Mathematical Optimization Techniques for Resource Allocation and Spatial Multiplexing in Spectrum Sharing Networks

---

by

Jie Tang

A Doctoral Thesis submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy (PhD), at Loughborough University.

September 2012

Advanced Signal Processing Group,
School of Electronic, Electrical and Systems Engineering,
Loughborough University, Loughborough
Leicestershire, UK, LE11 3TU.

CERTIFICATE OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this thesis, that the original work is my own except as specified in acknowledgements or in footnotes, and that neither the thesis nor the original work contained therein has been submitted to this or any other institution for a degree.

..... (Signed)

Jie Tang (Candidate)

I dedicate this thesis to my dear parents Jianhua Tang and Jianhua Guo and my beloved wife Jieqian Zhou.

Abstract

Due to introduction of smart phones with data intensive multimedia and interactive applications and exponential growth of wireless devices, there is a shortage for useful radio spectrum. Even though the spectrum has become crowded, many spectrum occupancy measurements indicate that most of the allocated spectrum is underutilised. Hence radically new approaches in terms of allocation of wireless resources are required for better utilization of radio spectrum.

This has motivated the concept of opportunistic spectrum sharing or the so-called cognitive radio technology that has great potential to improve spectrum utilization. The cognitive radio technology allows an opportunistic user namely the secondary user to access the spectrum of the licensed user (known as primary user) provided that the secondary transmission does not harmfully affect the primary user. This is possible with the introduction of advanced resource allocation techniques together with the use of wireless relays and spatial diversity techniques.

In this thesis, various mathematical optimization techniques have been developed for the efficient use of radio spectrum within the context of spectrum sharing networks. In particular, optimal power allocation techniques and centralised and distributed beamforming techniques have been developed. Initially, an optimization technique for subcarrier and power allocation has been proposed for an Orthogonal Frequency Division Multiple Access (OFDMA) based secondary wireless network in the presence of multiple primary users. The solution is based on integer linear programming with multiple interference leakage and transmission power constraints. In order to enhance the spectrum efficiency further, the work has been extended to allow multiple secondary users to occupy the same frequency band under a multiple-input and multiple-output (MIMO) framework. A sum rate maximization technique based on uplink-downlink duality and dirty paper coding has been developed for the MIMO based OFDMA network. The work has also been extended to handle fading scenarios based on maximization of ergodic capacity. The optimization techniques for MIMO network has been extended to a spectrum sharing network with relays. This has the advantage of extending the coverage of the secondary network and assisting the primary network in return for the use of the primary spectrum. Finally, instead of considering interference mitigation, the recently emerged concept of interference alignment has been used for the resource allocation in spectrum sharing networks. The performances of all these new algorithms have been demonstrated using MATLAB based simulation studies.

Contents

Statement of Originality

The contributions of this thesis are mainly on the development of various resource allocation and spatial multiplexing techniques for spectrum sharing networks. The following aspects of this thesis are believed to be originals:

- Resource allocation and admission control techniques were proposed for an Orthogonal Frequency Division Multiple Access (OFDMA)-based cognitive radio network in Chapter 3. In particular, an integer linear programming based optimization technique was proposed to optimally allocate subcarriers, power and bits to secondary users while satisfying the data rate and bit error rate (BER) requirements for each secondary user [5]. Furthermore, to solve a problem where the number of users seeking access to the network exceeds the available resources, a suboptimal optimization algorithm based on integer programming was proposed to admit as many secondary users as possible while allocating subcarriers and bits to each admitted user in such a way that the interference leakage to the primary users is below a specific threshold [7].

- A weighted sum rate maximization and rate balancing techniques have been proposed for Multiple-Input and Multiple-Output Orthogonal Frequency Division Multiplexing (MIMO-OFDM) based spectrum sharing broadcast channels [1]. This problem has been solved by converting the MIMO-OFDM channel into block diagonal form and using the principle of broadcast channel - multiple access channel (BC-MAC) duality. In addition, the sum rate maximization problem was also solved for spectrum sharing MIMO BC under Rayleigh fading environment using multiple auxiliary variables, Karush-Kuhn-Tucker (KKT) optimality conditions and BC-MAC duality [11]. The work has also been extended to consider rate balancing in spectrum sharing

MIMO based wireless relay network [10].

- Beamforming and power control for an overlay cognitive radio network has been proposed using second order cone programming in [2].

- In Chapter 6, interference alignment algorithms have been proposed for a multi-cell multi-user MIMO Gaussian interfering broadcast channels (MIMO-IFBC) [3]. A grouping method already known in the literature has been extended to a multiple-cells scenario to jointly design transmit and receiver beamforming vectors using a closed-form expression. A new approach using the principle of BC-MAC duality has been proposed to perform interference alignment while maximizing capacity of users in each cell. To further enhance the performance, a MIMO cognitive radio network with a MIMO relay that opportunistically accesses the same frequency band as that of a MIMO primary network has been proposed in [4].

The novelty of the contributions is supported by the following international journals and conference papers.

Journal Papers

1. J. Tang, K. Cumanan and S. Lambotharan, "Sum Rate Maximization Technique for Spectrum Sharing MIMO OFDM Broadcast Channels", IEEE Transactions on Vehicular Technology, Vol. 60, no. 4, pp. 1960-1964, May, 2011.

2. J. Tang and S. Lambotharan, "Beamforming and Temporal Power Optimization for an Overlay Cognitive Radio Relay Network", IET Signal Processing, Vol. 5, no. 6, pp. 582-588, Sept., 2011.

3. J. Tang and S. Lambotharan, "Interference Alignment Techniques for

MIMO Multi-Cell Interfering Broadcast Channels", accepted for publication in IEEE Transactions on Communications, 2012.

4. J. Tang, S. Lambotharan and S. Pomeroy, "Interference Cancelation and Alignment Techniques for MIMO Cognitive Relay Networks", submitted to IEEE Transactions on Vehicular Technology, 2012.

Conference Papers

5. J. Tang, R. Yogachandran and S. Lambotharan, "Adaptive Bit Loading and Subcarrier Allocation Techniques for OFDM Based Cognitive Radio Systems", IEEE International Conference on Communication Technology (ICCT), Nanjing, China, Nov. 2010.

6. W. Wang, N. K. Zong, J. Tang and S. Lambotharan, "Soft Information Improvement for PN sequence iterative acquisition", IEEE International Conference on Computational Intelligence and Security (CIS), Nanning, China, 11-14 Dec. 2010.

7. J. Tang, and S. Lambotharan, "A Suboptimal User Maximization Algorithm for an OFDMA Based Cognitive Radio Network", IEEE Vehicular Technology Conference (VTC), Budapest, Hungary, May. 2011.

8. K. Cumanan, J. Tang and S. Lambotharan, "Rate Balancing Based Linear Transceiver Design for Multiuser MIMO System With Multiple Linear Transmit Covariance Constraints", IEEE International Conference on Communicatons (ICC), Kyoto, Japan, May. 2011.

9. J. Tang, Amod J.G. Anandkumar and S. Lambotharan, "Opportunistic MIMO Multi-Cell Interference Alignment Techniques", International Workshop on Cognitive Radio and Smart Antennas (CORSA), Bangalore, India, Dec. 2011.

10. J. Tang, and S. Lambotharan, "Optimal Resource Allocation Technique for Spectrum Sharing MIMO Based Wireless Relay Systems",

IEEE International Conference on Communications (ICC), Ottawa, CA, Jun. 2012.

11. J. Tang, G. Bournaka and S. Lambotharan, "Sum Rate Maximization for Spectrum Sharing Multiuser MIMO Network under Rayleigh Fading", IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Bali, Indonesia, July, 2012.

Acknowledgements

I AM DEEPLY INDEBTED to my supervisor Professor Sangarapillai Lambotharan for his kind interest, generous support and inspiring supervision throughout the past three years. I have benefitted tremendously from his rare insight, his ample intuition, his boundless patience and his profound knowledge. This thesis would never have been written without his tireless and patient mentoring. It is my very great privilege to have been one of his research students.

I am extremely thankful to Dr Simon Pomeroy for providing me with valuable comments on the IEEE Transactions on Vehicular Technology journal work that was submitted in July 2012. Also, I would like to thank Professor Jonathan Chambers for his support and encouragement.

I would like to thank the Loughborough University, as well as the School of Electronic, Electrical and Systems Engineering, for providing outstanding training and facilities to research students. I am very grateful to the Department Scholarships Awards for the financial support of this work.

I would like to express my appreciations to many members from ASPG, who have generously helped me to develop my knowledge through wide-ranging discussions. I would like to extend my appreciations to my friends for making my stay at Loughborough pleasant.

I wish to take this opportunity to thank my parents Jianhua Tang and Jianhua Guo for their moral support throughout my studies.

Last, but most importantly, I wish to express my deepest gratitude and love to my beloved wife, Jieqian Zhou, for her endless love and support. She is the reason for my blissful life at United Kingdom. I would like to dedicate this thesis to my parents and wife.

Jie Tang

September, 2012

List of Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
AWGN	Additive White Gaussian Noise
B3G	Beyond 3G
BC	Broadcast Channel
BER	Bit Error Rate
BnB	Branch and Bound
BPSK	Binary Phase Shift Keying
BS	Base Station
CDMA	Code Division Multiple Access
COFDM	Coded OFDM
CR	Cognitive Radio
CRN	Cognitive Radio Network

CSI	Channel State Information
CSCG	Circularly Symmetric Complex Gaussian
DAB	Digital Audio Broadcasting
DECT	Digital Enhanced Cordless Telecommunication
DFT	Discrete Fourier Transform
DoF	Degrees of Freedom
DPC	Dirty Paper Coding
DVB-T	Digital Video Broadcasting Terrestrial
EDGE	Enhanced Data rates for GSM Evolution
EV-DO	Evolution-Data Optimised
EV-DV	Evolution-Data/Voice
FDD	Frequency Division Duplexing
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FM	Frequency Modulation
GPRS	General Packet Radio Service
GSM	Global System for Mobile
IA	Interference Alignment
ICI	Inter-Cell Interference
IDFT	Inverse Discrete Fourier Transform

IFFT	Inverse Fast Fourier Transform
ILP	Integer Linear Programming
IP	Integer Programming
IS	Interim Standard
IUI	Inter-User Interference
LTE	Long Term Evolution
MA	Margin Adaptive
MAC	multiple access channel
MIMO	Multiple-Input-Multiple-Output
MISO	Multiple-Input-Single-Output
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
QAM	Quadrature Amplitude Modulation
PDF	Power Density Function
PNB	Primary Network Basestation
PU	Primary User
QoS	Quality-of-Service
QPSK	Quadrature Phase Shift Keying
RA	Rate Adaptive
RS	Relay Station
SDMA	Space Division Multiple Access

SDP	Semi Definite Programming
SIC	Successive Interference Cancellation
SIMO	Single-Input-Multiple-Output
SINR	Signal-to-Noise-plus-Interference Ratio
SNB	Secondary Network Basestation
SNR	Signal to Noise Ratio
SISO	Single-Input-Single-Output
SVD	Singular Value Decomposition
SU	Secondary User
TDMA	Time Division Multiple Access
TDMA-SCDMA	Time Division Synchronous Code Division Multiple Access
WiMax	Worldwide Interpretability for Microwave Access
W-CDMA	Wideband Code Division Multiple Access

List of Symbols

Scalar variables are denoted by plain lower-case letters, (i.e., x), vectors by bold-face lower-case letters, (i.e., $\mathbf{x}$), and matrices by upper-case bold-face letters, (i.e., $\mathbf{X}$). Some frequently used notations are as follows:

$\mathrm{E}\{\cdot\}$	Statistical expectation
$(\cdot)^T$	Transpose
$(\cdot)^H$	Hermitian transpose
$\|.\|$	Euclidean norm
$(.)^{-1}$	Matrix inverse
$\mathbf{1}_M$	$M \times 1$ vector of ones
$\mathbf{0}_M$	$M \times 1$ vector of zeros
$\mathbf{1}_{M,N}$	$M \times N$ matrix of ones
$\mathbf{0}_{M,N}$	$M \times N$ matrix of zeros
$\mathrm{tr}\{\cdot\}$	Trace operator
$\odot$	Hadamard product
$\mathbf{I}$	Identity matrix
$\mathbf{diag}(\mathbf{x})$	Diagonal matrix with vector $\mathbf{x}$

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Evolution of Wireless Communication Systems

Due to fast development of smart phones and associated multimedia and interactive applications, wireless communication systems have been experiencing an explosive growth. The cellular wireless communication system and wireless local area network (WLAN) is the most successful wireless application, nowadays, which is also an important element for globally ubiquitous wireless connections. In the 90s, wired local area network (LAN) was the dominated technology worldwide. However, at the end of 90s, the idea of wireless LAN came up and acted as an attractive technology to expend the wired LAN system. The major advantage of this technology is that the users do not need to connect to the network using cables thus it improves the flexibility of the systems.

The first WLAN standard was approved by the Institute of Electrical and Electronic Engineering (IEEE) 802.11 group [2]. After that, IEEE 802.11 has proposed two new standards 802.11a [3] and 802.11b [4] in order to satisfy the demand for high data throughput on limited channel bandwidth. Nowadays, WLANs (802.11a/g/n) are being utilized everywhere such as offices, universities and homes as they are of low cost and easy access.

The mobile wireless communication standards evolve rapidly. Based on analogue technologies, the first generation (1G) of mobile systems were developed in the 1980s. These first generation (1G) networks relied on analog

Frequency Modulation (FM), where each user was assigned a separate downlink and uplink FM channel. This method of disjoint frequency sharing is called Frequency Division Multiple Access (FDMA) with Frequency Division Duplexing (FDD). However, due to the low quality of analogue systems, digital technologies based on second generation system were introduced.

Due to the capacity limitation of 1G cellular systems, these were phased out by the second generation (2G) cellular systems in the early 1990s. There are three major 2G standards, Interim Standard (IS)-95, IS-136 in the United States, and Global System for Mobile (GSM) in Europe. The most widely used 2G standard in the world today, with more than 4.4 billion subscribers, is GSM. Unlike 1G cellular system that relied exclusively on FDMA/FDD and analog FM, 2G standards use digital modulation formats and time division multiple access (TDMA)- FDD (TDMA/FDD) technique. The enhanced versions of 2G GSM standards with higher data-rate are known as General Packet Radio Service (GPRS) and Enhanced Data rates for GSM Evolution (EDGE) for GSM [5]. These improved 2G cellular systems are usually referred to as 2.5G systems [6].

The increasing demands of multimedia introduced the development of the third generation mobile technology (3G), which is based on three different access technologies. The major technology of 3G standards include Wideband Code Division Multiple Access (W-CDMA), Time Division - Synchronous Code Division Multiple Access (TD-SCDMA), Digital Enhanced Cordless Telecommunication (DECT), CDMA2000 including CDMA2000 1x, CDMA2000 3x, CDMA2000 Evolution-Data Optimised (EV-DO) and CDMA2000 Evolution-Data/Voice (EV-DV) [7].

However, as the data throughput of the 3G system is unable to meet the ever increasing requirements in terms of high data rate multimedia applications, beyond 3G (B3G) and fourth generation (4G) standards were proposed. The major candidates are Third Generation Partnership Project

(3GPP) covering GSM and W-CDMA specifications [8], Third Generation Partnership Project 2 (3GPP2) exploiting CDMA2000 specifications [9] and the Worldwide Interpretability for Microwave Access (WiMAX) forum developed from IEEE 802.16 standard [10]. By adding multiple antennas at both the transmitters and the receivers, Long Term Evolution (LTE) [11] is being designed to provide improved spectral efficiency and capacity without significantly increasing the operating costs.

By utilizing multiple antennas at both the transmitter and the receiver, Multiple-Input and Multiple-Output (MIMO) architecture was introduced to provide spatial diversity gain and to increase the reliability of the wireless links. The potential to dynamically form arrays of antennas by aggregating those on several distinct users offers the possibility of significantly increasing the flexibility and availability of MIMO technology, without needing complex radio frequency hardware to manage large numbers of antennas on individual units.

Multi-carrier modulation techniques are an attractive solution for providing high data throughput. The basic idea of multi-carrier modulation is to devide the whole channel bandwidth into a number of smaller subbands. The data stream modulated on each sub-band is facing narrowband thus improving the system's performance [10]. However, how to separate the data stream modulated on each of the sub-carrier is one of the major problems. Orthogonal Frequency Division Multiplexing (OFDM) is therefore an attractive multi-carrier modulation technology used nowadays. Orthogonal Frequency Division Multiple Access (OFDMA) is a multiple access scheme based on OFDM. The basic idea of OFDMA is to allocate different number of subchannels to different users in order to support different quality of service (QoS) for several users [12]. OFDMA was assumed in the Long Term Evolution of the Third Generation Partnership Project. Several existing specifications have chosen MIMO-OFDM(A) based air-interfaces, for

example WiMAX based on IEEE 802.16e standard [13] and 3GPP LTE [8].

The technology development in wireless communication throughout the generations has boosted the growth of wireless applications. However, the limited availability of spectrum resources has brought the necessity of efficient spectrum utilization in wireless communications [14]. Traditionally, frequency bands are divided into various sub-bands and each sub-band is licensed to operators by spectrum regulatory bodies (i.e., by OFCOM in United Kingdom). The continuous growth in wireless applications have caused spectrum crisis and saturation in the frequency allocation table. Hence, spectrum shortage has become one of the key issues in spectrum allocation. On the other hand, different spectrum measurements showed that most of the time the licensed frequency bands are under-utilized [14, 15]. This motivates the concept of cognitive radio technology that has great capabilities to improve spectrum utilization. Hence, the focus of this thesis is on various spectrum efficient resource allocation techniques for wireless network that incorporate MIMO, OFDMA and cognitive radio technologies.

1.2 Motivation for Spectrum Sharing Cognitive Radio Networks

According to the studies by the Federal Communication Commission (FCC), the usage of the allocated spectrum is under-utilized [14]. This report indicates that the spectrum resources allocation policy may be inefficient. Cognitive Radio (CR) [16] is a promising technology which is capable of solving the inefficient spectrum usage problem. The idea of cognitive radio is to allow secondary user (SU, also known as unlicensed user) in the network to share the spectrum resources which are allocated to some primary users (PU, also known as licensed user). The spectrum can be accessed by secondary users when it is not being used by the primary users. The spectrum slots that are not being used are called as spectrum holes as shown in Fig.

1.1. Spectrum holes represent the potential opportunities for non-interfering (safe) use of spectrum and can be considered as multidimensional regions within frequency, time, and space. The spectrum utilization efficiency could

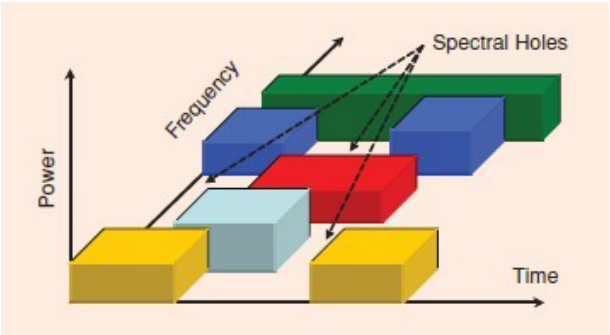

Figure 1.1. A conceptual illustration of spectrum utilization over time and frequency. Within a certain geographical region and at a certain time, some frequency bands are not used by legacy systems.

be greatly improved by using this technology. As a result, cognitive radio has been an interesting research topic recently. Sharing the licensed spectrum by secondary users improves the overall spectrum utilization and at the same time the transmission power of secondary user causes interference to primary user. Therefore, secondary user network should be designed in a way to allocate its radio resources to satisfy its own QoS requirements while ensuring that the interference caused to the primary users is below the predefined threshold level. The main functions of a cognitive radio network (CRN) are spectrum sensing and exploitation of available spectrum by adjusting the transmission parameters such as frequency allocation and transmission power.

In order to use the licensed spectrum, cognitive radio networks should detect the under-utilized licensed frequency bands. The performance of the detection schemes is mainly affected by channel fading and shadowing. There are difficulties in differentiating the attenuated primary signal from a white noise spectrum. This spectrum sensing problem has been widely studied and different spectrum sensing schemes have been proposed to improve the

detection performance [17, 18]. Spectrum detection techniques can be classified based on the type of detection techniques employed at the receiver: energy detection [19], coherent detection [20] and cyclostationary feature detection [21]. Energy detection is optimal when the information on the primary signal is limited. Coherent detection can be efficiently employed when the primary pilot signal is known, whereas a cyclostationary detector has the potential to distinguish the primary signal energy from the local noise energy.

There are three fundamental operational modes of the cognitive radio network, namely, interweave, overlay, and underlay [22]. Interweave method is known as an opportunistic spectrum access (OSA) method, whereby the secondary user transmits over the spectrum which is originally allocated to a primary user when the primary user transmission is detected to be off, as shown in Fig. 1.2. The overlay and the underlay methods allow the cogni-

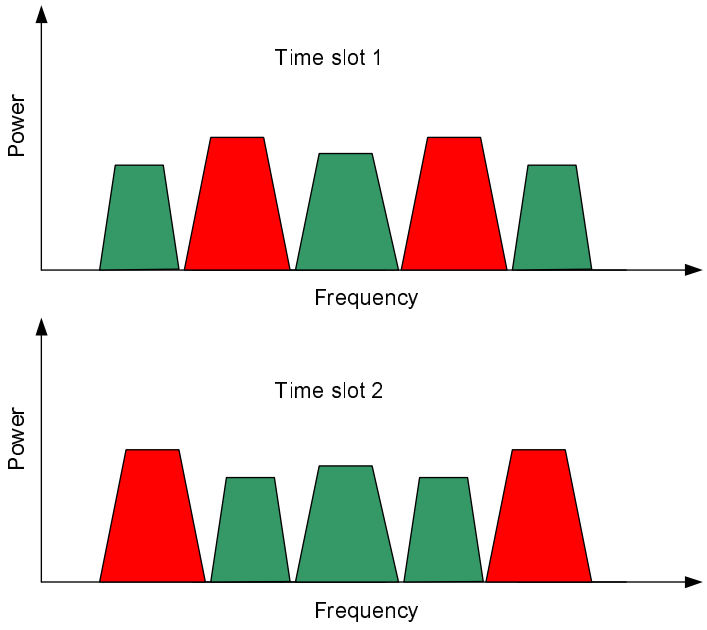

Figure 1.2. Interweave spectrum scheme. Green and red represent the spectrum occupied by the primary users and secondary users respectively.

tive radio to operate simultaneously with the primary user. Overlay method

is based on the cognitive relay idea presented in [23], the secondary users coexist with primary users and use part of the transmission power to relay the primary users' signals to the primary receiver. This assistance will offset the interference caused by the secondary user transmissions at the primary users' receiver. Hence, there is no loss in primary users' signal-to-noise ratio by secondary users spectrum access. The underlay method regulates the interference power level at the secondary user receiver introduced by the primary user transmission to be below a predefined threshold which is also known as interference temperature [24] as shown in Fig. 1.3.

In the interweave approach, identifying spectrum holes in the absence

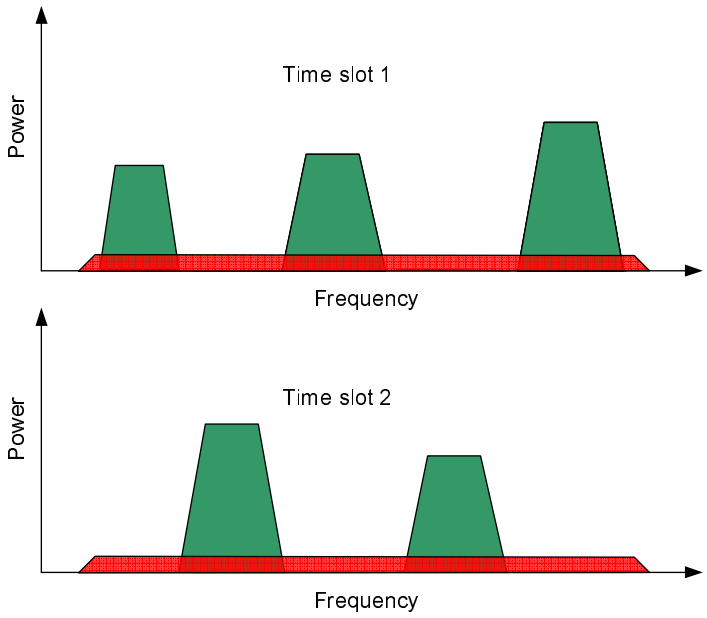

Figure 1.3. Underlay spectrum paradigm. Green and red represent the spectrum occupied by the primary users and the secondary users respectively.

of cooperation between primary and secondary networks is very challenging. For example, a secondary transmitter could be in the shadow region of the primary transmitter which will falsely indicate (to the secondary transmitter) availability of spectrum. The secondary transmission based on this false indication may harm the primary receivers. This hidden terminal problem

is deemed to be very challenging and a limiting factor for the employment of interweave cognitive radio networks. Hence, the underlay scheme and the overlay scheme seem more realistic and easy to implement compared to the other schemes. The resource allocation techniques for underlay and overlay cognitive network is the main focus of this thesis.

1.3 Thesis Outline

In wireless communications, adaptive resource allocation and spatial multiplexing techniques significantly enhance the spectrum utilization. The techniques developed for conventional wireless networks cannot be directly applied to a cognitive radio network due to the additional interference constraints on the primary users. Hence the work in this thesis mainly focuses on the resource allocation and spatial multiplexing techniques for cognitive radio networks using various mathematical optimization techniques.

The Chapter 2 provides a survey on resource allocation and spatial multiplexing techniques used in conventional wireless networks. Initially, static and adaptive resource allocation techniques for wireless networks are introduced and different types of adaptive resource allocation techniques followed by resource allocation using multiple antennas techniques such as beamforming and spatial multiplexing are discussed briefly. Following on from this, the capacity of a single-user wireless channel is studied where the transmitter and/or receiver have a single antenna. Capacity of a single-user system where the transmitter and the receiver have multiple antennas as well as the capacity of a multiuser network are also summarized. Finally, an interference alignment (IA) technique is introduced as an efficient multiuser capacity achieving scheme at high signal to noise ratio (SNR) regime.

The novel contributions of this thesis are provided in Chapters 3, 4, 5 and 6. Resource allocation techniques for OFDMA based cognitive radio

networks are introduced in Chapter 3. Resource allocation problem is first considered for an OFDMA-based cognitive radio network. This problem can be formulated within an integer linear programming framework and solved using branch and bound method. The solution to this problem optimally allocates the subcarriers, power and bits to secondary users while satisfying the data rate and bit error rate requirements for each secondary user. It ensures that the interference leakage to the primary users is always less than a specific threshold. Admission control for an OFDMA based cognitive radio network is then studied in Chapter 3. A suboptimal optimization algorithm based on integer programming is proposed to admit as many secondary users as possible in the network while allocating subcarriers and bits to each admitted user in such a way that the interference leakage to primary users is below a specific threshold.

Chapter 4 studied the capacity optimization problems for spectrum sharing networks. A weighted sum rate maximization and rate balancing problems have been investigated first for spectrum sharing MIMO OFDM based broadcast channels. This problem has been solved by converting the MIMO OFDM channel into block diagonal form and using the principle of MAC-BC duality. Then a sum rate maximization problem for spectrum sharing MIMO broadcast channels under Rayleigh fading has been studied. In contrast to previous problem on MIMO CR networks, fading channels are assumed and the ergodic capacity region has been studied. Based on the multiple auxiliary variables, the Karush-Kuhn-Tucker optimality conditions and the broadcast channel-multiple access channel (BC-MAC) duality, an iterative algorithm has been developed to solve the equivalent problem using a Lagrangian formulation and it is shown that the proposed algorithm converges to the setting as defined by the optimization problem, i.e., the interference and the power limits are met whilst maximizing the sum capacity. Finally, a weighted sum rate maximization and rate balancing problem has been proposed for a spec-

trum sharing MIMO based wireless relay network. The aim is to maximize the sum rate of the wireless relay network whilst ensuring the interference leakage to the primary user terminals during two time slots are below a specific value. This problem is solved by asymmetrically allocating the power to different time slots and using the principle of MAC-BC duality.

Chapter 3 and 4 focus on underlay cognitive radio network where the secondary users ensure that interference leakage to the primary users is below an acceptable level. However, in the overlay cognitive radio network, the secondary users coexist with primary users and use part of the transmission power to relay the primary users' signals to the primary receiver. This assistance will offset the interference caused by the secondary user transmissions at the primary users' receiver and hence no loss in primary users' signal-to-noise ratio by secondary users spectrum access. Chapter 5 focuses on beamforming and power control for an overlay cognitive radio network. Co-existence of a secondary network with a primary network under an overlay framework has been considered. The secondary network serves multiple users in the same frequency band as of the primary network, however, in order to compensate the interference leakage to the primary user terminals, the secondary network acts as a relay to forward the primary user signals. The interference and noise level at the primary terminals during various time slots are different, therefore, the primary network needs to allocate resources asymmetrically during various time slots for the optimal performance. Hence, a joint spatial and temporal resource allocation technique has been proposed to enhance the overall system power saving while satisfying the data rate or the signal to interference plus noise ratio (SINR) requirement of the primary and the secondary users.

Interference mitigation techniques have become an important part of wireless network design. Interference alignment (IA) technique has been proposed recently as an efficient capacity achieving scheme at high SNR

regime. Interference alignment techniques are studied for various spectrum sharing networks in Chapter 6. Interference alignment algorithms are first proposed for a multi-cell MIMO Gaussian interfering broadcast channels. A grouping method already known in the literature has been extended to this multiple-cells scenario and transmit and receiver beamforming vectors are jointly designed using a closed-form expression without iterative computation. Based on the grouping method, a new approach using the principle of MAC-BC duality is proposed to perform interference alignment while maximizing capacity of users in each cell. The proposed approach has been shown to outperform the extension of the grouping method in terms of capacity and basestation complexity. A MIMO cognitive radio network with a MIMO relay that opportunistically accesses the same frequency band as that of a MIMO primary network is then considered. In particular both interference cancelation and interference alignment techniques have been investigated to enhance the achievable degrees of freedom for the MIMO cognitive radio network. The degrees of freedom obtained by the cognitive radio network in the presence of a MIMO relay has been shown to be higher than that could be obtained without a relay. The analyses considered both sufficient and insufficient number of antennas at the relay in terms of the ability to separate and decode both the primary and secondary transmitted signals.

Conclusions are drawn in Chapter 7. A brief summary is also provided and potential future directions are identified.

Chapter 2

RESOURCE ALLOCATION AND SPATIAL MULTIPLEXING TECHNIQUES FOR WIRELESS COMMUNICATION SYSTEM

2.1 Introduction

The exponential growth of wireless communication systems opened up new challenges in the network design. One of the major challenges faced by wireless communication systems is the increasing demand for spectral resources created by high data rate services such as rich multimedia and interactive services. Designing wireless communication systems with high spectrum efficiency and utilization is one of the most important research objectives. In the wireless communications, various resource allocation techniques have been proposed to utilize the wireless scare resources efficiently. These tech-

niques involve strategies and algorithms for controlling transmission power, frequency allocation, modulation scheme and error coding. The main objective of the resource allocation scheme is to make the best use of the scarce radio resources to increase spectrum efficiency as much as possible [25].

Multi-carrier modulation techniques are an attractive solution for providing higher data throughput on limited bandwidth channels. The basic idea of multi-carrier modulation is to divide the wide band channel into multiple narrow band subchannels. Hence, the frequency selective wide band channel is converted into frequency non-selective multiple subchannels, which means the data stream modulated on each sub-band is facing narrowband thus improving the system's performance.

Multi-antenna techniques can be used in wireless systems to achieve higher spectral efficiency in terms of higher throughput, more users per cell and improved coverage [26]. This spectral efficiency can be achieved by deploying multiple antennas at the transmitter and/or at the receiver. However, for the cognitive radio network, in addition to the data rate consideration, the interference at the receiving points of the primary network should be maintained below a target value. Hence, resource allocation in cognitive radio network should control the interference to primary user. One of the promising approaches to mitigate interference is antenna array processing technique such as beamforming techniques. Beamforming is a general signal processing technique used in the physical layer of a communication channel to control the directionality of the transmission or the reception of a signal using an array of antennas. Beamforming takes advantage of interference to change the directionality of the array thus improves the system's performance.

Interference mitigation techniques have become an important part of wireless network design. An interference alignment (IA) technique has been proposed recently in [27] as an efficient multiuser capacity achieving scheme

in a high signal to noise ratio (SNR) regime. The fundamental concept of IA is to align all interference signals in a particular subspace at each receiver so that an interference-free orthogonal subspace can be solely allocated for data transmission. Since the work of [27], IA techniques have attracted significant interests and various algorithms have been proposed and analyzed, for example, multiple-input multiple-output (MIMO) interference network [28,29], X network [30], and cellular network [31,32].

In this chapter, multi-carrier modulation techniques as well as multi-antenna techniques for general wireless communication system are introduced. It is followed by beamforming techniques, including receiver beamforming techniques and transmitter beamforming techniques. Spatial multiplexing techniques are introduced to further improve the system performance against fading. The capacity of a single-user wireless channel is studied where the transmitter and/or receiver have a single antenna. Capacity of a single-user system where the transmitter and the receiver have multiple antennas as well as the capacity of a multiuser system are also treated. The duality of Gaussian multiple-access and broadcast channels are introduced for the wireless communication system. Finally, interference alignment techniques and the related research literatures are reviewed.

2.2 Multi-carrier Modulation Technology

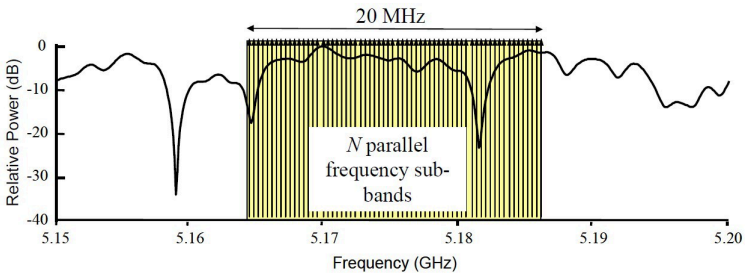

Figure 2.1. Multicarrier modulations

Multi-carrier modulation technique is an attractive solution for provid-

ing high data throughput for band limited channels, as shown in Fig. 2.1. The basic idea of multi-carrier modulation is to divide the whole channel bandwidths into a number of smaller sub-bandwidth. The data stream modulated on each sub-band is facing narrowband thus improving the system's performance [23]. However, how to separate the data stream modulated on each of the sub-carrier is one of the major problems. Orthogonal Frequency Division Multiplexing (OFDM) is therefore a dominated multi-carrier modulation technology used nowadays. In an OFDM system the subcarriers are constrained by the inverse Fast Fourier transform (IFFT) to lie in a contiguous block with constant spacing. These constraints ensure that the subcarrier frequencies are orthogonal at the transmitter. This is vital since it allows the data sent on each subcarrier to be separated mathematically at the receiver.

2.2.1 Introduction to OFDM

OFDM systems modulate their transmit data onto N frequency parallel subcarriers using an IFFT block. The value of N is constrained to a power of two. The subcarriers are made orthogonal by forcing the frequency spacing to be the reciprocal of the OFDM symbol period. This ensures that the carriers are spaced at the Nyquist minimum, which in turn maximises the bandwidth efficiency. Since the subcarriers are separated in the receiver by performing mathematically orthogonal operation as shown in 2.2.1. There is no need for guard frequencies between the subcarriers. As shown in Fig. 2.2, the orthogonal subcarriers actually overlap.

$$\int_0^T \cos(2\pi f_n t)\cos(2\pi f_k t)\mathrm{dt} = 0 \begin{cases} 0, & \text{For } n \neq k; \\ R, & \text{For } n = k. \end{cases} \tag{2.2.1}$$

To ensure that the orthogonality of each subcarrier is maintained, an integer number of cycles in the receive Fast Fourier transform (FFT) win-

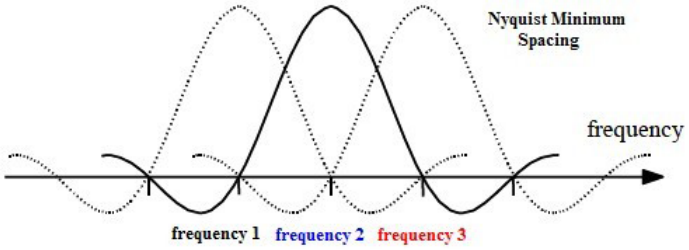

Figure 2.2. Subcarrier orthogonality

dow is required. Individual subcarriers in an OFDM system result in time-domain tones in the OFDM symbol period when transformed via the IFFT. As shown in Fig. 2.3, the first subcarrier results in a single cycle across the time-domain symbol period (the blue curve). The second subcarrier results in two complete cycles (the red curve). The fourth subcarrier results in four complete cycles (the green curve). In practice the amplitude and the phase of each of subcarrier are modulated by a sample taken from the complex modulation constellation. It should be noted that the vector of complex constellation symbols is formed in the frequency domain.

As the number of subcarriers N increases, for a fixed bandwidth the

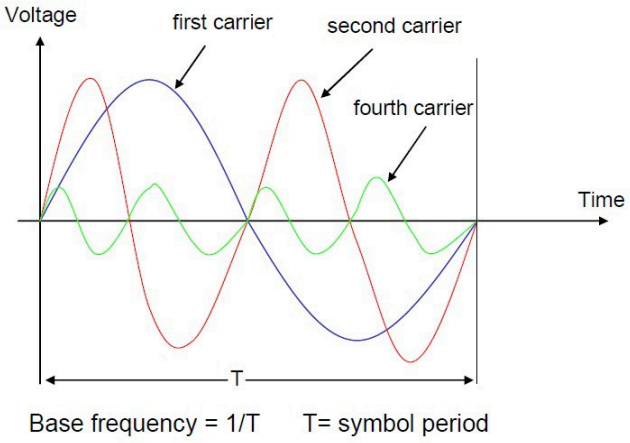

Figure 2.3. Individual sub-carriers in the time-domain post IFFT

subcarrier spacing reduces and the OFDM symbol period increases. This increased OFDM symbol period (which is N times the equivalent of single carrier system) makes the system robust to delay spread. To maintain

orthogonality in an OFDM system the following relationship is required

$$\frac{1}{T_s} = \triangle f, \tag{2.2.2}$$

where $\triangle f$ represents the sub-carrier spacing and T_s represents the unextended OFDM symbol period. The term unextended here is used for clarity since later in the report a guard interval is added to each OFDM symbol to extend its symbol period (it is called the extended symbol period). If an N-point IFFT is used then the total bandwidth $W = N \triangle f$.

Coded OFDM (COFDM) is formed by combining OFDM and forward error correction (FEC) code in order to further exploit the frequency and time diversity of the system. It was first employed in the Digital Audio Broadcasting (DAB) and Digital Video Broadcasting Terrestrial (DVB-T) standards [23]. Nowadays, 802.11a employs COFDM as its core transmission technology.

2.2.2 Guard Interval

A guard interval, also known as a cyclic extension or cyclic prefix, is a simple technique that protects the OFDM time-domain signal from loss of orthogonality due to channel delay spread. The length of the required guard interval must be sufficient to cover all of the significant delay taps. In practice the Guard Interval is often set at around twice the delay spread to provide an additional margin for symbol timing error in the receiver. The extended OFDM symbol is shown in Fig. 2.4

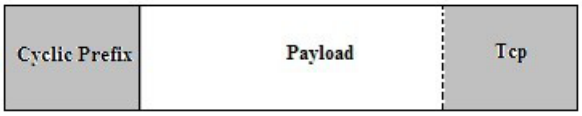

Figure 2.4. Extended OFDM symbol

2.2.3 Mathematical Transformation

Mathematically, the guard interval is used to convert the linear convolution of the channel into a perceived cyclic convolution. This is why the process is known as cyclic extension. After removal of the guard interval, the resulting convolution as seen at the receiver appears cyclic. The linear channel convolution process is shown in the two equations below (Fig. 2.5). In the left-hand equation an unextended symbol period is used. It is assumed $N = 5$ for the simplicity of the illustration. In the right-hand equation the OFDM symbol has been extended at the transmitter. The last two samples have been copied and appended at the start of the symbol (the appended section shown in green). Clearly the actual channel convolution process remains linear. However, after removing the cyclic prefix at the receiver, the convolution matrix can be rewritten as shown in the lower part of the equation. The presence of the cyclic prefix results in an upper triangle of new entries in the channel matrix. Clearly, each row in this matrix can be generated from a right cyclic shift of the row above. This form of convolution is perceived as cyclic and is denoted by the channel matrix H_c. It is also called circulant matrix

Mathematically, it is known that a circulant matrix can be diagonalised

$$y = H_L x = \begin{bmatrix} h_0 & & & & \\ h_1 & h_0 & & & \\ h_2 & h_1 & h_0 & & \\ & h_2 & h_1 & h_0 & \\ & & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Linear Convolution (unextended)

$$y = H_L x = \begin{bmatrix} h_0 & & & & & & \\ h_1 & h_0 & & & & & \\ h_2 & h_1 & h_0 & & & & \\ & h_2 & h_1 & h_0 & & & \\ & & h_2 & h_1 & h_0 & & \\ & & & h_2 & h_1 & h_0 & \\ & & & & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_4 \\ x_5 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Linear Convolution (extended)

$$y = H_C x = \begin{bmatrix} h_0 & & & h_2 & h_1 \\ h_1 & h_0 & & & h_2 \\ h_2 & h_1 & h_0 & & \\ & h_2 & h_1 & h_0 & \\ & & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Figure 2.5. Mathematical representation of the channel convolution

by pre and post multiplying it with unitary matrices that characterize the inverse discrete Fourier transform (IDFT) and discrete Fourier transform (DFT). This is described as follows:

$$\triangle = U^{-1} H_c U = F^{-1} H_c F, \tag{2.2.3}$$

Addition and removal of a cyclic prefix renders the resulting channel matrix circulant. The IFFT(F^{-1}) and FFT(F) processes represent a unitary transform pair. Hence, the resulting frequency domain matrix is diagonalized, i.e., none of the subcarriers interfere with each other. The above explains why the process is known as a cyclic extension. The received signal could be expressed in frequency domain using multiplications instead of convolution process in time domain, as shown below:

$$R(f) = H_c(f,t)X(f) + N(f), \tag{2.2.4}$$

2.2.4 Overview of OFDMA

Earlier OFDM is used as a modulation scheme for wireless systems, where all the subchannels of OFDM are assigned to a single user at any given time (i.e., IEEE 802.11a/g). Later TDMA and FDMA has been used with OFDM in order to support multiple users. As mentioned earlier, this kind of static resource allocation cannot provide a good performance. The drawback of this static resource allocation is that multi user diversity is not exploited (i.e., the fact that the different users has different channel channel gain is not used). OFDMA has been developed to exploit the multi user diversity where multiple users are allowed to transmit simultaneously on the different subchannels per OFDM symbol. Since the probability that all users experience worst channel gain in a particular subchannel is typically quite low. Hence, adaptive resource allocation algorithms can be developed in order to

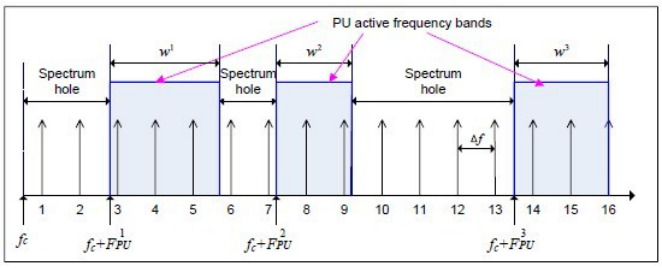

Figure 2.6. Primary users' active frequency bands, spectrum holes and cognitive radio network OFDM subchannels.

assure that subchannel is assigned to a user who has higher channel gain on that subchannel compared to other users.

2.2.5 Related Works on Resource Allocation for OFDMA Based Wireless Network

Adaptive resource allocation techniques allocate radio resources to various users according to users channel gain and QoS requirements. The problem of assigning the subchannels and transmission power to different users in an OFDMA system has been intensively studied over the past decade (i.e., [33–38] and references therein). All of these studies can be divided into two categories namely Margin Adaptive (MA) and Rate Adaptive (RA) resource allocation problems [33–35]. The objective of the MA problem is to minimize the total transmission power subject to users' individual data rate constraint, Bit Error Rate (BER) requirements while the objective of the RA is to maximize system data throughput subject to a total transmission power constraint. However, the conventional resource allocation techniques cannot trivially be extended to cognitive radio networks due to the additional interference constraints imposed by primary users. Hence, in literature, various resource allocation techniques have been used to allocate radio resources to secondary users while maintaining the interference leakage to the primary users is below the threshold. In this section, various algorithms proposed for cognitive radio networks based on OFDMA technique and/or multiple

antennas technique are reviewed.

The OFDMA has been recognized as one of the best candidates for the resource allocation in cognitive radio networks because of its natural ability to utilize different portions of the spectrum [15]. In [39–42], resource allocation techniques have been studied for an OFDMA based cognitive radio networks. In [39], power is allocated to each subchannel, by considering the received interference as a fairness metric. This problem has been solved using Lagrangian dual function. On the other hand, adaptive power loading has been investigated for OFDM-based cognitive radio network in [40]. This scheme is developed based on a non-integer Lagrangian formulation. However, power allocation problem together with subchannel allocation problem will not provide a closed-form solution [33]. Hence, joint subchannel, transmission power and data bit allocation problem for OFDMA based cognitive radio network is studied in [43]. Due to high computational complexity of determining optimal solution for OFDMA based resource allocation problems, low complexity algorithms have been proposed in [41, 42]. The work in [41, 42] consider a single user cognitive radio network and maximizes the cognitive radio network throughput by allocating data bits to various subchannels.

2.3 Beamforming Techniques

Beamforming is a general signal processing technique used in the physical layer of a communication channel to control the directionality of the transmission or the reception of a signal using an array of antennas [44]. Beamforming takes advantage of interference to change the directionality of the array. When transmitting, a beamformer controls the amplitude and the relative phase of the signal at each transmitter, in order to create a constructive radiation pattern and destructive interference in the wavefront.

When receiving, the signals from different antennas are combined in such a way that the expected pattern of radiation is preferentially observed. In the following subsection, receiver beamforming techniques and transmitter beamforming techniques are introduced.

2.3.1 Receiver Beamforming Techniques

In the receiver beamforming design, the objective is to estimate the desired signal in the presence of noise and interference. A narrowband beamformer is depicted in Figure 2.7. The beamformer output can be written as

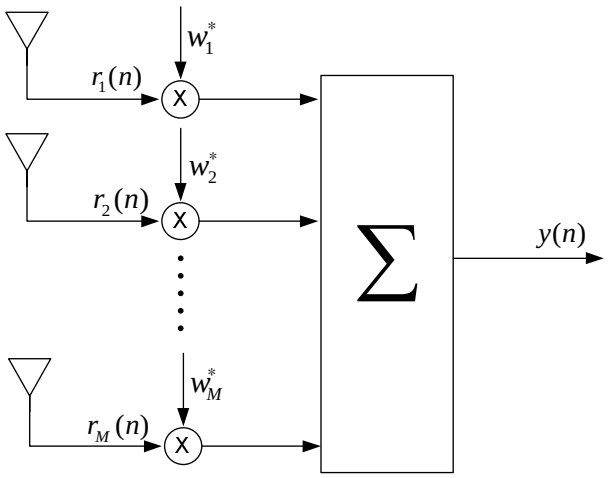

Figure 2.7. A narrowband beamformer

$$y(n) = \mathbf{w}^H \mathbf{r}(n), \tag{2.3.1}$$

where n is the time index, $\mathbf{r}(n) = [r_1(n)\ r_2(n)\ \cdots\ r_M(n)]^T$ is the $M \times 1$ vector of array observations and $\mathbf{w} = [w_1\ w_2\ \cdots\ w_k]^T$ is the complex beamforming weight vector. The array observation vector is given by

$$\mathbf{r}(n) = \mathbf{d}(n) + \mathbf{i}(n) + \boldsymbol{\eta}(n), \tag{2.3.2}$$

where $\mathbf{d}(n)$, $\mathbf{i}(n)$ and $\boldsymbol{\eta}(n)$ are the desired signal, interference and receiver noise respectively. If the desired signal is a far field point source and has a

time-invariant wavefront, then it can be written as

$$\mathbf{d}(n) = s(n)\mathbf{a}_s(\theta), \tag{2.3.3}$$

where $s(n)$ is the source signal and $\mathbf{a}_s(\theta)$ is an $M \times 1$ steering vector in the direction of θ or the channel response between the transmitter and the array of antennas.

The beamforming design can be classified into two categories, namely, data independent beamformer, and statistically optimum beamformer [44]. In the data independent beamformer design, the beamforming weight vector is obtained independent of array observations to present a specific response for all signal and interference scenarios. The beamforming weight vector in statistically optimum design is chosen based on the statistics of the array observations to optimize the array response.

2.3.2 Transmitter Beamforming Techniques

Beamforming at the transmitter (downlink beamformer) is substantially different in several aspects than that at the receiver. In the latter, the design will only determine the performance of a specific user whereas the transmit beamformer will affect not only the desired user but also all the users in the coverage area. Hence the downlink beamforming design should ideally take into consideration the system level performance, i.e. all the users in the reception area rather than a specific user. Another fundamental difference is the channel knowledge. For the receiver beamformer design, the receiver could estimate the channel coefficients using the training signal. For transmitter beamformer design, the channel knowledge could be made available to the transmitter by sending the estimates of the channel state information (CSI) from the receiver through a finite rate feedback channel [45–48] .

The focus of this subsection is on multiuser downlink beamformers. The

transmitter beamformers can be designed to satisfy quality of service requirements for each user. The target SINR for each user is considered as QoS requirement. Consider a wireless network basestation equipped with N_t transmit antennas serving K users. Each user is equipped with a single antenna. The signal transmitted by the basestation is given by

$$\mathbf{x}(n) = \tilde{\mathbf{W}}\mathbf{s}(n), \tag{2.3.4}$$

where $s(n) = [s_1(n) \cdots s_K(n)]^T$, $s_k(n) \quad (k = 1, 2, \cdots, K)$ is the symbol intended for the k^{th} user, $\tilde{\mathbf{W}} = [\mathbf{w}_1 \ \cdots \ \mathbf{w}_K]$ and $w_k \in \mathbb{C}^{N_t \times 1}$ is the transmit beamforming weight vector for the k^{th} user. The received signal at the k^{th} user can be written as

$$y_k(n) = \mathbf{h}_k^H \mathbf{x}(n) + \eta_k(n), \tag{2.3.5}$$

where $\mathbf{h}_k$ is the channel coefficient vector between the basestation and the k^{th} user and $\eta_k(n)$ is the receiver noise. The downlink beamforming problem based on the SINR requirement can be formulated as the minimization of the transmitted power at the basestation subject to each user SINR being greater than a target γ_i [49, 50].

$$\begin{aligned} \min_{\mathbf{w}_i} \quad & \sum_{i=1}^{K} \|\mathbf{w}_i\|_2^2 \\ \text{subject to} \quad & \frac{\|\mathbf{w}_i^H \mathbf{h}_i\|_2^2}{\sum_{k=1, k\neq i}^{K} \|\mathbf{w}_k^H \mathbf{h}_i\|_2^2 + \sigma_i^2} \geq \gamma_i \quad i = 1, \cdots, K, \end{aligned} \tag{2.3.6}$$

where σ_i^2 is the noise variance at the i^{th} receiver. The problem in (2.3.6) is a non-convex problem. Nevertheless, this problem can be converted into an semidefinite programming (SDP) with Lagrangian relaxation and it can be solved efficiently using convex optimization toolboxes [51–53]. However, it is quite difficult to predict in advance whether the problem in (2.3.6) with

a given set of target SINRs and total transmit power at the basestation is feasible.

