UNIVERSITY OF CALIFORNIA
RIVERSIDE

On Duality of MIMO Relays and Performance Limits of Full-Duplex MIMO Radios

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Ali Cagatay Cirik

March 2014

Dissertation Committee:

Professor Yingbo Hua, Chairperson
Professor Ilya Dumer
Professor Ertem Tuncel

The Dissertation of Ali Cagatay Cirik is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

First and foremost, I would like to thank and express my gratitude to my advisor, Professor Yingbo Hua for his support, valuable guidance and patience during my PhD period. His thoughtful feedbacks, careful comments and criticism, and knowledgeable suggestions, have significantly enhanced my thought processes. His office door was always open (after 2.00pm :)) whenever I needed his advice on my research problems. If I can be half the advisor and researcher he is, I'd consider myself a great success. I will always be indebted to him for all the things he has done for me.

I would like to thank all my colleagues in Laboratory of Signals, Systems and Networks, where I have had the great opportunity to meet high talented people. Qian Gao, Yiming Ma, Shenyang Xu, Armen Gholian, you made this lab a great place to work.

I would also like to thank my collaborators, Dr. Yue Rong from Curtin University; Prof. Martin Haardt and Jianshu Zhang from Ilmenau University of Technology, and Dr. Rui Wang from the Chinese University of Hong Kong, who all worked with me closely. I would also like to thank the jury members Prof. Ilya Dumer and Prof. Ertem Tuncel for agreeing to read this thesis and participate in my public defense.

I would also like to thank to my friends who made Riverside my second home. Roger Mousalli, Mehran Kafai, I was happy to have you guys as my housemates. Thank you for your friendship, and more importantly thank you for all the delicious Lebanese and Persian foods you cooked:) Selda Ors, Busra Celikkaya, Muzaffer Akbay, Omer Cetinkale, Doruk Sart, Onur Turkcu. Thank you all for the "crazy" parties we threw in Riverside, the adventures in Las Vegas, San Diego, LA... I will always miss our Sunday barbecues,

and Saturday Xbox nights:). And i want to thank Makbule Koksal, for opening her house, and feeding me with the delicious Turkish dishes she cooked, and Abdurrahman Koksal for fixing my old car every time it breaks down.

Finally, I would like to extend my deepest gratitude to my beloved parents, Mustafa and Elmas, and my brother Kubilay, for their constant love, unconditional support throughout the long path of my academic journey. I owe them all I have ever accomplished. And special note to my father Mustafa Cirik. I want to thank you for being such a great father. Please never give up, and fight like hell.

What Cancer Cannot Do

Cancer is so limited...

It cannot cripple love.

It cannot shatter hope.

It cannot corrode faith.

It cannot destroy peace.

It cannot kill friendship.

It cannot suppress memories.

It cannot silence courage.

It cannot invade the soul.

It cannot steal eternal life.

It cannot conquer the spirit.

Author Unknown

To my parents.

ABSTRACT OF THE DISSERTATION

On Duality of MIMO Relays and Performance Limits of Full-Duplex MIMO Radios

by

Ali Cagatay Cirik

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2014
Professor Yingbo Hua, Chairperson

In the first part of this thesis, linear transmitters and receivers (i.e., transceivers) are designed for bi-directional and/or relay multiple input multiple output (MIMO) full-duplex (FD) systems. The transmitters and receivers are designed under imperfect channel state information (CSI) and transmitter/receiver impairments at the FD nodes. Different metrics, like ergodic sum-rate maximization, weighted sum-rate maximization, sum mean-squared-error (MSE) minimization and maximum per-node MSE minimization are considered subject to individual and/or total power constraints in the system.

The proposed sum-rate maximization algorithms exploit both spatial and temporal freedoms of the source covariance matrices of the MIMO links between transmitters and receivers to achieve a higher achievable sum-rate. It is observed through simulations that the algorithms reduce to a FD scheme when the nominal self-interference is low, or to a half-duplex (HD) scheme when the nominal self-interference is high.

As for the MSE based transceiver designs, we studied the sum-MSE and Min-Max MSE transceiver design problems for a FD MIMO bi-directional system that suffers from

self-interference under the imperfect CSI knowledge and limited dynamic ranges at the transmitters and receivers. Since the globally optimal solution is difficult to obtain due to the non-convex nature of the problems, algorithms that iterate between transmit precoding and receive filtering matrices while keeping the other fixed are proposed. It is shown in simulations that sum-MSE minimization scheme achieves the minimum sum MSE over two FD nodes, and the Min-Max MSE minimization scheme almost achieves the same MSE for the two FD nodes.

In the second part of this thesis, we establish the uplink-downlink duality in terms of signal-to-interference-plus-noise ratio (SINR), MSE, and capacity for uplink and downlink multi-hop amplify-and-forward (AF) MIMO relay channels, which is a generalization of several previously established uplink-downlink duality results. And an interesting perspective to the relation of the uplink-downlink duality based on the Karush-Kuhn-Tucker (KKT) conditions of sum-MSE transceiver optimization problems for uplink and downlink multi-hop AF MIMO relay channels is provided.

Contents

List of Figures **xiii**

List of Tables **xv**

1 Introduction **1**
1.1 Outline of the Thesis 3

I Performance Limits of Full-Duplex MIMO Radios **5**

2 Literature Survey on Full-Duplex Radios **6**
2.1 Experimental Works 9
2.2 Theoretical Works 19
2.2.1 Self-Interference Cancellation 19
2.2.2 Multi-User 23
2.2.3 Cognitive Radios 25
2.3 Performance Analysis 26
2.3.1 Transceiver Design 27
2.3.2 Outage Performance 33
2.3.3 Diversity 34
2.3.4 Relay/Antenna Selection 35
2.3.5 Other issues 36

3 Achievable Rates of Full-Duplex MIMO Radios in Fast Fading Channels with Imperfect Channel Estimation **40**
3.1 Introduction 41
3.2 System Model for a FD Bi-Directional Link 46
3.3 Achievable Rates 50
3.4 Maximization of the Sum Ergodic Mutual Information 55
3.4.1 Gradient Projection Approach 55
3.4.2 Approximation of Sum Ergodic Mutual Information 59
3.5 Full-Duplex Relay Systems 61

3.5.1 Maximization of the Ergodic Mutual Information of the FD Relay System . . . 63
3.6 Simulation Results . . . 64
3.7 Appendix 3.A: Parameters in (3.8) . . . 71
3.8 Appendix 3.B: Parameters in (3.15) . . . 71

4 Weighted-Sum-Rate Maximization for Bi-directional Full-Duplex MIMO Systems 74
4.1 Introduction . . . 74
4.2 System Model . . . 76
4.3 Weighted-Sum-Rate Maximization . . . 80
4.3.1 MSE Weight Design . . . 81
4.3.2 Sum-power constrained transceiver design . . . 82
4.3.3 Individual-power-constrained transceiver design . . . 83
4.4 Simulation Results . . . 85
4.5 Appendix 4.A: Covariance of (4.3) . . . 91
4.6 Appendix 4.B: Derivation of (4.23) . . . 93

5 MSE Based Transceiver Designs for Bi-directional Full-Duplex MIMO Systems with Residual Self-Interference 96
5.1 Introduction . . . 97
5.2 System Model . . . 98
5.3 Sum-MSE Minimization . . . 102
5.4 Min-Max MSE Minimization . . . 105
5.5 Simulation Results . . . 107
5.6 Appendix 5.A: Vector Forms . . . 111

II On Duality of MIMO Relays 113

6 Literature Survey on Uplink-Downlink Duality 114
6.1 SINR Duality . . . 115
6.2 MSE Duality . . . 116
6.3 Capacity Duality . . . 117

7 On MAC-BC Duality of Multi-hop MIMO Relay Channel with Imperfect Channel Knowledge 119
7.1 Introduction . . . 120
7.1.1 Contributions of Our Work . . . 120
7.2 System Model . . . 122
7.2.1 Multi-hop BC MIMO Relay System . . . 123
7.2.2 Multi-hop MAC MIMO Relay System . . . 126
7.3 Channel Model . . . 128
7.4 MAC-BC Duality . . . 130

7.5 Appendix 7.A: Proof of Theorem 1 . . . 133
7.5.1 Step 1 . . . 134
7.5.2 Step 2 . . . 139
7.5.3 Step 3 . . . 143
7.6 Appendix 7.B: Proof of Theorem 2 . . . 144
7.7 Appendix 7.C: SINR-MSE Relation . . . 146

8 On Uplink-Downlink Sum-MSE Duality of Multi-hop MIMO Relay Channel 148
8.1 Introduction . . . 149
8.1.1 Contributions of This Work . . . 150
8.2 System Model . . . 152
8.2.1 Uplink MIMO Relay System . . . 152
8.2.2 Downlink MIMO Relay System . . . 155
8.3 Uplink-Downlink Duality . . . 159
8.3.1 The KKT Conditions of the Uplink Problem . . . 159
8.3.2 The KKT Conditions of the Downlink Problem . . . 161
8.3.3 Sum-MSE Uplink-Downlink Duality . . . 163
8.4 Numerical Examples . . . 165
8.5 Appendix 8.A: Proof of Lemma 1 . . . 166
8.6 Appendix 8.B: Proof of Theorem 1 . . . 170

9 Conclusion 174
9.1 Summary of Contributions . . . 174
9.2 Future Work . . . 178

Bibliography 180

List of Figures

3.1 The signal flow diagram of a bi-directional full-duplex MIMO system. The node on the left has its transmitter denoted as transmitter 1 and its receiver as receiver 2. The node on the right has its transmitter denoted as transmitter 2 and its receiver as receiver 1. Each of the receivers and transmitters has N antennas. 47
3.2 The signal flow diagram of a two-hop full-duplex MIMO relay system. . . . 62
3.3 Ergodic mutual information comparison of the FD2 and HD systems with different number of antennas versus INR. Here SNR = 20dB, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30$dB. 67
3.4 Ergodic mutual information comparison of the FD2 and HD systems with different channel estimation errors versus INR. Here $N = 2$, SNR = 20dB, $\sigma_t^2 = -30$dB. 67
3.5 Ergodic mutual information comparison of the FD2, FD1, and HD systems versus INR for different SNR_2 values. Here $N = 3$, $\text{SNR}_1 = 20$dB, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30$dB. 69
3.6 Ergodic mutual information comparison of the FD2 and HD systems versus INR for different σ_e^2 and σ_t^2 values. Here $N = 3$, SNR = 20dB. 69
3.7 Ergodic mutual information comparison of the FD2 and HD systems for different INR values versus SNR. Here $N = 3$, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30$dB. . . 70
3.8 Ergodic mutual information comparison of the FD2 and HD systems with different σ_t^2 values versus SNR. Here $N = 3$, INR = 20dB, $\sigma_e^2 = 0.01$. 70
3.9 Ergodic mutual information comparison of the FD2, FD1, and HD relay systems versus INR_1. Here $N = 3$, $\text{SNR}_1 = 20$dB, $\text{SNR}_2 = 10$dB, $\text{INR}_2 = 0$dB, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30$dB. 72

4.1 Sum-rate comparison of the FD2, FD1, and HD systems versus INR. Here $N = 2$, SNR = 20dB, $\sigma_e^2 = 0$, $\kappa = \beta = -40$dB, $\mu = 0.25$. 88
4.2 Sum-rate comparison of the FD2 and HD systems with different σ_e^2 versus INR for individual power constrained problem. Here $N = 2$, SNR = 20dB, $\kappa = \beta = -40$dB and $\mu = 0.25$. 88

4.3 Sum-rate comparison of the FD2 and HD systems with different N versus INR for sum-power constrained problem. Here SNR = 20dB, $\sigma_e^2 = 0.01$ $\kappa = \beta = -40$dB and $\mu = 0.25$. 89
4.4 Sum-rate comparison of the FD2 and HD systems for different INR values versus SNR for individual power constrained problem. Here $N = 2$, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB and $\mu = 0.25$. 90
4.5 Sum-rate comparison of the FD2 and HD systems for different INR values versus SNR for sum-power constrained problem. Here $N = 2$, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB and $\mu = 0.25$. 90
4.6 Cumulative distribution of the sum-rate. Here $N = 2$, SNR = 20dB, INR = 15dB, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB, $\mu = 0.25$. 91

5.1 A bi-directional full-duplex MIMO system 99
5.2 MSE comparison of the Total and MinMax algorithms with different channel estimation errors versus SNR. Here INR = 20dB, $N = 2$, $\kappa = \beta = -40$dB. . 108
5.3 MSE comparison of the Total and MinMax algorithms with different $\kappa = \beta$ values versus SNR. Here $N = 2$, INR = 20dB, $\sigma_e^2 = 0.01$. 109
5.4 MSE of the Total algorithm with different INR values versus SNR. Here $N = 2$, $\sigma_e^2 = 0.01$ and $\kappa = \beta = -40$dB. 109
5.5 MSE distribution of Total and MinMax algorithms. The schemes $1-3$ correspond to Total and the schemes $4-6$ correspond to MinMax algorithms. For each algorithm, the first two bars are the achieved user MSEs and the third bar is the sum-MSE. Here $N = 3$, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB, SNR = 20dB and INR = 10dB. 111

7.1 Multi-hop BC AF MIMO relay system 125
7.2 Multi-hop MAC AF MIMO relay system 128

8.1 Uplink multi-hop AF MIMO relay system. 153
8.2 Downlink multi-hop AF MIMO relay system. 156
8.3 Example 1: MSE versus P. $K = 3$, $M = 2$, $N = 10$, $P_L^{UL} = P_1^{DL} = 20$dB. . 166
8.4 Example 2: BER versus P. $K = 3$, $M = 2$, $N = 10$, $P_L^{UL} = P_1^{DL} = 20$dB. . 167

List of Tables

2.1 Experimental works on full-duplex systems 11
2.2 Theoretical works on full-duplex systems 20
2.3 Performance analysis on full-duplex systems 27

3.1 Procedure of the projected gradient power allocation approach 59

4.1 WSR maximization algorithm 86

5.1 Sum-MSE minimization algorithm 104
5.2 Min-Max MSE minimization algorithm 108

6.1 Existing MAC-BC duality results 116

Chapter 1

Introduction

To meet the growing demand for high data services, low latencies and enhanced bandwidth efficiency in wireless communication systems, FD radios, which can transmit and receive at the same time and the same frequency were proposed recently to increase the spectral efficiency and as a solution to spectrum scarcity problem.

However, the challenge of designing FD radios is that due to simultaneous transmission and reception, the signal transmitted from the transmitter antennas of a FD node is received (coupled) at the receiver antennas of the same FD node, which leads to strong self-interference. The strong self-interference saturates the front-end of the receiver (low-noise-amplifier, mixer, etc.), and reduces the dynamic range of the analog-to-digital converter, which in turn increases the quantization noise for the desired signal. So to make FD radios feasible, the self-interference must be canceled up to a some degree before it saturates the front-end of the receiver. Recent experimental results have designed some analog domain algorithms to cancel the self-interference.

But it is not possible to cancel the self-interference perfectly due to channel estimation errors, transmitter/receiver impairments (oscillator phase noise, the non-linearities in the power amplifiers, etc.), so residual error always exists in the digital domain (baseband) after the analog-to-digital converter (ADC). So to improve the performance of the FD systems, in the first part of this thesis, we have developed efficient transceiver design algorithms for the optimization problems related to FD MIMO systems under this residual self interference in the digital domain.

MIMO relay systems have attracted much attention for next generation wireless communication systems as promising techniques for extending the cell coverage, increasing the system capacity, reducing power consumption (increasing power efficiency) and reducing the overall path loss, since they lead to cooperative diversity and spatial diversity. MIMO relays can be used in both the uplink (from users to a base station) and the downlink (from a base station to users) communication systems.

Duality properties between uplink and downlink communication systems have gained much interest in signal processing and information theory. Due to the coupled structure of the transmitted signals in the downlink channel, the optimization problems associated with the downlink system are usually difficult to solve. The key technique used to overcome this difficulty is to transform the downlink problem into an uplink problem via a so-called uplink-downlink duality relationship. Since the uplink channel has a simpler mathematical structure, less coupling of variables, it is usually more efficient to solve the optimization problems associated with the dual uplink system. Therefore, in the second part of this thesis, we have established the uplink-downlink duality in terms of SINR,

MSE, and capacity for uplink and downlink multi-hop AF MIMO relay channels, which is a generalization of several previously established uplink-downlink duality results.

1.1 Outline of the Thesis

- Chapter 2 summarizes the last five years of research efforts on FD systems, and provides a comprehensive overview of this exciting research field.

- Chapter 3 focuses on the exact and asymptotic closed form ergodic mutual information maximization of FD MIMO bi-directional and FD MIMO relay radio systems under a fast fading channel model. Since the problem is non-convex, a gradient projection algorithm is developed to optimize the power allocation vectors with the knowledge of statistical CSI at the transmitters.

- Chapter 4 addresses the transmit filter design for weighted-sum-rate maximization problem in FD MIMO bi-directional systems subject to sum-power constraint or individual power constraint. An iterative alternating algorithm to find a local optimum was proposed based on the relationship between weighted-sum-rate maximization and weighted-minimum-mean-squared-error minimization problems.

- Chapter 5 focuses on transceiver design for FD MIMO bi-directional systems based on the minimization MSE and the maximum per-node MSE optimization problems subject to individual power constraints at each node through an iterative alternating algorithm, which is proven to converge to at least a local optimal solution.

- Chapter 6 summarizes the works on uplink-downlink duality, and provides a comprehensive overview of this research field.

- Chapter 7 establishes the SINR, MSE and capacity duality between multiple access (MAC) and broadcast (BC) multi-hop AF MIMO relay systems under an imperfect channel state model, which is a generalization of several previously established MAC-BC duality results. Duality is established under both the total power constraint of the system, and individual power constraints at each node of the system.

- Chapter 8 provides an interesting perspective to the relation of the uplink-downlink duality based on the Karush-Kuhn-Tucker (KKT) conditions of sum-MSE transceiver optimization problems for uplink and downlink multi-hop AF MIMO relay channels.

- Finally, Chapter 9 presents the concluding remarks and some directions for future work.

Part I

Performance Limits of Full-Duplex MIMO Radios

Chapter 2

Literature Survey on Full-Duplex Radios

This survey concerns radio frequency (RF) wireless communication systems or simply called radios. A radio can be used as a wireless relay between two other radios, which we call a relay system. Wireless relays have attracted a great deal of attention for next generations of wireless communication systems as they can reduce the overall path loss and transmission power consumption and they also can increase cell coverage and capacity. Two radios can be used to communicate directly with each other, which we call a bi-directional system. Conceptually, most networks require a bi-directional communication channel between communicating nodes.

The notion of a communication channel allowing bi-directional data transfer is called duplexing. For duplex communication to be feasible, interference between transmissions and receptions should be eliminated. Currently, the isolation/orthogonality between

the two directions of communication is achieved using independence in either time (Time-Division Duplexing (TDD)) or frequency (Frequency-Division Duplexing (FDD)), which are named as half-duplex (HD) systems. But the problem of using multiple time slots in TDD, and multiple frequency bands can be mitigated by designing a radio that can send and receive data at the same time and at the same frequency (a single carrier frequency), simultaneously, which is named as full-duplex (FD) systems.

But the problem with designing FD radios is that as signals attenuates (the power of the signal decreases) while propagating through the air, the signal transmitted from its own transmitter antennas of a node is much stronger (around 100dB) than the transmitted signals from other nodes. So the challenge is to eliminate the strong self-interference at the receiver antennas of the node before it saturates the front-end of the receiver. In theory, self-interference cancellation should be easy to solve. In particular, since the node knows the signal transmitted from its own transmit antenna, it can subtract it from the signal received at its receive antenna. Unfortunately, solving this problem in practice is difficult. Strong self interference saturates the front-end of the receiver, specifically low-noise-amplifier (LNA), the mixer and the analog to digital converter (ADC), thus the receiver cannot decode the packet after cancelling the self-interference. Particularly, since the desired received signal is hundreds of thousands of times weaker than the self-interference signal, the output of the ADC may contain no information about the desired signal. The reason for that is that since ADC has a finite number of quantization levels (because of finite resolution), it scales its quantization levels to match the level of the self-interference. As a result, even if all the self-interference is cancelled after the ADC, the receiver will not be able to process the

output of the ADC to decode the intended packet. Therefore, to avoid the saturation of the front-end of the receiver, self-interference should be cancelled (or at least reduced) in the analog domain before it reaches to any component in the receive chain.

Since HD systems require the partition of the time or frequency resources, it is not efficient in terms of the spectral efficiency and the data rate. Therefore, the potential advantages of FD radios over HD radios have recently motivated active research in several different aspects, ranging from information theory and signal processing based on mathematical models to hardware experimentation and real system demonstration [1, 2]. FD systems has the potential to complement and sustain the evolution of 5G technologies and can be utilized in wireless communication systems in multiple ways, including increased link capacity, novel relay solutions, and enhanced interference coordination, etc. [3].

The objective of this chapter is to summarize the last five years of research efforts, so as to provide a comprehensive overview of this exciting research field[1]. Focus will be on experimental self-interference techniques (Section 2.1). Self-interference mitigation techniques using transmit beamforming, multi-user systems and cognitive radio systems under FD model is discussed in Section 2.2. Finally, performance analysis for FD systems is discussed in Section 2.3 and alternative categorizations of the available FD systems will be discussed in Section 2.3.5. We believe that the cited papers (as well as the references therein) will serve as a good starting point for further reading.

[1] A wide array of techniques proposed in the literature for self-interference suppression are surveyed in [1].

2.1 Experimental Works

The possible benefits of FD wireless systems have recently led researchers to explore on designing FD radios. As mentioned above, the main challenge in FD systems is the strong self-interference signal caused by the node‘s own transmission, which makes it impossible to decode the desired signal, since the self-interference saturates the front-end of the receiver, and limits the dynamic range of the ADC in the receiving chain, which in turn increases the quantization noise for the desired signal. Recent experimental results have designed some algorithms to cancel the self-interference and shown that FD system is a feasible option for future wireless communications. The main techniques used to cancel the self-interference are:

- **Antenna separation** is a simple passive self-interference cancellation method, where the self-interference is attenuated due to the path-loss between the transmitting and receiving antennas on the FD node. The more separation between transmitting and receiving antennas, the better cancellation is achieved, but this will increase the size of the wireless terminals.

- **Analog cancellation** is an active cancellation mechanism where a canceling signal is sent through RF chains to cancel the self-interference signal at the analog domain of the receive antenna;

- **Digital cancellation** is an active cancellation mechanism which uses the knowledge of the self-interference signal to cancel the self-interference signal in the digital domain (baseband) after the ADC.

Some analog cancellation techniques use noise canceling chips [4] to subtract the self-interference signal from the received signal [5]. But practical noise cancellation circuits can only handle a dynamic range of at most 30dB [5]. Similarly, self-interference cancellation can be implemented in the digital domain after ADC, but as mentioned before existing ADCs have finite resolution that increases the quantization noise for the desired signal. Motivated by these limitations, antenna cancellation been explored as a self-interference cancellation technique. Antenna cancellation uses the fact that the signal power decreases as the distance between the transmit and receive antennas increases [6]. However, to eliminate the self-interference, one needs large antenna separation, which in turn increases the size of the wireless node impractically. A promising antenna cancellation technique to implement practical FD radios is proposed in [7]. Although promising, the technique in [7] has three major limitations. The first is that it requires three antennas for the self-interference cancellation: two transmit, and one receive. FD doubles throughput, but with three antennas, MIMO HD system can triple throughput. Furthermore, having two transmit antennas creates slight null regions of destructive interference in the far field. The second limitation is a "bandwidth constraint", so it does not support wideband signals such as WiFi. Finally, [7] introduces a third, practical limitation: it requires manual tuning, so it cannot automatically adapt to realistic environments.

All three limitations are addressed in [8], which designs a two antenna (one transmit and one receive) FD radio using adaptive cancellation and signal inversion through a Balun (balanced/unbalanced) circuit, which does not have a bandwidth constraint. Fur-

Table 2.1: Experimental works on full-duplex systems

	Experimental		
	Antenna	Analog	Digital
	[5, 6, 7, 11, 12, 13, 14] [16, 22, 26, 50, 51, 52]	[5]-[10], [15]-[21], [23, 24, 26] [28]-[31], [43, 48, 52]	[6]-[8], [16, 17, 20, 24, 26, 28] [29], [32]-[36], [40, 45, 46, 48]

thermore, [8] presents an automatic tuning algorithm to adapt the realistic environments automatically and quickly. However, Balun method still has a critical bandwidth limitation, since it can only cancel the self-interference perfectly at a single frequency, and the cancellation performance degrades as the bandwidth of the self-interference signal increases. Based on the observation that the delay of the echo path equivalently operates as a phase rotation of the baseband transmit signal, in [9], an analog domain self-interference cancelation scheme for orthogonal frequency-division multiplexing (OFDM) systems is proposed. Thus, the baseband transmit signal is rotated at each subcarrier, and the rotated baseband signal is used to emulate RF self-interference signals for analog cancelation. Under an ideal RF radio, where the effect of RF impairments are ignored, a single channel 20MHz FD OFDM wireless system based on analog domain self-interference cancellation is developed and mathematically analyzed in [10].

Other antenna placement techniques to cancel the interference are also discussed in [11, 12, 13, 14], which are, similar to [7], designed to create a null at a single frequency, and the performance degrades for wideband signals. As a result, beamforming nulling based designs in [7, 11, 12] require the antennas to be either symmetrically spaced or placed in a way that transmitted equal-amplitude and opposite-phased signals cancel each other at the receive antenna. In [13], a compact FD MIMO relay antenna domain isolation technique

is proposed to enhance the isolation using loops for field suppression. [14] proposed the method of self-interference cancellation for FD single-channel MIMO systems based on antenna mutual-coupling model.

The works [5, 7, 8] have assumed single-tap cancellation, which has been shown to have a psuedo-convex optimization surface [8] and the gradient descent algorithms proposed in [5, 7, 8] converge to global optima. However, in practice self-interference channel is emulated using multiple taps. Therefore, the algorithms proposed in [5, 7, 8] are sub-optimal in a multiple-tap environment, and they converge to local optima. Therefore, in [15], an analytical solution for the global optimum is derived, which achieves significant wideband cancellation in multiple-tap self-interference channel in the analog domain. Particularly, an optimal algorithm for tuning the attenuation and phase-shift parameters of a multiple-tap self-interference cancellation technique is proposed.

In [6, 16], a canceling signal is generated by using MIMO RF chains and added at the receive antenna using RF adder in the analog domain. While this technique does not require two transmit antennas as in [7], it also works better at a very narrow bandwidth (0.625 MHz) and the performance degrades for wideband signals. The authors in [17] presented wideband implementation (100MhZ) of the Rice architecture [6, 16]. In [18], the authors implement a real-time OFDM-based FD radio.[2] The open-loop techniques used in [6, 16, 18] largely depends on accurate self-interference channel estimation, and thus are directly sensitive to transmitter/receiver impairments. Since self-interference cancellation is not perfect due to these impairments, additional cancellation methods, like digital domain

[2]In [18], the authors also propose the first FD medium access protocol (MAC), FD-MAC, in multi-node networks.

cancellation and separate transmit/receive antenna have to be incorporated to cancel the residual interference. In [19], without using additional antennas, the authors propose the implementation of an analog domain closed-loop adaptive echo cancellation technique for wideband signals, which is robust to impairments.

In [20], a baseband transmitter output based echo cancellation approach with a 2-stage iterative echo canceller at the receiver is proposed. The reference samples for the self-interference cancellation are obtained from the transmit power amplifier output. The method has been shown to increase the channel capacity by a factor between 1.4 and 1.8, depending on the distance between communicating nodes.

All of the proposed systems above require at least two antennas. Using separate transmit and receive antennas may provide a high level of isolation but the same doubling of capacity (achieved by ideal FD mode) could also be achieved by using the two antennas for transmission or reception as in HD mode. [21] first proposed a single antenna FD system that achieves 40dB self-interference isolation and low insertion loss. Different from multiple antenna FD designs in [7, 8, 11, 22], a novel FD design in [23], Picasso, achieves self-interference cancellation using a single antenna and passive attenuation/delay components which do not leak interference into adjacent spectrum. Furthermore, unlike [8], Picasso does not require constant tuning. But Picasso only allows the radio to simultaneous transmission and reception on different adjacent channels, not simultaneous transmission and reception on the same channel, which is more difficult to deal with. In [24], the authors propose novel analog and digital cancellation techniques that can achieve 110dB self interference cancellation for WiFi signals, and it works with the high bandwidths as 80MHz. Particularly, they

design and implement a FD WiFi radio that uses a single antenna and achieves close to the ideal doubling of capacity under all SNR regimes. But to emulate the self-interference channel, the discrete number of parallel, independently-tunable delay lines is used in [24], which increases the complexity of the FD design. Since the self-interference channel may vary rapidly, the delay lines have to tuned frequently. Hence, the performance of the self-interference cancellation [24] decreases in the analog domain for strongly frequency selective channels.

Furthermore, [7] supports only a single data stream, and can not support multiple streams, which is a requirement for future wireless standards. [25] discusses the possibility of extending FD designs to support multiple streams (or MIMO systems) and support asymmetric throughput on forward and reverse links. [26] proposes the first design for a 20 MHz FD MIMO OFDM WiFi radio, which achieves high rate and extended range, which can support most of the indoor WiFi deployments. The authors also present an integrated physical and MAC design that is compatible with IEEE 802.11x standard, enabling accelerated adoption of FD wireless. In [27], the authors consider a more generalized scenario, and examine the performance of a random-access time-slotted wireless network consisting of a single access point and a mix of HD and FD nodes. Unlike the SISO case in [24], the authors in [28] present the implementation of the first in-band FD MIMO WiFi radio that can achieve the ideal doubling of the capacity, and the complexity scales linearly with the number of antennas. The authors in [29] presents the implementation of "FlexRadio", the first system that enables flexible selection of the number of RF chains for simultaneous transmission and reception dynamically based on network topology.

The experimental works shown in [5, 7, 8, 16, 18] either applied a frequency-domain transmit beamforming method or assumed frequency-flat channels. In [30], the authors present a time-domain transmit beamforming method for broadband self-interference cancellation at the RF front-end, which does not have the prefix-region problem associated with the frequency selective frequency-domain transmit beamforming methods.[3]

In all of the aforementioned works, for example [6, 7, 24], the self-interference is cancelled at the analog RF stage to prevent the saturation of the front-end of the receiver. The work [31] shows that adding a complementary analog baseband self-interference cancellation stage to analog RF stage can significantly improve the performance of total cancellation achieved, and thus enhances the performance of decoding of the desired signal.

In [32], the authors presented efficient algorithms to cancel the self-interference in the digital domain for frequency selective channels. In the proposed method, the canceling signal is fed back prior to ADC, so larger dynamic range and resolution is achieved. A digital domain self-interference cancellation for wideband passband signals is proposed in [33], which is based on adaptive auxiliary transmission.

In [34], an adaptive digital interference cancellation method based on least mean square (LMS) algorithm is proposed. Particularly, the adaptive filter can adjust its coefficients to minimize delay, amplitude and phase offsets between the self-interference signal and the reconstructed signal based on the changes in the self-interference channel. Another adaptive filter that can retrieve the desired signal from the interfered signal in the digital domain by proper selection of the filter parameters is proposed in [35]. The authors in [36]

[3]The theory proposed in [30] also applies for single antenna systems.

propose a robust baseband self-interference cancellation method in the digital domain for wireless OFDM communication systems, based on near- and far-end channel estimation, which have not been considered before.

Moreover, recent works have shown that the radio impairments such as transmitter and receiver phase noise in the local oscillators of the FD node limits the amount of self-interference that can be mitigated [37, 38, 39, 40, 41, 42]. For example, in [37], an analytical model that incorporates the phase noise in the transmit and receive chain, quantization noise, and receiver additive white gaussian noise (AWGN) noise is derived. It is shown that among the three impairments mentioned, oscillator phase noise in the transmit and receive chain is the main bottleneck that limits the amount of self-interference that can be canceled actively. Different from the analytical model in [37], the main reason that limits the performance of FD systems is investigated experimentally in [38], and it is shown, again, that the main bottleneck in the self-interference cancellation is the local oscillator phase noise in the transmit and receive chain. However, a very small phase-noise assumption, separate transmitter and receiver oscillators, and narrow-band signals are assumed in the analysis [38]. In [39], both common shared oscillator and two independent oscillators under wideband signals is considered. A digital-domain self-interference cancellation technique for FD OFDM systems under the transmitter/receiver oscillator phase noise is presented [40]. In [41], the phase noise estimation and mitigation in FD systems under the oscillator phase noise at the transmitter and receiver side has been examined analytically and experimentally. In [42], the performance of a amplify-and-forward (AF) FD OFDM repeater link in the presence oscillator phase noise resulting in inter-carrier interference is investigated. Two

different repeater designs are considered for comparison: One using a single oscillator signal for down- and up-conversion; and the other employing two separate oscillators. In [43], the authors present a new analog-digital hybrid method whose performance is no longer limited by the transmission noise. This method also involves a blind system identification and equalization algorithm for finding the optimal parameters of the cancellation filter.

In addition to phase noise, the amplifiers and mixers cause nonlinear distortion, which can limit the amount of self-interference cancellation. For example, it is shown in [44] that, due to the large power difference between the self-interference and the desired signal, nonlinear distortion in the transmitted and received signals becomes one of the key factors that limit amount of self-interference that can be canceled. The effects of nonlinear distortion on the performance of self-interference cancellation has been ignored in most of the existing works. Recently, the effects of nonlinear distortion on the self-interference cancellation for a FD node, and its compensation, have been studied in [24, 45, 46]. Moreover, in [47], a comprehensive analysis and detailed system calculations of the effects of the nonlinear distortion in the transmitter power amplifier (PA) and the receiver chain of the FD node is analyzed. In [48], the performance of FD MIMO OFDM transceivers with subtractive self-interference cancellation in analog and/or digital domain is analyzed under the non-ideal ADCs. In particular, the trade-off between ADC resolution, maximum transmit power, minimum physical isolation and sufficient signal to self-interference ratio needed to avoid receiver saturation is investigated.

In [49], thermal noise, one of the factors that effect the performance of the self-interference cancellation, is analyzed for wideband FD wireless system, and a power differ-

ence method to mitigate the impact of thermal noise on the FD OFDM Radio is proposed.

It is not known whether passive suppression will also encounter fundamental bottlenecks (phase noise, nonlinear distortion, etc.) mentioned above and what those bottlenecks may be. Thus, the performance of self-interference cancellation for FD systems can be improved by examining and understanding the limitations of passive suppression techniques. Therefore, directional isolation and cross-polarization are used in [50] to suppress the self-interference in a passive way for outdoor-to-indoor relay systems.[4] The authors in [22] use directional antennas for passive suppression. Moreover, in [51] a measurement-based study is conducted on the performance of the three passive suppression techniques: directional isolation, absorptive shielding, and cross-polarization. The study demonstrates that more than 70dB of passive suppression can be achieved in certain environments.

Prior works, except [30], have assumed that the transmission and reception at the FD node are synchronized, so that channel estimation for the self-interference channel and the channel with the other nodes can be done orthogonally in time. However, in random-access networks, each node makes the decision of a packet transmission independently, so the packet transmitted by a node do not need to be synchronized with the packet received. In [52], the authors study two FD asynchronous communication techniques (i) where the start of packet transmission of a node precedes start of the packet reception from the same node, and (ii) where the start of packet reception precedes the start of packet transmission. And it is shown that the first technique performs better, and performs close to the performance of synchronous FD.

[4]The first channel measurements for FD MIMO relays are also reported in [50].

2.2 Theoretical Works

In addition to experimental results discussed in Section 2.1, there have been also some analytical studies that show the benefits FD systems over HD systems in the presence of self-interference.

2.2.1 Self-Interference Cancellation

The effect and cancellation of self-interference was first studied in analog repeaters [53], [54]. The optimized gain control to mitigate the effect of self-interference is first introduced in [55] and the benefit of limiting the gain at the relay station depending on the self-interference level was demonstrated.

