Coverage and Throughput Enhancement in MIMO Cellular Networks via Cooperative Relaying

by

Obumneme Godfrey Okeke

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Communications

Department of Electrical and Computer Engineering

University of Alberta

Abstract

One of the major challenges of broadband cellular systems is delivering high data rates to the users, especially those at the cell edges. In this thesis, transceiver designs and scheduling algorithms for enhancing cell-edge user experience in emerging and future broadband wireless cellular networks are investigated. In particular, the use of fixed infrastructure-based non-regenerative and altruistic (i.e., not having their own data to transmit) relays for coverage extension (at high data rates) in cellular networks with multiple antenna transmitters and receivers is considered. In order to mitigate the high interference jeopardizing the high data rate promise of multiple-input multiple-output (MIMO) spatial multiplexing in cellular networks, coordinated multipoint (CoMP) transmission/reception is employed.

Leveraging *MIMO, CoMP and relay* techniques in a unified manner, we jointly design the input covariance and relay beamforming matrices to maximize the system data rate. For the single-cell MIMO downlink, uplink-downlink duality is first proved and then employed to design the downlink parameters from more tractable uplink channel. For the multi-cell MIMO network, interference mitigation is incorporated into the relay designs in order to decouple the correlated relay transmissions due to interference from unintended sources. Furthermore, we propose novel low complexity joint user-relay selection/scheduling and/or association strategies to further improve the system performance, which lead to low complexity and overhead. Simulation results demonstrate the effectiveness of the proposed schemes.

The favorable performance and low complexity of the proposed schemes make them very attractive for possible implementation in emerging and future broadband wireless networks.

Dedicated to
The Almighty God
&
My Parents, who are always understanding, but may not understand everything in this thesis

Acknowledgements

My deepest gratitude goes to my supervisors, Dr. Witold A. Krzymień and Dr. Yindi Jing for their unparalleled guidance and support throughout this work.

Also, I will like to say a very big thank you to the Telecommunications Research Laboratories (TRTech), TELUS Communications, Mitacs, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and Rohit Sharma Professorship for their funding of my work. In particular, I would like to thank Dr. Jordan Melzer for his support and useful feedback especially on our joint work during my Mitacs internship with TELUS.

The extensive numerical simulations and computations involved in this work would not have been possible without the computing resources of Information Services and Technology (IST) at the University of Alberta, TRTech, WestGrid and Compute/Calcul Canada. Thank you very much.

Also, I wish to thank members of my PhD final examination committee: Dr. Xianbin Wang, Dr. Yindi Jing, Dr. Hai Jiang, Dr. Chintha Tellambura and Dr. Witold A. Krzymień. Your valuable comments and feedback are highly appreciated.

Special thanks to my research colleagues at the University of Alberta and TRTech. Your friendship and infectious camaraderie throughout my Ph.D. studies will always be remembered.

Finally, to my parents, siblings, and friends for their support, patience and understanding throughout this seemingly unending journey. I am forever indebted to you all.

Contents

List of Tables

List of Figures

List of Symbols

$\mathbf{0}$ All zero matrix

$\alpha_{r,j}$ Distance-dependent channel gain between the rth relay station and the jth base station

$\beta_{i,r}$ Distance-dependent channel gain between the ith user and the rth relay station

C Number of cells/BSs within a coordinated cluster

$|\mathcal{X}|$ Cardinality of set $\mathcal{X}$

$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ Complex Gaussian random variable with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$

$\mathbf{D^B}$ Relay amplification/beamforming matrix in the single-cell downlink

$|\mathbf{A}|$ Determinant of matrix $\mathbf{A}$

$\text{diag}(x_1, x_2, ..., x_K)$ Diagonal matrix with diagonal elements $x_1, x_2, ..., x_K$

$\mathbf{D^M}$ Relay amplification/beamforming matrix in the single-cell uplink

$\mathbf{X} \cdot \mathbf{Y}$ Element-wise multiplication of $\mathbf{X}$ and $\mathbf{Y}$

$\mathbf{D}_r$ Amplification/beamforming matrix of the rth relay station in the multi-cell network

$\mathbb{E}[\cdot]$ Expectation operator

$\|\mathbf{A}\|_F$ Frobenius norm of matrix $\mathbf{A}$

$\hat{\mathbf{G}}_r$ Aggregate channel matrix of the users assigned to the rth relay station in the multi-cell network

$\mathbf{G}_i$ Access link channel matrix from the relay station to the ith user in a single-cell single-relay downlink

$\mathbf{G}_{i,r}$ Access link channel matrix from the ith user to the rth relay station in a multi-cell multi-relay network

$\mathbf{G}_k(i,j)$ Channel gain from Antenna j of the RS to Antenna i of User k in the single-cell downlink or the gain from Antenna i of the transmitter User k to Antenna j of the relay in the single-cell uplink

$\mathbf{G}_r$ Aggregate channel matrix of the access link channel of the rth relay station in the multi-cell network

$\nabla_{\mathbf{X}}$ Gradient with respect to $\mathbf{X}$

$\tilde{\mathbf{G}}_r$ Aggregate channel matrix of the users not assigned to the rth relay station in the multi-cell network

$\mathbf{H}$ Relay link channel matrix from the base station to the relay station in the single-cell single-relay downlink

$(.)^{\mathsf{H}}$ Hermitian (conjugate) transposition of a matrix

$\mathbf{H}_r$ Aggregate channel matrix of the relay link channel of the rth relay station in the multi-cell network

$\mathbf{H}_{r,j}$ Relay link channel matrix from the rth relay to the jth base station in a multi-cell multi-relay network

$\mathbf{I}$ Identitiy matrix

$\bar{\mathbf{I}}$ User-relay association matrix

$I(x;y)$ Mutual information between x and y

K Number of active users

$\tilde{K}_r$ Number of users not assigned to the rth relay station

K_r Number of users assigned to the rth relay station

K_t Total number of users in the system

$\Lambda_{\mathbf{D}}$ Power allocation matrix of the relay station in the single-cell single-relay network

$\bar{\mathbf{\Lambda}}_{\mathbf{D}}$ Aggregate power allocation matrix of the relay stations in the multi-cell multi-relay network

$\Lambda_{\mathbf{D}_r}$ Power allocation matrix of the rth relay station in the multi-cell multi-relay network

M_B Total number of base station antennas per cluster in a multi-cell network

M_b Number of antennas at the base station/antenna group in a single-cell network

$M_{B,j}$ Number of antennas at the jth base station in a multi-cell network

M_R Total number of relay station antennas per cluster in a multi-cell network

$M_{R,r}$ Number of antennas at the rth relay station in a multi-relay network

M_r Number of antennas at the relay station in a single relay network

$\mathbf{n}_B$ Aggregate noise vector of the coordinated base stations in the multi-cell network

$\bar{N}_r$ Total number of tranmit antennas of the users served by the rth relay station

$\mathbf{n}_d$ Noise vector at the base station in the uplink

N_i Number of antennas at the ith user

$\mathbf{n}_i$ Noise vector at ith user in the downlink

$\mathbf{n}_R$ Noise vector at the relay station in the single-cell network

$\mathbf{n}_{R,r}$ Noise vector at the rth relay station in the multi-cell network

N_U Total number of antennas at active users in the multi-cell network

$\mathcal{O}$ "Big-O" ("order of") notation

$P_R^{\mathbf{B}}$ Average transmitted power at the relay station in the single-cell downlink

$P_R^{\mathbf{M}}$ Average transmitted power at the relay station in the single-cell uplink

$P_{R,r}$ Average transmitted power of the rth relay station in the multi-cell network

P_T Total average transmitted power at the base station and relay station in the downlink or at the users and relay station in the single-cell uplink

$P_T^{\mathbf{B}}$ Average transmitted power at the base station in the single-cell downlink

$P_T^{\mathbf{M}}$ Average transmitted power at the users in the single-cell uplink

$P_{U,i}$ Average transmitted power at User i in the multi-cell uplink

$\mathbf{Q_i}$ Covariance matrix of the ith user in the uplink channel

$\mathbf{Q_u}$ Effective covariance matrix of the users in the uplink channel

R Number of active relay stations

$\mathbb{R}_+$ Set of positive real numbers

R_i^B Achievable rate of the ith user in the single-cell downlink

R_i^M Achievable rate of the ith user in the single-cell uplink

R_{sum}^B Network sum rate of the single-cell downlink

R_{sum}^M Network sum rate of the single-cell uplink

$\mathbf{\Sigma_i}$ Covariance matrix of the ith user in the downlink channel

$\mathbf{\Sigma_x}$ Effective covariance matrix of the users in the downlink channel

$\mathcal{S}_r$ Set of users served by the rth relay station

$\mathcal{S}_r^{\perp}$ Set of users not served by the rth relay station

$Tr(\mathbf{A})$ Trace of matrix $\mathbf{A}$

$(.)^{\mathbf{T}}$ Matrix transposition

$\mathbf{u}$ Effective transmitted signal vector by the users in the uplink channel

$\mathbf{u}_i$ Transmitted signal vector by the ith user in the uplink channel

$U_{[ij]}$ The (i,j)-th element of $\mathbf{U}$

$\mathbf{x}$ Effective transmitted signal vector to the users in the downlink channel

$\mathbf{x}_i$ Transmitted signal vector to User i in the downlink channel

$\mathbf{x}_R$ Transmitted signal vector by the relay station in the single-cell network

$\mathbf{x}_{R,r}$ Transmitted signal vector by the rth relay station in the multi-cell network

$\mathbf{y}_i$ Received signal vector at user i

$\mathbf{y}_R$ Received signal vector by the relay station in the downlink channel

$\mathbf{y}_{R,r}$ Received signal vector by the rth relay station in the multi-cell network

List of Abbreviations

3G Third generation

3GPP 3rd generation partnership project

4G 4th generation

5G 5th generation

AF Amplify-and-forward

b/s Bits per second

b/s/Hz Bits per second per hertz

B4G Beyond 4G

BC Broadcast channel

BD Block diagonalization

BRC Broadcast relay channel

BS Base station

CCI co-channel interference

CoMP Coordinated multipoint

CQI Channel quality indicator

CRS Common reference signal (CRS)

dB Decibel

DoF Degrees of freedom

DPC Dirty paper coding

eICIC Enhanced intercell interference coordination

eNB Evolved node B

FDD Frequency division duplexing

HetNets Heterogeneous networks

Hz Hertz

i.i.d. independent and identically distributed

ICI Inter-cell interference

IMT-Advanced International mobile telecommunications - advanced

ITU-R International telecommunication union - radiocommunication sector

JPC Joint power constraint

LTE Long-term evolution

LTE-Advanced Long-term evolution advanced

MAC Multiple access channel

MARC Multiple access relay channel

MIMO Multiple-input multiple-output

MMSE Minimum mean square error

MS Mobile station

MUI Multiuser interference

OFDM orthogonal frequency division multiplexing

P2P Point-to-point

PF Proporional fairness

PMI Precoding matrix indicator

PSD Posititve semi definite

QoS Quality of service

RF Radio frequency

RHS Right hand side

RI Rank indicator

RNs Relay nodes

RSs Relay stations

SDMA Space division multiple access

SINR Signal-to-interference plus noise ratio

SISO Single-input single-output

SNR Signal-to-noise ratio

SPC Separate power constraint

SU-MIMO Single-user multiple-input multiple-output

SVD Singular value decomposition

SZF Successive zero forcing

SZF-DPC successive zero forcing dirty paper coding

TDD Time division duplexing

TDM Time division multiplexing

TDMA Time division multiple access

THP Tomlinson-Harashima precoding

UE User equipemnt

ULA Uniform linear array

VP Vector pertubation

ZF Zero forcing

ZF-DPC Zero-forcing dirty paper coding

ZFB Zero forcing beamforming

Chapter 1

Introduction

1.1 Motivation

In the past few years, the primary focus of wireless cellular network design has shifted from the traditional voice traffic to data traffic. This is due to the ever-increasing proliferation and popularity of "smart" wireless devices (such as smart phones, tablets, etc). As a result, telecommunication operators are currently witnessing enormous increase in demand for high data rate mobile services, which require significantly higher spectral resources than conventional cell phones. These operators are therefore faced with providing ubiquitous high quality of service (QoS) and efficient mobile applications in a cost effective manner in order to meet their users' needs.

In October 2007, the International Telecommunication Union - Radiocommunication Sector (ITU-R) defined the International Mobile Telecommunications - Advanced (IMT-Advanced), also known as 4th generation (4G), specifications [2] for mobile broadband telecommunication systems. Peak data rates of 1 Gb/s for low mobility and 100 Mb/s for high mobility (with up to 100 MHz supported spectrum bandwidth) were established [3]. In response to the ITU's invitation of candidate proposals for IMT-Advanced, the 3rd Generation Partnership Project (3GPP) launched a study item (March 2008) on long-term evolution advanced (LTE-Advanced) to meet and possibly surpass the 4G requirements [4–10]. LTE-Advanced is a major enhancement of the 3GPP's earlier releases (Release 8 and 9) called long-term evolution (LTE) and was recognized by the ITU-R as 4G cellular system in November 2010 [11]. With 4G systems currently being deployed, attention has shifted to the beyond 4G (B4G) systems, the most common of which is currently known as 5th generation (5G) [12–14].

One of the major challenges confronting these emerging and future broadband cellular systems is delivering high data rates to the user equipment (UE), also known as mobile station (MS), especially those at the cell edges [5–8]. This is primarily due

to very high path-loss and interference, resulting in very low signal-to-interference plus noise ratio (SINR). The use of multiple transmit and/or receive antennas necessary to enable MIMO spatial multiplexing has been shown to significantly increase the data rates of wireless systems. However, MIMO spatial multiplexing is achievable at high signal-to-noise ratio (SNR) [15, 16].

On the other hand, a well designed and spectrally-efficient cellular network (especially with frequency reuse of 1) is interference-limited, hence experiences very low SINR, especially at the cell edges. As a result, cellular networks are hindered from leveraging the high data rate promise of MIMO spatial multiplexing. Several approaches have been developed to reduce inter-cell interference (ICI), or its effect in order to increase the SINR. The most popular of these ICI mitigation approaches is base station (BS) cooperation, also known as coordinated multi-point (CoMP) transmission/reception or network MIMO [17–21]. When multiple BSs, also known as evolved Node Bs (eNBs), cooperate (under the most effective form of CoMP known as joint transmission), their antennas together form a large virtual MIMO array thereby enabling them to co-process their transmissions on the downlink or receptions on the uplink (thanks to the high-speed, low-latency X2 cabled backhaul links between the BSs), such that ICI is eliminated or mitigated. This in turn will result in high SINRs [17,18], thereby enabling the cellular networks to leverage MIMO spatial multiplexing gains.

Furthermore, emerging and future broadband wireless networks are expected to operate in a wide range of higher carrier frequencies with different propagation characteristics, and as such experience limited service coverage. This is a result of higher path-losses compared to earlier generation cellular systems, e.g., third generation (3G). An ideal solution to extend cellular coverage would be to shrink the cell sizes (by deploying more BSs) so that users are closer to the BSs. However, the cost and other problems with such approach make it less attractive. Deployment of fixed infrastructure-based relay nodes (RNs), also known an as relay stations (RSs), within the cellular network has emerged as very attractive and cost effective technique for coverage extension in broadband cellular networks [22, 23]. RSs are low-power nodes with wireless backhaul connections, that can aid communications between BSs and users, unlike the BSs, which are connected via expensive high-capacity cabled backhaul [22–24]. Hence, RSs are cheaper as well as offer higher deployment flexibility to the operators compared to the BSs.

Moreover, in emerging broadband cellular networks, it is expected that there will be several RSs and MSs within a cell or cluster. However, not all RSs and MSs will be able (or be allowed) to participate during data transmission. Selecting subsets of

advantageous relays and users (during the control signaling phase) for cooperation and communication (during the data transmission phase) will not only enhance the system performance by leveraging the inherent cooperative and multiuser diversity in the network, but also reduce interference, energy consumption, signaling overhead and design complexity. However, in order to have any practical relevance, such user-relay selection and/or association schemes must be of acceptable complexity and fit within a small fraction of the coherence time of the channel.

Enhanced MIMO, CoMP, and relaying have already been adopted as the key enablers toward cellular broadband targets [7, 25]. For instance, while relaying and enhanced MIMO techniques form part of the LTE-Advanced Release 10, CoMP is one of the key features of Release 11. Unfortunately, very few results have been reported so far on the joint design of transmission schemes that leverage the benefits of them when applied simultaneously [20, 26, 27].

1.2 Research Goals and Summary of Contributions

The focus of this work is to investigate transceiver schemes and scheduling algorithms that enhance users' experience in emerging and future wireless broadband cellular networks, especially for users at the cell edges.

To achieve this, we adopt the more attractive non-regenerative, also known as amplify-and-forward (AF) relaying for coverage extension and MIMO spatial multiplexing for data rate improvement. That is, the MSs, RSs, and BSs are equipped with multiple antennas, unlike most systems considered in the literature. Moreover, due to the presently still insufficient isolation between the RS transmit and receive antenna circuitry as required in full-duplex relaying, we consider a more practical half-duplex relaying. Thus, the source and RS transmissions are time division multiplexed (TDM). In order to mitigate the high ICI jeopardizing the high data rate promise of MIMO spatial multiplexing in cellular networks, CoMP transmission/reception is employed. Leveraging the benefits of MIMO, CoMP, and relaying in a unified manner, we investigate the joint design of the input covariance and relay beamforming matrices to maximize the network sum rate. Also, owing to the availability of more users and relays than the system can support at a time, we investigate practical user-relay selection/scheduling and association strategies, which lead to low complexity and overhead.

The main contributions of this dissertation are summarized as follows:

- Firstly, we have investigated single-cell single-relay multi-user AF relay downlink with multi-antenna BS, RS and MSs. Prior to this work, various studies in

the literature had considered systems with single-antenna users/MSs and their proposed schemes are not applicable to systems with multi-antenna MSs.

In particular, we jointly designed the transmit covariance matrices at the BS and beamforming matrix at the RS in order to maximize the system sum-rate. We considered both the joint and separate power constraints at the RS and BS. Owing to the highly complex sum rate optimization problem for the downlink, we studied a more tractable dual uplink channel by employing uplink-downlink duality. We then derived the mapping from the resulting covariance matrices for the uplink to the desired covariance matrices for the downlink. Compared with two known single-antenna-user schemes [1], simulations show that our scheme outperformed the *all-pass relay* design and performed similarly to the *SVD-relay* design [1]. Moreover, our design performed close to the sum-rate upper bound with the performance gap decreasing with increasing number of antennas at the users. Finally, we demonstrated (via simulation) that having more antennas at the RS than the BS (compared to having more antennas at the BS than the RS) is desirable for best system performance.

- Secondly, we have investigated multi-cell cooperative cellular AF relay uplink with several multi-antenna BSs, RSs and MSs. We have made the practical assumption that each transmitter has its own individual transmit power. The MSs' transmit covariance matrices and RSs' beamforming matrices are jointly designed to maximize the network sum-rate. However, the second hop channels from the RSs to the BSs are coupled due to the amplifying and forwarding of interference by the RSs. In order to undo the coupling of the relay channels, we incorporated interference mitigation mechanism into the RSs' beamforming designs. Moreover, for a fixed number of scheduled users and relays we proposed a user-relay association scheme, by which different users were assigned to different relays for service in order to enhance the system performance. Simulation results showed the effectiveness of our proposed designs under various system configurations and channel conditions, with better performance in networks with a large number of users.

- Lastly, we have investigated user-relay selection and association in multi-user multi-relay MIMO wireless cellular networks. We have proposed a novel low-complexity joint scheme, which simultaneously selects multiple relays and users as well as assigns different selected users to different selected relays for service. Our proposed joint scheme utilizes only the amplitude of the complex channel gains between the nodes (i.e., full channel knowledge is not needed), which leads

to reduced feedback and overhead in comparison to schemes that require full channel knowledge. Furthermore, the complexity of the scheme scales linearly with the total number of relays, the total number of users and the number of selected users. Simulation results demonstrate the superiority of our proposed scheme compared to (i) a scheme with neither user-relay selection nor user-relay association, (ii) a scheme with user-relay association, but no user-relay selection, and (iii) two 2-step schemes which select the relays in the first step and then select and assign the users to the selected relays in the second step. Most importantly, the proposed joint scheme achieves a reasonable percentage of the performance of the exhaustive search scheme. The favorable performance and low complexity of our proposed joint scheme make it very attractive for possible implementation in emerging and future broadband wireless relay networks.

1.3 Organization of the Thesis

Chapter 2 provides some background details on the key concepts and techniques employed in this work. In particular, we provide an overview of the wireless systems as well as briefly discuss key technologies for the emerging broadband cellular networks employed in this work.

Chapter 3 details our proposed joint design of the transmit covariance matrices and relay beamforming matrix that maximizes the sum rate of a single-cell MIMO relay downlink, comprising one BS, one RS and K cell-edge MSs, each with multiple antennas [28, 29]. The system model, problem formulation, proposed sum-rate optimization algorithm and the mapping from the MARC covariance matrices to the BRC covariance matrices are presented. We conclude the chapter with some numerical results, notably the impact of the relative number of antennas at the BS and the RS on the network performance.

Chapter 4 discusses our proposed interference-aware joint design of the relays' beamforming and users' transmit covariance matrices that maximize the network sum rate of a multi-cell MIMO cellular relay uplink comprising of C BSs, R RSs and K MSs, each with multiple antennas [30, 31]. The system model, sum-rate formulation and an iterative algorithm for obtaining the network sum rate were first presented followed by our proposed user-relay association scheme. Furthermore, we briefly touch on CSI issues (acquisition and feedback/overhead reduction) peculiar to the proposed joint user-relay beamforming design and association scheme. Simulation setup and numerical results under various system configurations and channel conditions are also presented.

Chapter 5 focuses on joint design of user-relay selection/scheduling and association strategy in multi-user multi-relay wireless relay networks with multi-antenna transmitters and receivers [32]. Our formulated integer programming optimization problem for the joint user-relay selection and association scheme is first presented. Owing to the complexity of the formulated problem, we propose a practical low-complexity greedy algorithm to solve the highly complex optimization problem. Numerical results demonstrating the effectiveness of the proposed scheme conclude the chapter.

Chapter 6 summarizes the main contributions of this dissertation and proposes possible directions for future work.

Chapter 2

Background and Key Concepts

In this chapter, we briefly outline some background and key concepts helpful for understanding of materials treated in this thesis before delving into the major results of this work in subsequent chapters.

2.1 Classical Wireless Communication Systems

Dated back to the beginning of radio age over a century ago with the invention of radiotelegraph by Guglielmo Marconi, the wireless industry is poised for rapid growth owing to recent advances in radio technology. As a result, new and improved mobile services at lower costs are constantly being created, resulting in increase in the number of subscribers and airtime usage.

The two fundamental features of a wireless communication system that make it challenging and interesting are *fading* and *interference* [33, 34]. *Fading* is the time variation of the channel gains due to small-scale effects of multi-paths as well as large-scale effects such as path-loss (due to distance attenuation) and shadowing (due to large obstacles e.g., buildings, hills, etc). Unlike in the wired world where each transmitter-receiver pair can be viewed as an isolated point-to-point link, wireless users communicate over the air, thereby experience significant *interference* among themselves. Such *interference* can be between signals from a single transmitter to multiple receivers (e.g., cellular downlink), between transmitters communicating with a common receiver (e.g., cellular uplink) or between different transmitter-receiver pairs (e.g., users in different cells).

Reliability and throughput (a.k.a. data rate) are two common measures of wireless communication systems. A popular measure of reliability of a communication channel is the detection error probability, which has been shown to decay exponentially with signal-to-noise ratio (SNR) in the additive white Gaussian noise (AWGN) channel

but decays only inversely with SNR in fading channel. Thus, significantly higher power is required for a reliable communication over a fading channel compared to the AWGN channel. For reliable communication over a fading channel, an information symbol can be meant to pass through multiple independent propagation paths. Since the probability that at least one of the paths does not experience deep fade increases as more paths are used, the reliability of the channel is improved. This technique is commonly known as *diversity* and can be obtained over time (e.g., via coding and interleaving), frequency (if the channel is frequency-selective), space (e.g., by using multiple transmit or receive antennas), macro-diversity (e.g., having two base stations receive signals from a mobile device), etc. Traditionally, when the focus of wireless system design is on increasing the reliability of the air interface, fading and interference were viewed as nuisances to be countered. However, with the focus shifted towards increasing the throughput, fading is now viewed as an opportunity to be exploited. Following his formulation of information theory to characterize the limits of reliable communication in 1948, Claude Shannon showed the ground-breaking result that communication at a strictly positive rate with as small detection error probability as desired is possible, provided that the rate is not above the channel capacity (the maximum throughput at which reliable communication is possible).

The capacity (in b/s) of an AWGN channel depends on the channel bandwidth W and SNR (i.e., the ratio of the received signal power to the noise power) as follows

$$C = W \ \log_2 \ (1 + SNR) \tag{2.1}$$

In order to increase the channel capacity, the SNR and/or bandwidth has to be increased. At low SNR, capacity increases linearly with transmit power, whereas, at high SNR, capacity increases only logarithmically with transmit power. Thus, there is a diminishing return in capacity with increasing transmit power. On the other hand, bandwidth is a very expensive and scarce resource; as a result novel approaches towards increasing channel capacity are continually being pursued.

2.2 Overview of Cellular Networks

Cellular network is one of the most important and fastest-growing wireless systems. In a cellular network, a given geographical area is partitioned into *cells*, where each cell is served by a BS. Each BS uses the allocated wireless frequency spectrum to communicate with one or more mobile users assigned to it within its coverage area. A cellular system can be categorized as the uplink (or reverse link) when multiple users transmit to a BS receiver or downlink (or forward link) when a BS transmits to

multiple users. Furthermore, a cellular network can operate in either the frequency division duplexing (FDD) mode or time division duplexing (TDD) mode. In FDD systems, the uplink and downlink transmissions occur on different frequency/bandwidth resources, where as in TDD systems, uplink and downlink transmissions occur at different times over the same bandwidth/frequency.

A wireless cellular system designer is faced with a number of challenges such as complex time-varying wireless environment (with fading and multipath), limited availability of radio spectrum, as well as meeting ever-increasing demand for high data rates, better quality of service, high network capacity, high service coverage, and fewer dropped calls. All these often-conflicting demands call for innovative techniques to enhance users' experience.

A common measure of cellular network performance is *spectral efficiency* (measured in bits per second per hertz (b/s/Hz)), and defined as the *throughput* (i.e., the total data rate transmitted or received by a BS, measured in bits per second (b/s)), over a given channel bandwidth (measured in hertz (Hz)). Subject to constraints on the radiated power, the wireless system engineer's job is to design the network to provide reliable communication over the allocated spectral resources by utilizing the BS assets. These spectral resources are reused at different cells in order to improve spectral efficiency, which can translate to higher data rate to a user, increased coverage, increased reliability and reduced network cost to the operator [35]. Universal frequency reuse where the total allocated channel bandwidth is reused in every cell is commonly used in contemporary cellular networks to improve spectral efficiency.

With multiple users assigned to each BS, transmissions from a cell cause inter-cell interference (ICI), also known as co-channel interference (CCI), at the receivers of others cells, thereby degrading the SINR. Thus, ICI inhibits ubiquitous user experience of high data rate envisaged in the emerging and next generation broadband cellular networks. That is, the data rate achieved by users at the cell edge is only a small fraction of the peak rate. As a result, there is a non-uniform SINR distribution within the network. For instance, in a typical homogeneous network with inter-site distance of 500 m, while over 20 dB SINR was achievable at the cell center, only about 5 dB or lower SINR was achievable in most areas of the cell due to ICI [21].

The enormous performance improvements of emerging and future broadband cellular systems over other legacy cellular standards (e.g., 3G) are realized with advanced wireless techniques such as MIMO, CoMP, relaying, etc. [25]. In what follows, we provide brief background on some of these techniques employed in this work.

2.3 MIMO Wireless Systems

Starting in mid-1990s [15, 36, 37], the use of multiple antennas at the transmitter and/or receiver of a wireless system has been shown to deliver enormous performance improvement, at no extra cost of bandwidth or power [16]. Such systems are commonly referred to as MIMO wireless systems, also known as space-time (ST) wireless systems [16]. The performance improvement of MIMO is achieved by leveraging the new resource (space) provided by multiple antennas at the transmitters and/or receivers.

2.3.1 Single-User MIMO (SU-MIMO) Systems

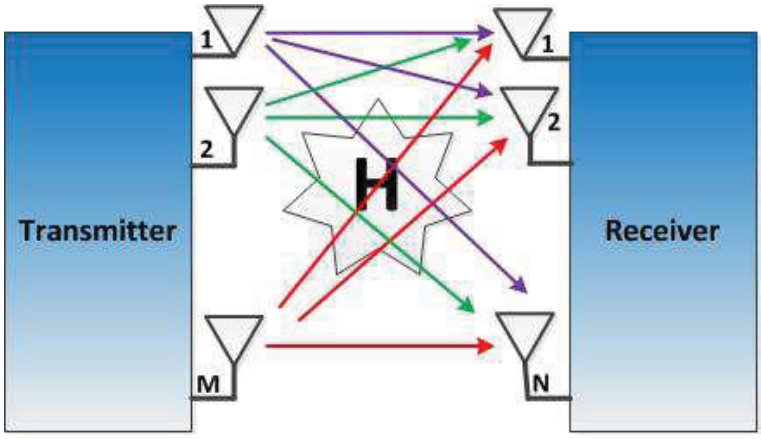

Figure 2.1: A single-user MIMO system.

For a general single-user wireless communication system, different antenna configurations exist, namely: SISO (single-input single-output) which is the familiar wireless configuration with single transmit and receive antenna, SIMO (single-input multiple-output) with single transmit antenna and multiple receive antennas, MISO (multiple-input single-output) with multiple transmit antennas and single receive antenna and MIMO (multiple-input multiple-output) with multiple transmit antennas and multiple receive antennas. Figure 2.1 depicts a single-user MIMO (SU-MIMO), also known as point-to-point (P2P) MIMO, with M antennas at the transmitter and N antennas at the receiver. The complex channel gains between the transmit and receive antennas is captured by an $M \times N$ channel matrix $\mathbf{H}$.

MIMO benefits manifest themselves in the form of increased received SNR (*array gain*), increased link reliability (*diversity gain*), increased data rate (*multiplexing*

gain) and reduced interference (*interference reduction gain*) [16] as briefly outlined next.

- **Array gain**

  Array gain is the average increase in the SNR at the receiver due to the coherent combining effect of multiple antennas at the receiver or transmitter or both [16, 38]. In multiple transmit antenna systems (i.e., MISO, MIMO), channel knowledge at the transmitter is required to leverage array gain. Array gain increases resistance to noise and improves the coverage range of a wireless system.

- **Diversity gain**

  In a wireless channel, the power of the signal fluctuates (i.e., fades) as it propagates through the air. When this fluctuation is significant, the channel is said to be in a deep fade. In general, diversity (characterized by the number of independent fading branches, also known as *diversity order*) is used to combat fading in wireless systems. In particular, MIMO diversity leverages the spatial degrees of freedom (DoF) provided by the multiple antennas. In a MIMO system, multiple (ideally independent) copies of the transmitted signal for each path can be transmitted or received to create a more robust and reliable link. With adequate spacing between the antenna elements, channels of different antennas experience independent fading, such that the probability of all channels being in deep fade is largely reduced. So with a high probability, at least one copy of the signal can be detected at the receiver.

  MIMO diversity can be categorized as receive antenna or transmit antenna diversity. Receive antenna diversity can be used in SIMO systems [39]. Similarly, transmit antenna diversity is applicable to MISO systems [40, 41], with or without channel knowledge at the transmitter. For instance, [42–44] rely on coding across the antennas (space) and time to extract transmit diversity without channel knowledge at the transmitter.

  In MIMO systems, spatial diversity can be harnessed by combining the transmit antenna diversity and the receive antenna diversity. The maximum possible diversity order, given by the product of the number of transmit and receive antennas (i.e., MN), is achievable when the channels between the transmit-receive antenna pairs experience independent fading.

- **Multiplexing gain**

  Spatial multiplexing (SM) allows for simultaneous transmission of independent data streams to multi-antenna users. It has been shown that MIMO spatial multiplexing scales the data rate (or capacity) by the minimum of the number of antennas at the transmitter and receiver (i.e., $\min(M, N)$) at no extra cost of bandwidth and power [16]. Spatial multiplexing is only possible in MIMO systems [15, 36, 37], and is achievable in the high SINR regimes. In order to achieve full spatial multiplexing in single-user MIMO systems, rich-scattering environment is required.

- **Interference Reduction/Avoidance Gain**

  Interference in wireless networks stems from multiple users sharing the same time and frequency resources. By exploiting the spatial dimension of MIMO systems, separation between wireless users can be increased thereby reducing interference. Also, MIMO spatial dimensions can be leveraged at the transmitter by focusing the signal energy towards the direction of the intended user while minimizing interference to other users. Interference reduction allows the use of aggressive reuse factors and improves network capacity.

It is usually not possible to leverage all MIMO gains simultaneously. This is often due to the conflicting demands on the spatial degrees of freedom (i.e., number of antennas). In particular, a fundamental trade-off exists between spatial diversity and spatial multiplexing in MIMO systems [45]. However, combining some of the benefits will result in increased capacity, reliability, and coverage.

2.3.2 Multiple-User MIMO (MU-MIMO) Systems

With multiple antennas at BS, multiple users can communicate simultaneously (with the BS) in the same frequency channel by leveraging the differences in spatial signatures at the BS antenna array induced by spatial separation of the users. This is commonly known as space division multiple access (SDMA). The major attraction of SDMA is the possibility of channel reuse within a cell in order to increase spectral efficiency [16].

While SU-MIMO requires rich scattering environment to achieve full MIMO spatial multiplexing gain, such requirement is not necessary in multi-user (MU)-MIMO. The latter is due to sufficient spatial separation of users. More so, compared to the full CSI case, while SU-MIMO suffers minor capacity penalty when CSI knowledge is not available, MU-MIMO channel suffers much larger capacity penalty when CSI is

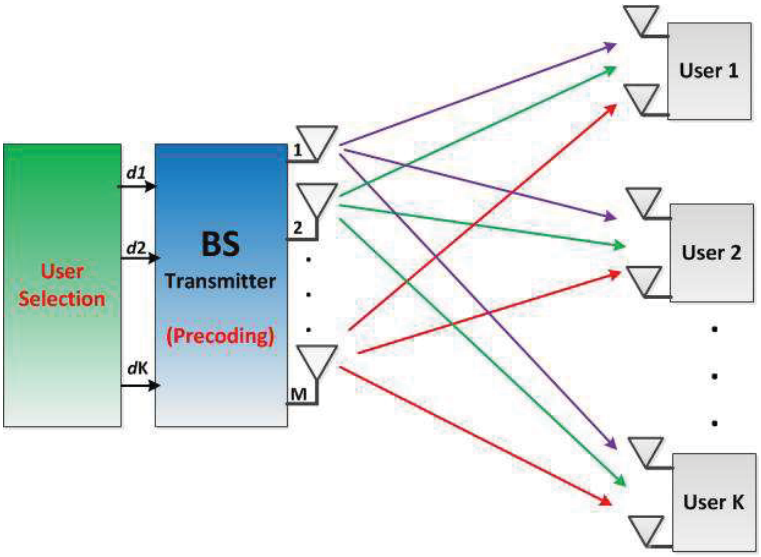

Figure 2.2: A MIMO broadcast channel (downlink).

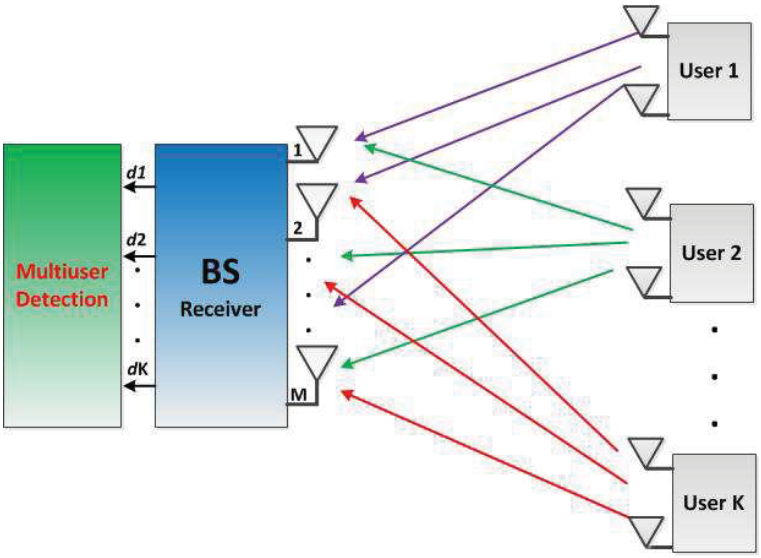

Figure 2.3: A MIMO multiple access channel (uplink).

not available. Hence, robust channel estimation is crucial in order to leverage MIMO spatial multiplexing gain in multi-user wireless cellular networks.

There exist two types of MU-MIMO channels: the Gaussian MIMO broadcast channel (BC), where a multi-antenna BS simultaneously transmits different data streams to multiple users in the downlink, and the Gaussian MIMO multiple access channel (MAC), where multiple users simultaneously transmit to a multi-antenna BS in the uplink [35]. Figures 2.2 and 2.3 depict a MIMO-BC and a MIMO-MAC respectively. Unlike the SU-MIMO channel with scalar capacity, a MU-MIMO channel with K users is defined by a K-dimensional *capacity region*, consisting of the achievable rate vectors $(R_1, R_2, ..., R_K)$ of the K users, with $R_k \geq 0$ being the achievable rate of the kth user. An important scalar performance metric for the MU-MIMO channels is the sum-capacity defined as the maximum achievable total data rate.

The sum-capacity of the MIMO-MAC is achievable by having all users transmit with full power and employing successive interference cancelation at the BS, with user information successively decoded and canceled one after the other in the presence of noise and interference from yet to be canceled users. On the other hand, the MIMO BC sum-rate capacity is achievable with dirty paper coding (DPC) [46–48], a non-linear and complex encoding of the user information in an ordered manner such that each user experiences interference only from users encoded after it.

