FAIR LICENSED SPECTRUM SHARING BETWEEN TWO MNOS USING RESOURCE
OPTIMIZATION IN MULTI-CELL MULTI-USER MIMO NETWORKS

by

Tianchen Wang

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

© Copyright 2021 by Tianchen Wang

Fair Licensed Spectrum Sharing Between Two MNOs Using Resource Optimization in Multi-Cell Multi-User MIMO Networks

Tianchen Wang

Master of Applied Science

Graduate Department of Electrical and Computer Engineering

University of Toronto

2021

Abstract

Spectrum sharing has been proposed to both improve spectrum utilization efficiency in modern wireless communications and provide mobile network operators (MNOs) with required spectrum at times of increased traffic. In this thesis, we investigate the use of multi-user multiple input, multiple output (MIMO) techniques to enable spectrum sharing. Specifically, we present a fair spectrum sharing system between two MNOs in multi-cell multi-user MIMO networks. We impose fairness by ensuring that each MNO receives spectrum in proportion to the amount it contributes. We formulate a constrained optimization problem to determine resource allocation and user scheduling across two MNOs. Since the problem is non-convex, we develop an effective solution through fractional programming and block coordinate descent. In addition, we incorporate channel estimation into this spectrum sharing framework in a time division duplex (TDD) system. We propose a cluster based pilot allocation mechanism where the pilot sequences are assigned to user terminals (UTs) within each newly organized cluster, which is determined by a BS-centric clustering algorithm and a modified K-means clustering algorithm. Our numerical results illustrate that the proposed spectrum sharing scheme can achieve up to a 60% improvement in terms of the average user rate among the two operators while ensuring that neither MNO is exploited for participating in the sharing mechanism, within a diagonally overlapped cell deployment. This improvement is in relation to the baseline of each MNO using multi-user MIMO communications on its own.

Acknowledgments

Due to the COVID-19 pandemic, most of the work in this thesis is done remotely at home, which is more challenging than normal, due to the inefficiency in online communications. This thesis would not have been possible without the tremendous support I received from many people during the period.

First and foremost, I would like to give my deepest gratitude to my supervisor, Prof. Raviraj Adve, for his trust in me to offer me such a valuable opportunity two years ago, for his guidance in my research throughout my graduate studies, and for his criticism and advice on the mistakes I made, without which this thesis would not have been possible. I have learnt much from his persistence and tireless pursuit in diving deep into the research areas.

I would like to thank my thesis committee members Prof. Shahrokh Valaee, Prof. Ben Liang, and Prof. Mark Jeffrey, for their considerable time spent on reading my thesis and insightful comments provided that improve this thesis.

I feel blessed to have such great colleagues in the group. Specifically, I would like to thank Ahmad Khan, Hadeel Elayan and Hussein Ammar for their kind help and guidance in my graduate coursework and teaching assistantship during the limited time in office before the pandemic. Moreover, I value the help from Hussein and Ahmad for their timely and detailed emails in answering the questions I had for their work, during the remote working time.

Last but not least, I would like to thank my beloved, Ziou Wang, who I spent most of the time with in the past two years, for her company, encouragement and support. I wish to thank my parents Youzhao Wang and Jun Wang, for the endless love and support they gave me over the past seven years when I study abroad alone. This thesis is dedicated to them.

Contents

1	Introduction	1
1.1	Background	1
1.2	Literature Review	2
1.3	Motivation and Contributions	6
1.4	Thesis Organization	7
1.5	Notations	8
2	SINR-based Fair licensed Spectrum Sharing Between Two MNOs	9
2.1	Introduction	9
2.2	System Model	9
2.3	Problem Formulation	11
2.4	Proposed Solution Approach	14
2.5	Simulation Results and Performance Evaluation	24
2.5.1	Diagonally Overlapped Cell Deployment	26
2.5.2	Co-located Cell Deployment	37
2.6	Summary	45
3	Channel Estimation in Spectrum Sharing Between Two MNOs	46
3.1	Introduction	46
3.2	Proposed Approach	47
3.2.1	Uplink Training and Channel Estimation	47

3.2.2	User re-clustering	48
3.3	Simulation Results and Performance Evaluation	53
3.3.1	Diagonally Overlapped Cell Deployment	54
3.3.2	Co-located Cell Deployment	62
3.4	Summary	68
4	SLINR-based Fair licensed Spectrum Sharing Between Two MNOs	69
4.1	Introduction	69
4.2	Problem Formulation	70
4.3	Proposed Solution Approach	72
4.4	Simulation Results and Performance Evaluation	74
4.4.1	Diagonally Overlapped Cell Deployment	74
4.4.2	Co-located Cell Deployment	75
4.5	Summary	79
5	Conclusions and Future Work	80
	Bibliography	83

List of Tables

2.1	System parameter settings used.	24
-----	---	----

List of Figures

2.1	Example bipartite graph between users and beams	20
2.2	Convergence of the objective function in (2.4) using Algorithm 1.	25
2.3	Cell deployment for operators ($\mathcal{B} = 9$).	26
2.4	Average user rate among 9 cells of both operators (Diagonally overlapped cell deployment).	27
2.5	Average user rate among 9 cells of each operator (Diagonally overlapped cell deployment).	28
2.6	Rate gain compared with no spectrum sharing.	29
2.7	Average number of users scheduled per cell (Diagonally overlapped cell deployment).	31
2.8	Distribution of frequency slot allocation among users, fair spectrum sharing ($N_2 = 4$, Diagonally overlapped cell deployment).	33
2.9	Distribution of frequency slot allocation among users, fair / full spectrum sharing ($N_2 = 16$, Diagonally overlapped cell deployment).	33
2.10	Distribution of frequency slot allocation among users, no spectrum sharing ($N_2 = 4$, Diagonally overlapped cell deployment).	33
2.11	Rate changes among users ($N_2 = 4$, Diagonally overlapped cell deployment).	35
2.12	Rate changes among users ($N_2 = 16$, Diagonally overlapped cell deployment).	36
2.13	Average user rate among 9 cells of both operators (Co-located cell deployment).	37
2.14	Average user rate among 9 cells of each operator (Co-located cell deployment).	38
2.15	Average number of users scheduled per cell (Co-located cell deployment).	41

2.16	Distribution of frequency slot allocation among users, fair spectrum sharing ($N_2 = 4$, Co-located cell deployment).	42
2.17	Distribution of frequency slot allocation among users, fair / full spectrum sharing ($N_2 = 16$, Co-located cell deployment).	42
2.18	Distribution of frequency slot allocation among users, no spectrum sharing ($N_2 = 4$, Co-located cell deployment).	42
2.19	Rate changes among users ($N_2 = 4$, Co-located cell deployment).	43
2.20	Rate changes among users ($N_2 = 16$, Co-located cell deployment).	44
3.1	User clustering with only BS-centric user clustering (Diagonally overlapped cell deployment).	51
3.2	User clustering with BS-centric user clustering (as in the red circle) and K- means clustering (as in the blue circle) (Diagonally overlapped cell deployment).	52
3.3	User clustering with BS-centric user clustering and K-means clustering (Co- located cell deployment).	52
3.4	Intra-operator channel estimation NMSE (Diagonally overlapped cell deploy- ment).	55
3.5	Inter-operator channel estimation NMSE (Diagonally overlapped cell deploy- ment).	56
3.6	Average user rate among 9 cells of both operators with channel estimation (Diagonally overlapped cell deployment).	57
3.7	Average user rate among 9 cells of each operator with channel estimation (Diagonally overlapped cell deployment).	58
3.8	Achievable sum rate of each operator with channel estimation (Diagonally overlapped cell deployment).	61
3.9	Intra-operator channel estimation NMSE (Co-located cell deployment).	63
3.10	Inter-operator channel estimation NMSE (Co-located cell deployment).	63
3.11	Average user rate among 9 cells of both operators with channel estimation (Co-located cell deployment).	64

3.12	Average user rate among 9 cells of each operator with channel estimation (Co-located cell deployment).	66
3.13	Achievable sum rate of each operator with channel estimation (Co-located cell deployment).	67
4.1	Average user rate among 9 cells of both operators (SLINR).	75
4.2	Average user rate among 9 cells of each operator (SLINR).	76
4.3	Average user rate among 9 cells of both operators (SLINR).	77
4.4	Average user rate among 9 cells of each operator (SLINR).	78

Chapter 1

Introduction

1.1 Background

Along with the development of modern wireless communication systems, the number of mobile devices has increased explosively over the years. However, spectrum, which plays a fundamental role in wireless communications, remains a scarce resource [1]. Nonetheless, the spectrum is under-utilized, partly because of licensing spectrum on an exclusive basis by regulators for a wide range of wireless communication systems, including mobile network operators (MNOs) and non-mobile systems (e.g., radar, military, etc.). In this regard, allowing (or requiring) the wireless communication systems to share their licensed spectrum is an innovative way to approach the above problem. In addition, sharing spectrum could allow the wireless communication systems access to extra spectrum in times of increased traffic, and thus reduces the cost in acquiring additional licensed spectrum via auction. The methods of sharing the licensed spectrum are named licensed spectrum sharing.

Licensed spectrum sharing can be categorized into *heterogeneous* spectrum sharing and *homogeneous* spectrum sharing [1]. In heterogeneous sharing, the spectrum is shared between different wireless communication systems (i.e., mobile and non-mobile systems). Licensed-Assisted Access (LAA) and Licensed Shared Access (LSA) are two typical heterogeneous spectrum sharing schemes, where the spectrum is shared between MNOs and an incumbent spectrum user. Licensed-Assisted Access (LAA) is an LTE-unlicensed spectrum sharing

scheme which allows MNOs to incorporate the 5 GHz unlicensed band together with their licensed spectrum to deliver wireless service to the end users. Licensed Shared Access (LSA) is a spectrum sharing concept which enables MNOs (i.e., licensee operators (LOs)) to access additional spectrum resource from other spectrum owners (i.e., incumbent operators (IOs)) on a licensed shared base [2].

On the contrary, homogeneous sharing denotes the spectrum sharing within the same communication system (i.e., sharing among multiple MNOs). Homogeneous spectrum sharing among operators can be further classified into mutual renting and spectrum pooling, when spectrum is the only shared resource. For mutual renting, each MNO can rent their unutilized spectrum to other MNOs for additional revenue, and thus increases the spectrum utilization efficiency.

In the case of spectrum pooling, in addition to the spectrum exclusively allocated to specific MNOs, there is assumed to exist an unallocated band spectrum operated by the national regulatory authority which is to be shared by MNOs. Alternatively, multiple operators agree to use their aggregate spectrum in a shared manner, where there is a resource manager to manage spectrum allocation among the operators. The topic covered in this thesis falls into this category of spectrum pooling where two MNOs agree to share spectrum, but exploits concepts from multi-user multiple input, multiple output (MIMO) networks to simultaneously use the same spectrum resources.

1.2 Literature Review

We begin here with a comprehensive review of the relevant literature of licensed spectrum sharing.

In [3], the authors propose an auction scheme through Lyapunov optimization for the LTE-unlicensed users to share the same unlicensed spectrum with Wi-Fi in an opportunistic manner. The method maximizes the overall social welfare and guarantees bounded drop rate and delay for the participating BSs. In [4], the authors apply game theory to study spectrum

sharing within unlicensed bands among multiple strategic operators. The dynamics of the spectrum sharing market are investigated through an entry game. A dynamic sharing scheme among participating operators is proposed to work with the time-varying traffic.

Regarding mutual renting, in [5], the authors investigate how secondary network operators (i.e., buyers) should acquire spectrum access from primary network operators (i.e., sellers) based on the characteristics (i.e., cost and quality) of the available bands. Two optimization problems are formulated to study the questions on how SNOs could minimize cost while ensuring targeted quality-of-service (QoS), and how SNOs could achieve maximal utility gain with limited budget, respectively. However, mutual renting is an opportunistic spectrum sharing mechanism which is beneficial to relieve temporal spectrum shortage, yet cannot improve the whole system throughput in the long run.

The spectrum pooling approach can be either cooperative or non-cooperative. The difference between the two schemes is whether the coordination of the spectrum allocation among MNOs is real-time or not. With non-cooperative spectrum pooling, the operators access the spectrum pool following predefined regulations or protocols. In [6], the authors propose three resource sharing models between two operators by overflow management. The proposed models reduce the blocking probability for both operators with low traffic intensity. If only one operator suffers from heavy traffic, it can attain a reduction in blocking probability as a tradeoff with a marginal increase in the blocking probability of the other operator. The work in [6] is revisited in [7], where the authors propose two different rules in deciding the amount of spectrum to be shared based on the same model. The degradation of blocking probability for the operator with lower traffic intensity (the lender) during the sharing is minimized. However, the scenario when the traffic intensities for both operator are high is not investigated in either work.

Game theory is a useful tool to investigate the behaviour of MNOs when the spectrum sharing is non-cooperative. In [8], the authors investigate spectrum sharing amongst operators for device-to-device communications. They formulate a non-cooperative game to model the bargain of spectrum contributions of each participating operator to the shared spectrum

pool. A unique equilibrium can be achieved by the proposed Jacobi-play strategy update and all operators can achieve performance gains compared with the non-sharing case. However, the gain and cost is not guaranteed to be proportional (i.e., although an operator could improve its performance, it may need to contribute more spectrum to the pool than others). In this thesis, we attempt to address this potential disparity.

The authors in [9] apply matching game theory to study non-orthogonal spectrum assignment among operators to maximize the sum rate of all operators. In the matching game, the distribution of transmit power at each BS is investigated using Q-learning. The study reveals that spectrum allocation contributes more to the gain in the sum rate than power allocation. However, the authors do not study the benefit brought to each operator (i.e., operators may not be incentivized to join in spectrum sharing).

In [10], the authors used non-cooperative repeated games to model the interactions between two operators in a small (femto) cell. At each term, each operator proposes the number of component carriers (CCs) to borrow and lease among the K CCs in the spectrum pool. The CCs can be used exclusively by one operator once the number of CCs it borrows is no more than the number of CCs the other operator is willing to grant. The proposed model outperforms the case where there is no coordination between the operators. The same repeated game model as in [10] is used to study spectrum sharing in multi-operator heterogeneous networks in [11]. The spectrum is shared between the micro-cells and femto-cells of different operators. However, in [10] and [11], it is assumed that the operators have similar loads and are rational during the games, which is not realistic.

In regard to cooperative spectrum pooling, the spectrum is allocated to different operators in a centralized manner through a central controller or a resource manager. In [12], the authors propose an algorithm for the spectrum manager to allocate the spectrum between two networks within a single cell, which improves the effective spectrum utilisation while maintaining the quality of service (QoS) of each network under a fluctuating traffic environment. The work shows that sharing spectrum within a single cell by multiple operators could improve the spectrum utilisation efficiency compared with fixed spectrum allocation schemes.

However, this work does not guarantee fairness between the two networks in spectrum usage.

In [13], the authors investigate the spectrum sharing among MNOs at very short timescales with an instantaneous, dynamic sharing framework. A spectrum sharing manager regulates the resource allocation among the BSs of different MNOs. The proposed sharing scheme achieves increased spectral efficiency and reduced over-the-air latency compared to existing technologies (i.e., Wi-Fi). However, only a general coordination protocol is proposed without considering the wireless transmission specifications (e.g., path loss). In [14], the authors form a resource allocation problem that investigates the allocation of spectrum and BS density in an ultra-dense network (UDN) to minimize the cost of all operators while maintaining the average user rate. They illustrate the tradeoff between spectrum usage and the density of operating BSs during the peak hours. However, the case when both spectrum and BS density are limited is not studied.

Of relevance to this thesis, the authors in [15] reveal the superiority of non-orthogonal spectrum sharing over orthogonal spectrum sharing in terms of user throughput. Orthogonal spectrum sharing refers to the scenario when a certain frequency band cannot be simultaneously accessed by multiple operators. On the other hand, with non-orthogonal spectrum sharing, operators can transmit signals to their users through the same frequency band concurrently, which brings out the issue of inter-operator interference. In [16], the authors investigate the potential gain of non-orthogonal spectrum sharing with a multiple input, single-output model. The authors conclude that only a limited improvement can be achieved through non-orthogonal spectrum sharing. However, the authors assume that BSs from both operators are co-located in each cell, which is unlikely in a real world BS deployment.

Fairness among operators in accessing spectrum is another concern in spectrum sharing and of importance in this thesis. In [17], the author propose a scheme for MNOs to share their spectrum within indoor small cells. Spectrum from all MNOs assemble a spectrum pool to be shared fairly by all users in a small cell. The amount each MNO pays for the spectrum is proportional to the spectrum usage of its own users. The performance of the proposed scheme is limited by the co-channel interference; under maximum interference, the

spectrum sharing scheme does not improve the performance from the no sharing case.

