School of Electrical Engineering and Computing
Department of Electrical and Computer Engineering

Signal Processing Algorithms for Multiuser MIMO Relay Communication Systems

Muhammad Ruhul Amin Khandaker

This thesis is presented for the degree of
Doctor of Philosophy
of
Curtin University

December 2012

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature: .....

Date: .....

To

my wife Rumi and my parents

who have always been the most supportive and caring towards me.

Acknowledgements

First and foremost, I especially like to express my gratitude to my thesis supervisor Dr. Yue Rong for his constant support and encouragement during the entire period of my PhD studies. His invaluable guidance and timely correspondence directed me to reach the goal. He was constantly patient and caring in his instructive and research support, which guided me to pursue the right path.

I express my appreciations to all members of my thesis committee for helping me to complete the thesis. The valuable comments and suggestions from Dr. Roy Howard, Chairperson of my thesis committee, Dr. Yee-Hong Leung, associate supervisor of my thesis, and Prof. Sven Nordholm, Head of the Department, especially during the group seminars, helped me improve a lot. Also, the unstinting support I got from Prof. Syed Islam, former Head of the Department, helped me overcome any difficult time.

Next, I would like to thank my research peers at the Communication Technology and Signal Processing (CSP) group of the Department for the good times and the friendly help they have given me during my time here.

I am grateful to Curtin University for the financial and administrative support they have provided me during the course of my doctoral study.

I would like to acknowledge my wife, Rumi, who deserves a doctorate herself for the patience, sacrifice, and care she has shown during the entire period of my PhD study. I must not forget the uncountable moments she spent after-hours with me in the laboratory. I would also like to thank my only sister and brothers for their generous understanding, encouragement and prayer.

Most of all, I am indebted to my unconditionally loving parents, whose moral support in my research endeavours has made this dissertation possible. I would like to finish off by thanking my Creator for making my life bountiful.

Abstract

The increasing demand for mobile applications such as streaming media, software updates, and location-based services involving group communications has prompted the need for wireless communication technologies that can support reliable high data rates. However, wireless channel is subject to signal fading that severely degrades the system spectral efficiency. By exploiting the spatial diversity, multiple-input multiple-output (MIMO) techniques can provide both theoretically attractive and technically practical solutions to combat channel fading. Moreover, in the case of long source-destination distance, single or multiple MIMO relay node(s) is necessary to combat the pathloss and/or shadowing effects of wireless channel and relay signals from the source to the destination.

In this thesis, we focus on multiuser MIMO relay systems. We first present joint source, relay, and receiver optimization algorithms for the uplink system based on the minimum mean-squared error (MMSE) criterion subjecting to individual power constraints at the source and the relay nodes. The proposed algorithms outperform the existing techniques in terms of both MSE and bit-error-rate (BER). Next, in the downlink system, we consider multicasting multiple data streams among a group of users with the aid of a relay node, where all the nodes are equipped with multiple antennas. The downlink system performance is optimized subjecting to both power constraints at the source and the relay nodes and quality-of-service (QoS) constraints at the receivers.

Then we present the duality between the uplink and the downlink of a multi-hop MIMO relay system. Based on this duality, we propose an optimal design of the source precoding matrix and relay amplifying matrices for multi-hop MIMO relay system with a nonlinear dirty paper coding (DPC)-based transmitter at the source node. The proposed nonlinear transmitter

algorithm outperforms the existing decision feedback equalizer (DFE)-based receiver schemes.

Multiuser MIMO relaying is then considered in an interference system where a group of transmitters communicate simultaneously with their desired destination nodes with the aid of multiple relay nodes, all equipped with multiple antennas. Transmit-relay-receive beamforming technique is exploited to minimize the total source and relay transmit power in conjunction with transmit power control such that a minimum QoS threshold is maintained at each receiver. The proposed scheme generalizes the existing single-hop MIMO interference systems and the single-antenna, dual-hop interference relay systems to the dual-hop interference MIMO relay systems with any number of source, relay, and destination nodes, all equipped with multiple antennas.

The above algorithms are developed assuming that the instantaneous channel state information (CSI) knowledge of both the source-relay link and the relay-destination link is available at the scheduler. However, in practical relay communication systems, the instantaneous CSI is unknown, and therefore, has to be estimated. Hence, we finally propose a bandwidth efficient MIMO channel estimation algorithm that provides the destination node with full knowledge of all channel matrices involved in a dual-hop MIMO communication. The proposed approach attains smaller channel estimation error and is applicable for both one-way and two-way MIMO relay systems.

Author's Note

Parts of this thesis and concepts from it have been previously published in the following journal and/or conference papers.

Journal Papers

[1] M. R. A. Khandaker and Y. Rong, "Precoding Design for MIMO Relay Multicasting", *IEEE Trans. Signal Process.*, submitted, Jul. 2012.

[2] M. R. A. Khandaker and Y. Rong, "Joint Transceiver Optimization for Multiuser MIMO Relay Communication Systems", *IEEE Trans. Signal Process.*, to appear, 2012.

[3] M. R. A. Khandaker and Y. Rong, "Interference MIMO Relay Channel: Joint Power Control and Transceiver-Relay Beamforming", *IEEE Trans. Signal Process.*, to appear, 2012.

[4] Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel Estimation of Dual-Hop MIMO Relay Systems via Parallel Factor Analysis", *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2224-2233, Jun. 2012.

[5] Y. Rong and M. R. A. Khandaker, "On Uplink-Downlink Duality of Multi-Hop MIMO Relay Channel", *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1923-1931, Jun. 2011.

Conference Papers

[1] M. R. A. Khandaker and Y. Rong, "Multicasting MIMO Relay Optimization Based on Min-Max MSE Criterion", *Proc. IEEE ICCS*, Singapore, Nov. 21-23, 2012.

[2] M. R. A. Khandaker and Y. Rong, "Joint Power Control and Beamforming for Peer-to-Peer MIMO Relay Systems", in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Nov. 9-11, 2011.

[3] M. R. A. Khandaker and Y. Rong, "Joint Power Control and Beamforming for Interference MIMO Relay Channel", in *Proc. 17th Asia-Pacific Conference on Communications (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

[4] Y. Rong and M. R. A. Khandaker, "Channel Estimation of Dual-Hop MIMO Relay Systems Using Parallel Factor Analysis", in *Proc. 17th Asia-Pacific Conference on Communications (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

[5] M. R. A. Khandaker and Y. Rong, "Joint Source and Relay Optimization for Multiuser MIMO Relay Communication Systems", in *Proc. 4th Int. Conf. Signal Processing Communication Systems (ICSPCS'2010)*, Gold Coast, Australia, Dec. 13-15, 2010.

[6] M. R. A. Khandaker and Y. Rong, "Dirty Paper Coding Based Optimal MIMO Relay Communications", in *Proc. 16th Asia-Pacific Conference on Communications (APCC'2010)*, Auckland, New Zealand, Nov. 1-3, 2010, pp. 328-333. (**Best Paper Award**)

[7] M. R. A. Khandaker and Y. Rong, "Performance Measure of Multi-User Detection Algorithms for MIMO Relay Network", in *Proc. 10th Postgraduate Electrical Engineerig and Computing Symposium (PEECS'2009)*, Perth, Australia, Oct. 1, 2009.

Contents

List of Figures

List of Tables

List of Acronyms

AF	amplify-and-forward
AWGN	additive white Gaussian noise
BALS	bilinear alternating least-squares
BC	broadcast channel
BER	bit-error-rate
BS	base station
CCI	cochannel interference
CSI	channel state information
DFE	decision feedback equalizer
DFT	discrete Fourier transform
DPC	dirty paper coding
EVD	eigenvalue decomposition
HD	high definition
i.i.d.	independent and identically distributed
LMMSE	linear minimal mean-squared error
LOS	line-of-sight
LS	least-squares
LTE	long-term evolution

MAC	multiaccess channel
MI	mutual information
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
MMSE	minimal mean-squared error
MSE	mean-squared error
MU-MIMO	multiuser MIMO
MUI	multiple user interference
NAF	naive amplify-and-forward
NMSE	normalized mean-squared error
NPC	no power control
P2P	point-to-point
PARAFAC	parallel factor
PMF	pseudo match-and-forward
PSD	positive semi-definite
PZF	partial zero-forcing
QCQP	quadratically constrained quadratic programming
QoS	quality-of-service
QPSK	quadrature phase-shift keying
RFID	radio-frequency identification
SDP	semidefinite programming
SDR	semidefinite relaxation
SIC	successive interference cancellation

SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
SOCP	second-order cone programming
SU-MIMO	single-user MIMO
SVD	singular value decomposition
TST	two-stage training
UB	upper-bound
WLS	weighted least-squares

Chapter 1

Introduction

In next generation wireless systems, it is expected that multiple users equipped with multiple antennas will communicate simultaneously with the base station (BS) equipped with multiple receive antennas [1–4]. This thesis aims at developing advanced signal processing algorithms for multiuser multiple-input multiple-output (MIMO) relay communication systems. In this introductory chapter, we briefly present necessary background on multiuser MIMO relay systems and overview the contributions of the thesis.

1.1 Overview of MIMO Communication Systems

Due to the emerging demand on new multimedia applications, next-generation wireless systems are expected to support higher data rates compared to the existing systems. The bandwidth limitation of wireless communication prompts the need for spectrally efficient methods of communication. MIMO systems can provide an effective means to increase system spectral efficiency through exploiting multiple antennas at transmitting and receiving ends. By spatially multiplexing several data streams onto the MIMO channel, the system can provide an additional *degree-of-freedom* which leads to an increase in the channel capacity [1, 2, 5–9]. In a rich scattering environment, the capacity of such a MIMO channel with N transmit and N receive antennas is N times that of a conventional single-antenna channel [1]. On the other hand, the random scattering in wireless channels results in multipath fading which is traditionally a pitfall of wireless transmission. A key feature of MIMO systems is the ability to turn multipath propagation into a benefit for the user [5, 10–12]. The performance improvements

resulting from the use of MIMO systems are due to the following unique features of MIMO configuration [1, 8].

- Array gain that is achievable through processing at the transmitter and the receiver and results in an increase in average receive signal-to-noise ratio (SNR) due to a coherent combining effect. Transmit and/or receive array gain requires channel knowledge at the transmitter and receiver, respectively, and depends on the number of transmit and receive antennas.

- Diversity gain that is favorable to mitigating the fading effects of wireless channels. Diversity techniques rely on transmitting the signal over multiple independently fading paths in time, frequency, or space. Spatial diversity is preferred over time/frequency diversity as it does not incur any cost in transmission time or bandwidth. A MIMO system with N_{T} transmit and N_{R} receive antennas can achieve $N_{\mathrm{T}}N_{\mathrm{R}}$th-order diversity.

- Spatial multiplexing gain that is achievable by transmitting independent data streams from different antennas. Under rich scattered channel conditions, the receiver can separate the different streams, and the capacity scales linearly with $N_{\min}$ where $N_{\min}$ is the minimum of the number of transmit antennas N_{T} and the number of receive antennas N_{R} [1, 10].

- When multiple antennas are used, spatial filters preserve the signals coming from a certain spatial location, while suppressing signals from other spatial locations (interferences). Therefore, MIMO transceivers can separate signals which differ in spatial signatures, just as a conventional filter which can separate signals occupying different frequency band. Interference suppression can also be implemented at the transmitter, where the goal is to minimize the interference power sent toward the cochannel users while delivering the signal to the desired user.

However, it is important to note that in general it may not be possible to exploit all the leverages of MIMO technology simultaneously due to conflicting demands on the spatial degrees of freedom between diversity gain and spatial multiplexing gain. The degree to which these conflicts are resolved depends upon the signaling scheme and transceiver design [8]. A tradeoff between the two gains can be an effective solution for practical applications [13].

MIMO technologies have evolved to become an inherent component in the next-generation wireless standards including the cellular systems, long-term evolution (LTE) systems, and the IEEE 802.xx family of standards 802.16e, 802.16j, 802.16m, and 802.11n [14]. Future enhancements to these standards will use MIMO techniques to achieve required data rates in the order of hundreds of Mbps and spectral efficiencies in the order of tens of bps/Hz. Potential MIMO applications include, but are not limited to, large files backup, high definition (HD) video streaming, online interactive gaming, home entertainment, radio-frequency identification (RFID), digital home and so on. MIMO systems can reliably connect video devices, computer networking devices, broadband connections, phone lines, music and storage devices, etc.

1.2 MIMO Relay Communication Systems

In the case of long transmitter-receiver distance, some intermediate nodes, which we term as relay nodes, are necessary to efficiently combat the pathloss of wireless channels. Relays can receive and forward data from source to destination. Since they implement a subset of BS functions, relays are a low cost and low complexity solution to meet the requirement of higher capacity, extended coverage and improved link reliability [15, 16]. In a cellular environment, relays can be deployed in areas with significant shadowing effects [17], such as inside buildings, mines and tunnels as well as areas far from the base station at the cell edge. Recently, relaying and cooperative communications are being considered for several industrial standards such as IEEE802.16j [18].

However, conventional relay techniques are mainly developed for networks with single antenna nodes. Incorporating relays in a MIMO network can significantly extend the coverage and improve the link reliability of the network [19–28] compared with point-to-point (P2P) MIMO systems. Both regenerative [19, 20] and non-regenerative [21–28] relaying strategies have been considered in the existing works. For the non-regenerative strategy, the relay nodes only amplify and retransmit their received signals. While in regenerative strategies, the relay nodes first decode the received signals, then re-encode and forward the signals to the destination node. Thus, for MIMO relay networks, the complexity of the non-regenerative strategy is much lower than that of the regenerative strategy, since decoding and encoding multiple data streams involves great computational efforts. Consequently, compared with regenerative relays, non-regenerative relays are more cost-effective, and have a less end-to-end delay, thus it has attracted

much research interest [21–28]. However, a hybrid relaying concept for MIMO relay channels has also been proposed [20], which combines the benefits of regenerative and non-regenerative relaying schemes.

1.3 Multiuser MIMO Communication Systems

In a large wireless network, multiple users may concurrently intend to communicate with the same base station. Thus in a practical wireless system, multiple user interference (MUI) inevitably occurs. MUI refers to the interference from unintended users to the user of interest. MUI is a major factor which limits the capacity and performance of wireless communication systems [29, 30]. While the MUI caused by any one user is generally small, as the number of interferers or their power increases, MUI becomes significant. MUI thus has to be taken into account in system design and schemes to mitigate its effect need to be developed. The conventional single-user detectors do not take into account the existence of MUI. In single-user detection algorithms, each user is detected separately without regard for other users. The MUI should be taken into consideration in the detection process which is termed as multiuser detection (also referred to as interference cancelation). In multiuser detectors, information about multiple users is used jointly to better detect each individual user's signal. The utilization of multiuser detection algorithms in MIMO systems has the potential to provide significant additional benefits for wireless communication systems [29, 31, 32].

Recently, multiuser MIMO (MU-MIMO) techniques are being extensively investigated [1, 3, 4, 31, 33–41] because of several key advantages over single-user MIMO (SU-MIMO) communications [31].

- MU-MIMO schemes allow for a significant improvement in multiple access capacity which is proportional to the minimum of the number of BS antennas and the sum of the number of users times the number of antennas per user.

- Geographical separation of the users can afford some means to combat antenna correlation in MIMO setup. Additionally, line-of-sight (LOS) propagation, which causes severe performance degradation in single-user spatial multiplexing schemes, is no longer a problem in MU-MIMO setting.

- MU-MIMO allows the spatial multiplexing gain at the BS to be obtained without the need for multiple antenna terminals.

- The reduction of MUI on the uplink may translate to some reduction in the required transmit power of the mobiles. Alternatively, the same transmit power may be used to extend the size of the coverage region.

Table 1.1: An overview of MIMO research topics with key references

	P2P systems	Relay systems
SU-MIMO	[1, 2, 5–13, 42]	[19–28, 43–47]
MU-MIMO	[1, 3, 4, 31, 33, 34]	Focus of this thesis

1.4 Thesis Overview and Contributions

In next generation wireless systems, multiple users equipped with multiple antennas will transmit simultaneously to the base station with multiple receive antennas and vice versa [3, 4]. However, in the case of long transmitter-receiver distance, relay node(s) is necessary to efficiently combat the pathloss of wireless channel. MIMO relays are particularly useful in extending the network coverage and improving the link reliability of the network. In this thesis, advanced signal processing algorithms for multiuser MIMO relay communication systems are presented considering non-regenerative relaying schemes (Table 1.1). The joint source and relay optimization problems for multiuser MIMO relay systems are highly nonconvex, in general. One of the major contributions of this thesis is transforming the nonconvex problems into suitable forms which can be efficiently solved using standard convex optimization techniques. In Chapter 2, we address the joint transceiver and relay design problem for uplink multiuser MIMO relay systems. Chapter 3 studies multicast beamforming problems for downlink multiuser MIMO relay systems. The uplink-downlink duality of multi-hop MIMO relay systems is investigated in Chapter 4 in order to design nonlinear dirty paper coding (DPC)-based transmitters from decision feedback equalizer (DFE)-based receivers. Chapter 5 studies the joint power control and beamforming problem for interference (multiuser) MIMO relay systems. In Chapter 6, we propose a novel channel estimation algorithm for dual-hop MIMO relay systems using the parallel factor (PARAFAC) analysis. Chapter 7 summarizes the thesis and gives the outlook to some interesting future works.

Chapter 2: Uplink Multiuser MIMO Relay Communication Systems
In this chapter, we develop two iterative methods to solve the highly nonconvex joint source, relay, and receiver optimization problem for uplink multiuser MIMO relay communication systems based on minimal mean-squared error (MMSE) criterion. In the first approach, we iteratively optimize the source, relay, and receiver matrices. In order to reduce the overall computational complexity, we propose a simplified algorithm in the second approach. Compared with the existing techniques for uplink multiuser MIMO relay systems, the proposed algorithms perform much better in terms of bit-error-rate (BER) and mean-squared error (MSE) of signal waveform estimation at the destination node.

Chapter 2 is based on the journal publication:

- M. R. A. Khandaker and Y. Rong, "Joint Transceiver Optimization for Multiuser MIMO Relay Communication Systems", *IEEE Trans. Signal Process.*, to appear, 2012.

and two conference publications:

- M. R. A. Khandaker and Y. Rong, "Joint Source and Relay Optimization for Multiuser MIMO Relay Communication Systems", in *Proc. 4th Int. Conf. Signal Processing Communication Systems (ICSPCS'2010)*, Gold Coast, Australia, Dec. 13-15, 2010.

- M. R. A. Khandaker and Y. Rong, "Performance Measure of Multi-User Detection Algorithms for MIMO Relay Network", in *Proc. 10th Postgraduate Electrical Engineerig and Computing Symposium (PEECS'2009)*, Perth, Australia, Oct. 1, 2009.

Chapter 3: Multicasting MIMO Relay Communication Systems
The increasing demand for mobile applications such as streaming media, software updates, and location-based services involving group communications has triggered the need for wireless multicasting technology. The broadcasting nature of the wireless channel makes it naturally suitable for multicasting applications, since a single transmission may be simultaneously received by a number of users. However, wireless channel is subject to signal fading. By exploiting the spatial diversity, multi-antenna techniques can be applied to combat channel fading [1]. Hence, in this chapter, we focus on the joint transmit and relay precoding design problems for multicasting multiple data streams

in the downlink multiuser MIMO relay systems based on two design criteria. In the first scheme, we aim at minimizing the maximal MSE of the signal waveform estimation among all receivers subjecting to power constraints at the transmitter and the relay node. In the second scheme, we propose a total source and relay transmission power minimization strategy subjecting to quality-of-service (QoS) constraints.

The material in Chapter 3 is based on the journal submission:

- M. R. A. Khandaker and Y. Rong, "Precoding Design for MIMO Relay Multicasting", *IEEE Trans. Signal Process.*, submitted, Jul. 2012.

and the conference publication:

- M. R. A. Khandaker and Y. Rong, "Multicasting MIMO Relay Optimization Based on Min-Max MSE Criterion", *Proc. IEEE ICCS*, Singapore, Nov. 21-23, 2012.

Chapter 4: Duality in Multi-Hop MIMO Relay Channel

There is an interesting duality between the uplink and the downlink in a MIMO system, and by exploiting this duality, one can map each receive architecture for the uplink to a corresponding transmit architecture for the downlink. We define duality as the achievement of identical signal-to-interference-plus-noise ratios (SINRs) at the uplink and the downlink systems with the same amount of total network transmission power.

Recently, the uplink-downlink duality of two-hop amplify-and-forward (AF) MIMO relay systems has been derived in [48]. We extend the uplink-downlink duality results in [48–51] to multi-hop AF-MIMO relay systems with any number of hops and any number of antennas at each node. In particular, we show that the duality can be achieved by two approaches. First, with the same total network transmission power constraint, one simply applies Hermitian transposed uplink relay amplifying matrices at relay nodes in the downlink system. Second, with transmission power constraint at each node of the relay network, one can use scaled and Hermitian transposed uplink relay amplifying matrices in the downlink system, with scaling factors obtained by switching power constraints at different nodes of the uplink system. As an application of the uplink-downlink duality, we propose an optimal design of the source precoding matrix and relay amplifying matrices for multi-hop MIMO relay system with a nonlinear DPC transmitter at the source node.

Chapter 4 is based on the journal publication:

- Y. Rong and M. R. A. Khandaker, "On Uplink-Downlink Duality of Multi-Hop MIMO Relay Channel", *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1923-1931, Jun. 2011.

and the conference publication:

- M. R. A. Khandaker and Y. Rong, "Dirty Paper Coding Based Optimal MIMO Relay Communications", in *Proc. 16th Asia-Pacific Conference on Communications (APCC'2010)*, Auckland, New Zealand, Nov. 1-3, 2010, pp. 328-333. (**Best Paper Award**)

Chapter 5: Interference MIMO Relay Systems

In a large wireless network with many nodes, multiple source-destination links must share a common frequency band concurrently to ensure a high spectral efficiency of the whole network [1]. In such network, cochannel interference (CCI) is one of the main impairments that degrades the system performance. Developing schemes that mitigate the CCI is therefore important. We consider an interference MIMO relay system where multiple source nodes communicate with their desired destination nodes concurrently with the aid of distributed relay nodes all equipped with multiple antennas. We aim at minimizing the total source and relay transmit power such that a minimum SINR threshold is maintained at each receiver. An iterative joint power control and beamforming algorithm is developed to achieve this goal. The proposed algorithm exploits transmit-relay-receive beamforming technique to mitigate the interferences from the unintended sources in conjunction with transmit power control.

The material in Chapter 5 is based on one journal publication:

- M. R. A. Khandaker and Y. Rong, "Interference MIMO Relay Channel: Joint Power Control and Transceiver-Relay Beamforming", *IEEE Trans. Signal Process.*, to appear, 2012.

and two conference publications:

- M. R. A. Khandaker and Y. Rong, "Joint Power Control and Beamforming for Peer-to-Peer MIMO Relay Systems", in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Nov. 9-11, 2011.

- M. R. A. Khandaker and Y. Rong, "Joint Power Control and Beamforming for Interference MIMO Relay Channel", in *Proc. 17th Asia-Pacific Conference on Communications (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

Chapter 6: Channel Estimation of Dual-Hop MIMO Relay System

The optimal source precoding matrix and relay amplifying matrix developed in Chapter 2 to Chapter 5 require the instantaneous channel state information (CSI). However, in practical relay communication systems, the instantaneous CSI is unknown, and therefore, has to be estimated at the destination node. We develop a novel channel estimation algorithm for two-hop MIMO relay systems using the PARAFAC analysis. The proposed algorithm provides the destination node with full knowledge of all channel matrices involved in the communication. The number of training data blocks required in the proposed channel estimation algorithm can be less than the number of relay nodes (antennas). In particular, we show that when the number of relay nodes (antennas) is smaller than the number of antennas at the source node and the destination node, as few as two training data blocks are sufficient to estimate all channels. Thus, the proposed algorithm has a higher spectral efficiency compared to the existing techniques. Moreover, the initial estimation of channel matrices is improved by a linear MMSE algorithm, which yields a smaller estimation error.

Chapter 6 is based on the journal publication:

- Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel Estimation of Dual-Hop MIMO Relay Systems via Parallel Factor Analysis", *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2224-2233, Jun. 2012.

and the conference publication:

- Y. Rong and M. R. A. Khandaker, "Channel Estimation of Dual-Hop MIMO Relay Systems Using Parallel Factor Analysis", in *Proc. 17th Asia-Pacific Conference on Communications (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

1.5 Notations

The notations used in this thesis are as follows: Lower case letters are used to denote scalars, e.g. s, n. Bold face lower case letters denote vectors, e.g. $\mathbf{s}$, $\mathbf{n}$. Bold face upper case letters are reserved for matrices, e.g. $\mathbf{S}$, $\mathbf{N}$. For matrices, $(\cdot)^T, (\cdot)^*, (\cdot)^H, (\cdot)^{-1}$, and $(\cdot)^\dagger$ denote transpose, conjugate, Hermitian transpose, inverse, and pseudo-inverse operations, respectively. rank$(\cdot)$ and tr$(\cdot)$ denote the rank and trace of matrices, respectively. E$[\cdot]$ represents the statistical expectation and blkdiag$(\cdot)$ stands for a block-diagonal matrix. An N dimensional identity matrix is denoted as either $\mathbf{I}_N$ or $\mathbf{I}$. Note that the scope of any variable in each chapter is limited to that particular chapter.

Chapter 2

Uplink Multiuser MIMO Relay Communication Systems

In this chapter, we address the optimal source, relay, and receive matrices design for linear non-regenerative uplink multiuser MIMO relay communication systems. The MMSE of the signal waveform estimation at the destination node is adopted as our design criterion. After a review of existing works in Section 2.1, the system model of an uplink multiuser MIMO relay network is introduced in Section 2.2. Two iterative methods to solve the highly nonconvex joint source, relay, and receiver optimization problem are proposed in Section 2.3. Simulation results are presented in Section 2.4 to justify the significance of the proposed algorithms before summarizing the chapter in Section 2.5. The proof of Theorem 2.1 is given in Section 2.A.

2.1 Overview of Existing Techniques

In next generation wireless systems, multiple users equipped with multiple antennas will transmit simultaneously to the base station with multiple receive antennas and vice versa [3, 4]. Transceiver design for multiuser MIMO systems has been studied in [3]. The capacity of multiuser MIMO systems was investigated for flat fading channels in [4] using real channel measurement data.

Incorporating relays in a MIMO network can significantly extend the coverage and improve the link reliability of the network [19, 21–28]. The capacity of a single-user non-regenerative MIMO relay channel has been studied in [21]. In [22] and [23], the optimal

relay amplifying matrix maximizing the mutual information (MI) between source and destination was derived assuming that the source covariance matrix is an identity matrix. In [24] and [25], MMSE-based approaches for MIMO relay systems have been studied. In [26], an iterative tri-step source precoder, relay amplifying matrix and destination equalizer design algorithm has been proposed for a single-user MIMO relay system with channel uncertainties. A unified framework was developed in [27] and [28] to jointly optimize the source precoding matrix and the relay amplifying matrices for a broad class of frequently used objective functions in MIMO relay system design.

For a multiuser MIMO relay system, the achievable sum rate has been derived in [52] using non-regenerative relaying scheme. In [53], both non-regenerative and regenerative relays have been considered in a multiuser MIMO network without optimizing the power loading schemes at the relay and the source nodes. An adaptive relay power allocation algorithm has been developed in [54] to mitigate the self-interference. An MMSE-based joint filter design has been proposed for a multiuser non-regenerative MIMO relay system in [55]. All these works [52]-[55] assume that each user is equipped with a single antenna. Several recent works have addressed multiuser MIMO relay systems where users also have multiple antennas. In [56], the optimal source and relay matrices were developed to maximize the source-destination MI. The non-regenerative MIMO relay technique has been applied to multi-cellular (interference) systems in [57]. The joint source and relay optimization problem has been addressed in [41] for multiple-antenna users using the MMSE criterion. The authors in [58] addressed the joint transceiver and relay design problem in a downlink (broadcast) multiuser system.

The main contribution of this chapter is the joint source, relay, and receiver optimization for multiuser MIMO relay communication systems under the MMSE criterion where all nodes (users, relay, and destination) are equipped with possibly different number of multiple antennas. In contrast to [58], we consider an uplink (multiaccess) multiuser MIMO relay system. Note that although we consider the joint transceiver design problem for an uplink system, transceivers in a downlink system can be obtained by exploiting the uplink-downlink duality of MIMO relay channel [48, 59]. This problem has not been addressed in [19]-[57]. In particular, [19]-[28] considered the transceiver and/or relay design problems for single-user MIMO relay systems whereas [52]-[55] considered multiuser MIMO relay design problems with single-antenna transmitters. The problems addressed in [56] and [57] are also different from our problem. In this chapter, we derive the optimal structure of the source precoding matrix of each user and the

relay amplifying matrix to jointly minimize the MSE of the signal waveform estimation at the destination node in a multiuser MIMO relay system. The original optimization problem is highly nonconvex and a closed-form solution is intractable. To overcome this difficulty, we develop a Tri-Step iterative algorithm to jointly optimize the source, relay, and receive matrices through solving convex subproblems. It is shown that this algorithm is guaranteed to converge to (at least) a locally optimal solution. Note that the Tri-Step algorithm is not presented in [41].

To reduce the computational complexity of the Tri-Step algorithm, we develop a simplified Bi-Step algorithm, where the source and relay matrices are optimized in an alternating fashion. The receive matrix is not updated in each iteration, and instead, it is obtained as an MMSE receiver after the convergence of the Bi-Step algorithm. We show that for given source precoding matrices, the optimal relay amplifying matrix diagonalizes the source-relay-destination channel. While for fixed relay matrix and source matrices of all other users, the source matrix of each user has a beamforming structure. Simulation results demonstrate that both the proposed Tri-Step and Bi-Step iterative algorithms perform much better than existing techniques in terms of both MSE and BER. Moreover, it is shown that compared with the Tri-Step algorithm, the Bi-Step algorithm requires less number of iterations till convergence with only a small degradation in MSE and BER. Such performance-complexity tradeoff is very important for practical multiuser MIMO relay communication systems. We would like to mention that such Bi-Step algorithm is not considered in [58].

2.2 Uplink Multiuser MIMO Relay System Model

We consider a two-hop multiuser MIMO relay communication system as illustrated in Fig. 2.1 where K users transmit information to the same destination node with the aid of one relay node. The ith user, $i = 1, \cdots, K$, the relay and the destination nodes are equipped with N_i, N_r, and N_d antennas, respectively. We denote $N_b = \sum_{i=1}^{K} N_i$ as the total number of independent data streams from all users, and assume that $N_b \leq \min(N_r, N_d)$, since otherwise the system cannot support N_b independent data streams simultaneously. For simplicity, as in [22]-[56], a linear non-regenerative strategy is applied at the relay node to process and forward the received signal.

We assume that the relay node works in the practical half-duplex mode. Thus, the communication between the users and the destination is completed in two time slots.

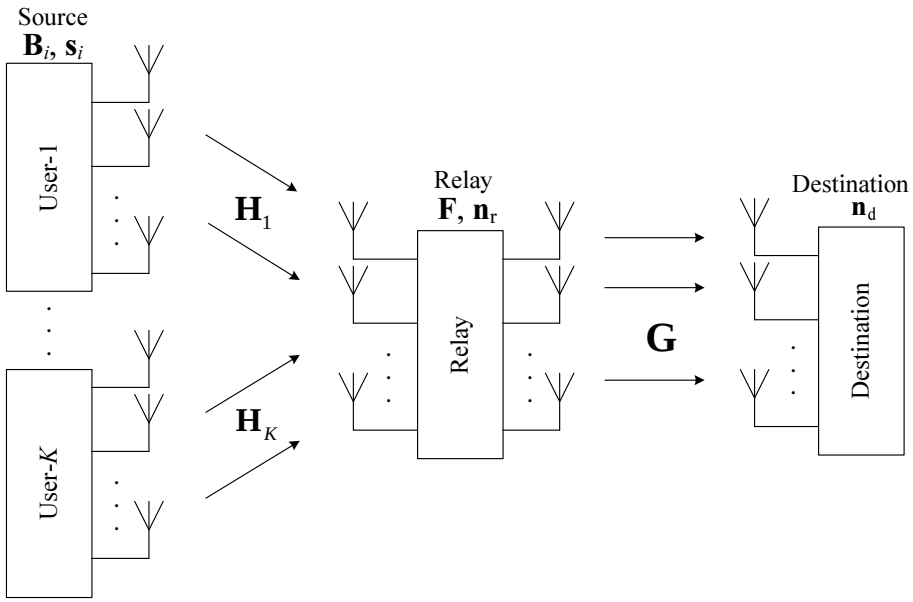

Figure 2.1: Block diagram of a K-user linear non-regenerative MIMO relay communication system.

In the first time slot, the $N_i \times 1$ modulated signal vector $\mathbf{s}_i$ is linearly precoded at the ith user by the $N_i \times N_i$ source precoding matrix $\mathbf{B}_i$. The precoded signal vector

$$\mathbf{x}_i = \mathbf{B}_i \mathbf{s}_i \tag{2.1}$$

is transmitted to the relay node from the ith user. The received signal vector at the relay node can be written as

$$\mathbf{y}_r = \sum_{i=1}^{K} \mathbf{H}_i \mathbf{x}_i + \mathbf{n}_r \tag{2.2}$$

where $\mathbf{H}_i$ is the $N_r \times N_i$ MIMO channel matrix between the ith user and the relay, $\mathbf{y}_r$ and $\mathbf{n}_r$ are the received signal and the additive Gaussian noise vectors at the relay node, respectively. Substituting (2.1) into (2.2), we have

$$\mathbf{y}_r = \sum_{i=1}^{K} \mathbf{H}_i \mathbf{B}_i \mathbf{s}_i + \mathbf{n}_r = \bar{\mathbf{H}} \mathbf{s} + \mathbf{n}_r \tag{2.3}$$

where $\bar{\mathbf{H}} \triangleq [\mathbf{H}_1 \mathbf{B}_1, \cdots, \mathbf{H}_K \mathbf{B}_K]$ is the equivalent multiaccess MIMO channel matrix of the source-relay link and $\mathbf{s} \triangleq \left[\mathbf{s}_1^T, \cdots, \mathbf{s}_K^T\right]^T$ is the equivalent transmitted signal vector. We assume that $\mathrm{E}\left[\mathbf{s}\mathbf{s}^H\right] = \mathbf{I}_{N_b}$.

In the second time slot, the users remain silent and the relay node multiplies (linearly precodes) the received signal vector $\mathbf{y}_r$ by an $N_r \times N_r$ relay amplifying matrix $\mathbf{F}$ and transmits the signal vector

$$\mathbf{x}_r = \mathbf{F}\mathbf{y}_r \tag{2.4}$$

to the destination node. The received signal vector at the destination node can be written as

$$\mathbf{y}_d = \mathbf{G}\mathbf{x}_r + \mathbf{n}_d \tag{2.5}$$

where $\mathbf{G}$ is the $N_d \times N_r$ MIMO channel matrix between the relay and the destination nodes, $\mathbf{y}_d$ and $\mathbf{n}_d$ are the received signal and the additive Gaussian noise vectors at the destination node, respectively.

Substituting (2.3) and (2.4) into (2.5), we obtain

$$\begin{aligned}\mathbf{y}_d &= \mathbf{G}\mathbf{F}\sum_{i=1}^{K}\mathbf{H}_i\mathbf{B}_i\mathbf{s}_i + \mathbf{G}\mathbf{F}\mathbf{n}_r + \mathbf{n}_d \\ &= [\mathbf{G}\mathbf{F}\mathbf{H}_1\mathbf{B}_1, \cdots, \mathbf{G}\mathbf{F}\mathbf{H}_K\mathbf{B}_K]\,\mathbf{s} + \mathbf{G}\mathbf{F}\mathbf{n}_r + \mathbf{n}_d = \mathbf{H}\mathbf{s} + \mathbf{n}\end{aligned} \tag{2.6}$$

where $\mathbf{H} \triangleq [\mathbf{G}\mathbf{F}\mathbf{H}_1\mathbf{B}_1, \cdots, \mathbf{G}\mathbf{F}\mathbf{H}_K\mathbf{B}_K] = \mathbf{G}\mathbf{F}\bar{\mathbf{H}}$ is the equivalent MIMO channel matrix of the source-relay-destination link, and $\mathbf{n} \triangleq \mathbf{G}\mathbf{F}\mathbf{n}_r + \mathbf{n}_d$ is the equivalent noise vector at the destination. We assume that the channel matrices $\mathbf{H}_i, i = 1, \cdots, K$, and $\mathbf{G}$ are all quasi-static, i.e., the channel matrices are constant throughout a block of transmission and known to the relay and the destination nodes. In practice, the CSI of $\mathbf{G}$ can be obtained at the destination node through standard training method. The relay node can have the CSI of $\mathbf{H}_i, i = 1, \cdots, K$, through channel training, and obtain the CSI of $\mathbf{G}$ by a feedback from the destination node. The quasi-static channel model is valid in practice when the mobility of all communicating nodes is relatively slow. As a result, we can obtain the necessary CSI with a reasonably high precision during the channel training period. The relay node calculates the optimal source matrices $\{\mathbf{B}_i\} \triangleq \{\mathbf{B}_i, i = 1, \cdots, K\}$, and the relay matrix $\mathbf{F}$, and forwards $\mathbf{B}_i$ to the ith source node and forwards $\mathbf{F}$ and $\mathbf{H}_i, i = 1, \cdots, K$ to the destination node. Note that individual users do not require any channel knowledge. This is a very important assumption for multiuser communication since in a multiuser scenario the users are distributed and cannot cooperate. We assume that all noises are independent and identically distributed (i.i.d.) complex circularly symmetric Gaussian noise with zero mean and unit variance.

Due to its simplicity, a linear receiver is used at the destination node to retrieve the transmitted signals. Denoting $\mathbf{W}$ as an $N_d \times N_b$ weight matrix, the estimated signal vector $\hat{\mathbf{s}}$ is given by

$$\hat{\mathbf{s}} = \mathbf{W}^H \mathbf{y}_d \,. \tag{2.7}$$

2.3 Proposed Source, Relay, and Receive Matrices Design Algorithm

In this section we develop the optimal source precoding matrices $\{\mathbf{B}_i\}$, the relay amplifying matrix $\mathbf{F}$, and the destination receive matrix $\mathbf{W}$ to minimize the MSE of the signal waveform estimation. Using (2.6) and (2.7), the MSE of the signal waveform estimation at the destination is given by

$$\begin{aligned} \text{MSE} &= \text{tr}\left(\text{E}\left[(\hat{\mathbf{s}}-\mathbf{s})(\hat{\mathbf{s}}-\mathbf{s})^H\right]\right) \\ &= \text{tr}\left(\left(\mathbf{W}^H\mathbf{H}-\mathbf{I}_{N_b}\right)\left(\mathbf{W}^H\mathbf{H}-\mathbf{I}_{N_b}\right)^H + \mathbf{W}^H\mathbf{C}_n\mathbf{W}\right) \end{aligned} \tag{2.8}$$

where $\mathbf{C}_n$ is the equivalent noise covariance matrix given by

$$\mathbf{C}_n = \text{E}\left[\mathbf{n}\mathbf{n}^H\right] = \text{E}\left[\left(\mathbf{G}\mathbf{F}\mathbf{n}_r + \mathbf{n}_d\right)\left(\mathbf{G}\mathbf{F}\mathbf{n}_r + \mathbf{n}_d\right)^H\right] = \mathbf{G}\mathbf{F}\mathbf{F}^H\mathbf{G}^H + \mathbf{I}_{N_d}.$$

From (2.4), the power of the signal transmitted by the relay node can be expressed as

$$\text{tr}\left(\text{E}\left[\mathbf{x}_r\mathbf{x}_r^H\right]\right) = \text{tr}\left(\mathbf{F}\left(\sum_{i=1}^{K}\mathbf{H}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{H}_i^H + \mathbf{I}_{N_r}\right)\mathbf{F}^H\right). \tag{2.9}$$

From (2.8) and (2.9), the joint source, relay, and receive matrices optimization problem for the linear non-regenerative multiuser MIMO relay system can be formulated as

$$\min_{\{\mathbf{B}_i\},\mathbf{F},\mathbf{W}} \quad \text{tr}\left(\left(\mathbf{W}^H\mathbf{H}-\mathbf{I}_{N_b}\right)\left(\mathbf{W}^H\mathbf{H}-\mathbf{I}_{N_b}\right)^H + \mathbf{W}^H\mathbf{C}_n\mathbf{W}\right) \tag{2.10a}$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{F}\left(\sum_{i=1}^{K}\mathbf{H}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{H}_i^H + \mathbf{I}_{N_r}\right)\mathbf{F}^H\right) \le P_r \tag{2.10b}$$

$$\text{tr}\left(\mathbf{B}_i\mathbf{B}_i^H\right) \le P_i, \qquad i = 1, \cdots, K \tag{2.10c}$$

where (2.10b) and (2.10c) are the constraints for the transmission power at the relay and the ith user, respectively, and $P_r > 0$, $P_i > 0$ are the power budget available at the relay and the ith source node, respectively. The optimization problem (2.10) is highly

nonconvex and a closed-form solution to this problem is intractable. In the following, we develop two iterative algorithms namely the Tri-Step and the Bi-Step algorithms to optimize the source, relay, and receive matrices. In the Tri-Step algorithm, the source, relay, and receive matrices are optimized iteratively through solving convex sub-problems. In the Bi-Step algorithm, the source and relay matrices are optimized alternatingly and the MMSE receive matrix is calculated after the convergence of the source and relay matrices. In particular, the relay matrix is optimized by the Lagrange multiplier method in the Tri-Step algorithm, and by the majorization theory in the Bi-Step algorithm. The optimal source matrices are obtained by solving semidefinite programming (SDP) problem in the Bi-Step algorithm, and by solving quadratically constrained quadratic programming (QCQP) problem in the Tri-Step algorithm.

2.3.1 Iterative optimization of source, relay, and receive matrices (Tri-Step Algorithm)

This algorithm starts at a random $\mathbf{F}$ and $\{\mathbf{B}_i\}$ satisfying (2.10b) and (2.10c). In each iteration, the source, relay, and receive matrices are updated alternatingly through solving convex subproblems. Firstly, with given $\mathbf{F}$ and $\{\mathbf{B}_i\}$, the optimal $\mathbf{W}$ is obtained by solving the unconstrained convex problem (2.10a), since $\mathbf{W}$ does not appear in constraints (2.10b) and (2.10c). The solution is the well-known MMSE receiver given by [60]

$$\mathbf{W} = \left(\mathbf{H}\mathbf{H}^H + \mathbf{C}_n\right)^{-1}\mathbf{H}. \tag{2.11}$$

Secondly, with given $\mathbf{W}$ and $\{\mathbf{B}_i\}$, $\mathbf{F}$ can be updated by solving the following problem

$$\min_{\mathbf{F}} \quad \mathrm{tr}\left(\left(\bar{\mathbf{G}}\mathbf{F}\bar{\mathbf{H}} - \mathbf{I}_{N_b}\right)\left(\bar{\mathbf{G}}\mathbf{F}\bar{\mathbf{H}} - \mathbf{I}_{N_b}\right)^H + \bar{\mathbf{G}}\mathbf{F}\mathbf{F}^H\bar{\mathbf{G}}^H\right) \tag{2.12a}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{I}_{N_r}\right)\mathbf{F}^H\right) \le P_r \tag{2.12b}$$

where $\bar{\mathbf{G}} \triangleq \mathbf{W}^H\mathbf{G}$ is the equivalent relay-destination MIMO channel. Using the Lagrange multiplier method, we obtain $\mathbf{F}$ from (2.12) as

$$\mathbf{F} = \bar{\mathbf{G}}^H(\bar{\mathbf{G}}\bar{\mathbf{G}}^H + \mu\mathbf{I}_{N_b})^{-1}\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{I}_{N_r})^{-1} \tag{2.13}$$

where $\mu \ge 0$ is the Lagrange multiplier associated with the power constraint (2.12b). Interestingly, (2.13) can be viewed as $\mathbf{F} = \mathbf{F}_2\mathbf{F}_1$, where $\mathbf{F}_1 = \bar{\mathbf{H}}^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{I}_{N_r})^{-1}$ is the weight matrix of the MMSE receiver for the equivalent first-hop multiaccess MIMO

channel at the relay node given in (2.3), and $\mathbf{F}_2 = \bar{\mathbf{G}}^H(\bar{\mathbf{G}}\bar{\mathbf{G}}^H+\mu\mathbf{I}_{N_b})^{-1}$ can be viewed as the transmit precoding matrix for the effective second-hop MIMO system $\mathbf{y} = \bar{\mathbf{G}}\mathbf{F}_2\mathbf{x}+\mathbf{v}$, where $\mathbf{x}$ is the transmitted signal vector with $\mathrm{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_{N_b}$, and $\mathbf{v}$ is the noise vector with covariance matrix $\mathbf{C}_v = \mathrm{E}[\mathbf{v}\mathbf{v}^H]$. In this MIMO system, the MSE of estimating $\mathbf{x}$ is given by $\mathrm{tr}\big(\mathrm{E}[(\mathbf{y}-\mathbf{x})(\mathbf{y}-\mathbf{x})^H]\big) = \mathrm{tr}\big(\big(\bar{\mathbf{G}}\mathbf{F}_2-\mathbf{I}_{N_b}\big)\big(\bar{\mathbf{G}}\mathbf{F}_2-\mathbf{I}_{N_b}\big)^H+\mathbf{C}_v\big)$. The optimal $\mathbf{F}_2$ that minimizes the MSE can be obtained by solving the following problem

$$\min_{\mathbf{F}_2} \ \mathrm{tr}\big(\big(\bar{\mathbf{G}}\mathbf{F}_2-\mathbf{I}_{N_b}\big)\big(\bar{\mathbf{G}}\mathbf{F}_2-\mathbf{I}_{N_b}\big)^H\big) \qquad \text{s.t. } \mathrm{tr}\big(\mathbf{F}_2\mathbf{F}_2^H\big) \le P_x$$

where P_x is the transmission power constraint. Using the Lagrange multiplier method to solve the problem above, we obtain $\mathbf{F}_2 = \bar{\mathbf{G}}^H(\bar{\mathbf{G}}\bar{\mathbf{G}}^H+\mu\mathbf{I}_{N_b})^{-1}$.

The Lagrange multiplier μ in (2.13) can be found from the following complementary slackness condition

$$\mu\big(\mathrm{tr}\big(\mathbf{F}\big(\bar{\mathbf{H}}\bar{\mathbf{H}}^H+\mathbf{I}_{N_r}\big)\mathbf{F}^H\big)-P_r\big) = 0. \tag{2.14}$$

Assuming $\mu = 0$, we have the following $\mathbf{F}$ from (2.13)

$$\mathbf{F} = \bar{\mathbf{G}}^H(\bar{\mathbf{G}}\bar{\mathbf{G}}^H)^{-1}\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H+\mathbf{I}_{N_r})^{-1}. \tag{2.15}$$

Since in this case (2.14) is already satisfied, if $\mathbf{F}$ in (2.15) satisfies the constraint (2.12b), then (2.15) is the solution to the problem (2.12). Otherwise, there must be $\mu > 0$, and from (2.14) we can see that $\mathrm{tr}\big(\mathbf{F}\big(\bar{\mathbf{H}}\bar{\mathbf{H}}^H+\mathbf{I}_{N_r}\big)\mathbf{F}^H\big) = P_r$ must hold. In this case, μ can be obtained from (2.12b) by solving the following nonlinear equation

$$\mathrm{tr}\big(\bar{\mathbf{G}}^H(\bar{\mathbf{G}}\bar{\mathbf{G}}^H+\mu\mathbf{I}_{N_b})^{-1}\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H+\mathbf{I}_{N_r})^{-1}\bar{\mathbf{H}}(\bar{\mathbf{G}}\bar{\mathbf{G}}^H+\mu\mathbf{I}_{N_b})^{-1}\bar{\mathbf{G}}\big) = P_r. \tag{2.16}$$

Let us now define the singular value decomposition (SVD) of $\bar{\mathbf{G}} \triangleq \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$, where the dimensions of $\mathbf{U}$, $\mathbf{\Lambda}$, $\mathbf{V}$ are $N_b \times N_b$, $N_b \times N_r$, and $N_r \times N_r$, respectively. Then we have from (2.16) that

$$\mathrm{tr}\big(\mathbf{\Lambda}(\mathbf{\Lambda}^2+\mu\mathbf{I}_{N_b})^{-1}\mathbf{U}^H\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H+\mathbf{I}_{N_r})^{-1}\bar{\mathbf{H}}\mathbf{U}(\mathbf{\Lambda}^2+\mu\mathbf{I}_{N_b})^{-1}\mathbf{\Lambda}\big) = P_r. \tag{2.17}$$

Denoting $\mathbf{\Phi} \triangleq \mathbf{U}^H\bar{\mathbf{H}}^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H+\mathbf{I}_{N_r})^{-1}\bar{\mathbf{H}}\mathbf{U}$, (2.17) can be equivalently written as

$$\sum_{i=1}^{N_b}\frac{\lambda_i^2\Phi_{i,i}}{\big(\lambda_i^2+\mu\big)^2} = P_r \tag{2.18}$$

where λ_i and $\Phi_{i,i}$ are the ith main diagonal elements of $\mathbf{\Lambda}$ and $\mathbf{\Phi}$, respectively. Since the left-hand side of (2.18) is a monotonically decreasing function of $\mu > 0$, it can be efficiently solved using the bisection method [61].