A very useful duality between the MAC and BC capacity regions exists [49], which provides valuable insight and a tool for evaluating the performance of the BC from that of the MAC. In particular, the duality relationship establishes the fact that the dirty paper rate region of the BC with BS power constraint P is equal to the union of the dual MAC capacity regions, with the union taken over all individual user transmit power constraints that sum to P. In addition to this relationship between the BC and MAC rate regions, an elegant set of matrix transformations to find the BC covariance matrices from the MAC covariances (and vice versa) that achieve the same sum rate exists [49].

Although MU-MIMO is very useful in improving the average spectral efficiency of a cell, the spectral efficiency of a user at the cell-edge can be degraded if MU-MIMO is exclusively employed. As a result, support for dynamic switching between SU-MIMO and MU-MIMO in order to balance the average user spectral efficiency and the cell-edge user spectral efficiency has been proposed for LTE-Advanced Rel-10 [50]. In particular, a transparent MU-MIMO framework in which a user is unaware of whether or not it is scheduled alone or with other users in the same time-frequency resource is adopted in LTE-Advanced Rel-10.

In terms of antenna configurations at BS, highly-correlated antenna setups, e.g.,

uniform linear array (ULA) and small antenna spacing, which create narrow antenna beams beneficial for SDMA are desirable in MU-MIMO unlike for the SU-MIMO where widely spaced and cross-polarized antennas are preferred [50].

2.3.3 Multiple-User MIMO Precoding

With the availability of channel state information at the receiver (CSIR), a SU-MIMO receiver can post-process its received signal such that each antenna element receives interference-free signal. Similarly, in the MIMO MAC (i.e., uplink), the BS receiver can readily employ the capacity-achieving successive interference cancelation strategy in order to successfully decode each user's interference-free signal. However, for the MIMO BC (i.e., downlink), the users are spatially separated, and hence cannot co-process their received signals such that each user sees interference-free signal, as for the SU-MIMO and the MIMO-MAC. Moreover, the mobile receiver is much smaller and less sophisticated to handle the demanding task of interference cancelation. As a result, interference-cancelation task is relegated to the BS transmitter in MIMO-BC. With channel state information at the transmitter (CSIT), the BS can pre-process its transmission such that each user can receive interference-free signal (or at least with reduced interference), thereby increasing the SINRs. This signaling technique is commonly known as precoding (or beamforming) in MIMO literature.

While DPC is the optimal (i.e., capacity-achieving) strategy for MIMO-BC, it requires highly complex multi-dimensional vector quantization [51–53], and hence is very difficult to implement in practical systems. The impracticality of DPC is further exacerbated by the time-varying nature of wireless channels [35]. For one-dimensional quantization, DPC can be implemented using a non-linear Tomlinson-Harashima precoding (THP) [35]. However, THP suffers from significant performance losses at low SINRs. In addition to DPC and THP, another common non-linear MIMO precoding scheme is vector perturbation (VP) [54–57].

Owing to the high complexity of non-linear precoding schemes, linear precoding or beamforming approaches are more attractive due to their relatively lower complexity. Linear precoding schemes are designed to remove (at the transmitter) some or all of the interference between the users, which hence can limit the available DoF in the system. However, for networks with a large number of users, the sum rates of linear beamforming schemes approach that of DPC asymptotically [58]. The most common linear precoding scheme is the zero forcing (ZF) also known as channel inversion or zero forcing beamforming (ZFB) for networks with single-antenna users [48] or block diagonalization (BD) [59] for networks with multi-antenna users. Minimum

mean square error (MMSE) precoding, also known as regularized channel inversion, is another common linear precoding scheme for MU-MIMO systems [60]. Both ZF and MMSE precoding require a great deal of power to perform channel inversion when the MIMO channel is ill-conditioned and are both sensitive to channel estimation errors.

ZF precoding is designed to create orthogonal non-interfering channels between single-antenna users. A direct extension of ZF precoding to systems with multi-antenna users is to consider each antenna as a separate user and then apply ZF. However, such approach is sub-optimal since each user can coordinate its multiple antennas to jointly process the received signals of all its antennas. Block diagonalization (BD) extends the idea of ZF to MIMO systems with multiple-antenna users by focusing on removing interference between users but not interference between different antenna elements of a given user. Unlike BD, which completely removes all the interference between the users, successive ZF (SZF) removes only the interference a user experiences from users encoded after it (i.e., interference from users encoded before the user remains).

The performance gap between the linear schemes and DPC increases with increasing SNR. Zero-forcing DPC (ZF-DPC) [48] combines the inherent structures of ZF (by relaxing the complete removal of interference between the users) and DPC, to improve the performance of ZF for MIMO systems with single-antenna users [48]. A similar technique for MIMO systems with multi-antenna users is the successive ZF DPC (SZF-DPC) [61]. In both the ZF-DPC and SZF-DPC, some of the interference between the users is removed via ZF and some via DPC. Moreover, both schemes approach the capacity of DPC at low SNRs. To achieve this near-capacity performance, optimal ordering of the users for interference removal and very small transmit powers are required [48, 61]. In general, MIMO precoding schemes are designed to achieve a particular design goal, for instance, power minimization subject to SINR constraint, sum rate maximization subject to transmit power constraint, and so on.

MIMO precoding techniques are already included in recent wireless standards. For instance, while codebook-based precoding is adopted for SU-MIMO in LTE Rel 8, non codebook-based precoding is adopted for both the SU-MIMO and MU-MIMO in Rel 10 [50].

2.3.4 User Scheduling in Multiple-User MIMO

In multi-user single-input single-output (SISO) systems, where the BS and users are each equipped with single antenna, the optimal (capacity-achieving) strategy is to transmit to a single user at a time using the total available power [62]. Due to chan-

nel fluctuations in the system, multiuser diversity can be leveraged by selecting the user with the best channel during each transmission interval. As a result, different users can be scheduled for different transmission intervals in a time division multiple access (TDMA) fashion. This is possible because it is very unlikely that the same user will have the best channel for different transmission intervals if all the users' random channels are statistically identical in their distribution. However, in MU-MIMO, multi-user diversity is reduced due to the decrease in channel fluctuations as a result of multiple antennas at the nodes, a phenomenon widely known as *channel hardening* in MIMO literature [63]. As a result, scheduling a single user in each transmission interval by employing TDMA is sub-optimal, unlike in SISO systems. Instead, multiple users can be simultaneously scheduled for service during each transmission interval by leveraging the MIMO spatial dimensions, for instance, through DPC or beamforming. In a MIMO broadcast channel with M transmit antennas, and N receive antennas at each of the K users, it has been shown that the sum rate of DPC and beamforming scales as $M \log(\log KN)$, whereas that of TDMA only scales as $min\{M, N\} \log(\log K)$ at high SNR [64].

For cellular systems in general, the number of users that wants to communicate with a BS is usually larger than the BS can support. This is also the case in MU-MIMO system employing linear beamforming, where the number of antennas at the BS is usually less than the number of users within the coverage of the BS. Consequently, not all users requesting service can be served at the same time. User scheduling, where only a subset of users is selected for service, has been extensively studied in the literature, e.g., [58, 65–73] and references therein. Typically, user scheduling in MU-MIMO systems is implemented as an optimization problem, the most common of which is the maximization of the system aggregate throughput (i.e., sum rate), or a lower complexity metric also intended to maximize the throughput. The optimal user scheduling method is via exhaustive search. However, because of its rapidly increasing complexity with the number of users, lower-complexity schemes are crucial. While user scheduling primarily focuses on exploiting multi-user diversity, it can also be used to combat interference between the users. For instance, in order to maximize the system sum rate, it has been shown that selecting users with good channel gains as well as large spatial separation (in order to reduce interference) is optimal [74]. That is, the selection of users can be used to reduce multiuser interference (MUI) similar to precoding or beamforming. In [58, 67], scheduling algorithms aimed at reducing MUI by selecting users which are partially orthogonal were proposed. Most user scheduling algorithms in the literature are designed to maximize the sum-capacity, for example, [58] for ZFB, [67, 68] for BD, [75] for DPC, and [76] for SZF-DPC. However, such

schemes are generally unfair to users with poor channel conditions as they favour users with "best" channel gains. In order to ensure fairness among the users, user-scheduling algorithms based on proportional fairness (PF) criterion, where users with the highest ratio of current rate to their prior average rate are selected, have been proposed [77, 78].

2.4 CSI Issues in MIMO Systems

As have seen so far, the knowledge of the CSI at the receiver (CSIR) and/or transmitter (CSIT) is very crucial to harness the gains of MIMO technologies. In practical MIMO systems, CSIR is readily obtained by channel training. That is sending *pilot signals* (also known as *reference signals* [79] in the 3GPP's standards) by the transmitter over time or frequency resources that is orthogonal to the data signal. On the other hand, CSIT can be obtained via feedback from the receivers in FDD systems or through channel reciprocity in TDD systems. The accuracy of CSI estimations, and hence the system performance, depends greatly on the degree of channel variations as well as the percentage of resources devoted to channel estimation [35]. For optimal MIMO channel training, mutually orthogonal training sequences, one for each antenna element and with equal power are required [80]. While TDD-based CSIT yields better performance compared to the FDD counterpart, implementation of TDD requires perfect time synchronization of the downlink and uplink transmissions among all the BSs [80]. That is, all BSs must transmit at the same time and receive at the same time. As such, TDD systems are more restrictive than FDD systems.

For instance, for FDD systems, the CSI feedback mechanisms in LTE Rel-8 and LTE-Advanced Rel-10 are based on a well established and tested implicit feedback framework, where each user measures the downlink channel and reports a set of CSI recommended MIMO transmission formats, namely: the rank indicator (RI), precoding matrix indicator (PMI) and channel quality indicator (CQI), to the BS. Whereas PMI and RI jointly represent the spatial directions of the MIMO channel, CQI indicates the strength of the corresponding spatial directions. Unlike the downlink common reference signal (CRS) used for CSI measurement in LTE Rel-8, low-duty-cycle low-density CSI reference signal (CSI-RS) which allows higher reuse factor than the Rel-8 CRS was adopted in LTE-Advanced Rel-10 [50, 79].

2.5 Coordinated Multipoint (CoMP)

Inter-cell interference (ICI) is well-known to hinder ubiquitous user experience of high data rate in well-designed and spectrally-efficient broadband cellular networks. That is, data rates achievable at the cell-edge are only a fraction of the peak data rate [21]. As such, ICI management is a critical issue for emerging and future broadband cellular systems.

Coordinated multipoint (CoMP) transmission/reception, also known as *network MIMO*, in which several BSs operate in a coordinated manner has emerged as a very attractive technology for combating ICI [17–20]. When multiple BSs are coordinated, their antennas together form a large virtual MIMO array, thereby allowing them to co-process their transmissions (in the downlink) or receptions (in the uplink) such that ICI is mitigated or eliminated. This in turn will result in high SINRs [17, 18] necessary to leverage MIMO multiplexing gains.

CoMP can be employed both in the downlink and uplink. In the uplink, multiuser detection is employed across multiple BSs in order to jointly detect the users' signals, similar to the conventional (uncoordinated) multi-user MIMO receiver. For the downlink, both global CSIT and global data of all users (in case of joint processing variant of CoMP) are required at each BS, thereby resulting in enormous backhaul bandwidth demand [35]. Furthermore, tight carrier phase and symbol synchronization among the coordinated BSs are required so that the transmitted signals from the BSs coherently arrive in phase at each user. As a result, downlink CoMP poses more challenges to cellular system designers compared to uplink CoMP.

While coordination over the whole network will (in principle) completely eliminate ICI, such approach demands very large backhaul capacity due to the CSI and (in the case of joint processing CoMP) data exchange involved. It would also suffer from overwhelming signaling load for channel estimation. Additionally, the coordination of distant BSs is not necessary (and feasible), because of very large path-loss. Limited coordination (a.k.a., *clustered network MIMO*), whereby only a subset of nearby BSs (or cells) are coordinated to form a cluster, has emerged as a more attractive approach to CoMP deployment in broadband cellular systems [81–83]. However, *clustered network MIMO* converts inter-cell interference to inter-cluster interference, which degrades the SINRs of the cluster-edge MSs. This cluster edge effect can be "averaged out" over time by periodically reassigning the BSs that form a cluster such that MSs near the edge of a cluster during one interval are closer to the interior of the cluster area in another interval [84, 85]. In particular, dynamic clustering by which the BSs that form a cluster are chosen to serve a scheduled set of users based on the

fading channel conditions is of great interest [86, 87].

Depending on the degree and type of information (i.e., CSI and/or data) shared among the coordinated BSs, CoMP can be categorized as joint processing (JP) or coordinated beamforming/scheduling (CS/CB). Furthermore, it has been shown that data sharing constitutes most of the bandwidth demand compared to CSI sharing. In the following, we briefly describe each of the two main CoMP categories.

2.5.1 Joint Processing

Under *joint processing* (JP), both global CSI and data of all users are shared among the coordinated BSs. The major drawback of JP is the enormous signaling over-head due to data sharing. This form of CoMP is basically a large distributed MIMO array, hence also know as MIMO cooperation in some literature, for instance [20].

For the downlink CoMP, JP can be further categorized as joint transmission (JT) or fast cell selection (FCS), also known as transmission point selection (TPS) [19]. Under JT, each user's data is simultaneously transmitted from all the coordinated BSs (as depicted in Figure 2.4), and as such ICI (within the coordinated cluster) can be completely eliminated by jointly precoding the users' data. In this case, interference can be said to be exploited. JT is particularly useful to improving the cell-edge user experience by converting an interfering signal into a desired one. Depending on whether coherent combining of signals from multiple BSs is targeted, JT can be coherent or non-coherent, and aims to improve the overall system throughput or similar system-wide performance metric. For the TPS, although all users' data are available at all the BSs, each user receives its data only from one of the BSs (as shown in Figure 2.5). TPS exploits changes in the channel fading conditions to dynamically schedule each user to its most appropriate BS for service.

2.5.2 Coordinated Beamforming/Scheduling

Under coordinated beamforming/scheduling (CB/CS), the coordinated BSs share only CSI, but not the actual data (see Figure 2.6). Thus, the data for a given user is only available at a single BS. In the downlink, for instance, the BSs can coordinate their transmissions to their respective target users such that minimal interference are seen at other (un-targeted) users. Usually, ICI can only be coordinated, but cannot be exploited as in MIMO cooperation. Thus, this form of CoMP is also referred to as interference coordination [20], and is more practical than JP due to its relatively lower signaling over-head and backhaul requirements.

A hybrid of both JP and CS/CB is called partial interference cancelation, which

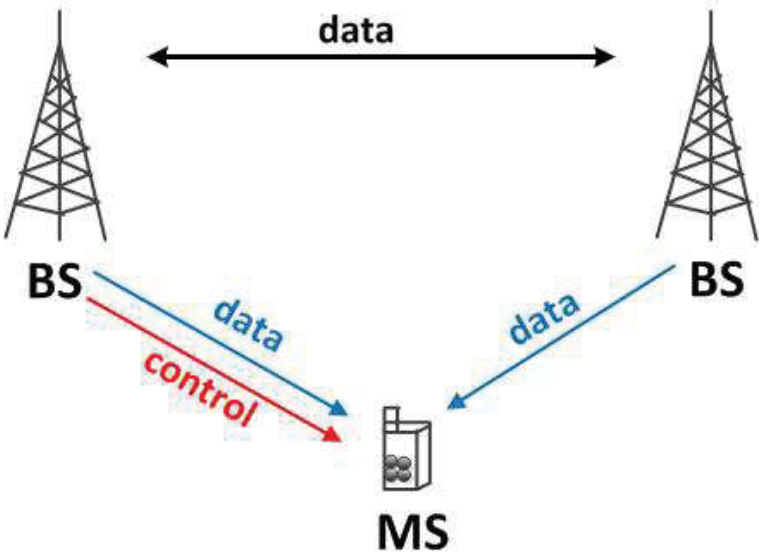

Figure 2.4: Joint Transmission.

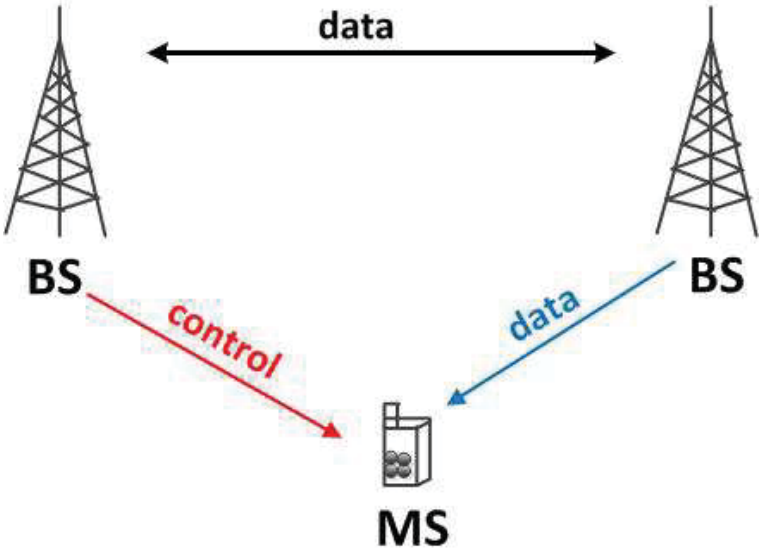

Figure 2.5: Fast Cell Selection.

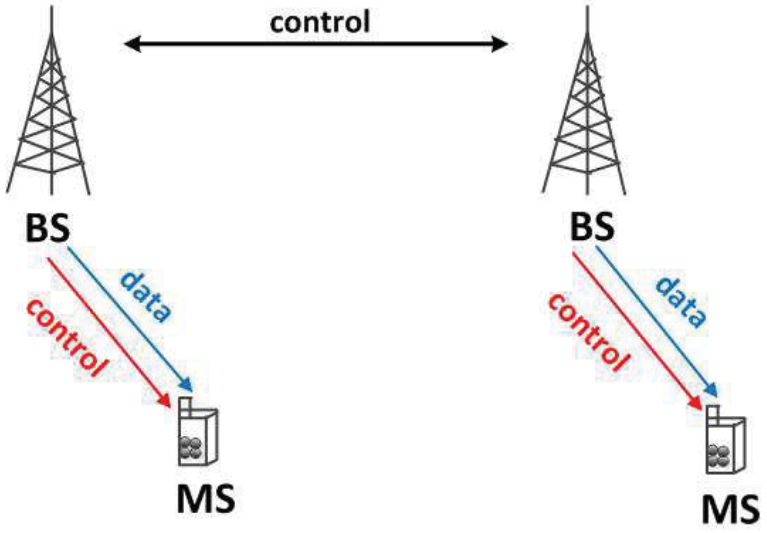

Figure 2.6: Coordinated Beamforming/Scheduling.

balances the performance-complexity tradeoff of JP and CS/CB by allowing partial interference cancelation [20]. In this case, the CSI is shared among the cooperating BSs, while only a subset of users' data is shared. For instance, in the downlink, each user will receive its signal from one or more BSs, but not from all BSs in the cluster as in JP.

CoMP was originally considered in the study item of LTE-Advanced [88] to meet the 4G requirements [2]. However, there was no consensus on its gain under realistic conditions, and hence it was not included in LTE Rel-10. On identifying some performance gains after further studies, a work item was approved to specify CoMP in the LTE Rel-11 [21]. Focusing on systems with backhaul of low latency and sufficient capacity, the 3GPP standardization group defined four different CoMP scenarios in their LTE-Advanced study, namely: coordination between sectors of the same macro BS (where no backhaul is needed), coordination between cells belonging to different radio sites from a macro BS, coordination between a macro cell and low power nodes (usually RRHs) within the macro coverage, each node with its own cell identity, and coordination between a macro cell and low power nodes (usually RRHs) within the macro coverage, all nodes with the same cell identity [19, 21, 89].

In 3GPP framework, downlink CoMP transmission is categorized into dynamic point selection (DPS), dynamic point blanking (DPB), joint transmission (JT), and coordinated scheduling/beamforming (CS/CB) [21]. With DPS, a BS is dynamically selected on per sub-frame basis according to the instantaneous channel conditions, while with DPB, the BSs that constitute the most interference to a user is dynamically

identified and muted (i.e., no signal is transmitted from the BSs).

2.6 Relaying

Coverage extension at high data rates is one of the major issues of emerging broadband cellular networks. This is due to the high path-loss inherent in such systems as a result of very wide range of operating carrier frequencies with different propagation characteristics and coverage levels. As such, several users especially those at the cell-edge experience very low SINRs. A brute force solution would be to shrink the cell sizes (by deploying more BSs) so that users are closer to the BSs in order to achieve higher SINRs [22]. However, the cost of cabled fiber backhaul, site acquisition, maintenance and other problems associated with such approach make it less attractive to cellular operators.

2.6.1 Fixed Infrastructure-Based Relays

A very attractive and cost effective solution to extend radio coverage in wireless cellular systems is the use of fixed infrastructure-based relays, also known as RSs or RNs [22–24]. Unlike the BSs which are connected via highly expensive high-speed and low-latency cabled backhaul, RSs are low-power nodes with wireless backhaul connections (usually sharing the radio frequency (RF) spectrum with user traffic) that can aid communications between BSs and users through multihop communications [90, 91]. As a result, RSs are simpler, less expensive, and much easier to deploy [6, 22–24, 92].

In addition to coverage extension to cell-edge users, RSs can also be used in many other scenarios; inside tunnels, trains, etc (where BSs cannot be deployed), during the initial stage of network deployment (when system capacity is not yet an issue), during a concert, after the damage of the physical infrastructure, and so on. Also, there can be one or more RSs between a BS and an MS as depicted in Figure 2.7. With the help of RSs, the radio link between the BS and an MS is divided into two hops. In the LTE terminology, the BS-RS and the RS-MS links are commonly known as the relay (or backhaul) and access links respectively [6,92]. Both links are expected to have better propagation conditions than the direct link from the BS to the MSs [92].

Within a given cell/sector, the relay link and the access link may or may not share the same frequency channels. For instance, an RS is said to be *in-band* when its relay and access links operate in the same frequency band and *out-of-band* when its relay

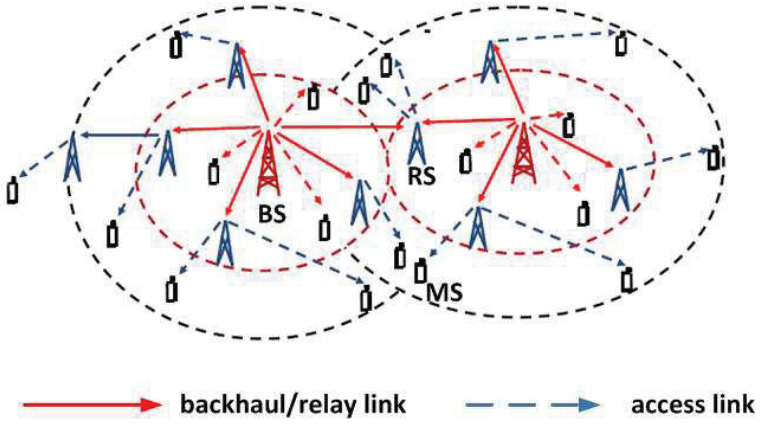

Figure 2.7: A 2-cell cellular relay network.

and access links operate in different frequency bands [23, 24]. Hence, in-band relays are more spectrally efficient compared to out-of-band relays.

Furthermore, a *half-duplex* RS can transmit or receive, but not at the same time, while a *full-duplex* RS can transmit and receive at the same time. Full-duplex relaying is still under investigation for deployment in practical systems due to its highly complex (and still not satisfactory for cellular deployment) antenna/RF hardware implementation. On the other hand, half-duplex relaying is more practical due to its much simpler antenna/RF hardware implementation. However, half-duplex relaying incurs a rate penalty (also known as half-duplex loss) as a result of the two or more transmission phases/hops required to transmit a signal from the source to the destination. Also, relays can be categorized as *one-way* (receiving and transmitting in one direction, uplink or downlink) or *two-way* (receiving and transmitting in both directions, uplink and downlink) [93–96]. Two-way relaying eliminates the half-duplex loss of half-duplex relaying by using analog network coding to transmit and receive two messages in two transmission phases [94–96].

Depending on whether a relay is authorized to manage its cell resources (i.e., can generate its cell control messages), relays can also be categorized as *non-transparent* as per IEEE 802.16 (or *Type 1* as per 3GPP) or *transparent* as per IEEE 802.16 (or *Type 2* as per 3GPP) [5, 6]. Type 1 RS has a distinct cell ID and generates its own control messages. Hence it is mainly suited for service coverage extension. On the other hand, a Type 2 RS shares cell ID and control messages with its serving BS (also known as donor eNodeB, DeNB in 3GPP parlance), hence mainly suited for increasing the overall system throughput [5, 6]. For instance, in the 3GPP's Release 10 (LTE-Advanced), a work item on relays was created to support Type 1 relay (with out-of-band, or Type 1a and in-band, or Type 1b variants) for at least coverage

extension, while in the 802.16m, only fixed, two-hop non-transparent in-band relay with distributed scheduling is included [6]. Both 4G technologies (3GPP's LTE-Advanced and IEEE 802.16m) support fixed and two-hop relays [6, 92], with strict backward compatibility requirement that the RS must appear as a BS to the user terminals.

While much simpler relays (analog repeaters) have long been deployed in commercial wireless systems, more advanced relays with signal processing capabilities are envisaged as a powerful alternative to improve coverage, capacity, and reliability of wireless networks [97, 98]. Several of such advanced relaying schemes have been extensively studied in [90, 91, 96, 99–104] and references therein. The two most popular relaying protocols are the *amplify-and-forward* (AF), in which the RS only linearly processes (amplifies) its received signal before transmission [1, 93, 103–109], and the *decode-and-forward* (DF), in which the RS decodes its received signal, then re-encodes it before transmission [97–100, 110–112]. AF relays are simpler and more attractive, but face the problem of noise-amplification, hence they are mainly suited for high SINR environments [8]. With fixed power at all stations, the performance of an AF relay largely depends on the relative position of the relay with respect to the source and the destination [98], with the optimum performance when it is equidistant from both and worse performance otherwise. On the other hand, DF relays offer some advantages such as the possibility of independent rate adaptation and/or scheduling in the access and relay links. In a point-to-point relay network, the performance of DF relays largely depends on the robustness of the relay (BS-RS) link, with the optimum performance when the relay link is robust enough such that there is no decoding error. However, DF relays are much more complex and prone to error propagation and delays. For instance, the decoding delays of DF relays have been shown to be much longer than an LTE sub-frame duration of 1 ms [6, 8].

Mainly due to their attractive features, a lot of research on relay networks has focused on AF-based multi-antenna relay networks [1, 93, 103–109, 113]. For example, [93, 103–107] studied beamforming in MIMO AF-relay channels, including the full and partial CSI cases. More so, MIMO techniques have been combined with AF-relaying in MIMO AF relay channels, as in [113, 114] for single-user MIMO AF-relay channels, and in [1, 93, 108, 109] for multiuser MIMO AF-relay channels. Earlier studies for AF-based MIMO broadcast relay channel (BRC) focused on networks with single-antenna users [1, 108, 109]. More recently, a few results have been reported on multi-user multi-relay networks with multiple antennas at all nodes [115–117], as envisaged in emerging cellular networks [7–10]. However, these studies have so far focused on systems with equal numbers of sources and destinations.

While AF relays are more attractive and simpler than DF relays, they are prone to noise-plus-interference amplification. This is because when the relays simply forward their received signals, they also forward noise and interference, thereby degrading the SINRs in the system. Interestingly, multi-antenna RSs can leverage their spatial dimensions and potentially process their received signals to mitigate or eliminate interference before forwarding the signals to the intended receivers. More so, suitable signal processing will allow the forwarded signals to be independent across the set of RSs. As a simple example, consider that the coverage area for a given RS does not include the entire set of MSs the umbrella BS covers. The RS can then cancel or simply not forward any portion of its received signal not intended for MSs within its coverage area.

2.6.2 Relay Selection

Cooperative communication, by which several nodes cooperate to leverage cooperative diversity, can be used to combat fading, and thereby improving link reliability in wireless systems [98, 118]. In emerging relay-equipped broadband wireless cellular systems, it is envisaged that several relays will be deployed within the coverage of a BS (i.e., within a cell) to extend radio coverage to users with poor coverage from the BS. However, deploying additional nodes (i.e., relays) within the network will result in increased interference, signaling overhead, design complexity and synchronization bottlenecks. Moreover, not all relays will be able (or be allowed) to cooperate during data transmission.

Relay selection, a technique by which subsets of advantageous relays are selected for cooperation can be used to enhance the network performance as well as reduce interference, signaling overhead and design complexity. A common performance metric for relay selection schemes is the diversity order, that is, the degrees of freedom due to the availability of multiple relays in the network. A well designed relay selection scheme achieves full diversity order and has low complexity and overhead.

Several relay selection schemes have been extensively reported in the literature, e.g., [119–131] and references therein. Generally, a relay selection scheme can be categorized as either *single-relay selection*, where only one of the available relays is selected, or *multiple-relay selection*, where more than one of the available relays are selected. Earlier studies on relay selection had focused on *single-relay selection* for networks with single source-destination pair, popularly known as *single-user networks* [120, 124, 126], and later shifted to networks with multiple source-destination pairs, popularly known as *multiple-user networks* [127–129, 132].

For single-relay selection in multi-user multi-relay networks, a 2-step user selection and relay selection for networks with AF relays [127] and DF relays [128] has been recently investigated. In both schemes, the user with the best direct link quality is selected in the first step, while the relay with the best end-to-end SNR to the selected user is selected in the second step. That is, a single user with its best relay is selected at any given time for cooperation, hence there is no competition among the nodes in order to select whom to cooperate with.

Recently, attention has shifted to *multiple-relay selection* in multi-user multi-relay networks [129, 130, 132, 133], where for instance, multiple users simultaneously select the relays. Such approach gives rise to user competition, thereby complicating the relay selection design. In [132], grouping and partner selection in non-altruistic (i.e., relay has its own data) and DF-based wireless relay networks was considered. For each user, the relays with the best channels to the user are selected. In [133], a relay selection strategy that maximizes the minimum achievable rate of all the users is proposed, with focus on the optimality of the scheme. The complexity of the scheme scales linearly in the number of users and quadratically in the number of relays. In [134], a more general system setup in which a user can be helped by multiple relays and each relay can help multiple users was studied using game theory. However, the users and relays were assumed to employ orthogonal channels in order to simplify the analysis. More recently, a multi-user multi-relay network where each relay can help at most one user, while each user can only be helped by a single relay, was studied [129], again assuming orthogonality of users' channels. A new measure of optimality which guarantees uniqueness of the optimal relay selection strategy as well as taking into consideration the performance of all the users was proposed. Relay selection with resource allocation in multiuser multi-relay OFDMA networks was considered in [130]. While the afore mentioned relay selection schemes are for networks with the same type of relays (e.g., AF or DF), [131] recently studied relay selection for networks with different relay types (e.g., AF and DF).

Unfortunately, the afore mentioned schemes are for wireless relay networks with single-antenna nodes. However, transceiver nodes are envisaged to be equipped with multiple antennas in the emerging broadband wireless networks [8], the focus of this work.

Chapter 3

Joint Source-Relay Beamforming in Non-Regenerative MIMO Broadcast Relay Networks

3.1 Introduction

In this chapter, a single-cell AF broadcast relay channel (BRC) with multi-antenna transmitters and receivers is considered. Applying the optimal capacity-achieving strategy for MIMO downlink, DPC at the BS and linear-processing at the RS, our goal is to find the input covariance matrices at the BS and the beamforming matrix at the RS that jointly maximize the network sum rate.

Due to the high complexity and non-convexity of the sum rate optimization problem, we employ uplink-downlink duality result in [135] to transform the BRC problem into a more tractable dual multiple access relay channel (MARC) problem, corresponding to a single-cell uplink. While this MIMO-AF duality was proved for single-antenna-user networks and stated to hold for multi-antenna-user cases, we first provide a detailed proof for the multi-antenna user case and then apply it to our system. However, even for the dual MARC, the sum-rate optimization problem is still non-convex, unlike for the conventional MAC network (without relays) [49, 136]. To make the problem tractable, the relay beamforming matrix is matched to the left and right singular vectors of the access/forward links (RS-to-MS) and relay/backward link (BS-to-RS) channels. With this RS beamforming structure, we propose an iterative algorithm for the sum-rate maximization for the dual MIMO MARC. The proposed scheme follows an alternating-minimization procedure [137–139] over the input covariance matrices at the transmitter and the beamforming matrix at the relay. The proposed algorithm is proved to be convergent.

Furthermore, we derive the mapping from the resulting covariance matrices for

the MARC to the desired covariance matrices for the BRC. Compared with two known single-antenna-user schemes proposed in [1], simulation results show that our proposed scheme outperforms the *all-pass relay* scheme. The proposed scheme also performs similarly to the *SVD-relay* scheme. More so, its achievable sum rate is close to the sum-rate upper bound with the performance gap decreasing with increasing number of antennas per user. It is further observed that improper choice of the ratio between the numbers of antennas at the RS and the BS can result in severe performance degradation due to a degenerate (low-rank) channel condition similar to the well-known pin-hole channel degeneracy for single-user MIMO systems in certain propagation environments, e.g. tunnels [16]. As a result, having more antennas at the RS than at the BS (compared to having more antennas at the BS than at the RS) is desirable for best system performance.

The remaining of this chapter is organized as follows. The system models (the BRC and its dual MARC) and the problem formulation are given in Section 3.2. Sections 3.3 and 3.4 give the main results of this work: the proposed sum-rate optimization algorithm and the mapping from the MARC covariance matrices to the BRC covariance matrices, respectively. Simulation results are presented in Section 3.5 and Section 3.6 concludes the paper. Throughout this chapter, superscripts **B** and **M** are used for the BRC and the MARC parameters, respectively.

3.2 System Models

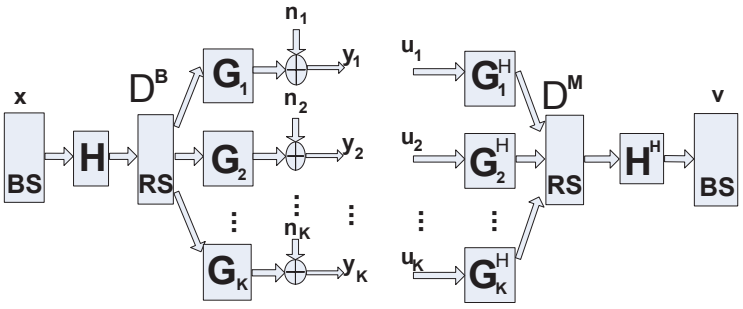

Figure 3.1: A MIMO BRC (left) and its dual MARC (right) networks.

3.2.1 MIMO Broadcast Relay Channel

Consider a single-cell MIMO BRC as depicted by the left-half of Figure 3.1 comprising one BS, one RS, and K users (i.e., MSs). The BS simultaneously transmits to the

users with the aid of the RS. The BS, RS, and User i are equipped with M_b, M_r, and N_i antennas respectively. The BS-to-RS (relay link) and the RS-to-user i (access link) channels are denoted as $\mathbf{H} \in \mathbb{C}^{M_r \times M_b}$ and $\mathbf{G}_i \in \mathbb{C}^{N_i \times M_r}$, respectively. We consider a flat block-fading channel model, where the channels remain constant over a block length and are independent from block to block. Note that frequency-selective broadband channels can be modeled as multiple flat narrow-band sub-channels when orthogonal frequency division multiplexing (OFDM) transmission is used. Hence, any narrow-band technique can be applied in each of the OFDM sub-channels [16]. We also assume that the channels are perfectly known to the BS, RS, and users.

Let $\mathbf{x}_i \in \mathbb{C}^{M_b \times 1}$ denote the codeword transmitted to User i with covariance matrix given by $\mathbf{\Sigma_i} \triangleq \mathbb{E}\{\mathbf{x}_i \mathbf{x}_i^H\}$, where $^{\mathsf{H}}$ denotes the Hermitian. Considering the half-duplex relaying mode, transmission of a block takes place in two orthogonal phases.

Employing DPC, the BS transmits $\mathbf{x} = \sum_{i=1}^{K} \mathbf{x}_i$ to the RS in the first transmission phase, where $\mathbf{x}_i$s are statistically independent Gaussian vectors [48, 49]. Hence, the covariance matrix of $\mathbf{x}$ is given by $\mathbf{\Sigma_x} \triangleq \mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \sum_{i=1}^{K} \mathbf{\Sigma}_i$, where $\mathbb{E}$ is the expectation operator.

In the first transmission phase, the received signal at the RS $\mathbf{y}_R$ is given by

$$\mathbf{y}_R = \mathbf{H}\mathbf{x} + \mathbf{n_R}, \tag{3.1}$$

where $\mathbf{n}_R \in \mathbb{C}^{M_r \times 1}$ is the noise vector at the RS whose elements are independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$. With power constraint $P_T^{\mathbf{B}}$ at the BS, we have

$$\sum_{i=1}^{K} Tr\left\{\mathbf{\Sigma}_i\right\} = P_T^{\mathbf{B}}. \tag{3.2}$$

where Tr is the trace of a square matrix.

