNR of each pilot, $\vartheta=\sqrt{\sum_{m=1}^{M_{\mathrm{S}}}|u_m|^2}$ and $\{u_m\}$ is i.i.d. $\mathscr{CN}(0,1)$.

Then,

$$\begin{aligned}\mathbb{E}\left\{\left\|\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\|^2\right\}&=\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\right\|^2\right\}+\mathbb{E}\left\{\frac{\hat{\mathbf{h}}_{11n}^{\mathrm{SS}\dagger}}{\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\|}\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}}\tilde{\mathbf{h}}_{11n}^{\mathrm{SS}\dagger}\frac{\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}}{\|\hat{\mathbf{h}}_{11n}^{\mathrm{SS}}\|}\right\}\\&=\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}\mathbb{E}\{\vartheta^2\}+1-\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}.\end{aligned}\tag{A.2}$$

Combining (A.1) and (A.2), we can get

$$\begin{aligned}\mathrm{var}\left\{\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\}&=\mathbb{E}\left\{\left\|\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\|^2\right\}-\left|\mathbb{E}\left\{\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11n}^{\mathrm{SS}}\right\}\right|^2\\&=\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}}\mathrm{var}\{\vartheta\}+1-\frac{\beta_{11n}^{\mathrm{SS}}}{\tau_{1n}^{\mathrm{S}}},\end{aligned}\tag{A.3}$$

where $\mathbb{E}\{\vartheta\}=\frac{\Gamma(M_{\mathrm{S}}+1/2)}{\Gamma(M_{\mathrm{S}})},\mathbb{E}\{\vartheta^2\}=M_{\mathrm{S}}$ and $\mathrm{var}\{\vartheta\}=M_{\mathrm{S}}-\mathbb{E}^2\{\vartheta\}$. Here, $\Gamma(\cdot)$ is the Gamma function.

We have analyzed the uplink training and downlink data transmission of the SN and omit the similar analysis for the PN because of limited space and get some conclusions directly.

When $n = k$ ($K_{\mathrm{P}} - K_{\mathrm{T}} + 1 \leq k \leq K_{\mathrm{P}}$), namely the n-th CU in the 1-st cell use the same pilot with the k-th PU in adjacent cell, according to the decomposition of MMSE, we have

$$\frac{\hat{\mathbf{h}}_{j1n}^{\mathrm{SP}}}{\sqrt{\beta_{j1n}^{\mathrm{SP}}}} = \frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}}{\sqrt{\beta_{jjk}^{\mathrm{PP}}}} \Rightarrow \frac{\hat{\mathbf{h}}_{j1k}^{\mathrm{SP}}}{\sqrt{\beta_{j1k}^{\mathrm{SP}}}} = \frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}}{\sqrt{\beta_{jjk}^{\mathrm{PP}}}}. \tag{A.4}$$

According to (A.4) and $\hat{\mathbf{h}}_{j1k}^{\mathrm{SP}} = \hat{\mathbf{h}}_{j1k}^{\mathrm{PP}} + \tilde{\mathbf{h}}_{j1k}^{\mathrm{SP}}$, we have

$$\mathbf{h}_{j1k}^{\mathrm{SP}} = \sqrt{\frac{\beta_{j1k}^{\mathrm{SP}}}{\beta_{jjk}^{\mathrm{PP}}}}\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}} + \tilde{\mathbf{h}}_{j1k}^{\mathrm{SP}}. \tag{A.5}$$

Since $\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}} \sim \mathscr{CN}\left(\mathbf{0}, \frac{\beta_{jjk}^{\mathrm{PP}}}{\tau_{jk}^{\mathrm{P}}}\mathbf{I}_{M_{\mathrm{P}}}\right)$, we can get $\tilde{\mathbf{h}}_{j1k}^{\mathrm{SP}} \sim \mathscr{CN}\left(0, \mathbf{I}_{M_{\mathrm{P}}} - \frac{\beta_{j1k}^{\mathrm{SP}}}{\tau_{jk}^{\mathrm{P}}}\mathbf{I}_{M_{\mathrm{P}}}\right)$, where $\tau_{jk}^{\mathrm{P}} = \left(1/\gamma_{\mathrm{p}} + \sum_{l=1}^{L}\beta_{jlk}^{\mathrm{PP}}\right)$ $(k \leq K_{\mathrm{P}} - K_{\mathrm{T}})$, $\tau_{jk}^{\mathrm{P}} = \left(1/\gamma_{\mathrm{p}} + \sum_{l=2}^{L}\beta_{jlk}^{\mathrm{PP}} + \beta_{j1k}^{\mathrm{SP}}\right)$ $(K_{\mathrm{P}} - K_{\mathrm{T}} + 1 \leq k \leq K_{\mathrm{P}})$.