1.3 Motivation and Contributions

Although importance of spectrum sharing is well documented, a cooperative spectrum sharing scheme that considers fairness between MNOs with real-world interference patterns has not been investigated. Furthermore, few works have investigated how the advent of physical layer technologies, such as cooperative MIMO systems, can help encourage spectrum sharing and enable maximizing the efficiency of spectrum usage. In addition, to the best of the author's knowledge, an investigation of channel estimation under a spectrum sharing framework (especially non-orthogonal spectrum sharing) is absent in the literature. In this regard, the main contributions of this paper are:

- Within the realm of cooperative spectrum pooling, we investigate non-orthogonal spectrum sharing under a deployment where BSs from different MNOs are not co-located for a multi-cell multi-user MIMO network. We then propose a spectrum sharing optimization framework with inter-operator coordinated beamforming between two MNOs which allows for simultaneous use by both while making fairness a key factor in determining the usage of the shared spectrum between the operators. Afterwards, we develop an algorithm based on fractional programming to manage resource allocation and user scheduling among the two MNOs. These well-established techniques provide smooth convergence to stationary points of the optimization problem. Lastly we show that both MNOs will have an incentive to participate in and benefit from the proposed spectrum sharing scheme.
- We consider the impact of channel estimation to the proposed fair licensed spectrum sharing framework, in a time division duplex (TDD) system. We incorporate channel estimation into the framework developed earlier, where UTs send pre-designed pilot sequences to the BSs per cell based. Moreover, we propose a cluster based pilot allocation mechanism where the pilot sequences are assigned to UTs within each newly organized

cluster, which is determined by a BS-centric clustering algorithm and a modified K-means clustering algorithm. Finally, we illustrate that once both operators can provide similar number of frequency slots, both of them can attain comparable gain to the case when perfect channel information is known, and thus have an incentive to take part in the proposed spectrum sharing framework.

- We briefly look into the distributive resource allocation and user scheduling at each MNO in spectrum sharing, with minimal information dissemination. Specifically, we propose a signal to leakage plus interference and noise ratio (SLINR) based optimization framework at each MNO. The results illustrate that the gain of spectrum sharing in this framework depends much on the cell deployment and the inter-operator interference power to the scheduled UTs.

It is worth mentioning that, in this thesis, spectrum is the **only** resource considered to be shared among MNOs. Allowing infrastructure sharing (e.g., user offloading, data transmission through BSs of the other MNO) is out of scope of this work.

1.4 Thesis Organization

The thesis is organized as follows: in Chapter 2, we present a centralized, coordinated fair spectrum sharing framework between two MNOs assuming perfect knowledge of the channel information. We formulate an optimization problem for joint resource allocation and user scheduling, which is solved through an iterative algorithm based on fractional programming. We incorporate channel estimation into the proposed spectrum sharing framework in Chapter 3. In Chapter 4, we briefly study the feasibility of decentralized spectrum sharing, where the resource allocation and user scheduling of each operator are determined individually with the same optimization framework presented in Chapter 2. Chapter 5 ends the thesis with some conclusions as well as discussions on future work.

1.5 Notations

Within this thesis, we use the lower case letters to denote scalars (e.g., a), bold lower-case letters to denote vectors (e.g., $\mathbf{v}$). We use bold upper-case letters to denote matrices (e.g., $\mathbf{D}$). For a vector $\mathbf{x}$, $\|\mathbf{x}\|$ refers to its Euclidean norm. For a matrix $\mathbf{A}$, $\mathbf{A}^{-1}$ denotes the inverse of the matrix, and $\mathbf{A}^\dagger$ refers to its conjugate transpose. With regard to sets of numbers, $\mathbb{R}$ denotes the set of real numbers, $\mathbb{C}$ denotes the set of complex numbers, and $\mathbb{Z}$ denotes the set of integers. We use the symbol $Re\{\cdot\}$ to indicate the real part of a complex term. Finally, $\mathbf{I}$ denotes an identity matrix.

Chapter 2

SINR-based Fair licensed Spectrum Sharing Between Two MNOs

2.1 Introduction

In this chapter, we deal with the problem of coordinated resource allocation and user scheduling between two MNOs, in the case of cooperative spectrum pooling. To ensure that both operators have the incentive to participate in spectrum sharing, we consider fairness within the resource allocation to prevent any of the operator from being exploited. The chapter starts from the description of the system model under investigation. We then formulate a constrained optimization problem to characterize centralized resource allocation and user scheduling in the system; this formulation is followed by our approach to solve it iteratively through fractional programming. In the last section of this chapter, we present simulation results and compare the system performance among various spectrum sharing mechanisms.

2.2 System Model

We consider the downlink of a multi-cell multi-user MIMO network comprising two operators sharing the same coverage area, denoted as OP1 and OP2 respectively. The spectrum of OP1 and OP2 can be divided into N_1 and N_2 frequency slots, each of equal bandwidth

individually. The operating area of each operator is divided into non-overlapping cells. Each operator has a base station (BS) located at the center of each cell to transmit signals to its own users within the cell using the frequency slots. We consider wraparound with a cluster of $\mathcal{B}$ cells to avoid network-edge issues and achieve more realistic results [18]. Finally, we assume there are K_1 users among all $\mathcal{B}$ cells of OP1 and K_2 users among all cells of OP2.

For either operator, each BS is equipped with M transmit antennas, and each user is equipped with one receive antenna. We use $\mathbf{H}_{1i,1i,j} = [\mathbf{h}_{1i,1i,j}^1, \mathbf{h}_{1i,1i,j}^2, \dots, \mathbf{h}_{1i,1i,j}^{K_1}] \in \mathbb{C}^{M \times K_1}$ to denote the downlink channel between the BS at cell i of OP1 and users at cell i of OP1 in the j^{th} frequency slot ($1 \leq j \leq N_1 + N_2$). Here, $\mathbf{h}_{1i,1i,j}^k \triangleq \sqrt{\beta_{1i,1i,j}^k} \cdot \mathbf{g}_{1i,1i,j}^k \in \mathbb{C}^{M \times 1}$, where $\beta_{1i,1i,j}^k = (\frac{d_{1i,1i}^k}{d_0})^{-\alpha}$ accounts for the path-loss between the BS and the k^{th} user belonging to OP1 and $\mathbf{g}_{1i,1i,j}^k \in \mathbb{C}^{M \times 1}$ accounts for the associated small-scale fading. Within the expression of path-loss, $d_{1i,1i}^k$ denotes the corresponding distances between the BS and the k^{th} user at the i^{th} cell, d_0 is the reference distance and α is the pass-loss exponent [19]. $g_{1i,1i,j}^{m,k}$, which denotes the small-scale fading between the m^{th} antenna and the k^{th} user, is a zero-mean, unit-variance complex Gaussian random variable, i.e., $g_{1i,1i,j}^{m,k} \sim \mathcal{CN}(0, 1)$. $\mathbf{h}_{2i,2i,j}^k$, $\mathbf{h}_{1i,2i,j}^k$ and $\mathbf{h}_{2i,1i,j}^k$ are similar, where the first subscript indicates which operator the BS belongs to while the second indicates the operator associated with the user¹.

With M transmit antennas per BS, within each cell, using linear processing, at most M users can be scheduled simultaneously in a single frequency slot [20]. Thus, there are up to M beamformers to be designed at each BS to facilitate the signal transmission to the scheduled users within each frequency slot. We use $\mathbf{v}_{1i,j}^m \in \mathbb{C}^{M \times 1}$ to denote the m^{th} beamformer at cell i with frequency slot j . Moreover, we use $\mathbf{u}_{1i,j}^m \in \mathbb{Z}^{K_1 \times 1}$ to illustrate the user scheduling variables on the m^{th} beam through the j^{th} frequency slot in the i^{th} cell. Specifically, $\mathbf{u}_{1i,j}^m$ is a length- K_1 vector of comprising $K - 1$ zeros and at most a single one indicating the scheduled user (i.e., if the k^{th} user is scheduled, then only the k^{th} entry of the vector $\mathbf{u}_{1i,j}^m$ is 1).

¹As mentioned, to focus on the performance upper bound, here we assume knowledge of all channels at a central controller. Investigations on estimating the channels and disseminating this information is presented in the following chapters of the thesis

2.3 Problem Formulation

Spectrum sharing between the two operators is implemented through spectrum pooling comprising the $N_1 + N_2$ frequency slots, which can be accessed by both operators in all cells. We use a central controller to manage the resource allocation and user scheduling among all cells from both operators. However, to impose fairness, the total number of frequency slots used by each operator within each cell simultaneously is restricted, i.e., the operator contributing less to the spectrum pool is restricted to using fewer frequency slots than the other.

Fairness can be guaranteed with an agreement between the participating operators on the usage of the shared spectrum. In our work, we let OP1 to be the operator that provides more frequency slots ($N_1 \geq N_2$) within the system, as such it has the right to access the whole spectrum pool (i.e., altogether $N_1 + N_2$ frequency slots). We impose proportional spectrum access fairness, i.e., OP2 can only use $\frac{N_2}{N_1}(N_2 + N_1)$ of the frequency slots. It is worth mentioning that, the proportional fairness here is simply one of many possible fairness metrics; any fairness condition, which incentivizes both operators to participate in spectrum sharing, can be applied.

We use $s_{1,i,j}, s_{2,i,j} \in \{0, 1\}$ to indicate that whether the j^{th} frequency slot is assigned to OP1 and OP2 respectively in the i^{th} cell. Note that, importantly, the same frequency slot *may be used by both operators*. Furthermore, it is not necessary for OP2 to use the same frequency slots in all its $\mathcal{B}$ cells. For each user, the *intra-operator* interference arises from the beams from all BSs within the $\mathcal{B}$ cells other than the beam on which the user is scheduled. Importantly, since both operators can access all slots, *inter-operator* interference arises from all $\mathcal{B}$ cells of the *other* operator.

Accordingly, the intra-operator interference to the m^{th} scheduled user through the j^{th} frequency slot in the i^{th} cell for OP1 and OP2 could be formulated respectively as:

$$I_{1,1_i,j} = \sum_{\substack{i'=1 \\ (i',m') \neq (i,m)}}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1_i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 \quad (2.1a)$$

$$I_{2,2_i,j} = \sum_{\substack{i'=1 \\ (i',m') \neq (i,m)}}^{\mathcal{B}} \sum_{m'=1}^M s_{2,i',j} |(\mathbf{H}_{2,i',2_i,j} \mathbf{u}_{2,i,j}^m)^\dagger \mathbf{v}_{2,i',j}^{m'}|^2 \quad (2.1b)$$

Here, $(\mathbf{H}_{1,i',1_i,j} \mathbf{u}_{1,i,j}^m)$ is a $M \times 1$ vector denoting the channel between the BS in the $(i')^{th}$ cell of the same operator and the scheduled user. For each user, only the inter-operator interference from the $\mathcal{B}$ closest BSs of the other operator is considered. Thus, the inter-operator interference shares a similar form as the intra-operator interference in (2.1):

$$I_{2,1_i,j} = \sum_{i''=1}^{\mathcal{B}} \sum_{m''=1}^M s_{2,i'',j} |(\mathbf{H}_{2,i'',1_i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{2,i'',j}^{m''}|^2 \quad (2.2a)$$

$$I_{1,2_i,j} = \sum_{i''=1}^{\mathcal{B}} \sum_{m''=1}^M s_{1,i'',j} |(\mathbf{H}_{1,i'',2_i,j} \mathbf{u}_{1,2,j}^m)^\dagger \mathbf{v}_{1,i'',j}^{m''}|^2 \quad (2.2b)$$

In this case, all the beams from the BSs of the other operator across all $\mathcal{B}$ cells cause interference. Note that if the j^{th} frequency slot is not allocated to the $(i')^{th}$ cell, e.g., $s_{1,i',j} = 0$, then there is no interference from this cell to all the other cells via this, the j^{th} frequency slot.

Using these interference expressions, the signal-to-interference-plus-noise ratio (SINR) for the user scheduled on the m^{th} beam on the j^{th} frequency slot in the i^{th} cell of OP1 is given by

$$\Gamma_{1,i,j}^m = \frac{s_{1,i,j} |(\mathbf{H}_{1,i,1_i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\sigma^2 + I_{1,1_i,j} + I_{2,1_i,j}} \quad (2.3)$$

with an associated data rate of $R_{1,i,j}^m = W \cdot \log(1 + \Gamma_{1,i,j}^m)$. A similar expression can be written for users of OP2 as well. Consequently, the sum-rate of the whole system is given by $\sum_{i,j,m} (R_{1,i,j}^m + R_{2,i,j}^m)$. Individual weights for the rate can be easily incorporated [21].

The central controller that manages the beamforming for each frequency slot in each cell and schedules users for both operators to maximize the sum-rate of all users within the system. Then, the objective function at hand is:

$$f_o(\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2) = \sum_{i=1}^B \sum_{j=1}^{N_1+N_2} \sum_{m=1}^M \left(\log\left(1 + \frac{s_{1,i,j} |(\mathbf{H}_{1,1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\sigma^2 + I_{1,1,i,j} + I_{2,1,i,j}}\right) \right. \\ \left. + \log\left(1 + \frac{s_{2,i,j} |(\mathbf{H}_{2,2,i,j} \mathbf{u}_{2,i,j}^m)^\dagger \mathbf{v}_{2,i,j}^m|^2}{\sigma^2 + I_{2,2,i,j} + I_{1,2,i,j}}\right) \right) \quad (2.4)$$

Optimal resource allocation and user scheduling, therefore, requires the solution to the following constrained optimization problem:

$$\max_{\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2} f_o(\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2) \quad (2.5a)$$

$$\text{subject to} \quad \sum_{m=1}^M \|\mathbf{v}_{1,i,j}^m\|^2 \leq P, \quad \forall i, j; \quad (2.5b)$$

$$\sum_{m=1}^M \|\mathbf{v}_{2,i,j}^m\|^2 \leq P, \quad \forall i, j; \quad (2.5c)$$

$$\sum_{k=1}^{K_1} (u_{1,i,j}^m)_k = 1, \quad \forall i, j, m; \quad (2.5d)$$

$$\sum_{k=1}^{K_2} (u_{2,i,j}^m)_k = 1, \quad \forall i, j, m; \quad (2.5e)$$

$$\sum_{m=1}^M (u_{1,i,j}^m)_k \leq 1, \quad \forall i, j, k; \quad (2.5f)$$

$$\sum_{m=1}^M (u_{2,i,j}^m)_k \leq 1, \quad \forall i, j, k; \quad (2.5g)$$

$$\sum_{j=1}^{N_1+N_2} s_{1,i,j} \leq N_1 + N_2, \quad \forall i; \quad (2.5h)$$

$$\sum_{j=1}^{N_1+N_2} s_{2,i,j} \leq \lfloor \frac{N_2}{N_1} \cdot (N_1 + N_2) \rfloor, \quad \forall i \quad (2.5i)$$

$$(u_{1,i,j}^m)_k, (u_{2,i,j}^m)_k, s_{1,i,j}, s_{2,i,j} \in \{0, 1\} \quad (2.5j)$$

Among the constraints in (2.5), (2.5b) and (2.5c) impose the power constraint on each frequency slot from both operators. As the power spectral density (PSD) is determined by the regulator, we assume a maximal power of P per frequency slot is used commonly for

both operators. The constraints (2.5d) and (2.5e) enforce that each beam is allocated to at most one user. Moreover, the constraints (2.5f) and (2.5g) ensure that one user can only be assigned to at most one beam within each frequency slot. (2.5h) and (2.5i) restrict the spectrum usage by each operator to guarantee the fairness between the two participating operators (while (2.5h) is redundant, we include it for completeness.)

2.4 Proposed Solution Approach

Solving the above constrained optimization problem is challenging since the objective function (2.5a) is non-convex on the beamformer $\mathbf{v}$ and discontinuous on $\mathbf{u}$ and $\mathbf{s}$. Fortunately, we can formulate an iterative optimization scheme to tackle the problem based on the well-established technique of fractional programming [22, 21]. By means of fractional programming, the objective function (2.5a) can be made concave. Fractional programming uses the fact that the following sum-of-logarithms optimization problem:

$$\max_{\mathbf{x}} f(\mathbf{x}) = \sum_i \log \left(1 + \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})} \right)$$

has an equivalent sum-of-ratios form:

$$\max_{\mathbf{x}, \gamma} f_r(\mathbf{x}, \gamma) = \sum_i \left(\log(1 + \gamma_i) - \gamma_i + \frac{(1 + \gamma_i)A_i(\mathbf{x})}{A_i(\mathbf{x}) + B_i(\mathbf{x})} \right) \quad (2.6)$$

where $\gamma_i = \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})}$ is an auxiliary variable [22].