To overcome this infeasibility issue, this problem can be formulated into a more attractive framework based on a max-min fairness approach where the worst-case user SINR is maximized while using the available total transmit power [54]. This is known as the SINR balancing technique and it can be stated as [54–57]

$$\max_{\mathbf{U},\mathbf{p}} \quad \min_{1\leq i\leq K} \quad \frac{\mathrm{SINR}_i(\mathbf{U},\mathbf{p})}{\gamma_i}$$
$$\text{subject to} \quad \mathbf{1}^T\mathbf{p} \leq \mathrm{P}_{\max}, \tag{2.3.7}$$

where $\mathbf{U} = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_K]$, $\|\mathbf{u}_k\|_2 = 1$, and $\mathbf{p} = [p_1 \ \cdots \ p_K]^T$. Here $\mathbf{u}_k \in \mathbb{C}^{N_t\times 1}$ and p_k are the transmit beamforming weight vector and the corresponding power allocation for the k^{th} user respectively. In [54], an iterative algorithm has been proposed using the uplink-downlink duality, where the solution balances the ratio between the achieved SINR and the target SINR for all users while using all the transmit power available at the basestation. These techniques have been extended to cognitive radio network in [58].

2.4 Spatial Multiplexing Techniques

The technique using multiple antennas at both the receiver and the transmitter is known as Multiple-Input-Multiple-Output (MIMO) array processing. Compare to the use of only multiple receive antennas or multiple transmit antennas [59–61], the MIMO array processing can further improves the system performance by adding additional diversity against fading. Spatial multiplexing can be obtained by decomposing the MIMO channel matrix into various independent spatial subchannels that are used to transmit different data streams independently. This has the potential to increase the data rate up to a factor that is the same as the rank of the MIMO matrix as compared to the single antenna system [62,63]. Consider a point-to-point MIMO chan-

nel with M_t transmit antennas and M_r receive antennas as shown in Fig. 2.8. The received signal is given by

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \boldsymbol{\eta}(n), \tag{2.4.1}$$

where $\mathbf{y} = [y_1(n) \ \cdots \ y_{M_r}(n)]^T$ and $y_r(n)$ is the received signal at the r^{th} receiver antenna. $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ and h_{ij} is the channel gain between the i^{th} transmitter antenna and j^{th} receiver antenna. $\mathbf{x}(n) \in \mathbb{C}^{M_t \times 1}$ and $\boldsymbol{\eta}(n) \in \mathbb{C}^{M_r \times 1}$ are the transmitted symbol vector and the noise vector at the receiver end respectively. It is assumed that the channel gain matrix $\mathbf{H}$ is known to

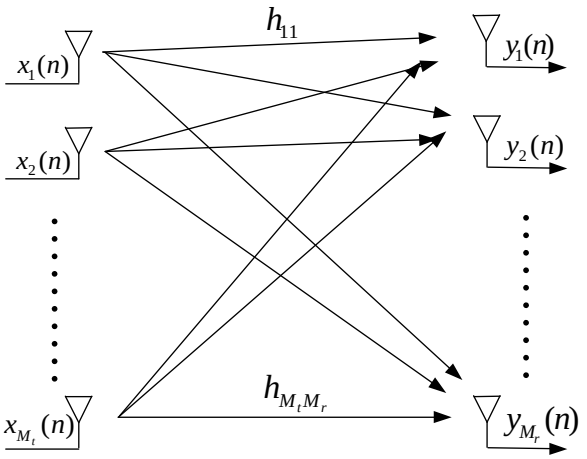

Figure 2.8. A MIMO system with M_t transmit antennas and M_r receive antennas.

both the transmitter and the receiver. The MIMO channel matrix $\mathbf{H}$ can be decomposed using the Singular Value Decomposition (SVD) as [64]

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H, \tag{2.4.2}$$

where $\mathbf{U} \in \mathbb{C}^{M_r \times M_r}$ and $\mathbf{V} \in \mathbb{C}^{M_t \times M_t}$ are unitary left and right singular matrices of $\mathbf{H}$. $\boldsymbol{\Sigma} \in \mathbb{R}^{M_r \times M_t}$ is a diagonal matrix with the singular values (υ_i) of $\mathbf{H}$. R_H number of singular values are nonzero, so that R_H is the rank of matrix $\mathbf{H}$. The singular value satisfies the property $\upsilon_i = \sqrt{\lambda_i}$, where λ_i is the i^{th} eigenvalue of $\mathbf{H}\mathbf{H}^H$. These MIMO spatial subchannels are obtained

using linear transformation of the input signal and the output signal through transmit precoding $\mathbf{V}$ and receiver shaping $\mathbf{U}^H$. In transmit precoding, the modulated symbol stream is precoded as

$$\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}, \tag{2.4.3}$$

where $\tilde{\mathbf{x}}$ is the modulated symbol stream. Similarly, the received signal is shaped as

$$\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y} \tag{2.4.4}$$

Such transmit precoding and receiver shaping decompose the MIMO channel into R_H number of independent single-input single-output (SISO) channels as follows:

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{U}^H(\mathbf{H}\mathbf{x} + \boldsymbol{\eta}) \\ &= \mathbf{U}^H\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathbf{H}}\mathbf{V}\tilde{\mathbf{x}} + \mathbf{U}^H\boldsymbol{\eta} \\ &= \boldsymbol{\Sigma}\tilde{\mathbf{x}} + \tilde{\boldsymbol{\eta}} \end{aligned} \tag{2.4.5}$$

where $\tilde{\boldsymbol{\eta}} = \mathbf{U}^H\boldsymbol{\eta}$. The resulting parallel spatial subchannels are shown in Fig. 2.9. They are independent from each other in the sense that signals through each spatial subchannels do not interfere with each other. Hence

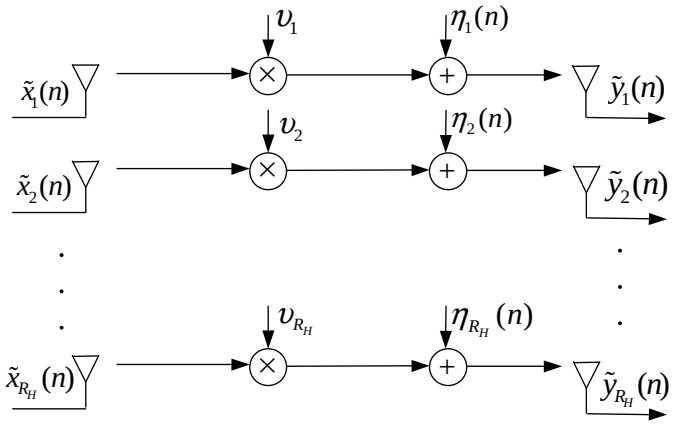

Figure 2.9. Parallel decomposition of the MIMO channel.

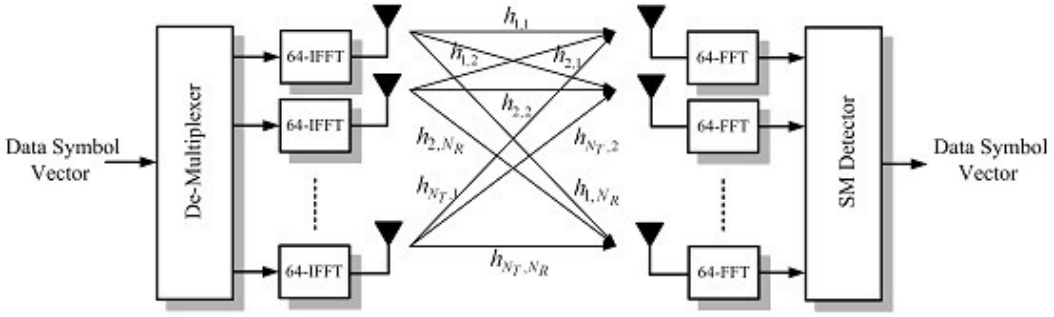

Figure 2.10. Spatial multiplexing architecture.

this MIMO channel can support up to R_H times the data rate of a SISO channel.

2.4.1 System Architecture for MIMO OFDM Based Spatial Multiplexing

The MIMO can also be combined with OFDM for better performance at lower complexity because a multipath channel in the time domain can be transformed into flat channel in each subcarrier of the OFDM. For a MIMO system with N_T transmit antennas and N_R receive antennas, the channel equation can be expressed as $N_R \times N_T$ matrix for each OFDM subcarrier. Each element is denoted by $h_{i,j,k}$, where i denotes the i^{th} transmitting antenna, j denotes the j^{th} receiving antenna and k denotes the k^{th} subcarrier. For simplicity, the subscript of the k^{th} subcarrier is ignored. Therefore, the MIMO channel matrix shown in Fig. 2.10 can be expressed as:

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{2,1} & \dots & h_{N_T,1} \\ h_{1,2} & h_{2,2} & \dots & h_{N_T,2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1,N_R} & h_{2,N_R} & \dots & h_{N_T,N_R} \end{bmatrix} \tag{2.4.6}$$

The transmit signal vector from N_T antennas is shown below:

$$x = [x_1 x_2 \cdots x_{N_T}] \tag{2.4.7}$$

Thus the received signal vector can be expressed using matrix multiplication in the frequency domain, as shown below:

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{N_R} \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{2,1} & \dots & h_{N_T,1} \\ h_{1,2} & h_{2,2} & \dots & h_{N_T,2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1,N_R} & h_{2,N_R} & \dots & h_{N_T,N_R} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_T} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{N_T} \end{bmatrix} \tag{2.4.8}$$

The MIMO technique explained earlier can be applied to each subcarrier.

2.5 Capacity of Wireless Networks

The capacity limits dictate the maximum data rates that can be transmitted over wireless channels with asymptotically small error probability, assuming no constraints on delay or complexity of the encoder and the decoder. In this section the capacity of a single-user wireless channel is first studied where the transmitter and/or receiver have a single antenna. Capacity of a single-user system where the transmitter and the receiver have multiple antennas as well as the capacity of a multiuser system are also treated subsequently.

2.5.1 Capacity of SISO AWGN Channel

Consider a discrete-time additive white Gaussian noise (AWGN) channel with channel input/output relationship $y[i] = x[i] + n[i]$, where $x[i]$ is the channel input at time i, $y[i]$ is the corresponding channel output, and $n[i]$ is a white Gaussian noise random process. Assume a channel bandwidth B and transmit power P. The channel SNR, the power in $x[i]$ divided by the power in $n[i]$, is given by $\gamma = P/(N_0 B)$, where N_0 is the noise power spectral

density. The capacity of this channel is given by the Shannon's well-known formula [65]:

$$C = B\log_2(1+\gamma) \ \text{ bit/s}, \tag{2.5.1}$$

Shannon capacity formula can be extended to fading channels with the knowledge of the channel at the receiver for an average power constraint P as follows [66]

$$C = \int_0^\infty B\log_2(1+\gamma)p(\gamma)d(\gamma), \tag{2.5.2}$$

where $p(\gamma)$ is the probability density function of γ. Note that this formula is a probabilistic average, i.e. Shannon capacity is equal to Shannon capacity for an AWGN channel with SNR γ, given by $B\log_2(1+\gamma)$, averaged over the distribution of γ. In this case, the Shannon capacity is called ergodic capacity.

2.5.2 MIMO Channel Capacity

This section focuses on the Shannon capacity for a MIMO channel. The capacity of a MIMO channel is an extension of the mutual information formula for a SISO channel to a matrix channel. Specifically, the capacity is given in terms of the mutual information between the channel input vector $\mathbf{x}$ and the output vector $\mathbf{y}$ as

$$C = \max_{p(x)} I(\mathbf{X};\mathbf{Y}) = \max_{p(x)}[H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X})], \tag{2.5.3}$$

where $H(\mathbf{Y})$ and $H(\mathbf{Y})|\mathbf{X}]$ are the entropy in $\mathbf{y}$ and $\mathbf{y}|\mathbf{x}$. The definition of entropy yields that $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{N})$, the entropy of the noise. Since this noise $\mathbf{n}$ has fixed entropy independent of the channel input, maximizing mutual information is equivalent to maximizing the entropy in $\mathbf{y}$.

The mutual information of $\mathbf{y}$ depends on its covariance matrix, which

for the narrowband MIMO model is given by

$$\mathbf{R_y} = E[\mathbf{y}\mathbf{y}^H] = \mathbf{H}\mathbf{R_x}\mathbf{H}^H + \mathbf{I_{M_r}}, \tag{2.5.4}$$

where $\mathbf{R_x}$ is the covariance of the MIMO channel input and it is the MIMO channel matrix. It turns out that for all random vectors with a given covariance matrix $\mathbf{R_y}$, the entropy of $\mathbf{y}$ is maximized when $\mathbf{y}$ is a zero-mean circularly-symmetric complex Gaussian (ZMCSCG) random vector [67]. But $\mathbf{y}$ is only ZMCSCG if the input $\mathbf{x}$ is ZMCSCG, and therefore this is the optimal distribution on $\mathbf{x}$. This yields $H(\mathbf{y}) = B\log_2\det[\pi e\mathbf{R_y}]$ and $H(\mathbf{n}) = B\log_2\det[\pi e\mathbf{I_{M_r}}]$, resulting in the mutual information [65]

$$\mathbf{I(X;Y)} = B\log_2\det[\mathbf{I_{M_r}} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H], \tag{2.5.5}$$

The MIMO capacity is achieved by maximizing the mutual information (2.5.5) over all input covariance matrices $\mathbf{R_x}$ satisfying the power constraint:

$$C = \max_{\mathbf{R_x}:\mathbf{Tr}(\mathbf{R_x})=\rho} B\log_2\det[\mathbf{I_{M_r}} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H], \tag{2.5.6}$$

where $\det[\mathbf{A}]$ denotes the determinant of the matrix $\mathbf{A}$. Clearly the optimization relative to $\mathbf{R_x}$ will depend on whether or not $\mathbf{H}$ is known at the transmitter.

2.5.3 Duality of Gaussian Multiple-Access and Broadcast Channels

This section focuses on the duality of Gaussian multiple-access and broadcast channels. In telecommunications and computer networks, a multiple access method allows several terminals connected to the same multi-point transmission medium to transmit and receive signals. In a multiple access channel (MAC) , multiple senders transmit independent messages over a

shared physical medium to one or several destination nodes. This requires a channel access scheme, including a media access control protocol combined with a multiplexing scheme. This channel model has applications in the uplink of the cellular networks. The broadcast channel (BC) is a point to multipoint, unidirectional (downlink) channel used in the cellular networks. The Gaussian multiple access channel (MAC) and the Gaussian BC have two fundamental differences. In the MAC, each transmitter has an individual power constraint, whereas in the BC there is only a single power constraint on the transmitter. In addition, signal and interference come from different transmitters in the MAC and are therefore multiplied by different channel gains (known as the nearfar effect) before being received. In the BC, the entire received signal comes from the same source and therefore has the same channel gain. It has been shown in [1] that the scalar Gaussian MAC and BC are duals of each other and as a result the BC and the MAC share the same capacity region.

Capacity Region of the MAC

From [68], the capacity region of a Gaussian MAC with channel gains $\mathbf{h} = (h_1, \cdots, h_K)$ and power constraints $\bar{\mathbf{P}} = (\bar{P}_1, \cdots, \bar{P}_K)$ is given by

$$C_{MAC}(\bar{\mathbf{P}}; \mathbf{h}) = \{\mathbf{R} : \sum_{j \in S} R_j \leq \frac{1}{2} \log(1 + \frac{1}{\sigma^2} \sum_{j \in S} h_j \bar{P}_j), \tag{2.5.7}$$

The capacity region of the constant MAC is a K-dimensional polyhedron, and successive decoding with interference cancelation can achieve all corner points of the capacity region [68]. As an example, the capacity region of a MAC with two users is shown in Fig. 2.11

Capacity Region of the BC

The capacity region of a Gaussian BC with channel gains $\mathbf{h} = (h_1, \cdots, h_K)$ and power constraints $\bar{\mathbf{P}}$ over all power allocations such that $\sum_{j=1}^{K} P_j^B = \bar{P}$ is given by [68]

$$C_{BC}(\bar{P}; \mathbf{h}) = \{\mathbf{R} : R_j \leq \frac{1}{2}(1 + \frac{h_j P_j^B}{\sigma^2 + h_j \sum_{k=1}^{K} P_k^B \mathbf{1}[h_k > h_j]})j = 1, \cdots, K\}, \tag{2.5.8}$$

Any rate vector taking the form of (2.5.8) with equality lies on the boundary of the capacity region.

Authors in [1] have proven that the capacity region of a constant Gaussian BC with power constraint $\bar{P}$ is equal to the union of capacity regions of the dual MAC with power constraints $(P_1, \cdots, P_K)$ such that $\sum_{j=1}^{K} P_j = \bar{P}$, as shown in (2.5.9) and Fig. 2.11.

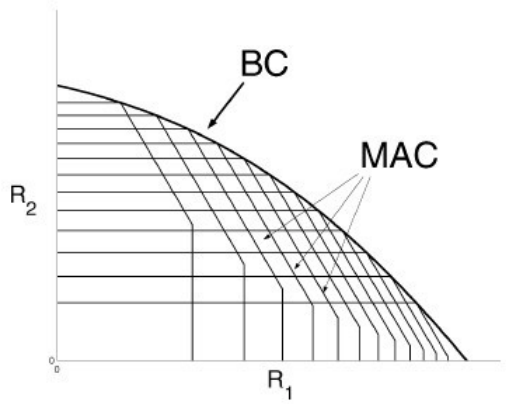

Figure 2.11. Constant BC capacity in terms of the dual MAC [1]. In this example, two users are considered in the system, R_1 and R_2 are the capacity for user 1 and user 2 respectively.

$$C_{BC}(\bar{P}; \mathbf{h}) = \bigcup_{\{\mathbf{P}:\mathbf{1}\cdot\mathbf{P}=\bar{P}\}} C_{MAC}(\mathbf{P}; \mathbf{h}), \tag{2.5.9}$$

Furthermore, the authors in [1] have proven that the capacity region of a constant Gaussian MAC is equal to the intersection of the capacity regions

of the scaled dual BC over all possible channel scalings for any vector of constants $\alpha > 0$, as shown in (2.5.10) and Fig. 2.12.

$$C_{MAC}(\mathbf{P};\mathbf{h}) = \bigcap_{\alpha>0} C_{BC}(\mathbf{1}\cdot\frac{\mathbf{P}}{\alpha};\alpha\mathbf{h}), \tag{2.5.10}$$

Using the above results, this duality has been extended to multiuser MIMO Gaussian BC channels in [69]. As it is difficult to directly solve the capacity problem in the downlink, this BC-MAC duality technique is very useful for solving optimization problems. Under a single sum power constraint, several efficient algorithms have been developed to obtain the

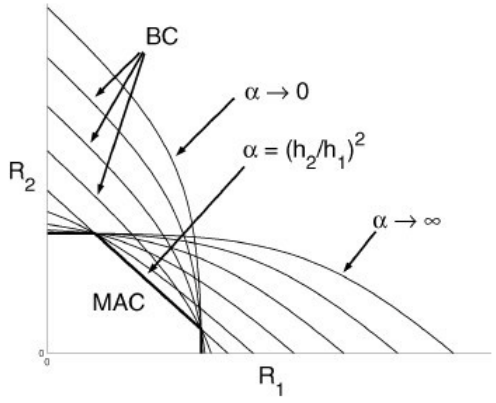

Figure 2.12. Constant MAC capacity in terms of the dual BC [1]. In this example, two users are considered in the system, R_1 and R_2 are the capacity for user 1 and user 2 respectively.

optimal solutions for the MIMO-BC sum rate problems. In [70], the authors used the principle of duality to transform the BC problem into a convex MAC problem and proposed an iterative algorithm for the optimum transmitter design. A numerical algorithm based on Lagrangian dual decomposition and iterative water-filling has been proposed in [71] for the computation of the sum capacity for the Gaussian vector BC. A weighted sum rate optimization problem was studied in [72] using iterative water-filling algorithm. A novel minimax BC-MAC duality was proposed in [73] for a Gaussian mutual information minimax optimization problem. It has been shown that the

uplink-downlink duality between BC and MAC channels can be obtained as a special case of this minimax duality. Based on this duality, an interior point algorithm was proposed in [74] to solve the MIMO-BC weighted sum rate maximization problem with per-antenna power constraints. It has been shown that the dual of the downlink problem is equivalent to the uplink problem with a different noise covariance matrix. The duality relationship also holds for the achievable rate regions of the respective uplink and downlink channels. Moreover, based on the duality, a weighted sum rate maximization algorithm for the cognitive radio MIMO BC has been proposed in [75].

2.6 Interference Alignment Techniques for Wireless Network

An interference alignment (IA) technique has been proposed recently in [27] as an efficient multiuser capacity achieving scheme in a high signal to noise ratio (SNR) regime. The fundamental concept of IA is to align all interference signals in a particular subspace at each receiver so that an interference-free orthogonal subspace can be solely allocated for data transmission. Since the work of [27], IA techniques have attracted significant interests and various algorithms have been proposed and analyzed, for example, multiple-input multiple-output (MIMO) interference network [28, 29], X network [30], and cellular network [31, 32].

The interference alignment problem is explained as follows. Consider the K-user MIMO interference channel where the k^{th} transmitter and receiver are equipped with $M^{[k]}$ and $N^{[k]}$ antennas respectively. Note that the antennas could represent symbol extensions in time or frequency as well. However, if the antennas correspond to symbol extensions over orthogonal dimensions (time, frequency slots) then the channel matrices will have a

diagonal structure. The channel is defined as:

$$\mathbf{Y}^{[k]}(n) = \sum_{l=1}^{K} \mathbf{H}^{[k,l]}(n)\mathbf{X}^{[l]}(n) + \mathbf{Z}^{[k]}(n) \tag{2.6.1}$$

where, at the n^{th} channel use, $\mathbf{Y}^{[k]}(n)$ and $\mathbf{Z}^{[k]}(n)$ are the $N^{[k]} \times 1$ received signal vector and the zero mean unit variance circularly symmetric additive white Gaussian noise vector (AWGN) at receiver k, $\mathbf{X}^{[k]}(n)$ is the $M^{[k]} \times 1$ signal vector transmitted by transmitter l, and $\mathbf{H}^{[k,l]}(n)$ is the $N^{[k]} \times M^{[k]}$ matrix of channel coefficients between transmitter l and receiver k. The transmit power at transmitter l is $E[||\mathbf{X}^{[l]}||^2] = P^{[l]}$. Let $d^{[k]} \leq \min M^{[k]}, N^{[k]}$, denote the degrees of freedom for user k's message. As in [29], the degrees of freedom is defined as the pre-log factor of the sum rate. This is one of the key metrics used for assessing the performance of a multiple antenna based system at high SNR regime, which is defined as

$$d \triangleq \lim_{SNR \to \infty} \frac{C_{\Sigma}(SNR)}{\log(SNR)}, \tag{2.6.2}$$

where $C_{\Sigma}(SNR)$ denotes the sum capacity that can be achieved for a given SNR.

Precoding at Transmitter:

let $\mathbf{V}^{[k]}$ be an $M^{[k]} \times d^{[k]}$ matrix whose columns are the orthonormal basis of the transmitted signal space of user K. Mathematically, the transmitted signal vector of user K is given by:

$$\mathbf{X}^{[k]} = \sum_{i=1}^{d^{[k]}} \mathbf{V}_d^{[k]} X_d^{[k]} = \mathbf{V}^{[k]}\bar{\mathbf{X}}^{[k]}, \bar{\mathbf{X}}^{[k]} \sim (0, \frac{P^{[k]}}{d^{[k]}}\mathbf{I}_{d^{[k]}}) \tag{2.6.3}$$

Each element of the $d^{[k]} \times 1$ vector $\bar{\mathbf{X}}^{[k]}$ represents an independently encoded Gaussian codebook symbol with power $\frac{P^{[k]}}{d^{[k]}}$ that is beamformed with the corresponding beamformer of $\mathbf{V}^{[k]}$.

Interference Suppression at Receiver:

Let $\mathbf{U}^{[k]}$ be an $N^{[k]} \times d^{[k]}$ matrix whose columns are the orthonormal basis of the interference-free desired signal subspace at receiver k. The k^{th} receiver filters its received signal to obtain:

$$\bar{\mathbf{Y}}^{[k]} = \mathbf{U}^{[k]H}\mathbf{Y}^{[k]} \tag{2.6.4}$$

If interference is aligned into the null space of $\mathbf{U}^{[k]}$ then the following condition must be satisfied:

$$\mathbf{U}^{[k]H}\mathbf{H}^{[k,l]}\mathbf{V}^{[l]} = 0, \ \forall l \neq k \tag{2.6.5}$$

$$\text{rank}\{\mathbf{U}^{[k]H}\mathbf{H}^{[k,k]}\mathbf{V}^{[k]}\} = d^{[k]} \tag{2.6.6}$$

In other words the desired signals are received through a $d^{[k]} \times d^{[k]}$ full rank channel matrix.

$$\bar{\mathbf{H}}^{[k,k]} \triangleq \mathbf{U}^{[k]H}\mathbf{H}^{[k,k]}\mathbf{V}^{[k]} \tag{2.6.7}$$

while the interference is completely eliminated. The effective channel for user k is then expressed as:

$$\bar{\mathbf{Y}}^{[k]} = \bar{\mathbf{H}}^{[k,k]}\bar{\mathbf{X}}^{[k]} + \bar{\mathbf{Z}}^{[k]} \tag{2.6.8}$$

where $\bar{\mathbf{Z}}^{[k]} \sim N(0, \mathbf{I})$ is the effective $d^{[k]} \times 1$ AWGN vector at receiver k. The rate achieved on this channel is:

$$R_k^b = \log|\mathbf{I}_{d^{[k]}} + \frac{P^{[k]}}{d^{[k]}}\bar{\mathbf{H}}^{[k,k]}\bar{\mathbf{H}}^{[k,k]H}| \tag{2.6.9}$$

$$= d^{[k]}\log(P^{[k]}) + o\log(P^{[k]}) \tag{2.6.10}$$

Thus, $d^{[k]}$ degrees of freedom are achieved by user k.

Related Works on Interference Alignment Techniques

The interference alignment has been proposed to achieve optimal degrees of freedom (DoF) in a single-input single-output (SISO) interference channel (IC) in [27]. It was shown that interference alignment can achieve the optimal DoF of $\frac{K}{2}$, in a K-user time varying interference channel. In [28], the authors provided examples of iterative algorithms that utilize the reciprocity of wireless networks to achieve interference alignment with only local channel knowledge at each node. The work in [76] proposed the alignment of multi-user interference at each receiver based on a carefully constructed signal structure, which was referred to as interference alignment in signal space. For the interference alignment in signal space, transmit precoding technique is used to align the multi-user interferences in the same interference space which is orthogonal to the desired signal space at each receiver. Furthermore, the authors in [29] provided an inner-bound and an outer-bound for the total number of degrees of freedom for the K user MIMO Gaussian time varyings interference channels with M antennas at each transmitter and N antennas at each receiver. For the case of K user $M \times N$ MIMO interference channel, authors in [29] showed that the total number of degrees of freedom is equal to $\min(M, N)K$ if $K \leq R$ and $\min(M, N)\frac{R}{R+1}K$ if $K > R$, where $R = \frac{\max(M,N)}{\min(M,N)}$. An interference alignment scheme was provided in [77] for a deterministic K-user interference channel.

Employing interference alignment techniques in cognitive radio networks, authors in [78] proposed a linear precoder for IA and power allocation which maximizes the individual data rate of the secondary link so that an opportunistic point-to-point MIMO link can reuse, without generating any additional interference, the same frequency band as that of the primary link. Furthermore, the authors in [79] extended this work to multiple cognitive secondary transmitters. In their proposed scheme, the precoding matrices of the SUs have been jointly designed such that no interference is introduced

at the PU receiver. In [80], a capacity achieving scheme has been proposed by transmiting along the spatial directions (SD) associated with the singular values of the channel matrix using a water-filling based power allocation (PA) technique.

There are several results that show significant improvement in SNR gains by employing relays within the context of interference cancelation [81–83]. The DoF is increased with the use of relays, which is particularly attractive when the source and the destination nodes are not directly connected. Authors in [84] characterized the DoF for a multihop two-user interference channel with layered relays. However, if the direct channel coefficients between the source and the destination nodes are non-zero, then from the perspective of DoF, it has been shown that the use of half-duplex relays does not contribute to any gain in DoF for the two-user interference channel [85], regardless of the number of relay nodes. Remarkably, this is true even if the relay nodes are equipped with multiple antennas. Since the conventional half-duplex relays are not contributing to increase the DoF of a two-user interference channel, another form of relaying introduced by El Gamal and Hassanpour was exploited in [86], known as instantaneous relaying (relay-without-delay). El Gamal and Hassanpour [86] have shown that the memoryless instantaneous relay channel can potentially achieve a higher capacity than a conventional relay. By employing this memoryless instantaneous relay channel, the authors in [87] showed that the two-user interference channel with an instantaneous relay can achieve $\frac{3}{2}$ DoF. An instantaneous relay increases the DoF for the two-user interference channel as compared to a network without relays.

The original contributions presented from Chapter 3 to Chapter 7 in this thesis have been built based on the MIMO, OFDMA and interference alignment techniques explained in this chapter.

Chapter 3

RESOURCE ALLOCATION TECHNIQUES FOR OFDMA BASED SPECTRUM SHARING NETWORKS

In this chapter, resource allocation techniques for OFDMA based cognitive radio systems are proposed. OFDMA is one of the best candidates for cognitive radio networks because of its natural ability to utilize different portions of the spectrum. Initially, an optimization technique for subcarrier and power allocation has been proposed for OFDMA-based cognitive radio network. This can be formulated within an integer linear programming (ILP) framework and solved using branch and bound method. The solution to this problem optimally allocates subcarriers, power and bits to secondary users while satisfying the data rate and BER requirements for each secondary user. It ensures that the interference leakage to the primary users is always less than a specific threshold. To consider a problem where the number of users seeking access to the network exceeds the available resources, user admission control techniques are introduced. A suboptimal optimization algorithm based on integer programming (IP) is proposed to admit as many secondary users as possible in the network while allocating subcarriers and

bits to each admitted user in such a way that the interference leakage to primary users is below a specific threshold.

3.1 Introduction

There are two classes of radio resource allocation problems, namely, the margin adaptive problem and the rate adaptive problem [37]. The objective of margin adaptive problem is to minimize the overall transmission power subject to data rate constraints, whereas the objective of rate adaptive problem is to maximize the overall data throughput subject to a transmission power constraint. The resource allocation algorithms in OFDM and OFDMA-based systems have been studied for many different scenarios in [33,37,39,40,88,89].

In [37] and [33], a joint subchannel, bit and power allocation problem has been studied for OFDMA and OFDM based conventional wireless systems (i.e., without interference constraint). Resource allocation problem in [37] was formulated into an integer linear programming framework and solved by branch and bound method. In [33], subchannels and bits are allocated to different users based on greedy algorithm. In [88] subchannels are allocated under the assumption of fixed power constraint for each user. In real scenarios, however, power and frequency allocations are intertwined. A suboptimal resource allocation algorithm is presented for MIMO-OFDM-based uplink system in [89]. The algorithm in [89] schedules the users based on their spatial separability in each subchannel and then allocates bits and power to each user. Grouping the users according to their spatial signature has been shown to reduce inter user interference [89]. Resource allocation algorithms developed for conventional wireless systems cannot be used directly to a cognitive radio network due to the additional interference constraint to primary users. The amount of interference introduced to primary users by a secondary network basestation depends on the power allocated to each subchannels and the corresponding subchannel gain between the secondary

network basestation and the primary users (interference subchannel gain).

Resource allocation problem for an OFDM-based cognitive radio network was studied in [39] and [40]. In [39], transmission power was allocated to each subchannel, by considering the received interference as a fairness metric. This problem has been solved using Lagrangian dual function. On the other hand, adaptive power loading has been investigated in [40]. This scheme was developed based on a non-integer Lagrangian formulation. However, disjoint power allocation and subchannel allocation in a resource allocation problem will not provide an optimal solution [33]. Therefore, in this chapter an integer linear programming approach is considered for joint subchannel, bit and power allocation in cognitive radio network, which is different from the work presented in [39] and [40].

3.2 Optimal Adaptive Bit Loading and Subcarrier Allocation Techniques for OFDMA Based Cognitive Radio Systems

In this section, the problem of adaptive radio resource allocation for an OFDMA-based CRN is studied. An adaptive bit loading optimization problem is formulated based on maximization of total throughput under interference power constraint to PUs, individual data rate constraints for SUs and total transmission power constraint at the SNB. This radio resource allocation algorithm are solved using integer linear programming. This algorithm optimally assigns subcarriers and bits for SUs while satisfying SUs' QoS requirements. The proposed technique considers interference leakage to and from multiple primary users and secondary users in order to optimize the throughput of the secondary network while keeping interference to primary network below a specific value. The simulation framework explicitly considers baseband filter response of the primary network and its impact on the interference leakage to the secondary network. Simulation results have been

provided later in this section to validate the performance of the algorithm.

3.2.1 System Model

In a system that both secondary user and primary user exist in side-by-side bands with different access technologies, mutual interference is introduced for both networks thus reducing the performance of both networks [90]. Therefore a sidelobe suppression mechanism has been proposed in [91] in order to minimize adjacent channel interference. More power can be loaded into far apart subcarriers for a given interference threshold by using this nulling mechanism to the subcarriers which exist adjacent to primary user's band. A network with K secondary users and L primary users is considered. The primary users are under the coverage region of the cognitive radio network. It is assumed that the basestation of the cognitive radio employs OFDMA scheme that both secondary user and primary user exist in side-by-side bands. The number of subcarriers available in the downlink is denoted by N. In spectrum domain, side-by-side cognitive radio radio model is considered as presented in [90]. In this model, multiple primary users 1, 2, ..., L, occupies multiple frequency bands of bandwidth B_1, B_2, ..., B_L in Hz. Secondary users sense the unoccupied band which is located on each side of the primary user bands for transmission. The bandwidth for cognitive radio transmission is divided into N subcarrier based OFDMA system while the bandwidth for each secondary user subcarrier is $\triangle f$.

In the downlink transmission scenario, there are three types of channels. The gain matrix between the secondary network secondary network basestation (SNB) and SUs is now defined as

$$\mathbf{H}_{\text{ss}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{K1} & \alpha_{K2} & \dots & \alpha_{KN} \end{bmatrix}, \mathbf{H}_{\text{ss}} \in R^{K \times N},$$

where α_{kn} represents the magnitude of the channel gain of the n^{th} subcarrier for k^{th} secondary user. The gain matrix between the primary network basestation (PNB) and the secondary users which denoted as $\mathbf{H}_{ps}$, and the gain matrix between secondary network basestation and primary users which denoted as $\mathbf{H}_{sp}$ is defined as follows

$$\mathbf{H}_{\text{ps}} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{L1} & \beta_{K2} & \dots & \beta_{LK} \end{bmatrix}, \mathbf{H}_{\text{ps}} \in R^{L\times K},$$

$$\mathbf{H}_{\text{sp}} = \begin{bmatrix} \zeta_1 & \zeta_2 & \dots & \zeta_L \end{bmatrix}, \mathbf{H}_{\text{sp}} \in R^{1\times L},$$

where β_{lk} represents the magnitude of the channel gain between the PNB and the k^{th} secondary user on l^{th} band. It is assumed that the subcarrriers are non-sensitive thus the channel gain on all subcarriers is the same for each secondary user. However, this work presented here can be extended to different fading channel without loss of generality. ζ_l represents the magnitude of the channel gain between secondary network basestation and l^{th} primary user.

As mentioned in [90], due to the coexistence of primary user and secondary user in side by side bands, there are two types of interference in this system. One is introduced by the primary users' bands into the secondary users' subcarriers, and the other is introduced by the secondary users' subcarriers into the primary users' bands. The same model is employed as in [90] in this section.

A. Interference introduced to primary user

Due to the side lobes of the secondary users' OFDM signal, the power spectrum density (PSD) of the k^{th} secondary user on n^{th} subcarrier can be

presented as [90]

$$\phi_i(f) = P_{k,n}T_s(\frac{\sin \pi f T_s}{\pi f T_s})^2, \tag{3.2.1}$$

where $P_{k,n}$ is the total transmit power for the k^{th} secondary user on the n^{th} subcarrier, T_s is the symbol duration. The interference introduced by the secondary user transmitter to the l^{th} primary user is written as below

$$I_{k,n}^{(l)}(d_{nl}, P_{k,n}) = |\zeta_l|^2 P_{k,n}T_s\delta(n,l), \tag{3.2.2}$$

where

$$\delta(n,l) = \int_{d_{nl}-B_l/2}^{d_{nl}+B_l/2} (\frac{\sin \pi f T_s}{\pi f T_s})^2 df, \tag{3.2.3}$$

and d_{nl} represents the distance in frequency between n^{th} subcarrier of the secondary user and the l^{th} band of primary user, B_l represents l^{th} primary users' bandwidth.

B. Interference introduced to secondary user

The reception of an OFDM symbol is performed using an N-ponits FFT function. This reveals that the signal is windowed in the time domain. In the frequency domain, it is a convolution process by the fourier transform of the received signal and the window function. Therefore, the power spectrum density after N-FFT processing can be expressed by the following expected value of the periodogram [90]:

$$E\{I_N(\omega)\} = \frac{1}{2\pi N}\int_{-\pi}^{\pi} \phi_{PU}(e^{j\omega})(\frac{\sin(\omega-\psi)N/2}{\sin(\omega-\psi)/2})^2 d\psi, \tag{3.2.4}$$

where ω represents the normalized angular. $\phi_{PU}(e^{j\omega})$ is the power spectrum density of the primary user's signal. An elliptically filtered white noise process with an amplitude P_{PU} is assumed for the primary user's signal [90]. As a result, the interference introduced by the l^{th} primary user's signal to the k^{th} secondary user on the n^{th} subcarrier will be the integration of the

power spectrum density of the primary user signal across the n^{th} subcarrier denoted as

$$J_{k,n}^{(l)}(d_{nl}, P_{PU}) = |\beta_{l,k}|^2 \int_{d_{nl}-\Delta f/2}^{d_{nl}+\Delta f/2} E\{I_N(\omega)\} dw. \tag{3.2.5}$$

3.2.2 Problem Statement

The general resource allocation algorithm for OFDMA-based cognitive radio network performs subcarrier and modulation scheme selection for each secondary user at the basestation. In the problem formulation, four types of modulation schemes BPSK, 4-QAM, 16-QAM and 64-QAM are considered. The minimum signal power required in the subcarrier to achieve a given BER at the receiver for BPSK modulation scheme is given by [92]

$$P_{\text{r, BPSK}}(\rho_k) = N_{\phi}^{k,n}[\text{erfc}^{-1}(2\rho_k)]^2, \tag{3.2.6}$$

where

$$\text{erfc}(x) = \frac{2}{\sqrt{\Pi}} \int_x^{\infty} e^{-t^2} dt, \tag{3.2.7}$$

The parameter ρ_k is an upper bound on the BER at the k^{th} secondary user's receiver. $N_{\phi}^{k,n}$ is the single-sided noise PSD for k^{th} secondary user on n^{th} subcarrier, which is shown below:

$$N_{\phi}^{k,n} = \frac{\sum_{l=1}^{L} J_{k,n}^{(l)}}{\Delta f} + N_0, \tag{3.2.8}$$

where $J_{k,n}^{(l)}$ denotes the interference introduced by the l^{th} primary user's band to the k^{th} secondary user on n^{th} subcarrier, Δf denotes the subcarrier spacing and N_0 is the noise PSD. Due to the central limit theorem, the sum of the interference can be assumed to be normal distributed.

For M-ary QAM, the required power for a given BER can be written

as [92]

$$P_{\mathrm{r,\ M\text{-}QAM}}(c,\rho_k,n)=\frac{2(M-1)N_{\phi}^{k,n}}{3}\left[\mathrm{erfc}^{-1}\left(\frac{\rho_k}{4}\right)\right]^2, \tag{3.2.9}$$

where parameter c denotes the number of bits per symbol, M is the modulation order given by $M=2^c$.

The required transmission power at the secondary network basestation (in energy per symbol) to achieve a certain received power at the k^{th} secondary user terminal over the n^{th} subcarrier for a given modulation scheme and BER target is given by the following equation

$$P_{k,n}=\frac{P_{\mathrm{r}}(c_{k,n},\rho_k)}{\alpha_{kn}^2}, \tag{3.2.10}$$

where $c_{k,n}$ denotes the number of bits allocated for the k^{th} secondary user over the n^{th} subcarrier.

From these definitions, the objective of the original RA problem can be stated as maximizing the total user data rate subject to individual data rate constraints to the each secondary users and interference leakage constraints to the primary users as follows:

$$\max_{c_{k,n}\in\{0,1,2,4,6\}} \quad \sum_{n=1}^{N}\sum_{k=1}^{K} c_{k,n}, \tag{3.2.11}$$

$$\text{s.t.} \quad \sum_{n=1}^{N}\sum_{k=1}^{K} |\zeta_l|^2 P_{k,n} T_s \delta(n,l) \le \gamma_l, \tag{3.2.12}$$

$$l=1,2,\ldots,L,$$

$$\sum_{n=1}^{N} c_{k,n} \ge r_k,\ k=1,2,\ldots,K, \tag{3.2.13}$$

$$\sum_{n=1}^{N}\sum_{k=1}^{K} P_{k,n} \le P, \tag{3.2.14}$$

$$c_{k,n}=0 : c_{k',n}\neq 0,$$

$$\forall k\neq k', \quad (k=1,2,\ldots,K). \tag{3.2.15}$$

The objective function in (3.2.11) represents the total throughput of the system, where $c_{k,n}$ denotes the number of bits allocated for the k^{th} secondary user over the n^{th} subcarrier. Due to the coexistence of the primary users and the secondary users in side by side bands, the first constraint ensures that the interference leakage to the l^{th} primary user is always below a certain threshold γ_l. The power allocated for the k^{th} secondary user in the n^{th} subcarrier is denoted by $P_{k,n}$ and P denotes the maximum total transmission power. Constraint (3.2.13) ensures that each secondary user achieves a predefined data rate. Due to the mutual exclusive secondary user allocation, it is possible to allocate more than one subcarrier for one secondary user as they can not be accompanied together in one subcarrier. Therefore the last constraint performs mutual exclusive secondary user allocation in each subcarrier. Allocating radio resources to secondary users without considering the QoS requirement in (3.2.13) may cause unfairness between secondary users. Because secondary users that are adjacent to the primary user's band have very high interference introduced by primary users, certain data rate should be guaranteed to these users using the constraint in (3.2.13). Due to the non-linear property of (3.2.9), RA problem introduced in (3.2.11)-(3.2.15) becomes non-linear in nature. Since the solution of this problem should be mutual exclusive secondary user allocation, this problem becomes a non-convex problem [93]. In the next subsection, integer linear programming formulation of the problem is proposed.

3.2.3 Integer Linear Programming Problem Formulation

Similar to the work presented in [37], the original optimization problem can be formulated into ILP framework. The original problem is a non-deterministic polynomial time (NP) hard problem. First, a binary vector $\mathbf{x}$

of size $NKC \times 1$ is defined as follows:

$$\mathbf{x} = [(\mathbf{x}_N^1)^T \ (\mathbf{x}_N^2)^T \ \dots \ (\mathbf{x}_N^N)^T]^T \ \in \ \{0,1\}^{NKC\times 1}, \tag{3.2.16}$$

where $\mathbf{x}_N^n = [\mathbf{x}_{1,n}^T \ \dots \ \mathbf{x}_{K,n}^T]^T \ \in \ \{0,1\}^{KC\times 1}$ and $\mathbf{x}_{k,n} = [x_{k,n,1} \ \dots \ x_{k,n,C}]^T \in \{0,1\}^{C\times 1}$. $x_{k,n,c}$ being equal to one represents that the n^{th} subcarrier is assigned to the k^{th} user which transmits c number of bits per symbol. In order to ensure no more than one user is allocated in each subcarrier only one of the element of $\mathbf{x}_N^n$ should be equal to one and rest of them should be zeros. Similarly, the transmission power vector $\mathbf{p}$ is defined as

$$\mathbf{p} = [(\mathbf{p}_N^1)^T \ (\mathbf{p}_N^2)^T \ \dots \ (\mathbf{p}_N^N)^T]^T \ \in \ \mathcal{C}^{NKC\times 1}, \tag{3.2.17}$$

where $\mathbf{p}_N^n = [\mathbf{p}_{1,n}^T \, \mathbf{p}_{2,n}^T \ \dots \ \mathbf{p}_{K,n}^T]^T \ \in \ \mathcal{C}^{KC\times 1}$ and $\mathbf{p}_{k,n} = [p_{k,n,1} \ \dots \ p_{k,n,C}]^T \in \mathcal{C}^{C\times 1}$, $p_{k,n,c}$ is the required transmit power for k^{th} secondary user over n^{th} subcarrier to transmit c number of bits [37]. The original problem can now be formulated into ILP as follows

$$\min_{x} \quad -\mathbf{d}^T\mathbf{x}, \tag{3.2.18}$$

$$\text{s.t.} \quad \mathbf{K}[\mathbf{A}_c(\mathbf{p}\odot\mathbf{x})] \le \boldsymbol{\xi}, \tag{3.2.19}$$

$$\mathbf{A}_u\mathbf{x} \ge \mathbf{R}, \tag{3.2.20}$$

$$\mathbf{p}^T\mathbf{x} \le P, \tag{3.2.21}$$

$$\mathbf{0}_N \le \mathbf{A}_c\mathbf{x} \le \mathbf{1}_N, \tag{3.2.22}$$

$$x_i \ \in \ \{0,1\}, \quad i = 1,2,\dots,NKC, \tag{3.2.23}$$

where $(.)^T$ denotes the transpose operation, $\odot$ denotes the entrywise product. $\boldsymbol{\xi} = [\xi_1 \ \xi_2 \ \dots \ \xi_L]^T \in C^{L\times 1}$, where $\xi_l = \frac{\gamma_l}{|\zeta_l|^2 T_s}$. $\mathbf{d} = [\mathbf{e}^T \ \dots \ \mathbf{e}^T]^T \ \in \ Z^{NKC\times 1}$, $\mathbf{e} = [1 \ 2 \ 4 \ 6]^T \ \in \ Z^{4\times 1}$, where $\mathbf{e}$ represents the number of bits required for four different modulation schemes considered in this section.

$$\mathbf{K} = \begin{bmatrix} \delta_{11} & \delta_{12} & \dots & \delta_{1N} \\ \delta_{21} & \delta_{22} & \dots & \delta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{L1} & \delta_{K2} & \dots & \delta_{LN} \end{bmatrix}, \mathbf{K} \in R^{L \times N},$$

$$\mathbf{A_c} = \begin{bmatrix} \mathbf{1}_{KC}^T & \mathbf{0}_{KC}^T & \dots & \mathbf{0}_{KC}^T \\ \mathbf{0}_{KC}^T & \mathbf{1}_{KC}^T & \dots & \mathbf{0}_{KC}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{KC}^T & \mathbf{0}_{KC}^T & \dots & \mathbf{1}_{KC}^T \end{bmatrix}, \mathbf{I} \in \{0,1\}^{N \times NKC},$$

$$\mathbf{1}_{KC} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \{1\}^{KC \times 1}, \quad \mathbf{0}_{KC} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \{0\}^{KC \times 1},$$

$$\mathbf{A_u} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_1 & \dots & \mathbf{e}_1 \\ \mathbf{e}_2 & \mathbf{e}_2 & \dots & \mathbf{e}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}_K & \mathbf{e}_K & \dots & \mathbf{e}_K \end{bmatrix}, \quad \left\{ \begin{array}{llll} \mathbf{e}_1 = [\mathbf{e}^T & \mathbf{0}_C^T & \dots & \mathbf{0}_C^T] \in Z^{1 \times KC}, \\ \mathbf{e}_2 = [\mathbf{0}_C^T & \mathbf{e}^T & \dots & \mathbf{0}_C^T] \in Z^{1 \times KC}, \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}_K = [\mathbf{0}_C^T & \mathbf{0}_C^T & \dots & \mathbf{e}^T] \in Z^{1 \times KC}. \end{array} \right\},$$

$\mathbf{A_u} \in Z^{K \times NKC}$ and $\mathbf{R} = [r_1\ r_2\ \dots\ r_K]^T \in \mathcal{C}^{K \times 1}$.

The above problem can be solved by using Branch and Bound method. This is a well known method for solving the class of integer linear programming and mixed integer programming [94].

3.2.4 Simulation Results

In order to validate the performance of the proposed bit loading and subcarrier allocation algorithms, the optimal bit loading and subcarrier allocation in a cognitive radio network is evaluated. The value of both L and K are assumed to be 2. There are 64 subchannels. The value of Ts is set to be

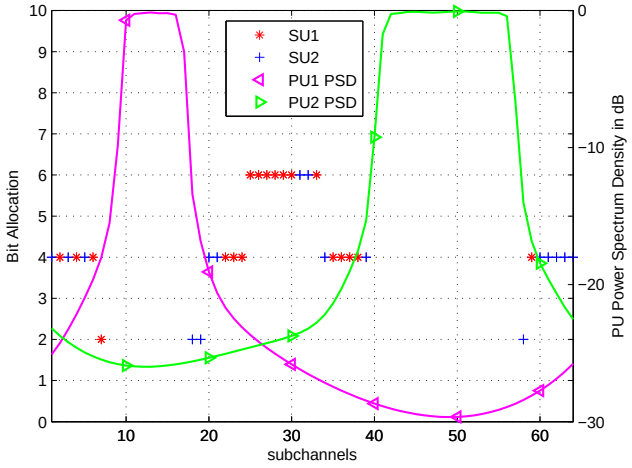

Figure 3.1. Bit loading and subcarrier allocation for secondary users as well as the power spectrum density after the N-FFT processing of primary users.

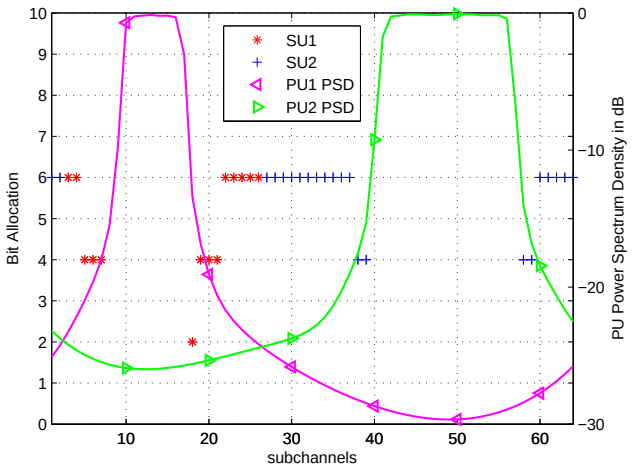

Figure 3.2. Bit loading and subcarrier allocation for secondary users as well as the power spectrum density after the N-FFT processing of primary users. Channel gain β_{lk} is different for secondary users.