In the second transmission phase, the RS uses beamforming matrix $\mathbf{D^B}$ to process its received signal $\mathbf{y}_R$, and subsequently transmits

$$\begin{aligned} \mathbf{x}_R &\triangleq \mathbf{D^B}\mathbf{y}_R \\ &= \mathbf{D^B}\mathbf{H}\mathbf{x} + \mathbf{D^B}\mathbf{n_R} \end{aligned} \tag{3.3}$$

to the K users. The signal vector received by User i, $\mathbf{y}_i$, is therefore given by

$$\mathbf{y}_i = \sum_{j<i} \mathbf{G}_i\mathbf{D^B}\mathbf{H}\mathbf{x}_j + \mathbf{G}_i\mathbf{D^B}\mathbf{H}\mathbf{x}_i + \sum_{j>i} \mathbf{G}_i\mathbf{D^B}\mathbf{H}\mathbf{x}_j + \mathbf{G}_i\mathbf{D^B}\mathbf{n_R} + \mathbf{n}_i, \tag{3.4}$$

where $\mathbf{n}_i \in \mathbb{C}^{N_i \times 1}$ is the noise vector at User i whose elements are i.i.d. $\mathcal{CN}(0,1)$. With power constraint $P_R^{\mathbf{B}}$ at the RS, we have

$$Tr\left\{\mathbf{D^B}\left(\sum_{i=1}^{K} \mathbf{H}\mathbf{\Sigma}_i\mathbf{H}^H + \mathbf{I}\right)\left(\mathbf{D^B}\right)^H\right\} = P_R^{\mathbf{B}}. \tag{3.5}$$

Applying the result of the extension of Costa's *writing on dirty paper* [46] to the Gaussian vector channel case, (Lemma 1 of [140,141]), the BS chooses the codeword $\mathbf{x}_i$ for User i with non-causal knowledge of the interference caused by the already chosen codewords $\{\mathbf{x}_j\}_{j<i}$. More precisely, the BS pre-cancels the backward interference $\sum_{j<i} \mathbf{G}_i\mathbf{D^B}\mathbf{H}\mathbf{x}_j$ caused by the already encoded Users j for $j < i$, while the ith decoder treats the forward interference $\sum_{j>i} \mathbf{G}_i\mathbf{D^B}\mathbf{H}\mathbf{x}_j$ caused by Users j for $j > i$ as additional noise. Hence, the first summation on the right hand side of (3.4) constituting the known interference does not affect the achievable rate of User i, denoted by $R_i^{\mathbf{B}}$. Obviously, the encoding order of the users' messages is important.

For a decreasing encoding order (i.e., User K encoded first, User $K-1$ second, ..., and User 1 last), the achievable rate of User i is given by

$$\begin{aligned} R_i^B &\triangleq I(\mathbf{x}_i;\mathbf{y}_i|\{\mathbf{x}_j\}_{j<i}) \\ &= \frac{1}{2}\log_2 \frac{\left|\mathbf{I}+\mathbf{G}_i\mathbf{D^B}\left(\sum_{j=1}^{i}\mathbf{H}\mathbf{\Sigma}_j\mathbf{H}^H+\mathbf{I}\right)(\mathbf{D^B})^H\mathbf{G}_i^H\right|}{\left|\mathbf{I}+\mathbf{G}_i\mathbf{D^B}\left(\sum_{j=1}^{i-1}\mathbf{H}\mathbf{\Sigma}_j\mathbf{H}^H+\mathbf{I}\right)(\mathbf{D^B})^H\mathbf{G}_i^H\right|}, \end{aligned} \tag{3.6}$$

where $I(x;y)$ is the mutual information between the transmited signal x and the received signal y and $|\mathbf{A}|$ is the determinant of matrix $\mathbf{A}$.

Thus, the network sum rate R_{sum}^B is given by

$$R_{sum}^B = \sum_{i=1}^{K} R_i^{\mathbf{B}}. \tag{3.7}$$

We wish to note that by removing the contributions of the relay in (3.7), that is, the $\frac{1}{2}$ and the $\mathbf{I}$ terms within the brackets in (3.6) while setting $\mathbf{G}_i\mathbf{D^B} = \mathbf{I}$, we obtain the sum rate expression of a conventional MIMO BC (i.e., without relay). While each of the numerator and the denominator for the equivalent MIMO BC is concave in the input covariance matrices, the corresponding sum rate expression is non-concave.

Similarly, the sum rate expression of (3.7) is non-concave in $\mathbf{D^B}$ and $\mathbf{\Sigma}_i$s. Hence its maximization is non-convex and difficult to solve. More so, for a fixed $\mathbf{D^B}$, (3.7) is non-concave in $\mathbf{\Sigma}_i$s and vice versa. This is not surprising since the sum-rate expression for the BC even without relays is not concave in $\mathbf{\Sigma}_i$s [49]. The concavity of a log-determinant function is given by Lemma 3.1.

Lemma 3.1 $\log|\mathbf{A}|$ *is concave on the convex set of positive semi definite (PSD) matrices.*

Proof. See Theorem 7.6.7 on Page 466 of [142]. ■

3.2.2 Dual Multiple Access Relay Channel

Now consider the dual MARC depicted on the right hand side (RHS) of Figure 3.1 arrived at by reversing the directions of the transmissions in the BRC. That is, the users transmit to the BS through a fixed infrastructure-based RS. The user i-to-RS and RS-to-BS MIMO channel matrices are therefore given by $\mathbf{G}_i^H$ and $\mathbf{H}^H$, respectively. Observe that the channel coefficients of the BRC and the dual MARC are the same. For instance, $\mathbf{G}_k(i,j)$ is the channel coefficient from Antenna j of the RS to Antenna i of User k in the BRC as well as the coefficient from Antenna i of the transmitter User k to Antenna j of the RS in the MARC.

Let $\mathbf{u}_i \in \mathbb{C}^{N_i \times 1}$ denote the $N_i \times 1$ transmit vector by User i with covariance matrix defined as $\mathbf{Q_i} \triangleq \mathbb{E}\{\mathbf{u}_i\mathbf{u}_i^H\}$, where $\mathbf{u}_i$s are again statistically independent Gaussian vectors. In the first transmission phase, all users simultaneously transmit to the RS. The RS linearly-processes its received signals from the users with beamforming matrix $\mathbf{D^B}$ and subsequently forwards the amplified signals to the BS in the second phase.

Following similar steps in the BRC, the BS, in the second phase, receives

$$\begin{aligned}\mathbf{v} &= \sum_{i=1}^{K} \mathbf{H}^H\mathbf{D}^{\mathbf{M}}\mathbf{G}_i^H\mathbf{u}_i + \mathbf{H}^H\mathbf{D}^{\mathbf{M}}\mathbf{n_R} + \mathbf{n_d} \\ &= \mathbf{H}^H\mathbf{D}^{\mathbf{M}}\mathbf{G}^H\mathbf{u} + \mathbf{H}^H\mathbf{D}^{\mathbf{M}}\mathbf{n_R} + \mathbf{n_d},\end{aligned} \tag{3.8}$$

where $\mathbf{G}^H \triangleq [\mathbf{G}_1^H, \mathbf{G}_2^H, ..., \mathbf{G}_K^H]$ and $\mathbf{u} \triangleq [\mathbf{u}_1^T, \mathbf{u}_2^T, ...\mathbf{u}_K^T]^T$ are the composite user-RS channels and the transmit signal vector respectively and $^\mathsf{T}$ denotes the transpose. $\mathbf{n_R} \in \mathbb{C}^{M_r \times 1}$ and $\mathbf{n}_d \in \mathbb{C}^{M_b \times 1}$ are the noise vectors at the RS and BS respectively, whose elements are i.i.d. $\mathcal{CN}(0,1)$.

If the sum power across the users and the power of the RS are respectively constrained by $P_T^{\mathbf{M}}$ and $P_R^{\mathbf{M}}$, we have

$$\sum_{i=1}^{K} Tr\left\{\mathbf{Q}_i\right\} = P_T^{\mathbf{M}} \tag{3.9}$$

and

$$Tr\left\{\mathbf{D}^{\mathbf{M}}\left(\sum_{i=1}^{K}\mathbf{G}_i^H\mathbf{Q}_i\mathbf{G}_i + \mathbf{I}\right)\left(\mathbf{D}^{\mathbf{M}}\right)^H\right\} = P_R^{\mathbf{M}}. \tag{3.10}$$

Under a fixed covariance matrix of the effective transmit vectors defined as

$$\mathbf{Q_u} \triangleq \mathbb{E}\{\mathbf{u}\mathbf{u}^H\}, \tag{3.11}$$

the sum rate defined as

$$\mathbf{R}_{sum}^{\mathbf{M}} \triangleq \sum_{i=1}^{K} R_i^{\mathbf{M}} = I(\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_K; \mathbf{v}) \tag{3.12}$$

with User i's rate given by

$$R_i^{\mathbf{M}} = I(\mathbf{u}_i; \mathbf{v} | \{\mathbf{u}_j\}_{j<i}) \tag{3.13}$$

is achievable using a decision-feedback structure with the knowledge of the already decoded codewords $\{\mathbf{u}_j\}_{j<i}$ available before decoding $\mathbf{u}_i$ [141, 143].

Furthermore, we note here that for the uplink-downlink duality to hold, the decoding order of the users' messages in the dual uplink (MARC) must be the reverse of the encoding order in the original downlink (BRC). Therefore, given the decreasing encoding order in the BRC, the decoding of the users signal in the dual MARC should follow an increasing order. That is, User 1 decoded first, User 2 second,..., and User K last. Consequently, the achievable rate for User i, R_i^M, in the MARC is given by

$$\begin{aligned} R_i^{\mathbf{M}} &\triangleq I(\mathbf{u}_i; \mathbf{v} | \{\mathbf{u}_j\}_{j<i}) \\ &= \frac{1}{2}\log_2 \frac{\left|\mathbf{I} + \mathbf{H}^H \mathbf{D}^{\mathbf{M}} \left(\sum_{j=i}^{K} \mathbf{G}_j^H \mathbf{Q}_j \mathbf{G}_j + \mathbf{I}\right) (\mathbf{D}^{\mathbf{M}})^H \mathbf{H}\right|}{\left|\mathbf{I} + \mathbf{H}^H \mathbf{D}^{\mathbf{M}} \left(\sum_{j=i+1}^{K} \mathbf{G}_j^H \mathbf{Q}_j \mathbf{G}_j + \mathbf{I}\right) (\mathbf{D}^{\mathbf{M}})^H \mathbf{H}\right|}, \end{aligned} \tag{3.14}$$

while the sum-rate R_{sum}^M is given by

$$\begin{aligned} \mathbf{R}_{sum}^{\mathbf{M}} &\triangleq \sum_{i=1}^{K} R_i^{\mathbf{M}} = I(\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_K; \mathbf{v}) \\ &= \frac{1}{2}\log_2 \frac{\left|\mathbf{I} + \mathbf{H}^H \mathbf{D}^{\mathbf{M}} \left(\sum_{j=1}^{K} \mathbf{G}_j^H \mathbf{Q}_j \mathbf{G}_j + \mathbf{I}\right) (\mathbf{D}^{\mathbf{M}})^H \mathbf{H}\right|}{\left|\mathbf{I} + \mathbf{H}^H \mathbf{D}^{\mathbf{M}} (\mathbf{D}^{\mathbf{M}})^H \mathbf{H}\right|}. \end{aligned} \tag{3.15}$$

Note that $\mathbf{I} + \mathbf{H}^H \mathbf{D}^{\mathbf{M}} (\mathbf{D}^{\mathbf{M}})^H \mathbf{H}$ is the covariance matrix of the total effective noise at the BS. Similar to the BRC sum-rate expression in (3.7), (3.15) is non-concave in $\mathbf{D}^{\mathbf{M}}$ and $\mathbf{Q}_i$s, hence its maximization is also a non-convex problem. However, unlike in the BRC case, for a fixed RS beamforming matrix $\mathbf{D}^{\mathbf{M}}$ (satisfying the corresponding power constraints), the MARC sum rate (3.15) is concave in $\mathbf{Q}_i$s. These follow from Lemma 1 as well as similar argument provided in Theorem 1 of [143] for conventional MAC (without relays).

3.2.3 Power Constraints

Since there are two transmitters in the network (one in each transmission phase), the system sum-rate can be investigated under two possible power constraints with respect

to the source and the RS namely; the joint power constraint (JPC) and separate power constraint (SPC). Let the total network power be given by $P_T \triangleq P_T^{\mathbf{B}}+P_R^{\mathbf{B}} = P_T^{\mathbf{M}}+P_R^{\mathbf{M}}$.

- Under the JPC, the sum-rate maximization is performed subject to the sum of the average transmit powers at the source and RS given by

$$\sum_{i=1}^{K} Tr\left\{\mathbf{\Sigma}_i\right\} + Tr\left\{\mathbf{D^B}\left(\sum_{i=1}^{K}\mathbf{H\Sigma}_i\mathbf{H}^H + \mathbf{I})(\mathbf{D^B}\right)^H\right\} = P_T \tag{3.16}$$

  for the BRC, and

$$\sum_{i=1}^{K} Tr\left\{\mathbf{Q}_i\right\} + Tr\left\{\mathbf{D^M}\left(\sum_{i=1}^{K}\mathbf{G}_i^H\mathbf{Q}_i\mathbf{G}_i + \mathbf{I}\right)\left(\mathbf{D^M}\right)^H\right\} = P_T \tag{3.17}$$

  for the MARC.

- While for the SPC, the sum-rate maximization is performed subject to the individual average powers at the source and RS given by

$$\sum_{i=1}^{K} Tr\left\{\mathbf{\Sigma}_i\right\} = P_T^{\mathbf{B}} \tag{3.18}$$

$$Tr\left\{\mathbf{D^B}\left(\sum_{i=1}^{K}\mathbf{H\Sigma}_i\mathbf{H}^H + \mathbf{I}\right)\left(\mathbf{D^B}\right)^H\right\} = P_R^{\mathbf{B}} \tag{3.19}$$

  for the BRC, and

$$\sum_{i=1}^{K} Tr\left\{\mathbf{Q}_i\right\} = P_T^{\mathbf{M}} \tag{3.20}$$

$$Tr\left\{\mathbf{D^M}\left(\sum_{i=1}^{K}\mathbf{G}_i^H\mathbf{Q}_i\mathbf{G}_i + \mathbf{I}\right)\left(\mathbf{D^M}\right)^H\right\} = P_R^{\mathbf{M}} \tag{3.21}$$

  for the MARC.

3.2.4 Uplink-Downlink Duality for MIMO AF Relay Channels

Here, we present the duality relationship between the BRC and its dual MARC for multiuser MIMO AF relay channels. This duality gives the conditions, under which the uplink (MARC) and the downlink (BRC) channels expend the same total power (i.e., $P_T \triangleq P_T^{\mathbf{B}} + P_R^{\mathbf{B}} = P_T^{\mathbf{M}} + P_R^{\mathbf{M}}$) as well as achieve the same rate pairs (i.e., $R_i^{\mathbf{B}} = R_i^{\mathbf{M}}$), hence the same sum rate. Uplink-downlink duality [49, 136] has been

extensively used to obtain the sum rate of the non-convex BC by solving a more tractable convex sum rate problem for the dual MAC in conventional MAC and BC systems without relays. However, unlike the MAC-BC duality [49, 136], which does not have the RS beamforming problem, the uplink-downlink duality for MIMO AF-relay channels [135] includes the relationship between the RS beamforming matrices in the BRC and MARC networks, $\mathbf{D^B}$ and $\mathbf{D^M}$ as outlined below.

Theorem 2 of [135] states that given the downlink and uplink channels described by (3.4) and (3.8) with RS beamforming matrices $\mathbf{D^B} = \mathbf{D}$ and $\mathbf{D^M} = c\mathbf{D}^H$ ($c \in \mathbb{R}_+$), the following statements are true:

1. Uplink-downlink duality holds when the total source and relay powers are switched in the dual network, i.e., $P_T^{\mathbf{B}} = P_R^{\mathbf{M}}$, $P_R^{\mathbf{B}} = P_T^{\mathbf{M}}$ and
$$c^2 = \frac{P_R^{\mathbf{M}}}{Tr\left\{\mathbf{D}^H\left(\sum_{i=1}^{K}\mathbf{G}_i^H\mathbf{Q}_i\mathbf{G}_i + \mathbf{I}\right)\mathbf{D}\right\}}. \tag{3.22}$$

2. Under a total network power constraint, uplink-downlink duality holds when $\mathbf{D}$ and $\mathbf{D}^H$ are the relaying matrices used in the downlink and uplink respectively. In other words, uplink-downlink duality holds when $c = 1$ and $P_T^{\mathbf{B}} + P_R^{\mathbf{B}} = P_T^{\mathbf{M}} + P_R^{\mathbf{M}} \triangleq P_T$.

3. The value of c given in (1) and (2) are the only cases where uplink-downlink duality holds for a total network power constraint.

The above duality results also hold with reverse RS beamforming matrices (i.e., $\mathbf{D^B} = c\mathbf{D}^H$; $\mathbf{D^M} = \mathbf{D}$). Cases 1 and 2 above are applicable to the SPC and JPC respectively.

While the authors of [135] proved that the MIMO AF uplink-downlink duality for the single-antenna user networks and stated that it holds for the multi-antenna users case, in Appendix A (also of [29]), we provided a detailed proof for the multi-antenna user case.

3.3 Sum-Rate Maximization

The optimization problems to maximize the BRC sum rate (3.7) and the MARC sum rate (3.15) are non-convex, hence quite challenging. However, for a fixed RS beamforming matrix satisfying the corresponding power constraints, the dual MARC sum rate (3.15) is convex with respect to $\mathbf{Q}_i$s. This is different from the BRC sum rate (3.7), which is non-convex with respect to $\mathbf{\Sigma}_i$s. Hence, we revert to the more tractable dual MARC sum rate of (3.15) (instead of the original BRC sum rate of (3.7)),

and design an iterative alternating-minimization-based algorithm that maximizes the network sum-rate. With the duality result outlined earlier, we choose $\mathbf{D^M} = \mathbf{D}$ and $\mathbf{D^B} = c\mathbf{D}^H$; so that $R_i^{\mathbf{B}} = R_i^{\mathbf{M}}$ and $P_T \triangleq P_T^{\mathbf{B}} + P_R^{\mathbf{B}} = P_T^{\mathbf{M}} + P_R^{\mathbf{M}}$.

However, for fixed $\mathbf{Q}_i$s, the optimization with respect to $\mathbf{D}$ is non-convex. In the following, we outline the approach used in obtaining the MARC sum-rate w.r.t. the beamforming matrix $\mathbf{D}$.

3.3.1 Sum-Rate Maximization: The SPC Case

In this sub-section, we consider the SPC case in detail, while the JPC case is considered in the next sub-section.

Recall that $\mathbf{G}^H \triangleq [\mathbf{G}_1^H, \mathbf{G}_2^H, ..., \mathbf{G}_K^H]$. Define a block diagonal matrix of the users' covariance matrices as $\mathbf{Q} \triangleq \text{diag}[\mathbf{Q}_1, \mathbf{Q}_2, ..., \mathbf{Q}_K]$. Without loss of generality, we assume that all users are equipped with equal number of antennas (i.e., $N_i = N$). Hence, the set of all possible $\mathbf{Q}$, denoted as $\mathcal{Q}$, is a convex set of $KN \times KN$ PSD block-diagonal matrices. Consequently, the sum-rate objective function of (3.15) can be re-written as

$$\begin{aligned} \mathbf{R}_{sum}^{\mathbf{M}} &= \frac{1}{2}\log_2 \frac{\left|\mathbf{I} + \mathbf{H}^H\mathbf{D}\left(\mathbf{G}^H\mathbf{Q}\mathbf{G} + \mathbf{I}\right)\mathbf{D}^H\mathbf{H}\right|}{\left|\mathbf{I} + \mathbf{H}^H\mathbf{D}\mathbf{D}^H\mathbf{H}\right|} \\ &\triangleq \frac{1}{2} f(\mathbf{Q}, \mathbf{D}). \end{aligned} \tag{3.23}$$

Thus, the sum rate maximization problem is given by

$$\begin{aligned} \max_{\mathbf{Q},\mathbf{D}} \quad & \frac{1}{2} f(\mathbf{Q}, \mathbf{D}) \\ s.t. \quad & Tr\{\mathbf{Q}\} = P_R^B; \mathbf{Q} \succeq \mathbf{0} \\ and \quad & Tr\{\mathbf{D}(\mathbf{G}^H\mathbf{Q}\mathbf{G} + \mathbf{I})\mathbf{D}^H\} = P_T^B. \end{aligned} \tag{3.24}$$

Let the singular value decomposition (SVD) of the channel matrices be $\mathbf{H} = \mathbf{U_H}\mathbf{\Sigma_H}\mathbf{V_H}^H$ and $\mathbf{G} = \mathbf{U_G}\mathbf{\Sigma_G}\mathbf{V_G}^H$. We constrain the RS beamforming matrix to have the following structure:

$$\mathbf{D} = \mathbf{U_H}\mathbf{\Sigma_D}\mathbf{V_G}^H, \tag{3.25}$$

where $\mathbf{\Sigma_D}$ is an $M_r \times M_r$ square and diagonal matrix. This RS beamforming structure has been shown to be capacity-optimal for single-user MIMO AF relay channels [113, 114]. It also enables us to express the sum-rate objective function (3.24) in terms

of concave functions as shown below, based on which our scheme is proposed. By employing (3.25), the objective function (3.24) can be re-written as

$$f(\mathbf{Q}, \Lambda_{\mathbf{D}}) = \log_2 \left| \mathbf{I} + {\mathbf{V_G}}^H \left(\mathbf{G}^H \mathbf{Q} \mathbf{G} + \mathbf{I} \right) \mathbf{V_G} \mathbf{\Lambda_H} \mathbf{\Lambda_D} \right| - \log_2 \left| \mathbf{I} + \mathbf{\Lambda_H} \mathbf{\Lambda_D} \right|, \quad (3.26)$$

where $\Lambda_{\mathbf{D}} = {\mathbf{\Sigma_D}}^2 \succeq \mathbf{0}$ and $\mathbf{\Lambda_H} = \mathbf{\Sigma_H} {\mathbf{\Sigma_H}}^T$ are square and diagonal matrices. Moreover, $\mathbf{\Lambda_D} \in \mathcal{D}$ can be seen as the relay power-allocation or eigenmode matrix, where $\mathcal{D}$ is a convex set of $M_r \times M_r$ diagonal matrices with non-negative elements.

To further simplify (3.26), we apply the concept of directional derivative. For a differentiable matrix-valued function $g(\mathbf{X_0})$ and any $\mathbf{X}$, the first order Taylor series expansion on an open interval of $\|\mathbf{X}\|$ is given by

$$g(\mathbf{X}) \triangleq g(\mathbf{X_0}) + Tr\{\nabla_{\mathbf{X_0}}\, g(\mathbf{X_0})^T (\mathbf{X} - \mathbf{X_0})\}, \quad (3.27)$$

where $Tr\{\nabla_{\mathbf{X_0}}\, g(\mathbf{X_0})^T (\mathbf{X} - \mathbf{X_0})\}$ is the directional derivative of $g(\mathbf{X_0})$ in the direction of $\mathbf{X} - \mathbf{X_0}$. Unlike the gradient, directional derivative does not expand dimension of g [144].

Let $\mathbf{Z_0} \triangleq \mathbf{I} + \mathbf{\Lambda_H} \mathbf{\Lambda_0}$. Applying the first-order Taylor expansion of the log-determinant function about the operating point $\mathbf{\Lambda_0}$ to the second term of (3.26), we obtain a lower bound on the sum-rate as follows

$$\begin{aligned} f(\mathbf{Q}, \Lambda_{\mathbf{D}}) \gtrsim\ & \log_2 \left| \mathbf{I} + {\mathbf{V_G}}^H \left(\mathbf{G}^H \mathbf{Q} \mathbf{G} + \mathbf{I} \right) \mathbf{V_G} \mathbf{\Lambda_H} \mathbf{\Lambda_D} \right| \\ & - \log_2 |\mathbf{Z_0}| - Tr\{{\mathbf{Z_0}}^{-1} \mathbf{\Lambda}_H (\mathbf{\Lambda_D} - \mathbf{\Lambda_0})\} \\ \triangleq\ & \tilde{f}(\mathbf{Q}, \Lambda_{\mathbf{D}}), \end{aligned} \quad (3.28)$$

where we have used $\nabla_{\mathbf{X}} \log |\mathbf{AX} + \mathbf{B}| = [(\mathbf{AX} + \mathbf{B})^{-1} \mathbf{A}]^T$, the gradient of a log-determinant function [144] to obtain (3.28). From the concavity of the log-determinant function [145], the inequality in (3.28) holds since $\log |\mathbf{I} + \mathbf{A}_0| + Tr\left\{(\mathbf{I} + \mathbf{A}_0)^{-1} (\mathbf{A} - \mathbf{A}_0)\right\}$ is a global over-estimator of $\log |\mathbf{I} + \mathbf{A}|$ for a fixed $\mathbf{A}_0 \succeq \mathbf{0}$. Furthermore, the above lower bound approximation (3.28) is tight when $\Lambda_{\mathbf{0}}$ is close to $\Lambda_{\mathbf{D}}$, with equality when $\Lambda_{\mathbf{0}} = \Lambda_{\mathbf{D}}$; and holds for any numbers of transmit and receive antennas in the network. Based on (3.28), we then define a new optimization problem (on maximizing the sum-rate lower bound) as

$$\begin{aligned} \max_{\mathbf{Q}, \mathbf{\Lambda_D}} \quad & \frac{1}{2} \tilde{f}(\mathbf{Q}, \Lambda_{\mathbf{D}}) \\ s.t. \quad & Tr\{\mathbf{Q}\} = P_R^B; \mathbf{Q} \succeq \mathbf{0}, \Lambda_{\mathbf{D}} \succeq \mathbf{0} \\ and \quad & Tr\{{\mathbf{V_G}}^H \left(\mathbf{G}^H \mathbf{Q} \mathbf{G} + \mathbf{I} \right) \mathbf{V_G} \mathbf{\Lambda_D}\} = P_T^B. \end{aligned} \quad (3.29)$$

While (3.29) is still non-convex in $\mathbf{Q}$ and $\mathbf{\Lambda_D}$ jointly, it is convex in $\mathbf{Q}$ and $\mathbf{\Lambda_D}$, separately unlike (3.24).

Further to the foregoing, we propose Algorithm 1 for solving the optimization problem of (3.29). The proposed algorithm is iterative in nature and follows an alternating minimization (or projection) procedure [137–139] over the input covariance matrices and the RS power-allocation matrix. For a given iteration, the algorithm maximizes the sum-rate lower bound (3.28) with respect to either $\mathbf{Q}$ or $\Lambda_{\mathbf{D}}$, while the other is fixed. In the next iteration, the optimization is then carried out over the fixed variable in the last iteration, while the optimized variable in the last iteration is now fixed. This process is repeated until convergence.

We initialize Algorithm 1 with the input covariance matrices $\mathbf{Q}$ (in lieu of $\Lambda_{\mathbf{D}}$) at Step 1. Note that as is common practice with non-convex optimization problems, our algorithm can be initialized with different initial values so that the value that results in the highest sum rate will be selected. Then, at Steps 3 and 4, the objective function is iteratively maximized over the RS power allocation matrix $\Lambda_{\mathbf{D}}$ and input covariance matrices $\mathbf{Q}$ respectively. This alternating procedure is primarily aimed at making $\Lambda_{\mathbf{0}}$ as close as possible to $\Lambda_{\mathbf{D}}$ at which point the lower bound approximation of (3.28) is tight. Recall that when $\Lambda_{\mathbf{0}} = \Lambda_{\mathbf{D}}$, then $\tilde{f}(\mathbf{Q}, \Lambda_{\mathbf{D}}) = f(\mathbf{Q}, \Lambda_{\mathbf{D}})$. Consequently, the sum-rate lower bound $\tilde{f}(\mathbf{Q}, \Lambda_{\mathbf{D}})$ well approximates (to the first order) the actual sum-rate $f(\mathbf{Q}, \Lambda_{\mathbf{D}})$ on the termination of our algorithm.

Due to the approximation in (3.28), the solution of our algorithm based on (3.29) is sub-optimal to the optimization with $f(\mathbf{Q}, \Lambda_{\mathbf{D}})$ as the objective directly. Hence, the achievable sum-rate of our scheme is a lower bound on the system sum rate.

Algorithm 1 : Sum-Rate Optimization.

1. **Initialize**: $\mathbf{\Lambda_0}$=$\mathbf{0}_{M_r \times M_r}$, $\mathbf{Q}^{(0)} \triangleq diag\{\mathbf{Q}_1^{(0)}, \mathbf{Q}_2^{(0)}, ..., \mathbf{Q}_K^{(0)}\}$.
2. Set iteration count to k.
3. Obtain $\mathbf{\Lambda_D}^{(k)} = \arg\max_{\mathbf{\Lambda_D} \in \mathcal{D}} \tilde{f}(\mathbf{Q}^{(k-1)}, \mathbf{\Lambda_D})$ from (3.29).
4. Obtain $\mathbf{Q}^{(k)} = \arg\max_{\mathbf{Q} \in \mathcal{Q}} \tilde{f}(\mathbf{Q}, \mathbf{\Lambda_D}^{(k)})$ from (3.29).
5. **if** $\tilde{f}(\mathbf{Q}^{(k)}, \mathbf{\Lambda_D}^{(k)}) - \tilde{f}(\mathbf{Q}^{(k-1)}, \mathbf{\Lambda_D}^{(k)}) < \mu$, break
6. **else**, $\mathbf{\Lambda_0} = \mathbf{\Lambda_D}^{(k)}$, $k = k + 1$ and return to Step 3.
7. Compute $\mathbf{D}^* = \mathbf{U_H} \mathbf{\Lambda_D}^{(k)\frac{1}{2}} \mathbf{V_G}^H$ from (3.25).
8. Obtain $\mathbf{Q}^* = \arg\max_{\mathbf{Q} \succeq \mathbf{0}; \mathbf{Q} \in \mathcal{Q}} f(\mathbf{Q}, \mathbf{D}^*)$ from (3.24).

3.3.2 Sum-Rate Maximization: The JPC Case

Following similar approach to that in the SPC case, the lower bound on the sum-rate maximization for the JPC case can be written as

$$\begin{aligned} \max_{\mathbf{Q},\mathbf{\Lambda_D}} \quad & \frac{1}{2}\tilde{f}(\mathbf{Q},\Lambda_{\mathbf{D}}) \\ s.t. \quad & Tr\{\mathbf{Q}\} + Tr\{{\mathbf{V_G}}^H \left(\mathbf{G}^H\mathbf{Q}\mathbf{G}+\mathbf{I}\right)\mathbf{V_G}\mathbf{\Lambda_D}\} = P_T, \\ & \mathbf{Q}\succeq\mathbf{0}, \mathbf{\Lambda_D}\succeq\mathbf{0} \end{aligned} \tag{3.30}$$

As a result, **Algorithm 1** can be straightforwardly adapted to the JPC case with the respective alternating optimizations at Steps 3 and 4 adjusted subject to the total network power.

3.4 Uplink-to-Downlink Covariance Matrices Transformation

In this work, we also derived the MARC-BRC covariance matrix transformation for mapping the resulting MARC covariances $\mathbf{Q}_i$s back to the desired BRC covariances $\mathbf{\Sigma}_i$s as given in Theorem 3.2 below. Unlike the conventional MAC-to-BC covariance matrix transformation for MIMO networks without relays [49], the input covariance matrix transformation treated here also depends on the relationship between the relay beamforming matrices in the MARC and BRC, for MU-MIMO AF relay channels. Denoting the covariance matrices of the interference-plus-noise at User i in the BRC (denominator of (3.6)) and the MARC (denominator of (3.14)) as $\mathbf{B}_i$ and $\mathbf{M}_i$ respectively, we have

$$\mathbf{B}_i \triangleq \mathbf{I} + \mathbf{G}_i\mathbf{D}^{\mathbf{B}}\left(\sum_{j=1}^{i-1}\mathbf{H}\mathbf{\Sigma}_j\mathbf{H}^H + \mathbf{I}\right)(\mathbf{D}^{\mathbf{B}})^H\mathbf{G}_i^H \tag{3.31}$$

and

$$\mathbf{M}_i \triangleq \mathbf{I} + \mathbf{H}^H\mathbf{D}^{\mathbf{M}}\left(\sum_{j=i+1}^{K}\mathbf{G}_j^H\mathbf{Q}_j\mathbf{G}_j + \mathbf{I}\right)(\mathbf{D}^{\mathbf{M}})^H\mathbf{H}. \tag{3.32}$$

We then express the SVD of User i's *effective* channel matrix as $\mathbf{M}_i^{-\frac{1}{2}}\mathbf{H}^H\mathbf{D}\mathbf{G}_i^H\mathbf{B}_i^{-\frac{1}{2}} = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$, with $\mathbf{\Lambda}_i$ being diagonal and square.

Theorem 3.2 *In multiuser MIMO AF relay channels given by* (3.4) *and* (3.8) *with relay beamforming matrices* $\mathbf{D}^{\mathbf{M}} = \mathbf{D}$ *and* $\mathbf{D}^{\mathbf{B}} = c\mathbf{D}^H$*, the BRC covariance matrices*

$\mathbf{\Sigma}_i$s can be generated from the MARC covariance matrices $\mathbf{Q}_i$s as

$$\mathbf{\Sigma}_i = \frac{1}{c^2}\mathbf{M}_i^{-\frac{1}{2}}\mathbf{U}_i\mathbf{V}_i^H\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}\mathbf{V}_i\mathbf{U}_i^H\mathbf{M}_i^{-\frac{1}{2}}. \tag{3.33}$$

On the other hand, with reverse beamforming matrices in the MARC and BRC (i.e., $\mathbf{D}^{\mathbf{M}} = c\mathbf{D}^H$ and $\mathbf{D}^{\mathbf{B}} = \mathbf{D}$), the BRC covariance matrices can be recovered from the MARC's using

$$\mathbf{\Sigma}_i = c^2\mathbf{M}_i^{-\frac{1}{2}}\bar{\mathbf{U}}_i\bar{\mathbf{V}}_i^H\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}\bar{\mathbf{V}}_i\bar{\mathbf{U}}_i^H\mathbf{M}_i^{-\frac{1}{2}}, \tag{3.34}$$

with the SVD of user i*'s* effective *channel matrix given by $\mathbf{M}_i^{-\frac{1}{2}}\mathbf{H}^H\mathbf{D}^H\mathbf{G}_i^H\mathbf{B}_i^{-\frac{1}{2}} = \bar{\mathbf{U}}_i\bar{\mathbf{\Lambda}}_i\bar{\mathbf{V}}_i^H$. Again, $\mathbf{\Lambda}_i$ is a diagonal and square matrix.*

Proof. See Appendix B. ■

Please note that the above SVDs are precisely the economy-size decompositions.

3.5 Simulation Results

The simulation results of our single-cell MIMO BRC system model are presented. We compare our proposed design with a sum-rate upper bound and other MIMO BRC schemes proposed for single-antenna user networks only. Recall that there are no existing designs for MIMO AF relay channels with multi-antenna users other than our work at the time of this work.

From our BRC system model (3.1) and (3.4), it holds that $I(\mathbf{x};\mathbf{y}_R) \geq I(\mathbf{x};\mathbf{y}_1,\mathbf{y}_2,...,\mathbf{y}_K)$ [94]. Hence, an upper bound on the system sum-rate is the capacity of the first hop (BS-to-RS) channel given by

$$\begin{aligned} &\max_{f(\mathbf{x})} I(\mathbf{x};\mathbf{y}_R) = \max_{\mathbf{\Sigma_x}} \frac{1}{2}\log|\mathbf{I} + \mathbf{H}\mathbf{\Sigma_x}\mathbf{H}^H| \\ &s.t.\ Tr(\mathbf{\Sigma_x}) \leq P_T^{\mathbf{B}};\ \mathbf{\Sigma_x} \succeq \mathbf{0}, \end{aligned} \tag{3.35}$$

where $f(\mathbf{x})$ is the probability distribution of the transmitted signal vector $\mathbf{x}$ [16]. The solution of (3.35) is given by the well-known water-filling technique [16], [15]. We use this upper bound as a benchmark.

Similar to our system model, authors of [1] considered a MIMO BRC in which the BS employs DPC, while the RS operates in the AF mode. Unlike our system model, they considered a BRC network in which the users are each equipped with single antenna. In their *all-pass relay* design, the RS beamforming matrix is chosen as $\mathbf{D}^{\mathbf{B}} = g\mathbf{I}_{M_r\times M_r}$, where $g \geq 0$ is the relay gain. For the *SVD-relay* design, the RS matrix

structure is similar to ours except that the QR decomposition of the RS-to-users channel matrix is used instead of the SVD. That is, with the QR-decomposition of the permuted rows of RS-users channel given as $\bar{\mathbf{G}} \triangleq \Pi\mathbf{G} = \mathbf{L}_G\mathbf{Q}_G$ (where $\mathbf{L}_G$ is a lower triangular matrix, while $\mathbf{Q}_G$ is unitary), the RS matrix is chosen as $\mathbf{D^B} = \mathbf{Q}_G^H\mathbf{K}\mathbf{U_H}^H$, where $\mathbf{K}$ is a diagonal matrix akin to $\mathbf{\Sigma_D}$ in (3.25). These two schemes (unlike ours) involve exhaustive search over all possible user orders, hence these schemes can be impractical for systems with a large number of single-antenna users.

In our simulations, we plot the average sum rate of a MIMO BRC network with M_b antennas at the BS, M_r antennas at the RS, and K users each equipped with an equal number of antennas (i.e., $N_i = N$). We use fixed initial point, where users are allocated equal power. That is, $\mathbf{Q}^{(0)} = \frac{P_T^M}{KN}\mathbf{I}$. The tolerance level is set to $\mu = 10^{-5}$ in Algorithm 1. Furthermore, we used `CVX`, a package for specifying and solving convex programs [146] in our simulations.

Firstly, we consider the effect of source-relay power allocation on the network performance. Let α denote the percentage of the total network power P_T allotted to the source to that to the relay. That is $P_T^M = P_R^B = \alpha P_T$ and $P_R^M = P_T^B = (1-\alpha)P_T$. For instance, $\alpha = 0.3$ implies that 30% and 70% of the total network power are allocated to the source in the MARC (i.e., RS in the BRC) and the RS in the MARC (i.e., BS in the BRC), respectively. In Figure 3.2, the average system sum rate is plotted against the total network power for different α. It can be observed that equal power sharing between the source and the relay gives the highest sum-rate. Consequently, the rest of the results for the SPC case are generated with equi-power allocation at the source and RS (i.e., $\alpha = 0.5$). That is $P_T^M = P_R^B = P_R^M = P_T^B$.

Figure 3.3 depicts the system sum rate for various MIMO BRC schemes for both the JPC and SPC cases. For single-antenna-user networks ($N = 1$), we compare the *proposed SVD design* with the *all-pass relay* and *SVD-relay* designs of [1]. It can be observed that our design outperforms the *all-pass relay* design and performs similarly to the *SVD-relay* design. The latter can be attributed to the similar RS beamforming matrix structure in both designs. Furthermore, the JPC case outperforms (albeit marginally) the SPC case, as expected. This is due to the tighter, but more realistic power constraint with the SPC.