Proof. By introducing an auxiliary variable $\gamma_i = \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})}$, the original sum-of-logarithms optimization problem (2.4) can be rewritten as:

$$\begin{aligned} \max_{\mathbf{x}, \gamma} \quad & f(\gamma) = \sum_i \log(1 + \gamma_i) \\ \text{subject to} \quad & \gamma_i = \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})} \end{aligned} \quad (2.7)$$

The Lagrange dual function of the optimization problem (2.7) is:

$$\begin{aligned} g(\boldsymbol{\nu}) &= \max_{\mathbf{x}, \boldsymbol{\gamma}} \mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\nu}) \\ &= \max_{\mathbf{x}, \boldsymbol{\gamma}} \sum_i \left(\log(1 + \gamma_i) - \nu_i \left(\gamma_i - \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})} \right) \right) \end{aligned} \quad (2.8)$$

Since $\mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\nu})$ in (2.8) is concave differentiable over γ_i when $\mathbf{x}$ is fixed, the optimal γ_i^* can be achieved by setting $\frac{\partial \mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\nu})}{\partial \gamma_i} = 0$, which yields $\gamma_i^* = \frac{1}{\nu_i} - 1$. Replacing the optimal γ_i^* back into (2.8), yields:

$$g(\boldsymbol{\nu}) = \sum_i \left(\log \left(\frac{1}{\nu_i} \right) - \nu_i \left(\frac{1}{\nu_i} - 1 - \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})} \right) \right) \quad (2.9)$$

The dual optimization problem of (2.7) is then:

$$\min_{\boldsymbol{\nu}} g(\boldsymbol{\nu}) = \min_{\boldsymbol{\nu}} \sum_i \left(\log \left(\frac{1}{\nu_i} \right) - \nu_i \left(\frac{1}{\nu_i} - 1 - \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})} \right) \right) \quad (2.10)$$

Since $g(\boldsymbol{\nu})$ in (2.10) is convex differentiable over ν_i , the optimal ν_i^* can be achieved by setting $\frac{\partial g(\boldsymbol{\nu})}{\partial \nu_i} = 0$, which yields $\nu_i^* = \frac{B_i(x)}{A_i(x) + B_i(x)}$.

Since there is only one equality constraint in (2.7), Slater's condition is satisfied and thus strong duality holds, which implies:

$$\begin{aligned} \max_{\mathbf{x}, \boldsymbol{\gamma}} f(\boldsymbol{\gamma}) &= \max_{\mathbf{x}, \boldsymbol{\gamma}} \mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\nu}^*) \\ &= \max_{\mathbf{x}, \boldsymbol{\gamma}} \sum_i \left(\log(1 + \gamma_i) - \gamma_i + \frac{(1 + \gamma_i)A_i(\mathbf{x})}{A_i(\mathbf{x}) + B_i(\mathbf{x})} \right) \end{aligned} \quad (2.11)$$

□

Thus, the objective function (2.4) can be converted into the form of sum-of-ratios as:

$$\begin{aligned}
 f_r(\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2) = & \sum_{i=1}^{\mathcal{B}} \sum_{j=1}^{N_1+N_2} \sum_{m=1}^M \left(\log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m + \log(1 + \gamma_{2,i,j}^m) - \gamma_{2,i,j}^m \right. \\
 & + \frac{(1 + \gamma_{1,i,j}^m) s_{1,i,j} |(\mathbf{H}_{1i,1i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\sigma^2 + \sum_{i'=1}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1i',1i,j} \mathbf{u}_{1,i',j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + I_{2,1i,j}} \\
 & \left. + \frac{(1 + \gamma_{2,i,j}^m) s_{2,i,j} |(\mathbf{H}_{2i,2i,j} \mathbf{u}_{2,i,j}^m)^\dagger \mathbf{v}_{2,i,j}^m|^2}{\sigma^2 + \sum_{i'=1}^{\mathcal{B}} \sum_{m'=1}^M s_{2,i',j} |(\mathbf{H}_{2i',2i,j} \mathbf{u}_{2,i',j}^m)^\dagger \mathbf{v}_{2,i',j}^{m'}|^2 + I_{1,2i,j}} \right)
 \end{aligned} \tag{2.12}$$

where $I_{2,1i,j}$ and $I_{1,2i,j}$ are specified in (2.2a) and (2.2b) respectively, and,

$$\begin{aligned}
 \gamma_{1,i,j}^m = & \frac{s_{1,i,j} |(\mathbf{H}_{1i,1i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\sigma^2 + \sum_{\substack{i'=1 \\ (i',m') \neq (i,m)}}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1i',1i,j} \mathbf{u}_{1,i',j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2} \\
 & + \sum_{i''=1}^{\mathcal{B}} \sum_{m''=1}^M s_{2,i'',j} |(\mathbf{H}_{2i'',1i,j} \mathbf{u}_{1,i'',j}^m)^\dagger \mathbf{v}_{2,i'',j}^{m''}|^2
 \end{aligned} \tag{2.13}$$

The value of $\gamma_{2,i,j}^m$ can be obtained similarly as $\gamma_{1,i,j}^m$ in (2.13). Note that the new objective function $f_r(\cdot)$ is not convex at this point, as the variables (i.e., $\mathbf{s}$, $\mathbf{u}$, $\mathbf{v}$) appear in both the numerators and denominators among the fractions. Thus, $f_r(\cdot)$ needs to be further transformed.

Quadratic Transform for Multiple-Ratio Fractional Programming: the following sum-of-ratio optimization problem:

$$\max_{\mathbf{x}} \quad \sum_i \frac{A_i(\mathbf{x})}{B_i(\mathbf{x})} \tag{2.14a}$$

$$\text{subject to} \quad \mathbf{x} \in \mathcal{X} \tag{2.14b}$$

is equivalent to the convex optimization problem:

$$\max_{\mathbf{x}, \mathbf{y}} \quad \sum_i (2y_i \sqrt{A_i(\mathbf{x})} - y_i^2 B_i(\mathbf{x})) \tag{2.15a}$$

$$\text{subject to} \quad \mathbf{x} \in \mathcal{X} \tag{2.15b}$$

where $\mathbf{y}$ is an auxiliary variable. Here, $A_i(\mathbf{x})$ is a non-negative function and $B_i(\mathbf{x})$ is a

positive function. With the optimal auxiliary variable $\mathbf{y}^*$, both (2.14) and (2.15) share the same optimal solution $\mathbf{x}^*$ as well as the same optimum of the objective functions.

The detailed proof of the preceding quadratic transform is presented in [22]. Thus, by incorporating the quadratic transform to the sum of ratios terms, the objective function in (2.12) can be reformulated correspondingly as:

$$\begin{aligned}
f_q(\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2, \gamma_1, \gamma_2, \mathbf{y}_1, \mathbf{y}_2) = & \sum_{i=1}^B \sum_{j=1}^{N_1+N_2} \sum_{m=1}^M \left\{ \left(\log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m \right. \right. \\
& + 2Re\{(y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,1,j} \mathbf{u}_{1,i,j}^m)\} - |y_{1,i,j}^m|^2 \sigma^2 \\
& - |y_{1,i,j}^m|^2 \left(\sum_{i'=1}^B \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \sum_{i''=1}^B \sum_{m''=1}^M s_{2,i'',j} |(\mathbf{H}_{2,i'',1,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{2,i'',j}^{m''}|^2 \right) \Big) \\
& + \left(\log(1 + \gamma_{2,i,j}^m) - \gamma_{2,i,j}^m + 2Re\{y_{2,i,j}^m \sqrt{(1 + \gamma_{2,i,j}^m)} s_{2,i,j} (\mathbf{v}_{2,i,j}^m)^\dagger (\mathbf{H}_{2,i,2,j} \mathbf{u}_{2,i,j}^m)\} - |y_{2,i,j}^m|^2 \sigma^2 \right. \\
& \left. \left. - |y_{2,i,j}^m|^2 \left(\sum_{i'=1}^B \sum_{m'=1}^M s_{2,i',j} |(\mathbf{H}_{2,i',2,j} \mathbf{u}_{2,i,j}^m)^\dagger \mathbf{v}_{2,i',j}^{m'}|^2 + \sum_{i''=1}^B \sum_{m''=1}^M s_{1,i'',j} |(\mathbf{H}_{1,i'',2,j} \mathbf{u}_{2,i,j}^m)^\dagger \mathbf{v}_{1,i'',j}^{m''}|^2 \right) \right) \right\} \quad (2.16)
\end{aligned}$$

Here, f_q is a concave function of $\mathbf{y}$, such that the optimal solution for $y_{1,i,j}^m$ can be achieved by setting $\frac{\partial f_q}{\partial y_{1,i,j}^m} = 0$ with all other variables fixed, i.e.,

$$y_{1,i,j}^m = \frac{s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,1,j} \mathbf{u}_{1,i,j}^m) \sqrt{(1 + \gamma_{1,i,j}^m)}}{\left(\sigma^2 + \sum_{i'=1}^B \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \sum_{i''=1}^B \sum_{m''=1}^M s_{2,i'',j} |(\mathbf{H}_{2,i'',1,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{2,i'',j}^{m''}|^2 \right)} \quad (2.17)$$

A corresponding expression for $y_{2,i,j}^m$ can be attained. The objective function f_q is concave in $\mathbf{v}$ as well. Thus, with fixed $\mathbf{s}$, $\mathbf{u}$, γ and $\mathbf{y}$, the optimal solution of $\mathbf{v}$ can be achieved by solving the following constrained convex optimization problem:

$$\max_{\mathbf{v}_1, \mathbf{v}_2} f_q(\mathbf{v}_1, \mathbf{v}_2) \quad (2.18a)$$

$$\text{subject to} \quad (2.5b), (2.5c) \quad (2.18b)$$

The method of Lagrange multipliers allows us to solve the convex optimization problem (2.18),

where the Lagrangian over $\mathbf{v}$ is

$$\mathcal{L}_q(\mathbf{v}_1, \mathbf{v}_2) = f_q(\mathbf{v}_1, \mathbf{v}_2, \dots) - \sum_{i,j} \left(\lambda_{1,i,j} \left(\sum_{m=1}^M (\mathbf{v}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m - P \right) + (\lambda_{2,i,j} \sum_{m=1}^M (\mathbf{v}_{2,i,j}^m)^\dagger \mathbf{v}_{2,i,j}^m - P) \right) \quad (2.19)$$

By solving $\frac{\partial \mathcal{L}_q}{\partial (\mathbf{v}_{1,i,j}^m)^\dagger} = 0$, we can achieve a closed form solution for the vector $\mathbf{v}_{1,i,j}^m$:

$$\mathbf{v}_{1,i,j}^m = s_{1,i,j} \mathbf{D}_{1,i,j}^{-1} \sqrt{(1 + \gamma_{1,i,j}^m)} (y_{1,i,j}^m)^\dagger (\mathbf{H}_{1,1,i,j} \mathbf{u}_{1,i,j}^m) \quad (2.20)$$

where

$$\begin{aligned} \mathbf{D}_{1,i,j} = & \lambda_{1,i,j} \mathbf{I} + s_{1,i,j} \sum_{i'=1}^B \sum_{m'=1}^M |y_{1,i',j}^{m'}|^2 (\mathbf{H}_{1,1,i',j} \mathbf{u}_{1,i',j}^{m'}) (\mathbf{H}_{1,1,i',j} \mathbf{u}_{1,i',j}^{m'})^\dagger \\ & + s_{1,i,j} \sum_{i''=1}^B \sum_{m''=1}^M |y_{2,i'',j}^{m''}|^2 (\mathbf{H}_{1,2,i'',j} \mathbf{u}_{2,i'',j}^{m''}) (\mathbf{H}_{1,2,i'',j} \mathbf{u}_{2,i'',j}^{m''})^\dagger \end{aligned} \quad (2.21)$$

The Lagrange multiplier $\lambda_{1,i,j}$ can be obtained through bi-section search on the power. The optimal $\mathbf{v}_{2,i,j}^m$ can be obtained similarly.

Since the interference to each scheduled user depends only on the beamformers at each BS regardless of the users scheduled in the beams, the optimization of user scheduling can be decoupled from the optimization of beamformers. Importantly, the user scheduling variable $\mathbf{u}$ can be optimized individually keeping the other variables fixed. Specifically, optimal user scheduling $\mathbf{u}^*$ can be achieved through the following integer optimization problem with respect to $\mathbf{u}$:

$$\max_{\mathbf{u}_1, \mathbf{u}_2} f_q(\mathbf{u}_1, \mathbf{u}_2) \quad (2.22a)$$

$$\text{subject to} \quad (2.5d), (2.5e), (2.5f), (2.5g) \quad (2.22b)$$

where all the other variables are kept fixed. To help simplify the objective function (2.22a), we introduce a new function $\mathcal{R}(\cdot)$ to include all terms and variables within f_q with regard

to the i^{th} cell and the j^{th} frequency slot of a single operator,

$$\begin{aligned}
\mathcal{R}(\mathbf{u}_{1,i,j}) = & \sum_{m=1}^M \left\{ \log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m - |y_{1,i,j}^m|^2 \sigma^2 \right. \\
& + 2Re\{(y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,1,i,j} \mathbf{u}_{1,i,j}^m)\} \\
& - |y_{1,i,j}^m|^2 \left(\sum_{i'=1}^B \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 \right. \\
& \left. \left. + \sum_{i''=1}^B \sum_{m''=1}^M s_{2,i'',j} |(\mathbf{H}_{2,i'',1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{2,i'',j}^{m''}|^2 \right) \right\}
\end{aligned} \tag{2.23}$$

such that f_q could be re-written as:

$$f_q(\mathbf{u}_1, \mathbf{u}_2) = \sum_{i=1}^B \sum_{j=1}^{N_1+N_2} (\mathcal{R}(\mathbf{u}_{1,i,j}) + \mathcal{R}(\mathbf{u}_{2,i,j})) \tag{2.24}$$

Since $\mathcal{R}(\mathbf{u}_{1,i,j})$ and $\mathcal{R}(\mathbf{u}_{2,i,j})$ are independent of the user scheduling in the other cells and frequency slots of both operators, the optimal solution for (2.22a) can be obtained by solving $\max_{\mathbf{u}_{1,i,j}} \mathcal{R}(\mathbf{u}_{1,i,j})$ and $\max_{\mathbf{u}_{2,i,j}} \mathcal{R}(\mathbf{u}_{2,i,j})$, $\forall i, j$ separately with constraints (2.5d) (2.5e) (2.5f) (2.5g). In fact, each of the above constrained optimization problems on $\mathcal{R}(\cdot)$ can be described as a maximum weighted bipartite matching problem where we need to obtain the optimal one-to-one matches between the beams and the users, as shown in Figure 2.1.

In the complete bipartite graph, the nodes in the right denote the M beams and the nodes in the left denote the users to be scheduled. The weight of the edge between the k^{th} user node and the m^{th} beam node is $W_{k,m}$. Considering the optimization for the i^{th} cell and j^{th} frequency slot of OP1 we have,

$$\begin{aligned}
W_{k,m} = & \log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m - |y_{1,i,j}^m|^2 \sigma^2 + 2Re\{(y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,1,i,j} \hat{\mathbf{u}}^k)\} \\
& - |y_{1,i,j}^m|^2 \left(\sum_{i'=1}^B \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1,i,j} \hat{\mathbf{u}}^k)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \sum_{i''=1}^B \sum_{m''=1}^M s_{2,i'',j} |(\mathbf{H}_{2,i'',1,i,j} \hat{\mathbf{u}}^k)^\dagger \mathbf{v}_{2,i'',j}^{m''}|^2 \right)
\end{aligned} \tag{2.25}$$

Unlike in (2.23), $\hat{\mathbf{u}}^k$ is used instead of $\mathbf{u}_{1,i,j}^m$ in (2.25). Here, $\hat{\mathbf{u}}^k \in \mathbb{Z}_2^{K_1 \times 1}$ is a unit vector with the k^{th} element equaling to one. Using this vector, $W_{k,m}$ represents the contribution to the

Figure 2.1: Example bipartite graph between users and beams

objective function $\mathcal{R}(\cdot)$ from the m^{th} beam if the k^{th} user is scheduled with it.

The maximum weighted bipartite matching problem has been extensively studied in the realm of combinatorics [23]. It can be solved effectively through algorithms with polynomial-time complexity, e.g., the Hungarian algorithm [24] and the auction algorithm [25]. As in [21], with the Hungarian algorithm, the proposed bipartite matching problem can be solve in polynomial time with a complexity of $\mathcal{O}(M^2K)$. Because of the M one-to-one matches between the user nodes and the beam nodes, the solution satisfies the constraints (2.5d) (2.5e) (2.5f) (2.5g), i.e., each beam must have a user assigned and one user can only be scheduled with at most one beam. Based on the solution, the optimal $(\mathbf{u}_{1,i,j}^n)^*$ can be achieved by setting one to the k^{th} term from a zero vector of length K_1 , if the k^{th} user is scheduled within the m^{th} beam.