Thirdly, with given $\mathbf{W}$ and $\mathbf{F}$, we reformulate the problem (2.10) as a QCQP problem [61] to update $\mathbf{b}_i, i = 1, \cdots, K$, where $\mathbf{b}_i = \text{vec}(\mathbf{B}_i)$ stands for a vector obtained by stacking all column vectors of $\mathbf{B}_i$ on top of each other. Let $\mathbf{A}_i \triangleq \mathbf{W}^H\mathbf{G}\mathbf{F}\mathbf{H}_i$ and $\mathbf{A}_{ii}$ be a matrix containing the $\left(\sum_{j=1}^{i-1} N_j + 1\right)$-th to $\left(\sum_{j=1}^{i} N_j\right)$-th rows of $\mathbf{A}_i$. Using the identity of $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$ [62], where $\otimes$ denotes the matrix Kronecker product, we obtain that

$$\begin{aligned}
\text{tr}\big(\mathbf{W}^H\mathbf{G}\mathbf{F}\mathbf{H}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{F}^H\mathbf{G}^H\mathbf{W}\big) &= \text{tr}\big(\mathbf{A}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{A}_i^H\big) \\
&= \mathbf{b}_i^H\big(\mathbf{I}_{N_i} \otimes (\mathbf{A}_i^H\mathbf{A}_i)\big)\mathbf{b}_i, \quad i = 1, \cdots, K \\
\text{tr}\left(\mathbf{W}^H\mathbf{G}\mathbf{F}[\mathbf{H}_1\mathbf{B}_1, \cdots, \mathbf{H}_K\mathbf{B}_K]\right) &= \text{tr}([\mathbf{A}_1\mathbf{B}_1, \cdots, \mathbf{A}_K\mathbf{B}_K]) \\
&= \sum_{i=1}^{K} \text{tr}(\mathbf{A}_{ii}\mathbf{B}_i) \\
&= \sum_{i=1}^{K} \big(\text{vec}(\mathbf{A}_{ii}^T)\big)^T\mathbf{b}_i.
\end{aligned}$$

Thus the MSE in (2.8) can be expressed as

$$\begin{aligned}
\text{MSE} = \text{tr}\Bigg(&\mathbf{W}^H\mathbf{G}\mathbf{F}\left(\sum_{i=1}^{K}\mathbf{H}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{H}_i^H\right)\mathbf{F}^H\mathbf{G}^H\mathbf{W} - \mathbf{W}^H\mathbf{G}\mathbf{F}[\mathbf{H}_1\mathbf{B}_1, \cdots, \mathbf{H}_K\mathbf{B}_K] \\
&- \left(\mathbf{W}^H\mathbf{G}\mathbf{F}\left[\mathbf{H}_1\mathbf{B}_1, \cdots, \mathbf{H}_K\mathbf{B}_K\right]\right)^H + \mathbf{I}_{N_b} + \mathbf{W}^H(\mathbf{G}\mathbf{F}\mathbf{F}^H\mathbf{G}^H + \mathbf{I}_{N_d})\mathbf{W}\Bigg) \\
= &\sum_{i=1}^{K}\mathbf{b}_i^H\left(\mathbf{I}_{N_i} \otimes (\mathbf{A}_i^H\mathbf{A}_i)\right)\mathbf{b}_i - \sum_{i=1}^{K}\big(\text{vec}(\mathbf{A}_{ii}^T)\big)^T\mathbf{b}_i - \sum_{i=1}^{K}\mathbf{b}_i^H\text{vec}(\mathbf{A}_{ii}^H) + t \\
\triangleq &\ \mathbf{b}^H\mathbf{A}\mathbf{b} - \mathbf{c}^H\mathbf{b} - \mathbf{b}^H\mathbf{c} + t
\end{aligned} \tag{2.19}$$

where $t \triangleq \text{tr}(\mathbf{I}_{N_b} + \mathbf{W}^H(\mathbf{G}\mathbf{F}\mathbf{F}^H\mathbf{G}^H + \mathbf{I}_{N_d})\mathbf{W})$, $\mathbf{A} \triangleq \text{blkdiag}\big(\mathbf{I}_{N_1} \otimes (\mathbf{A}_1^H\mathbf{A}_1), \cdots, \mathbf{I}_{N_K} \otimes (\mathbf{A}_K^H\mathbf{A}_K)\big)$, $\mathbf{b} \triangleq \left[\mathbf{b}_1^T, \cdots, \mathbf{b}_K^T\right]^T$, and $\mathbf{c} \triangleq \left[\big(\text{vec}(\mathbf{A}_{11}^H)\big)^T, \cdots, \big(\text{vec}(\mathbf{A}_{KK}^H)\big)^T\right]^T$. Now the MSE in (2.19) can be equivalently rewritten as

$$\begin{aligned}
\text{MSE} &= \mathbf{b}^H\mathbf{A}^{\frac{1}{2}}\mathbf{A}^{\frac{1}{2}}\mathbf{b} - \mathbf{c}^H\mathbf{A}^{-\frac{1}{2}}\mathbf{A}^{\frac{1}{2}}\mathbf{b} - \mathbf{b}^H\mathbf{A}^{\frac{1}{2}}\mathbf{A}^{-\frac{1}{2}}\mathbf{c} + \mathbf{c}^H\mathbf{A}^{-\frac{1}{2}}\mathbf{A}^{-\frac{1}{2}}\mathbf{c} - \mathbf{c}^H\mathbf{A}^{-1}\mathbf{c} + t \\
&= \big(\mathbf{b}^H\mathbf{A}^{\frac{1}{2}} - \mathbf{c}^H\mathbf{A}^{-\frac{1}{2}}\big)\big(\mathbf{A}^{\frac{1}{2}}\mathbf{b} - \mathbf{A}^{-\frac{1}{2}}\mathbf{c}\big) - \mathbf{c}^H\mathbf{A}^{-1}\mathbf{c} + t
\end{aligned}$$

where $\mathbf{A}^{\frac{1}{2}}\mathbf{A}^{\frac{1}{2}} = \mathbf{A}$ and $\mathbf{A}^{\frac{1}{2}} = \mathbf{A}^{\frac{H}{2}}$. Note that we can ignore the term $t - \mathbf{c}^H\mathbf{A}^{-1}\mathbf{c}$ while optimizing $\mathbf{b}$ with given $\mathbf{W}$ and $\mathbf{F}$, since it is free of the optimization variable $\mathbf{b}$. Assuming $\mathbf{C}_i \triangleq \mathbf{F}\mathbf{H}_i, i = 1, \cdots, K$, the relay transmit power constraint in (2.10b) can be rewritten as

$$\mathbf{b}^H\mathbf{C}\mathbf{b} \leq P_r - \text{tr}\big(\mathbf{F}\mathbf{F}^H\big)$$

where $\mathbf{C} \triangleq \text{blkdiag}\left(\mathbf{I}_{N_1} \otimes (\mathbf{C}_1^H \mathbf{C}_1), \cdots, \mathbf{I}_{N_K} \otimes (\mathbf{C}_K^H \mathbf{C}_K)\right)$. Thus the optimization problem (2.10) can be equivalently rewritten as the following QCQP problem

$$\min_{\mathbf{b}} \quad \left(\mathbf{A}^{\frac{1}{2}}\mathbf{b} - \mathbf{A}^{-\frac{1}{2}}\mathbf{c}\right)^H \left(\mathbf{A}^{\frac{1}{2}}\mathbf{b} - \mathbf{A}^{-\frac{1}{2}}\mathbf{c}\right) \tag{2.20a}$$

$$\text{s.t.} \quad \mathbf{b}^H \mathbf{C} \mathbf{b} \leq P_r - \text{tr}\left(\mathbf{F}\mathbf{F}^H\right) \tag{2.20b}$$

$$\mathbf{b}^H \mathbf{D}_i \mathbf{b} \leq P_i, \qquad i = 1, \cdots, K \tag{2.20c}$$

where $\mathbf{D}_i \triangleq \text{blkdiag}\left(\tilde{\mathbf{D}}_{i1}, \tilde{\mathbf{D}}_{i2}, \cdots, \tilde{\mathbf{D}}_{iK}\right)$ with $\tilde{\mathbf{D}}_{ii} = \mathbf{I}_{N_i}$ and $\tilde{\mathbf{D}}_{ij} = \mathbf{0}$, $j = 1, \cdots, K, j \neq i$. The QCQP problem (2.20) can be efficiently solved by the disciplined convex programming toolbox CVX [63] where interior-point method-based solvers such as SeDuMi and SDPT3 are called internally. Since all subproblems (2.10a), (2.12), and (2.20) are convex, the solution to each subproblem is optimal. Thus, the value of the objective function (2.10a) decreases (or at least maintains) after each iteration. Moreover, the objective function is lower bounded by at least zero.

Now, assuming that $\mathbf{W}_0$, $\{\mathbf{B}_{i,0}\}$, and $\mathbf{F}_0$ are the optimal solution for each subproblem, we have

$$\text{tr}(\nabla_{\mathbf{W}} J(\mathbf{X}_0)^T (\mathbf{W} - \mathbf{W}_0)) \geq 0 \tag{2.21}$$

$$\text{tr}(\nabla_{\mathbf{B}_i} J(\mathbf{X}_0)^T (\mathbf{B}_i - \mathbf{B}_{i,0})) \geq 0 \tag{2.22}$$

$$\text{tr}(\nabla_{\mathbf{F}} J(\mathbf{X}_0)^T (\mathbf{F} - \mathbf{F}_0)) \geq 0 \tag{2.23}$$

where $\mathbf{X}_0 \triangleq [\mathbf{W}_0, \{\mathbf{B}_{i,0}\}, \mathbf{F}_0]$ and $\nabla_{\mathbf{A}} J(\mathbf{X}_0)$ is the gradient of the objective function (10) along the direction of $\mathbf{A} \in \{\mathbf{W}, \{\mathbf{B}_i\}, \mathbf{F}\}$ at $\mathbf{X}_0$. Summing up (2.21)-(2.23), we obtain $\text{tr}(\nabla J(\mathbf{X}_0)^T (\mathbf{X} - \mathbf{X}_0)) \geq 0$, indicating that $\mathbf{X}_0$ is a stationary point of (10). Moreover, it can be seen that $\mathbf{X}_0$ must be on the edge of the feasible set specified by inequalities in (2.10b) and (2.10c) (i.e., (2.10b) and (2.10c) must be satisfied with equality at $\mathbf{X}_0$, since otherwise, one can simply scale $\mathbf{F}_0$ and $\mathbf{B}_{i,0}$ such that the value of (2.10a) is decreased without violating (2.10b) and (2.10c)). This indicates that $\mathbf{X}_0$ cannot be a saddle point and is indeed the local-optimal solution. Therefore, the proposed iterative algorithm monotonically converges to (at least) a locally optimal solution. The procedure of solving the problem (2.10) using the proposed Tri-Step iterative algorithm is listed in Table 2.1, where $\|\cdot\|_1$ denotes the matrix maximum absolute column sum norm, ε is a small positive number close to zero and the superscript (n) denotes the number of iterations.

Table 2.1: Procedure of solving the problem (2.10) by the Tri-Step algorithm

1. Initialize the algorithm with $\mathbf{B}_i^{(0)} = \sqrt{P_i/N_i}\ \mathbf{I}_{N_i}$, $i = 1, \cdots, K$, and $\mathbf{F}^{(0)} = \sqrt{P_r/\mathrm{tr}(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{I}_{N_r})}\ \mathbf{I}_{N_r}$; Set $n = 0$.
2. Update $\mathbf{W}^{(n)}$ using $\{\mathbf{B}_i^{(n)}\}$ and $\mathbf{F}^{(n)}$ as in (2.11).
3. Update $\mathbf{F}^{(n+1)}$ as in (2.13) using given $\mathbf{W}^{(n)}$ and $\{\mathbf{B}_i^{(n)}\}$.
4. Solve the subproblem (2.20) using known $\mathbf{F}^{(n+1)}$ and $\mathbf{W}^{(n)}$ to obtain $\mathbf{B}_i^{(n+1)}$, $i = 1, \cdots, K$.
5. If $\max_i \left\|\mathbf{B}_i^{(n+1)} - \mathbf{B}_i^{(n)}\right\|_1 \leq \varepsilon$, then end.
   Otherwise, let $n := n + 1$ and go to step 2.

2.3.2 Simplified source and relay matrices design (Bi-Step Algorithm)

In this subsection, we propose an iterative source and relay matrices design algorithm which has a smaller computational complexity than the Tri-Step algorithm developed in the previous subsection. In particular, using the MMSE receiver (2.11) at the destination node, the MSE of the signal waveform estimation (2.8) becomes a function of $\{\mathbf{B}_i\}$ and $\mathbf{F}$ as

$$\mathrm{MMSE} = \mathrm{tr}\left(\left[\mathbf{I}_{N_b} + \mathbf{H}^H \mathbf{C}_n^{-1} \mathbf{H}\right]^{-1}\right). \tag{2.24}$$

Thus, the joint source and relay optimization problem is given by

$$\min_{\{\mathbf{B}_i\},\mathbf{F}} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_b} + \mathbf{H}^H \mathbf{C}_n^{-1} \mathbf{H}\right]^{-1}\right) \tag{2.25a}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}\left(\sum_{i=1}^{K} \mathbf{H}_i \mathbf{B}_i \mathbf{B}_i^H \mathbf{H}_i^H + \mathbf{I}_{N_r}\right)\mathbf{F}^H\right) \leq P_r \tag{2.25b}$$

$$\mathrm{tr}\left(\mathbf{B}_i \mathbf{B}_i^H\right) \leq P_i, \qquad i = 1, \cdots, K. \tag{2.25c}$$

In this Bi-Step algorithm, we update the source and the relay matrices in an alternating fashion. In each iteration, for given source matrices $\{\mathbf{B}_i\}$ satisfying (2.25c), we optimize the relay matrix $\mathbf{F}$ by solving the following problem

$$\min_{\mathbf{F}} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_b} + \mathbf{H}^H \mathbf{C}_n^{-1} \mathbf{H}\right]^{-1}\right) \tag{2.26a}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}\left(\sum_{i=1}^{K} \mathbf{H}_i \mathbf{B}_i \mathbf{B}_i^H \mathbf{H}_i^H + \mathbf{I}_{N_r}\right)\mathbf{F}^H\right) \leq P_r. \tag{2.26b}$$

Then using this $\mathbf{F}$, we solve the problem (2.25) (with only $\{\mathbf{B}_i\}$ as the optimization variables) to obtain optimal source precoding matrices $\{\mathbf{B}_i\}$. Finally, the receive matrix $\mathbf{W}$ is obtained as (2.11) using the value of $\{\mathbf{B}_i\}$ and $\mathbf{F}$ at the convergence point.

Let us now define the following SVDs

$$\bar{\mathbf{H}} = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{V}_s^H, \qquad \mathbf{G} = \mathbf{U}_r \mathbf{\Lambda}_r \mathbf{V}_r^H$$

where the dimensions of $\mathbf{U}_s$, $\mathbf{\Lambda}_s$, $\mathbf{V}_s$ are $N_r \times N_r$, $N_r \times N_b$, $N_b \times N_b$, respectively, and the dimensions of $\mathbf{U}_r$, $\mathbf{\Lambda}_r$, $\mathbf{V}_r$ are given as $N_d \times N_d$, $N_d \times N_r$, $N_r \times N_r$, respectively. We assume that the main diagonal elements of $\mathbf{\Lambda}_s$ and $\mathbf{\Lambda}_r$ are arranged in a decreasing order. Based on Theorem 1 in [27], the optimal structure of $\mathbf{F}$ obtained from solving the problem (2.26) is given by

$$\mathbf{F} = \mathbf{V}_{r,1} \mathbf{\Lambda}_f \mathbf{U}_{s,1}^H \tag{2.27}$$

where $\mathbf{\Lambda}_f$ is an $N_b \times N_b$ diagonal matrix, $\mathbf{V}_{r,1}$ and $\mathbf{U}_{s,1}$ contain the leftmost N_b columns from $\mathbf{V}_r$ and $\mathbf{U}_s$, respectively.

It can be seen from (2.27) that the optimal $\mathbf{F}$ diagonalizes the equivalent source-relay-destination MIMO channel $\mathbf{H}$. Substituting (2.27) back into (2.26a) and (2.26b), we obtain the problem of optimizing $\mathbf{\Lambda}_f$ as

$$\min_{\{\lambda_{f,i}\}} \quad \sum_{i=1}^{N_b} \left(1 + \frac{\lambda_{s,i}^2 \lambda_{r,i}^2 \lambda_{f,i}^2}{1 + \lambda_{r,i}^2 \lambda_{f,i}^2}\right)^{-1} \tag{2.28a}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_b} \lambda_{f,i}^2 (\lambda_{s,i}^2 + 1) \le P_r, \qquad \lambda_{f,i} \ge 0, \quad i = 1, \cdots, N_b \tag{2.28b}$$

where $\lambda_{s,i}$, $\lambda_{f,i}$, and $\lambda_{r,i}$ are the ith main diagonal elements of $\mathbf{\Lambda}_s$, $\mathbf{\Lambda}_f$, and $\mathbf{\Lambda}_r$, respectively. The problem (2.28) has a water-filling solution which is given by

$$\lambda_{f,i} = \frac{1}{\lambda_{r,i}} \left[\frac{1}{\lambda_{s,i}^2 + 1} \left(\frac{\lambda_{s,i} \lambda_{r,i}}{\left[(\lambda_{s,i}^2 + 1)\nu\right]^{\frac{1}{2}}} - 1 \right)^+ \right]^{\frac{1}{2}}, \qquad i = 1, \cdots, N_b \tag{2.29}$$

where for a real-valued number x, $(x)^+ \triangleq \max(x, 0)$, and $\nu > 0$ is the solution to the nonlinear problem of

$$\sum_{i=1}^{N_b} \frac{1}{\lambda_{r,i}^2} \left(\frac{\lambda_{s,i} \lambda_{r,i}}{\left[(\lambda_{s,i}^2 + 1)\nu\right]^{\frac{1}{2}}} - 1 \right)^+ = P_r. \tag{2.30}$$

Since (2.30) is a monotonically decreasing function of ν, it can be efficiently solved using the bisection method [61].

Using the identity of $\mathrm{tr}\big(\big[\mathbf{I}_m+\mathbf{A}_{m\times n}\mathbf{B}_{n\times m}\big]^{-1}\big)=\mathrm{tr}\big(\big[\mathbf{I}_n+\mathbf{B}_{n\times m}\mathbf{A}_{m\times n}\big]^{-1}\big)+m-n$, for a given feasible $\mathbf{F}$, the objective function (2.24) can be rewritten as

$$\begin{aligned}\mathrm{MMSE}&=\mathrm{tr}\left(\left[\mathbf{I}_{N_d}+\mathbf{H}\mathbf{H}^H\mathbf{C}_n^{-1}\right]^{-1}\right)+N_b-N_d\\&=\mathrm{tr}\left(\left[\mathbf{I}_{N_d}+\mathbf{C}_n^{-\frac{1}{2}}\mathbf{H}\mathbf{H}^H\mathbf{C}_n^{-\frac{1}{2}}\right]^{-1}\right)+N_b-N_d\\&=\mathrm{tr}\left(\left[\mathbf{I}_{N_d}+\mathbf{C}_n^{-\frac{1}{2}}\mathbf{G}\mathbf{F}\sum_{i=1}^{K}\mathbf{H}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{F}^H\mathbf{G}^H\mathbf{C}_n^{-\frac{1}{2}}\right]^{-1}\right)+N_b-N_d\\&=\mathrm{tr}\left(\left[\mathbf{I}_{N_d}+\sum_{i=1}^{K}\tilde{\mathbf{H}}_i\mathbf{Q}_i\tilde{\mathbf{H}}_i^H\right]^{-1}\right)+N_b-N_d\end{aligned}$$

where $\tilde{\mathbf{H}}_i\triangleq\mathbf{C}_n^{-\frac{1}{2}}\mathbf{G}\mathbf{F}\mathbf{H}_i$ and $\mathbf{Q}_i=\mathbf{B}_i\mathbf{B}_i^H$ is the source covariance matrix of the ith user. In the following, we focus on optimizing $\mathbf{Q}_i$. Once we obtain the optimal $\mathbf{Q}_i$, the optimal $\mathbf{B}_i$ is calculated as $\mathbf{B}_i=\mathbf{\Theta}_i\mathbf{\Lambda}_i^{\frac{1}{2}}\mathbf{\Phi}_i$, where $\mathbf{\Theta}_i\mathbf{\Lambda}_i\mathbf{\Theta}_i^H$ is the eigenvalue decomposition (EVD) of $\mathbf{Q}_i$, and $\mathbf{\Phi}_i$ is an arbitrary $N_i\times N_i$ unitary matrix. Considering the transmission power constraints in (2.25b) and (2.25c), the source covariance matrices $\{\mathbf{Q}_i\}\triangleq\{\mathbf{Q}_i,i=1,\cdots,K\}$ can be optimized by solving the following problem

$$\min_{\{\mathbf{Q}_i\}}\quad\mathrm{tr}\left(\left[\mathbf{I}_{N_d}+\sum_{i=1}^{K}\tilde{\mathbf{H}}_i\mathbf{Q}_i\tilde{\mathbf{H}}_i^H\right]^{-1}\right)\tag{2.31a}$$

$$\text{s.t.}\quad\mathrm{tr}\left(\sum_{i=1}^{K}\mathbf{Q}_i\mathbf{\Psi}_i\right)\le\bar{P}_r\tag{2.31b}$$

$$\mathrm{tr}(\mathbf{Q}_i)\le P_i,\qquad\mathbf{Q}_i\succcurlyeq 0,\qquad i=1,\cdots,K\tag{2.31c}$$

where $\mathbf{\Psi}_i\triangleq\mathbf{H}_i^H\mathbf{F}^H\mathbf{F}\mathbf{H}_i$, $\bar{P}_r\triangleq P_r-\mathrm{tr}(\mathbf{F}\mathbf{F}^H)$, and for a matrix $\mathbf{A}$, $\mathbf{A}\succcurlyeq 0$ means that $\mathbf{A}$ is a positive semi-definite (PSD) matrix.

Let us now introduce a PSD matrix $\mathbf{X}$ that satisfies

$$\left[\mathbf{I}_{N_d}+\sum_{i=1}^{K}\tilde{\mathbf{H}}_i\mathbf{Q}_i\tilde{\mathbf{H}}_i^H\right]^{-1}\preccurlyeq\mathbf{X}\tag{2.32}$$

where for two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{B}\succcurlyeq\mathbf{A}$ means that $\mathbf{B}-\mathbf{A}\succcurlyeq 0$. By using (2.32) and the Schur complement [61], the problem (2.31) can be equivalently converted to the

following SDP problem

$$\min_{\{\mathbf{Q}_i\},\mathbf{X}} \quad \mathrm{tr}\,(\mathbf{X}) \tag{2.33a}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{X} & \mathbf{I}_{N_d} \\ \mathbf{I}_{N_d} & \mathbf{I}_{N_d} + \sum_{i=1}^{K} \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H \end{bmatrix} \succcurlyeq 0 \tag{2.33b}$$

$$\mathrm{tr}\left(\sum_{i=1}^{K} \mathbf{Q}_i \mathbf{\Psi}_i \right) \leq \bar{P}_r \tag{2.33c}$$

$$\mathrm{tr}(\mathbf{Q}_i) \leq P_i, \qquad \mathbf{Q}_i \succcurlyeq 0, \qquad i = 1, \cdots, K\,. \tag{2.33d}$$

We use the CVX software package [63] to solve the problem (2.33). Now the original source and relay matrices optimization problem (2.25) can be solved by an iterative technique as shown in Table 2.2.

Table 2.2: Procedure of solving the problem (2.25) by the Bi-Step algorithm

1. Initialize the algorithm with $\mathbf{Q}_i^{(0)} = P_i/N_i \mathbf{I}_{N_i}$, $i = 1, \cdots, K$; Set $n = 0$.
2. Solve the subproblem (2.26) using given $\mathbf{Q}_i^{(n)}$, $i = 1, \cdots, K$, to obtain $\mathbf{F}^{(n)}$ as in (2.27).
3. Solve the subproblem (2.33) using known $\mathbf{F}^{(n)}$ to obtain $\mathbf{Q}_i^{(n+1)}$, $i = 1, \cdots, K$.
4. If $\max_i \left\| \mathbf{Q}_i^{(n+1)} - \mathbf{Q}_i^{(n)} \right\|_1 \leq \varepsilon$, then end.
   Otherwise, let $n := n + 1$ and go to step 2.

Since the problem (2.28) is a convex optimization problem, the conditional update of $\mathbf{F}^{(n)}$ will not increase (2.28a) and hence the objective function (2.25a). Similarly, the problem (2.33) is also convex, and the conditional update of $\mathbf{Q}_i^{(n)}$ cannot increase (2.33a) and hence the value of (2.25a). Therefore, each conditional update of $\mathbf{F}^{(n)}$ and $\mathbf{Q}_i^{(n)}$ may either decrease or maintain but cannot increase the objective function (2.25a). Note that the constraints in the problem (2.25) are always satisfied with every conditional update. Similar to the justification for the Tri-Step algorithm, a monotonic convergence of $\mathbf{F}^{(n)}$ and $\mathbf{Q}_i^{(n)}$ towards (at least) a locally optimal solution follows directly from this observation.

The numerical solution to the problem (2.33) does not provide sufficient insight to the structure of the optimal $\mathbf{Q}_i$. Interestingly, by solving the problem (2.31) applying the Lagrange multiplier method, we obtain the following theorem for the structure of the optimal $\mathbf{Q}_i$.

Theorem 2.1 *The optimal source covariance matrix* $\mathbf{Q}_i$ *for the* i*th user as the solution to the problem (2.31) has the following general beamforming structure*

$$\mathbf{Q}_i = \mathbf{V}_{h_i}\mathbf{\Lambda}_{h_i,1}^{-1}\mathbf{U}_{h_i,1}^H(\mathbf{V}_i\mathbf{J}_i\mathbf{V}_i^H - \mathbf{D}_i)^\sharp\mathbf{U}_{h_i,1}\mathbf{\Lambda}_{h_i,1}^{-1}\mathbf{V}_{h_i}^H, \qquad i = 1, \cdots, K \tag{2.34}$$

where $\mathbf{D}_i \triangleq \mathbf{I}_{N_d} + \sum_{j=1, j\neq i}^K \tilde{\mathbf{H}}_j\mathbf{Q}_j\tilde{\mathbf{H}}_j^H$, $(\cdot)^\sharp$ stands for the projection to the set of $N_d \times N_d$ PSD matrices, $\tilde{\mathbf{H}}_i = [\mathbf{U}_{h_i,1} \quad \mathbf{U}_{h_i,2}][\mathbf{\Lambda}_{h_i,1} \quad \mathbf{0}]^T\mathbf{V}_{h_i}^H$ and $\mathbf{K}_i^{-1}\tilde{\mathbf{H}}_i^H = \mathbf{U}_i[\mathbf{\Sigma}_i \quad \mathbf{0}]\mathbf{V}_i^H$ are the SVDs of $\tilde{\mathbf{H}}_i$ and $\mathbf{K}_i^{-1}\tilde{\mathbf{H}}_i^H$, respectively, and $\mathbf{J}_i \triangleq \text{blkdiag}(\mathbf{\Sigma}_i, \mathbf{\Delta}_{i,2})$. Here $\mathbf{K}_i\mathbf{K}_i^H = \lambda_1\mathbf{\Psi}_i + \lambda_2\mathbf{I}_{N_i}$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ are the Lagrange multipliers, and $\mathbf{\Delta}_{i,2}$ is an $(N_d - N_i) \times (N_d - N_i)$ diagonal matrix.

PROOF: See Appendix 2.A. □

The unknown Lagrange multipliers λ_1 and λ_2 in (2.34) can be found by solving the dual optimization problem associated with the problem (2.35) in Appendix 2.A. Note that the optimal structure of the source covariance matrices in (2.34) can be viewed as $\mathbf{Q}_i = \tilde{\mathbf{H}}_i^\dagger\mathbf{\Lambda}_i\big(\tilde{\mathbf{H}}_i^\dagger\big)^H, i = 1, \cdots, K$, where $\mathbf{\Lambda}_i \triangleq (\mathbf{V}_i\mathbf{J}_i\mathbf{V}_i^H - \mathbf{D}_i)^\sharp$ is the power-loading matrix. Note that (2.34) indicates that the power distribution at each user needs to be adapted to the current power levels of all other users. The pseudo-inverse in $\mathbf{Q}_i, i = 1, \cdots, K$, indicates that the source covariance matrix of the ith user needs to match the corresponding source-relay-destination channel.

In summary, matrices $\mathbf{W}$, $\mathbf{F}$, and $\{\mathbf{B}_i\}$ are optimized in each iteration of the Tri-Step algorithm, where the major computation task lies in solving the QCQP problem (2.20). The amount of computation required for updating $\mathbf{W}$ and $\mathbf{F}$ is negligible compared with that of solving the QCQP problem. The complexity order of solving the problem (2.20) using the interior point method [64] is $\mathcal{O}\big((\sum_{i=1}^K N_i^2)^3\big)$.

In each iteration of the Bi-Step algorithm, $\mathbf{F}$ and $\{\mathbf{B}_i\}$ are optimized. Here the major computation task is solving the SDP problem (2.33), which has a complexity order of $\mathcal{O}\big((\sum_{i=1}^K N_i^2)^{3.5}\big)$ using the interior point method [64]. Therefore, the per iteration computational complexity of the Bi-Step algorithm is slightly higher than that of the Tri-Step algorithm. However, the overall computational complexity of both iterative algorithms also depends on the number of iterations they need till convergence, which will be studied in Section 2.4 (see Table 2.3). Note that in the Tri-Step algorithm, the source, relay, and receive matrices are optimized iteratively while in the Bi-Step algorithm, the source and relay matrices are optimized alternatingly and the MMSE receive matrix is calculated after the convergence of the source and relay matrices. Thus the Bi-Step approach has a fast convergence than the Tri-Step algorithm.

2.4 Numerical Examples

In this section, we study the performance of the proposed optimal multiuser MIMO relay algorithms through numerical simulations. For simplicity, we consider a system with two users. The extension to $K > 2$ users is straight-forward. The two users, relay and destination nodes are all equipped with multiple antennas. We simulate a flat Rayleigh fading environment where the channel matrices have entries with zero mean and variances σ_g^2/N_r, $\sigma_{h,i}^2/N_i$, for $\mathbf{G}$, $\mathbf{H}_i$, $i = 1, 2$, respectively. We define

$$\mathrm{SNR}_{\mathrm{r-d}} \triangleq \frac{\sigma_g^2 P_r N_d}{N_r}, \qquad \mathrm{SNR}_{\mathrm{s_i-r}} \triangleq \frac{\sigma_{h,i}^2 P_i N_r}{N_i}, \quad i = 1, 2$$

as the SNR of the relay-destination and user-i-relay links, $i = 1, 2$, respectively. For simplicity, we assume $N_1 = N_2 = N_s$ and $\mathrm{SNR}_{\mathrm{s_1-r}} = \mathrm{SNR}_{\mathrm{s_2-r}} = \mathrm{SNR}_{\mathrm{s-r}}$ throughout the simulations. All simulation results are averaged over 1000 independent channel realizations.

We compare the performance of the proposed Tri-Step and Bi-Step algorithms with the naive amplify-and-forward (NAF) algorithm, and the pseudo match-and-forward (PMF) algorithm in terms of both MSE and BER. For the Tri-Step algorithm, the procedure in Table 2.1 is carried out to obtain the optimal relay and source matrices, whereas for the Bi-Step algorithm, the steps defined in Table 2.2 are followed. For both algorithms, we use the CVX Matlab toolbox for disciplined convex programming [63] to find the optimal source precoding matrices. For the NAF scheme, we use $\mathbf{B}_i = \sqrt{P_i/N_i}\,\mathbf{I}_{N_i}$, $i = 1, 2$, and $\mathbf{F} = \sqrt{P_r/\mathrm{tr}(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{I}_{N_r})}\,\mathbf{I}_{N_r}$. For the PMF algorithm, the same $\mathbf{B}_i$ in the NAF algorithm is taken and $\mathbf{F} = \sqrt{P_r/\mathrm{tr}((\bar{\mathbf{H}}\mathbf{G})^H(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{I}_{N_r})\bar{\mathbf{H}}\mathbf{G})}\,(\bar{\mathbf{H}}\mathbf{G})^H$. Both the NAF and the PMF algorithms use the MMSE receiver at the destination node.

In the first two examples, we compare the performance of the proposed algorithms with the other two approaches in terms of MSE normalized by the number of data streams (NMSE) for $N_s = 2$, $N_r = 4$, and $N_d = 4$. Fig. 2.2 shows the MSE performance of all tested algorithms versus $\mathrm{SNR}_{\mathrm{s-r}}$ with $\mathrm{SNR}_{\mathrm{r-d}} = 20$dB, whereas Fig. 2.3 illustrates the MSE performance of tested algorithms versus $\mathrm{SNR}_{\mathrm{r-d}}$ for an $\mathrm{SNR}_{\mathrm{s-r}}$ fixed at 20dB. Our results clearly demonstrate the better performance of the proposed iterative joint source and relay optimization algorithms. It can be seen that the proposed optimal algorithms consistently yield the lowest MSE over the entire $\mathrm{SNR}_{\mathrm{s-r}}$ and $\mathrm{SNR}_{\mathrm{r-d}}$ region. The NAF and PMF algorithms have much higher MSE compared with the proposed schemes even at very high SNR. Note that the MSE performance of both the Tri-Step algorithm and the Bi-Step algorithm are almost similar to each other.

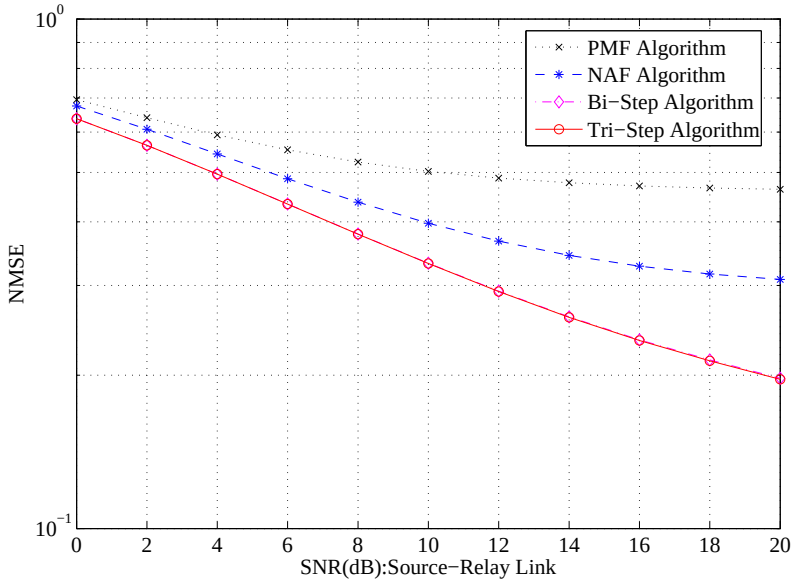

Figure 2.2: Example 1: Normalized MSE versus $\mathrm{SNR_{s-r}}$. $N_s = 2$, $N_r = 4, N_d = 4$, $\mathrm{SNR_{r-d}} = 20\mathrm{dB}$.

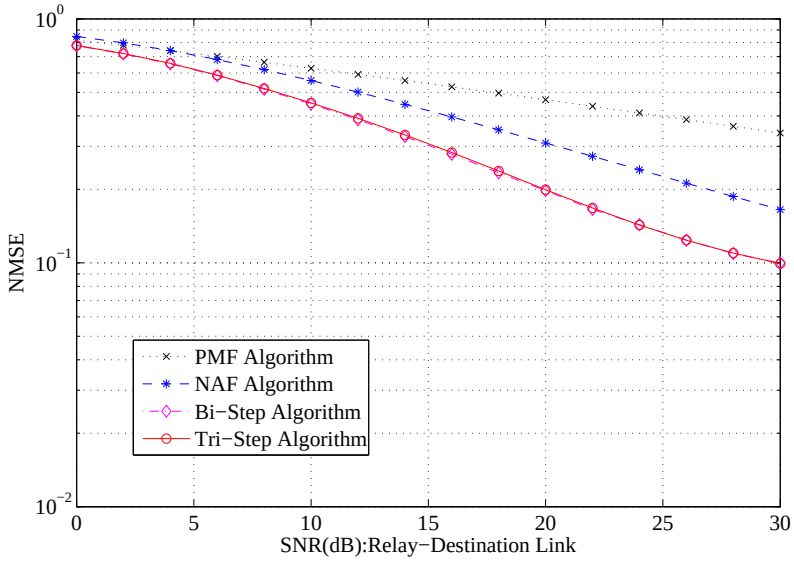

Figure 2.3: Example 2: Normalized MSE versus $\mathrm{SNR_{r-d}}$. $N_s = 2$, $N_r = 4, N_d = 4$, $\mathrm{SNR_{s-r}} = 20\mathrm{dB}$.

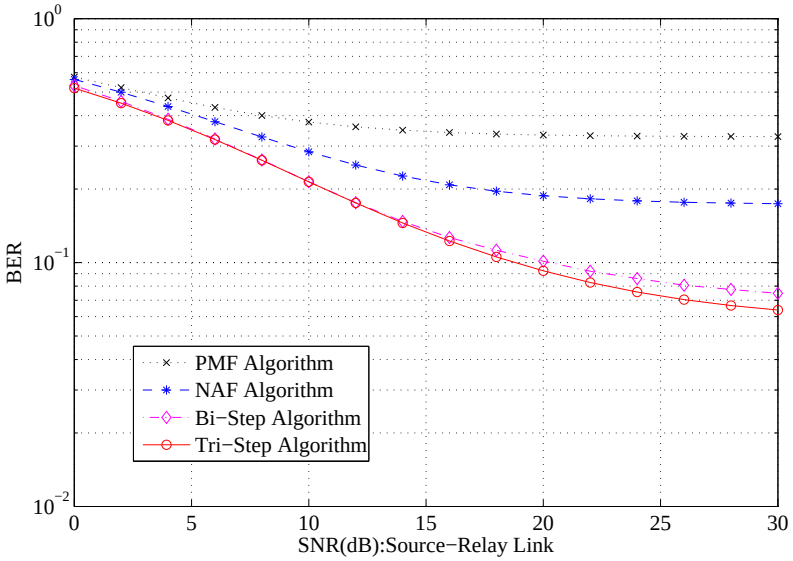

Figure 2.4: Example 3: BER versus $\mathrm{SNR_{s-r}}$. $N_s = 3$, $N_r = 6, N_d = 6$, $\mathrm{SNR_{r-d}} = 20\mathrm{dB}$.

In the next example, we compare the performance of the four algorithms in terms of BER. QPSK signal constellations are used to modulate the transmitted signals. We set $N_s = 3$, $N_r = 6$, $N_d = 6$, and transmit 3000 randomly generated bits from each user in each channel realization. Fig. 2.4 shows the BER performance of all algorithms versus $\mathrm{SNR_{s-r}}$ for $\mathrm{SNR_{r-d}} = 20\mathrm{dB}$.

In the fourth example, we compare the BER performance of the algorithms varying the SNR in the relay-destination channel. This time we set $N_s = 2$, $N_r = 6$, $N_d = 8$, and transmit 3000 randomly generated bits from each user in each channel realization using QPSK signal constellations. Fig. 2.5 shows the BER performance of the algorithms versus $\mathrm{SNR_{r-d}}$ for $\mathrm{SNR_{s-r}} = 20\mathrm{dB}$. Note that in contrast to other three schemes, the PMF algorithm requires $N_b = N_d$, and thus, its performance cannot be included in Fig. 2.5.

It can be seen from Fig. 2.4 and Fig. 2.5 that the proposed joint source and relay optimization algorithms obtain the lowest BERs compared with the other approaches. Interestingly, the BER performance of the Tri-Step algorithm is slightly better than that of the Bi-Step algorithm, especially at the high SNR region. The reason is that in the Tri-Step algorithm, we update the receiver weight matrix at each iteration in addition to the source and relay matrices. Note that since QPSK constellations are used to modulate the transmitting signals, it is possible that the bit error of the two

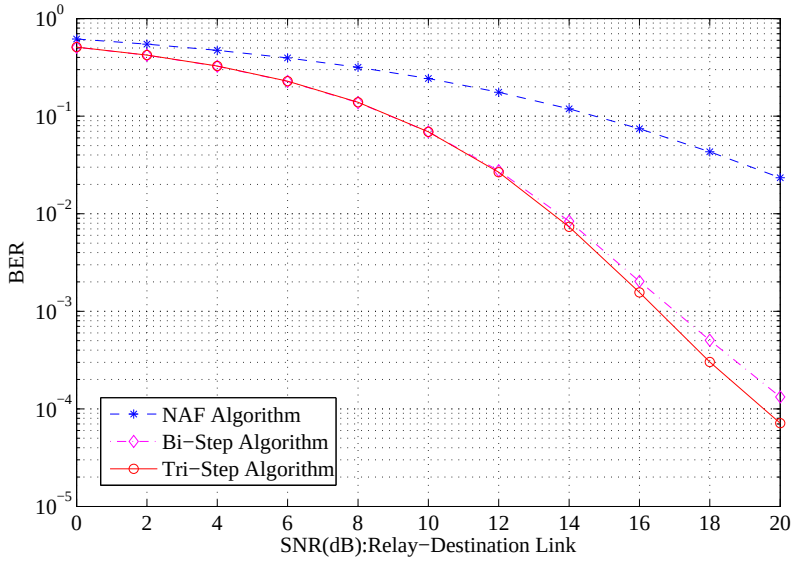

Figure 2.5: Example 4: BER versus $\mathrm{SNR_{r-d}}$. $N_s = 2$, $N_r = 6, N_d = 8$, $\mathrm{SNR_{s-r}} = 20$dB.

algorithms are different when the receiver demodulates the transmitted signals, even though the errors of signal waveform estimation are almost identical. Although the Tri-Step algorithm performs better than the Bi-Step algorithm, the former algorithm requires a larger number of iterations than the latter one to converge to the same ε. For comparison, the number of iterations both algorithms require in a typical run to converge up to $\varepsilon = 10^{-3}$ are listed in Table 2.3. Here we set $N_s = 2$, $N_r = 6$, $N_d = 6$ and $\mathrm{SNR_{r-d}} = 20$dB. Therefore, based on the per iteration complexity of two algorithms discussed in Section 2.3 and the number of iterations they need to converge, the overall computational complexity of the Bi-Step algorithm is smaller than that of the Tri-Step algorithm when the number of antennas at each user is small (which is the case in practical uplink multiuser communication systems). Such performance-complexity tradeoff is very important for practical multiuser MIMO relay communication systems.

Table 2.3: Iterations required till convergence in the proposed algorithms

$\mathrm{SNR_{s-r}}$ (dB)	0	5	10	15	20
Bi-Step algorithm	2	3	3	5	6
Tri-Step algorithm	6	6	10	19	22

Note that both the Bi-Step and Tri-Step approaches achieve an error floor at the

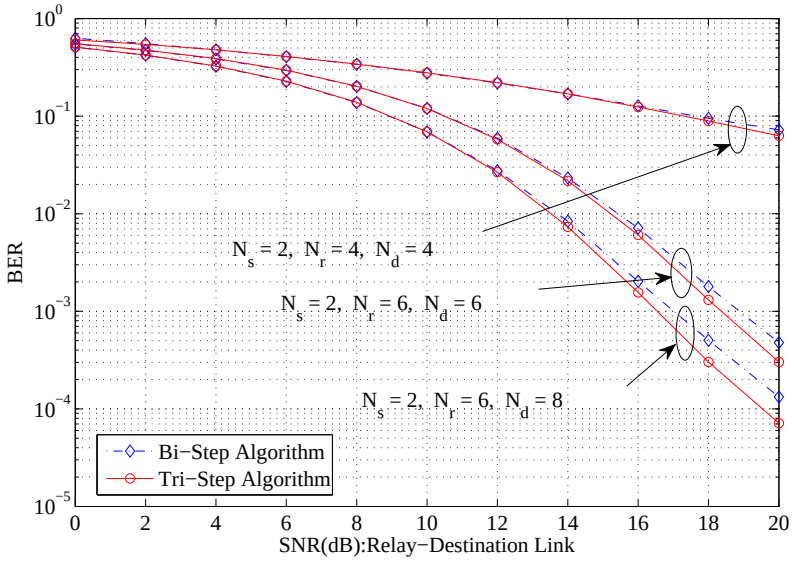

Figure 2.6: Example 5: BER versus $\mathrm{SNR_{r-d}}$. Varying number of antennas, $\mathrm{SNR_{s-r}}$ = 20dB.

BER of 6×10^{-2} in Fig. 2.4, but no error floor can be seen in Fig. 2.5. The reason is that the MIMO relay system simulated in Fig. 2.4 is fully-loaded while the system in Fig. 2.5 is under-loaded in terms of data streams. In particular, the total number of data streams $N_b = 6 = N_r = N_d$ in Fig. 2.4 and $N_b = 4 < N_r, N_d$ in Fig. 2.5. Thus the system in Fig. 2.5 has a higher spatial diversity order, which overcomes the saturation effect of the BER curve observed in Fig. 2.4.

In the last example, we compare the BER performance of the proposed algorithms for different number of antennas at the relay and the destination nodes with a fixed number of antennas at the source nodes. Fig. 2.6 compares the BER performance of the proposed algorithms versus $\mathrm{SNR_{r-d}}$ for $\mathrm{SNR_{s-r}}$ = 20dB with different number of antennas. It can be clearly seen from Fig. 2.6 that as we increase the number of antennas at the relay and/or destination node(s), the performance of the proposed algorithms improve significantly.

2.5 Chapter Summary

In this chapter, we developed the optimal structure of the source precoding matrices and the relay amplifying matrix in a multiuser MIMO relay network to jointly minimize

the MSE of the signal waveform estimation. We proposed two iterative algorithms to optimize the source, relay, and receive matrices. Simulation results demonstrate that the proposed algorithms outperform the existing techniques in terms of both MSE and BER.

2.A Proof of Theorem 2.1

To determine the structure of the optimal source covariance matrix $\mathbf{Q}_i$ for the ith user, we rewrite the problem (2.31) with given $\mathbf{Q}_j$, $j = 1, \cdots, K$, $j \neq i$ as

$$\min_{\mathbf{Q}_i} \quad \mathrm{tr}\left(\left[\mathbf{D}_i + \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H\right]^{-1}\right) \tag{2.35a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{Q}_i \mathbf{\Psi}_i) \leq \tilde{P}_r \tag{2.35b}$$

$$\mathrm{tr}(\mathbf{Q}_i) \leq P_i, \qquad \mathbf{Q}_i \succcurlyeq 0 \tag{2.35c}$$

where $\tilde{P}_r \triangleq \bar{P}_r - \mathrm{tr}\left(\sum_{j=1, j\neq i}^{K} \mathbf{Q}_j \mathbf{\Psi}_j\right)$. The Lagrangian function associated with the problem (2.35) is given by

$$\mathcal{L} \;=\; \mathrm{tr}\left(\left[\mathbf{D}_i + \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H\right]^{-1}\right) + \lambda_1 \left(\mathrm{tr}(\mathbf{Q}_i \mathbf{\Psi}_i) - \tilde{P}_r\right) + \lambda_2 \left(\mathrm{tr}(\mathbf{Q}_i) - P_i\right)$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the Lagrange multipliers. Making the derivative of $\mathcal{L}$ with respect to $\mathbf{Q}_i$ be zero, we obtain

$$\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_i} = -\tilde{\mathbf{H}}_i^H \left(\mathbf{D}_i + \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H\right)^{-2} \tilde{\mathbf{H}}_i + \lambda_1 \mathbf{\Psi}_i + \lambda_2 \mathbf{I}_{N_i} = 0. \tag{2.36}$$

By introducing an invertible matrix $\mathbf{K}_i$ with $\mathbf{K}_i \mathbf{K}_i^H = \lambda_1 \mathbf{\Psi}_i + \lambda_2 \mathbf{I}_{N_i}$, (2.36) becomes

$$\mathbf{K}_i^{-1} \tilde{\mathbf{H}}_i^H \left(\mathbf{D}_i + \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H\right)^{-2} \tilde{\mathbf{H}}_i \mathbf{K}_i^{-H} = \mathbf{I}_{N_i}. \tag{2.37}$$

Obviously, (2.37) is valid if and only if

$$\mathbf{K}_i^{-1} \tilde{\mathbf{H}}_i^H = \mathbf{P}_i \left(\mathbf{D}_i + \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H\right) \tag{2.38}$$

where $\mathbf{P}_i$ is an $N_i \times N_d$ semi-unitary matrix with $\mathbf{P}_i \mathbf{P}_i^H = \mathbf{I}_{N_i}$.

Let us introduce the following SVD and EVD

$$\mathbf{K}_i^{-1} \tilde{\mathbf{H}}_i^H = \mathbf{U}_i [\mathbf{\Sigma}_i \quad \mathbf{0}] \mathbf{V}_i^H, \qquad \mathbf{D}_i + \tilde{\mathbf{H}}_i \mathbf{Q}_i \tilde{\mathbf{H}}_i^H = [\mathbf{L}_{i,1} \quad \mathbf{L}_{i,2}] \mathrm{blkdiag}(\mathbf{\Delta}_{i,1}, \mathbf{\Delta}_{i,2}) \mathbf{L}_i^H \tag{2.39}$$

where the dimensions of $\mathbf{U}_i$, $\mathbf{V}_i$, $\mathbf{L}_i$ are $N_i \times N_i$, $N_d \times N_d$, and $N_d \times N_d$, respectively, $\mathbf{L}_{i,1}$ and $\mathbf{L}_{i,2}$ contain the leftmost N_i columns and the rightmost $N_d - N_i$ columns of

$\mathbf{L}_i$, respectively, and $\mathbf{\Sigma}_i$, $\mathbf{\Delta}_{i,1}$, $\mathbf{\Delta}_{i,2}$ are $N_i \times N_i$, $N_i \times N_i$, and $(N_d - N_i) \times (N_d - N_i)$ diagonal matrices, respectively. Substituting (2.39) back into (2.38), we have

$$\mathbf{U}_i[\mathbf{\Sigma}_i \quad \mathbf{0}]\mathbf{V}_i^H = [\mathbf{P}_i\mathbf{L}_{i,1}\mathbf{\Delta}_{i,1} \quad \mathbf{P}_i\mathbf{L}_{i,2}\mathbf{\Delta}_{i,2}]\mathbf{L}_i^H. \tag{2.40}$$

Equation (2.40) holds if and only if $\mathbf{P}_i = \mathbf{U}_i\mathbf{L}_{i,1}^H$, $\mathbf{\Delta}_{i,1} = \mathbf{\Sigma}_i$, and $\mathbf{L}_i = \mathbf{V}_i$. Thus, from (2.39) we have that

$$\mathbf{D}_i + \tilde{\mathbf{H}}_i\mathbf{Q}_i\tilde{\mathbf{H}}_i^H = \mathbf{V}_i\mathbf{J}_i\mathbf{V}_i^H \tag{2.41}$$

where $\mathbf{J}_i \triangleq \text{blkdiag}(\mathbf{\Sigma}_i, \mathbf{\Delta}_{i,2})$. Let us introduce the SVD of $\tilde{\mathbf{H}}_i$ as

$$\tilde{\mathbf{H}}_i = [\mathbf{U}_{h_i,1} \quad \mathbf{U}_{h_i,2}][\mathbf{\Lambda}_{h_i,1} \quad \mathbf{0}]^T\mathbf{V}_{h_i}^H \tag{2.42}$$

where the dimensions of $\mathbf{U}_{h_i,1}$, $\mathbf{U}_{h_i,2}$, $\mathbf{V}_{h_i}$ are $N_d \times N_i$, $N_d \times (N_d - N_i)$, and $N_i \times N_i$, respectively, $\mathbf{\Lambda}_{h_i,1}$ is an $N_i \times N_i$ diagonal matrix. By substituting (2.42) back into (2.41) and solving (2.41) for $\mathbf{Q}_i$, we have $\mathbf{Q}_i = \mathbf{V}_{h_i}\mathbf{\Lambda}_{h_i,1}^{-1}\mathbf{U}_{h_i,1}^H(\mathbf{V}_i\mathbf{J}_i\mathbf{V}_i^H - \mathbf{D}_i)\mathbf{U}_{h_i,1}\mathbf{\Lambda}_{h_i,1}^{-1}\mathbf{V}_{h_i}^H$. Finally, taking into account the constraint (2.35c), we obtain (2.34).