While the self-interference cancellation is limited mostly to subtractive cancellation in single-input-single-output (SISO) FD systems, the increased degrees of freedom (spatial diversity) offered by MIMO systems result in a range of new solutions to mitigate the self-interference, mostly using transmit beamforming (spatial suppression) techniques. In [56], the authors propose to direct the self-interference of a decode-and-forward (DF) relay in the FD mode to the least harmful spatial dimensions, which is an extension of [55] to MIMO systems. The authors of [57] analyze a wide range of self-interference mitigation techniques when the relay has multiple antenna, including natural isolation, time-domain cancellation and spatial domain suppression (antenna subset selection, null-space projection, and minimum mean-squared error (MMSE) filtering). The techniques in [57] apply to general protocols including AF and DF. Null-space projection and MMSE filtering are

Table 2.2: Theoretical works on full-duplex systems

	Theoretical		
	Theoretical Cancellation	Multi-User	Cognitive Radios
	[55]-[82]	[83]-[96]	[97]-[102]

also considered for spatial self-interference suppression in [58, 59, 60]. In [61], the authors consider spatial-domain suppression using optimal eigenbeamforming that minimizes the power of the residual self-interference signal by pointing the transmit and receive beams to the minimum eigenmodes of the self-interference channel, which covers the null-space projection [57]-[60] as a special case. These spatial domain interference nulling techniques are possible when there are enough degrees of freedom (d.o.f.). In [62], a scheme maximizing the signal-to-interference ratio at the relay input and output is proposed, which may be effective when the d.o.f. are not enough. In [63], interference nulling algorithm is performed by an optimization of the relay processing vectors over the continuous domain, which was shown to have better performance than the singular value decomposition based method [57]. In [64], the authors study spatial self-interference cancellation methods usable for MIMO space division duplexing eigenmode transmission. In [65], the authors propose an adaptive gradient descent method to mitigate the self-interference signal for FD DF MIMO relays which can track temporal variations of the self-interference channel. In [66], two self-interference suppression schemes for cellular FD communication systems are proposed: zero forcing (ZF) precoding, extended regularized channel inversion (RCI), which can jointly leverage the advantages of large-scale MIMO and FD communication. Spatial-domain suppression decreases the degrees of freedom (multiplexing order), on the

other hand time-domain cancellation leads to residual self-interference, and thus increased decoding error. Therefore, the choice (trade-off) between two self-interference mitigation techniques: spatial-domain suppression, time-domain cancellation were considered in [67] for bi-directional systems.The authors in [68] consider the same problem in [67] under transmitter noise as the number of antennas grows large.

Some other spatial domain solutions that exploit the multiple antennas at a FD node have been proposed so that FD systems can eliminate or avoid the interference and improve the link quality [69, 70, 71, 72, 73]. Space-time equalization [69], precoding/decoding [70], or eigenbeam selection [71] techniques were applied for self-interference mitigation.[5] [72] designed transmit and receive filters for FD relay systems in order to cancel the self-interference and double the achievable rate of the HD relay system by using the orthogonal complement concept. Using the same system model in [63], the usage of asymmetric complex Gaussian signaling in FD DF relay systems is proposed to cancel the self-interference signal and increase the throughput [73] by searching for the optimum weights that increase the smaller signal-to-noise ratio (SNR) between source-relay and relay-destination link such that the self-interference is perfectly nulled. In [74], an overview of beamforming and power allocation for both FD and HD MIMO relays operating in DF or AF mode are provided.

Most of the self-interference cancellation techniques for FD relays assume an instantaneous self-interference channel and apply spatial processing to mitigate the self-interference. However, in practice, a non-negligible delay always occurs because of the

[5] In [70], the authors propose a new FD system that shares antennas at the relay node for both transmission and reception, rather than separating the antenna resources into transmit and receive antenna sets.

analog transmit/receive filters in the front-ends of relays, so spatio-temporal mitigation techniques should be applied to cancel the self-interference. By exploiting knowledge of the autocorrelation of the useful signal, a blind adaptive feedback cancellation method for MIMO AF FD relays has been proposed in [75], which can mitigate the self-interference for frequency-selective channels as well as perform the channel equalization of the source-relay channel. This extends previous work on SISO [76] and MISO [77] relays to the more general MIMO case.[6] However, one drawback of the design from [76, 77] is that the adaptation of the beamformer requires the knowledge the angle of arrival (AoA) of the incoming source signal. So by extending the methods from [76] and [77], the authors in [78] presents a spatio-temporal adaptive feedback suppressor which does not require AoA knowledge.

Coupling wave cancellation schemes using adaptive filters have been proposed in [54, 79], however, the schemes in [54, 79] require small gains at the relay amplifier in order to have stable adaptive filters. In [80], taking advantage of the fact that the coupling waves to be cancelled at the relay node consist of its own transmitted signals, a beamforming method using not only received signals at actual antenna elements but also virtual coupling wave paths (generated with finite impulse response (FIR) filters) is proposed. The required degrees of freedom to cancel the coupling waves can be achieved by only increasing the number of FIR filters at the relay node without the need to increase the number of antennas. This approach is effective with the small number of paths of between the base station and the relay node. So, in [81], a beamforming method for multi-path FD relay systems

[6]The authors in [76] have presented a blind, second-order statistics based, low complex adaptive feedback cancellation technique for AF FD relays, which can restore the spectral shape of the desired signal effectively. This was extended in [77] to relays with multiple receive antennas (MISO), that can combine the temporal and spatial processing effectively to improve the performance of the self-interference cancellation.

with frequency domain equalization (FDE) based transmission is proposed. Particularly, the relay node uses a blind coupling wave canceller for each antenna that controls the beam weight in order to maximize the SNR at the output of the FDE. In [82], a MMSE based tapped filter is employed at the destination node to mitigate the self interference, where a blind channel estimation used to compute the optimum filter weights, is performed at the destination.

2.2.2 Multi-User

More recently, FD has been studied in the context of multiuser MIMO relay systems, where FD relays are used to improve the cell coverage and the cell-edge throughput [83]. Although most of the works on the FD have focused on self-interference cancellation, a multiple user MIMO relaying scenario has been studied in [84]-[93]. In [84], block diagonalization (BD) based precoding and power allocation is proposed for FD MIMO relay systems, where each relay node receives data streams from the base station (BS) and transmit the decoded data streams to mobile users (MSs) simultaneously. In [85] and [86], the BS serves both the MSs inside its coverage and the FD relay through multiuser MIMO transmission using BD. On the other hand, the FD relay simultaneously helps the BS forward its data to the MSs out of the BS coverage, which in turn cause interference on the MSs inside the BS coverage, and thus the FD relay was designed to deal with both the interference to the MSs within the BS coverage and the self-interference at its own receive antennas.

A centralized approach has been considered in [84]-[86], in which the BS has the

global channel state information, and based on this knowlege, it computes the its own and relay node's beamforming vector, which is fed to the relay node. Unlike the centralized approach, in [87], a distributed FDR relay is presented for multiuser MIMO relaying systems based on the preliminary results in [88].

Moreover, in [86], a single relay case is considered. So in [89], it is extended to multiple MIMO source, FD relay, destination nodes, where zero forcing approach for single stream transmission is employed. In [90], the authors study self-interference cancellation techniques for MIMO multiuser systems in terms of the achievable sum rate. Capacity of FD MIMO downlink relay system under BD precoding is analyzed in [91], a power allocation method for the multiuser MIMO relay system based on BD is proposed in [92].

In [93], the authors consider the total throughput and energy efficiency maximization problem under a sum power constraint in the downlink channel (DL) and per-user power constraints in the uplink channel (UL) for FD multiuser system. An iterative joint optimization approach that simultaneously optimizes linear precoders of the DL channel and a power allocation strategy of the UL channel using a convex relaxation method is proposed.

The authors in [94] study the network capacity of FD bi-directional communication systems in ad-hoc networks under the delay and outage constraints. It is concluded that the increase in the network capacity through FD mode outweigh the increase in interference caused by the bi-directional use of spatial resources.

In [95], cellular (multiuser) FD networks is considered in which UL and UL operate simultaneously, and the authors identify the scenarios where the interference from UL users

to DL users can be managed to increase the degrees of freedom over the conventional HD cellular networks. In [96], the authors derive the achievable transmission rate of a cooperative wireless network in which users transmit signals to a common destination through a FD AF and FD DF relay node, and an optimal transmission scheduling algorithm is also presented to maximize the minimum transmission rate among a set of users.

2.2.3 Cognitive Radios

A key challenge in two-tier cognitive radio networks (CRN) is the ability of secondary users (SU) to sense and learn the activity of the primary users (PU). With HD radios, once the SU begin to transmit, they cannot learn about the PU activity till they have finished their transmission. As a result, SU transmissions have to be very conservative to ensure due protection to ongoing PU transmissions.

Recent work [97] studies the employment of FD CRN that are able to transmit and sense simultaneously by using the antenna cancellation technique proposed in [7]. As it is mentioned in Section 2.1, the main disadvantage of the scheme in [7] is that it requires manual tuning of RF circuits, so it cannot automatically adapt to realistic environments, which is not practical. Therefore, in [98], the authors study the increase in the achievable rate and transmission range in FD CRN using directional multi-reconfigurable antennas [6] to enable simultaneous transmission and sensing.

In [99], a continuous time Markov chain is used to develop a FD spectrum sensing technique for non-time-slotted CRN. The probabilities of detection and false alarm were derived with random arrival/departure of primary users' traffic and the effect of bandwidth

and antennas placement error on the performance of non-time-slotted CRNs is analyzed. In [100], the authors try to improve the performance PU detection and SU throughput, where the SU can perform the self-interference cancellation to transmit and receive/sense simultaneously. In [101], an optimal power allocation scheme that minimize the outage probability in FD relay CRN is presented and in addition the outage probability of the SU in the noise-limited and interference-limited environments is analyzed.

In [102], the authors study the achievable primary-cognitive rate region by focusing on the cognitive rate maximization problem. Particularly, a primary system and a cognitive system in a cellular network cooperates in the sense that the FD cognitive BS help the primary system relay its data using AF or DF modes, and in return it can transmit its own cognitive signal. It is shown that the cooperation between the primary and cognitive system substantially increases rate region (system spectral efficiency) and the opportunities for a SU to access the primary system.

2.3 Performance Analysis

In addition to the investigation of signal processing techniques in the presence of self-interference discussed in Section 2.2, the performance evaluation of the FD techniques in terms of outage probability, diversity, etc. becomes an active research area.

Table 2.3: Performance analysis on full-duplex systems

	Performance Analysis	
	Transceiver Design	[103]-[126]
	Outage	[120], [141]-[152], [157], [161, 162, 164, 166]
	Diversity	[137], [141], [153]-[159], [163]
	Relay/Antenna-Selection	[120], [153], [160]-[167]
	Other	[168]-[188]

2.3.1 Transceiver Design

It is shown in [103] that the FD relays under self-interference is indeed feasible and can provide higher capacity than the HD systems. The the average end-to-end capacity of the HD and the FD AF relays in a SISO frequency flat channel without a direct link is compared in [104] and the corresponding rate-interference trade-off in [105]. The studies on SISO relays [104, 105] demonstrate that FD systems can perform better than HD systems and improve the system spectral efficiency even in the presence of residual self-interference, and can achieve theoretical doubled symbol rate of the HD counterparts provided that residual self-interference is not strong. This motivates to develop and analyze the FD MIMO relays. In [106], the results in [104] are extended to the MIMO channels, in which the optimal relay transformation matrix that maximizes the capacity of the channel between the source and the destination under average power constraint is derived and find the conditions where FD outperforms HD in terms of SNR. The same authors derive the optimal FD AF relay under the presence of LNA and ADC limitations [107]. A non-convex quadratically constrained quadratic programming (QCQP) problem is formulated and the closed form optimal solution is derived by reducing and partitioning the constraint set. Furthermore,

unlike [104], which has not considered DF relays and spatial diversity in their analysis, [108] investigates the performance of FD relays for AF and DF with spatial diversity.

While most of the theoretical works on the FD MIMO systems studied the moderate interference levels which do not saturate the front-end of the receiver, the performance of the FD systems is limited by the quantization error caused by the limited dynamic range of the ADC. By exploiting both spatial and temporal freedoms of the source covariance matrices of the MIMO links, the authors of [109, 110] maximize the lower bound of the achievable rates for FD MIMO relay channels and FD bi-directional MIMO channels for slow fading channels using gradient projection (GP) method under transmitter and receiver distortions, respectively. The work in [111] later extended the findings in [110] to fast fading channels, where instantaneous channel state information (CSI) is not known at the transmitters. In the absence of instantaneous CSI at the transmitting nodes, the knowledge of some statistical properties (mean, variance) of the CSI is used for designing optimal power schedules. Using the same system model in [110], the authors in [112] consider the weighted sum-rate (WSR) maximization problem subject to total power constraint of the FD system. Based on the relationship between WSR and weighted minimum-mean-squared-error (WMMSE) problem, a low complexity iterative alternating algorithm is proposed. The technique in [109, 110] requires global CSI knowledge. However, in practice it is difficult to obtain the self-interference channel of the other node accurately because of the heavy feedback information required by the large dynamic range. Therefore to resolve this feedback problem, in [113], the authors consider a low complexity iterative distributed source covariance matrix design to maximize the sum-rate in the absence of the knowledge of the

other node's self-interference channel for bi-directional FD MIMO systems.

Moreover, the authors in [114] worked on the optimal transmitting filter design strategies that maximize the achievable sum rate of the MIMO and MISO bi-directional systems, and a two-fold sum rate gain compared to HD system is achieved under small residual self-interference. The same authors extend this bi-directional problem to an AF FD relay system in [115], and propose different convex optimization based suboptimal schemes for different system settings. Since sum-rate maximization problem is non-convex [114]-[115], the same authors propose a sub-optimal solution which maximizes the signal-to-leakage-plus-noise ratio [116]. The same authors also study robust transmit strategies in [117] to minimize the total transmit power of a FD bi-directional MIMO system subject to self-interference constraints and total SINR requirements under a deterministic imperfect channel error model. The authors in [118] study the impact of residual self-interference on sum-rate performance under two situations: CSI is available only at receiver, and CSI is available at both transmitter and receiver. In [119], the instantaneous achievable rates for the FD and HD two-way AF relay systems are derived and compared. In two-way FD AF relay system, all the nodes, including the two source nodes are assumed to operate in the FD mode. Optimal transmit power allocation is further studied to increase the instantaneous achievable rate. Most of the work in the literature does not optimize the precoding and decoding filters jointly, which has resulted in suboptimal end-to-end performance. Hence, in [120], the authors consider joint design of precoding and decoding at the source, the relay and the destination in order to maximize the achievable rate. Exact as well as asymptotic expressions for the outage probability were derived which give insights about the diversity

order and the array gain.

The authors in [121] consider MMSE based FD systems, in which optimal beamforming matrices at the source and destination node, and optimal relay processing matrices are computed under the transmit power constraints at the source and relay nodes. The relays in [121] were not arbitrarily distributed. Hence, the authors in [122] extend the same problem in [121] to AF MIMO HD/FD wireless distributed relay networks.

While the are so many efforts on the cancellation of the self-interference, FD has not been studied in detail for network protocols. To this end, the authors in [123] study a quality-of-service (QoS) based power allocation for a bi-directional FD systems, in which a resource allocation scheme for FD and HD relay networks in [124] is proposed. Unlike the sub-optimal QoS driven dynamic power allocation scheme in [123], three optimal dynamic power allocation schemes are proposed in [125] to maximize the achievable sum-rate of the FD bi-directional systems. It is shown that the optimal capacity with FD can not always achieve the two-fold gain compared to the HD mode, so it may be better to employ a hybrid transmission mode than using only FD mode. In [126], a method to compute the achievable gains (average sum rate and average network congestion) at the network layer is provided by using a network utility maximization framework when the nodes in the network apply self-interference cancellation techniques with different degree of accuracy.

Since the theoretical self-interference signal model in [109, 110] is intractable and requires approximations, self interference signal is modeled as an increase in noise floor in [127]. In [127], FD and HD systems are compared in terms of the capacity under the constraint that the total amount of analog radio hardware is bounded. Under this constraint,

it is not clear if these radios should be used to increase the MIMO multiplexing gain, or should be used to cancel the self-interference. It is shown in [127] in some practical SNR regions using these radios to cancel the self-interference is more beneficial since the resulting FD system after the self-interference cancellation performs better than the HD systems and almost always outperforms HD in large SNR regions. To perform the self-interference cancellation, FD mode requires additional resources such as antennas and RF chains. In [128], using a realistic residual self-interference model adopted from [16], the achievable rates and degrees of freedom for MIMO HD/FD DF are computed for two cases: The relay has the same number of antennas and the same number of RF chains as in the HD mode. An upper bound on the the performance of the FD systems can be obtained under perfect self-interference cancellation as in [129], which is very optimistic in practice.

As mentioned above, theoretical self-interference signal model in [109, 110] is intractable and requires approximations. In addition, all radio impairments are combined together in one system parameter, which simplifies system, but makes it difficult to identify the performance loss associated with specific transmitter impairments. In [37], for SISO narrow-band FD systems, the authors analytically study the achievable rate gain region as a piecewise linear approximation in log-domain, which is shown that to closely match the exact region under a signal model capturing the performance loss effects of oscillator phase noise, LNA, mixer noise, and ADC quantization noise. A local-area FD cellular radio system which serves the UL and DL channels simultaneously by frequency reuse is considered in [130]. In particular, achievable rate regions of the UL and DL were studied for arbitrary channel inputs and interference in the large system limit by using the replica method. The

ergodic capacity of bi-directional FD systems under channel estimation errors is considered using two combining schemes: maximal-ratio combining and optimum combining is investigated in [131].

Information-theoretic models for the self-interference have not been studied in detail in the literature. Recently, in [132], the deterministic approximate capacity method has been employed for a FD relaying system with an unknown self-interference channel gain. In contrast, [133] random coding methods have been employed for a FD Z–channel with side information.

Because of the trade-off between the self-interference in the FD mode and rate loss in the HD mode, the combination of both FD and HD modes is an efficient approach to improve the performance of the FD and HD systems alone [134, 135, 136, 137]. A hybrid scheme that combines the FD and HD systems for a relay channel is proposed in [134]. The duration of each mode is optimized, and has been shown to achieve higher end-to-end throughput than that of simple switching between FD and HD modes. In [135], single-user hybrid FD-HD transmission scheduling [134] is extended to multi-user hybrid FD-HD transmission scheduling, and has been shown to achieve higher end-to-end throughput than the equal opportunity scheme. A dynamic selection between FD/HD with AF/DF relaying for an orthogonal frequency division multiple access (OFDMA) system under a general optimization problem is studied in [136]. In [137], the authors study the switching boundaries between FD and HD modes with transmit power adaption to maximize the instantaneous and average spectral efficiency for both AF and DF relaying systems. However, the analysis in [137] considered only a single link. In [138], the authors study the performance of FD

multi-hop networks in an indoor environment under practical and realistic self-interference level. FD and HD multi-hop relaying in terms of coverage range is compared in [139], and the effects of FD mode on delay and throughput in interference-limited multi-hop networks under the absence of CSI at transmitters is studied in [140].

2.3.2 Outage Performance

Bit error rate (BER) analysis for a FD relay system under different self-interference statistics and binary phase shift keying (BPSK) modulation is studied in [141]. The optimal duplex selection based on the outage probability of a FD DF relay scheme with a cooperative direct link is investigated in [142]. The authors in [143] propose a physical layer network code for SISO FD two-way relay channel, where BER is derived for all the FD nodes (two sources and the relay node). In [144], the outage probability of a multi-hop FD DF relay system has been derived, and the performance of FD and HD systems are compared under a metric named path-loss-to-interference ratio (PLIR). In [145], the closed form expressions for outage probability for both FD and HD relay systems are derived under Nakagami-m fading channels. The authors in [146] derive the closed-form approximate expressions for the outage probability and throughput of a FD Block Markov relaying scheme under independent non-identically distributed Nakagami-m fading. In [147], the outage probability of a FD AF relay system with a MMSE decision feedback equalizer at the destination is studied under a a multi-tap channel created by the direct link and self-interference. But, no relay power optimization was addressed in [147]. Therefore, a relay transmit power optimization scheme is proposed in [148] to minimize the outage probability of FD DF

cooperative relay system. In [149], additional power savings is obtained by an incremental selective DF protocol, while keeping the same outage performance as in [148]. The outage performance of an FD DF system with block Markov encoding, is studied in [150]. The outage probability and system throughput of two-way HD and one-way FD relay systems based on physical-layer network coding is compared in [151], and show that one-way FD relay system can perform better than the two-way HD relay system. The outage probability and ergodic capacity performance of two-way FD AF relay channels are derived in [152].

2.3.3 Diversity

The achieved diversity gain for the FD relay systems is discussed in [137, 141, 153]. In, the authors analyze the error performance of a FD AF cooperative network is analyzed in [141] and shown to have an error floor and thus a zero diversity gain. In order to increase the diversity gain, the authors in [137] propose a hybrid relaying scheme that dynamically switches between FD and HD operation, and this hybrid scheme is combined with opportunistic relay selection in [153].

In [154], two distributed linear convolutional space-time coding schemes are proposed for FD asynchronous cooperative communications: For the complete self-interference cancellation and for partial self-interference cancellation. And both schemes have been shown to achieve full asynchronous cooperative diversity. In [154], the cross-talk interference does not exist, because only one relay node is considered. In [155], a partial distributed linear convolutional space-time coding is proposed for a cooperative network with two FD AF relays, in which loop interference and cross-talk coexist. Unlike the symbol-by-symbol

transmission, FD precoding techniques to ensure diversity (at least one) for single antenna relay networks in block Rayleigh fading channels are presented under a practical and real imperfect self-interference cancellation in [156]. The authors in [157] study the outage probability and the diversity-multiplexing trade-off for a distributed Alamouti code implementation of a FD cooperative network.

Distributed antenna systems (DAS) have attracted much attention to enhance the spectral efficiency and coverage. In [158, 159], a fiber-connected DAS is used as a distributed relay antenna system (DRAS) to support FD transmissions with its self-interference cancellation capability. The throughput and the diversity order of the DRAS are analyzed for Rayleigh fading channels, and it has been shown that FD DRAS outperforms HD relay systems in terms of throughput with large number of antennas at the relay node and at high SNR regions.

2.3.4 Relay/Antenna Selection

The closed-form expressions for the average capacity and the BER are derived in [160] under a relay selection technique based on the best instantaneous SNR over two hops for a FD DF multi-relay network. This relay selection technique assumes a constant self-interference and global CSI knowledge, which result in high power consumption and additional network delays. Therefore, the closed-form expressions for outage and average capacity under partial relay selection based on the best instantaneous SNR over only one-hop (the source-relay link) have been derived for FD AF and FD DF relay systems in [161] and in [162], respectively. In [153], the exact and approximate expressions of the

outage probability under different relay selection algorithms based on the the assumption of availability of different instantaneous information for FD AF cooperative communication systems are derived. Relay selection is also studied in [163] for FD AF cooperative systems, and has been shown to result in a zero diversity order despite the relay selection process. Apart from relay selection, relay location optimization is an effective technique to enhance the system performance. In [164], the relay location optimization based on the minimization of the outage probability for FD DF relay systems is proposed.

The complexity of implementing MIMO radios can be significantly reduced with antenna selection (AS) technique, which requires fewer RF chains than antenna elements. Four low complexity AS schemes at the transmitter side of the relay node to mitigate self-interference are proposed in [165]. In [166], several low complexity AS schemes for MIMO FD relay systems are proposed and the outage performance of these schemes is analyzed. In [120], the outage probability of several optimal and sub-optimal AS schemes that maximize the end-to-end SNR at the destination node are studied. In [167], a transmit-receive antenna pair selection based on two system performance criteria: Maximum sum-rate, and minimum symbol-error rate is proposed for bi-directional FD communication systems.

2.3.5 Other issues

In this section, we discuss the other systems that FD is applied, like femtocells, multi-cells, cross-layer optimization, etc.

FD and HD systems have been compared mostly in terms of outage probability and

sum-rate to evaluate the performance of the system on physical-layer, but these works have not considered the performance of the systems on network-layer, including queue states. Thus, the authors in [168] compared the queuing performance of FD and HD systems under Finite-state Markov chain models with adaptive modulation, and have shown that FD scheme outperforms HD scheme in the low-SNR region but is outperformed by HD scheme as the SNR increases.

Recently, physical layer secure transmission schemes have gained interest to prevent eavesdropping and security attacks by a malicious user. In FD mode, an eavesdropper can act as both a jammer and a classical eavesdropper. In [169], a multi-antenna FD active eavesdropper which can jam and eavesdrop simultanesously is considered. The active FD eavesdropper optimizes its beamforming weights to minimize the achievable secrecy rate of the system under self-interference. In [170], a FD destination node that can jam and receive simultaneously to improve the secrecy rate performance is investigated. In [171], the effect of jamming on the design of two-hop cooperative FD/HD AF relaying is examined. The authors in [172] consider the sum secrecy rate in FD wiretap system under an imperfect CSI model subject to individual transmit power constraints.

Harvest-use (HU) is an energy harvesting (EH) technique where the received energy cannot be stored and must be used immediately to increase the lifetime of energy-constrained networks. [173] has considered the maximization of channel capacity of FD/HD AF cooperative relaying systems where the relay nodes deploy a HU architecture. The closed form expression of optimal time division between EH and relaying transmission that maximize the channel capacity is derived. It has been shown that FD mode is a better choice

for HU-based systems, since it offers more efficient balance between EH and relaying time than HD mode.

In [174], the authors derive the closed-form expressions of the outage probability and spectral efficiency a FD relay within a cooperative femtocell underlaid in a macro-cell network. It has been shown that FD mode improves the spectral efficiency of the nodes inside the femtocell compared to the HD mode.

In [175], the authors propose distributed FD architecture via wireless side channels for interference management. Particularly, a three-node FD SISO network, where a BS serves UL and DL channels simultaneously, is considered. An orthogonal wireless side channel between the UL and DL channels is used to reduce the impact of interference from UL to DL channel to increase the multiplexing gains. The SISO case in [175] is extended to the MIMO three-node FD network in [176], where the impact of interference from UL to DL on diversity-multiplexing trade-off (DMT) under wireless-side-channel is studied. Similarly, the same authors extend the application of wireless side-channels to a larger multiuser interference network to characterize the generalized degrees of freedom per user per antenna of the system in all SNR regimes [177].

In [178], the authors study the theoretical performance of a FD multi-cell model based on stochastic geometry and show that FD multi-cell systems increase capacity over traditional cellular systems.

The authors in [179, 180] study the cross-layer optimization to choose distributed end-to-end routes that maximize the total profit of users, and that minimize the network power consumption subject to rate constraints in FD ad-hoc networks. In [181], a MAC

technique that takes advantage of the features of FD radios to eliminate the hidden terminal problem and to improve the efficiency WLAN is proposed. The nodes in [181] are fully hidden or fully conflicting, so the interference the nodes experience has not been considered. A more efficient MAC protocol for FD networks that allows partially interfering nodes to cooperate based on interference levels they experience is proposed in [182]. Moreover, when secondary packet transmission is not present, a busy tone is transmitted in [181] to prevent the hidden terminal problem, which wastes the consumption of energy. Therefore, a MAC protocol which improves the energy and bandwidth efficiency for FD radios was proposed in [183]. In [184], an asynchronous MAC protocol for FD multi-hop networks is proposed. The joint routing and power allocation problem for a wireless FD network under residual self-interference is studied in [185]. And in [186], a routing protocol to reduce the hidden terminal problem in FD multi-hop networks is proposed. The authors in [187], propose a MAC protocol for MIMO FD wireless to increase the opportunity to adopt to the MIMO FD operations in WLAN is proposed.

In [188], the authors study the system throughput and packet delay of a token-based MAC scheme for unmanned aerial vehicle ad-hoc networks with FD radios and capability of multipacket reception. The MAC scheme has been formulated and solved under perfect and imperfect CSI as a combinatorial and a discrete stochastic optimization problem, respectively.

Chapter 3

Achievable Rates of Full-Duplex MIMO Radios in Fast Fading Channels with Imperfect Channel Estimation

This chapter studies the theoretical performance of two FD MIMO radio systems: a FD bi-directional communication system and a FD relay system. We focus on the effect of a (digitally manageable) residual self-interference, imperfect channel estimation (with independent and identically distributed (i.i.d.) Gaussian channel estimation error) and transmitter noise. We assume that the instantaneous channel state information (CSI) is not used for the transmitters and an imperfect CSI is used for the receivers. To maximize

the system ergodic mutual information, which is a non-convex function of power allocation vectors at the nodes, a gradient projection algorithm is developed to optimize the power allocation vectors. This algorithm exploits both spatial and temporal freedoms of the source covariance matrices of the MIMO links between transmitters and receivers to achieve higher sum ergodic mutual information. It is observed through simulations that the algorithm reduces to a FD scheme when the nominal self-interference is low, or to a HD scheme when the nominal self-interference is high. In addition to an exact closed-form ergodic mutual information expression, we introduce a much simpler asymptotic closed-form ergodic mutual information expression, which in turn simplifies the computation of the power allocation vectors.

3.1 Introduction

This thesis concerns radio frequency (RF) wireless communication systems or simply called radios. A radio can be used as a wireless relay between two other radios, which we call a relay system. Two radios can be used to communicate directly with each other, which we call a bi-directional system.

Wireless relays have attracted a great deal of attention for next generations of wireless communication systems as relays can reduce the overall path loss and transmission power consumption and they also can increase cell coverage and capacity. A conventional wireless relay is HD, which transmits and receives using two different channels (in time or frequency). A FD relay can transmit and receive using a single frequency at the same time and is more spectrally efficient [103, 104].

Bi-directional communication is commonly required in virtually all modern communication systems, where two terminals exchange information with each other. Currently, all bi-directional systems are HD, which requires two different channels for two opposite directions. A FD bi-directional system uses a single frequency at the same time for both directions and is twice as spectrally efficient [68, 131].

Among the earliest works on FD radio is studied in [189], where a narrowband (200KHz) FD radio testbed was reported. This research effort stayed almost dormant until the work [69] published ten years later. It was then followed by the hardware-based research activities as well as the theoretical research activities mentioned in Chapter 2.

A fundamental enabler for FD radios is known as the self-interference cancelation. When a FD radio transmits, it causes self-interference which must be canceled satisfactorily. The cancelation can be done by different methods, to different degrees, and at different stages along the receiving chain of a FD radio. Cancelation of interference before the interference-corrupted signal is digitized is called analog cancelation. One important advantage of analog cancelation is that the desired (weak) signal from a remote radio will be less saturated with the receiver noise (including the receiver quantization noise).

A simple testbed for analog cancelation was reported in [7] where two transmit antennas were used to create a null at a receive antenna. A demonstration of analog cancelation using an analog circuit was shown in [8]. Analog cancelation using real-time channel estimation was reported in [16]. Analog cancelation for a single antenna used for both reception and transmission was demonstrated in [21, 23, 24, 30]. The works shown in [7, 8, 16, 21, 23] assume that the interference channel is allpass. Broadband analog

cancelation for frequency-selective interference channels was demonstrated in [18, 24, 30]. The amount of cancelation demonstrated on hardware varies and depends on many possible factors in the hardware systems.

The theoretical works shown in [56, 57, 58, 59, 61, 62, 63, 74] all exploit multiple antennas for analog interference cancelation. The key idea among all these theoretical works for analog interference cancelation is based on a well-known concept of array processing, which is often referred to as transmit beamforming. The basic idea of this approach is that the self-interference can be cancelled at the front-end of the receiver by generating a cancellation signal based on the transmit signal in the baseband.

We assume that an (imperfect) analog interference cancelation or passive suppression has been implemented in the FD radios and the residual self-interference can be handled digitally in the baseband. We focus on a theoretical performance of the FD radios under the effect of the residual self-interference. The contributions shown in this chapter are closely related to [109] and [110]. One of the differences between this work and those two is that we consider fast fading channels and they considered slow fading channels. Fast fading channel results from such a fast varying environment where the channel coherence time is much less than a coding and channel estimation delay requirement. For each residual self-interference channel, we also apply the fast fading channel model. This is because the self-interference channel (even if through an RF circulator for a single antenna) still depends on the positions of the nearby moving reflectors. Consequently, we use an ergodic mutual information to measure the system performance. Note that unlike slow fading channels assumed in [109, 110] where instantaneous CSI can be estimated with reasonable accuracy,

here we do not assume any *instantaneous* CSI feedback from the receiver. Instead, we assume that the receiver feeds the transmitter with *statistical* CSI (the mean and variance of the CSI) and the knowledge of the statistics of the CSI is used at the transmitter to design optimal power schedules.

Since computing the closed form expression of the ergodic mutual information for fast fading channels is intractable, unlike [109]-[110], we assume that the variances of the transmission noise and the receiver noise do not depend on the variance of the transmitted signal and the received signal, respectively. Such an assumption is reasonable, since recent experimental results presented in [16] suggest that the residual self-interference of a point-to-point FD system is additive, noise-like and its variance does not depend on the variance of the transmitted signal [118]. In addition, the approximation of the effects of nonlinearities in [109]-[110] is valid only if higher order nonlinearities are contributing significantly [44], which is not the model we are considering in this chapter. This invariant transmission noise model has been commonly used in other papers [44, 57, 58, 66, 68, 130]. [1]

By exploiting both spatial and temporal freedoms of the source covariance matrices of the MIMO links, the authors of [109] and [110] maximize the lower bound of the achievable rates for FD MIMO relay channels and FD bi-directional MIMO channels for slow fading channels using gradient projection (GP) method under transmitter and receiver distortions, respectively. In this chapter, we develop algorithms useful to reveal a lower bound on the

[1] Note that the baseband cancellation is only possible when the residual self-interference is small. Subject to a small self-interference, it is appropriate to model the transmission noise variance and receiver noise variance as independent of the variance of the transmitted and received signal, respectively. This is because that the impact of these noises is much smaller than the self-interfering "signal". Note that the power of the transmission noise is typically $30 - 40$dB below that of the transmitted signal. The model we use is completely reasonable for a small dynamic range commonly encountered in baseband processing, and this chapter only claims the applicability in this situation.

ergodic mutual information of a FD bi-directional MIMO system and a FD MIMO relay system under a transmitter distortion model for fast fading channels where the instantaneous CSI is not known at the transmitters and imperfectly known at the receivers. In particular, using statistical CSI at the transmitters, we optimize the power allocation vectors at the nodes to maximize the ergodic mutual information of the FD systems subject to power constraints at the nodes under transmitter impairments. We develop a GP method to solve these non-convex optimization problems.

Moreover, based on [190], we introduce a simpler asymptotic closed-form expression for the ergodic mutual information of these FD systems, which is shown to be an accurate approximation even for systems with a small number of antennas. This expression simplifies the computation of the non-convex power allocation problem. It is shown through numerical simulations that at a high self-interference power level (when the INR is above the transmission SNR), the optimal power schedule reduces to the HD mode and at a low self-interference power level (when the INR is below the transmission SNR), the optimal power schedule switches to the FD mode.

This chapter is organized as follows. In Section 3.2, the system model of FD bi-directional MIMO system is discussed. In Section 3.3, we formulate the exact closed form of the lower bound ergodic mutual information expression for the FD bi-directional MIMO system. In Section 3.4, we maximize the sum ergodic mutual information subject to per node average power constraints using the GP method, and a simple asymptotic closed-form ergodic mutual information expression is introduced as well. In Section 3.5, the system model of FD MIMO relay system is discussed. In Section 3.6, simulation results are

provided to validate the performance of the algorithms.

The following notations are used in this chapter. Matrices and vectors are denoted by bold capital and lowercase letters, respectively. For matrices and vectors, $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate transpose, respectively. $\mathbf{E_H}\{\cdot\}$ stands for the statistical expectation with respect to the channel matrix $\mathbf{H}$; $\mathbf{I}_N$ denotes an $N \times N$ identity matrix; $\text{tr}\{\cdot\}$ stands for matrix trace; $|\cdot|$ is the determinant; $\|\cdot\|$ is the Euclidean norm of a vector and the Frobenius-norm of a matrix; $(\cdot)'$ denotes the first order derivative; $diag\{a_1, \cdots, a_n\}$ denotes a diagonal matrix with the diagonal elements given by $a_1, \cdots, a_n$. $\mathcal{CN}\left(\mu, \sigma^2\right)$ denotes complex Gaussian distribution with mean μ and variance σ^2.

3.2 System Model for a FD Bi-Directional Link

In this section, we describe the system model of a FD bi-directional MIMO system. (A FD MIMO relay system is discussed in Section 3.5.) We assume that each node has N physical antennas, but also has N virtual transmit antennas and N virtual receive antennas at any given time. The reason for using the word "virtual" is because a physical antenna can be used for simultaneous receiving and transmitting at the same carrier frequency [21]-[30].[2] Also note that even for a single physical antenna, there is still a self-interference channel between the virtual transmit antenna and the virtual receive antenna, and the response of this (circuit) channel is still affected by the reflectors around the physical antenna. The number of virtual antennas may correspond to the number of front-ends. A two front-end relay case was studied in [137].

[2] A FD WiFi radio that uses a single antenna was designed and implemented in [24].

Similar to [109] and [110], we partition the data transmission period under consideration or control into two time slots, since the benefit when the number of time slots is larger than the number of links is not significant [191]. The partition of the data transmission follows the concept of space-time power scheduling for multiple concurrent co-channel links shown in [191]. Particularly, the use of two distinct time slots gives the freedom to switch between FD and HD signaling depending on the power of the self-interference channel, while one time slot forces FD signaling, regardless of the power of the self-interference channel. This is similar to the MIMO interference channel in [191] and FD systems in [109, 110]. Particularly, the data transmission period is partitioned into two non-equal-length slots normalized to $\tau \in [0, 1]$ and $1-\tau$, respectively, and τ can be optimized using a grid search [109]. For convenience, we define $\tau(1) \triangleq \tau$ and $\tau(2) \triangleq 1 - \tau$.