With regard to the optimal selection of frequency slots $\mathbf{s}^*$ that maximizes f_q , we observe that f_q can be rearranged into a linear form with respect to $\mathbf{s}$ when the other variables are fixed:

$$f_q(\mathbf{s}_1, \mathbf{s}_2) = \sum_{i=1}^B \sum_{j=1}^N \mathcal{Q}_{1,i,j} s_{1,i,j} + \mathcal{Q}_{2,i,j} s_{2,i,j} + const \quad (2.26)$$

where

$$\begin{aligned}
\mathcal{Q}_{1,i,j} = & \sum_{m=1}^M 2\text{Re}\{(y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,j} \mathbf{u}_{1,i,j}^m)\} \\
& - \sum_{i'=1}^B \sum_{m'=1}^M |y_{1,i',j}^{m'}|^2 \sum_{m=1}^M |(\mathbf{H}_{1,i',j} \mathbf{u}_{1,i',j}^{m'})^\dagger \mathbf{v}_{1,i,j}^m|^2 \\
& - \sum_{i''=1}^B \sum_{m''=1}^M |y_{2,i'',j}^{m''}|^2 \sum_{m=1}^M |(\mathbf{H}_{1,2,i'',j} \mathbf{u}_{2,i'',j}^{m''})^\dagger \mathbf{v}_{1,i,j}^m|^2
\end{aligned} \tag{2.27}$$

Importantly, this implies that the optimizations on $\mathbf{s}_1$ and $\mathbf{s}_2$ can be independent, since the objective function (2.26) is linear in both $\mathbf{s}_1$ and $\mathbf{s}_2$. Meanwhile, the constraints (2.5h) and (2.5i) regulate the usage of the spectrum for the two operators individually. Then the optimal $\mathbf{s}_1^*$ and $\mathbf{s}_2^*$ can be achieved by solving the following optimization problems individually:

$$\max_{\mathbf{s}_1} \quad \sum_{i=1}^B \sum_{j=1}^N \mathcal{Q}_{1,i,j} s_{1,i,j} \tag{2.28a}$$

$$\text{subject to} \quad (2.5h), s_{1,i,j} \in \{0, 1\} \tag{2.28b}$$

and

$$\max_{\mathbf{s}_2} \quad \sum_{i=1}^B \sum_{j=1}^N \mathcal{Q}_{2,i,j} s_{2,i,j} \tag{2.29a}$$

$$\text{subject to} \quad (2.5i), s_{2,i,j} \in \{0, 1\} \tag{2.29b}$$

(2.28) and (2.29) can be solved effectively through integer programming. However, the optimal $\mathbf{s}_1$ and $\mathbf{s}_2$ can also be achieved in a greedy manner per cell. We first sort the frequency slots in descending order with respect to the values of $\mathcal{Q}$, and then set the corresponding $s = 1$ if $\mathcal{Q}$ is positive for the first N_c sorted frequency slots. Here, N_c denotes the maximum number of frequency slots assigned to an operator, i.e.,

$$N_c = \begin{cases} N_1 + N_2, & \text{OP1} \\ \lfloor \frac{N_2}{N_1} \cdot (N_1 + N_2) \rfloor, & \text{OP2} \end{cases} \tag{2.30}$$

Claim. *For the problems in (2.28) and (2.29), i.e., keeping all other variables fixed, the frequency slot selection variable $\mathbf{s}^*$ obtained through the proposed greedy algorithm above is the globally optimal solution.*

Proof. Assume there exists one frequency slot selection $\mathbf{s}'$ such that it reaches a superior outcome than $\mathbf{s}^*$. This would imply

$$\sum_{j=1}^N Q_j s'_j > \sum_{j=1}^N Q_j s_j^* \quad (2.31)$$

The relationship between the left term and the right term in (2.31) can be understood through the sign of the difference between them, i.e., $\sum_{j=1}^N Q_j s'_j - \sum_{j=1}^N Q_j s_j^*$, which could be re-formulated as the following,

$$\sum_{\substack{j=1, \\ Q_j > 0}}^N Q_j (s'_j - s_j^*) + \sum_{\substack{j=1, \\ Q_j < 0}}^N Q_j (s'_j - s_j^*) \quad (2.32)$$

In (2.32), the terms with positive and negative values of Q are separated. The first scenario to be considered is when the number of positive Q is no more than N_c , i.e., $\sum_{j=1}^N Q_j (Q_j > 0) \leq N_c$. In this case, $\mathbf{s}_j^* = 1$ only if $Q_j > 0$ according to the proposed greedy approach. Then, $\sum_{\substack{j=1, \\ Q_j > 0}}^N Q_j (s'_j - s_j^*) = \sum_{\substack{j=1, \\ Q_j > 0}}^N Q_j (s'_j - 1) \leq 0$, since $s'_j \leq 1$ and $Q_j > 0$. Identically, $\sum_{\substack{j=1, \\ Q_j < 0}}^N Q_j (s'_j - s_j^*) = \sum_{\substack{j=1, \\ Q_j < 0}}^N Q_j (s'_j - 0) \leq 0$, since $s'_j \geq 0$ and $Q_j < 0$ here. Thus, the value of (2.32) is not positive when $\sum_{j=1}^N Q_j (Q_j > 0) \leq N_c$.

The other scenario is when there are more than N_c positive Q s, i.e., $\sum_{j=1}^N Q_j (Q_j > 0) > N_c$. Similar as the previous case, $\mathbf{s}_j^* = 1$ only if $Q_j > 0$, such that $\sum_{\substack{j=1, \\ Q_j < 0}}^N Q_j (s'_j - s_j^*) \leq 0$. Moreover, since $\mathbf{s}_j^* = 1$ only if its corresponding Q_j is among the top N_c largest, $\sum_{\substack{j=1, \\ Q_j > 0}}^N Q_j (s'_j - s_j^*) = \sum_{\substack{j=1, \\ Q_j > 0}}^N Q_j s'_j - \sum_{\substack{j=1, \\ N_c \text{ largest } Q_j}}^{N_c} Q_j s_j^* \leq 0$. Hence, the value of (2.32) is not positive as well when $\sum_{j=1}^N Q_j (Q_j > 0) > N_c$. In conclusion, $\sum_{j=1}^N Q_j s'_j - \sum_{j=1}^N Q_j s_j^* \leq 0$, which contradicts (2.31). Therefore, there is no solution better than $\mathbf{s}^*$. $\square$

Based on the optimization of each individual variable as indicated above, we can gather all the optimization steps into an iterative optimization algorithm, as presented in Algorithm 1.

The algorithm implements block coordinate descent to determine the resource allocation and user scheduling within the system. In each step, one type of variable is optimized keeping the others fixed. Since the objective function is upper bounded and each step can only improve on the value of the objective function, f_q , the algorithm converges in a non-decreasing manner.

Algorithm 1: Joint Resource Allocation and User Scheduling

```

1: Initialize  $\mathbf{s}, \mathbf{u}, \mathbf{v}$ 
2: while the objective function  $f_q$  in (2.16) does not converge do
3:   for  $i = 1, 2, \dots, \mathcal{B}$  do
4:     Update  $\gamma_i$  using (2.13).
5:     Update  $\mathbf{y}_i$  using (2.17).
6:     Update  $\mathbf{v}_i$  using (2.20) and update  $\lambda_i$  with bisection search.
7:     Update  $\mathbf{u}_i$  using Hungarian Algorithm
8:     Update  $\mathbf{s}_i$  by solving (2.28) and (2.29).
9:   end for
10: end while

```

In Algorithm 1, for a single iteration, the time complexity of Step 4 - 6 is $\mathcal{O}(\mathcal{B}NM)$, as the corresponding parameters (i.e., γ , y , and v) are updated for every user in all cells with all frequency slots. The Hungarian algorithm is applied in user scheduling in each cell and with each frequency slot, resulting in the total time complexity of $\mathcal{O}(\mathcal{B}NM^2K)$ for Step 7. In Step 8, for each cell, the value of $\mathcal{Q}$ is calculated per frequency slot. The values of $\mathcal{Q}$ s are then sorted in the greedy algorithm used to determine the optimal values of $\mathbf{s}$. Thus in total, the time complexity for Step 8 is $\mathcal{O}(\mathcal{B}(N + N \log(N)))$. Consequently, the computational complexity of the proposed algorithm is given by $\mathcal{O}(\mathcal{B}NM + \mathcal{B}NM^2K + \mathcal{B}(N + N \log(N)))$ for every iteration.

To summarize our contribution here, we have formulated an optimization problem that encourages spectrum sharing while imposing fairness constraints. The problem is solved using block coordinate descent as listed in Algorithm 1. The output of the algorithm is, in each cell, the frequency slots used by each operator, the users (of each operator) scheduled and the associated beamforming weights (which, implicitly includes power allocation).

2.5 Simulation Results and Performance Evaluation

In this section, we illustrate the gains made possible by our resource allocation scheme. The numerical values of the parameters used in the simulations are given in Table 2.1. We fix the number of frequency slots provided by OP1 at 16 (i.e., $N_1 = 16$), and vary the number of frequency slots provided by OP2 from 4 up to 16; this allows us to investigate the impact of unequal contributions to the spectrum pool. Note that, as per the constraint in (2.5i), if it contributes fewer than 4 frequency slots, OP2 would not receive any additional frequency slots, i.e., operators must contribute some minimum amount for a fair spectrum sharing to be possible. The number of users from OP1 and OP2 within each cell is set to be M times the number of frequency slots provided by OP1 and OP2 respectively (i.e., $K_1 = M \times N_1$ and $K_2 = M \times N_2$). In our case, we assume each BS is deployed with $M = 8$ antennas.

Table 2.1: System parameter settings used.

Cell size per side	2000 m
BS transmit power	$P = 43$ dBm
Frequency slot bandwidth	1.25 MHz
Noise power spectral density	- 174 dBm/Hz
Path-loss exponent	$\alpha = 3.7$
Reference distance	$d_0 = 0.4$ m

To evaluate the performance of the proposed spectrum sharing mechanism, we compare the *average* user rate (defined as (sum rate)/total number of users among $\mathcal{B}$ cells) achieved through the proposed framework with the following cases:

- *No spectrum sharing*: Each operator uses its own spectrum exclusively as in the conventional wireless communications framework. The resource allocation and user scheduling for each operator are achieved *individually* using the algorithm presented in [21].
- *Spectrum sharing with same number of frequency slots as provided*: The spectrum pool is divided into two disjoint sets comprising N_1 and N_2 frequency slots respectively, which are then allocated to the operators in accordance with the number of frequency slots they provide. Note that, in this case, the same frequency slots are allocated across all $\mathcal{B}$ cells, such that there is no inter-operator interference (as in the no spectrum sharing

case). Effectively, in this scheme, the operators exchange their frequency slots in order to achieve maximum rate across all cells.

- *Full spectrum sharing:* Both operators can transmit signals over all frequency slots together in this case, ignoring the contributions to the spectrum pool. In this case, the constraints (2.5i) is turned into $\sum_{j=1}^{N_1+N_2} s_{2,i,j} \leq N_1 + N_2, \quad \forall i$, as the constraint (2.5h).

Due to the discrete variables (i.e., $\mathbf{u}$, $\mathbf{s}$) within the optimization problem, the optimal solution is sensitive to the initialization of the variables. Based on a suggestion in [21], we use a greedy initialization for $\mathbf{u}$ based on the channels between the BSs and the users, i.e., the M users with the best M channels are scheduled within each frequency slot. The beamforming vector $\mathbf{v}$ for each scheduled user is matched to the channel, i.e., proportional to the corresponding channel and normalized to satisfy the total power constraint per frequency slot.

Figure 2.2: Convergence of the objective function in (2.4) using Algorithm 1.

The objective function has an upper bound and each step in the algorithm improves its value, hence, the algorithm converges. In our experience, the proposed algorithm converges within 20 iterations of updating the variables among the $\mathcal{B}$ cells of both operators; Figure 2.2

represents a typical convergence curve for one realization of channels and users, i.e., one run of the algorithm.

2.5.1 Diagonally Overlapped Cell Deployment

We firstly look into the improvement made possible by the proposed spectrum sharing schemes with the diagonally overlapped cell deployment, where the BSs of one operator stand at the corners of the cells of the other operator, as illustrated in Figure 2.3.

Figure 2.3: Cell deployment for operators ($\mathcal{B} = 9$).

With wraparound, a cluster of $\mathcal{B}$ cells is surrounded by another $\mathcal{B} - 1$ cloned clusters, and so on (Figure 2.3 uses $\mathcal{B} = 9$). Thus, the cellular layout of each operator consists of infinitely duplicated clusters with identical distributions of users, as well as channels between each BS and each user. The distribution of users within each cell of a cluster is independent of each other. In addition, all wireless channels between BSs and users in a cluster are assumed independent.

Figure 2.4 plots the average user rate among all users of both operators across all cells. Specifically, the figure plots the average user rate as a function of the number of frequency slots contributed by OP2 (N_2). We observe that the proposed spectrum sharing mechanism

Figure 2.4: Average user rate among 9 cells of both operators (Diagonally overlapped cell deployment).

can increase the user average rate by up to 60%, compared to the case where there is no spectrum sharing, when both operators make equal contributions to the spectrum pool ($N_1 = N_2 = 16$). When the number of frequency slots provided by the two operators are not equal, e.g., when $N_2 = 8$, fair spectrum sharing can still increase the achievable user rate by approximately 25%. Importantly, our algorithm, while implementing fairness, achieves 80% of the average user rate of full spectrum sharing (without fairness constraints). In addition, due to increased frequency diversity, the performance of spectrum sharing increases as the number of frequency slots in the spectrum pool increases; by sharing, the operators gain the additional benefit of an increased number of frequency slots assigned to both operators. Importantly, as the “Spectrum Sharing Same Usage” case shows, frequency diversity by itself provides little gain over the “No Sharing” case. This implies that merely reorganizing spectrum occupation provides limited gains.

The average user rate among users from OP1 and OP2 are presented in Figure 2.5a and Figure 2.5b, respectively. Moreover, the gain of fair spectrum sharing and full spectrum sharing over no spectrum sharing in terms of the user average rate of each operator is given in Figure 2.6. Interestingly, under the proposed fair spectrum sharing scheme, it is OP1, not

(a) Average user rate among 9 cells of OP1.

(b) Average user rate among 9 cells of OP2.

Figure 2.5: Average user rate among 9 cells of each operator (Diagonally overlapped cell deployment).

(a) Rate gain with fair spectrum sharing.

(b) Rate gain with full spectrum sharing.

Figure 2.6: Rate gain compared with no spectrum sharing.

OP2, that experiences a higher gain compared with the full sharing case. This is because, under fair spectrum sharing, OP1 will still have nearly exclusive access to several frequency slots in each cell due to the restrictions on the usage of OP2. In addition, OP1 always achieves a better gain than OP2. When there is a wide difference between the number of frequency slots provided by OP2 (N_2) and OP1 (N_1), the average user rate of OP2 *decreases* compared to the “No Sharing” case. This is because under this circumstance, there is more benefit towards the objective function f_o through minimizing the inter-operator interference from OP2 to OP1, compared to maximizing the rate contributed by OP2. This suggests OP1 has an incentive to participate even if it contributes more than OP2. Nonetheless, as the contribution of the spectrum pool from OP2 expands, its gain in average user rate raises accordingly. Thus, the proposed spectrum sharing scheme encourages the powerless operator (i.e., OP2 in our system) to provide as much spectrum as possible. On the other hand, the proposed spectrum sharing framework prevents the powerful operator (i.e., OP1 in our system) to get exploited by the powerless operator, especially compared with the full spectrum sharing mechanism.

When one operator is allocated with the whole spectrum pool (i.e., OP1 in fair spectrum sharing and both operators in full spectrum sharing), the plot of the gain in average user rate at the other operator is consistent with the plot of the gain in frequency slots, as illustrated in Figure 2.6. This indicates that the decrease in the average user rate per frequency slot due to the inter-operator interference is finite, such that the powerless operator could estimate the gain from the proposed spectrum sharing mechanism through the drop in average user rate per frequency slot and the raise in number of frequency slots; this observation helps it determine whether to participate in spectrum sharing especially when it can merely provide much fewer frequency slots than the powerful operator.

Figure 2.7a and Figure 2.7b shows the average number of scheduled users per cell for OP1 and OP2 respectively. Although there is general improvement for fair spectrum sharing and full spectrum sharing, the growth in average scheduled users is less than the enhancement in average user rate. This implies that the proposed spectrum sharing framework mostly

(a) Average scheduled users of OP1.