Chapter 3

Multicasting MIMO Relay Communication Systems

In many practical communication systems, the transmitter often needs to send common message to a group of receivers simultaneously, which is commonly referred to as multicast broadcasting. In this chapter, we focus on multicasting in the downlink multiuser MIMO relay systems where one transmitter multicasts common message to multiple receivers with the aid of a relay node, and all nodes are equipped with multiple antennas. In Section 3.1, we give a brief overview of the known multicasting techniques exploiting spatial diversity. The system model of a two-hop multicasting MIMO relay network is introduced in Section 3.2. The joint transmit and relay precoding matrices design algorithms are developed in Section 3.3 for the general case of multiple data streams multicasting based on two design criteria. In the first scheme, we aim at minimizing the maximal MSE of the signal waveform estimation among all receivers subjecting to power constraints at the transmitter and the relay node. In the second scheme, we propose a total transmission power minimization strategy subjecting to QoS constraints. In Section 3.4, we study the joint transmit and relay beamforming for the special case of single data stream multicasting through two-hop MIMO relay systems. Interestingly, we show that for the case of single data stream multicasting, the relay precoding matrix optimization problem can be equivalently converted to the transmit beamforming problem for single-hop multicasting systems. Section 3.5 shows the simulation results which justify the significance of the proposed algorithms under various scenarios. The chapter is briefly summarized in Section 3.6.

3.1 Existing Multicasting Techniques

The broadcasting nature of the wireless channel makes it naturally suitable for multicasting applications, since a single transmission may be simultaneously received by a number of users. Recently, wireless multicasting technology has triggered great interest among researchers across the world, due to the increasing demand for mobile applications such as streaming media, software updates, and location-based services involving group communications. However, wireless channel is subject to signal fading. By exploiting the spatial diversity, multi-antenna techniques can be applied to combat channel fading [1]. Next generation wireless standards such as WiMAX 802.16m and 3GPP LTE-Advanced have already included technologies which enable better multicasting solutions based on multi-antenna and beamforming techniques [65].

The problem of designing optimal beamforming vectors for multicasting is hard in general, mainly due to its nonconvex nature. The authors of [66] have designed transmit beamformers for physical layer multicasting using rank relaxations. The information theoretic capacity of the multi-antenna multicasting channel is studied in [67] with a particular focus on the scaling of the capacity and achievable rates as the number of antennas and/or users approaches infinity. The asymptotic capacity limits of multi-antenna physical layer multicasting has been studied in [68] based on antenna subset selection. The effect of channel spatial correlation on the multicasting capacity has been investigated in [69]. The authors of [70] investigated transmit precoding design for multi-antenna multicasting systems where the CSI is obtained via limited feedback. The authors of [71] considered transmit covariance design for a secrecy rate maximization problem under a multicasting scenario, where a multi-antenna transmitter delivers a common confidential message to multiple single-antenna receivers in the presence of multiple multi-antenna eavesdroppers.

The works in [66]-[71] solved the max-min SNR/rate beamforming problems with the aid of semidefinite relaxation (SDR) and rank-one approximation (which is suboptimal, in general). In [72], a stochastic beamforming strategy is proposed for multi-antenna physical layer multicasting where the randomization is guided by SDR, but without the need of rank-one approximation. The authors of [72] adopted an achievable rate perspective, and showed that the gaps between the rates of the proposed stochastic beamformers and the optimal multicasting capacity are no greater than 0.8314 bps/Hz. The fundamental limit of the max-min beamforming is that as the number of users grows to infinity, the achievable rate decreases to zero [67]. To cure this problem, a

joint beamforming and admission control problem has been considered in [73] and [74], where a subset of users is selected so that certain QoS requirements can be satisfied. An iterative transmit beamforming strategy has been proposed in [74] for multiple cochannel multicasting groups to minimize the total power transmitted by the antenna array subjecting to SINR constraints at the receivers. It has been shown that the problem can be approximated by a second-order cone programming (SOCP) problem which does not require rank relaxations.

While the works in [66]-[74] investigated multicasting systems with single-antenna receivers, recently multi-antenna receivers have been considered for multicasting systems [75]-[76]. In particular, coordinated beamforming techniques have been investigated in [75] to facilitate physical layer multicasting with multi-antenna receivers, and a generalized form of block diagonalization has been proposed to make orthogonal transmissions to distinct multicasting antenna groups. In [77], non-iterative nearly optimal transmit beamformers are designed for wireless link layer multicasting with real-valued channels, and for complex-valued channels an upper-bound on the multicasting rate is derived. The scaling of the achievable rate for increasing number of users is investigated in [76] for MIMO multicasting where the transmission is coded at the application layer over a number of channel realizations.

The above works [65]-[76] considered single-hop multicasting systems. However, in the case of long transmitter-receiver distance, relay node(s) is necessary to efficiently combat the pathloss of wireless channel. The concept of multiuser peer-to-peer relay network has been generalized to that of a multi-group multicasting relay network in [78] and a distributed beamforming algorithm is proposed to minimize the total relay power where each node is equipped with a single antenna. In [79], the authors investigated multicast scheduling with multiple sessions and multiple channels where the base station may multicast data in two sessions using MIMO simultaneously through the same channel leading to a higher multicasting rate than single-session transmissions, and the users are allowed to cooperatively help each other on orthogonal channels. The authors in [80] studied the lower-bound for the outage probability of cooperative multi-antenna multicasting schemes based on the AF strategy where the users are equipped with a single antenna.

In this chapter, we consider a two-hop multicasting MIMO relay system where one transmitter multicasts common message to multiple receivers with the aid of a relay

node. The transmitter, relay node, and receivers are all equipped with multiple antennas. To the best of our knowledge, such two-hop multicasting MIMO relay system has not been investigated in existing works. For the sake of the implementation simplicity, we choose the AF relaying strategy. We consider the joint transmit and relay precoding design problem based on two criteria. In the first scheme, we aim at minimizing the maximal MSE of the signal waveform estimation among all receivers subjecting to power constraints at the transmitter and the relay node. In order to facilitate low-power transmissions, in the second scheme, we propose a total transmitter and relay transmission power minimization strategy subjecting to QoS constraints in terms of the MSE of the signal waveform estimation at each receiver. Both problems are highly nonconvex with matrix variables and the exactly optimal solutions are very difficult to obtain. By exploiting the optimal structure of the relay precoding matrix, we propose low-complexity solutions to both problems under some mild approximation. In particular, we show that under (moderately) high first-hop SNR assumption, both problems can be solved using standard SDP techniques. Numerical simulations demonstrate the effectiveness of the proposed algorithms. Note that both algorithms support multicasting multiple data streams in contrast to the existing single data stream multicasting schemes [65]-[79]. Interestingly, we show that for the special case of single data stream multicasting, the relay precoding matrix optimization problem can be equivalently converted to the transmit beamforming problem for single-hop broadcasting systems. In this chapter, for notational convenience, we consider a narrow-band single-carrier system. However, our results can be straightforwardly generalized to each subcarrier of a broadband multi-carrier multicasting MIMO relay system.

3.2 System Model

We consider a two-hop MIMO multicasting system with L receivers as illustrated in Fig. 3.1. The transmitter and the relay node are equipped with $N_{\rm s}$ and $N_{\rm r}$ antennas, respectively. For the sake of notational simplicity, we assume that each receiver has $N_{\rm d}$ antennas. The algorithms developed in this chapter can be straightforwardly extended to multicasting systems where receivers have different number of antennas. The transmitter multicasts its information-carrying symbols to all receivers with the aid of a relay node. The direct links between the transmitter and the receivers are not considered since we assume that these direct links undergo much larger path attenuations

compared with the links via the relay node.

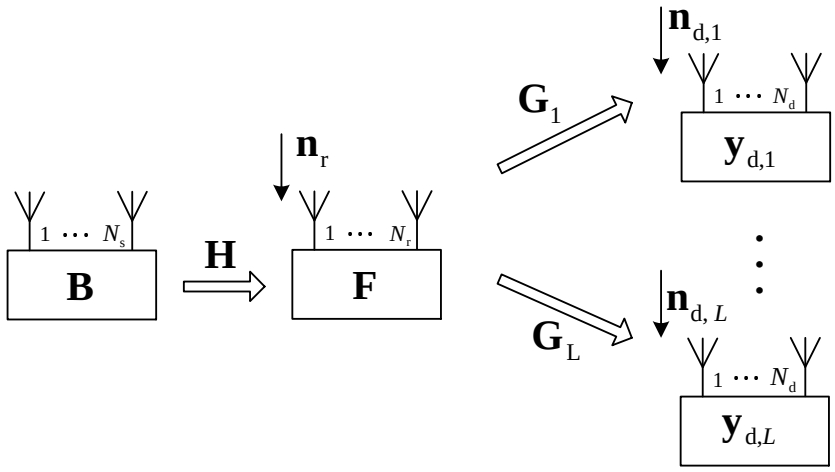

Figure 3.1: Block diagram of a two-hop multicasting MIMO relay system.

We assume that the relay node works in half-duplex mode. Thus the communication between the transmitter and receivers is accomplished in two time slots. In the first time slot, the transmitter linearly precodes an $N_\mathrm{b} \times 1$ modulated signal vector $\mathbf{s}$ (common message to all receivers) by an $N_\mathrm{s} \times N_\mathrm{b}$ precoding matrix $\mathbf{B}$ and transmits the precoded vector $\mathbf{x} = \mathbf{Bs}$ to the relay node. We assume that $\mathrm{E}[\mathbf{ss}^H] = \mathbf{I}_{N_\mathrm{b}}$. The received signal vector at the relay node is given by

$$\mathbf{y}_\mathrm{r} = \mathbf{HBs} + \mathbf{n}_\mathrm{r} \tag{3.1}$$

where $\mathbf{H}$ is the $N_\mathrm{r} \times N_\mathrm{s}$ MIMO channel matrix between the transmitter and the relay node, $\mathbf{y}_\mathrm{r}$ and $\mathbf{n}_\mathrm{r}$ are the $N_\mathrm{r} \times 1$ received signal and additive Gaussian noise vectors introduced at the relay node, respectively.

In the second time slot, the transmitter remains silent and the relay node multiplies (linearly precodes) the received signal vector $\mathbf{y}_\mathrm{r}$ by an $N_\mathrm{r} \times N_\mathrm{r}$ relay precoding matrix $\mathbf{F}$ and multicasts the precoded signal vector

$$\mathbf{x}_\mathrm{r} = \mathbf{F}\mathbf{y}_\mathrm{r} \tag{3.2}$$

to all L receivers. From (3.1) and (3.2), the received signal vector at the ith receiver can be written as

$$\mathbf{y}_{\mathrm{d},i} = \mathbf{G}_i\mathbf{F}\left(\mathbf{HBs} + \mathbf{n}_\mathrm{r}\right) + \mathbf{n}_{\mathrm{d},i} \triangleq \mathbf{A}_i\mathbf{s} + \mathbf{n}_i, \qquad i = 1, \cdots, L \tag{3.3}$$

where $\mathbf{G}_i$ is the $N_{\rm d} \times N_{\rm r}$ MIMO channel matrix between the relay node and the ith receiver and $\mathbf{n}_{{\rm d},i}$ is the additive Gaussian noise vector at the ith receiver. Here $\mathbf{A}_i \triangleq \mathbf{G}_i\mathbf{FHB}$ is the equivalent MIMO channel between the transmitter and the ith receiver, and $\mathbf{n}_i \triangleq \mathbf{G}_i\mathbf{F}\mathbf{n}_{\rm r} + \mathbf{n}_{{\rm d},i}$ is the equivalent noise vector at the ith receiver. All noises are independent and identically distributed (i.i.d.) complex circularly symmetric Gaussian noise with zero mean and unit variance.

We assume that all channels are quasi-static, i.e., the channel matrices $\mathbf{H}$ and $\mathbf{G}_i, i = 1, \cdots, L$, are constant throughout a block of transmission. In practice, the CSI of $\mathbf{G}_i, i = 1, \cdots, L$, can be obtained at the ith receiver through standard training methods. The relay node can have the CSI of $\mathbf{H}$ through channel training, and obtain the CSI of $\mathbf{G}_i, i = 1, \cdots, L$, by a feedback from the ith receiver. The quasi-static channel model is valid in practice when the mobility among all communicating nodes is relatively slow. Therefore, we can obtain the necessary CSI with a reasonably high precision during the channel training period. The relay node calculates the optimal transmit matrix $\mathbf{B}$ and the relay matrix $\mathbf{F}$, and forwards $\mathbf{B}$ to the transmitter and forwards $\mathbf{F}$ and $\mathbf{H}$ to all receivers. Note that the transmitter does not need any channel knowledge and each receiver only needs the CSI of its own channel with the relay and that of the first-hop channel. This is a very important assumption for multicasting communication since in a multicasting scenario the receivers are distributed and cannot cooperate.

We aim at improving the system performance through optimizing the transmit and relay precoding matrices. Usually, the system performance is quantified by its QoS and the resources it consumes. The most commonly used QoS metrics include the MSE of the signal waveform estimation, BER, system capacity and the output SNR. Interestingly, the aforementioned QoS measures can be expressed in terms of MSE [42]. On the other hand, resources that a multicasting system consumes include the spectrum and transmission power. In the next section, we study two types of optimization problems for two-hop multicasting MIMO relay systems. The first problem deals with minimizing the MSE of the signal waveform estimation subjecting to transmission power constraints at the transmitter and the relay node. While the second problem investigates how to achieve given MSEs using a minimal possible total network transmission power.

3.3 Proposed Transmitter and Relay Design Algorithms

Due to its simplicity, a linear receiver is used at each receiver to retrieve the transmitted signals. Denoting $\mathbf{W}_i$ as an $N_\mathrm{d} \times N_\mathrm{b}$ weight matrix at the ith receiver, the estimated signal vector $\hat{\mathbf{s}}_i$ is given by

$$\hat{\mathbf{s}}_i = \mathbf{W}_i^H \mathbf{y}_{\mathrm{d},i}, \qquad i = 1, \cdots, L\,. \tag{3.4}$$

From (3.4), the MSE of the signal waveform estimation at the ith receiver is given by

$$\begin{aligned} E_i &= \mathrm{tr}\big(\mathrm{E}\left[(\hat{\mathbf{s}}_i - \mathbf{s})(\hat{\mathbf{s}}_i - \mathbf{s})^H\right]\big) \\ &= \mathrm{tr}\big((\mathbf{W}_i^H\mathbf{A}_i - \mathbf{I}_{N_\mathrm{b}})(\mathbf{W}_i^H\mathbf{A}_i - \mathbf{I}_{N_\mathrm{b}})^H + \mathbf{W}_i^H\mathbf{C}_i\mathbf{W}_i\big)\,, \qquad i = 1, \cdots, L \end{aligned} \tag{3.5}$$

where $\mathbf{C}_i \triangleq \mathrm{E}[\mathbf{n}_i\mathbf{n}_i^H] = \mathbf{G}_i\mathbf{F}\mathbf{F}^H\mathbf{G}_i^H + \mathbf{I}_{N_\mathrm{d}}$ is the covariance matrix of $\mathbf{n}_i$. Obviously, the power consumed by the transmitter is $\mathrm{tr}(\mathbf{B}\mathbf{B}^H)$. And from (3.1) and (3.2), the transmission power consumed by the relay node is given by $\mathrm{tr}(\mathbf{F}(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_\mathrm{r}})\mathbf{F}^H)$. In the following, we consider two design strategies for optimizing the transmit and relay matrices. The first optimization strategy is to minimize the maximal MSE among all receivers subjecting to power constraints at the transmitter and the relay node. The second strategy minimizes the total network transmission power subjecting to QoS constraints.

3.3.1 Min-max MSE-based transmitter and relay design

Given the power constraints at the transmitter and the relay node, we aim at minimizing the maximal MSE of the signal waveform estimations among all receivers. This problem formulation is important when the power consumption is a strict system constraint that cannot be relaxed. In this case, the transmit, relay, and receive matrices optimization problem can be formulated as

$$\min_{\mathbf{B},\mathbf{F},\{\mathbf{W}_i\}} \quad \max_i \quad E_i \tag{3.6a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{F}(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_\mathrm{r}})\mathbf{F}^H) \le P_\mathrm{r} \tag{3.6b}$$

$$\mathrm{tr}(\mathbf{B}\mathbf{B}^H) \le P_\mathrm{s} \tag{3.6c}$$

where $\{\mathbf{W}_i\} \triangleq \{\mathbf{W}_i, i = 1, \cdots, L\}$, (3.6b) and (3.6c) are the transmission power constraints at the relay node and the transmitter, respectively, and $P_\mathrm{r} > 0$, $P_\mathrm{s} > 0$ are the corresponding power budgets.

For *any* given $\mathbf{B}$ and $\mathbf{F}$, the receiver $\mathbf{W}_i$ minimizing E_i in (3.5) is the linear MMSE filter [60] and given by

$$\mathbf{W}_i = \left(\mathbf{A}_i\mathbf{A}_i^H + \mathbf{C}_i\right)^{-1}\mathbf{A}_i, \qquad i = 1, \cdots, L. \tag{3.7}$$

By substituting (3.7) back into (3.5), we have

$$\begin{aligned} E_i &= \mathrm{tr}\Big(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{A}_i^H\mathbf{C}_i^{-1}\mathbf{A}_i\right]^{-1}\Big) \\ &= \mathrm{tr}\Big(\Big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{B}^H\mathbf{H}^H\mathbf{F}^H\mathbf{G}_i^H\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^H\mathbf{G}_i^H + \mathbf{I}_{N_\mathrm{d}}\right)^{-1}\mathbf{G}_i\mathbf{F}\mathbf{H}\mathbf{B}\Big]^{-1}\Big), i = 1, \cdots, L. \end{aligned} \tag{3.8}$$

Therefore, we can equivalently rewrite the problem (3.6) as

$$\min_{\mathbf{B},\mathbf{F}} \quad \max_i \quad \mathrm{tr}\Big(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{A}_i^H\mathbf{C}_i^{-1}\mathbf{A}_i\right]^{-1}\Big) \tag{3.9a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{F}(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_\mathrm{r}})\mathbf{F}^H) \le P_\mathrm{r} \tag{3.9b}$$

$$\mathrm{tr}(\mathbf{B}\mathbf{B}^H) \le P_\mathrm{s}. \tag{3.9c}$$

The min-max problem (3.9) is highly nonconvex with matrix variables, and an exactly optimal solution is very hard to obtain with a reasonable computational complexity (non-exhaustive searching). In the following, we propose a low complexity solution to the problem (3.9).

It can be shown similar to [81] that for any $\mathbf{B}$, the optimal $\mathbf{F}$ for *each link* with the input-output relationship given by (3.3) has the generic structure of

$$\mathbf{F} = \mathbf{T}\mathbf{D}^H \tag{3.10}$$

where $\mathbf{D} = (\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_\mathbf{r}})^{-1}\mathbf{H}\mathbf{B}$. Interestingly, $\mathbf{D}$ can be viewed as the weight matrix of the linear MMSE receiver for the first-hop MIMO channel at the relay node given by (3.1), and $\mathbf{T}$ can be treated as the transmit precoding matrix for the effective second-hop MIMO multicasting system $\mathbf{y}_i = \mathbf{G}_i\mathbf{T}\mathbf{x} + \mathbf{v}_i$, $i = 1, \cdots, L$, where $\mathbf{x}$ is the transmitted signal vector and $\mathbf{v}_i$ is the noise vector.

Using the optimal $\mathbf{F}$ in (3.10), the MSE of the signal waveform estimation at the ith receiver in (3.8) can be equivalently rewritten as the sum of two individual MSEs

$$E_i = \mathrm{tr}\Big(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\right]^{-1}\Big) + \mathrm{tr}\Big(\left[\mathbf{R}^{-1} + \mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\right]^{-1}\Big), \quad i = 1, \cdots, L \tag{3.11}$$

where

$$\mathbf{R} = \mathbf{B}^H\mathbf{H}^H(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_\mathrm{r}})^{-1}\mathbf{H}\mathbf{B}. \tag{3.12}$$

Note that the first term in (3.11) $\mathrm{tr}\big(\big[\mathbf{I}_{N_\mathrm{b}}+\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\big]^{-1}\big)$ is actually the MSE of estimating the signal vector $\mathbf{s}$ from the received signal vector (3.1) at the relay node using the linear MMSE receiver $\mathbf{D}$, while the second term in (3.11) $\mathrm{tr}\big(\big[\mathbf{R}^{-1}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\big)$ can be viewed as the increment of the MSE introduced by the second-hop. Interestingly, matrix $\mathbf{R}$ in (3.12) is in fact the covariance matrix of $\mathbf{z}\triangleq\mathbf{D}^H\mathbf{y}_\mathrm{r}$ as $\mathbf{R}=\mathrm{E}[\mathbf{z}\mathbf{z}^H]=\mathbf{D}^H\mathrm{E}[\mathbf{y}_\mathrm{r}\mathbf{y}_\mathrm{r}^H]\mathbf{D}$. It can be seen from (3.11) that the effect of noise in the first-hop is reflected by $\mathbf{I}_{N_\mathrm{b}}$ in the first term and that of the second-hop is reflected by $\mathbf{R}^{-1}$ in the second term. Using the optimal structure of $\mathbf{F}$ in (3.10), the relay power consumption $\mathrm{tr}(\mathbf{F}(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H+\mathbf{I}_{N_\mathrm{r}})\mathbf{F}^H)$ is equivalent to $\mathrm{tr}(\mathbf{T}\mathbf{R}\mathbf{T}^H)$. Therefore, the problem (3.9) can be equivalently rewritten as

$$\min_{\mathbf{B},\mathbf{T}}\ \max_{i}\quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}}+\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\big]^{-1}\Big)+\mathrm{tr}\Big(\big[\mathbf{R}^{-1}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\Big) \tag{3.13a}$$

$$\text{s.t.}\quad \mathrm{tr}(\mathbf{T}\mathbf{R}\mathbf{T}^H)\le P_\mathrm{r} \tag{3.13b}$$

$$\mathrm{tr}(\mathbf{B}\mathbf{B}^H)\le P_\mathrm{s}. \tag{3.13c}$$

By applying the matrix inversion lemma $(\mathbf{A}+\mathbf{B}\mathbf{C}\mathbf{D})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}\mathbf{A}^{-1}\mathbf{B}+\mathbf{C}^{-1}\right)^{-1}\mathbf{D}\mathbf{A}^{-1}$, matrix $\mathbf{R}$ can be rewritten as

$$\begin{aligned}\mathbf{R}&=\mathbf{B}^H\mathbf{H}^H\left(\mathbf{I}_{N_\mathrm{r}}-\mathbf{H}\mathbf{B}\left(\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}+\mathbf{I}_{N_\mathrm{b}}\right)^{-1}\mathbf{B}^H\mathbf{H}^H\right)\mathbf{H}\mathbf{B}\\&=\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\left(\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}+\mathbf{I}_{N_\mathrm{b}}\right)^{-1}.\end{aligned} \tag{3.14}$$

An interesting observation from (3.14) is that with increasing first-hop SNR, the term $\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}$ approaches to infinity. And at a (moderately) high SNR level, there is $\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\gg\mathbf{I}_{N_\mathrm{b}}$. Thus we can approximate $\mathbf{R}$ as $\mathbf{I}_{N_\mathrm{b}}$ for the high SNR case [82]. As a consequence, $\mathrm{tr}\big(\big[\mathbf{R}^{-1}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\big)$ in (3.13a) is upper-bounded by $\mathrm{tr}\big(\big[\mathbf{I}_{N_\mathrm{b}}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\big)$, $i=1,\cdots,L$, and the tightness of this bound increases with the first-hop SNR. Therefore, the problem (3.13) can be approximated as

$$\min_{\mathbf{B},\mathbf{T}}\ \max_{i}\quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}}+\mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\big]^{-1}\Big)+\mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\Big) \tag{3.15a}$$

$$\text{s.t.}\quad \mathrm{tr}(\mathbf{T}\mathbf{T}^H)\le P_\mathrm{r} \tag{3.15b}$$

$$\mathrm{tr}(\mathbf{B}\mathbf{B}^H)\le P_\mathrm{s}. \tag{3.15c}$$

We would like to mention that since $\mathrm{tr}(\mathbf{T}\mathbf{T}^H)>\mathrm{tr}(\mathbf{T}\mathbf{R}\mathbf{T}^H)$, if $\mathrm{tr}(\mathbf{T}\mathbf{T}^H)=p$, then $\mathrm{tr}(\mathbf{T}\mathbf{R}\mathbf{T}^H)<p$. This indicates that due to the approximation in (3.15b), the transmission power available at the relay node is not fully utilized in the case of the low first-hop SNR. We can simply scale the relay matrix obtained from solving the problem (3.15) to

compensate such loss and make the best use of the available power budget at the relay node.

Interestingly, it can be seen from the problem (3.15) that $\mathbf{T}$ does not affect the first term of the objective function (3.15a) and $\mathbf{B}$ is irrelevant to the second term of (3.15a). This fact implies that the objective function (3.15a) and the constraints (3.15b) and (3.15c) are decoupled with respect to the optimization variables $\mathbf{B}$ and $\mathbf{T}$. In this case, matrix $\mathbf{B}$ can be determined independent of $\mathbf{T}$, and vice-versa, which greatly simplifies the design of the transmit and relay matrices. Therefore, with the (relatively) high first-hop SNR assumption, the problem (3.15) can be decomposed into the following transmit precoding matrix optimization problem

$$\min_{\mathbf{B}} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\big]^{-1}\Big) \tag{3.16a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{B}\mathbf{B}^H) \le P_\mathrm{s} \tag{3.16b}$$

and the relay precoding matrix optimization problem

$$\min_{\mathbf{T}} \quad \max_{i} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\Big) \tag{3.17a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{T}\mathbf{T}^H) \le P_\mathrm{r}. \tag{3.17b}$$

Let $\mathbf{H} = \mathbf{U}_\mathrm{h}\mathbf{\Lambda}_\mathrm{h}\mathbf{V}_\mathrm{h}^H$ denote the SVD of $\mathbf{H}$, where the dimensions of $\mathbf{U}_\mathrm{h}$, $\mathbf{\Lambda}_\mathrm{h}$, and $\mathbf{V}_\mathrm{h}$ are $N_\mathrm{r} \times N_\mathrm{r}$, $N_\mathrm{r} \times N_\mathrm{s}$, and $N_\mathrm{s} \times N_\mathrm{s}$, respectively. We assume that the main diagonal elements of $\mathbf{\Lambda}_\mathrm{h}$ are arranged in decreasing order. According to Lemma 2 in [81], the transmitter optimization problem (3.16) has a closed form solution with the optimal structure of $\mathbf{B}$ given by

$$\mathbf{B} = \mathbf{V}_{\mathrm{h},1}\mathbf{\Lambda}_\mathrm{b} \tag{3.18}$$

where $\mathbf{V}_{\mathrm{h},1}$ contains the leftmost N_b columns of $\mathbf{V}_\mathrm{h}$ and $\mathbf{\Lambda}_\mathrm{b}$ is an $N_\mathrm{b} \times N_\mathrm{b}$ diagonal power loading matrix. Substituting the optimal $\mathbf{B}$ in (3.18) back into (3.16) and using the Lagrangian multiplier method [61], we find that the ith diagonal element of $\mathbf{\Lambda}_\mathrm{b}$ is given by

$$\lambda_{\mathrm{b},i} = \left[\frac{1}{\lambda_{\mathrm{h},i}}\Big(\sqrt{\lambda_{\mathrm{h},i}/\mu} - 1\Big)^+\right]^{\frac{1}{2}}, \qquad i = 1, \cdots, N_\mathrm{b}$$

where $(x)^+ \triangleq \max(x, 0)$, $\lambda_{\mathrm{h},i}$ is the ith diagonal element of $\mathbf{\Lambda}_\mathrm{h}$, and $\mu > 0$ is the Lagrangian multiplier and the solution to the nonlinear equation of $\sum_{i=1}^{N_\mathrm{b}} \frac{1}{\lambda_{\mathrm{h},i}}\big(\sqrt{\lambda_{\mathrm{h},i}/\mu} - 1\big)^+ = P_\mathrm{s}$.

By introducing $\mathbf{T}\mathbf{T}^H \triangleq \mathbf{Q}$, the problem (3.17) can be equivalently rewritten as

$$\min_{\mathbf{Q}} \quad \max_{i} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{d}} + \mathbf{G}_i\mathbf{Q}\mathbf{G}_i^H\big]^{-1}\Big) + N_\mathrm{b} - N_\mathrm{d} \tag{3.19a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{Q}) \leq P_\mathrm{r} \tag{3.19b}$$

$$\mathbf{Q} \succcurlyeq 0. \tag{3.19c}$$

By introducing $\big[\mathbf{I}_{N_\mathrm{d}} + \mathbf{G}_i\mathbf{Q}\mathbf{G}_i^H\big]^{-1} \preccurlyeq \mathbf{Y}_i, i = 1, \cdots, L$, and a real-valued slack variable t, the problem (3.19) can be equivalently transformed to

$$\min_{t,\mathbf{Q},\{\mathbf{Y}_i\}} \quad t \tag{3.20a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{Y}_i) \leq t, \qquad i = 1, \cdots, L \tag{3.20b}$$

$$\mathrm{tr}(\mathbf{Q}) \leq P_\mathrm{r} \tag{3.20c}$$

$$\begin{pmatrix} \mathbf{Y}_i & \mathbf{I}_{N_\mathrm{d}} \\ \mathbf{I}_{N_\mathrm{d}} & \mathbf{I}_{N_\mathrm{d}} + \mathbf{G}_i\mathbf{Q}\mathbf{G}_i^H \end{pmatrix} \succcurlyeq 0, \qquad i = 1, \cdots, L \tag{3.20d}$$

$$t \geq 0, \qquad \mathbf{Q} \succcurlyeq 0 \tag{3.20e}$$

where $\{\mathbf{Y}_i\} \triangleq \{\mathbf{Y}_i, i = 1, \cdots, L\}$ and we use the Schur complement to obtain (3.20d). Note that in the above formulation, t provides an MSE upper-bound for the relay-receiver channels. The problem (3.20) is an SDP problem which can be efficiently solved by the disciplined convex programming toolbox CVX [63], where interior-point method-based solvers such as SeDuMi or SDPT3 are called internally, at a maximal complexity order of $\mathcal{O}\big((N_\mathrm{r}^2 + L + 1)^{3.5}\big)$ [64]. Since most of the computation task in solving the problem (3.16) involves performing SVD and calculating the power loading parameters, the computation overhead is negligible compared with that of solving the problem (3.20). Note that the problem (3.15) can also be directly formulated as an SDP problem which can be solved using interior point-based solvers at a complexity order that is at most $\mathcal{O}\big((N_\mathrm{s}^2 + N_\mathrm{r}^2 + L + 2)^{3.5}\big)$. Therefore, solving the decoupled transmit and relay precoding problems (3.16) and (3.17) has a much smaller computational complexity compared with directly solving the problem (3.15).

3.3.2 Minimal total power-based transmitter and relay design

In this scheme, we aim at minimizing the total transmitter and relay power consumption satisfying the target QoS requirements at all receivers. This criterion is important when certain QoS level must be maintained at each receiver. We choose the MSE of the signal waveform estimation at the receiver as the QoS metric. The multicasting MIMO

relay system tries to satisfy the given required QoS (maximal allowable MSE) with the minimal total transmission power. Thus the optimization problem can be written as

$$\min_{\mathbf{B},\mathbf{F}} \quad \mathrm{tr}(\mathbf{F}(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_\mathrm{r}})\mathbf{F}^H) + \mathrm{tr}(\mathbf{B}\mathbf{B}^H) \tag{3.21a}$$

$$\text{s.t.} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{A}_i^H\mathbf{C}_i^{-1}\mathbf{A}_i\big]^{-1}\Big) \le \varepsilon_i, \qquad i = 1, \cdots, L \tag{3.21b}$$

where ε_i is the maximal allowable MSE at the ith receiver. Similar to the problem (3.9), the problem (3.21) is highly nonconvex with matrix variables.

It can be shown similar to [81] that the optimal structure of $\mathbf{F}$ for the problem (3.21) is given by (3.10). Then, the problem (3.21) can be equivalently converted to the following problem

$$\min_{\mathbf{B},\mathbf{T}} \quad \mathrm{tr}(\mathbf{T}\mathbf{R}\mathbf{T}^H) + \mathrm{tr}(\mathbf{B}\mathbf{B}^H) \tag{3.22a}$$

$$\text{s.t.} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\big]^{-1}\Big) + \mathrm{tr}\Big(\big[\mathbf{R}^{-1} + \mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\Big) \le \varepsilon_i, \quad i = 1, \cdots, L. \tag{3.22b}$$

Similar to Section 3.3.1, we have $\mathrm{tr}\big(\big[\mathbf{R}^{-1}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\big) \le \mathrm{tr}\big(\big[\mathbf{I}_{N_\mathrm{b}}+\mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\big)$, and at a relatively high first-hop SNR, the problem (3.22) can be approximated (relaxed) to the problem of

$$\min_{\mathbf{B},\mathbf{T}} \quad \mathrm{tr}(\mathbf{T}\mathbf{T}^H) + \mathrm{tr}(\mathbf{B}\mathbf{B}^H) \tag{3.23a}$$

$$\text{s.t.} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{B}^H\mathbf{H}^H\mathbf{H}\mathbf{B}\big]^{-1}\Big) + \mathrm{tr}\Big(\big[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{T}\big]^{-1}\Big) \le \varepsilon_i, \qquad i = 1, \cdots, L. \tag{3.23b}$$

By introducing $\mathbf{B}\mathbf{B}^H \triangleq \mathbf{P}$ and $\big[\mathbf{I}_{N_\mathrm{r}} + \mathbf{H}\mathbf{P}\mathbf{H}^H\big]^{-1} \preccurlyeq \mathbf{X}$, the problem (3.23) can be transformed to the following SDP problem

$$\min_{\mathbf{P},\mathbf{Q},\mathbf{X},\{\mathbf{Y}_i\}} \quad \mathrm{tr}(\mathbf{P}) + \mathrm{tr}(\mathbf{Q}) \tag{3.24a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{X}) + \mathrm{tr}(\mathbf{Y}_i) \le \varepsilon_i + N_\mathrm{r} - N_\mathrm{b}, \qquad i = 1, \cdots, L \tag{3.24b}$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{I}_{N_\mathrm{r}} \\ \mathbf{I}_{N_\mathrm{r}} & \mathbf{I}_{N_\mathrm{r}} + \mathbf{H}\mathbf{P}\mathbf{H}^H \end{pmatrix} \succcurlyeq 0 \tag{3.24c}$$

$$\begin{pmatrix} \mathbf{Y}_i & \mathbf{I}_{N_\mathrm{d}} \\ \mathbf{I}_{N_\mathrm{d}} & \mathbf{I}_{N_\mathrm{d}} + \mathbf{G}_i\mathbf{Q}\mathbf{G}_i^H \end{pmatrix} \succcurlyeq 0, \qquad i = 1, \cdots, L \tag{3.24d}$$

$$\mathbf{P} \succcurlyeq 0, \qquad \mathbf{Q} \succcurlyeq 0 \tag{3.24e}$$

where the Schur complement is used to obtain (3.24c). Note that unlike the problem (3.15), the problem (3.24) is not decoupled due to the constraint (3.24b) which couples

$\mathrm{tr}(\mathbf{X})$ and $\mathrm{tr}(\mathbf{Y}_i)$. However, we can use the CVX software package [63] to solve the SDP problem (3.24) at a complexity order that is at most $\mathcal{O}\big((N_\mathrm{s}^2+N_\mathrm{r}^2+L)^{3.5}\big)$ and is usually much less [66].

We would like to mention that for a point-to-point two-hop MIMO relay system, it has been shown in [82] through numerical examples that the high first-hop SNR approximation provides negligible performance loss in all SNR range in comparison to the optimal designs. For the multicasting MIMO relay system addressed in this chapter, the exactly optimal solution for the transmit and relay precoding matrices are intractable. However, by using the high first-hop SNR approximation, the nearly optimal transmit and relay matrices can be designed with a significantly reduced computational complexity.

3.4 Transmitter and Relay Optimization for Single Data Stream Multicasting

In this section, we derive the solution for the min-max MSE and the total power minimization problems for single data stream multicasting in two-hop MIMO relay systems. In the case where the transmitter multicasts a single data stream s, i.e., $N_\mathrm{b}=1$, the received signal vector at the ith receiver can be written as

$$\mathbf{y}_{\mathrm{d},i}=\mathbf{G}_i\mathbf{F}\left(\mathbf{H}\mathbf{b}s+\mathbf{n}_\mathrm{r}\right)+\mathbf{n}_{\mathrm{d},i},\qquad i=1,\cdots,L \tag{3.25}$$

where $\mathbf{b}$ is the $N_\mathrm{s}\times 1$ transmit beamforming vector. In the following, we solve the min-max MSE and the total network transmission power minimization problems for this system. In particular, we show that the relay precoding matrix design for the MSE minimization problem is equivalent to the transmit beamforming problem for single-hop multicasting [66] and the power minimization problem can be solved by a combination of one-dimensional searching and SDR techniques.

3.4.1 Min-max MSE-based transmitter and relay design

Using a linear MMSE receiver $\mathbf{w}_i$ ((3.7) with $N_\mathrm{b}=1$) at the ith receiver to estimate s from (3.25), it can be shown similar to (3.8) that the MSE of the signal waveform estimation can be written as

$$e_i=\left[1+\mathbf{b}^H\mathbf{H}^H\mathbf{F}^H\mathbf{G}_i^H\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^H\mathbf{G}_i^H+\mathbf{I}_{N_\mathrm{d}}\right)^{-1}\mathbf{G}_i\mathbf{F}\mathbf{H}\mathbf{b}\right]^{-1},\qquad i=1,\cdots,L. \tag{3.26}$$

Similar to Section 3.3.1, the worst-case MSE minimization problem for a single data stream multicasting MIMO relay system can be formulated as

$$\min_{\mathbf{b},\mathbf{F}} \quad \max_{i} \quad \left[1+\mathbf{b}^H\mathbf{H}^H\mathbf{F}^H\mathbf{G}_i^H\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^H\mathbf{G}_i^H+\mathbf{I}_{N_\mathrm{d}}\right)^{-1}\mathbf{G}_i\mathbf{F}\mathbf{H}\mathbf{b}\right]^{-1} \tag{3.27a}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{F}(\mathbf{H}\mathbf{b}\mathbf{b}^H\mathbf{H}^H+\mathbf{I}_{N_\mathrm{r}})\mathbf{F}^H) \le P_\mathrm{r} \tag{3.27b}$$

$$\mathbf{b}^H\mathbf{b} \le P_\mathrm{s}. \tag{3.27c}$$

We would like to mention that for the single data stream multicasting case, minimizing the MSE of the signal waveform estimation is equivalent to maximizing the receive SNR, and thus, the transmitter-receiver MI.

Let $\mathbf{G}_i = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$ denote the SVD of $\mathbf{G}_i, i = 1,\cdots,L$, where the dimensions of $\mathbf{U}_i$, $\mathbf{\Lambda}_i$, and $\mathbf{V}_i$ are $N_\mathrm{d}\times N_\mathrm{d}$, $N_\mathrm{d}\times N_\mathrm{r}$, and $N_\mathrm{r}\times N_\mathrm{r}$, respectively. We assume that the main diagonal elements of $\mathbf{\Lambda}_i$ are arranged in decreasing order. According to the unified framework developed in [27], the optimal transmitter forms a "beam" along the direction of the strongest singular value of the source-relay channel and hence, the transmit beamforming vector is given by

$$\mathbf{b} = \mathbf{v}_{\mathrm{h},1}\sqrt{P_\mathrm{s}} \tag{3.28}$$

where $\mathbf{v}_{\mathrm{h},1}$ is the leftmost column of $\mathbf{V}_\mathrm{h}$. Moreover, the optimal relay precoding matrix is given by $\mathbf{F} = \alpha_i\mathbf{v}_{i,1}\mathbf{u}_{\mathrm{h},1}^H$ for the ith receiver [27], where $\alpha_i, i = 1,\cdots,L$, is the power loading factor at the relay node, $\mathbf{u}_{\mathrm{h},1}$ and $\mathbf{v}_{i,1}$ are the leftmost columns of $\mathbf{U}_\mathrm{h}$ and $\mathbf{V}_i, i = 1,\cdots,L$, respectively.

Although the optimal $\mathbf{F}$ for the ith receiver depends on $\mathbf{G}_i$, it can be seen from the discussion above that for *each user*, $\mathbf{F}$ has the generic optimal structure of

$$\mathbf{F} = \mathbf{f}\mathbf{u}_{\mathrm{h},1}^H \tag{3.29}$$

where vector $\mathbf{f}$ remains to be optimized. Using (3.28) and (3.29), the MSE in (3.26) can be rewritten as

$$e_i = \left[1+P_\mathrm{s}\lambda_{\mathrm{h},1}^2\mathbf{f}^H\mathbf{G}_i^H\left(\mathbf{G}_i\mathbf{f}\mathbf{f}^H\mathbf{G}_i^H+\mathbf{I}_{N_\mathrm{d}}\right)^{-1}\mathbf{G}_i\mathbf{f}\right]^{-1}, \qquad i = 1,\cdots,L \tag{3.30}$$

where $\lambda_{\mathrm{h},1}$ is the largest singular value of $\mathbf{H}$. From (3.28) and (3.29), the relay transmit power (3.27b) becomes

$$\mathrm{tr}\big(P_\mathrm{s}\lambda_{\mathrm{h},1}^2\mathbf{f}\mathbf{f}^H+\mathbf{f}\mathbf{f}^H\big) = \mathbf{f}^H\mathbf{f}\big(P_\mathrm{s}\lambda_{\mathrm{h},1}^2+1\big).$$

Applying the matrix inversion lemma, (3.30) can be equivalently rewritten as

$$e_i = \left[1 + \frac{P_{\rm s}\lambda_{{\rm h},1}^2 \mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f}}{\mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f} + 1}\right]^{-1} = \left[1 + P_{\rm s}\lambda_{{\rm h},1}^2 - \frac{P_{\rm s}\lambda_{{\rm h},1}^2}{\mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f} + 1}\right]^{-1}, \; i = 1, \cdots, L. \quad (3.31)$$

Note that minimizing the MSE in (3.31) is equivalent to maximizing $\mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f}$ which in fact determines the SNR at the ith receiver in a single-hop (relay-receiver channel in this case) MIMO multicasting system. Therefore, the min-max MSE problem (3.27) is equivalent to the following max-min SNR optimization problem

$$\max_{\mathbf{f}} \quad \min_{i} \quad \mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f} \quad (3.32a)$$

$$\text{s.t.} \quad \mathbf{f}^H \mathbf{f} \le \frac{P_{\rm r}}{P_{\rm s}\lambda_{{\rm h},1}^2 + 1}. \quad (3.32b)$$

Interestingly, problem (3.32) is essentially a max-min transmit beamforming problem for a single-hop multicasting system [66], which can be efficiently solved using standard techniques such as rank relaxations at a maximal complexity order of $\mathcal{O}\big((N_{\rm r}^2 + L + 1)^{3.5}\big)$ [66].

3.4.2 Minimal total power-based transmitter and relay design

For single data stream multicasting, the design of the transmit beamformer and the relay precoding matrix that minimize the total network transmission power, subjecting to constraints on the MSE of the signal waveform estimation at each user (3.26), can be written as the following problem

$$\min_{\mathbf{b},\mathbf{F}} \quad \text{tr}(\mathbf{F}(\mathbf{H}\mathbf{b}\mathbf{b}^H\mathbf{H}^H + \mathbf{I}_{N_{\rm r}})\mathbf{F}^H) + \mathbf{b}^H\mathbf{b} \quad (3.33a)$$

$$\text{s.t.} \quad \left[1 + \mathbf{b}^H\mathbf{H}^H\mathbf{F}^H\mathbf{G}_i^H\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^H\mathbf{G}_i^H + \mathbf{I}_{N_{\rm d}}\right)^{-1}\mathbf{G}_i\mathbf{F}\mathbf{H}\mathbf{b}\right]^{-1} \le \varepsilon_i, \qquad i = 1, \cdots, L. \quad (3.33b)$$

It can be shown similar to [83] that the optimal transmit beamformer and relay precoding matrix as the solution to problem (3.33) has the generic structure of $\mathbf{b} = \mathbf{v}_{{\rm h},1}\sqrt{P_1}$ and $\mathbf{F} = \mathbf{f}\mathbf{u}_{{\rm h},1}^H$, respectively, where $P_1 > 0$ is the power required at the transmitter for multicasting a single data stream and remains to be optimized.

Based on the optimal structure of $\mathbf{b}$ and $\mathbf{F}$, and using (3.31), the QoS-constrained problem (3.33) can be equivalently converted to the following problem

$$\min_{P_1,\mathbf{f}} \quad P_1 + \left(P_1\lambda_{{\rm h},1}^2 + 1\right)\mathbf{f}^H\mathbf{f} \quad (3.34a)$$

$$\text{s.t.} \quad \frac{P_1\lambda_{{\rm h},1}^2 \mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f}}{\mathbf{f}^H \mathbf{G}_i^H \mathbf{G}_i \mathbf{f} + 1} \ge \eta_i, \qquad i = 1, \cdots, L \quad (3.34b)$$

where $\eta_i \triangleq \frac{1}{\varepsilon_i} - 1$ can be viewed as the minimal SNR required at the ith receiver. The problem (3.34) is still nonconvex. However, it can be seen from (3.34b) that for a given P_1, there is $\mathbf{f}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{f} \geq \beta_i$ for $\beta_i \triangleq \frac{\eta_i}{P_1\lambda_{\mathrm{h},1}^2 - \eta_i} > 0$, $i = 1, \cdots, L$. Thus, for a given P_1, the problem (3.34) reduces to

$$\min_{\mathbf{f}} \quad \mathbf{f}^H\mathbf{f} \tag{3.35a}$$

$$\text{s.t.} \quad \mathbf{f}^H\mathbf{G}_i^H\mathbf{G}_i\mathbf{f} \geq \beta_i, \qquad i = 1, \cdots, L. \tag{3.35b}$$

The problem (3.35) is equivalent to the minimal transmission power beamforming problem for a single-hop multicasting system [66]. Note that the problem (3.35) can be solved using standard SDR techniques at a complexity order that is at most $\mathcal{O}\big((N_{\mathrm{r}}^2 + L)^{3.5}\big)$ [66].

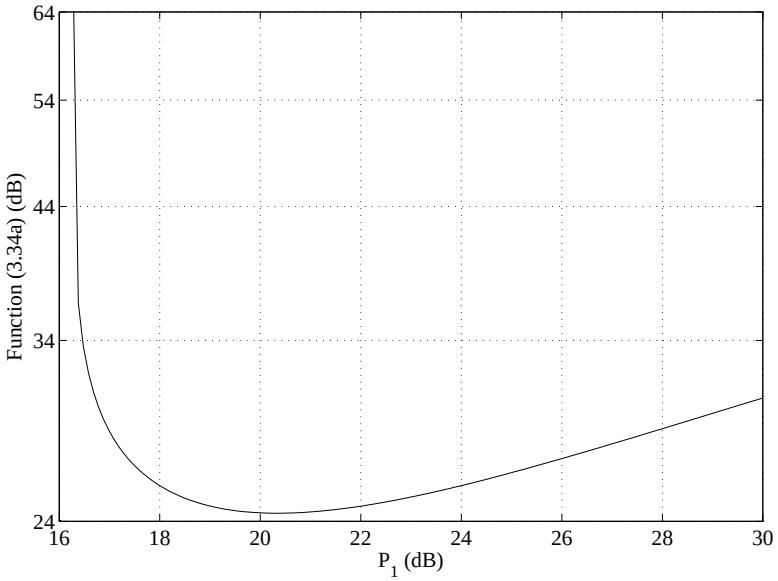

Figure 3.2: Function (3.34a) versus P_1. $L = 6$, $N_{\mathrm{s}} = 6, N_{\mathrm{r}} = 3, N_{\mathrm{d}} = 3$, $\lambda_{\mathrm{h},1} = 1.5345$, $\eta_i = \eta = 20$dB.

Now we show some insights on the optimal P_1 by considering the objective function (3.34a). Obviously, it can be seen from the definition of β_i that $P_1\lambda_{\mathrm{h},1}^2 > \eta_i, i = 1, \cdots, L$. In other words, the lower-bound of the feasible region of P_1 is $\max_i(\eta_i/\lambda_{\mathrm{h},1}^2)$. A plot of the objective function (3.34a) over the range of feasible values of P_1 is generated in Fig. 3.2 for the case of $L = 6$, $N_{\mathrm{s}} = 6, N_{\mathrm{r}} = 3, N_{\mathrm{d}} = 3, \lambda_{\mathrm{h},1} = 1.5345$, and $\eta_i = \eta$ is set to be 20dB. Here for each P_1, the problem (3.35) is solved to obtain the optimal

f. It can be observed from Fig. 3.2 that the objective function (3.34a) is a unimodal function of P_1, i.e., the function has only one local minimum [61]. An efficient method of locating the minimal value of a unimodal function is the golden section searching (GSS) algorithm [84]. Hence, the optimal P_1 for the problem (3.34) can be obtained by applying the GSS technique.