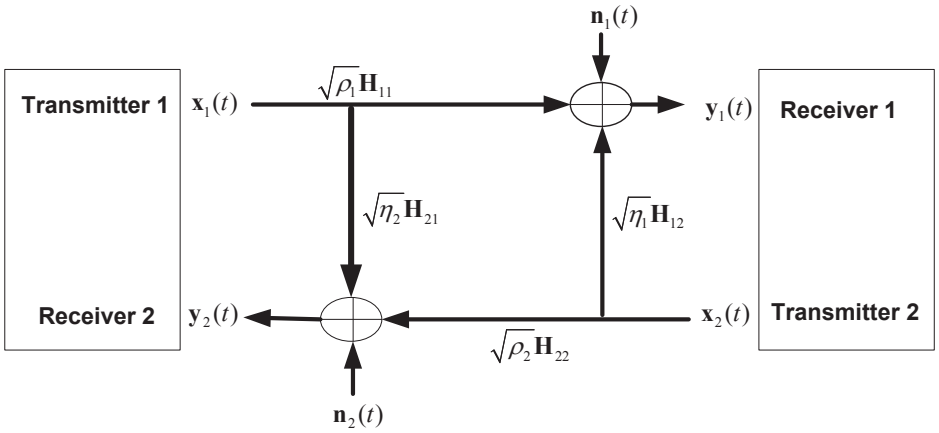

Figure 3.1: The signal flow diagram of a bi-directional full-duplex MIMO system. The node on the left has its transmitter denoted as transmitter 1 and its receiver as receiver 2. The node on the right has its transmitter denoted as transmitter 2 and its receiver as receiver 1. Each of the receivers and transmitters has N antennas.

As illustrated in Fig. 3.1, the receiver $i \in \{1, 2\}$ receives signals from both transmitters via MIMO channels $\mathbf{H}_{ij} \in \mathbb{C}^{N \times N}$. Here, $\mathbf{H}_{ii}$ is the channel for ith transmitter-receiver

pair between the two nodes, and $\mathbf{H}_{ij}, j \in \{1,2\}$ and $j \neq i$ denotes the self-interference channel from transmitter j to receiver i. All the channel matrices are assumed to be mutually independent and the entries of each matrix are i.i.d. circular complex Gaussian variables with zero mean and unit variance. We adopt the channel error model used for the FD systems in [57, 58, 59, 62, 131] and [136], where the receiver $i \in \{1,2\}$ is provided with some partial information of the channel, $\mathbf{H}_{ij}$, $j = 1,2$, and with this imperfect CSI, the receiver i performs MMSE estimation of $\mathbf{H}_{ij}$. Let us denote the MMSE estimation as $\tilde{\mathbf{H}}_{ij}$, and the estimation error as $\Delta\mathbf{H}_{ij} = \mathbf{H}_{ij} - \tilde{\mathbf{H}}_{ij}$, where $\tilde{\mathbf{H}}_{ij}$ and $\Delta\mathbf{H}_{ij}$ are uncorrelated, and the entries of $\Delta\mathbf{H}_{ij}$ are zero mean circularly symmetric complex Gaussian with variance $\sigma^2_{e,ij}$, as opposed to non-i.i.d. channel estimation errors in [109, 110]. Note that $\sigma^2_{e,ij}$ is assumed to be known to both the transmitter and receiver [192]. We will assume that the channel matrices remain constant over two consecutive time slots, but change randomly over an interval of many multiples of two time slots. We will design the power schedule to maximize an ergodic system mutual information which is averaged over the statistical distribution of the channel matrices. This mutual information is achievable (approximately) over the interval of many multiples of two time slots. Therefore, our theory is valid for "fast fading" channels, i.e., the time delay due to encoding and decoding over many multiples of two time slots is tolerable.

The quality of transmitted signals suffer from non-linear distortions in the power amplifier, phase noise, and IQ-imbalance [193]. The measurement results by [194] indicate that an i.i.d. additive Gaussian noise model accurately describes the sum of all such residual transmitter impairments. Such an assumption has also been commonly used in other FD

papers [44, 57, 58, 66, 68, 130].

We consider a FD bi-directional MIMO system that suffers from self-interference. The $N \times 1$ received signal vector at the ith receiver can be written as

$$\begin{aligned}
\mathbf{y}_i(t) &= \sqrt{\rho_i}\mathbf{H}_{ii}\left(\mathbf{x}_i(t)+\mathbf{c}_i(t)\right)+\sqrt{\eta_i}\mathbf{H}_{ij}\left(\mathbf{x}_j(t)+\mathbf{c}_j(t)\right)+\mathbf{n}_i(t) \\
&= \sqrt{\rho_i}\tilde{\mathbf{H}}_{ii}\mathbf{x}_i(t)+\sqrt{\rho_i}\Delta\mathbf{H}_{ii}\mathbf{x}_i(t)+\sqrt{\rho_i}\mathbf{H}_{ii}\mathbf{c}_i(t)+\sqrt{\eta_i}\tilde{\mathbf{H}}_{ij}\mathbf{x}_j(t) \\
&\quad +\sqrt{\eta_i}\Delta\mathbf{H}_{ij}\mathbf{x}_j(t)+\sqrt{\eta_i}\mathbf{H}_{ij}\mathbf{c}_j(t)+\mathbf{n}_i(t), \quad i,j\in\{1,2\} \text{ and } j\neq i \qquad (3.1)
\end{aligned}$$

where ρ_i denotes the average power gain of the ith transmitter-receiver link, η_i denotes the average power gain of the self-interference channel, $\mathbf{x}_i(t) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{Q}_i(t)\right)$ is the signal vector transmitted by node i within time slot t, $\mathbf{x}_j(t) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{Q}_j(t)\right)$ is the self-interference vector from the transmitter $j,\ j \neq i$ within time slot t, and $\mathbf{n}_i(t) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}_N\right)$ is the receiver noise which is additive white Gaussian noise (AWGN) vector. We assume that $\mathbf{n}_i(t)$ is independent of $\mathbf{x}_i(t)$ and $\mathbf{x}_j(t)$.

In (3.1), $\sqrt{\rho_i}\mathbf{c}_i(t)$ denotes the transmission noise from the ith transmitter, where $\mathbf{c}_i(t) \sim \mathcal{CN}\left(\mathbf{0}, \sigma_t^2\mathbf{I}_N\right),\ i = 1, 2$. Note that the transmit noise in (3.1) is $\sqrt{\rho_i}\mathbf{c}_i(t)$, not $\mathbf{c}_i(t)$ alone. And since the signal power is $\rho_i P_i$, while the transmission noise power is $\rho_i\sigma_t^2$, the transmitter noise depends on the power level. Here P_i is the averaged transmit power from the ith transmitter. In particular, incorporating $\sqrt{\rho_i}$ into $\mathbf{c}_i(t)$, we have the same transmission noise model as [44, 57, 58, 66, 68, 130, 194].

The receiver $i \in \{1, 2\}$ knows the interfering signal $\mathbf{x}_j(t)$ from transmitter $j \in \{1, 2\},\ j \neq i$, so the self-interference term $\sqrt{\eta_i}\tilde{\mathbf{H}}_{ij}\mathbf{x}_j(t)$ can be subtracted from $\mathbf{y}_i(t)$ [109,

110].

$$\begin{aligned}\tilde{\mathbf{y}}_i(t) &\triangleq \mathbf{y}_i(t) - \sqrt{\eta_i}\tilde{\mathbf{H}}_{ij}\mathbf{x}_j(t) \\ &= \sqrt{\rho_i}\tilde{\mathbf{H}}_{ii}\mathbf{x}_i(t) + \mathbf{v}_i(t) \end{aligned} \tag{3.2}$$

where

$$\mathbf{v}_i(t) = \sqrt{\rho_i}\Delta\mathbf{H}_{ii}\mathbf{x}_i(t) + \sqrt{\rho_i}\mathbf{H}_{ii}\mathbf{c}_i(t) + \sqrt{\eta_i}\Delta\mathbf{H}_{ij}\mathbf{x}_j(t) + \sqrt{\eta_i}\mathbf{H}_{ij}\mathbf{c}_j(t) + \mathbf{n}_i(t) \tag{3.3}$$

is the total noise in $\tilde{\mathbf{y}}_i(t)$. The covariance matrix of $\mathbf{v}_i(t)$ can be written as

$$\begin{aligned}\tilde{\mathbf{\Sigma}}_i(t) &= \mathtt{E}\left\{\mathbf{v}_i(t)\mathbf{v}_i(t)^H|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right\} \\ &= \rho_i\mathtt{E}_{\Delta\mathbf{H}_{ii}}\left\{\Delta\mathbf{H}_{ii}\mathbf{Q}_i(t)\Delta\mathbf{H}_{ii}^H\right\} + \rho_i\sigma_t^2\mathtt{E}_{\Delta\mathbf{H}_{ii}}\left\{\mathbf{H}_{ii}\mathbf{H}_{ii}^H\right\} + \eta_i\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\Delta\mathbf{H}_{ij}\mathbf{Q}_j(t)\Delta\mathbf{H}_{ij}^H\right\} \\ &\quad + \eta_i\sigma_t^2\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\mathbf{H}_{ij}\mathbf{H}_{ij}^H\right\} + \mathbf{I}_N \\ &= \rho_i\sigma_{e,ii}^2\mathrm{tr}\left\{\mathbf{Q}_i(t)\right\}\mathbf{I}_N + \rho_i\sigma_t^2\left(\tilde{\mathbf{H}}_{ii}\tilde{\mathbf{H}}_{ii}^H + \sigma_{e,ii}^2N\mathbf{I}_N\right) + \eta_i\sigma_{e,ij}^2\mathrm{tr}\left\{\mathbf{Q}_j(t)\right\}\mathbf{I}_N \\ &\quad + \eta_i\sigma_t^2\left(\tilde{\mathbf{H}}_{ij}\tilde{\mathbf{H}}_{ij}^H + \sigma_{e,ij}^2N\mathbf{I}_N\right) + \mathbf{I}_N, \quad i,j\in\{1,2\} \text{ and } j\neq i \end{aligned} \tag{3.4}$$

where the first expectation is taken with respect to $\mathbf{x}_i(t)$, $\mathbf{x}_j(t)$ and $\mathbf{n}_i(t)$, and here we have used the identity of $\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\Delta\mathbf{H}_{ij}\mathbf{A}\Delta\mathbf{H}_{ij}^H\right\}= \sigma_{e,ij}^2\mathrm{tr}\{\mathbf{A}\}\mathbf{I}_N$, where the entries of $\Delta\mathbf{H}_{ij}$ are i.i.d. with $\mathcal{CN}(0,\sigma_{e,ij}^2)$ and $\mathbf{A}\in\mathbb{C}^{N\times N}$ is a known matrix.

3.3 Achievable Rates

In this section, we formulate the ergodic mutual information expression for the FD bi-directional MIMO system when the transmitters do not have instantaneous CSI and the receivers have imperfect instantaneous CSI, i.e., $\mathbf{H}_{ii}$ is unknown at the transmitter i

but partially known at the receiver i. As a result of the channel estimation errors and transmitter impairments in (3.3), the noise $\mathbf{v}_i(t)$ is generally non-Gaussian. To the best of our knowledge, the exact mutual information of MIMO channels with channel estimation errors is still an open problem even for point-to-point MIMO systems [192, 195]. However, assuming $\mathbf{v}_i(t)$ as Gaussian, which is the worst noise distribution from the perspective of mutual information, we can obtain the lower bound [195], which was also used in [109, 110].

For a given time-sharing parameter τ, the lower bound of the sum mutual information of the system averaged over two time slots can be written as

$$I\left(\mathbf{Q}_1, \mathbf{Q}_2\right) \;=\; \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\log_2\left|\mathbf{I}_N + \rho_i\tilde{\mathbf{H}}_{ii}\mathbf{Q}_i(t)\tilde{\mathbf{H}}_{ii}^H\tilde{\mathbf{\Sigma}}_i(t)^{-1}\right|$$

where $\mathbf{Q}_i \triangleq \left[\mathbf{Q}_i^T(1), \mathbf{Q}_i^T(2)\right]^T$, $i = 1, 2$. Then, a lower bound of the ergodic sum mutual information of the system averaged over two time slots can be written as

$$\bar{I}\left(\mathbf{Q}_1, \mathbf{Q}_2\right) = \underbrace{\sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\mathrm{E}_{\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}}\left\{\log_2\left|\mathbf{I}_N + \rho_i\tilde{\mathbf{H}}_{ii}\mathbf{Q}_i(t)\tilde{\mathbf{H}}_{ii}^H\tilde{\mathbf{\Sigma}}_i(t)^{-1}\right|\right\}}_{\bar{I}_i(\mathbf{Q}_1,\mathbf{Q}_2)} \tag{3.5}$$

To derive a closed-form expression for the ergodic sum mutual information (3.5), we use the eigendecomposition of $\mathbf{Q}_i(t)$, which can be written as $\mathbf{Q}_i(t) = \mathbf{U}_i(t)\mathbf{D}_i(t)\mathbf{U}_i(t)^H$, $i = 1, 2$, where $\mathbf{U}_i(t)$ is the unitary matrix of eigenvectors, and

$$\mathbf{D}_i(t) = diag\left\{d_{i1}(t), d_{i2}(t), \ldots, d_{iN}(t)\right\},\ i = 1, 2$$

is a diagonal matrix of all eigenvalues. For convenience, we will use the column vectors $\mathbf{d}_1(t)$ and $\mathbf{d}_2(t)$ defined as

$$\mathbf{d}_i(t) = \left[d_{i1}(t), d_{i2}(t), \ldots, d_{iN}(t)\right]^T,\ i = 1, 2.$$

Now we can rewrite (3.5) as

$$\bar{I}\left(\mathbf{D}_1,\mathbf{D}_2\right)=\sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\mathtt{E}_{\hat{\tilde{\mathbf{H}}}_{ii},\hat{\tilde{\mathbf{H}}}_{ij}}\left\{\log_2\left|\mathbf{I}_N+\rho_i\hat{\tilde{\mathbf{H}}}_{ii}\mathbf{D}_i(t)\hat{\tilde{\mathbf{H}}}_{ii}^H\hat{\tilde{\mathbf{\Sigma}}}_i(t)^{-1}\right|\right\} \tag{3.6}$$

where

$$\begin{aligned}
\mathbf{D}_i &\triangleq \left[\mathbf{D}_i^T(1),\mathbf{D}_i^T(2)\right]^T, \quad i=1,2\\
\hat{\tilde{\mathbf{H}}}_{ii} &\triangleq \tilde{\mathbf{H}}_{ii}\mathbf{U}_i(t), \quad i=1,2\\
\hat{\tilde{\mathbf{H}}}_{ij} &\triangleq \tilde{\mathbf{H}}_{ij}\mathbf{U}_j(t), \quad (i,j)\in\{1,2\} \text{ and } j\neq i\\
\hat{\tilde{\mathbf{\Sigma}}}_i(t) &\triangleq \rho_i\sigma_{e,ii}^2\mathrm{tr}\left\{\mathbf{D}_i(t)\right\}\mathbf{I}_N+\rho_i\sigma_t^2\left(\hat{\tilde{\mathbf{H}}}_{ii}\hat{\tilde{\mathbf{H}}}_{ii}^H+\sigma_{e,ii}^2N\mathbf{I}_N\right)+\eta_i\sigma_{e,ij}^2\mathrm{tr}\left\{\mathbf{D}_j(t)\right\}\mathbf{I}_N\\
&\quad+\eta_i\sigma_t^2\left(\hat{\tilde{\mathbf{H}}}_{ij}\hat{\tilde{\mathbf{H}}}_{ij}^H+\sigma_{e,ij}^2N\mathbf{I}_N\right)+\mathbf{I}_N, \quad i,j\in\{1,2\} \text{ and } j\neq i.
\end{aligned}$$

Since $\tilde{\mathbf{H}}_{ii}$ has i.i.d. Gaussian entries and $\mathbf{U}_i(t)$ is unitary, the statistics of $\hat{\tilde{\mathbf{H}}}_{ii}$ is identical to that of $\tilde{\mathbf{H}}_{ii}$ [196, Lemma 5]. Therefore, for notational simplicity, in the sequel we will drop the hats on the matrices. Thus, the ergodic sum mutual information expression (3.6) can be rewritten as

$$\begin{aligned}
\bar{I}\left(\mathbf{d}_1,\mathbf{d}_2\right) &= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\mathtt{E}_{\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}}\left\{\log_2\left|\mathbf{I}_N+\rho_i\tilde{\mathbf{H}}_{ii}\mathbf{D}_i(t)\tilde{\mathbf{H}}_{ii}^H\tilde{\mathbf{\Sigma}}_i(t)^{-1}\right|\right\}\\
&= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\left[\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\rho_i\tilde{\mathbf{H}}_{ii}\mathbf{D}_i(t)\tilde{\mathbf{H}}_{ii}^H+\tilde{\mathbf{\Sigma}}_i(t)\right|\right\}-\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\tilde{\mathbf{\Sigma}}_i(t)\right|\right\}\right]\\
&= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\left[\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\rho_i\tilde{\mathbf{H}}_{ii}\left(\mathbf{D}_i(t)+\sigma_t^2\mathbf{I}_N\right)\tilde{\mathbf{H}}_{ii}^H+\eta_i\sigma_t^2\tilde{\mathbf{H}}_{ij}\tilde{\mathbf{H}}_{ij}^H+c_i(t)\mathbf{I}_N\right|\right\}\right.\\
&\qquad\left.-\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\rho_i\sigma_t^2\tilde{\mathbf{H}}_{ii}\tilde{\mathbf{H}}_{ii}^H+\eta_i\sigma_t^2\tilde{\mathbf{H}}_{ij}\tilde{\mathbf{H}}_{ij}^H+c_i(t)\mathbf{I}_N\right|\right\}\right]\\
&= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\left[\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\tilde{\mathbf{H}}_i\mathbf{\Lambda}_i(t)\tilde{\mathbf{H}}_i^H+\mathbf{I}_N\right|\right\}\right.\\
&\qquad\left.-\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\tilde{\mathbf{H}}_i\bar{\mathbf{\Lambda}}_i(t)\tilde{\mathbf{H}}_i^H+\mathbf{I}_N\right|\right\}\right]
\end{aligned} \tag{3.7}$$

where

$$
\begin{aligned}
\tilde{\mathbf{H}}_i &= \left[\tilde{\mathbf{H}}_{ii}, \tilde{\mathbf{H}}_{ij}\right] \\
c_i(t) &\triangleq \rho_i \sigma_{e,ii}^2 \mathbf{1}_N^T \mathbf{d}_i(t) + \rho_i \sigma_t^2 \sigma_{e,ii}^2 N + \eta_i \sigma_{e,ij}^2 \mathbf{1}_N^T \mathbf{d}_j(t) \\
&\quad + \eta_i \sigma_t^2 \sigma_{e,ij}^2 N + 1, \ i,j \in \{1,2\} \text{ and } j \neq i \\
\mathbf{d}_i &\triangleq \left[\mathbf{d}_i(1)^T, \mathbf{d}_i(2)^T\right]^T, \quad i = 1,2 \\
\tilde{\mathbf{\Sigma}}_i(t) &\triangleq \rho_i \sigma_{e,ii}^2 \mathrm{tr}\left\{\mathbf{D}_i(t)\right\} \mathbf{I}_N + \rho_i \sigma_t^2 \left(\tilde{\mathbf{H}}_{ii} \tilde{\mathbf{H}}_{ii}^H + \sigma_{e,ii}^2 N \mathbf{I}_N\right) + \eta_i \sigma_{e,ij}^2 \mathrm{tr}\left\{\mathbf{D}_j(t)\right\} \mathbf{I}_N \\
&\quad + \eta_i \sigma_t^2 \left(\tilde{\mathbf{H}}_{ij} \tilde{\mathbf{H}}_{ij}^H + \sigma_{e,ij}^2 N \mathbf{I}_N\right) + \mathbf{I}_N, \quad i,j \in \{1,2\} \text{ and } j \neq i \\
\mathbf{\Lambda}_i(t) &\triangleq diag\left\{\boldsymbol{\lambda}_{1,i}^T(t), \boldsymbol{\lambda}_{2,i}^T(t)\right\} \quad i = 1,2. \\
\bar{\mathbf{\Lambda}}_i(t) &\triangleq diag\left\{\bar{\boldsymbol{\lambda}}_{1,i}^T(t), \boldsymbol{\lambda}_{2,i}^T(t)\right\} \quad i = 1,2. \\
\boldsymbol{\lambda}_{1,i}(t) &= \rho_i \frac{\mathbf{d}_i(t) + \sigma_t^2 \mathbf{1}_N}{c_i(t)} \\
\boldsymbol{\lambda}_{2,i}(t) &= \eta_i \frac{\sigma_t^2}{c_i(t)} \mathbf{1}_N \\
\bar{\boldsymbol{\lambda}}_{1,i}(t) &= \rho_i \frac{\sigma_t^2}{c_i(t)} \mathbf{1}_N.
\end{aligned}
$$

Here $\mathbf{1}_N$ is an $N \times 1$ column vector of ones. Note that $\mathbf{1}_N^T \mathbf{d}_i(t),\ (i,t) = 1,2$ is the power consumed at the ith node at time slot t and it is not fixed and changes with respect to self-interference power as we will see in the simulations, whereas $\sum_{t=1}^{2} \mathbf{1}_N^T \mathbf{d}_i(t)$ is the total power consumed by the node i and it is fixed.

The expression $\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\tilde{\mathbf{H}}_i \mathbf{\Lambda}_i(t) \tilde{\mathbf{H}}_i^H + \mathbf{I}_N\right|\right\}$ in (3.7) can be viewed as the ergodic mutual information of a point-to-point MIMO channel with $2N$ transmit and N receive antennas. A closed-form expression for the ergodic mutual information of such a system has been shown in [197], where a determinant representation for the distribution of quadratic forms of a complex Gaussian matrix has been used. Using the results in [197], (3.7) can be

equivalently expressed as

$$\bar{I}\left(\mathbf{d}_1, \mathbf{d}_2\right) = \log_2(e) \sum_{i=1}^{2} \sum_{t=1}^{2} \tau(t) \left[\sum_{n=0}^{N-1} \sum_{k=1}^{2N} \left(c_{tikn}\left(\mathbf{\Lambda}_i(t)\right) Q(n, \lambda_{tik}) - c_{tikn}\left(\bar{\mathbf{\Lambda}}_i(t)\right) Q(n, \bar{\lambda}_{tik}) \right) \right] \tag{3.8}$$

where $c_{tikn}\left(\mathbf{\Lambda}_i(t)\right)$ and $Q\left(n, \lambda_{tik}\right)$ are defined in Appendix 3.A. Here $\lambda_{tik} \triangleq \left[\mathbf{\Lambda}_i(t)\right]_{k,k}$ and $\bar{\lambda}_{tik} \triangleq \left[\bar{\mathbf{\Lambda}}_i(t)\right]_{k,k}$, $k = 1, \ldots, 2N$ denote the (k,k)th element of matrix $\mathbf{\Lambda}_i(t)$ and $\bar{\mathbf{\Lambda}}_i(t)$, respectively. In (3.40) of Appendix 3.A, $S_1(x) \triangleq \int_x^\infty e^{-\tau}/\tau d\tau$ is the exponential integral function of order 1 [198].

As shown in (3.8), the ergodic sum mutual information is now expressed as a finite summation involving rational functions and exponential integration functions of the power scheduling vectors $\mathbf{d}_i(t)$, $(i,t) \in \{1,2\}$, of both transmitting nodes. The exponential integration function is available in many software such as MATLAB and Mathematica. Thus, (3.8) is easy to compute. Note that (3.8) is derived under the assumption that all λ_{tik}, $k = 1, \ldots, 2N$, have distinct values. Under the condition that some of them are identical, the closed-form ergodic sum mutual information expression can be obtained by deriving the limit of (3.8) with respect to those common values of λ_{tik} using L'Hospital's rule. However, for numerical evaluation, it is sufficient to slightly and randomly perturb these identical values of λ_{tik}, since all functions are continuous and λ_{tik} is deterministic [199]. The same assumption holds for $\bar{\lambda}_{tik}$ as well.

3.4 Maximization of the Sum Ergodic Mutual Information

In this section, we aim at maximizing the sum ergodic mutual information (3.5) by choosing the transmit covariance matrices $\mathbf{Q}_1(t)$ and $\mathbf{Q}_2(t)$, $t = 1, 2$ subject to per node average power constraints and subsequently optimize the time-sharing parameter τ. Note that we consider fast fading channels in which the instantaneous CSI is assumed to be unknown at the transmitting nodes. When the knowledge of the instantaneous CSI is absent, statistical properties of the CSI is necessary for designing optimal power schedules. The optimization problem can be formulated as

$$\max_{\mathbf{Q}_1, \mathbf{Q}_2, \tau(t)} \quad \sum_{i=1}^{2} \bar{I}_i \left(\mathbf{Q}_1, \mathbf{Q}_2\right) \tag{3.9}$$

$$\text{s.t.} \quad \sum_{t=1}^{2} \tau(t) \mathrm{tr} \left\{\mathbf{Q}_i(t)\right\} \leq P_i, \quad i = 1, 2 \tag{3.10}$$

$$\mathbf{Q}_i(t) \geq 0, \quad \forall i, t \in \{1, 2\} \tag{3.11}$$

where $\bar{I}_i \left(\mathbf{Q}_1, \mathbf{Q}_2\right)$ is given in (3.5) and P_i is the averaged transmit power from the ith transmitter.

3.4.1 Gradient Projection Approach

For a fixed τ, the optimal $\mathbf{d}_1$ and $\mathbf{d}_2$ can be obtained by solving the following problem

$$\max_{\mathbf{d}_1, \mathbf{d}_2} \quad \bar{I}(\mathbf{d}_1, \mathbf{d}_2) \tag{3.12}$$

$$\text{s.t.} \quad \sum_{t=1}^{2} \tau(t) \|\mathbf{d}_i(t)\|_1 = P_i, \quad i = 1, 2 \tag{3.13}$$

$$\mathbf{d}_i \geq 0, \quad i = 1, 2 \tag{3.14}$$

where $\bar{I}(\mathbf{d}_1, \mathbf{d}_2)$ is given in (3.8) and (3.13) is the power constraint at the ith transmitter. Here $\|.\|_1$ denotes the sum norm (or l_1 norm) of a vector. For a vector $\mathbf{x}, \mathbf{x} \geq 0$ means that each entry of $\mathbf{x}$ is nonnegative.

The objective function (3.12) is highly non-convex and does not have a clear structure. We can develop numerical algorithms based on nonlinear programming techniques to obtain a locally optimal solution to the problem (3.12)-(3.14). We choose the GP method [200], which is an extension of the unconstrained steepest descent method to the convex constrained problems. The GP method is simple, efficient, and guarantees the convergence to a stationary point, provided that proper step sizes are chosen.

There are two important steps in the GP algorithm: the computation of the gradient of the objective function, and the projection of the updated optimization variable onto the convex set specified by constraint functions. To apply the GP method to solve the problem (3.12)-(3.14), we first take gradient steps for $\mathbf{d}_1$ and $\mathbf{d}_2$, and then project the updated $\mathbf{d}_1$ and $\mathbf{d}_2$ onto the constraint set specified by (3.13) and (3.14). The gradient of the objective function (3.12) with respect to $d_{lm}(t),\ l = 1, 2,\ m = 1, \ldots, N,\ t = 1, 2$, is given by

$$\begin{aligned}\frac{\partial \bar{I}(\mathbf{d}_1, \mathbf{d}_2)}{\partial d_{lm}(t)} &= \tau(t) \log_2(e) \sum_{i=1}^{2} \Bigg[\sum_{n=0}^{N-1} \sum_{k=1}^{2N} \left(c'_{tikn} \left(\mathbf{\Lambda}_i(t) \right) Q(n, \lambda_{tik}) + c_{tikn} \left(\mathbf{\Lambda}_i(t) \right) Q'(n, \lambda_{tik}) \right. \\ &\quad \left. - c'_{tikn} \left(\bar{\mathbf{\Lambda}}_i(t) \right) Q(n, \bar{\lambda}_{tik}) - c_{tikn} \left(\bar{\mathbf{\Lambda}}_i(t) \right) Q'(n, \bar{\lambda}_{tik}) \right) \Bigg]. \end{aligned} \tag{3.15}$$

The parameters in (3.15) are given in Appendix 3.B.

Let us first consider the gradient steps of the ith transmitter-receiver pair, $i \in$

$\{1,2\}$, and denote the $2N \times 1$ vector of gradient as

$$\mathbf{g}_i \triangleq \left[\frac{\partial \bar{I}(\mathbf{d}_1,\mathbf{d}_2)}{\partial d_{i1}(1)},\ldots,\frac{\partial \bar{I}(\mathbf{d}_1,\mathbf{d}_2)}{\partial d_{iN}(2)}\right]^T, \qquad i=1,2. \tag{3.16}$$

Then taking a step along the positive gradient direction, the power allocation vector is updated as

$$\hat{\mathbf{d}}_i = \bar{\mathbf{d}}_i + s\mathbf{g}_i, \qquad i=1,2$$

where s is a scalar of step size, and $\bar{\mathbf{d}}_i$ is the previous power allocation vector.

The next step of the GP algorithm is to project $\hat{\mathbf{d}}_i$ onto the feasible region of power vector constraints (3.13)-(3.14). The projection operation is basically searching for a point $\tilde{\mathbf{d}}_i$ in the region of (3.13)-(3.14), which has a minimum Euclidean distance to the point $\hat{\mathbf{d}}_i$. Thus, the optimization problem for the projection operation can be written as

$$\min_{\tilde{\mathbf{d}}_i} \qquad \left\|\tilde{\mathbf{d}}_i - \hat{\mathbf{d}}_i\right\|^2 \tag{3.17}$$

$$\text{s.t.} \qquad \sum_{t=1}^{2}\tau(t)\|\tilde{\mathbf{d}}_i(t)\|_1 = P_i, \qquad \tilde{\mathbf{d}}_i \geq 0,\ i=1,2. \tag{3.18}$$

The problem (3.17)-(3.18) is convex and can be efficiently solved by the Lagrange multiplier method. It turns out that the problem (3.17)-(3.18) has a water-filling solution which is given by

$$\tilde{d}_{ik}(t) = \left[\hat{d}_{ik}(t) - \frac{\tau(t)\mu}{2}\right]^+,\ k=1,\ldots,N,\ (i,t)=1,2. \tag{3.19}$$

where $\mu \geq 0$ is the Lagrange multiplier, and for a real scalar x, $[x]^+ \triangleq \max\{x,0\}$. The Lagrange multiplier μ can be obtained by substituting (3.19) back into (3.18) and solving the following nonlinear equation

$$\sum_{t=1}^{2}\sum_{k=1}^{N}\tau(t)\left[\hat{d}_{ik}(t) - \frac{\tau(t)\mu}{2}\right]^+ = P_i,\ i=1,2. \tag{3.20}$$

We can use the bisection method to solve (3.20), since the left hand side of (3.20) is a piecewise linear function and monotonically decreasing with respect to μ.

At the kth iteration, the power allocations vectors are updated as

$$\bar{\mathbf{d}}_i^{(k+1)} = \bar{\mathbf{d}}_i^{(k)} + \delta^{(k)}\left(\tilde{\mathbf{d}}_i^{(k)} - \bar{\mathbf{d}}_i^{(k)}\right), \qquad i = 1, 2 \tag{3.21}$$

$$\tilde{\mathbf{d}}_i^{(k)} = \text{proj}\left[\bar{\mathbf{d}}_i^{(k)} + s^{(k)}\mathbf{g}_i^{(k)}\right], \qquad i = 1, 2. \tag{3.22}$$

where proj[.] stands for the projection operation in (3.17)-(3.18), $\delta^{(k)}$ and $s^{(k)}$ are scalars of step size and can be chosen according to the Armijo rule [200]. In this rule, $s^{(k)} = s$ is a constant throughout the iterations, and $\delta^{(k)} = \theta^{m_k}$, where m_k is the minimal nonnegative integer that satisfies the following inequality

$$\bar{I}\left(\bar{\mathbf{d}}^{(k+1)}\right) - \bar{I}\left(\bar{\mathbf{d}}^{(k)}\right) \geq \sigma\theta^{m_k}\sum_{i=1}^{2}\left(\mathbf{g}_i^{(k)}\right)^T\left(\tilde{\mathbf{d}}_i^{(k)} - \bar{\mathbf{d}}_i^{(k)}\right) \tag{3.23}$$

where σ and θ are constants and $\bar{\mathbf{d}}^{(k)} = \left[(\bar{\mathbf{d}}_1^{(k)})^T, (\bar{\mathbf{d}}_2^{(k)})^T\right]^T$. According to [200], usually σ is chosen close to 0, and a proper choice of θ is from 0.1 to 0.5.

The steps of (3.21) and (3.22) are performed for both nodes and continue until vector $\bar{\mathbf{d}}^{(k)}$ converges. The GP algorithm using the Armijo rule along the feasible direction guarantees such a convergence [200] and the convergence criterion is given as

$$\max\text{abs}\left\{\bar{\mathbf{d}}^{(k+1)} - \bar{\mathbf{d}}^{(k)}\right\} \leq \epsilon \tag{3.24}$$

where max abs{.} denotes the maximal absolute value among all elements of a vector and ϵ is a positive constant close to 0. The procedure of applying the GP technique to solve the problem (3.12)-(3.14) is summarized in Table I. Subsequently, we optimize over $\tau \in [0, 1]$ using a grid-search [109].

Table 3.1: Procedure of the projected gradient power allocation approach

1) Initialize power allocation vectors $\bar{\mathbf{d}}_i$.
Choose step sizes. Set $k = 0$.
2) Set $k := k + 1$.
Calculate the gradient of (3.8) $\mathbf{g}_i^{(k)}$ from (3.15) using (3.42)-(3.46).
Let $\hat{\mathbf{d}}_i^{(k)} = \bar{\mathbf{d}}_i^{(k)} + s\mathbf{g}_i^{(k)}$.
Project $\hat{\mathbf{d}}_i^{(k)}$ to obtain $\tilde{\mathbf{d}}_i^{(k)}$ using (3.19).
Update $\bar{\mathbf{d}}_i^{(k)}$ using (3.21) and (3.22).
3) If convergent, end.
Else go to step 2.

For the bi-directional case, the ergodic mutual information (3.8) and (3.25) are functions of averaged signal-to-noise ratio (SNR) and nominal interference-to-noise ratio (INR). Under the same INR for all interfering links, the desired link with the higher SNR gets the whole data transmission slot, i.e. $\tau = 1$ and the link with the lower SNR does not transmit, i.e. $\tau = 0$. In other words, the optimal τ is either one or zero depending on the average SNR. Though we presented a general transmission protocol and solved the optimization problem as a function of τ, this time-slot allocation is not fair for the bi-directional case, so we assumed $\tau = 0.5$ in our simulations for the bi-directional system.

3.4.2 Approximation of Sum Ergodic Mutual Information

In this subsection, we introduce a much simpler expression of $\bar{I}(\mathbf{d}_1, \mathbf{d}_2)$ than the one in (3.8), which in turn simplifies the computation in solving the problem (3.12)-(3.14). This simplification is based on an asymptotical form of $\bar{I}(\mathbf{d}_1, \mathbf{d}_2)$ when $N \to \infty$ as proposed in [190]. The proof of this asymptotical form is as follows: In [201], SNR at the output of an MMSE receiver is shown. And using the results in [201], the authors in [202] obtain

the asymptotic capacity of an optimum receiver for randomly spread CDMA in fading channels. With a simple SNR normalization and by applying [202, Theorem IV.1], the asymptotic capacity of MIMO architectures impaired by AWGN as well as spatially colored interference can be easily found as the number of antennas go to infinity as shown in Appendix of [190]. Applying the result in [190], the sum ergodic mutual information in (3.7) can be approximated as

$$\begin{aligned}
\bar{I}\left(\mathbf{d}_1,\mathbf{d}_2\right) &= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\mathtt{E}_{\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}}\left\{\log_2\left|\mathbf{I}_N+\rho_i\tilde{\mathbf{H}}_{ii}\mathbf{D}_i(t)\tilde{\mathbf{H}}_{ii}^H\tilde{\mathbf{\Sigma}}_i(t)^{-1}\right|\right\} \\
&= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\left[\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\tilde{\mathbf{H}}_i\mathbf{\Lambda}_i(t)\tilde{\mathbf{H}}_i^H+\mathbf{I}_N\right|\right\}-\mathtt{E}_{\tilde{\mathbf{H}}_i}\left\{\log_2\left|\tilde{\mathbf{H}}_i\bar{\mathbf{\Lambda}}_i(t)\tilde{\mathbf{H}}_i^H+\mathbf{I}_N\right|\right\}\right] \\
&= \sum_{i=1}^{2}\sum_{t=1}^{2}\tau(t)\left[\sum_{k=1}^{2N}\log_2\left(\frac{1+N\alpha_{i,1}(t)\lambda_{tik}}{1+N\alpha_{i,2}(t)\bar{\lambda}_{tik}}\right)\right. \\
&\qquad\left.+N\log_2\left(\frac{\alpha_{i,2}(t)}{\alpha_{i,1}(t)}\right)+N\left(\alpha_{i,1}(t)-\alpha_{i,2}(t)\right)\log_2 e\right] \qquad (3.25)
\end{aligned}$$

where λ_{tik} and $\bar{\lambda}_{tik}$ is defined in (3.8) and $0<\alpha_{i,1}(t),\alpha_{i,2}(t)<1$ satisfies the following nonlinear equation

$$\alpha_{i,1}(t)+\sum_{k=1}^{2N}\frac{\alpha_{i,1}(t)\lambda_{tik}}{N\alpha_{i,1}(t)\lambda_{tik}+1}=1. \qquad (3.26)$$

$$\alpha_{i,2}(t)+\sum_{k=1}^{2N}\frac{\alpha_{i,2}(t)\bar{\lambda}_{tik}}{N\alpha_{i,2}(t)\bar{\lambda}_{tik}+1}=1. \qquad (3.27)$$

We can use the bisection method to compute $\alpha_{i,1}(t)$, since the left hand side of (3.26) is monotonically increasing functions of $\alpha_{i,1}(t)$. Same argument also holds for $\alpha_{i,2}(t)$. It is shown in the simulations that (3.25) is an accurate approximation of (3.8) even when N is as small as three.