(b) Average scheduled users of OP2.

Figure 2.7: Average number of users scheduled per cell (Diagonally overlapped cell deployment).

increases the rate of the users which are already scheduled in the no spectrum sharing case, rather than scheduling many more users within each cell. This is due to our focus on the sum-rate.

The detailed information of user scheduling with the proposed spectrum sharing framework when the number of frequency slots provided by OP2 is minimal (i.e., $N_2 = 4$) and maximal (i.e., $N_2 = 16$) are illustrated in Figure 2.8 and 2.9, respectively, compared to the case of no spectrum sharing which is presented in Figure 2.10. Within these plots, each dot and the filled circle represents a user within each cell that is not scheduled or scheduled, respectively. The colors of the filled circles reflect the number of frequency slots used to transmit signals to the users. The corresponding amount of the spectrum increases as the color changes from blue to red. In general, the system will allocate more frequency slots to users that are closer to the BSs. The users that are closer to the corners of the cells are less likely be scheduled especially when both operators contribute equally to the spectrum pool, when inter-operator interference exists within every frequency slot among the cells.

The rate changes with respect to each user for *fair spectrum sharing*, *full spectrum sharing* and *spectrum sharing with same number of frequency slots* compared with the case of *no spectrum sharing* within the simulation are illustrated in Figure 2.11 and Figure 2.12. In Figure 2.11, OP2 contributes fewer frequency slots than OP1 (i.e., $N_2 = 4$). While in Figure 2.12, OP2 provides the same number of frequency slots as OP1 (i.e., $N_2 = 16$). Each dot within the figures represents a user within the system. The black dots depict the users without rate changes, and the filled circles with warm colors (i.e., from red to yellow) and cold colors (i.e., blue) represent the users with rate improvement and reduction, respectively. The lighter the color and larger the circle, the greater the rate changes. For the case of *spectrum sharing with same number of frequency slots*, the number of users with increased rate and the number of users with decreased rate are comparable, which leads to the minor change of the overall average user rate.

In regards to *fair spectrum sharing*, when $N_2 = 4$, the users of OP1 who are close to the BSs achieve most of the rate improvement. While most scheduled users of OP2 suffer from

Figure 2.8: Distribution of frequency slot allocation among users, fair spectrum sharing ($N_2 = 4$, Diagonally overlapped cell deployment).

Figure 2.9: Distribution of frequency slot allocation among users, fair / full spectrum sharing ($N_2 = 16$, Diagonally overlapped cell deployment).

Figure 2.10: Distribution of frequency slot allocation among users, no spectrum sharing ($N_2 = 4$, Diagonally overlapped cell deployment).

rate loss, as stated above, due to the reduction of inter-operator interference to OP1. It is worth noting that some of the users of OP1 near the corner of its corresponding cell (i.e., closer to the BS of OP2) still get scheduled or achieve rate gain in this case. On the other hand, when both operators provide the same amount of spectrum, the users near the corners of the cell are not scheduled, as shown in Figure 2.12. Towards *full spectrum sharing*, most of the scheduled users of OP2 gain extra rate when $N_2 = 4$. However, for OP1, only the users close to the BSs achieve a slight rate improvement in most cases.

Figure 2.11: Rate changes among users ($N_2 = 4$, Diagonally overlapped cell deployment).

Figure 2.12: Rate changes among users ($N_2 = 16$, Diagonally overlapped cell deployment).

2.5.2 Co-located Cell Deployment

We also explore the potential opportunities created by spectrum sharing with the co-located cell deployment, where the cells of the two operators are overlapped. The BSs of the two operators are co-located at the center of the cells.

Figure 2.13: Average user rate among 9 cells of both operators (Co-located cell deployment).

Figure 2.13 shows the average user rates among all users of both operators across all cells, achieved with the different spectrum sharing mechanisms. We observe that under this circumstance, the proposed fair spectrum sharing mechanism can merely achieve approximately 10% gain over the “No Sharing” case, due to the drastically increased interference at each UT. Since if the two operators are accessing the same frequency slot simultaneously, the interference power at the UTs is likely to be comparable to the signal power received, thus the SINR is decreased. In addition, the improvement of the “Full Spectrum Sharing” case over the “Fair Spectrum Sharing” case is limited compared with in diagonally overlapped cell deployment. The “Full Spectrum Sharing” case performs even worse than “Fair Spectrum Sharing” when OP2 provides much less spectrum than OP1 (e.g., $N_2 = 4$). This implies that the gain from the increased amount of spectrum cannot offset the loss caused by the

(a) Average user rate among 9 cells of OP1.

(b) Average user rate among 9 cells of OP2.

Figure 2.14: Average user rate among 9 cells of each operator (Co-located cell deployment).

increased interference.

The average user rate among users from OP1 and OP2 are presented in Figure 2.14a and Figure 2.14b, respectively. With the proposed fair spectrum sharing framework, the gain of OP1 remains consistent regardless of the number of frequency slots provided by OP2. When OP2 provides fewer frequency slots, although the increase in the amount of total usable spectrum is not significant, OP1 will still have exclusive access to several frequency slots in each cell. On the other hand, as the quantity of spectrum provided by OP2 increases, the exclusively accessed frequency slots diminish, though the aggregate frequency slots in the pool increases. Note these exclusively accessed frequency slots are taken over by shared frequency slots where the UTs may receive severe interference. In this case, OP1 is willing to take part in the proposed spectrum sharing framework no matter how much spectrum is contributed from OP2. Furthermore, to enjoy the benefit from spectrum sharing, OP2 has to provide similar amount of frequency slots to OP1, which prevents OP2 from free-riding. However, the powerless operator, which cannot provide as much spectrum as a powerful operator, will be discouraged from participating in the spectrum sharing structure as it will not attain any gain under this circumstance.

The average number of scheduled users per cell for OP1 and OP2 are illustrated in Figure 2.15a and Figure 2.15b, respectively. Contrary to the diagonally overlapped cell deployment, we observe a significant improvement for spectrum sharing in terms of the number of scheduled users. This is because for each UT, the most significant interference is likely to come from a BS at the center of the UT's cell. For those UTs which are scheduled in the case of no spectrum sharing (e.g., the UTs closer to the BSs), their corresponding channel qualities deteriorate drastically in the case of spectrum sharing. Thus the enhancement in assigning more frequency slots to these users that are already scheduled when there is no spectrum sharing, is below the improvement from scheduling more users towards the sum rate.

The detailed information of user scheduling with the proposed spectrum sharing framework when the number of frequency slots provided by OP2 is minimal (i.e., $N_2 = 4$) and

maximal (i.e., $N_2 = 16$) are illustrated in Figure 2.16 and 2.17, respectively, compared to the case of no spectrum sharing which is presented in Figure 2.18. With spectrum sharing, the users that are closer to the corners and edges of the cells get more chances to be scheduled especially when both operators contribute equally to the spectrum pool, compared with the case of no spectrum sharing. Moreover, fewer frequency slots are allocated to users that are closer to the BSs in the case of spectrum sharing than no sharing, when the operators provide equal amount of the spectrum.

In addition, the rate changes with respect to each user for *fair spectrum sharing*, *full spectrum sharing* and *spectrum sharing with same number of frequency slots* compared with the case of *no spectrum sharing* within the simulation are illustrated in Figure 2.19 and Figure 2.20. In general, the users which are closed to the BSs suffer from the most significant drop in rate. When OP2 provides too few frequency slots to the spectrum pool (i.e., $N_2 = 4$ in our simulations), almost no UT will gain a rate improvement across the cells.

(a) Average scheduled users of OP1.

(b) Average scheduled users of OP2.

Figure 2.15: Average number of users scheduled per cell (Co-located cell deployment).

Figure 2.16: Distribution of frequency slot allocation among users, fair spectrum sharing ($N_2 = 4$, Co-located cell deployment).

Figure 2.17: Distribution of frequency slot allocation among users, fair / full spectrum sharing ($N_2 = 16$, Co-located cell deployment).

Figure 2.18: Distribution of frequency slot allocation among users, no spectrum sharing ($N_2 = 4$, Co-located cell deployment).

Rate improvement on users, spectrum sharing with the same number of frequency slots

Rate improvement on users, fair spectrum sharing

Rate improvement on users, full spectrum sharing

Figure 2.19: Rate changes among users ($N_2 = 4$, Co-located cell deployment).

Rate improvement on users, spectrum sharing with the same number of frequency slots

Rate improvement on users, fair spectrum sharing

Rate improvement on users, full spectrum sharing

Figure 2.20: Rate changes among users ($N_2 = 16$, Co-located cell deployment).

2.6 Summary

In this chapter, we developed a fair spectrum sharing system between two MNOs with inter-operator coordinated beamforming in multi-cell multi-user MIMO networks. Our goal is to understand how an optimization framework and MIMO technologies help in enhancing the benefits of spectrum sharing. In this regard, we formulate an optimization problem to determine the resource allocation and user scheduling within the system. We develop an algorithm to tackle the problem through fractional programming and block coordinate descent. The simulation results illustrate that, with the diagonally overlapped cell deployment, both MNOs could achieve up to 60% improvement in terms of the average user rate with the proposed fair spectrum sharing framework. While with the co-located cell deployment, more users that are closer to the cell edges and corners could have more opportunities to get scheduled with the proposed fair spectrum sharing framework. Interference suppression allows for both operators to use the same frequency slots simultaneously; indeed, the superiority of spectrum sharing mainly comes from the increase in the number of frequency slots used by each operator. Importantly, our results show that both operators have an incentive to take part in the spectrum sharing mechanism proposed, when the gap between the amount of spectrum they provide is not large.

Chapter 3

Channel Estimation in Spectrum Sharing Between Two MNOs

3.1 Introduction

In modern wireless communication networks, channel state information (CSI) is vital in the design of pre-coders or beamforming vectors, as well as for decisions on user scheduling. However, achieving timely and accurate CSI is challenging due to the effects of interference and environmental noise. Among various channel estimation techniques in the literature, time division duplex (TDD) is a promising approach to acquire accurate channel estimation in MIMO networks (especially massive MIMO networks) [26]. In TDD systems, one coherence transmission interval is divided into an uplink (UL) pilot training phase and a data transmission phase (including uplink and downlink data transmissions). During the UL training phase, user terminals (UTs) transmit pre-designed pilot sequences (which are known to both UTs and BSs) to the BSs, and then the BSs can estimate the UL channels based on the received pilot signals. The corresponding downlink (DL) channels can be acquired as well based on channel reciprocity, i.e., the DL and UL channels are the same within the coherence time for the same frequency spectrum [27]. The estimated channel information is thereafter used in the design of precoders or beamformers in the data transmission phase.

In this chapter, we look into channel estimation in the realm of spectrum sharing. Specifically, we investigate that whether the fair spectrum sharing framework proposed in Chapter 2 can still achieve the desired gains, when the knowledge of channel state information remains unknown at the beginning, within a TDD system.

3.2 Proposed Approach

3.2.1 Uplink Training and Channel Estimation

We incorporate the channel estimation mechanism proposed in [28] into our spectrum sharing system. During the UL training period, each UT within the system transmits a pilot sequence of length τ_p across all frequency slots, denoted by $\boldsymbol{\psi} \in \mathbb{C}^{1 \times \tau_p}$. Note that τ_p should be greater than or equal to the number of UTs within each cell. We consider the case that the pilot reuse factor is 1, i.e., one pilot set $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1^T \ \boldsymbol{\psi}_2^T \ \dots \ \boldsymbol{\psi}_{\tau_p}^T]^T \in \mathbb{C}^{\tau_p \times \tau_p}$ is shared among all cells from both operators. The τ_p pilot sequences are orthogonal between each other and $\boldsymbol{\psi}_i^\dagger \boldsymbol{\psi}_i = 1$. Each UT within each cell is assigned one unique pilot sequence from the τ_p pilots in $\boldsymbol{\Psi}$ to prevent intra-cell interference. Then, for the BS located in the i^{th} cell of OP1, the received signal during the training phase for OP1 through the j^{th} frequency slot, $\mathbf{Y}_{1,i,j}$ ($M \times \tau_p$ matrix), can be expressed as

$$\mathbf{Y}_{1,i,j} = \sqrt{p_t} \sum_{i'=1}^B \sum_{k_1=1}^{K_1} \sqrt{\beta_{1i,1i'}^{k_1}} \mathbf{g}_{1i,1i',j}^{k_1} \boldsymbol{\psi}_{1,i'}^{k_1} + \sqrt{p_t} \sum_{i''=1}^B \sum_{k_2=1}^{K_2} \sqrt{\beta_{1i,2i''}^{k_2}} \mathbf{g}_{1i,2i'',j}^{k_2} \boldsymbol{\psi}_{2,i''}^{k_2} + \mathbf{W}_{1,i,j} \quad (3.1)$$

The first term in (3.1) denotes the signals received from all users of OP1, and the second term denotes the signals received from all users of OP2. $\mathbf{W}_{1,i,j}$ ($M \times \tau_p$ matrix) denotes the additive white Gaussian noise (AWGN). In more detail, $\boldsymbol{\psi}_{1,i'}^{k_1}$ refers to the pilot sequence transmitted by the k_1^{th} user from the $(i')^{th}$ cell of OP1. $\beta_{1i,1i'}^{k_1} = (\frac{d_0}{d_{1i,1i'}^{k_1}})^\alpha$ is the large scale fading between the BS (i^{th} cell) and the UT ($(i')^{th}$ cell, k_1^{th} user), where $d_{1i,1i'}^{k_1}$ denotes the corresponding distance. $\mathbf{g}_{1i,1i',j}^{k_1} \sim \mathcal{CN}(0, I_M)$, which is an independent and identically distributed (i.i.d.) zero-mean, circularly-symmetric complex Gaussian random variable, denotes

the corresponding small scale fading. It is worth mentioning that we assume the large scale fading terms are known at each BS, whereas the small scale fading remains unknown to be estimated. $\beta_{1_i,2_{i''}}^{k_2}$, $\mathbf{g}_{1_i,2_{i''},j}^{k_2}$ and $\psi_{2,i''}^{k_2}$ denote the large scale fading, small scale fading, and the transmitted pilot sequence, respectively, of the k_2^{th} UT in the $(i'')^{th}$ cell of OP2. The received signals at the BSs of OP2 is similar to (3.1).

Due to the orthogonality of different pilot sequences, the channels between each BS and the UTs who share the same pilot sequence can be filtered out through multiplying the received signals at the BSs by the corresponding pilot sequence, e.g.,

$$\begin{aligned} \mathbf{y}_{1,i,j}^k &= \mathbf{Y}_{1,i,j} \cdot (\psi^k)^\dagger \\ &= \sqrt{p_t} \sum_{i'} \sqrt{\beta_{1_i,1_{i'}}^k} \mathbf{g}_{1_i,1_{i'},j}^k + \sqrt{p_t} \sum_{i''} \sqrt{\beta_{1_i,2_{i''}}^k} \mathbf{g}_{1_i,2_{i''},j}^k + \mathbf{W}_{1,i,j} (\psi^k)^\dagger \end{aligned} \quad (3.2)$$

$\mathbf{y}_{1,i,j}^k$ ($M \times 1$ vector) contains the information of channels between the BS of the i^{th} cell of OP1 and the UTs from both operators who transmit the k^{th} pilot in the pilot set Ψ . Given $\mathbf{y}_{1,i,j}^k$, the MMSE estimation of the small scale fading term $\mathbf{g}_{1_i,1_{i'},j}^k$ is

$$\hat{\mathbf{g}}_{1_i,1_{i'},j}^k = \sqrt{p_t \beta_{1_i,1_{i'}}^k} \left(p_t \sum_{i'} \beta_{1_i,1_{i'}}^k I_M + p_t \sum_{i''} \beta_{1_i,2_{i''}}^k I_M + \sigma^2 I_M \right)^{-1} \mathbf{y}_{1,i,j}^k \quad (3.3)$$

Afterwards, $\hat{\mathbf{g}}_{1_i,1_{i'},j}^k$ is used in the beamformer design and user scheduling of the system. The channel estimation error due to other UTs sharing the same pilot sequence is known as pilot contamination.

3.2.2 User re-clustering

In the above pilot assignment scheme, the pilots are assigned per cell. One limitation raised with such a mechanism is that when the number of UTs affiliated with the two operators are different, the number of pilot sequences (or pilot length τ_p) should be at least the larger number of UTs per cell of the two operators, which is disadvantageous for the operator with fewer UTs. This is because its training period is extended compared with when there is no spectrum sharing. Pilot contamination could be another issue which restricts the perfor-

mance of the channel estimation framework. Since without pre-designed pilot assignment schemes (i.e., when the pilot sequences are randomly assigned to UTs), the closely located UTs from different operators assigned with the same pilot sequence may strongly interfere with each other.