Therefore

$$\begin{aligned}
\mathbb{E}\left\{\left|\mathbf{h}_{j1n}^{\mathrm{SP}\dagger}\mathbf{w}_{jjk}^{\mathrm{PP}}\right|^2\right\} &= \mathbb{E}\left\{\left|\left(\sqrt{\frac{\beta_{j1k}^{\mathrm{SP}}}{\beta_{jjk}^{\mathrm{PP}}}}\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}} + \tilde{\mathbf{h}}_{j1k}^{\mathrm{SP}}\right)^{\dagger}\frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}}{\|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\|}\right|^2\right\} \\
&= \mathbb{E}\left\{\frac{\beta_{j1k}^{\mathrm{SP}}}{\beta_{jjk}^{\mathrm{PP}}}\left\|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\right\|^2\right\} + \mathbb{E}\left\{\frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}\dagger}}{\|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\|}\tilde{\mathbf{h}}_{j1k}^{\mathrm{SP}}\tilde{\mathbf{h}}_{j1k}^{\mathrm{SP}\dagger}\frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}}{\|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\|}\right\} \\
&= \frac{\beta_{j1k}^{\mathrm{SP}}}{\tau_{jk}^{\mathrm{P}}}\mathbb{E}\{\varepsilon^2\} + 1 - \frac{\beta_{j1k}^{\mathrm{SP}}}{\tau_{jk}^{\mathrm{P}}},
\end{aligned} \tag{A.6}$$

where $\varepsilon = \sqrt{\sum_{m=1}^{M_{\mathrm{P}}}|u_m|^2}$. $\mathbb{E}\{\varepsilon\} = \frac{\Gamma(M_{\mathrm{P}}+1/2)}{\Gamma(M_{\mathrm{P}})}$, $\mathbb{E}\{\varepsilon^2\} = M_{\mathrm{P}}$ and $\mathrm{var}\{\varepsilon\} = M_{\mathrm{P}} - \mathbb{E}^2\{\varepsilon\}$.

When $n \neq k$

$$\mathbb{E}\left\{\left|\mathbf{h}_{j1n}^{\mathrm{SP}\dagger}\mathbf{w}_{jjk}^{\mathrm{PP}}\right|^2\right\} = \mathbb{E}\left\{\frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}\dagger}}{\|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\|}\mathbf{h}_{j1n}^{\mathrm{SP}}\mathbf{h}_{j1n}^{\mathrm{SP}\dagger}\frac{\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}}{\|\hat{\mathbf{h}}_{jjk}^{\mathrm{PP}}\|}\right\} = 1. \tag{A.7}$$

$$\mathbb{E}\left\{\left|\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\mathbf{w}_{11i}^{\mathrm{SS}}\right|^2\right\} = \mathbb{E}\left\{\frac{\hat{\mathbf{h}}_{11i}^{\mathrm{SS}\dagger}}{\|\hat{\mathbf{h}}_{11i}^{\mathrm{SS}}\|}\mathbf{h}_{11n}^{\mathrm{SS}}\mathbf{h}_{11n}^{\mathrm{SS}\dagger}\frac{\hat{\mathbf{h}}_{11i}^{\mathrm{SS}}}{\|\hat{\mathbf{h}}_{11i}^{\mathrm{SS}}\|}\right\} = 1. \tag{A.8}$$

We finish the proof.