Figure 3.4 further compares the performance gap between the JPC and SPC for different system configurations. Interestingly, the marginal performance gap between the JPC and SPC is maintained for low number of antennas at the users despite the relatively higher number of antennas at the BS and RS . On the contrary, increasing the number of antennas at the users shows an appreciable performance gap between the JPC and SPC despite the relatively low antenna numbers at the BS and RS.

For networks with multi-antenna users (e.g., $N = 5, 7$), we compare the proposed design with the *AF upper bound* as shown in Figure 3.5 for the SPC case. The *all-pass relay* and *SVD-relay* designs do not apply since they are only for single-antenna-user systems. As can be seen, the network sum-rate with $N = 5$ is slightly better compared with $N = 1$. More so, there is no appreciable sum-rate improvement with $N = 7$ compared to $N = 5$.

Next, we investigate the impact of the relative numbers of antennas at the BS and the RS on the system performance.

Figure 3.6 shows the average sum-rate for a fixed number of antennas at the BS ($M_b = 2$) and increasing number of antennas at the relay, M_r. It can be observed that the network sum-rate increases with increasing number of antennas at the relay. This could be attributed to the increasing number of spatial degrees of freedom (DoF) at the RS, which is beneficial for both transmission steps of the network and can be further leveraged to optimally allocate power across the relay eigenmodes.

On the other hand, Figure 3.7 shows the average sum-rate of our system for a fixed number of antennas at the relay ($M_r = 2$) and increasing number of antennas at the BS, M_b. It is observed that the sum rate improvement with increasing M_b is marginal compared to the earlier result of Figure 3.6. For instance, at 15dB SNR, the sum-rate improvement with increasing M_b (from 2 to 16) is 41%, while this improvement is 98% when M_r increases from 2 to 16. This is due to the fact that increasing the number of antennas at the BS only helps the first step of the transmission, while the sum rate of the network depends on both transmission steps.

Figure 3.8, depicting the average sum rate versus the number of antennas at 20 decibel (dB) SNR, further corroborates the above observations. It is evident that the sum-rate improvement with increasing number of antennas at the BS (for fixed M_r) is marginal compared to increasing number of antennas at the RS (for fixed M_b). Also, for a fixed M_r, the network sum rate saturates quickly with increasing value of M_b. This marginal sum-rate increase can be attributed to the bottleneck at the RS, which limits the number of spatial DoFs offered by the multiple antennas at the BS. This phenomenon can be likened to the pin-hole channel degeneracy for single-user MIMO systems in certain propagation environments (e.g., tunnels, street canyons), which results in a degenerate (low-rank) channel condition due to spatially "frozen" signatures [16]. Hence, it can be inferred that having more antennas at the RS than at the BS is desirable for better system performance.

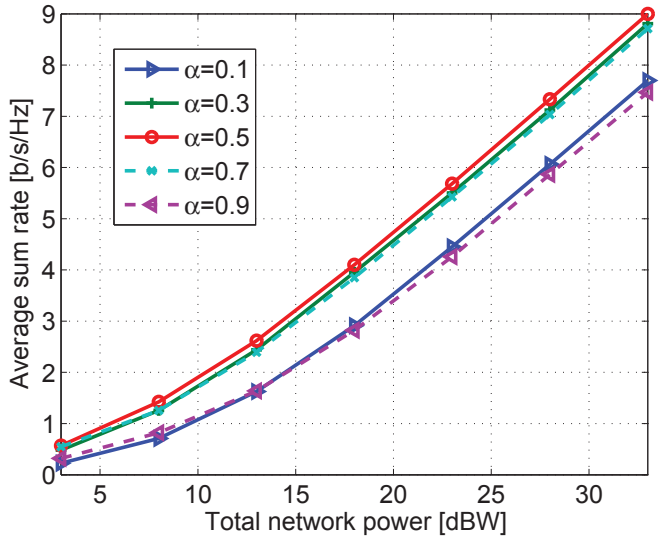

Figure 3.2: Average sum rate of a MIMO BRC network with $M_b = 2$ antennas at the BS, $M_r = 3$ antennas at the RS, $K = 5$ users, and $N = 1$ antenna per user for different power allocations at the source and relay (SPC case).

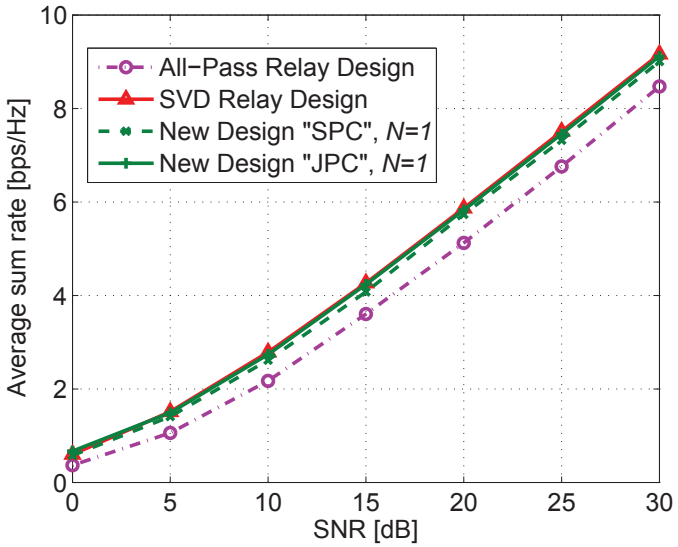

Figure 3.3: Average sum rate of a MIMO BRC network with $M_b = 2$ antennas at the BS, $M_r = 3$ antennas at the RS, $K = 5$ users, and $N = 1$ antenna per user; proposed scheme vs single-antenna schemes of [1].

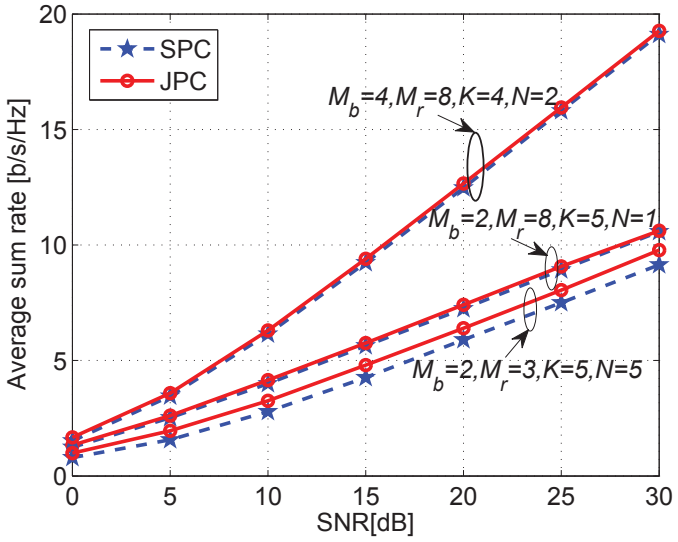

Figure 3.4: Average sum rate of a MIMO BRC network for various system configurations; JPC vs SPC.

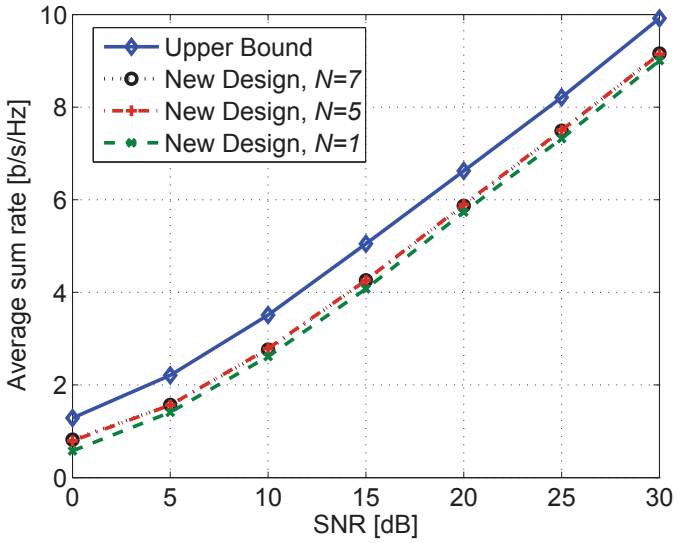

Figure 3.5: Average sum rate of a MIMO BRC network with $M_b = 2$ antennas at the BS, $M_r = 3$ antennas at the RS, $K = 5$ users, and $N = 1, 5, 7$ antennas per user, for the SPC case.

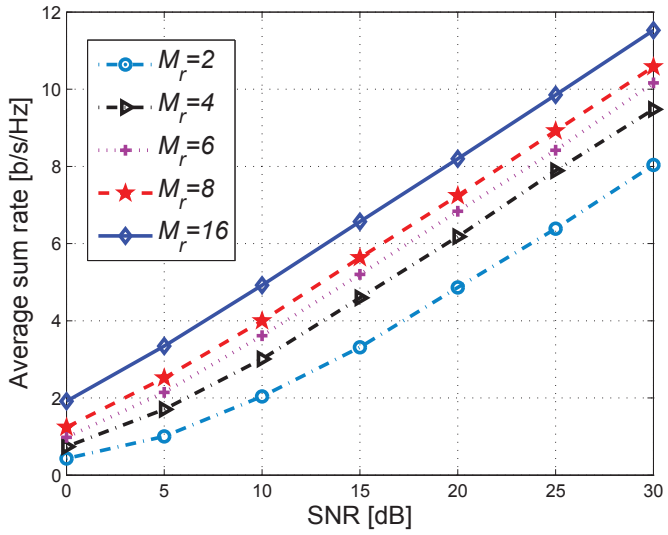

Figure 3.6: Average sum rate of a MIMO BRC network with fixed $M_b = 2$ antennas at the BS, increasing M_r antennas at the RS, $K = 5$ users, and $N = 1$ antenna per user for the SPC case.

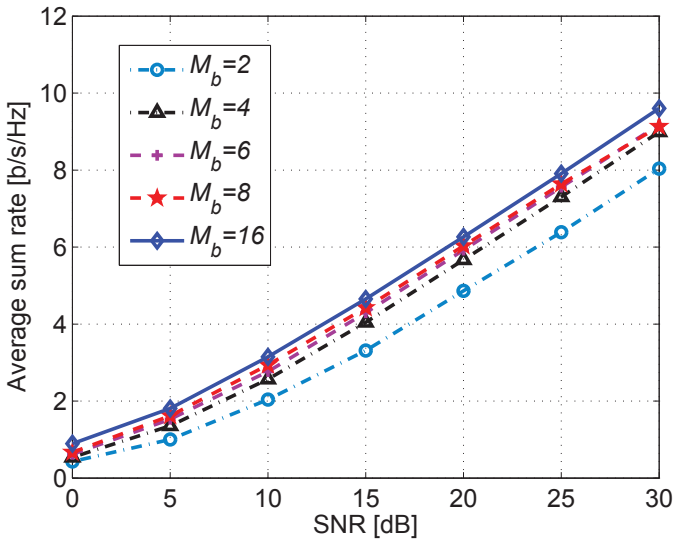

Figure 3.7: Average sum rate of a MIMO BRC network with increasing M_b antennas at the BS, fixed $M_r = 2$ antennas at the RS, $K = 5$ users, and $N = 1$ antenna per user for the SPC case.

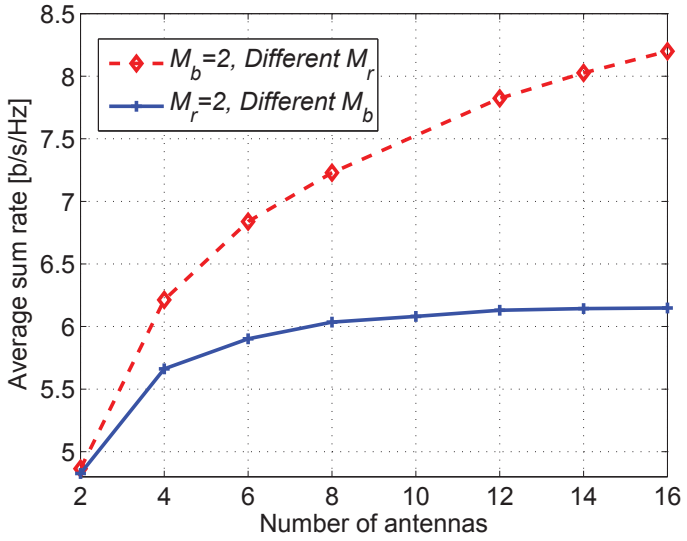

Figure 3.8: Average sum rate of a MIMO BRC network for various number of antennas at the BS and the RS, $K = 5$ users, and $N = 1$ antenna per user for the SPC case at 20 dB SNR.

3.6 Conclusion

In this chapter, we have studied a MIMO broadcast relay channel (BRC), in which a fixed infrastructure-based half-duplex amplify-and-forward multi-antenna RS is used to extend radio coverage (at high data rate) to the cell-edge users in a single-cell downlink. Unlike various schemes in the literature in which users are each equipped with a single antenna, a more general system setup with multi-antenna users has been considered. Applying dirty paper coding at the BS, we aimed to find the users' input covariance matrices at the BS and the beamforming matrix at the RS that jointly maximize the system sum-rate. Due to the high complexity of the problem, uplink-downlink duality is used to transform the problem to a more tractable sum-rate optimization for the dual MARC corresponding to the single-cell uplink. Unlike the MAC without relays, the MARC sum-rate optimization is non-convex. To make the problem tractable, we match the relay beamforming matrix to the right and left singular vectors of its forward/access and backward/relay channels. With this matrix structure, we design an iterative alternating-minimization based algorithm for evaluating the MARC sum rate. Furthermore, the transformation for recovering the desired BRC covariance matrices from the resulting MARC covariance matrices is derived. In comparison with two existing single-antenna-user schemes, simulation results show that the proposed design outperforms the *all-pass relay* design and performs similarly to the *SVD-relay* design. The proposed design also performs close to

a sum-rate upper bound with the performance gap decreasing with increasing number of antennas per user. We further observed that having more antennas at the RS than at the BS (compared to having more antennas at the BS than the RS) is desirable for better system performance.

Chapter 4

Interference Mitigation and User-Relay Association in Multi-Cell MIMO Relay Networks

4.1 Introduction

In Chapter 3, we investigated joint source and relay beamforming design for single-cell single-relay MIMO wireless relay cellar downlink with an AF relay. However, in emerging and future broadband cellular networks, multiple relays are expected to be deployed within a cell.

In this chapter, we consider multi-cell MIMO cellular network, in which multiple MSs communicate with multiple BSs through multiple fixed infrastructure-based in-band AF RSs. Each node is equipped with multiple antennas. Considering the interference-limited regime of the network and with cell-edge user experience enhancement in mind, we study the scenario in which there is no direct link between the BSs and the MSs. Hence, the MSs can only be served with the aid of the RSs. Moreover, due to the presently still insufficient isolation between the RS transmit and receive antenna circuitry as required in full-duplex relaying, we consider a more practical half-duplex relaying.

With co-ordination among the BSs (forming a cluster) and linear processing at the RSs, we aim to jointly design the transmit covariance matrices at the MSs and the beamforming matrices at the RSs such that the system sum rate is maximized. While the above set-up represents an UL-CoMP (albeit with multiple RSs), the resulting optimization problem is non-convex and high-dimensional, unlike for the conventional uplink channels (without RSs). Thus, the sum rate maximization problem of the multi-cell MARC studied in this work is quite complex and challenging.

Furthermore, the existence of multiple AF RSs gives rise to cross-coupling of the

RSs' channels. This is due to the forwarding of the amplified versions of the received signals (including noise and interference) by the RSs, thereby degrading the SINR. Interestingly, multi-antenna AF RSs can utilize their spatial dimension to process their received signals in order to mitigate interference before forwarding the signals to the intended receivers. For example, consider that the coverage area for a given RS does not include the entire set of MSs the umbrella BS covers. The RS can cancel or simply not forward any portion of its received signal not intended for the MSs within its coverage area. Moreover, suitable signal processing at the RSs will allow the forwarded signals to be independent across the set of RSs.

Owing to the high likelihood of the lack of coordination among the RSs in practical systems (due to low-capacity wireless backhaul), we resort to distributed beamforming designs at the RSs, whereby each RS processes its received signals based on its local relay-link (BS-RS) and access-link (RS-MSs) CSI only. That is, there is no sharing of data and/or CSI among the RSs unlike for the BSs. We first seek to decouple the RSs' channels by incorporating interference management mechanism into the RSs' beamforming design. Secondly, we match each RS's beamforming matrix to the corresponding left and right singular vectors of the effective relay and access links, such that the end-to-end channel is diagonalized. With this beamforming structure we then propose an iterative algorithm, which follows alternating minimization procedure (over the transmit covariance matrices at the MSs and the beamforming matrices at the RSs) to maximize the system sum rate. To further enhance the system performance, we propose a low complexity CSI-based association scheme, by which different MSs are assigned to different RSs for service.

Simulation results demonstrate the effectiveness of our designs. In particular, the proposed CSI-based "channel aware" association scheme outperforms a blind "random" MS-RS association scheme, with an increasing performance gap as the number of users increases. Most importantly, the proposed "channel-aware" scheme achieves a reasonable percentage of the performance of the highly complex optimal exhaustive search MS-RS association strategy.

The remaining of this chapter is organized as follows. We present the system model in Section 4.2. Section 4.3 introduces the problem formulation and proposed joint source-relay beamforming design, while Section 4.4 details the proposed MS-RS association scheme. Section 4.5 touches on the CSI issues peculiar to the proposed schemes. Simulation parameters and results are presented in Section 4.6. Section 4.7 concludes the paper.

4.2 System Model

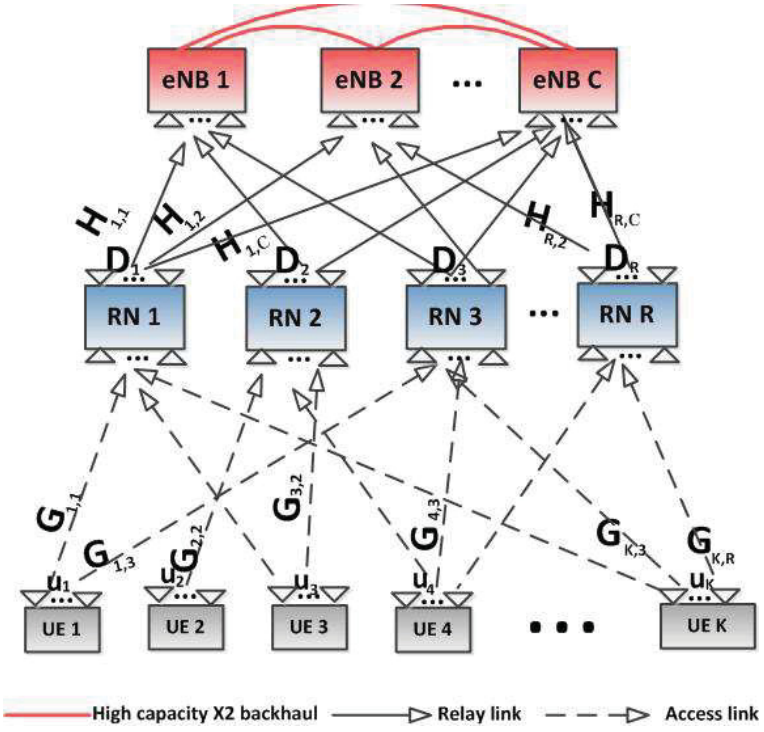

Figure 4.1: A multi-cell MIMO cellular MARC network.

Figure 4.1 depicts a multi-cell MARC comprising C cooperating cells/BSs serving K scheduled MSs with the aid of R fixed AF RSs. The jth BS, rth RS, and ith MS are equipped with $M_{B,j}$, $M_{R,r}$, and N_i antennas, respectively. Thus, the total numbers of antennas over all the BSs, RSs, and MSs are given by $M_B = \sum_{j=1}^{C} M_{B,j}$, $M_R = \sum_{r=1}^{R} M_{R,r}$ and $N_U = \sum_{i=1}^{K} N_i$, respectively. The MIMO channel matrix from the ith MS to the rth RS is denoted by $\mathbf{G}_{i,r} \in \mathbb{C}^{M_{R,r} \times N_i}$ and that from the rth RS to the jth BS is $\mathbf{H}_{r,j} \in \mathbb{C}^{M_{B,j} \times M_{R,r}}$. We assume a flat quasi-static fading channel model (constant over a block length and independent from block to block).

Furthermore, we assume that the BSs employ joint processing variant of CoMP in order to mitigate inter-cell interference (ICI). While coordination over the whole network will (in principle) completely eliminate ICI, such approach demands very large backhaul capacity due to the channel state information (CSI) and (in case of joint processing CoMP) data exchange involved, and also suffers from overwhelming signaling load for channel estimation. Additionally, coordination of distant nodes is not necessary (and not feasible), because of very large path loss. As a result, we employ limited coordination (a.k.a. clustered network MIMO), whereby only a subset of BSs (or cells) are coordinated [81, 82]. Limited coordination, however,

converts inter-cell interference to inter-cluster interference, which degrades the SINRs of the cluster-edge MSs. This cluster edge effect can be "averaged out" over time by periodically reassigning the BSs that form a cluster [84, 85] such that MSs near the edge of a cluster during one interval are closer to the interior of the cluster area in another interval.

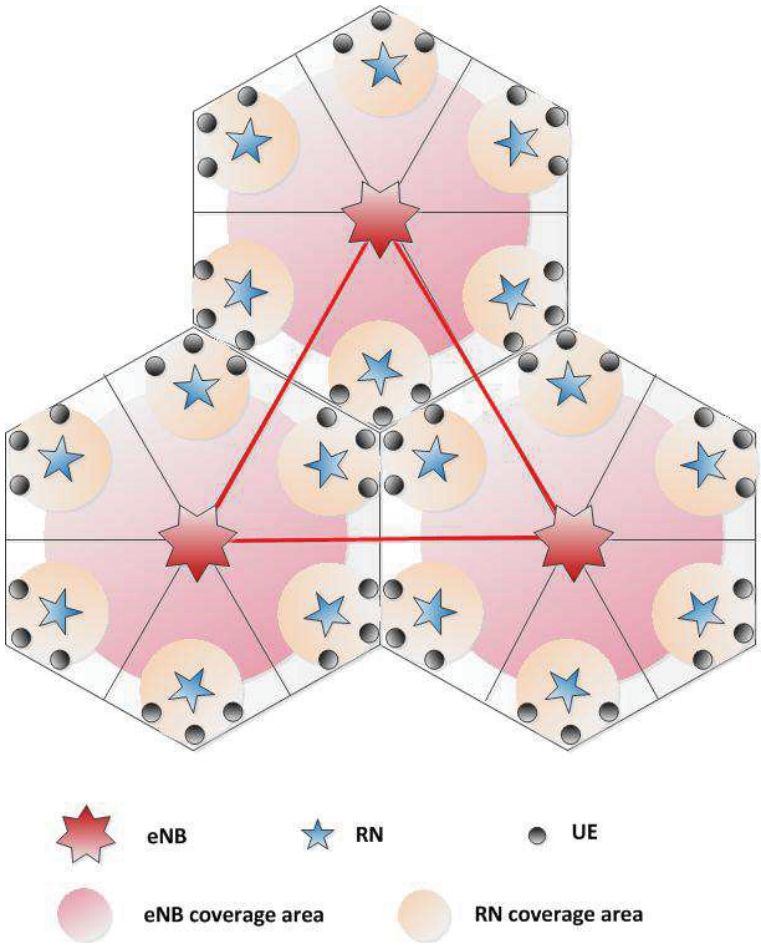

Figure 4.2: A 3-cell cooperative MIMO relay network

Figure 4.2 illustrates a 3-cell cellular relay network, where $C = 3$ BSs are coordinated (i.e., the triangular area comprising 3 contiguous sectors of the cells at the center) to jointly detect the transmitted signals by the K scheduled cell-edge MSs with the aid of AF RSs. As depicted, the MSs are outside the coverage of the BSs and therefore rely on the RSs for service. Also, a MS within the coverage of an RS causes interference at other RSs, herein called inter-relay interference (IRI). By simply forwarding their received signals, the AF RSs also forward noise and IRI, thereby degrading the network SINR. Moreover, even with uncorrelated transmitted signals by the MSs, the transmitted signals by the RSs are correlated due to IRI,

thereby complicating the RSs' beamforming design, as will be shown later. However, with multiple antennas at the RSs, each RS can cancel IRI (interference from non-associated MSs) and forward only useful signals (from associated MSs), thereby simplify the beamforming design as well as enhance the network SINR [147, 148].

Due to the low-capacity of the (wireless) backhaul between the RSs, which makes it difficult for RSs to cooperate, we resort to distributed beamforming designs at the RSs. Thus, each RS knows only its local channels' (both relay and access links) CSI in order to design its beamforming matrix. On the contrary, the BSs are assumed to know all the relay (BS-RS) and access (RS-MS) links CSI, thanks to the high-speed low-latency backhaul links between the BSs. That is, there is no sharing of data or CSI among the RSs unlike for the BSs. The MSs do not need any CSI knowledge for their transmissions.

Given the half-duplex TDM mode of operation of RSs, the transmission of a symbol block takes place in two hops. In the first hop, each MS transmits to its serving or "best" RS, while causing interference at the non-serving RSs. The RSs simultaneously forward amplified versions of their received signals to the C BSs (i.e., CoMP-BS) in the second hop.

Let $\mathcal{S}_r$ denote the set of $K_r \leq K$ MSs served by the rth RS and $\mathcal{S}_r^{\perp}$ the set of $\tilde{K}_r = K - K_r$ interfering MSs not served by the rth RS. Thus, $K = \sum_{r=1}^{R} K_r$. Also, let $\mathbf{u}_i \in \mathbb{C}^{N_i \times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_i)$ be the transmitted signal vector of the ith MS, with covariance matrix defined as $\mathbf{Q}_i = \mathbb{E}\{\mathbf{u}_i \mathbf{u}_i^H\}$. The composite transmitted signal vector by the MSs is given by $\mathbf{u} \triangleq [\mathbf{u}_1^T \ \mathbf{u}_2^T \ ... \ \mathbf{u}_K^T]^T$, with covariance matrix $\mathbf{Q_u} \triangleq \mathbb{E}\{\mathbf{u}\mathbf{u}^H\} = diag(\mathbf{Q}_1, \mathbf{Q}_2, ..., \mathbf{Q}_K)$, which is a diagonal matrix of the MSs' transmit covariance matrices. With average transmit power $P_{U,i}$ at the ith MS, we have

$$Tr\{\mathbf{Q}_i\} \leq P_{U,i}. \tag{4.1}$$

The received signal by the rth RS in the first hop is given by

$$\mathbf{y}_{R,r} = \sum_{i=1}^{K} \mathbf{G}_{i,r} \mathbf{u}_i + \mathbf{n}_{R,r}, \tag{4.2}$$

where $\mathbf{n}_{R,r} \in \mathbb{C}^{M_{R,r} \times 1}$ is the noise vector at the rth RS whose elements are i.i.d. $\mathcal{CN}(0, 1)$.

Expressing the received signal by the rth RS in terms of the "desired" and "interference" components, equation (4.2) becomes

$$\mathbf{y}_{R,r} = \sum_{i=1, i \in \mathcal{S}_r}^{K_r} \mathbf{G}_{i,r} \mathbf{u}_i + \sum_{j=1, j \in \mathcal{S}_r^{\perp}}^{\tilde{K}_r} \mathbf{G}_{j,r} \mathbf{u}_j + \mathbf{n}_{R,r} \tag{4.3}$$

$$= \hat{\mathbf{G}}_r \hat{\mathbf{u}}_r + \tilde{\mathbf{G}}_r \tilde{\mathbf{u}}_r + \mathbf{n}_{R,r},$$

where $\hat{\mathbf{G}}_r \in \mathbb{C}^{M_{R,r} \times \bar{N}_r}$ and $\tilde{\mathbf{G}}_r \in \mathbb{C}^{M_{R,r} \times (N_U - \bar{N}_r)}$ are the composite channel matrices of the desired and interference signals at the rth RS respectively. That is, the total number of transmit antennas across the MSs served by the rth RS is given by $\bar{N}_r = \sum_{i=1, i \in \mathcal{S}_r}^{K_r} N_i$. Similarly, $\hat{\mathbf{u}}_r \in \mathbb{C}^{\bar{N}_r \times 1}$ and $\tilde{\mathbf{u}}_r \in \mathbb{C}^{(N_U - \bar{N}_r) \times 1}$ are the corresponding transmitted signal vectors from the K_r desired MSs and $\tilde{K}_r$ interfering MSs to the rth RS with covariance matrices respectively defined as: $\hat{\mathbf{Q}}_r \triangleq \mathbb{E}\{\hat{\mathbf{u}}_r \hat{\mathbf{u}}_r^H\} = diag\{\mathbf{Q}_i,\ i \in \mathcal{S}_r\}$ and $\tilde{\mathbf{Q}}_r \triangleq \mathbb{E}\{\tilde{\mathbf{u}}_r \tilde{\mathbf{u}}_r^H\} = diag\{\mathbf{Q}_j,\ j \in \mathcal{S}_r^{\perp}\}$.

The rth RS linearly processes its received signal with beamforming matrix $\mathbf{D}_r$ and then transmits

$$\mathbf{x}_{R,r} = \mathbf{D}_r \mathbf{y}_{R,r} \tag{4.4}$$

to the CoMP-BS in the second hop. With average transmit power $P_{R,r}$ at the rth RS, we have

$$Tr\left\{\mathbf{D}_r \left(\sum_{i=1}^{K} \mathbf{G}_{i,r} \mathbf{Q}_i \mathbf{G}_{i,r}^H + \mathbf{I}\right) \mathbf{D}_r^H\right\} \leq P_{R,r}. \tag{4.5}$$

The $M_B \times 1$ received signal vector by the CoMP-BS is given by

$$\begin{aligned} \mathbf{y} &\triangleq \sum_{r=1}^{R} \mathbf{H}_r \mathbf{x}_{R,r} + \mathbf{n}_B \\ &= \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\sum_{i=1}^{K} \mathbf{G}_{i,r} \mathbf{u}_i + \mathbf{n}_{R,r}\right) + \mathbf{n}_B, \end{aligned} \tag{4.6}$$

where $\mathbf{H}_r \triangleq [\mathbf{H}_{r,1}^T\ \mathbf{H}_{r,1}^T\ ...\ \mathbf{H}_{r,C}^T]^T$ is the aggregate matrix of the rth RS's relay link (RS-BS) MIMO channel.

$\mathbf{n}_B \in \mathbb{C}^{M_B \times 1}$ is the noise vector at the CoMP-BS, whose elements are i.i.d. $\mathcal{CN}(0,1)$.

As evident in Equations (4.2) and (4.6), the 1st and 2nd hop channels are MACs. However, unlike the 1st hop channel comprising R Gaussian MACs with uncorrelated MS transmissions, $\big(i.e., \mathbb{E}\{\mathbf{u}_i \mathbf{u}_j^H\} = \mathbf{0}, \forall i \neq j\big)$, the 2nd hop channel is a Gaussian MAC with correlated RS transmissions, $\big(i.e., \mathbb{E}\{\mathbf{x}_{R,r} \mathbf{x}_{R,l}^H\} \neq \mathbf{0}, \forall r \neq l\big)$. The later is due to the forwarded interference (i.e., IRI) by the RSs from their non-serving MSs, causing coupling of the RSs' channels.

4.3 Problem Formulation and System Design

In this section, we present a joint design of MSs' transmit covariance and RS's beamforming matrices based on the formulated sum rate problem as outlined below. Unlike

other relay beamforming designs in the literature, our design takes into consideration interference mitigation at the RSs in order to decouple the RSs' channel due to the forwarded interference.

4.3.1 Network Sum Rate Formulation

Assuming successive decoding, the optimal (capacity-achieving) strategy for MAC, the CoMP-BS decodes the ith MS's signal $\mathbf{u}_i$ with non-causal knowledge of the already decoded MSs' signals $\{\mathbf{u}_j\}_{j<i}$ [141], [143]. We assume that the noises at the receivers are independent with the transmitted signals by the MSs. With an increasing decoding order $\pi'\{i\} = \{1, 2, ..., K\}$, (i.e., MS 1 is decoded first, MS 2 second,..., and MS K last), the achievable rate for the ith MS is given by

$$
\begin{aligned}
R_i &\triangleq I(\mathbf{u}_i; \mathbf{y} | \{\mathbf{u}_j\}_{j<i}) \qquad (4.7)\\
&= \frac{1}{2}\log_2 \frac{\left| \mathbf{I} + \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \mathbf{D}_r^H \mathbf{H}_r^H + \sum_{r=1}^{R}\sum_{l=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\sum_{j=i}^{K} \mathbf{G}_{j,r} \mathbf{Q}_j \mathbf{G}_{j,l}^H \right) \mathbf{D}_l^H \mathbf{H}_l^H \right|}{\left| \mathbf{I} + \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \mathbf{D}_r^H \mathbf{H}_r^H + \sum_{r=1}^{R}\sum_{l=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\sum_{j=i+1}^{K} \mathbf{G}_{j,r} \mathbf{Q}_j \mathbf{G}_{j,l}^H \right) \mathbf{D}_l^H \mathbf{H}_l^H \right|}
\end{aligned}
$$

Thus, the sum rate is given by

$$
\begin{aligned}
\mathbf{R}_{sum} &\triangleq I(\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_K; \mathbf{y}) \qquad (4.8)\\
&= \frac{1}{2}\log_2 \frac{|\mathbf{R_{yy}}|}{|\mathbf{R_{nn}}|},
\end{aligned}
$$

where $\mathbf{R_{nn}}$ and $\mathbf{R_{yy}}$ are the auto-covariance matrices of the effective noise and received signals given by

$$
\mathbf{R_{nn}} = \mathbf{I} + \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \mathbf{D}_r^H \mathbf{H}_r^H \qquad (4.9)
$$

and

$$
\begin{aligned}
\mathbf{R_{yy}} &= \mathbf{I} + \sum_{r=1}^{R}\sum_{l=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\mathbf{G}_r \mathbf{Q_u} \mathbf{G}_l^H \right) \mathbf{D}_l^H \mathbf{H}_l^H \qquad (4.10)\\
&= \mathbf{I} + \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\mathbf{G}_r \mathbf{Q_u} \mathbf{G}_r^H + \mathbf{I} \right) \mathbf{D}_r^H \mathbf{H}_r^H \qquad (4.11)\\
&\quad + \sum_{r=1}^{R}\sum_{l=1,l\neq r}^{R} \mathbf{H}_r \mathbf{D}_r \left(\mathbf{G}_r \mathbf{Q_u} \mathbf{G}_l^H \right) \mathbf{D}_l^H \mathbf{H}_l^H
\end{aligned}
$$

The last term of equation (4.11) shows the cross-coupling at the RSs due to the forwarded interference by the RSs.

4.3.2 Relay Beamforming Design with Interference Cancelation

To design the RSs' beamforming matrices, we aim to jointly decouple the coupled RSs' channels as well as ensure distributed linear processing at the RSs. The later implies that each RS's beamforming matrix should only depend on its local CSI to the MSs and BSs (with no knowledge of the CSI from other RSs to the MSs and BSs). This is necessitated by the improbability of cooperation among the RSs due to their low-capacity wireless backhaul.

Re-writing (4.11) in terms of interference and desired signals, we obtain

$$\mathbf{R_{yy}} = \mathbf{I} + \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\hat{\mathbf{G}}_r \hat{\mathbf{Q}}_r \hat{\mathbf{G}}_r^H + \tilde{\mathbf{G}}_r \tilde{\mathbf{Q}}_r \tilde{\mathbf{G}}_r^H + \mathbf{I} \right) \mathbf{D}_r^H \mathbf{H}_r^H \tag{4.12}$$
$$+ \sum_{r=1}^{R} \sum_{l=1, l\neq r}^{R} \mathbf{H}_r \mathbf{D}_r \Big(\hat{\mathbf{G}}_r \mathbb{E}\{\hat{\mathbf{u}}_r \tilde{\mathbf{u}}_l^H\} \tilde{\mathbf{G}}_l^H + \tilde{\mathbf{G}}_r \mathbb{E}\{\tilde{\mathbf{u}}_r \hat{\mathbf{u}}_l^H\} \hat{\mathbf{G}}_l^H + \tilde{\mathbf{G}}_r \mathbb{E}\{\tilde{\mathbf{u}}_r \tilde{\mathbf{u}}_l^H\} \tilde{\mathbf{G}}_l^H \Big) \mathbf{D}_l^H \mathbf{H}_l^H .$$

Let us denote the composite channel matrix of the rth RS as

$$\acute{\mathbf{G}}_r \triangleq [\hat{\mathbf{G}}_r \ \tilde{\mathbf{G}}_r]. \tag{4.13}$$

The pseudo-inverse of $\acute{\mathbf{G}}_r$ is defined as

$$\grave{\mathbf{G}}_r = (\acute{\mathbf{G}}_r^H \acute{\mathbf{G}}_r)^{-1} \acute{\mathbf{G}}_r^H \triangleq \begin{bmatrix} \check{\mathbf{G}}_r \\ \dot{\mathbf{G}}_r \end{bmatrix} \tag{4.14}$$

for $r = 1, 2, ..., R$. Taking the QR decomposition of $\check{\mathbf{G}}_r \in \mathbb{C}^{\bar{N}_r \times M_{R,r}}$, we obtain

$$\check{\mathbf{G}}_r = \check{\mathbf{R}}_r \check{\mathbf{Q}}_r, \tag{4.15}$$

where $\check{\mathbf{R}}_r \in \mathbb{C}^{\bar{N}_r \times \bar{N}_r}$ is a lower triangular matrix, while $\check{\mathbf{Q}}_r \in \mathbb{C}^{\bar{N}_r \times M_{R,r}}$ (with $\check{\mathbf{Q}}_r \check{\mathbf{Q}}_r^H = \mathbf{I}$) is a matrix whose rows form an orthonormal basis for $\check{\mathbf{G}}_r$.