To reduce the interference between adjacent UTs during uplink training, the UTs from **both** operators need to be organized into clusters. The pilot sequences are then to be assigned per cluster instead of per cell. There is no hard constraint on the cluster size, though the selection of cluster size will affect the accuracy of the estimation and the performance of the whole system. In this way, the size of the pilot set τ_p can be more flexible and is not limited by the number of users in each cell. The effect of cluster size will be investigated in the next section.

Within this work, firstly the UTs are clustered centering on the BSs from different locations, as in a conventional cell deployment in wireless communications, where the BSs are located at the center of each cell¹. This ensures that for each UT within these groups, the uplink training interference can only come from other UTs that are farther away from its corresponding BS, which reduces the probability that the UT suffers higher interference power than the signal itself. The detailed algorithm in assigning UTs to different clusters is presented in Algorithm 2.

In Algorithm 2, C_c denotes the number of BS-centric clusters, which equals to the number of different locations of BSs. T is the cluster size (The length of pilot sequences $\tau_p \geq T$). In general, the UTs are assigned to different clusters based on the ranks of closeness to the clusters centroids, i.e., the BSs. The ranks of UTs are achieved in terms of the distances towards the centroid of the cluster, for all clusters, at the beginning of the algorithm. Moreover, a pointer p_i is initialized to point to the UT which is closest to the BS of the i^{th} cluster. The clustering process is described by the loop between Step 3 and Step 12, where one UT (which the pointer p_i points to) is classified into each cluster during each iteration until reaching the ceiling of the cluster size. If the pointer p_i points to a UT that has already been assigned

¹In the algorithm below, we rank UTs by their large-scale fading experienced. This would also address the situation where the BS is not at the center of a cell.

Algorithm 2: BS-centric user clustering

-
- 1: For each cluster, rank all UTs based on the large-scale fading parameters between the UTs and the BS, in ascending order.
 - 2: Initialize pointers $p_i = 1, i \in [1, C_c]$, which points to the closest UT of each BS in each group.
 - 3: **for** τ from 1 to T **do**
 - 4: Sort the clusters indexes based on the value of pointer p_i , in descending order.
 - 5: **for** each cluster (with index i) in the order of sorted clusters **do**
 - 6: **while** the UT which the pointer p_i points to (i.e., the p_i^{th} closed UT) is already assigned to any other cluster **do**
 - 7: $p_i \leftarrow p_i + 1$
 - 8: **end while**
 - 9: Assign the UT which the pointer p_i points to to the i^{th} cluster.
 - 10: $p_i \leftarrow p_i + 1$
 - 11: **end for**
 - 12: **end for**
-

to another cluster, the pointer will keep moving to the next candidate (i.e., the next closest UT) until it finds an available UT, as the loop between Step 6 and Step 8. To prevent one pointer from moving much faster than the others, which leads to imbalanced ranks of UTs to be assigned among all clusters, the cluster with a larger value of p_i is prioritized to get the UT assigned. For instance, consider the case when $p_i = p$ and $p_j = p + 1$, and they point to the same UT (i.e., the UT which is both the p^{th} closest to the i^{th} BS and the $(p + 1)^{th}$ closest to the j^{th} BS). If the UT is assigned to the i^{th} BS first, then p_j has to increase, which further enlarges the gap between p_i and p_j . On the other hand, assigning the UT to the j^{th} BS first will likely to get p_i and p_j back to the same starting line. The prioritization is achieved through sorting the cluster indexes in Step 4 of the algorithm.

Figure 3.1 offers an example of the acquired clusters from the BS-centric clustering algorithm with the diagonally overlapped cell deployment. Here, we assume $\mathcal{B} = 9$. Thus there are altogether 18 clusters. A BS of OP1 is located at the *center* of each of its cells, and a BS of OP2 is located at the *corner* of each cell of OP1. The different clusters of UTs are distinguished by different colors. Among the data points, the circular ones represent UTs of OP1, and the triangular ones represent UTs of OP2. In general, the UTs are assigned to the cluster of the closest BS in the network area.

Figure 3.1: User clustering with only BS-centric user clustering (Diagonally overlapped cell deployment).

If the cluster size is small, or with co-located cell deployment where BSs of the two operators overlap, only part of the UTs can be grouped into the C_c BS-centric clusters. The remaining UTs are thus clustered through a modified K-means algorithm proposed in [29]. As opposed to the standard K-means algorithm, the modified algorithm can be applied to clustering with cluster size constraints, as presented in Algorithm 3.

Algorithm 3: Modified K-means Algorithm

- 1: Initialize cluster centroids $\mu_1, \mu_2, \dots, \mu_K$ randomly.
 - 2: **repeat**
 - 3: $c_j = \{u_i : \|u_i - \mu_j\|^2 \leq \|u_i - \mu_{j'}\|^2, \forall j' \neq j, |c_j| < T\}$;
 - 4: $\mu_j = \frac{1}{|c_j|} \sum_{u_i \in c_j} u_i$;
 - 5: $1 \leq j \leq K$
 - 6: **until** Convergence
-

Assume K clusters are needed here to group the remaining UTs (i.e., the unclustered UTs from Algorithm 2). Firstly, in Step 1 of Algorithm 3, the centroids of the K clusters are initialized with the randomly selected K distinct UT locations. The UT assignment is done in Step 3, when each UT is assigned to the cluster with the closest centroid. It

Figure 3.2: User clustering with BS-centric user clustering (as in the red circle) and K-means clustering (as in the blue circle) (Diagonally overlapped cell deployment).

Figure 3.3: User clustering with BS-centric user clustering and K-means clustering (Co-located cell deployment).

is worth mentioning that if the size of the targeting cluster reaches its maximum (i.e., T), the algorithm will keep moving to the next closest candidate until an available cluster (i.e., $|c_j| < T$) is found. Afterwards, in Step 4, the centroid of each cluster is updated as the mean of the UTs locations within the cluster. The algorithm keeps alternating between the assignment step and the update step until convergence (i.e., the changes of centroids are below a predefined threshold).

Figure 3.2 provides an example on the user clustering with a small cluster size that not all UTs can be clustered with merely BS-centric clustering, in the case of diagonally overlapped cell deployment. The legends used here are the same as in Figure 3.1. The UTs clustered through BS-centric clustering are shown by data points with larger sizes, as depicted in the red circle. The UTs clustered through the modified K-means algorithm are shown by the smaller data points, as illustrated in the blue circle. Similarly, an example of user clustering in the case of co-located cell deployment is presented in Figure 3.3.

3.3 Simulation Results and Performance Evaluation

In this section, we evaluate the performances of the channel estimation schemes presented in the previous section. The comparisons are made among the following scenarios:

- *No spectrum sharing*: Each operator owns an independent pilot set. The length of the pilot sequences from different operators could be different, but it must be at least the number of UTs per cell. The pilots are assigned to UTs randomly in each cell. The resource allocation and user scheduling are then determined by each operator itself with the individually estimated channels.
- *Spectrum sharing*: One pilot set is used for the whole system (i.e., both OP1 and OP2). The length of the pilot sequences must be at least the larger number of UTs per cell between the two operators. The pilots are assigned to UTs randomly in each cell of both operators. The resource allocation and user scheduling of the two operators are achieved jointly with the fair spectrum sharing framework presented in Chapter 2 based

on the estimated channels.

- *Spectrum sharing with pilot assignment:* Everything is the same as the *Spectrum sharing* case, except the pilot sequences are assigned to UTs with pre-defined regulations or protocols. Within this work, we employ an intuitive pilot assignment mechanism that the pilots are assigned according to the ranking of distances between the BS and the UTs.
- *Spectrum sharing with clustering:* The UTs of both operators are re-clustered using the methods proposed in Section 3.2.2. One pilot set is used for the whole system, and the length of the pilot sequences must be larger than or equal to the cluster size. Note there is no restrictions on the cluster size. The pilots are assigned to UTs randomly in each cluster.

The simulation environment is set to be the same as in Chapter 2, i.e., the numerical values of the parameters used are illustrated in Table 2.1; the number of users from OP1 and OP2 within each cell is set to be M times the number of frequency slots provided by OP1 and OP2 respectively (i.e., $K_1 = M \times N_1$ and $K_2 = M \times N_2$); and we assume each BS is deployed with $M = 8$ antennas.

3.3.1 Diagonally Overlapped Cell Deployment

In this part, we evaluate the performance of channel estimation in spectrum sharing, in the case of diagonally overlapped cell deployment, where the BSs of one operator stand at the corners of the cells of the other operator, as described in Section 2.2.

First, we evaluate the effectiveness of the channel estimation methods in terms of the normalized mean squared error (NMSE) per cell, $\text{NMSE} = E_k[\frac{\|\hat{\mathbf{h}}^k - \mathbf{h}^k\|_2^2}{\|\mathbf{h}^k\|_2^2}]$, as defined in [30]. Here, $\mathbf{h}^k \in \mathbb{C}^{M \times 1}$ denotes the channel between a BS and the k^{th} UT in a cell (the BS and the UT may not belong to the same cell). In this study, we assume the size of the pilot set (length of each pilot sequence) used in spectrum sharing is fixed at $\tau_p = K_1$, since $K_1 \geq K_2$. The cluster size T is set to K_1 as well, i.e., $T = \tau_p$. The NMSE in estimating the intra-

operator channels and the inter-operator channels are shown in Figure 3.4 and Figure 3.5 respectively.

Figure 3.4: Intra-operator channel estimation NMSE (Diagonally overlapped cell deployment).

In Figure 3.4, the lines near 0 dB (as highlighted by the red rectangle) denote the estimation errors of the inter-cell channels, i.e., the channels between a BS and UTs from different cells. Such channels can hardly be estimated due to severe interference, since the same pilot sequence is shared with at least one other UT that is closer to the BS. The NMSEs in estimating the intra-cell channels for OP1 and OP2 are depicted with the blue curves and orange curves respectively. With spectrum sharing, the estimation errors increase compared with the “No Sharing” case due to the inter-operator interference. Amongst the different channel estimation methods in spectrum sharing, “Spectrum Sharing with Pilot Assignment” achieves the best performance. This is because this method ensures minimal inter-operator interference towards UTs which are adjacent to the BSs during the uplink training period (the sources of interference are around cell corners). The “Spectrum Sharing with Clustering” case achieves the worst performance, since within each cell, although the UTs inside the

cluster are prevented from devastating interference, achieving accurate channel information for UTs outside the cluster is nearly impossible. In addition, the NMSE for OP1 is smaller than the NMSE for OP2, when OP2 provides less spectrum in spectrum sharing. This is because, the size of the pilot set $\tau_p = K_1 > K_2$, such that UTs of OP2 will only use part of the pilot sequences from the pilot set, which reduces the interference towards OP1.

Figure 3.5: Inter-operator channel estimation NMSE (Diagonally overlapped cell deployment).

Regarding to the performance on inter-operator channel estimation, as illustrated in Figure 3.5, only channels between UTs and their closest BS (from the other operator) can be estimated. The NMSEs in estimating the channels between UTs and farther BSs are nearly 0 dB, as highlighted by the red rectangle. Contrary to intra-operator channel estimation, the “Spectrum Sharing with Clustering” case achieves the best performance, as there are UTs from the other operator in the cluster associated with each BS. These channels can be estimated with guaranteed accuracy due to the limited interference from merely UTs in the other clusters. In the “Spectrum Sharing with Pilot Assignment” case, even the channel between a BS and a closely located UT from different operators can suffer from severe

interference, which leads to its poor channel estimation performance. This is because each BS locates at the intersection of four cells from the other operator, such that possibly more than one UTs from different cells which are adjacent to this BS could share the same pilot sequence. In addition, with the pilot assignment scheme used, the channel estimation gets worse with the decrease of number of UTs per cell from OP2. This is because the pilots are assigned to UTs in sequence based on the ranking of distances to the BS, such that when OP2 has much fewer UTs per cell than OP1, all UTs of OP2 in each cell are likely to undergo significant interference from UTs of OP1.

Figure 3.6: Average user rate among 9 cells of both operators with channel estimation (Diagonally overlapped cell deployment).

Next, we evaluate the system performance in terms of the average user rate with the estimated channel state information. The average user rates among 9 cells of both operators for the different channel estimation schemes proposed are illustrated in Figure 3.6. The average user rates for the individual operator are shown in Figure 3.7. The solid lines represent the average rates achieved with perfect knowledge of the channels, while the dashed

(a) Average user rate among 9 cells of OP1 with channel estimation.

(b) Average user rate among 9 cells of OP2 with channel estimation.

Figure 3.7: Average user rate among 9 cells of each operator with channel estimation (Diagonally overlapped cell deployment).

lines denote the average rates achieved with the estimated channels. We observe that with the estimated channel information, the achieved average rate drops by approximately 20% compared with when accurate CSI is known, under the “No Sharing” case. The decrease in average rate under the “Spectrum Sharing” case ranges from 25% to 30%, which is slightly larger than the “No Sharing” case due to the greater channel estimation error. Thus, with imperfect channel knowledge, the increase in performance with the fair spectrum sharing framework proposed over no spectrum sharing is up to 50%. This is smaller than the 60% improvement under perfect channel knowledge. Note that for the discussion above, the pilots are assigned randomly to UTs in each cell for both the “No Sharing” and the “Spectrum Sharing” case during the channel estimation process. However, once the pilots are assigned to UTs following some pre-developed pilot assignment protocols, the improvement brought by spectrum sharing can be further enhanced. As in the simulation, with the intuitive pilot assignment scheme, the “Spectrum Sharing with Pilot Assignment” case is able to gain a 60% improvement over the “No Sharing” case. This increase is comparable to the improvement under perfect knowledge of CSI.

Interestingly, the improvement in average user rate is positively correlated with the channel estimation accuracy of the intra-operator, intra-cell channels, i.e, the channels of transmitting signals. The “Spectrum Sharing with Pilot Assignment” case, which achieves the most accurate estimation on these channels, gives the most significant improvement in average user rate. While the “Spectrum Sharing with Clustering” case provides the least improvement. This implies that within this spectrum sharing framework, the intra-operator, intra-cell interference are more crucial than the inter-operator interference in determining the overall system performance.

In these discussions, the size of the pilot set (or length of each pilot sequence) is fixed at $\tau_p = K_1$, which is the number of UTs per cell for OP1, among the different channel estimation methods. With longer pilot sequences (and hence larger size of the orthogonal pilot set), the performance of channel estimation can be further improved due to the reduced pilot contamination. However, as previously mentioned, both channel estimation and data

transmission should be executed within a single coherent transmission period. The data transmission phase will be compressed if the time spent in channel estimation gets increased. To investigate the impact of the length of the pilot sequence, we evaluate the achievable sum rate (ASR) as defined in [31],

$$ASR = \left(\frac{T_{coherent} - \tau_p}{T_{coherent}} \right) \sum_{i,j,m} \log(1 + \Gamma_{i,j}^m) \quad (3.4)$$

achieved among the different channel estimation scenarios. In (3.4), $T_{coherent}$ denotes the length of the whole coherent transmission period, τ_p stands for the duration of the channel estimation phase (length of the pilot sequences), and thus $T_{coherent} - \tau_p$ is the length of the data transmission period. The pre-log term, $\frac{T_{coherent} - \tau_p}{T_{coherent}}$, refers to the proportion that data transmission accounts for in each coherent transmission period. The second half in (3.4), $\sum_{i,j,m} \log(1 + \Gamma_{i,j}^m)$, represents the sum rate of either OP1, OP2 or the whole system, depending on the specifications of parameters.

The ASRs achieved, for individual operators, among “No Sharing”, “Spectrum Sharing”, “Spectrum Sharing with Clustering” and “Spectrum Sharing with Pilot Assignment” with respect to the length of the pilot sequences are illustrated in Figure 3.8. Here, we assume the length of the coherent transmission time is $T_{coherent} = 256$. Regarding to the length of pilot sequences, for both the “Spectrum Sharing” case and the “Spectrum Sharing with Pilot Assignment” case, it should be at least K_1 , i.e., $\tau_p \geq K_1 = 128$, as OP1 has more UTs per cell. For the “No Sharing” case, the length of pilot sequences should be at least the number of UTs per cell for each operator individually (i.e., $\tau_p \geq K_1$ for OP1 and $\tau_p \geq K_2$ for OP2). There are no restrictions on length of pilot sequences for the “Spectrum Sharing with Clustering” case, as it is the cluster size that is upper-bounded by the size of the pilot set. In simulations, we set $\tau_p \geq \min(K_2)$.

In Figure 3.8, the plots in blue, orange, yellow and purple denote the cases when OP2 provides $N_2 = 4, 8, 12, 16$ frequency slots to the spectrum pool, respectively. Among the channel estimation methods, “Spectrum Sharing with Clustering” always achieves the highest

(a) Achievable sum rate of OP1 with channel estimation.