Appendix B

Since the process of the PU's SINR is similar with that of the CU, we directly get the PU's SINR as follows

$$\mathrm{SINR}_{jk} = \frac{(1/\tau_{jk}^{\mathrm{P}}) p_{\mathrm{t}} {\beta_{jjk}^{\mathrm{PP}}}^2 \mathbb{E}^2\{\varepsilon\}}{I_1' + I_2' + I_3' + I_4' + \sigma^2}. \tag{B.1}$$

We assume two cases:

(1) The first case: when $k \leq K_{\mathrm{P}} - K_{\mathrm{T}}$

$$\begin{aligned}
I_1' &= p_{\mathrm{t}} \beta_{jjk}^{\mathrm{PP}} \left(\frac{\beta_{jjk}^{\mathrm{PP}}}{\tau_{jk}^{\mathrm{P}}} \mathrm{var}\{\varepsilon\} + 1 - \frac{\beta_{jjk}^{\mathrm{PP}}}{\tau_{jk}^{\mathrm{P}}} \right) \\
I_2' &= \sum_{l \neq j}^{L} \beta_{ljk}^{\mathrm{PP}} p_{\mathrm{t}} \left(\frac{\beta_{ljk}^{\mathrm{PP}}}{\tau_{lk}^{\mathrm{P}}} \mathbb{E}\{\varepsilon^2\} + 1 - \frac{\beta_{ljk}^{\mathrm{PP}}}{\tau_{lk}^{\mathrm{P}}} \right) \\
I_3' &= \sum_{l=2}^{L} \sum_{i \neq k}^{K_{\mathrm{P}}} \beta_{ljk}^{\mathrm{PP}} p_{\mathrm{t}} + \sum_{m \neq k}^{K_{\mathrm{P}} - K_{\mathrm{T}}} \beta_{1jk}^{\mathrm{PP}} p_{\mathrm{t}} \\
I_4' &= \sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}} \beta_{1jk}^{\mathrm{PS}} P_{1n}
\end{aligned}$$

According to (4.10c), we have

$$\sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}} P_{1n} \leq \frac{\left(\frac{(1/\tau_{jk}^{\mathrm{P}}) p_{\mathrm{t}} {\beta_{jjk}^{\mathrm{PP}}}^2 \mathbb{E}^2\{\varepsilon\}}{\eta_{jk}} - I_1' - I_2' - I_3' \right)}{\beta_{1jk}^{\mathrm{PS}}}. \tag{B.2}$$

Combining (4.10b) and (B.2), we can get

$$\sum_{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1}^{K_{\mathrm{P}}} P_{1n} \leq \min\{I_{\mathrm{P}}, P_{\max}\}, \tag{B.3}$$

where

$I_{\mathrm{P}} = \min\left\{[\mathscr{I}_{jk}]_{L\times(K_{\mathrm{P}}-K_{\mathrm{T}})}\right\}$, $\mathscr{I}_{jk} = \frac{1}{\beta_{1jk}^{\mathrm{PS}}}\left(\frac{(1/\tau_{jk}^{\mathrm{P}})p_{\mathrm{t}}{\beta_{jjk}^{\mathrm{PP}}}^2\mathbb{E}^2\{\varepsilon\}}{\eta_{jk}} - I_1' - I_2' - I_3'\right)$

(2) The second case: when $K_{\mathrm{P}} - K_{\mathrm{T}} + 1 \le k \le K_{\mathrm{P}}$

$$
\begin{aligned}
I_1' &= p_{\mathrm{t}}\beta_{jjk}^{\mathrm{PP}}\left(\frac{\beta_{jjk}^{\mathrm{PP}}}{\tau_{jk}^{\mathrm{P}}}\mathrm{var}\{\varepsilon^2\} + 1 - \frac{\beta_{jjk}^{\mathrm{PP}}}{\tau_{jk}^{\mathrm{P}}}\right)\\
I_2' &= \sum_{l\neq j}^{L}\beta_{ljk}^{\mathrm{PP}}p_{\mathrm{t}}\left(\frac{\beta_{ljk}^{\mathrm{PP}}}{\tau_{lk}^{\mathrm{P}}}\mathbb{E}\{\varepsilon^2\} + 1 - \frac{\beta_{ljk}^{\mathrm{PP}}}{\tau_{lk}^{\mathrm{P}}}\right)\\
I_3' &= \sum_{l=2}^{L}\sum_{i\neq k}^{K_{\mathrm{P}}}\beta_{ljk}^{\mathrm{PP}}p_{\mathrm{t}} + \sum_{m=k}^{K_{\mathrm{P}}-K_{\mathrm{T}}}\beta_{1jk}^{\mathrm{PP}}p_{\mathrm{t}}\\
I_4' &= \sum_{\substack{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ n\neq k}}^{K_{\mathrm{P}}}\beta_{1jk}^{\mathrm{PS}}P_{1n} + \beta_{1jk}^{\mathrm{PS}}P_{1k}\left(\frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}}\mathbb{E}\{\vartheta^2\} + 1 - \frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}}\right)
\end{aligned}
$$