3.5 Numerical Examples

In this section, we study the performance of the proposed two-hop multicasting MIMO relay optimization algorithms through numerical simulations. The transmitter, relay node, and receivers are equipped with $N_{\rm s}$, $N_{\rm r}$, and $N_{\rm d}$ antennas, respectively. We simulate a flat Rayleigh fading environment where the channel matrices have entries with zero mean and variances $1/N_{\rm s}$ and $1/N_{\rm r}$, for $\mathbf{H}$ and $\mathbf{G}_i$, $i = 1, \cdots, L$, respectively. All simulation results are averaged over 500 independent channel realizations.

We compare the performance of the proposed min-max MSE algorithm in Section 3.3.1 with the NAF algorithm and the PMF algorithm in terms of both MSE and BER. For the NAF scheme, we use

$$\mathbf{B} = \sqrt{P_{\rm s}/N_{\rm s}}\,\mathbf{I}_{N_{\rm s}}, \qquad \mathbf{F} = \sqrt{P_{\rm r}/{\rm tr}(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_{\rm r}})}\,\mathbf{I}_{N_{\rm r}}.$$

For the PMF algorithm, the same $\mathbf{B}$ in the NAF algorithm is taken and

$$\mathbf{F} = \sqrt{P_{\rm r}/{\rm tr}((\mathbf{H}\mathbf{G})^H(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{I}_{N_{\rm r}})\mathbf{H}\mathbf{G})}\,(\mathbf{H}\mathbf{G})^H$$

where we randomly pick $\mathbf{G}$ from among the relay-receiver channels $\mathbf{G}_i$, $i = 1, \cdots, L$. Both the NAF and the PMF algorithms adopt the MMSE receiver at all users.

In the first example, we compare the performance of the proposed algorithm with the other two approaches in terms of the MSE normalized by the number of data streams for $L = 4$ and $N_{\rm b} = N_{\rm s} = N_{\rm r} = N_{\rm d} = 3$. Fig. 3.3 shows the MSE performance of all tested algorithms versus $P_{\rm s}$ with $P_{\rm r} = 20$dB, whereas Fig. 3.4 illustrates the MSE performance of tested algorithms versus $P_{\rm r}$ for a $P_{\rm s}$ fixed at 20dB. For the proposed algorithm, we plot the NMSE of the user with the worst channel (Worst) and the average of all the users (Avg.) along with the MSE upper-bound (UB), which is t in (3.20a). Our results clearly demonstrate the better performance of the proposed joint transmitter and relay optimization algorithm. It can be seen that the proposed algorithm consistently yields the lowest average MSE over the entire $P_{\rm s}$ and $P_{\rm r}$ region. The worst-user MSE is always

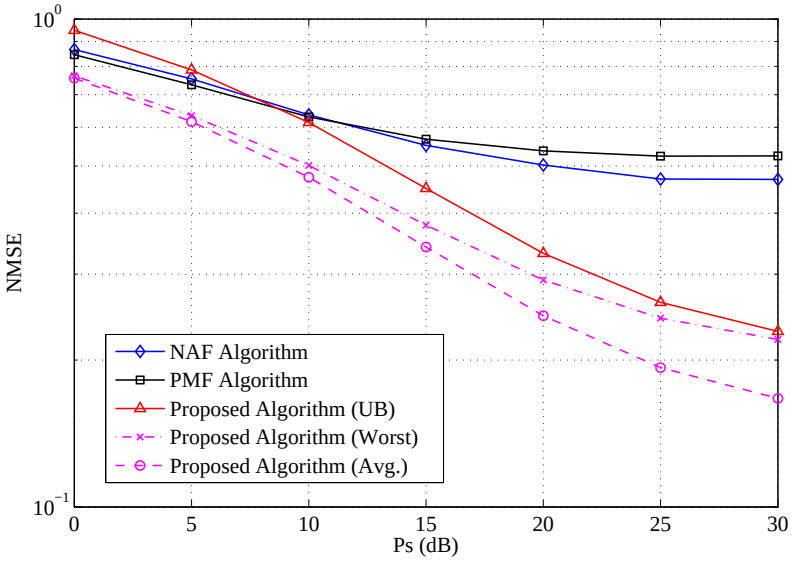

Figure 3.3: Example 1: Normalized MSE versus $P_{\rm s}$. $L = 4$, $N_{\rm b} = N_{\rm s} = N_{\rm r} = N_{\rm d} = 3$, $P_{\rm r} = 20$dB.

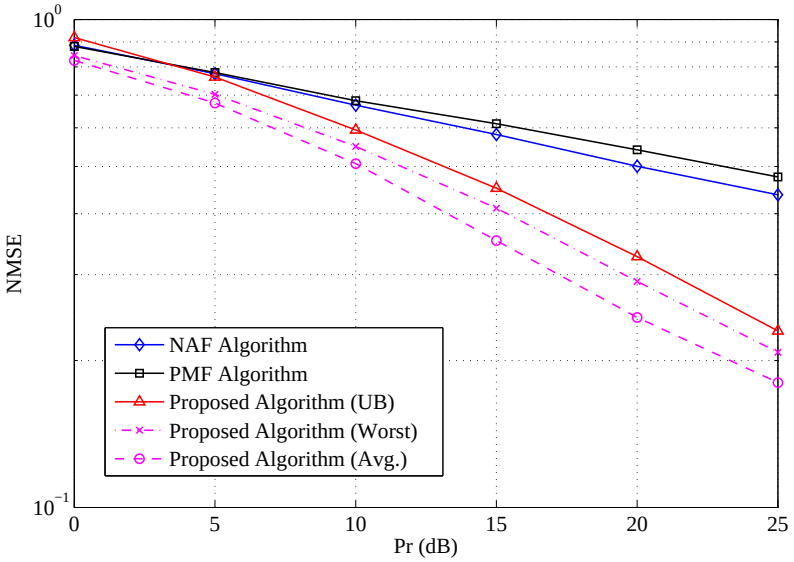

Figure 3.4: Example 1: Normalized MSE versus $P_{\rm r}$. $L = 4$, $N_{\rm b} = N_{\rm s} = N_{\rm r} = N_{\rm d} = 3$, $P_{\rm s} = 20$dB.

better than the MSE upper-bound. The NAF and PMF algorithms have much higher MSE compared with the proposed scheme even with very high transmission power.

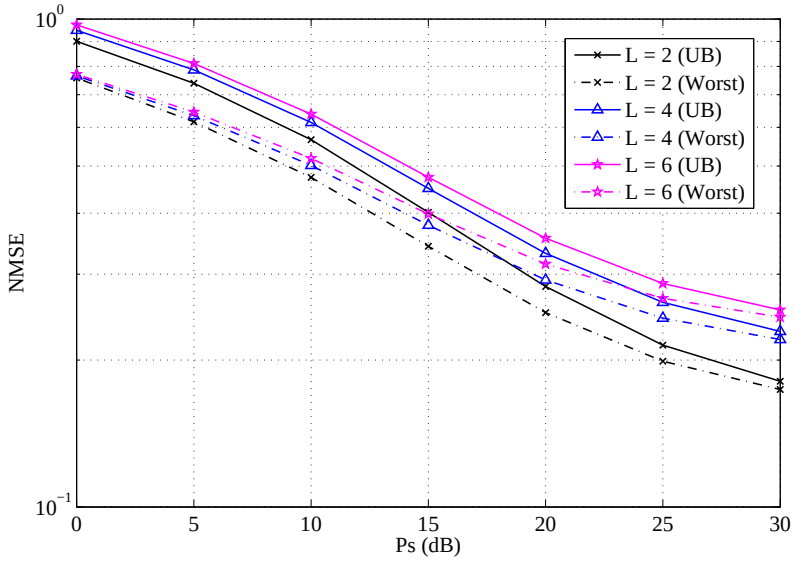

Figure 3.5: Example 2: Normalized MSE versus P_{s}. Varying number of receivers, $N_{\mathrm{b}} = N_{\mathrm{s}} = N_{\mathrm{r}} = N_{\mathrm{d}} = 3$, $P_{\mathrm{r}} = 20$dB.

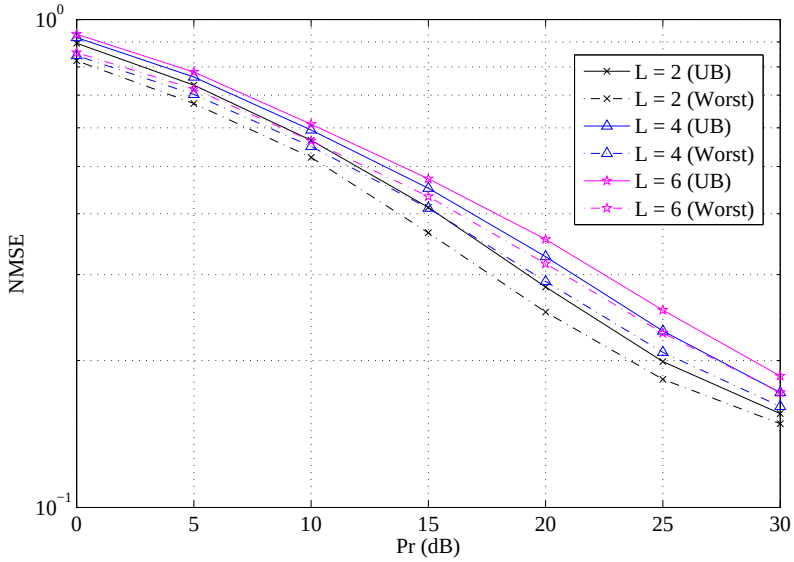

Figure 3.6: Example 2: Normalized MSE versus P_{r}. Varying number of receivers, $N_{\mathrm{b}} = N_{\mathrm{s}} = N_{\mathrm{r}} = N_{\mathrm{d}} = 3$, $P_{\mathrm{s}} = 20$dB.

In the second example, we compare the MSE performance of the proposed algorithm for different number of receivers. We set $N_{\mathrm{b}} = N_{\mathrm{s}} = N_{\mathrm{r}} = N_{\mathrm{d}} = 3$. Fig. 3.5 shows the

MSE upper-bound and the worst-user MSE of the proposed algorithm versus P_{s} with $P_{\mathrm{r}} = 20\mathrm{dB}$, whereas Fig. 3.6 illustrates the same performance versus P_{r} for $P_{\mathrm{s}} = 20\mathrm{dB}$. It can be clearly seen from Figs. 3.5 and 3.6 that as the number of receivers increases, the MSE upper-bound and the worst-user MSE keep increasing. This is reasonable since it is more likely to find a worse relay-receiver channel among the increased number of users and we choose the worst-user MSE as the objective function.

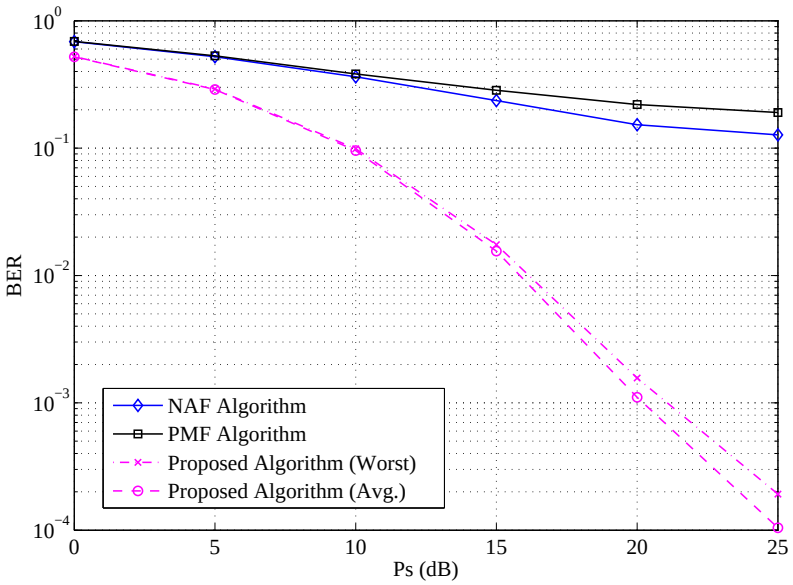

Figure 3.7: Example 3: BER versus P_{s}. $L = 2$, $N_{\mathrm{b}} = 2$, $N_{\mathrm{s}} = 4$, $N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $P_{\mathrm{r}} = 20\mathrm{dB}$.

In the next example, we compare the performance of the min-max MSE algorithm with the NAF and PMF algorithms in terms of BER. QPSK signal constellations are used to modulate the transmitted signals. We set $L = 2$, $N_{\mathrm{b}} = 2$, $N_{\mathrm{s}} = 4$, $N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, and multicast $N_{\mathrm{b}} \times 1000$ randomly generated bits from the transmitter in each channel realization. Fig. 3.7 shows the BER performance of all tested algorithms versus P_{s} with $P_{\mathrm{r}} = 20\mathrm{dB}$, whereas Fig. 3.8 illustrates the BER performance of tested algorithms versus P_{r} for a P_{s} fixed at 20dB. It can be seen from Figs. 3.7 and 3.8 that the proposed joint transmitter and relay optimization algorithm obtains the lowest BER compared with the other approaches. Even the worst-user BER of the proposed algorithm is always much better than that of the NAF and the PMF schemes.

In the next example, we compare the average- and the worse-user BERs of the proposed algorithm for different number of receivers. This time we set $N_{\mathrm{b}} = 2$, $N_{\mathrm{s}} = 4$,

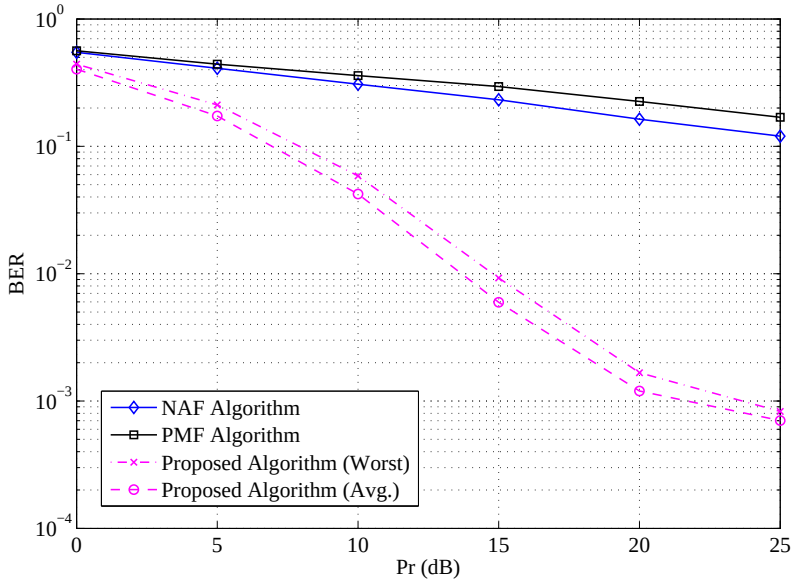

Figure 3.8: Example 3: BER versus P_{r}. $L = 2$, $N_{\mathrm{b}} = 2$, $N_{\mathrm{s}} = 4$, $N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $P_{\mathrm{s}} = 20\mathrm{dB}$.

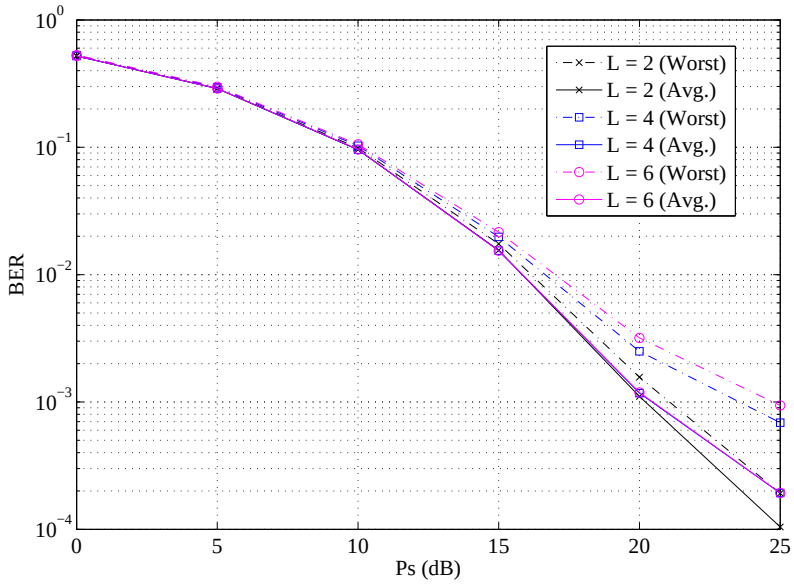

Figure 3.9: Example 4: BER versus P_{s}. Varying number of receivers, $N_{\mathrm{b}} = 2$, $N_{\mathrm{s}} = 4$, $N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $P_{\mathrm{r}} = 20\mathrm{dB}$.

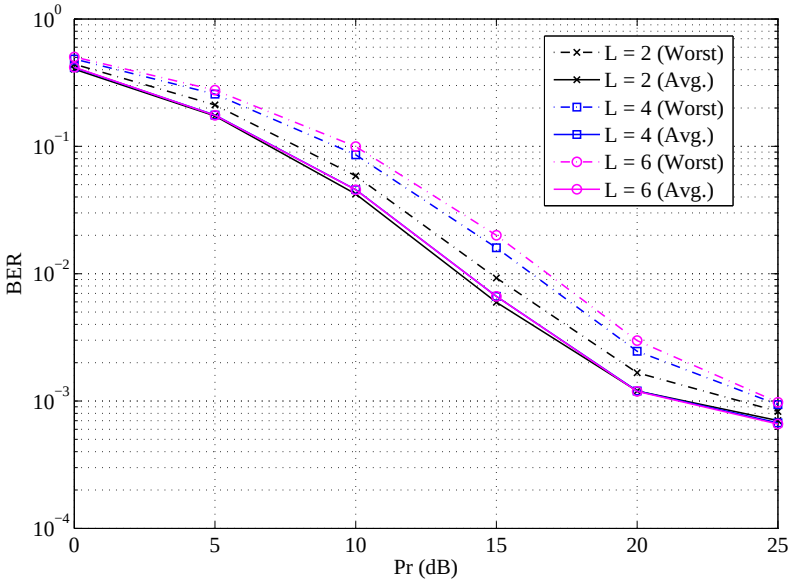

Figure 3.10: Example 4: BER versus $P_{\rm r}$. Varying number of receivers, $N_{\rm b} = 2$, $N_{\rm s} = 4$, $N_{\rm r} = 2$, $N_{\rm d} = 4$, $P_{\rm s} = 20$dB.

$N_{\rm r} = 2$, and $N_{\rm d} = 4$. Fig. 3.9 shows the BER performance of the proposed algorithm versus $P_{\rm s}$ with $P_{\rm r} = 20$dB, whereas Fig. 3.10 illustrates the BER performance versus $P_{\rm r}$ for $P_{\rm s} = 20$dB for different number of receivers. It can be clearly seen from Figs. 3.9 and 3.10 that as we increase the number of receivers, the worst-user BER keeps increasing which is analogous to the MSE performance that we observed in Figs. 3.5 and 3.6. However, the average BERs of the users are almost similar for different number of receivers. Note that the varying $P_{\rm r}$ affects the effective noise at the receiving nodes in Figs. 3.8 and 3.10. Thus there are some error floors of BER in Figs. 3.8 and 3.10, but no error floor can be seen in Fig. 3.9.

In the fifth example, we study the performance of the total power minimization algorithm proposed in Section 3.3.2 for different number of receivers. For simplicity, we assume $\varepsilon_i = \varepsilon$, $i = 1, \cdots, L$. Fig. 3.11 shows the total network transmission power versus the target MSE threshold ε for $N_{\rm b} = N_{\rm s} = N_{\rm r} = N_{\rm d} = 3$. Here the UB refers to (3.24a). From Fig. 3.11, it is obvious that more transmission power is required to multicast to a larger number of receivers assuring the same minimal MSE threshold at each receiver. The reason is that the proposed algorithm applies transmission power to satisfy the MSE requirement at all receivers. As the number of users increases, more transmission power is needed to combat the worst relay-receiver channel.

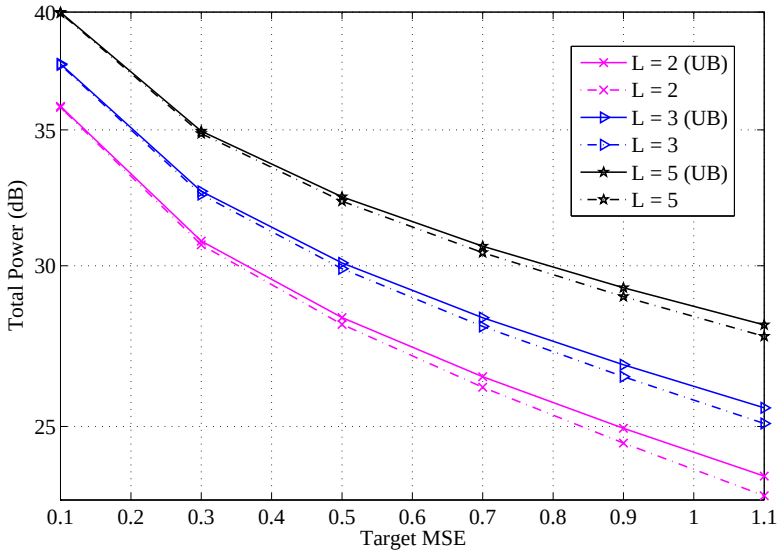

Figure 3.11: Example 5: Total power versus target MSE ε. Varying number of receivers, $N_{\mathrm{b}} = N_{\mathrm{s}} = N_{\mathrm{r}} = N_{\mathrm{d}} = 3$.

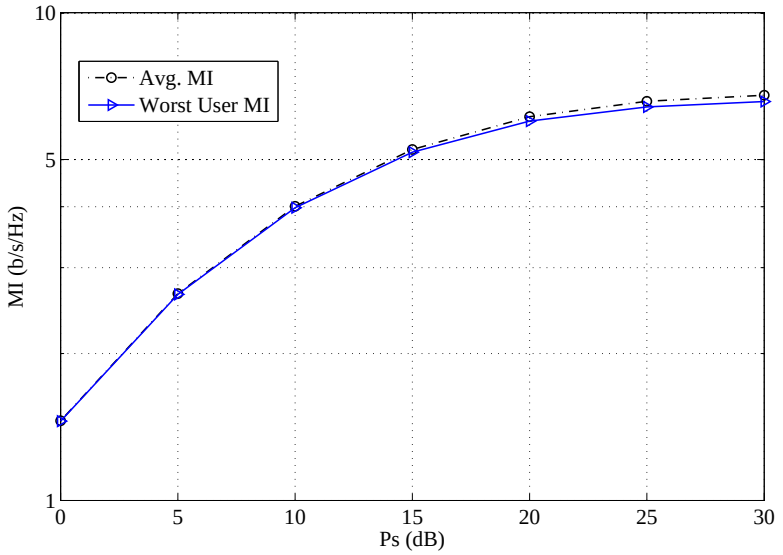

Figure 3.12: Example 6: Per-user MI versus P_{s}. $L = 6$, $N_{\mathrm{b}} = 1$, $N_{\mathrm{s}} = 6$, $N_{\mathrm{r}} = 3$, $N_{\mathrm{d}} = 3$, $P_{\mathrm{r}} = 20\mathrm{dB}$.

In the next two examples, we study the MI performance of the proposed single data stream two-hop multicasting MIMO relay algorithm developed in Section 3.4.1. The

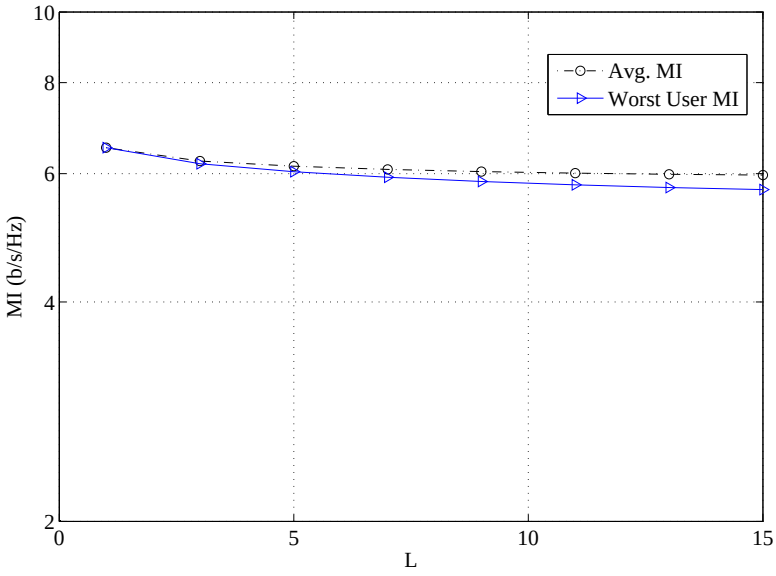

Figure 3.13: Example 7: Per-user MI versus number of users L. $N_{\rm b} = 1$, $N_{\rm s} = 6$, $N_{\rm r} = 3$, $N_{\rm d} = 3$, $P_{\rm s} = P_{\rm r} = 20$dB.

MI is computed as $\log_2(e_i^{-1})$, $i = 1 \cdots, L$, where e_i is obtained from (3.31). Firstly, the average MI and the worst user MI of the proposed algorithm versus $P_{\rm s}$ are shown in Fig. 3.12 for $L = 6$, $N_{\rm s} = 6$, $N_{\rm r} = 3$, $N_{\rm d} = 3$, and $P_{\rm r} = 20$dB. From Fig. 3.12, we can see that even the worst-channel user can have a comparable MI performance with the average MI. In Fig. 3.13, the MI performance of the proposed algorithm versus the number of receivers L is plotted with $N_{\rm s} = 6$, $N_{\rm r} = 3$, $N_{\rm d} = 3$, and $P_{\rm s} = P_{\rm r} = 20$dB. Note that with the increasing number of receivers, the MI keeps decreasing as is also observed in [67]. It can also be seen from Fig. 3.13 that the worst user MI decreases slowly with increasing L. It has been shown in [67] that the multicast rate converges to zero when the number of users L approaches to infinity.

In the last example, the total powers required for multicasting a single data stream among different number of receivers are compared. We use the GSS algorithm [84] to find the optimal P_1 for each target SNR threshold. The optimal relay precoding matrix is obtained by solving the problem (3.35) as is described in Section 3.4.2. For simplicity, we assume $\eta_i = \eta$, $i = 1, \cdots, L$. Fig. 3.14 illustrates the total powers required for $L = 2, 4$, and 6 versus target SNR threshold η with $N_{\rm s} = 6$, $N_{\rm r} = 3$, and $N_{\rm d} = 3$. It is notable that as we increase the target SNR requirements, the gap among the total powers required for different number of receivers reduces.

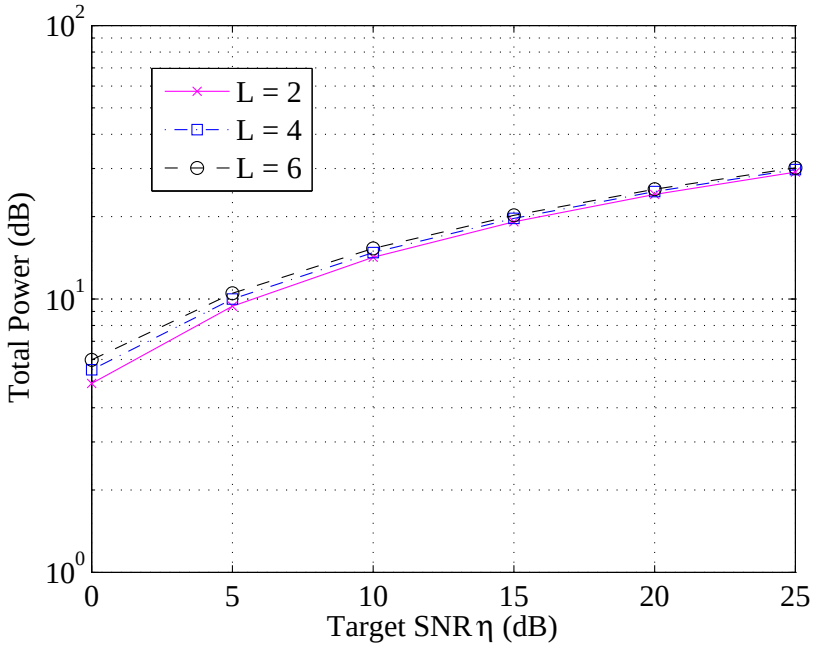

Figure 3.14: Example 8: Total power versus target SNR η. Varying number of receivers, $N_{\mathrm{b}} = 1$, $N_{\mathrm{s}} = 6$, $N_{\mathrm{r}} = 3$, $N_{\mathrm{d}} = 3$.

3.6 Chapter Summary

In this chapter, we considered a two-hop multicasting MIMO relay system with multi-antenna nodes and proposed transmit and relay precoding matrices based on two design criteria. Firstly, the worst-case MSE is minimized subjecting to power constraints at the transmitter and the relay node. Secondly, we propose a total network transmission power minimization strategy subjecting to QoS constraints. Under some mild approximation, we show that the problems can be solved with a significantly smaller computational complexity. In addition, we show that for the special case of single data stream multicasting, the relay precoding matrix optimization problem can be equivalently converted to the transmit beamforming problem for single-hop multicasting systems. Simulation results demonstrate that the proposed transmitter and relay design outperforms the existing techniques.

Chapter 4

Duality in Multi-Hop MIMO Relay Channel

In this chapter, we establish the duality between uplink and downlink multi-hop AF-MIMO relay channels with any number of hops and any number of antennas at each node. We show that in the downlink relay system, SINRs identical to those in the uplink relay system, and vice versa, can be achieved under certain conditions. As an application of the uplink-downlink duality, we propose an optimal design of the source precoding matrix and relay amplifying matrices for multi-hop MIMO relay system with a DPC transmitter at the source node.

The rest of this chapter is organized as follows. Existing works on the duality of MIMO systems are briefly reviewed in Section 4.3. In Section 4.2, we introduce the model of uplink and its downlink dual for the multi-hop AF-MIMO relay communication systems. The duality theorems are proven in Section 4.3. An optimal multi-hop MIMO relay system with a DPC-based transmitter at the source node is developed in Section 4.4. In Section 4.5, we show some numerical examples. Section 4.6 briefly summarizes the chapter. The proofs of Theorem 4.1 and 4.2 are listed in Sections 4.A-4.B, respectively.

4.1 Existing Works on Duality of MIMO Systems

Recently, the uplink-downlink duality of two-hop AF-MIMO relay systems has been derived in [48]. It is shown that for any relay amplifying matrix used in the uplink

channel, duality holds when a scaled Hermitian transpose of this matrix is employed in the downlink channel, where the scaling factor is obtained by switching the transmission power constraints at the source and the relay nodes. This result can be seen as a generalization of the well-known duality result for single-hop MIMO systems [49], [50]. For a multi-hop AF-MIMO relay network with single antenna source and destination nodes, the uplink-downlink duality has been established in [51].

In this chapter, we extend the uplink-downlink duality results in [48–51] to multi-hop AF-MIMO relay systems with any number of hops and any number of antennas at each node. We define duality as the achievement of identical SINRs at the uplink and the downlink systems with the same amount of total network transmission power. The reasons of considering SINR are two folds: First, SINR is an important parameter in communication system in the sense that it directly determines the QoS of each data stream. Second, many other parameters such as the achievable data rate and the MSE of signal estimation are closely related to SINR [42]. In particular, we show that for any number of hops, duality can be achieved by two approaches. First, if there is only total network transmission power constraint and no power constraint at individual nodes, then duality holds if $\mathbf{F}_l$ and $\mathbf{F}_{L-l}^H$, $l = 1, \cdots, L-1$, are used as the relay amplifying matrices at the lth relay node of the downlink and the uplink MIMO relay systems, respectively. Here L is the number of hops of the relay network. Second, with transmission power constraint at each node of the relay network, duality can be achieved by applying $c_l\mathbf{F}_l$ and $\mathbf{F}_{L-l}^H$ respectively as the amplifying matrices at the lth relay node of the downlink and the uplink relay systems, $l = 1, \cdots, L-1$, where the scaling factor c_l is obtained by switching the power constraints at the lth node of the downlink system and the $(L+1-l)$-th node of the uplink system, $l = 1, \cdots, L$. For both approaches, the source precoding matrix and the destination receiving matrix in the downlink system are swapped with the destination receiving matrix and the source precoding matrix at the uplink system, respectively.

Furthermore, we prove that the two approaches developed above are not only valid for relay systems with linear transceivers at the source and the destination nodes, but also hold if a successive interference cancellation (SIC)-based receiver is used at the destination node of the uplink MIMO relay system, and a transmitter based on DPC [85] is employed at the source node of the downlink MIMO relay channel. Interestingly, we show that the two duality approaches can be extended to multiuser AF-MIMO relay systems with any number of multi-antenna users.

As an application of the uplink-downlink duality theorem, we propose an optimal design of the source precoding matrix and relay amplifying matrices for multi-hop AF-MIMO relay systems with a DPC-based transmitter at the source node, by exploiting the results obtained for the dual uplink relay system [86]. Simulation results demonstrate that the optimal DPC-based MIMO relay system has a better BER performance compared with the optimal relay system using the SIC receiver, because the SIC receiver suffers from the error propagation effect, while the DPC transmitter does not.

4.2 Multi-Hop MIMO Relay System Model

We consider a wireless communication system with one source node, one destination node, and $L-1$ ($L \geq 2$) relay nodes. We assume that due to the propagation path-loss, the signal transmitted by the lth node can only be received by its direct forward node, i.e., the $(l+1)$-th node. Thus, signals transmitted by the source node pass through L hops until they reach the destination node. We also assume that the number of antennas at each node is N_l, $l = 1, \cdots, L+1$, and the number of source symbols in each transmission is N_b. A linear non-regenerative relay matrix is used at each relay node to amplify and forward the received signals. The system block diagrams of downlink and uplink multi-hop AF-MIMO relay systems are shown in Fig. 4.1.

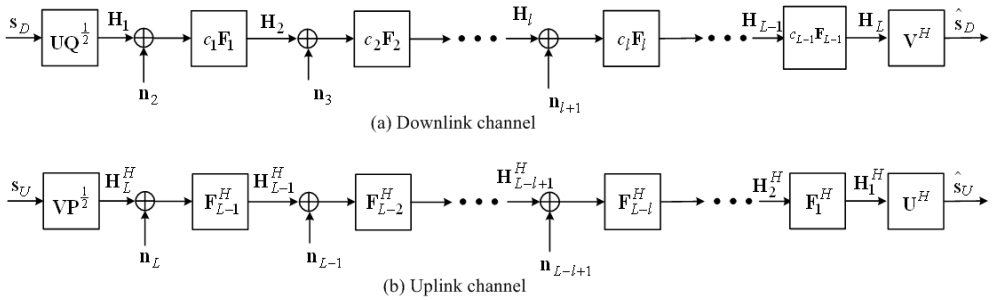

Figure 4.1: Block diagrams of downlink and uplink multi-hop AF-MIMO relay systems.

We would like to mention that for AF-MIMO relay systems with linear transceivers at the source and destination nodes, there should be $N_b \leq \min(N_1, N_2, \cdots, N_{L+1})$ in order to support N_b data streams in one transmission. However, if a nonlinear transmitter is installed at the source node or a nonlinear receiver is installed at the destination node of a MIMO relay system, N_b can be greater than $\min(N_1, N_2, \cdots, N_{L+1})$ [86].

In the case of downlink communication, the $N_b \times 1$ source symbol vector $\mathbf{s}_D = [s_1, s_2, \cdots, s_{N_b}]^T$ at the source node is linearly precoded by the $N_1 \times N_b$ matrix $\mathbf{U}\mathbf{Q}^{\frac{1}{2}}$, where $\mathbf{Q} = \text{diag}(q_1, q_2, \cdots, q_{N_b})$ and $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_b}]$ with $\|\mathbf{u}_i\|_2 = 1$, $i = 1, \cdots, N_b$. Here $\text{diag}(\cdot)$ denotes a diagonal matrix, $\|\cdot\|_2$ stands for the vector Euclidean norm, and q_i, $i = 1, \cdots, N_b$, is the power assigned to the ith data stream. We assume that $\text{E}[\mathbf{s}_D \mathbf{s}_D^H] = \mathbf{I}_{N_b}$. The $N_1 \times 1$ linearly precoded symbol vector $\mathbf{x}_1^D = \mathbf{U}\mathbf{Q}^{\frac{1}{2}}\mathbf{s}_D$ is transmitted by the source node. The $N_l \times 1$ signal vector received at the lth node is written as

$$\mathbf{y}_l^D = \mathbf{H}_{l-1}\mathbf{x}_{l-1}^D + \mathbf{n}_l, \qquad l = 2, \cdots, L+1 \tag{4.1}$$

where $\mathbf{H}_l$, $l = 1, \cdots, L$, is the $N_{l+1} \times N_l$ MIMO channel matrix between the $(l+1)$-th and the lth node, i.e., the lth hop, $\mathbf{n}_l$ is the $N_l \times 1$ i.i.d. additive white Gaussian noise (AWGN) vector at the lth node, and $\mathbf{x}_{l-1}^D$ is the $N_{l-1} \times 1$ signal vector transmitted by the $(l-1)$-th node. We assume that all noises are complex circularly symmetric with zero mean and unit variance.

Using the AF scheme, the input-output relationship at node l is given by

$$\mathbf{x}_l^D = c_{l-1}\mathbf{F}_{l-1}\mathbf{y}_l^D, \qquad l = 2, \cdots, L \tag{4.2}$$

where $c_{l-1}\mathbf{F}_{l-1}$ is the $N_l \times N_l$ amplifying matrix at node l (the $(l-1)$-th relay node), $c_l > 0$ is a scaling coefficient which is important for studying the uplink-downlink duality [48] and will be explained later. Combining (4.1) and (4.2), the received signal vector at the lth node of the downlink MIMO relay channel is given by

$$\mathbf{y}_2^D = \mathbf{H}_1\mathbf{U}\mathbf{Q}^{\frac{1}{2}}\mathbf{s}_D + \mathbf{n}_2 \tag{4.3}$$

$$\mathbf{y}_{l+1}^D = \mathbf{H}_l \bigotimes_{m=l-1}^{1} (c_m\mathbf{F}_m\mathbf{H}_m)\mathbf{U}\mathbf{Q}^{\frac{1}{2}}\mathbf{s}_D + \sum_{k=2}^{l}\bigotimes_{m=l}^{k}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\mathbf{n}_k + \mathbf{n}_{l+1},\ l = 2, \cdots, L \tag{4.4}$$

where for matrices $\mathbf{A}_i$, $\bigotimes_{i=l}^{k}(\mathbf{A}_i) \triangleq \mathbf{A}_l \cdots \mathbf{A}_k$. At the destination node, a linear receiver with an $N_{L+1} \times N_b$ weight matrix $\mathbf{V}$ is used to estimate the source symbol vector $\mathbf{s}_D$. The estimated symbol vector $\hat{\mathbf{s}}_D$ is given by

$$\hat{\mathbf{s}}_D = \mathbf{V}^H\mathbf{y}_{L+1}^D. \tag{4.5}$$

Since scaling the columns of $\mathbf{V}$ does not change the SINRs at the destination node, we assume that $\|\mathbf{v}_i\|_2 = 1$, $i = 1, \cdots, N_b$.

For the uplink MIMO relay system, the communication direction is reversed, and the roles of the source node and the destination node are swapped. The channel matrices are replaced by the Hermitian transpose of channel matrices in the downlink channel. Now the source node applies $\mathbf{V}\mathbf{P}^{\frac{1}{2}}$ to precode the uplink source symbol vector $\mathbf{s}_U$, where $\mathbf{P} = \text{diag}(p_1, p_2, \cdots, p_{N_b})$, and p_i, $i = 1, \cdots, N_b$, is the power assigned to the ith data stream. The lth node, $l = 2, \cdots, L$, uses $\mathbf{F}_{L+1-l}^H$ to amplify and forward received signals. Similar to (4.3) and (4.4), the received signal vector at the lth node of the uplink MIMO relay system can be written as

$$\mathbf{y}_2^U = \mathbf{H}_L^H \mathbf{V}\mathbf{P}^{\frac{1}{2}}\mathbf{s}_U + \mathbf{n}_L \tag{4.6}$$

$$\mathbf{y}_{L+2-l}^U = \bigotimes_{m=l}^{L-1}(\mathbf{H}_m^H\mathbf{F}_m^H)\mathbf{H}_L^H\mathbf{V}\mathbf{P}^{\frac{1}{2}}\mathbf{s}_U + \sum_{k=l}^{L-1}\bigotimes_{m=l}^{k}(\mathbf{H}_m^H\mathbf{F}_m^H)\mathbf{n}_{k+1} + \mathbf{n}_l, \quad l = 1, \cdots, L-1. \tag{4.7}$$

Finally, the destination node applies $\mathbf{U}$ to estimate the transmitted symbol vector with

$$\hat{\mathbf{s}}_U = \mathbf{U}^H\mathbf{y}_{L+1}^U. \tag{4.8}$$

We would like to note that at this point, for both the downlink and uplink systems, there is no specific design for $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{F}_l$, $l = 1, \cdots, L-1$.

In this chapter, the CSI requirement is the same as that in [86]. Basically, we assume that the source node has the CSI knowledge of the first-hop channel, the destination node knows the receiver weight matrix and each relay node knows the CSI of its backward channel and its forward channel. In practice, the backward CSI can be obtained through standard training methods. The forward CSI required at one relay node is exactly the backward CSI at its direct forward relay node, and thus can be obtained by a feedback from its direct forward relay node.

4.3 Uplink-Downlink Duality

In this section, we investigate the duality between the uplink and the downlink multi-hop AF-MIMO relay systems with any number of hops and any number of antennas at each node. It can be seen from Section 4.2 that given an uplink MIMO relay system, constructing its dual downlink MIMO relay system boils down to determining the appropriate relay scaling factors c_l, $l = 1, \cdots, L-1$, and the source power loading matrices $\mathbf{Q}$. The following two theorems establish the uplink-downlink duality property of multi-hop MIMO relay communication system with any number of hops.

Theorem 4.1 *If linear transceivers are used at the source and destination nodes of the uplink and the downlink systems, and there is no specific transmission power constraint at each node, then for any $L \geq 2$, the uplink-downlink duality can be achieved by $c_l = 1, l = 1, \cdots, L-1$. With transmission power constraint at individual nodes, duality is attained by setting $\rho_{L+1-l}^U = \rho_l^D, l = 1, \cdots, L$, and $c_l, l = 1, \cdots, L-1$, are obtained by transmission power constraints. Here $\rho_l^U \geq 0$ and $\rho_l^D \geq 0$, $l = 1, \cdots, L$, are the power budgets at the lth node of the uplink and the downlink systems, respectively.*

PROOF: See Appendix 4.A. □

It can be seen from Theorem 4.1 that if there is only total network transmission power constraint and no power constraint at individual nodes, then duality holds if $\mathbf{F}_l$ and $\mathbf{F}_{L-l}^H$, $l = 1, \cdots, L-1$, are used as the amplifying matrix at the lth relay node of the downlink and the uplink MIMO relay systems, respectively. However, in some practical applications, there is transmission power constraint at each node of the relay network. In such case, as suggested by Theorem 4.1, duality can be achieved by applying $c_l\mathbf{F}_l$ and $\mathbf{F}_{L-l}^H$ respectively as the amplifying matrix at the lth relay node of the downlink and the uplink relay systems, $l = 1, \cdots, L-1$, where the scaling factor c_l is determined by switching the power constraints at the lth node of the downlink system and the $(L+1-l)$-th node of the uplink system, $l = 1, \cdots, L$. It is worth noting that Theorem 4.1 holds for any transceiver matrices $\mathbf{U}$, $\mathbf{V}$, and relay amplifying matrices $\mathbf{F}_l$, $l = 1, \cdots, L-1$. However, for the uplink MIMO relay system, if a linear MMSE receiver is used, the optimal $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{F}_l$, $l = 1, \cdots, L-1$, are derived in [28].

Interestingly, Theorem 4.1 includes the results in [48]-[51] as special cases. It extends the uplink-downlink duality results from single-hop MIMO systems and two-hop AF-MIMO relay systems to multi-hop AF-MIMO relay systems with any number of hops and any number of antennas at each node.

Theorem 4.2 *If a DPC-based transmitter is used at the source node of the downlink MIMO relay system, and an SIC-based receiver is employed at the destination node of the uplink MIMO relay system, the uplink-downlink duality can be achieved by $c_l = 1, l = 1, \cdots, L-1$, when there is no specific transmission power constraint at each node. With transmission power constraint at individual nodes, duality can be attained by setting $\rho_{L+1-l}^U = \rho_l^D, l = 1, \cdots, L$, and $c_l, l = 1, \cdots, L-1$, are obtained by transmission power constraints.*

PROOF: See Appendix 4.B. □

Theorem 4.2 extends the duality results in Theorem 4.1 to the scenario where nonlinear transceivers are used at the source node of the downlink channel and the destination node of the uplink channel, respectively. Similar to Theorem 4.1, Theorem 4.2 holds for any transceiver matrices $\mathbf{U}$, $\mathbf{V}$, and relay amplifying matrices $\mathbf{F}_l$, $l = 1, \cdots, L-1$. However, if a nonlinear MMSE-SIC receiver is used at the uplink MIMO relay system, the optimal $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{F}_l$, $l = 1, \cdots, L-1$, can be found in [86]. Interestingly, both Theorem 4.1 and Theorem 4.2 also hold for multiuser MIMO relay scenario as explained below. In a broadcast channel (BC), a central station broadcasts information through L hops to M users each having R_m antennas, $m = 1, \cdots, M$, while in a multiaccess channel (MAC), M users each having T_m antennas, $m = 1, \cdots, M$, send information to a central station via L hops. The BC channel can be equivalently treated as a downlink multi-hop MIMO relay channel by grouping all users to form a "super" destination node with $N_{L+1} = \sum_{m=1}^{M} R_m$ antennas. Accordingly, one can view the MAC channel as an uplink multi-hop MIMO channel with $N_1 = \sum_{m=1}^{M} T_m$ antennas at the source node. Obviously, duality holds for the BC and MAC, provided that $\mathbf{V}$ (the destination receiving matrix in the BC and the source precoding matrix in the MAC) is chosen as a block diagonal matrix.

4.4 DPC-Based Optimal Multi-Hop MIMO Relay Design

In a DPC-based multi-hop MIMO relay system, the source symbol vector $\mathbf{s}_D$ is formed by encoding the information-bearing symbols μ_i, $i = 1, \cdots, N_b$, successively by removing the interference from the symbols already encoded. Thus, compared with a MIMO relay system using a linear transmitter at the source node, the DPC-based relay system should have a better BER performance.

The DPC-based optimal multi-hop MIMO relay design problem can be formulated as

$$\min_{\mathbf{U},\mathbf{Q},\{c_l\mathbf{F}_l\}} \quad f(\mathbf{U}, \mathbf{Q}, \{c_l\mathbf{F}_l\}) \tag{4.9a}$$

$$\text{s.t.} \quad P_l^D \le \rho_l^D, \qquad l = 1, \cdots, L \tag{4.9b}$$

where $f(\cdot)$ stands for a unified objective function [27], and P_l^D, $l = 1, \cdots, L$, is the power consumed by the lth node in the downlink system given by (4.27)-(4.29) in Appendix 4.A. Function $f(\cdot)$ includes a broad class of frequently used objective functions

in MIMO system design such as the negative source-destination mutual information, and the MSE of the signal waveform estimation at the destination. Directly solving the problem of (4.9) is difficult for a unified objective function $f(\cdot)$ and MIMO relay systems with a DPC transmitter, even for a single-hop (point-to-point) MIMO system [42]. Based on our knowledge, the problem of (4.9) has not been directly solved in literature. Now we apply Theorem 4.2 to optimize the DPC-based MIMO relay system. First, the optimization problem for the dual SIC-based uplink MIMO relay system can be written as

$$\min_{\mathbf{V},\mathbf{P},\{\mathbf{F}_{L-l}^H\}} \quad f(\mathbf{V},\mathbf{P},\{\mathbf{F}_{L-l}^H\}) \tag{4.10a}$$

$$\text{s.t.} \quad P_l^U \le \rho_{L+1-l}^D, \qquad l = 1,\cdots,L \tag{4.10b}$$

where P_l^U, $l = 1,\cdots,L$, is the power consumed by the lth node in the uplink system given by (4.22)-(4.24) in Appendix 4.A. Applying the majorization theory [87], the problem of (4.10) has been solved in [86] for a broad class of objective functions $f(\cdot)$. The optimal feed-forward matrix $\mathbf{W}$ at the destination node of the uplink MIMO relay system has also been developed in [86]. Note that for a single-hop MIMO system, the uplink-downlink duality has been applied in [42] to optimize the DPC-based transceiver design.

In this section, we take the optimal $\mathbf{V}$, $\mathbf{P}$, and $\mathbf{F}_l$, $l = 1,\cdots,L-1$, from [86]. Based on Theorem 4.2, the optimal source precoding matrix in the DPC-based downlink relay system can be written as $\mathbf{U}\mathbf{Q}^{\frac{1}{2}}$, where $\mathbf{U} = [\bar{\mathbf{w}}_1, \bar{\mathbf{w}}_2, \cdots, \bar{\mathbf{w}}_{N_b}]$, and $\bar{\mathbf{w}}_i$, $i = 1,\cdots,N_b$, are obtained by scaling columns of $\mathbf{W}$ such that $\|\bar{\mathbf{w}}_i\|_2 = 1$, $i = 1,\cdots,N_b$. The optimal amplifying matrix at the lth relay node of the DPC-based relay system is $c_l\mathbf{F}_l$, where the scalar c_l is determined by the transmission power budget ρ_{l+1}^D as explained later. The optimal receiving matrix at the destination node of the DPC-based relay system is $\mathbf{V}\mathbf{\Gamma}$, where $\mathbf{\Gamma} = \text{diag}(\gamma_1, \gamma_2, \cdots, \gamma_{N_b})$ with γ_i chosen such that $\gamma_i\sqrt{q_i}\,\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{m=L-1}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_i = 1$, $i = 1,\cdots,N_b$.