(b) Achievable sum rate of OP2 with channel estimation.

Figure 3.8: Achievable sum rate of each operator with channel estimation (Diagonally overlapped cell deployment).

ASR with the length of pilot sequences $\tau_p < K_1$. The curves of ‘Spectrum Sharing with Clustering’ are parabolic with the peak appearing in the middle, which illustrates the trade-off between the pre-log term and the sum rate. When τ_p is small, the pre-log term gets large. Yet the channel estimation error increases, which decreases the sum rate achieved. On the other hand, if τ_p is large, although a favorable sum rate can be attained thanks to the accurate channel estimation, the pre-log term is small. In addition, when OP2 provides as many frequency slots as OP1, i.e., $N_2 = N_1$, “Spectrum Sharing with Pilot Assignment” can approximate to the highest ASR at $\tau_p = K_1$, as a result of its superior sum rate achieved. However, when there is a gap between the number of frequency slots provided by the two operators (e.g., $N_2 \leq 0.5N_1$), no channel estimation method in spectrum sharing can achieve a higher ASR at the weaker operator (OP2 here) over no spectrum sharing. This will discourage the operator from participating the spectrum sharing mechanism proposed if it owns far fewer frequency slots than the other operator.

3.3.2 Co-located Cell Deployment

In this sections, we evaluate the performance of channel estimation in spectrum sharing, in the case of co-located cell deployment, where the cells of the two operators are overlapped.

As in the previous section, we start by evaluating the NMSE for the different channel estimation methods. The NMSEs in estimating the intra-operator channels and the inter-operator channels are shown in Figure 3.9 and Figure 3.10, respectively. It is worth mentioning that with this cell deployment, the estimated *intra-operator* channels and the estimated *inter-operator* channels at each UT can not be distinguished, as its corresponding large-scale fading factors at the BSs from the two operators are identical on average. This leads to the essentially identical plots of the channel estimation errors between Figure 3.9 and Figure 3.10 (i.e., the differences in between are negligible). We observe that the overall minimum estimation error between the two operators is achieved in the “Spectrum Sharing with Clustering” case, amongst the different channel estimation methods in spectrum sharing. The “Spectrum Sharing with Pilot Assignment” case has the worst performance in

Figure 3.9: Intra-operator channel estimation NMSE (Co-located cell deployment).

Figure 3.10: Inter-operator channel estimation NMSE (Co-located cell deployment).

terms of NMSE. Since with this pilot assignment scheme, the two UTs from different operators within the same cell, which share the same pilot sequence, are likely to be adjacent between each other. Moreover, in this case, the difference of estimation errors between the two operators is significant, when the gap between the number of frequency slots provided by the two operators is large. As we assume the operator that provides less spectrum has fewer UTs, such that the corresponding UT density per cell is small. Thus, for each UT from the powerless operator (OP2 in our case), the UT from the powerful operator which shares the same pilot is likely to be closer to the BS, which causes severe interference and thereafter affect the estimation accuracy during the channel estimation process.

Figure 3.11: Average user rate among 9 cells of both operators with channel estimation (Co-located cell deployment).

With the estimated channel information, the achieved average user rates among 9 cells of both operators for the different channel estimation schemes proposed are illustrated in Figure 3.11. The achieved average user rates at each operator are illustrated in Figure 3.12. We observe that no spectrum sharing mechanism can outperform the “No Sharing” case with

channel estimation under co-located cell deployment in terms of average user rate. This is because, in this case, under spectrum sharing, accurate channel information is crucial in the coordination between the beamforming of the two operators to reduce the inter-operator interference introduced. However, in terms of ASR, the “Spectrum Sharing with Clustering” case can achieve an improved performance at the peak, as the gain in the pre-log term outweighs the loss in sum rate (loss in estimation accuracy). Unfortunately, as shown in Figure 3.13, the results demonstrate that the powerless operator does not have the incentive to participate in spectrum sharing, as the ASR of OP2 surpasses the “No Sharing” case only when it can provide similar amount of spectrum as OP1.

(a) Average user rate among 9 cells of OP1 with channel estimation.

(b) Average user rate among 9 cells of OP2 with channel estimation.

Figure 3.12: Average user rate among 9 cells of each operator with channel estimation (Co-located cell deployment).

(a) Achievable sum rate of OP1 with channel estimation.

(b) Achievable sum rate of OP2 with channel estimation.

Figure 3.13: Achievable sum rate of each operator with channel estimation (Co-located cell deployment).

3.4 Summary

In this chapter, we study the impact of channel estimation on the fair licensed spectrum sharing framework proposed in Chapter 2, in a TDD system. We incorporated a channel estimation technique into our framework, where UTs send pre-designed pilot sequences to the BSs. As the cell based pilot allocation scheme is inferior for the powerless operator, we propose a cluster based pilot allocation mechanism where the pilot sequences are assigned to UTs within each newly organized cluster. The simulation results illustrate that, although the cluster-based pilot allocation scheme achieves an improvement over the conventional cell-based allocation in terms of achievable sum rate, the operator that can merely provide much less spectrum than the other operator will be unlikely to take part in the spectrum sharing framework. On the other hand, however, once both operators can provide similar number of frequency slots, channel estimation does not affect the gain attained through the spectrum sharing mechanism presented in Chapter 2.

Chapter 4

SLINR-based Fair licensed Spectrum Sharing Between Two MNOs

4.1 Introduction

In the previous chapters, we investigate the SINR-based spectrum sharing between two MNOs, which requires a centralized controller to manage the resource allocation and user scheduling within the system. However, the SINR-based optimization framework requires significant overhead of information exchange between the controller and the MNOs (e.g., channel state information, beamforming information, user scheduling decisions, etc.), which is hard to realize in practice. In the literature, a signal to leakage and noise ratio (SLNR) based approach, which uses leakage as a proxy for interference, becomes influential in determining the resource allocation in a decentralized manner with merely local information [32][33]. In addition to the SLNR-based approach, the authors in [34] propose a signal to leakage plus interference and noise ratio (SLINR) based approach to tackle the problem of resource allocation and user scheduling in a multi-cell multi-user MIMO network, where they replace the inter-cell interference with leakage, but keep the intra-cell interference to minimize the effect of co-channel interference. The SLINR-based framework presented in [34] requires no information exchange between BSs.

In this chapter, we briefly look into a SLINR-based spectrum sharing mechanism where each MNO can determine the resource allocation by itself, which also requires minimum information exchange between the MNOS.

4.2 Problem Formulation

In this chapter, we apply the identical system model as in Chapter 2. In Chapter 2, the objective is to maximize the sum rate of the whole system, which requires the information of SINR (2.3) at each UT. The only term in (2.3) that contains information of the other operator is the inter-operator interference $I_{2,i',1,i,j}$, which includes the channel information $\mathbf{H}$, the beamformers $\mathbf{v}$, the user scheduling $\mathbf{u}$ and the frequency selection $\mathbf{s}$. This gives us the motivation to replace this inter-operator interference term with inter-operator leakage in the problem formulation. The inter-operator leakage from the m^{th} beam in the i^{th} cell of OP1 to all UTs of OP2 through the j^{th} frequency slot, $L_{1,i,2,j}^m$, is defined as¹:

$$L_{1,i,2,j}^m = \sum_{i''=1}^B s_{1,i,j} \|\mathbf{H}_{1,i,2,i'',j}^\dagger \mathbf{v}_{1,i,j}^m\|_2^2 \quad (4.1)$$

Here, no information from OP2 is needed for OP1 assuming perfect knowledge of the channel information between the i^{th} BS of OP1 and users of OP2, i.e., $\mathbf{H}_{1,i,2,i'',j} \in \mathbb{C}^{M \times K_2}$ (Note that the channel information is acquired at the BSs). Thus, we define the signal to leakage plus interference and noise ratio (SLINR) at the m^{th} scheduled user through the j^{th} frequency slot in the i^{th} cell of OP1, $\Gamma_{1,i,j}^{m(LI)}$, as

$$\Gamma_{1,i,j}^{m(LI)} = \frac{s_{1,i,j} |(\mathbf{H}_{1,i,1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\underbrace{\sigma^2 + \sum_{\substack{i'=1 \\ (i',m') \neq (i,m)}}^B \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2}_{\text{Intra-operator interference}} + \epsilon L_{1,i,2,j}^m} \quad (4.2)$$

¹Here, we assume perfect knowledge of the channels at each operator. The CSI is the only information that shared between the operators

The numerator in (4.2) represents the received signal power at the user, and the denominator consists of the power of noise, intra-operator interference and inter-operator leakage. Note that $L_{1,i,2,j}^m$ is a general leakage to all UTs of OP2, which is larger in power than merely the leakage to the scheduled UTs. Thus, we use a factor ϵ ($0 \leq \epsilon \leq 1$) to control the weight of the leakage term accounts for in the denominator. The effect of ϵ will be presented in Section 4.4.

Without information from the other operator, we cannot maximize the system sum-rate as in Chapter 2. Instead, we define a pseudo-rate as a proxy at each scheduled user as $R_{1,i,j}^{m(LI)} = W \cdot \log(1 + \Gamma_{1,i,j}^{m(LI)})$. Accordingly, the objective function to be maximized for OP1 (similar for OP2) is the pseudo-sum-rate:

$$f_o^{LI}(\mathbf{s}_1, \mathbf{u}_1, \mathbf{v}_1) = \sum_{i=1}^B \sum_{j=1}^{N_1+N_2} \sum_{m=1}^M \log(1 + \Gamma_{1,i,j}^{m(LI)}) \quad (4.3)$$

It is worth noting that the performance is still evaluated through the SINR-based sum-rate. Thus, the optimal resource allocation and user scheduling at OP1 (similar for OP2) can be achieved through solving the following constrained optimization problem:

$$\max_{\mathbf{s}_1, \mathbf{u}_1, \mathbf{v}_1} f_o(\mathbf{s}_1, \mathbf{u}_1, \mathbf{v}_1) \quad (4.4a)$$

$$\text{subject to} \quad \sum_{m=1}^M \|\mathbf{v}_{1,i,j}^m\|^2 \leq P_1, \quad \forall i, j; \quad (4.4b)$$

$$\sum_{k=1}^{K_1} (u_{1,i,j}^m)_k = 1, \quad \forall i, j, m; \quad (4.4c)$$

$$\sum_{m=1}^M (u_{1,i,j}^m)_k \leq 1, \quad \forall i, j, k; \quad (4.4d)$$

$$\sum_{j=1}^{N_1+N_2} s_{1,i,j} \leq N_c, \quad \forall i; \quad (4.4e)$$

$$(u_{1,i,j}^m)_k, s_{1,i,j} \in \{0, 1\} \quad (4.4f)$$

The constraints in (4.4) are the same as those in (2.5). N_c is the maximal number of frequency slots allowed to access as defined in (2.30).

4.3 Proposed Solution Approach

To solve the optimization problem (4.4), we follow the same procedure with identical techniques as in Section 2.4. At first, the objective function (4.4a) is turned into the form of sum-of-ratios as:

$$f_r^{(LI)}(\mathbf{s}_1, \mathbf{u}_1, \mathbf{v}_1) = \sum_{i=1}^{\mathcal{B}} \sum_{j=1}^{N_1+N_2} \sum_{m=1}^M \left(\log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m \right. \\ \left. + \frac{(1 + \gamma_{1,i,j}^m) s_{1,i,j} |(\mathbf{H}_{1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\sigma^2 + \sum_{\substack{i'=1 \\ (i',m') \neq (i,m)}}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',j} \mathbf{u}_{1,i',j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \epsilon L_{1,i,j}^m} \right) \quad (4.5)$$

where,

$$\gamma_{1,i,j}^m = \frac{s_{1,i,j} |(\mathbf{H}_{1,i,j} \mathbf{u}_{1,i,j}^m)^\dagger \mathbf{v}_{1,i,j}^m|^2}{\sigma^2 + \sum_{\substack{i'=1 \\ (i',m') \neq (i,m)}}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',j} \mathbf{u}_{1,i',j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \epsilon \sum_{i''=1}^{\mathcal{B}} s_{1,i,j} \|\mathbf{H}_{1,i,2,i'',j}^\dagger \mathbf{v}_{1,i,j}^m\|_2^2} \quad (4.6)$$

Next, through quadratic transform to the sum-of-ratios terms, (4.5) can be further convexified as:

$$f_q^{(LI)}(\mathbf{s}_1, \mathbf{u}_1, \mathbf{v}_1, \boldsymbol{\gamma}_1, \mathbf{y}_1) = \sum_{i=1}^{\mathcal{B}} \sum_{j=1}^{N_1+N_2} \sum_{m=1}^M \left\{ \left(\log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m \right. \right. \\ \left. + 2Re\{(y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,j} \mathbf{u}_{1,i,j}^m)\} - |y_{1,i,j}^m|^2 \sigma^2 \right. \\ \left. \left. - |y_{1,i,j}^m|^2 \left(\sum_{i'=1}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',j} \mathbf{u}_{1,i',j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \sum_{i''=1}^{\mathcal{B}} s_{1,i,j} \|\mathbf{H}_{1,i,2,i'',j}^\dagger \mathbf{v}_{1,i,j}^m\|_2^2 \right) \right) \right\} \quad (4.7)$$

Note that $f_q^{(LI)}$ is concave on $\boldsymbol{\gamma}_1$ and $\mathbf{y}_1$. Thus, the optimum $(y_{1,i,j}^m)^*$ can be acquired by solving $\frac{\partial f_q^{(LI)}}{\partial y_{1,i,j}^m} = 0$, with the other parameters unchanged. The corresponding expression is as follows:

$$y_{1,i,j}^m = \frac{s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,j} \mathbf{u}_{1,i,j}^m) \sqrt{(1 + \gamma_{1,i,j}^m)}}{\sigma^2 + \sum_{i'=1}^{\mathcal{B}} \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',j} \mathbf{u}_{1,i',j}^m)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 + \sum_{i''=1}^{\mathcal{B}} s_{1,i,j} \|\mathbf{H}_{1,i,2,i'',j}^\dagger \mathbf{v}_{1,i,j}^m\|_2^2} \quad (4.8)$$

Similarly as in Chapter 2, the optimal beamforming vector $\mathbf{v}_{1,i,j}^m$ can be attained by solving the following constrained convex optimization problem:

$$\max_{\mathbf{v}} \quad f_q^{(LI)}(\mathbf{v}_1) \quad (4.9a)$$

$$\text{subject to} \quad (4.4b) \quad (4.9b)$$

which could be solved by means of Lagrange multipliers. The corresponding closed form solution to $\mathbf{v}_{1,i,j}^m$ is:

$$\mathbf{v}_{1,i,j}^m = s_{1,i,j} \mathbf{D}_{1,i,j}^{-1} \sqrt{(1 + \gamma_{1,i,j}^m)} (y_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,1,i,j} \mathbf{u}_{1,i,j}^m) \quad (4.10)$$

where

$$\mathbf{D}_{1,i,j} = \lambda_{1,i,j} \mathbf{I} + s_{1,i,j} \left(\sum_{i'=1}^B \sum_{m'=1}^M |y_{1,i',j}^{m'}|^2 (\mathbf{H}_{1,i,1,i',j} \mathbf{u}_{1,i',j}^{m'}) (\mathbf{H}_{1,i,1,i',j} \mathbf{u}_{1,i',j}^{m'})^\dagger + \sum_{i''=1}^B |y_{1,i,j}^m|^2 \mathbf{H}_{1,i,2,i'',j} \mathbf{H}_{1,i,2,i'',j}^\dagger \right) \quad (4.11)$$

$\lambda_{1,i,j}$ is the Lagrange multiplier regarding to the constraint (4.4b).