According to (4.10c), we have

$$
\sum_{\substack{n=K_{\mathrm{P}}-K_{\mathrm{T}}+1\\ n\neq k}}^{K_{\mathrm{P}}} P_{1n} + \xi_{1k}P_{1k} \le \mathscr{I}_{jk}, \tag{B.4}
$$

where $\xi_{1k} = \left(\frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}}\mathbb{E}\{\vartheta^2\} + 1 - \frac{\beta_{1jk}^{\mathrm{PS}}}{\tau_{1k}^{\mathrm{S}}}\right)$.

Appendix C

After ZF precoding, the SINR of the MU m can be written as:

$$\mathrm{SINR}_m^{\mathrm{M}} = \frac{P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\mathbf{w}_m\right|^2}{\sum\limits_{k=1}^{K} P_{\mathrm{S}}\beta_{k,m}^{\mathrm{S,M}}\left|\mathbf{h}_{k,m}^{\mathrm{S,M}}\mathbf{v}_k\right|^2+\delta^2}. \tag{C.1}$$

According to Jensen's inequality and the convexity of $\log_2(1+\frac{1}{x})$, we can get as follows:

$$R_m^{\mathrm{M}} \geq \underline{R}_m^{\mathrm{M}} \triangleq \left(1-\frac{1+M}{S}\right)\log_2\left(1+\left(\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_m^{\mathrm{M}}}\right\}\right)^{-1}\right). \tag{C.2}$$

Then, we have:

$$\begin{aligned}
\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_m^{\mathrm{M}}}\right\} &= \mathbb{E}\left\{\frac{\sum\limits_{k=1}^{K} P_{\mathrm{S}}\beta_{k,m}^{\mathrm{S,M}}\left|\mathbf{h}_{k,m}^{\mathrm{S,M}}\mathbf{v}_k\right|^2+\delta^2}{P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\mathbf{w}_m\right|^2}\right\} \\
&= \left(\sum_{k=1}^{K} P_{\mathrm{S}}\beta_{k,m}^{\mathrm{S,M}}\mathbb{E}\left|\mathbf{h}_{k,m}^{\mathrm{S,M}}\mathbf{v}_k\right|^2+\delta^2\right)\mathbb{E}\left\{\frac{1}{P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\mathbf{w}_m\right|^2}\right\} \\
&= \left(\sum_{k=1}^{K} P_{\mathrm{S}}\beta_{k,m}^{\mathrm{S,M}}\mathbb{E}\left\{\frac{\mathbf{h}_{k,k}^{\mathrm{S,S}}}{\|\mathbf{h}_{k,k}^{\mathrm{S,S}^H}\|}\mathbf{h}_{k,m}^{\mathrm{S,M}^H}\mathbf{h}_{k,m}^{\mathrm{S,M}}\frac{\mathbf{h}_{k,k}^{\mathrm{S,S}^H}}{\|\mathbf{h}_{k,k}^{\mathrm{S,S}^H}\|}\right\}+\delta^2\right)\frac{1}{P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}}\mathbb{E}\left\{\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\frac{\overline{\mathbf{w}}_{\mathbf{m}}}{\|\overline{\mathbf{w}}_m\|}\right|^{-2}\right\}
\end{aligned} \tag{C.3}$$

Since $\mathbf{h}_{k,k}^{\mathrm{S,S}}$ is independent of $\mathbf{h}_{k,m}^{\mathrm{S,M}}$, we have

$\mathbb{E}\left\{\frac{\mathbf{h}_{k,k}^{\mathrm{S,S}}}{\|\mathbf{h}_{k,k}^{\mathrm{S,S}^H}\|}\mathbf{h}_{k,m}^{\mathrm{S,M}^H}\mathbf{h}_{k,m}^{\mathrm{S,M}}\frac{\mathbf{h}_{k,k}^{\mathrm{S,S}^H}}{\|\mathbf{h}_{k,k}^{\mathrm{S,S}^H}\|}\right\}=1$, then,