Since $\grave{\mathbf{G}}_r \acute{\mathbf{G}}_r = \mathbf{I}$ (provided that $M_{R,r} \geq N_U$) and $\check{\mathbf{R}}_r$ is invertible,

$$\check{\mathbf{G}}_r \tilde{\mathbf{G}}_r \triangleq \check{\mathbf{R}}_r \check{\mathbf{Q}}_r \tilde{\mathbf{G}}_r = \mathbf{0} \tag{4.16}$$

which implies that $\check{\mathbf{Q}}_r \tilde{\mathbf{G}}_r = \mathbf{0}$. That is, for complete de-correlation of the correlated RSs' transmissions, the number of antennas at each RS should be at least the total number of transmit antennas at the K scheduled MSs.

By designing the rth RS's beamforming matrix $\mathbf{D}_r$ as a linear combination of $\check{\mathbf{Q}}_r$, that is,

$$\mathbf{D}_r \triangleq \mathbf{A}_r \check{\mathbf{Q}}_r, \tag{4.17}$$

the auto-covariance matrix of (4.12) becomes

$$\mathbf{R_{yy}} = \mathbf{I} + \sum_{r=1}^{R} \mathbf{H}_r \mathbf{D}_r \left(\hat{\mathbf{G}}_r \hat{\mathbf{Q}}_r \hat{\mathbf{G}}_r^H + \mathbf{I} \right) \mathbf{D}_r^H \mathbf{H}_r^H. \tag{4.18}$$

Hence, the above beamforming structure at the rth RS in (4.17) not only eliminates interference from its non-serving MSs, but also eliminates the cross-coupling of the RSs' channels. We also need to mention that the above interference-cancelation approach at the RSs follows the well-known zero-forcing interference-mitigation strategy employed in multi-user MIMO systems with multiple-antenna MSs (i.e., block-diagonalization) [59]. Furthermore, in what follows, we design $\mathbf{D}_r$ (via $\mathbf{A}_r$) such that the effective access and relay links of the rth RS's channels are diagonalized, thereby simplifying the MIMO receiver processing [16].

Note that the maximum number of end-to-end independent streams that can be transmitted via the rth RS is given by

$$L_r \triangleq \min\{M_B, M_{R,r}, \bar{N}_r\}. \tag{4.19}$$

By the definition of N_U, when $M_B \geq N_U$, we have $M_B \geq \bar{N}_r$. Hence, when $M_B \geq \bar{N}_r$, $L_r = \bar{N}_r$.

Taking the SVDs of the effective relay and access links of the rth RS as

$$\check{\mathbf{Q}}_r \hat{\mathbf{G}}_r = \mathbf{U}_r \mathbf{\Lambda}_r \mathbf{V}_r^H \tag{4.20}$$

and

$$\mathbf{H}_r = \mathbf{U}_{\mathbf{H}_r} \mathbf{\Sigma}_{\mathbf{H}_r} \mathbf{V}_{\mathbf{H}_r}^H, \tag{4.21}$$

we choose the rth RS's beamforming matrix as

$$\mathbf{D}_r = \mathbf{V}_{\bar{\mathbf{H}}_r} \mathbf{\Lambda}_{\mathbf{D}_r}^{1/2} \mathbf{U}_r^H \check{\mathbf{Q}}_r, \tag{4.22}$$

where ${\Lambda_{\mathbf{D}_r}}^{1/2} \in \mathbb{C}^{\bar{N}_r \times \bar{N}_r}$ is the power allocation matrix of the rth RS. $\mathbf{V}_{\bar{\mathbf{H}}_r} \in \mathbb{C}^{\bar{M}_{R,r} \times L_r}$ contains the first L_r columns of $\mathbf{V}_{\mathbf{H}_r}$. Furthermore, we denote the eigen-decomposition of the input covariance matrix of $\hat{\mathbf{Q}}_r$ as

$$\hat{\mathbf{Q}}_r = \mathbf{V}_r \mathbf{\Lambda}_{\mathbf{F}_r} \mathbf{V}_r^H, \tag{4.23}$$

where $\mathbf{\Lambda}_{\mathbf{F}_r} \triangleq diag\left(\mathbf{\Lambda}_{\mathbf{F}_{u,1}}, \mathbf{\Lambda}_{\mathbf{F}_{u,2}}, \ldots, \mathbf{\Lambda}_{\mathbf{F}_{u,Kr}} \right)$ is a diagonal matrix of the power allocation matrices of the K_r MSs served by the rth RS.

4.3.3 Sum Rate Optimization and Algorithm

Employing the above matrix structures (4.22) and (4.23), and re-writing the network sum rate in terms of the MSs and RSs power allocation matrices, we obtain

$$\begin{aligned}\mathbf{R}_{sum} &\triangleq \sum_{i=1}^{K} R_i = I(\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_K; \mathbf{y}) \\ &= \frac{1}{2}\log_2 \frac{\left|\mathbf{I}+\sum_{r=1}^{R} \mathbf{U}_{\bar{\mathbf{H}}_r}\mathbf{\Sigma}_{\bar{\mathbf{H}}_r}\left(\mathbf{\Lambda}_r\mathbf{\Lambda}_{\mathbf{F}_r}+\mathbf{I}\right)\mathbf{\Lambda}_{\mathbf{D}_r}\mathbf{\Sigma}^T_{\bar{\mathbf{H}}_r}\mathbf{U}^H_{\bar{\mathbf{H}}_r}\right|}{\left|\mathbf{I}+\sum_{r=1}^{R} \mathbf{U}_{\bar{\mathbf{H}}_r}\mathbf{\Sigma}_{\bar{\mathbf{H}}_r}\mathbf{\Lambda}_{\mathbf{D}_r}\mathbf{\Sigma}^T_{\bar{\mathbf{H}}_r}\mathbf{U}^H_{\bar{\mathbf{H}}_r}\right|}.\end{aligned} \tag{4.24}$$

$\mathbf{U}_{\bar{\mathbf{H}}_r}$ and $\mathbf{\Sigma}_{\bar{\mathbf{H}}_r}$ contain the first L_r columns of $\mathbf{U}_{\mathbf{H}_r}$ and the corresponding L_r singular values respectively. More compactly, the network sum rate (4.24) can be written as

$$\begin{aligned}\mathbf{R}_{sum} &= \frac{1}{2}\log_2\left|\mathbf{I}+\bar{\mathbf{U}}_{\mathbf{H}}\bar{\mathbf{\Sigma}}_{\mathbf{H}}\left(\bar{\mathbf{\Lambda}}\bar{\mathbf{\Lambda}}_{\mathbf{F}}+\mathbf{I}\right)\bar{\mathbf{\Lambda}}_{\mathbf{D}}\bar{\mathbf{\Sigma}}^T_{\mathbf{H}}\bar{\mathbf{U}}^H_{\mathbf{H}}\right| \\ &\quad -\frac{1}{2}\log_2\left|\mathbf{I}+\bar{\mathbf{U}}_{\mathbf{H}}\bar{\mathbf{\Sigma}}_{\mathbf{H}}\bar{\mathbf{\Lambda}}_{\mathbf{D}}\bar{\mathbf{\Sigma}}^T_{\mathbf{H}}\bar{\mathbf{U}}^H_{\mathbf{H}}\right| \\ &\triangleq \frac{1}{2}f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}),\end{aligned} \tag{4.25}$$

where

$$\bar{\mathbf{U}}_H \triangleq \left[\mathbf{U}_{\bar{\mathbf{H}}_1}, \mathbf{U}_{\bar{\mathbf{H}}_2}, \ldots, \mathbf{U}_{\bar{\mathbf{H}}_R}\right] \tag{4.26}$$

$$\bar{\mathbf{\Sigma}}_{\mathbf{H}} \triangleq diag\left(\mathbf{\Sigma}_{\bar{\mathbf{H}}_1}, \mathbf{\Sigma}_{\bar{\mathbf{H}}_2}, \ldots \mathbf{\Sigma}_{\bar{\mathbf{H}}_R}\right) \tag{4.27}$$

$$\bar{\mathbf{\Lambda}} \triangleq diag\left(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \ldots \mathbf{\Lambda}_R\right) \tag{4.28}$$

$$\bar{\mathbf{\Lambda}}_{\mathbf{F}} \triangleq diag\left(\mathbf{\Lambda}_{\mathbf{F}_1}, \mathbf{\Lambda}_{\mathbf{F}_2}, \ldots \mathbf{\Lambda}_{\mathbf{F}_R}\right) \tag{4.29}$$

$$\bar{\mathbf{\Lambda}}_{\mathbf{D}} \triangleq diag\left(\mathbf{\Lambda}_{\mathbf{D}_1}, \mathbf{\Lambda}_{\mathbf{D}_2}, \ldots \mathbf{\Lambda}_{\mathbf{D}_R}\right). \tag{4.30}$$

Thus, the sum rate maximization problem over the MSs and RSs power allocation matrices is formulated as

$$\begin{aligned}\max_{\mathbf{\Lambda}_{\mathbf{F}_{u,i}}, \mathbf{\Lambda}_{\mathbf{D}_r}} \quad & \frac{1}{2}f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}) \\ \text{s.t.} \quad & Tr(\mathbf{\Lambda}_{\mathbf{F}_{u,i}}) \leq P_{U,i}\ ; \mathbf{\Lambda}_{\mathbf{F}_{u,i}} \succeq \mathbf{0} \\ \text{and} \ & Tr\{\left(\mathbf{\Lambda}_r\mathbf{\Lambda}_{\mathbf{F}_r}+\mathbf{I}\right)\mathbf{\Lambda}_{\mathbf{D}_r}\} \leq P_{R,r}\ ; \mathbf{\Lambda}_{\mathbf{D}_r} \succeq \mathbf{0} \\ & \text{for } i = 1, ..., K;\ r = 1, ..., R.\end{aligned} \tag{4.31}$$

The sum rate maximization problem of (4.31) is non-convex, and hence challenging. However, for a fixed power-allocation matrix $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$ satisfying the power constraint at

the RSs, the objective function, $\frac{1}{2}f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}})$ defined in (4.25) is concave with respect to $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$. This is because $\log|\cdot|$ is concave on the convex set of PSD matrices [142]. Thus, for a given $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$, the optimization problem with respect to $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ is convex. On the contrary, for a fixed $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ satisfying the power constraint at the MSs, the objective function of (4.31) is a difference of concave (DC) functions in $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$. Unfortunately, we cannot claim concavity of this function.

Applying the concept of directional derivative and Taylor expansion on the second term of (4.25), for the tractability of the problem, we seek to find a concave lower bound (on the system sum rate) with respect to $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$ as outlined in the following.

Let $\mathbf{Y_0} = \mathbf{I} + \mathbf{A_H}\bar{\mathbf{\Lambda}}_{\mathbf{0}}$, where $\mathbf{A_H} = \bar{\mathbf{\Sigma}}_{\mathbf{H}}^{T}\bar{\mathbf{U}}_{\mathbf{H}}^{H}\bar{\mathbf{U}}_{\mathbf{H}}\bar{\mathbf{\Sigma}}_{\mathbf{H}}$. Applying the first-order Taylor expansion of the log-determinant function (about the operating point $\bar{\mathbf{\Lambda}}_0$) to the second term of (4.25), we obtain a lower bound on the network sum rate as

$$\begin{aligned} f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}) &\gtrsim \log_2\left|\mathbf{I} + \mathbf{A_H}\left(\bar{\mathbf{\Lambda}}\bar{\mathbf{\Lambda}}_{\mathbf{F}} + \mathbf{I}\right)\bar{\mathbf{\Lambda}}_{\mathbf{D}}\right| - \log_2|\mathbf{Y_0}| - Tr\{\mathbf{Y_0}^{-1}\mathbf{A_H}(\bar{\mathbf{\Lambda}}_{\mathbf{D}} - \bar{\mathbf{\Lambda}}_{\mathbf{0}})\} \\ &\triangleq \tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}). \end{aligned} \tag{4.32}$$

We have used $\nabla_{\mathbf{X}}\log_2|\mathbf{AX}+\mathbf{B}| = [(\mathbf{AX}+\mathbf{B})^{-1}\mathbf{A}]^T$, the gradient of a log-determinant function [144] to obtain (4.32). From the concavity of the log-determinant function [145], the inequality in (4.32) holds since $\log|\mathbf{I}+\mathbf{A}_0| + Tr(\mathbf{I}+\mathbf{A}_0)^{-1}(\mathbf{A}-\mathbf{A}_0)$ is a global over-estimator of $\log|\mathbf{I}+\mathbf{A}|$ for a fixed $\mathbf{A}_0 \succeq \mathbf{0}$. Note that the lower bound in (4.32) is tight when $\bar{\mathbf{\Lambda}}_{\mathbf{0}}$ is close to $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$, with equality (i.e., $\tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}) = f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}})$) when $\bar{\mathbf{\Lambda}}_{\mathbf{0}} = \bar{\mathbf{\Lambda}}_{\mathbf{D}}$.

Based on the sum rate lower bound (4.32), we then define a new optimization problem as

$$\begin{aligned} \max_{\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}} \quad & \frac{1}{2}\tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}) \\ \text{s.t.} \quad & Tr(\mathbf{\Lambda}_{\mathbf{F}_{u,i}}) \leq P_{U,i}\ ; \mathbf{\Lambda}_{\mathbf{F}_{u,i}} \succeq \mathbf{0} \\ \text{and} \ & Tr\{(\mathbf{\Lambda}_r\mathbf{\Lambda}_{\mathbf{F}_r} + \mathbf{I})\,\mathbf{\Lambda}_{\mathbf{D}_r}\} \leq P_{R,r}\ ; \mathbf{\Lambda}_{\mathbf{D}_r} \succeq \mathbf{0} \\ \text{for} \ \ & i = 1, ..., K;\ r = 1, ..., R. \end{aligned} \tag{4.33}$$

While (4.33) is still non-convex in $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ and $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$ jointly, it is convex in $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ and $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$ individually unlike (4.31). Hence, we propose Algorithm 2 for solving the optimization problem of (4.33). The proposed algorithm is iterative in nature and follows alternating projection procedure [137, 139] over the MSs' and the RSs' power-allocation matrices, $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ and $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$, respectively. For a given iteration, the algorithm maximizes the sum rate lower bound (4.32) with respect to either $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ or $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$, while the other is fixed. In the next iteration, the optimization is carried out over the fixed variable in

the last iteration, while the optimized variable in the last iteration is now fixed. This process is repeated until convergence. This alternating procedure is primarily aimed at making $\bar{\mathbf{\Lambda}}_{\mathbf{0}}$ as close as possible to $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$, at which point the lower bound in (4.32) is tight. Recall that $\tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}) = f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}})$ when $\bar{\mathbf{\Lambda}}_{\mathbf{0}} = \bar{\mathbf{\Lambda}}_{\mathbf{D}}$.

Due to the approximation in (4.32), the solution of (4.33) is sub-optimal to the optimization in (4.31), where $f(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}})$ is the objective function. Hence, the sum rate achieved by the obtained solution is a lower bound on the actual system sum rate. Since the sum rate lower bound objective function of (4.33) is non-negative and concave in $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ and $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$, hence monotonically non-decreasing (i.e., it increases or stays the same after each iteration) and terminates when the sum-rate improvement is less than a pre-defined tolerance level $\mu > 0$ (Step 5), our proposed algorithm is convergent. However, we cannot claim convergence to the global optimal solution due to the non-convexity of the problem. As is common practice with non-convex optimization problems, our algorithm can be initialized with different initial values so that the value that results in the highest sum rate will be selected. Algorithm 1 requires global CSI. Thus the optimization can be conducted at one of the BSs, which are assumed to have global CSI. The BS then transmits the solutions of $\bar{\mathbf{\Lambda}}_{\mathbf{D}}$ and $\bar{\mathbf{\Lambda}}_{\mathbf{F}}$ to the RSs and MSs.

Algorithm 2 : Maximization of Sum Rate Lower Bound.

1. **Initialize**:
   $\bar{\mathbf{\Lambda}}_0 = \mathbf{0}_{\bar{N}_i \times \bar{N}_i}$, $\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(0)} \triangleq diag\,\{\mathbf{\Lambda}_{\mathbf{F}u,1}^{(0)}\ \mathbf{\Lambda}_{\mathbf{F}u,2}^{(0)}\ ...\ \mathbf{\Lambda}_{\mathbf{F}u,K}^{(0)}\}$.
2. Set iteration count to k.
3. Obtain $\bar{\mathbf{\Lambda}}_{\mathbf{D}}^{(k)} = \arg\max_{\bar{\mathbf{\Lambda}}_{\mathbf{D}}} \tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(k-1)}, \bar{\mathbf{\Lambda}}_{\mathbf{D}})$, using standard convex optimization routines.
4. Obtain $\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(k)} = \arg\max_{\bar{\mathbf{\Lambda}}_{\mathbf{F}}} \tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}^{(k)})$, using standard convex optimization routines.
5. **if** $\tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(k)}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}^{(k)}) - \tilde{f}(\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(k-1)}, \bar{\mathbf{\Lambda}}_{\mathbf{D}}^{(k)}) < \mu$, **break** ;
6. **else**, $\bar{\mathbf{\Lambda}}_0 = \bar{\mathbf{\Lambda}}_{\mathbf{D}}^{(k)}$, $k = k+1$ and return to Step 2.
7. Compute $\mathbf{D}_i^* = \mathbf{V}_{\bar{\mathbf{H}}_i} \mathbf{\Lambda}_{\mathbf{D}_i}^{1/2} \mathbf{U}_i^H \check{\mathbf{Q}}_i$ from (4.22).

4.4 User-Relay Association

In the emerging and future broadband relay-equipped broadband MIMO cellular networks, it is envisaged that there will be several RSs and MSs within a cell (or cluster). Assigning subsets of MSs to different RSs for service would improve the system performance (e.g., throughput and coverage). One key challenge with such system would be the design of practical and efficient MS-RS association schemes. With respect to the proposed joint design of Section 4.3, such MS-RS association scheme decides which MSs' transmissions will be forwarded (useful signal transmissions) or muted (interference transmissions) by which RS.

An MS-RS association strategy can be either "centralized" or "distributed" in nature. In a centralized scheme, the BS makes the association decision based on the received global CSI from all the MSs and RSs, whereas in a distributed association scheme, each RS selects the MSs within its coverage area based on the local CSI from the MSs. Naturally, a centralized association scheme should outperform its distributed counterpart, albeit at the cost of higher signaling overhead and complexity.

Association of MSs to the RSs can also be categorized as "static" or "dynamic". In the static case, the number of MSs served by a given RS is fixed for different transmission intervals, while in the dynamic association, the number of MSs served by an RS may vary from one transmission interval to another. In general, a dynamic association scheme will likely outperform a static one, since there is higher possibility that each MS will be served by (i.e., associated with) its "best" RS. However, dynamic schemes pose challenges in design and practicability. On the other hand, static association schemes are usually simpler, but may suffer some performance degradation.

We consider a low complexity and centralized static association scheme, where each RS serves at least one of the K scheduled MSs, while each MS is served by only one of the R RSs within the cluster. Recall that $\mathcal{S}_i$ denotes the set of $K_i \leq K$ MSs served by the ith RS while $\mathcal{S}_i^{\perp}$ denotes the set of $\tilde{K}_i = K - K_i$ MSs not served by the ith RS (that is, served by other RSs). Thus, $K = \sum_{i=1}^{R} K_i$ with $K_i \geq 1$.

For throughput maximizing association, the CoMP-BS needs to calculate the achievable system sum rate for all possible MS-RS associations, and then broadcast the association that results in the highest sum rate to the RSs. Unfortunately, the sum rate maximization problem for the multi-cell MARC with multiple AF RSs under consideration is a very complex (non-convex, multi-variable, and high dimensional matrix) optimization problem as shown earlier. Also, the complexity of the sum rate optimal association solution based on exhaustive search grows exponentially with network size. Moreover, note that the MS-RS association ought to be done before the

actual signal transmissions, thereby making it extremely difficult if not impractical to implement in the already delay-sensitive relay-equipped communication systems.

For a low complexity solution, we deploy an MS-RS association framework based on utility matrix $\mathbf{U}$ and association matrix $\bar{\mathbf{I}}$. The central node (in this case CoMP-BS) generates a K by R binary association matrix $\bar{\mathbf{I}}$ (with only 0s and 1s) and broadcasts it to the RSs. Each RS then checks for the MSs it's assigned to (that is, the indices with 1s in the RS's column). The binary matrix $\bar{\mathbf{I}}$ is obtained by manipulating a K by R utility matrix $\mathbf{U}$, with rows and columns corresponding to the MSs and RSs IDs respectively. $U_{[ij]} \geq 0$ is a pre-defined system utility achieved when the ith MS is served by the jth RS over the two-hop channels. In other words, the CoMP-BS manipulates $\mathbf{U}$ to obtain $\bar{\mathbf{U}}$ with only one non-zero element in each row and at least one non-zero element (i.e., $K_i \geq 1$) in each column, such that the sum of all elements of $\bar{\mathbf{U}}$ is maximum.

Analytically, the outlined MS-RS association strategy can be modeled as the following optimization problem:

$$\begin{aligned} &\max \sum_{i=1}^{K}\sum_{j=1}^{R} \quad \bar{U}_{[ij]} \\ &where \quad \bar{\mathbf{U}} = \mathbf{U} \cdot \bar{\mathbf{I}} \\ &s.t. \quad \sum_{i=1}^{K} \bar{I}_{[ij]} = K_j \\ &and \quad \sum_{j=1}^{R} \bar{I}_{[ij]} = 1 \end{aligned} \tag{4.34}$$

where $\mathbf{U}\cdot\bar{\mathbf{I}}$ is the element-wise multiplication of the utility matrix $\mathbf{U}$ and the association matrix $\bar{\mathbf{I}}$. The optimization problem of (4.34) is an NP-hard integer programming problem, hence very difficult to solve. As a result, we propose a low complexity greedy-based algorithm to solve (4.34), as outlined in Algorithm 3.

Next, we consider the design of the utility matrix $\mathbf{U}$. In addition to the earlier mentioned challenges, achievable rate is not appropriate for our formulation in (4.34) since the change of one MS-RS association will affect the achievable rate of all MSs. Hence, we propose a CSI-only based utility matrix $\mathbf{U}$ which depends on the instantaneous channel conditions, herein called the "channel-aware" MS-RS association scheme.

Specifically, we choose

$$U_{[ij]} = \frac{||\mathbf{G}_{i,j}||_F}{\min(\mathbf{N}_i, \mathbf{M}_{R,j}, \mathbf{M}_B)}, \tag{4.35}$$

where $\|\mathbf{A}\|_F$ is the Frobenius norm of matrix $\mathbf{A}$.

In this design, we consider only the MS-RS channels instead of the end-to-end channel comprising the MS-RS channels, RS beamforming matrix and the RS-BS channels. This is motivated by the fact that RS-to-CoMP-BS channel for each RS is the same irrespective of the MSs served by the RS.

Algorithm 3 : A Centralized "Channel-Aware" MS-RS Association Scheme.

1. Given $\mathbf{U}$, determine the number of MSs to be served by each RS (i.e., K_j for the jth RS)

   Initialize:

   Set the number of MSs served by each RS to zero (i.e., $K_i^{(0)} = 0, \forall i = 1, 2, ..., R$)

   Set iteration count to k.

2. Find the MS-RS association with the best channel quality (i.e., the row and column indices of $\mathbf{U}$ with the largest entry).

3. Check if the given RS has been assigned up to its maximum number of MSs to be served.

   If "no",
   ✓assign the MS to the RS, then
   ✓increase the number of MSs assigned to the RS by 1 (i.e., $K_i^{(k)} = K_i^{(k-1)} + 1$),
   ✓remove the MS from the set of MSs yet to be assigned to an RS (i.e., set the corresponding row of $\mathbf{U}$ to zeros).

   elseif "yes",
   ✓remove the RS from the set of RSs to be assigned MSs (i.e., set the corresponding column of $\mathbf{U}$ to zeros).

   end "if",

4. Increase the iteration count by one (i.e., k=k+1).

5. Repeat Steps 2 to 4 until all the K scheduled MSs have been assigned (i.e., sum of elements of $\bar{\mathbf{I}}$ equals K).

So far, we have assumed that only K cell-edge MSs are available and requesting service within the cooperating cluster. However, in LTE-Advanced networks with a large number of users, it's most likely to have more cell-edge MSs, say K_t within the cluster than can be served at the same time. Selecting and associating K out of the K_t MSs ($K \leq K_t$) will further enhance the system performance as shown in our simulation results. Several factors such as the number of antennas at the nodes (MSs, RSs, and BSs), constraint on system parameters in order to achieve certain design

goals (e.g., $M_{R,i} \geq N_u$ to decouple RSs' channels as in this work), etc., determine the actual value of K.

4.5 Discussions on CSI Acquisition and Overhead for the Joint MS-RS Design and Association

While CSI acquisition is in general beyond the scope of this work, we would like to mention a few things about CSI acquisition peculiar to this work. One of the major challenges faced by practical wireless relay networks is the enormous feedback overhead and complexity encountered during CSI acquisition.

Channel training and estimation in MIMO AF relay-equipped wireless networks for the receiver to obtain both the transmitter-relay and the relay-receiver CSI (similar to the CSI requirement at the CoMP-BS for MS-RS association in this work) have been investigated in the literature. While the relay-receiver CSI is readily obtained by pilot training (from the relay) using standard MIMO training schemes, obtaining the transmitter-relay CSI at the receiver is usually more complicated [149–151].

Both the joint MS covariance and RS beamforming design as well as the MS-RS association scheme proposed in this chapter require CSI at the RS and CoMP-BS. Whereas each RS needs to know only its local relay and access links' CSI, the CoMP-BS requires CSI of both the access and relay links. In what follows, we present a simple joint CSI acquisition procedure (for both the RS beamforming and MS-RS association), which leverages the inherent feature of cellular networks.

Similarly to the hand-off mechanism in cellular networks, each MS measures and reports its channel conditions to the serving RS. Each RS in turn reports its channel conditions (for the channel between the RS and its serving BS, as well as for the channel between itself and each MS) to the serving BS. The channel quality indicator (CQI) is then shared among the BSs within the cluster (CoMP-BS). This is done periodically in order to track the channel variations within the cluster.

For the MS-RS association, the CoMP-BS generates a K by R binary matrix $\bar{\mathbf{I}}$ and broadcasts it to the RSs. Each RS then checks for the MSs it is assigned to (that is, the indices with 1s in the RS's column). Each RS uses the already-available local and access link CSI to design its beamforming matrix for its assigned MSs. Thus, a one-time channel training is sufficient for both the MS-RS association and beamforming design before the actual data transmission, thereby reducing the feedback overhead and processing time.

4.6 Simulation Results

In this section, we present the simulation results. The proposed schemes are employed for a coordinated multi-point cellular uplink with focus on cell-edge user experience.

4.6.1 Network Setup and Channel Model

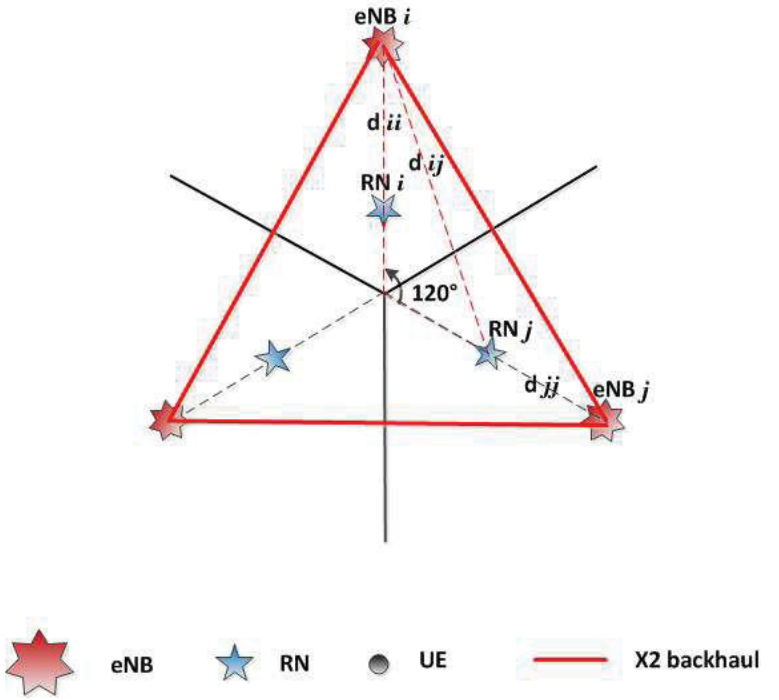

Figure 4.3: Cell geometry for path-loss evaluation.

A 3-cell hexagonal network with sectorization is considered as depicted in Figures 4.2 and 4.3. Each cell comprises 1 BS with $M_{B,i}$ antennas per sector and 6 RSs (one per sector) with $M_{R,i}$ antennas at each RS. We consider the triangular area comprising 3 contiguous sectors at the center of the cooperating cells (a.k.a., the sector of interest). That is, sectors of three adjoining cells are coordinated to form a cluster. We neglect the inter-cluster and/or inter-sector interference from outside the sector of interest. We assume that each RS is scheduled to serve equal number of MSs out of the K scheduled MSs within the coordinated cluster (i.e., $K_i = K_j \triangleq \frac{K}{R}, i \neq j$). In the following, we first provide a brief overview of our channel model, before presenting the simulation results.

We account for both Rayleigh fading and path-loss in our channel models (unless otherwise stated). For the path-loss, the 3GPP's LTE-Advanced channel model proposed for the evaluation of relay-equipped heterogeneous deployments with baseline

parameters given in Table A.2.1.1.2-3 of [152] is used. An outdoor urban deployment with fixed infrastructure at carrier frequency of 2 GHz and cell radius of 1 km is considered.

We model the MIMO channel from the ith MS to the rth RS as $\mathbf{G}_{i,r} \triangleq \sqrt{\frac{\beta_{i,r}}{N_{01}}}\bar{\mathbf{G}}_{i,r}$, where $\bar{\mathbf{G}}_{i,r} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Similarly, the MIMO channel from the rth RS to the jth BS is modeled as $\mathbf{H}_{r,j} \triangleq \sqrt{\frac{\alpha_{r,j}}{N_{02}}}\bar{\mathbf{H}}_{r,j}$, where $\bar{\mathbf{H}}_{r,j} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. N_{01} and N_{02} are the noise powers at the RSs and BSs respectively, while $\beta_{i,r}$ and $\alpha_{r,j}$ are the distance and carrier frequency dependent power gains (inverses of the path-losses) of the respective channels. $\bar{\mathbf{G}}_{i,r}$ and $\bar{\mathbf{H}}_{r,j}$ capture the small-scale Rayleigh fading. Shadowing is not accounted for. The normalization of the channels by the received noise powers follows from effective channel relation for MIMO systems [29,49]. Owing to the cell-edge user experience focus of this work, all the K scheduled MSs are assumed to be located at the cell edge. That is, $\beta_{i,r} = \beta_{j,r}, \forall i \neq j$. Thus, with each BS located at the center of its cell, the $\beta_{i,r}$ and $\alpha_{r,j}$ values depend on the position of the RS within the cell.

Figure 4.3 depicts the cell-geometry used to determine the path-loss values. d_{ii} [km] denotes the distance between the RS and the BS both in the ith cell while d_{ij} [km] denotes the distance between the BS in the ith cell and the RS in the jth cell. We assume that each RS is located along the straight line joining the BS and MSs as depicted in Figure 4.2, with the distance of each RS to its serving BS (i.e., the BS in the same cell as the RS) the same for all RSs during each 2-hop transmission interval, (i.e., $d_{ii} = d_{jj}, \forall i \neq j$). Equations (4.36), (4.37), and (4.38) provide the expression for obtaining the channels power gains [152].

$$-\alpha_{ii} \triangleq 100.7 + 23.5 \ \log_{10} \ d_{ii} \tag{4.36}$$

$$-\alpha_{ij} \triangleq 100.7 + 23.5 \ \log_{10} \ d_{ij} \tag{4.37}$$

$$-\beta_{ii} = -\beta_{ij} \triangleq 145.4 + 37.5 \ \log_{10} \ (1 - d_{ii}) \tag{4.38}$$

Table I gives the system parameters used in our simulations (unless otherwise stated).

4.6.2 Numerical Results

In this sub-section, we study the system behaviour with respect to the proposed joint MS transmit covariances and RS beamforming design as well as compare the proposed "channel-aware" MS-RS association scheme with a naive "random" association

Table 4.1: System parameters used in simulation

RS transmit power	$P_{R,r}$
MS transmit power	$P_{U,i}$
RS noise power (N_{01})	-109 dBm
BSnoise power (N_{02})	-109 dBm
Cell radius	1 km
Sectors per cell	6
RS per cell	6
RS per sector	1
Carrier frequency	2 GHz
Environment	Outdoor
Relaying protocol	amplify-and-forward
RS type	in-band

scheme and the optimal exhaustive search association strategy. In the random case, we assume that the first K_i MSs are served by the 1st RS, the next K_i MSs served by the 2nd RS, and so on. Using the "random" association strategy (rather than one based on signal strengths) as a reference is reasonable, since all MSs experience the same path losses to their associated RSs (see Fig. 4.2). For the exhaustive search scheme, we consider all possible MS-RS associations to obtain the association that yields the highest sum rate. Thus, the exhaustive search has very high computational complexity especially for networks with large number of MSs and RSs. Furthermore, we assume the same transmit powers and numbers of antennas at similar nodes. That is, $M_{B,i} = M_B$ antennas at each BS, $M_{R,i} = M_R$ antennas at each RS, and $N_i = N$ antennas at each of the K MSs and $P_{U,i} = P_{U,j} = P_U, i \neq j$ and $P_{R,r} = P_{R,l} = P_R, r \neq l$.

The average sum rate (averaged over 500 channel realizations) with iteration-refinement level $\mu = 10^{-5}$ as specified in *Algorithm* 2 is shown in all simulations, unless otherwise stated. As is common practice with non-convex optimization problems, our algorithm can be initialized with different initial values and the value that results in the highest sum rate is selected. However, in our simulations, fixed initial value assuming equal power allocation at the MSs is used. That is, $\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(0)} = \frac{P_U}{N}\mathbf{I}$ in *Algorithm* 2. We used `CVX`, a package for specifying and solving convex programs [146], in all our simulations.

First, we consider networks in which only $K = 3$ scheduled MSs within the coordinated cluster are available for service in a Rayleigh fading environment. In partic-

ular, Figure 4.4 shows the average network sum rate versus the RS transmit power at fixed MS transmit power $P_U = 0$ dB. For the same system configuration, our proposed *channel-aware* scheme achieves a reasonable percentage of the performance of the optimal *exhaustive search* MS-RS association strategy. Also, the proposed "channel-aware" MS-RS association scheme outperforms the "random" scheme with larger performance gap at lower antenna configurations. This could be attributed to "*channel hardening*", a reduction in channel variations or diversity due to multi-antennas at MIMO nodes. Furthermore, for a fixed RS transmit power, the network sum rate increases with the number of antennas at each node. Moreover, at high user transit power $P_U = 20$ dB, similar performance trends (for low user transmit power $P_U = 0$ dB) between the schemes are observed, albeit with higher multiplexing gain (indicated by the slope of the sum-rate curves) as depicted in Figure 4.5.

Similarly, Figure 4.6 depicts the average network sum rate versus the MS transmit power at fixed RS transmit power $P_R = 0$ dB. For the same system configuration, the proposed "channel-aware" scheme outperforms the "random" scheme with larger performance gap at low antenna configurations. Also, the proposed *channel-aware* MS-RS association scheme performs close to the optimal exhaustive search scheme. Furthermore, for a fixed MS transmit power, the network sum rate increases with the number of antennas at each node. Moreover, at high relay transmit power $P_R = 20$ dB, similar performance trends (with low relay transmit power) are observed as shown in Figure 4.7. There is higher multiplexing gain (indicated by the slope of the sum-rate curves (Figs. 4.6 and 4.7) with increasing MS transmit power compared to the RS transmit power (Figs. 4.4 and 4.5).

Next, we consider networks in which $K = 3$ out of K_t subset of MSs within the coordinated cluster are associated with $R = 3$ RSs for service in a Rayleigh fading environment. Owing to the unacceptably high complexity involved, the exhaustive search scheme is not considered in this case. Comparing the proposed "channel-aware" and "random" MS-RS association schemes, all the earlier observations for networks with $K = K_t = 3$ MSs and $R = R_t = 3$ RSs still hold. However, there is a wider performance gap between the proposed "channel-aware" scheme and the "random" scheme for the same transmit power and system configurations (e.g., Fig. 4.8 vs Fig. 4.4, Fig. 4.11 vs Fig. 4.7, etc). Hence, the proposed user-relay association scheme is more effective in networks with large number of users

Finally, Figure 4.12 depicts the average network sum rate versus the RS-BS distance for a multi-cell MIMO AF MARC in a large-scale fading environment. The MS and RS powers are fixed at $P_U = -6$ dB and $P_R = 0$ dB respectively. Furthermore, $K = 3$ out of K_t MSs are associated with the $R = 3$ RSs. As can be seen, the network

sum rate increases as the RSs are located closer to the MSs (i.e., the cell-edge). Also, the channel-aware MS-RS association scheme again outperforms the random scheme. There is wider performance gap with larger user pool (i.e., larger K_t).

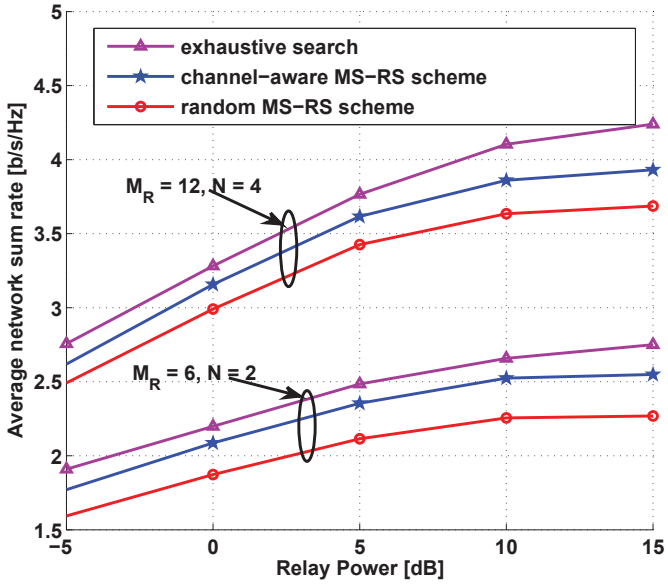

Figure 4.4: Average network sum rate versus relay transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment at fixed MS transmit power $P_U = 0$ dB.