For the discrete variables, the user scheduling $\mathbf{u}_1$ can be determined through Hungarian algorithm as described in Chapter 2, with $W_{k,m}$ defined in (2.25) replaced by

$$\begin{aligned} W_{k,m} = & \log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m - |y_{1,i,j}^m|^2 \sigma^2 + 2 \operatorname{Re} \{ (y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} s_{1,i,j} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,1,i,j} \hat{\mathbf{u}}^k) \} \\ & - |y_{1,i,j}^m|^2 \sum_{i'=1}^B \sum_{m'=1}^M s_{1,i',j} |(\mathbf{H}_{1,i',1,i,j} \hat{\mathbf{u}}^k)^\dagger \mathbf{v}_{1,i',j}^{m'}|^2 \end{aligned} \quad (4.12)$$

Likewise, the selection of frequency slots $\mathbf{s}_1$ can be obtained by solving the optimization problem defined in (2.28), with $\mathcal{Q}_{1,i,j}$ replaced by

$$\begin{aligned} \mathcal{Q}_{1,i,j} = & \sum_{m=1}^M 2 \operatorname{Re} \{ (y_{1,i,j}^m)^\dagger \sqrt{(1 + \gamma_{1,i,j}^m)} (\mathbf{v}_{1,i,j}^m)^\dagger (\mathbf{H}_{1,i,1,i,j} \mathbf{u}_{1,i,j}^m) \} \\ & - \sum_{i'=1}^B \sum_{m'=1}^M |y_{1,i',j}^{m'}|^2 \sum_{m=1}^M |(\mathbf{H}_{1,i,1,i',j} \mathbf{u}_{1,i',j}^{m'})^\dagger \mathbf{v}_{1,i,j}^m|^2 - \sum_{m=1}^M |y_{1,i,j}^m|^2 \sum_{i''=1}^B \|\mathbf{H}_{1,i,2,i'',j}^\dagger \mathbf{v}_{1,i,j}^m\|_2^2 \end{aligned} \quad (4.13)$$

Consequently, combining all of the above steps, the iterative optimization process at each operator is summarized in the following algorithm, which is similar to Algorithm 1:

Algorithm 4: Resource Allocation and User Scheduling at Each Operator

```

1: Initialize  $\mathbf{s}, \mathbf{u}, \mathbf{v}$ 
2: while the objective function  $f_q^{(LI)}$  in (4.7) does not converge do
3:   for  $i = 1, 2, \dots, \mathcal{B}$  do
4:     Update  $\gamma_i$  using (4.6).
5:     Update  $\mathbf{y}_i$  using (4.8).
6:     Update  $\mathbf{v}_i$  using (4.10) and update  $\lambda_i$  with bisection search.
7:     Update  $\mathbf{u}_i$  using Hungarian Algorithm with  $W_{k,m}$  replaced by (4.12)
8:     Update  $\mathbf{s}_i$  by solving (2.28) with  $\mathcal{Q}_{1,i,j}$  replaced by (4.13).
9:   end for
10: end while

```

4.4 Simulation Results and Performance Evaluation

In this section, we look into the improvement made possible by the SLINR based spectrum sharing mechanism. The simulation environment used is exactly the same as in Chapter 2.

4.4.1 Diagonally Overlapped Cell Deployment

Firstly, we investigate the case of diagonally overlapped cell deployment.

The achieved average user rates of both operators are illustrated in Figure 4.1. Surprisingly, the SLINR-based spectrum sharing can reach 95% the performance as the coordinated SINR-based spectrum sharing, when the scaling factor $\epsilon = 0$. While the averaged user rate achieved when $\epsilon = 1$ is below the case of no spectrum sharing. This is because the total leakage power (4.1), which consists of the interference to all UTs of the other operator, is much larger than merely the real leakage to the scheduled UTs of the other operator. Since most of the scheduled UTs from the other operator are not adjacent to the BS, as illustrated in Figure 2.9. Importantly, although not shown in the results, we observed that the maximal average user rate with SLINR-based spectrum sharing is achieved with an ϵ close to 0. In addition, under this cell deployment structure, the improvement from the coordination between

Figure 4.1: Average user rate among 9 cells of both operators (SLINR).

the operators towards the sum rate is insignificant, since the inter-operator interference is limited.

The average user rates achieved for each operator are given in Figure 4.2. For the average user rates of OP1, we observe that the performance degradation of SLINR-based spectrum sharing ($\epsilon = 1$) is smaller when OP2 provides less spectrum. This is because the aggregate leakage in this case is smaller as we assume there are fewer UTs per cell when OP2 provides fewer frequency slots. Among the average rates of OP2, there is negligible improvement of SINR-based spectrum sharing over SLINR-based spectrum sharing, when there is a huge gap between the amount of spectrum provided by the two operators. This also implies that under this circumstance, increasing the sum-rate at OP1 will be prioritised regardless of the interference to OP2, in the optimization of SINR-based spectrum sharing, as discussed in Chapter 2.

4.4.2 Co-located Cell Deployment

Under co-located cell deployment, the achieved average user rates of both operators are illustrated in Figure 4.3 and the average user rates achieved for each operator are given in

(a) Average user rate among 9 cells of OP1 (SLINR).

(b) Average user rate among 9 cells of OP2 (SLINR).

Figure 4.2: Average user rate among 9 cells of each operator (SLINR).

Figure 4.4. In this case, the performance of SLINR-based spectrum sharing is much worse than SINR-based spectrum sharing. This implies that coordinated beamforming and user scheduling are critical in improving the system performance in an environment with heavy interference between the operators.

Figure 4.3: Average user rate among 9 cells of both operators (SLINR).

(a) Average user rate among 9 cells of OP1 (SLINR).

(b) Average user rate among 9 cells of OP2 (SLINR).

Figure 4.4: Average user rate among 9 cells of each operator (SLINR).

4.5 Summary

In this chapter, we simply investigate whether there is any possibility to deploy spectrum sharing with minimum information sharing between the operators. On the basis of a SINR-based spectrum sharing mechanism presented in Chapter 2, we propose a SLINR-based spectrum sharing optimization framework at each operator to determine the resource allocation and user scheduling individually. The simulation results reveal that coordination is necessary among operators if communications in one operator will strongly interfere the communications in another operator, while considerable improvement can be achieved without coordination once the inter-operator interference power is limited. In addition, we assume perfect channel knowledge between the operators in the discussions of this Chapter. Estimating the channels with minimum information dissemination is a challenging topic to be further studied in the future.

Chapter 5

Conclusions and Future Work

In this thesis, we investigate the topics of fair licensed spectrum sharing between two MNOs in multi-cell multi-user MIMO networks, including the coordinated resource allocation and beamforming, the methods of channel estimation, and reduction of inter-operator information exchange.

First, we developed a fair spectrum sharing system with inter-operator coordinated beamforming, where we formulate an optimization problem to determine the resource allocation and user scheduling between the two MNOs. We then developed an algorithm to tackle the problem through fractional programming and block coordinate descent. The simulation results reveal that, both operators can achieve visible improvement in average user rate with diagonally overlapped cell deployment; and both operators can have more users (especially users near the boundary of the cells) scheduled with co-located cell deployment. In conclusion, both operators have an incentive to participate in the spectrum sharing mechanism proposed, if the difference between the amount of spectrum provided by the two operators is not large.

Next, we investigated the impact of channel estimation within the proposed spectrum sharing framework, in a time division duplex (TDD) system. In addition to the conventional cell based pilot allocation, we develop a cluster based pilot allocation scheme depending on the cell layout, to balance the gap between number of UTs per cell of the two operators.

However, with channel estimation, the operator that provides much less spectrum than the other will hardly get performance improvement, which is different from the case when perfect knowledge of channels is ready at the BSs. Therefore, with basic channel estimation mechanisms, operators are willing to participate in, and benefit from the spectrum sharing framework, only when they can provide similar number of frequency slots.

Finally, we briefly look into whether the resource allocation and user scheduling in spectrum sharing can be implemented distributively at each operator, with minimal information dissemination. Specifically, we replace the inter-operator interference terms used in SINR-based spectrum sharing with the inter-operator leakage power. The results show that the gain in this case depends much on the cell deployment and the inter-operator interference power to the scheduled UTs.

Due to time limitations, the coverage of the research presented in this thesis is limited and far from complete. There are several interesting topics that can be further explored with the spectrum sharing mechanism proposed:

- *Real-world cell deployments:* Within this work, the fair spectrum sharing framework is simulated in two cases, i.e., the co-located cell deployment, which is the cell deployment with the most severe inter-operator interference, as well as the diagonally overlapped cell deployment, where the inter-operator interference is the least assuming fixed cell size. The cell deployment in the real world is in between the two cases, where there are BSs stand at the same location, and BSs that are far from each other. How much gain in spectrum sharing with real-world cell deployments remains an unanswered question.
- *Non-uniform user distribution:* In this thesis, we assume the UTs are uniformly distributed within each cell. However, there might be some hotspots in each cell where the user density are larger than somewhere else. In this setting, the channel estimation and interference pattern are likely to change significantly.
- *Infrastructure sharing:* Within the co-located cell deployment, the intra-operator interference and the inter-operator, intra-cell interference can hardly be eliminated simulta-

neously, as there may be more scheduled UTs in each frequency slot than the number of antennas at the BS. Thus, sharing the infrastructure (e.g., the BSs) may help further improve the system performance.

- *Spectrum reorganization*: In this work, we assume all spectrum from the operators are used in spectrum sharing. Thus, using part of the spectrum in spectrum sharing to improve the rate achieved at certain UTs, while keeping the remaining spectrum to other UTs exclusively is an interesting scheme to look at (e.g., under co-located cell deployment, for each operator, spectrum sharing is applied to UTs that are closer to cell edges with part of the aggregate spectrum, while UTs that are adjacent to the BSs access the remaining spectrum exclusively).

Bibliography

- [1] R. H. Tehrani et al. “Licensed Spectrum Sharing Schemes for Mobile Operators: A Survey and Outlook”. In: *IEEE Comm. Surveys & Tutorials* 18.4 (2016).
- [2] M. D. Mueck, V. Frascolla, and B. Badic. “Licensed shared access — State-of-the-art and current challenges”. In: *2014 1st International Workshop on Cognitive Cellular Systems (CCS)*. 2014, pp. 1–5.
- [3] Z. Jiang and S. Mao. “Interoperator Opportunistic Spectrum Sharing in LTE-Unlicensed”. In: *IEEE Transactions on Vehicular Technology* 66.6 (2017), pp. 5217–5228.
- [4] F. Teng, D. Guo, and M. L. Honig. “Sharing of Unlicensed Spectrum by Strategic Operators”. In: *IEEE Journal on Selected Areas in Communications* 35.3 (2017), pp. 668–679.
- [5] M. Asaduzzaman, R. Abozariba, and M. Patwary. “Dynamic Spectrum Sharing Optimization and Post-Optimization Analysis With Multiple Operators in Cellular Networks”. In: *IEEE Transactions on Wireless Communications* 17.3 (2018), pp. 1589–1603.
- [6] R. Abozariba, M. Asaduzzaman, and M. Patwary. “Radio Resource Sharing Framework for Cooperative Multioperator Networks With Dynamic Overflow Modeling”. In: *IEEE Transactions on Vehicular Technology* 66.3 (2017), pp. 2433–2447.
- [7] S. Dhavane and M. Z. A. Khan. “Impact of spectrum sharing rules in multioperator communication system”. In: *2018 10th International Conference on Communication Systems Networks (COMSNETS)*. 2018, pp. 447–449.
- [8] B. Cho et al. “Co-Primary Spectrum Sharing for Inter-Operator Device-to-Device Communication”. In: *IEEE Journal on Selected Areas in Communications* 35.1 (2017), pp. 91–105.
- [9] T. Sanguanpuak et al. “Multi-Operator Spectrum Sharing for Small Cell Networks: A Matching Game Perspective”. In: *IEEE Transactions on Wireless Communications* 16.6 (2017), pp. 3761–3774.

- [10] B. Singh et al. “Co-primary inter-operator spectrum sharing over a limited spectrum pool using repeated games”. In: *2015 IEEE International Conference on Communications (ICC)*. 2015, pp. 1494–1499.
- [11] B. Singh et al. “Repeated spectrum sharing games in multi-operator heterogeneous networks”. In: *2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. 2015, pp. 221–228.
- [12] Y. -. Chung and Z. Tsai. “Performance evaluation of dynamic spectrum sharing for two wireless communication networks”. In: *IET Communications* 4.4 (2010), pp. 452–462.
- [13] J. Jeon et al. “Coordinated Dynamic Spectrum Sharing for 5G and Beyond Cellular Networks”. In: *IEEE Access* 7 (2019), pp. 111592–111604.
- [14] G. P. Koudouridis and P. Soldati. “Joint network density and spectrum sharing in multi-operator collocated ultra-dense networks”. In: *2018 7th International Conference on Modern Circuits and Systems Technologies (MOCAST)*. 2018, pp. 1–4.
- [15] R. Litjens et al. “System-Level Assessment of Non-Orthogonal Spectrum Sharing via Transmit Beamforming”. In: *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*. 2013, pp. 1–6. DOI: 10.1109/VTCSpring.2013.6692778.
- [16] J. Lindblom and E. G. Larsson. “Does non-orthogonal spectrum sharing in the same cell improve the sum-rate of wireless operators?” In: *2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. 2012, pp. 6–10. DOI: 10.1109/SPAWC.2012.6292981.
- [17] R. K. Saha. “Nationwide Spectrum Sharing of Mobile Network Operators with Indoor Small Cells”. In: *DySPAN*. 2019, pp. 1–2.
- [18] A. K. Dinnis and J. S. Thompson. “The Effects of Including Wraparound When Simulating Cellular Wireless Systems with Relaying”. In: *2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring*. 2007, pp. 914–918. DOI: 10.1109/VETECS.2007.197.
- [19] Andrea Goldsmith. *Wireless Communications*. 2005.
- [20] Zukang Shen et al. “Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization”. In: *IEEE Transactions on Signal Processing* 54.9 (2006), pp. 3658–3663. DOI: 10.1109/TSP.2006.879269.
- [21] A. A. Khan, R. Adve, and W. Yu. “Optimizing Downlink Resource Allocation in Multiuser MIMO Networks via Fractional Programming and the Hungarian Algorithm”. In: *IEEE Transactions on Wireless Comm.* (2020).

- [22] K. Shen and W. Yu. “Fractional Programming for Communication Systems—Part II: Uplink Scheduling via Matching”. In: *IEEE Transactions on Signal Processing* 66.10 (2018), pp. 2631–2644. DOI: 10.1109/TSP.2018.2812748.
- [23] Robin J Wilson. *Introduction to Graph Theory*. USA: John Wiley & Sons, Inc., 1986. ISBN: 0470206160.
- [24] H. W. Kuhn. “The Hungarian method for the assignment problem”. In: *Naval Research Logistics (NRL)* 52.1 (Feb. 2005), pp. 7–21. DOI: 10.1002/nav.20053. URL: <https://ideas.repec.org/a/wly/navres/v52y2005i1p7-21.html>.
- [25] D. Bertsekas. “The auction algorithm: A distributed relaxation method for the assignment problem”. In: *Annals of Operations Research* 14 (1988), pp. 105–123.
- [26] Olakunle Elijah et al. “A Comprehensive Survey of Pilot Contamination in Massive MIMO—5G System”. In: *IEEE Communications Surveys Tutorials* 18.2 (2016), pp. 905–923. DOI: 10.1109/COMST.2015.2504379.
- [27] Thomas L. Marzetta. “How Much Training is Required for Multiuser MIMO?” In: *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*. 2006, pp. 359–363. DOI: 10.1109/ACSSC.2006.354768.
- [28] Jubin Jose et al. “Pilot Contamination and Precoding in Multi-Cell TDD Systems”. In: *IEEE Transactions on Wireless Communications* 10.8 (2011), pp. 2640–2651. DOI: 10.1109/TWC.2011.060711.101155.
- [29] Nuwan Ganganath, Chi-Tsun Cheng, and Chi K. Tse. “Data Clustering with Cluster Size Constraints Using a Modified K-Means Algorithm”. In: *2014 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*. 2014, pp. 158–161. DOI: 10.1109/CyberC.2014.36.
- [30] Javad Mirzaei, Raviraj S. Adve, and Shahram Shahbazpanahi. “Semi-Blind Time-Domain Channel Estimation for Frequency-Selective Multiuser Massive MIMO Systems”. In: *IEEE Transactions on Communications* 67.2 (2019), pp. 1045–1058. DOI: 10.1109/TCOMM.2018.2875724.
- [31] Cunhua Pan et al. “The Non-Coherent Ultra-Dense C-RAN Is Capable of Outperforming Its Coherent Counterpart at a Limited Fronthaul Capacity”. In: *IEEE Journal on Selected Areas in Communications* 36.11 (2018), pp. 2549–2560. DOI: 10.1109/JSAC.2018.2874138.
- [32] Xin Xia et al. “Leakage-based user scheduling in MU-MIMO broadcast channel”. In: *Science in China Series F: Information Sciences* 52 (2009), pp. 2259–2268.

- [33] Mirette Sadek, Alireza Tarighat, and Ali Sayed. “A Leakage-Based Precoding Scheme for Downlink Multi-User MIMO Channels”. In: *Wireless Communications, IEEE Transactions on* 6 (June 2007), pp. 1711–1721. DOI: 10.1109/TWC.2007.360373.
- [34] Tyler Gamvrelis, Ahmad A Khan, and Raviraj S Adve. “SLNR- and SLINR-Based Downlink Optimization in MU-MIMO Networks”. In: *IEEE Transactions on Wireless Communications* (2021). Under submission, to appear.