With the simplified closed-form expression (3.25), the problem (3.12)-(3.14) can be solved by the GP method similar to the one developed in Section 3.4.1. We only need the

gradient of the objective function (3.25) with respect to $d_{lm}(t),\ l = 1, 2,\ m = 1, \ldots, N,\ t = 1, 2$, which is given by

$$\frac{\partial \bar{I}\left(\mathbf{d}_1, \mathbf{d}_2\right)}{\partial d_{lm}(t)} = \tau(t) \sum_{i=1}^{2} \sum_{k=1}^{2N} \left[\frac{N\alpha_{i,1}(t)\lambda'_{tik}}{1 + N\alpha_{i,1}(t)\lambda_{tik}} - \frac{N\alpha_{i,2}(t)\bar{\lambda}'_{tik}}{1 + N\alpha_{i,2}(t)\bar{\lambda}_{tik}} \right] \tag{3.28}$$

where λ'_{tik} and $\bar{\lambda}'_{tik}$ are defined in (3.45) and (3.46) in Appendix 3.B, respectively. Note that since $\alpha_{i,1}(t)$ and $\alpha_{i,2}(t)$ are coefficients and are not functions of $d_{lm}(t)$, they can be treated as constants in the gradient expression [203].

3.5 Full-Duplex Relay Systems

In this section, we study the performance of a DF FD relay system that suffers from self-interference, where all nodes are equipped with multiple antennas. The source node transmits signal streams to the destination node via the relay node and the direct link as shown in Fig. 3.2. We assume that the instantaneous CSI is not used by the transmitters and an imperfect CSI is used by the receiver. It can be seen from Fig. 3.2 that the system model of a FD relay is similar to that of the bi-directional FD system in Fig. 3.1.

For the relay system, we still assume that the relay uses N transmit antennas and N receive antennas in either FD mode or HD mode. For relay, the direction of reception is generally different from the direction of transmission. If directional antennas are used, the transmit antennas and the receive antennas should face different directions. And hence, even if the HD mode is considered, the relay still should use N antennas for transmission and N antennas for reception at any given time. For power efficiency, directional antennas are a much better choice than omnidirectional antennas.

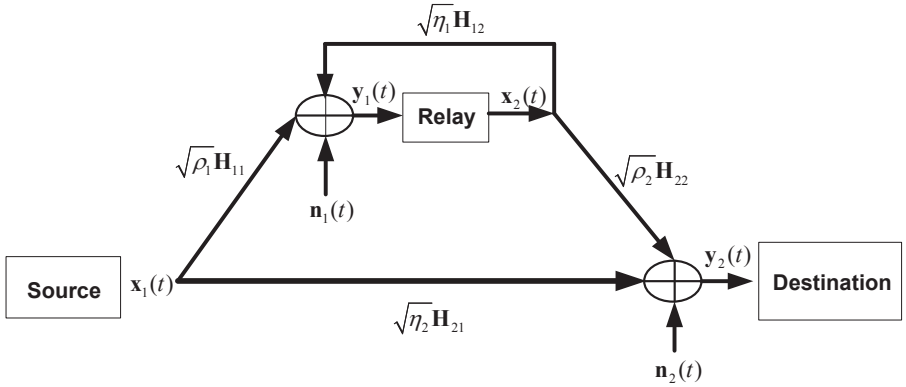

Figure 3.2: The signal flow diagram of a two-hop full-duplex MIMO relay system.

After the partial self-interference cancelation at the relay node, the received signal at the relay node and the destination is given by

$$\mathbf{y}_R(t) = \tilde{\mathbf{y}}_1(t) \tag{3.29}$$

$$\mathbf{y}_D(t) = \tilde{\mathbf{y}}_2(t) + \sqrt{\eta}_2\tilde{\mathbf{H}}_{21}\mathbf{x}_1(t) \tag{3.30}$$

where $\tilde{\mathbf{y}}_i(t),\ i = 1, 2$ is defined in (3.2). Unlike the relay node, where the partial self-interference cancellation is possible, the destination node cannot cancel the interference term $\sqrt{\eta}_2\tilde{\mathbf{H}}_{21}\mathbf{x}_1(t)$ resulting from the direct link, but adds it to the total noise $\mathbf{v}_2(t)$ in (3.3). (If the direct link is strong, the optimal scheme may switch to direct transmission as shown in [137]. But we do not consider this scenario). For fixed τ, the lower bound of the averaged ergodic mutual information of the DF FD relay system over two time slots can be written as [204]

$$\bar{I}\left(\mathbf{Q}_1, \mathbf{Q}_2\right) = \min\left\{\bar{I}_1\left(\mathbf{Q}_1, \mathbf{Q}_2\right), \bar{I}_2\left(\mathbf{Q}_1, \mathbf{Q}_2\right)\right\} \tag{3.31}$$

where $\bar{I}_i\left(\mathbf{Q}_1, \mathbf{Q}_2\right),\ i = 1, 2$ is defined in (3.5). The only difference is that the covariance

matrix of the total noise $\tilde{\mathbf{\Sigma}}_2(t)$ (3.4) in $\bar{I}_2\left(\mathbf{Q}_1, \mathbf{Q}_2\right)$ has the additional term $\eta_2 \tilde{\mathbf{H}}_{21} \mathbf{Q}_1(t) \tilde{\mathbf{H}}_{21}^H$ because of the additional term in (3.30).

3.5.1 Maximization of the Ergodic Mutual Information of the FD Relay System

In this subsection, we aim at maximizing the ergodic mutual information (3.31) by choosing the transmit covariance matrices $\mathbf{Q}_1(t)$ and $\mathbf{Q}_2(t)$, $t = 1, 2$, subject to per link power constraints and subsequently optimize over τ. Similar to Section 3.4, we consider fast fading channels in which the statistical CSI is assumed to be known at the transmitting nodes to design the optimal power schedules. This problem can be formulated as

$$\max_{\mathbf{Q}_1, \mathbf{Q}_2} \quad \min\left\{\bar{I}_1\left(\mathbf{Q}_1, \mathbf{Q}_2\right), \bar{I}_2\left(\mathbf{Q}_1, \mathbf{Q}_2\right)\right\} \tag{3.32}$$

$$\text{s.t.} \quad \sum_{t=1}^{2} \tau(t) \mathrm{tr}\left\{\mathbf{Q}_i(t)\right\} \leq P_i, \quad i = 1, 2, \tag{3.33}$$

$$\mathbf{Q}_i(t) \geq 0, \quad i = 1, 2. \tag{3.34}$$

Applying the link equalizing algorithm proposed in [109], $\min\left\{\bar{I}_1\left(\mathbf{Q}_1, \mathbf{Q}_2\right), \bar{I}_2\left(\mathbf{Q}_1, \mathbf{Q}_2\right)\right\}$, the objective function in (3.32), can be replaced with a ζ-weighted sum-rate problem, i.e., $\zeta \bar{I}_1\left(\mathbf{Q}_1, \mathbf{Q}_2\right) + (1-\zeta) \bar{I}_2\left(\mathbf{Q}_1, \mathbf{Q}_2\right)$, where ζ is computed using bisection method (see Section IV-A of [109] for more details about the link-equalizing algorithm). Therefore, the ζ-weighted sum-rate optimization problem can be expressed as

$$\max_{\mathbf{Q}_1, \mathbf{Q}_2} \quad \sum_{i=1}^{2} \zeta(i) \bar{I}_i\left(\mathbf{Q}_1, \mathbf{Q}_2\right) \tag{3.35}$$

$$\text{s.t.} \quad \sum_{t=1}^{2} \tau(t) \mathrm{tr}\left\{\mathbf{Q}_i(t)\right\} \leq P_i, \quad i = 1, 2, \tag{3.36}$$

$$\mathbf{Q}_i(t) \geq 0, \quad i = 1, 2. \tag{3.37}$$

where $\zeta(1) = \zeta$ and $\zeta(2) = 1 - \zeta$. Since the optimization problem (3.35)-(3.37) has a similar structure with (3.9)-(3.11), GP method proposed in Section 3.4.1 can be applied to solve (3.35)-(3.37). Note that at each bisection step to compute ζ, GP method is used. The closed-form ergodic mutual information expression of the relay system can be obtained similar to (3.8). Due to the additional term in (3.30), the only modification required is on the term $\boldsymbol{\lambda}_{2,2}(t)$ in $\boldsymbol{\Lambda}_2(t)$ and $\bar{\boldsymbol{\Lambda}}_2(t)$, which is modified as

$$\boldsymbol{\lambda}_{2,2}(t) \quad = \quad \eta_2 \frac{\mathbf{d}_1(t) + \sigma_t^2 \mathbf{1}_N}{c_2(t)}.$$

Similarly the gradient of the objective function (3.35) can be obtained similar to (3.15). The only modification is on the terms λ'_{t2k} and $\bar{\lambda}'_{t2k}$, which are given at the top of next page.

3.6 Simulation Results

In this section, we study the performance of the proposed FD MIMO bi-directional communication system through numerical simulations as a function of the averaged SNR, the nominal INR, the number of antennas N, the channel estimation errors $\sigma_{e,ij}^2$ and the transmitter impairments σ_t^2. For all simulation examples, we set the same channel estimation error for all links, i.e., $\sigma_{e,ij}^2 = \sigma_e^2,\ i,j \in \{1,2\}$. The Armijo parameters are selected as $\sigma = 0.1$, $\theta = 0.5$, and the stopping threshold of the GP algorithm is chosen as $\epsilon = 10^{-5}$. For simplicity, we focus on the case of $\eta_1 = \eta_2 = \eta$ and the same average transmit power for each node (i.e., $P_i = N,\ i = 1,2$). Thus, the averaged SNR for all de-

$$\lambda'_{t2k} = \begin{cases} -\rho_2^2 c_2(t)^{-2}\sigma_{e,22}^2\left(d_{2k}(t)+\sigma_t^2\right), & l=2 \text{ and } k\neq m \text{ and } k\leq N \\ \frac{\rho_2}{c_2(t)} - \rho_2^2 c_2(t)^{-2}\sigma_{e,22}^2\left(d_{2k}(t)+\sigma_t^2\right), & l=2 \text{ and } k=m \text{ and } k\leq N \\ -\rho_2\eta_2 c_2(t)^{-2}\sigma_{e,22}^2\left(d_{1k}(t)+\sigma_t^2\right), & l=2 \text{ and } k>N \\ -\rho_2\eta_2 c_2(t)^{-2}\sigma_{e,21}^2\left(d_{2k}(t)+\sigma_t^2\right), & l\neq 2\ (l=1) \text{ and } k\leq N \\ -\eta_2^2 c_2(t)^{-2}\sigma_{e,21}^2\left(d_{1(k-N)}(t)+\sigma_t^2\right), & l\neq i \text{ and } k\neq N+m \text{ and } k>N \\ \frac{\eta_2}{c_2(t)} - \eta_2^2 c_2(t)^{-2}\sigma_{e,21}^2\left(d_{1(k-N)}(t)+\sigma_t^2\right), & l\neq i \text{ and } k=N+m \text{ and } k>N \end{cases}.$$

$$\bar{\lambda}'_{t2k} = \begin{cases} -\rho_2^2 c_2(t)^{-2}\sigma_{e,22}^2\sigma_t^2, & l=2 \text{ and } k\leq N \\ -\rho_2\eta_2 c_2(t)^{-2}\sigma_{e,22}^2\left(d_{1(k-N)}(t)+\sigma_t^2\right), & l=2 \text{ and } k>N \\ -\rho_2\eta_2 c_2(t)^{-2}\sigma_{e,21}^2\sigma_t^2, & l\neq 2\ (l=1) \text{ and } k\leq N \\ -\eta_2^2 c_2(t)^{-2}\sigma_{e,21}^2\left(d_{1(k-N)}(t)+\sigma_t^2\right), & l\neq i \text{ and } k\neq N+m \text{ and } k>N \\ \frac{\eta_2}{c_2(t)} - \eta_2^2 c_2(t)^{-2}\sigma_{e,21}^2\left(d_{1(k-N)}(t)+\sigma_t^2\right), & l\neq i \text{ and } k=N+m \text{ and } k>N \end{cases}.$$

---

sired links is defined as $\text{SNR}_i = \rho_i N,\ i=1,2$ and the nominal INR for all interfering links $\text{INR}_i = \text{INR} = \eta N,\ i=1,2$. Since the nominal INR and the averaged $\text{SNR}_i,\ i=1,2$ are quasi static, we assume that their values can be obtained with relatively high precision, so we treat them as deterministic parameters. To optimize the HD scheme, we use the GP method to solve the problem (3.9)-(3.11) with the HD constraint of $\mathbf{Q}_1(2) = \mathbf{Q}_2(1) = \mathbf{0}$. Note that the HD scheme is invariant to INR. To show the importance of using two time slots, we compare our FD system using two data transmission slots (FD2) with the FD system using

only one data transmission slot (FD1). In the FD1 scheme, the same source covariance matrices are used for both time slots, i.e., $\mathbf{Q}_1(1) = \mathbf{Q}_1(2)$ and $\mathbf{Q}_2(1) = \mathbf{Q}_2(2)$. Since the GP algorithm only converges to a locally optimal solution, we use the output of the HD scheme as the initialization of the FD scheme. For the maximization problem (3.9)-(3.11), the time-sharing coefficient τ can be optimized over the grid $\tau \in \{0.1, 0.2, \ldots, 0.9\}$ [109].

In the first example, we compare the exact and approximate closed-form expressions of the lower bound ergodic mutual information of the FD2 system using (3.8) and (3.25), respectively, for different number of antennas. We set $\text{SNR}_i = \text{SNR} = 20\text{dB}$, $i = 1, 2$, $\sigma_e^2 = 0.01$ and $\sigma_t^2 = -30\text{dB}$. It can be seen from Fig. 3.3 that the ergodic mutual information increases with the number of antennas. Note that the ergodic mutual information of the FD2 system is always equal to or greater than that of the HD system (the reason is explained in Fig. 3.5). It can also be seen from Fig. 3.3 that the asymptotic closed-form expression for the ergodic mutual information is an accurate approximation even when the number of antennas is as small as $N = 3$. Unless otherwise stated, hereafter we adopt the asymptotic closed-form ergodic mutual information expression, since it has a much lower computational complexity.

In the second example, we investigate the role of channel estimation errors on the lower bound of the ergodic mutual information (3.25). Here we set $N = 2$, $\text{SNR}_i = \text{SNR} = 20\text{dB}$, $i = 1, 2$ and $\sigma_t^2 = -30\text{dB}$. It can be seen from Fig. 3.4 that as the channel estimation errors increases, the ergodic mutual information of both the FD2 and HD systems decreases. The gap between ergodic mutual information curves diminishes as σ_e^2 increases.

In the next example, we investigate the impact of INR on the ergodic mutual

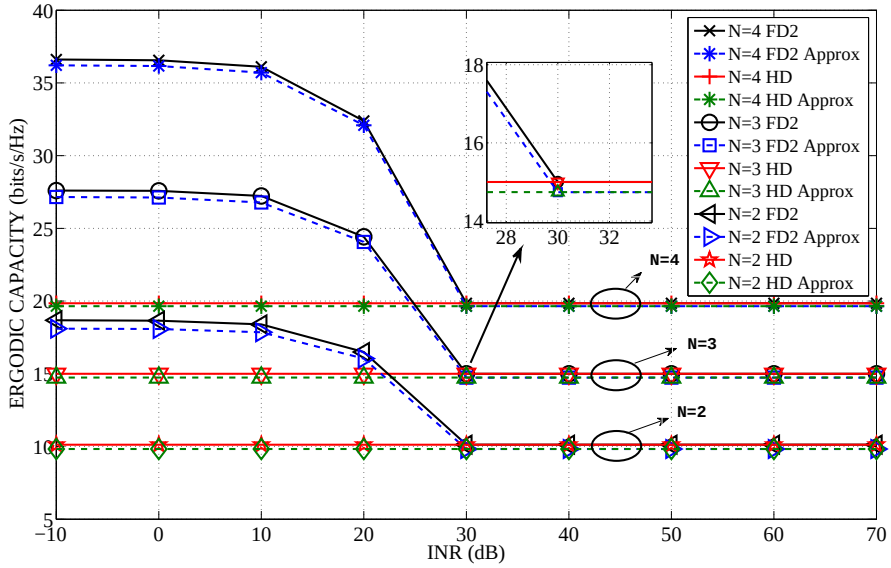

Figure 3.3: Ergodic mutual information comparison of the FD2 and HD systems with different number of antennas versus INR. Here SNR = 20dB, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30$dB.

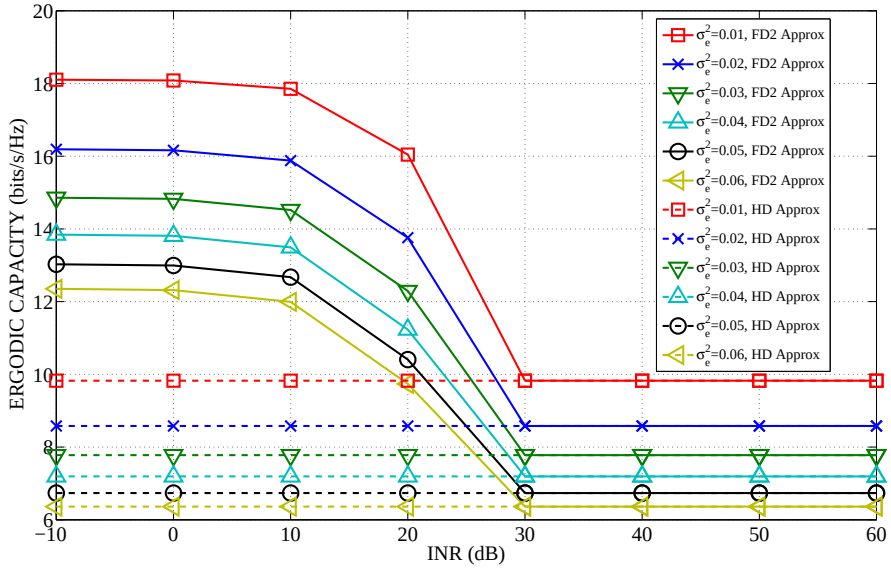

Figure 3.4: Ergodic mutual information comparison of the FD2 and HD systems with different channel estimation errors versus INR. Here $N = 2$, SNR = 20dB, $\sigma_t^2 = -30$dB.

information of the FD2, FD1, and HD schemes with $N = 3$, $\text{SNR}_2 = 20\text{dB}$, $\sigma_e^2 = 0.01$ and $\sigma_t^2 = -30\text{dB}$ for different SNR_1 values. As expected, it can be observed from Fig. 3.5 that the HD scheme is invariant to INR. For the low-to-mid values of INR, the FD2 scheme has the FD system behavior and it switches to the HD scheme at the high values of INR. The FD1 scheme performs similar to the FD2 scheme at low-to-mid values of INR, but its performance drops below that of the HD scheme for larger values of INR. The use of two distinct data time slots gives the freedom to switch to the HD signaling when the power of the self-interference channel is high (where the HD scheme is optimal), while the FD1 system forces FD signaling at each time slot, regardless of the strength of the self-interference channel.

In our fourth example, we examine the INR that FD2 converges to HD. Fig. 3.6 demonstrates that the behavior of convergence depends on σ_t^2 and σ_e^2 values.

In our fifth example, we examine the ergodic mutual information of the FD2 and HD systems versus $\text{SNR}_i = \text{SNR},\ \text{i} = 1, 2$ for various fixed values of INR. We choose $N = 3$, $\sigma_e^2 = 0.01$ and $\sigma_t^2 = -30\text{dB}$. It can be observed from Fig. 3.7 that at low INR, the system operates in the FD mode for all values of SNR, since SNR mostly dominates INR. At high INR, the system operates in the HD mode at low values of SNR (since INR dominates SNR), but switches to the FD mode as SNR increases, since SNR starts to dominate INR.

In our sixth example, we examine the ergodic mutual information of the FD2 and HD systems versus $\text{SNR}_i = \text{SNR},\ \text{i} = 1, 2$ for various values of σ_t^2. We choose $N = 3$, $\text{INR} = 20\text{dB}$ and $\sigma_e^2 = 0.01$. It can be observed from Fig. 3.8 that as σ_t^2 decreases, the ergodic mutual information increases and the gap between the curves diminishes.

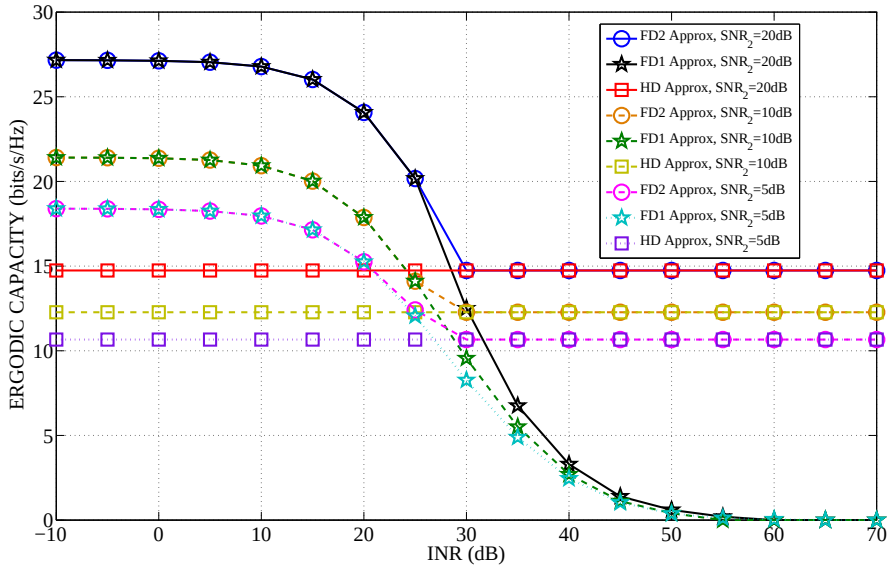

Figure 3.5: Ergodic mutual information comparison of the FD2, FD1, and HD systems versus INR for different SNR_2 values. Here $N = 3$, $\mathrm{SNR}_1 = 20\mathrm{dB}$, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30\mathrm{dB}$.

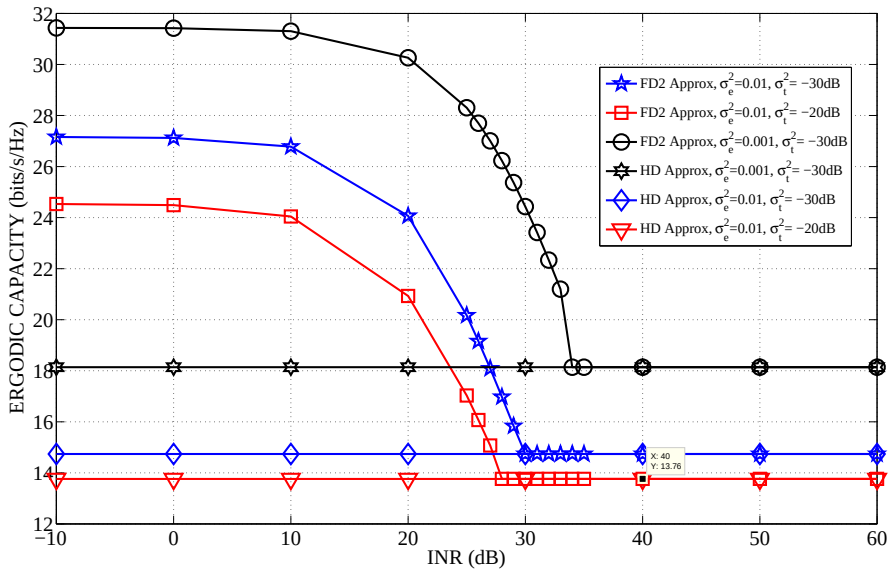

Figure 3.6: Ergodic mutual information comparison of the FD2 and HD systems versus INR for different σ_e^2 and σ_t^2 values. Here $N = 3$, $\mathrm{SNR} = 20\mathrm{dB}$.

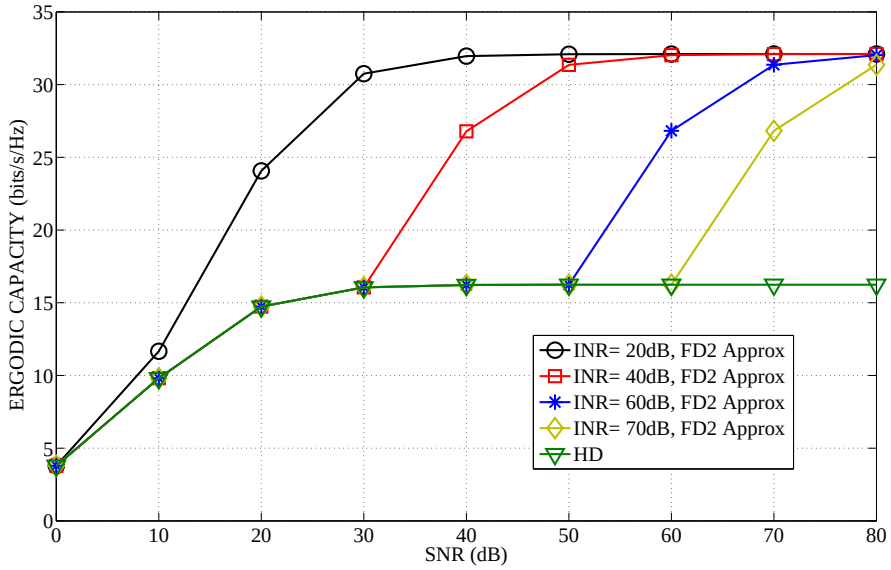

Figure 3.7: Ergodic mutual information comparison of the FD2 and HD systems for different INR values versus SNR. Here $N = 3$, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30$dB.

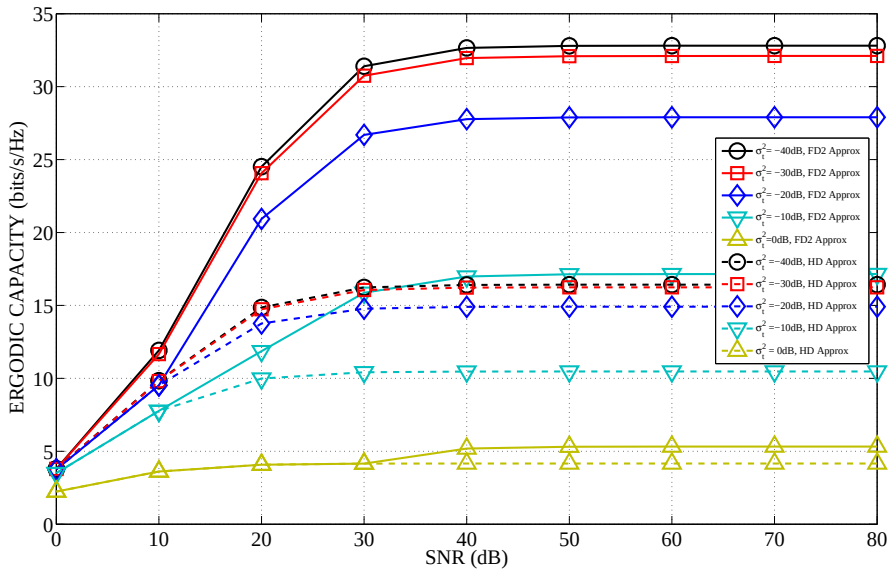

Figure 3.8: Ergodic mutual information comparison of the FD2 and HD systems with different σ_t^2 values versus SNR. Here $N = 3$, INR = 20dB, $\sigma_e^2 = 0.01$.

In our last example, we consider MIMO FD relay systems. We obtain similar results as MIMO bi-directional FD system as shown in Fig. 3.9. In particular, the relay node operates in the FD mode when the self-interference is weak, and as the self-interference increases, we observe a transition of the relay node to the HD mode. Similar to [109], we can also observe that, compared to using fixed value $\tau = 0.5$, the optimization of τ gives a small rate improvement.

3.7 Appendix 3.A: Parameters in (3.8)

See (3.38)-(3.40) for the parameters $c_{tikn}\left(\mathbf{\Lambda}_i(t)\right)$ and $Q\left(n, \lambda_{tik}\right)$ in (3.8).

$$c_{tikn}\left(\mathbf{\Lambda}_i(t)\right) = \frac{(-1)^{N-n-1}}{n!}\lambda_{tik}^{N-1}\left(\prod_{h\neq k}^{2N}(\lambda_{tik}-\lambda_{tih})\right)^{-1} b_{tikn}\left(\mathbf{\Lambda}_i(t)\right) \tag{3.38}$$

$$b_{tikn}\left(\mathbf{\Lambda}_i(t)\right) = \begin{cases} \sum_{1\leq j_1<\ldots<j_{N-n-1}\leq 2N}^{j_r\neq k} \lambda_{tij_1}\ldots\lambda_{tij_{N-n-1}}, & n=0,\ldots,N-2 \\ 1, & n=N-1 \end{cases} \tag{3.39}$$

$$Q(n,\lambda_{tik}) = \int_0^\infty \ln(1+x)x^n e^{-(x/\lambda_{tik})}\,\mathrm{d}x \tag{3.40}$$

$$= \sum_{r=0}^{n}\frac{n!(-1)^{(n-r)}}{(n-r)!}\lambda_{tik}^{r+1}e^{1/\lambda_{tik}}S_1\left(\frac{1}{\lambda_{tik}}\right) + \sum_{r=1}^{n}\sum_{s=0}^{r-1}\sum_{h=0}^{r-s-1}\frac{n!(-1)^{(n-r)}\lambda_{tik}^{h+s+2}}{(n-r)!(r-s-h-1)!(r-s)} \tag{3.41}$$

3.8 Appendix 3.B: Parameters in (3.15)

See (3.42)-(3.46) shown at the bottom of the next page for the definition of the parameters in (3.15).

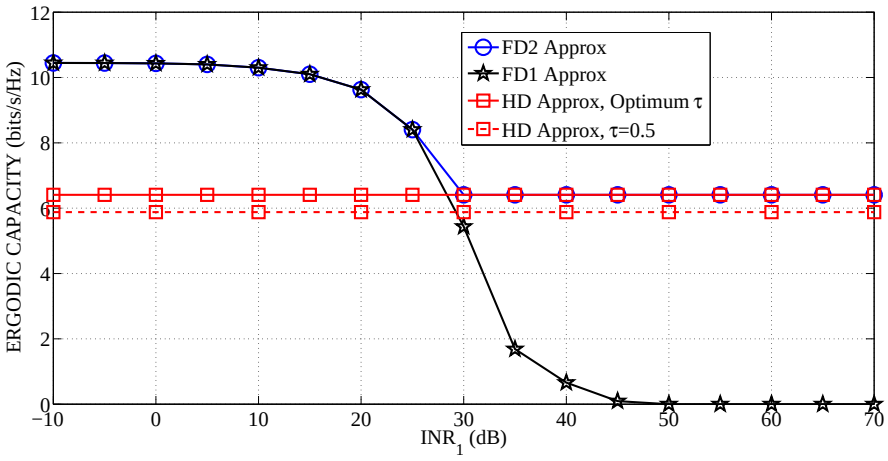

Figure 3.9: Ergodic mutual information comparison of the FD2, FD1, and HD relay systems versus INR_1. Here $N = 3$, $\mathrm{SNR}_1 = 20\mathrm{dB}$, $\mathrm{SNR}_2 = 10\mathrm{dB}$, $\mathrm{INR}_2 = 0\mathrm{dB}$, $\sigma_e^2 = 0.01$, $\sigma_t^2 = -30\mathrm{dB}$.

$$\bar{\lambda}'_{tik} = \begin{cases} -\rho_i^2 c_i(t)^{-2}\sigma_{e,ii}^2\sigma_t^2, & l = i \text{ and } k \le N \\ -\rho_i\eta_i c_i(t)^{-2}\sigma_{e,ii}^2\sigma_t^2, & l = i \text{ and } k > N \\ -\rho_i\eta_i c_i(t)^{-2}\sigma_{e,ij}^2\sigma_t^2, & l \ne i \ (l = j) \text{ and } k \le N \\ -\eta_i^2 c_i(t)^{-2}\sigma_{e,ij}^2\sigma_t^2, & l \ne i \ (l = j) \text{ and } k > N \end{cases} . \tag{3.46}$$

$$
\begin{aligned}
c'_{tikn}\left(\mathbf{\Lambda}_i(t)\right) &= \frac{(-1)^{N-n-1}(N-1)}{n!}\lambda_{tik}^{N-2}\lambda'_{tik}\left(\prod_{h\neq k}^{2N}(\lambda_{tik}-\lambda_{tih})\right)^{-1} b_{tikn}\left(\mathbf{\Lambda}_i(t)\right) \\
&\quad -\frac{(-1)^{N-n-1}}{n!}\lambda_{tik}^{N-1}\sum_{j=1,j\neq k}^{2N}\left[\left(\prod_{h\neq k}^{2N}(\lambda_{tik}-\lambda_{tih})\right)^{-1}(\lambda_{tik}-\lambda_{tij})^{-1}\left(\lambda'_{tik}-\lambda'_{tij}\right)\right] \\
&\quad \times b_{tikn}\left(\mathbf{\Lambda}_i(t)\right)+\frac{(-1)^{N-n-1}}{n!}\lambda_{tik}^{N-1}\left(\prod_{h\neq k}^{2N}(\lambda_{tik}-\lambda_{tih})\right)^{-1} b'_{tikn}\left(\mathbf{\Lambda}_i(t)\right) \qquad (3.42)
\end{aligned}
$$

$$
\begin{aligned}
Q'(n,\lambda_{tik}) &= \sum_{r=0}^{n}\frac{n!(-1)^{(n-r)}}{(n-r)!}\lambda_{tik}^{r}\lambda'_{tik}\left[(r+1)e^{1/\lambda_{tik}}S_1\left(\frac{1}{\lambda_{tik}}\right)-\frac{1}{\lambda_{tik}}e^{1/\lambda_{tik}}S_1\left(\frac{1}{\lambda_{tik}}\right)+1\right] \\
&\quad +\sum_{r=1}^{n}\sum_{s=0}^{r-1}\sum_{h=0}^{r-s-1}\frac{n!(-1)^{(n-r)}(h+s+2)\lambda_{tik}^{h+s+1}}{(n-r)!(r-s-h-1)!(r-s)}\lambda'_{tik} \qquad (3.43)
\end{aligned}
$$

$$
b'_{tikn}\left(\mathbf{\Lambda}_i(t)\right) = \begin{cases}
\sum_{1\leq j_1<\ldots<j_{N-n-1}\leq 2N}^{j_r\neq k}\lambda'_{tij_1}\lambda_{tij_2}\ldots\lambda_{tij_{N-n-1}} & \\
+\sum_{1\leq j_1<\ldots<j_{N-n-1}\leq 2N}^{j_r\neq k}\lambda_{tij_1}\lambda'_{tij_2}\ldots\lambda_{tij_{N-n-1}} & \\
+\ldots+\sum_{1\leq j_1<\ldots<j_{N-n-1}\leq 2N}^{j_r\neq k}\lambda_{tij_1}\lambda_{tij_2}\ldots\lambda'_{tij_{N-n-1}}, & n=0,\ldots,N-2 \\
0, & n=N-1
\end{cases} \qquad (3.44)
$$

$$
\lambda'_{tik} = \begin{cases}
-\rho_i^2 c_i(t)^{-2}\sigma_{e,ii}^2\left(d_{ik}(t)+\sigma_t^2\right), & l=i \text{ and } k\neq m \text{ and } k\leq N \\
\frac{\rho_i}{c_i(t)}-\rho_i^2 c_i(t)^{-2}\sigma_{e,ii}^2\left(d_{ik}(t)+\sigma_t^2\right), & l=i \text{ and } k=m \text{ and } k\leq N \\
-\rho_i\eta_i c_i(t)^{-2}\sigma_{e,ii}^2\sigma_t^2, & l=i \text{ and } k>N \\
-\rho_i\eta_i c_i(t)^{-2}\sigma_{e,ij}^2\left(d_{ik}(t)+\sigma_t^2\right), & l\neq i\ (l=j) \text{ and } k\leq N \\
-\eta_i^2 c_i(t)^{-2}\sigma_{e,ij}^2\sigma_t^2, & l\neq i\ (l=j) \text{ and } k>N
\end{cases} . \qquad (3.45)
$$

Chapter 4

Weighted-Sum-Rate Maximization for Bi-directional Full-Duplex MIMO Systems

4.1 Introduction

With a given level of residual self-interference, what could be the best possible performance of a bi-directional link between two FD radios? This is one of the important questions of our interest.