$$\begin{aligned}\mathbb{E}\left\{\left|\mathbf{h}_{0,m}^{\mathrm{M,M}}\frac{\overline{\mathbf{w}}_{\mathbf{m}}}{\|\overline{\mathbf{w}}_m\|}\right|^{-2}\right\}=&\mathbb{E}\left\{\|\overline{\mathbf{w}}_m\|^2\right\}=\mathbb{E}\left\{\left[\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right]_{m,m}\right\}\\ =&\frac{1}{M}\mathbb{E}\left\{\mathrm{tr}\left\{\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right\}\right\}\overset{(a)}{=}\frac{1}{N_{\mathrm{M}}-M},\end{aligned}\tag{C.4}$$

where (a) is obtained by using the identity [63],

$$\mathbb{E}\left\{\mathrm{tr}\left(\mathbf{W}^{-1}\right)\right\}=\frac{m}{n-m},\tag{C.5}$$

where $\mathbf{W}\sim\mathscr{W}_m(n,\mathbf{I}_n)$ is an $m\times n$ central complex Wishart matrix with $n(n>m)$ degrees of freedom. Then, we can get:

$$\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_m^{\mathrm{M}}}\right\}=\frac{\sum\limits_{k=1}^{K}P_{\mathrm{S}}\beta_{k,m}^{\mathrm{S,M}}+\delta^2}{P_{\mathrm{M}}\beta_{0,m}^{\mathrm{M,M}}(N_{\mathrm{M}}-M)}.\tag{C.6}$$

Therefore, we finish the proof. ∎

For (6.9), similar to (C.1),

$$R_k^{\mathrm{S}}\geq\underline{R}_k^{\mathrm{S}}\triangleq\left(1-\frac{1+M}{S}\right)\log_2\left(1+\left(\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_k^{\mathrm{S}}}\right\}\right)^{-1}\right).\tag{C.7}$$

Then, we have:

$$
\begin{aligned}
\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_k^{\mathrm{S}}}\right\} &= \mathbb{E}\left\{\frac{\sum_{i\neq k}^{K} P_{\mathrm{S}}\beta_{i,k}^{\mathrm{S,S}}\left|\mathbf{h}_{i,k}^{\mathrm{S,S}}\mathbf{v}_i\right|^2+\sum_{m=1}^{M} P_{\mathrm{M}}\beta_{0,k}^{\mathrm{M,S}}\left|\mathbf{h}_{0,k}^{\mathrm{M,S}}\mathbf{w}_m\right|^2+\delta^2}{P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}\left|\mathbf{h}_{k,k}^{\mathrm{S,S}}\mathbf{v}_k\right|^2}\right\} \\
&= \left(\sum_{i\neq k}^{K} P_{\mathrm{S}}\beta_{i,k}^{\mathrm{S,S}}\mathbb{E}\left\{\frac{\mathbf{h}_{i,i}^{\mathrm{S,S}}}{\|\mathbf{h}_{i,i}^{\mathrm{S,S}H}\|}\mathbf{h}_{i,k}^{\mathrm{S,S}H}\mathbf{h}_{i,k}^{\mathrm{S,S}}\frac{\mathbf{h}_{i,i}^{\mathrm{S,S}H}}{\|\mathbf{h}_{i,i}^{\mathrm{S,S}H}\|}\right\}\right. \\
&\quad \left.+\sum_{m=1}^{M} P_{\mathrm{M}}\beta_{0,k}^{\mathrm{M,S}}\mathbb{E}\left\{\frac{\overline{\mathbf{w}}_m^H}{\|\overline{\mathbf{w}}_m\|}\mathbf{h}_{0,k}^{\mathrm{M,S}H}\mathbf{h}_{0,k}^{\mathrm{M,S}}\frac{\overline{\mathbf{w}}_{\mathbf{m}}}{\|\overline{\mathbf{w}}_m\|}\right\}+\delta^2\right) \\
&\quad \frac{1}{P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}}\mathbb{E}\left\{\left(\frac{\mathbf{h}_{k,k}^{\mathrm{S,S}}}{\|\mathbf{h}_{k,k}^{\mathrm{S,S}H}\|}\mathbf{h}_{k,k}^{\mathrm{S,S}H}\mathbf{h}_{k,k}^{\mathrm{S,S}}\frac{\mathbf{h}_{k,k}^{\mathrm{S,S}H}}{\|\mathbf{h}_{k,k}^{\mathrm{S,S}H}\|}\right)^{-1}\right\} \\
&= \left(\sum_{i\neq k}^{K} P_{\mathrm{S}}\beta_{i,k}^{\mathrm{S,S}}+P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}+\delta^2\right)\frac{1}{P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}}\mathbb{E}\left\{\|\mathbf{h}_{k,k}^{\mathrm{S,S}}\|^{-2}\right\} \\
&= \frac{1}{P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}\left(N_{\mathrm{S}}-1\right)}\left(\sum_{i\neq k}^{K} P_{\mathrm{S}}\beta_{i,k}^{\mathrm{S,S}}+P_{\mathrm{S}}\beta_{k,k}^{\mathrm{S,S}}+\delta^2\right)
\end{aligned}
\tag{C.8}
$$

We finish the proof.