1.2 second and f, B1 and B2 have been assigned the value of 1Hz, 8Hz and 16Hz, respectively. Therefore, secondary network basestation could not allocate subcarriers from 9 to 16 and from 41 to 56 to secondary user as they are occupied by primary users already. The value of amplitude P_{PU} is assumed to be 1W. The required BER and data rate for each secondary user have been set to 0.01 and 32 bits per OFDM symbol, respectively. The

background noise PSD is set to be 10^{-4}. The upper bound on the interference power for primary user 1 and primary user 2 are set to $2 * 10^{-4}$ and $4 * 10^{-4}$ respectively. The total power for cognitive radio network are set to 0.5W.

The channel gain α_{kn}, β_{lk} and ζ_l is assumed to be $\sqrt{5}$, 1 and 1 respectively. The number of bits allocated for each subcarrier is indicated using a star mark, "*" for secondary user 1 and using a plus mark, "+" for secondary user 2. PSD after the N-FFT processing is indicated using a left triangle mark, "$\triangleleft$" for primary user 1 and using a right triangle mark, "$\triangleright$" for primary user 2. Fig. 3.1 reveals that secondary network basestation allocates fewer bits into the subcarriers that are adjacent to the primary users, for example subcarrier 7, 18, 39, 58 and 59. Specifically, there are no bits

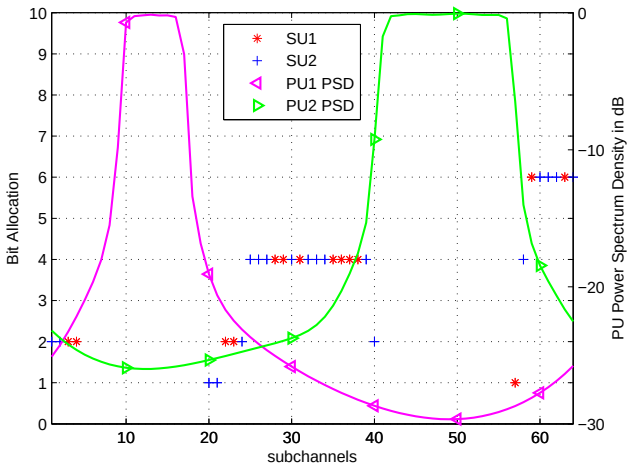

Figure 3.3. Bit loading and subcarrier allocation for secondary users as well as the power spectrum density after the N-FFT processing of primary users. Channel gain ζ_l is different for primary users

allocated to subcarrier 8, 17, 40 and 57 due to the high mutual interference. Furthermore, secondary network basestation allocates more bits (64-QAM) to the subcarriers from 25 to 33 due to the low interference as they exist far apart from primary users' band. The channel gain β_{11} and β_{22} are modified to 0.5. These settings reveal that interference leakage from primary user

1 to secondary user 1 is fewer than secondary user 2 while the interference leakage from primary user 2 to secondary user 2 is fewer than secondary user 1. Fig. 3.2 shows the subcarriers adjacent to primary user 1 are allocated to secondary user 1 while subcarriers adjacent to primary user 2 are allocated to secondary user 2. Finally, the channel gain ζ_1 and ζ_2 are modified to 2.5 and 0.1 respectively. These setting reveal the interference from secondary network basestation to primary user 1 is greater than primary user 2. Fig. 3.3 shows the subcarriers near primary user 1 are allocated fewer bits while subcarriers near primary user 2 are allocated more bits.

3.3 Suboptimal User Maximization Algorithm

In this section, admission control techniques at the physical layer for OFDM based cognitive radio systems are introduced. A suboptimal optimization algorithm based on IP is proposed to admit as many secondary users as possible in the network while allocating subcarriers and bits to each admitted user in such a way that the interference leakage to primary users is below a specific threshold. The case that the number of users seeking access to the network exceeds the available resources is considered. The aim is to optimally choose a subset of users from a larger set of users and to jointly allocate power and subcarriers to each admitted users to satisfy certain constraints. The proposed suboptimal algorithm performs very closely to the optimal algorithm while having a lower complexity.

3.3.1 Problem Formulation

A downlink underlay network with K secondary users, L primary users and N subcarriers is considered. Different from the previous section, both secondary user and primary user exist in the same band with different access technologies is employed in this section. The channel gain matrix between

the secondary network basestation and the secondary users is defined as $\mathbf{H}^{\mathrm{s}} \in R^{K\times N}$, where H^s_{kn} represents the magnitude of the channel gain of the n^{th} subcarrier for the k^{th} secondary user. The channel gain matrix between secondary network basestation and primary users is defined as $\mathbf{H}^{\mathrm{p}} \in R^{L\times N}$, where H^p_{kn} represents the magnitude of the channel gain between the secondary network basestation and the l^{th} primary user at the n^{th} subcarrier.

For simplicity, only four modulation schemes BPSK, 4-QAM, 16-QAM and 64-QAM is employed as in the previous section. The objective of the optimization problem is stated as maximizing the total number of users subject to individual data rate requirement for each admitted user and the interference leakage constraints to the primary users as follows:

$$\max_{\delta_{k,n}\in\{0,1\}} \sum_{k=1}^{K} (\delta_{k,1}|\delta_{k,2}\cdots|\delta_{k,N}), \tag{3.3.1}$$

$$\text{s.t.} \sum_{n=1}^{N}\sum_{k=1}^{K} (H^p_{kn})^2 \delta_{k,n} P_{k,n} \le \gamma_l,\ l = 1,2,\ldots,L, \tag{3.3.2}$$

$$\sum_{n=1}^{N} \delta_{k,n} c_{k,n} \ge r_k(\delta_{k,1}|\delta_{k,2}\cdots|\delta_{k,N}), \tag{3.3.3}$$

$$k = 1,2,\ldots,K,\quad c_{k,n} \in \{1,2,4,6\},$$

$$\sum_{n=1}^{N}\sum_{k=1}^{K} \delta_{k,n} P_{k,n} \le P_{max}, \tag{3.3.4}$$

$$\delta_{k,n} = 0 : \delta_{k',n} \ne 0, \forall k \ne k',\ k = 1,\ldots,K. \tag{3.3.5}$$

The notation | denotes *OR* operation, $\delta_{k,n} \in \{0,1\}$ denotes whether the n^{th} subcarrier is allocated to the k^{th} secondary user, hence $(\delta_{k,1}|\delta_{k,2}\cdots|\delta_{k,N})$ indicates whether k^{th} user is admitted in at least one of the N subcarriers. Therefore the cost function in (3.3.1) indicates the number of users admitted in the network. In (3.3.2), $P_{k,n}$ is the power allocated for the k^{th} secondary user over the n^{th} subcarrier. Due to the coexistence of primary user and secondary user in the frequency bands, this constraint ensures that the inter-

ference leakage to the l^{th} primary user is below the threshold γ_l. Constraint (3.3.3) ensures that each admitted secondary user achieves a predefined data rate. Constraint (3.3.4) ensures that the power consumed is below the maximum total transmission power P_{max}. It is possible to allocate more than one subcarrier for one secondary user, however multiple users cannot be accompanied together in one subcarrier. Therefore (3.3.5) performs mutual exclusive secondary user allocation in each subcarrier. This problem cannot be formulated into IP framework, because of the logical summation (i.e. OR) of the indication symbols $\delta_{k,n}$. If an secondary user has employed more than one subcarriers, for example subcarrier 1 and subcarrier 2, the normal summation of $\delta_{k,n}$ for this user is two, but the number of users admitted in the network should be counted as one. Therefore, maximization of the summation of $\delta_{k,n}$ will not serve the purpose. Therefore, $(\delta_{k,1}|\delta_{k,2}\cdots|\delta_{k,N})$ is used to formulate the objective function. Unfortunately, OR operation is not convex. Besides, due to the non-linear and non-convex property of (3.3.3), the original problem introduced in (3.3.1)-(3.3.5) is non-convex [93]. A suboptimal technique is proposed to solve this problem.

3.3.2 Algorithms Using Integer Programming

Similar to the previous section, $\mathbf{x}$ is defined as the binary vector and $\mathbf{p}$ as the transmission power vector. A matrix $\mathbf{Y}$ of size $K \times NKC$ is introduced and defined as follows:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_1 & \cdots & \mathbf{y}_1 \\ \mathbf{y}_2 & \mathbf{y}_2 & \cdots & \mathbf{y}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_K & \mathbf{y}_K & \cdots & \mathbf{y}_K \end{bmatrix}, \quad \left\{ \begin{array}{llll} \mathbf{y}_1 = [\mathbf{y}^T & \mathbf{0}_C^T & \ldots & \mathbf{0}_C^T] \in Z^{1\times KC}, \\ \mathbf{y}_2 = [\mathbf{0}_C^T & \mathbf{y}^T & \ldots & \mathbf{0}_C^T] \in Z^{1\times KC}, \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_K = [\mathbf{0}_C^T & \mathbf{0}_C^T & \ldots & \mathbf{y}^T] \in Z^{1\times KC}. \end{array} \right\},$$

where $\mathbf{y} = [1\ 1\ 1\ 1]^T \in Z^{4\times 1}$. Hence $\mathbf{Y}^T\mathbf{x}$ indicates number of subcarriers allocated to each user. For example, subcarrier allocation vector

$\mathbf{Y}^T\mathbf{x} = [2\ 1\ 1\ 1\ 0]^T$ indicates that the first user has been allocated two subcarriers, the second, third and fourth user have been allocated one subcarrier each and the fifth user has no subcarrier allocation.

First, the least favorable users are removed from the optimization problem by selecting the best possible candidates out of K users. This is a combinatorial optimization problem, however, a suboptimum technique called min-norm algorithm is proposed to choose the best users. The basic hypothesis is that for the best users, the number of subcarriers needed should be more or less the same. If a particular user needs many more subcarriers than the others, it means this user may not be an optimal selection. Hence $\mathbf{Y}^T\mathbf{x}$ is used to develop the suboptimum approach. This vector of length K should have either zeros (for the users not allowed in the network) or integers with small variance (less fluctuations). Hence, minimise the variance of $\mathbf{Y}^T\mathbf{x}$ is proposed. However, since number of users requesting the channels is greater than the number of subcarriers, all the subcarriers should be used. Hence the mean of $\mathbf{Y}^T\mathbf{x}$ should be one (because $sum(\mathbf{Y}^T\mathbf{x}) = N$ regardless of the optimization). Hence the minimization of variance of $\mathbf{Y}^T\mathbf{x}$ is equivalent to the minimization of the norm of $\mathbf{Y}^T\mathbf{x}$. However, the minimum norm of the subcarrier allocation vector is zero, resulting into zero throughput. Hence, minimize the norm of $\mathbf{Y}^T\mathbf{x}$ while keeping the throughput of the network over a certain threshold is required. As a result, another variable t is introduced to account for the throughput of the network. The objective of the problem is to minimize $\|\mathbf{Y}^T\mathbf{x}\| - \alpha t$ while keeping the throughput greater than t, where α is a constant weight which combines the minimization of vector norm $\|\mathbf{Y}^T\mathbf{x}\|$ and the maximization of data throughput into a single objective function. Therefore, for the first phase, the following optimization is proposed

$$\min_{x,t} \quad \|\mathbf{Y}^T\mathbf{x}\| - \alpha t \tag{3.3.6}$$

$$\text{s.t.} \quad \mathbf{H}[\mathbf{A}_c(\mathbf{p} \odot \mathbf{x})] \leq \boldsymbol{\xi} \tag{3.3.7}$$

$$\mathbf{d}^T\mathbf{x} \geq t, \ t \geq 0 \tag{3.3.8}$$

$$\mathbf{p}^T\mathbf{x} \leq P_{max} \tag{3.3.9}$$

$$\mathbf{0}_N \leq \mathbf{A}_c\mathbf{x} \leq \mathbf{1}_N. \tag{3.3.10}$$

The aim is to admit as many users as possible while minimizing the number of subcarriers taken by any of the admitted users and maximising the total throughput. The notation $\odot$ denotes the entrywise dot Hadamard product. $\mathbf{H} = \mathbf{H}_\mathrm{p} \odot \mathbf{H}_\mathrm{p} \in R^{L\times N}$, $\boldsymbol{\xi} = [\xi_1 \ \xi_2 \ \ldots \ \xi_L]^T \in C^{L\times 1}$. $\mathbf{d} = [(\mathbf{d}^1)^T \ \ldots \ (\mathbf{d}^N)^T]^T \in Z^{NKC\times 1}$, $\mathbf{d}^n = [(\mathbf{d}_{1,n})^T \ \ldots \ (\mathbf{d}_{K,N})^T]^T \in Z^{KC\times 1}$ and $\mathbf{d}_{k,n} = [1 \ 2 \ 4 \ 6]^T \in Z^{4\times 1}$, where $\mathbf{d}_{k,n}$ represents the number of bits required for four different modulation schemes considered in this letter. $\mathbf{0}_N = [0 \ 0 \ \ldots \ 0] \in \{0\}^{N\times 1}$ and $\mathbf{1}_N = [1 \ 1 \ \ldots \ 1] \in \{1\}^{N\times 1}$. The matrix $\mathbf{A}_c$ is same as that is defined in [37].

The above IP problem can be efficiently solved using Branch and Bound (BnB) method. By satisfying the total throughput constraint, the system will automatically choose the best users for the subcarriers. These selected users will be reserved in a table and other users who are not admitted in the first phase will be removed. In the reservation table, an efficiency factor for each reserved user is calculated as

$$\eta_k = \frac{\sum_{n=1}^{N} \mathbf{d}_{k,n}^T \mathbf{x}_{k,n}}{\sum_{n=1}^{N} \mathbf{p}_{k,n}^T \mathbf{x}_{k,n}} \tag{3.3.11}$$

The users are arranged in the table in the descending order of efficiency, i.e, the users with the best data rate/power efficiency will be placed on the top of the table. In the second phase, whether these reserved users satisfy the user data rate constraint are tested. If not, the algorithm will remove iteratively one user at a time starting with the least efficient user. Hence

the second phase of the proposed algorithm is

$$\max_{x} \quad \mathbf{d}^T\mathbf{x} \tag{3.3.12}$$

$$\text{s.t.} \quad \mathbf{H}[\mathbf{A}_c(\mathbf{p} \odot \mathbf{x})] \leq \boldsymbol{\xi} \tag{3.3.13}$$

$$\mathbf{A}_u\mathbf{x} \geq \mathbf{r} \tag{3.3.14}$$

$$\mathbf{p}^T\mathbf{x} \leq P_{max} \tag{3.3.15}$$

$$\mathbf{0}_N \leq \mathbf{A}_c\mathbf{x} \leq \mathbf{1}_N \tag{3.3.16}$$

where the matrix $\mathbf{A}_u$ is the same as that is defined in [37], vector $\mathbf{r}$ is defined as $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_K]^T \in \mathcal{C}^{K\times 1}$, r_k is the data rate requirement for the k^{th} user. Constraint (3.3.14) ensures that each secondary user will satisfy its QoS in terms of data rate. This iteration will continue until the data rate requirement, total power constraint and the interference thresholds are satisfied. The remaining users are the maximum number of users that could be allocated in the network.

3.3.3 Complexity Analysis

The optimal algorithm for user selection and resource allocation is a combinatorial problem. For the optimal algorithms, first, all possible combinations of the admitted users are required to check and the corresponding feasibility for each combination is required to be tested. For example, the feasibility for all $\binom{K}{N}$ combinations needs to be tested. If all combinations are infeasible, the network will not be able to serve all the users. Then the feasibility for $N-1$ users chosen out of K users is required to be tested. The number of users needs to reduce one by one until the feasibility is satisfied. In the worst case, the feasibility starting from all N users down to one user is required to check. In this case, there are $\sum_{k=1}^{N}\binom{K}{k}$ combinations. In each combination, integer linear programming needs to operate, the complexity of which is $O((kNC)^3)$ per iteration per branch of BnB method. The num-

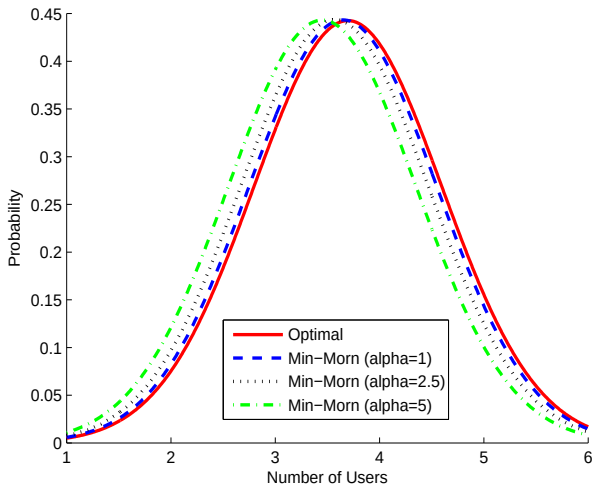

Figure 3.4. PDF of the number of secondary users served by the network for different α.

ber of iterations is approximately $O(\sqrt{kNC}log(1/\varepsilon))$ to achieve ε optimal solution, and it is required kNC number of branches in the worst case. It is generally known in the optimization community that the algorithm would converge typically in less than 10 iterations [95]. Hence the overall complexity of the optimum method is $\sum_{k=1}^{N}\binom{K}{k}O((kNC)^4)$. The proposed min-norm algorithm is a two phase program. The complexity of the first phase is $O((KNC)^4)$. For the second phase, (3.3.12)-(3.3.16) is required to run N times for the worst case, the complexity of which is $\sum_{k=1}^{N}O((kNC)^4)$. Hence the complexity of the suboptimal algorithm is substantially lower than that of the optimal algorithm, however, the sub optimal algorithm performs very closely to the optimal algorithm.

3.3.4 Simulation Results

A network with 2 primary users, 12 secondary users and 6 subcarriers is first considered. The required BER and the data rate for each secondary user have been set to 0.01 and 4 bits/Hz/symbol, respectively. The background noise power spectrum density for secondary user is set to 0.01. The upper

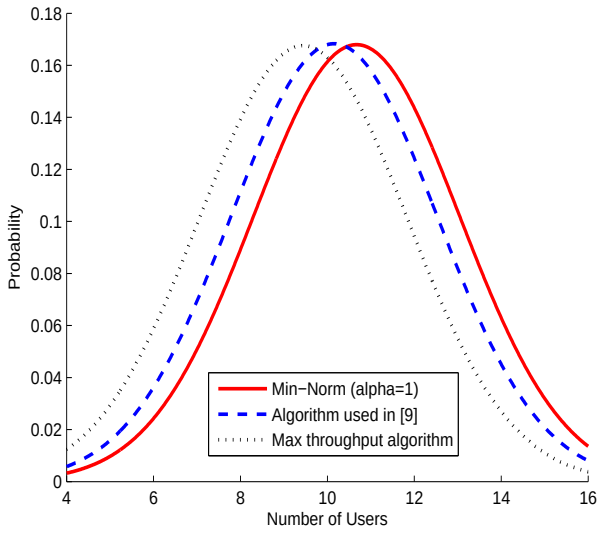

Figure 3.5. PDF of the number of secondary users served by the network for three different algorithms.

bound on the the interference for each primary user has been set to 0.01. The total available power for the secondary network basestation is set to 0.1. Multiple channel gains between the secondary network basestation and primary users and that between secondary network basestation and secondary users have been generated using independent Gaussian random variables. The average channel gain between the secondary network basestation and primary users is equal to unity while such a gain between the secondary network basestation and secondary users is equal to six. The performance of the optimal scheme is used as a bench mark for comparison. A Monte carlo experiment is performed with random channels and recorded the number of secondary users chosen by the Min-Norm algorithm and the optimal algorithm.

The probability density function (PDF) for the number of users is shown in Fig. 3.4. The proposed Min-Norm algorithm is performing very closely to the optimal resource allocation algorithm. The mean user value is 3.81 for the optimal algorithm and 3.74 for the Min-Norm algorithm with $\alpha = 1$. The performance of the proposed algorithm degrades slightly as α

increases. This is because as α increases, more emphasis is put on total data throughput maximization than user maximization. However, as can be seen in Fig. 3.4, the algorithm is less sensitive to variations in α. As there are no relevant existing works for cognitive radio network, the work of [96] is extended to cognitive radio network and compared the result with the Min-Norm algorithm. The first phase of the proposed algorithm have been also modified by considering maximization of sum throughput of the objective function in (3.3.6) and compared with the result. A network with 2 primary users, 100 secondary users and 16 subcarriers is considered. As can be seen in Fig. 3.5, the Min-Norm algorithm outperforms other two algorithms.

3.4 Conclusion

Resource allocation techniques for OFDMA based cognitive radio systems have been studied. Initially, the problem of adaptive radio resource allocation for an OFDMA-based cognitive radio network has been considered. An adaptive bit loading optimization problem has been formulated which is based on maximization of the total throughput under interference power constraint to primary users, individual data rate constraints for secondary users and the total transmission power constraint at the secondary network basestation. This radio resource allocation algorithm has been solved using integer linear programming. The proposed technique considers interference leakage to and from multiple primary users and secondary users in order to optimize the throughput of the secondary network while keeping interference to primary network below a specific value as well as satisfying secondary users' QoS requirements. The simulation framework explicitly considered baseband filter response of the primary network and its impact on the interference leakage to the secondary network. Simulation results demonstrate the ability of the algorithm to take advantage of the mutual

interference between primary users and secondary users, to maximize data throughput for secondary users. Then user admission control techniques have been studied. The problem of user maximization and resource allocation for an OFDMA-based cognitive radio network have been studied. The proposed technique maximizes the number of users served by the network while keeping interference to primary users below a threshold. The original problem is transformed into an integer programming based problem and a low complexity suboptimal algorithm was developed. The proposed suboptimal algorithm performs very closely to the optimal algorithm while having a lower complexity.

Chapter 4

SUM RATE MAXIMIZATION FOR SPECTRUM SHARING MULTIUSER MIMO NETWORKS

In this chapter, the capacity optimization problem for spectrum sharing networks is studied. A weighted sum rate maximization and rate balancing problems for spectrum sharing MIMO-OFDM based broadcast channels has been investigated first. This problem has been solved by converting the MIMO-OFDM channel into block diagonal form and using the principle of BC-MAC duality. The algorithm in its dual form has been solved using sub-gradient methods. Fading channels and the ergodic capacity region maximization problem is another interesting research area. A sum rate maximization problem for spectrum sharing MIMO BC under Rayleigh fading has then been studied. Based on multiple auxiliary variables, KKT optimality conditions and BC-MAC duality, this problem has been solved using an iterative technique. To further improve the performance of the network, relay technology has been introduced. A weighted sum rate maximization and rate balancing problem for a spectrum sharing MIMO based wireless relay network has been investigated. The aim is to maximize the sum rate

of the wireless relay network whilst ensuring the interference leakage to the primary user terminals during two time slots are below a specific value. This problem has been solved by asymmetrically allocating the power to different time slots and using the principle of BC-MAC duality. The simulation results demonstrated the convergence of the algorithm and the simultaneous satisfaction of maximum power and interference constraints.

4.1 Introduction

The capacity problem for a single-user MIMO Gaussian channel was first studied in [97]. Later, this work was extended to multiple users in [69]. One of the important precoding techniques for BCs is the dirty paper coding [98]. The principle of dirty paper coding (DPC) has been used to determine the achievable capacity region of MIMO-BC channels in [99]. It was shown in [100] that the sum rate MIMO-BC capacity is equal to the maximum sum rate of the achievable region of the BC channels for the case of two users. Furthermore, using the duality between Gaussian MACs and Gaussian BCs, it was shown in [1] that the achievable capacity region of the BC is the same as that of the dual MAC, and vice versa. In [69], this duality has been extended to multiuser MIMO Gaussian BC channels. Under a single sum power constraint, several efficient algorithms have been developed to obtain the optimal solutions for the MIMO-BC sum rate problems. A novel minimax BC-MAC duality was proposed in [73] for a Gaussian mutual information minimax problem. It has been shown that the uplink-downlink duality between BC and MAC channels can be obtained as a special case of this minimax duality. Based on this duality, an interior point algorithm was proposed in [74] to solve the MIMO-BC weighted sum rate maximization problem with per-antenna power constraints. It has been shown that the dual of the downlink is equivalent to the uplink problem with a different noise

covariance matrix. The duality relationship also holds for the achievable rate regions of the respective uplink and downlink channels. Furthermore, based on the duality, a weighted sum rate maximization algorithm for the CR-MIMO-BC has been proposed in [75]. A rate balancing technique for multiuser MIMO-OFDM network has been proposed in [101] by converting the MIMO-OFDM channels into a block diagonal form.

Fading channels and the ergodic capacity region maximization problems are another interesting research area. In this situation, the aim is to find the optimum transmit covariance matrices, i.e., the optimum transmit directions and power allocation for various users to maximize the ergodic capacity. There are two different feedback strategies, mean feedback and covariance feedback. It was shown in [102] that for both the mean and covariance feedbacks, the optimal transmit covariance matrix and the channel covariance matrix have the same eigenvectors for a MISO system. Furthermore, the authors in [103] proved the same results for a MIMO system with covariance feedback. The optimal transmit directions are the eigenvectors of the channel covariance matrix. This principle was proven in [104] for the mean feedback case, where the eigenvectors of the optimal transmit covariance matrix were shown to be the same as the right singular vectors of the channel mean matrix for a MIMO system. For the MIMO-MAC systems with partial CSI at the transmitters, the authors in [105] have proven that all the users should transmit in the direction of the eigenvectors of their own channel matrices, both for covariance and mean feedback cases. The optimum power allocation strategies for a single-user MIMO and a multi-user MIMO-MAC with partial CSI have been proposed in [106]. In this chapter, capacity optimization problems have been investigated for various spectrum sharing networks. Rate balancing techniques are also introduced to ensure fairness among users.

4.2 Sum Rate Maximization Technique for Spectrum Sharing MIMO-OFDM Broadcast Channels

In this section, a weighted sum rate maximization and rate balancing technique has been proposed for MIMO-OFDM-based spectrum sharing broadcast channels. The aim is to maximize the sum rate of the secondary users whilst ensuring interference leakage to the primary user terminals is below a specific value and each secondary user attains a specific portion of the total data rate. This problem has been solved by converting the MIMO-OFDM channel into a block diagonal form and using the principle of BC-MAC duality. The algorithm in its dual form has been solved using sub-gradient methods. The simulation results demonstrated the convergence of the algorithm and the simultaneous satisfaction of maximum power and interference constraints.

4.2.1 System Model Combining MIMO-OFDM and Cognitive Radio Network

A downlink cognitive radio network is considered for the system model. The basestation is equipped with N_t transmit antennas. Each secondary user $k \in \{1, 2, \ldots, K\}$ has N_r receive antennas. The K secondary users share the same spectrum as of the primary users. Each primary user is equipped with one receive antenna. An OFDM transmission scheme with N subcarriers is employed. The length of the cyclic prefix is assumed to be longer than the length of the channel so that no intersymbol interference occurs. The orthogonality between the subcarriers is assumed to be preserved during the transmission process. The received signal from the secondary network basestation to the k^{th} secondary user on the n^{th} subcarrier can be written as

$$\mathbf{y}_{n,k} = \mathbf{H}_{n,k}\mathbf{x}_n + \mathbf{n}_{n,k}, \tag{4.2.1}$$

where $\mathbf{H}_{n,k} \in C^{N_r \times N_t}$ is the channel matrix of user k on the n^{th} subcarrier and $\mathbf{n}_{n,k} \in C^{N_r \times 1}$ is a zero mean CSCG distributed random noise vector. Noise processes for different subcarriers are assumed to be uncorrelated. As shown in [101], a MIMO-OFDM system with frequency selective channels can be viewed as a MIMO non-frequency selective system using a block diagonal matrix as shown below,

$$\mathbf{H}_k = \mathbf{diag}[\mathbf{H}_{1,k}\ \mathbf{H}_{2,k} \cdots \mathbf{H}_{N,k}],$$

Using this model the received signal can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{4.2.2}$$

where $\mathbf{H} = [\mathbf{H}_1^T\ \mathbf{H}_2^T \cdots \mathbf{H}_K^T]^T$, $\mathbf{y} = [\mathbf{y}_1^T\ \mathbf{y}_2^T \cdots \mathbf{y}_K^T]^T$, $\mathbf{y}_k = [\mathbf{y}_{1,k}^T\ \mathbf{y}_{2,k}^T \cdots \mathbf{y}_{N,k}^T]^T$, $\mathbf{n} = [\mathbf{n}_1^T\ \mathbf{n}_2^T \cdots \mathbf{n}_K^T]^T$, $\mathbf{n}_k = [\mathbf{n}_{1,k}^T\ \mathbf{n}_{2,k}^T \cdots \mathbf{N,k}_K^T]^T$ and $\mathbf{x} = [\mathbf{x}_1^T\ \mathbf{x}_2^T \cdots \mathbf{x}_K^T]^T$.

4.2.2 Problem Statement

Assuming dirty paper coding is used at the transmitter, the data rate R_k^b for the k^{th} user can be written as

$$R_k^b = \log \frac{|\mathbf{I}_{N\times N_r} + \sum_{i=k}^{K} \mathbf{H}_i \mathbf{Q}_i^b \mathbf{H}_i^H|}{|\mathbf{I}_{N\times N_r} + \sum_{i=k+1}^{K} \mathbf{H}_i \mathbf{Q}_i^b \mathbf{H}_i^H|}, \tag{4.2.3}$$

where $\mathbf{Q}_k^b \in C^{NN_t \times NN_t}$ is the transmit covariance matrix for the k^{th} secondary user, $\mathbf{Q}_k^b \succeq 0$ indicates that $\mathbf{Q}_k^b$ is a semidefinite matrix. Using the above definition of data rate, the weighted sum rate maximization problem for the CR-MIMO-OFDM-BC can be written as

Problem 4.1 (Main Problem):

$$\max_{\{\mathbf{Q}_k^b\}k=1,\cdots,K:\mathbf{Q}_k^b\succeq 0} \quad \sum_{k=1}^{K} \mu_k R_k^b \tag{4.2.4}$$

$$\text{s.t.} \sum_{k=1}^{K} tr(\mathbf{Q}_k^b) \le P_t, \tag{4.2.5}$$

$$\sum_{k=1}^{K} \mathbf{h}_0^H tr(\mathbf{Q}_k^b)\mathbf{h}_0 \le P_I, \tag{4.2.6}$$

where μ_k denotes the weight used for the rate of the k^{th} secondary user. P_t and P_I represent the sum power constraint at the secondary network basestation and the interference threshold for the primary user. $\mathbf{h}_0 \in C^{NN_t \times 1}$ denotes the channel gain vector between the secondary network basestation and the primary user.

Applying the duality between Gaussian MAC and Gaussian BC as in [1] and [69], the general MIMO sum rate optimization problem can be solved. However, due to the additional interference constraint for the primary users and the OFDM block structure, the optimization problem for this CR-MIMO-OFDM-BC is non-convex [93] and it is difficult to solve it directly. Similar to [75], by introducing auxiliary dual variables for the interference power constraint and the sum power constraint, the original optimization problem in (4.2.4)-(4.2.6) is transformed to the following equivalent problem,

Problem 4.2 (Equivalent Problem):

$$\min_{\alpha\geq 0,\beta\geq 0} \max_{\{\mathbf{Q}_k^b\}k=1,\cdots,K:\mathbf{Q}_k^b\succeq 0} \sum_{k=1}^{K} \mu_k R_k^b \tag{4.2.7}$$

$$\text{s.t.} \quad \alpha(\sum_{k=1}^{K} tr(\mathbf{Q}_k^b) - P_t) + \beta(\sum_{k=1}^{K} \mathbf{h}_0^H tr(\mathbf{Q}_k^b)\mathbf{h}_0 - P_I) \leq 0, \tag{4.2.8}$$

where α and β are the auxiliary dual variables for the interference power constraint and the sum power constraint respectively. Constraint (4.2.8) is still non-convex due to the multiplication of dual variables with the optimization variables $\mathbf{Q}_k^b$. However, if α and β are set as constants, the constraint in (4.2.8) becomes convex. Hence, α and β are set as constants and write both constraints into a single constraint as described in [75] as follows

$$\alpha(\sum_{i=1}^{K} tr(\mathbf{Q}_k^b)) + \beta(\sum_{i=1}^{K} \mathbf{h}_0^H tr(\mathbf{Q}_k^b)\mathbf{h}_0) \leq \alpha P_t + \beta P_I, \tag{4.2.9}$$

By applying the BC-MAC duality principle as stated in [75], the equivalent problem shown above can be transformed to the dual MAC problem as follows

Problem 4.3 (CR-MIMO-OFDM-MAC):

$$\max_{\{\mathbf{Q}_k^m\}k=1,\cdots,K:\mathbf{Q}_k^m\succeq 0} \quad \sum_{k=1}^{K}\mu_k R_k^m \tag{4.2.10}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} tr(\mathbf{Q}_k^m)\sigma^2 \leq \alpha P_t + \beta P_I, \tag{4.2.11}$$

where $\mathbf{Q}_k^m \in C^{NN_r \times NN_r}$ is the transmit covariance matrix for the k^{th} secondary user in the MAC setup, σ^2 is the noise variance for all the secondary users, R_k^m is the rate achieved by the k^{th} secondary user in the MAC setup. Due to the duality between the MIMO-BC and the MIMO-MAC, the optimal encoding order for the MIMO-BC employing DPC is the reverse of the decoding order for the MIMO-MAC using successive interference cancelation (SIC) scheme. Hence R_k^m is written as

$$R_k^m = \log \frac{|\mathbf{A} + \sum_{i=1}^{k} \mathbf{H}_i^H \mathbf{Q}_i^m \mathbf{H}_i|}{|\mathbf{A} + \sum_{i=1}^{k-1} \mathbf{H}_i^H \mathbf{Q}_i^m \mathbf{H}_i|}, \tag{4.2.12}$$

where the matrix $\mathbf{A} \in C^{NN_t \times NN_t}$ is defined as [107]

$$\mathbf{A} = \alpha \mathbf{h}_0 \mathbf{h}_0^H + \beta \mathbf{I}_{N \times N_t}, \tag{4.2.13}$$

Due to the additional interference power constraint, this problem cannot be solved using the conventional BC-MAC duality as presented in [1] and [69]. Besides, due to the OFDM diagonal block setting, the algorithm used in [75] cannot be directly applied to solve the problem. Therefore, in the following section, an efficient algorithm is proposed to solve Problem 4.3.

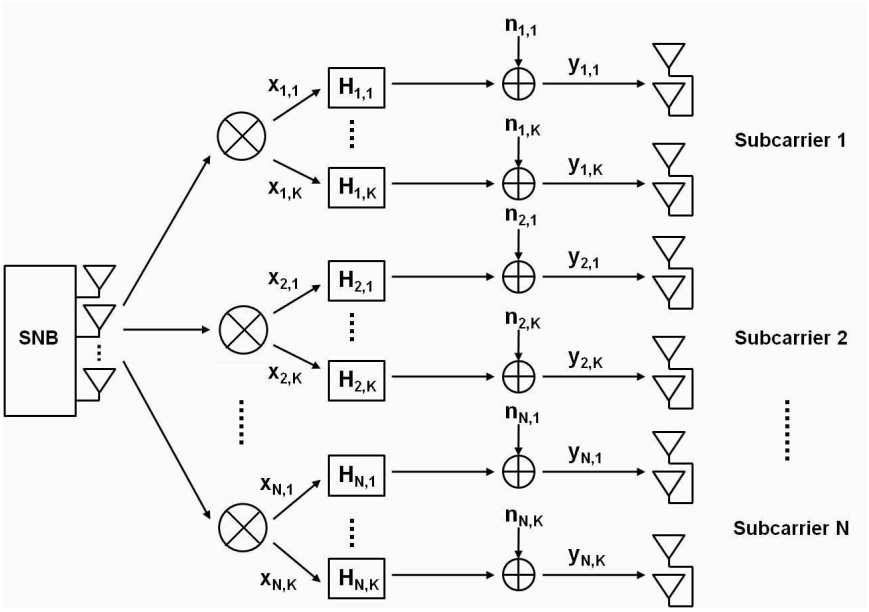

Figure 4.1. The system model for cognitive radio-MIMO-OFDM-BC (Problem 4) with $n_{i,k} \sim N(0, \sigma^2 \mathbf{I}_{Nr})$

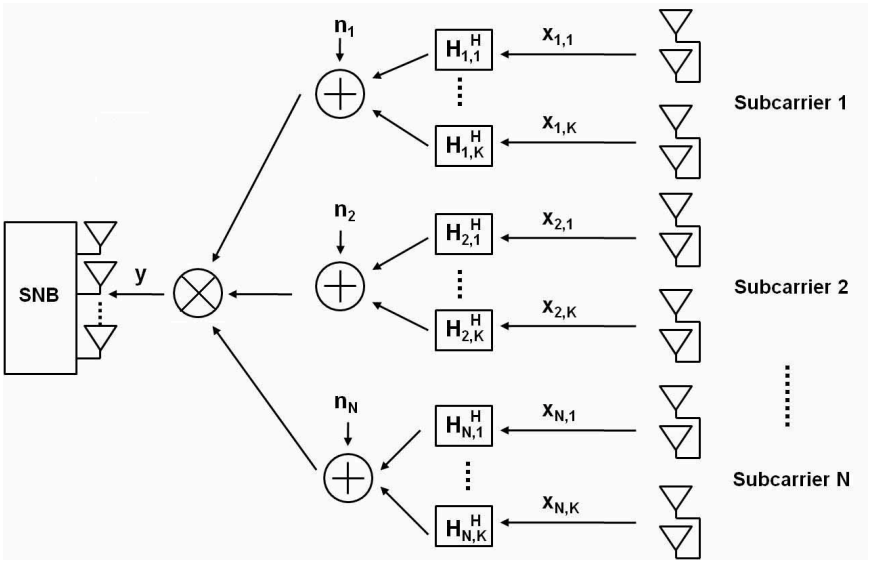

Figure 4.2. The system model for the dual cognitive radio-MIMO-OFDM-MAC (Problem 5) with $n_i \sim N(0, \alpha \mathbf{h}_{i,0} \mathbf{h}_{i,0}^H + \beta \mathbf{I}_{N_t})$

4.2.3 Dual MAC Weighted Sum Rate Maximization Problem for CR-MIMO-OFDM-BC

It has been shown in [101] that the optimal covariance matrices $\mathbf{Q}_k^m$ for the weighted sum rate maximization problem for the virtual MIMO-OFDM-

MAC have a block diagonal structure matching the structure of their respective channels $\mathbf{H}_k$ as shown below

$$\mathbf{Q}_k^m = diag[\mathbf{Q}_{1,k}^m \ \mathbf{Q}_{2,k}^m \cdots \mathbf{Q}_{N,k}^m],$$

Hence the original problem using the block diagonal structure can be decoupled for different subcarriers. This will also decrease the complexity dramatically. The Problem 4.2 and the Problem 4.3 are now decoupled as described in Problem 4.4 and Problem 4.5 and as shown in Fig. 4.1 and Fig. 4.2.

Problem 4.4 (Decoupled CR-MIMO-OFDM-BC):

$$\min_{\alpha\geq 0,\beta\geq 0} \max_{\{\mathbf{Q}_{n,k}^b\}k=1,\cdots,K:\mathbf{Q}_{n,k}^b\succeq 0} \sum_{k=1}^{K}\sum_{n=1}^{N}\mu_k R_{n,k}^b \tag{4.2.14}$$

$$\text{s.t.}\alpha(\sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^b) - P_t)$$

$$+\beta(\sum_{k=1}^{K}\sum_{n=1}^{N}\mathbf{h}_{n,0}^H tr(\mathbf{Q}_{n,k}^b)\mathbf{h}_{n,0} - P_I) \leq 0, \tag{4.2.15}$$

Problem 4.5 (Decoupled CR-MIMO-OFDM-MAC):

$$\max_{\{\mathbf{Q}_{n,k}^m\}k=1,\cdots,K.\ n=1,\cdots,N\ :\mathbf{Q}_{n,k}^m\succeq 0} \sum_{i=1}^{K}\sum_{n=1}^{N}\mu_k R_{n,k}^m \tag{4.2.16}$$

$$\text{s.t.}\sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^m)\sigma^2 \leq \alpha P_t + \beta P_I, \tag{4.2.17}$$

where $R_{n,k}^m$ is the rate achieved by the k^{th} secondary user on the n^{th} subcarrier for the MAC, $\mathbf{Q}_{n,k}^m \in C^{N_r\times N_r}$ denotes the transmit covariance matrix for the k^{th} secondary user on the n^{th} subcarrier for the MAC, $\mathbf{h}_{n,0} \in C^{N_t\times 1}$ denotes the channel gain vector on the n^{th} subcarrier between the secondary network basestation and the primary user. As a result, the general procedure for solving this problem is described in the following pseudo of code.

1) Initialization: $\alpha^{(1)}$, $\beta^{(1)}$, $n = 1$ **Repeat** 2) a. Find the optimal solution $\mathbf{Q}_{n,k}^m$ of the decoupled CR-MIMO-OFDM-MAC Problem 4.5; b. Map $\mathbf{Q}_{n,k}^m$ back to $\mathbf{Q}_{n,k}^b$ of the CR-MIMO-OFDM-BC problem and solve Problem 4.4; c. Update $\alpha^{(n)}$ and $\beta^{(n)}$ using the subgradient method; d. $n = n + 1$ 3) **Stop at convergence.**

Table 4.1. Pseudo-Code Description

The exact methods for solving the problem, in particular, 2(a)-2(c) in the pseudocode will be explained in the subsequent sections.

4.2.4 The Optimization for the Solution of $\mathbf{Q}_{n,k}^m$ in Problem 4.5

According to (4.2.12), the objective function in (4.2.16) can be rewritten as

$$f(\mathbf{Q}_{1,1}^m, \cdots, \mathbf{Q}_{N,K}^m) = \sum_{k=1}^{K}\sum_{n=1}^{N} \eta_k \log|\mathbf{A}_n + \sum_{i=1}^{k} \mathbf{H}_{n,i}^H \mathbf{Q}_{n,i}^m \mathbf{H}_{n,i}|, \qquad (4.2.18)$$

where $\mathbf{A}_n = \alpha \mathbf{h}_{n,0}\mathbf{h}_{n,0}^H + \beta \mathbf{I}_{N_t}$, $\eta_k = \mu_k - \mu_{k+1}, \mu_{K+1} = 0$.

In the following, based on [108] and [75], an iterative algorithm based on the primal dual method is proposed. First of all, problem 4.5 can be rewrote as follows:

$$\max_{\{\mathbf{Q}_{n,k}^m\}k=1,\cdots,K.\ n=1,\cdots,N\ :\mathbf{Q}_{n,k}^m \succeq 0} f(\mathbf{Q}_{1,1}^m, \cdots \mathbf{Q}_{N,K}^m) \qquad (4.2.19)$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^m)\sigma^2 \leq P_{TOTAL}, \qquad (4.2.20)$$

where $P_{TOTAL} = \alpha P_t + \beta P_I$. Due to the semidefinite matrix constraint $\mathbf{Q}_{n,k}^m \succeq 0$, the eigenvalue $q_{n,k,j}$ of transmit covariance matrix $\mathbf{Q}_{n,k}^m$ should be

non-negative, i.e., $q_{n,k,j} \geq 0$. The Lagrangian duality function is

$$\begin{aligned} L(\mathbf{Q}_{1,1}^m, \cdots, \mathbf{Q}_{N,K}^m, \lambda, \delta_{n,k,j}) &:= f(\mathbf{Q}_{1,1}^m, \cdots, \mathbf{Q}_{N,K}^m) \\ -\lambda(\sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^m)\sigma^2 - P) &+ \sum_{k=1}^{K}\sum_{n=1}^{N}\sum_{j=1}^{M} \delta_{n,k,j} q_{n,k,j}, \end{aligned} \tag{4.2.21}$$

where λ and $\delta_{n,k,j}$ are the Lagrangian multipliers associated with the total power constraint and the positive eigenvalues constraints respectively. The dual objective function of (4.2.19) is

$$g(\lambda) = \max_{\mathbf{Q}_{n,k}^m \succeq 0} \; L(\mathbf{Q}_{1,1}^m, \cdots \mathbf{Q}_{N,K}^m, \lambda), \tag{4.2.22}$$

Since the above problem is convex, the duality gap between the original problem and the dual problem is zero. The dual problem is given by

$$\min_{\lambda} g(\lambda) \quad s.t. \quad \lambda \geq 0, \tag{4.2.23}$$

Let us explain the proposed algorithm. An initial value for λ has been chosen and solve the optimization problem in (4.2.22). λ is then updated according to the steepest descent direction of $g(\lambda)$ until the convergence of the algorithm. Similar to [109], λ is updated via the gradient of equation (4.2.21) for each subcarrier. The gradient is determined as follows

$$\nabla_{\mathbf{Q}_{n,k}} L := \frac{\partial f(\mathbf{Q}_{1,1}^m, \cdots, \mathbf{Q}_{N,K}^m)}{\partial \mathbf{Q}_{n,k}^m} - \lambda \mathbf{I}_{Nr}, \tag{4.2.24}$$

where

$$\begin{aligned} &\frac{\partial f(\mathbf{Q}_{1,1}^m, \cdots, \mathbf{Q}_{N,K}^m)}{\partial \mathbf{Q}_{n,k}^m} \\ &= \sum_{j=k}^{K} \eta_j \mathbf{H}_{n,k} (\mathbf{A}_n + \sum_{i=1}^{j} \mathbf{H}_{n,i}^H \mathbf{Q}_{n,i}^m \mathbf{H}_{n,i})^{-1} \mathbf{H}_{n,k}^H, \end{aligned} \tag{4.2.25}$$

The transmit covariance matrix $\mathbf{Q}_{n,k}^m$ is first updated using the gradient (4.2.24) as follows

$$\mathbf{Q}_{n,k}^m(n) = \mathbf{Q}_{n,k}^m(n-1) + t\nabla_{\mathbf{Q}_{n,k}} L, \tag{4.2.26}$$

where t is a small step size. The eigenvalue decomposition is then performed to the gradient $\nabla_{\mathbf{Q}_{n,k}} L$, and the transmit covariance matrix $\mathbf{Q}_{n,k}^m$ is updated by using only the positive eigenvalues and the corresponding eigenvectors as

$$\mathbf{Q}_{n,k}^m(n) = \sum_j [q_{n,k,j}]^+ \upsilon_{n,k,j} \upsilon_{n,k,j}^H, \tag{4.2.27}$$

where $q_{n,k,j}$ and $\upsilon_{n,k,j}$ are the j^{th} eigenvalue and the corresponding eigenvector of $\mathbf{Q}_{n,k}^m$ respectively. $[q_{n,k,j}]^+$ denotes $max(q_{n,k,j}, 0)$. Applying the method of [75], it can be shown that that the subgradient of $g(\lambda)$ is $P_{TOTAL} - \sum_{k=1}^K \sum_{n=1}^N tr(\mathbf{Q}_{n,k}^m)\sigma^2$.

The transmit covariance matrix $\mathbf{Q}_{n,k}^m$ is updated using (4.2.24)-(4.2.27). After transmit covariance matrix $\mathbf{Q}_{n,k}^m$ converges, the power consumed is compared with P_{TOTAL}. If $P_{TOTAL} > \sum_{k=1}^K \sum_{n=1}^N tr(\mathbf{Q}_{n,k}^m)\sigma^2$, the value of λ should be increased, and decreased otherwise. This process will continue until $g(\lambda)$ converges.

4.2.5 Mapping MAC Optimization Solution to BC Solution

Once the optimal solution for the decoupled dual MAC optimization problem has been found, the MAC covariance matrix is required to map back to the BC covariance matrix. As the block diagonal structure is decoupled into linear structure, the cognitive radio-MIMO-OFDM-BC share the same capacity region as of the CR-MIMO-OFDM-MAC. Thus the conventional BC-MAC transformation algorithm has been extended to the CR-MIMO-

1) Initialization setting: $\mathbf{H}_k = diag[\mathbf{H}_{1,k}\ \mathbf{H}_{2,k} \cdots \mathbf{H}_{N,k}]$
$\mathbf{Q}_k^m = diag[\mathbf{Q}_{1,k}^m\ \mathbf{Q}_{2,k}^m \cdots \mathbf{Q}_{N,k}^m]$ for $k = 1, 2, \cdots K$
2) a. **For** $k = 1, 2, \cdots, K$
b. Define $\hat{\mathbf{H}}_k = \mathbf{H}_k \mathbf{A}^{-1/2}$,
$M_k = \mathbf{I}_{N \times N_r} + \sum_{j=k+1}^{K} \hat{\mathbf{H}}_j^H \mathbf{Q}_j^m \hat{\mathbf{H}}_j$,
$N_k = \mathbf{I}_{N \times N_t} + \hat{\mathbf{H}}_j (\sum_{j=k+1}^{K} \mathbf{Q}_j^m) \hat{\mathbf{H}}_j^H$
c. Define the effective channel as $\breve{\mathbf{H}}_k = \mathbf{M}_k^{-1/2} \hat{\mathbf{H}}_k \mathbf{N}_k^{-1/2}$
d. Apply Singular Value Decomposition to the effective channel $\breve{\mathbf{H}}_k = \mathbf{R}_k \mathbf{D}_k \mathbf{L}_k$
e. Define a new matrix
$\mathbf{J}_k = \mathbf{A}^{-1/2} \mathbf{M}_k^{-1/2} \mathbf{R}_k \mathbf{L}_k^H \mathbf{N}_k^{1/2}$
f. $\mathbf{Q}_k^b = \mathbf{J}_k \mathbf{Q}_k^m \mathbf{J}_k^H$
g. **End for**
3) Decouple the BC transmit covariance matrix $\mathbf{Q}_k^b$ from the diagonal structure: $\mathbf{Q}_k^b = diag[\mathbf{Q}_{1,k}^b\ \mathbf{Q}_{2,k}^b \cdots \mathbf{Q}_{N,k}^b]$ for $K = 1, 2, \cdots K$

Table 4.2. BC-MAC transformation algorithm to the CR-MIMO-OFDM framework

OFDM framework as in Table 4.2 [107].