Now the task is to obtain the unknown quantities $q_i, i = 1,\cdots,N_b$, and $c_l, l = 1,\cdots,L-1$. Towards this goal, we assume for the moment that $\mathbf{\Gamma} = \mathbf{I}_{N_b}$, since scaling the receiving vector $\mathbf{v}_i$ does not change SINR_i^D, $i = 1,\cdots,N_b$, in (4.40) of Appendix 4.B. With a DPC encoder at the source node, the information-bearing symbol μ_i is encoded into s_i by treating $\sum_{j=1}^{i-1}\sqrt{q_j}\,\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{m=L-1}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_j s_j$ as the interference known at the transmitter. Note that at this stage the interference term $\sum_{j=i+1}^{N_b}\sqrt{q_j}\,\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{m=L-1}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_j s_j$ is unknown. At the destination node, after

applying the linear filter $\mathbf{V}$ and the DPC decoder, the estimated information-bearing symbol can be written as

$$\hat{\mu}_i = \sqrt{q_i}\,\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{m=L-1}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_i\mu_i + \sum_{j=i+1}^{N_b}\sqrt{q_j}\,\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{m=L-1}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_j s_j + \mathbf{v}_i^H\Big(\sum_{l=2}^{L}\bigotimes_{m=L}^{l}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\mathbf{n}_l + \mathbf{n}_{L+1}\Big),\quad i=1,\cdots,N_b$$

with the output SINR β_i as

$$\beta_i = \frac{d_{i,i}q_i}{\sum_{j=i+1}^{N_b} d_{i,j}q_j + e_i + \mathbf{g}_i^T\boldsymbol{\theta}},\qquad i=1,\cdots,N_b \tag{4.11}$$

where $\boldsymbol{\theta}$ is an $(L-1)\times 1$ vector with elements $\theta_l = 1/\prod_{m=1}^{l}c_m^2$, $l=1,\cdots,L-1$, $\mathbf{g}_i$ is an $(L-1)\times 1$ vector with elements given by $g_{i,L-1}=1$, $g_{i,l}=\mathbf{v}_i^H\bigotimes_{m=L}^{l+2}(\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=l+2}^{L}\times(\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\mathbf{v}_i$, $l=1,\cdots,L-2$, and

$$d_{i,j} \triangleq \left|\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{m=L-1}^{1}(\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_j\right|^2,\qquad j=i,\cdots,N_b$$
$$e_i \triangleq \mathbf{v}_i^H\bigotimes_{m=L}^{2}(\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=2}^{L}(\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\mathbf{v}_i.$$

Collecting all equations in (4.11) for $i=1,\cdots,N_b$, we obtain the following systems of linear equations

$$\boldsymbol{\Phi}\mathbf{q} + \mathbf{G}^T\boldsymbol{\theta} = -\mathbf{e} \tag{4.12}$$

where $\mathbf{q} \triangleq [q_1,q_2,\cdots,q_{N_b}]^T$, $\mathbf{e} \triangleq [e_1,e_2,\cdots,e_{N_b}]^T$, $\boldsymbol{\Phi}$ is an $N_b\times N_b$ upper-triangle matrix with elements of $\phi_{i,j}=0$, $j=1,\cdots,i-1$, $\phi_{i,i}=-d_{i,i}/\beta_i$, $\phi_{i,j}=d_{i,j}$, $j=i+1,\cdots,N_b$, $i=1,\cdots,N_b$, and $\mathbf{G}$ is an $(L-1)\times N_b$ matrix whose ith column is given by $\mathbf{g}_i$ in (4.11).

From the transmission power consumed at the relay nodes (4.27) and (4.28) in Appendix 4.A, we have

$$\frac{P_{l+1}^D}{\prod_{m=1}^{l}c_m^2} = \mathbf{a}_l^T\mathbf{q} + \xi_{l,1} + \sum_{k=1}^{l-1}\frac{\xi_{l,k+1}}{\prod_{m=1}^{k}c_m^2},\qquad l=2,\cdots,L-1 \tag{4.13}$$

$$P_2^D/c_1^2 = \mathbf{a}_1^T\mathbf{q} + \xi_{1,1} \tag{4.14}$$

where $\mathbf{a}_l$ is an $N_b\times 1$ vector with elements $a_{l,i}=\bar{\mathbf{w}}_i^H\bigotimes_{m=1}^{l}(\mathbf{H}_m^H\mathbf{F}_m^H)\bigotimes_{m=l}^{1}(\mathbf{F}_m\mathbf{H}_m)\bar{\mathbf{w}}_i$, $i=1,\cdots,N_b$, and

$$\xi_{l,k} \triangleq \begin{cases}\mathrm{tr}\big(\mathbf{F}_l\bigotimes_{m=l}^{k+1}(\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=k+1}^{l}(\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\mathbf{F}_l^H\big), & k=1,\cdots,l-1\\ \mathrm{tr}(\mathbf{F}_l\mathbf{F}_l^H), & k=l\end{cases}.$$

Equation (4.13) can be rewritten as

$$\sum_{k=1}^{l-1} \frac{\xi_{l,k+1}}{\prod_{m=1}^{k} c_m^2} - \frac{P_{l+1}^D}{\prod_{m=1}^{l} c_m^2} = -\mathbf{a}_l^T \mathbf{q} - \xi_{l,1}, \quad l = 2, \cdots, L-1. \tag{4.15}$$

Collecting all equations in (4.15) for $l = 2, \cdots, L-1$, and together with (4.14), we obtain the following system of equations

$$\boldsymbol{\Psi}\boldsymbol{\theta} = -\mathbf{A}^T\mathbf{q} - \boldsymbol{\xi} \tag{4.16}$$

where $\boldsymbol{\Psi}$ is an $(L-1) \times (L-1)$ lower-triangle matrix with elements of $\psi_{l,k} = \xi_{l,k+1}$, $k = 1, \cdots, l-1$, $\psi_{l,l} = -P_{l+1}^D$, $\psi_{l,k} = 0$, $k = l+1, \cdots, L-1$, $l = 1, \cdots, L-1$, $\mathbf{A}$ is an $N_b \times (L-1)$ matrix whose lth column is given by $\mathbf{a}_l$ in (4.14) and (4.15), and $\boldsymbol{\xi} \triangleq [\xi_{1,1}, \xi_{2,1}, \cdots, \xi_{L-1,1}]^T$.

Now (4.12) and (4.16) form a system of linear equations of $\boldsymbol{\theta}$ and $\mathbf{q}$. Solving (4.12) and (4.16), we obtain

$$\mathbf{q} = (\boldsymbol{\Phi} - \mathbf{G}^T\boldsymbol{\Psi}^{-1}\mathbf{A}^T)^{-1}(\mathbf{G}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\xi} - \mathbf{e}) \tag{4.17}$$

$$\boldsymbol{\theta} = -\boldsymbol{\Psi}^{-1}\big[\boldsymbol{\xi} + \mathbf{A}^T(\boldsymbol{\Phi} - \mathbf{G}^T\boldsymbol{\Psi}^{-1}\mathbf{A}^T)^{-1}(\mathbf{G}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\xi} - \mathbf{e})\big]. \tag{4.18}$$

Note that $\boldsymbol{\Psi}^{-1}$ always exists because $\boldsymbol{\Psi}$ is a lower-triangle matrix, and $\boldsymbol{\Phi} - \mathbf{G}^T\boldsymbol{\Psi}^{-1}\mathbf{A}^T$ is invertible since $\boldsymbol{\Phi}$ is an upper-triangle matrix and $\boldsymbol{\Psi}^{-1}$ is a lower-triangle matrix. Finally, c_l can be obtained from (4.18) with $c_1 = \sqrt{1/\theta_1}$ and $c_l = \sqrt{\theta_{l-1}/\theta_l}$, $l = 2, \cdots, L-1$.

4.5 Numerical Examples

In this section, we study the performance of the proposed DPC-based source precoding matrix and relay amplifying matrices through numerical simulations. To precode the ith information-bearing symbol μ_i into the source symbol $s_i, i = 1, \cdots, N_b$, we apply the Tomlinson-Harashima coding technique [88] and [89], which is a simple but suboptimal implementation of the DPC scheme. The Tomlinson-Harashima scheme makes use of the modulo operation to remove the interference from the preceding symbols without increasing the transmission power at the source node. Accordingly, the length of the modulo Δ is chosen to preserve the transmission power consistency.

The source, destination, and all relay nodes are equipped with multiple antennas. We simulate a flat Rayleigh fading environment where the channel matrices have i.i.d.

entries with zero mean and variances σ_l^2/N_l for $\mathbf{H}_l$, $l = 1, \cdots, L$. For each channel realization, 1000 QAM-modulated symbols with Gray mapping are transmitted at each data stream, and all simulation results are averaged over 500 independent channel realizations. We define $\text{SNR}_l \triangleq \sigma_l^2 \rho_l^D N_{l+1}/N_l$ as the SNR of the lth hop, $l = 1, \cdots, L$. The BER performance of the proposed optimal DPC-based relay system is compared with that of the optimal relay system using the SIC receiver [86]. As a benchmark, we also show the performance of the fictitious genie-aided SIC-based relay system, where the error propagation at each layer of the SIC receiver is eliminated by a genie.

In the first example, we simulate a relay system with $L = 2$ hops and choose $N_l = 5$, $l = 1, 2, 3$, and $N_b = 5$. The symbols are modulated by the 16-QAM constellations. Fig. 4.2 shows BERs of all systems versus SNR_1 for $\text{SNR}_2 = 20$dB. It can be seen from Fig. 4.2 that the optimal DPC-based relay system has a better BER performance compared with the relay system using the SIC receiver, since the latter system suffers from error propagation.

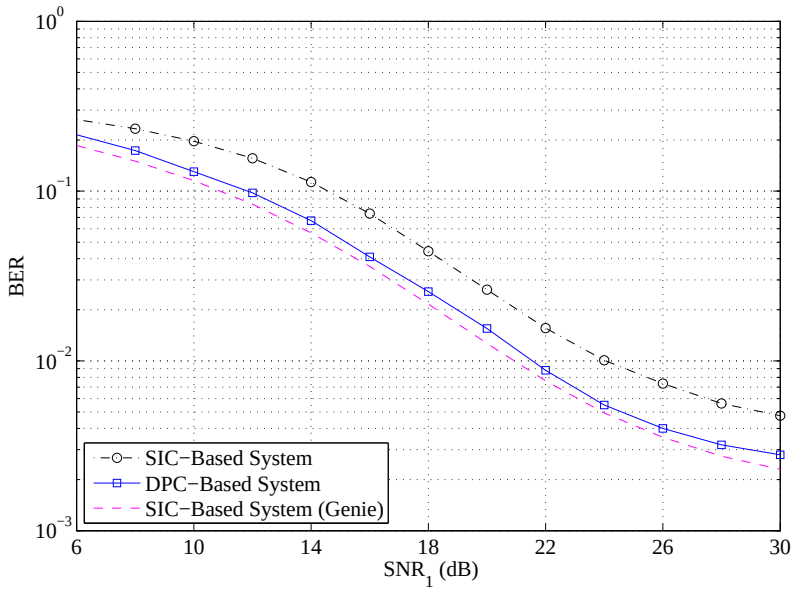

Figure 4.2: Example 1: Two hops. $N_l = 5$, $l = 1, 2, 3$, $N_b = 5$, 16-QAM, $\text{SNR}_2 = 20$dB.

In the second example, a multi-hop MIMO relay system with $L = 5$ and $N_l = 5$, $l = 1, \cdots, 6$, and $N_b = 4$ is simulated. Each hop has the same SNR, i.e., $\text{SNR}_l = \text{SNR}$, $l = 1, \cdots, 5$. The 64-QAM constellations are used to modulate the symbols. Fig. 4.3 displays the BER performance of all three systems versus SNR. Obviously, for multi-hop systems, the DPC-based relay design outperforms the relay system using the SIC

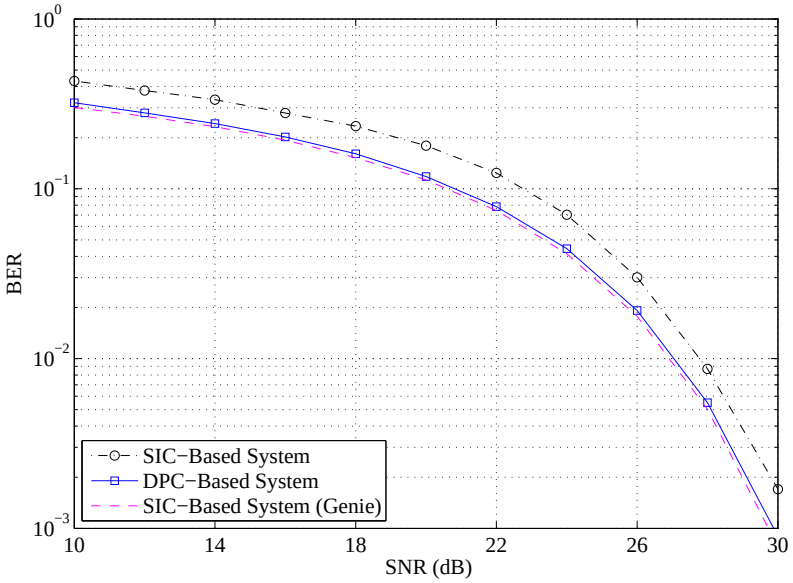

Figure 4.3: Example 2: Five hops. $N_l = 5$, $l = 1, \cdots, 6$, $N_b = 4$, 64-QAM, $\text{SNR}_l = \text{SNR}$, $l = 1, \cdots, 5$.

receiver by removing the error propagation effect. From Figs. 4.2 and 4.3, we observe a slight SNR loss of the DPC-based relay system compared with the genie-aided SIC-based relay system. This is mainly due to the inherent power loss and modulo loss of the Tomlinson-Harashima precoder [90].

The reason for the error floor effect displayed in Fig. 4.2 for both the SIC-based system and the DPC-based system is that the relay system is fully-loaded in the sense that $N_l = N_b$, $l = 1, 2, 3$. It is well-known that for a one-hop (point-to-point) MIMO system with an SIC receiver, the diversity order for the first decoded stream is only one if the transmitter and the receiver have the same number of antennas. Compared with that in Fig. 4.2, the system shown in Fig. 4.3 is under-loaded since $N_l > N_b$, $l = 1, \cdots, 6$. Thus, the error floor effect is not observed in Fig. 4.3.

4.6 Chapter Summary

In this chapter, we have established the uplink-downlink duality of multi-hop AF-MIMO relay systems with any number of hops and any number of antennas at each node, which generalizes several previously established results. Based on such duality, we proposed an optimal design of the source precoding matrix and relay amplifying matrices for

multi-hop AF-MIMO relay systems with a DPC-based transmitter at the source node. Simulation results show that the optimal DPC-based MIMO relay system has a lower BER than the optimal relay system using the SIC receiver.

4.A Proof of Theorem 4.1

The basic idea of the proof is to show under which conditions of $\mathbf{P}$, $\mathbf{Q}$, and c_l, $l = 1, \cdots, L-1$, the uplink and downlink channels achieve identical SINRs. The proof is conducted in three steps. First, for both the uplink MIMO relay channel (4.6)-(4.8) and the downlink MIMO relay channel (4.3)-(4.5), we write the SINR for each data stream and the required total transmission power. Second, we rewrite the total transmission power of the downlink system based on the definition of duality that both channels should achieve identical SINRs. Finally, we find under which $\mathbf{P}$, $\mathbf{Q}$, and c_l, $l = 1, \cdots, L-1$, the total transmission power consumed by both systems is identical.

Based on (4.7), the SINRs of data streams at the destination node of the uplink MIMO relay channel are given by

$$\mathrm{SINR}_i^U = \frac{\left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1} (\mathbf{H}_l^H \mathbf{F}_l^H) \mathbf{H}_L^H \mathbf{v}_i\right|^2 p_i}{\sum_{\substack{j=1 \\ j\neq i}}^{N_b} \left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1} (\mathbf{H}_l^H \mathbf{F}_l^H) \mathbf{H}_L^H \mathbf{v}_j\right|^2 p_j + \mathbf{u}_i^H \sum_{l=1}^{L-1} \Big(\bigotimes_{m=1}^{l} (\mathbf{H}_m^H \mathbf{F}_m^H) \bigotimes_{m=l}^{1} (\mathbf{F}_m \mathbf{H}_m) \Big) \mathbf{u}_i + 1}, \quad i = 1, \cdots, N_b. \tag{4.19}$$

The transmission power P_l^U consumed by the lth node, $l = 1, \cdots, L$, in the uplink relay system can be calculated using (4.6) and (4.7) as

$$\begin{aligned} P_{L+2-l}^U &= \mathbf{F}_{l-1}^H \mathrm{E}\big[\mathbf{y}_{L+2-l}^U (\mathbf{y}_{L+2-l}^U)^H\big] \mathbf{F}_{l-1} \\ &= \mathrm{tr}\left(\bigotimes_{m=l}^{L} (\mathbf{F}_{m-1}^H \mathbf{H}_m^H) \mathbf{V}\mathbf{P}\mathbf{V}^H \bigotimes_{m=L}^{l} (\mathbf{H}_m \mathbf{F}_{m-1}) + \mathbf{F}_{l-1}^H \left(\sum_{k=l}^{L-1} \bigotimes_{m=l}^{k} (\mathbf{H}_m^H \mathbf{F}_m^H) \right. \right. \\ &\qquad \left. \left. \bigotimes_{m=k}^{l} (\mathbf{F}_m \mathbf{H}_m) + \mathbf{I}_{N_l} \right) \mathbf{F}_{l-1} \right), \qquad l = 2, \cdots, L-1 \end{aligned} \tag{4.22}$$

$$P_2^U = \mathrm{tr}\big(\mathbf{F}_{L-1}^H (\mathbf{H}_L^H \mathbf{V}\mathbf{P}\mathbf{V}^H \mathbf{H}_L + \mathbf{I}_{N_L}) \mathbf{F}_{L-1}\big) \tag{4.23}$$

$$P_1^U = \mathrm{tr}(\mathbf{P}). \tag{4.24}$$

The total transmission power consumed by the uplink MIMO relay system can be ob-

tained by the sum of (4.22)-(4.24) and written as

$$P_T^U = \sum_{i=1}^{N_b} p_i \Big(1 + \mathbf{v}_i^H \sum_{l=2}^{L} \Big(\bigotimes_{m=L}^{l} (\mathbf{H}_m \mathbf{F}_{m-1}) \bigotimes_{m=l}^{L} (\mathbf{F}_{m-1}^H \mathbf{H}_m^H) \Big) \mathbf{v}_i \Big) + \sum_{l=2}^{L-1} \mathrm{tr}\Big(\mathbf{F}_{l-1}^H \sum_{k=l}^{L-1} \Big(\bigotimes_{m=l}^{k} (\mathbf{H}_m^H \mathbf{F}_m^H) \bigotimes_{m=k}^{l} (\mathbf{F}_m \mathbf{H}_m) \Big) \mathbf{F}_{l-1} \Big) + \sum_{l=1}^{L-1} \mathrm{tr}(\mathbf{F}_l^H \mathbf{F}_l). \quad (4.25)$$

Similarly, using (4.4), the SINRs of data streams at the destination node of the downlink relay channel are given by

$$\mathrm{SINR}_i^D = \frac{P_{s,i}}{P_{in,i}}, \qquad i = 1, \cdots, N_b, \quad (4.26)$$

where $P_{s,i} \triangleq \Big| \mathbf{v}_i^H \mathbf{H}_L \bigotimes_{l=L-1}^{1} (c_l \mathbf{F}_l \mathbf{H}_l) \mathbf{u}_i \Big|^2 q_i$ and $P_{in,i} \triangleq \sum_{\substack{j=1 \\ j \neq i}}^{N_b} \Big| \mathbf{v}_i^H \mathbf{H}_L \bigotimes_{l=L-1}^{1} (c_l \mathbf{F}_l \mathbf{H}_l) \mathbf{u}_j \Big|^2 q_j + \mathbf{v}_i^H \sum_{l=2}^{L} \Big(\bigotimes_{m=L}^{l} (c_{m-1} \mathbf{H}_m \mathbf{F}_{m-1}) \bigotimes_{m=l}^{L} (c_{m-1} \mathbf{F}_{m-1}^H \mathbf{H}_m^H) \Big) \mathbf{v}_i + 1$. The transmission power P_l^D consumed by the lth node, $l = 1, \cdots, L$, in the downlink system can be obtained from (4.3) and (4.4) as

$$\begin{aligned} P_{l+1}^D &= \mathbf{F}_l \mathrm{E}\big[\mathbf{y}_{l+1}^D (\mathbf{y}_{l+1}^D)^H \big] \mathbf{F}_l^H \\ &= \mathrm{tr}\Bigg(\bigotimes_{m=l}^{1} (c_m \mathbf{F}_m \mathbf{H}_m) \mathbf{U} \mathbf{Q} \mathbf{U}^H \bigotimes_{m=1}^{l} (c_m \mathbf{H}_m^H \mathbf{F}_m^H) + c_l^2 \times \mathbf{F}_l \Bigg(\sum_{k=2}^{l} \bigotimes_{m=l}^{k} (c_{m-1} \mathbf{H}_m \mathbf{F}_{m-1}) \\ &\quad \bigotimes_{m=k}^{l} (c_{m-1} \mathbf{F}_{m-1}^H \mathbf{H}_m^H) + \mathbf{I}_{N_{l+1}} \Bigg) \mathbf{F}_l^H \Bigg), \quad l = 2, \cdots, L-1 \quad (4.27) \\ P_2^D &= \mathrm{tr}\big(c_1^2 \mathbf{F}_1 (\mathbf{H}_1 \mathbf{U} \mathbf{Q} \mathbf{U}^H \mathbf{H}_1^H + \mathbf{I}_{N_2}) \mathbf{F}_1^H \big) \quad (4.28) \\ P_1^D &= \mathrm{tr}(\mathbf{Q}). \quad (4.29) \end{aligned}$$

The total transmission power consumed by the downlink MIMO relay system is obtained by the sum of (4.27)-(4.29) and given by

$$\begin{aligned} P_T^D = &\sum_{i=1}^{N_b} q_i \Big(1 + \mathbf{u}_i^H \sum_{l=1}^{L-1} \Big(\bigotimes_{m=1}^{l} (c_m \mathbf{H}_m^H \mathbf{F}_m^H) \bigotimes_{m=l}^{1} (c_m \mathbf{F}_m \mathbf{H}_m) \Big) \mathbf{u}_i \Big) \\ &+ \sum_{l=2}^{L-1} \mathrm{tr}\Big(c_l^2 \mathbf{F}_l \sum_{k=2}^{l} \Big(\bigotimes_{m=l}^{k} (c_{m-1} \mathbf{H}_m \mathbf{F}_{m-1}) \bigotimes_{m=k}^{l} (c_{m-1} \mathbf{F}_{m-1}^H \mathbf{H}_m^H) \Big) \mathbf{F}_l^H \Big) \\ &+ \sum_{l=1}^{L-1} \mathrm{tr}(c_l^2 \mathbf{F}_l \mathbf{F}_l^H). \end{aligned} \quad (4.30)$$

To achieve identical SINRs at the uplink and the downlink systems, we should have $\mathrm{SINR}_i^U = \mathrm{SINR}_i^D$, $i = 1, \cdots, N_b$. Please note that we do not assume that all data

streams have identical SINR, i.e., it is possible that $\mathrm{SINR}_i^U \neq \mathrm{SINR}_j^U$, for $i \neq j$. Using (4.19) and (4.26) we obtain from $\sum_{i=1}^{N_b} \mathrm{SINR}_i^U = \sum_{i=1}^{N_b} \mathrm{SINR}_i^D$ that

$$\begin{aligned}
&\sum_{i=1}^{N_b}\Bigg(\prod_{l=1}^{L-1}c_l^2\, q_i\Big(\sum_{\substack{j=1\\ j\neq i}}^{N_b}\Big|\mathbf{u}_i^H\bigotimes_{l=1}^{L-1}(\mathbf{H}_l^H\mathbf{F}_l^H)\mathbf{H}_L^H\mathbf{v}_j\Big|^2 p_j\\
&+\mathbf{u}_i^H\sum_{l=1}^{L-1}\Big(\bigotimes_{m=1}^{l}(\mathbf{H}_m^H\mathbf{F}_m^H)\bigotimes_{m=l}^{1}(\mathbf{F}_m\mathbf{H}_m)\Big)\mathbf{u}_i+1\Big)\Bigg)\\
&=\sum_{i=1}^{N_b}p_i\Bigg(\sum_{\substack{j=1\\ j\neq i}}^{N_b}\Big|\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{l=L-1}^{1}(c_l\mathbf{F}_l\mathbf{H}_l)\mathbf{u}_j\Big|^2 q_j+\mathbf{v}_i^H\\
&\times\sum_{l=2}^{L}\Big(\bigotimes_{m=L}^{l}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=l}^{L}(c_{m-1}\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\Big)\mathbf{v}_i+1\Big).
\end{aligned}\tag{4.31}$$

By using the identity of

$$\sum_{i=1}^{N_b}\sum_{j=1,j\neq i}^{N_b} q_i\Big|\mathbf{u}_i^H\bigotimes_{l=1}^{L-1}(c_l\mathbf{H}_l^H\mathbf{F}_l^H)\mathbf{H}_L^H\mathbf{v}_j\Big|^2 p_j=\sum_{i=1}^{N_b}\sum_{j=1,j\neq i}^{N_b} p_i\Big|\mathbf{v}_i^H\mathbf{H}_L\bigotimes_{l=L-1}^{1}(c_l\mathbf{F}_l\mathbf{H}_l)\mathbf{u}_j\Big|^2 q_j$$

we obtain from (4.31) that

$$\begin{aligned}
&\prod_{l=1}^{L-1}c_l^2\sum_{i=1}^{N_b}q_i\Big(\mathbf{u}_i^H\sum_{l=1}^{L-1}\Big(\bigotimes_{m=1}^{l}(\mathbf{H}_m^H\mathbf{F}_m^H)\bigotimes_{m=l}^{1}(\mathbf{F}_m\mathbf{H}_m)\Big)\mathbf{u}_i+1\Big)\\
&\quad=\sum_{i=1}^{N_b}p_i\Big(\mathbf{v}_i^H\sum_{l=2}^{L}\Big(\bigotimes_{m=L}^{l}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=l}^{L}(c_{m-1}\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\Big)\mathbf{v}_i+1\Big).
\end{aligned}\tag{4.32}$$

Substituting (4.32) back into (4.30), we can rewrite P_T^D as

$$\begin{aligned}
P_T^D=&\,\mathrm{tr}\Bigg(\sum_{l=2}^{L}\prod_{m=l}^{L}c_{m-1}^2\bigotimes_{m=l}^{L}(\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\mathbf{V}\mathbf{P}\mathbf{V}^H\bigotimes_{m=L}^{l}(\mathbf{H}_m\mathbf{F}_{m-1})+\mathbf{P}\Bigg)\\
&+\mathrm{tr}\Bigg(\sum_{l=1}^{L-2}\Big(1-\prod_{m=l+1}^{L-1}c_m^2\Big)\bigotimes_{m=1}^{l}(c_m\mathbf{H}_m^H\mathbf{F}_m^H)\bigotimes_{m=l}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\mathbf{U}\mathbf{Q}\mathbf{U}^H\\
&+\Big(1-\prod_{m=1}^{L-1}c_m^2\Big)\mathbf{U}\mathbf{Q}\mathbf{U}^H\Bigg)+\sum_{l=2}^{L-1}\mathrm{tr}\Big(c_l^2\mathbf{F}_l\sum_{k=2}^{l}\Big(\bigotimes_{m=l}^{k}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\\
&\bigotimes_{m=k}^{l}(c_{m-1}\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\Big)\mathbf{F}_l^H\Big)+\sum_{l=1}^{L-1}\mathrm{tr}(c_l^2\mathbf{F}_l\mathbf{F}_l^H).
\end{aligned}\tag{4.33}$$

For notational simplicity, let us denote

$$a_l\triangleq\mathrm{tr}\Big(c_l^2\mathbf{F}_l\sum_{k=2}^{l}\Big(\bigotimes_{m=l}^{k}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=k}^{l}(c_{m-1}\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\Big)\mathbf{F}_l^H\Big),l=2,\cdots,L-1.\tag{4.34}$$

Then with some manipulations, we have

$$\begin{aligned}\sum_{l=2}^{L-1} a_l &= \sum_{l=2}^{L-1}\prod_{m=l+1}^{L-1} c_m^2 a_l + \sum_{l=2}^{L-1}\Big(1-\prod_{m=l+1}^{L-1} c_m^2\Big)a_l \\ &= \sum_{l=2}^{L-1}\sum_{k=l}^{L-1}\prod_{m=k+1}^{L-1} c_m^2 \operatorname{tr}\Big(c_k^2\mathbf{F}_k \bigotimes_{m=k}^{l}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1}) \\ &\quad \bigotimes_{m=l}^{k}(c_{m-1}\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\mathbf{F}_k^H\Big) + \sum_{l=2}^{L-1}\Big(1-\prod_{m=l+1}^{L-1} c_m^2\Big)a_l \\ &= \sum_{l=2}^{L-1}\sum_{k=l}^{L-1}\prod_{m=k+1}^{L-1} c_m^2 \operatorname{tr}\Big(c_{l-1}^2\mathbf{F}_{l-1}^H \bigotimes_{m=l}^{k}(c_m\mathbf{H}_m^H\mathbf{F}_m^H) \\ &\quad \bigotimes_{m=k}^{l}(c_m\mathbf{F}_m\mathbf{H}_m)\mathbf{F}_{l-1}\Big) + \sum_{l=2}^{L-1}\Big(1-\prod_{m=l+1}^{L-1} c_m^2\Big)a_l \\ &= \sum_{l=2}^{L-1}\prod_{m=l}^{L} c_{m-1}^2 \sum_{k=l}^{L-1}\operatorname{tr}\Big(\mathbf{F}_{l-1}^H \bigotimes_{m=l}^{k}(\mathbf{H}_m^H\mathbf{F}_m^H) \\ &\quad \bigotimes_{m=k}^{l}(\mathbf{F}_m\mathbf{H}_m)\mathbf{F}_{l-1}\Big) + \sum_{l=2}^{L-1}\Big(1-\prod_{m=l+1}^{L-1} c_m^2\Big)a_l \end{aligned} \tag{4.35}$$

and

$$\sum_{l=1}^{L-1}\operatorname{tr}(c_l^2\mathbf{F}_l\mathbf{F}_l^H) = \sum_{l=2}^{L}\prod_{m=l}^{L} c_{m-1}^2\operatorname{tr}(\mathbf{F}_{l-1}^H\mathbf{F}_{l-1}) + \sum_{l=1}^{L-1}\Big(1-\prod_{m=l+1}^{L-1} c_m^2\Big)\operatorname{tr}(c_l^2\mathbf{F}_l\mathbf{F}_l^H). \tag{4.36}$$

Substituting (4.35) and (4.36) back into (4.33) and after rearranging terms, we can rewrite P_T^D as

$$\begin{aligned} P_T^D = &\sum_{l=2}^{L-1}\prod_{m=l}^{L} c_{m-1}^2 \operatorname{tr}\Bigg(\bigotimes_{m=l}^{L}(\mathbf{F}_{m-1}^H\mathbf{H}_m^H)\mathbf{V}\mathbf{P}\mathbf{V}^H\bigotimes_{m=L}^{l}(\mathbf{H}_m\mathbf{F}_{m-1}) \\ &+\mathbf{F}_{l-1}^H\Big(\sum_{k=l}^{L-1}\bigotimes_{m=l}^{k}(\mathbf{H}_m^H\mathbf{F}_m^H)\bigotimes_{m=k}^{l}(\mathbf{F}_m\mathbf{H}_m)+\mathbf{I}_{N_l}\Big)\mathbf{F}_{l-1}\Bigg) \\ &+c_{L-1}^2\operatorname{tr}\Big(\mathbf{F}_{L-1}^H(\mathbf{H}_L^H\mathbf{V}\mathbf{P}\mathbf{V}^H\mathbf{H}_L+\mathbf{I}_{N_L})\mathbf{F}_{L-1}\Big)+\operatorname{tr}(\mathbf{P}) \\ &+\sum_{l=2}^{L-2}\Big(1-\prod_{m=l+1}^{L-1} c_m^2\Big)\operatorname{tr}\Bigg(\bigotimes_{m=1}^{l}(c_m\mathbf{H}_m^H\mathbf{F}_m^H)\bigotimes_{m=l}^{1}(c_m\mathbf{F}_m\mathbf{H}_m)\ \mathbf{U}\mathbf{Q}\mathbf{U}^H \\ &+c_l^2\mathbf{F}_l\Big(\sum_{k=2}^{l}\bigotimes_{m=l}^{k}(c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\bigotimes_{m=k}^{l}(c_{m-1}\mathbf{F}_{m-1}^H\mathbf{H}_m^H)+\mathbf{I}_{N_{l+1}}\Big)\mathbf{F}_l^H\Bigg) \\ &+\Big(1-\prod_{m=2}^{L-1} c_m^2\Big)\operatorname{tr}\Big(c_1^2\mathbf{F}_1(\mathbf{H}_1\mathbf{U}\mathbf{Q}\mathbf{U}^H\mathbf{H}_1^H+\mathbf{I}_{N_2})\mathbf{F}_1^H\Big)+\Big(1-\prod_{m=1}^{L-1} c_m^2\Big)\operatorname{tr}(\mathbf{U}\mathbf{Q}\mathbf{U}^H). \end{aligned} \tag{4.37}$$

Using the expressions of P_l^U in (4.22)-(4.24) and P_l^D in (4.27)-(4.29), $l = 1, \cdots, L$, (4.37) can be rewritten as

$$\begin{aligned} P_T^D &= \sum_{l=2}^{L-1} \prod_{m=l}^{L} c_{m-1}^2 P_{L+2-l}^U + c_{L-1}^2 P_2^U + P_1^U + \sum_{l=2}^{L-2} \Big(1 - \prod_{m=l+1}^{L-1} c_m^2\Big) P_{l+1}^D + \sum_{l=1}^{2} \Big(1 - \prod_{m=l}^{L-1} c_m^2\Big) P_l^D \\ &= \sum_{l=1}^{L-1} \prod_{m=l}^{L-1} c_m^2 P_{L+1-l}^U + P_1^U + \sum_{l=1}^{L-1} \Big(1 - \prod_{m=l}^{L-1} c_m^2\Big) P_l^D. \end{aligned}$$

Since the uplink and downlink systems should consume the same amount of total transmission power, we have

$$P_T^D - P_T^U = \sum_{l=1}^{L-1} \Big(\prod_{m=l}^{L-1} c_m^2 - 1\Big)(P_{L+1-l}^U - P_l^D) = 0. \tag{4.38}$$

Obviously, for any $L \geq 2$, (4.38) is true if $\prod_{m=l}^{L-1} c_m^2 = 1$ for $l = 1, \cdots, L-1$, which is equivalent to $c_l = 1, l = 1, \cdots, L-1$. Thus, the first part of Theorem 4.1 (without transmission power constraint at each node) is proven. Moreover, (4.38) also holds if $P_{L+1-l}^U = P_l^D$, $l = 1, \cdots, L-1$. Then we have $P_1^U = P_L^D$ due to $P_T^D = P_T^U$. With transmission power constraint at individual nodes, there is $P_l^D \leq \rho_l^D$ and $P_l^U \leq \rho_l^U$, $l = 1, \cdots, L$. Obviously, to optimize the system performance, all available power should be exploited, i.e., $P_l^D = \rho_l^D$ and $P_l^U = \rho_l^U$, $l = 1, \cdots, L$. Thus, we have $\rho_{L+1-l}^U = \rho_l^D, l = 1, \cdots, L$, and the second part of Theorem 4.1 (with transmission power constraint at individual nodes) is proven.

4.B Proof of Theorem 4.2

With an SIC receiver, the source symbols are detected successively with the last symbol detected first and the first symbol detected last, and the interference from detected symbols is subtracted. Therefore, the SINRs of data streams at the destination node can be written as

$$\mathrm{SINR}_i^U = \frac{\Big|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1} (\mathbf{H}_l^H \mathbf{F}_l^H) \mathbf{H}_L^H \mathbf{v}_i\Big|^2 p_i}{\sum_{j=1}^{i-1} \Big|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1} (\mathbf{H}_l^H \mathbf{F}_l^H) \mathbf{H}_L^H \mathbf{v}_j\Big|^2 p_j + \mathbf{u}_i^H \sum_{l=1}^{L-1} \Big(\bigotimes_{m=1}^{l} (\mathbf{H}_m^H \mathbf{F}_m^H) \bigotimes_{m=l}^{1} (\mathbf{F}_m \mathbf{H}_m)\Big) \mathbf{u}_i + 1}, \quad i = 1, \cdots, N_b. \tag{4.39}$$

For a MIMO relay system employing a DPC transmitter at the source node, the information-bearing symbols are encoded successively with the first symbol encoded

first and the last symbol encoded last, and the interference from encoded symbols is removed. Thus, the SINRs of data streams at the destination node of the DPC-based downlink MIMO relay channel is given by (4.40)

$$\mathrm{SINR}_i^D = \frac{P_{s,i}}{P_{in,i}}, \qquad i = 1, \cdots, N_b, \tag{4.40}$$

where $P_{s,i} \triangleq \left|\mathbf{v}_i^H \mathbf{H}_L \bigotimes_{l=L-1}^{1} (c_l \mathbf{F}_l \mathbf{H}_l) \mathbf{u}_i\right|^2 q_i$ and $P_{in,i} \triangleq \sum_{j=i+1}^{N_b} \left|\mathbf{v}_i^H \mathbf{H}_L \bigotimes_{l=L-1}^{1} (c_l \mathbf{F}_l \mathbf{H}_l) \mathbf{u}_j\right|^2 q_j +$ $\mathbf{v}_i^H \sum_{l=2}^{L} \left(\bigotimes_{m=L}^{l} (c_{m-1} \mathbf{H}_m \mathbf{F}_{m-1}) \bigotimes_{m=l}^{L} (c_{m-1} \mathbf{F}_{m-1}^H \mathbf{H}_m^H) \right) \mathbf{v}_i + 1$. Using (4.39) and (4.40), and the identity of

$$\sum_{i=1}^{N_b} \sum_{j=1}^{i-1} q_i \left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1} (c_l \mathbf{H}_l^H \mathbf{F}_l^H) \mathbf{H}_L^H \mathbf{v}_j\right|^2 p_j = \sum_{i=1}^{N_b} \sum_{j=i+1}^{N_b} p_i \left|\mathbf{v}_i^H \mathbf{H}_L \bigotimes_{l=L-1}^{1} (c_l \mathbf{F}_l \mathbf{H}_l) \mathbf{u}_j\right|^2 q_j$$

we obtain from $\sum_{i=1}^{N_b} \mathrm{SINR}_i^U = \sum_{i=1}^{N_b} \mathrm{SINR}_i^D$ the expression of P_T^D as in (4.33), and the steps in (4.34)-(4.38) remain valid. Thus Theorem 4.2 is proven.

Chapter 5

Interference MIMO Relay Systems

In this chapter, we consider an interference MIMO relay system where multiple source nodes communicate with their desired destination nodes concurrently with the aid of distributed relay nodes all equipped with multiple antennas. We provide a brief overview of existing works on interference systems in Section 5.1. In Section 5.2, the system model of an interference MIMO relay network is introduced. An iterative joint power control and beamforming algorithm is developed in Section 5.3 to minimize the total source and relay transmit power such that a minimum SINR threshold is maintained at each receiver. Section 5.4 shows the simulation results which justify the effectiveness of the proposed algorithm under various scenarios. The chapter is summarized in Section 5.5.

5.1 Overview of Existing Works on Interference Systems

In a large wireless network with many nodes, multiple source-destination links must share a common frequency band concurrently to ensure a high spectral efficiency of the whole network [1]. In such network, CCI is one of the main impairments that degrades the system performance. Developing schemes that mitigate the CCI is therefore important.

By exploiting the spatial diversity, multi-antenna technique provides an efficient approach to CCI minimization [1, 2]. When each source node has a single antenna and the destination nodes are equipped with multiple antennas, a joint power control and

receiver beamforming scheme is developed in [34] to meet the SINR threshold with the minimal transmission power. A joint transmit-receive beamforming and power control algorithm is proposed in [91], when the source nodes also have multiple antennas. Due to the transmit diversity, the total transmit power required in [91] is less than that in [34].

In addition to the transmit and/or receive beamforming considered in [34] and [91], distributed/network beamforming technique [92] can further increase the reliability of the communication link even if the direct path between the transmitter and the receiver is subject to serious degradation, especially for long-distance communication. The network beamforming scheme stems from the idea of cooperative diversity [93]-[94], where users share their communication resources such as bandwidth and transmit power to assist each other in data transmission. The optimal relay matrix design has been recently studied for the MIMO broadcast channel [95] and the point-to-point MIMO relay channel [22], [23]. In [96], a decentralized relay beamforming technique has been developed considering a network of one transmitter, one receiver, and several relay nodes each having a single antenna. In [97], a wireless ad hoc network consisting of multiple source-destination pairs and multiple relay nodes, each having a single antenna, is considered, where the network beamforming scheme is used to meet the SINR threshold at all links with the minimal total transmission power consumed by all relay nodes. Relay beamformers are designed in [98] for multiple-antenna relay nodes with single-antenna source-destination pairs. The non-regenerative MIMO relay technique has been applied to multi-cellular (interference) systems in [57] where transceiver beamformers are designed using the partial zero-forcing (PZF) technique.

However, it is assumed in [57, 97, 98] that each source node uses its maximum available transmit power. Such assumption not only raises the system transmit power consumption, but also increases the interference from one user to all other users. This indicates that the beamforming and the power control problem should be considered jointly as in [34] and [91].

In this chapter, we consider a two-hop interference MIMO relay system consisting of L source-destination pairs communicating with the aid of K relay nodes to enable successful communication over a long distance. Each of the source, relay and destination nodes is equipped with (possibly different number of) multiple antennas. The amplify-and-forward scheme is used at each relay node due to its practical implementation

simplicity. In fact, these relay nodes assist in CCI mitigation by performing distributed network beamforming[1].

We aim at developing a joint power control and beamforming algorithm such that the total transmission power consumed by all source nodes and relay nodes are minimized while maintaining the SINR at each receiver above a minimum threshold value. Compared with [96]-[98], we not only use the network beamforming technique at the relay nodes, but also apply the joint transmit-receive beamforming technique for multiple-antenna users to mitigate the CCI. In contrast to [57], we develop an iterative technique to solve the total power minimization problem rather than using the suboptimal PZF approach. Moreover, transmit power control is used in our algorithm to minimize the total transmit power and the interference to other users, which is not considered in [57, 96–98].

A two-tier iterative algorithm is proposed to jointly optimize the source, relay and receive beamformers, and the source transmission power. We update the relay beamformer in the outer loop using fixed source power, transmit beamformers, and receive beamformers. Since the relay beamforming optimization problem is nonconvex, we use the SDR technique to transform the problem into an SDP problem which can be efficiently solved by interior point-based methods. Then in each iteration of the inner loop, we optimize the receive beamformers first with fixed transmit and relay beamformers and source power. Next, we update the source power such that the target SINR is just met with given transmit, relay and receive beamformers. Finally in the inner loop, we update the transmit beamformers with known transmit power, relay beamformers, and receive beamformers. Numerical simulations are carried out to evaluate the performance of the proposed algorithm.

5.2 Interference MIMO Relay System Model

We consider a two-hop interference MIMO relay system with L source-destination pairs as illustrated in Fig. 5.1. Each source node communicates with its corresponding destination node with the aid of a network of K distributed relays in order to enable

[1] Although the relay beamforming matrices are optimized by a central processing unit in our algorithm, the relay beamforming operation is indeed distributed in the sense that the relays are geographically distributed and they perform beamforming only using their own received signal without exploiting the information on the received signals at other relay nodes.

successful communication over a long distance. The direct links between the source nodes and the destination nodes are not considered as they undergo much larger path attenuations compared with the links via relays. The source and destination nodes of the lth link are equipped with $N_{\mathrm{s},l}$ and $N_{\mathrm{d},l}$ antennas, respectively, whereas the kth relay node is mounted with $N_{\mathrm{r},k}$ antennas. Note that the sum of the transmitting antennas of all source nodes must be smaller than the sum of the relay antennas. Also, for single-stream transmission from each source node, the number of receiving antennas at each destination node must at least equal to the total number of transmitters L in order to suppress $L-1$ independent interferences simultaneously.

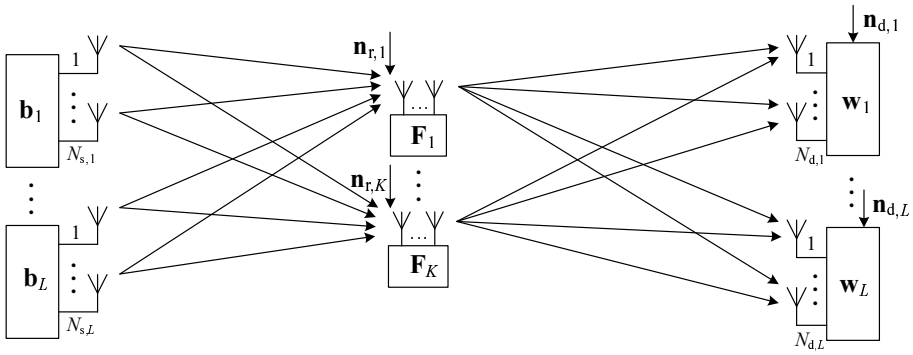

Figure 5.1: Block diagram of an interference MIMO relay system.

We assume that all relay nodes work in half-duplex mode as in [96]-[98]. Thus the communication between the source-destination pairs is completed in two time slots. In the first time slot, the lth source node transmits an $N_{\mathrm{s},l} \times 1$ signal vector $\mathbf{b}_l s_l$, where s_l is the information-carrying symbol and $\mathbf{b}_l$ is the transmit beamforming vector. The received signal vector at the kth relay node is given by

$$\mathbf{y}_{\mathrm{r},k} = \sum_{l=1}^{L} \mathbf{H}_{k,l}\mathbf{b}_l s_l + \mathbf{n}_{\mathrm{r},k}, \qquad k = 1, \cdots, K$$

where $\mathbf{H}_{k,l}$ is the $N_{\mathrm{r},k} \times N_{\mathrm{s},l}$ MIMO channel matrix between the lth transmitting node and the kth relay node and $\mathbf{n}_{\mathrm{r},k}$ is the $N_{\mathrm{r},k} \times 1$ additive Gaussian noise vector at the kth relay node.

In the second time slot, the kth relay node multiplies its received signal vector by an $N_{\mathrm{r},k} \times N_{\mathrm{r},k}$ complex matrix $\mathbf{F}_k$ and transmits the amplitude- and phase-adjusted version of its received signal. Thus the $N_{\mathrm{r},k} \times 1$ signal vector $\mathbf{x}_{\mathrm{r},k}$ transmitted by the kth relay node is given by

$$\mathbf{x}_{\mathrm{r},k} = \mathbf{F}_k \mathbf{y}_{\mathrm{r},k}, \qquad k = 1, \cdots, K. \tag{5.1}$$

The received signal at the lth destination node is obtained as the weighted sum of the received signals at each antenna element of that node, and is given by

$$\begin{aligned} y_{\mathrm{d},l} &= \mathbf{w}_l^H \left(\sum_{k=1}^{K} \mathbf{G}_{l,k}\mathbf{x}_{\mathrm{r},k} + \mathbf{n}_{\mathrm{d},l} \right) \\ &= \mathbf{w}_l^H \left(\sum_{k=1}^{K} \mathbf{G}_{l,k}\mathbf{F}_k \left(\sum_{m=1}^{L} \mathbf{H}_{k,m}\mathbf{b}_m s_m + \mathbf{n}_{\mathrm{r},k} \right) + \mathbf{n}_{\mathrm{d},l} \right), \qquad l = 1, \cdots, L \end{aligned} \tag{5.2}$$

where $\mathbf{G}_{l,k}$ is the $N_{\mathrm{d},l} \times N_{\mathrm{r},k}$ MIMO channel matrix between the kth relay node and the lth destination node, $\mathbf{w}_l$ and $\mathbf{n}_{\mathrm{d},l}$ are the $N_{\mathrm{d},l} \times 1$ receive beamforming weight vector and the additive Gaussian noise vector at the lth destination node, respectively. We assume that all noises are independent and identically distributed (i.i.d.) complex Gaussian noise with zero mean and variance σ_{n}^2.

Let us introduce the following definitions

$$\begin{aligned} \tilde{\mathbf{h}}_l &\triangleq \left[(\mathbf{H}_{1,l}\mathbf{b}_l)^T, \cdots, (\mathbf{H}_{K,l}\mathbf{b}_l)^T \right]^T \in \mathcal{C}^{\bar{N}_{\mathrm{r}} \times 1}, \quad l = 1, \cdots, L \\ \tilde{\mathbf{G}}_l &\triangleq [\mathbf{G}_{l,1}, \cdots, \mathbf{G}_{l,K}] \in \mathcal{C}^{N_{\mathrm{d},l} \times \bar{N}_{\mathrm{r}}}, \quad l = 1, \cdots, L \\ \mathbf{F} &\triangleq \mathrm{blkdiag}\,(\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_K) \in \mathcal{C}^{\bar{N}_{\mathrm{r}} \times \bar{N}_{\mathrm{r}}} \\ \tilde{\mathbf{n}}_{\mathrm{r}} &\triangleq \left[\mathbf{n}_{\mathrm{r},1}^T, \cdots, \mathbf{n}_{\mathrm{r},K}^T \right]^T \in \mathcal{C}^{\bar{N}_{\mathrm{r}} \times 1} \end{aligned}$$

where $\bar{N}_{\mathrm{r}} \triangleq \sum_{k=1}^{K} N_{\mathrm{r},k}$. Here $\tilde{\mathbf{h}}_l$ can be viewed as the effective first-hop channel vector between s_l and all relay nodes, $\tilde{\mathbf{G}}_l$ is the MIMO channel matrix between all relay nodes and the lth receiver, $\mathbf{F}$ is the effective block-diagonal relay precoding matrix, and $\tilde{\mathbf{n}}_{\mathrm{r}}$ is a vector containing the noises at all relay nodes. Using these definitions, (5.2) can be rewritten as

$$\begin{aligned} y_{\mathrm{d},l} &= \mathbf{w}_l^H \tilde{\mathbf{G}}_l \mathbf{F} \left(\sum_{m=1}^{L} \tilde{\mathbf{h}}_m s_m + \tilde{\mathbf{n}}_r \right) + \mathbf{w}_l^H \mathbf{n}_{\mathrm{d},l} \\ &= \mathbf{w}_l^H \left(\sum_{m=1}^{L} \boldsymbol{\psi}_{ml} s_m + \mathbf{n}_l \right), \qquad l = 1, \cdots, L \end{aligned} \tag{5.3}$$

where $\boldsymbol{\psi}_{ml} \triangleq \tilde{\mathbf{G}}_l \mathbf{F} \tilde{\mathbf{h}}_m$ is the equivalent vector channel response between the mth source node and the lth destination node, and $\mathbf{n}_l \triangleq \tilde{\mathbf{G}}_l \mathbf{F} \tilde{\mathbf{n}}_r + \mathbf{n}_{\mathrm{d},l}$ is the equivalent noise vector at the lth receiver.