4.7 Conclusion

In this chapter, we have studied a multi-cell MIMO MARC, in which R half-duplex fixed infrastructure-based in-band AF multi-antenna RSs are used to extend radio coverage between K cell-edge MSs and C cooperative BSs. Unlike various schemes considered in the literature in which some (or all) of the nodes are equipped with a single antenna, a more general system setup with multiple antennas at all nodes has been studied. With cooperation among the BSs and distributed linear processing at the RSs, we jointly design the MS transmit covariance and the RS beamforming matrices to maximize the system sum rate. Unlike for the conventional MAC without relays, the investigated multi-cell MARC sum-rate maximization is a complex high-dimensional non-convex optimization problem. Furthermore, the RS-BS channels are cross-coupled due to the amplified and forwarded interference by the AF RSs. In order to decouple the RSs' channels as well as eliminate the interference experienced by

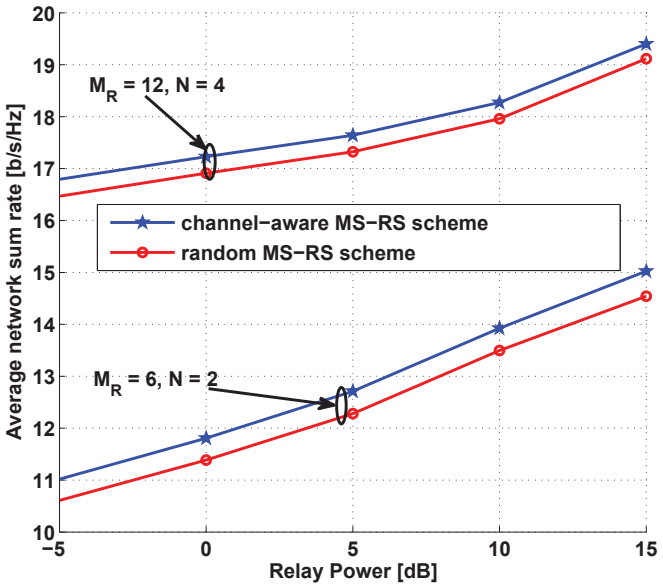

Figure 4.5: Average network sum rate versus relay transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment at fixed MS transmit power $P_U = 20$ dB.

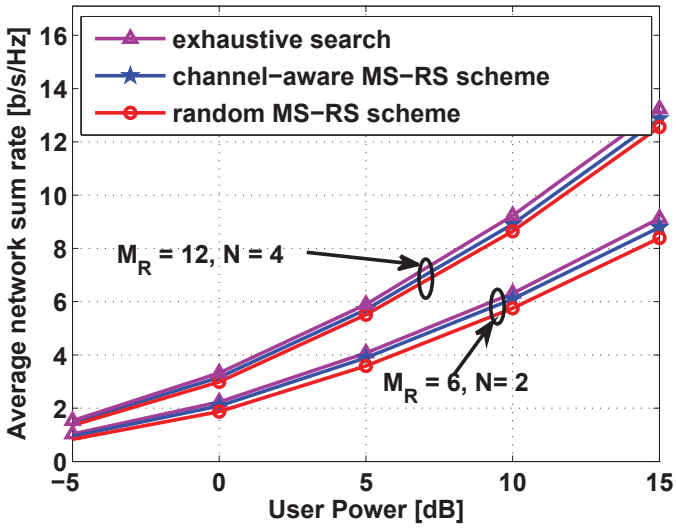

Figure 4.6: Average network sum rate versus user transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment at fixed RS transmit power $P_R = 0$ dB.

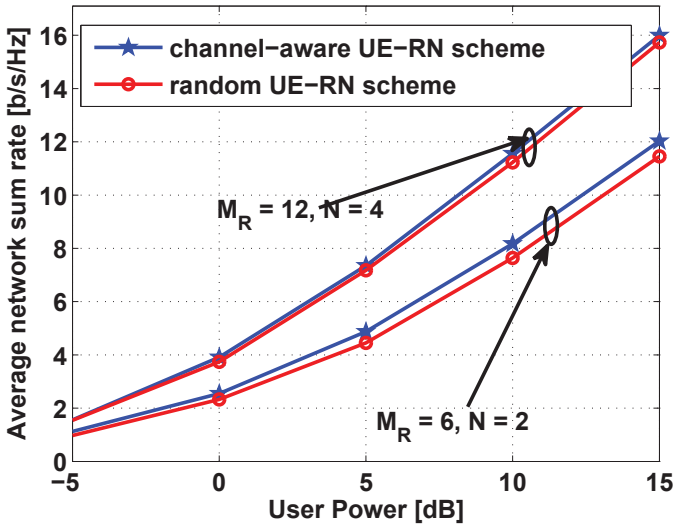

Figure 4.7: Average network sum rate versus user transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment at fixed RS transmit power $P_R = 20$ dB.

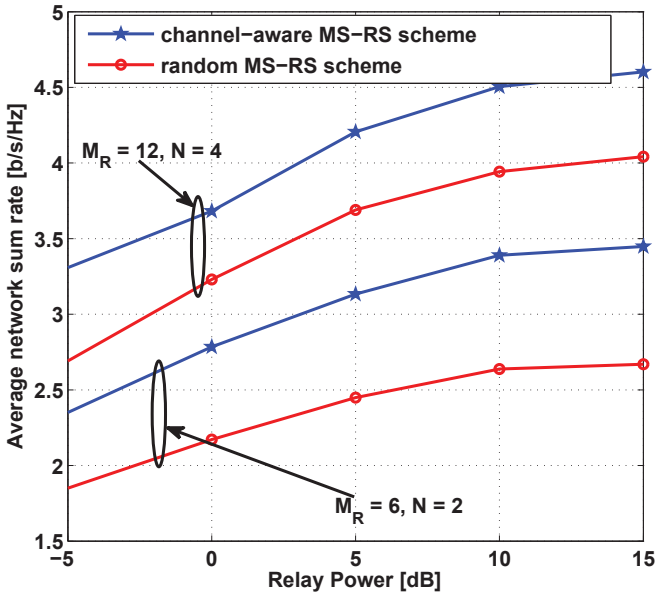

Figure 4.8: Average network sum rate versus relay transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment at fixed MS transmit power $P_U = 0$ dB.

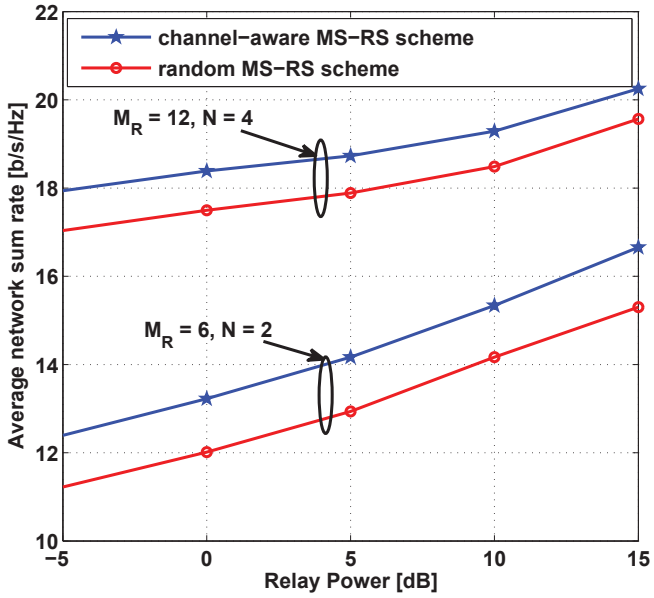

Figure 4.9: Average network sum rate versus relay transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment at fixed MS transmit power $P_U = 20$ dB.

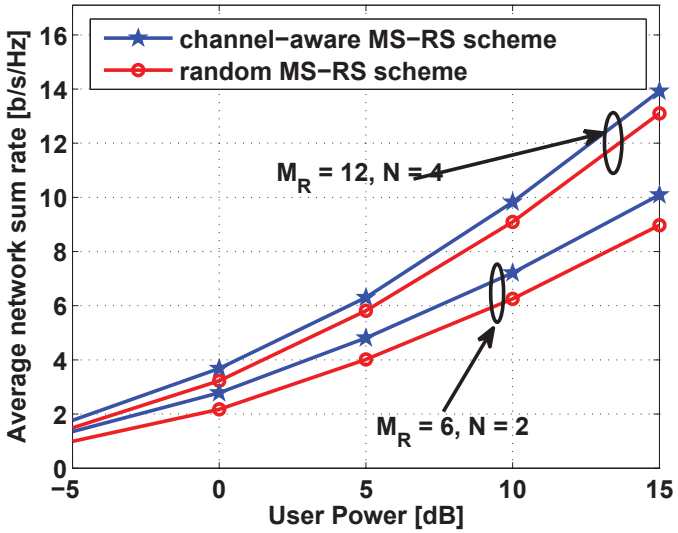

Figure 4.10: Average network sum rate versus user transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment with $K = 3$ out of $K_t = 100$ MSs associated with $R = 3$ RSs for service at fixed RS transmit power $P_R = 0$ dB.

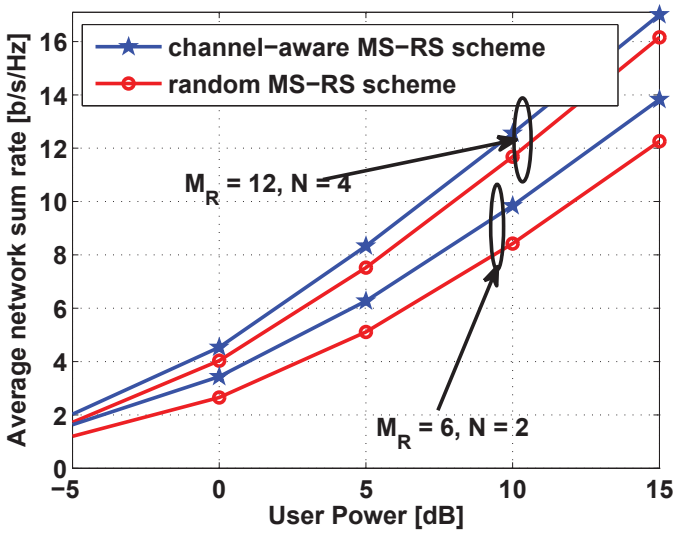

Figure 4.11: Average network sum rate versus user transmit power for a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, M_R antennas at each RS, and N antennas at each MS in a Rayleigh fading environment with $K = 3$ out of $K_t = 100$ MSs associated with $R = 3$ RSs for service at fixed RS transmit power $P_R = 20$ dB.

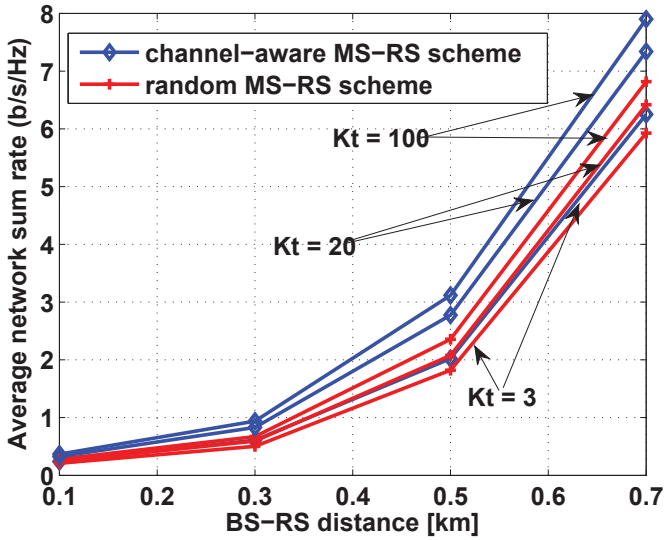

Figure 4.12: Average network sum rate of a 3-cell, 3-relay and 3-user multiple access relay network with $M_B = 2$ antennas at each BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS for fixed MS and RS powers ($P_U = -6$ dB, $P_R = 0$ dB) and different RS-BS distances in a large-scale fading environment.

each RS from its non-served MSs, we have incorporated interference pre-cancelation into the RS beamforming designs. Furthermore, we have matched each RS's beamforming matrix to the right and left singular vectors of the relay and access link matrices of the RSs such that the end-to-end channel has been diagonalized. With the beamforming matrix structure, an iterative alternating-minimization algorithm has been proposed. Moreover, we have proposed a low complexity CSI-based MS-RS association scheme for assigning the MSs to their serving RSs. Simulation results demonstrate the effectiveness of the proposed schemes for various system configurations and channel conditions, especially in relay-equipped broadband cellular networks with a large number of users.

Chapter 5

Joint User-Relay Selection and Association in Multi-User MIMO Relay Networks

5.1 Introduction

In Chapter 4, we investigated user-relay beamforming as well as user-relay association in multi-cell multi-relay AF relay uplink with multi-antenna nodes. That is, subsets of available users and relays are blindly (i.e., randomly) selected for cooperation, while the selected users are strategically assigned to the selected relays in order to improve the system performance.

However, in practical multi-user multi-relay systems, selecting subsets of "advantageous" relays and users (e.g., with the best channels) for cooperation and communication will further enhance the system performance by leveraging the inherent cooperative and multi-user diversity in the network. Also, such approach will reduce interference, signaling overhead and design complexity. In order to have any practical relevance, user-relay selection and/or association schemes must have acceptable complexity and fit within a small fraction of the coherence time of the channel.

Relay selection and user scheduling have been separately and extensively reported in the literature. For example, relay selection for networks with the same type of relays (e.g., AF) was studied in [119–130], while [131] considered networks with different types of relays (e.g., AF and DF). Similarly, user-scheduling in MIMO networks were investigated in [58, 65–69]. Recently, 2-step *single-user single-relay* selection schemes in which the user with the best channel is selected in the first step while the relay with the best channel to the selected user is selected in the 2nd step, for networks with single-antenna nodes have been studied [127, 128, 131]. That is, a single user with its best relay is selected for cooperation. However, transceiver nodes are envisaged to

have multiple antennas in the emerging wireless broadbands [8]. To the best of our knowledge, there is no work till date on the joint selection and association of multi-antenna relays and users in order to enhance the network performance by leveraging the inherent multi-user and multi-relay diversities in the network.

In this chapter, we investigate user-relay selection/scheduling and association in multi-user multi-relay wireless networks with multi-antenna transmitters and receivers. In particular, we propose a novel low complexity joint scheme, which simultaneously selects multiple relays and users for cooperation as well as assigns different selected users to different selected relays for service. That is, three different tasks (i.e., relay selection, user selection/scheduling, user-relay selection and association) are performed concurrently. The proposed scheme utilizes only the Frobenius norms of the MIMO channel matrices between the nodes (thus, partial instead of full CSI is needed), which is more practical and easier to implement. Furthermore, the complexity of the scheme scales linearly with the product of the total number of relays, the total number of users and the number of selected users. Compared with (i) a scheme with neither user-relay selection nor user-relay association (i.e., both the users and relays are selected randomly) (ii) a scheme with user-relay association but no user-relay selection (i.e., both the users and relays are selected randomly while the users are assigned to the relays based on their channel gains) (iii) two schemes each with 2-step user-relay selection and association (where the relays with best channels are selected in the first step while the users with best channels to the selected relays are then selected and associated to the relays in the second step), the proposed joint scheme offers superior performance. The favorable performance and low complexity of the proposed scheme make it very attractive for possible implementation in emerging and future broadband cellular networks

The remaining of this chapter is organized as follows. In Section 5.2, the system model is presented while Section 5.3 details the proposed low complexity joint user-relay selection and association scheme. Simulation results are presented in Section 5.4 and Section 5.5 concludes the chapter.

5.2 System Model

We consider a single-cell MARC with non-regenerative and altruistic (not having their own data to transmit) multi-antenna AF relays. It comprises a BS, also know as evolved node B (eNB), serving K scheduled MSs with the aid of R fixed in-band half-duplex AF RSs as depicted in Figure 5.1. That is, there is no direct link between the MSs and the BS. The $K \leq K_t$ MSs and $R \leq R_t$ RSs are selected from K_t MSs and

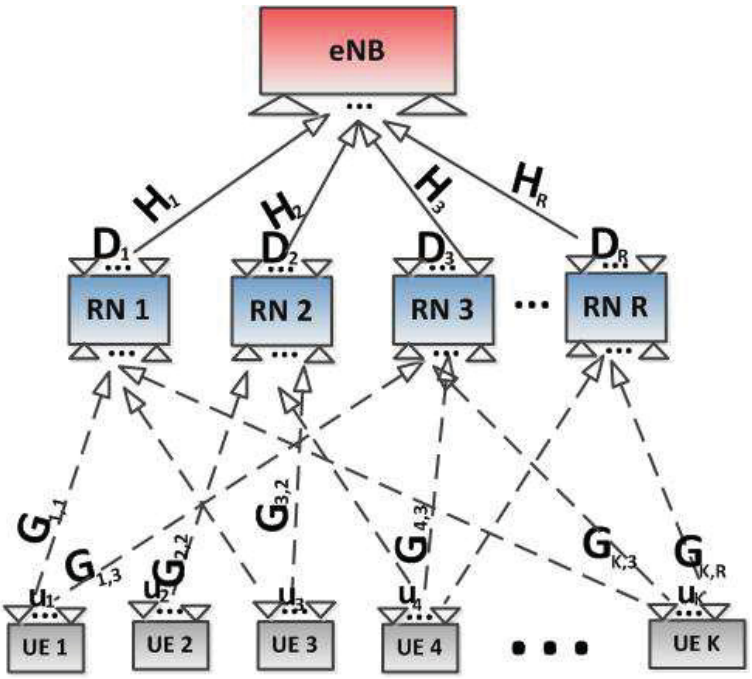

Figure 5.1: A single-cell multi-user multi-relay MIMO relay uplink (MARC).

R_t RSs, respectively, for data transmission. The BS is equipped with M_B antennas, while the rth RS and MS have $M_{R,r}$ and N_r antennas, respectively. The channel matrix from the ith MS to the rth RS is denoted by $\mathbf{G}_{i,r} \in \mathbb{C}^{M_{R,r} \times N_i}$ and that from the rth RS to the BS is $\mathbf{H}_r. \in \mathbb{C}^{M_B \times M_{R,r}}$.

With the half-duplex RSs, the transmission of a symbol block takes place in two hops. In the first hop, each MS transmits to its associated RS while causing interference at other RSs. The RSs simultaneously forward amplified versions of their received signals to the BS in the second hop.

Let $\mathcal{S}_r$ denote the set of $K_r \leq K$ selected MSs associated with (or served by) the rth RS and $\mathcal{S}_r^{\perp}$ the set of $\tilde{K}_r = K - K_r$ selected MSs not associated with the rth RS. Thus, $K = \sum_{r=1}^{R} K_r$. Also, $\mathbf{u}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_i)$ is the transmitted signal by the ith MS, with covariance matrix $\mathbf{Q}_i = \mathbb{E}\{\mathbf{u}_i \mathbf{u}_i^H\}$.

The received signal by the rth RS in the first hop is given by:

$$\begin{aligned} \mathbf{y}_{R,r} &= \sum_{i=1, i \in \mathcal{S}_r}^{K_r} \mathbf{G}_{i,r} \mathbf{u}_i + \sum_{j=1, j \in \mathcal{S}_r^{\perp}}^{\tilde{K}_r} \mathbf{G}_{j,r} \mathbf{u}_j + \mathbf{n}_{R,r} \\ &= \hat{\mathbf{G}}_r \hat{\mathbf{u}}_r + \tilde{\mathbf{G}}_r \tilde{\mathbf{u}}_r + \mathbf{n}_{R,r}, \end{aligned} \tag{5.1}$$

where $\mathbf{n}_{R,r}$ is the noise vector at the rth RS whose elements are i.i.d. $\mathcal{CN}(0,1)$. $\hat{\mathbf{G}}_r$

and $\tilde{\mathbf{G}}_r$ are the composite channel matrices of the desired and interference signals at the rth RS respectively. Similarly, $\hat{\mathbf{u}}_r$ and $\tilde{\mathbf{u}}_r$ are the corresponding transmitted signal vectors from the K_r desired MSs and $\tilde{K}_r$ interfering MSs to the rth RS. The rth RS amplifies its received signal with beamforming matrix $\mathbf{D}_r$ and then forwards $\mathbf{x}_{R,r} = \mathbf{D}_r \mathbf{y}_{R,r}$ to the BS in the second hop.

The received signal vector by the BS is given by

$$\mathbf{y} \triangleq \sum_{r=1}^{R} \mathbf{H}_r \mathbf{x}_{R,r} + \mathbf{n}_B, \tag{5.2}$$

where $\mathbf{n}_B$ is the noise vector at the BS, whose elements are i.i.d. $\mathcal{CN}(0,1)$.

We adopt the joint source and relay beamforming design of Chapter 4 and [31], where each AF RS beamforming matrix is designed to pre-cancel the interference from unintended sources while forwarding the useful signals to the destination, as briefly outlined in the sequel.

Denote the access link channel of the rth RS as

$$\acute{\mathbf{G}}_r \triangleq [\hat{\mathbf{G}}_r \ \tilde{\mathbf{G}}_r], \tag{5.3}$$

with pseudo-inverse

$$\grave{\mathbf{G}}_r = (\acute{\mathbf{G}}_r^H \acute{\mathbf{G}}_r)^{-1} \acute{\mathbf{G}}_r^H \triangleq \begin{bmatrix} \check{\mathbf{G}}_r \\ \dot{\mathbf{G}}_r \end{bmatrix}. \tag{5.4}$$

The QR decomposition of $\check{\mathbf{G}}_r$ is given by

$$\check{\mathbf{G}}_r = \check{\mathbf{R}}_r \check{\mathbf{Q}}_r, \tag{5.5}$$

where $\check{\mathbf{R}}_r$ is a lower triangular matrix, and $\check{\mathbf{Q}}_r$ is unitary. Taking the SVDs of the effective channels of the rth RS as

$$\check{\mathbf{Q}}_r \hat{\mathbf{G}}_r = \mathbf{U}_r \mathbf{\Lambda}_r \mathbf{V}_r^H \tag{5.6}$$

and

$$\mathbf{H}_r = \mathbf{U}_{\mathbf{H}_r} \mathbf{\Sigma}_{\mathbf{H}_r} \mathbf{V}_{\mathbf{H}_r}^H, \tag{5.7}$$

the rth RS's beamforming matrix is designed as

$$\mathbf{D}_r = \mathbf{V}_{\bar{\mathbf{H}}_r} \mathbf{\Lambda}_{\mathbf{D}_r}^{1/2} \mathbf{U}_r^H \check{\mathbf{Q}}_r, \tag{5.8}$$

where $\mathbf{V}_{\bar{\mathbf{H}}_r}$ contains the first $L_r \triangleq \min\{M_B, M_{R,r}, \bar{N}_r\}$ columns of $\mathbf{V}_{\mathbf{H}_r}$ while $\mathbf{\Lambda}_{\mathbf{D}_r}^{1/2}$ is the RS's power allocation matrix. Since $\grave{\mathbf{G}}_r \acute{\mathbf{G}}_r = \mathbf{I}$ and $\check{\mathbf{R}}_r$ is invertible, $\check{\mathbf{G}}_r \tilde{\mathbf{G}}_r \triangleq \check{\mathbf{R}}_r \check{\mathbf{Q}}_r \tilde{\mathbf{G}}_r = \mathbf{0}$, which implies that $\check{\mathbf{Q}}_r \tilde{\mathbf{G}}_r = \mathbf{0}$. Thus, $\mathbf{D}_r \tilde{\mathbf{G}}_r = \mathbf{0}$, that is, the RS beamforming pre-cancels the interference from its unassigned MSs before forwarding

to the destination. The input covariance matrix is given by $\hat{\mathbf{Q}}_r \triangleq \mathbb{E}\{\hat{\mathbf{u}}_r\hat{\mathbf{u}}_r^H\} = \mathbf{V}_r\mathbf{\Lambda}_{\mathbf{F}_r}\mathbf{V}_r^H$, where $\mathbf{\Lambda}_{\mathbf{F}_r} \triangleq diag\left(\mathbf{\Lambda}_{\mathbf{F}_{u,1}}, \mathbf{\Lambda}_{\mathbf{F}_{u,2}}, \ldots, \mathbf{\Lambda}_{\mathbf{F}_{u,K_r}}\right)$ is a diagonal matrix of the power allocation matrices of the K_r MSs assigned to the rth RS. $\mathbf{\Lambda}_{\mathbf{D}_r}$'s and $\mathbf{\Lambda}_{\mathbf{F}_r}$'s are optimized using alternating-maximization algorithm based on the sum rate lower bound.

5.3 Joint User-Relay Selection and Association

We consider a centralized user-relay selection and association scheme in which the BS makes the selection and association decisions and communicates them to all the nodes. We assume that R (out of R_t) RSs and K (out of K_t) MSs are to be selected with each selected RS serving at least one selected MS. Without loss of generality, we also assume that the maximum number of MSs each of the selected RSs can serve is the same (i.e., $K_1 = \cdots = K_R = K_{max}$), and each MS is associated with only one RS for service, which implies $R \leq K$.

Ideally, the selection and association scheme that maximizes the system throughput is desirable. However, to obtain such solution, the BS needs to calculate the achievable sum rate for all possible MS-RS selections and associations (via exhaustive search). Unfortunately, for each possible user-relay selection and association, the user and relay beamforming design is highly complex, as detailed in Chapter 4. Moreover, with exhaustive search the number of possible selections and associations grows fast with the network size, which means that the number of user and relay beamforming optimization problems that need to be solved grows fast. Since the MS-RS selection and association ought to be done before data transmission, the throughput-maximization selection and association (even without exhaustive search) is extremely difficult if not impractical to implement in the already delay-sensitive multi-hop communication networks.

As a result, we resort to a sub-optimal, but low complexity approach. We formulate an optimization problem based on a utility matrix $\mathbf{U}$ and association matrix $\bar{\mathbf{I}}$. The row index and column index of $\mathbf{U}$ and $\bar{\mathbf{I}}$ correspond to the MS and RS ID, respectively. $\mathbf{U}$ is a pre-defined utility matrix whose (i,r)th entry ($U_{[ir]} \geq 0$) is the system utility achieved, when the ith MS is served by the rth RS. The association matrix $\bar{\mathbf{I}}$ is a binary matrix (with only 0s and 1s), where an (i,r)th entry of 1 implies that the ith MS and the rth RS are selected and associated with each other. Thus, the number of 1s in the rth column gives the number of MSs associated with the rth RS while the number of 1s in the ith row gives the number of RSs the ith MS is associated with. Moreover, $\bar{\mathbf{U}} = \mathbf{U} \cdot \bar{\mathbf{I}}$ is element-wise multiplication of $\mathbf{U}$ and $\bar{\mathbf{I}}$.

Given $\mathbf{U}$, we aim to design $\bar{\mathbf{I}}$ with at most one 1 in each row and at most K_{max} 1s in each column, such that the sum of all entries of $\bar{\mathbf{U}}$ is maximized.

Let $\mathcal{U}$ and $\mathcal{U}^{\perp}$ denote the set of the K selected and $K_t - K$ non-selected MSs, while $\mathcal{R}$ and $\mathcal{R}^{\perp}$ denote the set of the R selected and $R_t - R$ non-selected RSs, respectively. The joint MS-RS selection and association strategy can be formulated as the following optimization problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^{K_t} \sum_{r=1}^{R_t} \bar{U}_{[ir]} \\ s.t. \quad & C1 : \sum_{i \in \mathcal{U}} \bar{I}_{[ir]} = K_r, \\ & C2 : \sum_{i \in \mathcal{U}^{\perp}} \bar{I}_{[ir]} = 0, \\ & C3 : \sum_{r \in \mathcal{R}} \bar{I}_{[ir]} = 1, \\ & C4 : \sum_{r \in \mathcal{R}^{\perp}} \bar{I}_{[ir]} = 0. \end{aligned} \tag{5.9}$$

Problem (5.9) is an NP-hard integer programming problem, hence difficult to solve. Consequently, we propose a low complexity greedy algorithm (*Algorithm* 4) to solve (5.9). Notice that the problem formulation in (5.9) and Algorithm 4 work for any utility matrix.

Next, we specify the design of the utility matrix $\mathbf{U}$, the choice of which greatly impacts the performance of the scheme. Achievable rate is not appropriate for our formulation in (5.9) since the change of one MS-RS association will affect the achievable rate of all MSs. Consequently, we propose a CSI-only based utility matrix $\mathbf{U}$, which depends on the Frobenius norms of the channel matrices. Thus, our proposed scheme requires partial instead of full CSI, hence more practical and easy to implement[1].

Recall that $\mathbf{H}_r$ and $\mathbf{G}_r \triangleq [\mathbf{G}_{1,r}\ \mathbf{G}_{2,r}\ ...\ \mathbf{G}_{K_t,r}]$ are, respectively, the aggregate relay and access links' MIMO channels for the rth RS, where $\mathbf{G}_{i,r}$ is the MIMO channel from the ith MS to the rth RS. Entries of $\mathbf{U}$ are chosen as

$$U_{[ir]} = \frac{||\mathbf{H}_r||_F\ ||\mathbf{G}_{i,r}||_F}{||\mathbf{H}_r||_F + ||\mathbf{G}_{i,r}||_F}. \tag{5.10}$$

With this utility matrix design, the RSs and MSs whose harmonic mean of the amplitudes of their relay and access links' complex channel gains are largest are selected

[1]We should clarify that while only the channel amplitude information is needed for the user-relay selection and association in the control signaling phase, full CSI (i.e., both amplitude and phase information) is needed for the beamforming/precoding designs during data transmission.

and associated with each other for cooperation. Thus, the utility values depend on both the relay and access links channels. Compared with utility designs depending on either the relay or access links' channels alone, the proposed utility is expected to yield better performance, since an MS may have a good channel to a given RS, but the RS's channel to the BS may be very bad (or vice versa), hence degrading the 2-hop channel for the MS if associated with the RS.

Next, we consider the worst case complexity of the proposed joint scheme in terms of the number of comparisons at Steps 1 and 2 of *Algorithm* 4. For each value of k in the "***for***" loop, the required number of worst-case comparisons in Step 1 is $K_t R_t$. Similarly, the number of worst-case comparisons in Step 2 is R_t. Hence, for the K iterations, a total of $K(K_t R_t + R_t)$ comparisons are needed. Thus, the worst case complexity of the proposed joint selection and association scheme is $\mathcal{O}(KK_t R_t)$. That is, it scales as the product of the number of selected MSs, total number of MSs, and total number of RSs.

Algorithm 4 : Joint MS-RS Selection and Association Scheme.

Given: $\mathbf{U}$, $R = |\mathcal{R}|$, $K = |\mathcal{U}|$, and $K_{\max}$.
Initialize: $R = 0$, $K_r^{(0)} = 0, \forall r = 1, 2, ..., R_t$, $\bar{\mathbf{I}}^{(0)} = \mathbf{0}$.
for *k=1:K*

1. Find the MS with the largest utility value to each RS (i.e., the row index with the largest entry in each column of $\mathbf{U}$), denoted as $U_{i_1,1}, U_{i_2,2}, ...U_{i_{R_t},R_t}$.

2. Find the MS-RS pair with the largest utility value at Step 1 (i.e., the column index with the largest entry of $\mathbf{U}$).

   That is, $U_{i^*,r^*} = \max(U_{i_1,1}, U_{i_2,2}..., U_{i_{R_t},R_t})$.

3. Select and assign the i^*th MS to the r^*th RS. Set the i^*th row and r^*th column of $\bar{\mathbf{I}}$ to 1 (i.e., $\bar{\mathbf{I}}_{[i,r]} = 1$).

4. Increase the number of MSs assigned to the r^*th RS by 1 (i.e., $K_r^{(k)} = K_r^{(k-1)} + 1, r \in \mathcal{R}^{(k)}$).

5. Remove the i^*th MS from the set of MSs yet to be selected and assigned to an RS (i.e., set the i^*th row of $\mathbf{U}$ to zeros).

6. If the r^*th RS has been assigned its maximum number of MSs, remove it from the yet to be selected RSs (i.e., set the r^*th column of $\mathbf{U}$ to zero).

7. If R RSs has been selected, discard all yet to be selected RSs (i.e., set columns of $\mathbf{U}$ corresponding to all non-selected RSs to zero).

end

5.3.1 Possible Implementation

As noted earlier in Section 4.5, CSI acquisition is beyond the scope of this work. However, we would like to mention a few things about CSI acquisition peculiar to this work. One of the major challenges faced by practical wireless relay networks is the enormous feedback overhead and complexity encountered during CSI acquisition. Channel training and estimation in MIMO AF relay-equipped wireless networks for the receiver to obtain both the transmitter-relay and the relay-receiver CSI (similar to the CSI requirement at the BS for joint MS-RS association and selection in this work) have been investigated in the literature. While the relay-receiver CSI is readily obtained by pilot training (from the relay) using standard MIMO training schemes, obtaining the transmitter-relay CSI at the receiver is usually more complicated [149–151].

For practical implementation of the proposed joint user-relay selection and association scheme, each node should be assigned a unique identification (ID) number. In any wireless network, a node identifier is assigned to each user by the MAC layer upon initialization or admission to the network. Similar mechanism can be used to assign IDs to the RSs and MSs.

Similarly to hand-off mechanism in cellular networks, each MS reports its channel conditions to the nearest RS. Each RS in turn reports its channel conditions (i.e., between the RS and the BS as well as between itself and each MS) to the BS. The BS then implements Algorithm 4 to generate the binary association matrix $\bar{\mathbf{I}}$ and broadcasts it to the RSs. Each RS then checks, if it has been selected (i.e., if there is a 1 in its column) and for the MSs it is assigned to (i.e., the indices with 1s in the RS's column). Recall that the MS-RS selection and association need to be completed within a small fraction of the coherence time of the channel before data transmission.

By having the selected RSs "turned on" and the non-selected RSs "turned off", interference, energy consumption, signalling overhead, and design complexity are notably reduced.

5.4 Simulation Results

In this section, we present the simulation results of our proposed joint user-relay selection and association scheme. Without loss of generality, we assume the same numbers of antennas and transmit powers (normalized by the noise power) at similar nodes. That is M_B antennas at the BS, $M_{R,r} = M_R$ antennas at each RS, $N_i = N$ antennas at each MS, as well as $P_{U,i} = P_U$ and $P_{R,r} = P_R$ transmit powers at each

MS and RS, respectively.

In Sub-section 5.4.1, we briefly outline other selection and association schemes which we compare our proposed scheme with, while Sub-section 5.4.2 shows the numerical results.

5.4.1 Comparison with Other Schemes

1. **No Selection and No Association Scheme**

   We first consider a "*no selection, no association*" scheme which randomly selects K MSs and R RSs respectively for cooperation, and then sequentially assign subsets of selected MSs to different selected RSs for service. For example, the first K_1 MSs to the first RS, the next K_2 MSs to the 2nd RS and so on. This scheme does not take into consideration the channel conditions between the nodes, hence is expected to perform worse than CSI-based schemes. The worst-case complexity (in terms of the number of comparisons) of the "*no selection, no association*" is $\mathcal{O}(1)$.

2. **No Selection and Yes Association Scheme**

   Similarly to the "*no selection, no association*" presented earlier, the "*no selection, yes association*" scheme randomly selects K MSs and R RSs respectively for cooperation. That is, it does not consider the channel gains between the nodes to select the MSs and RSs. However, it takes into consideration the channel strengths to assign the MSs to the RSs. Specifically, the CSI-based user-relay association scheme of Algorithm 3 proposed in Chapter 4 and [31], where the MS with the best channel to a given RS is assigned to the RS (one at a time), is employed. While Algorithm 3 of Chapter 4 uses the utility matrix (4.35), the same result is obtained when the utility matrix (5.10) is used instead. Notice that only the $||\mathbf{G}_{i,r}||_F$ part of (5.10) actually matters. This is because for a given RS, the RS-BS link is the same irrespective of the MS associated with the RS.

   The worst-case complexity (in terms of the number of comparisons) of the "*no selection, yes association*" scheme is $\mathcal{O}(K^2R)$. That is, it scales quadratically in the number of selected users and linearly in the number of selected relays. Hence, our proposed joint user-relay selection and association scheme has higher complexity compared to the "*no selection, yes association*" as should be expected. This is especially so when $K \ll K_t$ and/or $R \ll R_t$. However, when $K \approx K_t$ and $R \approx R_t$, the worst-case complexities of both schemes are similar.

3. **2-Step Selection and Association Schemes**

Here, we consider two 2-step user-relay selection and association schemes namely the *harmonic* and the *max-min* schemes. Each scheme (*harmonic* and *max-min*) selects R "best" RSs out of the available R_t RSs in the first step, and then selects and associates K out of K_t available MSs to the R selected RSs in the second step.

Specifically, the *harmonic* scheme is based on the harmonic relay selection (Policy I) of [120, 126] in which the RS with the largest harmonic-mean of the access (MS-RS) and relay (RS-BS) links' channel gains is selected. However, the harmonic scheme of [120,126] is a *single-relay selection* scheme and only for networks with single-antenna transmitters and receivers. In order to accommodate the *multi-relay selection* for networks with multi-antenna transmitters and receivers considered in this work, the R RS with the largest channel gains are selected (one at a time) using the following criterion:

$$\max_{j} \left\{ \frac{||\mathbf{H}_r||_F \; ||\mathbf{G}_r||_F}{||\mathbf{H}_r||_F + ||\mathbf{G}_r||_F} \right\}. \tag{5.11}$$

Similarly, the *max-min* scheme is based on the *best-worst* relay selection (Policy II) of [120, 126] in which the RS whose worst access link or relay link channel gain is the largest is selected. However, the *best-worst* relay selection scheme of [120,126] is a *single-relay selection* scheme and only for networks with single antennas nodes. In order to accommodate the *multi-relay selection* for networks with multi-antenna nodes considered in this work, the first R RS with largest channel gains are selected (one at a time) as follows:

$$\max_{r} \left\{ \min\{||\mathbf{H}_r||_F, ||\mathbf{G}_r||_F\} \right\}. \tag{5.12}$$

Both the *harmonic* and *best-worst single-relay selection* schemes of [120, 126] were proved to achieve full diversity orders in [126, 129].