4.2.6 The Complete Solution of the Original Problem

The complete description of the algorithm is now presented to solve Problem 4.4. The inner part of Problem 4.4 can be rewritten as

$$g(\alpha, \beta) = \max_{\{\mathbf{Q}_{n,k}^m\}k=1,\cdots,K:\mathbf{Q}_{n,k}^m \succeq 0} \sum_{k=1}^{K} \sum_{n=1}^{N} \mu_k R_{n,k}^b \tag{4.2.28}$$

$$\text{s.t.} \quad \alpha(\sum_{k=1}^{K} \sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^m) - P_t)$$

$$+\beta(\sum_{k=1}^{K} \sum_{n=1}^{N} \mathbf{h}_{n,0}^H tr(\mathbf{Q}_{n,k}^m) \mathbf{h}_{n,0} - P_I) \leq 0, \tag{4.2.29}$$

After solving the above optimization problem, it is required to minimize it with respect to the auxiliary dual variables α and β as follows

$$\min g(\alpha, \beta) \quad \text{s.t. } \alpha \geq 0, \beta \geq 0, \tag{4.2.30}$$

1) Initialization: $\alpha^{(1)}$, $\beta^{(1)}$, $n = 1$
Repeat
2) a. Find the optimal solution of the decoupled dual MAC Problem 4.5 using the subgradient update method;
b. Find the solution of the BC problem via the MAC-to-BC mapping algorithm;
c. Update $\alpha^{(n)}$ and $\beta^{(n)}$ using subgradient method (4.2.31) and (4.2.32);
d. $n = n + 1$
3) **Stop** when $|P_t - \sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^{b})| \leq \epsilon$ and $|P_I - \sum_{k=1}^{K}\sum_{n=1}^{N} \mathbf{h}_{n,0}^{H} tr(\mathbf{Q}_{n,k}^{b})\mathbf{h}_{n,0}| \leq \epsilon$ are satisfied simultaneously.

Table 4.3. Proposed iterative resource allocation algorithm

Hence, the remaining task is to determine the optimal α and β. It has been shown in [75] that the subgradient of $g(\alpha, \beta)$ is $[P_t - \sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^{b}), P_I - \sum_{k=1}^{K}\sum_{n=1}^{N} \mathbf{h}_{n,0}^{H} tr(\mathbf{Q}_{n,k}^{b})\mathbf{h}_{n,0}]$. It has been shown in [110] that with a constant step size, the subgradient algorithm converges to a solution which is in the close proximity to the optimal solution. Hence, α and β can be updated as follows

$$\alpha^{(n+1)} = \alpha^{(n)} + t(\sum_{k=1}^{K}\sum_{n=1}^{N} tr(\mathbf{Q}_{n,k}^{b}) - P_t) \tag{4.2.31}$$

$$\beta^{(n+1)} = \beta^{(n)} + t(\sum_{k=1}^{K}\sum_{n=1}^{N} \mathbf{h}_{n,0}^{H} tr(\mathbf{Q}_{n,k}^{b})\mathbf{h}_{n,0} - P_I), \tag{4.2.32}$$

The proposed iterative algorithm for resource allocation of CR-MIMO-OFDM-BC problem can be summarized as Table 4.3.

4.2.7 Simulation Results

In order to illustrate the effectiveness of the proposed algorithm, several simulation results are provided in this section. The elements of the channel matrix $\mathbf{H}_{n,k}$ are assumed to be CSCG variables with zero mean and unitary variance. The components of interference channel $\mathbf{h}_{n,0}$ are also assumed to be CSCG variables with zero mean and unitary variance. The total power

constraint P_t and the interference power threshold P_I are set to be 10 and

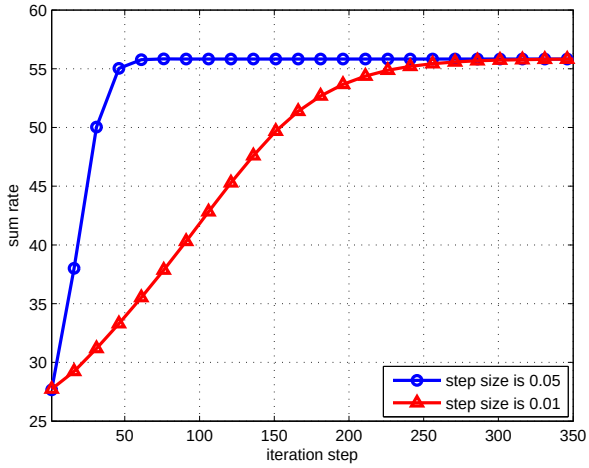

Figure 4.3. Convergence of the proposed algorithm ($N_t = 5$, $N_r = 3$, $K = 5$, $N = 16$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

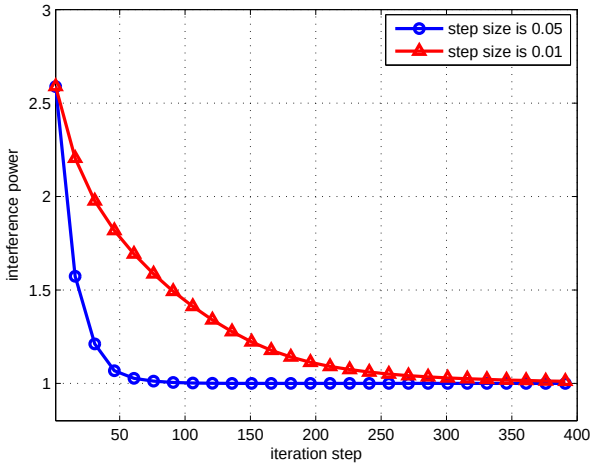

Figure 4.4. Evolution of interference of the proposed iterative algorithm ($N_t = 5$, $N_r = 3$, $K = 5$, $N = 16$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

1 respectively. The noise variances at the secondary users are also set to unity.

A CR-BC-MIMO-OFDM network with five secondary users and one primary user is considered in the simulation. Five transmit antennas, three receive antennas and 16 subcarriers are assumed in the system. The data rates for all the secondary users are weighted equally. Fig. 4.3 depicts the

sum rate versus the number of iterations of the proposed algorithm for two

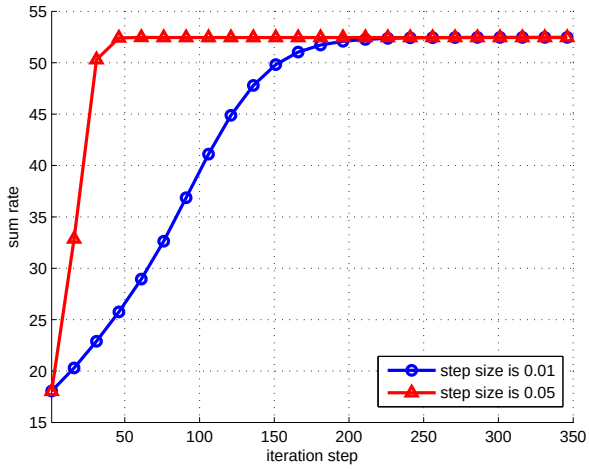

Figure 4.5. Convergence behavior of the proposed algorithm ($N_t = 5$, $N_r = 3$, $K = 5$, $N = 16$, $P_I = 1$, $P_t = 10$ and $\mu_1 = 5$, $\mu_1 = 4$, $\mu_1 = 3$, $\mu_1 = 2$, $\mu_1 = 1$).

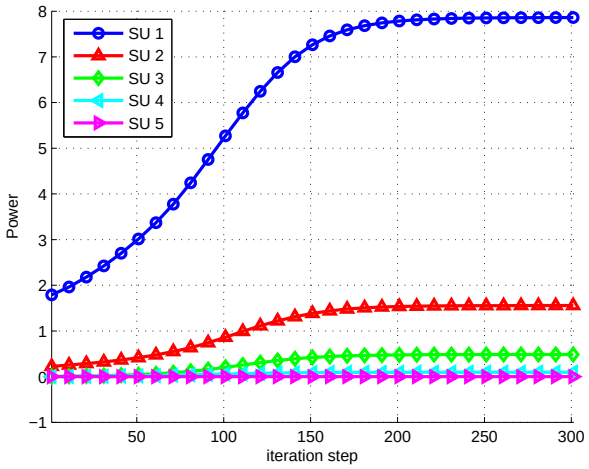

Figure 4.6. Power allocation of the proposed algorithm for different secondary users ($N_t = 5$, $N_r = 3$, $K = 5$, $N = 16$, $P_I = 1$, $P_t = 10$ and $\mu_1 = 5$, $\mu_1 = 4$, $\mu_1 = 3$, $\mu_1 = 2$, $\mu_1 = 1$).

different step sizes $t = 0.05$ and $t = 0.01$. It is clear from Fig. 4.3 that the step size affects the convergence speed of the proposed algorithm. Fig. 4.4 depicts the interference power at the primary user against the number of iterations. As seen, the figure that the interference power approaches to $P_I = 1$ when the proposed algorithm converges.

Different data rate weights for the secondary users are also considered in this simulation. Data rate weights for different users are set as $\mu_1 = 5,\ \mu_1 =$

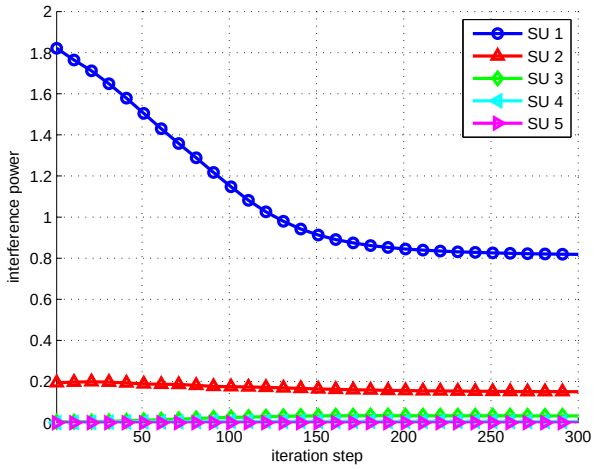

Figure 4.7. Interference behavior of the proposed algorithm for different secondary users ($N_t = 5,\ N_r = 3,\ K = 5,\ N = 16,\ P_I = 1,\ P_t = 10$ and $\mu_1 = 5,\ \mu_1 = 4,\ \mu_1 = 3,\ \mu_1 = 2,\ \mu_1 = 1$).

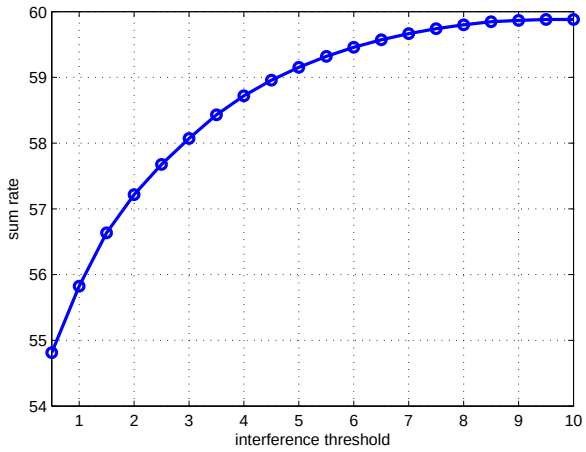

Figure 4.8. Comparison of the optimal achievable rates obtained by the proposed algorithm for different interference thresholds ($N_t = 5,\ N_r = 3,\ K = 5,\ N = 16,\ P_I = 1,\ P_t = 10$ and $\mu_i = 1$ for all i).

$4,\ \mu_1 = 3,\ \mu_1 = 2,\ \mu_1 = 1$. Fig. 4.5 depicts the weighted sum rate versus the number of iterations of the proposed algorithm for step sizes $t = 0.05$ and $t = 0.01$. The sum rate is lower as compared to that presented in Fig. 4.3. This is because the system may loose user diversity gain due to priority

settings. Similar to Fig. 4.3, the step size affects the convergence speed of the proposed algorithm. The power consumed by all the secondary users against the number of iterations is depicted in Fig. 4.6. As seen, secondary user 1 consumes most of the power and settles at eight. Secondary user 2 consumes only 1.6 due to a lower priority in terms of the data rate. Fig.

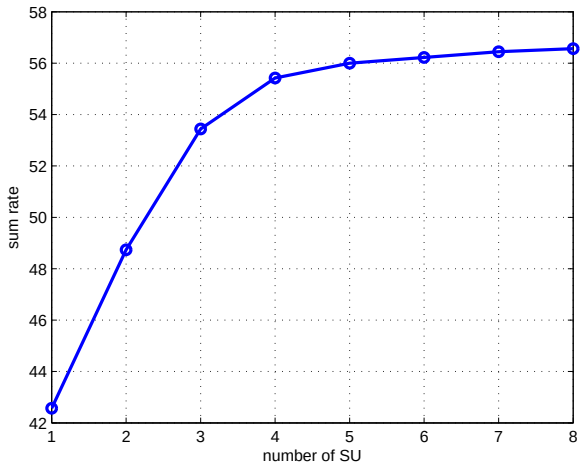

Figure 4.9. Comparison of the optimal achievable rates obtained by the proposed algorithm for different number of users admitted in the network ($N_t = 5$, $N_r = 3$, $K = 5$, $N = 16$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

4.7 depicts the interference power at the primary user against the number of iterations. It can be seen from the figure that secondary user 1 introduces most of the interference to the primary user and approaches 0.81.

In Fig. 4.8, all the settings are fixed as before, but the interference threshold is modified from 0.1 to 10. As seen, the sum rate increases dramatically when the interference threshold is set to a higher value. The interference threshold is then fixed to one but varied the number of secondary users served by the network, from one to eight. Due to the user diversity gain, as shown in Fig. 4.9 the sum rate increases dramatically until the number of secondary users reaches five. After that, the sum rate increases slightly and approaches an asymptotic value. This is because any gains achieved by user diversity is overwhelmed by the mutual interference between users for large

number of users. Hence the sum rate increases marginally.

4.3 Sum Rate Maximization for Spectrum Sharing Multiuser MIMO Network under Rayleigh Fading

In this section, a sum rate maximization problem for the CR-MIMO-BC under Rayleigh fading with partial CSI in the form of covariance feedback has been investigated. In contrast to previous section on MIMO CR networks, in this section, the ergodic capacity region has been studied based on fading channels. In this situation, the aim is to find the optimum transmit covariance matrices, i.e., the optimum transmit directions and power allocation for various users to maximize the ergodic capacity. Based on multiple auxiliary variables, KKT optimality conditions and BC-MAC duality, an iterative algorithm has been developed to solve the equivalent problem using Lagrangian optimization. Simulation results demonstrate the proposed algorithm converges to a globally optimal solution.

4.3.1 System Model and Problem Statement

A downlink underlay cognitive radio network has been considered. The SNB is equipped with N_t transmit antennas. Each SU $k \in \{1, 2, \ldots, K\}$ has N_r receive antennas. All the SUs share the same spectrum as of the PUs. Each PU is equipped with one receive antenna. The downlink channel between the transmitter and the user k is represented by a matrix $\mathbf{H}_k \in C^{N_t \times N_r}$. It is assumed that the receivers of SUs know the downlink CSI, while the SNB has only the statistical model of the channel. The statistical model that has been considered in this section is the partial CSI in terms of covariance matrix feedback. In this model, there exists correlation between the signals transmitted by or received at different antenna elements. For each user, the

downlink channel is modeled as [111],

$$\mathbf{H}_k = (\mathbf{\Phi}_k^b)^{1/2}\mathbf{Z}_k^b(\mathbf{\Sigma}_k^b)^{1/2}, \tag{4.3.1}$$

where the transmit antenna correlation matrix, $\mathbf{\Sigma}_k^b$, is the correlation between the signals transmitted for user k and the SNB, and the receive antenna correlation matrix, $\mathbf{\Phi}_k^b$, is the correlation between the signals received by the N_r receive antennas of user k. The entries of $\mathbf{Z}_k^b$ are i.i.d., zero-mean, unity-variance complex Gaussian random variables. In this section, it is assumed that the SNB transmitter does not have any physical restrictions hence there is sufficient spacing between the antenna elements, such that the signals transmitted from different antenna elements are uncorrelated. As a result, the transmit antenna correlation matrix is assumed to be an identity matrix, i.e., $\mathbf{\Sigma}_k^b = \mathbf{I}$. Therefore, the channel for the user k is written as

$$\mathbf{H}_k = (\mathbf{\Phi}_k^b)^{1/2}\mathbf{Z}_k^b, \tag{4.3.2}$$

The matrix $\mathbf{\Phi}_k^b$ is defined as the channel covariance feedback matrix for user k. Essentially, this covariance feedback model has been used in [103] and [104].

Assuming dirty paper coding at the transmitter and the encoding order is $(1, \cdots, K)$, i.e., the codeword of user 1 is encoded first, the average data rate R_k^b for the k^{th} user can be written as [69]

$$R_k^b = E_{\mathbf{H}}[\log \frac{|\mathbf{I}_{N_r} + \sum_{i=k}^{K} \mathbf{H}_i\mathbf{Q}_i^b\mathbf{H}_i^H|}{|\mathbf{I}_{N_r} + \sum_{i=k+1}^{K} \mathbf{H}_i\mathbf{Q}_i^b\mathbf{H}_i^H|}], \tag{4.3.3}$$

where $E_{\mathbf{H}}$ is the expectation operator with respect to the channel variations in the matrix $\mathbf{H} = [\mathbf{H}_1^T, \cdots, \mathbf{H}_K^T]$, and $|\cdot|$ is the determinant operator. Let $\mathbf{Q}_k^b = E[\mathbf{x}_k\mathbf{x}_k^H]$ be the transmit covariance matrix for the k^{th} SU, $\mathbf{x}_k$ is the signal transmitted by the k^{th} user at the SNB, $\mathbf{Q}_k^b \in C^{N_t \times N_t}$. Hence $\mathbf{Q}_k^b$ is

a semidefinite matrix, i.e., $\mathbf{Q}_k^b \succeq 0$. Using the above definition for data rate, the sum rate maximization problem is written as

$$\max_{\{\mathbf{Q}_k^b\}k=1,\cdots,K:\mathbf{Q}_k^b\succeq 0} \quad \sum_{k=1}^{K} R_k^b \tag{4.3.4}$$

$$\text{s.t.} \sum_{k=1}^{K} tr(\mathbf{Q}_k^b) \leq P_t, \tag{4.3.5}$$

$$E_{\mathbf{h0}}[\sum_{k=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0] \leq P_I, \tag{4.3.6}$$

where P_t and P_I represent the sum power constraint at the SNB and the interference threshold for the PU. $\mathbf{h}_0 \in C^{N_t \times 1}$ denotes the channel gain vector between the SNB and the PU.

4.3.2 Dual MAC Optimization Problem

The general MIMO sum rate maximization problem can be solved by applying the duality between the Gaussian MACs and the Gaussian BCs as in [1] and [69]. However, due to the additional interference constraint for the primary users and the expectation operator with respect to the channel variations, the optimization problem for this CR-MIMO-BC is non-convex [93] and it is difficult to solve it directly. Similar to [75], by introducing auxiliary dual variables for the interference power constraint and the sum power constraint, the original problem in (4.3.4)-(4.3.6) is transformed to the following equivalent problem,

$$\min_{\alpha\geq 0,\beta\geq 0} \max_{\{\mathbf{Q}_k^b\}k=1,\cdots,K:\mathbf{Q}_k^b\succeq 0} \quad \sum_{k=1}^{K} R_k^b \tag{4.3.7}$$

$$\text{s.t.} \quad \alpha(\sum_{k=1}^{K} tr(\mathbf{Q}_k^b) - P_t) \tag{4.3.8}$$

$$+\beta E_{\mathbf{h0}}[(\sum_{k=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0] - P_I) \leq 0,$$

where α and β are the auxiliary dual variables for the interference power constraint and the sum power constraint respectively. The constraint in

(4.3.8) is still non-convex due to the multiplication of dual variables with the optimization variables $\mathbf{Q}_k^b$. However, if α and β are set to constants, the constraints in (4.3.8) becomes convex. Hence, α and β are set to constants and both constraints are written into a single constraint as described in [75] as follows

$$\alpha(\sum_{i=1}^{K} tr(\mathbf{Q}_k^b)) + \beta E_{\mathbf{h0}}(\sum_{i=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0) \leq \alpha P_t + \beta P_I, \tag{4.3.9}$$

By applying the BC-MAC duality as described in [75], the equivalent problem shown above can be transformed to the dual MAC problem as follows

$$\max_{\{\mathbf{Q}_k^m\}k=1,\cdots,K:\mathbf{Q}_k^m \succeq 0} \quad \sum_{k=1}^{K} R_k^m \tag{4.3.10}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} tr(\mathbf{Q}_k^m)\sigma^2 \leq \alpha P_t + \beta P_I, \tag{4.3.11}$$

where $\mathbf{Q}_k^m \in C^{N_r \times N_r}$ is the transmit covariance matrix for the k^{th} SU in the virtual MAC setup, σ^2 is the noise variance for all the SUs, R_k^m is the rate achieved by the k^{th} SU in the virtual MAC setup. Due to the duality between the MIMO-BC and the MIMO-MAC, the optimal encoding order for the MIMO-BC employing DPC is the reverse of the decoding order for the MIMO-MAC using successive interference cancelation scheme. Besides, due to BC-MAC duality, the SNB acts as a receiver. Hence the virtual uplink channel for user k has been written as follows

$$\mathbf{H}_k^m = \mathbf{H}_k^H = ((\mathbf{\Phi}_k^b)^{1/2}\mathbf{Z}_k^b)^H, \tag{4.3.12}$$

As a result, the rate achieved by the k^{th} user R_k^m in the virtual MAC setup is written as

$$R_k^m = E_{\mathbf{H}}[\log \frac{|\mathbf{A} + \sum_{i=1}^{k} \mathbf{H}_i^H \mathbf{Q}_i^m \mathbf{H}_i|}{|\mathbf{A} + \sum_{i=1}^{k-1} \mathbf{H}_i^H \mathbf{Q}_i^m \mathbf{H}_i|}], \tag{4.3.13}$$

where the matrix $\mathbf{A} \in C^{N_t \times N_t}$ is defined as [107]

$$\mathbf{A} = \alpha E_{\mathbf{h0}}[\mathbf{h}_0 \mathbf{h}_0^H] + \beta \mathbf{I}_{N \times N_t} \tag{4.3.14}$$

Substituting (4.3.13) into the dual MAC problem in (4.3.10)-(4.3.11) , the problem is transformed to the following problem

$$\max_{\{\mathbf{Q}_k^m\} k=1,\cdots,K : \mathbf{Q}_k^m \succeq 0} E_{\mathbf{H}}[\log|\mathbf{A} + \sum_{i=1}^{K} \mathbf{H}_i^H \mathbf{Q}_i^m \mathbf{H}_i|] \tag{4.3.15}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} tr(\mathbf{Q}_k^m)\sigma^2 \leq \alpha P_t + \beta P_I \tag{4.3.16}$$

As the antennas at the SNB have been assumed to be uncorrelated, it is assumed that $E_{\mathbf{h0}}[\mathbf{h}_0 \mathbf{h}_0^H]$ is an identity matrix. Let $\mathbf{\Phi}_k = \mathbf{U}_{\mathbf{\Phi}_k} \mathbf{\Lambda}_{\Phi_k} \mathbf{U}_{\mathbf{\Phi}_k}^H$ the spectral decomposition for the channel covariance matrix of user k. It has been shown in [105] that the optimum transmit covariance matrix $\mathbf{Q}_k$ for user k has the form of $\mathbf{Q}_k^m = \mathbf{U}_{\mathbf{\Phi}_k} \mathbf{\Lambda}_{\mathbf{Q}_k} \mathbf{U}_{\mathbf{\Phi}_k}^H$, for all users [105]. This means that each user transmits along the directions of its own channel covariance matrix. Hence (4.3.15) can be rewritten as

$$C_{MAC} = \max_{\mathbf{Q}_k^m \succeq 0} E_{\mathbf{H}}[\log|\mathbf{A} + \sum_{i=1}^{K} \mathbf{Z}_k \mathbf{\Lambda}_{\mathbf{Q}_k} \mathbf{\Lambda}_{\mathbf{\Phi}_k} \mathbf{Z}_k^H|] \tag{4.3.17}$$

$$= \max_{\mathbf{Q}_k^m \succeq 0} E_{\mathbf{H}}[\log|\mathbf{A} + \sum_{i=1}^{K} \sum_{j=1}^{N_r} \lambda_{ij}^Q \lambda_{ij}^\Phi \mathbf{z}_{ij} \mathbf{z}_{ij}^H|] \tag{4.3.18}$$

where $\mathbf{z}_{ij}$ is the j^{th} column of $\mathbf{Z}_i$, λ_{ij}^Q and λ_{ij}^Φ are the eigenvalues of Λ_{Q_i} and Λ_{Φ_i} respectively. An iterative algorithm similar to [106] is provided to determine the optimum eigenvalues of all users. By writing the Lagrangian for (4.3.18) and using the identity proven in [105] as

$$\frac{\partial}{\partial x} \log|\mathbf{C} + x\mathbf{D}| = \text{tr}[(\mathbf{C} + x\mathbf{D})^{-1}]\mathbf{D} \tag{4.3.19}$$

The j^{th} KKT condition for user k can be obtained as

$$G_{kj}(\lambda^Q) \triangleq E[\frac{\lambda_{kj}^{\Phi}\mathbf{z}_{kj}^H\mathbf{B}_{kj}^{-1}\mathbf{z}_{kj}}{1+\lambda_{kj}^{Q}\lambda_{kj}^{\Phi}\mathbf{z}_{kj}^H\mathbf{B}_{kj}^{-1}\mathbf{z}_{kj}}] \leq \nu_k \tag{4.3.20}$$

where $\lambda^Q = [\lambda_1^Q, \cdots, \lambda_K^Q]$, $\lambda_k^Q = [\lambda_{k1}^Q, \cdots, \lambda_{kN_r}^Q]$ is the vector containing the eigenvalue of user k, and ν_k is the Lagrange multiplier corresponding to the user k, $\mathbf{B}_{kj} = \mathbf{B} - \lambda_{kj}^Q\lambda_{kj}^{\Phi}\mathbf{z}_{kj}^H\mathbf{z}_{kj}$ and $\mathbf{B} = \mathbf{A} + \sum_{i=1}^{K}\sum_{j=1}^{N_r}\lambda_{ij}^Q\lambda_{ij}^{\Phi}\mathbf{z}_{ij}\mathbf{z}_{ij}^H$. Due to the KKT conditions, if the optimum λ_{kj}^Q is non-zero, the inequalities in (4.3.20) are satisfied with equality. Furthermore, if the optimum λ_{kj}^Q is zero, the inequalities in (4.3.20) are satisfied with strict inequality. λ_{kj}^Q cannot be solved directly from (4.3.20) because of the expectation operator. Hence, both sides of (4.3.20) are multiplied by λ_{kj}^Q,

$$\lambda_{kj}^Q G_{kj}(\lambda^Q) = \lambda_{kj}^Q\nu_k \tag{4.3.21}$$

As discussed above, for all λ_{kj}^Q, (4.3.21) is satisfied with equality. Some artificial fixed points have been created while obtaining (4.3.21) from (4.3.20). By summing over all antennas for any user k, the Lagrange multiplier ν_k corresponding to the user k can be determined. Then the fixed point equations can be obtained by inserting this ν_k into (4.3.21). Note that, the fixed point equations are satisfied by the optimum power values of user k,

$$f_{kj}(\lambda^Q) \triangleq \lambda_{kj}^Q = \frac{\lambda_{kj}^Q G_{kj}(\lambda^Q)}{\sum_i \lambda_{ki}^Q G_{ki}(\lambda^Q)} \tag{4.3.22}$$

where the right hand side of (4.3.22) is defined as $f_{kj}(\lambda^Q)$. Note that, $f_{kj}(\lambda^Q)$ depends on all of the eigenvalues. The power allocation is updated using the following method,

$$\lambda_k^Q(n+1) = f_k(\lambda_1^Q(n), \lambda_2^Q(n), \cdots, \lambda_K^Q(n)) \tag{4.3.23}$$

where $f_k = [f_{k1}, \cdots, f_{kN_r}]$ is the vector valued update function of user k. By assuming that the eigenvalues of the rest of the users are fixed, this algorithm finds the optimum eigenvalues for a given user. Hence, the algorithm moves to another user, after the convergence of (4.3.23). The transmit covariance matrix for the user k is computed using the following method,

$$\mathbf{Q}_k^m = \mathbf{U}_{\mathbf{\Phi}_k} \mathbf{\Lambda}_{\mathbf{Q}_k} \mathbf{U}_{\mathbf{\Phi}_k}^H \tag{4.3.24}$$

where λ_{kj}^Q is the j^{th} diagonal element of the matrix $\mathbf{\Lambda}_{\mathbf{Q_k}}$. Once the optimal solution for the dual MAC optimization problem has been found, the MAC covariance matrix is required to map back to the BC covariance matrix. As shown in [75] the CR-MIMO-BC share the same capacity region as of the CR-MIMO-MAC. Hence the conventional BC-MAC transformation algorithm has been extended to the CR-MIMO framework as Table 4.4.

The complete algorithm to solve the problem in (4.3.7)-(4.3.8) is described now. The inner part of the problem in (4.3.7)-(4.3.8) can be rewritten as follows

$$g(\alpha, \beta) = \max_{\{\mathbf{Q}_k^b\}k=1,\cdots,K:\mathbf{Q}_k^b \succeq 0} \quad \sum_{k=1}^{K} R_k^b \tag{4.3.25}$$

$$\text{s.t.} \quad \alpha(\sum_{k=1}^{K} tr(\mathbf{Q}_k^b) - P_t) \tag{4.3.26}$$

$$+\beta E_{\mathbf{h0}}[(\sum_{k=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0] - P_I) \leq 0$$

After solving the above optimization problem, the next step is to minimize it through the auxiliary dual variables α and β as follows

$$\min g(\alpha, \beta) \quad \text{s.t. } \alpha \geq 0, \beta \geq 0, \tag{4.3.27}$$

Hence, the remaining task is to determine the optimal α and β. Extending the principle used in [75], the subgradient of $g(\alpha, \beta)$ can be written as

1) Define $\hat{\mathbf{H}}_k = \mathbf{H}_k \mathbf{A}^{-1/2}$,
$\mathbf{M}_k = \mathbf{I}_{N_r} + E_{\hat{\mathbf{H}}}[\sum_{j=k+1}^{K} \hat{\mathbf{H}}_j^H \mathbf{Q}_j^m \hat{\mathbf{H}}_j]$,
$\mathbf{N}_k = \mathbf{I}_{N_t} + E_{\hat{\mathbf{H}}}[\hat{\mathbf{H}}_j (\sum_{j=k+1}^{K} \mathbf{Q}_j^m) \hat{\mathbf{H}}_j^H]$
2) Define the effective channel as
$\breve{\mathbf{H}}_k = \mathbf{M}_k^{-1/2} \hat{\mathbf{H}}_k \mathbf{N}_k^{-1/2}$
3) Apply Singular Value Decomposition to the effective channel $\breve{\mathbf{H}}_k = \mathbf{R}_k \mathbf{D}_k \mathbf{L}_k$
4) Define a new matrix
$\mathbf{J}_k = E_{\mathbf{R}}[\mathbf{A}^{-1/2} \mathbf{M}_k^{-1/2} \mathbf{R}_k \mathbf{L}_k^H \mathbf{N}_k^{1/2}]$
5) $\mathbf{Q}_k^b = \mathbf{J}_k \mathbf{Q}_k^m \mathbf{J}_k^H$

Table 4.4. BC-MAC transformation algorithm to the CR-MIMO framework

$[P_t - \sum_{k=1}^{K} tr(\mathbf{Q}_k^b), P_I - E_{\mathbf{h0}}(\sum_{k=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0)]$. It has been shown in [110] that with a constant step size, the subgradient algorithm converges to a value that is within a small range of the optimum solution. Hence, α and β can be updated as follows

$$\alpha^{(n+1)} = \alpha^{(n)} + t(\sum_{k=1}^{K} tr(\mathbf{Q}_k^b) - P_t) \tag{4.3.28}$$

$$\beta^{(n+1)} = \beta^{(n)} + t(E_{\mathbf{h0}}[\sum_{k=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0] - P_I) \tag{4.3.29}$$

This procedure is continued until $|P_t - \sum_{k=1}^{K} tr(\mathbf{Q}_k^b)| \leq \epsilon$ and $|P_I - E_{\mathbf{h0}}[\sum_{k=1}^{K} \mathbf{h}_0^H \mathbf{Q}_k^b \mathbf{h}_0]| \leq \epsilon$ are satisfied simultaneously, where ϵ is a small positive value.

Furthermore, as $\mathbf{Q}_k^b$ is not a function of the interference channel $\mathbf{h}_0$, the primary user does not need to feedback the instantaneous channel state information but it is required to feedback only the statistical information, i.e. the covariance matrix of the channel it sees from the SNB.

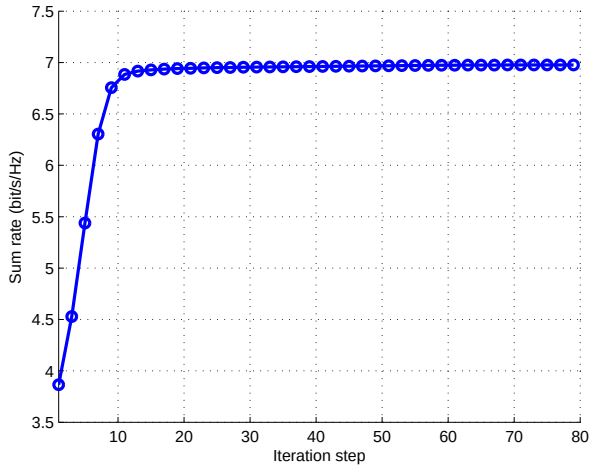

Figure 4.10. Convergence behavior of sum rate iterative algorithm ($N_t = 5$, $N_r = 3$, $K = 5$, $P_I = 1$ and $P_t = 10$).

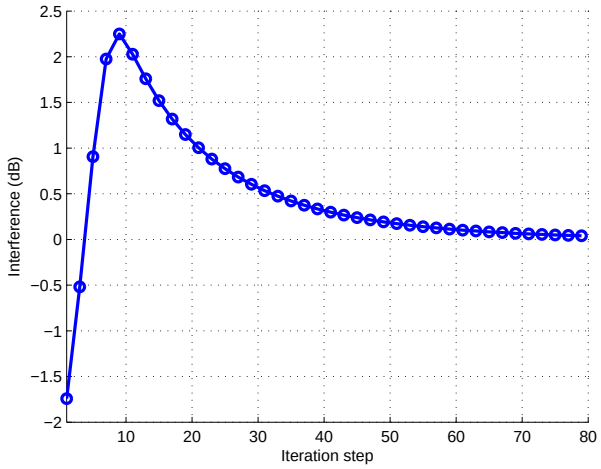

Figure 4.11. Evolution of the iterative algorithm: PU interference versus iteration number ($N_t = 5$, $N_r = 3$, $K = 5$, $P_I = 1$ and $P_t = 10$)
.

4.3.3 Simulation Results

In order to illustrate the effectiveness of the proposed algorithm, several simulation results have been provided. The interference channel components of $\mathbf{h}_0$ are assumed to be CSCG variables with zero mean and unitary variance. The total power constraint P_t and the interference power threshold P_I are set to 10 and 1 respectively. The noise variances at the SUs are set to unity.

A CR-BC-MIMO network under Rayleigh fading with partial CSI with

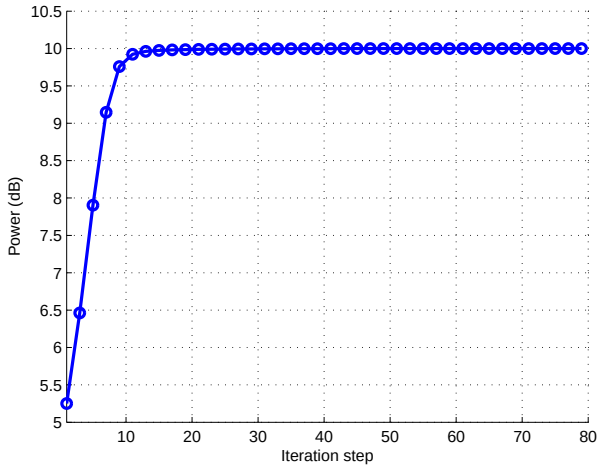

Figure 4.12. Evolution of the proposed iterative algorithm: The power at SNB reaches the maximum limit of 10 after convergence ($N_t = 5$, $N_r = 3$, $K = 5$, $P_I = 1$ and $P_t = 10$).

five SUs and one PU is considered. Five transmit antennas and three receive antennas have been chosen in this simulation. All the SUs are assumed to share the same spectrum as of PU. Fig. 4.10 depicts the sum rate versus the number of iterations of the proposed algorithm for a step size $t = 0.01$. It can be seen from the figure that the sum rate approaches to 7 at the convergence of the algorithm. Fig. 4.11 depicts the interference power at the PU versus the number of iterations. As seen, the interference power approaches the threshold of 1 at approximately 80 iterations. Fig. 4.12 depicts the power consumption at the SNB. It can be seen from the figure that the power consumption approaches the maximum limit of 10 at the convergence of the algorithm.

4.4 An Optimal Resource Allocation Technique for Spectrum Sharing MIMO Wireless Relay Network

Relay communications have recently emerged as a powerful spatial diversity technique that can improve the performance over the conventional point-to-

point transmissions, such as throughput enhancement, coverage extension, and power saving. An important work on relay communications was performed by Cover and Gamal [112]. Combining relay technique and cognitive radio techniques, cognitive radio relaying has attracted a great deal of attention lately [113–115]. In this section, a weighted sum rate maximization and rate balancing problem for a spectrum sharing MIMO based wireless relay network has been investigated. The aim is to maximize the sum rate of the wireless relay network whilst ensuring the interference leakage to the primary user terminals during two time slots are below a specific value. This problem has been solved by asymmetrically allocating the power to different time slots and using the principle of BC-MAC duality. The algorithm in its dual form has been solved using sub-gradient methods. The simulation results demonstrate the convergence of the algorithm and the simultaneous satisfaction of maximum power and the interference constraints.

4.4.1 System Model

A downlink three-node spectrum sharing MIMO based wireless relay network has been considered, where a source, which is a secondary network basestation, communicates with the destinations (SUs) through a relay station, as shown in Fig. 4.13. Each node in this network cannot transmit and receive at the same time. The SNB and the relay station (RS) are equipped with N_t antennas. Each SU $k \in \{1, 2, \ldots, K\}$ has N_r antennas. The K SUs share the same spectrum as of the PUs. Each PU is equipped with one receive antenna. The downlink transmission between the SNB and the SUs in the relay network consists of two time slots. In the first time slot, the source nod sends the data to the relay which is considered as a downlink problem. The relay receives and decodes the signal. In the second time slot, the relay re-encodes the signal and forwards it to the destination (SUs), which is also a downlink resource allocation problem. Finally, the destination decodes the

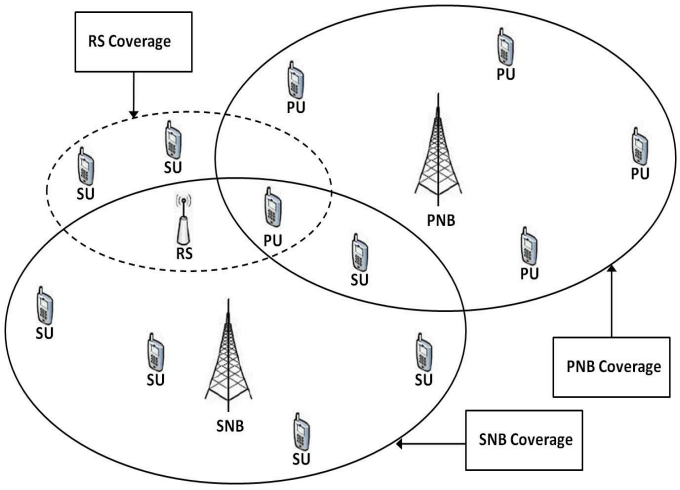

Figure 4.13. Wireless cognitive relay network system model (SNB: secondary network basestation; PNB: primary network basestation; RS: relay station)

signal from the relay. In this section, it is assumed that transmission durations for the consecutive time slots are symmetric. The total transmission duration from the source to the destination is the summation of two time slots. The total transmit power in the relay network is P_T. During the two time slots, both the SNB and RS introduce interference to PUs. For simplicity, one PU is considered in the network, as shown in Fig. 4.14, but extension to multiple PUs is straight forward.

Assuming dirty paper coding at the RS transmitter and the encoding order is $(1, \cdots, K)$, i.e., the codeword of user 1 is encoded first, the data rate R_k^B (from relay to k^{th} destination) for the k^{th} user can be written as [69]

$$R_k^B = \log \frac{|\mathbf{I}_{N_r} + \sum_{i=k}^{K} \mathbf{H}_i \mathbf{Q}_i^B \mathbf{H}_i^H|}{|\mathbf{I}_{N_r} + \sum_{i=k+1}^{K} \mathbf{H}_i \mathbf{Q}_i^B \mathbf{H}_i^H|}, \tag{4.4.1}$$

where $\mathbf{Q}_k^B = E\{\mathbf{x}_k \mathbf{x}_k^H\}$ is the transmit covariance matrix for the k^{th} SU and $\mathbf{x}_k$ is the signal transmitted by the k^{th} user in the downlink at the RS. $\mathbf{H}_k \in C^{N_r \times N_t}$ is the channel matrix from RS to the k^{th} user. The capacity

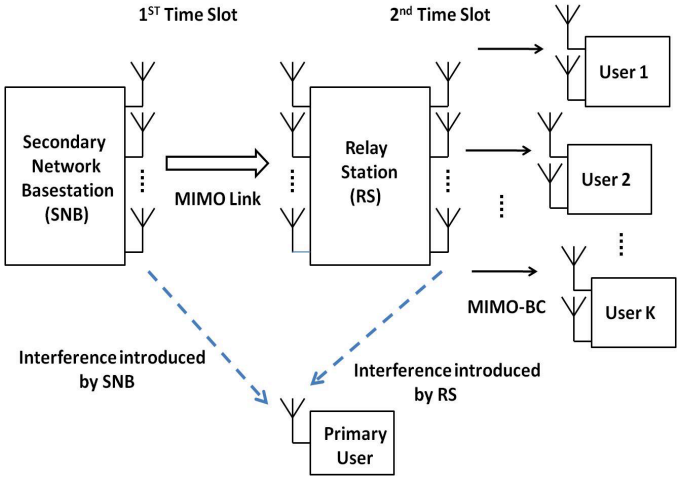

Figure 4.14. The system model for the K-SU MIMO cognitive relay system having N_t transmit antennas and N_r receive antennas at each SU.

achieved at the RS (from SNR) is defined as

$$R_{RS} = \log |\mathbf{I}_{N_t} + \mathbf{H}_0 \mathbf{S}_0^B \mathbf{H}_0^H| \tag{4.4.2}$$

where $\mathbf{S}_0^B = E\{\mathbf{s}_0 \mathbf{s}_0^H\}$ is the transmit covariance matrix at the SNB and $\mathbf{s}_0$ is the signal transmitted to RS. $\mathbf{H}_0 \in C^{N_t \times N_t}$ is the channel matrix from SNB to RS. Hence, the capacities of the two consecutive time slots are defined as follows

$$C_{1st} = R_{RS} = \log |\mathbf{I}_{N_t} + \mathbf{H}_0 \mathbf{S}_0^B \mathbf{H}_0^H| \tag{4.4.3}$$

$$C_{2nd} = \sum_{k=1}^{K} \log \frac{|\mathbf{I}_{N_r} + \sum_{i=k}^{K} \mathbf{H}_i \mathbf{Q}_i^B \mathbf{H}_i^H|}{|\mathbf{I}_{N_r} + \sum_{i=k+1}^{K} \mathbf{H}_i \mathbf{Q}_i^B \mathbf{H}_i^H|} \tag{4.4.4}$$

4.4.2 Resource Allocation Scheme for Spectrum Sharing MIMO Based Wireless Relay Systems

A. Problem Formulation

It is assumed that the data received by the RS are decoded and retransmitted to the destinations (users). Hence, the capacity from the source to the destination is $C = \min\{C_{1st}, C_{2nd}\}$, where $C = \min\{x, y\}$ takes the smaller

value of x and y. The objective function of the symmetric resource allocation problem is formulated as follows

$$\max_{\{\mathbf{S}_0^B\}\mathbf{S}_0^B\succeq 0;\ \{\mathbf{Q}_k^B\}k=1,\cdots,K:\mathbf{Q}_k^B\succeq 0} (\min\{C_{1st}, C_{2nd}\}) \tag{4.4.5}$$

subject to

$$\mathrm{tr}(\mathbf{S}_0^B) + \sum_{k=1}^{K} \mathrm{tr}(\mathbf{Q}_k^B) \leq P_t, \tag{4.4.6}$$

$$\mathbf{h}_1^H \mathbf{S}_0^B \mathbf{h}_1 \leq P_I, \tag{4.4.7}$$

$$\sum_{k=1}^{K} \mathbf{h}_2^H \mathbf{Q}_k^B \mathbf{h}_2 \leq P_I, \tag{4.4.8}$$

where $\mathbf{h}_1 \in C^{N_t \times 1}$ denotes the channel gain vector between the SNB and the PU in the first time slot, $\mathbf{h}_2 \in C^{N_t \times 1}$ denotes the channel gain vector between the RS and the PU in the second time slot. A sum power constraint P_t is used to control the SNB and RS transmission power. P_I denotes the interference threshold for the primary receiver. By solving the problem in (4.4.5)-(4.4.8), joint source-relay power allocation is obtained. It is straightforward to derive from (4.4.5) that C is maximized only when C_{1st} in (4.4.3) is equal to C_{2nd} in (4.4.4). Therefore, the optimization problem becomes

$$\max_{\{\mathbf{S}_0^B\}\mathbf{S}_0^B\succeq 0;\ \{\mathbf{Q}_k^B\}k=1,\cdots,K:\mathbf{Q}_k^B\succeq 0} C_{1st} + C_{2nd} \tag{4.4.9}$$

subject to (4.4.6)-(4.4.8) and

$$C_{1st} = C_{2nd} \tag{4.4.10}$$

The constraint in (4.4.10) is a non-convex constraint [93]. This is because the constraint in (4.4.10) can be written as $C_{1st} - C_{2nd} = 0$, and minus sign result into loss of convexity. Hence, in the next section, a sum rate maximization and rate balancing technique has been proposed to solve this

resource allocation problem using a different approach.

B. Algorithm Description

Since (4.4.10) is not a convex set, this constraint can be relaxed and a new variable P_1 is introduced which denotes the power allocation in the first time slot. Hence, the capacity between the two time slots can be balanced and $C_{1st} - C_{2nd} = 0$ can be achieved by using an appropriate power level P_1. As a result, the problem in (4.4.5)-(4.4.8) can be modified as follows

$$\max_{\{\mathbf{S}_0^B\}\mathbf{S}_0^B \succeq 0;\ \{\mathbf{Q}_k^B\}k=1,\cdots,K:\mathbf{Q}_k^B \succeq 0,\ P_1} C_{1st} + C_{2nd} \tag{4.4.11}$$

subject to

$$\mathrm{tr}(\mathbf{S}_0^B) \leq P_1, \tag{4.4.12}$$

$$\sum_{k=1}^{K} \mathrm{tr}(\mathbf{Q}_k^B) \leq P_t - P_1, \tag{4.4.13}$$

$$\mathbf{h}_1^H \mathbf{S}_0^B \mathbf{h}_1 \leq P_I, \tag{4.4.14}$$

$$\sum_{k=1}^{K} \mathbf{h}_2^H \mathbf{Q}_k^B \mathbf{h}_2 \leq P_I, \tag{4.4.15}$$

Let us explain the procedure of solving the above problem. First, $P_1^{(n)}$ is initialized to an arbitrary positive value. The problem in (4.4.11)-(4.4.15) is solved for optimal value of $\mathbf{S}_0^B$ and $\mathbf{Q}_k^B$. The capacity achieved during the two time slots C_{1st} and C_{2nd} are then compared. The value of $P_1^{(n)}$ is updated as $P_1^{(n+1)} = P_1^{(n)} - t(C_{1st} - C_{2nd})$, where t is a small positive step size. This process will continue until the $|C_{1st} - C_{2nd}| \leq \epsilon$, where ϵ is a small positive value.

Applying the duality between Gaussian MAC and Gaussian BC as in [1] and [69], the general MIMO sum rate optimization problem can be solved. However, due to additional interference constraint for the primary users (4.4.14) and (4.4.15), the optimization problem in (4.4.11) is non-convex [93].

Similar to [75], by introducing auxiliary dual variables for the interference power constraint and the sum power constraint, the original problem in (4.4.11)-(4.4.15) is transformed to the following equivalent problem

$$\min_{\alpha\geq 0,\beta\geq 0,\gamma\geq 0,\delta\geq 0} \max_{\{\mathbf{S}_0^B\}\mathbf{S}_0^B\succeq 0;\{\mathbf{Q}_k^B\}k=1,\cdots,K:\mathbf{Q}_k^B\succeq 0} C_{1st}+C_{2nd} \tag{4.4.16}$$

subject to

$$\alpha \mathrm{tr}(\mathbf{S}_0^B)+\beta\sum_{k=1}^{K}\mathrm{tr}(\mathbf{Q}_k^B)+\gamma(\mathbf{h}_1^H\mathbf{S}_0^B\mathbf{h}_1)+\delta(\sum_{k=1}^{K}\mathbf{h}_2^H\mathbf{Q}_k^B\mathbf{h}_2) \leq \alpha P_1+\beta(P_t-P_1)+\gamma P_I+\delta P_I, \tag{4.4.17}$$

where α, β, γ and δ are the auxiliary dual variables for the interference power constraints and the power constraint in two time slots respectively. Constraint (4.4.17) is still non-convex due to the multiplication of the dual variables with the optimization variables $\mathbf{S}_0$ and $\mathbf{Q}_k$. However, if α, β, γ and δ are set as constants, the constraint in (4.4.17) becomes convex. Hence, α, β, γ and δ are set as constants and write both constraints into a single constraint. By applying the BC-MAC duality principle as stated in [75], the problem shown in (4.4.16) and (4.4.17) is transformed to the dual MAC problem as follows

$$\max_{\{\mathbf{S}_0^M\}\mathbf{S}_0^M\succeq 0;\{\mathbf{Q}_k^M\}k=1,\cdots,K:\mathbf{Q}_k^M\succeq 0} C_{1st}^M+C_{2nd}^M \tag{4.4.18}$$

$$\text{s.t.}\{tr(\mathbf{S}_0^M)+\sum_{k=1}^{K}tr(\mathbf{Q}_k^M)\}\sigma^2 \leq \alpha P_1+\beta(P_t-P_1)+\gamma P_I+\delta P_I, \tag{4.4.19}$$

where $\mathbf{S}_0^M \in C^{N_t\times N_t}$ and $\mathbf{Q}_k^M \in C^{N_r\times N_r}$ denote the transmit covariance matrix by the RS and the k^{th} SU for the MAC respectively. C_{1st}^M and C_{2nd}^M are the rate achieved during the first and the second time slots for the MAC respectively. σ^2 is the noise variance for the RS and all the SUs. Due to the

duality between the MIMO-BC and the MIMO-MAC, the optimal encoding order for the MIMO-BC employing DPC is the reverse of the decoding order for the MIMO-MAC using successive interference cancelation scheme. Hence R_{RS}^{M} and R_{k}^{M} are written as

$$R_{RS}^{M} = \log|\mathbf{A}_1 + \mathbf{H}_0^H \mathbf{S}_0^M \mathbf{H}_0| \tag{4.4.20}$$

$$R_{k}^{M} = \log \frac{|\mathbf{A}_2 + \sum_{i=1}^{k} \mathbf{H}_i^H \mathbf{Q}_i^M \mathbf{H}_i|}{|\mathbf{A}_2 + \sum_{i=1}^{k-1} \mathbf{H}_i^H \mathbf{Q}_i^M \mathbf{H}_i|}, \tag{4.4.21}$$

where the matrix $\mathbf{A}_1 \in C^{N_t \times N_t}$ is defined as $\mathbf{A}_1 = \gamma \mathbf{h}_1 \mathbf{h}_1^H + \alpha \mathbf{I}_{N_t}$ and $\mathbf{A}_2 \in C^{N_t \times N_t}$ is defined as $\mathbf{A}_2 = \delta \mathbf{h}_2 \mathbf{h}_2^H + \beta \mathbf{I}_{N_t}$ [107]. Hence the problem in (4.4.16)-(4.4.17) and its dual form (4.4.18)-(4.4.19) can be solved using the subgradient algorithm and the CR-BC-MAC duality described in [75].