From (5.3), the total power of the received signal at the destination node of the lth link is given by

$$\mathrm{E}\left[y_{\mathrm{d},l} y_{\mathrm{d},l}^* \right] = \sum_{m=1}^{L} p_m \mathbf{w}_l^H \boldsymbol{\psi}_{ml} \boldsymbol{\psi}_{ml}^H \mathbf{w}_l + \mathbf{w}_l^H \mathbf{C}_l \mathbf{w}_l, \qquad l = 1, \cdots, L \tag{5.4}$$

where $\mathbf{C}_l \triangleq \sigma_{\mathrm{n}}^2 \tilde{\mathbf{G}}_l \mathbf{F}\mathbf{F}^H \tilde{\mathbf{G}}_l^H + \sigma_{\mathrm{n}}^2 \mathbf{I}_{N_{\mathrm{d},l}}$ is the covariance matrix of $\mathbf{n}_l$. Here we assume that $\mathrm{E}\left[|s_l|^2\right] = p_l$ is the transmit power of the lth information-carrying symbol. Based on (5.4), the SINR at the lth destination node is given by

$$\Gamma_l = \frac{p_l \mathbf{w}_l^H \boldsymbol{\psi}_{ll} \boldsymbol{\psi}_{ll}^H \mathbf{w}_l}{\sum_{m \neq l}^{L} p_m \mathbf{w}_l^H \boldsymbol{\psi}_{ml} \boldsymbol{\psi}_{ml}^H \mathbf{w}_l + \mathbf{w}_l^H \mathbf{C}_l \mathbf{w}_l}, \qquad l = 1, \cdots, L. \tag{5.5}$$

Using (5.1), the transmission power consumed by the kth relay node can be expressed as

$$P_{\mathrm{r},k} = \mathrm{tr}\big(\mathrm{E}\big[\mathbf{x}_{\mathrm{r},k} \mathbf{x}_{\mathrm{r},k}^H\big]\big) = \mathrm{tr}\big(\mathbf{F}_k \mathbf{R}_{\mathrm{y},k} \mathbf{F}_k^H\big), \qquad k = 1, \cdots, K \tag{5.6}$$

where $\mathbf{R}_{\mathrm{y},k} \triangleq \mathrm{E}\big[\mathbf{y}_{\mathrm{r},k} \mathbf{y}_{\mathrm{r},k}^H\big] = \sum_{l=1}^{L} p_l \mathbf{H}_{k,l} \mathbf{b}_l \mathbf{b}_l^H \mathbf{H}_{k,l}^H + \sigma_{\mathrm{n}}^2 \mathbf{I}_{N_{\mathrm{r},k}}$ is the covariance matrix of the received signal vector at the kth relay node. Using (5.6), the total transmit power consumed by the whole network can be expressed as

$$P_{\mathrm{T}} = \sum_{k=1}^{K} P_{\mathrm{r},k} + \sum_{l=1}^{L} p_l \mathbf{b}_l^H \mathbf{b}_l. \tag{5.7}$$

5.3 Joint Power Control and Beamforming

Let us define the relay beamforming vector $\mathbf{f}$ from the relay amplifying matrices $\mathbf{F}_1, \cdots, \mathbf{F}_K$ as

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \triangleq \mathrm{vec}(\mathbf{F}_1) \\ \vdots \\ \mathbf{f}_K \triangleq \mathrm{vec}(\mathbf{F}_K) \end{bmatrix} \in \mathcal{C}^{\tilde{N}_{\mathrm{r}} \times 1} \tag{5.8}$$

where $\tilde{N}_{\mathrm{r}} \triangleq \sum_{k=1}^{K} N_{\mathrm{r},k}^2$, and $\mathrm{vec}(\cdot)$ stands for a vector obtained by stacking all column vectors of a matrix on top of each other. In this section, we design the source transmit power vector $\mathbf{p} \triangleq [p_1, p_2, \cdots, p_L]^T$, the relay beamforming vector $\mathbf{f}$, transmit beamforming vectors $\{\mathbf{b}_l\} \triangleq \{\mathbf{b}_l, l = 1, \cdots, L\}$, and receive beamforming vectors $\{\mathbf{w}_l\} \triangleq \{\mathbf{w}_l, l = 1, \cdots, L\}$, such that a target SINR threshold $\gamma_l > 0, l = 1, \cdots, L$, is maintained at the lth destination node with the minimal P_{T}. The optimization problem can be written as

$$\min_{\mathbf{p}, \mathbf{f}, \{\mathbf{b}_l\}, \{\mathbf{w}_l\}} \quad P_{\mathrm{T}} \tag{5.9a}$$

$$\text{s.t.} \quad \Gamma_l \geq \gamma_l, \qquad l = 1, \cdots, L. \tag{5.9b}$$

The problem (5.9) is nonconvex due to the constraints in (5.9b). We propose a two-tier iterative algorithm to efficiently solve the problem (5.9). In the following, we solve corresponding subproblems to optimize each variable.

5.3.1 Receive beamforming

The optimal receive beamforming vectors $\mathbf{w}_l$, $l = 1, \cdots, L$, for fixed $\mathbf{p}$, $\mathbf{f}$, and $\{\mathbf{b}_l\}$ can be obtained such that it minimizes the noise-plus-interference power at the receiver under the condition of unity gain for the signal of interest, which can be written as

$$\min_{\mathbf{w}_l} \quad \sum_{m \neq l}^{L} p_m \mathbf{w}_l^H \boldsymbol{\psi}_{ml} \boldsymbol{\psi}_{ml}^H \mathbf{w}_l + \mathbf{w}_l^H \mathbf{C}_l \mathbf{w}_l \tag{5.10a}$$

$$\text{s.t.} \quad \mathbf{w}_l^H \boldsymbol{\psi}_{ll} = 1. \tag{5.10b}$$

The unity gain condition ensures that the desired signal is unaffected by beamforming. Using the Lagrangian multiplier method, the solution to the problem (5.10) is given by

$$\mathbf{w}_l = \frac{\boldsymbol{\Phi}_l^{-1} \boldsymbol{\psi}_{ll}}{\boldsymbol{\psi}_{ll}^H \boldsymbol{\Phi}_l^{-1} \boldsymbol{\psi}_{ll}} \tag{5.11}$$

where $\boldsymbol{\Phi}_l \triangleq \sum_{m \neq l}^{L} p_m \boldsymbol{\psi}_{ml} \boldsymbol{\psi}_{ml}^H + \mathbf{C}_l$ is the interference-plus-noise covariance matrix at the lth receiver.

5.3.2 Transmit power allocation

To obtain optimal $\mathbf{p}$ with given beamforming vectors $\mathbf{f}$, $\{\mathbf{b}_l\}$, and $\{\mathbf{w}_l\}$, we reformulate the problem (5.9) as

$$\min_{\mathbf{p}} \quad P_{\mathrm{T}} \tag{5.12a}$$

$$\text{s.t.} \quad \frac{p_l [\bar{\mathbf{H}}]_{l,l}}{\sum_{m \neq l}^{L} p_m [\bar{\mathbf{H}}]_{m,l} + \bar{n}_l} \geq \gamma_l, \quad l = 1, \cdots, L \tag{5.12b}$$

where $\bar{\mathbf{H}}$ is an $L \times L$ covariance matrix such that $[\bar{\mathbf{H}}]_{m,l} = \mathbf{w}_l^H \boldsymbol{\psi}_{ml} \boldsymbol{\psi}_{ml}^H \mathbf{w}_l$ and $\bar{n}_l \triangleq \mathbf{w}_l^H \mathbf{C}_l \mathbf{w}_l$. Here for a matrix $\mathbf{A}$, $[\mathbf{A}]_{i,j}$ indicates the (i,j)th element of $\mathbf{A}$. In the optimal power allocation, the transmit power of each user is set to the minimum required level such that the target SINR is just met [34], [91]. That is, the constraints in (5.12b) should hold with equality as

$$\frac{p_l [\bar{\mathbf{H}}]_{l,l}}{\sum_{m \neq l}^{L} p_m [\bar{\mathbf{H}}]_{m,l} + \bar{n}_l} = \gamma_l, \qquad l = 1, \cdots, L \tag{5.13}$$

which can be equivalently rewritten as

$$p_l = \frac{\gamma_l}{[\bar{\mathbf{H}}]_{l,l}} \left(\sum_{m \neq l}^{L} p_m [\bar{\mathbf{H}}]_{m,l} + \bar{n}_l \right), \qquad l = 1, \cdots, L. \tag{5.14}$$

Equation (5.14) can be written in matrix form as

$$\mathbf{p} = \breve{\mathbf{H}}\mathbf{p} + \mathbf{u} \tag{5.15}$$

where $[\breve{\mathbf{H}}]_{l,m} = \begin{cases} 0, & m = l \\ \gamma_l [\bar{\mathbf{H}}]_{m,l}/[\bar{\mathbf{H}}]_{l,l}, & m \neq l \end{cases}$, and $\mathbf{u}$ is an $L \times 1$ vector whose lth element is given by $\gamma_l \bar{n}_l/[\bar{\mathbf{H}}]_{l,l}$, $l = 1, \cdots, L$. From (5.15), it can be seen that the optimal solution to the problem (5.12) is given by

$$\mathbf{p} = (\mathbf{I}_L - \breve{\mathbf{H}})^{-1}\mathbf{u}. \tag{5.16}$$

5.3.3 Transmit beamforming

With given $\mathbf{p}, \mathbf{f}$ and $\{\mathbf{w}_l\}$, the optimal $\{\mathbf{b}_l\}$ can be obtained simply by swapping the roles of the transmitters and the receivers as in [99]. First we rewrite the objective function by substituting $P_{\mathrm{r},k}$ in (5.6) into (5.7) as

$$P_{\mathrm{T}} = \sum_{l=1}^{L} p_l \left[\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{Q}_{k,l}\mathbf{b}_l\mathbf{b}_l^H\right) + \mathbf{b}_l^H\mathbf{b}_l \right] + \sigma_{\mathrm{n}}^2 \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{F}_k\mathbf{F}_k^H\right) \tag{5.17}$$

where $\mathbf{Q}_{k,l} \triangleq \mathbf{H}_{k,l}^H\mathbf{F}_k^H\mathbf{F}_k\mathbf{H}_{k,l}$. Let us now denote $\mathbf{Q}_l \triangleq \sum_{k=1}^{K}\mathbf{Q}_{k,l}$ and $\tilde{\mathbf{b}}_l \triangleq \left(\mathbf{Q}_l + \mathbf{I}_{N_{\mathrm{s},l}}\right)^{\frac{1}{2}}\mathbf{b}_l$. Thus (5.17) can be equivalently written as

$$P_{\mathrm{T}} = \sum_{l=1}^{L} p_l \tilde{\mathbf{b}}_l^H\tilde{\mathbf{b}}_l + \sigma_{\mathrm{n}}^2 \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{F}_k\mathbf{F}_k^H\right). \tag{5.18}$$

Since the equivalent noise $\mathbf{n}_l$ at the lth destination node is non-white, we need to perform the pre-whitening operation before we swap the roles of the transmitting and the receiving nodes. After the pre-whitening and receive beamforming operations, the received signal at the lth destination node can be expressed as

$$\begin{aligned} \tilde{y}_{\mathrm{d},l} &= \mathbf{w}_l^H\mathbf{C}_l^{-\frac{1}{2}}\tilde{\mathbf{G}}_l\mathbf{F}\left(\sum_{m=1}^{L}\tilde{\mathbf{H}}_m\mathbf{b}_m s_m + \tilde{\mathbf{n}}_r\right) + \mathbf{w}_l^H\mathbf{C}_l^{-\frac{1}{2}}\mathbf{n}_{\mathrm{d},l} \\ &= \mathbf{w}_l^H\left(\sum_{m=1}^{L}\bar{\mathbf{G}}_{m,l}\tilde{\mathbf{b}}_m s_m + \mathbf{C}_l^{-\frac{1}{2}}\tilde{\mathbf{G}}_l\mathbf{F}\tilde{\mathbf{n}}_r + \mathbf{C}_l^{-\frac{1}{2}}\mathbf{n}_{\mathrm{d},l}\right), \quad l = 1, \cdots, L \end{aligned} \tag{5.19}$$

where $\tilde{\mathbf{H}}_l \triangleq \left[\mathbf{H}_{1,l}^T, \cdots, \mathbf{H}_{K,l}^T\right]^T$ is the equivalent MIMO channel between the lth source node and all relay nodes and $\bar{\mathbf{G}}_{m,l} \triangleq \mathbf{C}_l^{-\frac{1}{2}}\tilde{\mathbf{G}}_l\mathbf{F}\tilde{\mathbf{H}}_m\left(\mathbf{Q}_m + \mathbf{I}_{N_{\mathrm{s},m}}\right)^{-\frac{1}{2}}$.

It can be seen from (5.19) that the equivalent noise is now white, and the received SINR in the lth virtual link (where $\tilde{\mathbf{b}}_l^*$ is the receive beamforming vector and $\mathbf{w}_l^*$ becomes the transmit beamforming vector) can be expressed as

$$\tilde{\Gamma}_l = \frac{\tilde{p}_l \tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ll} \boldsymbol{\xi}_{ll}^H \tilde{\mathbf{b}}_l^*}{\sum_{m\neq l}^L \tilde{p}_m \tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ml} \boldsymbol{\xi}_{ml}^H \tilde{\mathbf{b}}_l^* + \tilde{\mathbf{b}}_l^T \tilde{\mathbf{b}}_l^*}, \qquad l = 1, \cdots, L. \tag{5.20}$$

Here $\boldsymbol{\xi}_{ml} \triangleq \bar{\mathbf{G}}_{l,m}^T \mathbf{w}_m^*$, $\tilde{p}_l$ is the transmit power in the lth virtual link. Note that since the noise in the original link is pre-whitened before we swap the roles of transmitters and receivers, the equivalent virtual link noise is also white with unit-variance. Thus, the corresponding noise power after the receive beamforming is given by $\tilde{\mathbf{b}}_l^T \tilde{\mathbf{b}}_l^*$ in (5.20).

The optimal $\{\tilde{\mathbf{b}}_l^*\}$ can be obtained from (5.20) by solving the following problem for each $l = 1, \cdots, L$

$$\min_{\tilde{\mathbf{b}}_l} \quad \sum_{m\neq l}^L \tilde{p}_m \tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ml} \boldsymbol{\xi}_{ml}^H \tilde{\mathbf{b}}_l^* + \tilde{\mathbf{b}}_l^T \tilde{\mathbf{b}}_l^* \tag{5.21a}$$

$$\text{s.t.} \quad \tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ll} = 1. \tag{5.21b}$$

The solution to this problem is given by

$$\tilde{\mathbf{b}}_l^* = \frac{\boldsymbol{\Theta}_l^{-1} \boldsymbol{\xi}_{ll}}{\boldsymbol{\xi}_{ll}^H \boldsymbol{\Theta}_l^{-1} \boldsymbol{\xi}_{ll}} \tag{5.22}$$

where $\boldsymbol{\Theta}_l \triangleq \sum_{m\neq l}^L \tilde{p}_m \boldsymbol{\xi}_{ml} \boldsymbol{\xi}_{ml}^H + \mathbf{I}_{N_{\mathrm{s},l}}$ is the noise-plus-interference covariance matrix at the lth receiver of the virtual link. The transmit power of the virtual link can be obtained as

$$\tilde{\mathbf{p}} \triangleq (\mathbf{I}_L - \breve{\mathbf{G}})^{-1} \tilde{\mathbf{u}} \tag{5.23}$$

where $[\breve{\mathbf{G}}]_{l,m} = \begin{cases} 0, & m = l \\ \gamma_l \tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ml} \boldsymbol{\xi}_{ml}^H \tilde{\mathbf{b}}_l^* / (\tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ll} \boldsymbol{\xi}_{ll}^H \tilde{\mathbf{b}}_l^*), & m \neq l \end{cases}$, and $[\tilde{\mathbf{u}}]_l \triangleq \frac{\gamma_l \tilde{\mathbf{b}}_l^T \tilde{\mathbf{b}}_l^*}{\tilde{\mathbf{b}}_l^T \boldsymbol{\xi}_{ll} \boldsymbol{\xi}_{ll}^H \tilde{\mathbf{b}}_l^*}$, $l = 1, \cdots, L$. Here for a vector $\mathbf{v}$, $[\mathbf{v}]_l$ stands for the lth element of $\mathbf{v}$.

5.3.4 Relay beamforming

In this subsection we optimize the relay amplifying matrices such that the total relay transmit power is minimized while satisfying the SINR constraints in (4.9b). First, (5.4) can be rewritten as

$$\begin{aligned} \mathrm{E}\left[y_{\mathrm{d},l} y_{\mathrm{d},l}^*\right] =& \sum_{m=1}^L p_m \mathrm{tr}\big(\tilde{\mathbf{G}}_l^H \mathbf{w}_l \mathbf{w}_l^H \tilde{\mathbf{G}}_l \mathbf{F} \tilde{\mathbf{h}}_m \tilde{\mathbf{h}}_m^H \mathbf{F}^H\big) + \sigma_{\mathrm{n}}^2 \mathrm{tr}\big(\tilde{\mathbf{G}}_l^H \mathbf{w}_l \mathbf{w}_l^H \tilde{\mathbf{G}}_l \mathbf{F} \mathbf{F}^H\big) + \sigma_{\mathrm{n}}^2 \mathbf{w}_l^H \mathbf{w}_l \\ =& \sum_{m=1}^L \mathrm{tr}\big(\mathbf{R}_{\mathrm{g},l} \mathbf{F} \mathbf{R}_{\mathrm{h},m} \mathbf{F}^H\big) + \sigma_{\mathrm{n}}^2 \mathrm{tr}\big(\mathbf{R}_{\mathrm{g},l} \mathbf{F} \mathbf{F}^H\big) + \sigma_{\mathrm{n}}^2 \mathbf{w}_l^H \mathbf{w}_l, \quad l = 1, \cdots, L \end{aligned} \tag{5.24}$$

where $\mathbf{R}_{\mathrm{g},l} \triangleq \tilde{\mathbf{G}}_l^H \mathbf{w}_l \mathbf{w}_l^H \tilde{\mathbf{G}}_l$, $l = 1, \cdots, L$, and $\mathbf{R}_{\mathrm{h},m} \triangleq p_m \tilde{\mathbf{h}}_m \tilde{\mathbf{h}}_m^H$, $m = 1, \cdots, L$. Using (5.24), the SINR of the lth link in (5.5) can be expressed as

$$\Gamma_l = \frac{\mathrm{tr}\big(\mathbf{R}_{\mathrm{g},l}\mathbf{F}\mathbf{R}_{\mathrm{h},l}\mathbf{F}^H\big)}{\mathrm{tr}\big(\mathbf{R}_{\mathrm{g},l}\mathbf{F}\big(\sum_{m\neq l}^{L}\mathbf{R}_{\mathrm{h},m} + \sigma_{\mathrm{n}}^2\mathbf{I}_{\tilde{N}_{\mathrm{r}}}\big)\mathbf{F}^H\big) + \sigma_{\mathrm{n}}^2\mathbf{w}_l^H\mathbf{w}_l}, \qquad l = 1, \cdots, L. \tag{5.25}$$

Applying the fact that $\mathrm{tr}(\mathbf{A}^H\mathbf{B}\mathbf{A}\mathbf{C}) = \mathrm{vec}(\mathbf{A})^H(\mathbf{C}^T \bigotimes \mathbf{B})\mathrm{vec}(\mathbf{A})$ [100], where $\bigotimes$ denotes the matrix Kronecker product, (5.25) can be expressed as

$$\Gamma_l = \frac{\mathrm{vec}(\mathbf{F})^H\big(\mathbf{R}_{\mathrm{h},l}^T \bigotimes \mathbf{R}_{\mathrm{g},l}\big)\mathrm{vec}(\mathbf{F})}{\mathrm{vec}(\mathbf{F})^H\big(\tilde{\mathbf{R}}_{\mathrm{h},l}^T \bigotimes \mathbf{R}_{\mathrm{g},l}\big)\mathrm{vec}(\mathbf{F}) + \sigma_{\mathrm{n}}^2\mathbf{w}_l^H\mathbf{w}_l}, \qquad l = 1, \cdots, L \tag{5.26}$$

where $\tilde{\mathbf{R}}_{\mathrm{h},l} \triangleq \sum_{m\neq l}^{L}\mathbf{R}_{\mathrm{h},m} + \sigma_{\mathrm{n}}^2\mathbf{I}_{\tilde{N}_{\mathrm{r}}}$. Let us now introduce the link between $\mathbf{f}$ in (5.8) and $\mathrm{vec}(\mathbf{F})$ as $\mathrm{vec}(\mathbf{F}) = \mathbf{D}_{\mathrm{F}}\mathbf{f}$, where $\mathbf{D}_{\mathrm{F}} \in \mathcal{R}^{\tilde{N}_{\mathrm{r}}^2 \times \tilde{N}_{\mathrm{r}}}$ is a matrix of ones and zeros and is constructed by observing the nonzero entries of $\mathrm{vec}(\mathbf{F})$. Note that $\mathbf{D}_{\mathrm{F}}$ does not depend on the exact numeric value of $\mathrm{vec}(\mathbf{F})$, instead it depends on the way the entries of $\mathbf{f}$ are taken to form $\mathrm{vec}(\mathbf{F})$. As an example, for a system with two relay nodes each having two antennas, there is $\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0}_{2\times 2} \\ \mathbf{0}_{2\times 2} & \mathbf{F}_2 \end{bmatrix}$ with $\mathbf{F}_1 = [\mathbf{f}_{1,1}, \mathbf{f}_{1,2}]$ and $\mathbf{F}_2 = [\mathbf{f}_{2,1}, \mathbf{f}_{2,2}]$, where $\mathbf{f}_{i,j}$, $i, j = 1, 2$, are 2×1 vectors and $\mathbf{0}_{m\times n}$ denotes an $m \times n$ matrix with all zero elements. In this case, we have

$$\mathrm{vec}(\mathbf{F}) = \big[\mathbf{f}_{1,1}^T, \mathbf{0}_{1\times 2}, \mathbf{f}_{1,2}^T, \mathbf{0}_{1\times 2}, \mathbf{0}_{1\times 2}, \mathbf{f}_{2,1}^T, \mathbf{0}_{1\times 2}, \mathbf{f}_{2,2}^T\big]^T, \qquad \mathbf{f} = \big[\mathbf{f}_{1,1}^T, \mathbf{f}_{1,2}^T, \mathbf{f}_{2,1}^T, \mathbf{f}_{2,2}^T\big]^T.$$

Therefore, to obtain $\mathrm{vec}(\mathbf{F}) = \mathbf{D}_{\mathrm{F}}\mathbf{f}$, matrix $\mathbf{D}_{\mathrm{F}}$ should be constructed as

$$\mathbf{D}_{\mathrm{F}} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0}_{4\times 2} & \mathbf{0}_{10\times 2} & \mathbf{0}_{2\times 2} \\ \mathbf{0}_{2\times 2} & \mathbf{I}_2 & \mathbf{I}_2 & \mathbf{0}_{12\times 2} \\ \mathbf{0}_{12\times 2} & \mathbf{0}_{10\times 2} & \mathbf{0}_{4\times 2} & \mathbf{I}_2 \end{pmatrix}.$$

Now (5.26) can be rewritten as

$$\Gamma_l = \frac{\mathbf{f}^H\mathbf{D}_{\mathrm{F}}^T\big(\mathbf{R}_{\mathrm{h},l}^T \bigotimes \mathbf{R}_{\mathrm{g},l}\big)\mathbf{D}_{\mathrm{F}}\mathbf{f}}{\mathbf{f}^H\mathbf{D}_{\mathrm{F}}^T\big(\tilde{\mathbf{R}}_{\mathrm{h},l}^T \bigotimes \mathbf{R}_{\mathrm{g},l}\big)\mathbf{D}_{\mathrm{F}}\mathbf{f} + \sigma_{\mathrm{n}}^2\mathbf{w}_l^H\mathbf{w}_l}, \qquad l = 1, \cdots, L. \tag{5.27}$$

From (5.8), we have $\mathbf{f}_k = \mathbf{D}_k\mathbf{f}, k = 1, \cdots, K$, with $\mathbf{D}_k \in \mathcal{R}^{N_{\mathrm{r},k}^2 \times \tilde{N}_{\mathrm{r}}}$ defined as $\mathbf{D}_k = [\mathbf{D}_{k,1}, \cdots, \mathbf{D}_{k,K}]$, where $\mathbf{D}_{k,k} = \mathbf{I}_{N_{\mathrm{r},k}^2 \times N_{\mathrm{r},k}^2}$ and $\mathbf{D}_{k,j} = \mathbf{0}_{N_{\mathrm{r},k}^2 \times N_{\mathrm{r},j}^2}$, $j = 1, \cdots, K$, $j \neq k$. By using the identity of $\mathrm{tr}(\mathbf{A}^H\mathbf{A}\mathbf{B}) = \mathrm{vec}(\mathbf{A})^H\big(\mathbf{B}^T \bigotimes \mathbf{I}_n\big)\mathrm{vec}(\mathbf{A})$ for $\mathbf{A}, \mathbf{B} \in \mathcal{C}^{n\times n}$ [100], the transmit power of the kth relay node in (5.6) can be expressed as

$$P_{\mathrm{r},k} = \mathbf{f}_k^H\big(\mathbf{R}_{\mathrm{y},k}^T \bigotimes \mathbf{I}_{N_{\mathrm{r},k}}\big)\mathbf{f}_k = \mathbf{f}^H\mathbf{D}_k^T\big(\mathbf{R}_{\mathrm{y},k}^T \bigotimes \mathbf{I}_{N_{\mathrm{r},k}}\big)\mathbf{D}_k\mathbf{f}, \qquad k = 1, \cdots, K. \tag{5.28}$$

Using (5.27) and (5.28), with given $\mathbf{p}$, $\{\mathbf{b}_l\}$ and $\{\mathbf{w}_l\}$, the problem (5.9) can be reformulated as the following nonconvex QCQP problem

$$\min_{\mathbf{f}} \quad \mathbf{f}^H\mathbf{A}\mathbf{f} \tag{5.29a}$$

$$\text{s.t.} \quad \mathbf{f}^H\mathbf{B}_l\mathbf{f} \geq \gamma_l\sigma_{\text{n}}^2\mathbf{w}_l^H\mathbf{w}_l, \qquad l = 1, \cdots, L \tag{5.29b}$$

where we introduce

$$\mathbf{A} \triangleq \sum_{k=1}^{K} \mathbf{D}_k^T(\mathbf{R}_{\text{y},k}^T \bigotimes \mathbf{I}_{N_{\text{r},k}})\mathbf{D}_k, \quad \mathbf{B}_l \triangleq \mathbf{D}_{\text{F}}^T(\mathbf{R}_{\text{h},l}^T \bigotimes \mathbf{R}_{\text{g},l} - \gamma_l\tilde{\mathbf{R}}_{\text{h},l}^T \bigotimes \mathbf{R}_{\text{g},l})\mathbf{D}_{\text{F}}, \quad l = 1, \cdots, L. \tag{5.30}$$

The problem (5.29) is non-convex, since $\mathbf{B}_l$ in (5.30) can be indefinite. In the following, we resort to the SDR technique [101]-[102] to solve the problem (5.29). By introducing $\mathbf{X} = \mathbf{f}\,\mathbf{f}^H$, the problem (5.29) can be equivalently rewritten as

$$\min_{\mathbf{X}} \quad \text{tr}(\mathbf{A}\mathbf{X}) \tag{5.31a}$$

$$\text{s.t.} \quad \text{tr}(\mathbf{B}_l\mathbf{X}) \geq \gamma_l\sigma_{\text{n}}^2\mathbf{w}_l^H\mathbf{w}_l, \quad l = 1, \cdots, L \tag{5.31b}$$

$$\mathbf{X} \succcurlyeq 0 \tag{5.31c}$$

$$\text{rank}(\mathbf{X}) = 1. \tag{5.31d}$$

Note that in the problem (5.31), the cost function is linear in $\mathbf{X}$, the trace constraints are linear inequalities in $\mathbf{X}$, and the PSD matrix constraint is convex. However, the rank constraint on $\mathbf{X}$ is not convex. Interestingly, the problem (5.31) can be solved by the SDR technique [101]-[102] as explained in the following. First we drop the rank constraint (5.31d) to obtain the following relaxed SDP problem which is convex in $\mathbf{X}$.

$$\min_{\mathbf{X}} \quad \text{tr}(\mathbf{A}\mathbf{X}) \tag{5.32a}$$

$$\text{s.t.} \quad \text{tr}(\mathbf{B}_l\mathbf{X}) \geq \gamma_l\sigma_{\text{n}}^2\mathbf{w}_l^H\mathbf{w}_l, \quad l = 1, \cdots, L \tag{5.32b}$$

$$\mathbf{X} \succcurlyeq 0. \tag{5.32c}$$

SDP problems like (5.32) can be conveniently solved by using interior point methods at a complexity order that is at most $\mathcal{O}((L + \tilde{N}_{\text{r}}^2)^{3.5})$ [66]. One can use, for example, the CVX MATLAB toolbox for disciplined convex programming [63] to obtain the optimal $\mathbf{X}$. Due to the relaxation, $\mathbf{X}_{\text{opt}}$ obtained by solving the problem (5.32) is not necessarily rank one in general. If it is, then its principal eigenvector (scaled by the square root of the principal eigenvalue of $\mathbf{X}_{\text{opt}}$) is the optimal solution $\mathbf{f}_{\text{opt}}$ to the

original problem (5.29). If $\text{rank}(\mathbf{X}_{\text{opt}}) \geq 3$ and $L \leq 4$, the recent results on Hermitian matrix rank-one decomposition in [103] can be used to generate the exact optimal $\mathbf{f}_{\text{opt}}$ for the problem (5.29) based on $\mathbf{X}_{\text{opt}}$. Otherwise, we may resort to alternative techniques such as randomization [101]-[102] to obtain a (suboptimal) $\mathbf{f}$ from $\mathbf{X}_{\text{opt}}$. Different randomization techniques have been studied in the literature [101]-[102]. The one we choose is summarized in Table 5.1. Note that using this approach, some of the constraints in (4.9b) may be violated after the randomization operation. However, a feasible relay beamforming vector can be obtained by simply scaling $\mathbf{f}$ so that all the constraints are satisfied.

Table 5.1: Randomization technique for semidefinite relaxation approach

1. Let $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$ be the eigenvalue decomposition of $\mathbf{X}$.
2. Choose an $\tilde{N}_{\text{r}} \times 1$ random vector $\mathbf{v}$ whose elements are independent random variables, uniformly distributed on the unit circle in the complex plane, i.e., $[\mathbf{v}]_i = e^{j\theta_i}, i = 1, \cdots, \tilde{N}_{\text{r}}$, where θ_i is independent and uniformly distributed on $[0, 2\pi)$.
3. Choose $\mathbf{f} = \mathbf{U}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{v}$ which ensures that $\mathbf{f}^H\mathbf{f} = \text{tr}(\mathbf{X})$.

Now the original total transmit power minimization problem (5.9) can be solved by an iterative algorithm as shown in Table 5.2. Here ε_i, $i = 1, 2$, are small positive numbers close to zero up to which convergence is acceptable, max stands for the maximal element of a vector, and the superscript (m) and $[n]$ denotes the number of iterations at the outer loop and the inner loop, respectively. It can be seen from Table II that the proposed algorithm iteratively optimizes two blocks of variables: (i) The relay weighting coefficients $\mathbf{f}$; (ii) The transmit beamformer vectors $\{\mathbf{b}_l\}$, the receive beamformer vectors $\{\mathbf{w}_l\}$, and the transmit power vector $\mathbf{p}$. With fixed $\mathbf{f}$, we solve the problem of optimizing $\{\mathbf{b}_l\}$, $\{\mathbf{w}_l\}$, and $\mathbf{p}$ through step (3) in Table 5.2. In fact, this problem is similar to the joint transceiver design problem in a single-hop MIMO interference channel [91]. Therefore, it can be shown similar to [91] that the inner iteration in step (3) converges to the optimal solution of $\{\mathbf{b}_l\}$, $\{\mathbf{w}_l\}$, and $\mathbf{p}$ for a given $\mathbf{f}$. With fixed $\{\mathbf{b}_l\}$, $\{\mathbf{w}_l\}$, and $\mathbf{p}$, we optimize $\mathbf{f}$ through step (2) in Table 5.2.

In numerical simulations we observe that the outer loop converges typically within 3 to 5 iterations, while the inner loop converges usually within 3 iterations. However, a rigorous analysis on whether the outer loop converges to a locally optimal solution

Table 5.2: Procedure of solving the problem (5.9) by the proposed iterative algorithm

1. Initialize the algorithm with an arbitrary forward link power vector $\mathbf{p}^{(0)}$, virtual link power vector $\tilde{\mathbf{p}}^{(0)}$, and randomly generated transmit beamforming vectors $\{\mathbf{b}_l^{(0)}\}$ and receive beamforming vectors $\{\mathbf{w}_l^{(0)}\}$; Set $m = 0$.

2. Solve the subproblem (5.32) using known $\{\mathbf{b}_l^{(m)}\}$, $\{\mathbf{w}_l^{(m)}\}$, and $\mathbf{p}^{(m)}$ to obtain $\mathbf{X}$.
   If $\text{rank}(\mathbf{X}) = 1$, obtain $\mathbf{f}^{(m)}$ as the principal eigenvector of $\mathbf{X}$ scaled by the square root of its principal eigenvalue.

   If $\text{rank}(\mathbf{X}) \geq 3$ and $L \leq 4$, use the approaches in [103] to obtain $\mathbf{f}^{(m)}$.

   Otherwise

   (a) Use the randomization technique in Table 5.1 to obtain $\mathbf{f}$.

   (b) Find the most violated constraint in the original problem (5.9) using such $\mathbf{f}$.

   (c) Scale $\mathbf{f}$ so that the most violated constraint is satisfied with equality to obtain $\mathbf{f}^{(m)}$.

3. Set $n = 0$, $\mathbf{p}^{[0]} = \mathbf{p}^{(m)}$, $\{\mathbf{b}_l^{[0]}\} = \{\mathbf{b}_l^{(m)}\}$, $\tilde{\mathbf{p}}^{[0]} = \tilde{\mathbf{p}}^{(m)}$, and

   (a) Solve the subproblem (5.10) using given $\mathbf{p}^{[n]}$, $\{\mathbf{b}_l^{[n]}\}$, and $\mathbf{f}^{(m)}$ to obtain $\{\mathbf{w}_l^{[n+1]}\}$ as in (5.11).

   (b) Solve the subproblem (5.12) with fixed $\mathbf{f}^{(m)}$, $\{\mathbf{b}_l^{[n]}\}$, and $\{\mathbf{w}_l^{[n+1]}\}$ to obtain power vector $\mathbf{p}^{[n+1]}$ as in (5.16).

   (c) Update the transmit beamforming vectors $\{\mathbf{b}_l^{[n+1]}\}$ by solving the subproblem (5.21) with given $\mathbf{f}^{(m)}$, $\{\mathbf{w}_l^{[n+1]}\}$, and $\tilde{\mathbf{p}}^{[n]}$.

   (d) Update the virtual link transmit power $\tilde{\mathbf{p}}^{[n+1]}$ with fixed $\{\mathbf{b}_l^{[n+1]}\}$, $\{\mathbf{w}_l^{[n+1]}\}$, and $\mathbf{f}^{(m)}$ as in (5.23).

   (e) If $\max\left|\mathbf{p}^{[n+1]} - \mathbf{p}^{[n]}\right| \leq \varepsilon_1$, then $\mathbf{p}^{(m+1)} = \mathbf{p}^{[n+1]}$, $\{\mathbf{b}_l^{(m+1)}\} = \{\mathbf{b}_l^{[n+1]}\}$, $\{\mathbf{w}_l^{(m+1)}\} = \{\mathbf{w}_l^{[n+1]}\}$, $\tilde{\mathbf{p}}^{(m+1)} = \tilde{\mathbf{p}}^{[n+1]}$; end of step 3.
   Otherwise, let $n := n + 1$ and go to step 3a.

4. If $\max\left|\mathbf{p}^{(m+1)} - \mathbf{p}^{(m)}\right| \leq \varepsilon_2$, then end.
   Otherwise, let $m := m + 1$ and go to step 2.

is difficult, due to the coupling between the optimization variables in (4.9b). We also observe that the proposed algorithm requires less iterations till convergence for lower target SINR thresholds. Moreover, it can be seen from Table 5.2 that the amount

of computations required for the convergence of the inner loop is much smaller than the computation involved in solving the SDP problem in the outer loop. Therefore, the overall computational complexity of the proposed algorithm can be estimated as $\mathcal{O}(c(L+\tilde{N}_{\mathrm{r}}^2)^{3.5})$ with c between 3 and 5.

Before moving to the next section, we would like to comment on several issues related to the implementation of the proposed algorithm in practice.

Remark 1: The channel state information (CSI) on $\{\mathbf{H}_{k,l}\} \triangleq \{\mathbf{H}_{k,l}, k=1,\cdots,K, l=1,\cdots,L\}$ and $\{\mathbf{G}_{l,k}\} \triangleq \{\mathbf{G}_{l,k}, l=1,\cdots,L, k=1,\cdots,K\}$ is required in the proposed algorithm. Since the perfect CSI is not available in a real communication system due to limited feedback and/or inaccurate channel estimation, robust designs can be considered in case of imperfect CSI. A worst-case based robust relay matrices design for interference relay system has been proposed in [98] where each source and destination node has a single antenna (i.e., only $\mathbf{f}$ needs to be optimized). However, when all source and destination nodes have multiple antennas, the worst-case based robust design becomes extremely challenging since the worst-case SINR Γ_l is a very complicated function of $\mathbf{f}$, $\{\mathbf{b}_l\}$, $\{\mathbf{w}_l\}$, and $\mathbf{p}$. Alternatively, we can try the statistically robust design [104], where we average over the mismatch between the true and the estimated CSI. However, the statistical expectation of Γ_l in (5.5) with respect to all channel matrices turns out to be an extremely complicated expression of the design variables $\mathbf{f}$, $\mathbf{p}$, $\{\mathbf{b}_l\}$, and $\{\mathbf{w}_l\}$. This makes the statistically robust design problem every difficult to solve. The impact of imperfect CSI on the performance of the proposed algorithm will be studied through numerical simulation in Section 6.4.

Remark 2: The procedure in Table 5.2 needs to be carried out by a central processing unit due to the requirement of the global CSI. With the advancement of modern chip design, the amount of computation $\mathcal{O}(c(L+\tilde{N}_{\mathrm{r}}^2)^{3.5})$ can be handled by the central processing unit. Nevertheless, it is interesting to investigate distributed algorithms that can solve the problem (5.9). In fact, the inner loop in step (3) of Table 5.2 is easier than step (2) for a distributed implementation. The reason is that in step (2), an SDP problem needs to be solved, which is difficult to be implemented in a distributed manner.

Remark 3: In practical applications, to meet the SINR requirements (4.9b), some nodes may require larger transmission power that exceeds their available limit. A possible way out to this problem is to identify the SINR constraints that produce the largest increase in terms of transmit power first, and then relax those constraints in order to reduce the required power using a perturbation analysis [105]. Alternatively, one may

apply an admission control algorithm first to maximize the number of links possibly served, and then perform optimal power allocation [106].

5.4 Numerical Examples

In this section, we study the performance of the proposed joint power control and beamforming algorithm for an interference MIMO relay system through numerical simulations where all nodes are equipped with multiple antennas. For simplicity, we assume $\gamma_l = \gamma, N_{\mathrm{s},l} = N_{\mathrm{s}}$, $N_{\mathrm{d},l} = N_{\mathrm{d}}$, $l = 1, \cdots, L$, and $N_{\mathrm{r},k} = N_{\mathrm{r}}$, $k = 1, \cdots, K$, in all simulations. All noises are i.i.d. complex circularly symmetric Gaussian noise with zero mean and unit variance (i.e., $\sigma_{\mathrm{n}}^2 = 1$). The channel matrices have entries generated as i.i.d. complex Gaussian random variables with zero mean and variances σ_{h}^2 and σ_{g}^2 for $\{\mathbf{H}_{k,l}\}$ and $\{\mathbf{G}_{l,k}\}$, respectively. All simulation results are averaged over 500 independent channel realizations.

For the proposed algorithm, the procedure in Table 5.2 is carried out in each simulation to obtain the power vector $\mathbf{p}$, transmit beamforming vectors $\{\mathbf{b}_l\}$, relay beamforming vector $\mathbf{f}$, and receive beamforming vectors $\{\mathbf{w}_l\}$. To initialize the algorithm in Table 5.2, we randomly generate the transmit and receive beamforming vectors $\{\mathbf{b}_l\}$ and $\{\mathbf{w}_l\}$, respectively, along with arbitrary transmit power vector $\mathbf{p}$ and virtual power vector $\tilde{\mathbf{p}}$.

In the first example, we compare the performance of the proposed joint power control and beamforming algorithm (Proposed TxRxBF) with the relay-only beamforming without power control (RoBF-NPC) scheme studied in [97], [98] and the conventional SVD-based transmit beamforming approach (SVD-based TxBF). For the SVD-based TxBF scheme, we choose $\mathbf{b}_l$ as the principal right singular vector of $\tilde{\mathbf{H}}_l$. Then we update the transmit power vector $\mathbf{p}$, relay beamforming vector $\mathbf{f}$ and receive beamformers $\{\mathbf{w}_l\}$ based on the proposed structure. We plot the total power consumed by all source nodes and relay nodes versus the target SINR threshold γ (dB). Two channel fading environments are simulated: (i) Both $\{\mathbf{H}_{k,l}\}$ and $\{\mathbf{G}_{l,k}\}$ have Rayleigh fading; (ii) Only $\{\mathbf{H}_{k,l}\}$ has Rayleigh fading while $\{\mathbf{G}_{l,k}\}$ has Ricean fading with a Ricean factor of 5. Fig. 5.2 shows the performance of all three algorithms for $L = 2$, $K = 15$, $N_{\mathrm{s}} = N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $\sigma_{\mathrm{h}}^2 = 15$, and $\sigma_{\mathrm{g}}^2 = 10$. It can be seen from Fig. 5.2 that the proposed algorithm requires significantly less total power compared with the other two schemes in both Rayleigh and Ricean fading environments.

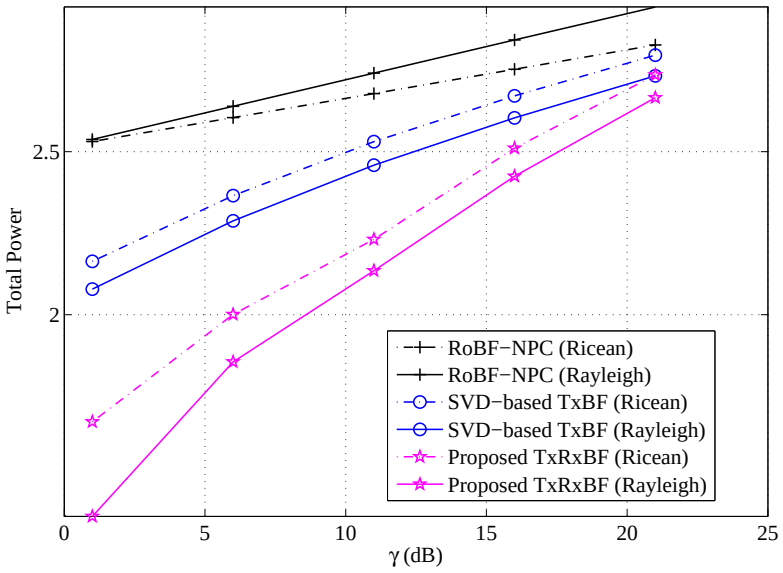

Figure 5.2: Total power versus target SINR. $L = 2$, $K = 15$, $N_{\rm s} = N_{\rm r} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$.

Note that the RoBF-NPC scheme performs better in Ricean fading channel whereas the performance of the other two approaches degrades under Ricean fading environment. This can be explained as follows. In the RoBF-NPC scheme, each transmitter and receiver has a single antenna as in [97] and [98], which indicates that the relay-destination channels $\{\mathbf{G}_{l,k}\}$ are in fact multiple-input single-output (MISO) channels. Therefore the LOS path component improves the system performance. For the other two schemes, the relay-destination channels are MIMO channels. In MIMO Ricean channels, the benefit of scattering environment reduces due to the LOS component. This weaker scattering component causes the performance degradation. Similar phenomenon has been observed in [107] for point-to-point MISO and MIMO Ricean channels.

In the second example, we vary the number of transmit antennas $N_{\rm s}$ to show the effect of transmit diversity with $L = 2$, $K = 8$, $N_{\rm r} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$. Fig. 5.3 indicates the significance of transmit beamforming in the proposed algorithm. It is obvious from Fig. 5.3 that with the increase in the spatial dimension of the transmit beamformers the performance of the proposed algorithm keeps improving.

In the next example, we study the performance of the proposed algorithm for different number of relays K with $L = 2$, $N_{\rm s} = N_{\rm r} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$. The total power required for $K = 10, 12$, and 15 versus γ (dB) is displayed in Fig. 5.4.

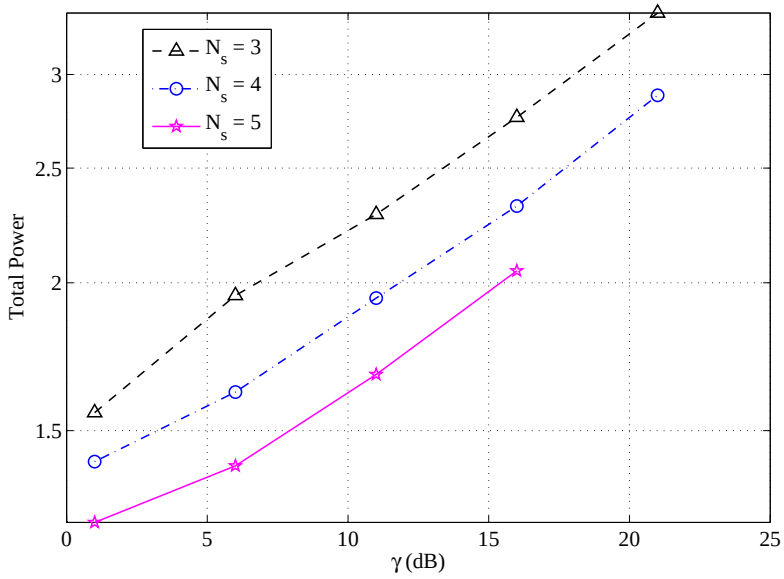

Figure 5.3: Total power versus target SINR for different number of transmit antennas. $L = 2$, $K = 8$, $N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $\sigma_{\mathrm{h}}^2 = 15$, and $\sigma_{\mathrm{g}}^2 = 10$.

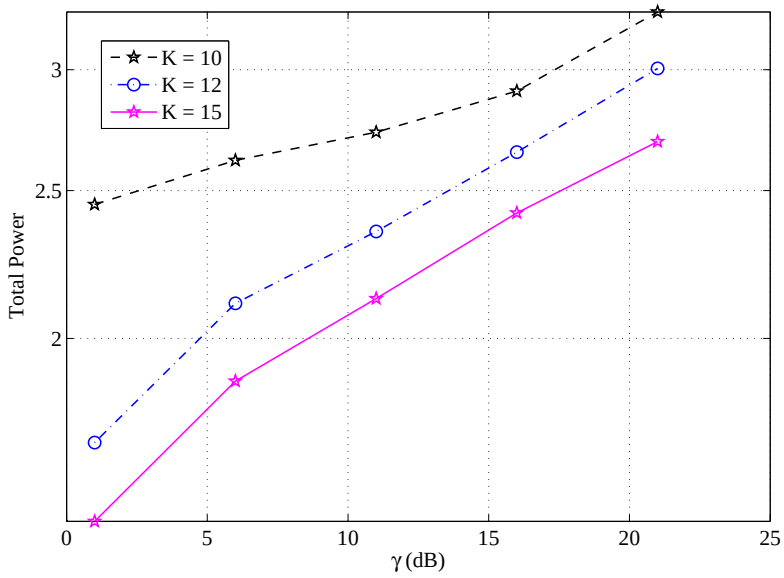

Figure 5.4: Total power versus target SINR for different number of relays. $L = 2$, $N_{\mathrm{s}} = N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $\sigma_{\mathrm{h}}^2 = 15$, and $\sigma_{\mathrm{g}}^2 = 10$.