The authors in [110] explored the above question. Following their previous work on FD relays in [109], they studied how to best utilize two FD radio nodes for bi-directional communications. They formulated the problem into a power allocation problem in two time or frequency slots, which is similar to the space-time power scheduling approach proposed

in [191]. The work in [111] later extended the findings in [110] to fast fading channels, under a simpler model, where channel state information (CSI) is not known at the transmitters and only imperfectly known at the receivers.

This chapter reports a further analysis of the performance of a bi-directional link between two FD MIMO radio nodes. We also cast the problem into a power allocation problem using two time or frequency slots. We develop algorithms useful for computing the maximum weighted sum-rate (WSR) of the link (also called the system). Two types of power constraints are considered. One is a power constraint on the total power of the system. The other is a power constraint on the power at each radio node. Under either constraint, the problem is non-convex. Following an approach used in [205] and [206], we turn the the WSR problem into a weighted minimum-mean-squared-error (WMMSE)problem, the latter of which is easier to solve.

The algorithms we develop determine a power schedule to best utilize the FD MIMO radio nodes for bi-directional communication. For instance, with the individual power constraint, the power schedule resulting from our algorithm reduces to the HD mode when the self-interference level is high, or otherwise to the FD mode when the self-interference level is low. Also for instance, with the total power constraint and when the self-interference level is high, the power schedule resulting from our algorithm not only switches to the HD mode but also opportunistically selects one of the two directions of communication in order to maximize the sum rate of the system. Most importantly, our algorithm can be used to determine the best power schedule for maximum WSR of the bi-directional link at any given self-interference level.

The following notations are used in this chapter. The matrices and the vectors are denoted as bold capital and lowercase letters, respectively. $(\cdot)^T$ is the transpose; $(\cdot)^H$ is the conjugate transpose. $\mathtt{E}\{\cdot\}$ means the statistical expectation; $\mathbf{I}_N$ is the N by N identity matrix; $\mathrm{tr}(\cdot)$ is the trace; $|\cdot|$ is the determinant; $\mathrm{diag}(\mathbf{A})$ is the diagonal matrix with the same diagonal elements as $\mathbf{A}$. $\mathcal{CN}\left(\mu, \sigma^2\right)$ denotes a complex Gaussian distribution with mean μ and variance σ^2. $[x]^+$ denotes $\max(x, 0)$.

4.2 System Model

In this chapter, we consider the same FD bi-directional MIMO system model covered in Fig. 3.1 of Chapter 3. Particularly, we partition the data transmission period into two time slots ($t = 1, 2$), consider the same channel estimation error model, and assume that two FD nodes have N transmit/receive antennas. But unlike the fast-fading model in Chapter 3, in this chapter, we consider slow-fading channels.

We also take into account the limited dynamic range (DR). Limited-DR is caused by non-ideal amplifiers, oscillators, analog-to-digital converters (ADCs), and digital-to-analog converters (DACs). To model the effects of limited DR, we used the same assumptions in [109, 110]. Particularly, at each receive antenna an additive white Gaussian "receiver distortion" with variance β times the energy of the undistorted receive signal on that receive antenna is applied and at each transmit antenna, an additive white Gaussian "transmitter noise" with variance κ times the energy of the intended transmit signal is applied.[1]

[1]Note that unlike the invariant transmitter distortion model assumed in Chapter 3, in this chapter we assume both transmitter and receiver distortion, which depend on the covariance matrices of the transmit

We consider a FD bi-directional MIMO system that suffers from self-interference. Thus, the receiver i receives a combination of the signals transmitted by both transmitters and noise. The $N \times 1$ received signal at the ith receiver is written as

$$\begin{aligned}
\mathbf{y}_i(t) &= \sqrt{\rho}_i \mathbf{H}_{ii} \left(\mathbf{x}_i(t) + \mathbf{c}_i(t)\right) + \sqrt{\eta}_i \mathbf{H}_{ij} \left(\mathbf{x}_j(t) + \mathbf{c}_j(t)\right) + \mathbf{e}_i(t) + \mathbf{n}_i(t) \\
&= \sqrt{\rho}_i \tilde{\mathbf{H}}_{ii} \mathbf{x}_i(t) + \sqrt{\rho}_i \Delta \mathbf{H}_{ii} \mathbf{x}_i(t) + \sqrt{\rho}_i \mathbf{H}_{ii} \mathbf{c}_i(t) + \sqrt{\eta}_i \tilde{\mathbf{H}}_{ij} \mathbf{x}_j(t) + \sqrt{\eta}_i \Delta \mathbf{H}_{ij} \mathbf{x}_j(t) \\
&\quad + \sqrt{\eta}_i \mathbf{H}_{ij} \mathbf{c}_j(t) + \mathbf{e}_i(t) + \mathbf{n}_i(t), \qquad i, j \in \{1, 2\} \text{ and } j \neq i
\end{aligned} \tag{4.1}$$

where $\mathbf{x}_i(t) = \mathbf{V}_i(t)\mathbf{d}_i(t)$ is $N \times 1$ signal vector transmitted by transmitter i with covariance matrix $\mathtt{E}\left\{\mathbf{x}_i(t)\mathbf{x}_i(t)^H\right\} = \mathbf{V}_i(t)\mathbf{V}_i(t)^H$ and $\mathbf{x}_j(t) = \mathbf{V}_j(t)\mathbf{d}_j(t)$ is $N \times 1$ signal vector transmitted vector from the transmitter j, $j \neq i$ with a covariance matrix $\mathtt{E}\left\{\mathbf{x}_j(t)\mathbf{x}_j(t)^H\right\} = \mathbf{V}_j(t)\mathbf{V}_j(t)^H$, which incurs self-interference at the ith receiver. $\mathbf{V}_k(t) \in \mathbb{C}^{N \times N}$ and $\mathbf{d}_k(t) \in \mathbb{C}^N$ represent precoding matrix and data streams at the node k, $k = 1, 2$, respectively. Note that, we assume $\mathtt{E}\left\{\mathbf{d}_i(t)\mathbf{d}_i(t)^H\right\} = \mathbf{I}_N$, $\mathtt{E}\left\{\mathbf{d}_i(t)\mathbf{d}_j(t)^H\right\} = \mathbf{0}$, $i \neq j$ and the number of streams is equal to the number of antennas, N. $\mathbf{n}_i(t) \in \mathbb{C}^N$ is the additive white Gaussian noise (AWGN) vector at the ith receiver with zero mean and unit covariance matrix, $\mathtt{E}\left\{\mathbf{n}_i(t)\mathbf{n}_i(t)^H\right\} = \mathbf{I}_N$ and it is uncorrelated to $\mathbf{x}_i(t)$ and $\mathbf{x}_j(t)$. ρ_i denotes the average gain of the ith transmitter-receiver link, and η_i denotes the average gain of the self-interference channel. $\mathbf{c}_i(t) \in \mathbb{C}^N$, $i = 1, 2$ is the transmitter noise at the ith transmitter, which models the effect of limited transmitter DR and closely approximates the effects of additive power-amplifier noise, non-linearities in the DAC and phase noise [109, 110]. The covariance matrix of $\mathbf{c}_i(t)$ is given by κ ($\kappa \ll 1$) times the energy of the intended signal at each transmit antenna, i.e. $\mathbf{c}_i(t) \sim \mathcal{CN}\left(\mathbf{0}, \kappa \text{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\right)$, and is independent of

and received signals, respectively.

$\mathbf{x}_i(t)$. $\mathbf{e}_i(t) \in \mathbb{C}^N,\ i = 1, 2$ is the additive receiver distortion at the ith receiver, which models the effect of limited receiver DR and closely approximates the combined effects of additive gain-control noise, non-linearities in the ADC and phase noise [109, 110]. The covariance matrix of $\mathbf{e}_i(t)$ is given by β ($\beta \ll 1$) times the energy of the undistorted received signal at each receive antenna, i.e. $\mathbf{e}_i(t) \sim \mathcal{CN}\left(\mathbf{0}, \beta \text{diag}\left(\mathbf{\Phi}_i(t)\right)\right)$. $\mathbf{\Phi}_i(t) = \text{Cov}\{\mathbf{u}_i(t)\}$, where $\mathbf{u}_i(t)$ is the ith receiver's undistorted received vector, i.e. $\mathbf{u}_i(t) = \mathbf{y}_i(t) - \mathbf{e}_i(t)$. $\mathbf{e}_i(t)$ is independent of $\mathbf{u}_i(t)$.

The receiver $i \in \{1, 2\}$ knows the interfering codewords $\mathbf{x}_j(t)$ from transmitter $j \in \{1, 2\},\ j \neq i$, so the self-interference term $\sqrt{\eta}_i \tilde{\mathbf{H}}_{ij}\mathbf{x}_j(t)$ is known and thus be cancelled. The interference canceled signal can then be written as

$$\begin{aligned} \tilde{\mathbf{y}}_i(t) &= \mathbf{y}_i(t) - \sqrt{\eta}_i \tilde{\mathbf{H}}_{ij}\mathbf{x}_j(t) \\ &= \sqrt{\rho}_i \tilde{\mathbf{H}}_{ii}\mathbf{x}_i(t) + \mathbf{v}_i(t) \end{aligned} \tag{4.2}$$

where $\mathbf{v}_i(t)$ is the unknown interference components of (4.2) after self-interference cancellation and given by

$$\begin{aligned} \mathbf{v}_i(t) &= \sqrt{\rho}_i \Delta\mathbf{H}_{ii}\mathbf{x}_i(t) + \sqrt{\rho}_i \mathbf{H}_{ii}\mathbf{c}_i(t) + \sqrt{\eta}_i \Delta\mathbf{H}_{ij}\mathbf{x}_j(t) \\ &\quad + \sqrt{\eta}_i \mathbf{H}_{ij}\mathbf{c}_j(t) + \mathbf{e}_i(t) + \mathbf{n}_i(t) \end{aligned} \tag{4.3}$$

We show in Appendix 4.A that the covariance matrix of $\mathbf{v}_i(t)$ can be approximated as

$$\begin{aligned} \tilde{\mathbf{\Sigma}}_i(t) \approx\ & \rho_i \kappa \tilde{\mathbf{H}}_{ii} \text{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\tilde{\mathbf{H}}_{ii}^H + \rho_i \sigma_e^2 \text{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\}\mathbf{I}_N \\ & + \eta_i \kappa \tilde{\mathbf{H}}_{ij} \text{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\tilde{\mathbf{H}}_{ij}^H + \eta_i \sigma_e^2 \text{tr}\left\{\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right\}\mathbf{I}_N \\ & + \beta\rho_i \text{diag}\left(\tilde{\mathbf{H}}_{ii}\mathbf{V}_i(t)\mathbf{V}_i(t)^H\tilde{\mathbf{H}}_{ii}^H\right) + \beta\eta_i \text{diag}\left(\tilde{\mathbf{H}}_{ij}\mathbf{V}_j(t)\mathbf{V}_j(t)^H\tilde{\mathbf{H}}_{ij}^H\right) \\ & + \mathbf{I}_N \end{aligned}$$

As a result of the channel estimation errors and limited dynamic ranges in (4.3), the noise $\mathbf{v}_i(t)$ is generally non-Gaussian. To the best of our knowledge, the exact capacity of MIMO channels with channel estimation errors is still an open problem even for point-to-point MIMO systems [192, 195]. However, assuming $\mathbf{v}_i(t)$ as Gaussian, we can obtain their useful lower bounds [195]. A lower bound of the achievable rate of the ith node at time t can be written as

$$I_i(t) = \log_2 \left| \mathbf{I}_N + \rho_i \tilde{\mathbf{H}}_{ii} \mathbf{V}_i(t) \mathbf{V}_i^H(t) \tilde{\mathbf{H}}_{ii}^H \tilde{\mathbf{\Sigma}}_i(t)^{-1} \right| \tag{4.4}$$

At time t, node i applies the linear receiver $\mathbf{R}_i(t),\ (i,\ t) \in \{1, 2\}$. That is

$$\begin{aligned} \hat{\mathbf{d}}_i(t) &= \mathbf{R}_i(t) \tilde{\mathbf{y}}_i(t) \qquad (4.5) \\ &= \sqrt{\rho_i} \mathbf{R}_i(t) \tilde{\mathbf{H}}_{ii} \mathbf{V}_i(t) \mathbf{d}_i(t) + \mathbf{R}_i(t) \mathbf{v}_i(t) \end{aligned}$$

We can now formulate the MSE of the ith transmitter-receiver pair at time slot t. Using (4.5), the MSE matrix of the ith receiver at time t can be written as

$$\begin{aligned} \mathbf{MSE}_i(t) &= \mathtt{E} \left\{ \left(\hat{\mathbf{d}}_i(t) - \mathbf{d}_i(t) \right) \left(\hat{\mathbf{d}}_i(t) - \mathbf{d}_i(t) \right)^H \right\} \qquad (4.6) \\ &= \left(\sqrt{\rho_i} \mathbf{R}_i(t) \tilde{\mathbf{H}}_{ii} \mathbf{V}_i(t) - \mathbf{I}_N \right) \left(\sqrt{\rho_i} \mathbf{R}_i(t) \tilde{\mathbf{H}}_{ii} \mathbf{V}_i(t) - \mathbf{I}_N \right)^H + \mathbf{R}_i(t) \tilde{\mathbf{\Sigma}}_i(t) \mathbf{R}_i(t)^H \end{aligned}$$

4.3 Weighted-Sum-Rate Maximization

WSR optimization scheme is formulated as follows

$$\max_{\mathbf{V}_i(t)} \quad \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mu_i(t)\, I_i(t) \tag{4.7}$$

$$\text{s.t} \quad \frac{1}{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\} \leq P_i,\ i=1,2 \tag{4.8}$$

$$\text{or s.t} \quad \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\} \leq P_T \tag{4.9}$$

where P_i is the power constraint at the ith transmitter, P_T is the total power constraint of the system and $\mu_i(t) \geq 0$ denotes the weight.

To understand the link between the WSR maximization and the WMMSE minimization problems in the FD bi-directional MIMO channels, we need to establish the relationship between the achievable rate and the error covariance matrix. This argument is parallel to the one given in [205] for the MIMO broadcast channel and in [206] for the MIMO interference channel. The MMSE receiver filter applied at node i at time t can be expressed as

$$\begin{aligned}\mathbf{R}_i^{opt}(t) &= \arg\min_{\mathbf{R}_i(t)} \mathbf{MSE}_i(t) \\ &= \sqrt{\rho_i}\mathbf{V}_i^H(t)\tilde{\mathbf{H}}_{ii}^H\left(\rho_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i(t)\mathbf{V}_i^H(t)\tilde{\mathbf{H}}_{ii}^H + \tilde{\mathbf{\Sigma}}_i(t)\right)^{-1}\end{aligned} \tag{4.10}$$

Plugging (4.10) in (4.6), we can write (4.6) as

$$\mathbf{E}_i(t) = \left(\mathbf{I}_N + \rho_i\mathbf{V}_i^H(t)\tilde{\mathbf{H}}_{ii}^H\tilde{\mathbf{\Sigma}}_i(t)^{-1}\tilde{\mathbf{H}}_{ii}\mathbf{V}_i(t)\right)^{-1} \tag{4.11}$$

where matrix inversion lemma $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1}\right)^{-1}\mathbf{DA}^{-1}$ is applied in the second equality.

Comparing (4.4) and (4.11), it is easy to see the relationship between the achievable rate and the error covariance matrix as

$$I_i(t) = \log_2 \left| \mathbf{E}_i(t)^{-1} \right| \tag{4.12}$$

4.3.1 MSE Weight Design

WMMSE problem can be formulated as

$$\min_{\mathbf{V}_i(t)} \quad \frac{1}{2} \sum_{i=1}^{2} \sum_{t=1}^{2} \mathrm{tr} \left\{ \mathbf{W}_i(t) \mathbf{E}_i(t) \right\} \tag{4.13}$$

$$\text{s.t} \quad \frac{1}{2} \sum_{t=1}^{2} \mathrm{tr} \left\{ \mathbf{V}_i(t) \mathbf{V}_i(t)^H \right\} \leq P_i, \ i = 1, 2 \tag{4.14}$$

$$\text{or s.t} \quad \frac{1}{2} \sum_{i=1}^{2} \sum_{t=1}^{2} \mathrm{tr} \left\{ \mathbf{V}_i(t) \mathbf{V}_i(t)^H \right\} \leq P_T \tag{4.15}$$

where $\mathbf{W}_i(t) \in \mathbb{C}^{N \times N}$ is a constant weight matrix associated with node i at time t.

The Lagrangian functions of the optimization problems (4.7)-(4.9) and (4.13)-(4.15) can be written as

$$\mathcal{L}_{WSR} = -\frac{1}{2} \sum_{i=1}^{2} \sum_{t=1}^{2} \mu_i(t) \, I_i(t) + \sum_{i=1}^{2} \lambda_i \left(\frac{1}{2} \sum_{t=1}^{2} \mathrm{tr} \left\{ \mathbf{V}_i(t) \mathbf{V}_i(t)^H \right\} - P_i \right) \tag{4.16}$$

$$\mathcal{L}_{WMMSE} = \frac{1}{2} \sum_{i=1}^{2} \sum_{t=1}^{2} \mathrm{tr} \left\{ \mathbf{W}_i(t) \mathbf{E}_i(t) \right\} + \sum_{i=1}^{2} \lambda_i \left(\frac{1}{2} \sum_{t=1}^{2} \mathrm{tr} \left\{ \mathbf{V}_i(t) \mathbf{V}_i(t)^H \right\} - P_i \right) \tag{4.17}$$

where Q selects the desired power constraint ($Q = 1$ for the sum power constraint and $Q = 0$ for the individual power constraint), λ and λ_i denote the Lagrange multipliers for the sum power constraint and individual power constraint at the ith node, respectively. The gradients of both Lagrangian functions (4.16) and (4.17) with respect to $\mathbf{V}_i(t)$ can be

written as

$$\begin{aligned}\frac{\partial \mathcal{L}_{WSR}}{\partial \mathbf{V}_i^*(t)} &= -\frac{1}{2\ln 2}\left(\sum_{i=1}^{2}\mu_i(t)\frac{\mathrm{tr}\left\{\mathbf{E}_i(t)\partial\mathbf{E}_i(t)^{-1}\right\}}{\partial\mathbf{V}_i^*(t)}\right)+\frac{\lambda_i}{2}\mathbf{V}_i(t)\\ \frac{\partial \mathcal{L}_{WMMSE}}{\partial \mathbf{V}_i^*(t)} &= -\frac{1}{2}\left(\sum_{i=1}^{2}\frac{\mathrm{tr}\left\{\mathbf{W}_i(t)\mathbf{E}_i(t)\partial\mathbf{E}_i(t)^{-1}\mathbf{E}_i(t)\right\}}{\partial\mathbf{V}_i^*(t)}\right)+\frac{\lambda_i}{2}\mathbf{V}_i(t) \qquad (4.18)\end{aligned}$$

where we have used the matrix derivative formulas $\partial \ln|\mathbf{X}| = \mathrm{tr}\left\{\mathbf{X}^{-1}\partial\mathbf{X}\right\}$ and $\partial\mathbf{X}^{-1} = \mathbf{X}^{-1}\partial\mathbf{X}\mathbf{X}^{-1}$.

Comparing (4.18) and (4.18), we can see that given transmit filters $\mathbf{V}_i(t),\ (i,t)\in\{1,2\}$ and MMSE error covariance matrices $\mathbf{E}_i(t),\ (i,t)\in\{1,2\}$, the gradient of WSR and the gradient of WMMSE problems are equal if the MSE-weights $\mathbf{W}_i(t),\ (i,t)\in\{1,2\}$ are chosen as:

$$\mathbf{W}_i(t) = \frac{\mu_i(t)}{\ln 2}\mathbf{E}_i(t)^{-1} \qquad (4.19)$$

Since the KKT-conditions of the WSR and WMMSE problems can be satisfied simultaneously with the choice of MSE-weights (4.19), we can solve the WSR problem (4.7)-(4.9) through solving WMMSE problem (4.13)-(4.15).

4.3.2 Sum-power constrained transceiver design

The problem to find the optimal transmit filters $\mathbf{V}_i(t)$ for fixed receive filters under the sum-power constraint of the system is formulated as below:

$$\min_{\mathbf{V}_i(t)} \quad \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{W}_i(t)\mathtt{E}\left\{\left(\mathbf{d}_i(t)-\alpha^{-1}\hat{\mathbf{d}}_i(t)\right)\left(\mathbf{d}_i(t)-\alpha^{-1}\hat{\mathbf{d}}_i(t)\right)^H\right\}\right\} \qquad (4.20)$$

$$\text{s.t} \quad \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\} \le P_T \qquad (4.21)$$

where $\mathbf{W}_i(t)$ is chosen according to (4.19) and α is a scaling parameter. Similar to [207], where the optimal transmit filters are computed for the unweighted case, the WMMSE transmit filter of (4.20)-(4.21) can be shown to be

$$\mathbf{V}_i(t) = \alpha \bar{\mathbf{V}}_i(t) \tag{4.22}$$

Here, $\alpha = \sqrt{\frac{P_T}{\frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\right\}}}$ and as shown in Appendix 4.B, $\bar{\mathbf{V}}_i(t)$ is computed as

$$\begin{aligned}\bar{\mathbf{V}}_i(t) &= \sqrt{\rho_i}\left(\mathbf{X}_i(t) + \frac{\frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{W}_i(t)\mathbf{R}_i(t)\mathbf{R}_i(t)^H\right\}}{P_T}\mathbf{I}_N\right)^{-1} \\ &\quad \times \tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i(t)^H\mathbf{W}_i(t)\end{aligned} \tag{4.23}$$

where $\mathbf{X}_i(t)$ is given by

$$\begin{aligned}\mathbf{X}_i(t) &= \rho_i\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i(t)^H\mathbf{W}_i(t)\mathbf{R}_i(t)\tilde{\mathbf{H}}_{ii} + \rho_i\kappa\mathrm{diag}\left(\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i(t)^H\mathbf{W}_i(t)\mathbf{R}_i(t)\tilde{\mathbf{H}}_{ii}\right) \\ &\quad + \rho_i\sigma_e^2\mathrm{tr}\left\{\mathbf{R}_i(t)^H\mathbf{W}_i(t)\mathbf{R}_i(t)\right\}\mathbf{I}_N + \rho_i\beta\tilde{\mathbf{H}}_{ii}^H\mathrm{diag}\left(\mathbf{R}_i(t)^H\mathbf{W}_i(t)\mathbf{R}_i(t)\right)\tilde{\mathbf{H}}_{ii} \\ &\quad + \eta_j\kappa\mathrm{diag}\left(\tilde{\mathbf{H}}_{ji}^H\mathbf{R}_j(t)^H\mathbf{W}_j(t)\mathbf{R}_j(t)\tilde{\mathbf{H}}_{ji}\right) + \eta_j\ \sigma_e^2\mathrm{tr}\left\{\mathbf{R}_j(t)^H\mathbf{W}_j(t)\mathbf{R}_j(t)\right\}\mathbf{I}_N \\ &\quad + \eta_j\beta\tilde{\mathbf{H}}_{ji}^H\mathrm{diag}\left(\mathbf{R}_j(t)^H\mathbf{W}_j(t)\mathbf{R}_j(t)\right)\tilde{\mathbf{H}}_{ji}\end{aligned} \tag{4.24}$$

4.3.3 Individual-power-constrained transceiver design

The problem to find the optimal transmit filters $\mathbf{V}_i(t)$ for fixed receive filters under the individual-power constraint at each node of the system is formulated as below:

$$\min_{\mathbf{V}_i(t)} \quad \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{W}_i(t)\mathtt{E}\left\{\left(\mathbf{d}_i(t)-\hat{\mathbf{d}}_i(t)\right)\left(\mathbf{d}_i(t)-\hat{\mathbf{d}}_i(t)\right)^H\right\}\right\} \tag{4.25}$$

$$\text{s.t} \quad \frac{1}{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\} \le P_i,\ i=1,2 \tag{4.26}$$

Taking the partial derivative of the Lagrange function of (4.25)-(4.26) with respect to the matrix $\mathbf{V}_i(t)$, we can obtain the optimal $\mathbf{V}_i(t)$ as

$$\mathbf{V}_i(t) = \sqrt{\rho_i}\left(\lambda_i \mathbf{I}_N + \mathbf{X}_i(t)\right)^{-1} \tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i(t)^H \mathbf{W}_i(t) \tag{4.27}$$

where $\mathbf{X}_i(t)$ is defined in (4.24). The values of the Lagrange multiplier $\lambda_i,\ i = 1, 2$ in (4.27) are calculated similar to [208] by taking the singular value decomposition of $\mathbf{X}_i(t) = \mathbf{U}_i(t)\mathbf{\Delta}_i(t)\left(\mathbf{U}_i(t)\right)^H$ and writing the power constraint in (4.26), after simple steps, as

$$\frac{1}{2}\sum_{t=1}^{2} \mathrm{tr}\left\{\mathbf{V}_i(t)\left(\mathbf{V}_i(t)\right)^H\right\} = \frac{1}{2}\rho_i \sum_{t=1}^{2}\sum_{k=1}^{N} \frac{g_{ik}(t)}{\left(\lambda_i + \Delta_{ik}(t)\right)^2} = P_i \tag{4.28}$$

where $g_{ik}(t)$ is the kth row and kth column of $\mathbf{U}_i(t)^H \tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i(t)^H \mathbf{W}_i(t) \mathbf{W}_i(t)^H \mathbf{R}_i(t) \tilde{\mathbf{H}}_{ii} \mathbf{U}_i(t)$ and $\Delta_{ik}(t)$ denotes the kth row and kth column element of the matrix $\mathbf{\Delta}_i(t)$. We can compute $\lambda_i,\ i = 1, 2$ from (4.28) numerically. If the values of the Lagrange multipliers $\lambda_i,\ i = 1, 2$ are negative, we assign $\lambda_i,\ i = 1, 2$ as zeros.

The iterative alternating algorithm for the WSR optimization problem (4.7)-(4.9) through WMMSE minimization problem is given in Table 4.1. The algorithm in Table 4.1 holds for both the sum-power constraint and the individual-power-constraint WSR problems. With the same logic in [205], by manipulating the WSR maximization problem of (4.7)-(4.9), we can include the MMSE weights and receive filters as new optimization variables. By showing that this new optimization problem converges monotonically, we can prove that the algorithm in Table 4.1 is guaranteed to converge to a local optimum. Since the WMMSE minimization (4.13)-(4.15) problem is not jointly convex over optimization variables, the proposed algorithm does not ensure to converge to the global optimal so-

lution. Because of this non-convexity of the optimization problems we are dealing with, we need to choose good initialization points to have a suboptimal solution with a good performance.

4.4 Simulation Results

In this section, we numerically investigate the WSR optimization problem for MIMO FD bi-directional systems as a function of signal-to-noise ratio (SNR), nominal interference to noise ratio (INR), number of antennas N and channel estimation errors σ_e^2. For brevity, we set the same average transmit power for each node $P_1 = P_2 = N$ and $P_T = 2N$. We also assumed $\rho_1 = \rho_2 = \text{SNR}/N$, $\eta_1 = \eta_2 = \text{INR}/N$, and $\mu_i(t) = \mu,\ (i,t) = 1,2$. To optimize the HD scheme, we solved the WSR optimization problem with the HD constraint that $\mathbf{V}_1(2) = \mathbf{V}_2(1) = \mathbf{0}$. Note that HD scheme is invariant to INR. To show the importance of using two time slots, we also compared our FD scheme with two data transmission slots (denoted as FD2) with the FD system with only one data transmission slot (denoted as FD1), where we assumed the same source covariance matrices for both time slots, $\mathbf{V}_1(1) = \mathbf{V}_1(2)$ and $\mathbf{V}_2(1) = \mathbf{V}_2(2)$. Since the optimization problems we are dealing with are non-convex, we need to choose good initialization points to have a suboptimal solution with a good performance. We choose 30 random initialization points and picked the one that gives the best performance. The results are averaged over 100 independent channel realizations.

In the first example, we investigate the impact of INR on the sum-rate of the

Table 4.1: WSR maximization algorithm

1) Set the iteration number $n = 0$ and initialize $\mathbf{V}_i^{[0]}(t),\ (i,t) \in \{1,2\}$
   and calculate $I_{sum}^{[0]} = \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mu_i(t)\, I_i(t)$
2) $n \leftarrow n+1$. Update $\mathbf{R}_i^{[n+1]}(t)$:
   $\mathbf{R}_i^{[n+1]}(t) = \sqrt{\rho_i}\left(\mathbf{V}_i^{[n]}(t)\right)^H \tilde{\mathbf{H}}_{ii}^H \left(\rho_i \tilde{\mathbf{H}}_{ii}\mathbf{V}_i^{[n]}(t)\left(\mathbf{V}_i^{[n]}(t)\right)^H \tilde{\mathbf{H}}_{ii}^H + \tilde{\mathbf{\Sigma}}_i^{[n]}(t)\right)^{-1}$
3) Calculate and update $\mathbf{W}_i^{[n+1]}(t)$:
   $\mathbf{W}_i^{[n+1]}(t) = \frac{\mu_i(t)}{\ln 2}\left(\mathbf{E}_i^{[n]}(t)\right)^{-1},\ (i,t) \in \{1,2\}$
4) Calculate and update $\mathbf{V}_i^{[n+1]}(t)$
   $\mathbf{V}_i^{[n+1]}(t) = \sqrt{\rho_i}\left(\lambda_i \mathbf{I}_N + \mathbf{X}_i^{[n+1]}(t)\right)^{-1} \tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i^{[n+1]}(t)^H \mathbf{W}_i^{[n+1]}(t)$
4) Repeat steps 2, 3 and 4 until convergence, i.e. $\left|I_{sum}^{[n+1]} - I_{sum}^{[n]}\right| \leq \epsilon$, where ϵ is some arbitrarily small value or a predefined number of iterations is reached.

FD2, FD1, and HD schemes with $N = 2$, SNR = 20dB, $\sigma_e^2 = 0$, $\beta = \kappa = -40$dB and $\mu = 0.25$. As expected, it can be observed from Fig. 4.1 that the HD scheme is invariant to INR. For the individual power constrained problem, at the low-to-mid values of INR, the FD2 scheme behaves as a FD system and it switches to the HD scheme at the high values of INR. The FD1 scheme performs similar to the FD2 scheme at low-to-mid values of INR, but its performance drops below that of the HD scheme for larger values of INR. The use of two distinct data time slots gives the freedom to switch to the HD signaling when the power of the self-interference channel is high (where the HD scheme is optimal), while the FD1 system forces FD signaling, regardless of the strength of the self-interference channel. As for the sum-power constrained problem, similar to individual power constrained problem, at the low-to-mid values of INR, the FD2 scheme behaves as a FD system and it converges towards the HD scheme at the high values of INR, but unlike the individual power constrained problem, the sum-rate of FD2 is never equal to the sum-rate of HD

scheme (always higher). The reason that the performance of FD2 is always higher than HD is that at high INR depending on the channel conditions, the total power is allocated to only one of the nodes which has a better channel. For example, if $\mathbf{H}_{ii}$ has more channel gains than $\mathbf{H}_{jj}$, then all the power is given to the ith transmitter. On the other hand, for HD case the two nodes have to transmit regardless of their channel conditions. The FD1 scheme performs similar to FD2 scheme for all values of INR and the sum-rate of FD1 does not drop to zero unlike the individual power constrained case.

In the second example, we investigate the role of channel estimation errors on the sum-rate of FD2 and HD systems for the individual power constrained problem. Here we set $N = 2$ and SNR = 20dB, $\beta = \kappa = -40$dB and $\mu = 0.25$. It can be seen from Fig. 4.2 that as the channel estimation errors increases, the sum-rate of both the FD2 and HD systems decreases. Note that the sum-rate of the FD2 system is always equal to or greater than that of the HD system (the reason is explained in Fig. 4.1). We have a similar result for the sum-power constraint problem, which we have not included to avoid the repetition.

In our third example, we investigate the role of antenna number N on the sum-rate of FD2 and HD systems for the sum-power constrained problem. We set SNR = 20dB and $\sigma_e^2 = 0.01$, $\beta = \kappa = -40$dB and $\mu = 0.25$. It can be seen from Fig. 4.3 that the sum-rate increases with the number of antennas. Note that the sum-rate of the FD2 system is always greater than that of the HD system (the reason is explained in Fig. 4.1). We have a similar result for the individual power constrained problem, which we have not included to avoid the repetition.

In our fourth example, we examine the sum-rate of the FD2 and HD systems versus

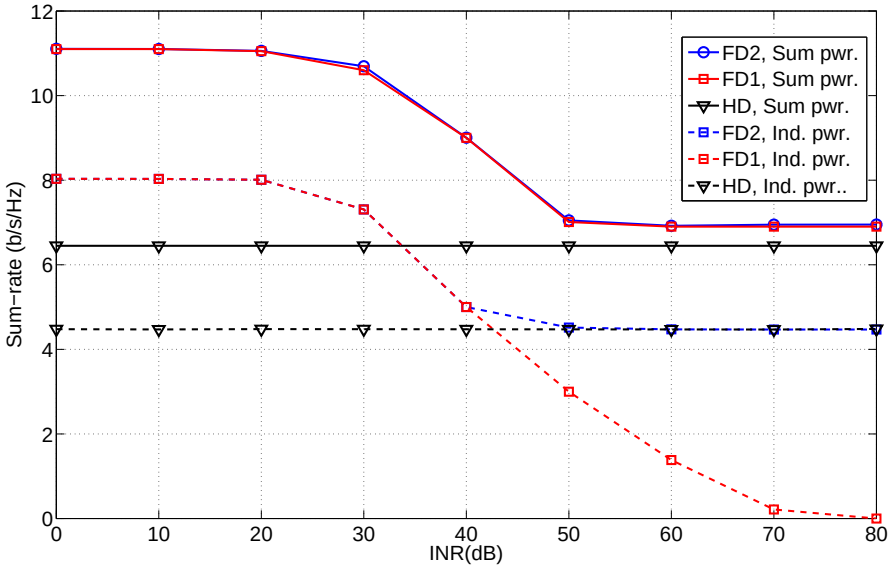

Figure 4.1: Sum-rate comparison of the FD2, FD1, and HD systems versus INR. Here $N = 2$, SNR $= 20$dB, $\sigma_e^2 = 0$, $\kappa = \beta = -40$dB, $\mu = 0.25$.

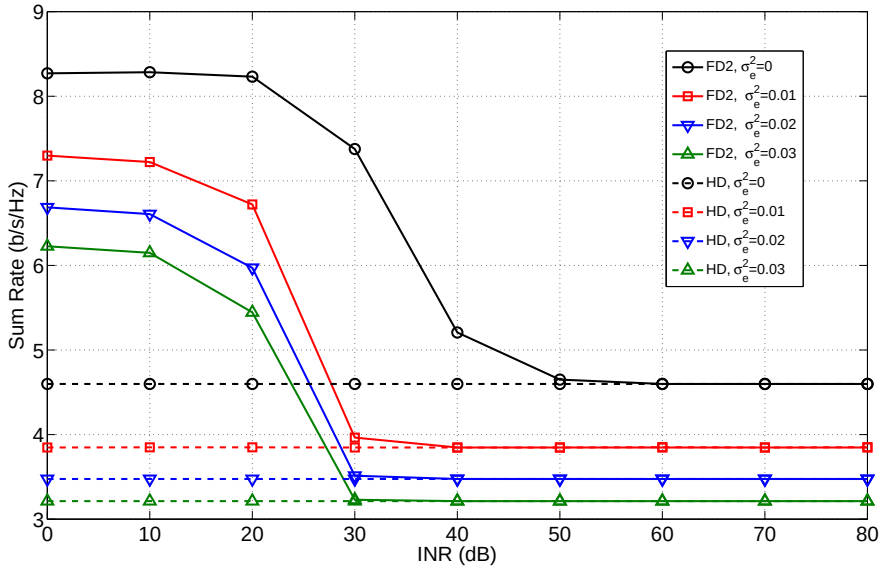

Figure 4.2: Sum-rate comparison of the FD2 and HD systems with different σ_e^2 versus INR for individual power constrained problem. Here $N = 2$, SNR $= 20$dB, $\kappa = \beta = -40$dB and $\mu = 0.25$.