For the case that PU has multiple antennas N_{pr}, the algorithm can be extended by writing the primary user channel in the matrix form as $\mathbf{H}_1 \in C^{N_{pr} \times N_t}$, $\mathbf{H}_2 \in C^{N_{pr} \times N_t}$, and the interference constraint in (4.4.17) as $\alpha \mathrm{tr}(\mathbf{S}_0) + \beta \sum_{k=1}^{K} \mathrm{tr}(\mathbf{Q}_k) + \gamma \ tr(\mathbf{H}_1^H \mathbf{S}_0 \mathbf{H}_1) + \delta \ tr(\sum_{k=1}^{K} \mathbf{H}_2^H \mathbf{Q}_k \mathbf{H}_2) \leq \alpha P_1 + \beta(P_t - P_1) + \gamma P_I + \delta P_I$. The algorithm presented here will remain the same afterwards.

The algorithm can be extended to multiple PUs by introducing additional auxiliary variables β_m for each primary user interference constraint and modifying (4.4.17) as $\alpha \mathrm{tr}(\mathbf{S}_0) + \beta \sum_{k=1}^{K} \mathrm{tr}(\mathbf{Q}_k) + \sum_{m=1}^{M} \gamma_m (\mathbf{h}_{1,m}^H \mathbf{S}_0 \mathbf{h}_{1,m}) + \sum_{m=1}^{M} \delta_m (\sum_{k=1}^{K} \mathbf{h}_{2,m}^H \mathbf{Q}_k \mathbf{h}_{2,m}) \leq \alpha P_1 + \beta(P_t - P_1) + \sum_{m=1}^{M} \gamma_m P_{I,m} + \sum_{m=1}^{M} \delta_m$ $P_{I,m}$, where $\mathbf{h}_{1,m}$ is the channel response from the SNB to the m^{th} PU, $\mathbf{h}_{2,m}$ is the channel response from the RS to the m^{th} PU and $P_{I,m}$ is the interference threshold for the m^{th} PU. The role of the auxiliary variables γ_m and δ_m is similar to that of γ and δ for the case of single PU. Hence the proposed algorithm can be modified to solve this problem.

In this section, the power constraint is on the total base station and relay power. In this case, it is possible to equate the data rates from the

base station to the relay and from the relay to the destination. However, for the case when the base station and the relay have individual maximum power constraints, the proposed resource allocation algorithm will still work, but the overall data rate should be expected to be lower than that for the case of combined power constraint. This is due to the inability to maintain equal data rate from base station and the relay.

C. Rate Balancing

Suppose the achievable rate region C_{2nd} in the second time slot for the problem in (4.4.5)-(4.4.8) is $\mathbf{r} \in C(\mathbf{H}, P_t, P_I)$. The technique investigated so far aims to maximize the sum rate with interference constraint. To ensure fairness among secondary users, a better criterion is to maximize the data rate while balancing rate achieved for each SU as [101]

$$\max_{\mathbf{r},\lambda} \lambda \quad s.t. \;\; \mathbf{r} = \lambda\boldsymbol{\rho}, \quad \mathbf{r} \in C(\mathbf{H}, P_t, P_I) \tag{4.4.22}$$

For example if $\boldsymbol{\rho} = \mathbf{1}_K$, all users attain the same data rate. For other values of $\boldsymbol{\rho}$, data rate is maximized while dividing the total rate to users according to the ratio defined by the vector $\boldsymbol{\rho}$. Fortunately, the dual problem can be shown as a weighted sum rate optimization of the following form [101]

$$\min_{\mu} \;\; \max_{\mathbf{r}} \;\; \sum_{k=1}^{K} \mu_k \frac{R_k}{\rho_k} \quad s.t. \;\; \sum_{k=1}^{K} \mu_k = 1 \tag{4.4.23}$$

where μ_k are the Lagrangian coefficients for the K constraints in (4.4.22), $\mathbf{r} = [R_1 \cdots R_K]$ is the rate vector and ρ_k is the k^{th} element of $\boldsymbol{\rho}$. The capacity in the second time slot is now modified as follows

$$C_{2nd} = \sum_{k=1}^{K} \mu_k \log \frac{|\mathbf{I}_{N_r} + \sum_{i=k}^{K} \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^H|}{|\mathbf{I}_{N_r} + \sum_{i=k+1}^{K} \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^H|} \tag{4.4.24}$$

Hence, for an initial setting of μ_k, $\max \sum_{k=1}^{K} (\frac{\mu_k}{\rho_k}) R_k$ can be solved using the method described earlier, and the μ_k can be updated using a subgradient method as in [101], i.e. $\mu_k^{(d)} = \mu_k^{(d-1)} - t(R_k - R_K)$. For the updated μ_k, the weighted sum rate problem is solved again and μ_k is computed again.

1) Initialization: P_1^n to positive values, $n = 1$
2) **repeat**
3) Initialization: $\mu_1^d \cdots \mu_K^d$ to positive values, $d = 1$,
4) **repeat**
5) Solve the problem in (4.4.16)-(4.4.17) using the subgradient method in [75],
6) Update the Lagrangian coefficient μ_k^d for the problem in (4.4.23), $d = d + 1$,
7) **Stop** when $\lvert R_k - R_{k-1}\rvert \leq \epsilon$ for $k = 2, \cdots, K$,
8) **Stop** when $\lvert C_{1st} - C_{2nd}\rvert \leq \epsilon$.

Table 4.5. Iterative rate maximization and balancing algorithm

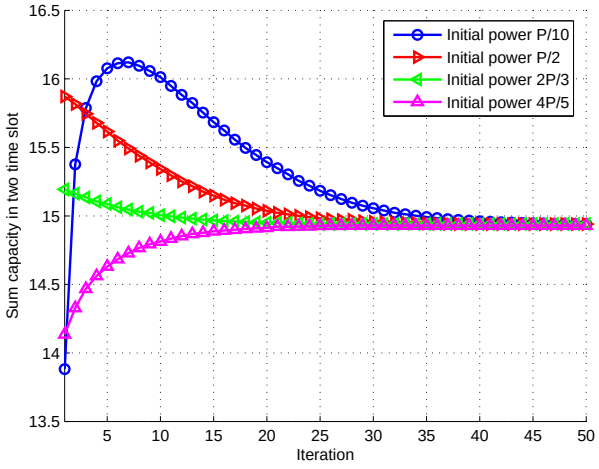

Figure 4.15. The convergence of the proposed algorithm ($N_t = 5$, $N_r = 3$, $K = 3$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

This is continued until convergence. The resulting algorithm will maximize the data rate while ensuring data rate balancing and interference leakage to primary user is below a target value. The proposed iterative algorithm for resource allocation of spectrum sharing MIMO based wireless relay network has been summarized as in Table 4.5.

4.4.3 Simulation Results

The elements of the channel matrix $\mathbf{H}_0$ and $\mathbf{H}_k$ and the interference channel vector $\mathbf{h}_1$ and $\mathbf{h}_2$ are assumed to be CSCG variables with zero mean and

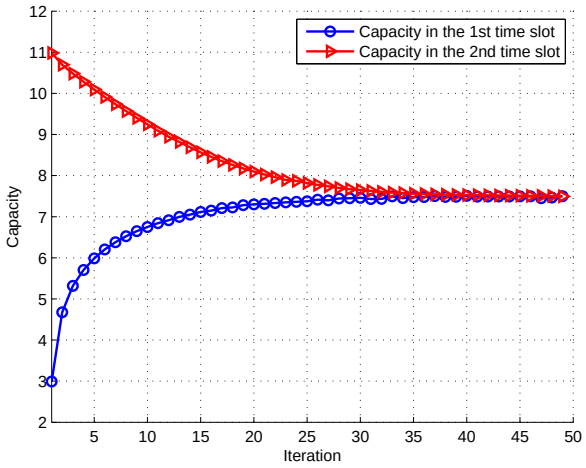

Figure 4.16. Comparison of the optimal achievable rates obtained by the proposed algorithm for different time slots ($N_t = 5$, $N_r = 3$, $K = 3$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

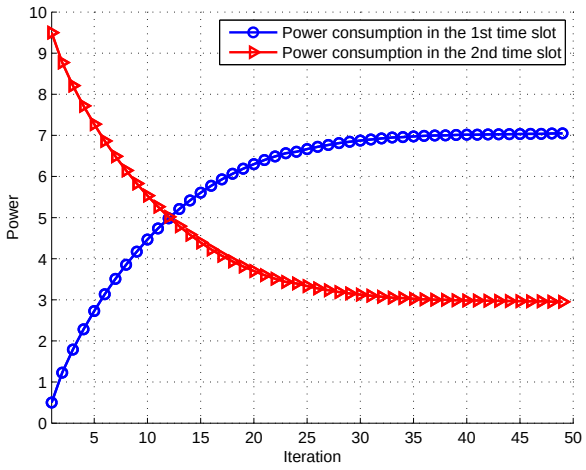

Figure 4.17. Comparison of the power allocation obtained by the proposed algorithm for different time slots ($N_t = 5$, $N_r = 3$, $K = 3$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

unity variance. The total power constraint P_t and the interference power threshold P_I are set to 10 and 1 respectively. The noise variances at the SUs have been set to unity.

A MIMO cognitive relay network with three SUs and one PU has been considered. Five antennas at the SNB and RS, three receive antennas at each SU has been chosen in this simulation. In the first simulation, the aim is to

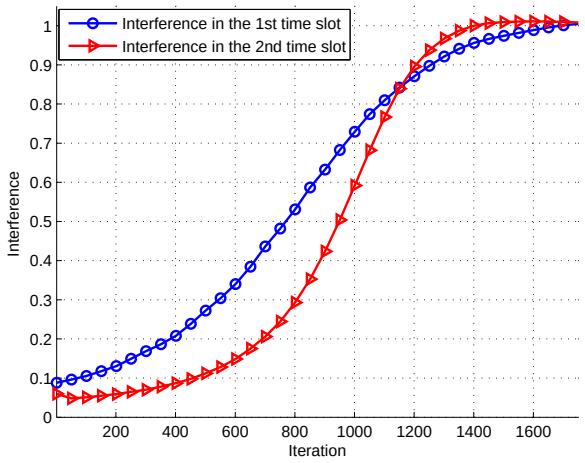

Figure 4.18. Comparison of the optimal achievable rates obtained by the proposed algorithm for different time slot ($N_t = 5$, $N_r = 3$, $K = 3$, $P_I = 1$, $P_t = 10$ and $\mu_i = 1$ for all i).

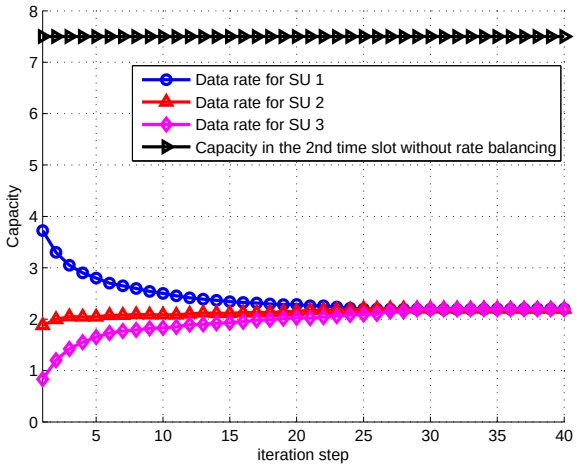

Figure 4.19. Rate balancing of the proposed algorithm for the case of three SU ($N_t = 5$, $N_r = 3$, $K = 3$, $P_I = 1$ and $P_t = 10$)

show the proposed algorithm converges to an identical setting regardless of the initializations in term of P_1 for a particular set of random channels. The initial values for the power allocation are varied in the first and the second time slots as ($P_1 = \frac{1}{10}P_t$, $P_1 = \frac{9}{10}P_t$), ($P_1 = \frac{1}{2}P_t$, $P_1 = \frac{1}{2}P_t$), ($P_1 = \frac{2}{3}P_t$, $P_1 = \frac{1}{3}P_t$) and ($P_1 = \frac{4}{5}P_t$, $P_1 = \frac{1}{5}P_t$). As seen in Fig. 4.15, the sum capacity of the two time slots approaches to 15 regardless of the initializations.

Fig. 4.16 depicts the capacity of both time slots against the number of iterations. As seen, after approximately 50 iteration, the capacity of both time slot approaches to 7.5 bits. Fig. 4.17 depicts the power consumption in both time slots against the number of iterations. As seen, the power consumption in the first time slot is approximately 7 whilst it is 3 in the second time slot. Fig. 4.18 examines the validity of the interference control of the proposed algorithm. As seen, the interference power in both the time slots approaches to $P_I = 1$ when the proposed algorithm converges.

Finally, the proposed algorithm can balance all the SUs data rate has been shown. All the elements of the data rate balancing vector $\boldsymbol{\rho}$ were set to one. Fig. 4.19 depicts the convergence of the data rate for all three users against the adaptation of the Lagrangian multiplier μ_k as explained in Section C. All the users attain equal data rate. The data rate without rate balancing constraints is also shown. The total sum rate in this case is 7.5 bits/s/Hz. With the rate balancing constraints, each user attains 2.2 bits/s/Hz. Hence the total rate is less than that of the scheme that does not use rate balancing constraint. However the rate balancing constraint ensures fairness among users.

4.5 Conclusion

The sum rate maximization problem for spectrum sharing networks has been studied. Weighted sum rate maximization and rate balancing techniques for MIMO-OFDM based spectrum sharing broadcast channels have been proposed first. The aim was to maximize the sum rate of the secondary users whilst ensuring interference leakage to the primary user terminals is below a specific value and each secondary user attains a specific portion of the total data rate. This problem has been solved by converting the MIMO-OFDM channel into a block diagonal form and using the principle of BC-MAC du-

ality. The algorithm in its dual form has been solved using sub-gradient methods. The simulation results demonstrated the convergence of the algorithm and the simultaneous satisfaction of maximum power and interference constraints. Fading channel and the ergodic capacity region maximization problem is another interesting research area. Hence, in contrast to previous section on MIMO CR networks, the ergodic capacity region has been studied based on fading channels. The sum rate maximization problem for spectrum sharing MIMO BC under Rayleigh fading has then been investigated. Based on the multiple auxiliary variables, the KKT optimality conditions and the BC-MAC duality, an iterative algorithm has been developed to solve the equivalent problem using the Lagrangian theory and showed that the proposed algorithm converges to the setting as defined by the optimization problem, i.e., the interference and the power limit are met whilst maximizing the sum capacity. Finally, to further exploit the performance of the network, a weighted sum rate maximization and rate balancing problem for a spectrum sharing MIMO based wireless relay network has been studied. The aim is to maximize the sum rate of the wireless relay network whilst ensuring the interference leakage to the primary user terminals during two time slots are below a specific value. This problem has been solved by asymmetrically allocating the power to different time slots and using the principle of BC-MAC duality. The algorithm in its dual form has been solved using sub-gradient methods. The simulation results demonstrated the convergence of the algorithm and the simultaneous satisfaction of maximum power and interference constraints.

Chapter 5

BEAMFORMING AND TEMPORAL POWER OPTIMIZATION FOR SPECTRUM SHARING NETWORKS

This chapter focuses on beamforming and power control for an overlay cognitive radio network. Co-existence of a secondary network with a primary network under an overlay framework has been considered. Beamformer design and power allocation approach has been proposed using an iterative optimization technique. The secondary network serves multiple users in the same frequency band as of the primary network, however, in order to compensate the interference leakage to the primary user terminals, the secondary network acts as a relay to forward the primary user signals. The interference and noise level at the primary terminals during various time slots are different, therefore, the primary network needs to allocate resources asymmetrically during various time slots for the optimal performance. Hence, a joint spatial and temporal resource allocation technique has been proposed

to enhance the overall system power saving while satisfying the data rate or the SINR requirement of the primary and the secondary users.

5.1 Introduction

The spectrum efficiency can be enhanced further by using spatial diversity techniques [23, 116, 117] and user cooperation diversity [118] and [119]. Communications based on user cooperation, namely cooperative communication, exploit the spatial diversity of multiuser systems without a need for using multiple antennas at the user terminals [120]. In cooperative communications, users relay messages to each other thereby providing multiple paths from sources to the destinations. Combining cooperations and cognitive radio techniques, cognitive radio relaying has attracted a great deal of attention lately [113–115]. Distributed beamforming has been recently proposed under a relay network scenario in [121–128]. Mitigation of interference is very important in cognitive radio networks to ensure primary users are not harmfully affected by the secondary user transmissions [129]. Beamforming can be used to mitigate interference leakage towards the primary users [75, 107, 130–132]. In [58], a SINR balancing beamforming technique for a CRN was proposed to maximize the worst case SU SINR while ensuring the interference leakage to the PUs is below a specific threshold.

There are three fundamental operational modes for the CRN, namely, interweave, overlay, and underlay methods [22]. The interweave method is known as opportunistic spectrum access, whereby the SU transmits over the spectrum which was originally allocated to PU when the PU transmission is detected to be off. The overlay and the underlay methods allow the SU to operate simultaneously with the PU. The underlay method regulates the interference power level at the PU receiver introduced by the SU transmission to be below a specific threshold which is known as "interference tempera-

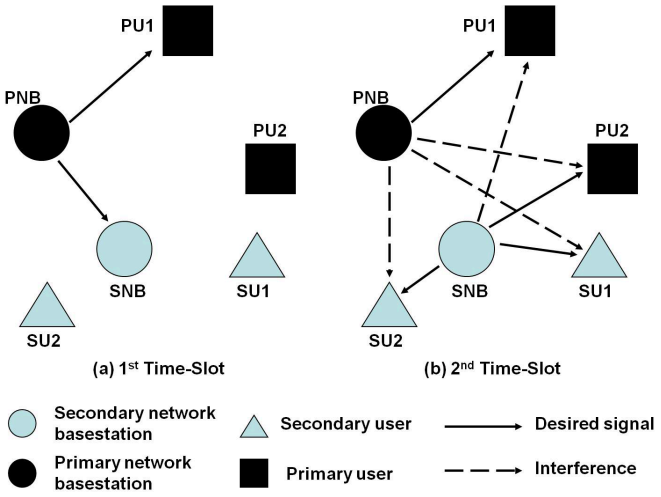

Figure 5.1. System model for the cognitive radio relay network (K=2).

ture". In the overlay method, the SU network helps the PU network to offset the interference caused by the SU transmission by assisting and relaying PU signals. Most of the published works on interference mitigation using beamforming considered only an underlay cognitive radio network [58, 75, 107]. In this section, beamforming and power control for an overlay cognitive radio network has been considered. The proposed algorithm minimizes the total power consumed by the network while satisfying each user's SINR requirement. An iterative algorithm based on second order cone programming has been proposed to greatly improve the overall power saving of the network.

5.2 System Model and Problem Statement

Primary and secondary downlink networks with two PUs, K SUs, one PNB and one SNB has been considered as shown in Fig. 5.1 (for the case K=2). Each of the PUs and SUs is equipped with a single antenna. The PNB and the SNB are equipped with multiple antennas. The SU network uses the licensed frequency of the PU for transmission of signals to its users. However, in order to compensate the interference leakage to the primary network, the SNB acts as a relay to assist PU transmission. Our aim is

to minimize the total power consumed by the network while satisfying each user's SINR requirement. It is assumed that both the PNB and the SNB have the perfect channel state information of all the channels between their basestations and the SUs as well as the PUs. Such information is exploited to determine the optimal beamforming as well as the optimum transmit power levels for the PUs and the SUs. It is assumed that the channels remain unchanged for a number of frames to allow feedback based spatial multiplexing and interference control. PU 1 is within the coverage region of PNB. It is assumed that PU 2 is out of the coverage region of PNB, thus its channel gain from the PNB is very low. As a result, the PNB will require very high power to transmit signals to PU 2 in the absence of cooperations from SNB. It is assumed that PU 2 is within the coverage region of SNB, and the SNB is able to help the PNB to transmit signal to PU 2. Therefore, SNB will act as a relay station for the primary network. This relaying requires multiple time slots. In this chapter, two time slots are employed as an example, however, extension to more than two time slots is straightforward.

During the first time slot, PNB transmits signals $s_{p1}(n)$ and $s_{p2}(n)$ to PU 1 and SNB respectively. $s_{p2}(n)$ is intended for the PU 2 through SNB. $\mathbb{E}\{|s_{p1}(n)|^2\} = \mathbb{E}\{|s_{p2}(n)|^2\} = 1$. The channels from PNB to PU 1 and SNB are denoted by $\mathbf{g}_1 \in C^{M_p \times 1}$ and $\mathbf{g}_{ps} \in C^{M_p \times 1}$ respectively, where M_p is the number of antennas at the PNB. Let $\mathbf{v}_{p1,1} \in C^{M_p \times 1}$ and $\mathbf{v}_{ps} \in C^{M_p \times 1}$ denote the beamforming vectors for the signals transmitted to PU 1 and SNB respectively in the first time slot. The received signals at the PU 1 and the SNB in the first time slot are given by

$$y_{p1,1}(n) = \mathbf{v}_{p1,1}^H \mathbf{g}_1 s_{p1}(n) + \mathbf{v}_{ps} \mathbf{g}_1 s_{p2}(n) + q_1(n), \tag{5.2.1}$$

$$y_{BS}(n) = \mathbf{v}_{ps}^H \mathbf{g}_{ps} s_{p1}(n) + \mathbf{v}_{1,1} \mathbf{g}_{ps} s_{p2}(n) + q_2(n), \tag{5.2.2}$$

where $q_1(n)$ and $q_2(n)$ are the additive white Gaussian noise components with variance σ_n^2. Hence, the received SINRs at the PU 1 and the SNB on the first time slot are given respectively as

$$\gamma_{p1,1} = \frac{\mathbf{v}_{p1,1}^H \mathbf{R}_{p1} \mathbf{v}_{p1,1}}{\mathbf{v}_{ps}^H \mathbf{R}_{p1} \mathbf{v}_{ps} + \sigma_n^2}, \tag{5.2.3}$$

$$\gamma_{BS} = \frac{\mathbf{v}_{ps}^H \mathbf{R}_{ps} \mathbf{v}_{ps}}{\mathbf{v}_{p1,1}^H \mathbf{R}_{ps} \mathbf{v}_{p1,1} + \sigma_n^2}, \tag{5.2.4}$$

where $\mathbf{R}_{p1} = \mathbb{E}\{\mathbf{g}_1 \mathbf{g}_1^H\}$, $\mathbf{R}_{ps} = \mathbb{E}\{\mathbf{g}_{ps} \mathbf{g}_{ps}^H\}$. The channel covariance matrices $\mathbf{R}_{p1}$ and $\mathbf{R}_{ps}$ can be of any arbitrary rank. However, for pencil beams $\mathbf{R}_{p1}$ and $\mathbf{R}_{ps}$ are of rank one. SNB needs to decode PU 1's data successfully, otherwise the relay (SNB) will not be able to forward the message to PU 2 in the second time slot. As a result, it is required to set the SINR target high enough to successfully decode the PNB's signal during the first time slot. The SINR target for SNB during the first time slot is denoted as δ_{BS}, thus this leads to the following constraint

$$\frac{\mathbf{v}_{ps}^H \mathbf{R}_{ps} \mathbf{v}_{ps}}{\mathbf{v}_{p1,1}^H \mathbf{R}_{ps} \mathbf{v}_{p1,1} + \sigma_n^2} \geq \delta_{BS}, \tag{5.2.5}$$

Power budget $P_{p,1}$ for the primary network at first time slot leads to another constraint as

$$\|\mathbf{v}_{p1,1}\|_2^2 + \|\mathbf{v}_{ps}\|_2^2 \leq P_{p,1}, \tag{5.2.6}$$

During the second time slot, PNB will continue to transmit data $s_{p1,2}(n)$ to PU 1, $\mathbb{E}\{|s_{p1,2}(n)|^2\} = 1$. The SNB will relay the signal it received from PNB during the first time slot to the PU 2. In addition, SNB should also transmit its own data to SUs. Therefore PU 2 joins in the secondary network and receive the data from SNB. PU 2 can be treated as the $(K+1)^{th}$ secondary user. Hence the SNB will transmit data $s_k(n)(k = 1, 2, \cdots K+1)$ to all the $K+1$ users, $\mathbb{E}\{|s_k(n)|^2\} = 1(k = 1, 2, \cdots K+1)$. In the second time

slot, mutual interference is introduced between the primary users and the secondary users. The channel between the SNB and the k^{th} SU is denoted by $\mathbf{h}_k \in C^{M_s \times 1}$, $k = 1, 2, \ldots, K+1$, where M_s is the number of antennas at the SNB. The interference channel between the PNB and the k^{th} SU is denoted as $\mathbf{g}_{sk} \in C^{M_p \times 1}$, $k = 1, 2, \ldots, K+1$. The interference channel between the SNB and the PU 1 is denoted as $\mathbf{h}_{p1} \in C^{M_s \times 1}$. Let $\mathbf{v}_{p1,2} \in C^{M_s \times 1}$ denote the transmit beamforming vector at the PNB for PU 1 in the second time slot. Let $\mathbf{v}_k \in C^{M_s \times 1}$ denote the transmit beamforming vector at the SNB for the k^{th} SU, $k = 1, 2, \ldots, K+1$. The received signal at PU 1 and the k^{th} SU in the second time slot is given by

$$y_{p1,2}(n) = \mathbf{v}_{p1,2}^H \mathbf{g}_1 s_{p1,2}(n) + \sum_{i=1}^{K+1} \mathbf{v}_i \mathbf{h}_{p1} s_k(n) + q_3(n), \tag{5.2.7}$$

$$y_{sk}(n) = \mathbf{v}_k^H \mathbf{h}_k s_k(n) + \mathbf{v}_{p1,2} \mathbf{g}_{sk} s_{p1,2}(n) + \sum_{i=1, i \neq k}^{K+1} \mathbf{v}_i \mathbf{h}_{p1} s_i(n) + q_{sk}(n). \tag{5.2.8}$$

where $q_3(n)$ and $q_{sk}(n)$ are the additive white Gaussian noise components with variance σ_n^2. Hence, the received SINR at the PU 1 and the k^{th} SU in the second time slot is given by

$$\gamma_{p1,2} = \frac{\mathbf{v}_{p1,2}^H \mathbf{R}_{p1} \mathbf{v}_{p1,2}}{\sum_{i=1}^{K+1} \mathbf{v}_i^H \mathbf{R}_{s,p1} \mathbf{v}_i + \sigma_n^2}, \tag{5.2.9}$$

$$\gamma_{sk} = \frac{\mathbf{v}_k^H \mathbf{R}_{sk} \mathbf{v}_k}{\mathbf{v}_{p1,2}^H \mathbf{R}_{p,sk} \mathbf{v}_{p1,2} + \sum_{i=1, i \neq k}^{K+1} \mathbf{v}_i^H \mathbf{R}_{sk} \mathbf{v}_i + \sigma_n^2}, \tag{5.2.10}$$

where $\mathbf{R}_{s,p1} = \mathbb{E}\{\mathbf{h}_{p1} \mathbf{h}_{p1}^H\}$, $\mathbf{R}_{p,sk} = \mathbb{E}\{\mathbf{g}_{sk} \mathbf{g}_{sk}^H\}$ and $\mathbf{R}_{sk} = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$. $\gamma_{p1,2}$ and γ_{sk} are the SINRs of PU 1 and the k^{th} SU respectively. Power budget for both the primary network and the secondary network in the second time slot leads to the following constraints as

$$\|\mathbf{v}_{p1,2}\|_2^2 \leq P_{p,2}, \tag{5.2.11}$$

$$\sum_{i=1}^{K+1} \|\mathbf{v}_i\|_2^2 \leq P_s, \tag{5.2.12}$$

where $P_{p,2}$ and P_s are the power budget for the primary network and the secondary network in the second time slot. PU 1 is the only user who receives useful signals in both time slots. The data rate requirement for PU 1 over the two time slots is set to ζ. The SINR targets for PU 1 at the first and the second time slots are denoted as $\delta_{p1,1}$ and $\delta_{p1,2}$. Therefore, it is required to satisfy the following equation:

$$\log_2(1+\delta_{p1,1}) + \log_2(1+\delta_{p1,2}) = \zeta \text{ bits/s/Hz}. \tag{5.2.13}$$

From these definitions, the objective of the optimization problem can be stated as the minimization of the total network's power consumption subject to the individual SINR constraints to PUs and SUs together with the power constraint as follows:

$$\min_{\mathbf{v}_{p1,1},\mathbf{v}_{ps},\mathbf{v}_{p1,2},\mathbf{v}_k,\delta_{p1,1},\delta_{p1,2}} \|\mathbf{v}_{p1,1}\|_2^2 + \|\mathbf{v}_{ps}\|_2^2 + \|\mathbf{v}_{p1,2}\|_2^2 + \sum_{i=1}^{K+1} \|\mathbf{v}_i\|_2^2, \tag{5.2.14}$$

$$\text{s.t.} \quad \frac{\mathbf{v}_{p1,1}^H \mathbf{R}_{p1} \mathbf{v}_{p1,1}}{\mathbf{v}_{ps}^H \mathbf{R}_{p1} \mathbf{v}_{ps} + \sigma_n^2} \geq \delta_{p1,1}, \tag{5.2.15}$$

$$\frac{\mathbf{v}_{ps}^H \mathbf{R}_{ps} \mathbf{v}_{ps}}{\mathbf{v}_{p1,1}^H \mathbf{R}_{ps} \mathbf{v}_{p1,1} + \sigma_n^2} \geq \delta_{BS}, \tag{5.2.16}$$

$$\frac{\mathbf{v}_{p1,2}^H \mathbf{R}_{p1} \mathbf{v}_{p1,2}}{\sum_{i=1}^{K+1} \mathbf{v}_i^H \mathbf{R}_{s,p1} \mathbf{v}_i + \sigma_n^2} \geq \delta_{p1,2}, \tag{5.2.17}$$

$$\frac{\mathbf{v}_k^H \mathbf{R}_{sk} \mathbf{v}_k}{\mathbf{v}_{p1,2}^H \mathbf{R}_{p,sk} \mathbf{v}_{p1,2} + \sum_{i=1,i\neq k}^{K+1} \mathbf{v}_i^H \mathbf{R}_{sk} \mathbf{v}_i + \sigma_n^2} \geq \delta_{sk} k = 1, 2, \ldots, K+1, \tag{5.2.18}$$

$$\|\mathbf{v}_{p1,1}\|_2^2 + \|\mathbf{v}_{ps}\|_2^2 \leq P_{p,1}, \tag{5.2.19}$$

$$\|\mathbf{v}_{p1,2}\|_2^2 \leq P_{p,2}, \tag{5.2.20}$$

$$\sum_{i=1}^{K+1} \|\mathbf{v}_i\|_2^2 \leq P_s, \tag{5.2.21}$$

$$\log_2(1+\delta_{p1,1}) + \log_2(1+\delta_{p1,2}) = \zeta. \tag{5.2.22}$$

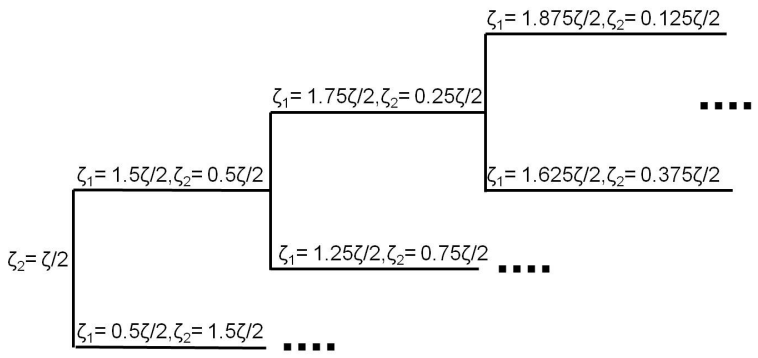

Figure 5.2. Data rate allocation procedure.

The objective function in (5.2.14) represents the total power consumption of both networks over the two time slots. Constraint (5.2.15) and (5.2.16) represent the SINR target for PU 1 and SNB respectively during the first time slot. Constraints (5.2.17)-(5.2.18) represents the SINR target for PU 1, k_{th} SU and the PU 2 (considered as $(K+1)^{th}$ SU) in the second time slot. Constraint (5.2.19) represents the power budget for the primary network during the first time slot. Constraint (5.2.20) and (5.2.21) represents the power budgets for the primary network and the secondary network during the second time slot. Constraint (5.2.22) represents the data rate requirement for the PU 1 during the two time slots. Since $\delta_{p1,1}$ in (5.2.15) and $\delta_{p1,2}$ in (5.2.17) are optimization variables, the constraints in (5.2.15) and (5.2.17) are non-convex [93]. Hence the overall problem is non-convex [93]. In the next section, a dual bi-section method has been proposed to avoid this non-convexity issue.

5.3 Iterative Algorithms using Convex Technique

As the constraints in (5.2.15) and (5.2.17) are non-convex, a method similar to the bi-section method has been proposed to solve the problem. The data rate target for PU 1 at the first time slot has been set to ζ_1. Hence the data

rate target for PU 1 at the second time slot should be set to $\zeta_2 = \zeta - \zeta_1$, where $\zeta_1 = \log_2(1 + \delta_{p1,1})$, $\zeta_2 = \log_2(1 + \delta_{p1,2})$. As a result, the constraint in (5.2.22) can be removed. Constraint (5.2.15) and (5.2.17) need to be updated. If the SINR target for the PU 1 during the first time slot is set to $\delta_{p1,1} = \lambda$, in order to maintain the total data rate at ζ, the SINR target for PU 1 at the second time slot should be set to:

$$\delta_{p1,2} = \frac{2^{\zeta}}{1+\lambda} - 1, \tag{5.3.1}$$

This is because $\log_2(1+\delta_{p1,1})+\log_2(1+\delta_{p1,2}) = \log_2(1+\lambda)+\log_2(\frac{2^{\zeta}}{1+\lambda}) = \zeta$ bits/s/Hz. First of all, identical data rate target is set for PU 1 at the first and the second time slots i.e. $\zeta_1 = \zeta_2 = \frac{\zeta}{2}$, the SINR target $\delta_{p1,1}$ for this step is set to $\lambda^{(0)} = 2^{\frac{\zeta}{2}} - 1$. The problem in (5.2.14)-(5.2.21) is solved and the optimum cost value P is saved in a memory. Then the data rate target for PU 1 during the first time slot is increased to $\frac{1.5\zeta}{2}$. Hence the data rate target for PU 1 at the second time slot should be set to $\frac{0.5\zeta}{2}$. For this step, SINR target for the first and the second time slots are $\lambda^{(1)} = \delta_{p1,1} = 2^{\frac{1.5\zeta}{2}} - 1$ and $\delta_{p1,2} = \frac{2^{\zeta}}{1+\lambda^{(1)}} - 1$ respectively. The problem (5.2.14)-(5.2.21) is solved again and the result is used to compare with the cost value saved in the memory. If the total power consumption is smaller than the one already saved in the memory, this new value will be saved in the memory and the data rate target for PU 1 during the first time slot will be increased to $\frac{1.75\zeta}{2}$, and the data rate target for PU 1 at the second time slot will be reduced to $\frac{0.25\zeta}{2}$ as shown in Fig. 5.2. The optimization in (5.2.14)-(5.2.21) is then performed again and the cost value is used to compare with the one saved in the memory. If the total power is smaller than the one already saved already in the memory, the data rate target for PU 1 during the first time slot will be increased to $\frac{1.875\zeta}{2}$, and the data rate target for PU 1 at the second time slot will be reduced to $\frac{0.125\zeta}{2}$. Otherwise, the data rate target for PU 1 during the first time slot will be reduced to $\frac{1.625\zeta}{2}$ and the data rate target for PU 1 at

the second time slot will be increased to $\frac{0.375\zeta}{2}$. This procedure as shown in Fig. 5.2 will continue until no further significant improvement on the power consumption is observed, i.e. until the following is satisfied

$$|P^{(q)} - P^{(q-1)}| < \epsilon, \tag{5.3.2}$$

where $P^{(q)}$ and $P^{(q-1)}$ indicates the total power consumed at the q^{th} and the $(q-1)^{th}$ step. ϵ represents the stoping criterion for the iteration. Defining a rank one matrix $\mathbf{V}_k = \mathbf{v}_k\mathbf{v}_k^H$, $\mathbf{V}_{p1,1} = \mathbf{v}_{p1,1}\mathbf{v}_{p1,1}^H$, $\mathbf{V}_{p1,2} = \mathbf{v}_{p1,2}\mathbf{v}_{p1,2}^H$ and $\mathbf{V}_{ps} = \mathbf{v}_{ps}\mathbf{v}_{ps}^H$ the problem in (5.2.14)-(5.2.21) can be solved using semidefinite programming with Lagrangian relaxation on the rank of the matrix $\mathbf{V}_k$, $\mathbf{V}_{p1,1}$, $\mathbf{V}_{p1,2}$ and $\mathbf{V}_{ps}$ [133]. For example, the constraint in (5.2.16) can be written as

$$\begin{aligned} tr(\mathbf{R}_{ps}\mathbf{V}_{ps}) - \delta_{BS}tr(\mathbf{R}_{ps}\mathbf{V}_{p1,1}) \geq \delta_{BS}\sigma_n^2, \\ \mathbf{V}_{ps} = \mathbf{V}_{ps}^H, \mathbf{V}_{p1,1} = \mathbf{V}_{p1,1}^H, rank(\mathbf{V}_{ps}) = 1, rank(\mathbf{V}_{p1,1}) = 1. \end{aligned} \tag{5.3.3}$$

Writing constraints (5.2.15), (5.2.17) and (5.2.18) also in this form, and relaxing the rank constraints, the problem can be converted into semidefinite programming form which is convex for a given $\delta_{p1,1}$ and $\delta_{p1,2}$. However when the covariance matrix can be written in the form of $\mathbf{R}_{p1} = \mathbf{g}_1\mathbf{g}_1^H$, $\mathbf{R}_{ps} = \mathbf{g}_{ps}\mathbf{g}_{ps}^H$, $\mathbf{R}_{s,p1} = \mathbf{h}_{p1}\mathbf{h}_{p1}^H$, $\mathbf{R}_{p,sk} = \mathbf{g}_{sk}\mathbf{g}_{sk}^H$ and $\mathbf{R}_{sk} = \mathbf{h}_k\mathbf{h}_k^H$, the constraints (5.2.15)-(5.2.18) can be written in the second order cone programming form whose solution is less complex as compared to the semidefinite programming. To convert the problem to second order cone programming without loss of generality, additional constraints $\mathbf{v}_{p1,1}^H\mathbf{g}_1 \geq 0$, $\mathbf{v}_{ps}^H\mathbf{g}_{ps} \geq 0$, $\mathbf{v}_{p1,2}^H\mathbf{g}_1 \geq 0$ and $\mathbf{v}_k^H\mathbf{h}_k \geq 0$ where $k = 1, 2, \ldots, K+1$ are added and the constraints (5.2.15)-(5.2.18) are formulated to the second order cone programming form (5.3.5)-(5.3.8). Hence for the case of rank one channel covariance matrices, the

optimization problem is solved in the q^{th} step of Fig. 5.2 as

$$\min_{\mathbf{v}_{p1,1},\mathbf{v}_{ps},\mathbf{v}_{p1,2},\mathbf{v}_{k(k=1,\cdots,K+1)}} \quad \|\mathbf{v}_{p1,1}\|_2^2 + \|\mathbf{v}_{ps}\|_2^2 + \|\mathbf{v}_{p1,2}\|_2^2 + \sum_{i=1}^{K+1} \|\mathbf{v}_i\|_2^2, \quad (5.3.4)$$

$$\text{s.t.} \quad (\mathbf{v}_{p1,1}^H \mathbf{g}_1)^2 \geq \lambda^{(q)}(\mathbf{v}_{ps}^H \mathbf{R}_{p1} \mathbf{v}_{ps} + \sigma_n^2), \quad (5.3.5)$$

$$(\mathbf{v}_{ps}^H \mathbf{g}_{ps})^2 \geq \delta_{BS}(\mathbf{v}_{p1,1}^H \mathbf{R}_{ps} \mathbf{v}_{p1,1} + \sigma_n^2), \quad (5.3.6)$$

$$(\mathbf{v}_{p1,2}^H \mathbf{g}_1)^2 \geq \quad (5.3.7)$$
$$(\frac{2^{\zeta}}{1+\lambda^{(q)}} - 1)(\sum_{i=1}^{K+1} \mathbf{v}_i^H \mathbf{R}_{s,p1} \mathbf{v}_i + \sigma_n^2),$$

$$(\mathbf{v}_k^H \mathbf{h}_k)^2 \geq \quad (5.3.8)$$
$$\delta_{sk}(\mathbf{v}_{p1,2}^H \mathbf{R}_{p,sk} \mathbf{v}_{p1,2} + \sum_{i=1, i\neq k}^{K+1} \mathbf{v}_i^H \mathbf{R}_{sk} \mathbf{v}_i + \sigma_n^2),$$

$$\|\mathbf{v}_{p1,1}\|_2^2 + \|\mathbf{v}_{ps}\|_2^2 \leq P_{p,1}, \quad (5.3.9)$$

$$\|\mathbf{v}_{p1,2}\|_2^2 \leq P_{p,2}, \quad (5.3.10)$$

$$\sum_{i=1}^{K+1} \|\mathbf{v}_i\|_2^2 \leq P_s, \quad (5.3.11)$$

$$\mathbf{v}_{p1,1}^H \mathbf{g}_1 \geq 0, \mathbf{v}_{ps}^H \mathbf{g}_{ps} \geq 0, \mathbf{v}_{p1,2}^H \mathbf{g}_1 \geq 0, \quad (5.3.12)$$

$$\mathbf{v}_k^H \mathbf{h}_k \geq 0, k = 1, 2, \ldots, K+1. \quad (5.3.13)$$

The original optimization problem in (5.2.14)-(5.2.22) was non-convex for two reasons. First in constraints (5.2.15) and (5.2.17), when the fractional form is expanded, the optimization variables $\delta_{p1,1}$ and $\delta_{p1,1}$ are multiplied with another set of optimization variables $\mathbf{v}_{ps}$ and $\mathbf{v}_i$ (i.e. the beamformer vectors). The multiplication of these optimization variables makes the constraints in (5.2.15) and (5.2.17) non-convex. This issue has been solved using the proposed dual bi-section method so that the parameters $\delta_{p1,1}$ and $\delta_{p1,1}$ are no longer optimization variables, but considered as constants in each step (node) of Fig. 5.2. Even after solving this issue, the constraints in (5.2.15)-(5.2.18) are non-convex due to the fractional quadratic form. For rank one channel covariance matrices, this has been solved by introducing additional constraints as in (5.3.12)-(5.3.13) and converting the constraints

in (5.2.15)-(5.2.18) in the second order cone form as in (5.3.5)-(5.3.8). However for general rank channel covariance matrices, this problem can be solved using semidefinite programming.

5.4 Simulation Results

In order to validate the performance of the proposed algorithm, a network with two PUs and two SUs has been considered. The channels have been generated using zero mean Gaussian random variables. Average channel gain for all the channels was set to one, except the channel from PNB to PU 2 which was set to 0.1 and the channel from the PNB to the SNB which was set to two. The additive white Gaussian noise variance σ_n^2 is 0.01. Power budgets for the primary network and the secondary network for both time slots $P_{p,1}$ and $P_{p,2}$ have been set to 1.5. The SINR target for PU 2 and the SUs would be varied depending on the target data rate requirement. However, in all simulation results, the SINR target δ_{BS} for the SNB for receiving PNB signal is set to two times higher than that of SUs. This is important for the SNB to decode the signal transmitted from the PNB error free. It should be noted that high SINR for SNB is possible to achieve as the channel between two basestations is likely to have much higher gain than the channels between BS and user terminals. The stoping criterion ϵ was set to 10^{-4}.

In the first simulation, the aim is to show the proposed algorithm converges to an identical setting regardless of the initializations in term of ζ_1 and ζ_2 for a particular set of random channels. For this simulation, the data rate requirements for both the primary users and the secondary users have been set to 4 bits/s/Hz. Therefore $\zeta = 4$ and $\delta_{sk} = 15$ for all $k = 1, ..., K+1$. Since the SINR target for the SNB is twice that of SINR of SUs, δ_{BS} has been set to 30. The initial values for the data rate allocation for PU 1 in

the first and the second time slots have been set to vary as ($\zeta_1 = 1$, $\zeta_2 = 3$),

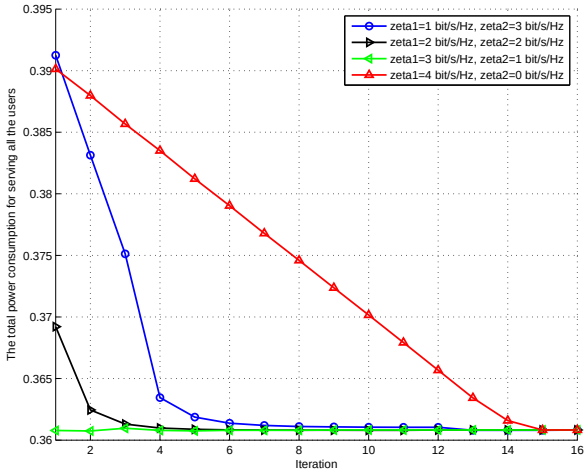

Figure 5.3. Convergence of the optimal joint optimization algorithm for various initializations in terms of the data rate for PU 1 in the first and the second time slots. The data rate target for PU 1 was set to 4 bits/s/Hz.

($\zeta_1 = 2$, $\zeta_2 = 2$), ($\zeta_1 = 3$, $\zeta_2 = 1$) and ($\zeta_1 = 4$, $\zeta_2 = 0$). As seen in Fig. 5.3, the total power consumption of both the basestations approaches to 0.361 regardless of the initializations.

In the proposed algorithm, the PU 1 receives data in both the first and

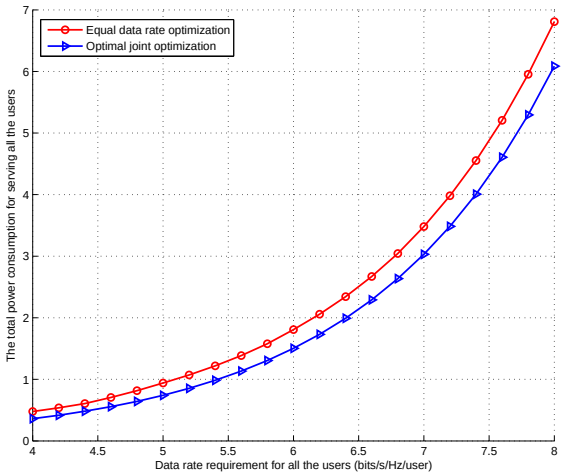

Figure 5.4. The total power consumption for serving PUs and SUs against various data rate requirements for users.

the second time slots. In the second time slot, the PU 1 receives signal from

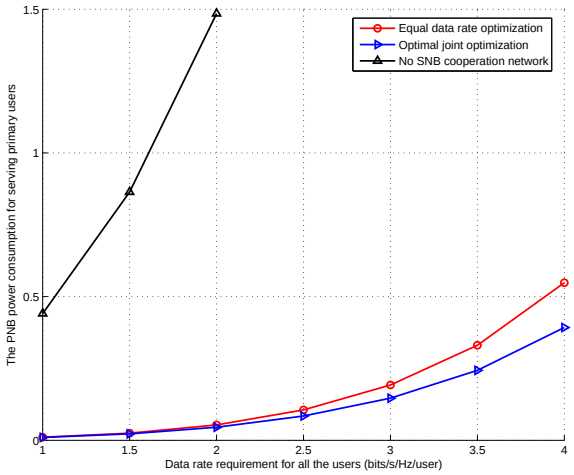

Figure 5.5. The power consumption of the PNB for serving PUs against various data rate requirements for PUs.

both PNB and SNB. Hence the PNB could allocate the resources in terms of power asymmetrically depending on the channel gains. Therefore the total power consumption of our proposed scheme is compared against a scheme that allocates equal data rate for PU 1 in both time slots (i.e., $\zeta_1 = \zeta_2$). The later scheme is called equal data rate optimization. The data rate targets for all other users have been changed from one to four bits/s/Hz. As expected, as seen in Fig. 5.4, the total power consumption for the proposed scheme that allocates date rate for PU 1 asymmetrically in the first and the second time slots is lower than that of the equal data rate allocation scheme.

The power consumption of the PNB for the proposed scheme and the equal data rate scheme are now investigated. These results are used to compare with a scheme where the PNB attempts to transmit signal to PU 2 without the help of SNB. The later scheme is called as "No SNB cooperation network". As seen in Fig. 5.5, without cooperation from the SNB, the PNB requires relatively very high power to meet its users' data rate requirements. This result confirms the advantage of subleasing the spectrum to the SU network in return of transmission power budget.

For a given transmission power budget for the PNB and the SNB, the

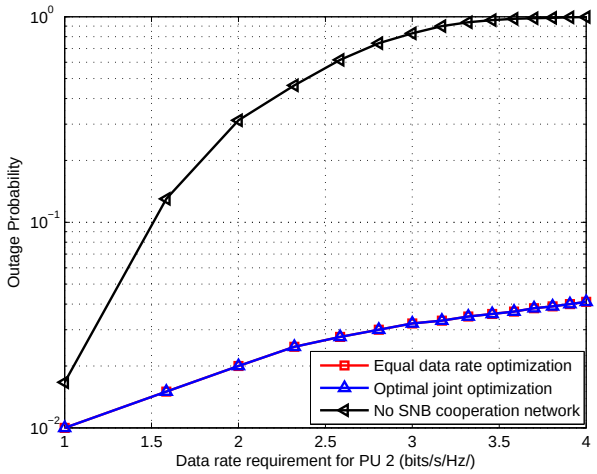

Figure 5.6. The outage probability of all three schemes against various data rate requirement for PU 2. The data rate target for PU 1, SU 1 and SU 2 was set to 4 bits/s/Hz.

optimization problem may not be feasible all the time. For example, when the target data rate is set too high, the optimization will turn out to be infeasible for most of the time. The occurrence of infeasibility is called as outage and this outage probability is used to compare all three schemes for various target data rate requirement for PU 2. The data rate requirement for PU 1, SU 1 and SU 2 has been set to 4 bits/s/Hz. As shown in Fig. 5.6, for the SNB cooperation based methods, the outage probability is very closer to zero while the outage probability approaches one for the scheme that does not have SNB cooperation.