Appendix D

Proof: According to the Nash theorem [64], NE exists if the following conditions hold:1) The action space of each player is convex and compact. 2) The utility function $U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})$ is concave with respect to (w.r.t.) $\mathbf{\Phi}_l$.

According to (6.30b), we can easily get that the action space $\mathbf{\Phi}_l$ satisfies the above 1). Next, we still need to prove that $U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})$ is concave w.r.t. $\mathbf{\Phi}_l$.

$$\begin{aligned} U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l}) &= \sum_{k=1}^{K_l} \widehat{U}_{lk}(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l}) = \sum_{k=1}^{K_l} \left(\widetilde{R}_{lk} - L_{lk} \right) \\ = \sum_{k=1}^{K_l} &\left(\log_2 \left(1 + \frac{\widetilde{\mathbf{h}}_{llk} \mathbf{\Phi}_{lk} \widetilde{\mathbf{h}}_{llk}^H}{\Xi_{lk} + \delta^2} \right) - \sum_{j \neq l}^{C} \sum_{i=1}^{K_j} m_j \mathbf{h}'_{lji} \mathbf{T}_{lk} \mathbf{\Phi}_{lk} \mathbf{T}_{lk}^H \mathbf{h}'^H_{lji} \right), \end{aligned} \tag{D.1}$$

where $\Xi_{lk} = \sum_{i \neq l}^{C} \sum_{j=1}^{K_i} \widetilde{\mathbf{h}}_{ilj} \mathbf{\Phi}_{ij} \widetilde{\mathbf{h}}_{ilj}^H$. Then, we have:

$$\frac{\partial^2 \widehat{U}_{lk}(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})}{\partial^2 \mathbf{\Phi}_{lk}} = -\frac{\Xi_{lk} + \delta_2}{\ln 2} \frac{\widetilde{\mathbf{h}}_{lk}^H \widetilde{\mathbf{h}}_{lk} \left(\widetilde{\mathbf{h}}_{lk}^H \widetilde{\mathbf{h}}_{lk} \right)^H}{(\Xi_{lk} + \delta_2 + \widetilde{\mathbf{h}}_{llk} \mathbf{\Phi}_{lk} \widetilde{\mathbf{h}}_{llk}^H)^2} \preceq 0. \tag{D.2}$$

According to (D.2), we can verify that $\widehat{U}_{lk}(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})$ is concave w.r.t. $\mathbf{\Phi}_{lk}$ [43], so the utility function $U_l(\mathbf{m}, \mathbf{\Phi}_l, \mathbf{\Phi}_{-l})$ is concave w.r.t. $\mathbf{\Phi}_l$. Therefore, (6.30) is a concave game and there exists one NE.

Appendix E

Proof: It is clear that matrix $\mathbf{Z}_{lk}$ is a symmetric matrix. Therefore, we only need to prove that matrix $\mathbf{Z}_{lk}$ is a full-rank matrix. Next, we prove it by contradiction. We assume that $\mathbf{Z}_{lk}$ is not a full-rank matrix and then, we can always find a vector $\mathbf{q}_{lk} \in \mathbb{C}^{(N_l-K_l+1)\times 1}$ such that $\mathbf{Z}_{lk}\mathbf{q}_{lk} = \mathbf{0}$ and $\widetilde{\mathbf{h}}_{llk}\mathbf{q}_{lk} \neq \mathbf{0}$. On this basis, we assume the optimal $\mathbf{\Phi}_{lk}^{*} = x\mathbf{q}_{lk}\mathbf{q}_{lk}^{H}(x \geq 0)$. Substituting this optimal solution the (6.36) and yielding:

$$\log_2\left(1+\frac{\widetilde{\mathbf{h}}_{llk}\mathbf{\Phi}_{lk}^{*}\widetilde{\mathbf{h}}_{llk}^{H}}{\Xi_{lk}+\delta^2}\right)-\mathrm{Tr}(\mathbf{Z}_{lk}\mathbf{\Phi}_{lk}^{*})=\log_2\left(1+\frac{x\widetilde{\mathbf{h}}_{llk}\mathbf{q}_{lk}\mathbf{q}_{lk}^{H}\widetilde{\mathbf{h}}_{llk}^{H}}{\Xi_{lk}+\delta^2}\right) \tag{E.1}$$