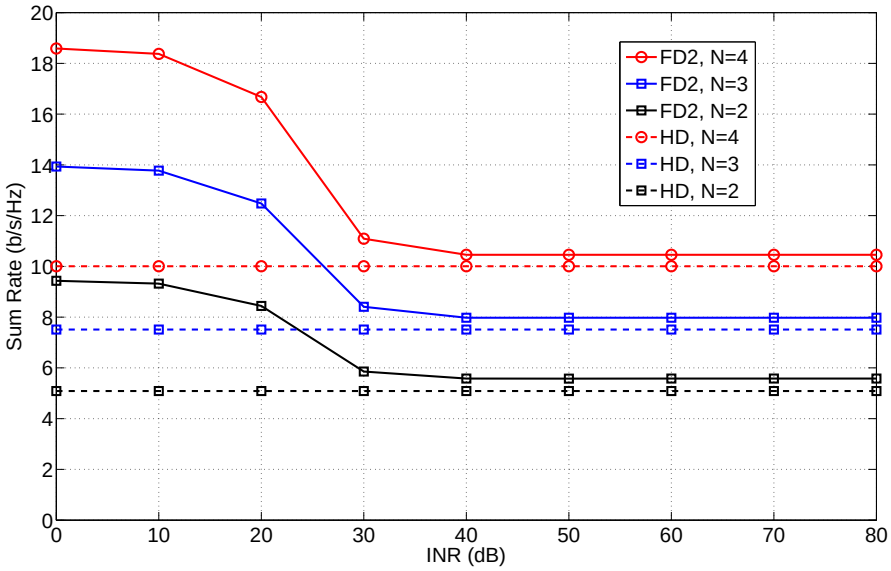

Figure 4.3: Sum-rate comparison of the FD2 and HD systems with different N versus INR for sum-power constrained problem. Here SNR = 20dB, $\sigma_e^2 = 0.01$ $\kappa = \beta = -40$dB and $\mu = 0.25$.

SNR for various fixed values of INR for both sum-power and individual power constrained problems. We choose $N = 2$ and $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB and $\mu = 0.25$. It can be observed from Fig. 4.4 and Fig. 4.5 that at low INR, the system operates in the FD mode for all values of SNR, since SNR mostly dominates INR. At high INR, the system operates in the HD mode for the individual power constrained problem and operates close to but better than HD mode for the sum-power constrained problem (as explained in Fig. 4.1) at low values of SNR (since INR dominates SNR), but switches to the FD mode as SNR increases, since SNR starts to dominate INR.

The cumulative distribution of the sum-rate of each of the tested schemes is plotted in Fig. 4.6 where $N = 2$, SNR = 20dB, INR = 15dB, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB and $\mu = 0.25$.

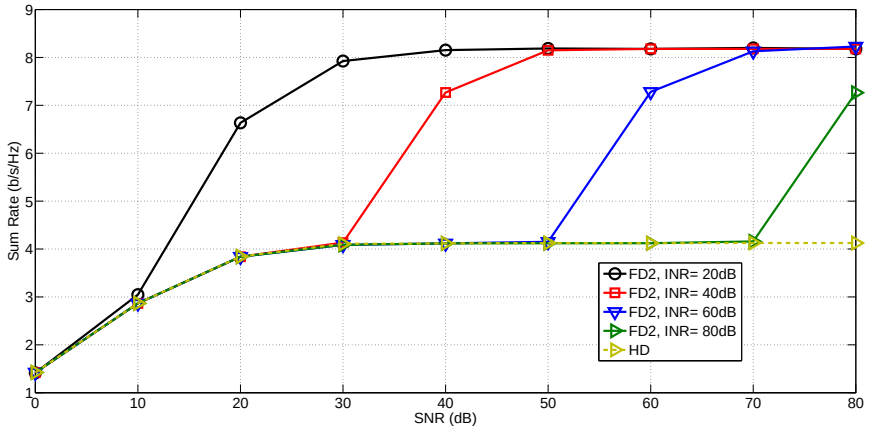

Figure 4.4: Sum-rate comparison of the FD2 and HD systems for different INR values versus SNR for individual power constrained problem. Here $N = 2$, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB and $\mu = 0.25$.

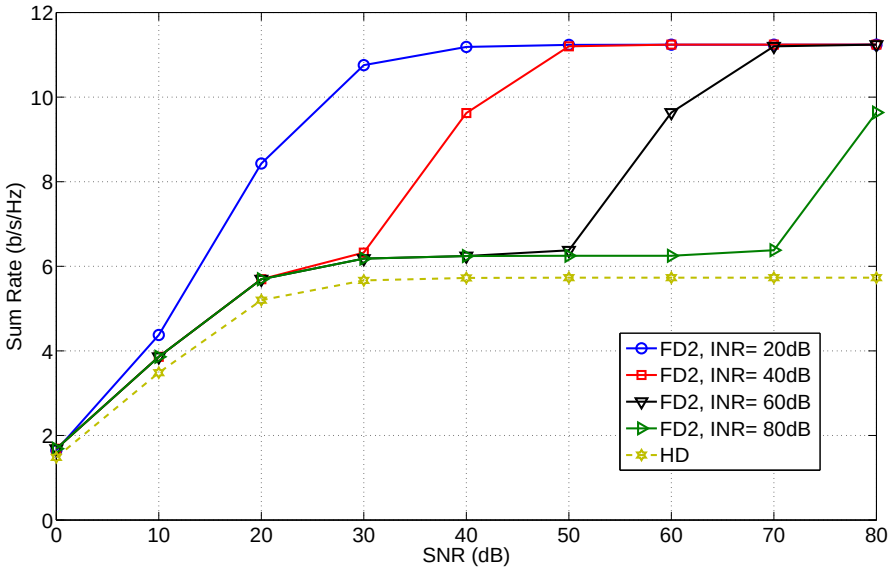

Figure 4.5: Sum-rate comparison of the FD2 and HD systems for different INR values versus SNR for sum-power constrained problem. Here $N = 2$, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB and $\mu = 0.25$.

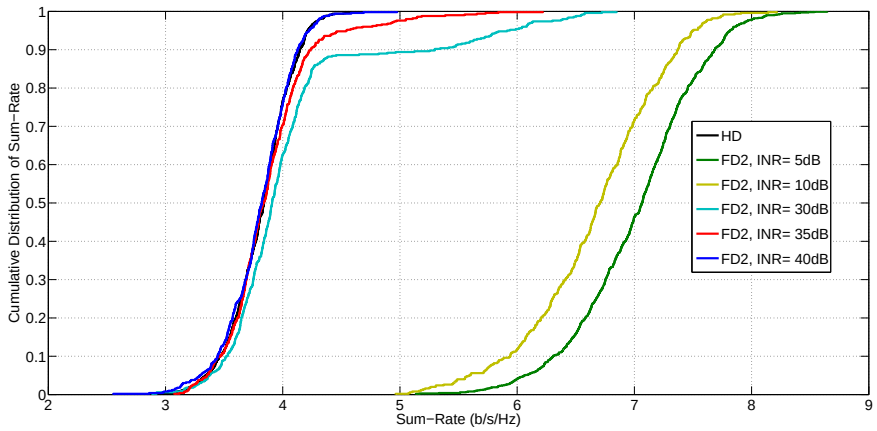

Figure 4.6: Cumulative distribution of the sum-rate. Here $N = 2$, SNR $= 20$dB, INR $=$ 15dB, $\sigma_e^2 = 0.01$, $\kappa = \beta = -40$dB, $\mu = 0.25$.

4.5 Appendix 4.A: Covariance of (4.3)

To calculate the covariance matrix of $\mathbf{v}_i(t)$, we first need to find the covariance of $\mathbf{e}_i(t)$, which is written as

$$
\begin{aligned}
&\mathtt{E}\left\{\mathbf{e}_i(t)\mathbf{e}_i(t)^H|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right\} \\
&= \beta\,\mathrm{diag}\left(\mathrm{Cov}\left(\mathbf{u}_i(t)|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right)\right) \\
&= \beta\,\mathrm{diag}\left(\mathtt{E}_{\Delta\mathbf{H}_{ii}}\left\{\rho_i\mathbf{H}_{ii}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H+\kappa\,\mathrm{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\right)\mathbf{H}_{ii}^H\right\}\right. \\
&\quad \left.+\,\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\eta_i\mathbf{H}_{ij}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H+\kappa\,\mathrm{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\right)\mathbf{H}_{ij}^H\right\}+\mathbf{I}_N\right) \\
&= \beta\,\mathrm{diag}\Big(\rho_i\tilde{\mathbf{H}}_{ii}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H+\kappa\,\mathrm{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\right)\tilde{\mathbf{H}}_{ii}^H \\
&\quad +\rho_i\sigma_e^2\,\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H+\kappa\,\mathrm{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\right\}\mathbf{I}_N \\
&\quad +\eta_i\tilde{\mathbf{H}}_{ij}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H+\kappa\,\mathrm{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\right)\tilde{\mathbf{H}}_{ij}^H+\mathbf{I}_N \\
&\quad +\eta_i\sigma_e^2\,\mathrm{tr}\left\{\mathbf{V}_j(t)\mathbf{V}_j(t)^H+\kappa\,\mathrm{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\right\}\mathbf{I}_N\Big)
\end{aligned}
\tag{4.29}
$$

Here, the first expectation is taken with respect to $\mathbf{d}_i(t)$, $\mathbf{d}_j(t)$, $\Delta\mathbf{H}_{ii}$, $\Delta\mathbf{H}_{ij}$ and $\mathbf{n}_i(t)$; and we used the identity $\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\mathbf{H}_{ij}\mathbf{A}\mathbf{H}_{ij}^H\right\} = \tilde{\mathbf{H}}_{ij}\mathbf{A}\tilde{\mathbf{H}}_{ij}^H + \sigma_e^2\mathrm{tr}(\mathbf{A})\mathbf{I}_N$ if the entries of $\Delta\mathbf{H}_{ij}$ are i.i.d with $\mathcal{CN}(\mathbf{0},\sigma_e^2)$ and $\mathbf{A}\in\mathbb{C}^{N\times N}$ is a known matrix.

Using $\beta \ll 1$ and $\kappa \ll 1$, (4.29) can be approximated as

$$\begin{aligned}
&\mathtt{E}\left\{\mathbf{e}_i(t)\mathbf{e}_i(t)^H|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right\}\\
&\approx \beta\left(\rho_i\mathrm{diag}\left(\tilde{\mathbf{H}}_{ii}\mathbf{V}_i(t)\mathbf{V}_i(t)^H\tilde{\mathbf{H}}_{ii}^H\right) + \rho_i\sigma_e^2\,\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\}\mathbf{I}_N\right.\\
&\left.+\,\eta_i\mathrm{diag}\left(\tilde{\mathbf{H}}_{ij}\mathbf{V}_j(t)\mathbf{V}_j(t)^H\tilde{\mathbf{H}}_{ij}^H\right) + \eta_i\sigma_e^2\,\mathrm{tr}\left\{\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right\}\mathbf{I}_N + \mathbf{I}_N\right) \qquad (4.30)
\end{aligned}$$

Using (4.3), the covariance matrix of $\mathbf{v}_i(t)$ can be written as

$$\begin{aligned}
\tilde{\mathbf{\Sigma}}_i(t) &= \mathtt{E}\left\{\mathbf{v}_i(t)\mathbf{v}_i(t)^H|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right\}\\
&= \rho_i\mathtt{E}_{\Delta\mathbf{H}_{ii}}\left\{\Delta\mathbf{H}_{ii}\mathbf{V}_i(t)\mathbf{V}_i(t)^H\Delta\mathbf{H}_{ii}^H\right\} + \kappa\rho_i\mathtt{E}_{\Delta\mathbf{H}_{ii}}\left\{\mathbf{H}_{ii}\mathrm{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\mathbf{H}_{ii}^H\right\}\\
&\quad+\eta_i\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\Delta\mathbf{H}_{ij}\mathbf{V}_j(t)\mathbf{V}_j(t)^H\Delta\mathbf{H}_{ij}^H\right\} + \kappa\eta_i\mathtt{E}_{\Delta\mathbf{H}_{ij}}\left\{\mathbf{H}_{ij}\mathrm{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\mathbf{H}_{ij}^H\right\}\\
&\quad+\mathtt{E}\left\{\mathbf{e}_i(t)\mathbf{e}_i(t)^H|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right\} + \mathbf{I}_N\\
&= \rho_i\sigma_e^2\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\}\mathbf{I}_N + \rho_i\kappa\left(\tilde{\mathbf{H}}_{ii}\mathrm{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\tilde{\mathbf{H}}_{ii}^H\right.\\
&\quad\left.+\,\sigma_e^2\mathrm{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\}\mathbf{I}_N\right) + \eta_i\sigma_e^2\mathrm{tr}\left\{\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right\}\mathbf{I}_N\\
&\quad+\eta_i\kappa\left(\tilde{\mathbf{H}}_{ij}\mathrm{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\tilde{\mathbf{H}}_{ij}^H + \sigma_e^2\mathrm{tr}\left\{\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right\}\mathbf{I}_N\right)\\
&\quad+\mathtt{E}\left\{\mathbf{e}_i(t)\mathbf{e}_i(t)^H|\tilde{\mathbf{H}}_{ii},\tilde{\mathbf{H}}_{ij}\right\} + \mathbf{I}_N \qquad (4.31)
\end{aligned}$$

By plugging (4.30) in (4.31), (4.31) can be written as

$$\begin{aligned}\tilde{\mathbf{\Sigma}}_i(t) &= \rho_i\kappa\tilde{\mathbf{H}}_{ii}\text{diag}\left(\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right)\tilde{\mathbf{H}}_{ii}^H + \rho_i\sigma_e^2\text{tr}\left\{\mathbf{V}_i(t)\mathbf{V}_i(t)^H\right\}(1+\kappa+\beta)\mathbf{I}_N \\ &+ \eta_i\kappa\tilde{\mathbf{H}}_{ij}\text{diag}\left(\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right)\tilde{\mathbf{H}}_{ij}^H + \eta_i\sigma_e^2\text{tr}\left\{\mathbf{V}_j(t)\mathbf{V}_j(t)^H\right\}(1+\kappa+\beta)\mathbf{I}_N \\ &+ \beta\rho_i\text{diag}\left(\tilde{\mathbf{H}}_{ii}\mathbf{V}_i(t)\mathbf{V}_i(t)^H\tilde{\mathbf{H}}_{ii}^H\right) + \beta\eta_i\text{diag}\left(\tilde{\mathbf{H}}_{ij}\mathbf{V}_j(t)\mathbf{V}_j(t)^H\tilde{\mathbf{H}}_{ij}^H\right) \\ &+ (\beta+1)\mathbf{I}_N \end{aligned} \tag{4.32}$$

Using $\beta \ll 1$ and $\kappa \ll 1$, (4.32) can be approximated as (4.4).

4.6 Appendix 4.B: Derivation of (4.23)

The Lagrangian function of the optimization problem (4.20)-(4.21) can be expressed as

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\text{tr}\Big\{\mathbf{W}_i(t)\Big(\mathbf{I}_N - \sqrt{\rho_i}\bar{\mathbf{V}}_i(t)^H\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i(t)^H \\ &\quad - \sqrt{\rho_i}\mathbf{R}_i(t)\tilde{\mathbf{H}}_{ii}\bar{\mathbf{V}}_i(t) + \mathbf{R}_i(t)\bar{\tilde{\mathbf{\Sigma}}}_i(t)\mathbf{R}_i(t)^H + \rho_i\mathbf{R}_i(t)\tilde{\mathbf{H}}_{ii}\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i(t)^H\Big)\Big\} \\ &\quad + \lambda\left(\frac{\alpha^2}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\text{tr}\left\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\right\} - P_T\right)\end{aligned} \tag{4.33}$$

where

$$\begin{aligned}\bar{\tilde{\mathbf{\Sigma}}}_i(t) &= \rho_i\kappa\tilde{\mathbf{H}}_{ii}\text{diag}\big(\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\big)\tilde{\mathbf{H}}_{ii}^H + \rho_i\sigma_e^2\text{tr}\big\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\big\}\mathbf{I}_N \\ &+ \eta_i\kappa\tilde{\mathbf{H}}_{ij}\text{diag}\big(\bar{\mathbf{V}}_j(t)\bar{\mathbf{V}}_j(t)^H\big)\tilde{\mathbf{H}}_{ij}^H + \eta_i\sigma_e^2\text{tr}\big\{\bar{\mathbf{V}}_j(t)\bar{\mathbf{V}}_j(t)^H\big\}\mathbf{I}_N \\ &+ \beta\rho_i\text{diag}\left(\tilde{\mathbf{H}}_{ii}\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\tilde{\mathbf{H}}_{ii}^H\right) + \alpha^{-2}\mathbf{I}_N \\ &+ \beta\eta_i\text{diag}\left(\tilde{\mathbf{H}}_{ij}\bar{\mathbf{V}}_j(t)\bar{\mathbf{V}}_j(t)^H\tilde{\mathbf{H}}_{ij}^H\right),\ (i,j)\in\{1,2\}, j\neq i\end{aligned} \tag{4.34}$$

Taking the partial derivative of (4.33) with respect to the matrix $\bar{\mathbf{V}}_i(t)$, we obtain

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \bar{\mathbf{V}}_i^*(t)} &= \frac{1}{2}\Big(-\sqrt{\rho_i}\tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i(t)^H \mathbf{W}_i(t) + \rho_i \tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i(t)^H \mathbf{W}_i(t)\mathbf{R}_i(t)\tilde{\mathbf{H}}_{ii}\bar{\mathbf{V}}_i(t) \\ &\quad + \frac{\mathrm{tr}\left\{\mathbf{R}_i(t)^H \mathbf{W}_i(t)\mathbf{R}_i(t)\partial\bar{\tilde{\mathbf{\Sigma}}}_i(t)\right\}}{\partial \bar{\mathbf{V}}_i^*(t)} + \frac{\mathrm{tr}\left\{\mathbf{R}_j(t)^H \mathbf{W}_j(t)\mathbf{R}_j(t)\partial\bar{\tilde{\mathbf{\Sigma}}}_j(t)\right\}}{\partial \bar{\mathbf{V}}_i^*(t)}\Big) \\ &\quad + \frac{\lambda\alpha^2}{2}\bar{\mathbf{V}}_i(t) \end{aligned} \tag{4.35}$$

Using (4.34) , we can write

$$\begin{aligned}\frac{\mathrm{tr}\left\{\mathbf{A}_i(t)\partial\bar{\tilde{\mathbf{\Sigma}}}_i(t)\right\}}{\partial \bar{\mathbf{V}}_i^*(t)} &= \Big(\rho_i\kappa \mathrm{diag}\left(\tilde{\mathbf{H}}_{ii}^H \mathbf{A}_i(t)\tilde{\mathbf{H}}_{ii}\right) + \rho_i\sigma_e^2 \mathrm{tr}\left\{\mathbf{A}_i(t)\right\}\mathbf{I}_N \\ &\quad + \rho_i\beta\tilde{\mathbf{H}}_{ii}^H \mathrm{diag}\left(\mathbf{A}_i(t)\right)\tilde{\mathbf{H}}_{ii}\Big)\bar{\mathbf{V}}_i(t) \end{aligned} \tag{4.36}$$

$$\begin{aligned}\frac{\mathrm{tr}\left\{\mathbf{A}_j(t)\partial\bar{\tilde{\mathbf{\Sigma}}}_j(t)\right\}}{\partial \bar{\mathbf{V}}_i^*(t)} &= \Big(\eta_j\kappa \mathrm{diag}\left(\tilde{\mathbf{H}}_{ji}^H \mathbf{A}_j(t)\tilde{\mathbf{H}}_{ji}\right) + \eta_j\sigma_e^2 \mathrm{tr}\left\{\mathbf{A}_j(t)\right\}\mathbf{I}_N \\ &\quad + \eta_j\beta\tilde{\mathbf{H}}_{ji}^H \mathrm{diag}\left(\mathbf{A}_j(t)\right)\tilde{\mathbf{H}}_{ji}\Big)\bar{\mathbf{V}}_i(t) \end{aligned} \tag{4.37}$$

where $\mathbf{A}_i(t) = \mathbf{R}_i(t)^H \mathbf{W}_i(t)\mathbf{R}_i(t),\ (i,t) \in \{1,2\}$. By plugging (4.36) and (4.37) into (4.35) and making it equal to zero, we can obtain the optimal $\bar{\mathbf{V}}_i(t)$ as

$$\bar{\mathbf{V}}_i(t) = \sqrt{\rho_i}\left(\mathbf{X}_i(t) + \lambda\alpha^2\mathbf{I}_N\right)^{-1}\tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i(t)^H \mathbf{W}_i(t) \tag{4.38}$$

where $\mathbf{X}_i(t)$ is defined in (4.24).

Taking the derivative of the Lagrange function (4.33) with respect to α, we obtain

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \alpha} &= -\alpha^{-3}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{W}_i(t)\mathbf{R}_i(t)\mathbf{R}_i(t)^H\right\} + \lambda\alpha\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\right\} \\ &= 0 \\ &\Rightarrow \lambda\alpha^2\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\right\} \\ &\quad = \alpha^{-2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{W}_i(t)\mathbf{R}_i(t)\mathbf{R}_i(t)^H\right\} \end{aligned} \tag{4.39}$$

The complementary slackness condition of (4.20)-(4.21) is:

$$\lambda\left(\alpha^2\frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\right\}-P_T\right)=0$$
$$\Rightarrow \lambda\alpha^2\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\bar{\mathbf{V}}_i(t)\bar{\mathbf{V}}_i(t)^H\right\}=2\lambda P_T \tag{4.40}$$

Plugging (4.39) into (4.40), we obtain

$$\lambda\alpha^2=\frac{\frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{2}\mathrm{tr}\left\{\mathbf{W}_i(t)\mathbf{R}_i(t)\mathbf{R}_i(t)^H\right\}}{P_T} \tag{4.41}$$

Substituting (4.41) into (4.38), we get the desired result (4.23).

Chapter 5

MSE Based Transceiver Designs for Bi-directional Full-Duplex MIMO Systems with Residual Self-Interference

We consider a FD bi-directional communication system between two nodes that suffer from self-interference, where the nodes are equipped with multiple antennas and instantaneous channel state information (CSI) at the nodes is imperfect. We focus on the effect of the residual self-interference due to independent and identically distributed (i.i.d.) channel estimation errors and limited dynamic ranges of the transmitters and receivers. We address the minimization of sum mean-squared error (MSE) and the maximum per-

node MSE optimization problems subject to power constraints at each node. For all the problems considered, we show that joint design of transceiver matrices can be obtained through efficient iterative alternating algorithms, which are guaranteed to converge to at least a local optimal solution.

5.1 Introduction

As discussed in Chapter 2, most of the works on FD systems have focused on the maximization of the achievable rate and MSE transceiver designs have not been studied in detail. In this chapter, we propose a joint and iterative transceiver design method for the MIMO bi-directional FD systems by taking the imperfect channel knowledge and limited dynamic ranges of the transmitter and receivers into account. We consider both the sum-MSE and the maximum per-node MSE as the objective functions to minimize subject to power constraints at both nodes. An iterative algorithm which optimizes the transmit precoding and receiving filter matrices alternatingly is proposed. At each iteration, the sum-MSE and maximum MSE decrease monotonically, and are guaranteed to converge to at least a local optimal solution.

The following notations are used in this chapter. The matrices and the vectors are denoted as bold capital and lowercase letters, respectively. $(\cdot)^T$ is the transpose; $(\cdot)^H$ is the conjugate transpose. $\mathtt{E}\{\cdot\}$ means the statistical expectation; $\mathbf{I}_N$ is the N by N identity matrix; $\mathrm{tr}(\cdot)$ is the trace; $|\cdot|$ is the determinant; $\mathrm{diag}(\mathbf{A})$ is the diagonal matrix with the same diagonal elements as $\mathbf{A}$. $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian distribution with mean μ and variance σ^2. $[x]^+$ denotes $\max(x, 0)$. $vec(\cdot)$ stacks the elements of a matrix

in one long column vector. The operator $\otimes$ denotes Kronecker product and $\perp$ denotes the statistical independence. $\|\cdot\|_2$ is the Euclidean norm of a vector. $\lfloor \mathbf{A}_i \rfloor_{i=1}^{K}$ denotes a tall matrix (or vector) obtained by stacking the matrices (or vectors) $\mathbf{A}_i,\ i = 1, \ldots, K$.

5.2 System Model

In this section, we describe the system model of a FD bi-directional MIMO system between two nodes as seen in Fig. 5.1. We consider the same channel estimation error model and transmitter/receiver distortion model adopted in Chapter 4. Similar to the model in Chapter 3 and Chapter 4, we assume that two FD nodes have N transmit/receive antennas. But unlike Chapter 3 and Chapter 4, in this chapter, we consider one-time slot data transmission, and assume that the channel matrices remain constant during one time slot, but change randomly at each time slot.

The data streams at the ith transmitter is denoted as $\mathbf{d}_i \in \mathbb{C}^N,\ i = 1, 2$. The transmit symbols are assumed to be complex, zero mean, independent and identically distributed with

$$\mathtt{E}\{\mathbf{d}_i\} = \mathbf{0} \tag{5.1}$$

$$\mathtt{E}\left\{\mathbf{d}_i \mathbf{d}_j^H\right\} = \begin{cases} \mathbf{I}_N & i = j, \\ \mathbf{0}_N & i \neq j. \end{cases} \tag{5.2}$$

The $N \times 1$ signal vector transmitted by transmitter i is given by

$$\mathbf{x}_i = \mathbf{V}_i \mathbf{d}_i, \quad i = 1, 2 \tag{5.3}$$

where $\mathbf{V}_i \in \mathbb{C}^{N \times N}$ represent precoding matrix. $\mathbf{x}_i$ is assumed to be Gaussian distributed

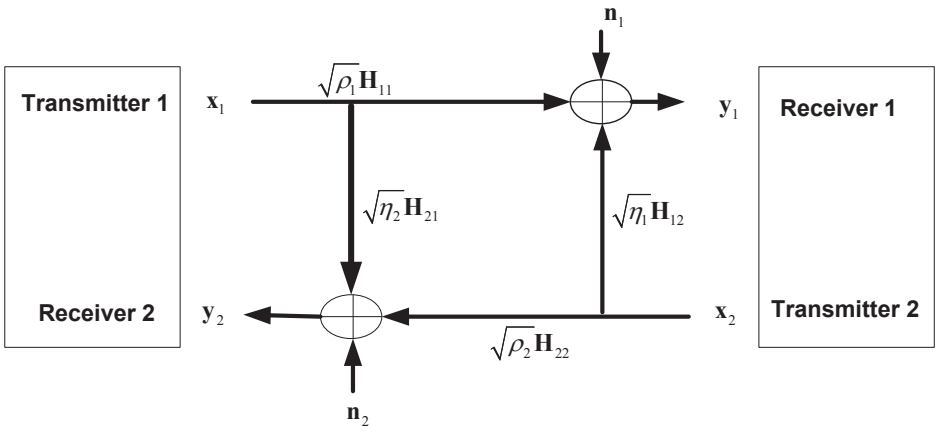

Figure 5.1: A bi-directional full-duplex MIMO system

with zero mean and covariance matrix $\mathtt{E}\left\{\mathbf{x}_i\mathbf{x}_i^H\right\} = \mathbf{V}_i\mathbf{V}_i^H$.

We consider a FD bi-directional MIMO system that suffers from self-interference. Thus, the receiver i receives a combination of the signals transmitted by both transmitters and noise. The $N \times 1$ received signal at the ith receiver is written as

$$\begin{aligned}\mathbf{y}_i &= \sqrt{\rho_i}\mathbf{H}_{ii}\left(\mathbf{x}_i + \mathbf{c}_i\right) + \sqrt{\eta_i}\mathbf{H}_{ij}\left(\mathbf{x}_j + \mathbf{c}_j\right) + \mathbf{e}_i + \mathbf{n}_i \\ &= \sqrt{\rho_i}\tilde{\mathbf{H}}_{ii}\mathbf{x}_i + \sqrt{\rho_i}\Delta\mathbf{H}_{ii}\mathbf{x}_i + \sqrt{\rho_i}\mathbf{H}_{ii}\mathbf{c}_i + \sqrt{\eta_i}\tilde{\mathbf{H}}_{ij}\mathbf{x}_j \\ &\quad + \sqrt{\eta_i}\Delta\mathbf{H}_{ij}\mathbf{x}_j + \sqrt{\eta_i}\mathbf{H}_{ij}\mathbf{c}_j + \mathbf{e}_i + \mathbf{n}_i,\ i = 1, 2 \end{aligned} \tag{5.4}$$

where $\mathbf{x}_j = \mathbf{V}_j\mathbf{d}_j$ is $N \times 1$ signal vector transmitted from the transmitter j, $j \neq i$ and is Gaussian distributed with zero mean and covariance matrix $\mathtt{E}\left\{\mathbf{x}_j\mathbf{x}_j^H\right\} = \mathbf{V}_j\mathbf{V}_j^H$, which incurs self-interference at the ith receiver. Note that, we assume the number of streams is equal to the number of antennas, N. $\mathbf{n}_i \in \mathbb{C}^N$ is the additive white Gaussian noise (AWGN) vector at the ith receiver with zero mean and unit covariance matrix, $\mathtt{E}\left\{\mathbf{n}_i\mathbf{n}_i^H\right\} = \mathbf{I}_N$ and it is uncorrelated to $\mathbf{x}_i$ and $\mathbf{x}_j$. ρ_i denotes the average power gain of the ith transmitter-

receiver link, and η_i denotes the average power gain of the self-interference channel.

$\mathbf{c}_k \in \mathbb{C}^N,\ k = 1, 2$ is the transmitter noise at the kth transmitter, which models the effect of limited transmitter DR and closely approximates the effects of additive power-amplifier noise, non-linearities in the DAC and phase noise [110]. The covariance matrix of $\mathbf{c}_k$ is given by κ ($\kappa \ll 1$) times the energy of the intended signal at each transmit antenna. In particular $\mathbf{c}_k$ can be modeled as [110]

$$\mathbf{c}_k \sim \mathcal{CN}\left(\mathbf{0}, \kappa \operatorname{diag}\left(\mathbf{V}_k \mathbf{V}_k^H\right)\right) \tag{5.5}$$

$$\mathbf{c}_k \perp \mathbf{x}_k \tag{5.6}$$

$\mathbf{e}_k \in \mathbb{C}^N,\ k = 1, 2$ is the additive receiver distortion at the kth receiver, which models the effect of limited receiver DR and closely approximates the combined effects of additive gain-control noise, non-linearities in the ADC and phase noise [110]. The covariance matrix of $\mathbf{e}_k$ is given by β ($\beta \ll 1$) times the energy of the undistorted received signal at each receive antenna. In particular, $\mathbf{e}_k$ can be modeled as [110]

$$\mathbf{e}_k \sim \mathcal{CN}\left(\mathbf{0}, \beta \operatorname{diag}\left(\mathbf{\Phi}_k\right)\right) \tag{5.7}$$

$$\mathbf{e}_k \perp \mathbf{u}_k \tag{5.8}$$

where $\mathbf{\Phi}_k = \operatorname{Cov}\{\mathbf{u}_k\}$ and $\mathbf{u}_k$ is the kth receiver's undistorted received vector, i.e. $\mathbf{u}_k = \mathbf{y}_k - \mathbf{e}_k$.

The receiver $i \in \{1, 2\}$ knows the interfering codewords $\mathbf{x}_j$ from transmitter $j \in \{1, 2\},\ j \neq i$, so the self-interference term $\sqrt{\eta_i}\tilde{\mathbf{H}}_{ij}\mathbf{x}_j$ is known and thus be cancelled [110].

The interference canceled signal can then be written as

$$\begin{aligned}\tilde{\mathbf{y}}_i &= \mathbf{y}_i - \sqrt{\eta_i}\tilde{\mathbf{H}}_{ij}\mathbf{x}_j \\ &= \sqrt{\rho_i}\tilde{\mathbf{H}}_{ii}\mathbf{x}_i + \mathbf{v}_i\end{aligned} \tag{5.9}$$

where $\mathbf{v}_i$ is the unknown interference components of (5.9) after self-interference cancellation and given by

$$\mathbf{v}_i = \sqrt{\rho_i}\Delta\mathbf{H}_{ii}\mathbf{x}_i + \sqrt{\rho_i}\mathbf{H}_{ii}\mathbf{c}_i + \sqrt{\eta_i}\Delta\mathbf{H}_{ij}\mathbf{x}_j + \sqrt{\eta_i}\mathbf{H}_{ij}\mathbf{c}_j + \mathbf{e}_i + \mathbf{n}_i. \tag{5.10}$$

Similar to the derivation in Appendix 4.A, we can show that the covariance matrix of $\mathbf{v}_i$ can be approximated as

$$\begin{aligned}\tilde{\mathbf{\Sigma}}_i \approx\ & \rho_i\kappa\tilde{\mathbf{H}}_{ii}\mathrm{diag}\left(\mathbf{V}_i\mathbf{V}_i^H\right)\tilde{\mathbf{H}}_{ii}^H + \rho_i\sigma_e^2\mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\}\mathbf{I}_N + \eta_i\kappa\tilde{\mathbf{H}}_{ij}\mathrm{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\tilde{\mathbf{H}}_{ij}^H \\ & + \eta_i\sigma_e^2\mathrm{tr}\left\{\mathbf{V}_j\mathbf{V}_j^H\right\}\mathbf{I}_N + \beta\rho_i\mathrm{diag}\left(\tilde{\mathbf{H}}_{ii}\mathbf{V}_i\mathbf{V}_i^H\tilde{\mathbf{H}}_{ii}^H\right) + \beta\eta_i\mathrm{diag}\left(\tilde{\mathbf{H}}_{ij}\mathbf{V}_j\mathbf{V}_j^H\tilde{\mathbf{H}}_{ij}^H\right) \\ & + \mathbf{I}_N,\ i,j \in \{1,2\} \text{ and } j \neq i.\end{aligned} \tag{5.11}$$

Node i applies the linear receiver $\mathbf{R}_i,\ i = 1, 2$. That is

$$\begin{aligned}\hat{\mathbf{d}}_i &= \mathbf{R}_i\tilde{\mathbf{y}}_i \\ &= \sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i\mathbf{d}_i + \mathbf{R}_i\mathbf{v}_i.\end{aligned} \tag{5.12}$$

We can now formulate the MSE of the ith transmitter-receiver pair. Using (5.12), the MSE matrix of the ith receiver can be written as

$$\begin{aligned}\mathbf{MSE}_i &= \mathtt{E}\left\{\left(\hat{\mathbf{d}}_i - \mathbf{d}_i\right)\left(\hat{\mathbf{d}}_i - \mathbf{d}_i\right)^H\right\} \\ &= \left(\sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i - \mathbf{I}_N\right)\left(\sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i - \mathbf{I}_N\right)^H + \mathbf{R}_i\tilde{\mathbf{\Sigma}}_i\mathbf{R}_i^H.\end{aligned} \tag{5.13}$$

5.3 Sum-MSE Minimization

Sum-MSE optimization scheme is formulated as follows

$$\min_{\mathbf{V}_i,\ \mathbf{R}_i} \quad \sum_{i=1}^{2} \mathrm{tr}\{\mathbf{MSE}_i\} \tag{5.14}$$

$$\text{s.t} \quad \mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \leq P_i,\ i=1,2 \tag{5.15}$$

where P_i is the power constraint at the ith transmitter.

Note that the sum-MSE function (5.14) is not jointly convex over transmit precoding matrices $\mathbf{V}_i$ and receiving filter matrices $\mathbf{R}_i$, but is component-wise convex over $\mathbf{V}_i$ and $\mathbf{R}_i$. Since it is not jointly convex, we can not apply the standard convex optimization methods to obtain the optimal solution. Therefore, we will employ an iterative algorithm that finds the efficient solutions of $\mathbf{V}_i,\ \mathbf{R}_i,\ i=1,2$ alternatingly based on the necessary conditions of the optimization problem (5.14)-(5.15).