5.5 Conclusion

A joint spatial and temporal resource allocation technique has been proposed for an overlay cognitive radio relaying network. By using cooperations in the cognitive radio network, the aim was to minimize the total power consumed by the networks while satisfying each user's SINR requirement. The original optimization problem was non-convex. A dual bi-section based second order

cone programming has been proposed to greatly improves the power saving as compared to a network without any cooperations. The simulation results demonstrated the convergence of the algorithm and power saving for both the primary and the secondary networks.

Chapter 6

INTERFERENCE ALIGNMENT TECHNIQUES FOR SPECTRUM SHARING NETWORKS

Interference mitigation techniques have become an important part of wireless network design. An interference alignment technique has been proposed recently in [27] as an efficient capacity achieving scheme at high SNR regime. The fundamental concept of interference alignment is to align the interference signals in a particular subspace at each receiver so that an interference-free orthogonal subspace can be solely allocated for data transmission. Since the work of [27], interference alignment techniques have attracted significant research interests and various algorithms have been proposed and analyzed. In this chapter, interference alignment algorithms have been proposed for a multi-cell multi-user MIMO Gaussian interfering broadcast channels (MIMO-IFBC) based on a hybrid interference alignment and BC-MAC duality based beamformer design. A grouping method already known in the literature has been extended to a multiple-cells scenario and jointly design transmit and receiver beamforming vectors using a closed-form expression without iterative computation. Based on the grouping method, a new ap-

proach using the principle of BC-MAC duality has been proposed to perform interference alignment while maximizing capacity of users in each cell. The algorithm in its dual form has been solved using interior point methods. It has been shown that the proposed approach outperforms the extension of the grouping method in terms of capacity and basestation complexity. To further exploit the performance of the network, a MIMO cognitive radio network with a MIMO relay that opportunistically accesses the same frequency band as that of a MIMO primary network has then been considered. In particular both interference cancelation and interference alignment techniques have been investigated to enhance the achievable DoF for the MIMO cognitive radio network. It has been shown that the DoF obtained by the cognitive radio network in the presence of a MIMO relay is higher than that could be obtained without a relay.

6.1 Introduction

Mitigation of interference is very important in cognitive radio networks to ensure that PUs are protected from the SU transmissions. Instead of considering interference mitigation, an interference alignment technique has been proposed recently in [27] as an efficient multiuser capacity achieving scheme in a high SNR regime. The fundamental concept of IA is to align all interference signals in a particular subspace at each receiver so that an interference-free orthogonal subspace can be solely allocated for data transmission. Since the work of [27], IA techniques have attracted significant interests and various algorithms have been proposed and analyzed, for example, MIMO interference network [28, 29], X network [30], and cellular network [31, 32].

Initially, the interference alignment has been proposed to achieve optimal DoF in a SISO interference channel (IC) [27]. It was shown that interference alignment can achieve the optimal DoF of $\frac{K}{2}$, in a K-user time varying

interference channel. In [28], the authors provided examples of iterative algorithms that utilize the reciprocity of wireless networks to achieve interference alignment with only local channel knowledge at each node. The work in [76] proposed the alignment of multi-user interference at each receiver based on a carefully constructed signal structure, which was referred to as interference alignment in signal space. For the interference alignment in signal space, transmit precoding technique is used to align the multi-user interferences in the same interference space which is orthogonal to the desired signal space at each receiver. Furthermore, the authors in [29] provided an inner-bound and an outer-bound for the total number of degrees of freedom for the K user MIMO Gaussian time varyings interference channels with M antennas at each transmitter and N antennas at each receiver. For the case of K user $M \times N$ MIMO interference channel, authors in [29] showed that the total number of degrees of freedom is equal to $\min(M,N)K$ if $K \leq R$ and $\min(M,N)\frac{R}{R+1}K$ if $K > R$, where $R = \frac{\max(M,N)}{\min(M,N)}$. An interference alignment scheme was provided in [77] for a deterministic K-user interference channel.

6.2 Interference Alignment Techniques for Multiple Input Multiple Output Multi-Cell Interfering Broadcast Channels

Multi-cell and multi-user downlink transmission schemes have been actively discussed for future generation cellular networks in [134]. The idea is to maximize the network capacity by efficiently mitigating interference. Authors in [31] proposed an interference alignment based scheme for cellular networks, namely subspace interference alignment. This is based on aligning interferences onto multi-dimensional subspace (instead of one dimension) for simultaneous alignments at multiple non-intended basestations (BS). In the multi-cell MIMO-IFBC, each BS supports multiple users within its cell. Therefore there exists two kinds of interference namely inter-user interference

(IUI) and inter-cell interference (ICI). To mitigate both IUI and ICI, authors in [135] proposed a zero-forcing (ZF) scheme for the IFBC with the aim of maximizing the sum rate performance in a MISO scenario. In [136], the ZF scheme for the MIMO-IFBC was extended to the case of multiple receiver antennas. Authors provided a precise expression of the spatial multiplexing gain for two mutually interfering MIMO broadcast channels using linear transceiver. Authors in [137] developed an interference alignment technique for a downlink cellular system which requires feedback only within each cell. The scheme provided substantial gain especially when the interference from a dominant interferer is significantly stronger than the remaining interference. Furthermore, for a two-cell MIMO-IFBC, the authors in [138] proposed a novel interference alignment technique for jointly designing the transmitter and receiver beamforming vectors using a closed-form expression without a need for iterative computation. It was shown both analytically and numerically that the proposed scheme achieves the optimal degrees of freedom. In this section, interference alignment algorithms have been proposed for a multi-cell multi-user MIMO-IFBC based on a hybrid interference alignment and BC-MAC duality based beamformer design. A grouping method already known in the literature has been extended to a multiple-cells scenario and jointly design transmit and receiver beamforming vectors using a closed-form expression without iterative computation. Based on the grouping method, a new approach using the principle of BC-MAC duality has been proposed to perform interference alignment while maximizing capacity of users in each cell. The proposed approach outperforms the extension of the grouping method in terms of capacity and basestation complexity.

6.2.1 System Model

The MIMO-IFBC model consists of a cellular network with L cells, each cell consists of K users. It is assumed that each user is equipped with N_r

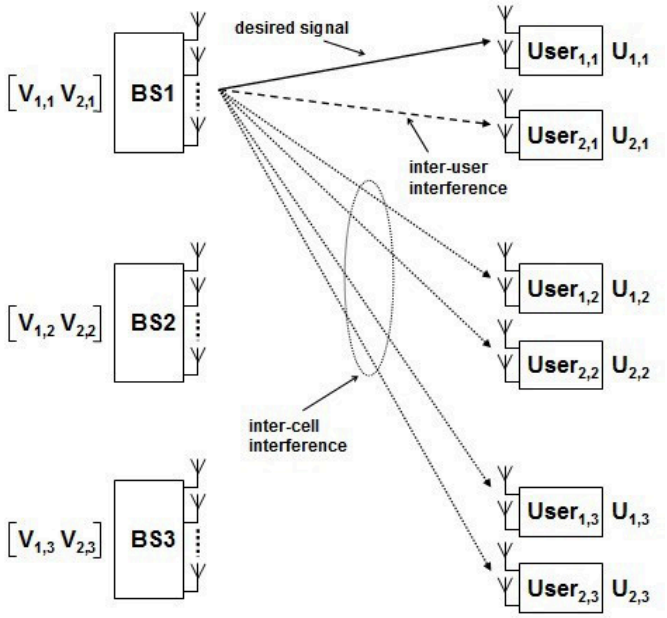

Figure 6.1. A multi-cell interference alignment scheme shown for the case of three cells and two users in each cell. In this example, BS 1 tries to convey data information to user 1 while introducing interference to other two cells.

antennas and each cell has one BS consisting of N_t antennas. The channel in each cell can be regarded as MIMO-IFBC. An example for the case of $L = 3$ and $K = 2$ is illustrated in Fig. 6.1. As shown in Fig. 6.1, the BS 1 sends data to user 1 while introducing both inter-user interference and inter-cell interference. Similarly, BS 2 and BS 3 introduce interference to other users. It is assumed that each BS aims to convey d_s data streams to its corresponding user, where $d_s \leq \min(N_t, N_r) = N_r$. The k^{th} user in the l^{th} cell is denoted as user $[k, l]$. The signal intended for the k^{th} user in the l^{th} cell is written as

$$\mathbf{x}^{[k,l]} = \sum_{i=1}^{d_s} \mathbf{v}_i^{[k,l]} s_i^{[k,l]} = \mathbf{V}^{[k,l]} \mathbf{s}^{[k,l]}, \tag{6.2.1}$$

where $s_i^{[k,l]}$ denotes the i^{th} transmitted symbol for the k^{th} user in the l^{th} cell, satisfying an average power constraint, $E[||\mathbf{x}^{[k,l]}||^2] \leq P^{[k,l]}$, and $\mathbf{v}_i^{[k,l]} \in C^{N_t \times 1}$ is the linear transmit beamforming vector corresponding to the symbol, $s_i^{[k,l]}$,

with a unity norm constraint, i.e., $||\mathbf{v}_i^{[k,l]}|| = 1$. The transmitter beamforming matrix for the user $[k,l]$ is written as $\mathbf{V}^{[k,l]} = [\mathbf{v}_1^{[k,l]}\ \mathbf{v}_2^{[k,l]}\ \cdots\ \mathbf{v}_{d_s}^{[k,l]}] \in C^{N_t \times d_s}$, and its corresponding data signal vector is denoted by $\mathbf{s}^{[k,l]} = [s_1^{[k,l]}\ s_2^{[k,l]}\ \cdots\ s_{d_s}^{[k,l]}]^T \in C^{d_s \times 1}$. Therefore the received signal of the k^{th} user in the l^{th} cell $\mathbf{y}^{[k,l]} \in C^{N_r \times 1}$ can be written as

$$\mathbf{y}^{[k,l]} = \sum_{i=1}^{L} \mathbf{H}_i^{[k,l]} \sum_{j=1}^{K} \mathbf{x}^{[j,i]} + \mathbf{n}^{[k,l]} \tag{6.2.2}$$

$$= \underbrace{\mathbf{H}_l^{[k,l]}\mathbf{V}^{[k,l]}\mathbf{s}^{[k,l]}}_{\text{desired signal}} + \underbrace{\sum_{j=1,j\neq k}^{K} \mathbf{H}_l^{[k,l]}\mathbf{V}^{[j,l]}\mathbf{s}^{[j,l]}}_{\text{inter-user interference}} + \underbrace{\sum_{i=1,i\neq l}^{L}\sum_{j=1}^{K} \mathbf{H}_i^{[k,l]}\mathbf{V}^{[j,i]}\mathbf{s}^{[j,i]}}_{\text{inter-cell interference}} + \mathbf{n}^{[k,l]},$$

where $\mathbf{n}^{[k,l]} \in C^{N_r \times 1}$ is the AWGN vector with variance σ^2 per entry at the receiver of user $[k,l]$, and $\mathbf{H}_i^{[k,l]}$ is the $N_r \times N_t$ channel matrix from the BS i to the user $[k,l]$. Each entry of the channel matrix $\mathbf{H}_i^{[k,l]}$ is generated using independently and identically distributed (i.i.d.) random variables according to $\mathcal{CN}(0,1)$. It is assumed that each channel is quasi-stationary and frequency flat fading. Each user decodes the desired signals coming from its corresponding BS by multiplying the received signal by a receiver beamforming matrix. The signal at the receiver for the user $[k,l]$ after receiver beamformer is written as

$$\tilde{\mathbf{y}}^{[k,l]} = \mathbf{U}^{[k,l]H}\mathbf{y}^{[k,l]} \tag{6.2.3}$$

$$= \mathbf{U}^{[k,l]H}\mathbf{H}_l^{[k,l]}\mathbf{V}^{[k,l]}\mathbf{s}^{[k,l]}$$

$$+ \mathbf{U}^{[k,l]H}\Big(\sum_{j=1,j\neq k}^{K} \mathbf{H}_l^{[k,l]}\mathbf{V}^{[j,l]}\mathbf{s}^{[j,l]} + \sum_{i=1,i\neq l}^{L}\sum_{j=1}^{K} \mathbf{H}_i^{[k,l]}\mathbf{V}^{[j,i]}\mathbf{s}^{[j,i]}\Big) + \tilde{\mathbf{n}}^{[k,l]},$$

where $\mathbf{U}^{[k,l]} = [\mathbf{u}_1^{[k,l]}\ \mathbf{u}_2^{[k,l]}\ \cdots\ \mathbf{u}_{d_s}^{[k,l]}] \in C^{N_r \times d_s}$ denotes the receiver beamforming matrix for the user $[k,l]$, and $\tilde{\mathbf{n}}^{[k,l]} = \mathbf{U}^{[k,l]H}\mathbf{n}^{[k,l]}$ is the effective noise component at the output of the beamformer which is distributed according to $\mathcal{CN}(0,1)$.

As in [29], the degrees of freedom for the multi-cell network is defined as the pre-log factor of the sum rate. This is one of the key metrics used for assessing the performance of a multiple antenna based system at high SNR regime, which is defined as

$$d \triangleq \lim_{SNR\to\infty} \frac{C_\Sigma(SNR)}{\log(SNR)} = \sum_{l=1}^{L}\sum_{k=1}^{K} d^{[k,l]}, \tag{6.2.4}$$

where $C_\Sigma(SNR)$ denotes the sum capacity that can be achieved for a given SNR, $C_\Sigma(SNR) = \sum_{l=1}^{L}\sum_{k=1}^{K} C^{[k,l]}$, where $C^{[k,l]}$ is the capacity achieved by the k^{th} user in the l^{th} cell and $d^{[k,l]}$ is the number of data streams transmitted to user $[k,l]$.

In order to decode the useful signal efficiently, both the ICI and IUI should be aligned into the interference space at the receiver. The desired signal should be linearly independent of the interference. Hence the signal space for the desired signal should be larger than or equal to the dimension of the data vector $d^{[k,l]}$. Both ICI and IUI are aligned into the sub-space which is orthogonal to $\mathbf{U}^{[k,l]}$. Therefore the following condition must be satisfied for the k^{th} user in the l^{th} cell:

$$\mathbf{U}^{[k,l]H}\mathbf{H}_i^{[k,l]}\mathbf{V}^{[j,i]} = 0,\ \forall i \neq l,\ j \in \{1,2,\cdots,K\} \tag{6.2.5}$$

$$\mathbf{U}^{[k,l]H}\mathbf{H}_l^{[k,l]}\mathbf{V}^{[m,l]} = 0,\ \forall m \neq k \tag{6.2.6}$$

$$\text{rank}\{\mathbf{U}^{[k,l]H}\mathbf{H}_l^{[k,l]}\mathbf{V}^{[k,l]}\} = d^{[k,l]} \tag{6.2.7}$$

In the next section, the interference alignment scheme has been extended using a grouping method proposed in [138] for two cells to jointly design transmitter and receiver beamforming vectors for multiple cells using a closed-form expression without a need for iterative computation. The extension as in the following section will be used later to compare other proposed interference alignment methods based on BC-MAC duality.

6.2.2 Extension of the Grouping Method

To maximize the sum rate performance of the MIMO-IFBC, the transmitter and the receiver beamforming matrices are usually designed by applying an iterative optimization algorithm as in [28]. The iterative scheme performs interference alignment implicitly and it normally requires a considerable number of iterations. In this section, the grouping method in [138] has been extended to the multi-cell scenario. This interference alignment scheme not only mitigates both ICI and IUI simultaneously in the multi-cell multi-user MIMO-IFBC, but also it does not require any iterative computation.

A simple example of $(N_t, N_r, K, L, d_s) = (10, 6, 2, 3, 2)$ has been employed to explain the algorithm. Suppose the BS l wants to transmit two sets of independent symbols $\mathbf{s}^{[1,l]} = [s_1^{[1,l]}\ s_2^{[1,l]}]$ and $\mathbf{s}^{[2,l]} = [s_1^{[2,l]}\ s_2^{[2,l]}]$ to user $[1,l]$ and user $[2,l]$ respectively. For this example, for the case of the first cell ($l = 1$), in order to transmit the symbol $\mathbf{s}^{[1,1]}$ without causing any interference to the users, the beamformer $\mathbf{V}^{[1,1]}$ should satisfy the following condition,

$$\mathbf{V}^{[1,1]} \subset \text{null}\ ([\underbrace{(\mathbf{U}^{[2,1]H}\mathbf{H}_1^{[2,1]})^H}_{\text{effective IUI channel}}\ \underbrace{(\mathbf{U}^{[1,2]H}\mathbf{H}_1^{[1,2]})^H\ (\mathbf{U}^{[2,2]H}\mathbf{H}_1^{[2,2]})^H}_{\text{effective ICI channel (Cell 2)}}$$

$$\underbrace{(\mathbf{U}^{[1,3]H}\mathbf{H}_1^{[1,3]})^H\ (\mathbf{U}^{[2,3]H}\mathbf{H}_1^{[2,3]})^H}_{\text{effective ICI channel (Cell 3)}}]^H) \tag{6.2.8}$$

where null($\mathbf{A}$) denotes an orthonormal basis for the null space of a matrix $\mathbf{A}$. The condition for the existence of the beamformer matrix $\mathbf{V}^{[1,1]}$ is that the dimension of the null space of the matrix $\mathbf{L} = [(\mathbf{U}^{[2,1]H}\mathbf{H}_1^{[2,1]})^H\ (\mathbf{U}^{[1,2]H}\mathbf{H}_1^{[1,2]})^H\ (\mathbf{U}^{[2,2]H}\mathbf{H}_1^{[2,2]})^H\ (\mathbf{U}^{[1,3]H}\mathbf{H}_1^{[1,3]})^H\ (\mathbf{U}^{[2,3]H}\mathbf{H}_1^{[2,3]})^H]^H$ is no less than d_s. The size of this matrix $\mathbf{L}$ is $Nt \times (5 \times d_s)$. For the example of $(N_t, N_r, K, L, d_s) = (10, 6, 2, 3, 2)$, the size of the matrix $\mathbf{L}$ is $N_t \times 10$. Since $d_s = 2$, it requires at least $N_t = 12$ antennas in order to have a null space of dimension at least 2. In general, suppose if a zero forcing based transmitter precoder is employed,

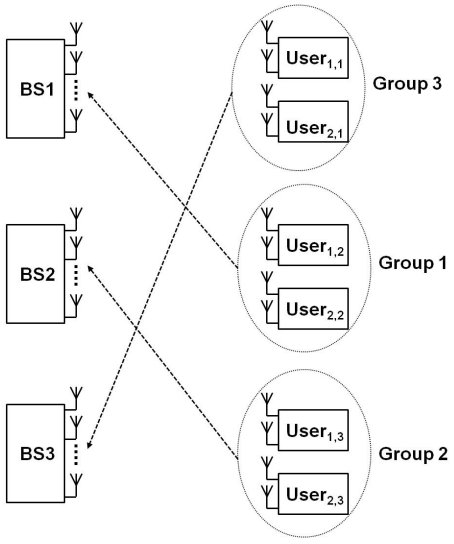

Figure 6.2. The extension of the grouping scheme shown for the case of three cells and two users in each cell.

for six users (two users in each cells) with two data streams for each user, it will require minimum of 12 antennas to cancel all ICI and IUI at the receiver of each user. However, the extension of grouping method requires only 10 transmitting antennas in order to remove both IUI and ICI simultaneously, thus reducing the hardware complexity at the BS. The transmitter and receiver beamforming design method has been presented using the following two-step scheme.

Step 1: Grouping the users and designing the receiver beamforming matrices

In the first step, similar to [138], the users are grouped to a particular interference space at different cells. The grouping procedure is explained in detail using the same example as in Fig. 6.2. The users $[1, 2]$ and $[2, 2]$ in the 2^{nd} cell are grouped together (Group 1) by designing the receiver beamforming matrices $\mathbf{U}^{[1,2]}$ and $\mathbf{U}^{[2,2]}$ such that the ICI channels from BS 1 are aligned in the same subspace. Hence from the view point of BS 1, users in the 2^{nd} cell are placed in the same interference space. In other words, the interference channels for users $[1, 2]$ and $[2, 2]$ from BS 1 span the same subspace as

follows,

$$\mathbf{G}_1 = \text{span}\{\mathbf{H}_1^{[1,2]H}\mathbf{U}^{[1,2]}\} = \text{span}\{\mathbf{H}_1^{[2,2]H}\mathbf{U}^{[2,2]}\}, \tag{6.2.9}$$

where span$(\cdot)$ denotes the subspace spanned by the column vectors of a matrix. The intersection subspace satisfying the condition (6.2.9) above can be determined by solving the following matrix equation [138],

$$\begin{bmatrix} \mathbf{I}_{N_t} & -\mathbf{H}_1^{[1,2]H} & \mathbf{0} \\ \mathbf{I}_{N_t} & \mathbf{0} & -\mathbf{H}_1^{[2,2]H} \end{bmatrix} \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{U}^{[1,2]} \\ \mathbf{U}^{[2,2]} \end{bmatrix} = \mathbf{F}_1\mathbf{X}_1 = \mathbf{0}, \tag{6.2.10}$$

where $\mathbf{G}_1$ accounts for the subspace spanned by the aligned effective interference channels from BS 1 to the user $[1,2]$ and user $[2,2]$ after applying the receiver beamforming. Since the size of the matrix $\mathbf{F}_1$ is $2N_t \times (N_t + 2N_r)$, the condition for existing null space is $2N_r > N_t$. For the example which has been considering, the size of the matrix $\mathbf{F}_1$ is 20×22. Therefore, it is possible to obtain receiver beamforming vectors for ICI channel alignment.

Generally, the size of the matrix $\mathbf{F}_1$ is $(2N_t) \times (N_t + 2N_r)$. When N_t and N_r are large, it can become cumbersome to find the solutions. Hence the sparsity in $\mathbf{F}_1$ is exploited to propose an efficient decomposition to reduce complexity. Instead of forming the $\mathbf{F}_1$ matrix, the matrix equation in (6.2.10) is decomposed into two matrix equations as follow,

$$\begin{bmatrix} \mathbf{I}_{N_t} & -\mathbf{H}_1^{[1,2]H} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}_{11} \\ \tilde{\mathbf{U}}^{[1,2]} \end{bmatrix} = \tilde{\mathbf{F}}_1\tilde{\mathbf{X}}_1 = \mathbf{0}, \tag{6.2.11}$$

$$\begin{bmatrix} \mathbf{I}_{N_t} & -\mathbf{H}_1^{[2,2]H} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}_{12} \\ \tilde{\mathbf{U}}^{[2,2]} \end{bmatrix} = \tilde{\mathbf{F}}_2\tilde{\mathbf{X}}_2 = \mathbf{0}, \tag{6.2.12}$$

where $\tilde{\mathbf{G}}_{11}$ and $\tilde{\mathbf{G}}_{12}$ account for the direction of the interference channels from BS 1 to the user $[1,2]$ and user $[2,2]$ respectively after applying the receiver beamforming matrices. First, it is required to determine the nullspaces $\tilde{\mathbf{X}}_1$ of matrix $\tilde{\mathbf{F}}_1$ and $\tilde{\mathbf{X}}_2$ of matrix $\tilde{\mathbf{F}}_2$. The nullspaces $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$ always exist because the size of both matrices $\tilde{\mathbf{F}}_1$ and $\tilde{\mathbf{F}}_2$ is $N_t \times (N_t + N_r)$. $\tilde{\mathbf{G}}_{11}$

and $\tilde{\mathbf{G}}_{12}$ are obtained from $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$. The aim is to determine $\mathbf{U}^{[1,2]}$ and $\mathbf{U}^{[2,2]}$ from $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$. For this, it is required to determine the intersection of subspaces spanned by $\tilde{\mathbf{G}}_{11}$ and $\tilde{\mathbf{G}}_{12}$ which will be equal to the subspace spanned by the columns of $\mathbf{G}_1$. i.e., It is required to determine the matrices (or vector if the dimension of $\mathbf{X}_1$ is 1) $\mathbf{C}_1$ and $\mathbf{C}_2$ such that

$$\tilde{\mathbf{G}}_{11}\mathbf{C}_1 = \tilde{\mathbf{G}}_{12}\mathbf{C}_2. \tag{6.2.13}$$

The matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ can be obtained as the nullspace of $[\tilde{\mathbf{G}}_{11} \ -\tilde{\mathbf{G}}_{12}]$ as follows,

$$[\tilde{\mathbf{G}}_{11} \ -\tilde{\mathbf{G}}_{12}] \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} = \bar{\mathbf{G}}\bar{\mathbf{C}} = \mathbf{0}. \tag{6.2.14}$$

The nullspace always exists because the size of the matrix $\bar{\mathbf{G}}$ is $N_t \times 2N_r$ (note $2N_r > N_t$). Once $\mathbf{C}_1$ and $\mathbf{C}_2$ are determined, the receiver beamforming matrices $\mathbf{U}^{[1,2]}$ and $\mathbf{U}^{[2,2]}$ are obtained from $\tilde{\mathbf{U}}^{[1,2]}$ and $\tilde{\mathbf{U}}^{[2,2]}$ as follows,

$$\mathbf{U}^{[1,2]} = \underbrace{\tilde{\mathbf{U}}^{[1,2]}\mathbf{C}_1}_{\text{receiver beamforer}} \tag{6.2.15}$$

$$\mathbf{U}^{[2,2]} = \underbrace{\tilde{\mathbf{U}}^{[2,2]}\mathbf{C}_2}_{\text{receiver beamforer}} \tag{6.2.16}$$

After applying the receiver beamforming matrices, the users $[1,2]$ and $[2,2]$ in the 2^{nd} cell are grouped together such that the ICI channels from BS 1 are aligned in the same subspace $\tilde{\mathbf{G}}_{11}$.

The complexity of the method in (6.2.10) and (6.2.11)-(6.2.16) mainly depends on the complexity of a matrix SVD and matrix multiplication. For a given $m \times n$ matrix, the required arithmetic operations to determine the singular values and the corresponding singular vectors are given by $\mathcal{O}(\min\{mn^2, m^2n\})$ [139]. Based on this, for a network with N_t transmit antennas at each BS and N_r receive antennas for each user, the number of arithmetic operations required for the first method in (6.2.10) is $\mathcal{O}(\min\{2N_t \times (N_t+2N_r)^2, (2N_t)^2 \times (N_t+2N_r)\})$. Since $2N_r > N_t$, the number

of arithmetic operations required for the first method is $\mathcal{O}(4N_t^3 + 8N_t^2N_r)$. For the second method described in (6.2.11)-(6.2.16), the number of arithmetic operations required involves performing SVD as in (6.2.11)-(6.2.12) and (6.2.14) and performing the matrix multiplication as in (6.2.13) and (6.2.15)-(6.2.16). Since the size of the matrices $\tilde{\mathbf{F}}_1$ and $\tilde{\mathbf{F}}_2$ is $N_t \times (N_t + Nr)$, the number of arithmetic operations required for the SVD as in (6.2.11) and (6.2.12) is $2\mathcal{O}(\min\{N_t \times (N_t + N_r)^2, N_t^2 \times (N_t + N_r)\})$. Since $N_t < N_t + N_r$, the number of arithmetic operations for performing SVD is $2\mathcal{O}(N_t^3 + N_t^2N_r)$. Since the size of the matrix $\bar{\mathbf{G}}$ is $N_t \times 2N_r$, the number of arithmetic operations required for the SVD as in (6.2.14) is $\mathcal{O}(2N_t^2N_r)$. Furthermore, the number of arithmetic operations required for the matrix multiplication in (6.2.13) and (6.2.15)-(6.2.16) is $2N_tN_r(2N_r - N_t) + 2N_r^2(2N_r - N_t)$. Hence, the overall complexity for the proposed low complexity algorithm is $2\mathcal{O}(N_t^3 + N_t^2N_r) + \mathcal{O}(2N_t^2N_r) + 2N_tN_r(2N_r - N_t) + 2N_r^2(2N_r - N_t)$. Since $N_r < N_t$, the overall complexity is bounded below $2\mathcal{O}(N_t^3 + N_t^2N_r) + 2\mathcal{O}(2N_t^2N_r)$, which is considerably lower than the complexity of the first method $\mathcal{O}(4N_t^3 + 8N_t^2N_r)$. This saving becomes more significant for $K \geq 3$.

When $K \geq 3$, as shown in the Appendix A, the complexity for the first algorithm depends on the complexity of a matrix SVD. Since the size of matrix $\mathbf{F}_1$ in this case is $KN_t \times (N_t + KN_r)$, the number of arithmetic operations required is $\mathcal{O}(K^2N_t^3 + K^3N_t^2N_r)$. For the proposed low complexity algorithm, the large matrix can be decomposed into K matrix equations and hence the number of arithmetic operations required is bounded below $K\mathcal{O}(N_t^3 + N_t^2N_r) + 2(K + \lfloor log_2K \rfloor)\mathcal{O}(2N_t^2N_r)$. Details and examples are shown in the Appendix A and Figure 6.1.

K	N_t	N_r	Matrix Decomposition	Matrix SVD
2	2	2	$8 \times O(16)$	O(96)
4	5	3	$4 \times O(200) + 12 \times O(150)$	O(6800)
4	10	6	$4 \times O(1600) + 12 \times O(1200)$	O(56800)

Table 6.1. Complexity of the proposed scheme

Similar to Group 1, all the users in cell 3 can be grouped to a particular interference space when seen from BS 2 (Group 2), and all the users in cell 1 to a particular interference space when seen from BS 3 (Group 3) as shown in Fig. 6.2.

$$\mathbf{G}_2 = \text{span}\{\mathbf{H}_2^{[1,3]H}\mathbf{U}^{[1,3]}\} = \text{span}\{\mathbf{H}_2^{[2,3]H}\mathbf{U}^{[2,3]}\} \tag{6.2.17}$$

$$\mathbf{G}_3 = \text{span}\{\mathbf{H}_3^{[1,1]H}\mathbf{U}^{[1,1]}\} = \text{span}\{\mathbf{H}_3^{[2,1]H}\mathbf{U}^{[2,1]}\} \tag{6.2.18}$$

As a result, all the users are put into three groups. From the view point of each BS, the interference to users in each group can be treated as interference to one destination. Hence in the example of $(N_t, N_r, K, L, d_s) = (10, 6, 2, 3, 2)$, each BS only requires 10 transmit antennas to remove both IUI and ICI at the same time. However a zero forcing based transmitter precoder will require minimum of 12 antennas to cancel all ICI and IUI at each user receiver. Hence, there is a saving in terms of complexity of the BS.

Please note that, for the case of the ICI in the first cell, if the technique used in [138] is directly employed for the three-cells example discussed above, the users $[1, 2]$ and $[2, 2]$ in the second cell, and the users $[1, 3]$ and $[2, 3]$ in the third cell are grouped together (e.g., Group 1). The receiver beamformer will be designed based on the interference space ($\mathbf{U}^{[1,2]}$, $\mathbf{U}^{[2,2]}$, $\mathbf{U}^{[1,3]}$, $\mathbf{U}^{[2,3]}$ are designed). However, since the receiver beamformers for all these users have already been designed, it is impossible to use the proposed grouping method to make the interference in BS 2 and BS 3 into a particular interference space. As a result, extra number of antennas are required by BS 2 and BS 3 to perform perfect interference alignment (to cancel out both the ICI and IUI), hence it increases the number of antennas required for each basestation and the complexity for each basestation. This situation becomes more complex for a large number of cells and users. However the extension of the grouping method solves this issue and reduces the complexity at the basestation while

maintaining perfect interference annihilation.

Step 2: Designing the transmit beamforming matrices

Since the effective ICI channels are aligned to each other, the BS 1 can treat the two different ICI channel vectors corresponding to user $[1,2]$ and user $[2,2]$ as a single ICI channel vector which spans d_s dimensional subspace as shown in (6.2.9). Hence the method in [138] is extended to the multi-cell model. The transmit beamforming vectors for the two users in the first cell $\mathbf{V}^{[1,1]}$ and $\mathbf{V}^{[2,1]}$ are designed as

$$\mathbf{V}^{[1,1]} \subset \text{null}\,([\mathbf{G}_1\ (\mathbf{U}^{[1,3]H}\mathbf{H}_1^{[1,3]})^H\ (\mathbf{U}^{[2,3]H}\mathbf{H}_1^{[2,3]})^H\ (\mathbf{U}^{[2,1]H}\mathbf{H}_1^{[2,1]})^H]^H) \tag{6.2.19}$$

$$\mathbf{V}^{[2,1]} \subset \text{null}\,([\mathbf{G}_1\ (\mathbf{U}^{[1,3]H}\mathbf{H}_1^{[1,3]})^H\ (\mathbf{U}^{[2,3]H}\mathbf{H}_1^{[2,3]})^H\ (\mathbf{U}^{[1,1]H}\mathbf{H}_1^{[1,1]})^H]^H) \tag{6.2.20}$$

Hence BS 1 can send the symbols $\mathbf{s}^{[1,1]}$ and $\mathbf{s}^{[2,1]}$ to the users $[1,1]$ and $[2,1]$ respectively without introducing any interference to users in the 2^{nd} and 3^{rd} cells. Similarly, the transmit beamforming vector at BS 2 and BS 3 are designed as

$$\mathbf{V}^{[1,2]} \subset \text{null}\,([\mathbf{G}_2\ (\mathbf{U}^{[1,1]H}\mathbf{H}_2^{[1,1]})^H\ (\mathbf{U}^{[2,1]H}\mathbf{H}_2^{[2,1]})^H\ (\mathbf{U}^{[2,2]H}\mathbf{H}_2^{[2,2]})^H]^H) \tag{6.2.21}$$

$$\mathbf{V}^{[2,2]} \subset \text{null}\,([\mathbf{G}_2\ (\mathbf{U}^{[1,1]H}\mathbf{H}_2^{[1,1]})^H\ (\mathbf{U}^{[2,1]H}\mathbf{H}_2^{[2,1]})^H\ (\mathbf{U}^{[1,2]H}\mathbf{H}_2^{[1,2]})^H]^H) \tag{6.2.22}$$

$$\mathbf{V}^{[1,3]} \subset \text{null}\,([\mathbf{G}_3\ (\mathbf{U}^{[1,2]H}\mathbf{H}_3^{[1,2]})^H\ (\mathbf{U}^{[2,2]H}\mathbf{H}_3^{[2,2]})^H\ (\mathbf{U}^{[2,3]H}\mathbf{H}_3^{[2,3]})^H]^H) \tag{6.2.23}$$

$$\mathbf{V}^{[2,3]} \subset \text{null}\,([\mathbf{G}_3\ (\mathbf{U}^{[1,2]H}\mathbf{H}_3^{[1,2]})^H\ (\mathbf{U}^{[2,2]H}\mathbf{H}_3^{[2,2]})^H\ (\mathbf{U}^{[1,3]H}\mathbf{H}_3^{[1,3]})^H]^H) \tag{6.2.24}$$

Hence, in the three-cell MIMO-IFBC with two users in each cell, and when each BS and the users are equipped with 10 and 6 antennas respectively, an effective interference alignment can be designed to achieve a DoF

of 12 (i.e., $\sum_{l=1}^{L}\sum_{k=1}^{K} d^{[k,l]} = 12$).

A general case with K users, L cells, and d_s data streams per user has now been studied. As an example, k^{th} user in the i^{th} cell has been considered. The transmit beamforming matrix should be obtained as follows,

$$\mathbf{V}^{[k,i]} \subset \text{null}\,([\underbrace{\mathbf{G}_i}_{\text{effective interference channels}} \underbrace{(\mathbf{U}^{[t(t=1,\cdots,K),s(s\neq i,i+1)]H}\mathbf{H}_i^{[t(t=1,\cdots,K),s(s\neq i,i+1)]})^H}_{\text{effective ICI channel}} \underbrace{(\mathbf{U}^{[t(t\neq k),i]H}\mathbf{H}_i^{[t(t\neq k),i]})^H}_{\text{effective IUI channel}}]^H) \tag{6.2.25}$$

Hence, the condition for the existence of $\mathbf{V}^{[k,i]}$ is that the above matrix should have a null space of dimension at least d_s. Since the size of above matrix is $[K(L-1)d_s] \times N_t$, the required minimum number of transmit antennas N_t is $[K(L-1)+1] \times d_s$.

The required minimum number of receive antennas N_r is now studied. Again, k^{th} user in the i^{th} cell is considered as an example. The method in (6.2.10) should be modified as follows,

$$\begin{bmatrix} \mathbf{I}_{N_t} & -\mathbf{H}_i^{[1,i+1]H} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I}_{N_t} & \mathbf{0} & -\mathbf{H}_i^{[2,i+1]H} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{I}_{N_t} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{H}_i^{[K,i+1]H} \end{bmatrix} \begin{bmatrix} \mathbf{G}_i \\ \mathbf{U}^{[1,i+1]} \\ \mathbf{U}^{[2,i+1]} \\ \vdots \\ \mathbf{U}^{[K,i+1]} \end{bmatrix} = \mathbf{F}_i\mathbf{X}_i = \mathbf{0}, \tag{6.2.26}$$

where $\mathbf{G}_i$ accounts for the direction of the aligned effective interference channels from BS i to all the users in the $(i+1)^{th}$ cell after applying the receiver beamforming matrices. Hence, the condition for the existence of the aligned effective interference channels $\mathbf{G}_i$ and the receiver beamforming matrices $\mathbf{U}^{[k,i+1]}$ is the existence of the nullspace of the matrix $\mathbf{F}_i$. Since the size of above matrix $\mathbf{F}_i$ is $(K \times N_t) \times (N_t + K \times N_r)$, the required minimum number of receive antennas N_r is $\frac{K-1}{K}N_t + \frac{d_s}{K}$. Since the required minimum

number of transmit antennas N_t is $[K(L-1)+1] \times d_s$, this result is substituted to state that the required minimum number of receive antennas N_r is $[(K-1)(L-1)+1] \times d_s$.

6.2.3 The Proposed Interference Alignment Scheme using BC-MAC Duality

The interference alignment scheme using the grouping method in section 6.2.2 can ensure zero IUI and ICI at the receiver of each user, hence this method is appropriate at high SNR region as the intention was to fully utilize the available degrees of freedom for transmission. However, for a moderate range of SNR, due to the perfect interference alignment, the interference alignment scheme using the grouping method proposed in section 6.2.2 may result in a lower network capacity. Also, due to the perfect interference alignment for the intra-cell users, the number of antennas required at the BS could still be comparably high. Therefore, interference alignment is proposed tp use only for the inter-cell users to ensure no ICI, but interference alignment is not performed for the intra-cell users. Instead, the IUI among users within each cell is treated by designing the beamformers to maximize capacity of each cell using the principle of BC-MAC duality.

For multiple-cells with multiple MIMO users, the basic approach of the proposed method is to write the beamformer matrix for the k^{th} user in the l^{th} cell as

$$\bar{\mathbf{V}}^{[k,l]} = \mathbf{V}^{[k,l]}\mathbf{W}_{k,l}, \tag{6.2.27}$$

where $\mathbf{V}^{[k,l]}$ ensures interference alignment to users in the $(l+1)^{th}$ cell, and $\mathbf{W}_{k,l}$ is a matrix that combines the column space of $\mathbf{V}^{[k,l]}$ optimally to maximize the intra-cell capacity. Hence, $\bar{\mathbf{V}}^{[k,l]}$ is the overall transmit beamforming matrix for user $[k,l]$ which is aimed at aligning interference of users in the $(l+1)^{th}$ cell and to maximize sum rate of cell l. The original problem

has become determining the optimal weight vector $\mathbf{W}_{k,l}$ to maximize the sum capacity in each cell l.

An example is provided to explain the proposed hybrid interference alignment scheme. The grouping method can still be used for the users as a first step of the algorithm described in Section 6.2.2, however the transmit beamforming matrix design should be modified as follows

$$\mathbf{T}_1 = \mathbf{V}^{[1,1]} = \mathbf{V}^{[2,1]} \subset \text{null}\ ([\mathbf{G}_1\ (\mathbf{U}^{[1,3]H}\mathbf{H}_1^{[1,3]})^H\ (\mathbf{U}^{[2,3]H}\mathbf{H}_1^{[2,3]})^H]^H) \tag{6.2.28}$$

$$\mathbf{T}_2 = \mathbf{V}^{[1,2]} = \mathbf{V}^{[2,2]} \subset \text{null}\ ([\mathbf{G}_2\ (\mathbf{U}^{[1,1]H}\mathbf{H}_2^{[1,1]})^H\ (\mathbf{U}^{[2,1]H}\mathbf{H}_2^{[2,1]})^H]^H) \tag{6.2.29}$$

$$\mathbf{T}_3 = \mathbf{V}^{[1,3]} = \mathbf{V}^{[2,3]} \subset \text{null}\ ([\mathbf{G}_3\ (\mathbf{U}^{[1,2]H}\mathbf{H}_3^{[1,2]})^H\ (\mathbf{U}^{[2,2]H}\mathbf{H}_3^{[2,2]})^H]^H) \tag{6.2.30}$$

where $\mathbf{T}_1$, $\mathbf{T}_2$ and $\mathbf{T}_3$ represent the base matrices (sub-spaces) for the beamformer design in (6.2.27) for the basestations BS 1, BS 2 and BS 3 respectively. It is clear that the transmit beamforming matrix for user $[1,1]$ will introduce IUI to user $[2,1]$, and vice versa, but there is no ICI introduced between users in other cells.

To explain the proposed algorithm, the same example of (Nt, Nr, K, L, d_s) $= (10, 6, 2, 3, 2)$ is employed. To design the beamformers, the interference alignment subspace $\mathbf{T}_1$, $\mathbf{T}_2$ and $\mathbf{T}_3$ are first determined using (6.2.28)-(6.2.30). Then it is required to find the optimal combination matrices as shown in (6.2.27) to maximize the sum rate of users in each cell. Considering the case of l^{th} cell, the BS l wants to transmit two independent symbol $\mathbf{s}^{[1,l]}$ and $\mathbf{s}^{[2,l]}$ to user $[1,l]$ and user $[2,l]$ respectively. Using (6.2.28), both the users $[1,l]$ and $[2,l]$ share the same base matrix $\mathbf{T}_l$. However, as the proposed interference alignment scheme only considers the inter-cell users to ensure no ICI, but interference alignment is not performed for the intra-cell users, the components $(\mathbf{U}^{[2,l]H}\mathbf{H}_l^{[2,l]})^H$ in (6.2.19) and $(\mathbf{U}^{[1,l]H}\mathbf{H}_l^{[1,l]})^H$ in

(6.2.20) should be excluded, as shown in (6.2.28). Hence the base matrix for cell l, $\mathbf{T}_l$ has four column vectors, e.g., $\mathbf{T}_l = [\mathbf{t}_{1,l}\ \mathbf{t}_{2,l}\ \mathbf{t}_{3,l}\ \mathbf{t}_{4,l}]$. Two weight matrices $\mathbf{W}_{1,l} \in\ C^{4\times 2}$ and $\mathbf{W}_{2,l} \in\ C^{4\times 2}$ are introduced for user $[1,l]$ and user $[2,l]$ respectively. Hence the transmit beamforming matrices for user $[1,l]$ and user $[2,l]$ are designed as a linear combination of the columns of $\mathbf{T}_l$ respectively as follows,

$$\bar{\mathbf{V}}^{[1,l]} = \mathbf{T}_l \mathbf{W}_{1,l}, \tag{6.2.31}$$

$$\bar{\mathbf{V}}^{[2,l]} = \mathbf{T}_l \mathbf{W}_{2,l}. \tag{6.2.32}$$

Hence $\bar{\mathbf{V}}^{[1,l]}$ and $\bar{\mathbf{V}}^{[2,l]}$ are the overall transmit beamforming matrices for users $[1,l]$ and $[2,l]$ respectively. Now it is required to determine the optimal weight vector $\mathbf{W}_{1,l} \in\ C^{4\times 2}$ and $\mathbf{W}_{2,l} \in\ C^{4\times 2}$ to maximize the sum rate for all the users in the cell l using the following optimization

$$\max_{\mathbf{W}_{k,l}} \quad \sum_{k=1}^{K} R^{[k,l]} \tag{6.2.33}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} P^{[k,l]} \leq P^{l}, \tag{6.2.34}$$

where $R^{[k,l]}$ denotes the data rate achieved by user k in the l^{th} cell, $P^{[k,l]}$ denotes the power consumption by user k in the l^{th} cell. P^l denotes the total power budget for the l^{th} cell. This optimization problem can not be solved directly using convex optimization techniques [93]. Therefore the BC-MAC duality technique is employed to solve the optimization problem.

The first step is to define the effective channel and introduce a virtual transmit covariance matrix. The duality between Gaussian MAC and Gaussian BC as in [69] and [1] is then applied to solve the MIMO sum rate optimization problem as described in section A. Finally, the optimization for the solution of virtual matrices in the MAC setting has been proposed

and this MAC solution needs to map it back to BC setting. In section B, a rate balancing technique is also used to maintain fairness among users.

A. Sum rate maximization

Dirty paper coding is employed at the BS. Assuming the encoding order is $(1, 2, \cdots, K)$, i.e., the codeword of user 1 is encoded first. The data rate $R_b^{[k,l]}$ for the k^{th} user in the l^{th} cell can be written as [69]

$$R_b^{[k,l]} = \log \frac{|\mathbf{I} + \sum_{i=k}^{K} \tilde{\mathbf{H}}_l^{[i,l]} \mathbf{Q}_b^{[i,l]} \tilde{\mathbf{H}}_l^{[i,l]H}|}{|\mathbf{I} + \sum_{i=k+1}^{K} \tilde{\mathbf{H}}_l^{[i,l]} \mathbf{Q}_b^{[i,l]} \tilde{\mathbf{H}}_l^{[i,l]H}|}, \tag{6.2.35}$$

where $\tilde{\mathbf{H}}_l^{[k,l]}$ denotes the effective channel from BS l to user $[k, l]$ which is written as

$$\tilde{\mathbf{H}}_l^{[k,l]} = \mathbf{U}^{[k,l]H} \mathbf{H}_l^{[k,l]} \mathbf{V}^{[k,l]}, \tag{6.2.36}$$

and $\mathbf{Q}_b^{[i,l]} = \{\mathbf{W}_{i,l} \mathbf{W}_{i,l}^H\}$ is the virtual transmit covariance matrix for the i^{th} user in the l^{th} cell, $\mathbf{Q}_b^{[i,l]} \succeq 0$. Hence the sum rate maximization problem in (6.2.33)-(6.2.34) can be written as

$$\max_{\mathbf{Q}_b^{[i,l]} \succeq 0} \quad \sum_{k=1}^{K} R_b^{[k,l]} \tag{6.2.37}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \text{tr}(\mathbf{Q}_b^{[k,l]}) \leq P_b^l, \tag{6.2.38}$$

where P_b^l denotes the total power budget for the l^{th} cell. The optimization problem in (6.2.37) and (6.2.38) is a sum rate maximization problem for MIMO-BC channel. However, this is not a convex problem, hence it can not be solved directly. The authors in [1] proved that the dirty paper capacity region of a MIMO BC channel with a power constraint P is equal to the capacity region of the dual MIMO MAC with the same sum power constraint

P. Hence, the problem in (6.2.37) and (6.2.38) can be solved by transforming the BC (i.e., downlink) problem to the dual MAC problem as follows

$$\max_{\mathbf{Q}_m^{[i,l]} \succeq 0} \quad \sum_{k=1}^{K} R_m^{[k,l]} \tag{6.2.39}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \text{tr}(\mathbf{Q}_m^{[k,l]}) \leq P_m^l, \tag{6.2.40}$$

where $R_m^{[k,l]}$ and $\mathbf{Q}_m^{[k,l]}$ denotes the rate achieved and the virtual transmit covariance matrix by the k^{th} user in the l^{th} cell for the MAC respectively. The problem in (6.2.39)-(6.2.40) is convex [69] and it can be solved for the optimal value of $\mathbf{Q}_m^{[k,l]}$ using interior point methods [93]. Due to the duality between the MIMO-BC and the MIMO-MAC, the optimal encoding order for the MIMO-BC employing DPC is the reverse of the decoding order for the MIMO-MAC using successive interference cancelation scheme. Hence $R_m^{[k,l]}$ is written as

$$R_m^{[k,l]} = \log \frac{|\mathbf{I} + \sum_{i=1}^{k} \tilde{\mathbf{H}}_l^{[i,l]H} \mathbf{Q}_m^{[i,l]} \tilde{\mathbf{H}}_l^{[i,l]}|}{|\mathbf{I} + \sum_{i=1}^{k-1} \tilde{\mathbf{H}}_l^{[i,l]H} \mathbf{Q}_m^{[i,l]} \tilde{\mathbf{H}}_l^{[i,l]}|}, \tag{6.2.41}$$

The authors in [69] derived a transformation that takes as inputs a set of MAC covariance matrices and a decoding order and outputs a set of BC covariance matrices with the same sum power as in MAC. The virtual beamforming matrix $\mathbf{Q}_m^{[k,l]}$ obtained in the MAC from (6.2.39)-(6.2.40) can be mapped to $\mathbf{Q}_b^{[k,l]}$ in the BC as follows [69],

$$\mathbf{Q}_b^{[k,l]} = \mathbf{B}^{[k,l]^{-1/2}} \mathbf{D}^{[k,l]} \mathbf{E}^{[k,l]^H} \mathbf{A}^{[k,l]^{1/2}} \mathbf{Q}_m^{[k,l]} \mathbf{A}^{[k,l]^{1/2}} \mathbf{E}^{[k,l]} \mathbf{D}^{[k,l]^H} \mathbf{B}^{[k,l]^{-1/2}} \tag{6.2.42}$$

where $\mathbf{A}^{[k,l]}$ and $\mathbf{B}^{[k,l]}$ are defined as follows

$$\mathbf{A}^{[k,l]} \triangleq \mathbf{I} + \tilde{\mathbf{H}}_l^{[k,l]} (\sum_{i=1}^{k-1} \mathbf{Q}_b^{[i,l]}) \tilde{\mathbf{H}}_l^{[k,l]H}, \tag{6.2.43}$$

$$\mathbf{B}^{[k,l]} \triangleq \mathbf{I} + \sum_{i=k+1}^{K} \tilde{\mathbf{H}}_l^{[i,l]H} \mathbf{Q}_m^{[i,l]} \tilde{\mathbf{H}}_l^{[i,l]}, \tag{6.2.44}$$

and the effective channel $\mathbf{B}^{[k,l]-1/2}\tilde{\mathbf{H}}_l^{[k,l]H}\mathbf{A}^{[k,l]-1/2}$ is decomposed using the SVD as $\mathbf{B}^{[k,l]^{-1/2}}\tilde{\mathbf{H}}_l^{[k,l]H}\mathbf{A}^{[k,l]^{-1/2}} = \mathbf{D}^{[k,l]}\mathbf{\Lambda}^{[k,l]}\mathbf{E}^{[k,l]}$, where $\mathbf{\Lambda}^{[k,l]}$ is a square diagonal matrix. Once the matrix $\mathbf{Q}_b^{[k,l]}$ is obtained, the optimal beamforming matrix can be determined using equation (6.2.31) and (6.2.32). This procedure is repeated for all the cells $l \in \{1, 2, \cdots, L\}$.