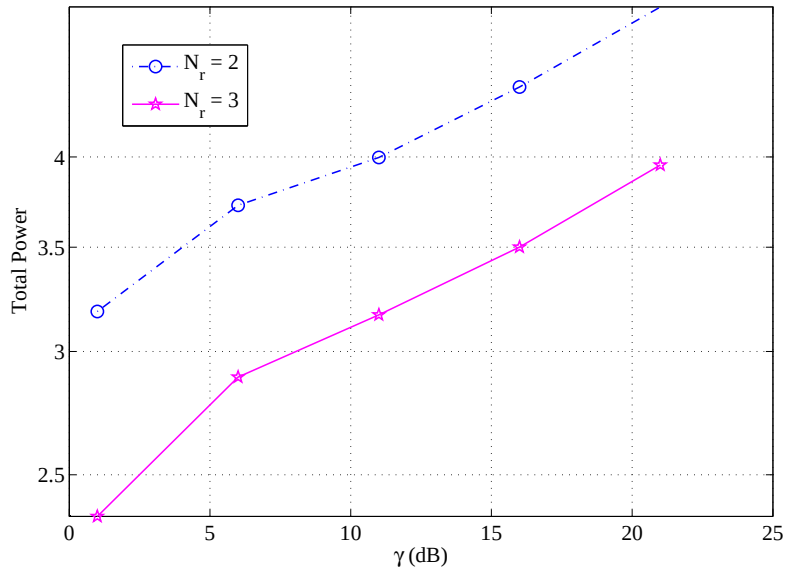

Figure 5.5: Total power versus target SINR for different number of relay antennas. $L = 3$, $K = 12$, $N_{\rm s} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$.

As expected, if we increase the number of relays the proposed algorithm requires less power since more relays provide more spatial diversity. We also show the impact of the number of relay antennas $N_{\rm r}$ in Fig. 5.5. This time, we set $L = 3$, $K = 12$, $N_{\rm s} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$ and the total power required for $N_r = 2$ and 3 versus γ (dB) is displayed. Note that with the increase in the number of relay antennas, the performance of the proposed scheme improves but at the same time, the computational complexity of solving the problem (5.32) significantly increases. Therefore, it is important to make a tradeoff between the performance and complexity based on the system requirements and the available resources.

In the next two examples, we study the impact of channel quality on the proposed algorithm. We assume that a larger variance of channel coefficients indicates a better channel. The impact of different $\sigma_{\rm h}^2$ and $\sigma_{\rm g}^2$ on the proposed algorithm is shown in Fig. 5.6 and Fig. 5.7, for $\sigma_{\rm g}^2 = 10$ and $\sigma_{\rm h}^2 = 10$, respectively. In these examples, we set $L = 2$, $K = 8$, $N_{\rm s} = N_{\rm d} = 4$, and $N_{\rm r} = 2$. A careful inspection of Figs. 5.6 and 5.7 reveals that the effect of channel variance of either hop is not homogeneous in general, but the results clearly demonstrate that the proposed algorithm performs better as the channel quality improves.

Next, we study the effect of channel interferences on the proposed algorithm. By

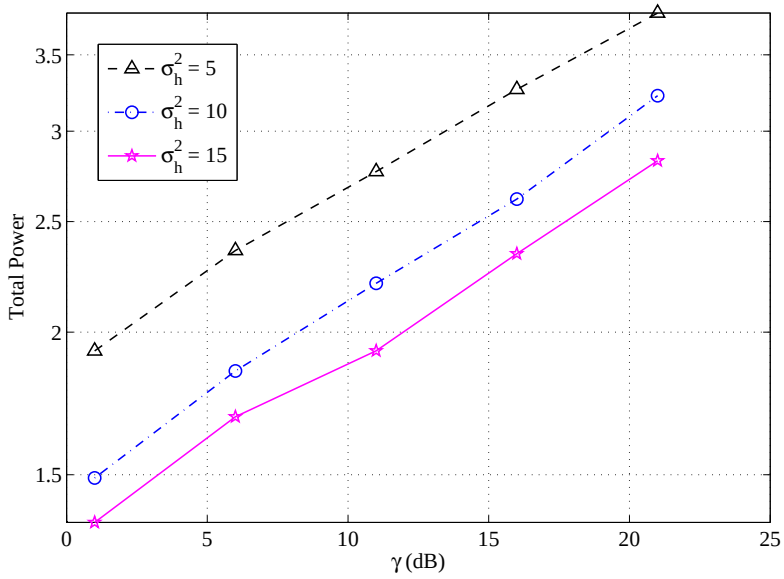

Figure 5.6: Effect of the first-hop channel quality. $L = 2$, $K = 8$, $N_{\mathrm{s}} = N_{\mathrm{d}} = 4$, $N_{\mathrm{r}} = 2$, and $\sigma_{\mathrm{g}}^2 = 10$.

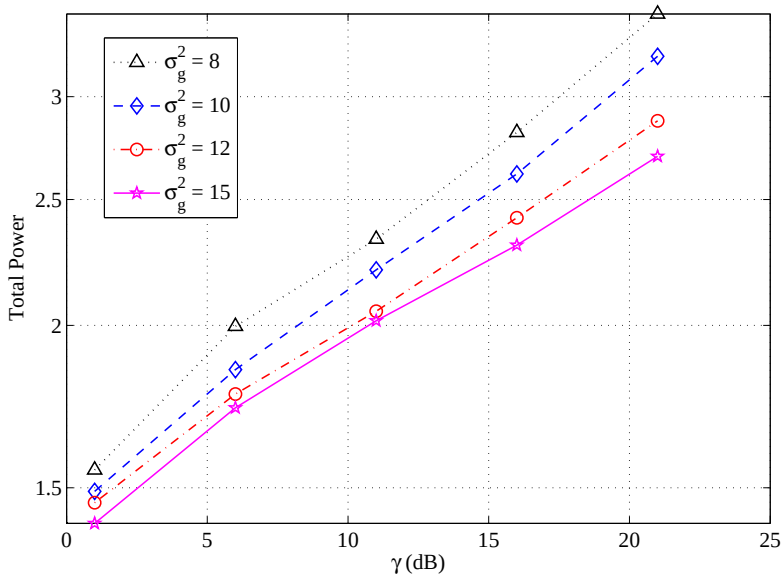

Figure 5.7: Effect of the second-hop channel quality. $L = 2$, $K = 8$, $N_{\mathrm{s}} = N_{\mathrm{d}} = 4$, $N_{\mathrm{r}} = 2$, and $\sigma_{\mathrm{h}}^2 = 10$.

increasing the number of source-destination pairs L, the interfering signal received at each destination node is also increased. The performance of the algorithm for different L is illustrated in Fig. 5.8 for $K = 12$, $N_{\rm s} = N_{\rm r} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$. From this figure it is clear that if there are more active users communicating simultaneously in the system, we need more power to achieve the same target SINR threshold γ.

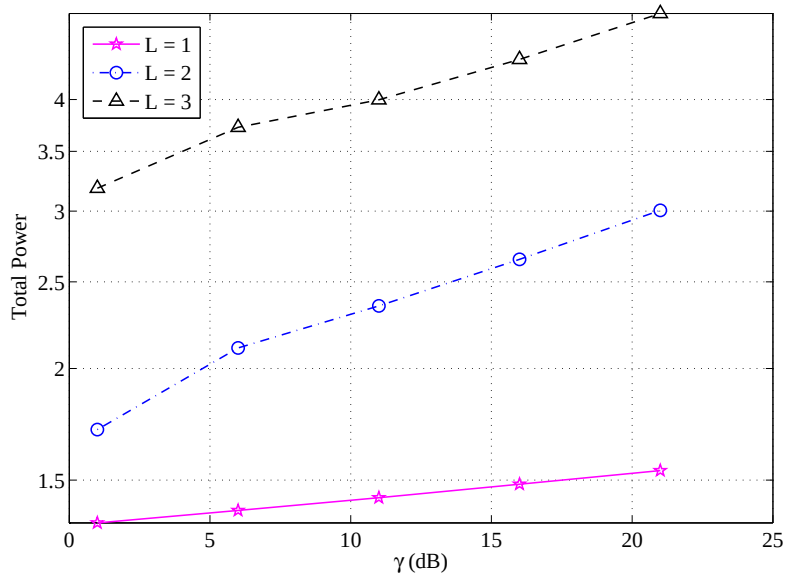

Figure 5.8: Total power versus target SINR for different number of users. $K = 12$, $N_{\rm s} = N_{\rm r} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 15$, and $\sigma_{\rm g}^2 = 10$.

In the last example, we study the impact of imperfect CSI on the performance of the proposed algorithm. The mismatch between the true CSI and the estimated CSI is modelled as complex Gaussian matrices with zero-mean and unit-variance entries. Fig. 5.9 shows the performance of all three algorithms for $L = 2$, $K = 12$, $N_{\rm s} = N_{\rm r} = 2$, $N_{\rm d} = 4$, $\sigma_{\rm h}^2 = 12$, and $\sigma_{\rm g}^2 = 10$. Clearly, the proposed algorithm outperforms the existing techniques with both perfect and imperfect CSI. Note that at very low (close to 0dB) target SINR, the RoBF scheme requires almost the same total power regardless the perfect CSI or imperfect CSI, because it does not involve any transmit/receive beamforming technique. The transmitters use their maximum available power budgets to transmit the signals. Thus, the total power varies only for relay beamforming resulting in little difference between total powers considering perfect and imperfect CSI at a low target SINR level.

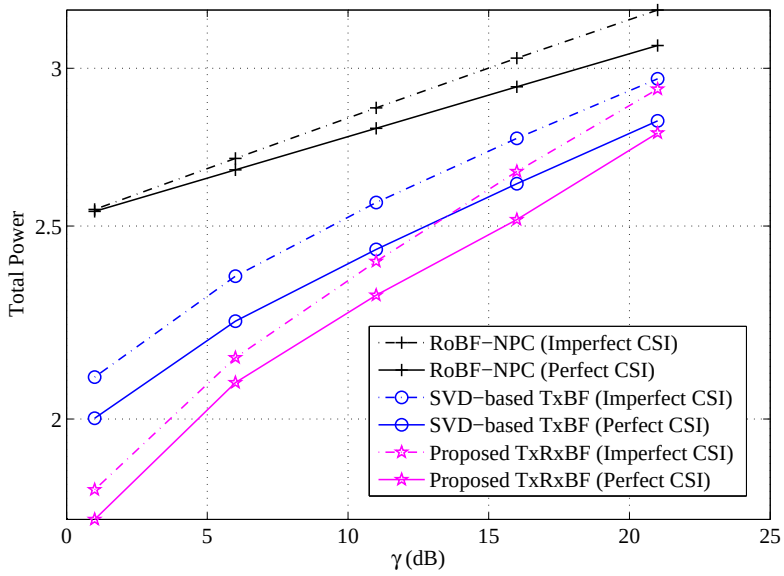

Figure 5.9: The impact of the CSI mismatch on the tested algorithms. $L = 2$, $K = 12$, $N_{\mathrm{s}} = N_{\mathrm{r}} = 2$, $N_{\mathrm{d}} = 4$, $\sigma_{\mathrm{h}}^2 = 12$, and $\sigma_{\mathrm{g}}^2 = 10$.

5.5 Chapter Summary

In this chapter, we considered a two-hop interference MIMO relay system with distributed relay nodes and developed an iterative technique to minimize the total transmit power consumed by all source and relay nodes such that a minimum SINR threshold is maintained at each receiver. The proposed algorithm exploits beamforming techniques at the source, relay, and destination nodes in conjunction with transmit power control. Simulation results demonstrate that the proposed power control and beamforming algorithm outperforms the existing techniques.

Chapter 6

Channel Estimation of Dual-Hop MIMO Relay System

In this chapter, we develop a novel channel estimation algorithm for two-hop MIMO relay systems using the parallel factor (PARAFAC) analysis. After a brief review of existing MIMO channel estimation techniques in Section 6.1, we introduce the PARAFAC data model of a two-hop AF-MIMO relay communication system in Section 6.2. The proposed channel estimation algorithm is developed in Section 6.3. The algorithm provides the destination node with full knowledge of all channel matrices involved in the communication. Compared with existing approaches, the proposed algorithm requires less number of training data blocks, yields smaller channel estimation error, and is applicable for both one-way and two-way MIMO relay systems with single or multiple relay nodes. In Section 6.4, we show some numerical examples to demonstrate the effectiveness of the PARAFAC-based channel estimation algorithm. Section 6.5 briefly summarizes the chapter.

6.1 Existing MIMO Channel Estimation Techniques

For the MIMO relaying algorithms developed in Chapters 2-5, the instantaneous CSI knowledge of both the source-relay link and the relay-destination link is required at the destination node to estimate the source signals. Moreover, in order to optimize the source and/or relay matrices, the instantaneous CSI knowledge of both links is needed to carry out the optimization procedure [20, 22, 25, 27, 35–37, 44, 59, 86]. When the direct

source-destination link is considered, the CSI knowledge of the direct link is also required at the destination node to estimate the source signals [108]. However, in practical relay communication systems, the instantaneous CSI is unknown, and therefore, has to be estimated. Recently, a tensor-based channel estimation algorithm is developed in [109] for a two-way MIMO relay system. Since the algorithm in [109] exploits the channel reciprocity in a two-way relay system, its application in one-way MIMO relay systems is not straightforward. In [110], a relay channel estimation algorithm using the least-squares (LS) fitting is proposed. The performance of the algorithm in [110] is further analyzed and improved by using the weighted least-squares (WLS) fitting in [111]. However, the number of training data blocks required in [110] and [111] is at least equal to the number of relay nodes (antennas), resulting in a low system spectral efficiency. For amplify-and-forward relay networks with single-antenna source, relay, and destination nodes, the optimal training sequence is developed in [112]. A superimposed training based channel estimation algorithm has been developed recently for OFDM modulated relay systems in [113]. The optimal training sequence is derived in [114] for a MIMO relay system with one multi-antenna relay node. However, for systems with distributed relay nodes which do not cooperate with each other, the result in [114] can not be used.

There are two major challenges in channel estimation for MIMO relay systems. Firstly, for most applications, the CSI on the compound source-relay-destination channel alone is not sufficient. In fact, the CSI of each hop is required at the destination node to perform signal retrieving and system optimization. Secondly, relay nodes (in particular, non-regenerative distributed relays) often have limited computation capacity. Thus, channel estimation is usually carried out at the destination node, not at the relay nodes [109–113]. In this chapter, we address these two challenges by proposing a novel MIMO relay channel estimation algorithm based on the PARAFAC analysis [115–117]. The proposed algorithm provides the destination node with full knowledge of all channel matrices involved in the communication. The contributions of this chapter can be summarized as follows. Firstly, compared with algorithms in [110] and [111] where the number of training data blocks should be at least equal to the number of relay nodes (antennas), the number of training data blocks required in the proposed algorithm can be less than the number of relay nodes (antennas). In particular, we show that when the number of relay nodes (antennas) is smaller than the number of antennas at the source node and the destination node, as few as two training data blocks are sufficient

to estimate all channels. Thus, the proposed algorithm has a higher spectral efficiency than those in [110] and [111]. Secondly, in this chapter, the initial estimation of channel matrices is improved by a linear MMSE (LMMSE) algorithm, which yields a smaller estimation error than the WLS fitting applied in [111]. Thirdly, in contrast to [109], the proposed algorithm is applicable for both one-way and two-way relay systems with single or multiple relay nodes.

In the proposed algorithm, the MIMO channel matrix of the direct source-destination link in one-way relay systems is estimated by the LS approach. For the source-relay-destination link in both one-way and two-way relay systems, we show that under a mild condition of the channel training data block length, the MIMO channel matrices of both hops can be estimated up to permutation and scaling ambiguities, which are inherent to the PARAFAC model. To remove the permutation ambiguity, we exploit the knowledge of the relay factors available at the destination node during the channel training period. Then by using a bilinear alternating least-squares (BALS) algorithm, the channel matrix of each hop can be estimated up to some scaling ambiguity, which can be resolved through normalization as in [110], [111], [116].

Since during the training period, the noise at the relay nodes is amplified and forwarded to the destination node, the effective noise vector at the destination node is non-white. Taking this fact into account, we propose an LMMSE approach to further improve the channel estimation, by exploiting the initial estimate of the relay-destination channel. We show that the proposed BALS and LMMSE algorithms can also be applied for channel estimation in two-way MIMO relay systems. Numerical examples demonstrate the effectiveness of the proposed PARAFAC-based channel estimation algorithm compared with existing techniques. We would like to mention that in this chapter, for notational convenience, we consider a narrowband single-carrier system. However, our algorithm can be straightforwardly applied to estimate the MIMO channel matrices in each subcarrier of a broadband multi-carrier relay communication system[1].

[1] In a multicarrier communication system, the spectral correlation among subcarriers can be exploited to reduce the computational complexity and improve the quality of channel estimation [118]. Exploiting such correlation in multicarrier MIMO relay channel estimation is an interesting future topic.

6.2 System Model

We consider a two-hop MIMO communication system where the source node transmits information to the destination node with the aid of R relay nodes as shown in Fig. 6.1. The source node and the destination node are equipped with $N_s \geq 2$ and $N_d \geq 2$ antennas, respectively, while the ith relay node has M_i antennas, $i = 1, \cdots, R$. Since several practical constraints such as power consumption, implementation costs and spatial efficiency make half-duplex relays more appealing for wireless applications than full-duplex relays, in this chapter, we consider half-duplex relays as in [109–114] (i.e., each relay node does not receive and transmit signals simultaneously). Thus, the communication process between the source and destination nodes is completed in two time slots. In the first time slot, the $N_s \times 1$ modulated signal vector $\mathbf{u}_s(t)$ is transmitted to all relay nodes and the destination node, and the received signal vectors are respectively given by

$$\begin{aligned} \mathbf{y}_{r,i}(t) &= \mathbf{H}_{sr,i}\mathbf{u}_s(t) + \mathbf{v}_{r,i}(t), \quad i = 1, \cdots, R \\ \mathbf{y}_d(t) &= \mathbf{H}_{sd}\mathbf{u}_s(t) + \mathbf{v}_d(t) \end{aligned} \tag{6.1}$$

where $\mathbf{y}_{r,i}(t)$ is an $M_i \times 1$ received signal vector at the ith relay node, $\mathbf{y}_d(t)$ is an $N_d \times 1$ received signal vector at the destination node, $\mathbf{H}_{sr,i}$ is the $M_i \times N_s$ MIMO fading channel matrix between the source node and the ith relay node, $\mathbf{H}_{sd}$ is the $N_d \times N_s$ MIMO source-destination channel matrix, $\mathbf{v}_{r,i}(t)$ is an $M_i \times 1$ noise vector at the ith relay node, and $\mathbf{v}_d(t)$ is the $N_d \times 1$ noise vector at the destination node. We assume that all noises are independent identically distributed (i.i.d.) complex Gaussian noise with zero mean and unit variance.

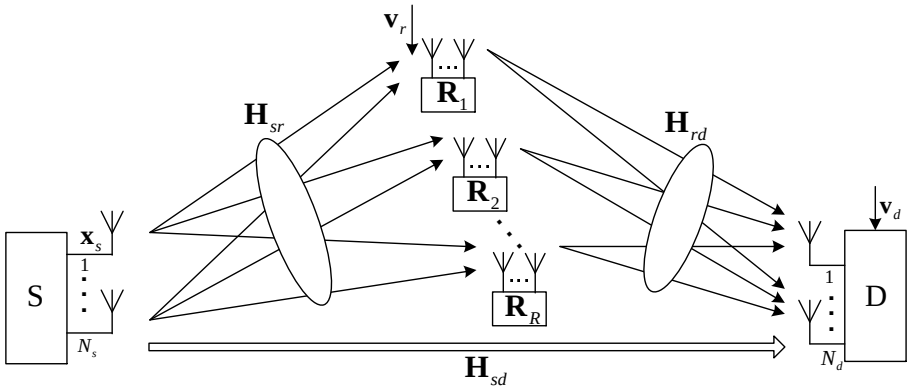

Figure 6.1: Two-Hop MIMO relay system with R relay nodes.

In the second time slot, the source node is silent, and each relay node amplifies the received signal vector with matrix $\mathbf{R}_i$ and forwards the amplified signals to the destination node. We assume that relay nodes are synchronized during transmission[2] as in [25], [110], and [111]. The received signal vector at the destination node is

$$\begin{aligned}\mathbf{y}_d(t+1) &= \sum_{i=1}^{R} \mathbf{H}_{rd,i}\mathbf{R}_i\mathbf{H}_{sr,i}\mathbf{u}_s(t) + \sum_{i=1}^{R} \mathbf{H}_{rd,i}\mathbf{R}_i\mathbf{v}_{r,i}(t) + \mathbf{v}_d(t+1) \\ &= \mathbf{H}_{rd}\mathbf{R}\mathbf{H}_{sr}\mathbf{u}_s(t) + \mathbf{H}_{rd}\mathbf{R}\mathbf{v}_r(t) + \mathbf{v}_d(t+1) \end{aligned} \tag{6.2}$$

where $\mathbf{H}_{rd,i}$ is the $N_d \times M_i$ MIMO fading channel matrix between the destination node and the ith relay node, and $\mathbf{v}_d(t+1)$ is an $N_d \times 1$ noise vector at the destination node at time $t+1$. Here $\mathbf{H}_{sr} \triangleq \left[\mathbf{H}_{sr,1}^T, \cdots, \mathbf{H}_{sr,R}^T\right]^T$ is the $M \times N_s$ ($M = \sum_{i=1}^{R} M_i$) MIMO channel from the source node to all relay nodes, $\mathbf{H}_{rd} \triangleq [\mathbf{H}_{rd,1}, \cdots, \mathbf{H}_{rd,R}]$ is the $N_d \times M$ channel matrix between all relay nodes and the destination node, $\mathbf{v}_r(t) \triangleq \left[\mathbf{v}_{r,1}^T(t), \cdots, \mathbf{v}_{r,R}^T(t)\right]^T$ is an $M \times 1$ vector stacking the noise at all relay nodes on top of each other, and $\mathbf{R} \triangleq \text{blkdiag}[\mathbf{R}_1, \cdots, \mathbf{R}_R]$ is an $M \times M$ block diagonal matrix containing all relay matrices. We assume that $\mathbf{H}_{sr}$, $\mathbf{H}_{rd}$, and $\mathbf{H}_{sd}$ have complex Gaussian entries with zero-mean and variances of $1/N_s$, $1/M$, $1/(8N_s)$, respectively[3]. Depending on the environment, the elements in each channel matrix can be independent or correlated [12]. We assume that the channel correlation knowledge is not available at the destination node and thus can not be exploited. All channels are quasi-static block fading which means they are constant over some time interval before changing to another realization. Combining (6.1) and (6.2), the received signals at the destination node over two time slots are given by

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{H}_{rd}\mathbf{R}\mathbf{H}_{sr} \\ \mathbf{H}_{sd} \end{bmatrix} \mathbf{u}_s(t) + \begin{bmatrix} \mathbf{H}_{rd}\mathbf{R}\mathbf{v}_r(t) + \mathbf{v}_d(t+1) \\ \mathbf{v}_d(t) \end{bmatrix}. \tag{6.3}$$

Due to its lower computational complexity, a linear receiver is used at the destination node to retrieve the transmitted signal vector $\mathbf{u}_s(t)$ [22, 25, 27, 120]. The estimated signal waveform vector is given by $\hat{\mathbf{u}}_s(t) = \mathbf{W}^H\mathbf{y}(t)$, where $\mathbf{W}$ is the $2N_d \times N_s$ weight

[2] If a blind synchronization technique is applied, relay synchronization and channel estimation can be jointly designed to improve the system performance [119]. While in pilot symbols-based synchronization methods, these pilot symbols can be exploited to assist channel estimation.

[3] The variances are set to normalize the effect of number of transmit antennas to the receive signal-to-noise ratio. The relay nodes are assumed to be of equal distance to the source and the destination nodes with a path loss factor of 3.

matrix. From (6.3), the MSE of the signal waveform estimation can be written as

$$e = \mathrm{tr}\big(\mathrm{E}\big[(\hat{\mathbf{u}}_s(t) - \mathbf{u}_s(t))(\hat{\mathbf{u}}_s(t) - \mathbf{u}_s(t))^H\big]\big). \tag{6.4}$$

Assuming that $\mathrm{E}\big[\mathbf{u}_s(t)\mathbf{u}_s(t)^H\big] = \mathbf{I}_{N_s}$, the receiver weight matrix which minimizes (6.4) is the Wiener filter given by [60]

$$\mathbf{W} = \left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \bar{\mathbf{C}}\right)^{-1}\bar{\mathbf{H}} \tag{6.5}$$

where

$$\bar{\mathbf{H}} \triangleq \begin{bmatrix}\mathbf{H}_{rd}\mathbf{R}\mathbf{H}_{sr}\\ \mathbf{H}_{sd}\end{bmatrix}, \bar{\mathbf{C}} \triangleq \begin{bmatrix}\mathbf{H}_{rd}\mathbf{R}\mathbf{R}^H\mathbf{H}_{rd}^H + \mathbf{I}_{N_d} & \mathbf{0}_{N_d\times N_d}\\ \mathbf{0}_{N_d\times N_d} & \mathbf{I}_{N_d}\end{bmatrix}. \tag{6.6}$$

Here $\mathbf{0}_{m\times n}$ denotes an $m\times n$ matrix with all zero entries. We assume that the destination node knows the relay amplifying matrix $\mathbf{R}$.

It can be clearly seen from (6.5) and (6.6) that in order to compute $\mathbf{W}$, the CSI knowledge of the compound channel $\bar{\mathbf{H}}$ alone is not sufficient. In fact, the CSI of $\mathbf{H}_{rd}$ is also needed at the destination node to obtain $\mathbf{W}$ in (6.5). Moreover, it has been shown in [108] that the CSI of $\mathbf{H}_{sr}$, $\mathbf{H}_{rd}$, and $\mathbf{H}_{sd}$ is required to optimize the source precoding matrix and the relay amplifying matrix.

It is shown in [114] that the CSI required above can be obtained through a two-stage training (TST) approach. At the first stage, $\mathbf{H}_{rd}$ is estimated by transmitting an $M \times L_1$ training sequence $\mathbf{S}_1$ from all R relay nodes to the destination node, where L_1 ($L_1 \geq M$) is the length of the training sequence. The received signal matrix at the destination node is given by $\mathbf{Y}_d = \mathbf{H}_{rd}\mathbf{S}_1 + \mathbf{V}_d(1)$, where $\mathbf{V}_d(1)$ is the noise matrix at the destination node. According to [121], the optimal $\mathbf{S}_1$ minimizing the MSE of channel estimation is orthogonal, i.e., $\mathbf{S}_1\mathbf{S}_1^H = \mathbf{I}_M$. Such $\mathbf{S}_1$ can be constructed, for example, from the normalized discrete Fourier transform (DFT) matrix [121]. The estimation of $\mathbf{H}_{rd}$ is given by

$$\hat{\mathbf{H}}_{rd} = \mathbf{Y}_d\mathbf{S}_1^H. \tag{6.7}$$

At the second stage, the source node transmits an $N_s \times L_2$ ($L_2 \geq N_s$) orthogonal training sequence $\mathbf{S}_2$ ($\mathbf{S}_2\mathbf{S}_2^H = \mathbf{I}_{N_s}$) to all relay nodes which then forward it to the destination node. From (6.3), the received signal matrix at the destination node is

$$\mathbf{Y} = \begin{bmatrix}\mathbf{H}_{rd}\mathbf{R}\mathbf{H}_{sr}\\ \mathbf{H}_{sd}\end{bmatrix}\mathbf{S}_2 + \begin{bmatrix}\mathbf{H}_{rd}\mathbf{R}\mathbf{V}_r + \mathbf{V}_d(2)\\ \mathbf{V}_d(3)\end{bmatrix} \tag{6.8}$$

where $\mathbf{V}_r$ is the noise matrix at the relay nodes, $\mathbf{V}_d(2)$ and $\mathbf{V}_d(3)$ are the noise matrices at the destination node. The estimation of the compound channel $\bar{\mathbf{H}}$ is obtained from

(6.8) as $\hat{\bar{\mathbf{H}}} = \mathbf{Y}\mathbf{S}_2^H$. Then an estimation of $\mathbf{H}_{sr}$ can be obtained from $\hat{\bar{\mathbf{H}}}$ as

$$\hat{\mathbf{H}}_{sr} = (\hat{\mathbf{H}}_{rd}\mathbf{R})^{\dagger}\hat{\bar{\mathbf{H}}}^{(1)} \tag{6.9}$$

where $\hat{\bar{\mathbf{H}}}^{(1)}$ contains the first N_d rows of $\hat{\bar{\mathbf{H}}}$. It can be seen from (6.9) that the error in estimating $\mathbf{H}_{sr}$ can be very large since $\hat{\mathbf{H}}_{sr}$ depends on $\hat{\mathbf{H}}_{rd}$, which is also an estimated matrix. To overcome this difficulty, in the following, we develop a PARAFAC analysis based algorithm to directly estimate all channel matrices $(\mathbf{H}_{sr}, \mathbf{H}_{rd}, \mathbf{H}_{sd})$ at the destination node.

6.3 Proposed Channel Estimation Algorithm

In order to estimate the channel matrices, training sequences are transmitted from the source node. The overall channel training period is divided into K time blocks (the minimal K required will be determined later). In each time block, the same $N_s \times L$ $(L \geq N_s)$ orthogonal channel training sequence $\mathbf{S}$ with $\mathbf{S}\mathbf{S}^H = \mathbf{I}_{N_s}$ is transmitted by the source node. In the kth time block, the ith relay node amplifies the received signal vector with a *diagonal* matrix $\mathbf{E}_{k,i}$ and forwards the amplified signal to the destination node[4]. Thus, the overall amplifying matrix from all relay nodes is $\mathbf{E}_k = \text{blkdiag}[\mathbf{E}_{k,1}, \cdots, \mathbf{E}_{k,R}]$, which is in fact a diagonal matrix. From (6.3), the received signal matrices at the destination node over K time blocks are given by

$$\mathbf{Y}_k \triangleq \begin{bmatrix} \mathbf{Y}_k^{(1)} \\ \mathbf{Y}_k^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{sr} \\ \mathbf{H}_{sd} \end{bmatrix} \mathbf{S} + \begin{bmatrix} \mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k} + \mathbf{V}_{d,k}^{(1)} \\ \mathbf{V}_{d,k}^{(2)} \end{bmatrix}, \quad k = 1, \cdots, K$$

where $\mathcal{D}_k\{\mathbf{F}\} \triangleq \mathbf{E}_k$, $\mathbf{F}$ is a $K \times M$ matrix whose kth row contains the amplifying factors of all M relay antennas at the kth time block, $\mathcal{D}_k\{\cdot\}$ is the operator that makes a diagonal matrix by selecting the kth row and putting it on the main diagonal while putting zeros elsewhere, $\mathbf{V}_{r,k}$ is the $M \times L$ noise matrix at the relay nodes during the kth time block, $\mathbf{V}_{d,k}^{(1)}$ and $\mathbf{V}_{d,k}^{(2)}$ are $N_d \times L$ noise matrices at the destination node during the kth time block, and $\mathbf{Y}_k^{(1)}$ and $\mathbf{Y}_k^{(2)}$ are matrices containing the first and the last N_d rows of $\mathbf{Y}_k$, respectively.

At the destination node, by multiplying both sides of (6.10) with $\mathbf{S}^H$, we obtain

$$\mathbf{Y}_k\mathbf{S}^H = \begin{bmatrix} \mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{sr} \\ \mathbf{H}_{sd} \end{bmatrix} + \begin{bmatrix} \mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k}\mathbf{S}^H + \mathbf{V}_{d,k}^{(1)}\mathbf{S}^H \\ \mathbf{V}_{d,k}^{(2)}\mathbf{S}^H \end{bmatrix}, \quad k = 1, \cdots, K. \tag{6.10}$$

[4]Diagonal relay amplifying matrix is only used for the purpose of channel estimation. During the normal communication period, however, the relay amplifying matrix does not need to be diagonal.

From (6.10), an LS estimate of $\mathbf{H}_{sd}$ is given by

$$\hat{\mathbf{H}}_{sd} = \frac{1}{K}\mathbf{Y}^{(2)}(\mathbf{1}_K \otimes \mathbf{S})^H = \mathbf{H}_{sd} + \frac{1}{K}\mathbf{V}_d^{(2)}(\mathbf{1}_K \otimes \mathbf{S})^H$$

where $\mathbf{Y}^{(2)} \triangleq \left[\mathbf{Y}_1^{(2)}, \mathbf{Y}_2^{(2)}, \cdots, \mathbf{Y}_K^{(2)}\right]$, $\mathbf{V}_d^{(2)} \triangleq \left[\mathbf{V}_{d,1}^{(2)}, \mathbf{V}_{d,2}^{(2)}, \cdots, \mathbf{V}_{d,K}^{(2)}\right]$, $\mathbf{1}_K$ denotes a $1\times K$ vector with all 1 elements, and $\otimes$ stands for the Kronecker matrix product [62]. In the following, we show how to estimate $\mathbf{H}_{rd}$ and $\mathbf{H}_{sr}$ at the destination node.

6.3.1 PARAFAC model and identifiability of channel matrices

Let us introduce

$$\tilde{\mathbf{X}}_k \triangleq \mathbf{Y}_k^{(1)}\mathbf{S}^H = \mathbf{X}_k + \mathbf{V}_k, \qquad k = 1, \cdots, K \tag{6.11}$$

$$\mathbf{X}_k \triangleq \mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{sr}, \qquad k = 1, \cdots, K \tag{6.12}$$

$$\mathbf{V}_k \triangleq \mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k}\mathbf{S}^H + \mathbf{V}_{d,k}^{(1)}\mathbf{S}^H, \; k = 1, \cdots, K \tag{6.13}$$

where $\mathbf{X}_k$ is the matrix-of-interest containing both $\mathbf{H}_{rd}$ and $\mathbf{H}_{sr}$, $\mathbf{V}_k$ is the effective noise matrix, and $\tilde{\mathbf{X}}_k$ is a noisy observation of $\mathbf{X}_k$. We would like to mention that $\mathbf{F}$ is chosen beforehand and is known at the destination node. The optimal $\mathbf{F}$ is very difficult to obtain for the PARAFAC-based channel estimation algorithm. Nevertheless, an intuitive way of designing $\mathbf{F}$ will be discussed later. By assembling the set of K matrices in (6.12) together along the direction of the index k (the third dimension), we obtain an $N_d \times N_s \times K$ three-way array $\underline{\mathbf{X}}$, whose (i,j,k)-th element is given by

$$x(i,j,k) = \sum_{m=1}^{M} h_{rd}(i,m)f(k,m)h_{sr}(m,j) \tag{6.14}$$

for all $i = 1, \cdots, N_d$, $j = 1, \cdots, N_s$, and $k = 1, \cdots, K$. Here $h_{rd}(i,m)$, $f(k,m)$, and $h_{sr}(m,j)$ stand for the (i,m)-th, (k,m)-th, and (m,j)-th elements of $\mathbf{H}_{rd}$, $\mathbf{F}$, and $\mathbf{H}_{sr}$, respectively. Equation (6.14) expresses $x(i,j,k)$ as a sum of M rank-1 triple products, which is known as the trilinear decomposition, or PARAFAC[5] analysis of $x(i,j,k)$ [115–117]. Correspondingly, assembling K matrices of $\tilde{\mathbf{X}}_k$ in (6.11) along the index k leads

[5]PARAFAC is a multi-way method originating from psychometrics [115] and has recently found applications in array signal processing [116] and communications [117]. Generalizing the concept of low-rank decomposition to higher way arrays or tensors, PARAFAC is instrumental in the analysis of data arrays indexed by three or more independent variables, just like singular value decomposition (SVD) is instrumental in ordinary matrix (two-way array) analysis. Unlike SVD, PARAFAC does not impose orthogonality constraints. The reason is that in contrast to low-rank matrix decomposition, low-rank decomposition of higher order tensorial data is essentially unique under certain conditions.

to a noise-contaminated $\underline{\mathbf{X}}$ given by $\underline{\tilde{\mathbf{X}}} = \underline{\mathbf{X}} + \underline{\mathbf{V}}$, where $\underline{\mathbf{V}}$ is obtained by assembling K noise matrices in (6.13).

Let us denote the Kruskal rank (or k-rank) [122] of a matrix $\mathbf{A}$ as $k_{\mathbf{A}}$, which is the maximum integer k, such that *any* k columns drawn from $\mathbf{A}$ are linearly independent. Note that Kruskal rank is always less than or equal to the conventional matrix rank. It can be easily checked that if $\mathbf{A}$ is full column rank, then it is also full Kruskal rank. It can be shown by using the identifiability theorem of the PARAFAC model in [116] and [122] that if

$$k_{\mathbf{H}_{rd}} + k_{\mathbf{F}} + k_{\mathbf{H}_{sr}} \geq 2M + 2 \tag{6.15}$$

then the triple $(\mathbf{H}_{rd}, \mathbf{F}, \mathbf{H}_{sr})$ is unique up to permutation and scaling ambiguities, i.e., if there exists any other triple $(\bar{\mathbf{H}}_{rd}, \bar{\mathbf{F}}, \bar{\mathbf{H}}_{sr})$ that gives rise to (6.12), then it is related to $(\mathbf{H}_{rd}, \mathbf{F}, \mathbf{H}_{sr})$ via

$$\bar{\mathbf{H}}_{rd} = \mathbf{H}_{rd}\mathbf{\Pi}\mathbf{\Delta}_1, \quad \bar{\mathbf{F}} = \mathbf{F}\mathbf{\Pi}\mathbf{\Delta}_2, \quad \bar{\mathbf{H}}_{sr}^T = \mathbf{H}_{sr}^T\mathbf{\Pi}\mathbf{\Delta}_3 \tag{6.16}$$

where $\mathbf{\Pi}$ is an $M \times M$ permutation matrix, and $\mathbf{\Delta}_i$, $i = 1, 2, 3$, are $M \times M$ diagonal (complex) scaling matrices satisfying

$$\mathbf{\Delta}_1\mathbf{\Delta}_2\mathbf{\Delta}_3 = \mathbf{I}_M. \tag{6.17}$$

Inequality (6.15) establishes the sufficient condition for the identifiability of $(\mathbf{H}_{rd}, \mathbf{F}, \mathbf{H}_{sr})$. Since $\mathbf{F}$ is chosen beforehand (e.g., based on the DFT matrix as shown later), one can guarantee that $\mathbf{F}$ has full k-rank. Moreover, both $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$ are random matrices, and hence have full k-rank. Therefore, in such case, condition (6.15) becomes

$$\min(N_d, M) + \min(K, M) + \min(N_s, M) \geq 2M + 2. \tag{6.18}$$

From (6.18), the identifiability condition can be summarized in the following theorem.

Theorem 6.1 *The PARAFAC model (6.14) is identifiable only if $N_s \geq 2$, $N_d \geq 2$, and $2 \leq M \leq N_s + N_d - 2$. Moreover, for different N_s, N_d, and M, the lower bound of K satisfying (6.18) is given by*

$$K \geq \begin{cases} 2M + 2 - N_s - N_d & M \geq N_s, N_d \\ M + 2 - N_d & N_d \leq M \leq N_s \\ M + 2 - N_s & N_s \leq M \leq N_d \\ 2 & M \leq N_s, N_d \end{cases}. \tag{6.19}$$

For all four cases in (6.19), the lower bound of K is no greater than M.

PROOF: The proof can be done by expanding the three $\min(\cdot)$ operators in (6.18).

- If $M \geq N_s, N_d \longrightarrow \min(K, M) \geq 2M + 2 - N_d - N_s \longrightarrow K \geq 2M + 2 - N_d - N_s$, and $M \geq 2M+2-N_d-N_s \longrightarrow M \leq N_s+N_d-2$. This together with $M \geq N_s, N_d$, we have $N_s, N_d \geq 2$ and $2 \leq M \leq N_s + N_d - 2$. Since $M \leq N_s + N_d - 2$, it holds that $2M + 2 - N_s - N_d \leq M$;

- If $N_d \leq M \leq N_s \longrightarrow \min(K, M) \geq M + 2 - N_d \longrightarrow K \geq M + 2 - N_d$, and $M \geq M + 2 - N_d \longrightarrow N_d \geq 2$. This together with $N_d \leq M \leq N_s$, we have $N_s, N_d \geq 2$ and $2 \leq M \leq N_s + N_d - 2$. Since $N_d \geq 2$, there is $M + 2 - N_d \leq M$;

- If $N_s \leq M \leq N_d \longrightarrow \min(K, M) \geq M + 2 - N_s \longrightarrow K \geq M + 2 - N_s$, and $M \geq M + 2 - N_s \longrightarrow N_s \geq 2$. This together with $N_s \leq M \leq N_d$, we have $N_s, N_d \geq 2$ and $2 \leq M \leq N_s+N_d-2$. Since $N_s \geq 2$, it holds that $M+2-N_s \leq M$;

- If $M \leq N_s, N_d \longrightarrow \min(K, M) \geq 2 \longrightarrow K \geq 2$, and $M \geq 2$. This together with $M \leq N_s, N_d$, we have $N_s, N_d \geq 2$ and $2 \leq M \leq N_s + N_d - 2$.

Summarizing the four cases above, we obtain the necessary conditions for identifiability in the PARAFAC model (6.14) as $N_s \geq 2$, $N_d \geq 2$, and $2 \leq M \leq N_s + N_d - 2$. The lower bound of K in each case, which is less than or equal to M, is also given above. □

Interestingly, it is shown in Theorem 6.1 that under the mild condition of $N_s, N_d \geq 2$ and $2 \leq M \leq N_s + N_d - 2$, the minimal K required in the proposed PARAFAC-based channel estimation algorithm can be less than M. While in [110] and [111], at least $K = M$ training data blocks are required to perform the channel estimation. Therefore, the proposed algorithm has a higher spectral efficiency than those in [110] and [111]. Moreover, Theorem 6.1 shows that if $N_d \geq M$ and $N_s \geq M$, then two training data blocks ($K = 2$) are sufficient to estimate both $\mathbf{H}_{rd}$ and $\mathbf{H}_{sr}$ at the destination node. We also observe that if (6.19) is satisfied, then it holds that $KN_d > M$ and $KN_s > M$. We would like to mention that since $N_s \geq 2$ and $N_d \geq 2$ are required, which implies that $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$ need to be matrices, the PARAFAC-based MIMO relay channel estimation algorithm can not be straightforwardly applied to relay systems with $N_s = 1$ and/or $N_d = 1$.

6.3.2 Bilinear alternating least-squares (BALS) fitting

In this subsection, we develop a BALS algorithm to estimate $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$ by carrying out the PARAFAC model fitting with known $\mathbf{F}$. First we show some rearrangements of three-way arrays $\underline{\mathbf{X}}$, $\underline{\mathbf{V}}$, and $\underline{\tilde{\mathbf{X}}}$ which will be used later.

By stacking K matrices of $\mathbf{X}_k$ in (6.12) on top of each other, we obtain

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_K \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{rd}\mathcal{D}_1\{\mathbf{F}\} \\ \vdots \\ \mathbf{H}_{rd}\mathcal{D}_K\{\mathbf{F}\} \end{bmatrix} \mathbf{H}_{sr} = (\mathbf{F} \odot \mathbf{H}_{rd})\mathbf{H}_{sr} \tag{6.20}$$

where $\odot$ stands for the Khatri-Rao (column-wise Kronecker) matrix product [62]. Correspondingly, stacking matrices $\tilde{\mathbf{X}}_k$ in (6.11) on top of each other gives rise to

$$\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_K \end{bmatrix} + \begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_K \end{bmatrix} = \mathbf{X} + \mathbf{V}. \tag{6.21}$$

By slicing $\underline{\mathbf{X}}$ perpendicular to the dimension of j, we obtain a set of N_s matrices $\mathbf{Z}_j = \mathbf{F}\mathcal{D}_j\{\mathbf{H}_{sr}^T\}\mathbf{H}_{rd}^T$, $j = 1, \cdots, N_s$. By stacking N_s matrices of $\mathbf{Z}_j$ on top of each other, we have

$$\mathbf{Z} \triangleq \begin{bmatrix} \mathbf{Z}_1 \\ \vdots \\ \mathbf{Z}_{N_s} \end{bmatrix} = \begin{bmatrix} \mathbf{F}\mathcal{D}_1\{\mathbf{H}_{sr}^T\} \\ \vdots \\ \mathbf{F}\mathcal{D}_{N_s}\{\mathbf{H}_{sr}^T\} \end{bmatrix} \mathbf{H}_{rd}^T = (\mathbf{H}_{sr}^T \odot \mathbf{F})\mathbf{H}_{rd}^T. \tag{6.22}$$

Similarly, by slicing $\underline{\tilde{\mathbf{X}}}$ perpendicular to the dimension of j and stacking the resulting matrices on top of each other, we have

$$\tilde{\mathbf{Z}} = \begin{bmatrix} \mathbf{Z}_1 \\ \vdots \\ \mathbf{Z}_{N_s} \end{bmatrix} + \begin{bmatrix} \mathbf{N}_1 \\ \vdots \\ \mathbf{N}_{N_s} \end{bmatrix} \tag{6.23}$$

where $\mathbf{N}_j$, $j = 1, \cdots, N_s$, are the slabs of $\underline{\mathbf{V}}$ along the dimension of j.

The BALS fitting starts at a random $\hat{\mathbf{H}}_{rd}$. In each iteration, we first update $\mathbf{H}_{sr}$ using the LS fitting with fixed $\mathbf{F}$ and $\hat{\mathbf{H}}_{rd}$. Using (6.20) and (6.21), we obtain an updated $\mathbf{H}_{sr}$ as

$$\hat{\mathbf{H}}_{sr} = \arg\min_{\mathbf{H}_{sr}} \left\| \tilde{\mathbf{X}} - (\mathbf{F} \odot \hat{\mathbf{H}}_{rd})\mathbf{H}_{sr} \right\| = (\mathbf{F} \odot \hat{\mathbf{H}}_{rd})^{\dagger}\, \tilde{\mathbf{X}} \tag{6.24}$$

where $\|\cdot\|$ denotes the matrix Frobenius norm. Then we update $\mathbf{H}_{rd}$ through the LS fitting with known $\mathbf{F}$ and $\hat{\mathbf{H}}_{sr}$, and obtain $\hat{\mathbf{H}}_{rd}$ using (6.22) and (6.23) as

$$\hat{\mathbf{H}}_{rd} = \arg\min_{\mathbf{H}_{rd}} \left\| \tilde{\mathbf{Z}} - (\hat{\mathbf{H}}_{sr}^T \odot \mathbf{F})\mathbf{H}_{rd}^T \right\| = \left[(\hat{\mathbf{H}}_{sr}^T \odot \mathbf{F})^{\dagger}\, \tilde{\mathbf{Z}} \right]^T. \tag{6.25}$$

Since the conditional update of matrices in (6.24) and (6.25) may either improve or maintain but can not worsen the current LS fit, a monotonic convergence of the BALS procedure to (at least) a locally optimal solution follows directly from this observation [116]. The procedure of the BALS fitting is listed in Table 6.1, where ε is a positive constant close to 0, and the matrix with superscript (n) denotes the estimated matrix at the nth iteration. Theoretically, for some particular data sets, the convergence of the BALS algorithm can be extremely slow. However, since both $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$ are random matrices, the probability that both matrices fall in such data sets is very small. For large values of N_s, M, and N_d, such probability is almost zero. It will be shown in Section 6.4 that the BALS algorithm typically converges in only a few iterations.

Table 6.1: Procedure of the BALS fitting

1. Initialize the algorithm with a given $\mathbf{F}$ and a random $\mathbf{H}_{rd}^{(0)}$; Set $\delta(0) = \infty$ and $n = 1$.
2. Update $\mathbf{H}_{sr}^{(n)}$ as (6.24) using $\mathbf{H}_{rd}^{(n-1)}$; Update $\mathbf{H}_{rd}^{(n)}$ as (6.25) using $\mathbf{H}_{sr}^{(n)}$; Calculate $\delta(n) = \|\tilde{\mathbf{X}} - (\mathbf{F} \odot \mathbf{H}_{rd}^{(n)})\mathbf{H}_{sr}^{(n)}\|$.
3. If $\big[\delta(n-1) - \delta(n)\big]/\delta(n) \leq \varepsilon$, then end.
   Otherwise, let $n := n + 1$ and go to step 2).

Since $\mathbf{F}$ is known, the BALS algorithm delivers an estimation of $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$ with only a scaling ambiguity $\mathbf{\Delta}_1$ at the convergence point, i.e., $\mathbf{\Pi} = \mathbf{I}_M$, $\mathbf{\Delta}_2 = \mathbf{I}_M$ in (6.16), and $\mathbf{\Delta}_3 = \mathbf{\Delta}_1^{-1}$ according to (6.17). This scaling ambiguity also exists in [110] and [111], and can be resolved through normalization as in [110], [111], [116].

The major computation task in the proposed BALS algorithm lies in the LS fittings in (6.24) and (6.25). Thus the per-iteration complexity of the BALS algorithm can be estimated as $\mathcal{O}(MKN_dN_s + M^3)$. The overall complexity of the BALS algorithm depends on the number of iterations and will be further discussed in Section 6.4. Note that the computational complexity of the LS-based algorithm in [110] can be estimated as $\mathcal{O}(MKN_dN_s + M^3 + MN_d^2N_s)$, where the three terms are from matrix multiplications, matrix inversion, and M matrix SVDs, respectively.

Now we present an intuitive choice of $\mathbf{F}$. By slicing $\underline{\tilde{\mathbf{X}}}$ perpendicular to the dimension of i and stacking the resulting matrices on top of each other, we have

$$\mathbf{P} = (\mathbf{H}_{rd} \odot \mathbf{H}_{sr}^T)\mathbf{F}^T + \mathbf{M}$$

where $\mathbf{M}$ is the corresponding noise matrix obtained by slicing $\underline{\mathbf{V}}$ perpendicular to the dimension of i and stacking the resulting matrices on top of each other. Let us denote $\mathbf{H}_{srd} \triangleq \mathbf{H}_{rd} \odot \mathbf{H}_{sr}^T$. Since $\mathrm{E}\big[\mathbf{H}_{srd}\mathbf{H}_{srd}^H\big] = \mathbf{I}_{N_s N_d}$, if $K \geq M$ and $\mathbf{M}$ have i.i.d. entries, the optimal $\mathbf{F}$ minimizing the MSE of a linear estimation of $\mathbf{H}_{srd}$ is unitary ($\mathbf{F}^H\mathbf{F} = \mathbf{I}_M$). However, it can be shown that the elements in $\mathbf{M}$ are correlated and the covariance matrix of $\mathbf{M}$ is a complicated function of $\mathbf{F}$. Thus, strictly speaking, a unitary $\mathbf{F}$ is not optimal in general. Nevertheless, such $\mathbf{F}$ is still a good choice especially when the signal-to-noise ratio is medium to high at channel training stage. In numerical simulations, we also find that the DFT matrix (which satisfies $\mathbf{F}^H\mathbf{F} = \mathbf{I}_M$) is a good choice for $\mathbf{F}$.