While the R RSs are selected based on Equations (5.11) and (5.12) for the *harmonic* and *max-min* schemes respectively, the K MSs are selected and assigned to the R selected RSs by solving Problem (5.9) without Condition C4. This is equivalent to employing Algorithm 4 with R equal to the number of selected RSs at initialization (instead of 0) and without Step 7.

Similarly to the proposed joint scheme, the worst case complexity for selecting the R relays using "min-heap" strategy [153] is $R_t log R$ while that for selecting

the K MSs and associating them to the R RSs is $K(RK_t+R)$. Thus, the worst-case complexities of the *harmonic* and *max-min* schemes scale as $R_t logR + K(RK_t + R)$.

4. **Optimal "Exhaustive Search" Scheme**

   For accurate evaluation of the performance loss due to the sub-optimality of the proposed joint scheme, we also consider the optimal exhaustive search strategy, which searches over all possible MS-RS selection and association combinations in order to find the exact combination that results in the highest sum rate. Hence, the exhaustive search scheme has a very high computational complexity, which is given by $\binom{R_t}{R}\binom{K_t}{K}$ K!.

5.4.2 Numerical Results

In this Sub-section, we evaluate the performance of our proposed joint scheme by simulation and compare it with other selection and/or association schemes. We assume that $R = 3$ RSs and $K = 3$ MSs are selected out of R_t RSs and K_t MSs, respectively.

The average sum rate (averaged over 500 channel realizations) in a Rayleigh fading environment is shown in all simulations, unless otherwise stated. For the user-relay beamforming designs, we employ *Algorithm* 2 of Chapter 4 (i.e., *Algorithm* 1 of [30, 31]), with iteration-refinement level $\mu = 10^{-5}$ and fixed initialization assuming equal power allocation at the MSs. That is, $\bar{\mathbf{\Lambda}}_{\mathbf{F}}^{(0)} = \frac{P_U}{N}\mathbf{I}$.

Firstly, we investigate the impact of utility matrix design on the system performance. Specifically, the proposed joint scheme is investigated using two different utility matrices. The first utility matrix design, herein called the *joint design - harmonic* employs the utility matrix design of (5.10), which is based on the harmonic mean of the access and relay links channel gains. The second design, herein called the *joint design - sum* employs a utility matrix design based the sum of the access and relay links' channel gains (i.e., $U_{[ir]} = ||\mathbf{H}_r||_F + ||\mathbf{G}_{i,r}||_F$). As evident in Figure 5.2, the harmonic-mean-based utility matrix design outperforms the sum-based utility matrix design. Thus, the utility matrix design impacts the performance of the proposed joint scheme. Henceforth, we adopt the harmonic-mean-based utility matrix design for the proposed joint scheme, which we henceforth refer to as "*joint design*" in subsequent numerical results.

Secondly, we compare the proposed joint scheme with the optimal exhaustive search scheme. Due to the very high computational complexity of the exhaustive search scheme, we consider low number of channel realizations (50) and reduced

simulation size ($K_t = R_t = 4$). As evident in Figure 5.3, the proposed scheme achieves a reasonable percentage of the exhaustive search performance.

Figure 5.4 compares our proposed *joint design* with various schemes outlined in Sub-section 5.4.1, except the exhaustive search (due to its very high computational complexity). As can be observed, the proposed joint scheme outperforms all the other schemes. Of the 2-step user-relay selection and association schemes, the harmonic scheme outperforms the max-min scheme. Furthermore, the *no selection yes association* outperforms the *no selection yes association* scheme. That is, associating different MSs with different RSs for service is beneficial, without both relay and user selection. Similar results are observed with higher user transmit power, albeit with higher multiplexing gain (i.e., slope of the sum rate curves) as shown in Figure 5.5.

Furthermore, Figures 5.6 and 5.7 illustrate the effects of the total number of users K_t on the network performance for fixed total number of relays ($R_t = 30$) and user transmit power ($P_U = 0$ dB). The proposed joint scheme outperforms the 2-step harmonic and max-min designs. However, the max-min scheme performed quite similarly to the harmonic scheme when the total number of users in the system is large. This is attributable to the higher probability of both the access and relay links of each relay having good channels with a large number of users. Similar behaviour can be observed at high user transmit power as depicted in Figures 5.8 and 5.9.

Figure 5.10 and 5.11 respectively depict the impact of total number of users and number of antennas on the network performance. Again, the proposed joint user-relay selection and association scheme outperforms both the 2-step (harmonic and max-min) schemes for the same systems configurations.

5.5 Conclusion

In this chapter, we have studied joint user-relay selection and association for MIMO MARC, in which R out of R_t fixed infrastructure-based half-duplex in-band multi-antenna AF RSs are selected to extend radio coverage between K out of K_t selected MSs and a BS. Unlike various schemes in the literature, in which either the RSs and the MSs are equipped with a single antenna and/or only a single RS and MS are selected, a more general system setup in which several multi-antenna RSs and MSs are selected and associated with each other, has been studied. Owing to the high complexity of the optimal solution, we proposed a novel low complexity CSI-based joint MS-RS selection and association scheme, which simultaneously selected the RSs and MSs, as well as assigned the selected MSs to different selected RSs for service. The proposed joint scheme leverages the inherent cooperative and multiuser diversi-

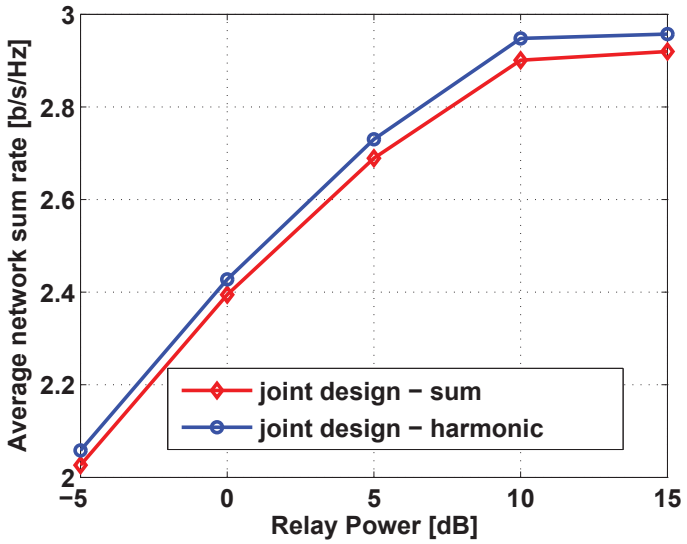

Figure 5.2: Average sum rate of a single-cell cellular relay uplink with user transmit power $P_U = 20$ dB, $M_B = 6$ antennas at BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 6$ and $K_t = 5$ RSs and MSs respectively, for different utility matrices.

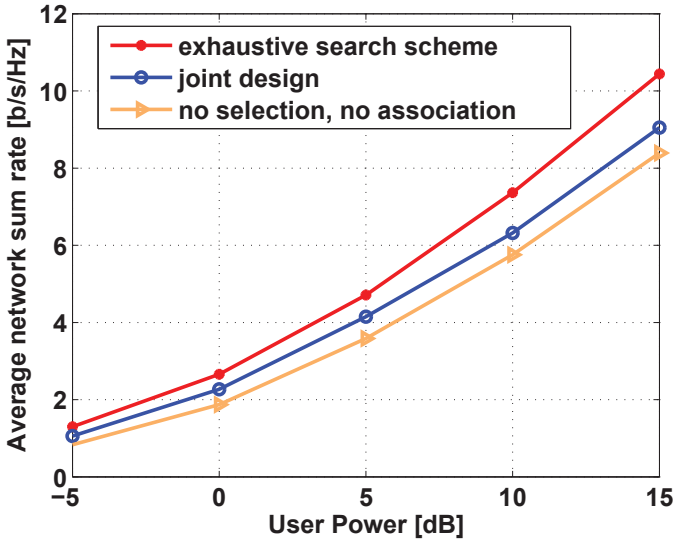

Figure 5.3: Average sum rate versus user transmit power of a single-cell cellular relay uplink with relay transmit power $P_U = 0$ dB, $M_B = 6$ antennas at BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 4$ and $K_t = 4$ RSs and MSs respectively, are selected for cooperation.

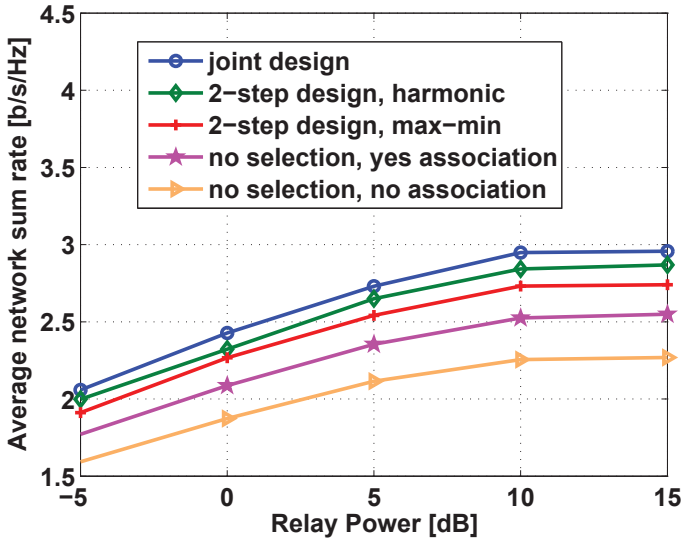

Figure 5.4: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 0$ dB, $M_B = 6$ antennas at BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 6$ and $K_t = 5$ RSs and MSs respectively, are selected for cooperation.

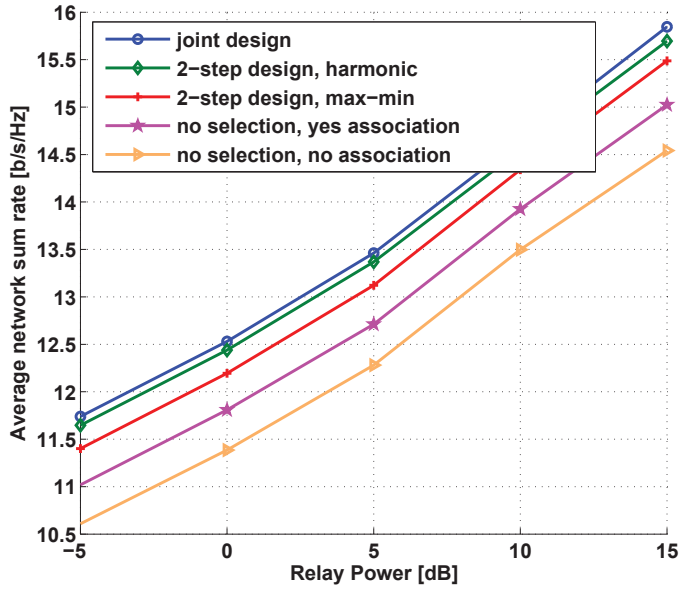

Figure 5.5: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 20$ dB, $M_B = 6$ antennas at BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 6$ and $K_t = 5$ RSs and MSs respectively, are selected for cooperation.

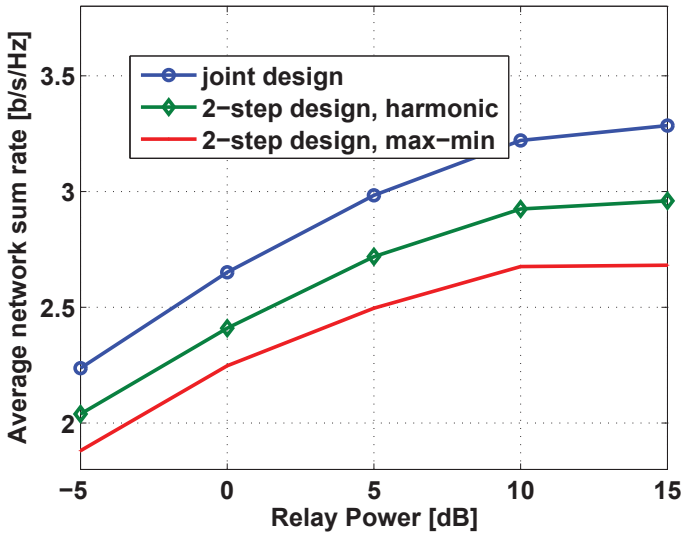

Figure 5.6: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 0$ dB, $M_B = 2$ antennas at each BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 30$ and $K_t = 5$ RSs and MSs respectively, are selected for cooperation.

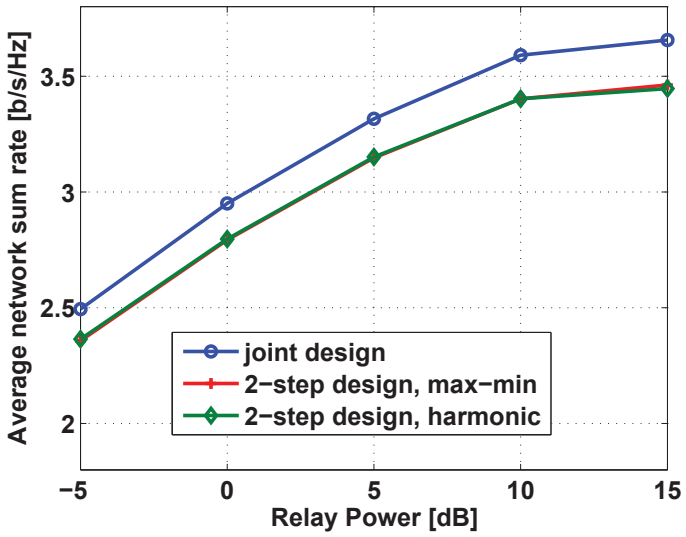

Figure 5.7: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 0$ dB, $M_B = 2$ antennas at each BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 30$ and $K_t = 100$ RSs and MSs respectively, are selected for cooperation.

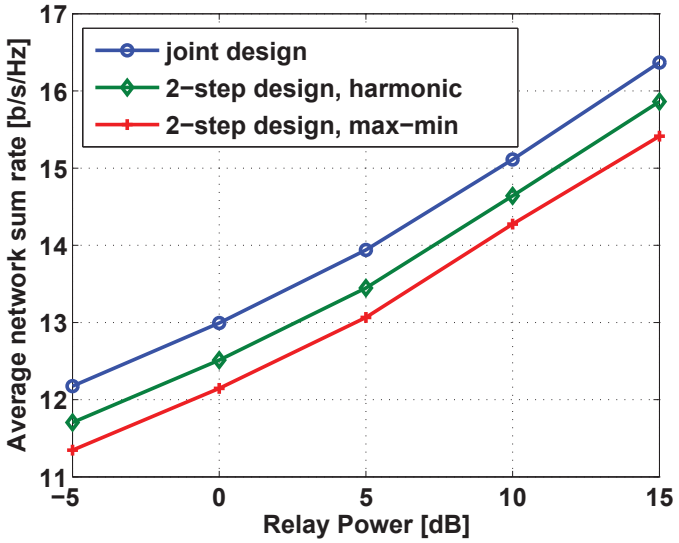

Figure 5.8: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 20$ dB, $M_B = 2$ antennas at each BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 30$ and $K_t = 5$ RSs and MSs respectively, are selected for cooperation.

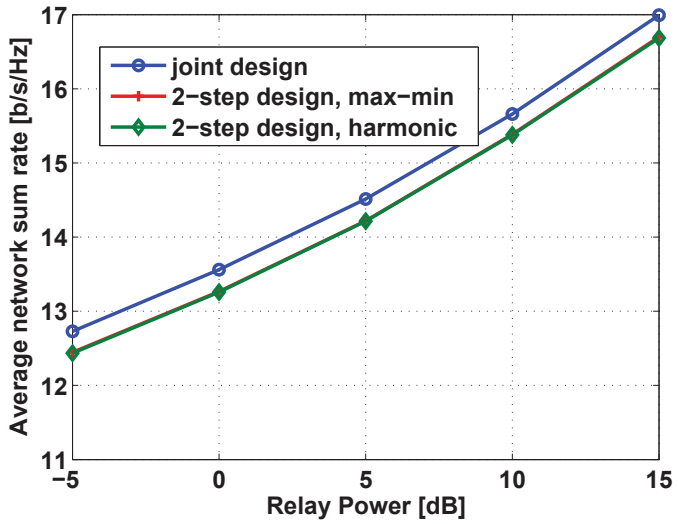

Figure 5.9: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 20$ dB, $M_B = 2$ antennas at each BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 30$ and $K_t = 100$ RSs and MSs respectively, are selected for cooperation.

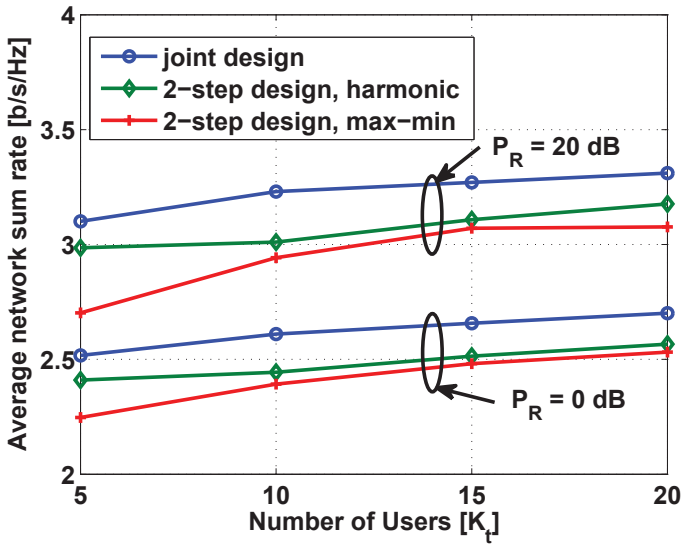

Figure 5.10: Average sum rate versus the number of users K_t of a single-cell cellular relay uplink with user transmit power $P_U = 0$ dB, $M_B = 6$ antennas at BS, $M_R = 6$ antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 30$ and K_t RSs and MSs respectively, are selected for cooperation.

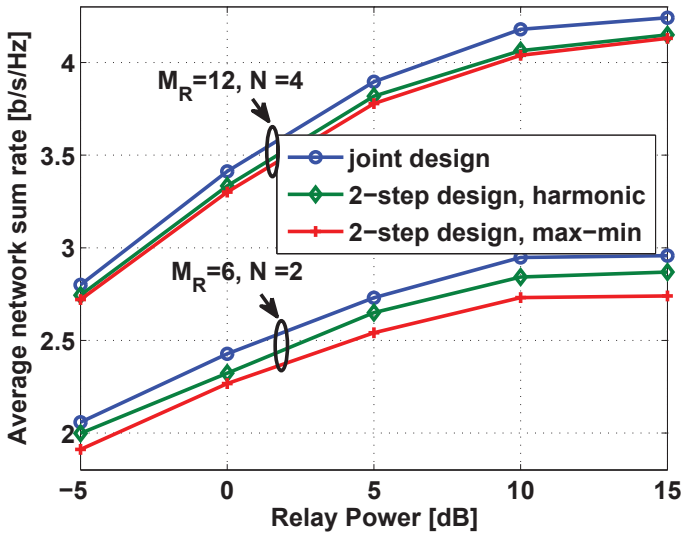

Figure 5.11: Average sum rate versus relay transmit power of a single-cell cellular relay uplink with user transmit power $P_U = 0$ dB, $M_B = 6$ antennas at BS, M_R antennas at each RS, and $N = 2$ antennas at each MS, in which $R = 3$ and $K = 3$ out of $R_t = 6$ and $K_t = 5$ RSs and MSs respectively, are selected for cooperation.

ties in the network to enhance the system performance and at the same time results in reduced feedback overhead and interference. Simulation results demonstrate the superiority of the proposed joint scheme compared to (i) a scheme with neither user-relay selection nor user-relay association, (ii) a scheme with user-relay association but no user-relay selection, and (iii) two 2-step schemes that select the RSs in the first step and then select and assign the MSs to the selected RSs in the second step. Moreover, associating different MSs with different RSs for service (without user and/or relay selection) offered better performance compared to a scheme without both user-relay selection and association. Most importantly, the proposed joint scheme achieves a reasonable percentage of the performance of the exhaustive search scheme. The favorable performance and relatively low complexity of the proposed scheme makes it very attractive for possible implementation in emerging and future broadband cellular networks. In spite of the emphasis on the uplink in this work, the proposed joint user-relay selection and association scheme is also applicable to the downlink of a multi-user multi-relay wireless network.

Chapter 6

Summary of Contributions and Future Work

In this chapter, we summarize the major contributions of this work as well as suggest possible directions for future work.

The focus of this work was to design transceiver schemes and scheduling algorithms for enhancing cell-edge user-experience in emerging broadband cellular networks. In particular, the use of fixed infrastructure-based multi-antenna in-band relays for coverage extension (at high data rates) was considered. We adopted the non-regenerative also known as amplify-and-forward relays owing to their more attractive features. In order to mitigate the high ICI jeopardizing the high data rate promise of MIMO spatial multiplexing in cellular networks, CoMP transmission/reception also known as *network MIMO* was employed. Leveraging these key technologies (MIMO, CoMP, and relaying) in a unified manner, we proposed transmit covariance matrices and relay beamforming schemes that enhance the cellular network performance. Due to the availability of more users and relays than the system can support simultaneously, we proposed practical user-relay selection and association schemes, which led to reduced interference, energy consumption, signaling overhead, and design complexity.

6.1 Summary of Contributions

The major contributions of this thesis are summarized as follows.

- In Chapter 3, we have jointly designed the input covariance matrices at the BS and beamforming matrix at the RS to maximize the sum rate of a single-cell BRC with multi-antenna transmitters and receivers. Due to the high complexity of the BRC sum-rate problem, we employed uplink-downlink duality relationship [135] in order to transform the BRC problem into the more tractable dual

MARC problem. While this duality was proved for single-antenna-user networks and claimed to hold for multi-antenna-user networks, we have provided a detailed proof for the multi-antenna user case. Furthermore, we derived the mapping of the resulting covariance matrices for the MARC to the desired covariance matrices for the BRC. Simulation results show that our design outperforms other existing schemes in the literature. Moreover, our proposed scheme performed close to the upper bound with the performance gap decreasing with increasing number of antennas at the users. It was shown that having more antennas at the RS than at the BS (compared to having more antennas at the BS than the RS) is crucial for best system performance.

- In Chapter 4, we have investigated the joint design of transmit covariance matrices and the RSs beamforming matrices to maximize the sum rate of a multi-cell MIMO-AF relay uplink with coordinated BSs. We jointly designed the input covariance matrices and the RSs' beamforming matrices to decouple the RSs' channels (by incorporating interference pre-cancelation into the design) as well as diagonalize the end-to-end channel (in order to simplify the MIMO receiver processing). With this beamforming design, an iterative alternating-maximization based algorithm to obtain the network sum rate was then proposed. Finally, we have proposed a low complexity CSI-based user-relay association scheme, by which MSs are assigned to the RSs for service. Simulation results have demonstrated the effectiveness of our proposed designs.

- In Chapter 5, we investigated user-relay selection and association strategy in multi-user multi-relay MIMO wireless network. We have proposed a novel low-complexity user-relay selection and association scheme, which simultaneously selects multiple relays and users for cooperation as well as assigns the selected users to different selected relays. That is, three different tasks (relay selection, user selection/scheduling, as well as user-relay association) are performed concurrently. Moreover, our proposed scheme utilizes only the channel gains between the nodes (i.e., partial instead of full CSI is needed), which is more practical and easier to implement. Compared with other existing schemes, the proposed scheme offers superior performance. The relatively low complexity and favorable performance of the proposed scheme makes it very attractive for implementation in emerging and future broadband cellular networks.

6.2 Future Work

In the following, we outline some areas of possible future research work.

- Investigation of the potential of exploiting the spatial dimensions of multi-antenna relays for interference mitigation (in addition to the primary task of coverage extension) in coordinated MIMO AF relay downlinks.

  In a coordinated MIMO downlink employing joint transmission (JT) variant of CoMP, the transmitted signals are jointly designed (via the precoding matrices) at the BSs within the coordinated cluster so that each user receives an interference-free signal. With the deployment of AF RSs within the network, extra care needs to be taken to design the precoding matrices at the BSs and beamforming matrices at the RSs. Employing JT to design the precoding matrices will be futile. This is because the MIMO AF relays will undo the interference-mitigation built into the precoding matrices as they amplify and forward their received signals (including interference) from the BSs to the users. This will in turn degrade the SINR of the system.

  We have considered the possibility of utilizing multi-antenna AF relays to mitigate interference from unintended sources (i.e., users) in a coordinated MIMO cellular uplink. Similar approach will be particularly useful in coordinated cellular downlinks with multiple MIMO AF relays. Coordinated beamforming/scheduling variant of CoMP, where interference-mitigation is coordinated among the BSs will be the most realistic approach. In particular, the task of interference mitigation can be shared between the BSs and RSs within the cluster, while utilizing the global CSI knowledge of other BSs.

- Further investigation of low complexity joint user-relay selection/scheduling and association schemes for multi-user multi-relay cellular networks, especially those employing half-duplex relays. This is particularly crucial due to the delay-sensitivity of half-duplex relay networks as a result of the multiple transmission phases/hops needed to convey information from the source to the destination.

  We have considered joint selection/scheduling and association of users and relays with best channel gains in this work. However, it has been shown that scheduling of users with best channel gains and large spatial separation (in order to reduce multi-user interference) is sum-rate optimal for MIMO broadcast channels without relays [74]. Similar consideration, that is, selection and association of relays and users with best as well as mutually orthogonal channels should be explored.

Also, fairness (e.g., a variant of proportional-fair user-scheduling for conventional multi-user MIMO networks without relays [77,78]) should be investigated in order to ensure better QoS for all users, unlike our proposed greedy scheduling and association schemes in which only users and relays with the best channel gains are selected for cooperation.

- Extension to heterogeneous networks (HetNets) comprising mixture of macro and pico cells/BSs. In order to provide flexible capacity expansion and/or traffic off-loading, networks with a mixture of cells of different types and power class (macro, pico, femto, relays, etc) widely known as HetNets are to be deployed in future broadband cellular systems. While the high power (e.g. macro) nodes will be deployed with wider coverage in mind, the low power nodes such as pico, femto and remote radio heads (RRHs) will be deployed for traffic offloading from the macro-cell [21, 154]. Co-existence of such systems will result in severe interference especially at the receivers of the low-power nodes (e.g., pico) due to transmissions of the high-power transmitter (e.g., macro). To effectively support the deployment of HetNets, enhanced intercell interference coordination(eICIC) has been proposed (e.g., in 3GPP's Release 10) to mitigate interference from macro to pico cells [21].

  Being an integral part of HetNets, RSs will be deployed along side macro and pico BSs. However, throughout this work, we have considered cellular relay networks with homogeneous (i.e., macro) cells/BSs. Thus, investigation of joint source-relay beamforming designs as well as user-relay selection/scheduling and association schemes for MIMO HetNets will be very desirable.

Bibliography

[1] C. Chae, T. Tang, R. W. Heath, Jr., and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Processing*, vol. 56, no. 2, pp. 727–738, February 2008.

[2] ITU-R Rep. M.2134, "Requirements related to technical performance for IMT-Advanced radio interface(s)," 2008.

[3] ITU-R Rec. M.1645, "Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000," 2003.

[4] 3GPP RP-080137, "Proposed SID on LTE-Advanced," *NTT DoCoMo, etc.*, 3GPP TSG RAN meeting 39.

[5] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMAX and LTE-Advanced mobile systems," *IEEE Commun. Mag.*, pp. 100–105, October 2009.

[6] K. Loa, C. Wu, S. Sheu, Y. Yuan, M. Chion, D. Huo, and L. Xu, "IMT-Advanced relay standards," *IEEE Commun. Mag.*, pp. 40–48, August 2010.

[7] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-Advanced - Evolving LTE towards IMT-Advanced," in *Proc. IEEE VTC Fall*, September 2008, pp. 1–5.

[8] E. Dahlman, S. Parkvall, and J. Skold, *4G LTE/LTE-Advanced for Mobile Broadband*. Academic Press, 2011.

[9] A. Ghosh, J. Zhang, J. Andrews, and R. Muhamed, *Fundamentals of LTE*. Upper Saddle River, NJ, USA: Prentice Hall, 2010.

[10] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*. Chichester, West Sussex, UK: John Wiley, 2011.

[11] "ITU paves way for Next-Generation 4G mobile technologies," *ITU Press Release, 21 Oct.*, 2010.

[12] J. Thompson, X. Ge, H.-C Wu, R. Irmer, H. Jiang, G. Fettweis, and S. Alamouti, "5G wireless communication systems: prospects and challenges," *IEEE Commun. Mag.*, vol. 52, no. 2, February 2014.

[13] J. Thompson, X. Ge, H-C Wu, R. Irmer, H. Jiang, G. Fettweis, and S. Alamouti, "5G wireless communication systems: prospects and challenges part 2," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014.

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

[15] I. Telatar, "Capacity of multi-antenna Gaussian channels," *Euro. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–596, November 1999.

[16] A. Paulraj, D. Gore, and R. Nabar, *Introduction to Space-Time Wireless Communications*. Cambridge, United Kingdom: Cambridge University Press, 2006.

[17] G. J. Foschini, K. Karakayali, and R. A. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," in *IEE Proc. - Commun.*, vol. 153, no. 4, August 2006, pp. 548–555.

[18] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, pp. 56–61, August 2006.

[19] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-Advanced: Deployment scenarios and operational challenges," *IEEE Commun. Mag.*, pp. 148–155, February 2012.

[20] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE J. Select. Areas Commun.*, vol. 28, no. 9, pp. 1–29, December 2007.

[21] S. Sun, Q. Gao, Y. Peng, Y. Wang, and L. Song, "Interference management through CoMP in 3GPP LTE-Advanced networks," *IEEE Wireless Commun.*, pp. 59–66, February 2013.

[22] R. Pabst, B. H. Walke, D. C. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80–89, Septemeber 2004.

[23] J. Sydir and R. Taori, "An evolved cellular system architecture incorporating relay stations," *IEEE Commun. Mag.*, pp. 115–121, June 2009.

[24] S. W. Peters and R. W. Heath, Jr., "The future of WiMAX: Multihop relaying with IEEE 802.16j," *IEEE Commun. Mag.*, pp. 104–111, January 2009.

[25] M. Rumney, *LTE and the Evolution to 4G Wireless: Design and Measurement Challenges. Second Edition*. Chichester, West Sussex, UK: John Wiley & Sons, 2013.

[26] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Commun. Mag.*, pp. 50–56, July 2009.

[27] O. Somekh, O. Simeone, V. Poor, and S. Shamai, "Cellular systems with non-regenerative relaying and cooperative base stations," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2654–2663, August 2010.

[28] G. O. Okeke, Y. Jing, and W. A. Krzymień, "Beamforming in MIMO broadcast relay networks with multiple antenna users," in *Proc. 45th Asilomar Conf. on Signals, Systems and Computers (ACSSC)-Invited Paper*, Pacific Grove, CA, USA, November 2011, pp. 1222–1226.

[29] G. Okeke, W. A. Krzymień, and Y. Jing, "Beamforming in non-regenerative MIMO broadcast relay networks," *IEEE Trans. Signal Processing*, vol. 60, no. 12, pp. 6641–6654, December 2012.

[30] G. O. Okeke, W. A. Krzymień, and Y. Jing, "Distributed beamforming in multi-cell cooperative MIMO cellular networks with non-regenerative relays: An LTE-Advanced framework," in *Proc. IEEE GLOBECOM*, Atlanta, GA, USA, December 2013, pp. 4240–4245.

[31] G. O. Okeke, Y. Jing, W. A. Krzymień, and J. Melzer, "Interference mitigation and user-relay association in multi-cell cooperative MIMO cellular uplink with non-regenerative relaying: An LTE-Advanced framework," *Submitted to IEEE Trans. Veh. Technol.*, July 2014.

[32] G. O. Okeke, W. A. Krzymień, Y. Jing, and J. Melzer, "A novel low-complexity joint user-relay selection and association for multi-user multi-relay MIMO uplink," *accepted for publication in IEEE Wireless Commun. Lett.*, February 2015.

[33] D. Tse and P. Viswanath, *Fundamentals of Wirelesss Communication*. Cambridge, United Kingdom: Cambridge University Press, 2005.

[34] A. Goldsmith, *Wirelesss Communications*. New York, NY, USA: Cambridge University Press, 2005.

[35] H. Huang, C. Papadias, and S. Venkatesan, *MIMO Communcations for Cellular Networks*. Spring City, NY, USA: Springer, 2012.

[36] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, 1996.

[37] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1998.

[38] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. Vincent Poor, *MIMO Wireless Communications*. New York, NY, USA: Cambridge University Press, 2007.

[39] W. Jakes, *Microwave Mobile Communications*. New York, NY, USA: Wiley, 1974.

[40] A. Wittneben, "Base station modulation diversity for digital SIMULCAST," in *Proc. IEEE VTC*, St. Louis, MO, May 1991, pp. 848–853.

[41] R. W. Heath Jr. and A. J. Paulraj, "Transmit diversity using decision directed antenna hopping," in *Proc. Comm. Theory Mini-Conference*, Vancouver, Canada, June 1999, pp. 141–145.

[42] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[43] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 744–765, March 1998.

[44] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[45] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[46] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[47] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *Proc. 38th Annu. Allerton Conf. Commun., Control and Comput.*, Monticello, IL, USA, October 2000.

[48] ——, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[49] 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, October 2003.

[50] L. Liu, R. Chen, S. Geirhofer, K. Sayana, Z. Shi, and Y. Zhou, "Downlink MIMO in LTE-Advanced: SU-MIMO vs. MU-MIMO," *IEEE Commun. Mag.*, pp. 140–147, February 2012.

[51] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3417–3432, October 2005.

[52] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice startegies for cancelling known interference," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3820–3833, November 2005.

[53] A. Khina and U. Erez, "On the robusteness of dirty paper coding," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1437–1446, May 2010.

[54] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication - part II: perturbation," *IEEE Trans. Commun.*, vol. 53, no. 31, pp. 537–544, March 2005.

[55] M. Mazrouei-Sebdani and W. A. Krzymień, "Vector perturbation precoding for network MIMO: sum rate, fair user scheduling and impact of backhaul delay," *IEEE Trans. Veh. Technol.*, vol. 61, no. 9, pp. 3946–3957, November 2012.

[56] ——, "On MMSE vector-perturbation precoding for MIMO broadcast channels with per-antenna group power constraints," *IEEE Trans. Signal Processing*, vol. 61, no. 15, pp. 3745–3751, August 2013.

[57] M. Mazrouei-Sebdani, W. A. Krzymień, and J. Melzer, "Massive MIMO with non-linear precoding: large-system analysis," *accepted for publication in IEEE Trans. Veh. Technol.*, March 2015.

[58] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Trans. J. Select. Areas Commun.*, vol. 24, no. 3, pp. 528–541, March 2006.

[59] Q. Spencer, A. Swindlehurst, and M. Haadrt, "Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 461–471, February 2004.

[60] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication - part I: channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, January 2005.

[61] A. D. Dabbagh and D. J. Love, "Precoding for muliple antenna Gaussian broadcast channels with successive zero-forcing," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3837–3850, July 2007.

[62] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communcations," in *Proc. IEEE Int. Conf. Commun. (ICC'95), Seattle, WA*, June 1995, pp. 331–335.

[63] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1893–1909, September 2004.

[64] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Trans. Commun.*, vol. 55, no. 1, p. 1115, January 2007.

[65] J. Wang, D. J. Love, and M. D. Zoltowski, "User selection with zero-forcing beamforming achieves asymptotically optimal sum rate," *IEEE Trans. Signal Processing*, vol. 56, no. 8, pp. 3713–3726, August 2008.

[66] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3857–3868, October 2005.

[67] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath Jr., and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Processing*, vol. 54, no. 9, pp. 3658–3663, September 2006.

[68] B. C. Lim, W. A. Krzymień, and C. Schlegel, "Efficient sum rate maximization and resource allocation in block-diagonalized space-division multiplexing," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 478–484, January 2009.

[69] J. Dai, C. Chang, Z. Ye, and Y. S. Hung, "An efficient greedy scheduler for zero-forcing dirty-paper coding," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1939–1943, July 2009.

[70] S. Sigdel, W. A. Krzymień, and M. Al-Shalash, "Greedy and progressive user scheduling for CoMP wireless networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, Canada, June 2012, pp. 5762–5767.

[71] H. Purmehdi, R. C. Elliott, and W. A. Krzymień, "Simulated annealing user scheduling for coordinated heterogeneous MIMO networks," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Comput.*, Pacific Grove, CA, USA, November 2012, p. 11571161.

[72] R. C. Elliott, S. Sigdel, and W. A. Krzymień, "Low complexity greedy, genetic, and hybrid user scheduling algorithms for multiuser mimo systems with successive zero-forcing," *Trans. Emerging Telecommun. Technol. (ETT)*, vol. 23, no. 7, pp. 604–617, November 2012.

[73] M. Eslami, R. C. Elliott, W. A. Krzymień, and M. Al-Shalash, "Location-assisted clustering and scheduling for coordinated homogeneous and heterogeneous cellular networks," *Trans. Emerging Telecommun. Technol. (ETT)*, vol. 24, no. 1, pp. 84–101, January 2013.

[74] R. W. Heath Jr., M. Airy, and A. J. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receivers," in *Proc. 35th Asilomar Conf. on Signals, Systems and Computers (ACSSC)*, November 2001, pp. 1194–1199.

[75] D. J. Mazzarese and W. A. Krzymień, "Scheduling algorithms and throughput maximization for downlink of packet-data cellular systems with multiple antennas at the base station," *Wireless Personal Commun.*, vol. 43, no. 2, pp. 215–260, October 2007.

[76] L. -N. Tran and E. -K. Hong, "Multiuser diversity for successive zero-forcing dirty paper coding: greedy scheduling algorithms and asymptotic performance analysis," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3411–3416, June 2010.

[77] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[78] V. K. N. Lau, "Proporional fair space-time scheduling for wirelesss communications," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1353–1360, August 2005.