The analysis for the number of antennas at both the transmitter and the receiver has now been provided. The required minimum number of transmit antennas N_t and receive antennas N_r is considered as the required number of antennas for the proposed hybrid IA scheme to perform interference alignment for the inter-cell users to ensure no ICI while maximizing network capacity of each cell in the moderate SNR regime. Consider a general case with K users, L cells, and d_s data stream per user. k^{th} user in the i^{th} cell is considered as an example. Since interference alignment is used only for the inter-cell users to ensure no ICI, but interference alignment is not performed for the intra-cell users, the transmit beamforming matrices should be obtained as follows,

$$\mathbf{V}^{[k,i]} \subset \text{null}([\mathbf{G}_i \quad \underbrace{(\mathbf{U}^{[t(t=1,\cdots,K),s(s\neq i,i+1)]H}\mathbf{H}_i^{[t(t=1,\cdots,K),s(s\neq i,i+1)]})^H}_{\text{effective ICI channel}}]^H) \tag{6.2.45}$$

Hence, the condition for the existence of $\mathbf{V}^{[k,i]}$ is that the above matrix should have a null space of dimension at least d_s. Since the size of above matrix is $[K(L-2)+1]d_s \times N_t$, the required minimum number of transmit antennas N_t is $[K(L-2)+2]d_s$. For the required minimum number of receive antennas N_r, the same grouping method has been employed as in section 6.2.2, hence the required minimum number of receive antennas N_r is $\frac{K-1}{K}N_t + \frac{d_s}{K}$. Since the required minimum number of transmit antennas

1) Group the users and design the receiver beamforming matrices $\mathbf{U}^{[k,l]}$, $k \in 1, 2, \cdots K,\ l \in 1, 2, \cdots L$;
2) Design the base matrices for the transmit beamforming using equation (6.2.28), (6.2.29) and (6.2.30), $k \in 1, 2, \cdots K,\ l \in 1, 2, \cdots L$;
3) Define the effective channel from BS l to user $[k, l]$ using equation (6.2.36) for $k \in 1, 2, \cdots K,\ l \in 1, 2, \cdots L$;
4) **For** $l = 1, 2, \cdots, L$, do the following:
5) Initialize: $\mu_1^d \cdots \mu_K^d$ to positive values, $d = 1$;
6) **repeat**
7) Determine the optimal solution $\mathbf{Q}_m^{[k,l]}$ for MAC problem as in Subsection A, $\mathbf{Q}_m^{[k,l]}$ is mapped to $\mathbf{Q}_b^{[k,l]}$ using BC-MAC duality;
8) Update the Lagrangian coefficient μ_k^d in (6.2.47), $d = d + 1$;
9) **Stop** when $|R_b^{[k,l]} - R_b^{[k-1,l]}| \leq \epsilon$ for $k = 2, \cdots, K$;
10) **End For**.

Table 6.2. Proposed interference alignment scheme

N_t is $[K(L-2)+2]d_s$, this result is substituted to determine the required minimum number of receive antennas N_r as $[(K-1)(L-2) + \frac{2K-1}{K}]d_s$.

B. Data rate balancing

Suppose the achievable rate region for the problem in (6.2.39)-(6.2.40) is $\mathbf{r} \in C(\tilde{\mathbf{H}}^{[l]}, P^l)$. The technique investigated so far aims to maximize the sum rate of users in each cell. To ensure fairness among the users, a better criterion is to maximize the data rate while balancing rate achieved for each user as in [101]

$$\max_{\mathbf{r},\gamma} \gamma \quad s.t. \ \ \mathbf{r} = \gamma \boldsymbol{\rho}, \quad \mathbf{r} \in C(\tilde{\mathbf{H}}^{[l]}, P^l) \tag{6.2.46}$$

where γ is the optimization variable and the cost function to reflect maximization of balanced data rate and $\boldsymbol{\rho}$ is a vector account for ratio of data

rats achieved for various users. For example if $\boldsymbol{\rho} = \mathbf{1}_K$, all users attain the same data rate. For other values of $\boldsymbol{\rho}$, data rate is maximized while scaling the total rate to users according to the ratio defined by the vector $\boldsymbol{\rho}$. Due to convexity of the capacity region [77], the optimization problem in (6.2.46) is convex. Hence the iterative algorithm for the solution of (6.2.46) can by found by exploiting Lagrangian duality. Furthermore, as rate vectors in the interior $C(\tilde{\mathbf{H}}^{[l]}, P^l)$ can always be found that lie on the straight line defined by the QoS constraints, strong duality holds [140]. As a consequence, (6.2.46) can be alternatively solved by solving the dual problem. Using the results in [101], the dual problem is a weighted sum rate optimization of the following form

$$\min_{\boldsymbol{\mu}} \ \max_{\mathbf{r}} \ \sum_{k=1}^{K} \mu_k \frac{R^{[k,l]}}{\rho_k} \quad s.t. \ \sum_{k=1}^{K} \mu_k = 1 \tag{6.2.47}$$

where μ_k are the Lagrangian coefficients for the k constraints in (6.2.46), $\mathbf{r} = [R^{[1,l]} R^{[2,l]} \cdots R^{[K,l]}]$ is the rate vector and ρ_k is the k^{th} element of $\boldsymbol{\rho}$. Hence, for an initial setting of μ_k, $\max \sum_{k=1}^{K} (\frac{\mu_k}{\rho_k}) R^{[k,l]}$ can be solved using the method described earlier (Section A), and the μ_k can be updated using a subgradient method as in [101], i.e. $\mu_k^{(d)} = \mu_k^{(d-1)} - t(R^{[k,l]} - R^{[K,l]})$, where t is a small positive step size. At each iteration, it is necessary to solve the weighted sum rate problem corresponding to the maximization step. For the updated μ_k, the sum rate optimization problem is solved again and μ_k is computed again. This is continued until convergence. The resulting algorithm will maximize the capacity for each cell while ensuring data rate balancing and interference alignment to all the users in other cells. The complete proposed interference alignment scheme using BC-MAC duality is now summarized in Table 6.2.

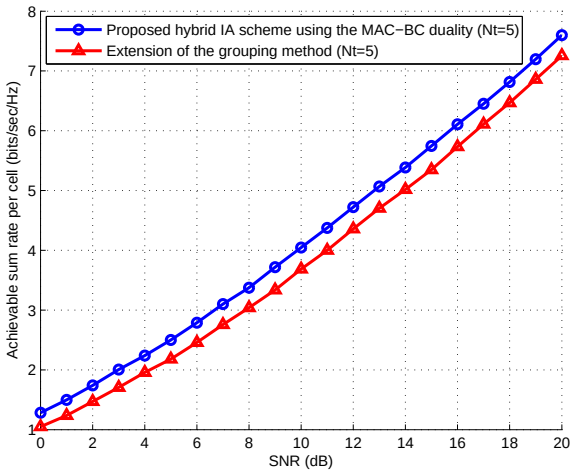

Figure 6.3. The achievable rates for the proposed hybrid interference alignment scheme and comparison to the extension of the grouping method (DoF = 6).

6.2.4 Simulation Results

In this section, the performance of the proposed hybrid interference alignment and BC-MAC duality based scheme is compared with the extension of the grouping method in terms of the sum rate. Two system configurations $(N_t, N_r, K, L, d_s) = (5, 3, 2, 3, 1)$ and $(N_t, N_r, K, L, d_s) = (10, 6, 2, 3, 2)$ have been considered. For these configurations, the proposed scheme is able to achieve 6 degrees of freedom and 12 degrees of freedom, respectively. To be specific, to achieve a degrees of freedom of 6, the extension of the grouping scheme requires $N_t = 5$ antennas for the example of $(N_t, N_r, K, L, d_s) = (5, 3, 2, 3, 1)$, but the conventional zero-forcing beamforming scheme requires $N_t = 6$ antennas for the case of $(N_t, N_r, K, L, d_s) = (6, 3, 2, 3, 1)$. Moreover, for a general case with K users, L cells, and d_s data stream per user, the required minimum number of transmit antennas N_t for the proposed hybrid interference alignment and BC-MAC duality based scheme is $[K(L-2)+2] \times d_s$. Therefore, the minimum number of antennas for each BS employing the proposed hybrid interference alignment and BC-MAC duality based scheme is 4 and 8. For the case that $(N_t, N_r, K, L, d_s) = (4, 3, 2, 3, 1)$,

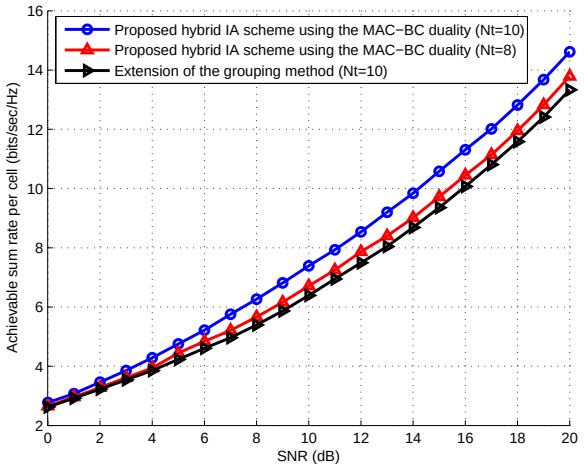

Figure 6.4. The achievable rates for the proposed hybrid interference alignment scheme with different number of antennas setting (DoF = 12).

and as an example, when cell l is considered, as 4 antennas are employed at BS l, the base matrix $\mathbf{T}_l$ for cell l as shown in (6.2.28)-(6.2.30) is a column vector, e.g., $\mathbf{T}_l = [\mathbf{t}_{1,l}]$. Thus the duality optimization problem for this case is a pure power allocation problem.

The performance of different number of transmit antennas for the proposed hybrid interference alignment and BC-MAC duality based scheme is compared. The elements of the channel matrix $\mathbf{H}_l^{[k,i]}$ is assumed to be circularly symmetric complex Gaussian variables with zero mean and unity variance. The total power constraint P^l is varied from 1 to 100. The noise variances at the users have been set to unity. First, the case that only a single data stream is transmitted for each user has been considered. There exists three cells, each containing two users with three receiver antennas. Fig. 6.3 depicts the sum rate versus SNR of the proposed algorithm and compares it with the grouping method. As seen, the proposed interference alignment scheme using the BC-MAC duality with five transmit antennas outperforms the extension of the grouping method with five transmit antennas.

The same system but with two data streams transmitted for each user

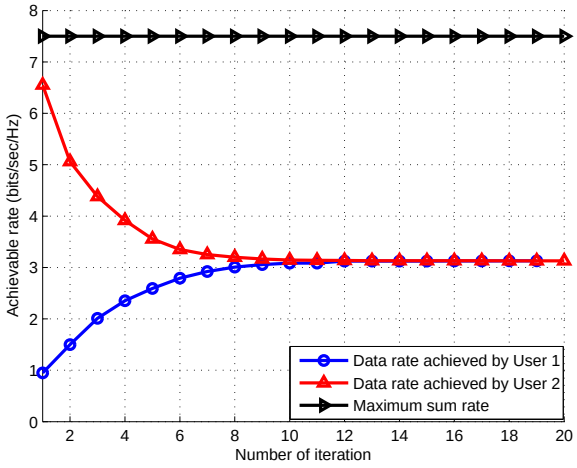

Figure 6.5. Rate balancing of the proposed hybrid interference alignment algorithm using the BC-MAC duality (DoF = 6).

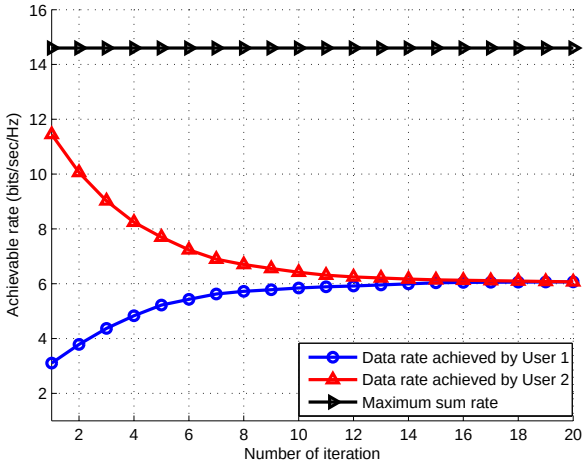

Figure 6.6. Rate balancing of the proposed hybrid interference alignment algorithm using the BC-MAC duality (DoF = 12).

has then been considered. For the proposed interference alignment scheme using BC-MAC duality, both 8 and 10 transmit antennas are considered for each base station. Fig. 6.4 depicts the sum rate versus SNR of the proposed algorithm and compares it with the grouping method but with 10 transmit antennas. As seen, the proposed interference alignment scheme, even with 8 transmit antennas, outperforms the extension of the grouping method with 10 transmit antennas. The reason is that although only eight

antennas are employed at the BS, the sum rate is maximized using the virtual beamforming matrices $\mathbf{Q}_m^{[k,l]}$, hence it outperforms the perfect interference alignment algorithm due to increase subspace dimension for mitigating intra-cell interference.

Simulation for rate balancing has been provided. The aim is to show

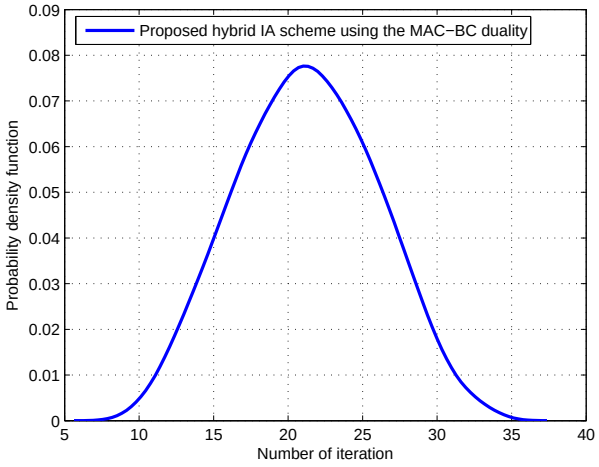

Figure 6.7. Probability density function for the number of iterations of the proposed hybrid interference alignment algorithm using the BC-MAC duality.

the proposed interference alignment algorithm can also balance all the users' data rate in each cell. All the elements of the data rate balancing vector $\boldsymbol{\rho}$ were set to one. Fig. 6.5 and Fig. 6.6 depict the convergence of the data rate for the two users against the adaptation of the Lagrangian multiplier μ_k as explained in the Section 6.2.3 B Section. All the users attain equal data rate. The data rate without rate balancing constraints is also shown. The total sum rate in this case is 7.5 bits/s/Hz and 14.6 bits/s/Hz for DoF = 6 and DoF = 12 respectively. With the rate balancing constraints, each user attains 3.05 bits/s/Hz and 6 bits/s/Hz for DoF = 6 and DoF = 12 respectively. Hence the total rate is less than that of the scheme that does not use rate balancing constraint. However, the rate balancing constraint ensures fairness among users. Finally, the probability density function has been

provided for the number of iterations required for the proposed interference alignment algorithm using the BC-MAC duality. As shown in Fig. 6.7, the algorithm converges with 21 iterations most of the time.

6.3 Interference Cancelation and Alignment Techniques for MIMO Cognitive Relay Networks

There are several results that show significant improvement in SNR gains by employing relays within the context of interference cancelation [81–83]. The DoF is increased with the use of relays, which is particularly attractive when the source and the destination nodes are not directly connected. Authors in [84] characterized the DoF for a multihop two-user interference channel with layered relays. However, if the direct channel coefficients between the source and the destination nodes are non-zero, then from the perspective of DoF, it has been shown that the use of half-duplex relays does not contribute to any gain in DoF for the two-user interference channel [85], regardless of the number of relay nodes. Remarkably, this is true even if the relay nodes are equipped with multiple antennas. Since the conventional half-duplex relays are not contributing to increase the DoF of a two-user interference channel, another form of relaying introduced by El Gamal and Hassanpour [86] has been exploited, known as instantaneous relaying (relay-without-delay). El Gamal and Hassanpour [86] have shown that the memoryless instantaneous relay channel can potentially achieve a higher capacity than a conventional relay. By employing this memoryless instantaneous relay channel, the authors in [87] showed that the two-user interference channel with an instantaneous relay can achieve $\frac{3}{2}$ DoF. An instantaneous relay increases the DoF for the two-user interference channel as compared to a network without relays. In this section, a MIMO cognitive radio network with a MIMO relay that opportunistically accesses the same frequency band as that of a

MIMO primary network has been considered. Both interference cancelation and interference alignment techniques have been investigated to enhance the achievable degrees of freedom for the MIMO cognitive radio network.

6.3.1 System Model

Coexistence of primary network and secondary network has been considered as shown in Fig. 6.8. The first transmitter-receiver pair (T_p, R_p) is for the primary network. The secondary link consists of a source node (T_s), an intermediate relay node (Relay) and a destination node (R_s). The secondary network is an opportunistic link that should not strictly cause interference to the primary network. It is assumed that primary transmitter and the receiver are equipped with M_p and N_p antennas respectively. The secondary transmitter and the receiver are equipped with M_s and N_s antennas respectively. It is assumed that the number of transmitting antennas is the same as that of the receiving antennas at the intermediate relay i.e, $M_r = N_r$. In this channel, both the primary and the secondary transmitters wish to send independent messages to their respective destinations and no explicit cooperation is assumed between them in the sense that there is no data message exchange between the transmitters. It is assumed that the channels between various nodes are quasi-stationary and fixed over the duration of a number of blocks.The channel matrices from the primary transmitter to the PU receiver, the relay and the SU receiver are denoted as $\mathbf{H}_{pp}$, $\mathbf{H}_{rp}$ and $\mathbf{H}_{sp}$ respectively. Similarly, the channel matrices from the secondary transmitter to the SU receiver, the relay and the PU receiver are denoted as $\mathbf{H}_{ss}$, $\mathbf{H}_{rs}$ and $\mathbf{H}_{ps}$ respectively. Finally, the channel transfer matrices from the intermediate relay to the PU receiver and the SU receiver are denoted as $\mathbf{H}_{pr}$ and $\mathbf{H}_{sr}$ respectively. Regarding channel knowledge assumptions at various nodes, it is assumed that the primary terminals (transmitter and receiver) have perfect knowledge of its own channel matrix $\mathbf{H}_{pp}$, while the secondary

terminals (transmitter and receiver) and the relay have perfect knowledge of all the channel transfer matrices. There are many cases that will make this assumption possible, for example, the channel reciprocity can be exploited to acquire CSI at the transmitters. Also, when the primary network subleases the spectrum for monitory purposes to the secondary network under a spectrum sharing arrangement, a degree of collaboration can be expected for acquiring CSI values.

It is assumed that the rank of the channel $\mathbf{H}_{pp}$ is $d_p = \min(M_p, N_p)$. The primary transmitter wishes to convey d_p data streams to its corresponding receiver. The underlying logic presented in this section will remain valid for case when the rank of the channel $\mathbf{H}_{pp}$ is less than $\min(M_p, N_p)$. The primary network aims to operate using its maximum possible DoF. Hence, the signal intended for the PU is written as

$$\mathbf{x}_p = \sum_{i=1}^{d_p} \mathbf{v}_p^i s_p^i = \mathbf{V}_p \mathbf{s}_p \tag{6.3.1}$$

where s_p^i denotes the i^{th} transmitted symbol at the primary transmitter, satisfying an average power constraint, $E\{||\mathbf{x}_p||^2\} \leq P_p$, and $\mathbf{v}_p^i \in C^{M_p \times 1}$ is a linear transmit beamforming vector corresponding to the symbol, s_p^i, with a unit norm constraint, i.e., $||\mathbf{v}_p^i|| = 1$. The transmitter beamforming matrix for the primary network is written as $\mathbf{V}_p = [\mathbf{v}_p^1\ \mathbf{v}_p^2\ \cdots\ \mathbf{v}_p^{d_p}] \in C^{M_p \times d_p}$, and its corresponding data signal vector is denoted as $\mathbf{s}_p = [s_p^1\ s_p^2\ \cdots\ s_p^{d_p}]^T \in C^{d_p \times 1}$. Similarly, it is assumed the secondary transmitter wishes to convey d_s data streams to its corresponding secondary receiver. Hence, the signal intended for the SU is written as

$$\mathbf{x}_s = \sum_{i=1}^{d_s} \mathbf{v}_s^i s_s^i = \mathbf{V}_s \mathbf{s}_s \tag{6.3.2}$$

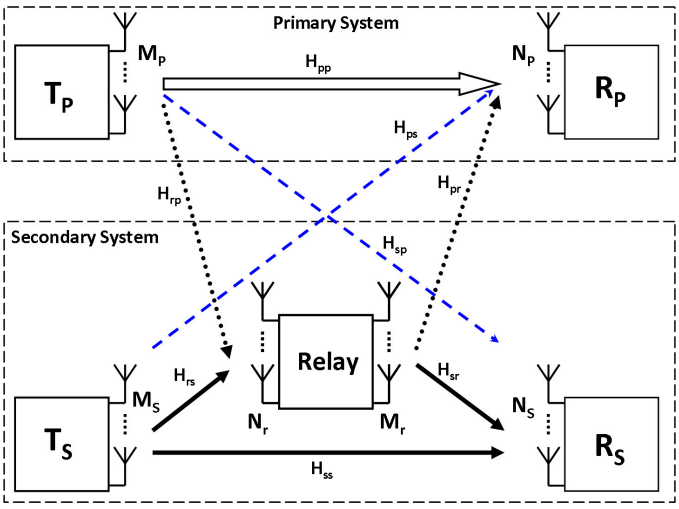

Figure 6.8. Cognitive radio network with a memoryless instantaneous relay. The first transmitter-receiver pair (T_p, R_p) is for the primary network. The second transmission link consists of a source node (T_s), an intermediate relay node (Relay) and a destination node (R_s).

where s_s^i denotes the i^{th} transmitted symbol at the secondary transmitter, satisfying an average power constraint, $E\{||\mathbf{x}_s||^2\} \leq P_s$, and $\mathbf{v}_s^i \in C^{M_s \times 1}$ is a linear transmit beamforming vector corresponding to the symbol, s_s^i, with a unit norm constraint, i.e., $||\mathbf{v}_s^i|| = 1$. The transmitter beamforming matrix for the SU is written as $\mathbf{V}_s = [\mathbf{v}_s^1\ \mathbf{v}_s^2\ \cdots\ \mathbf{v}_s^{d_p}] \in C^{M_s \times d_s}$, and its corresponding data signal vector is denoted as $\mathbf{s}_s = [s_s^1\ s_s^2\ \cdots\ s_s^{d_s}]^T \in C^{d_s \times 1}$. The relay operates in its full-duplex mode as similar to the work presented in [87]. It forwards the received data symbols from the sources to the destinations instantaneously. The received signal at the relay $\mathbf{y}_r \in C^{N_r \times 1}$ which is a combination of the primary and the secondary signals, can be written as

$$\mathbf{y}_r = \mathbf{H}_{rp}\mathbf{x}_p + \mathbf{H}_{rs}\mathbf{x}_s + \mathbf{n}_r \tag{6.3.3}$$

where $\mathbf{n}_r \in C^{N_r \times 1}$ is the AWGN vector with variance σ^2 per entry at the receiver of the relay. The received signal at the PU $\mathbf{y}_p \in C^{N_p \times 1}$ and the SU

receiver $\mathbf{y}_s \in C^{N_s \times 1}$ are written as

$$\mathbf{y}_p = \mathbf{H}_{pp}\mathbf{x}_p + \mathbf{H}_{ps}\mathbf{x}_s + \mathbf{H}_{pr}\mathbf{x}_r + \mathbf{n}_p \tag{6.3.4}$$

$$\mathbf{y}_s = \mathbf{H}_{ss}\mathbf{x}_s + \mathbf{H}_{sp}\mathbf{x}_p + \mathbf{H}_{sr}\mathbf{x}_r + \mathbf{n}_s \tag{6.3.5}$$

where $\mathbf{n}_p \in C^{N_p \times 1}$ and $\mathbf{n}_s \in C^{N_s \times 1}$ are the AWGN vectors with variance σ^2 per entry at the receivers of the primary and the secondary network respectively. $\mathbf{x}_r \in C^{M_r \times 1}$ is the signal transmitted from the intermediate relay.

The system model in Fig. 6.8 in the absence of any intermediate relays is known as MIMO IC. For this case, it was shown in [79] that the DoF of the cognitive network with M_s transmit antennas and N_s receive antennas operating in the presence of primary network with d_p DoF is upper bounded by

$$d_s \leq \min\{(M_s - d_p)^+, (N_s - d_p)^+\}. \tag{6.3.6}$$

In the subsequent subsections, an intermediate MIMO relay is used within the context of interference cancelation and interference alignment to enhance the above DoF further.

6.3.2 The Maximum Achievable Degrees of Freedom for Interference Cancelation Based MIMO Cognitive Relay Networks

In this section, the maximum achievable degrees of freedom for MIMO cognitive relay networks is established. To keep the problem formulation simple, the case of no symbol extension has been considered, however the same arguments can be extended to the case of symbol extension. The primary transmitter sends d_p independent data symbols to its corresponding receiver, where $d_p = \min\{M_p, N_p\}$. The primary transmitter chooses the transmit

beamforming matrix that maximizes its capacity as

$$\max_{\mathbf{V}_p \succeq 0} \quad \log |\mathbf{I} + \frac{1}{\sigma^2} \mathbf{H}_{pp} \mathbf{Q}_p \mathbf{H}_{pp}^H| \tag{6.3.7}$$

where σ^2 denotes the noise covariance at the PU receiver, $\mathbf{Q}_p = \mathbf{V}_p \mathbf{P}_p \mathbf{V}_p^H$ is the transmit beamforming matrix at the primary transmitter and $\mathbf{P}_p$ is the power allocation matrix. The PU receiver will receive simultaneously the interference signals from the secondary transmitter as well as the relay respectively. Hence, a relay transceiver is designed to cancel both these interferences.

Theorem 1: Suppose the number of antenna at the intermediate relay N_r is not less than the sum of DoF of the primary network d_p and the maximum DoF available for the secondary network without the presence of primary network $\min\{M_s, N_s\}$. The maximum achievable DoF of the cognitive radio network with M_s transmit antennas and N_s receive antennas operating in the presence of primary network with d_p DoF is $d_s \leq \min\{M_s, N_s\}$.

As the number of antennas at the primary network and the secondary network may be different, in this proof, four cases have been considered according to different combinations of antenna numbers. *Theorem 1* has been proven by considering all these four cases as follows,

Case A. $M_p \geq N_p, M_s \geq N_s$

In this case, the number of transmit antennas is not less than that of the corresponding receiver antennas at both the primary and the secondary network. The primary transmitter wishes to transmit d_p data streams to its corresponding receiver using the beamformer shown in (6.3.1). The secondary transmitter wishes to transmit $d_s = \min\{M_s, N_s\}$ data streams using the beamformer shown in (6.3.2) to its corresponding receiver. As the number of antennas at the relay is not less than the sum of DoF in the

primary network and the maximum DoF available in the secondary network in the absence of primary network, the case that the relay employs minimum number of antenna required has been considered for the proof, i.e., $N_r = M_r = d_p + \min\{M_s, N_s\}$. There are two different source signals, $\mathbf{H}_{rp}\mathbf{x}_p$ from the primary transmitter and $\mathbf{H}_{rs}\mathbf{x}_s$ from the secondary transmitter. As the channel transfer matrices $\mathbf{H}_{rp}$ and $\mathbf{H}_{rs}$ are realizations of a random matrix, the received source signal vectors are linearly independent with a probability approaching one at the relay receiver, i.e., $\mathbf{H}_{rp}\mathbf{x}_p$ and $\mathbf{H}_{rs}\mathbf{x}_s$ are not co-linear vectors. There are enough space to separate all the signals as such the relay sees $d_p + d_s$ independent data symbols from the two sources nodes.

First, the relay estimates $d_p + d_s$ data symbols from the two sources which are $\mathbf{s}_p$ and $\mathbf{s}_s$ respectively. The estimation of the symbols at the relay is written as

$$\mathbf{s}_r^1 = \mathbf{s}_p, \quad \mathbf{s}_r^2 = \mathbf{s}_s. \tag{6.3.8}$$

The relay employs two beamforming matrices $\mathbf{V}_r^p$ and $\mathbf{V}_r^s$ for carrying $\mathbf{s}_r^1$ and $\mathbf{s}_r^2$. In particular, the relay employs $\mathbf{V}_r^p$ and $\mathbf{V}_r^s$ to cancel $\mathbf{s}_p$ at the SU receiver and $\mathbf{s}_s$ at the PU receiver respectively. Hence, the relay beamforming matrices should satisfy

$$\mathbf{H}_{sr}\mathbf{V}_r^p = -\mathbf{H}_{sp}\mathbf{V}_p. \tag{6.3.9}$$

$$\mathbf{H}_{pr}\mathbf{V}_r^s = -\mathbf{H}_{ps}\mathbf{V}_s. \tag{6.3.10}$$

Therefore, $\mathbf{s}_p$ is canceled at the SU receiver while $\mathbf{s}_s$ is canceled at the PU receiver. For the case that $M_p \geq N_p$, $M_s \geq N_s$, there always exists two set of beamformers $\mathbf{V}_r^p$ and $\mathbf{V}_r^s$ to satisfy the conditions in (6.3.9) and (6.3.10), for which the proof is given in Appendix B. The next step is to verify whether each destination can decode the desired data symbols. From the relay transmission, the undesired interferences are annihilated at each desti-

nation, leaving only the desired signals which can be resolved in the receiver's dimensional space. As a result, each destination can decode the desired data symbols and thus $\min\{M_s, N_s\}$ DoF are achieved at the secondary network when $N_r = M_r \geq d_p + \min\{M_s, N_s\}$. This is interesting because it is higher than the upper bound of DoF $\min\{(M_s - d_p)^+, (N_s - d_p)^+\}$ reported for the case of two-user MIMO interference channel in the absence of relay node [79].

Remark 1: The relay beamformer $\mathbf{V}_r^p$ in this case not only cancels the interference to the SU receiver, but also it forwards the data symbols $\mathbf{s}_p$ to the PU receiver. The signal at the PU receiver is written as

$$\mathbf{y}_p = \mathbf{H}_{pp}\mathbf{V}_p\mathbf{s}_p + \mathbf{H}_{pr}\mathbf{V}_r^p\mathbf{s}_r^1 + \mathbf{n}_p \tag{6.3.11}$$

where $\mathbf{s}_r^1 = \mathbf{s}_p$. Therefore, to avoid the possibility of destructive addition of primary signals at the PU receiver, the beamformer $\mathbf{V}_r^p$ has to be designed such that the primary capacity is maximized through the effective primary channel $\tilde{\mathbf{H}}_p$ subject to a SU signal cancelation constraint as follows

$$\begin{aligned} \max_{\mathbf{G}\succeq 0} \quad & \log|\mathbf{I} + \frac{\tilde{\mathbf{H}}_p\mathbf{G}\tilde{\mathbf{H}}_p^H}{\sigma^2}| \\ \text{s.t.} \quad & \mathbf{H}_{sr}\mathbf{G} + \mathbf{T} = \mathbf{0} \end{aligned} \tag{6.3.12}$$

where the effective channel $\tilde{\mathbf{H}}_p = \mathbf{H}_{pr} - \mathbf{H}_{pp}(\mathbf{H}_{sp}^H\mathbf{H}_{sp})^{-1}\mathbf{H}_{sp}^H\mathbf{H}_{sr}$, $\mathbf{G} = \mathbf{V}_r^p\mathbf{V}_r^{pH}$, $\mathbf{T} = (\mathbf{H}_{sp}\mathbf{V}_p)(\mathbf{H}_{sp}\mathbf{V}_p)^H\mathbf{H}_{sr}(\mathbf{H}_{sr}^H\mathbf{H}_{sr})^{-1}$, $\mathbf{0}$ is an $N_p \times 1$ vector consisting of all zeros, σ^2 is the noise variance at the PU receiver. The above optimization problem is convex and hence it can be solved using convex techniques [93]. Details for the optimization problem in (6.3.12) and the proof for convexity is given in Appendix C. Hence by replacing the beamforming scheme from (6.3.9) to (6.3.12), the relay beamformer $\mathbf{V}_r^p$ not only cancels the interference to the SU receiver, but also helps the primary network to maximize its data rate.

Remark 2: The proposed beamforming scheme for the secondary transmitter and the relay focused on attaining maximum possible DoF gain. However, beamforming vectors at the secondary transmitter and the relay can also be designed to maximize its achievable data rate through the effective primary channel $\tilde{\mathbf{H}}_s$ while keeping the maximum possible DoF. To show this, the signal at the SU receiver is written as

$$\mathbf{y}_s = \mathbf{H}_{ss}\mathbf{V}_s\mathbf{s}_s + \mathbf{H}_{sr}\mathbf{V}_r^s\mathbf{s}_r^2 + \mathbf{n}_s \tag{6.3.13}$$

where $\mathbf{s}_r^2 = \mathbf{s}_s$. To maximize its achievable data rate, the beamforming vectors $\mathbf{V}_r^s$ is constructed as

$$\begin{aligned} \max_{\mathbf{F}\succeq 0} \quad & \log|\mathbf{I} + \frac{\tilde{\mathbf{H}}_s\mathbf{F}\tilde{\mathbf{H}}_s^H}{\sigma^2}| \\ \text{s.t.} \quad & \mathbf{H}_{pr}\mathbf{F} + \mathbf{R} = \mathbf{0} \end{aligned} \tag{6.3.14}$$

where the effective channel $\tilde{\mathbf{H}}_s = \mathbf{H}_{sr} - \mathbf{H}_{ss}(\mathbf{H}_{ps}^H\mathbf{H}_{ps})^{-1}\mathbf{H}_{ps}^H\mathbf{H}_{pr}$, $\mathbf{F} = \mathbf{V}_r^s\mathbf{V}_r^{sH}$, $\mathbf{R} = (\mathbf{H}_{ps}\mathbf{V}_s)(\mathbf{H}_{ps}\mathbf{V}_s)^H\mathbf{H}_{pr}(\mathbf{H}_{pr}^H\mathbf{H}_{pr})^{-1}$. Similar to Appendix C, the above optimization problem is convex and hence it can be solved using convex techniques [93]. Hence, by replacing the beamforming scheme in (6.3.10) to (6.3.14), the effective channel gain can be increased by coherently combining the desired symbols coming from the two different paths. As a remark, the interference cancelation technique not only annihilates interference at the primary destination (yielding required DoF gain), but it also maximizes its achievable data rate at the secondary destination (enhancing SNR gain).

Case B. $M_p < N_p, M_s \geq N_s$

In this case, the number of transmit antennas at the primary network is less than the number of receive antennas. However, the number of transmit antennas at the secondary network is not less than that of its receiver. A

beamforming matrix $\mathbf{V}_r^s$ may not always exist to satisfy equation (6.3.10). This is because $(d_s + d_p + M_s) > N_p$ does not always hold, and hence unlike in Case A, the existence of the nullspace of the effective channel matrix can not be proven. The interference cancelation scheme proposed previously can not be employed in this case. Hence a new interference cancelation approach is required.

The underlying idea is to cancel the interference at the PU at the point after the receive beaforming processing matrix $\mathbf{U}_p$. The interference cancelation scheme proposed in the previous section was to cancel the interference before the receiver beamforming matrix $\mathbf{U}_p$ (at the receiver antennas). Therefore equation (6.3.10) in the previous section should be modified to

$$\mathbf{U}_p^H \mathbf{H}_{pr} \mathbf{V}_r^s = -\mathbf{U}_p^H \mathbf{H}_{ps} \mathbf{V}_s. \tag{6.3.15}$$

Hence, the interference introduced from both the secondary transmitter and the relay at the signal space of the PU receiver is aimed to be canceled while aligning or ignoring the interference placed in the null space of the PU receiver. There always exists $\mathbf{V}_r^s$ to satisfy the condition in (6.3.15), for which the proof is given in Appendix D. Since $M_s \geq N_s$, by using the proposed interference cancelation scheme, there always exists $\mathbf{V}_r^p$ to satisfy the condition in (6.3.9), hence no interference is introduced to the secondary network. As a result, each destination can decode the desired data symbols and thus DoF of $\min\{M_s, N_s\}$ is achieved at the secondary network when $N_r = M_r \geq d_p + \min\{M_s, N_s\}$.

The same strategy employed in *Remark 2* is applicable here as well. The constraint in the optimization problem (6.3.14) should be replaced by (6.3.15). Since the constraint in (6.3.15) is affine, the optimization problem can be solved using convex techniques. Therefore, the proposed beamforming scheme will be able to obtain diversity gain while attaining maximum

DoF.

Case C. $M_p \geq N_p, M_s < N_s$

As similar to the case B, i.e., $M_p < N_p$, $M_s \geq N_s$, the interference cancelation constraint can be applied for the secondary network as follows,

$$\mathbf{U}_s^H \mathbf{H}_{sr} \mathbf{V}_r^p = -\mathbf{U}_s^H \mathbf{H}_{sp} \mathbf{V}_p. \tag{6.3.16}$$

The interference introduced from both the primary transmitter and the relay are canceled after the receiver beamforming matrix $\mathbf{U}_s$. There always exists $\mathbf{V}_r^p$ to satisfy the condition in (6.3.16), for which the proof is similar to the one presented in Appendix D. Since $M_p \geq N_p$, by using the proposed interference cancelation scheme, there always exists $\mathbf{V}_r^s$ to satisfy the condition in (6.3.10), hence no interference is introduced to the secondary network. As a result, each destination can decode the desired data symbols and thus $\min\{M_s, N_s\}$ DoF are achieved at the secondary network when $N_r = M_r \geq d_p + \min\{M_s, N_s\}$.

The strategy employed in *Remark 1* is also applicable here. The constraint in the optimization problem (6.3.12) should be replaced by (6.3.16). Since (6.3.16) is affine, the optimization problem can be solved using convex techniques. Therefore, the proposed beamforming scheme not only cancels the interference to the SU receiver, but also helps the primary network to transmit the data symbols to the PU receiver.

Case D. $M_p < N_p, M_s < N_s$

By employing the proposed interference cancelation scheme as in (6.3.15) and (6.3.16), the interference introduced to both the primary and the SU receivers is canceled. As a result, each destination can decode the desired data symbols and thus $\min\{M_s, N_s\}$ DoF is achieved at the secondary network

when $N_r = M_r \geq d_p + \min\{M_s, N_s\}$. By combining all four different cases in terms of different number of antennas settings, *Theorem 1* has been proven that the maximum achievable degrees of freedom for MIMO cognitive relay networks is $d_s \leq \min\{M_s, N_s\}$.

6.3.3 The Maximum Achievable Degrees of Freedom for Interference Alignment Based MIMO Cognitive Relay Networks

Suppose the primary link operates at its highest transmission rate in the absence of interference, and interference is not strictly allowed in the signal space at the PU receiver. Hence, the proposed interference cancelation scheme is unsuitable. This is because the desired signal $\mathbf{s}_p$ forwarded from the relay will cause inter-stream interference to the PU receiver. Hence the proposed interference cancelation technique can not be applied directly. As a result, IA scheme is now introduced to solve this problem. The basic idea is to ensure the received signal vectors at the PU receiver are aligned away from the used eigen modes at the PU receiver. Hence, no interference is introduced to the primary link and the primary network achieves the maximum capacity using this transmission technique. Next, the IA problem has been formulated for the cognitive relay network.

Following the results in [67], the optimal pre-processing and post-processing schemes for the primary network should be as follows. It is assumed that the primary transmitter is oblivious to the presence of the SUs. Let the SVD of the channel matrix $\mathbf{H}_{pp}$ is given by $\mathbf{H}_{pp} = \mathbf{U}_{pp}\mathbf{\Lambda}_{pp}\mathbf{V}_{pp}^H$, where $\mathbf{\Lambda}_{pp}$ is an $N_p \times M_p$ diagonal matrix consisting of the singular values of the channel matrix $\mathbf{H}_{pp}$. The matrices $\mathbf{U}_{pp}$ and $\mathbf{V}_{pp}$ contain the left and the right singular vectors of $\mathbf{H}_{pp}$, respectively. The primary transmitter and the receiver employs $\mathbf{V}_p = \mathbf{V}_{pp}$ and $\mathbf{U}_p = \mathbf{U}_{pp}$ as their precoding and postprocessing matrices respectively. The power allocation matrix $\mathbf{P}_{pp}$ is designed

to maximize the rate of the primary link under a power constraint as follows,

$$\begin{aligned} \max_{\mathbf{P}_{pp} \succeq 0} \quad & \log_2 |\mathbf{I}_{N_p} + \frac{1}{\sigma^2} \mathbf{H}_{pp} \mathbf{V}_{pp} \mathbf{P}_{pp} \mathbf{V}_{pp}^H \mathbf{H}_{pp}^H| \\ \text{s.t.} \quad & \text{tr}(\mathbf{P}_{pp}) \leq P_p \end{aligned} \tag{6.3.17}$$

where the nonzero elements of the $N_p \times M_p$ diagonal matrix $\mathbf{P}_{pp}$ are the power allocated to different symbols, i.e., $\{P_l\}_{l=1}^{d_p}$, and P_p is the total power available at the primary transmitter. The solution to (6.3.17) is the well-known water filling algorithm [80] where the power of the l^{th} symbol is given by water filling over the nonzero singular values λ_l of the matrix as

$$P_l = (\beta - \frac{\sigma^2}{\lambda_l^2})^+ \tag{6.3.18}$$

where β is the water level that should satisfy the power constraint. The primary transmitter employs a water-filling solution leaving unused eigenmodes at the PU receiver for possible IA.

The required conditions have now been investigated for IA. In order to ensure that both the secondary transmitter and the relay transmitter do not generate any interference to the PU receiver, the following conditions on the precoders of the SUs are required

$$\mathbf{U}_p^H \mathbf{H}_{ps} \mathbf{V}_s = \mathbf{0} \tag{6.3.19}$$

$$\mathbf{U}_p^H \mathbf{H}_{rs} \mathbf{V}_r = \mathbf{0} \tag{6.3.20}$$

Furthermore, to guarantee that the dimension of the desired signal space at the SU receiver is d_s (in order to enable the secondary link to achieve d_s DoF), the following condition on the secondary link is required as rank$\{\mathbf{U}_s^H \mathbf{H}_{ss} \mathbf{V}_s\} = d_s$.

Theorem 2: Suppose the primary network operates with its maximum possible capacity using SVD based encoding and decoding scheme, no in-

terference signal is allowed in its receiver signal space, and the number of antenna at the relay N_r is not less than the $\min\{M_s+d_p, N_s\}$. The maximum achievable DoF of the cognitive radio network with M_s transmit antennas and N_s receive antennas operating in the presence of primary network with d_p DoF is $d_s \leq \min\{M_s, (N_s - d_p)^+\}$.

Proof: Similar to *Theorem 1*, there are two different source signals, $\mathbf{H}_{rp}\mathbf{x}_p$ from the primary transmitter and $\mathbf{H}_{rs}\mathbf{x}_s$ from the secondary transmitter. As the channel transfer matrices $\mathbf{H}_{rp}$ and $\mathbf{H}_{rs}$ are the realization of a random matrix, as $d_s = \min\{M_s, (N_s - d_p)^+\}$, and as the number of antennas at the relay N_r is not less than the $\min\{M_s + d_p, N_s\}$, the vectors $\mathbf{H}_{rp}\mathbf{x}_p$ and $\mathbf{H}_{rs}\mathbf{x}_s$ are linearly independent with a probability reaching one at the relay receiver. The relay sees $d_p + d_s$ independent data symbols from the two source nodes.

First, the relay estimates $d_p + d_s$ data symbols from the two source nodes which are $\mathbf{s}_p$ and $\mathbf{s}_s$. The estimation of the symbols at the relay is written as

$$\mathbf{s}_r^1 = \mathbf{s}_p, \quad \mathbf{s}_r^2 = \mathbf{s}_s. \tag{6.3.21}$$

The relay employs two beamforming matrices for carrying $\mathbf{s}_r^1$ and $\mathbf{s}_r^2$. In particular, the relay employs $\mathbf{V}_r^p$ to align the interference at the PU receiver and employs $\mathbf{V}_r^s$ to cancel $\mathbf{s}_s$ at the PU receiver.

$\mathbf{V}_r^p$ has been designed to align the interference at the PU receiver. Define the effective channel from the relay to the PU receiver as $\bar{\mathbf{H}} \triangleq \mathbf{U}_p^H \mathbf{H}_{pr}$, which has a block structure [80]

$$\bar{\mathbf{H}} = \begin{bmatrix} \bar{\mathbf{H}}_1 \\ \bar{\mathbf{H}}_2 \end{bmatrix} \tag{6.3.22}$$

where the dimension of $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$ are $d_p \times M_r$ and $(N_p - d_p) \times M_r$ respectively. Hence, to satisfy the IA condition in (6.3.20), the transmit beam-

forming matrix at the relay should satisfy the condition as follows

$$\bar{\mathbf{H}}_1 \mathbf{V}_r^p = \mathbf{0}_{d_p \times L} \tag{6.3.23}$$

where L is the dimension of the null space of the matrix $\bar{\mathbf{H}}_1$. Hence, $\mathbf{V}_r^p$ can guarantee IA condition in (6.3.20). The relay transmitter has to avoid interfering with the signal space used by the primary transmitter. This is achieved by aligning interference from the relay with the unused receiver dimensions of the primary link. Similar to [80], it appears that L columns of matrix $\mathbf{V}_r^p$ belong to the null space of $\bar{\mathbf{H}}_1$. Therefore matrix $\mathbf{V}_r^p$ is in the space spanned by the last L columns of matrix $\mathbf{V}_{\bar{\mathbf{H}}_1}$, where $\bar{\mathbf{H}}_1 = \mathbf{U}_{\bar{\mathbf{H}}_1} \mathbf{\Lambda}_{\bar{\mathbf{H}}_1} \mathbf{V}_{\bar{\mathbf{H}}_1}^H$ is the SVD of $\bar{\mathbf{H}}_1$ with $\mathbf{U}_{\bar{\mathbf{H}}_1}$ and $\mathbf{V}_{\bar{\mathbf{H}}_1}$ two unitary matrices of respective sizes $d_p \times d_p$ and $M_r \times M_r$, $\mathbf{\Lambda}_{\bar{\mathbf{H}}_1}$ is a $d_p \times M_r$ matrix containing the elements $(\lambda^2_{\bar{\mathbf{H}}_{1,1}}, \cdots, \lambda^2_{\bar{\mathbf{H}}_{1,\min\{d_p,M_r\}}})$ on its main diagonal and zeros elsewhere, such that $\lambda^2_{\bar{\mathbf{H}}_{1,1}} \geq \cdots \geq \lambda^2_{\bar{\mathbf{H}}_{1,\min\{d_p,M_r\}}}$. Therefore, the beamforming matrix at the relay should be designed as

$$\mathbf{V}_r^p = \text{span}(\mathbf{v}_{\bar{\mathbf{H}}_1}^{(rank(\bar{\mathbf{H}}_1)+1)}, \cdots, \mathbf{v}_{\bar{\mathbf{H}}_1}^{(M_r)}) \tag{6.3.24}$$

where the column vector $\mathbf{v}_{\bar{\mathbf{H}}_1}^{(i)}$ represents the i^{th} column of the matrix $\mathbf{V}_{\bar{\mathbf{H}}_1}$. To accomplish this goal, the relay beamforming matrix $\mathbf{V}_r^s$ has been chosen as in Section 6.3.2. As a result, $\mathbf{s}_s$ is canceled at the PU receiver. No interference appears at the PU receiver. The received signal at the SU receiver is given by

$$\begin{aligned} \mathbf{y}_s &= \mathbf{H}_{sp}\mathbf{x}_p + \mathbf{H}_{ss}\mathbf{x}_s + \mathbf{H}_{sr}\mathbf{x}_r^1 + \mathbf{H}_{sr}\mathbf{x}_r^2 + \mathbf{n}_s \qquad (6.3.25) \\ &= (\mathbf{H}_{sp}\mathbf{V}_p + \mathbf{H}_{sr}\mathbf{V}_r^p)\mathbf{s}_p + (\mathbf{H}_{ss}\mathbf{V}_s + \mathbf{H}_{sr}\mathbf{V}_r^s)\mathbf{s}_s + \mathbf{n}_s \\ &= [\hat{\mathbf{H}}_1 \;\; \hat{\mathbf{H}}_2][\mathbf{s}_p^T \;\; \mathbf{s}_s^T]^T + \mathbf{n}_s \end{aligned}$$

where $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{H}}_2$ denote effective channel carrying the symbol vector $\mathbf{s}_p$ and $\mathbf{s}_s$ respectively. Since the dimension of $\mathbf{s}_p$ and $\mathbf{s}_s$ are $d_p \times 1$ and $\min\{M_s, (N_s - d_p)^+\} \times 1$ and the number of antennas at the SU receiver is N_s, there is enough spatial dimension for the SU receiver to receive and decode both the signals. Furthermore, from (6.3.25) it is seen that $\mathbf{s}_p$ and $\mathbf{s}_s$ can be decoded by applying zero-forcing since $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{H}}_2$ are linearly independent. As a result, each destination can decode two different independent data symbols and thus $\min\{M_s, (N_s - d_p)^+\}$ DoF is achieved by the secondary network. Therefore, it is higher than the upper bound $\min\{(M_s - d_p)^+, (N_s - d_p)^+\}$ DoF that can be achieved for the two-user MIMO interference channel without the assistance of the relay node.

Remark 3: Suppose at the PU receiver $\bar{\mathbf{H}}_1$ is a full row rank matrix $rank(\bar{\mathbf{H}}_1) = d_p$. Furthermore, suppose there is no adequate subspace at the PU receiver to enable the relay to perform IA, e.g., $M_r - rank(\bar{\mathbf{H}}_1) < length(\mathbf{s}_1^r) = d_p$. However, the beamforming vectors for carrying the data symbols $\mathbf{s}_r^1$ can be designed by sharing the column space spanned by the last $M_r - rank(\bar{\mathbf{H}}_1)$ columns of matrix $\mathbf{V}_{\bar{\mathbf{H}}_1}$ as

$$\mathbf{V}_r^p = [\mathbf{v}_{\bar{\mathbf{H}}_1}^{(rank(\bar{\mathbf{H}}_1)+1)} \cdots \mathbf{v}_{\bar{\mathbf{H}}_1}^{(M_r)}]\mathbf{W} \tag{6.3.26}$$

where $\mathbf{W} \in C^{(M_r - rank(\bar{\mathbf{H}}_1)) \times d_p}$ is a combination matrix.