The Lagrange function of the problem (5.14)-(5.15) can be written as:

$$\mathcal{L}(\mathbf{V}_i,\ \mathbf{R}_i, \lambda_i) = \sum_{i=1}^{2} \mathrm{tr}\{\mathbf{MSE}_i\} + \sum_{i=1}^{2} \lambda_i \left(\mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} - P_i\right)$$

where λ_i is the Lagrange multiplier associated with power constraints of transmitter i. The Karush-Kuhn-Tucker (KKT) conditions can be written as

$$\mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} - P_i \leq 0, \quad i=1,2 \tag{5.16}$$

$$\lambda_i \geq 0, \quad i=1,2 \tag{5.17}$$

$$\lambda_i \left(\mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} - P_i\right) = 0,\ i=1,2 \tag{5.18}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{V}_i^*} = \mathbf{0}, \quad \frac{\partial L}{\partial \mathbf{R}_i^*} = \mathbf{0}, \quad i=1,2. \tag{5.19}$$

Taking the partial derivative of $\mathcal{L}$ with respect to the matrix $\mathbf{V}_i$ and $\mathbf{R}_i$, we can obtain

$$\frac{\partial \mathcal{L}}{\partial \mathbf{V}_i^*} = \lambda_i \mathbf{V}_i - \sqrt{\rho_i}\tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i^H + \mathbf{X}_i \mathbf{V}_i \tag{5.20}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{R}_i^*} = -\sqrt{\rho_i}\mathbf{V}_i^H \tilde{\mathbf{H}}_{ii}^H + \rho_i \mathbf{R}_i \tilde{\mathbf{H}}_{ii}\mathbf{V}_i\mathbf{V}_i^H\tilde{\mathbf{H}}_{ii}^H + \mathbf{R}_i\tilde{\mathbf{\Sigma}}_i \tag{5.21}$$

where $\mathbf{X}_i$ in (5.20) is given by

$$\begin{aligned}\mathbf{X}_i = {} & \rho_i \tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i^H \mathbf{R}_i \tilde{\mathbf{H}}_{ii} + \rho_i \kappa \text{diag}\left(\tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i^H \mathbf{R}_i \tilde{\mathbf{H}}_{ii}\right) + \rho_i \sigma_e^2 \text{tr}\left\{\mathbf{R}_i \mathbf{R}_i^H\right\} \mathbf{I}_N \\ & + \rho_i \beta \tilde{\mathbf{H}}_{ii}^H \text{diag}\left(\mathbf{R}_i^H \mathbf{R}_i\right) \tilde{\mathbf{H}}_{ii} + \eta_j \kappa \text{diag}\left(\tilde{\mathbf{H}}_{ji}^H \mathbf{R}_j^H \mathbf{R}_j \tilde{\mathbf{H}}_{ji}\right) + \eta_j \ \sigma_e^2 \text{tr}\left\{\mathbf{R}_j \mathbf{R}_j^H\right\} \mathbf{I}_N \\ & + \eta_j \beta \tilde{\mathbf{H}}_{ji}^H \text{diag}\left(\mathbf{R}_j^H \mathbf{R}_j\right) \tilde{\mathbf{H}}_{ji}, \ i = 1, 2. \end{aligned} \tag{5.22}$$

From (5.20)–(5.21), the optimal $\mathbf{V}_i$ and $\mathbf{R}_i$, $i = 1, 2$ can be expressed as

$$\mathbf{V}_i = \sqrt{\rho_i}\left(\lambda_i \mathbf{I}_N + \mathbf{X}_i\right)^{-1} \tilde{\mathbf{H}}_{ii}^H \mathbf{R}_i^H \tag{5.23}$$

$$\mathbf{R}_i = \sqrt{\rho_i}\mathbf{V}_i^H \tilde{\mathbf{H}}_{ii}^H \left(\rho_i \tilde{\mathbf{H}}_{ii}\mathbf{V}_i\mathbf{V}_i^H\tilde{\mathbf{H}}_{ii}^H + \tilde{\mathbf{\Sigma}}_i\right)^{-1}. \tag{5.24}$$

As it is seen from (5.23)–(5.24) that the optimal transmit transmit precoding matrices and receiving filter matrices are coupled. Particularly, the optimal transmit precoding matrix $\mathbf{V}_i$ for transmitter i is dependent on the optimal receiving filter matrices of both two nodes, $\mathbf{R}_k$, $k = \{1, 2\}$ and vice versa. Therefore, we compute the transmit precoding and receive filtering matrices iteratively in an alternating fashion. Particularly, we update the transmit precoding matrices $\mathbf{V}_i$, $i = 1, 2$ from (5.23) when the receive filter matrices $\mathbf{R}_i$, $i = 1, 2$ are fixed, and then using $\mathbf{V}_i$, $i = 1, 2$ obtained at the previous step, we update the receiver filter matrices $\mathbf{R}_i$, $i = 1, 2$ from (5.24). The iterations continue until convergence or a pre-defined number of iterations is reached. The algorithm for the sum-MSE optimization problem (5.14)-(5.15) is given in Table 5.1.

Table 5.1: Sum-MSE minimization algorithm

1) Set the iteration number $n = 0$ and initialize $\mathbf{V}_i^{[0]},\ i = 1, 2$.

2) $n \leftarrow n + 1$. Update $\mathbf{R}_i^{[n+1]}$ using (5.24):

$$\mathbf{R}_i^{[n+1]} = \sqrt{\rho_i}\left(\mathbf{V}_i^{[n]}\right)^H \tilde{\mathbf{H}}_{ii}^H \left(\rho_i \tilde{\mathbf{H}}_{ii}\mathbf{V}_i^{[n]}\left(\mathbf{V}_i^{[n]}\right)^H \tilde{\mathbf{H}}_{ii}^H + \tilde{\mathbf{\Sigma}}_i^{[n]}\right)^{-1}, i = 1, 2.$$

3) Calculate and update $\mathbf{V}_i^{[n+1]}$ and $\lambda_i^{[n+1]}$ using (5.23) and (5.25), respectively:

$$\mathbf{V}_i^{[\tilde{\lambda}_i]} = \sqrt{\rho_i}\left(\tilde{\lambda}_i \mathbf{I}_N + \mathbf{X}_i^{[n+1]}\right)^{-1} \tilde{\mathbf{H}}_{ii}^H \left(\mathbf{R}_i^{[n+1]}\right)^H, \quad i = 1, 2.$$

$$\lambda_i^{[n+1]} = \left[\left\{\tilde{\lambda}_i | \text{ such that tr}\left\{\mathbf{V}_i^{[\tilde{\lambda}_i]}\left(\mathbf{V}_i^{[\tilde{\lambda}_i]}\right)^H\right\} = P_i\right\}\right]^+, \ i = 1, 2.$$

$$\mathbf{V}_i^{[n+1]} = \sqrt{\rho_i}\left(\lambda_i^{[n+1]} \mathbf{I}_N + \mathbf{X}_i^{[n+1]}\right)^{-1} \tilde{\mathbf{H}}_{ii}^H \left(\mathbf{R}_i^{[n+1]}\right)^H, \quad i = 1, 2.$$

4) Repeat steps 2, 3 until convergence or a predefined number of iterations is reached.

The values of the Lagrange multiplier $\lambda_i,\ i = 1, 2$ in step 3 of Table 5.1 are calculated by taking the singular value decomposition of $\mathbf{X}_i^{[n+1]} = \mathbf{U}_i^{[n+1]}\mathbf{\Delta}_i^{[n+1]}\left(\mathbf{U}_i^{[n+1]}\right)^H$ and writing the update as $\mathbf{V}_i^{[\tilde{\lambda}_i]} = \sqrt{\rho_i}\left(\tilde{\lambda}_i \mathbf{I}_N + \mathbf{X}_i^{[n+1]}\right)^{-1} \tilde{\mathbf{H}}_{ii}^H \left(\mathbf{R}_i^{[n+1]}\right)^H$ at each iteration. By plugging $\mathbf{V}_i^{[\tilde{\lambda}_i]}$ into the power constraint in (4.8) and after simple steps, (4.8) can be written as

$$\begin{aligned} \text{tr}\left\{\mathbf{V}_i^{[\tilde{\lambda}_i]}\left(\mathbf{V}_i^{[\tilde{\lambda}_i]}\right)^H\right\} &= \rho_i \sum_{k=1}^{N} \frac{g_{ik}^{[n+1]}}{\left(\tilde{\lambda}_i + \Delta_{ik}^{[n+1]}\right)^2} \\ &= P_i. \end{aligned} \tag{5.25}$$

where $g_{ik}^{[n+1]}$ denotes the kth element of $\left(\mathbf{U}_i^{[n+1]}\right)^H \tilde{\mathbf{H}}_{ii}^H \left(\mathbf{R}_i^{[n+1]}\right)^H \mathbf{R}_i^{[n+1]} \tilde{\mathbf{H}}_{ii} \mathbf{U}_i^{[n+1]}$ and $\Delta_{ik}^{[n+1]}$ denotes the kth element of the matrix $\mathbf{\Delta}_i^{[n+1]}$. We can compute $\tilde{\lambda}_i,\ i = 1, 2$ from (5.25) numerically. The values of the Lagrange multipliers $\lambda_i,\ i = 1, 2$ are equal to $\tilde{\lambda}_i,\ i = 1, 2$ if $\tilde{\lambda}_i,\ i = 1, 2$ is non-negative. Otherwise, we assign the Lagrange multipliers $\lambda_i,\ i = 1, 2$ as zeros.

Since the proposed sum-MSE algorithm monotonically decreases the sum-MSE over each iteration by updating the the transceivers in an alternating fashion, and the fact that MSE is bounded below (at least by zero), it is clear that the proposed algorithm sum-MSE algorithm is convergent and guaranteed to converge to a local minimum. Since sum-MSE optimization problem is not jointly convex, the proposed algorithm is not guaranteed to converge to a global optimum point. Therefore, good initialization points should be selected to ensure a suboptimal solution with a good performance.

5.4 Min-Max MSE Minimization

Unlike the minimum sum-MSE transceiver design discussed in Section 5.3, the Min-Max per-node MSE transceiver design ensures each receiver has the same MSE so that it introduces fairness among the two FD nodes. The Min-Max MSE optimization problem can be formulated as:

$$\min_{\mathbf{V}_i,\ \mathbf{R}_i} \max_{i=1,2} \quad \mathrm{tr}\{\mathbf{MSE}_i\} \tag{5.26}$$

$$\text{s.t} \quad \mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \leq P_i,\ i=1,2. \tag{5.27}$$

Similar to the sum-MSE optimization problem (5.14)-(5.15), the Min-Max MSE optimization problem is not jointly convex over transmit precoding matrices $\mathbf{V}_i$ and receive filtering matrices $\mathbf{R}_i$. Therefore we carry out the optimization procedure iteratively in an alternating fashion. In particular we find $\mathbf{V}_i$ or $\mathbf{R}_i$ while keeping the other one fixed. Note that since the Min-Max MSE optimization problem (5.26)–(5.27) is non-convex, it is not guaranteed to find the global optimum solution. For fixed receiver filtering matrices

$\mathbf{R}_i,\ i = 1, 2$, the Min-Max optimization problem can be written as:

$$\min_{\mathbf{V}_i} \max_{i=1,2} \quad \text{tr}\{\mathbf{MSE}_i\} \tag{5.28}$$

$$\text{s.t} \quad \text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \leq P_i,\ i = 1, 2. \tag{5.29}$$

With introduction of an auxiliary variable l which is an upper bound on the square root of $\text{tr}\{\mathbf{MSE}_i\}$ (i.e., $\sqrt{\text{tr}\{\mathbf{MSE}_i\}} \leq l\ \forall i \in \{1, 2\}$), the optimization problem (5.28)–(5.29) can be written as

$$\min_{\mathbf{V}_i,\, l} \quad l \tag{5.30}$$

$$\text{s.t} \quad \sqrt{\text{tr}\{\mathbf{MSE}_i\}} \leq l \quad i = 1, 2 \tag{5.31}$$

$$\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \leq P_i,\ i = 1, 2. \tag{5.32}$$

To solve the optimization problem in (5.30)-(5.32), we need to write $\text{tr}\{\mathbf{MSE}_i\}$ in vector form. As shown in Appendix 5.A, $\text{tr}\{\mathbf{MSE}_i\}$ can be written as

$$\begin{aligned}
\text{tr}\{\mathbf{MSE}_i\} &= \left\| \begin{array}{c}
\left[\mathbf{I}_N \otimes \left(\sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\right)\right] vec\left(\mathbf{V}_i\right) - vec\left(\mathbf{I}_N\right) \\
\sqrt{\rho_i\kappa}\left[\mathbf{I}_N \otimes \left(\left(\text{diag}\left(\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i^H\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\right)\right)^{1/2}\right)\right] vec\left(\mathbf{V}_i\right) \\
\sqrt{\rho_i\sigma_e^2}\sqrt{\text{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}} vec\left(\mathbf{V}_i\right) \\
\sqrt{\beta\rho_i}\left[\mathbf{I}_N \otimes \left(\left(\text{diag}\left(\mathbf{R}_i^H\mathbf{R}_i\right)\right)^{1/2}\tilde{\mathbf{H}}_{ii}\right)\right] vec\left(\mathbf{V}_i\right) \\
\sqrt{\eta_i\kappa}\left[\mathbf{I}_N \otimes \left(\left(\text{diag}\left(\tilde{\mathbf{H}}_{ij}^H\mathbf{R}_i^H\mathbf{R}_i\tilde{\mathbf{H}}_{ij}\right)\right)^{1/2}\right)\right] vec\left(\mathbf{V}_j\right) \\
\sqrt{\eta_i\sigma_e^2}\sqrt{\text{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}} vec\left(\mathbf{V}_j\right) \\
\sqrt{\beta\eta_i}\left[\mathbf{I}_N \otimes \left(\left(\text{diag}\left(\mathbf{R}_i^H\mathbf{R}_i\right)\right)^{1/2}\tilde{\mathbf{H}}_{ij}\right)\right] vec\left(\mathbf{V}_j\right) \\
\sqrt{\text{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}}
\end{array} \right\|_2^2 \\
&\triangleq \|\boldsymbol{\mu}_i\|_2^2
\end{aligned} \tag{5.33}$$

Then the Min-Max optimization problem of transmit precoding matrices (5.30)-(5.32) can be written as

$$\min_{\mathbf{V}_i,\, l} \quad l \tag{5.34}$$

$$\text{s.t} \quad \|\boldsymbol{\mu}_i\|_2 \leq l, \quad i = 1, 2 \tag{5.35}$$

$$\|vec\,(\mathbf{V}_i)\|_2 \leq \sqrt{P_i}, \quad i = 1, 2. \tag{5.36}$$

Since the objective function (5.34) is linear and the constraints (5.35)–(5.36) are second-order cones, the problem (5.34)–(5.36) is a SOCP problem [209] and can be efficiently solved by standard SOCP solver, e.g., SeDuMi [210].

As for the update of the optimal receiving filter matrices $\mathbf{R}_i$ under the fixed values of the transmit precoding matrices $\mathbf{V}_i$, it is shown in (5.24) that the optimal $\mathbf{R}_i,\ i = 1, 2$ is linear MMSE receiver. The iterative Min-Max MSE algorithm, as shown in Table 5.2, employs linear MMSE receiver to obtain the optimal receive filtering matrices $\mathbf{R}_i$, while keeping $\mathbf{V}_i$ fixed and solves (5.34)–(5.36) to obtain the optimal transmit precoding matrices $\mathbf{V}_i$, while keeping $\mathbf{R}_i$ fixed. Similar to the discussion on the convergence of the sum-MSE algorithm at the end of the Section 5.3, we can also argue that Min-Max MSE algorithm is guaranteed to converge to a local minimum.

5.5 Simulation Results

In this section, we numerically investigate the sum-MSE (Total) and Min-Max MSE (MinMax) optimization problems for MIMO FD bi-directional system and MIMO

Table 5.2: Min-Max MSE minimization algorithm

1) Set the iteration number $n = 0$ and initialize $\mathbf{V}_i^{[0]},\ i = 1, 2$.

2) $n \leftarrow n + 1$. Update $\mathbf{R}_i^{[n+1]}$ using (5.24):

$$\mathbf{R}_i^{[n+1]} = \sqrt{\rho_i}\left(\mathbf{V}_i^{[n]}\right)^H \tilde{\mathbf{H}}_{ii}^H \left(\rho_i \tilde{\mathbf{H}}_{ii}\mathbf{V}_i^{[n]}\left(\mathbf{V}_i^{[n]}\right)^H \tilde{\mathbf{H}}_{ii}^H + \tilde{\mathbf{\Sigma}}_i^{[n]}\right)^{-1}, i = 1, 2.$$

3) Solve (5.33)–(5.36) to get the optimal transmit precoding matrices, $\mathbf{V}_i^{[n+1]}$.

4) Repeat steps 2, 3 until convergence or a predefined number of iterations is reached.

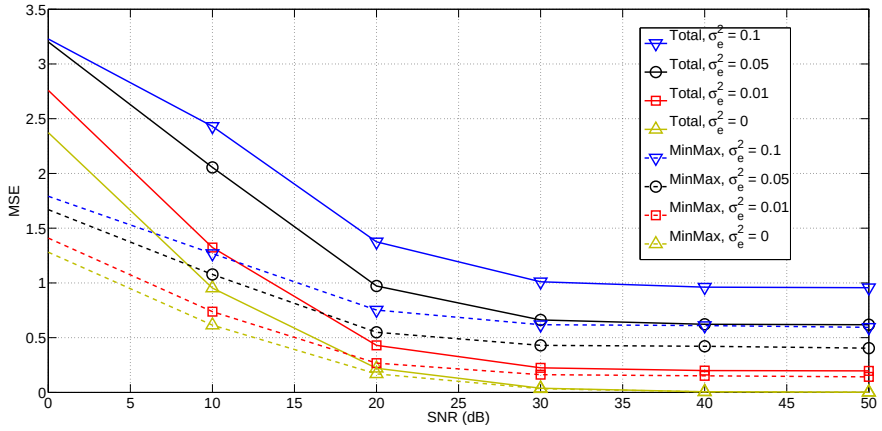

Figure 5.2: MSE comparison of the Total and MinMax algorithms with different channel estimation errors versus SNR. Here INR = 20dB, $N = 2,\ \kappa = \beta = -40$dB.

FD CRN as a function of signal-to-noise ratio (SNR), interference-to-noise ratio (INR), channel estimation errors σ_e^2 and dynamic range parameters κ and β. For brevity, we focused only on the case $\rho_1 = \rho_2 = \rho$ and $\eta_1 = \eta_2 = \eta$. We set the same transmit power cosntraint for each transmitter (i.e., $P_i = N,\ i = 1, 2$). Thus, the SNR for all desired links is defined as $\text{SNR}_i = \text{SNR} = \rho N,\ i = 1, 2$ and the INR for all interfering links $\text{INR}_i = \text{INR} = \eta N,\ i = 1, 2$. The tolerance (the difference between MSE of two iterations) of the proposed iterative algorithm is set to 10^{-4} and the maximum number of iterations is set to 1000 and the results are averaged over 1000 independent channel realizations.

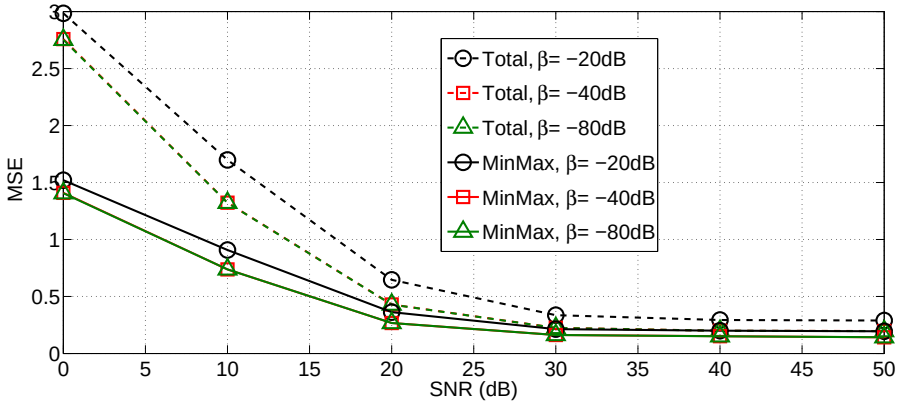

Figure 5.3: MSE comparison of the Total and MinMax algorithms with different $\kappa = \beta$ values versus SNR. Here $N = 2$, INR = 20dB, $\sigma_e^2 = 0.01$.

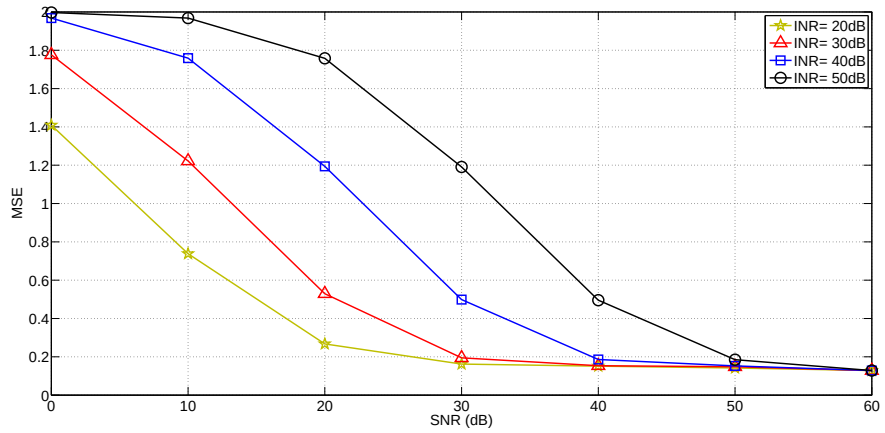

Figure 5.4: MSE of the Total algorithm with different INR values versus SNR. Here $N = 2$, $\sigma_e^2 = 0.01$ and $\kappa = \beta = -40$dB.

The CRN is installed within the service range of a primary network having $K = 2$ PUs. For simplicity, we set the same maximum allowed interfering power to the PUs (i.e., $\lambda = \lambda_k = -20\text{dB},\ k = 1, \ldots, K$) and $\mu = \mu_{ki},\ k = 1, \ldots, K$ and $i = 1, 2$.

Since the optimization problems we are dealing with are non-convex, we need to choose good initialization points to have a suboptimal solution with a good performance.

In [208], several reasonable choices such as right singular matrices and random matrices initialization have been proposed. In this chapter, we use right singular matrices initialization.

In our first example, we investigate the role of channel estimation errors on the MSE performance of both Total and MinMax algorithms. Here we set $N = 2$, INR = 20dB and $\kappa = \beta = -40$dB. It can be seen from Fig. 5.2 that as the channel estimation errors increases, the MSE of both the Total and MinMax algorithms increases. Also note that channel estimation error produces an irreducable error floor, i.e., MSE can not be further reduced by increasing SNR. The reason for the error floor is that at low SNR, channel estimation error is weaker than the noise, but at high SNR, it starts to dominate, and the performance is governed by channel estimation error [192].

In our second example, we examine the MSE performance of the Total and MinMax algorithms under transmitter and receiver impairments. Here we set $N = 2$, INR = 20dB, $\sigma_e^2 = 0.01$. It can be seen from Fig. 5.3 that as κ and β decrease, the MSE value also decreases and exhibits an error floor. Also note that at low κ and β values, the difference between the curves is indistinguishable.

Since both Total and MinMax algorithms have similar trends, in our third example, we investigate the role of INR on the MSE performance of Total algorithm only. Here we set $N = 2$, $\sigma_e^2 = 0.01$ and $\kappa = \beta = -40$dB. As shown in Fig. 5.4, as self-interference power INR decreases, the MSE also decreases and exhibit an error floor. Also note that the curves for different INRs converge at high SNR values.

The next example illustrates the MSE distribution among the two users and sum-

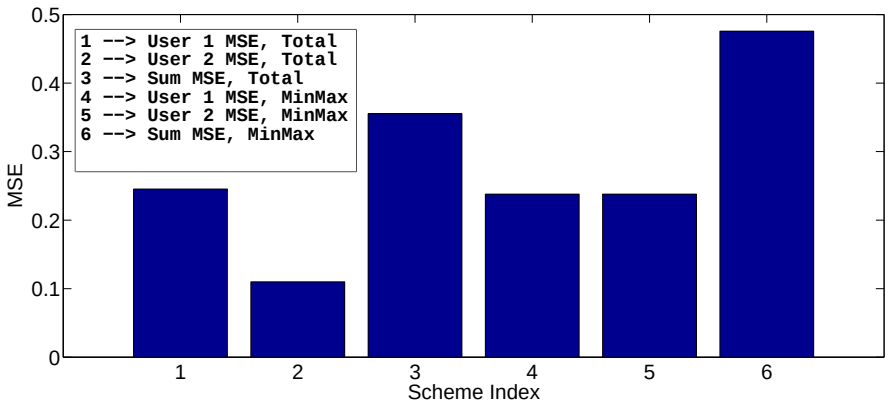

Figure 5.5: MSE distribution of Total and MinMax algorithms. The schemes $1-3$ correspond to Total and the schemes $4-6$ correspond to MinMax algorithms. For each algorithm, the first two bars are the achieved user MSEs and the third bar is the sum-MSE. Here $N=3$, $\sigma_e^2=0.01$, $\kappa=\beta=-40$dB, SNR $=20$dB and INR $=10$dB.

MSE for the Total and MinMax schemes out of one channel realization. Here we set $N=3$, $\sigma_e^2=0.01$, $\kappa=\beta=-40$dB, SNR $=20$dB and INR $=10$dB. We can see in Fig. 5.5 that the Total scheme achieves the minimum sum MSE and the MinMax scheme almost achieves the same MSE for the two users.

5.6 Appendix 5.A: Vector Forms

Using (5.13) and (5.11), we have

$$\begin{aligned}
\mathrm{MSE}_i &= \mathrm{tr}\{\mathbf{MSE}_i\} \qquad (5.37)\\
&= \mathrm{tr}\left\{\left(\sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i-\mathbf{I}_N\right)\left(\sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i-\mathbf{I}_N\right)^H\right\}+\mathrm{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}\\
&\quad+\rho_i\kappa\,\mathrm{tr}\left\{\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathrm{diag}\left(\mathbf{V}_i\mathbf{V}_i^H\right)\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i^H\right\}+\rho_i\sigma_e^2\mathrm{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}\mathrm{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\}\\
&\quad+\beta\rho_i\mathrm{tr}\left\{\mathbf{R}_i\mathrm{diag}\left(\tilde{\mathbf{H}}_{ii}\mathbf{V}_i\mathbf{V}_i^H\tilde{\mathbf{H}}_{ii}^H\right)\mathbf{R}_i^H\right\}+\eta_i\kappa\,\mathrm{tr}\left\{\mathbf{R}_i\tilde{\mathbf{H}}_{ij}\mathrm{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\tilde{\mathbf{H}}_{ij}^H\mathbf{R}_i^H\right\}\\
&\quad+\eta_i\sigma_e^2\mathrm{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}\mathrm{tr}\left\{\mathbf{V}_j\mathbf{V}_j^H\right\}+\beta\eta_i\mathrm{tr}\left\{\mathbf{R}_i\mathrm{diag}\left(\tilde{\mathbf{H}}_{ij}\mathbf{V}_j\mathbf{V}_j^H\tilde{\mathbf{H}}_{ij}^H\right)\mathbf{R}_i^H\right\}
\end{aligned}$$

Moreover, using the $vec(\cdot)$ operation and the identity $\|vec(\mathbf{A})\|_2^2 = \mathrm{tr}\left\{\mathbf{A}\mathbf{A}^H\right\}$, the MSE_i in (5.37) can be rewritten as

$$\begin{aligned}
\mathrm{MSE}_i \;=\; & \left\|vec\left(\sqrt{\rho_i}\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\mathbf{V}_i\right) - vec(\mathbf{I}_N)\right\|_2^2 + \rho_i\kappa\left\|vec\left(\left(\mathrm{diag}\left(\tilde{\mathbf{H}}_{ii}^H\mathbf{R}_i^H\mathbf{R}_i\tilde{\mathbf{H}}_{ii}\right)\right)^{1/2}\mathbf{V}_i\right)\right\|_2^2 \\
& + \rho_i\sigma_e^2\mathrm{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}\|vec(\mathbf{V}_i)\|_2^2 + \rho_i\beta\left\|vec\left(\left(\mathrm{diag}\left(\mathbf{R}_i^H\mathbf{R}_i\right)\right)^{1/2}\tilde{\mathbf{H}}_{ii}\mathbf{V}_i\right)\right\|_2^2 \\
& + \eta_i\kappa\left\|vec\left(\left(\mathrm{diag}\left(\tilde{\mathbf{H}}_{ij}^H\mathbf{R}_i^H\mathbf{R}_i\tilde{\mathbf{H}}_{ij}\right)\right)^{1/2}\mathbf{V}_j\right)\right\|_2^2 + \eta_i\sigma_e^2\mathrm{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\}\|vec(\mathbf{V}_j)\|_2^2 \\
& + \eta_i\beta\left\|vec\left(\left(\mathrm{diag}\left(\mathbf{R}_i^H\mathbf{R}_i\right)\right)^{1/2}\tilde{\mathbf{H}}_{ij}\mathbf{V}_j\right)\right\|_2^2 + \mathrm{tr}\left\{\mathbf{R}_i\mathbf{R}_i^H\right\} \qquad (5.38)
\end{aligned}$$

Using the identity $vec(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)vec(\mathbf{B})$, (5.38) can be written as (5.33).

Part II

On Duality of MIMO Relays

Chapter 6

Literature Survey on Uplink-Downlink Duality

Duality properties between uplink and downlink communication systems have gained much interest in signal processing and information theory. Due to the coupled structure of the transmitted signals in the downlink channel, the optimization problems associated with the downlink system are usually difficult to solve. The key technique used to overcome this difficulty is to transform the downlink problem into an uplink problem via a so-called uplink-downlink duality relationship. Since the uplink channel has a simpler mathematical structure, less coupling of variables, it is usually more efficient to solve the optimization problems associated with the dual uplink system.

To the best of our knowledge, MAC-BC duality results available in the literature can be divided into the following three main categories (also summarized in Table 6.1).

6.1 SINR Duality

The MAC-BC SINR duality for single-hop multiple-input single-output (MISO) systems was derived independently in [211] and [212]. Based on this SINR duality, sum-power minimization problem subject to minimum SINR requirements and the SINR balancing problem were solved in [212]. The SINR duality result for MISO systems was extended to MIMO systems with multi-antenna receivers/transmitters in [213] and [214]. Recently, the MAC-BC SINR duality for single-hop MIMO systems has been extended to two-hop AF MIMO relay systems, where all nodes in the system are equipped with multiple antennas [215]. It is shown in [215] that in a MAC system, SINRs identical to that of the BC system can be achieved by employing a scaled Hermitian transpose of the relay amplifying matrix used in the BC system, and the scaling factor is obtained by swapping the transmission power constraints at the source node and the relay node. This result generalizes the SINR duality established for single-hop MIMO systems in [211]-[214]. Recently, the authors of [216] extended the two-hop MAC-BC SINR duality results in [215] to multi-hop AF MIMO relay systems with any number of hops and any number of antennas at each node.

The aforementioned SINR duality results are established by assuming that the exact CSI is available in the system. However, in practical communication systems, the CSI knowledge is obtained through channel training/estimation. Due to limited length of training sequences, channel noise, quantization errors, outdated channel estimates, and/or time-varying nature of wireless channels, there is mismatch between the estimation and the

Table 6.1: Existing MAC-BC duality results

		MISO			**MIMO**		
		SH	TH	MH	SH	TH	MH
PCSI	SINR	[211, 212]	[215]	[216]	[213, 214]	[215]	[216]
	MSE	[218]	[225]	TT	[214, 219, 220]	[225]	TT
	Capacity	[211, 212]	[215, 227]	[216, 227]	[213, 214, 226]	[215]	[216]
ICSI	SINR	[217]	TT	TT	TT	TT	TT
	MSE	[217, 221]	[225]	TT	[222, 223, 224]	[225]	TT
	Capacity	[217]	TT	TT	TT	TT	TT

PCSI: Perfect channel state information (CSI), ICSI: Imperfect CSI

MIMO: Multi-input multi-output, MISO: Multi-input single-output

SH: Single-hop, TH: Two-hop, MH: Multi-hop

TT: This thesis

exact CSI, which may substantially degrade the system performance. Therefore, channel estimation errors should be taken into account for practical applications. This motivates the authors of [217] to establish the SINR duality under imperfect CSI for single-hop MISO systems, which generalizes the SINR duality with perfect CSI in [211] and [212].

6.2 MSE Duality

The MAC-BC MSE duality was first derived for MISO systems with a sum power constraint in [218] and then extended to MIMO systems in [214], [219], and [220]. It was observed in [219] that under a total power constraint, any MSE point that is achievable in the MAC system can also be attained in the BC system.

The MSE duality results obtained in [214] and [218]-[220] are based on the as-

sumption that the exact CSI is available in the system. For systems with imperfect CSI, the MSE duality has been established in [217] and [221] for single-hop MISO systems, and in [222] for single-hop MIMO systems. However, channel correlation among antenna elements is not considered in [217, 221], and [222]. Based on the Karush-Kuhn-Tucker (KKT) conditions associated with the optimization problems in the MAC and BC systems, the sum-MSE MAC-BC duality has been established in [223] for single-hop MIMO systems with imperfect CSI and antenna correlation at the base station. The MSE duality result in [223] is extended in [224], where stream-wise MSE duality is established by considering the imperfect CSI and antenna correlation at both the base station and the users. In fact, the duality results obtained in [224] can be viewed as the extension of those in [220] to the imperfect CSI case. Recently, the sum-MSE MAC-BC duality in single-hop MIMO systems under imperfect CSI has been extended to two-hop AF MIMO relay systems in [225].

6.3 Capacity Duality

The MAC-BC rate-region duality for single antenna terminals or single stream transmission with multi-antenna terminals can be readily derived from the SINR duality. In particular, the sum capacity duality was proven for MISO systems in [211] by showing that the achievable sum-rate with Costa precoding in the BC system is the same as the maximum sum-rate in the MAC system. The latter duality result was extended to MIMO systems supporting an arbitrary number of data streams per user in [226]. The authors of [227] derived the capacity duality for multi-hop AF MIMO relay systems with single antenna source and destination nodes, which generalizes the MAC-BC rate-region duality

results of [211] and [226].

Chapter 7

On MAC-BC Duality of Multi-hop MIMO Relay Channel with Imperfect Channel Knowledge

In this chapter, we establish the signal-to-interference-noise ratio (SINR) duality between multiple access (MAC) and broadcast (BC) multi-hop AF MIMO relay systems under an imperfect channel state model, which is a generalization of several previously established MAC-BC duality results. We show that identical SINRs in the MAC and BC systems can be achieved by two approaches. The first one is to use the Hermitian transposed MAC relay amplifying matrices at the relay nodes in the BC system, under the same total network transmission power constraint. The second one is to use the scaled and Hermitian transposed MAC relay amplifying matrices in the BC system, under the transmission power constraint at each node of the system, where the scaling factors are obtained by swapping

the power constraints of the nodes in the MAC system. Moreover, we derive the MAC-BC MSE and capacity duality properties based on the SINR duality. Numerical results show the utility of the duality results established.

7.1 Introduction

MIMO relays are important for wireless communication networks because they can be used to reduce the path loss, increase the power efficiency, and improve the network coverage. MIMO relays can be used for multiple access (MAC) – from multiple users to a base station. They can also be used for broadcast (BC) – from a base station to multiple users. In this chapter, we consider a chain of multi-hop AF MIMO relays for either MAC or BC. To achieve a desired performance for such a system, the transformation matrices applied at the source, the destination, and the relays need to be chosen properly. Our contribution in this chapter is about MAC-BC duality properties of the multi-hop MIMO relay system in terms of these transformation matrices under an imperfect channel state information (CSI) model.

7.1.1 Contributions of Our Work

To the best of our knowledge, so far no work has been done to prove the SINR (MSE, capacity) duality for multi-hop AF MIMO relay systems that consider imperfect CSI and antenna correlation at all nodes (see Table 6.1). In this chapter, we consider that the antennas of all the nodes in the system exhibit spatial correlations and the CSI at each hop is imperfect. We show that under imperfect CSI, stream-wise identical SINR (MSE,

capacity) can be achieved in multi-hop AF MIMO MAC and BC relay systems through two approaches. First, if there is only total network transmission power constraint and no power constraint at individual nodes, then duality can be achieved by employing $\mathbf{F}_l$ and $\mathbf{F}_{L-l}^H,\ l = 1, \ldots, L-1$, as the relay amplifying matrices at the lth relay node of the BC and the MAC MIMO relay systems, respectively, where 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 established when $c_l\mathbf{F}_l$ and $\mathbf{F}_{L-l}^H,\ l = 1, \ldots, L-1$, are employed as the amplifying matrices at the lth relay node of the BC and the MAC relay systems, respectively. In this case, the scaling factor $c_l > 0$ is obtained by swapping the power constraints at the lth node of the BC system and the $(L+1-l)$-th node of the MAC system, $l = 1, \ldots, L$.[1]

Furthermore, we prove that the two approaches developed above are not only valid for MIMO relay systems with linear transceivers at the source and the destination nodes, but also hold if a receiver employing successive interference cancellation (SIC) is used at the destination of the MAC MIMO relay system, and a transmitter employing dirty paper coding (DPC) is used at the source node of the BC MIMO relay channel. As an application of this MAC-BC duality, the complicated robust multi-hop MIMO BC system design problem under imperfect CSI can be efficiently solved by focusing on an equivalent multi-hop MIMO MAC problem.