[79] Y. Nam, Y. Akimoto, Y. Kim, M. Lee, K. Bhattad, and A. Ekpenyong, "Evolution of reference signals for LTE-Advanced systems," *IEEE Commun. Mag.*, pp. 132–138, February 2012.

[80] T. Marzetta, "BLAST training: estimating channel charateristics for high-capacity space-time wireless," in *Proc. Allerton Conference on Wireless Communications, Control and Computing*, September 1999, pp. 958–966.

[81] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. W. Heath, Jr., "Networked MIMO with clustered linear precoding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1910–1921, April 2009.

[82] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2983–2989, June 2009.

[83] S. Kaviani, O. Simeone, W. A. Krzymień, and S. Shamai (Shitz), "Linear precoding and equalization for network MIMO with partial cooperation," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2083–2096, June 2012.

[84] J. Y. Hwang, J. Kim, T. Kim, and Y. Han, "A periodic frequency band rotation scheme for multi-cell coordination clustering," *IEEE Commun. Lett.*, vol. 15, no. 9, pp. 956–958, September 2011.

[85] H. Purmehdi, R. C. Elliott, W. A. Krzymień, and J. Melzer, "Rotating clustering with simulated annealing user scheduling for coordinated heterogeneous MIMO cellular networks," in *Proc. IEEE ICC*, June 2014, pp. 5304–5309.

[86] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in *Proc. IEEE Intern. Conf. on Commun. (ICC)*, May 2008, pp. 4033–4037.

[87] N. Lee, R. W. Heath, Jr., D. Morales-Jimenez, and A. Lozano, "Base station cooperation with dynamic clustering in super-dense cloud-RAN," in *IEEE GLOBECOM Workshop*, Atlanta, GA, USA, December 2013, pp. 784 – 788.

[88] 3GPP TR 36.814 v9.0.0, "Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); further advancements for E-UTRA physical layer aspects (release 9)," TSG RAN.

[89] 3GPP TR 36.819 v11.1.0, "Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); coordinated multi-point operation for LTE physical layer aspects (release 11)," TSG RAN.

[90] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communications in wireles networks," *IEEE Commun. Mag.*, pp. 74–80, October 2004.

[91] Y. Li, "Distributed coding for cooperative wireless networks: an overview and recent advances," *IEEE Commun. Mag.*, pp. 71–77, August 2009.

[92] C. Hoymann, W. Chen, J. Montojo, A. Golitschek, C. Koutsimanis, and X. Shen, "Relaying operation in 3GPP LTE : Challenges and solutions," *IEEE Commun. Mag.*, pp. 156–162, February 2012.

[93] J. Zhang, F. Roemer, and M. Haardt, "Beamforming design for multi-user two-way relaying with MIMO amplify and forward relays," in *Proc. IEEE ICASSP*, May 2011, pp. 2824–2827.

[94] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York, NY, USA: Wiley, 1991.

[95] S. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Comparison of bi-directional relaying protocols," in *Proc. IEEE Sarnoff Symp.*, April 2008, pp. 1–5.

[96] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, Jr., "Relay architectures for 3GPP LTE-Advanced," *EURASIP Journal on Wirleless Communications and Networking*, vol. 2009, article ID 618787, 14 pages.

[97] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, p. 19271938, November 2003.

[98] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062 – 3080, December 2004.

[99] M. Janani, A. Hedayat, T. E. Hunter, and A. Nostratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *IEEE Trans. Signal Processing*, vol. 52, pp. 362–371, February 2004.

[100] R. U. Nabar, H. Bölcskei, and F. W. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, August 2004.

[101] T. M. Cover and A. El Gamal, "Capacity theorems for relay channels," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, September 1979.

[102] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of wireless relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 29–43, January 2005.

[103] J. Kim and D. Kim, "Performance of dual-hop amplify-and -forward beamforming and its equivalent systems in Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 729 –732, March 2010.

[104] Y. Zhao, R. Adve, and T. Lim, "Beamforming with limited feedback in amplify-and-forward cooperative networks," in *Proc. IEEE GLOBECOM*, November 2007, pp. 3457–3461.

[105] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inform. Theory*, vol. 55, no. 6, pp. 2499 –2517, June 2009.

[106] E. Koyuncu, Y. Jing, and H. Jafarkhani, "Distributed beamforming in wireless relay networks with quantized feedback," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1429 –1439, October 2008.

[107] E. Yilmaz and M. O. Sunay, "Amplify-and-forward capacity with transmit beamforning for MIMO multiple relay channels," in *Proc. IEEE GLOBECOM*, November 2007, pp. 3873–3877.

[108] W. Xu, X. Dong, and W. Lu, "Joint optimization for source and relay precoding under multiuser MIMO downlink channels," in *Proc. IEEE ICC*, May 2010, pp. 1–5.

[109] ——, "Joint precoding optimization for multiuser multi-antenna relaying downlinks using quadratic programming," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1228–1235, May 2011.

[110] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory.*, vol. 51, no. 12, pp. 4152–4172, December 2005.

[111] J. Zhao, M. Kuhn, A. Wittneben, and G. Bauch, "Cooperative transmission schemes for decode-and-forward relaying," in *Proc. IEEE PIMRC*, September 2007, pp. 1–5.

[112] M. N. Khormuji and E. G. Larsson, "Cooperative transmission based on decode-and-forward relaying with partial repetition coding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1716–1725, April 2009.

[113] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, April 2007.

[114] Z. Fang, Y. Hua, and J. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," in *Proc. 4th IEEE Workshop on Sensor Array and Multichannel Processing*, July 2006, pp. 239–243.

[115] M. Khandaker and Y. Rong, "Intereference MIMO relay channel: joint power control and transceiver-relay beamforming," *IEEE Trans. Signal Processing*, vol. 60, no. 12, pp. 6509–6518, December 2012.

[116] K. Truong, P. Sartori, and R. W. Heath Jr., "Cooperative algorithms for MIMO amplify-and-forward relay networks," *IEEE Trans. Signal Processing*, vol. 61, no. 5, pp. 1272–1287, March 2013.

[117] K. Truong and R. W. Heath Jr., "Joint transmit precoding for the relay interference broadcast channel," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1201–1215, March 2013.

[118] O. Oyman and A. J. Paulraj, "Design and analysis of linear distributed MIMO relaying algorithms," *IEE Proc. Commun.*, vol. 153, no. 4, pp. 565–572, August 2006.

[119] V. Sreng, H. Yanikomeroglu, and D. D. Falconer, "Relay seletion strategies in cellular networks with peer-to-peer relaying," in *Proc. IEEE Veh. Technol. Conf.*, October 2003.

[120] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 659–672, March 2006.

[121] Y. Zhao, R. Adve, and T. J. Lim, "Symbol error rate of selection amplify-and-forward relay systems," *IEEE Commun. Lett.*, vol. 10, pp. 757–759, November 2006.

[122] A. S. Ibrahim, A. Sadek, W. Su, and K. J. R. Liu, "Relay selection in multi-node cooperative communications: when to cooperate and whom to cooperate with?" in *Proc. IEEE GLOBECOM*, 2006, pp. 1–5.

[123] A. Bletsas, H. Shin, and M. Win, "Outage optimality of opportunistic amplify-and-forward relaying," *IEEE Commun. Lett.*, vol. 11, pp. 261–263, March 2007.

[124] Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: optimal power allocation vs selection," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 3114–3122, August 2007.

[125] A. S. Ibrahim, A. Sadek, W. Su, and K. J. R. Liu, "Cooperative communications with relay selection: when to cooperate and whom to cooperate with?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2814–2827, September 2008.

[126] Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1414 –1423, March 2009.

[127] L. Sun, T. Zhang, L. Lu, and H. Niu, "On the combination of cooperative diversity and multiuser diversity in multi-source multi-relay wireless networks," *IEEE Signal Processing Lett.*, vol. 17, no. 6, pp. 535–538, June 2010.

[128] H. Ding, J. Ge, D. B. da Costa, and Z. Jiang, "A new efficient low-complexity scheme for multi-source multi-relay cooperative networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 716–722, February 2011.

[129] S. Atapattu, Y. Jing, H. Jiang, and C. Tellambura, "Relay selection and performance analysis in multiple-user networks," *IEEE J. Select. Areas Commun.*, vol. 31, no. 8, pp. 1517–1529, August 2013.

[130] M. S. Alam, J. W. Mark, and X. Shen, "Relay selection and resource allocation for multi-user cooperative OFDMA newtorks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2193–2205, May 2013.

[131] M. Abouelseoud and A. Nosratinia, "Heterogeneous relay selection," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1735–1743, April 2013.

[132] A. Nosratinia and T. Hunter, "Grouping and partner selection in cooperative wireles networks," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 1–10, February 2007.

[133] S. Sharma, Y. Shi, Y. T. Hou, and S. Kompella, "An optimal algorithm for relay node assignmemnt in cooperative ad hoc networks," *IEEE /ACM Trans. Netw.*, vol. 19, no. 3, pp. 879–892, June 2010.

[134] D. Yang, X. Fang, and G. Xue, "HERA: an optimal relay assignment scheme for cooperative networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 245–253, February 2012.

[135] K. S. Gomadam and S. A. Jafar, "Duality of MIMO multiple access channel and broadcast channel with amplify-and-forward relays," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 211 –217, January 2010.

[136] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[137] I. Csiszár and G. Tusnády, "Information geometry and alternating minimization procedures," *Statistics and Decisions (Supplement Issue)*, no. 1, pp. 205–237, 1984.

[138] U. Niesen, D. Shah, and G. W. Wornell, "Adaptive alternating minimization algorithms," in *Proc. ISIT*, June 2007, pp. 1641–1645.

[139] ——, "Adaptive alternating minimization algorithms," *IEEE Trans. Inform. Theory*, vol. 55, no. 3, pp. 1423 – 1429, March 2009.

[140] W. Yu and J. Cioffi, "Trellis precoding for the broadcast channel," in *IEEE GLOBECOM*, November 2001, pp. 1344–1348.

[141] ——, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, September 2004.

[142] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY, USA: Cambridge University Press, 1985.

[143] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for Gaussian vector multiple access channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 145–152, January 2004.

[144] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*. Palo Alto, CA, USA: Meboo Publishing, 2005.

[145] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[146] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," ../../cvx, Apr. 2011.

[147] L. Li, Y. Jing, and H. Jafarkhani, "Interference cancellation at the relay for multi-user wireless cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 930–939, March 2011.

[148] ——, "Multisource transmission for wireless relay networks with linear complexity," *IEEE Trans. Signal Processing*, vol. 59, no. 6, pp. 2898–2912, June 2011.

[149] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4438–4446, September 2011.

[150] S. Sun and Y. Jing, "Channel training design in amplify-and-forward MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3380–3391, October 2011.

[151] ——, "Training and decodings for cooperative network with multiple relays and receive antennas," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1534–1544, June 2012.

[152] *3GPP TR 36.814, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects", V9.0.0*, March 2010.

[153] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms: Third Edition.* Cambridge, Massachusetts: The MIT Press, 2009.

[154] P. Bhat, S. Nagata, L. Campoy, I. Berberana, T. Derham, G. Liu, X. Shen, P. Zong, and J. Yang, "LTE-Advanced: an operator perspective," *IEEE Commun. Mag.*, pp. 104–114, February 2012.

Appendix A

Proof of Theorem 3.1

A.1 Uplink-Downlink Duality for MIMO AF-Based Relay Channels with Multi-Antenna Users

As mentioned in Sub-section 3.2.4, the authors of [135] proved uplink-downlink duality for MIMO AF relay channels with single-antenna users. Furthermore, they stated that the duality also holds for the multi-antenna-user case without proof. In the sequel, we provide a detailed proof of this duality for multi-user MIMO AF relay networks with multi-antenna users.

To prove the duality, we first decompose the users' input covariance matrices into their eigenmodes and then treat each sub-stream as a virtual single-antenna user. That is, User i's covariance matrix in the BRC and MARC are respectively re-written as

$$\begin{aligned} \mathbf{\Sigma}_i &\triangleq \mathbf{U}_i \mathbf{\Lambda}_i \mathbf{U}_i^H \\ &= \sum_{j=1}^{N_i} \mathbf{u}_{ij} \mathbf{u}_{ij}^H \lambda_{ij} \end{aligned} \tag{A.1}$$

and

$$\begin{aligned} \mathbf{Q}_i &\triangleq \mathbf{V}_i \mathbf{P}_i \mathbf{V}_i^H \\ &= \sum_{j=1}^{N_i} \mathbf{v}_{ij} \mathbf{v}_{ij}^H P_{ij}, \end{aligned} \tag{A.2}$$

where the jth column of $\mathbf{U}_i$ and $\mathbf{V}_i$ are denoted as $\mathbf{u}_{ij}$ and $\mathbf{v}_{ij}$, while the jth diagonal elements of $\mathbf{\Lambda}_i$ and $\mathbf{P}_i$ are denoted by Λ_{ij} and P_{ij} respectively. Hence, Λ_{ij} and P_{ij} are the power allocations to the jth stream of the ith user in the BRC and MARC respectively.

To prove the MARC-BRC duality, the following three steps are followed:

- With a given encoding order and relaying matrix (e.g., $\mathbf{D}^{\mathbf{B}} = \mathbf{D}$) in the BRC, we find the achievable rates with transmit precoding vector $\mathbf{u}_{ij}$ and receive combining vector $\mathbf{v}_{ij}^H$ for the jth stream of User i.

- With reverse decoding order (to the BRC encoding order), we then find the achievable rates for the dual MARC with relaying matrix (e.g., $\mathbf{D}^{\mathbf{M}} = c\mathbf{D}^H$, transmit precoding vector $\mathbf{v}_{ij}$, and receive combining vector $\mathbf{u}_{ij}^H$ for the jth stream of User i.

- Finally, we find the conditions for which both the BRC and MARC networks expend equal powers (i.e., $\Delta P \triangleq P_{total}^{\mathbf{M}} - P_{total}^{\mathbf{B}} = 0$) and at same time achieve the same rate pairs ($R_{sum}^{\mathbf{B}} = R_{sum}^{\mathbf{M}}$), the solution of which gives the MARC-BRC duality result.

Let denote the ranks of the input covariance matrices for the BRC and MARC as $r_i \triangleq \text{rank}\,\{\mathbf{\Sigma}_i\}$ and $\bar{r}_i \triangleq \text{rank}\,\{\mathbf{Q}_i\}$, with $r_i \leq \bar{r}_i \leq N_i$. This implies that zero power is allocated to user's sub-streams with indices beyond the rank of the input covariance matrices. In general, our approach follows the approach of [135], as outlined below.

For the BRC, the transmitted signal to user i can be re-written as

$$\mathbf{x}_i = \sum_{j=1}^{N_i} \mathbf{u}_{ij} \lambda_{ij}^{1/2} d_{ij}, \tag{A.3}$$

where $d_{ij} \sim \mathcal{CN}(0,1)$ is the data transmitted to the jth stream of User i.

From the BRC relation (3.4), in the first transmission phase we obtain

$$\mathbf{H}\mathbf{x}_i = \sum_{j=1}^{N_i} \mathbf{H}\mathbf{u}_{ij} \lambda_{ij}^{1/2} d_{ij}. \tag{A.4}$$

$\mathbf{H}\mathbf{u}_{ij}$ can be seen as the equivalent channel for the jth stream of User i in the first phase of the BRC. Hence, $\{\mathbf{H}\mathbf{u}_{ij}\}^H = \mathbf{u}_{ij}^H \mathbf{H}^H$ becomes the corresponding channel for the jth stream of User i in the second phase for the dual MARC.

Similarly, for the dual MARC, we have

$$\mathbf{G}_i^H \mathbf{u}_i = \sum_{j=1}^{N_i} \mathbf{G}_i^H \mathbf{v}_{ij} P_{ij}^{1/2} \bar{d}_{ij}, \tag{A.5}$$

where $\bar{d}_{ij} \sim \mathcal{CN}(0,1)$ is the data transmitted to the jth stream of User i. Similarly, $\mathbf{G}_i^H \mathbf{v}_{ij}$ can be seen as the equivalent channel for the jth stream of User i in the first

phase for the MARC, such that $\mathbf{v}_{ij}^H \mathbf{G}_i$ is the corresponding channel for the jth stream of User i in the second phase for the BRC.

As noted in Section 3.2, the optimal capacity-achieving strategies for the MAC and BC require ordering of users. Moreover, for the MAB-BC duality to hold, the decoding order in the MAC should the reverse encoding order in the BC. Hence, adopting an increasing encoding order $\pi\{i\} = \{1, 2, ...K\}$ for the BRC and decreasing decoding order $\pi'\{i\} = \{K, K-1..., 2, 1\}$ for the dual MARC, the received signals by the jth stream of User i (i.e., $y_{ij}^{\mathbf{B}}$ for the BRC and $y_{ij}^{\mathbf{M}}$ for the MARC) analogous to (3.4) and (3.8) are given by

$$\begin{aligned} y_{ij}^B &= \mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{H}\mathbf{u}_{ij} \lambda_{ij}^{1/2} d_{ij} + \sum_{m=j+1}^{N_i} \mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{H}\mathbf{u}_{im} \lambda_{im}^{1/2} d_{im} \\ &\quad + \sum_{t=i+1}^{K} \sum_{m=1}^{N_t} \mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{H}\mathbf{u}_{tm} \lambda_{tm}^{1/2} d_{tm} + \mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{n}_R + n_{ij}, \qquad \text{(A.6)} \\ y_{ij}^M &= \sum_{t=1}^{i-1} \sum_{m=1}^{N_t} c\mathbf{u}_{ij}^H \mathbf{H}^H \mathbf{D}^H \mathbf{G}_t^H \mathbf{v}_{tm} P_{tm}^{1/2} \bar{d}_{tm} + \sum_{m=1}^{j-1} c\mathbf{u}_{ij}^H \mathbf{H}^H \mathbf{D}^H \mathbf{G}_i^H \mathbf{v}_{im} P_{im}^{1/2} \bar{d}_{im} \\ &\quad + c\mathbf{u}_{ij}^H \mathbf{H}^H \mathbf{D}^H \mathbf{G}_i^H \mathbf{v}_{ij} P_{ij}^{1/2} \bar{d}_{ij} + c\mathbf{u}_{ij}^H \mathbf{H}^H \mathbf{D}^H \mathbf{n}_R + \bar{n}_{ij}, \qquad \text{(A.7)} \end{aligned}$$

where n_{ij},$\bar{n}_{ij} \sim \mathcal{CN}(0, 1)$ are the corresponding received noises for the BRC and MARC, respectively.

Consequently, the SNR of the jth stream of User i in the BRC ($\rho_{ij}^{\mathbf{B}}$) and MARC ($\rho_{ij}^{\mathbf{M}}$) are given by

$$\rho_{ij}^{\mathbf{B}} = \frac{\left|\mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{H}\mathbf{u}_{ij}\right|^2 \lambda_{ij}}{\sum_{m=j+1}^{N_i} \left|\mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{H}\mathbf{u}_{im}\right|^2 \lambda_{im} + \sum_{t=i+1}^{K} \sum_{m=1}^{N_t} \left|\mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\mathbf{H}\mathbf{u}_{tm}\right|^2 \lambda_{tm} + Tr\left\{\mathbf{D}^H \mathbf{G}_i^H \mathbf{v}_{ij} \mathbf{v}_{ij}^H \mathbf{G}_i \mathbf{D}\right\} + 1} \qquad \text{(A.8)}$$

$$\rho_{ij}^{\mathbf{M}} = \frac{\left|\mathbf{u}_{ij}^H (\mathbf{D}\mathbf{H})^H \mathbf{G}_i^H \mathbf{v}_{ij}\right|^2 P_{ij}}{\sum_{m=1}^{j-1} \left|\mathbf{u}_{ij}^H (\mathbf{D}\mathbf{H})^H \mathbf{G}_i^H \mathbf{v}_{im}\right|^2 P_{im} + \sum_{t=1}^{i-1} \sum_{m=1}^{N_t} \left|\mathbf{u}_{ij}^H (\mathbf{D}\mathbf{H})^H \mathbf{G}_t^H \mathbf{v}_{tm}\right|^2 P_{tm} + Tr\left\{\mathbf{D}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^H (\mathbf{D}\mathbf{H})^H\right\} + \frac{1}{c^2}}. \qquad \text{(A.9)}$$

Thus, the achievable sum rates for the BRC and MARC then respectively become

$$R_{sum}^{\mathbf{B}} = \sum_{i=1}^{K} \sum_{j=1}^{N_i} \left(1 + \rho_{ij}^B\right), \qquad \text{(A.10)}$$

$$R_{sum}^{\mathbf{M}} = \sum_{i=1}^{K}\sum_{j=1}^{N_i}\left(1+\rho_{ij}^{M}\right). \tag{A.11}$$

For both the BRC and MARC to achieve the same sum rate, the SNRs of the corresponding spatial streams must be equal. That is, $\rho_{ij}^{M} = \rho_{ij}^{B}$. Hence, equating (A.8) and (A.9), the power allocation for the jth stream of User i in the MARC that achieves the same rate in the BRC is given by

$$P_{ij} = \lambda_{ij} \times \tag{A.12}$$
$$\frac{\sum_{m=1}^{j-1}\left|\mathbf{u}_{ij}^{H}\mathbf{H}^{H}\mathbf{D}^{H}\mathbf{G}_{i}^{H}\mathbf{v}_{im}\right|^{2}P_{im} + \sum_{t=1}^{i-1}\sum_{m=1}^{N_t}\left|\mathbf{u}_{ij}^{H}\mathbf{H}^{H}\mathbf{D}^{H}\mathbf{G}_{t}^{H}\mathbf{v}_{tm}\right|^{2}P_{tm} + Tr\{\mathbf{D}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^{H}\mathbf{H}^{H}\mathbf{D}^{H}\} + \frac{1}{c^2}}{\sum_{m=j+1}^{N_i}\left|\mathbf{v}_{ij}^{H}\mathbf{G}_{i}\mathbf{D}\mathbf{H}\mathbf{u}_{im}\right|^{2}\lambda_{im} + \sum_{t=i+1}^{K}\sum_{m=1}^{N_t}\left|\mathbf{v}_{ij}^{H}\mathbf{G}_{i}\mathbf{D}\mathbf{H}\mathbf{u}_{tm}\right|^{2}\lambda_{tm} + Tr\{\mathbf{D}^{H}\mathbf{G}_{i}^{H}\mathbf{v}_{ij}\mathbf{v}_{ij}^{H}\mathbf{G}_{i}\mathbf{D}\} + 1}.$$

Now, let us consider the power constraints in the networks. For the BRC, the power constraints at the BS and RS are

$$\sum_{i=1}^{K} Tr(\mathbf{\Sigma}_i) = \sum_{i=1}^{K}\sum_{j=1}^{N_i}\lambda_{ij} = P_T^{\mathbf{B}} \tag{A.13}$$
$$Tr\left\{\mathbf{D}\left(\sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^{H}\mathbf{H}^{H}\lambda_{ij} + \mathbf{I}\right)\mathbf{D}^{H}\right\} = P_R^{\mathbf{B}}.$$

The total network power is given by

$$\begin{aligned} P_{total}^{\mathbf{B}} &\triangleq P_T^{\mathbf{B}} + P_R^{\mathbf{B}} \\ &= \sum_{i=1}^{K}\sum_{j=1}^{N_i}\lambda_{ij} + Tr\left(\mathbf{D}\left(\sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^{H}\mathbf{H}^{H}\lambda_{ij} + \mathbf{I}\right)\mathbf{D}^{H}\right). \end{aligned} \tag{A.14}$$

Similarly, for the MARC, the power constraints across the users and RS are given by

$$\sum_{i=1}^{K} Tr(\mathbf{Q}_i) = \sum_{i=1}^{K}\sum_{j=1}^{N_i} P_{ij} = P_T^{\mathbf{M}} \tag{A.15}$$
$$c^2 Tr\left\{\mathbf{D}^{H}\left(\sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{G}_{i}^{H}\mathbf{v}_{ij}\mathbf{v}_{ij}^{H}\mathbf{G}_{i}P_{ij} + \mathbf{I}\right)\mathbf{D}\right\} = P_R^{\mathbf{M}}$$

Hence, the total network power becomes

$$P_{total}^{\mathbf{M}} \triangleq P_T^{\mathbf{M}} + P_R^{\mathbf{M}} \tag{A.16}$$

$$
\begin{aligned}
&= \sum_{i=1}^{K}\sum_{j=1}^{N_i} P_{ij} + c^2 Tr\left(\mathbf{D}^H\left(\sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{G}_i^H\mathbf{v}_{ij}\mathbf{v}_{ij}^H\mathbf{G}_i P_{ij} + \mathbf{I}\right)\mathbf{D}\right) \\
&= c^2\sum_{i=1}^{K}\sum_{j=1}^{N_i} P_{ij}\left(1 + Tr\left(\mathbf{D}^H\mathbf{G}_i^H\mathbf{v}_{ij}\mathbf{v}_{ij}^H\mathbf{G}_i\mathbf{D}\right)\right) + c^2 Tr\{\mathbf{D}^H\mathbf{D}\} \\
&\quad +(1-c^2)\sum_{i=1}^{K}\sum_{j=1}^{N_i} P_{ij}.
\end{aligned}
$$

Using (A.12) and employing the relation

$$
\begin{aligned}
&\sum_{i=1}^{K}\sum_{j=1}^{N_i}\left(\sum_{m=1}^{j-1}\left|\mathbf{u}_{ij}^H\mathbf{H}^H\mathbf{D}^H\mathbf{G}_i^H\mathbf{v}_{im}\right|^2 P_{im}\lambda_{ij} + \sum_{t=1}^{i-1}\sum_{m=1}^{N_t}\left|\mathbf{u}_{ij}^H\mathbf{H}^H\mathbf{D}^H\mathbf{G}_t^H\mathbf{v}_{tm}\right|^2 P_{tm}\lambda_{ij}\right) \\
&= \sum_{i=1}^{K}\sum_{j=1}^{N_i}\left(\sum_{m=j+1}^{N_i}\left|\mathbf{v}_{ij}^H\mathbf{G}_i\mathbf{D}\mathbf{H}\mathbf{u}_{im}\right|^2\lambda_{im}P_{ij} + \sum_{t=i+1}^{K}\sum_{m=1}^{N_t}\left|\mathbf{v}_{ij}^H\mathbf{G}_i\mathbf{D}\mathbf{H}\mathbf{u}_{tm}\right|^2\lambda_{tm}P_{ij}\right),
\end{aligned}
\tag{A.17}
$$

the total MARC network power can be re-written as

$$
P_{total}^{\mathbf{M}} = \sum_{i=1}^{K}\sum_{j=1}^{N_i}\lambda_{ij}\left(1 + c^2 Tr\{\mathbf{D}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^H\mathbf{H}^H\mathbf{D}^H\}\right) + (1-c^2)\sum_{i=1}^{K}\sum_{j=1}^{N_i} P_{ij} + c^2 Tr\{\mathbf{D}^H\mathbf{D}\}. \tag{A.18}
$$

We now find the conditions for which both the MARC and BRC networks expend equal powers (i.e., $\Delta P \triangleq P_{total}^{\mathbf{M}} - P_{total}^{\mathbf{B}} = 0$) as follows;

$$
\begin{aligned}
\Delta P &\triangleq P_{total}^{\mathbf{M}} - P_{total}^{\mathbf{B}} && \text{(A.19)} \\
&= (c^2-1)Tr\left(\mathbf{D}\left(\mathbf{I} + \sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^H\mathbf{H}^H\lambda_{ij}\right)\mathbf{D}^H\right) && \text{(A.20)} \\
&\quad -(c^2-1)\sum_{i=1}^{K}\sum_{j=1}^{N_i} P_{ij} \\
&= (c^2-1)(P_R^{\mathbf{B}} - P_T^{\mathbf{M}}). && \text{(A.21)}
\end{aligned}
$$

Hence, for $\Delta P = 0$, the only solutions are

- $c = 1$
- $P_R^{\mathbf{B}} = P_T^{\mathbf{M}} \Leftrightarrow P_T^{\mathbf{B}} = P_R^{\mathbf{M}}$, and from (A.15),

$$
c^2 = \frac{P_R^{\mathbf{M}}}{Tr\left\{\mathbf{D}^H\left(\sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{G}_i^H\mathbf{v}_{ij}\mathbf{v}_{ij}^H\mathbf{G}_i P_{ij} + \mathbf{I}\right)\mathbf{D}\right\}}. \tag{A.22}
$$

Similarly, with the reverse RS beamforming matrices (i.e., $\mathbf{D^B} = c\mathbf{D}^H$ and $\mathbf{D^M} = \mathbf{D}$), it can be shown that

$$\begin{aligned} \Delta P &\triangleq P^{\mathbf{M}}_{total} - P^{\mathbf{B}}_{total} \\ &= (c^2-1)(P^{\mathbf{M}}_R - P^{\mathbf{B}}_T). \end{aligned} \tag{A.23}$$

Again, for $\Delta P = 0$, the only solutions are

- $c = 1$
- $P^{\mathbf{M}}_R = P^{\mathbf{B}}_T \Leftrightarrow P^{\mathbf{M}}_T = P^{\mathbf{B}}_R$, with

$$c^2 = \frac{P^{\mathbf{B}}_R}{Tr\left\{\mathbf{D}^H\left(\sum_{i=1}^{K}\sum_{j=1}^{N_i}\mathbf{H}\mathbf{u}_{ij}\mathbf{u}_{ij}^H\mathbf{H}^H\lambda_{ij} + \mathbf{I}\right)\mathbf{D}\right\}}. \tag{A.24}$$

Appendix B

Proof of Theorem 3.2

B.1 Mapping from the MARC to the BRC Input Covariance Matrices

In this Section, we prove the MARC-BRC covariance matrices transformation given by Theorem 3.2 of Section 3.4. In the following Sub-section, we first give some background on the concepts used in our derivation and then proceed with the derivation. In general, our derivation follows the approach of [49] for conventional MAC and BC networks, that is, without relays.

B.1.1 Background: Effective and Flipped Channels

Consider a single user MIMO system A with channel matrix $\mathbf{H}$, additive Gaussian noise with covariance matrix $\mathbf{R}_N$, and independent additive Gaussian interference with covariance matrix $\mathbf{R}_I$. The *effective* channel of this system is given by $(\mathbf{R}_N + \mathbf{R}_I)^{-\frac{1}{2}}\mathbf{H}$ [49]. Another system B with channel matrix $(\mathbf{R}_N + \mathbf{R}_I)^{-\frac{1}{2}}\mathbf{H}$ (i.e., the *effective* channel of B), additive white Gaussian noise with unit variance elements, and no interference achieves the same rate as system A [49].

On the other hand, system θ_1 with *effective* channel $\mathbf{G}$ achieves the same capacity as system θ_2 with *effective* channel $\mathbf{G}^H$. $\mathbf{G}^H$ is known as the *flipped* channel [49] [15]. This implies that for every transmit covariance matrix $\mathbf{\Sigma}$ for θ_1, there exists another covariance matrix $\bar{\mathbf{\Sigma}}$ for θ_2 with $Tr(\bar{\mathbf{\Sigma}}) \leq Tr(\mathbf{\Sigma})$ such that the rate achieved by $\mathbf{\Sigma}$ in θ_1 is the same as that achieved by $\bar{\mathbf{\Sigma}}$ in θ_2. That is, $\log|\mathbf{I}+\mathbf{G}\mathbf{\Sigma}\mathbf{G}^H| = \log|\mathbf{I}+\mathbf{G}^H\bar{\mathbf{\Sigma}}\mathbf{G}|$. It is further proved in [49] that $\bar{\mathbf{\Sigma}} = \mathbf{U}\mathbf{V}^H\mathbf{\Sigma}\mathbf{V}\mathbf{U}^H$ satisfies the above condition with the SVD (more precisely, the economy size decomposition) of the *effective* channel given by $\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$. That is, $\mathbf{\Lambda}$ is square and diagonal.

B.1.2 MARC-BRC Covariance Matrix Derivation

Similar to the approach for conventional MU-MIMO channels [49], we now derive the MARC-BRC covariance matrix transformation. For MU-MIMO AF relay networks, unlike [49], the input covariance matrix transformation also depends on the relationship between the relay beamforming matrices $\mathbf{D^M}$ in the MARC and $\mathbf{D^B}$ in the BRC.

Substituting the RS beamforming matrices $\mathbf{D^B} = \mathbf{D}$ for the BRC and $\mathbf{D^M} = c\mathbf{D}^H$ for the MARC into (3.31) and (3.32) respectively, the effective covariance matrices of noise-plus-interference seen by User i and denoted by $\mathbf{B}_i$ for the BRC and $\mathbf{M}_i$ for the MARC are given by

$$\mathbf{B}_i \triangleq \mathbf{I} + \mathbf{G}_i\mathbf{D}\left(\sum_{j=1}^{i-1}\mathbf{H}\mathbf{\Sigma}_j\mathbf{H}^H + \mathbf{I}\right)\mathbf{D}^H\mathbf{G}_i^H, \tag{B.1}$$

$$\mathbf{M}_i \triangleq \mathbf{I} + c^2\mathbf{H}^H\mathbf{D}^H\left(\sum_{j=i+1}^{K}\mathbf{G}_j^H\mathbf{Q}_j\mathbf{G}_j + \mathbf{I}\right)\mathbf{D}\mathbf{H}. \tag{B.2}$$

Then, User i's achievable rate for the MARC given by (3.14) can be re-written as

$$R_i^{\mathbf{M}} = \frac{1}{2}\log|\mathbf{I} + c^2\mathbf{M}_i^{-1}\mathbf{H}^H\mathbf{D}^H\mathbf{G}_i^H\mathbf{Q}_i\mathbf{G}_i\mathbf{D}\mathbf{H}|, \tag{B.3}$$

where we have used matrix identities: $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ and $|\mathbf{A}|^{-1} = |\mathbf{A}^{-1}|$.

Taking the square root of $\mathbf{M}_i^{-1}$, using the identity $|\mathbf{I}+\mathbf{AB}| = |\mathbf{I}+\mathbf{BA}|$, introducing $\mathbf{B}_i^{-\frac{1}{2}}\,\mathbf{B}_i^{\frac{1}{2}} = \mathbf{I}$, and then applying the *effective-flipped* channel relation given earlier, we obtain

$$\begin{aligned} R_i^{\mathbf{M}} &= \frac{1}{2}\log_2\left|\mathbf{I} + c^2\bar{\mathbf{H}}_i\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}\bar{\mathbf{H}}_i^H\right| \\ &= \frac{1}{2}\log_2\left|\mathbf{I} + c^2\bar{\mathbf{H}}_i^H\overline{\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}}\bar{\mathbf{H}}_i\right|, \end{aligned} \tag{B.4}$$

with $Tr\left(\overline{\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}}\right) \leq Tr\left(\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}\right)$. The economy size decomposition of User i's *effective* channel matrix $\bar{\mathbf{H}}_i \triangleq \mathbf{M}_i^{-\frac{1}{2}}\mathbf{H}^H\mathbf{D}^H\mathbf{G}_i^H\mathbf{B}_i^{-\frac{1}{2}}$ is given by

$$\mathbf{M}_i^{-\frac{1}{2}}\mathbf{H}^H\mathbf{D}^H\mathbf{G}_i^H\mathbf{B}_i^{-\frac{1}{2}} = \bar{\mathbf{U}}_i\bar{\mathbf{\Lambda}}_i\bar{\mathbf{V}}_i^H. \tag{B.5}$$

Similarly to (B.3) for the MARC, User i's achievable rate for the BRC given by (3.6) can be re-written as

$$R_i^{\mathbf{B}} = \frac{1}{2}\log_2\left|\mathbf{I} + \mathbf{B}_i^{-\frac{1}{2}}\mathbf{G}_i\mathbf{D}\mathbf{H}\mathbf{\Sigma}_i\mathbf{H}^H\mathbf{D}^H\mathbf{G}_i^H\mathbf{B}_i^{-\frac{1}{2}}\right|. \tag{B.6}$$

Equating User i's rates in the BRC and MARC given by (B.4) and (B.6) respectively, the BRC covariance $\mathbf{\Sigma}_i$ can be obtained from the MARC covariance $\mathbf{Q}_i$ using

$$\begin{aligned}\mathbf{\Sigma}_i &= c^2\mathbf{M}_i^{-\frac{1}{2}}\overline{\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}}\mathbf{M}_i^{-\frac{1}{2}} \\ &= c^2\mathbf{M}_i^{-\frac{1}{2}}\bar{\mathbf{U}}_i\bar{\mathbf{V}}_i^H\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}\bar{\mathbf{V}}_i\bar{\mathbf{U}}_i^H\mathbf{M}_i^{-\frac{1}{2}}.\end{aligned} \tag{B.7}$$

We have proved the MARC-BRC covariance matrix transformation for the case where the RS beamforming matrices for the BRC and MARC are given by

$$\mathbf{D}^{\mathbf{B}} = \mathbf{D}, \mathbf{D}^{\mathbf{M}} = c\mathbf{D}^H. \tag{B.8}$$

Alternatively, by reversing the relaying matrices, that is,

$$\mathbf{D}^{\mathbf{B}} = c\mathbf{D}^H, \mathbf{D}^{\mathbf{M}} = \mathbf{D}, \tag{B.9}$$

and following similar approach, User i's covariance matrix in the BRC $\mathbf{\Sigma}_i$ can be obtained from User i's covariance matrix in the MARC $\mathbf{Q}_i$ by using

$$\mathbf{\Sigma}_i = \frac{1}{c^2}\mathbf{M}_i^{-\frac{1}{2}}\mathbf{U}_i\mathbf{V}_i^H\mathbf{B}_i^{\frac{1}{2}}\mathbf{Q}_i\mathbf{B}_i^{\frac{1}{2}}\mathbf{V}_i\mathbf{U}_i^H\mathbf{M}_i^{-\frac{1}{2}}. \tag{B.10}$$

where the SVD (or more accurately economy size decomposition) of User i's *effective* channel matrix is given by

$$\mathbf{M}_i^{-\frac{1}{2}}\mathbf{H}^H\mathbf{D}\mathbf{G}_i^H\mathbf{B}_i^{-\frac{1}{2}} = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H. \tag{B.11}$$

Again, $\mathbf{\Lambda}_i$ is a diagonal and square matrix.