Since $M_r > d_p$, $\mathbf{V}_r^p$ exists with probability one. All the data symbols $\mathbf{s}_r^1$ arrive at the PU receiver which is orthogonal to the d_p dimension used by the primary link. As can be seen by (6.3.25) that $\mathbf{s}_p$ and $\mathbf{s}_s$ can be decoded by applying zero-forcing since $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{H}}_2$ in this case are still linearly independent. As a result, each destination can decode two different independent data symbols and thus $\min\{M_s, (N_s - d_p)^+\}$ DoF is achieved by the secondary network. Interestingly, when the number of antennas at the secondary transmitter is less than the rank of the first block $\tilde{\mathbf{H}}_1$ of the

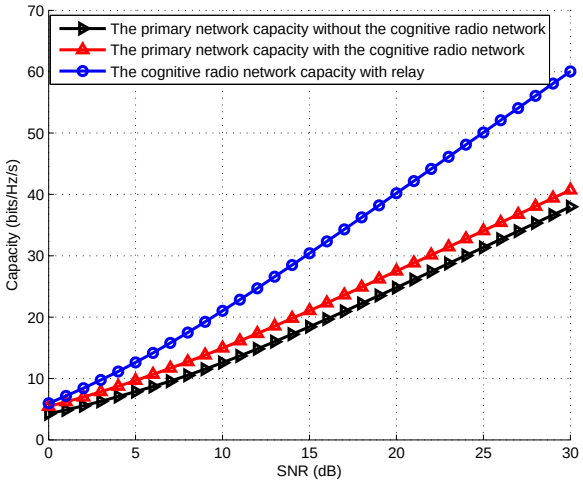

Figure 6.9. The achievable rates for the proposed interference cancelation scheme with a relay and comparison of the results to the rates achieved by a network without a relay ($M_p = 6$ $N_p = 4$, $N_r = M_r = 10$, $M_s = 8$, $N_s = 6$. The DoF achieved by the secondary network is 6).

effective channel $\tilde{\mathbf{H}}_{ps}$ ($\tilde{\mathbf{H}}_{ps} \triangleq \mathbf{U}_p^H \mathbf{H}_{ps}$), i.e., $M_s - rank(\tilde{\mathbf{H}}_1) < 0$, there is no interference subspace that exists for the secondary network to align its interference to the PU receiver. Hence no DoF is available for the secondary network for MIMO cognitive radio network without the involvement of the relay node.

6.3.4 Simulation Results

The performance of the proposed interference cancelation scheme and the proposed combined interference alignment and interference cancelation scheme have been studied in terms of achievable sum rate. The elements of all the channel matrices are assumed to be CSCG variables with zero mean and unity variance. The total power constraints for the primary transmitter, the cognitive transmitter and the relay are set to be identical and varied from 0dB to 30dB. The noise power at the PU receiver and the cognitive receiver are set to unity.

A MIMO cognitive relay network with a primary link operating with

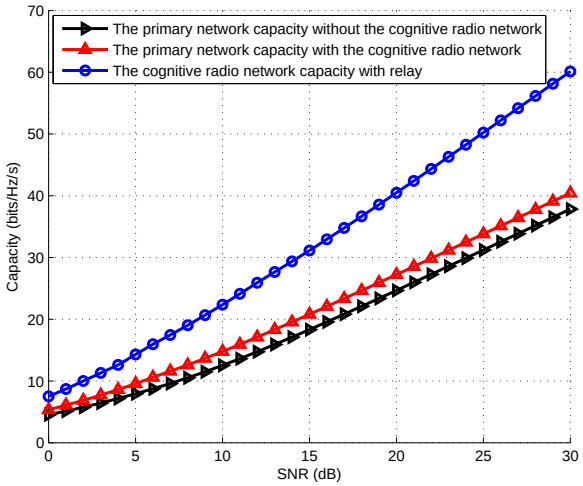

Figure 6.10. The achievable rates for the proposed interference cancelation scheme with a relay and comparison of the results to the rates achieved by a network without a relay ($M_p = 4$ $N_p = 6$, $N_r = M_r = 10$, $M_s = 8$, $N_s = 6$. The DoF achieved by the secondary network is 6).

DoF = 4 has been considered first. Four different cases in terms of different numbers of antenna settings in accordance with section 6.3.2 have been considered. The relay employs ten antennas for transmitting and receiving. For the first case ($M_p \geq N_p$, $M_s \geq N_s$), six transmit antennas and four receive antennas are employed for the primary link. The number of antennas at the secondary transmitter and the receiver is set to eight and six respectively. As seen in Fig. 6.9, the achievable rates for the cognitive radio network at 30 dB is 60 bits/Hz/s which indicates the DoF achieved is approximately six. If a relay is not employed, the cognitive radio network in this case can achieve only two DoF using the conventional IA scheme as shown in [79]. The achievable rates for the primary network with the cognitive relay network is higher than that of the primary network without a relay. This is because the interference cancelation technique not only annihilates interference at the primary destination, but it also assists the primary network to increase the effective channel gain. For the second case ($M_p < N_p$, $M_s \geq N_s$), four transmit antennas and six receiver antennas are employed at the primary

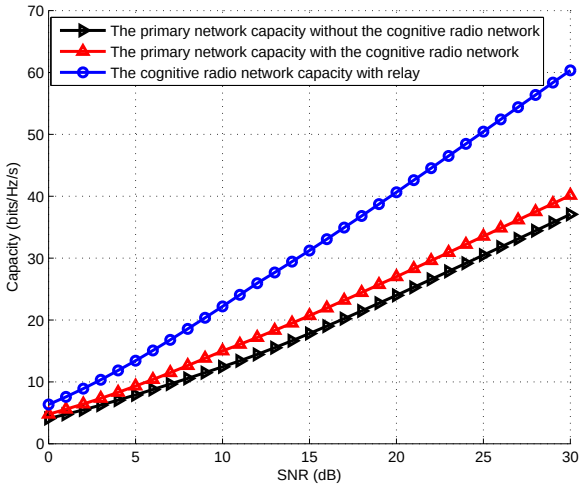

Figure 6.11. The achievable rates for the proposed interference cancelation scheme with a relay and comparison of the results to the rates achieved by a network without a relay ($M_p = 6$ $N_p = 4$, $N_r = M_r = 10$, $M_s = 6$, $N_s = 8$. The DoF achieved by the secondary network is 6).

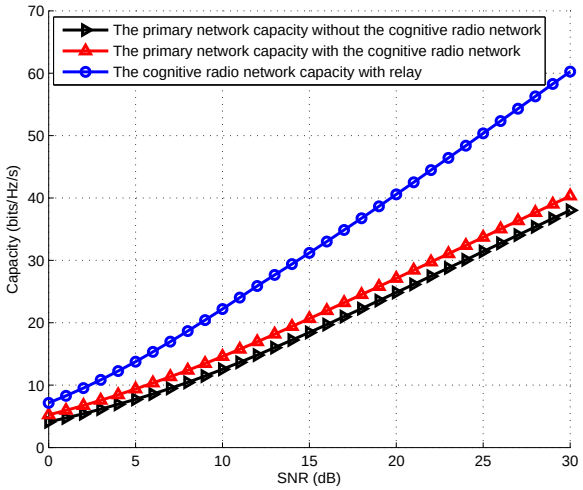

Figure 6.12. The achievable rates for the proposed interference cancelation scheme with a relay and comparison of the results to the rates achieved by a network without a relay ($M_p = 4$ $N_p = 6$, $N_r = M_r = 10$, $M_s = 6$, $N_s = 8$. The DoF achieved by the secondary network is 6).

link. The number of antennas for the secondary network remains the same as that of the first case. As seen in Fig. 6.10, the achievable rates for the cognitive radio network is 60 bits/Hz/s at 30 dB. This indicates that the DoF achieved is approximately six. For the third case ($M_p \geq N_p$, $M_s < N_s$)

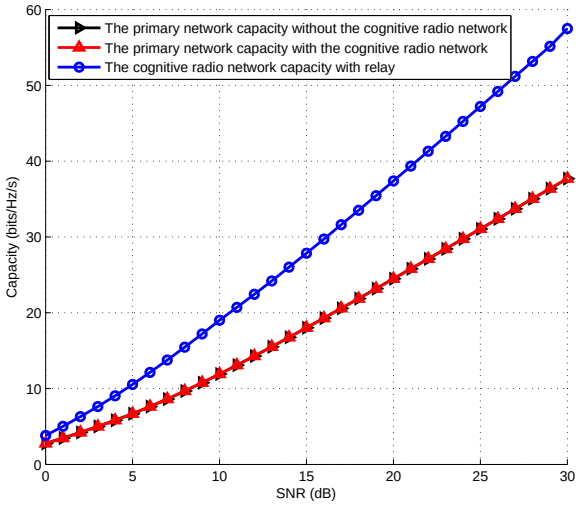

Figure 6.13. The achievable rates for the proposed combined interference alignment and interference cancelation scheme with a relay and comparison of the results to the rates achieved by a network without a relay ($M_p = 6$ $N_p = 4$, $N_r = M_r = 10$, $M_s = 6$, $N_s = 10$. The DoF achieved by the secondary network is 6).

and the fourth case ($M_p < N_p$, $M_s < N_s$), as seen in Fig. 6.11 and Fig. 6.12, the achievable rate for the cognitive radio network is 60 bits/Hz/s at 30 dB. This indicates that the DoF achieved for both the cases is approximately six. Hence, the simulation results for four cases support the result in *Theorem* 1.

The performance of the proposed combined interference alignment and interference cancelation scheme has then been investigated with six transmit antennas and four receiver antennas employed by the primary network. The maximum achievable DoF for the primary network in this case is four. The number of antennas at the secondary transmitter and the receiver is set to six and ten respectively. The relay employs ten antennas. Without employing a relay, the cognitive radio network can achieve only two DoF using the conventional IA scheme. However, as seen in Fig. 6.13, the achievable rates for the cognitive network is 58 bits/Hz/s at 30 dB. This indicates the DoF achieved is approximately six. This result is in accordance with *Theorem* 2. Interestingly, the achievable rate for the primary network with the support

of the cognitive radio relay is equal to that of the primary network without a relay. This is because the relay aligns the signal received from the primary transmitter in the unused space of the PU receiver. Hence there is no effect on the primary network in terms of achievable rates. Furthermore, the capacity of the secondary network is decreased as compared to the proposed IA scheme as shown in Fig. 6.11.

6.4 Conclusion

An interference alignment scheme has been proposed for a network with multiple cells and MIMO users under a Gaussian interference broadcast channel scenario. A grouping method has been extended to the multi-cell scenario to jointly design the transmitter and receiver beamforming vectors using a closed-form expression without a need for iterative computation. The grouping method can ensure no ICI and IUI at each user's receiver while reducing both the number of antennas and the complexity at the basestation as compared to the conventional zero-forcing beamforming scheme. After that, a hybrid interference alignment scheme has been proposed based on the principle of BC-MAC duality. This proposed scheme removes the ICI using interference alignment while maximizing the total capacity of the corresponding cell using BC-MAC duality. Since, interference alignment is not performed explicitly to all users in the network, but the users within each cell are dealt with using capacity maximization, the number of transmit antennas required is generally lower than the existing grouping method. A hybrid rate balancing and interference alignment technique was introduced to maintain fairness among users. The proposed technique is able to maximize the data rate while balancing the rate achieved by each user. To further enhance the performance of the network, a cognitive radio network with an instantaneous relay that allows the secondary transmitter to use the same frequency

band as that of the primary network without generating any interference to the PU within the context of interference cancelation and interference alignment has been considered. The achievable DoF was obtained analytically and shown to agree with the simulation results.

Chapter 7

SUMMARY, CONCLUSION AND FUTURE WORK

The thesis has four contributing chapters and the conclusion of each chapter is summarized below, followed by a discussion on future works.

7.1 Summary and Conclusions

This thesis has studied various resource allocation and spatial diversity techniques problems for spectrum sharing networks. Resource allocation techniques for OFDMA based cognitive radio networks have been introduced in Chapter 3. Resource allocation problem is first considered for an OFDMA-based cognitive radio network. This problem was formulated using an integer linear programming framework and solved using branch and bound method. The solution to this problem optimally allocated the subcarriers, power and bits to secondary users while satisfying the data rate and bit error rate requirements for each secondary user. It ensured that the interference leakage to the primary users was always less than a specific threshold. To admit as many secondary users as possible in the network while allocating subcarriers and bits to each admitted user in such a way that the interference leakage to primary users is below a specific threshold, admission control techniques for an OFDMA based cognitive radio network was also investigated in Chapter 3. A suboptimal optimization algorithm based on integer programming was

proposed to optimally choose a subset of users from a larger set of users and to jointly allocate power and subcarriers to each admitted users to satisfy certain constraints.

Chapter 4 studied the capacity optimization problems for spectrum sharing networks. A weighted sum rate maximization and rate balancing problems have been investigated for spectrum sharing MIMO OFDM based broadcast channels. The aim was to maximize the sum rate of the secondary users whilst ensuring interference leakage to the primary user terminals is below a specific value and each secondary user attains a specific portion of the total data rate. This problem was solved by converting the MIMO OFDM channel into block diagonal form and using the principle of MAC-BC duality. The algorithm in its dual form has been solved using sub-gradient methods. The simulation results demonstrated the convergence of the algorithm and the simultaneous satisfaction of maximum power and interference constraints. A sum rate maximization problem for spectrum sharing MIMO broadcast channels under Rayleigh fading has then been studied. The receiver side was assumed to have the perfect CSI, however the transmitter side has partial CSI in the form of channel covariance matrix feedback. Based on multiple auxiliary variables, KKT optimality conditions and broadcast channel-multiple access channel duality, an iterative algorithm was developed to solve the equivalent problem using the Lagrangian theory and it was shown that the proposed algorithm converges to the setting as defined by the optimization problem, i.e., the interference and the power limit are met whilst maximizing the sum capacity. Finally, a weighted sum rate maximization and rate balancing problem were proposed for a spectrum sharing MIMO based wireless relay network. The aim was to maximize the sum rate of the wireless relay network whilst ensuring the interference leakage to the primary user terminals during two time slots are below a specific value. This problem was solved by asymmetrically allocating the power to different time

slots and using the principle of MAC-BC duality.

Chapter 3 and 4 investigated resource allocation and spatial multiplexing problems for an underlay cognitive radio network. However, in the overlay cognitive radio network, the secondary users coexist with primary users and use part of the transmission power to relay the primary users' signals to the primary receiver. This assistance will offset the interference caused by the secondary user transmissions at the primary users' receiver and hence no loss in primary users' signal-to-noise ratio by secondary users spectrum access. Chapter 5 focused on beamforming and power control for an overlay cognitive radio network. Co-existence of a secondary network with a primary network under an overlay framework was considered. By using cooperation in the cognitive radio network, the aim was to minimize the total power consumed by the networks while satisfying each user's SINR requirement. Since the original optimization problem was non-convex, a dual bi-section based second order cone programming algorithm has been proposed and it has been shown that the proposed algorithm greatly improves the power saving as compared to a network without any cooperations. The simulation results demonstrate the convergence of the algorithm and power saving for both the primary and the secondary networks.

Interference mitigation techniques have become an important part of wireless network design. An interference alignment technique was proposed recently as an efficient capacity achieving scheme at high SNR regime. Interference alignment techniques were studied for various spectrum sharing networks in Chapter 6. An interference alignment algorithm was proposed for a multi-cell MIMO Gaussian interfering broadcast channels. A grouping method already known in the literature has been extended to this multiple-cells scenario and transmit and receiver beamforming vectors were jointly designed using a closed-form expression without a need for iterative computation. Based on the grouping method, a new approach using the principle of

MAC-BC duality was proposed to perform interference alignment while maximizing capacity of users in each cell. The proposed approach was shown to outperform the extension of the grouping method in terms of capacity and basestation complexity. A MIMO cognitive radio network with a MIMO relay that opportunistically accesses the same frequency band as that of a MIMO primary network was then considered. In particular both interference cancelation and interference alignment techniques have been investigated to enhance the achievable degrees of freedom for the MIMO cognitive radio network. The degrees of freedom obtained by the cognitive radio network in the presence of a MIMO relay has been shown to be higher than that could be obtained without a relay. The analyses considered both sufficient and insufficient number of antennas at the relay in terms of the ability to separate and decode both the primary and secondary transmitted signals.

All the techniques proposed in the thesis facilitate shared use of spectrum by multiple transmitters and networks without harmfully interfering each other. Hence the techniques proposed enhances the efficient use of the radio spectrum.

7.2 Future Work

The potential areas for future research have been recognized. Some of the algorithms proposed in this thesis, were developed based on the assumption that the secondary network basestation has perfect knowledge of the channel state information of the users. It is important to extend these algorithms for the case when only imperfect channel knowledge is available to the secondary network basestation. In addition to the evaluation of the physical layer performance for the secondary networks and primary networks for various channel fading environment, the performance evaluation of the networks based on the information theoretical methods will reveal more useful infor-

mation about the networks. Specifically, channel capacity under a spectrum sharing set up with interference constraints to primary users can be derived analytically. Outage capacity defines the rate that can be maintained in all fading states and is a more appropriate capacity notion in wireless systems that carry out real-time delay-sensitive applications. This capacity could be derived analytically in the future work for the techniques discussed in this thesis when only imperfect channel knowledge of the interference link is available to the secondary transmitters.

In Chapter 3, an adaptive radio resource allocation algorithm has been proposed based on BnB method. This algorithm optimally assigns subcarriers and bits for SUs while satisfying SUs' QoS requirements. However, when the number of secondary users and number of available frequency bands for secondary user access increase, the complexity of these algorithm also increases exponentially. This radio resource allocation problem is originally a non-convex problem and NP hard in terms of computational complexity. Therefore, different convex approximations could be investigated and the original problem can be formulated into one of the convex programming techniques.

In Chapter 5, only a single antenna was considered for the secondary receiver. MIMO systems have tremendous potential to increase the data throughput compared to a MISO system. Hence, in cognitive radio networks, employing multiple antennas at the receiver will significantly improve the data throughput and it will be very interesting to extend the work in Chapter 5 to a MIMO scenario. As a result, SOCP technique employed in Chapter 5 can be updated using an SDP framework considered spatial multiplexing to serve multiple secondary users simultaneously with target SINR while imposing constraints on primary user interference with a cost function on total transmit power. Furthermore, it would be very interesting to extend this technique with per antenna power constraints in addition to the aver-

age transmit power constraint. This problem will probably require an SDP relaxation to exclude the non-convex rank-one constraint. If the solution of the relaxed problem does not yield a rank-one solution, randomization techniques can be developed to compensate the Lagrangian relaxation on rank of the matrix. Moreover, in cognitive radio network secondary users and primary users can transmit their data simultaneously. Therefore, employing multiple antennas at the secondary user receivers will also be very beneficial in order to mitigate interference from the primary users. However, this problem will probably turn out to be a non-convex problem and a joint transmitter-receiver beamformer design is generally required. To overcome the non-convex issue, iterative methods can be exploited to design transmitter-receiver beamformers.

Chapter 6 considered a MIMO cognitive radio network with a MIMO relay that opportunistically accesses the same frequency band as that of a MIMO primary network. The analyses considered both sufficient and insufficient number of antennas at the relay in terms of the ability to separate and decode both the primary and secondary transmitted signals. It is assumed that there exists one SU and one PU in our system model. Therefore, the analyses can be extended to the case of multiple SUs and multiple PUs. Furthermore, we focus on the maximum achievable DoF for the MIMO Cognitive Relay Networks. This can be extended to explore the feasibility of interference alignment in signal vector space for this model.

Appendix

A. The case of multiple users in each cell $K \geq 3$

For the case that there are multiple users in each cell $K \geq 3$. For simplicity of explaining the algorithm, first we consider K is a even number. The case for which K is an odd number will be considered later. The matrix equation (6.2.10) for $K \geq 3$ will appear as,

$$\begin{bmatrix} \mathbf{I}_{N_t} & -\mathbf{H}_1^{[1,2]H} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I}_{N_t} & \mathbf{0} & -\mathbf{H}_1^{[2,2]H} & \cdots & \mathbf{0} \\ \vdots & \vdots & \mathbf{0} & \ddots & \vdots \\ \mathbf{I}_{N_t} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{H}_1^{[K,2]H} \end{bmatrix} \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{U}^{[1,2]} \\ \mathbf{U}^{[2,2]} \\ \vdots \\ \mathbf{U}^{[K,2]} \end{bmatrix} = \mathbf{F}_1 \mathbf{X}_1 = \mathbf{0}, \tag{7.2.1}$$

where $\mathbf{G}_1$ accounts for the subspace spanned by the aligned effective interference channels from BS 1 to all the users in the second cell after applying the receiver beamforming.

Similar to (6.2.11) and (6.2.12), we decompose the matrix equation in (7.2.1) into K matrix equations, and the k^{th} equation is shown as follows,

$$\begin{bmatrix} \mathbf{I}_{N_t} & -\mathbf{H}_1^{[k,2]H} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}_{1k} \\ \tilde{\mathbf{U}}^{[k,2]} \end{bmatrix} = \tilde{\mathbf{F}}_k \tilde{\mathbf{X}}_k = \mathbf{0}, \quad k = 1, 2, \cdots, K, \tag{7.2.2}$$

where $\tilde{\mathbf{G}}_{1k}$ accounts for the direction of the interference channels from BS 1 to the user $[k, 2]$ after applying the receiver beamforming matrices. We

find the nullspaces $\tilde{\mathbf{X}}_k$ of matrix $\tilde{\mathbf{F}}_k$. The nullspace $\tilde{\mathbf{X}}_k$ for $k = 1, 2, \cdots, K$ always exists because the size of the matrix $\tilde{\mathbf{F}}_k$ is $N_t \times (N_t + N_r)$. As a result, we can find the solution $\tilde{\mathbf{G}}_{1k}$ and $\tilde{\mathbf{U}}^{[k,2]}$, $k = 1, 2, \cdots, K$.

Since we have decomposed the matrix equation in (7.2.1) into K parts as in (7.2.2), there exists $\frac{K}{2}$ combination matrix pair $\mathbf{C}_{1k}$ and $\mathbf{C}_{1(\frac{K}{2}+k)}$ such that $\tilde{\mathbf{G}}_{1k}$ and $\tilde{\mathbf{G}}_{1(\frac{K}{2}+k)}$ span the same subspace as follows,

$$\tilde{\mathbf{G}}_{1k}\mathbf{C}_{1k} = \tilde{\mathbf{G}}_{1(\frac{K}{2}+k)}\mathbf{C}_{1(\frac{K}{2}+k)}, \quad k = 1, 2, \cdots, \frac{K}{2}. \tag{7.2.3}$$

The matrices $\mathbf{C}_{1k}$ and $\mathbf{C}_{1(\frac{K}{2}+k)}$ can be obtained as

$$[\tilde{\mathbf{G}}_{1k} \ -\tilde{\mathbf{G}}_{1(\frac{K}{2}+k)}] \begin{bmatrix} \mathbf{C}_{1k} \\ \\ \mathbf{C}_{1(\frac{K}{2}+k)} \end{bmatrix} = \hat{\mathbf{G}}_k\hat{\mathbf{C}}_k = \mathbf{0}, \quad k = 1, 2, \cdots, \frac{K}{2}. \tag{7.2.4}$$

Hence, $\tilde{\mathbf{G}}_{1k}\mathbf{C}_{1k}$ is the subspace from BS 1 to the user $[k, 2]$ and user $[(\frac{K}{2} + k), 2]$ respectively after applying the receiver beamforming matrices $\breve{\mathbf{U}}^{[k,2]} = \tilde{\mathbf{U}}^{[k,2]}\mathbf{C}_{1k}$ and $\breve{\mathbf{U}}^{[(\frac{K}{2}+k),2]} = \tilde{\mathbf{U}}^{[(\frac{K}{2}+k),2]}\mathbf{C}_{1(\frac{K}{2}+k)}$. We define this subspace as

$$\mathbf{S}_k = \tilde{\mathbf{G}}_{1k}\mathbf{C}_{1k} = \tilde{\mathbf{G}}_{1(\frac{K}{2}+k)}\mathbf{C}_{1(\frac{K}{2}+k)}, \quad k = 1, 2, \cdots, \frac{K}{2}. \tag{7.2.5}$$

As a result, we have $\frac{K}{2}$ subspaces $\mathbf{S} = [\mathbf{S}_1 \ \mathbf{S}_2 \ \cdots \ \mathbf{S}_{K/2}]$. Our next step is to find the intersection of these $\frac{K}{2}$ subspaces. Again if $\frac{K}{2}$ is even, we determine the combination matrices $\mathbf{Z}_{1k}$ and $\mathbf{Z}_{1(\frac{K}{4}+k)}$ such that subspaces $\mathbf{S}_k$ and $\mathbf{S}_{(\frac{K}{4}+k)}$ span the same space sa follows,

$$\mathbf{S}_{1k}\mathbf{Z}_{1k} = \mathbf{S}_{1(\frac{K}{4}+k)}\mathbf{Z}_{1(\frac{K}{4}+k)}, \quad k = 1, 2, \cdots, \frac{K}{4}. \tag{7.2.6}$$

We will continue this until intersection of all subspaces is determined. If at any stage, the number of subspaces is odd, the same procedure will be applied. For example, we consider $\mathbf{S} = [\mathbf{S}_1 \ \mathbf{S}_2 \ \mathbf{S}_3]$. In this case, we

first determine the intersection of subspaces $\mathbf{S}_1$ and $\mathbf{S}_2$ (let it call $\mathbf{\Theta}_1$), and intersection of $\mathbf{S}_2$ and $\mathbf{S}_3$ (let it call $\mathbf{\Theta}_2$). To find out $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, we solve the following matrix equations

$$[\tilde{\mathbf{S}}_{11} \ -\tilde{\mathbf{S}}_{12}] \begin{bmatrix} \mathbf{C}_{11} \\ \mathbf{C}_{12} \end{bmatrix} = \hat{\mathbf{G}}_1 \hat{\mathbf{C}}_1 = \mathbf{0} \tag{7.2.7}$$

$$[\tilde{\mathbf{S}}_{12} \ -\tilde{\mathbf{S}}_{13}] \begin{bmatrix} \mathbf{C}_{13} \\ \mathbf{C}_{14} \end{bmatrix} = \hat{\mathbf{G}}_2 \hat{\mathbf{C}}_2 = \mathbf{0} \tag{7.2.8}$$

Hence, there left two subspaces which are $\mathbf{\Theta}_1 = \tilde{\mathbf{S}}_{11}\mathbf{C}_{11}$ and $\mathbf{\Theta}_2 = \tilde{\mathbf{S}}_{12}\mathbf{C}_{13}$. We then determine the intersection of $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ by using the combination matrix pair, and hence all the users in the second cell are grouped together such that the ICI channels from BS 1 are aligned in the same subspace.

Since we decompose the problem in (7.2.1) into K matrix equations, the number of arithmetic operations required for the proposed algorithm depends on the complexity of matrix SVD for all the K matrix equations and also to determine intersection of the same subspace for all the K users. Hence, the overall complexity for the proposed low complexity algorithm is bounded below $K\mathcal{O}(N_t^3+N_t^2N_r)+2(K+\lfloor log_2 K \rfloor)\mathcal{O}(2N_t^2N_r)$, which is significantly lower than the complexity of the first method $\mathcal{O}(K^2N_t^3 + K^3N_t^2N_r)$.

B. Proof for the existence of $\mathbf{V}_r^p$ and $\mathbf{V}_r^s$ when $M_p \geq N_p$, $M_s \geq N_s$

We can rewrite equation (6.3.9) as

$$\mathbf{H}_{sr}\mathbf{V}_r^p + \mathbf{H}_{sp}\mathbf{V}_p = \mathbf{0}. \tag{7.2.9}$$

Defining $\mathbf{D} = [\mathbf{H}_{sr}\ \mathbf{H}_{sp}] \in C^{N_s \times (M_r + M_p)}$ and $\mathbf{B} = \begin{bmatrix} \mathbf{V}_r^p \\ \mathbf{V}_p \end{bmatrix}$, hence equation (7.2.9) can be written as

$$\mathbf{DB} = \mathbf{0}. \tag{7.2.10}$$

Hence matrix $\mathbf{B}$ spans the nullspace of matrix $\mathbf{D}$. As the dimension of the matrix $\mathbf{D}$ is $N_s \times (d_s + d_p + M_p)$, the condition for the existence of the nullspace for the matrix $\mathbf{D}$ is $d_s + d_p + M_p > N_s$. Since $M_s \geq N_s$ and $d_s = \min\{M_s, N_s\} = N_s$, $d_s + d_p + M_p > N_s$ is always hold. Hence, there exists a matrix $\mathbf{V}_r^p$ satisfying equation (6.3.9). Similar arguments can be used to prove the existence of $\mathbf{V}_r^s$ to satisfy equation (6.3.10).

C. Proof for the convexity for Remark 1

We design the beamformer $\mathbf{V}_r^p$ such that the primary capacity is maximized through the effective primary channel subject to a SU signal cancelation constraint as follows,

$$\begin{aligned} \max_{\mathbf{V}_r^p \succeq 0} \quad & \log |\mathbf{I} + \frac{(\mathbf{H}_{pp}\mathbf{V}_p + \mathbf{H}_{pr}\mathbf{V}_r^p)(\mathbf{H}_{pp}\mathbf{V}_p + \mathbf{H}_{pr}\mathbf{V}_r^p)^H}{\sigma^2}| \\ \text{s.t.} \quad & \mathbf{H}_{sr}\mathbf{V}_r^p + \mathbf{H}_{sp}\mathbf{V}_p = \mathbf{0} \end{aligned} \tag{7.2.11}$$

where $\mathbf{0}$ is an $N_p \times 1$ vector consisting of all zeros, σ^2 is the noise variance at the PU receiver. Unfortunately, the above optimization problem is non-convex and hence it can not be solved using convex techniques [93]. Here we introduce a method to transfer the non-convex problem to a convex optimization problem. The signal at the PU receiver is written as

$$\mathbf{y}_p = \mathbf{H}_{pp}\mathbf{V}_p\mathbf{s}_p + \mathbf{H}_{pr}\mathbf{V}_r^p\mathbf{s}_r^1 + \mathbf{n}_p \tag{7.2.12}$$

where $\mathbf{s}_r^1 = \mathbf{s}_p$. Since $\mathbf{H}_{sr}\mathbf{V}_r^p = -\mathbf{H}_{sp}\mathbf{V}_p$, after substitute $\mathbf{V}_p$, equation (7.2.12) can be written as

$$\mathbf{y}_p = (\mathbf{H}_{pr} - \mathbf{H}_{pp}(\mathbf{H}_{sp}^H\mathbf{H}_{sp})^{-1}\mathbf{H}_{sp}^H\mathbf{H}_{sr})\mathbf{V}_r^p\mathbf{s}_p + \mathbf{n}_p \tag{7.2.13}$$

Therefore, we define the effective channel $\tilde{\mathbf{H}}_p = \mathbf{H}_{pr} - \mathbf{H}_{pp}(\mathbf{H}_{sp}^H\mathbf{H}_{sp})^{-1}\mathbf{H}_{sp}^H\mathbf{H}_{sr}$. The objective function of the optimization problem (7.2.11) can be written as

$$\max_{\mathbf{G}\succeq 0} \quad \log|\mathbf{I} + \frac{\tilde{\mathbf{H}}_p\mathbf{G}\tilde{\mathbf{H}}_p^H}{\sigma^2}| \tag{7.2.14}$$

where $\mathbf{G} = \mathbf{V}_r^p\mathbf{V}_r^{pH}$ is a positive semidefinite matrix. The objective function is a log determine problem in terms of the positive semidefinite matrix $\mathbf{G}$ and hence it is a concave function. Then we define a matrix $\mathbf{T}$ as

$$\mathbf{T} = (\mathbf{H}_{sp}\mathbf{V}_p)(\mathbf{H}_{sp}\mathbf{V}_p)^H\mathbf{H}_{sr}(\mathbf{H}_{sr}^H\mathbf{H}_{sr})^{-1} \tag{7.2.15}$$

hence the constraint in optimization problem (7.2.11) can be written as

$$\mathbf{H}_{sr}\mathbf{G} + \mathbf{T} = \mathbf{0} \tag{7.2.16}$$

which is an affine function. Since we maximize a concave function subject to an affine function, the optimization problem is convex, which concludes the proof. The optimization problem can be solved using convex techniques. Once we have obtained the optimum solution $\mathbf{G}$, the beamformer $\mathbf{V}_r^p$ can be obtained.

D. Proof for the existence of $\mathbf{V}_r^s$ when $M_p < N_p$, $M_s \geq N_s$

We can rewrite equation (6.3.15) as follows

$$\mathbf{U}_p^H\mathbf{H}_{pr}\mathbf{V}_r^s + \mathbf{U}_p^H\mathbf{H}_{ps}\mathbf{V}_s = \mathbf{0}. \tag{7.2.17}$$

Similar to Appendix B, we define $\mathbf{S} = [\mathbf{U}_p^H \mathbf{H}_{pr}\ \mathbf{U}_p^H \mathbf{H}_{ps}]$ and $\mathbf{A} = \begin{bmatrix} \mathbf{V}_r^s \\ \mathbf{V}_s \end{bmatrix}$. The equation (7.2.17) can be written as

$$\mathbf{SA} = \mathbf{0}. \tag{7.2.18}$$

Hence columns of matrix $\mathbf{A}$ span the nullspace of matrix $\mathbf{S}$. The dimension of the matrix $\mathbf{S}$ is $d_p \times (d_s + d_p + M_p)$. Therefore, the condition for the existence of the nullspace of matrix $\mathbf{S}$ is $d_s + d_p + M_p > d_p$, which is always hold. Hence, there exists $\mathbf{V}_r^s$ satisfying equation (6.3.15).

References

[1] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 768 – 783, May 2004.

[2] "IEEE std 802.11a/d7.0-1999, part 11: wireless lan medium access control (MAC) and physical layer (PHY) specifications: High speed physical layer in the 5ghz band,"

[3] "CDMA industry is ready to meet future market needs head on - CDG and 3GPP2 collaborate to be the frist to deliver advanced ultra mobile broadband technology," Press Release, 15 August 2006.

[4] D. R. Bull, C. N. Canagarajal, and A. R. Nix., *Insight into Mobile Multimedia Communications.* Academic Press, 1999.

[5] Molkdar, W. Featherstone, and S. Lambotharan, "Enhanced general packet radio service (EGPRS), an overview," *IEE Electronics and Communication Magazine*, vol. 14(1), p. 2138, Feb. 2002.

[6] J. Proakis and M. Salehi, *Digital Communications.* McGraw-Hill Science, 5 ed., 2007.

[7] J. Korhonen, *Introduction to 3G Mobile Communications.* 2nd Edition, Artech House, 2003.

[8] "3GPP; technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); LTE physical layer - general de-

scription (release 8), 3GPP TS 36.201 V8.1.0," November 2007. [Available]: http://www.3gpp.org/ftp/Specs/html-info/36201.htm.

[9] "Evolution to UMB marks world's first IP-based mobile broadband solution to enable peak data rates of 288 Mbps with average latencies below 16mesc," Press Release, 24 September 2007, [Available]: http://www.3gpp2.org/Public-html/News/Release-UMBSpecification24SEP2007.pdf.

[10] "WiMAX forum mobile system profile, release 1.0 approved specification (revision 1.4.0: 2007-04-12).," [Available]: http://www.wimaxforum.org/technology/documents/.

[11] H. Holma, A. Toskala, K. Ranta-aho, and J. Pirskanen, "High-speed packet access evolution in 3GPP release 7," *IEEE Commun. Mag.*, vol. 45, no. 12, pp. 29 – 35, Dec. 2007.

[12] R. V. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications.* Artech House Publishers, Jan. 2000.

[13] "IEEE std 802.163-2005 and IEEE std 802.16-2004/cor1-2005, IEEE standard for local and metropolitan area networks, part 16: Air interface for fixed and mobile broadband wireless access systems, amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands," 28 February 2006.

[14] Federal Communications Commission, *ET Docket No 03-222 Notice of proposed rule making and order.* Dec. 2003.

[15] T. A. Weiss and F. K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, vol. 42, no. 3, pp. S8-S14, Mar. 2004.

[16] J. Mitola and G. Q. Maguire, "Cognitive radios: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[17] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 28–40, Feb. 2008.

[18] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* NJ, Prentice-Hall: Englewood Cliffs, 1998.

[20] D. Cabric, A. Tkachenco, and R. W. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," in *ACM 1st Int. Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, Aug. 2006.

[21] S. Enserink and D. Cochran, "A cyclostationary feature detector," in *Proc. Asilomar Conf. Sign., Syst. and Comp., Pacific Grove, CA*, pp. 806–810, Oct. 1994.

[22] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May, 2009.

[23] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications.* Cambridge University Press, 2003.

[24] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 73–79, Feb. 2005.

[25] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[26] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution-HSPA and LTE for Mobile Broadband.* ELSEVIER, 2 ed., 2008.

[27] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[28] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," preprint, available at arxiv.org/abs/0803.3816.

[29] T. Gou and S. A. Jafar, "Degrees of freedom of the K-user M×N MIMO interference channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.

[30] S. A. Jafar and S. Shamai(Shitz), "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inform. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[31] C. Suh and D. Tse, "Interference alignment for celluar networks," in *Proc. 46th Annual Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, Sep. 2008.

[32] V. Nagarajan and B. Ramamurthi, "Distributed cooperative precoder selection for interference alignment," *IEEE Trans. Veh. Tech.*, vol. 59, no. 9, pp. 4368–4376, Nov. 2010.

[33] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. sel. areas commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[34] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas in Commun.*, vol. 21, pp. 171 –178, Feb. 2003.

[35] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 1150 – 1158, Nov. 2003.

[36] S. K. Lai, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Adaptive trellis coded MQAM and power optimization for OFDM transmission," in *Proc. IEEE Vehicular Technology Conf. (VTC99)*, May Houston, TX, 1999.

[37] Z. Mao and X. Wang, "Efficient optimal and suboptimal radio resource allocation in OFDMA system," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 440–445, Feb. 2008.

[38] I. Kim, I. S. Park, and Y. H. Lee, "Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM," *IEEE Trans. Vehic. Tech.*, vol. 55, pp. 1195 – 1207, Jul. 2006.

[39] A. Attar, O. Holland, M. R. Nakhai, and A. H. Aghvami, "Inteference-limited resource allocation for cognitive radio in orthogonal frequency division multiplexing networks," *IET J. Commun.*, vol. 2, no. 6, pp. 806–814, 2008.

[40] G. Bansal, J. Hossain, and V. K. Bhargava, "Adaptive power loading for OFDM-based cognitive radio system," in *Proc. IEEE Int. Conf. Commun.*, pp. 5137–5142, Glasgow, Scotland, June 2007.

[41] S. Wang, "Efficient resource allocation algorithm for cognitive OFDM systems," *IEEE Commun. Lett.*, vol. 14, pp. 725 –727, Aug. 2010.

[42] Y. Zhang and C. Leung, "Resource allocation in an OFDM-based cognitive radio system," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1928-1931, Jul. 2009.

[43] Y. Rahulamathavan, K. Cumanan, L. Musavian, and S. Lambotharan, "Optimal subcarrier and bit allocation techniques for cognitive radio net-

works using integer linear programming," in *Proc. IEEE Statistical Signal Processing Workshop*, pp. 293 – 296, Cardiff, UK, Aug. 2009.

[44] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[45] D. J. Love and R. W. Heath, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[46] I. H. Kim and D. J. Love, "On the capacity and design of limited feedback multiuser MIMO uplink," *Linear Algebra Its Applications*, vol. 396, pp. 373–384, Feb. 2005.

[47] T. Pande, D. J. Love, and J. V. Krogmeier, "Reduced feedback MIMO-OFDM precoding and antenna selection," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2284–2293, may 2007.

[48] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Trans. Inform. Theory*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[49] M. Bengtsson and B. Ottersten, *Handbook of Antennas in Wireless Communications*. Optimal and Suboptimal Transmit Beamforming: Boca Raton, FL:CRC, Aug. 2001.

[50] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Allerton Conf. Commun. Control and Comp.*, pp. 987–996, Sept. 1999.

[51] J. Lofberg, "Yalmip: A toolbox for modelling and optimizationin MATLAB," in *Proc. IEEE Int. Symp. on Comp. Aided Control Sys. Design*, pp. 284–289, Taipei, Sept. 2004.

[52] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming." *Optimization Methods and Software*, Feb. 2007. Available [online]: http://www.stanford.edu/ boyd/cvx/V.1.0RC3.

[53] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11-12, pp. 625–653, 1999. Special issue on Interior Point Methods (CD supplement with software).

[54] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Vechicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[55] H. Boche and M. Schubert, "Solution of the SINR downlink beamforming problem," in *Proc. 36th Conf. Inform. Sci. Syst. (CISS)*, Princeton, USA, Mar. 2002.

[56] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE Vehicular Technology Conference (VTC), Fall*, Vancouver, Canada, Sept. 2002.

[57] H. Boche and M. Schubert, "A unifying theory for uplink and downlink multiuser beamforming," in *Proc. IEEE Intern. Zurich Seminar*, pp. 27–1–27–6, Zurich, Switzerland, Feb. 2002.

[58] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR balancing technique for downlink beamforming in cognitive radio networks," *IEEE Sig. Process. Lett.*, vol. 17, no. 2, pp. 133–136, Feb. 2010.

[59] D. Palomar and M. Lagunas, "Joint transmit-receive space-time equalization in spatially correlated MIMO channels: a beamforming approach," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 730–743, Jun 2003.

[60] D. Palomar, M. Lagunas, and J. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1179–1197, May 2004.

[61] D. Palomar, J. Cioffi, and M. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sept. 2003.

[62] E. Telatar, "Capacity of multi-antenna Gaussian channels," in *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[63] A. Goldsmith, *Wireless Communications.* Cambridge University Press, 2005.

[64] G. Strang, *Linear Algebra and its Applications.* Fourth edition: Thomson Brooks/Cole, 2006.

[65] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Sys. Tech. Journal*, pp. 379423, 623656, 1948.

[66] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, vol. IT-30, no. 1, pp. 44–53, Jan. 1984.

[67] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585 – 596, Nov. 1999.

[68] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[69] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658 – 2668, Oct. 2003.

[70] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1570 – 1580, May 2005.

[71] W. Yu, "Sum-capacity computation for the Gaussian vector broadcast channel via dual decomposition," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 754 – 759, Feb. 2006.

[72] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1640 – 1646, Aug. 2006.

[73] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 361 – 374, Feb. 2006.

[74] W. Yu, , and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2646 – 2660, Jun. 2007.

[75] R. Zhang, Y. C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE Trans. Wireless Commun.*, vol. 27, no. 5, pp. 699–712, Jun. 2009.

[76] R. Etkin and E. Ordentlich, "On the degrees-of-freedom of the K-user Gaussian interference channel," *IEEE Trans. Inform. Theory*, submitted.

[77] V. R. Cadambe, S. A. Jafar, and S. Shamai(Shitz), "Interference alignment on the deterministic channel and application to fully connected AWGN interference networks," *IEEE Trans. Inform. Theory*, vol. IT-55, no. 1, pp. 269–274, Jan. 2009.

[78] S. M. Perlaza, M. Debbah, S. Lasaulce, and J. M. Chaufray, "Opportunistic interference alignment in MIMO interference channels," in *Proc. IEEE Conf. Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5, Cannes, France, Sep. 2008.

[79] M. Amir, A. EI-Keyi, and M. Nafie, "Constrained interference alignment

and the spatial degrees of freedom of MIMO cognitive networks," *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2994 – 3004, May 2011.

[80] S. M. Perlaza, N. Fawaz, S. Lasaulce, and M. Debbah, "From spectrum pooling to space pooling : Opportunistic interference alignment in MIMO cognitive networks," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3728 – 3741, July 2011.

[81] Y. Tian and A. Yener, "The Gaussian interference relay channel with a potent relay," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Honolulu, Hawaii, USA, Dec. 2009.

[82] R. Dabora, I. Maric, and A. Goldsmith, "Relay strategies for interference forwarding," in *Proc. IEEE Inform. Theory Workshop (ITW)*, 2008.

[83] O. Sahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Washington D.C., USA, Nov. 2007.

[84] I. Shomorony and A. S. Avestimehr, "Two-unicast wireless networks: Characterizing the sum degrees of freedom," in *arXiv:1102.2498.*

[85] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation and full duplex operation," *IEEE Trans. Inform. Theory*, vol. 55, no. 5, pp. 2334–2344, May 2009.

[86] A. E. Gamal and N. Hassanpour, "Relay without delay," in *Proc. IEEE International Symp. on Inform. Theory (ISIT)*, Adelaide, Australia, Sep. 2005.

[87] N. Lee and S. A. Jafar, "Aligned interference neutralization and the degrees of freedom of the 2 user interference channel with instantaneous relay," *[Online]. Available: arxiv:1102.3833.*

[88] L. Cao and H. Zheng, "Distributed spectrum allocation via local bargaining," in *Proc. IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pp. 119–127, Santa Clara, California, USA, Sept. 2005.

[89] Y. J. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 13, no. 4, pp. 38 – 47, Aug. 2006.

[90] T. Weiss, J. Hillenbrand, A. Krohn, and F. K. Jondral, "Mutual interference in OFDM-based spectrum pooling systems," in *Proc. IEEE Vehicular Technol. Conf.*, vol. 4, pp. 1873–1877, Los Angeles, May. 2004.

[91] H. A. Mahmound and H. Arslan, "Sidelobes suppression in OFDM-based spectrum sharing systems using adaptive symbol transition," in *IEEE Commun. Lett.*, vol. 12, pp. 133–135, Feb. 2008.

[92] S. Haykin, *Communication System.* John Wiley and Sons, New york, 4 ed., 1994.

[93] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, Cambridge, UK, 2004.

[94] G. L. Nemhauser and L. A. Wolsey, *Integer and combinatorial optimization.* John Wiley and Sons, New york, 1998.

[95] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313 – 324, Feb. 2003.

[96] J. Gross, "Admission control based on OFDMA channel transformations," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1–11, Kos, June 2009.

[97] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311 – 335, 1998.

[98] M. Costa, "Writing on dirty paper," *IEEE. Trans. Inform. Theory*, vol. IT-29, no. 5, pp. 439 – 441, May 1983.

[99] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1691 – 1706, July 2003.

[100] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *Proc. 38th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 46 2000.

[101] P. Tejera, W. Utschick, J. A. Nossek, and G. Bauch, "Rate balancing in multiuser MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1370 – 1380, May 2009.

[102] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632 – 2639, Sep. 2001.

[103] H. Boche and E. Jorswieck, "On the optimality range of beamforming for MIMO systems with covariance feedback," *IEICE Trans. Commun.*, vol. E85-A, no. 11, pp. 2521 – 2528, Nov. 2002.

[104] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165 – 1175, Jul. 2004.

[105] A. Soysal and S. Ulukus, "Optimality of beamforming in fading MIMO multiple access channel," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1171 – 1183, Apr. 2009.

[106] A. Soysal and S. Ulukus, "Optimum power allocation for single-user MIMO and multi-user MIMO-MAC with partial CSI," *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1402 – 1412, Sep. 2007.

[107] R. Zhang and Y. C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Sig. Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[108] D. G. Luenberger, *Optimization by Vector Space Methods.* New york : John Wiley, 1969.

[109] H. Viswanathan, S. Venkatesan, and H. Huang, "Downlink capacity evaluation of cellular networks with known-interference cancellation," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 802 – 810, Jun. 2003.

[110] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," 2003, [Online]. Available: http://mit.edu/6.976/www/notes/sub gradmethod.

[111] C. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637 – 650, Mar. 2002.

[112] T. M. Cover and A. A. E. Gamal, "Capacity theorem for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 5, p. 572584, Sept. 1979.

[113] K. Lee and A. Yener, "Outage performance of cognitive wireless relay networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Francisco, CA, Nov. 2006.

[114] L. Musavian and S. Aissa, "Cross-layer analysis of cognitive radio relay networks under quality of service constraints," in *Proc. IEEE Vehicular Technol. Conf.*, pp. 1–5, Barcelona, Apr. 2009.

[115] L. Musavian, S. Aissa, and S. Lambotharan, "Effective capacity for

interference and delay constrained cognitive radio relay channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1698–1707, May 2010.

[116] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Sig. Process. Mag.*, vol. 14, no. 6, pp. 49–83, Nov. 1997.

[117] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications - a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[118] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1939–1948, Nov. 2003.

[119] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversitypart II: Implementation aspects and perfromance analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[120] Y. W. Hong, W. J. Huang, F. H. Chiu, and C. C. J. Kuo, "Cooperative communications resource constrained wireless networks," *IEEE Sig. Process. Mag.*, vol. 24, no. 3, pp. 47–57, May 2007.

[121] G. Zheng, K. K. Wong, A. Paulraj, and B. Ottersten, "Robust collaborative-relay beamforming," *IEEE Trans. Sig. Process.*, vol. 57, no. 8, pp. 3130–3143, Aug. 2009.

[122] G. Zheng, K. K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation," *IEEE Sig. Process. Lett.*, vol. 16, no. 4, pp. 257–260, Apr. 2009.

[123] M. M. Abdallah and H. C. Papadopoulos, "Beamforming algorithms for information relaying in wireless sensor networks," *IEEE Trans. Sig. Process.*, vol. 56, no. 10, pp. 4772–4784, Oct. 2008.

[124] R. Zhang, C. C. Chai, and Y. C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Trans. Sig. Process.*, vol. 57, no. 2, pp. 726–737, Jun. 2009.

[125] J. Jingon and A. H. Sayed, "Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems," *IEEE Trans. Sig. Process.*, vol. 58, no. 3, pp. 1833–1846, Mar. 2010.

[126] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, pp. III–473–III–476, Honolulu, HI, Apr. 2007.

[127] R. Krishna, Z. Xiong, and S. Lambotharan, "A cooperative MMSE relay strategy for wireless sensor networks," *IEEE Sig. Process. Lett.*, vol. 15, pp. 549–552, 2008.

[128] R. Krishna, K. Cumanan, Z. Xiong, and S. Lambotharan, "A novel cooperative relaying strategy for wireless networks with signal quantization," *IEEE Trans. Veh. Tech.*, vol. 59, no. 1, pp. 485–489, Jan. 2010.

[129] X. Hong, Z. Chen, C. X. Wang, S. Vorobyov, and J. Thompson, "Cognitive radio networks," *IEEE Veh. Tech. Mag.*, vol. 4, no. 4, pp. 76–84, Dec. 2009.

[130] L. Zhang, Y. C. Liang, Y. Xin, and V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143–4153, Aug. 2009.

[131] G. Zheng, S. Ma, K. K. Wong, and T. S. Ng, "Robust beamforming in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 570–576, Feb. 2010.

[132] G. Zheng, K. K. Wong, and B. Ottersten, "Robust cognitive beamform-

ing with bounded channel uncertainties," *IEEE Trans. Sig. Process.*, vol. 57, no. 12, pp. 4871–4881, Dec. 2009.

[133] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Allerton Conf. Commun. Control and Comp.*, pp. 987–996, Monticello, IL, Sep. 1999.

[134] 3GPP TR 36.814 V9.0.0, "Further advancements for E-UTRA physical layer aspects (release 9)," Oct. 2006.

[135] S. H. Park and I. Lee, "Degrees of freedom and sum rate maximization for two mutually interfering broadcast channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dresden, Germany, June 2009.

[136] J. Kim, S. H. Park, H. Sung, and I. Lee, "Sum rate analysis of two-cell MIMO broadcast channels: spatial multiplexing gain," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 2010.

[137] C. Suh, M. Ho, and D. Tse, "Downlink interference alignment," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2616–2626, Sept. 2011.

[138] W. Shin, N. Lee, J.-B. Lim, C. Shin, and K. Jang, "On the design of interference alignment scheme for two-cell MIMO interfering broadcst channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 437–442, Feb. 2011.

[139] M. P. Holmes, A. G. Gray, and C. L. Isbell, "Fast SVD for large-scale matrices," College of Computing, Georgia Institute of Technology, Atlanta, GA.

[140] H. Weingarten, Y. Steinberg, and S. S. (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.