Since $\widetilde{\mathbf{h}}_{llk}\mathbf{q}_{lk}\mathbf{q}_{lk}^{H}\widetilde{\mathbf{h}}_{llk}^{H} > 0$, the objective value is unbounded as x goes to infinity. Therefore, the original assumption that $\mathbf{Z}_{lk}$ not is a full-rank matrix is not correct. Furthermore, $\mathbf{Z}_{lk}$ is a positive definite matrix .

Appendix F

Proof: The concept of the standard function is defined in [65] to prove the uniqueness of the NE for a non-cooperative game. Here, we define $\mathbf{\Phi}_{lk} = \mathbf{\Phi}_{lk}(\Xi_{lk})$, namely other parameters ($\mathbf{m}$ and $\boldsymbol{\mu}_l$) other than Ξ_{lk} can be taken as given constants. Next, we need to prove that $\mathbf{\Phi}_{lk}(\Xi_{lk})$ is a standard function with respect to (w.r.t.) Ξ_{lk}, which must hold following: 1) positive: $\mathbf{\Phi}_{lk}(\Xi_{lk}) \succeq \mathbf{0}$; 2) monotonic; 3) scalable: $\lambda\mathbf{\Phi}_{lk}(\Xi_{lk}) \succ \mathbf{\Phi}_{lk}(\lambda\Xi_{lk})$ for any $\lambda > 1$.

First, we rewrite the $\mathbf{\Phi}_{lk}$ as follows:

$$\mathbf{\Phi}_{lk}(\Xi_{lk}) = \left(\frac{1}{\ln 2} - \frac{\Xi_{lk}+\delta^2}{\alpha_{lk}}\right)^+ \mathbf{\Theta}_{lk}, \tag{F.1}$$

where $\mathbf{\Theta}_{lk} = (\mathbf{a}_{lk}\hat{\mathbf{Z}}_{lk}^{-1})^H \mathbf{a}_{lk}\hat{\mathbf{Z}}_{lk}^{-1}$.

From (F.1), it is clear that $\mathbf{\Theta}_{lk} \succeq \mathbf{0}$ and $\mathbf{\Phi}_{lk}(\Xi_{lk})$ is a monotonically decreasing function w.r.t. Ξ_{lk}. Therefore, 1) and 2) hold. Next, we focus on proving 3).

$$\begin{aligned}
&\lambda\mathbf{\Phi}_{lk}(\Xi_{lk}) - \mathbf{\Phi}_{lk}(\lambda\Xi_{lk}) \\
=&\lambda\left(\frac{1}{\ln 2} - \frac{\Xi_{lk}+\delta^2}{\alpha_{lk}}\right)^+ \mathbf{\Theta}_{lk} - \left(\frac{1}{\ln 2} - \frac{\lambda\Xi_{lk}+\delta^2}{\alpha_{lk}^2}\right)^+ \mathbf{\Theta}_{lk} \\
=&(\lambda-1)\left(\frac{1}{\ln 2} - \frac{\delta^2}{\alpha_{lk}^2}\right)^+ \mathbf{\Theta}_{lk} \\
\succ&(\lambda-1)\left(\frac{1}{(\ln 2} - \frac{\Xi+\delta^2}{\alpha_{lk}^2}\right)^+ \mathbf{\Theta}_{lk} \\
=&(\lambda-1)\mathbf{\Phi}_{lk}(\Xi_{lk}) \\
\succeq&\mathbf{0}.
\end{aligned} \tag{F.2}$$

According to (F.2), it can be verified that the above 3) is held and $\mathbf{\Phi}_{lk}(\Xi_{lk})$ is a standard function w.r.t. Ξ_{lk}. Since a standard function will converge to a unique value, the NE for the non-cooperative game (6.30) is unique and we finish the proof.