6.3.3 Linear minimal mean-squared error (LMMSE) estimation

It can be seen from (6.13) that the covariance matrix of the effective noise $\mathbf{V}_k$ at the destination node is given by $\mathbf{C}_k \triangleq \mathrm{E}[\mathbf{V}_k\mathbf{V}_k^H] = N_s\big(\mathbf{H}_{rd}\mathcal{D}_k\{\mathbf{F}\}(\mathcal{D}_k\{\mathbf{F}\})^H\mathbf{H}_{rd}^H + \mathbf{I}_{N_d}\big)$, $k = 1, \cdots, K$. Obviously, $\mathbf{V}_k$ is non-white due to the channel $\mathbf{H}_{rd}$. Therefore, after an initial estimation of $\mathbf{H}_{rd}$ by the BALS algorithm in Section 6.3.2, an improved estimation of $\mathbf{H}_{sr}$ can be obtained by the LMMSE approach as $\breve{\mathbf{H}}_{sr} = \mathbf{T}_{sr}^H\tilde{\mathbf{X}}$, where $\mathbf{T}_{sr}$ is the $KN_d \times M$ weight matrix. The MSE of channel estimation can be written as

$$\mathrm{E}\Big[\mathrm{tr}\big((\breve{\mathbf{H}}_{sr} - \mathbf{H}_{sr})(\breve{\mathbf{H}}_{sr} - \mathbf{H}_{sr})^H\big)\Big] = \mathrm{tr}\Big(\big(\mathbf{T}_{sr}^H(\mathbf{F} \odot \hat{\mathbf{H}}_{rd}) - \mathbf{I}_M\big)\big(\mathbf{T}_{sr}^H(\mathbf{F} \odot \hat{\mathbf{H}}_{rd}) - \mathbf{I}_M\big)^H + \mathbf{T}_{sr}^H\hat{\mathbf{C}}\mathbf{T}_{sr}\Big) \tag{6.26}$$

where $\hat{\mathbf{C}} = \mathrm{blkdiag}\big[\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2, \cdots, \hat{\mathbf{C}}_K\big]$, and $\hat{\mathbf{C}}_k = N_s\big(\hat{\mathbf{H}}_{rd}\mathcal{D}_k\{\mathbf{F}\}(\mathcal{D}_k\{\mathbf{F}\})^H\hat{\mathbf{H}}_{rd}^H + \mathbf{I}_{N_d}\big)$, $k = 1, \cdots, K$, is an estimate of $\mathbf{C}_k$ using $\hat{\mathbf{H}}_{rd}$. The weight matrix minimizing (6.26) is given by

$$\mathbf{T}_{sr} = \big((\mathbf{F} \odot \hat{\mathbf{H}}_{rd})(\mathbf{F} \odot \hat{\mathbf{H}}_{rd})^H + \hat{\mathbf{C}}\big)^{-1}(\mathbf{F} \odot \hat{\mathbf{H}}_{rd}). \tag{6.27}$$

It will be seen in Section 6.4 that there is an obvious improvement in the estimation of $\mathbf{H}_{sr}$ by using (6.27) after the convergence of the BALS algorithm.

Similarly, we expect that the initial estimation of $\mathbf{H}_{rd}$ can be improved by the LMMSE approach. It can be shown from (6.23) that the covariance matrix of the noise $\mathbf{N}_j$, denoted as $\mathbf{\Theta}_j \triangleq \mathrm{E}[\mathbf{N}_j\mathbf{N}_j^H]$, $j = 1, \cdots, N_s$, is a diagonal matrix whose (k,k)-th diagonal element is given by $\sum_{m=1}^{M} \|f(k,m)\mathbf{h}_{rd,m}\|^2 + N_d$, where $\mathbf{h}_{rd,m}$ is the mth column of $\mathbf{H}_{rd}$. Thus, an improved LMMSE estimate of $\mathbf{H}_{rd}$ can be obtained as $\breve{\mathbf{H}}_{rd} = [\mathbf{T}_{rd}^H\tilde{\mathbf{Z}}]^T$, where $\mathbf{T}_{rd}$ is the $KN_s \times M$ weight matrix with

$$\mathbf{T}_{rd} = \big((\breve{\mathbf{H}}_{sr}^T \odot \mathbf{F})(\breve{\mathbf{H}}_{sr}^T \odot \mathbf{F})^H + \hat{\mathbf{\Theta}}\big)^{-1}(\breve{\mathbf{H}}_{sr}^T \odot \mathbf{F}). \tag{6.28}$$

Here $\hat{\mathbf{\Theta}} = \text{blkdiag}\big[\hat{\mathbf{\Theta}}_1, \hat{\mathbf{\Theta}}_2, \cdots, \hat{\mathbf{\Theta}}_{N_s}\big]$, and $[\hat{\mathbf{\Theta}}_j]_{k,k} = \sum_{m=1}^{M} \|f(k,m)\hat{\mathbf{h}}_{rd,m}\|^2 + N_d$, $k = 1, \cdots, K$, $j = 1, \cdots, N_s$, is an estimate of $[\mathbf{\Theta}_j]_{k,k}$ using $\hat{\mathbf{H}}_{rd}$.

We would like to mention that in [111], the WLS approach is used to improve the channel estimation after the LS algorithm. It will be shown in Section 6.4 that the LMMSE algorithm yields a smaller MSE of channel estimation (particularly for estimating $\mathbf{H}_{rd}$) than that of the WLS method in [111].

6.3.4 Extension to channel estimation in two-way MIMO relay systems

In the following, we show that the proposed algorithm can also be used for channel estimation in two-way MIMO relay systems.

In a two-way relay system, two users exchange their information through one or multiple relay nodes [123]. The received signal matrices at two users during the kth time block of the channel training period are given respectively by

$$\begin{aligned}\mathbf{Y}_{1,k} &= \mathbf{H}_{1,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{r,2}\mathbf{S}_2 + \mathbf{H}_{1,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{r,1}\mathbf{S}_1 \\ &\quad + \mathbf{H}_{1,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k} + \mathbf{V}_{1,k}, \quad k = 1, \cdots, K \qquad (6.29)\\ \mathbf{Y}_{2,k} &= \mathbf{H}_{2,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{r,1}\mathbf{S}_1 + \mathbf{H}_{2,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{r,2}\mathbf{S}_2 \\ &\quad + \mathbf{H}_{2,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k} + \mathbf{V}_{2,k}, \quad k = 1, \cdots, K \qquad (6.30)\end{aligned}$$

where $\mathbf{H}_{r,i}$, $i = 1, 2$, is the MIMO channel from user i to all relay nodes, $\mathbf{H}_{i,r}$, $i = 1, 2$, is the MIMO channel from all relay nodes to user i, and $\mathbf{V}_{i,k}$, $i = 1, 2$, is the noise matrix at user i during the kth time block.

The $N_i \times L$ training sequence $\mathbf{S}_i$ chosen by user i, $i = 1, 2$, in (6.29) and (6.30) is designed such that

$$\mathbf{S}_i\mathbf{S}_i^H = \mathbf{I}_{N_i}, \quad i = 1, 2, \qquad \mathbf{S}_1\mathbf{S}_2^H = \mathbf{0}_{N_1 \times N_2} \qquad (6.31)$$

where N_i is the number of antennas at user i. Note that $\mathbf{S}_1$ and $\mathbf{S}_2$ satisfying (6.31) can be easily constructed from the normalized DFT matrix with $L \geq N_1 + N_2$. Multiplying both sides of (6.29) with $\mathbf{S}_2^H$ and both sides of (6.30) with $\mathbf{S}_1^H$, we have

$$\mathbf{Y}_{1,k}\mathbf{S}_2^H = \mathbf{H}_{1,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{r,2} + \mathbf{H}_{1,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k}\mathbf{S}_2^H + \mathbf{V}_{1,k}\mathbf{S}_2^H, \quad k = 1, \cdots, K \quad (6.32)$$

$$\mathbf{Y}_{2,k}\mathbf{S}_1^H = \mathbf{H}_{2,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{H}_{r,1} + \mathbf{H}_{2,r}\mathcal{D}_k\{\mathbf{F}\}\mathbf{V}_{r,k}\mathbf{S}_1^H + \mathbf{V}_{2,k}\mathbf{S}_1^H, \quad k = 1, \cdots, K. \quad (6.33)$$

Now the proposed PARAFAC-based algorithm developed in Section 6.3.1 to Section 6.3.3 can be applied at user 1 to estimate $\mathbf{H}_{1,r}$ and $\mathbf{H}_{r,2}$ from (6.32) and at user 2 to estimate $\mathbf{H}_{2,r}$ and $\mathbf{H}_{r,1}$ from (6.33).

6.4 Numerical Examples

In this section, we study the performance of the proposed channel estimation algorithm through numerical simulations. In particular, we compare the proposed algorithm with the conventional TST scheme in Section 6.2, the LS-based algorithm in [110], and the WLS fitting algorithm in [111]. Note that the purpose of the WLS fitting in [111] is to improve the performance of the LS algorithm in [110]. To ensure a fair comparison, a factor of $\sqrt{K}$ is used to scale the training sequences $\mathbf{S}_1$ and $\mathbf{S}_2$ in the TST scheme such that the total energy spent on channel training is identical for all approaches. In the simulations, $\mathbf{F}$ is generated based on the DFT matrix, and the BALS algorithm is performed following the procedure in Table 6.1 with $\varepsilon = 1 \times 10^{-5}$. Similar to [110] and [111], the scaling ambiguity $\mathbf{\Delta}_1 = \mathbf{\Delta}_3^{-1}$ in the proposed algorithm is removed by assuming that the first column of $\mathbf{H}_{sr}$ contains all one elements[6]. For each channel realization, the normalized MSE (NMSE) of channel estimation for different algorithms is calculated as $\|\hat{\mathbf{H}}_{sr} - \mathbf{H}_{sr}\|^2 / \|\mathbf{H}_{sr}\|^2$ for the channel $\mathbf{H}_{sr}$, where $\hat{\mathbf{H}}_{sr}$ is the estimated value. The channel estimation errors of $\mathbf{H}_{rd}$ and $\mathbf{H}_{sd}$ are calculated in a similar way to that of $\mathbf{H}_{sr}$. All simulation results are averaged over 2000 independent channel realizations.

We consider a two-hop MIMO relay communication system with $M = 4$ single-antenna relay nodes, and the source and destination nodes are equipped with $N_s = N_d = 4$ antennas. Throughout the simulations, we use the minimal L, i.e., $L = N_s = 4$. The transmission power at the relay node is set to be 20dB above the noise level.

In the first example, we study the performance of the proposed algorithm and the TST approach with $K = 3$ where all channel matrices have i.i.d. complex Gaussian entries with zero-mean and variances of $1/N_s$, $1/M$, $1/(8N_s)$ for $\mathbf{H}_{sr}$, $\mathbf{H}_{rd}$, and $\mathbf{H}_{sd}$, respectively. Note that since $K < M$, the algorithms in [110] and [111] can not be applied in this case. The NMSE of both algorithms versus the source node transmission

[6]The scaling ambiguity is represented as $\hat{\mathbf{H}}_{sr} = \mathbf{\Delta}_3 \mathbf{H}_{sr}$ in (6.16) (with $\mathbf{\Pi} = \mathbf{I}_M$). Since the first column of $\mathbf{H}_{sr}$ contains all one elements, it can be seen that $[\mathbf{\Delta}_3]_{i,i} = \left[\hat{\mathbf{H}}_{sr}\right]_{i,1}$. Here $[\mathbf{A}]_{i,j}$ stands for the (i,j)-th element of matrix $\mathbf{A}$.

power P_s is shown in Fig. 6.2. Since the NMSE for the estimation of $\mathbf{H}_{rd}$ by the TST scheme does not change with P_s (see (6.7)), it is not displayed in Fig. 6.2 (neither in Figs. 6.4 and 6.6 later on). It can be seen that for the proposed algorithm, the NMSE of channel estimation decreases as P_s increases. As expected, the estimation of $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$ is improved by carrying out the additional MMSE estimation. At the low P_s level, the TST scheme is better than the proposed algorithm. While at medium to high P_s levels, the proposed algorithm significantly outperforms the TST scheme even without using the additional MMSE estimation. In fact, the TST scheme has an error floor in estimating $\mathbf{H}_{sr}$. The reason is that as can be seen from (6.9), the estimation of $\mathbf{H}_{sr}$ in the TST scheme is extracted from the estimation of the compound channel $\bar{\mathbf{H}}$ and the estimation of $\mathbf{H}_{rd}$. Thus, the accuracy of $\hat{\bar{\mathbf{H}}}$ and $\hat{\mathbf{H}}_{rd}$ has a great impact on the estimation of $\mathbf{H}_{sr}$. While in the proposed algorithm, $\mathbf{H}_{sr}$ is estimated together with $\mathbf{H}_{rd}$. We also observe from Fig. 6.2 that for the proposed algorithm, the NMSE of estimating $\mathbf{H}_{sd}$ is larger than that of $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$. This is due to the lower signal-to-noise ratio at the direct link as the source-destination distance is twice of the source-relay (or relay-destination) distance.

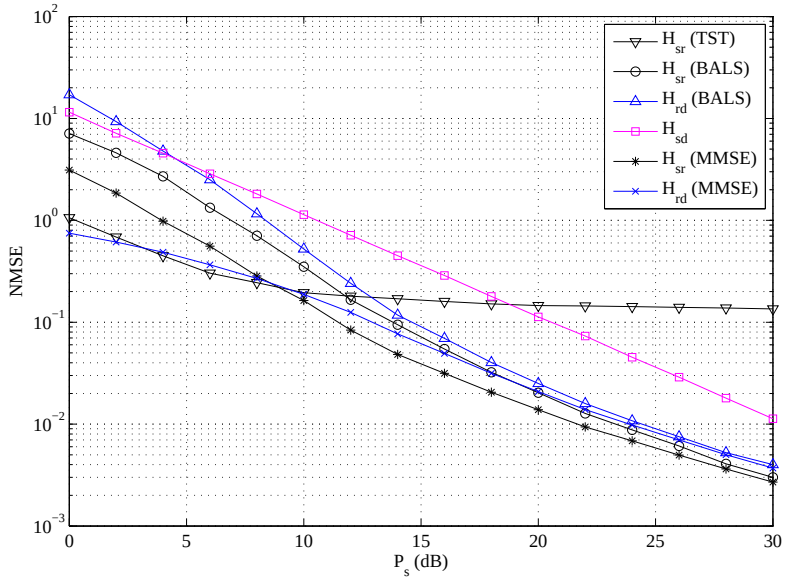

Figure 6.2: Example 1: Normalized MSE versus P_s for i.i.d. MIMO channels. $K = 3$.

The impact of channel estimation on the system BER performance in this example is shown in Fig. 6.3. QPSK constellations are used to modulate the source symbols, and 3000 randomly generated bits are transmitted for each channel realization. It can be

seen that at medium to high P_s levels, the proposed algorithm significantly outperforms the TST scheme even without using the additional MMSE estimation, and the TST scheme shows a high error floor. We also observe in Fig. 6.3 that around 1dB gain in P_s is obtained by using the additional MMSE estimation after the convergence of the BALS algorithm. At a BER of 1×10^{-4}, there is only around 2dB loss in P_s by using the estimated CSI obtained from the MMSE algorithm compared with the system using the perfect CSI.

In the second example, we consider correlated MIMO channels. Based on [12], we assume that $\mathbf{H}_l = \mathbf{Q}_l^{\frac{1}{2}} \mathbf{H}_l^{\mathrm{w}} \mathbf{C}_l^{\frac{T}{2}}$, where $l \in [sr, rd, sd]$ denotes the link index. Here $\mathbf{H}_{sr}^{\mathrm{w}}$, $\mathbf{H}_{rd}^{\mathrm{w}}$, and $\mathbf{H}_{sd}^{\mathrm{w}}$ are complex Gaussian random matrices having i.i.d. entries with zero mean and variances of $1/N_s$, $1/M$, $1/(8N_s)$, respectively, $\mathbf{Q}_l$ and $\mathbf{C}_l$ characterize the channel correlation at the receive side and the transmit side of link l, respectively. We adopt the commonly used exponential Toeplitz structure in [12] such that $[\mathbf{Q}_l]_{m,n} = \jmath_0(2\pi|m-n|/r_l)$ and $[\mathbf{C}_l]_{m,n} = \jmath_0(2\pi|m-n|/c_l)$, where $\jmath_0(\cdot)$ is the zeroth order Bessel function of the first kind, r_l and c_l stand for the correlation coefficients which depend on physical factors such as the angle of arrival spread, spacing between antenna elements, and the wavelength at the center frequency [12]. For the sake of simplicity, we choose $c_l = r_l = 2$ for all $l \in [sr, rd, sd]$. The NMSE and BER performance of different algorithms in this example are displayed in Fig. 6.4 and Fig. 6.5, respectively. It can be observed that similar to Fig. 6.2 and Fig. 6.3, the proposed algorithm performs better than the TST algorithm.

Table 6.2: Example 3: NMSE of the LS [110], the WLS [111], and the proposed BALS algorithm

P_s (dB)	0	4	8	12	16	20	24	28
BALS ($\mathbf{H}_{sr}$)	4.4316	1.2800	0.3212	0.0891	0.0296	0.0118	0.0050	0.0025
LS [110] ($\mathbf{H}_{sr}$)	4.4316	1.2800	0.3212	0.0891	0.0296	0.0118	0.0050	0.0025
BALS ($\mathbf{H}_{rd}$)	0.9208	0.3678	0.1359	0.0527	0.0206	0.0086	0.0038	0.0020
LS [110] ($\mathbf{H}_{rd}$)	0.9208	0.3679	0.1359	0.0527	0.0206	0.0086	0.0038	0.0020
WLS [111] ($\mathbf{H}_{rd}$)	0.9207	0.3678	0.1358	0.0526	0.0204	0.0084	0.0037	0.0020

In the third example, we simulate all algorithms with $K = 4$ and i.i.d. channel matrices. Since $K = M$, now we can compare the performance of the proposed algorithm with the algorithms developed in [110] and [111]. The NMSE of the LS algorithm in

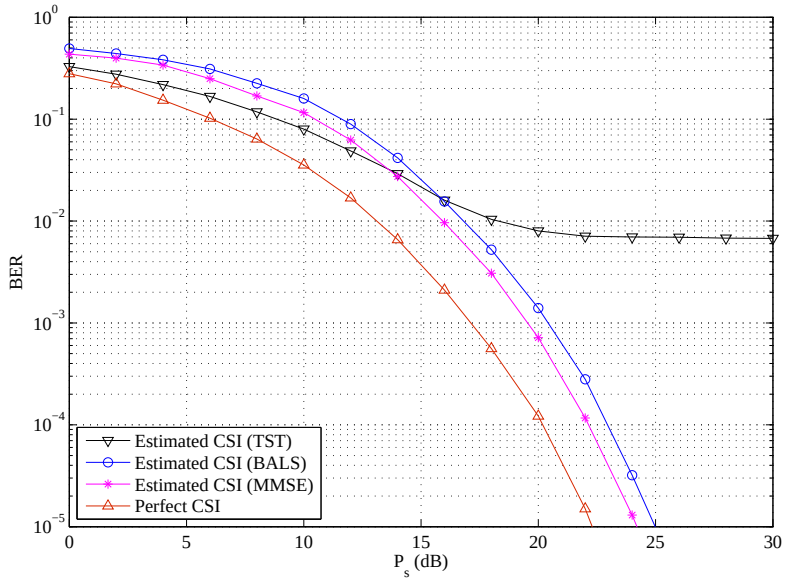

Figure 6.3: Example 1: BER versus P_s for i.i.d. MIMO channels. $K = 3$.

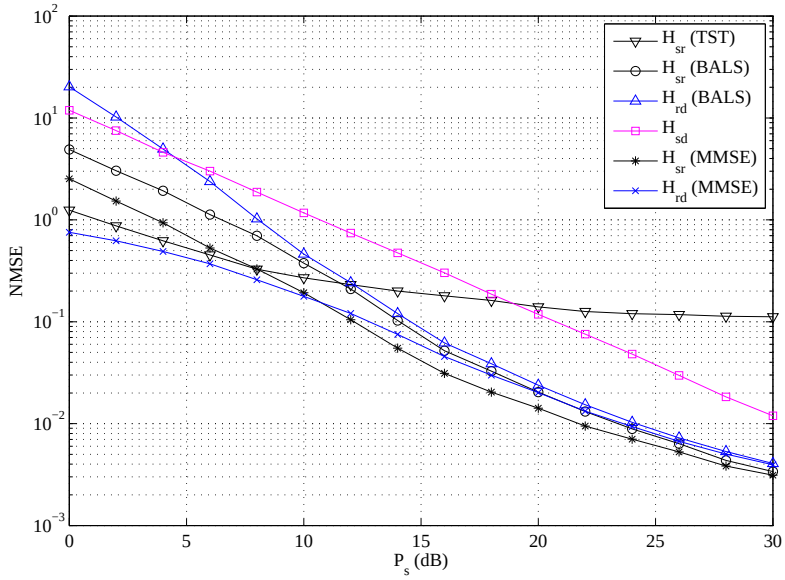

Figure 6.4: Example 2: Normalized MSE versus P_s for correlated MIMO channels. $K = 3$.

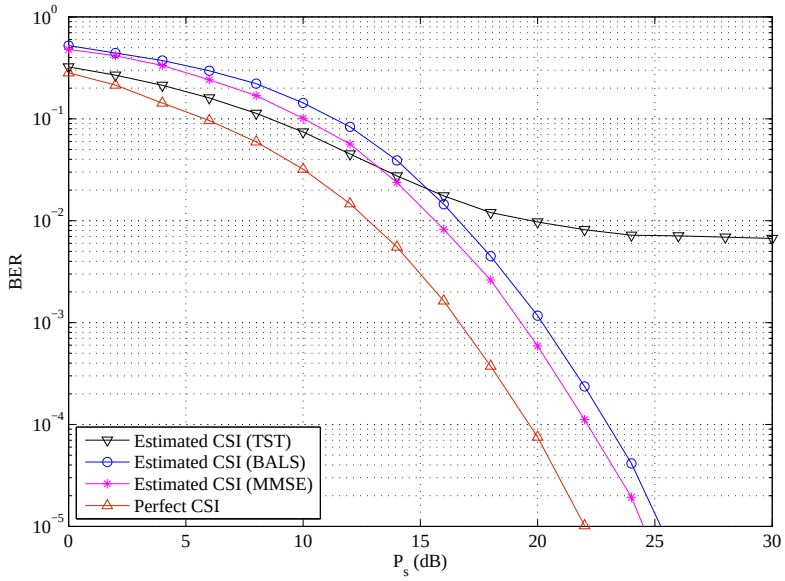

Figure 6.5: Example 2: BER versus P_s for correlated MIMO channels. $K = 3$.

[110], the additional WLS fitting in [111], and the proposed algorithm is shown in Table 6.2. It can be seen that the proposed BALS fitting yields the same NMSE as the LS approach. The NMSE of the proposed algorithm, the TST scheme, and the WLS fitting versus P_s is shown in Fig. 6.6, where we observe that as K is increased from 3 to 4, the NMSE of all algorithms is reduced compared with that in Fig. 6.2. Note that in Figs. 6.2 and 6.4, we used $K = 3$, while in Fig. 6.6, we used $K = 4$. According to Theorem 6.1, $K = 3$ is closer than $K = 4$ to the minimum $K\,(K = 2)$ that makes the PARAFAC model identifiable, which has more adverse effect on the estimation of $\mathbf{H}_{rd}$ than $\mathbf{H}_{sr}$. With an increased K, it is more likely that $\mathbf{H}_{rd}$(BALS) has smaller estimation error compared with that of $\mathbf{H}_{sr}$(BALS) in Fig. 6.6. Moreover, we see from Table 6.2 and Fig. 6.6 that the improvement in NMSE of the WLS fitting over the LS algorithm is obvious for the estimation of $\mathbf{H}_{sr}$, while the improvement for that of $\mathbf{H}_{rd}$ is negligible. The reason is that $\hat{\mathbf{C}}_k$ in (6.27) is non-diagonal, while $\hat{\mathbf{\Theta}}_j$ in (6.28) is diagonal. In contrast to the WLS fitting (Table 6.2), it can be seen from Fig. 6.6 that the MMSE approach greatly reduces the NMSE of estimating $\mathbf{H}_{rd}$. From Fig. 6.6 we also observe that the MMSE approach yields a smaller NMSE in estimating $\mathbf{H}_{sr}$ compared with the WLS fitting.

For this example, with a random initialization of $\mathbf{H}_{rd}$, the average and maximum number of iterations over 2000 independent channel realizations required by the pro-

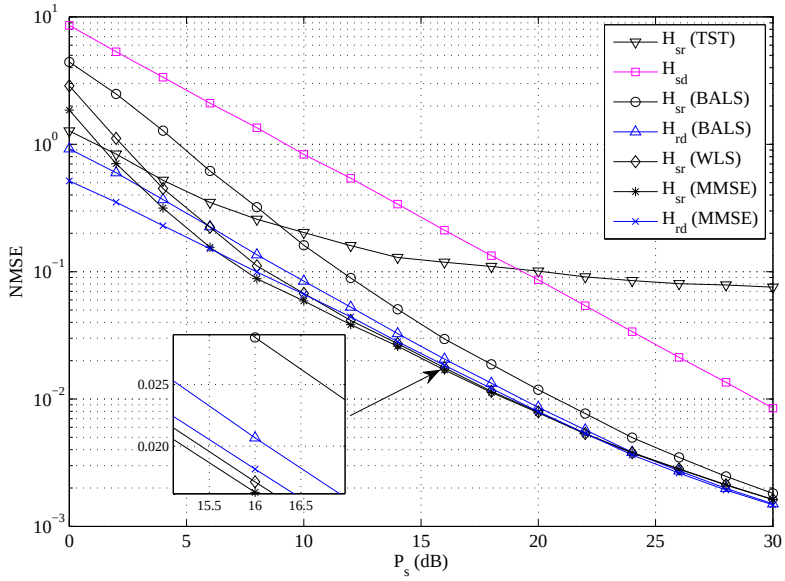

Figure 6.6: Example 3: Normalized MSE versus P_s for i.i.d. MIMO channels. $K = 4$.

posed BALS algorithm till convergence at different P_s level are listed in Table 6.3. Based on the analysis of the overall complexity of the LS-based algorithm and the per-iteration complexity of the BALS algorithm in Section 6.3.2, it can be seen from the second row of Table 6.3 that in average at medium and high P_s levels, the overall complexity of the proposed BALS algorithm is similar to that of the LS algorithm. When P_s is low, the complexity of the BALS algorithm is slightly higher than that of the LS algorithm. It can also be seen from the third row of Table 6.3 that at medium to high P_s levels (which is the P_s range in practical systems), the maximum number of iterations is only twice of or almost identical to the average ones.

Table 6.3: Iterations required till convergence by the proposed BALS algorithm

P_s (dB)	0	4	8	12	16	20	24	28
Iterations (average)	7	6	6	5	4	4	3	3
Iterations (maximum)	23	18	13	10	8	6	5	4

For the third example, the comparison of BERs among the system using different channel estimation algorithms is demonstrated in Fig. 6.7. We observe from Fig. 6.7 that as K is increased from 3 to 4, the BER of all algorithms is reduced compared with that in Fig. 6.3. Similar to Fig. 6.3, we see from Fig. 6.7 that the proposed

algorithm significantly outperforms the TST scheme at medium to high P_s levels, where the latter scheme shows a high error floor. The LS approach in [110] has the same BER performance as the proposed BALS approach, while the proposed MMSE algorithm performs slightly better than that of the WLS algorithm in [111]. At a BER of 1×10^{-4}, the loss in P_s using the estimated CSI from the MMSE algorithm is less than 2dB compared with the system using the perfect CSI. This is quite reasonable for practical systems.

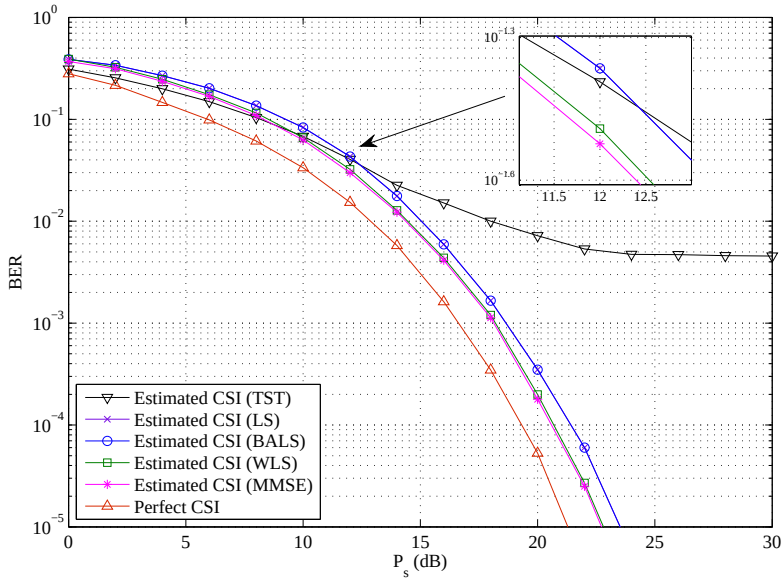

Figure 6.7: Example 3: BER versus P_s for i.i.d. MIMO channels. $K = 4$.

Based on the three numerical examples above, one can draw the following conclusions: (1) The proposed algorithm performs well in case of $K < M$ where the algorithms in [110] and [111] stop working; (2) When $K \geq M$, the proposed MMSE approach outperforms the algorithm in [111]; (3) The computational complexity of the proposed algorithm is similar to that of [110]; (4) The proposed algorithm performs well for both i.i.d. and correlated fading MIMO channels.

6.5 Chapter Summary

In this chapter, we have developed a novel PARAFAC-based channel estimation method for two-hop MIMO relay communication systems. The proposed algorithm provides the

destination node with full knowledge of all channel matrices involved in the communication. Compared with existing approaches, the proposed algorithm requires less number of training data blocks, yields smaller channel estimation error, and is applicable for both one-way and two-way MIMO relay systems. Simulation results demonstrate the effectiveness of the proposed channel estimation algorithm.

Chapter 7

Conclusions and Future Work

In next generation wireless systems, multiple users equipped with multiple antennas will transmit simultaneously to the base station with multiple receive antennas and vice versa. MIMO transceiver design taking multiuser interference into consideration is therefore important. In this thesis, we have developed several advanced algorithms for multiuser MIMO relay communication systems.

7.1 Concluding Remarks

Both uplink (MAC) and downlink (BC) multiuser MIMO relay systems have been investigated. In Chapter 2, the optimal source, relay, and receive matrices design problem has been considered for uplink multiuser MIMO relay communication systems based on MMSE criterion. We develop two iterative methods to solve the highly nonconvex joint source, relay, and receiver optimization problem. In particular, we show that for given source precoding matrices, the optimal relay amplifying matrix diagonalizes the source-relay-destination channel. While for fixed relay matrix and source matrices of all other users, the source matrix of each user has a general beamforming structure.

Then in Chapter 3, we consider multicasting in the downlink multiuser MIMO relay system where one transmitter multicasts common message to multiple receivers with the aid of a relay node. Joint transmit and relay precoding design problems are investigated for multicasting multiple data streams based on two design criteria. In the first scheme, we aim at minimizing the maximal MSE of the signal waveform estimation among all receivers subjecting to power constraints at the transmitter and the relay node. In the second scheme, we propose a total transmission power minimization strategy subjecting

to QoS constraints. We propose low complexity solutions for both problems under some mild approximation. In particular, we show that under (moderately) high first-hop SNR assumption, both problems can be formulated as standard SDP problems and can be efficiently solved using existing solvers. We also derive the analytical solution to both problems for the special case of single data stream multicasting and show that the relay precoding matrix optimization problem can be equivalently converted to the transmit beamforming problem for single-hop multicasting systems.

Then an interesting duality between the uplink and the downlink MIMO relay systems has been identified in Chapter 4. Applying this duality relationship, a DPC-based nonlinear transmitter has been designed from DFE-based receivers in a multi-hop MIMO relay system.

In Chapter 5, an interference MIMO relay system has been considered where multiple source nodes communicate with their desired destination nodes concurrently with the aid of distributed relay nodes all equipped with multiple antennas. An iterative joint power control and beamforming algorithm is developed to minimize the total source and relay transmit power such that a minimum SINR threshold is maintained at each receiver. In particular, we apply the semidefinite relaxation technique to transform the relay transmission power minimization problem into an SDP problem which can be efficiently solved by interior point-based methods. Simulation results demonstrate the effectiveness of the proposed algorithm.

Finally, we developed a PARAFAC-based channel estimation approach for dual-hop MIMO relay systems in Chapter 6. The proposed algorithm provides the destination node with full knowledge of all channel matrices involved in the communication. Compared with existing approaches, the proposed algorithm requires less number of training data blocks, yields smaller channel estimation error, and is applicable for both one-way and two-way MIMO relay systems with single or multiple relay nodes.

7.2 Future Works

In this thesis, we have developed a few advanced signal processing algorithms for multiuser MIMO relay systems. However, there are still many possibilities for extending this dissertation work. We have proposed a couple of iterative algorithms in Chapter 2 for the uplink multiuser MIMO relay systems. Any possible closed-form solution to the

problem can be an interesting future work since the complexity of iterative algorithms is comparatively higher than closed-form solutions.

Recently, there has been a growing interest on beamforming problems for multicasting in the downlink MIMO systems. We have extended the existing single-hop MIMO multicasting schemes to dual-hop MIMO multicasting systems in Chapter 3. We solved the min-max MSE problem for multicasting multiple data streams and the max-min rate problem for single-stream multicasting over two hops. However, the max-min rate problem for multiple-stream multicasting still remains open as a challenging problem.

It will also be interesting to investigate the performance of the one-way relaying algorithm for interference systems proposed in Chapter 5 for two-way relay networks. Additionally, the work can be extended from multiple peer-to-peer communications to multiple-group multicasting systems for multicasting different messages to each group. The duality relationships established in Chapter 4 can also be investigated for the interference system considered in Chapter 5.

Finally, the robust solution against channel uncertainties for each problem identified in the thesis is of course of practical interests.

Bibliography

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[2] A. Goldsmith, *Wireless Communications.* Cambridge University Press, 2005.

[3] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, pp. 214–226, 2004.

[4] F. Kaltenberger, M. Kountouris, L. Cardoso, R. Knopp, and D. Gesbert, "Capacity of linear multi-user MIMO precoding schemes with measured channel data," in *Proc. IEEE 9th Workshop on Signal Process. Adv. Wireless Commun.*, Recife, Brazil, Jul. 2008, pp. 580-584.

[5] E. Telatar, "Capacity of multi-antenna gaussian channels," *AT&T Bell Labs, Tech. Memo.*, Murray Hill, NJ, Jun. 1995.

[6] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[7] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 281–302, Apr. 2003.

[8] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications–a key to gigabit wireless," *Proc. IEEE*, vol. 92, pp. 198–218, Feb. 2004.

[9] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 684–702, Jun. 2003.

[10] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, pp. 41–59, Autumn 1996.

[11] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1998.

[12] D.-S. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 503–513, Mar. 2000.

[13] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, May 2003.

[14] M. Shafi, H. Huang, A. Hottinen, P. J. Smith, and R. A. Valenzuela, "MIMO systems and applications: Field experience, practical aspects, limitations and challenges," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 841–844, Guest Editorial, Aug. 2008.

[15] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Sep. 1979.

[16] R. Pabst, B. H. Walke, D. C. Schultz, D. C. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, pp. 80–89, Sep. 2004.

[17] I. Kang, W. Sheen, R. Chen, and S. L. C. Hsiao, "Throughput improvement with relay-augmented cellular architecture," *IEEE 802.16mmr-05 008*, http://www.wirelessman.org, Sep. 2005.

[18] V. Genc, S. Murphy, Y. Yu, and J. Murphy, "IEEE 802.16j relay-based wireless access networks: An overview," *IEEE Wireless Commun.*, vol. 15, pp. 56–63, Oct. 2008.

[19] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 29–43, Jan. 2005.

[20] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1774–1786, May 2007.

[21] B. Rankov and A. Wittneben, "On the capacity of relay-assisted wireless MIMO channels," in *Proc. 5th IEEE Workshop on Signal Processing Advances in Wireless Commun.*, Lisbon, Portugal, Jul. 2004, pp. 323-327.

[22] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398–1407, Apr. 2007.

[23] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Processing*, vol. 55, pp. 2593–2604, Jun. 2007.

[24] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, pp. 517–519, Jul. 2008.

[25] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, pp. 5062–5073, Oct. 2008.

[26] C. Xing, S. Ma, Y. C. Wu, and T. S. Ng, "Transceiver design for dual-hop nonregenerative MIMO-OFDM relay systems under channel uncertainties," *IEEE Trans. Signal Process.*, vol. 58, pp. 6325–6339, Dec. 2010.

[27] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837–4851, Dec. 2009.

[28] Y. Rong and Y. Hua, "Optimality of diagonalization of multihop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 6068–6077, Dec. 2009.

[29] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Info. Theory*, vol. 32, pp. 85–96, Jan. 1986.

[30] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Commun. Mag.*, pp. 124–136, Oct. 1996.

[31] D. Gesbert, M. Kountouris, R. W. Heath Jr., C.-B. Chae, and T. Sälzer, "From single-user to multiuser communications: Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, pp. 36–46, Sep. 2007.

[32] M. R. A. Khandaker and Y. Rong, "Performance measure of multi-user detection algorithms for MIMO relay network," in *Proc. 10th Postgraduate Electrical Engineerig and Computing Symposium (PEECS'2009)*, Perth, Australia, Oct. 1, 2009.

[33] R. L.-U. Choi, M. T. Ivrlač, R. D. Murch, and W. Utschick, "On strategies of multiuser MIMO transmit signal processing," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1936–1941, Nov. 2004.

[34] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, pp. 1313–1324, Oct. 1998.

[35] M. R. A. Khandaker and Y. Rong, "Precoding design for MIMO relay multicasting," *IEEE Trans. Signal Process.*, submitted, Jul. 2012.

[36] ——, "Joint transceiver optimization for multiuser MIMO relay communication systems," *IEEE Trans. Signal Process.*, to appear, 2012.

[37] ——, "Interference MIMO relay channel: Joint power control and transceiver-relay beamforming," *IEEE Trans. Signal Process.*, to appear, 2012.

[38] ——, "Multicasting MIMO relay optimization based on min-max MSE criterion," in *Proc. IEEE Int. Conf. Commun. Systems (ICCS 2012)*, Singapore, Nov. 21-23, 2012.

[39] ——, "Joint power control and beamforming for peer-to-peer MIMO relay systems," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Nov. 9-11, 2011.

[40] ——, "Joint power control and beamforming for interference MIMO relay channel," in *Proc. 17th Asia-Pacific Conf. Commun. (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

[41] ——, "Joint source and relay optimization for multiuser MIMO relay communication systems," in *Proc. 4th Int. Conf. Signal Process. Commun. Systems (ICSPCS'2010)*, Gold Coast, Australia, Dec. 13-15, 2010.

[42] D. P. Palomar and Y. Jiang, *MIMO Transceiver Design via Majorization Theory*. now Publishers, 2007.

[43] A. Toding, M. R. A. Khandaker, and Y. Rong, "Joint source and relay optimization for distributed MIMO relay system," in *Proc. 17th Asia-Pacific Conf. on Commun. (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

[44] ——, "Joint source and relay optimization for parallel MIMO relay networks," *EURASIP Journal Adv. Sig. Process.*, vol. 2012:174, Aug. 2012.

[45] ——, "Joint source and relay optimization for parallel MIMO relays using MMSE-DFE receiver," in *Proc. 16th Asia-Pacific Conf. on Commun. (APCC'2010)*, Auckland, New Zealand, Nov. 1-3, 2010, pp. 12-16.

[46] ——, "Optimal joint source and relay beamforming for parallel MIMO relay networks," in *Proc. 6th Int. Conf. Wireless Commun., Networking and Mobile Computing (WiCOM'2010)*, Chengdu, China, Sep. 23-25, 2010.

[47] Y. Rong and M. R. A. Khandaker, "Channel estimation of dual-hop MIMO relay system via parallel factor analysis," in *Proc. 17th Asia-Pacific Conf. Commun. (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.

[48] 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, pp. 211–217, Jan. 2010.

[49] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[50] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of gaussian multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, pp. 768–783, May 2004.

[51] S. A. Jafar, K. S. Gomadam, and C. Huang, "Duality and rate optimization for multiple access and broadcast channels with amplify-and-forward relays," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3350–3370, Oct. 2007.

[52] C.-B. 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, pp. 727–738, Feb. 2008.

[53] J. Yu, D. Liu, C. Yi, and G. Yue, "Relay-assisted MIMO multiuser precoding in fixed relay networks," in *Proc. Int. Conf. on Wireless Commun., Networking and Mobile Computing*, Shanghai, China, Sep. 2007, pp. 881-884.

[54] L. Weng and R. D. Murch, "Multi-user MIMO relay system with self-interference cancellation," in *Proc. IEEE WCNC*, Kowloon, China, Mar. 2007, pp. 958-962.

[55] G. Li, Y. Wang, T. Wu, and J. Huang, "Joint linear filter design in multi-user non-regenerative MIMO-relay systems," in *Proc. IEEE Int. Conf. Commun.*, Dresden, Germany, Jun. 2009, pp. 1-6.

[56] Y. Yu and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users," *IEEE Trans. Signal Process.*, vol. 58, pp. 2823–2835, May 2010.

[57] T. Taniguchi, N. B. Ramli, and Y. Karasawa, "Design of multiuser MIMO AF relay system with interference cancellation," in *Proc. 6th Int. Wireless Commun. Mobile Comput. Conf.*, Caen, France, Jun. 2010, pp. 1075-1080.

[58] S. Jang, J. Yang, and D. K. Kim, "Minimum MSE design for multiuser MIMO relay," *IEEE Commun. Lett.*, vol. 14, pp. 812–814, Sep. 2010.

[59] Y. Rong and M. R. A. Khandaker, "On uplink-downlink duality of multi-hop MIMO relay channel," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1923–1931, Jun. 2011.

[60] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cilffs, NJ: Prentice Hall, 1993.

[61] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U. K.: Cambridge University Press, 2004.

[62] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 772–781, Sep. 1978.

[63] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming (web page and software)." *http://cvxr.com/cvx*, April, 2010.

[64] Y. Nesterov and A. Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.

[65] M. A. Khojastepour, A. Salehi-Golsefidi, and S. Rangarajan, "Towards an optimal beamforming algorithm for physical layer multicasting," in *Proc. IEEE Inf. Theory Workshop*, Paraty, Brazil, Oct. 16-20, 2011, pp. 395 –399.

[66] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. T. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, pp. 2239–2251, Jun. 2006.

[67] N. Jindal and Z.-Q. Luo, "Capacity limits of multiple antenna multicast," in *Proc. IEEE ISIT*, Seattle, USA, Jul. 09-14, 2006, pp. 1841–1845.

[68] S. Y. Park and D. Love, "Capacity limits of multiple antenna multicasting using antenna subset selection," *IEEE Trans. Signal Process.*, vol. 56, pp. 2524–2534, Jun. 2008.

[69] S. Y. Park, D. J. Love, and D. H. Kim, "Capacity limits of multi-antenna multicasting under correlated fading channels," *IEEE Trans. Commn.*, vol. 58, pp. 2002–2013, Jul. 2010.

[70] E. Chiu and V. K. N. Lau, "Precoding design for multi-antenna multicast broadcast services with limited feedback," *IEEE Systems Journal*, vol. 4, pp. 550–560, Dec. 2010.

[71] Q. Li and W.-K. Ma, "Multicast secrecy rate maximization for MISO channels with multiple multi-antenna eavesdroppers," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 5-9, 2011.

[72] S. X. Wu and W.-K. Ma, "Multicast transmit beamforming using a randomize-in-time strategy," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 22-27, 2011, pp. 3376–3379.

[73] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control," *IEEE Trans. Signal Process.*, vol. 57, pp. 4882–4894, Dec. 2009.

[74] N. Bornhorst and M. Pesavento, "An iterative convex approximation approach for transmit beamforming in multi-group multicasting," in *Proc. IEEE 12th Int. Workshop Signal Process. Adv. Wireless Commun.*, San Francisco, USA, Jun. 26-29, 2011, pp. 426–430.

[75] D. Senaratne and C. Tellambura, "Beamforming for physical layer multicasting," in *Proc. IEEE WCNC*, Cancun, Mexico, Mar. 28-31, 2011, pp. 1776–1781.

[76] M. Kaliszan, E. Pollakis, and S. Stańczak, "Efficient beamforming algorithms for MIMO multicast with application-layer coding," in *Proc. IEEE ISIT*, St. Petersburg, Russia, Jul. 31-Aug. 5, 2011, pp. 928–932.

[77] M. A. Khojastepour, A. Khajehnejad, K. Sundaresan, and S. Rangarajan, "Adaptive beamforming algorithms for wireless link layer multicasting," in *Proc. IEEE PIMRC*, Toronto, Canada, Sep. 11-14, 2011, pp. 1994–1998.

[78] N. Bornhorst, M. Pesavento, and A. B. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Signal Process.*, vol. 60, pp. 221–232, Jan. 2012.

[79] H. Xu, J. Jin, and B. Li, "YMMV: Multiple session multicast with MIMO," in *Proc. IEEE GLOBECOM*, Texas, USA, Dec. 5-9, 2011.

[80] H. Zhang, X. You, G. Wu, and H. Wang, "Cooperative multi-antenna multicasting for wireless networks," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 6-10, 2010.

[81] Y. Rong, "Simplified algorithms for optimizing multiuser multi-hop MIMO relay systems," *IEEE Trans. Commun.*, vol. 59, pp. 2896–2904, Oct. 2011.

[82] C. Song, K.-J. Lee, and I. Lee, "MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2310–2319, Jul. 2010.

[83] Y. Rong, "Multi-hop non-regenerative MIMO relays - QoS considerations," *IEEE Trans. Signal Processing*, vol. 59, pp. 290–303, Jan. 2011.

[84] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, New York, 2007.

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

[86] Y. Rong, "Optimal linear non-regenerative multi-hop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2268–2279, Jul. 2010.

[87] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[88] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, pp. 138–139, Mar. 1971.

[89] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, pp. 774–780, Aug. 1972.

[90] S. Shamai and L. Laroia, "The inter-symbol interference channel: Lower bounds on capacity and channel precoding loss," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1388–1404, Sep. 1996.

[91] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 16–27, Jan. 2002.

[92] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Honolulu, HI, Apr. 2007, vol. 3.

[93] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity – Part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[94] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[95] R. Zhang, C. C. Chai, and Y.-C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Trans. Signal Process.*, vol. 57, pp. 726–737, Feb. 2009.

[96] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, pp. 4306–4316, Sep. 2008.

[97] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Trans. Signal Process.*, vol. 57, pp. 3053–3062, Aug. 2009.

[98] B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal Process.*, vol. 58, pp. 6355–6368, Dec. 2010.

[99] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

[100] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1991.

[101] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, "Quasi-ML multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Trans. Signal Process.*, vol. 50, pp. 912–922, Apr. 2002.

[102] P. Tseng, "Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation," *SIAM J. Optim.*, vol. 14, pp. 268–283, Jul. 2003.

[103] W. Ai, Y. Huang, and S. Zhang, "New results on hermitian matrix rank-one decomposition," *Math. Program.,* ser. A (2011) 128, pp. 253–283, Aug. 2009.

[104] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, pp. 2455–2460, May 2011.

[105] D. P. Palomar, M. Lagunas, and J. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Process.*, vol. 52, pp. 1179–1197, May 2004.

[106] K. Phan, T. Le-Ngoc, S. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 2535–2545, May 2009.

[107] G. Lebrun, M. Faulkner, M. Shafi, and P. J. Smith, "MIMO ricean channel capacity: An asymptotic analysis," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1343–1350, Jun. 2006.

[108] Y. Rong, "Optimal joint source and relay beamforming for MIMO relays with direct link," *IEEE Commun. Lett.*, vol. 14, pp. 390–392, May 2010.

[109] F. Roemer and M. Haardt, "Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Tran. Signal Process.*, vol. 58, pp. 5720–5735, Nov. 2010.

[110] P. Lioliou and M. Viberg, "Least-squares based channel estimation for MIMO relays," in *Proc. IEEE WSA*, Darmstadt, Germany, Feb. 2008, pp. 90-95.

[111] P. Lioliou, M. Viberg, and M. Coldrey, "Performance analysis of relay channel estimation," in *Proc. IEEE Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2009, pp. 1533-1537.

[112] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1907–1916, May 2008.

[113] F. Gao, B. Jiang, X. Gao, and X.-D. Zhang, "Superimposed training based channel estimation for OFDM modulated amplify-and-forward relay networks," *IEEE Trans. Commun.*, vol. 59, pp. 2029–2039, Jul. 2011.

[114] T. Kong and Y. Hua, "Optimal channel estimation and training design for MIMO relays," in *Proc. IEEE Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2010, pp. 663-667.

[115] R. A. Harshman, "Foundations of PARAFAC procedure: Models and conditions for an 'explanatory' multi-mode factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[116] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, pp. 2377–2388, Aug. 2000.

[117] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, pp. 810–823, Mar. 2000.

[118] H. Wang, Y. Lin, and B. Chen, "Data-efficient blind OFDM channel estimation using receiver diversity," *IEEE Trans. Signal Process.*, vol. 51, pp. 2613–2623, Oct. 2003.

[119] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Blind timing and carrier synchronization in decode and forward cooperative systems," in *Proc. IEEE Int. Conf. Commun.*, Kyoto, Japan, Jun. 5-9, 2011.

[120] I. Hammerström and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2798–2802, Aug. 2007.

[121] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, pp. 884–893, Mar. 2006.

[122] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Applicat.*, vol. 16, pp. 95–138, 1977.

[123] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 699–712, Jun. 2009.