In this chapter, we define the SINR as an estimated SINR when only an estimate of the channel response is available. An estimated SINR is a ratio of signal power coupled

[1]There is no loss of generality in assuming c_l to be a non-negative real number [215].

with the estimated channel over interference-plus-noise power plus the channel estimation error variance coupled with the signal power. Such a definition of SINR is consistent with those defined in prior works [228]-[231]. This estimated SINR is more meaningful than an exact SINR when the channel knowledge at a receiver is not exact. In fact, when there is only an estimated channel response, the performance of the receiver is directly governed by the estimated SINR instead of the exact SINR, in the sense that the channel estimation error affects the effective SINR as an additional noise.

The following notations are used in this chapter. Matrices and vectors are denoted as bold capital and lowercase letters, respectively. For matrices, $(.)^T$ and $(.)^H$ denote transpose and conjugate transpose, respectively. $\mathbb{E}[.]$ stands for the statistical expectation; $\mathbf{I}_N$ denotes an $N \times N$ identity matrix; $tr\{.\}$ stands for matrix trace; $\text{diag}\{a_1, \cdots, a_n\}$ denotes a diagonal matrix with the diagonal elements given by $a_1, \cdots, a_n$, and $\|.\|_2$ stands for the vector Euclidean norm. For matrices $\mathbf{A}_i$, $\bigotimes_{i=l}^{k}(\mathbf{A}_i) \triangleq \mathbf{A}_l \ldots \mathbf{A}_k$. For example, $\bigotimes_{i=1}^{3}(\mathbf{A}_i) \triangleq \mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$ and $\bigotimes_{i=3}^{1}(\mathbf{A}_i) \triangleq \mathbf{A}_3\mathbf{A}_2\mathbf{A}_1$. bd(.) stands for a block diagonal matrix, and $\Re\{.\}$ denotes the real part.

7.2 System Model

We consider a wireless communication system with one base station (BS), $L-1$ ($L \geq 2$) relay nodes and K user nodes, where the BS is equipped with N_1 antennas and the $(l-1)$-th relay node is equipped with N_l antennas, $l = 2, \ldots, L$. The ith user, $i = 1, \ldots, K$, transmits (receives) $N_b^{(i)}$ data streams using $N_{L+1}^{(i)}$ antennas in the MAC (BC) system. We denote $N_b = \sum_{i=1}^{K} N_b^{(i)}$ as the total number of independent data streams

from all users and $N_{L+1} = \sum_{i=1}^{K} N_{L+1}^{(i)}$ as the total number of antennas of all users. In order to support N_b data streams simultaneously, $N_b \leq \min(N_1, N_2, \ldots, N_{L+1})$ should be satisfied. 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 $\min(N_1, N_2, \ldots, N_{L+1})$ [216]. We assume that the signal transmitted by the lth node can only be received by the $(l+1)$-th node due to the propagation path-loss. Thus, there are L hops between the source and destination nodes. Each relay node works in half-duplex mode and employs a linear AF (non-regenerative) relay matrix to amplify and forward its received signals.

7.2.1 Multi-hop BC MIMO Relay System

In the multi-hop BC MIMO relay system shown in Fig. 7.1, the source symbol vector $\mathbf{s}_i^B = \left[s_{i,1}^B, s_{i,2}^B, \ldots, s_{i,N_b^{(i)}}^B\right]^T$ of size $N_b^{(i)} \times 1$ from the ith user is linearly precoded by matrix $\mathbf{U}_i\mathbf{Q}_i^{\frac{1}{2}} \in \mathcal{C}^{N_1 \times N_b^{(i)}}$, where $\mathbf{U}_i = \left[\mathbf{u}_i^{(1)}, \mathbf{u}_i^{(2)}, \ldots, \mathbf{u}_i^{(N_b^{(i)})}\right]$ with $\left\|\mathbf{u}_i^{(j)}\right\|_2 = 1$ and $\mathbf{Q}_i$ is defined as $\mathbf{Q}_i = \text{diag}\left\{q_i^{(1)}, q_i^{(2)}, \ldots, q_i^{(N_b^{(i)})}\right\}$ with $q_i^{(j)}$, $j = 1, \ldots, N_b^{(i)}$, $i = 1, \ldots, K$, being the power allocated to the jth data stream of the ith user. We assume that user symbols are independent and have unit-power, i.e., $\mathbb{E}\left[\mathbf{s}_i^B \left(\mathbf{s}_i^B\right)^H\right] = \mathbf{I}_{N_b^{(i)}}$. The BS transmits the $N_1 \times 1$ linearly precoded symbol vector $\mathbf{x}_1^B = \sum_{i=1}^{K} \mathbf{U}_i\mathbf{Q}_i^{\frac{1}{2}}\mathbf{s}_i^B = \mathbf{U}\mathbf{Q}^{\frac{1}{2}}\mathbf{s}^B$, where $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_K]$, $\mathbf{Q} = \text{bd}\left(\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_K\right)$, and $\mathbf{s}^B = \left[\left(\mathbf{s}_1^B\right)^T, \left(\mathbf{s}_2^B\right)^T, \ldots, \left(\mathbf{s}_K^B\right)^T\right]^T$. The $N_l \times 1$ received signal vector at the $(l-1)$-th relay node of the BC system can be written as

$$\mathbf{y}_l^B = \mathbf{H}_{l-1}\mathbf{x}_{l-1}^B + \mathbf{n}_l, \qquad l = 2, \ldots, L. \tag{7.1}$$

where $\mathbf{H}_l \in \mathcal{C}^{N_{l+1}\times N_l},\ l = 1, \ldots, L-1$, is the MIMO channel matrix between the $(l+1)$-th and the lth node, $\mathbf{x}_{l-1}^B \in \mathcal{C}^{N_{l-1}\times 1}$ is the signal vector transmitted by the $(l-1)$-th node, $l = 2, \ldots, L+1$, $\mathbf{n}_l \in \mathcal{C}^{N_l \times 1}$ is the independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vector at the $(l-1)$-th relay node, $l = 2, \ldots, L$. We assume that all noises are complex circularly symmetric with zero mean and unit variance. The transmitted signal vector at the $(l-1)$-th relay node is written as

$$\mathbf{x}_l^B = c_{l-1}\mathbf{F}_{l-1}\mathbf{y}_l^B, \qquad l = 2, \ldots, L. \tag{7.2}$$

where $c_{l-1}\mathbf{F}_{l-1} \in \mathcal{C}^{N_l \times N_l}$ is the amplifying matrix at the $(l-1)$-th relay node, and $c_l > 0$ is a scaling coefficient which is important for studying the MAC-BC duality [227]. Using (7.1) and (7.2), the received signal vector at the first and lth relay node is written, respectively as

$$\mathbf{y}_2^B = \mathbf{H}_1\mathbf{U}\mathbf{Q}^{\frac{1}{2}}\mathbf{s}^B + \mathbf{n}_2 \tag{7.3}$$

$$\mathbf{y}_{l+1}^B = \mathbf{H}_l \bigotimes_{m=l-1}^{1} (c_m\mathbf{F}_m\mathbf{H}_m)\,\mathbf{U}\mathbf{Q}^{\frac{1}{2}}\mathbf{s}^B + \mathbf{n}_{l+1} + \sum_{k=2}^{l}\bigotimes_{m=l}^{k} (c_{m-1}\mathbf{H}_m\mathbf{F}_{m-1})\,\mathbf{n}_k, \quad l = 2, \ldots, L-1. \tag{7.4}$$

The received signal vector at the ith user node is given by

$$\mathbf{y}_{L+1}^{(i)B} = \mathbf{H}_{L_i}c_{L-1}\mathbf{F}_{L-1}\mathbf{y}_L^B + \mathbf{n}_{L+1}^{(i)}, \qquad i = 1, \cdots, K. \tag{7.5}$$

where $\mathbf{H}_{L_i} \in \mathcal{C}^{N_{L+1}^{(i)}\times N_L}$ is the MIMO channel matrix between the ith user node and the $(L-1)$-th relay node and $\mathbf{n}_{L+1}^{(i)} \in \mathcal{C}^{N_{L+1}^{(i)}\times 1}$ is the i.i.d. AWGN vector at the ith user node.

A linear receiver matrix $\mathbf{V}_i \in \mathcal{C}^{N_{L+1}^{(i)}\times N_b^{(i)}}$ is applied at the ith user node to estimate the symbol vector $\mathbf{s}_i^B$, where the columns of $\mathbf{V}_i$ are assumed to satisfy $\left\|\mathbf{v}_i^{(j)}\right\|_2 = 1$, $j =$

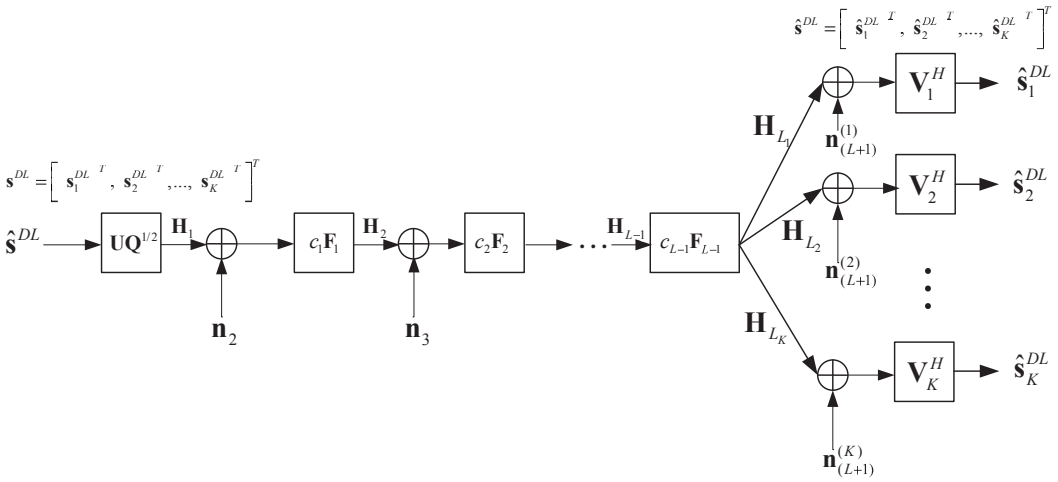

Figure 7.1: Multi-hop BC AF MIMO relay system

$1, \ldots, N_b^{(i)}$, $i = 1, \ldots, K$. The estimated symbol vector $\hat{\mathbf{s}}_i^B$ is expressed as

$$\hat{\mathbf{s}}_i^B = \mathbf{V}_i^H \mathbf{y}_{L+1}^{(i)B}. \tag{7.6}$$

Using (7.3)-(7.6), we have the decision variable of the jth data symbol of the ith user as

$$\begin{aligned}
\hat{s}_{i,j}^{B} \;=\; & \underbrace{\left(\mathbf{v}_i^{(j)}\right)^H \mathbf{H}_{L_i} \bigotimes_{m=L-1}^{1} \left(c_m \mathbf{F}_m \mathbf{H}_m\right) \mathbf{u}_i^{(j)} \left(q_i^{(j)}\right)^{\frac{1}{2}} s_{i,j}^{B}}_{\text{desired signal}} \\
& + \underbrace{\left(\mathbf{v}_i^{(j)}\right)^H \mathbf{H}_{L_i} \sum_{l=1, l\neq j}^{N_b^{(i)}} \bigotimes_{m=L-1}^{1} \left(c_m \mathbf{F}_m \mathbf{H}_m\right) \mathbf{u}_i^{(l)} \left(q_i^{(l)}\right)^{\frac{1}{2}} s_{i,l}^{B}}_{\text{intra-user interference}} \\
& + \underbrace{\left(\mathbf{v}_i^{(j)}\right)^H \mathbf{H}_{L_i} \sum_{k=1, k\neq i}^{K} \sum_{l=1}^{N_b^{(k)}} \bigotimes_{m=L-1}^{1} \left(c_m \mathbf{F}_m \mathbf{H}_m\right) \mathbf{u}_k^{(l)} \left(q_k^{(l)}\right)^{\frac{1}{2}} s_{k,l}^{B}}_{\text{inter-user interference}} \\
& + \underbrace{\left(\mathbf{v}_i^{(j)}\right)^H c_{L-1} \mathbf{H}_{L_i} \mathbf{F}_{L-1} \left(\sum_{k=2}^{L-1} \bigotimes_{m=L-1}^{k} \left(c_{m-1} \mathbf{H}_m \mathbf{F}_{m-1}\right) \mathbf{n}_k + \mathbf{n}_L \right)}_{\text{noise propagated from previous hops}} \\
& + \underbrace{\left(\mathbf{v}_i^{(j)}\right)^H \mathbf{n}_{L+1}^{(i)}}_{\text{noise at the destination}} \quad , \qquad j = 1, \ldots, N_b^{(i)}, \quad i = 1, \ldots, K. \qquad (7.7)
\end{aligned}$$

7.2.2 Multi-hop MAC MIMO Relay System

For the multi-hop MAC MIMO relay system shown in Fig. 7.2, the roles of the BS and user nodes at the BC MIMO relay system are swapped. The Hermitian transpose of the channel matrices used in the BC system are employed in the MAC system. The ith user node linearly precodes the symbol vector $\mathbf{s}_i^M = \left[s_{i,1}^M, s_{i,2}^M, \ldots, s_{i,N_b^{(i)}}^M\right]^T$ using the matrix $\mathbf{V}_i \mathbf{P}_i^{\frac{1}{2}}$, where $\mathbf{P}_i = \text{diag}\left\{p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(N_b^{(i)})}\right\}$ with $p_i^{(j)}$, $j = 1, \ldots, N_b^{(i)}$, $i = 1, \ldots, K$, being the power allocated to the jth data stream of the ith user. The lth node, i.e. $(l-1)$-th relay node, $l = 2, \ldots, L$, employs $\mathbf{F}_{L+1-l}^H$ to amplify and forward received signals. The received signal vector at the first and the $(L+1-l)$-th receiving node of the MAC system

is given, respectively, by

$$\mathbf{y}_2^M = \mathbf{H}_L^H \mathbf{V} \mathbf{P}^{\frac{1}{2}} \mathbf{s}^M + \mathbf{n}_L \tag{7.8}$$

$$\mathbf{y}_{L+2-l}^M = \bigotimes_{m=l}^{L-1} \left(\mathbf{H}_m^H \mathbf{F}_m^H\right) \mathbf{H}_L^H \mathbf{V} \mathbf{P}^{\frac{1}{2}} \mathbf{s}^M + \mathbf{n}_l + \sum_{k=l}^{L-1} \bigotimes_{m=l}^{k} \left(\mathbf{H}_m^H \mathbf{F}_m^H\right) \mathbf{n}_{k+1}, \quad l = 1, \ldots, L-1. \tag{7.9}$$

Here $\mathbf{s}^M = \left[\left(\mathbf{s}_1^M\right)^T, \ldots, \left(\mathbf{s}_K^M\right)^T\right]^T$, $\mathbf{V} = \mathrm{bd}\left(\mathbf{V}_1, \ldots, \mathbf{V}_K\right)$, $\mathbf{P} = \mathrm{bd}\left(\mathbf{P}_1, \ldots, \mathbf{P}_K\right)$, and $\mathbf{H}_L^H = \left[\mathbf{H}_{L_1}^H, \mathbf{H}_{L_2}^H \ldots, \mathbf{H}_{L_K}^H\right]$. A linear receiver matrix $\mathbf{U}_i$ is used at the BS to estimate the transmitted symbol vector of user i, and the estimated symbol vector $\hat{\mathbf{s}}_i^M$ is expressed as

$$\hat{\mathbf{s}}_i^M = \mathbf{U}_i^H \mathbf{y}_{L+1}^M. \tag{7.10}$$

Using (7.8)-(7.10), we have the decision variable of the jth data symbol of the ith user as

$$\begin{aligned} \hat{s}_{i,j}^M = & \underbrace{\left(\mathbf{u}_i^{(j)}\right)^H \bigotimes_{m=1}^{L-1} \left(\mathbf{H}_m^H \mathbf{F}_m^H\right) \mathbf{H}_{L_i}^H \mathbf{v}_i^{(j)} \left(p_i^{(j)}\right)^{\frac{1}{2}} s_{i,j}^M}_{\text{desired signal}} \\ & + \underbrace{\left(\mathbf{u}_i^{(j)}\right)^H \sum_{l=1, l\neq j}^{N_b^{(i)}} \bigotimes_{m=1}^{L-1} \left(\mathbf{H}_m^H \mathbf{F}_m^H\right) \mathbf{H}_{L_i}^H \mathbf{v}_i^{(l)} \left(p_i^{(l)}\right)^{\frac{1}{2}} s_{i,l}^M}_{\text{intra-user interference}} \\ & + \underbrace{\left(\mathbf{u}_i^{(j)}\right)^H \sum_{k=1, k\neq i}^{K} \sum_{l=1}^{N_b^{(k)}} \bigotimes_{m=1}^{L-1} \left(\mathbf{H}_m^H \mathbf{F}_m^H\right) \mathbf{H}_{L_k}^H \mathbf{v}_k^{(l)} \left(p_k^{(l)}\right)^{\frac{1}{2}} s_{k,l}^M}_{\text{inter-user interference}} \\ & + \underbrace{\left(\mathbf{u}_i^{(j)}\right)^H \sum_{k=1}^{L-1} \bigotimes_{m=1}^{k} \left(\mathbf{H}_m^H \mathbf{F}_m^H\right) \mathbf{n}_{k+1}}_{\text{noise propagated from previous hops}} + \underbrace{\left(\mathbf{u}_i^{(j)}\right)^H \mathbf{n}_1}_{\text{noise at the destination}}. \end{aligned} \tag{7.11}$$

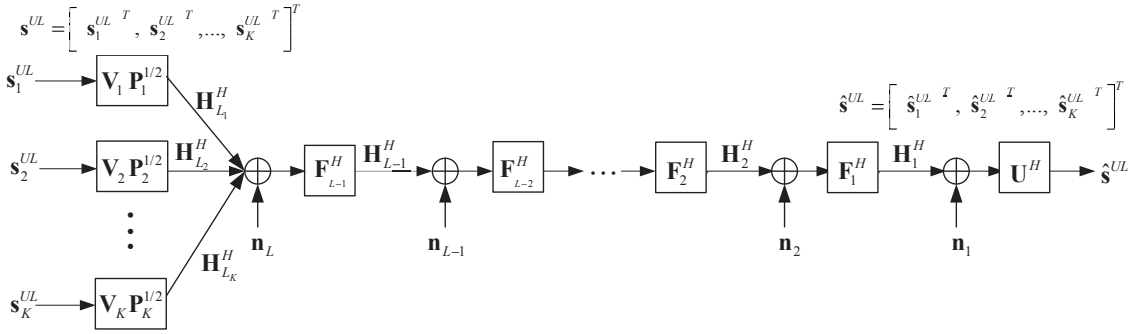

Figure 7.2: Multi-hop MAC AF MIMO relay system

7.3 Channel Model

Unlike [215], [216], and [227], where the exact CSI is perfectly known, in this chapter, we investigate the MAC-BC duality under imperfect CSI at each hop. There are two classes of models frequently used to model imperfect CSI: the Bayesian (stochastic) and the deterministic (or worst-case) models. In the stochastic model, the channel is usually modeled as a complex random matrix with normally distributed elements. The system design is then based on optimizing the stochastic measure of the system performance, such as the mean or outage performance under the assumption that the transmitter knows the mean and/or the covariance. On the other hand, the worst-case model assumes that the instantaneous channel, though not exactly known, lies in a known set of possible values. The error belongs to a predefined uncertainty region (with no inherent statistical assumption). In this case, the final objective is to optimize the worst system performance in this error region, which leads to a maximin formulation. In this chapter, we consider the stochastic

model, where the true CSI at each hop is modeled as

$$\mathbf{H}_l = \hat{\mathbf{R}}_l^{\frac{1}{2}} \mathbf{H}_{w,l} \mathbf{T}_l^{\frac{1}{2}}, \qquad l = 1, \ldots, L-1 \tag{7.12}$$

$$\mathbf{H}_{L_i} = \hat{\mathbf{R}}_{L_i}^{\frac{1}{2}} \mathbf{H}_{w,L_i} \mathbf{T}_{L_i}^{\frac{1}{2}}, \quad i = 1, \ldots, K \tag{7.13}$$

where the elements of $\mathbf{H}_{w,l}$ and $\mathbf{H}_{w,L_i}$ are i.i.d. zero mean circularly symmetric complex Gaussian random variables all with unit variance, $\mathbf{T}_l \in \mathcal{C}^{N_l \times N_l}$ and $\hat{\mathbf{R}}_l \in \mathcal{C}^{N_{l+1} \times N_{l+1}}$ are (from the perspective of BC system) antenna correlation matrices at the transmitter end of the lth node and the receiver end of the $(l+1)$-th node, respectively [232], [233]. Similarly, $\mathbf{T}_{L_i} \in \mathcal{C}^{N_L \times N_L}$ and $\hat{\mathbf{R}}_{L_i} \in \mathcal{C}^{N_{L+1}^{(i)} \times N_{L+1}^{(i)}}$ are antenna correlation matrices at the transmitter end of the $(L-1)$-th relay node and the ith user node, respectively.

We assume that channel estimation is performed on $\mathbf{H}_{w,l}$ and $\mathbf{H}_{w,L_i}$ using the orthogonal training method developed in [232]. Based on (7.12) and (7.13), the true channels $\mathbf{H}_l$, $\mathbf{H}_{L_i}$, and their minimum mean-squared error (MMSE) estimates $\tilde{\mathbf{H}}_l$, $\tilde{\mathbf{H}}_{L_i}$ are related as [232]

$$\mathbf{H}_l = \tilde{\mathbf{H}}_l + \mathbf{R}_l^{\frac{1}{2}} \mathbf{E}_{w,l} \mathbf{T}_l^{\frac{1}{2}} = \tilde{\mathbf{H}}_l + \mathbf{E}_l, \quad l = 1, \ldots, L-1$$

$$\mathbf{H}_{L_i} = \tilde{\mathbf{H}}_{L_i} + \mathbf{R}_{L_i}^{\frac{1}{2}} \mathbf{E}_{w,L_i} \mathbf{T}_{L_i}^{\frac{1}{2}} = \tilde{\mathbf{H}}_{L_i} + \mathbf{E}_{L_i}, \; i = 1, \ldots, K$$

where $\mathbf{R}_l \triangleq \left(\mathbf{I}_{N_{l+1}} + \sigma_{e,l}^2 \hat{\mathbf{R}}_l^{-1} \right)^{-1}$, $\mathbf{E}_l \triangleq \mathbf{R}_l^{\frac{1}{2}} \mathbf{E}_{w,l} \mathbf{T}_l^{\frac{1}{2}}$ is the estimation error of $\mathbf{H}_l$, $\mathbf{R}_{L_i} \triangleq \left(\mathbf{I}_{N_{L+1}^{(i)}} + \sigma_{e,L_i}^2 \hat{\mathbf{R}}_{L_i}^{-1} \right)^{-1}$ and $\mathbf{E}_{L_i} \triangleq \mathbf{R}_{L_i}^{\frac{1}{2}} \mathbf{E}_{w,L_i} \mathbf{T}_{L_i}^{\frac{1}{2}}$ is the estimation error of $\mathbf{H}_{L_i}$. Here, the entries of $\mathbf{E}_{w,l}$ and $\mathbf{E}_{w,L_i}$ are i.i.d. with $\mathcal{CN}(0, \sigma_{e,l}^2)$ and $\mathcal{CN}(0, \sigma_{e,L_i}^2)$, respectively. For the ease of explanation, let us take the BC system as an example. We assume that the lth node ($(l-1)$-th relay node), $l = 2, \ldots, L$, estimates its backward channel $\mathbf{H}_{l-1}$ through channel training and feeds the estimated CSI back to the $(l-1)$-th node without any error. Thus,

both the lth and the $(l-1)$-th nodes have the same CSI mismatch on $\mathbf{H}_{l-1}$. Similarly, the ith user node estimates its backward channel $\mathbf{H}_{L_i}$ through channel training and feeds the estimated CSI back to the $(L-1)$-th relay node without any error. In this chapter, we assume that $\mathbf{E}_{w,l}$ is unknown, but $\tilde{\mathbf{H}}_l, \mathbf{R}_l, \mathbf{T}_l$, and $\sigma^2_{e,l}$ are available at the lth and the $(l+1)$-th nodes, $l = 1, \ldots, L-1$. Similarly, we assume that $\mathbf{E}_{w,L_i}$ is unknown, but $\tilde{\mathbf{H}}_{L_i}, \mathbf{R}_{L_i}, \mathbf{T}_{L_i}$, and σ^2_{e,L_i} are available at the $(L-1)$-th relay node and ith user node, $i = 1, \ldots, K$.

7.4 MAC-BC Duality

In this section, we establish the duality between the MAC and BC multi-hop AF MIMO relay systems under imperfect CSI and antenna correlation at each hop. We define duality as the achievement of identical stream-wise SINR (MSE, capacity) at the MAC and the BC systems with the same amount of total network transmission power under imperfect CSI. In order to establish this duality, given a MAC MIMO relay system, we need to determine the scaling factors $c_l,\ l = 1, \ldots, L-1$, of the relay amplifying matrix $\mathbf{F}_l$ and the source power allocation matrix $\mathbf{Q}$ in the BC system.

Note that for the simplicity of presenting the proof of the duality results in this chapter, we analyze the duality for the single user case. Let us group all users together in BC and MAC to form one "super" destination node and one "super" source node with N_{L+1} antennas, respectively. The BC system and the MAC system can be equivalently viewed as a single-user downlink and uplink multi-hop MIMO relay system, respectively. The following theorem establishes the MAC-BC duality of multi-hop MIMO relay communication system under imperfect channel model.

Theorem 1. *Let $\mathbf{F}_l^H$ and $c_l\mathbf{F}_l$ be the relay amplifying matrices at the multi-hop MIMO MAC and BC systems, respectively. Under imperfect CSI, stream-wise identical SINRs (duality) in the MAC and BC systems can be achieved through the following two approaches:*

1. *With the transmission power constraint at individual nodes, duality holds by setting $P_{L+1-l}^M = P_l^B,\ l = 1, \ldots, L$, where P_l^B and P_l^M are the total transmitted powers at the lth BC and MAC node, respectively. The values of $c_l,\ l = 1, \ldots, L-1$, can be obtained from relay transmission power constraints of the BC. In other words, duality can be achieved by employing $\mathbf{F}_{L-l}^H$ and $c_l\mathbf{F}_l$ respectively as the relay amplifying matrix at the lth relay node of the MAC system and the BC system, $l = 1, \ldots, L-1$, and the scaling factor c_l is obtained by switching the power constraints at the lth node of the BC system and the $(L+1-l)$-th node of the MAC system, $l = 1. \ldots, L$.*

2. *Under a total network power constraint, MAC-BC duality holds when $\mathbf{F}_l$ and $\mathbf{F}_l^H,\ l = 1, \ldots, L-1$ are the relay amplifying matrices used in the BC and MAC, respectively. In other words, MAC-BC duality holds when $c_l = 1,\ l = 1, \ldots, L-1$.*

Proof. See Appendix 7.A. □

Theorem 1 includes the SINR duality results in [211]-[217] as special cases. It extends the MAC-BC SINR duality results to multi-hop AF MIMO relay systems under imperfect CSI, and thus generalizes all previous SINR duality results. Note that Theorem 1 holds for any linear transceiver matrices $\mathbf{U}, \mathbf{V}$, and linear relay amplifying matrices $\mathbf{F}_l,\ l = 1, \ldots, L-1$.

Theorem 2. *If the source node of the BC MIMO relay system employs DPC and the destination node of the MAC MIMO relay system employs SIC, then under imperfect CSI, stream-wise identical SINRs (duality) in the MAC and BC systems can be achieved through the same two approaches in Theorem 1.*

Proof. See Appendix 7.B. □

Theorem 2 extends the SINR duality results in Theorem 1 to the scenario where nonlinear transceivers are used at the source node of the BC system and the destination node of the MAC system. Similar to Theorem 1, Theorem 2 holds for any transceiver matrices $\mathbf{U}, \mathbf{V}$, and relay amplifying matrices $\mathbf{F}_l,\ l = 1, \ldots, L-1$. By choosing $\mathbf{V}$ (the destination receiving matrix in the BC and the source precoding matrix in the MAC) as a block diagonal matrix, i.e., $\mathbf{V} = \mathrm{bd}\,(\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_K)$, both Theorem 1 and Theorem 2 are applicable to multiuser MIMO relay scenario. Moreover, the two duality results are proved under the fair condition that MAC and BC systems consume the same amount of total transmission power.

We can also derive the MSE duality based on the SINR duality by using the relation of $\mathrm{MSE}_i^B = 1/(1 + \mathrm{SINR}_i^B),\ i = 1, \ldots, N_b$, as shown in Appendix 7.C, where MSE_i^B and SINR_i^B stand for the MSE and SINR of the ith data stream in the BC system, respectively. Thus identical SINR values in the MAC and BC systems imply identical MSE values. Therefore, it can be concluded from Theorem 1 that each MSE point achievable in the MAC system can be attained in the BC system. Clearly, the converse holds as well.

Based on the SINR duality $\mathrm{SINR}_i^M = \mathrm{SINR}_i^B,\ i = 1, \ldots, N_b$, where SINR_i^M is the SINR of the ith data stream of the MAC system. we can prove the MAC-BC capacity

duality as

$$\begin{aligned} C^M &= \sum_{i=1}^{N_b} \log_2\left(1+\mathrm{SINR}_i^M\right) \\ &= \sum_{i=1}^{N_b} \log_2\left(1+\mathrm{SINR}_i^B\right) \\ &= C^B \end{aligned} \tag{7.14}$$

where C^M and C^B are the capacities of the MAC and BC systems, respectively.

7.5 Appendix 7.A: Proof of Theorem 1

In order to establish the SINR duality for multi-hop AF MIMO relay systems under imperfect CSI, we have to show the conditions on $\mathbf{P}$, $\mathbf{Q}$, and $c_l,\ l = 1, \ldots, L-1$, that identical SINRs are achieved in both the MAC and BC systems. The proof consists of the following three main steps.

1. We write the total transmission power and the SINR of each stream for both the MAC and BC systems using (7.8)-(7.11) and (7.3)-(7.7), respectively.

2. We rewrite the total transmission power of the BC system obtained in Step 1 based on the definition of duality that both channels should achieve identical SINRs.

3. Using the final expression of the total transmission power of the BC system obtained in Step 2, we find the conditions on $\mathbf{P}$, $\mathbf{Q}$, and $c_l,\ l = 1, \ldots, L-1$, such that both the MAC and BC systems consume the same amount of total transmission power.

7.5.1 Step 1

In this step, we first write the total required transmission power for the MAC and BC systems. For this purpose, we first express the transmitted power at each node of both systems.

Using (7.8) and (7.9), the individual transmission power for all the transmitting nodes in the MAC system can be written as

$$P_1^M = tr\{\mathbf{P}\} \tag{7.15}$$

$$\begin{aligned} P_2^M &= \mathbb{E}\left[tr\{\mathbf{F}_{L-1}^H\left(\mathbf{H}_L^H\mathbf{V}\mathbf{P}\mathbf{V}^H\mathbf{H}_L + \mathbf{I}_{N_L}\right)\mathbf{F}_{L-1}\}\right] \\ &= tr\{\mathbf{F}_{L-1}^H\mathbf{F}_{L-1}\} \\ &\quad + tr\left\{\left(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{F}_{L-1}^H\}\mathbf{R}_L\right)\mathbf{V}\mathbf{P}\mathbf{V}^H\right\} \end{aligned} \tag{7.16}$$

$$\begin{aligned} P_{L+2-l}^M &= \mathbb{E}\left[tr\left\{\mathbf{F}_{l-1}^H\mathbf{y}_{L+2-l}^M\left(\mathbf{y}_{L+2-l}^M\right)^H\mathbf{F}_{l-1}\right\}\right] \\ &= tr\left\{\left(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr\left\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\right\}\mathbf{R}_L\right)\mathbf{V}\mathbf{P}\mathbf{V}^H\right\} \\ &\quad + \underbrace{\sum_{k=l}^{L-1} tr\left\{\mathbf{F}_{l-1}^H\mathbf{D}_{l-1}^{(k)}\mathbf{F}_{l-1}\right\}}_{\text{propagated noise power}} + \underbrace{tr\{\mathbf{F}_{l-1}^H\mathbf{F}_{l-1}\}}_{\text{noise power}},\ l = 2,\ldots,L-1. \end{aligned} \tag{7.17}$$

where $\mathbf{B}_m^{(n)}$ and $\mathbf{D}_m^{(n)}$ are recursive functions and given by

$$\mathbf{B}_m^{(n)} = \begin{cases} \mathbf{I}_{N_n}, & \text{if } m = n \\ \tilde{\mathbf{H}}_{m-1}\mathbf{F}_{m-2}\mathbf{B}_{m-1}^{(n)}\mathbf{F}_{m-2}^H\tilde{\mathbf{H}}_{m-1}^H + \sigma_{e,m-1}^2 tr\left\{\mathbf{T}_{m-1}\mathbf{F}_{m-2}\mathbf{B}_{m-1}^{(n)}\mathbf{F}_{m-2}^H\right\}\mathbf{R}_{m-1}, & \text{o.w.} \end{cases}$$

$$\mathbf{D}_m^{(n)} = \begin{cases} \mathbf{I}_{N_{n+1}}, & \text{if } m = n \\ \tilde{\mathbf{H}}_{m+1}^H\mathbf{F}_{m+1}^H\mathbf{D}_{m+1}^{(n)}\mathbf{F}_{m+1}\tilde{\mathbf{H}}_{m+1} + \sigma_{e,m+1}^2 tr\left\{\mathbf{R}_{m+1}\mathbf{F}_{m+1}^H\mathbf{D}_{m+1}^{(n)}\mathbf{F}_{m+1}\right\}\mathbf{T}_{m+1}, & \text{o.w.} \end{cases}$$

The total transmission power of the MAC system P_T^M can be calculated with the summation of all individual powers (7.15)-(7.17). We would like to note that since the exact CSI is unknown, the transmission power is averaged over the imperfect CSI through the expectation operations in (7.16) and (7.17) with respect to $\mathbf{E}_{w,l}$. Here, we have used the result of $\mathbb{E}\left[\mathbf{E}\mathbf{A}\mathbf{E}^H\right] = \sigma_e^2 tr\{\mathbf{A}\}\mathbf{I}$, when the entries of $\mathbf{E}$ are i.i.d. with $\mathcal{CN}(0,\sigma_e^2)$ and $\mathbf{A}$ is a given matrix [232].

Using (7.3) and (7.4), the individual transmission power P_l^B of the lth node, $l = 1,\ldots,L$, in the BC system can be written as

$$
\begin{aligned}
P_1^B &= tr\{\mathbf{Q}\}, && (7.18)\\
P_2^B &= \mathbb{E}\left[tr\{c_1^2\mathbf{F}_1\left(\mathbf{H}_1\mathbf{U}\mathbf{Q}\mathbf{U}^H\mathbf{H}_1^H + \mathbf{I}_{N_2}\right)\mathbf{F}_1^H\}\right]\\
&= tr\{c_1^2\mathbf{F}_1\mathbf{F}_1^H\} + c_1^2 tr\left\{\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{F}_1\tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{F}_1\}\mathbf{T}_1\right)\mathbf{U}\mathbf{Q}\mathbf{U}^H\right\}, && (7.19)\\
P_{l+1}^B &= \mathbb{E}\left[c_l^2 tr\left\{\mathbf{F}_l\mathbf{y}_{l+1}^B\left(\mathbf{y}_{l+1}^B\right)^H\mathbf{F}_l^H\right\}\right]\\
&= \left(\prod_{m=1}^{l} c_m^2\right) tr\left\{\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\right\}\mathbf{T}_1\right)\mathbf{U}\mathbf{Q}\mathbf{U}^H\right\}\\
&\quad + \underbrace{\sum_{k=2}^{l}\left(\prod_{m=k-1}^{l} c_m^2\right) tr\left\{\mathbf{F}_l\mathbf{B}_{l+1}^{(k)}\mathbf{F}_l^H\right\}}_{\text{propagated noise power}} + \underbrace{tr\{c_l^2\mathbf{F}_l\mathbf{F}_l^H\}}_{\text{noise power}},\ l = 2,\ldots,L-1. && (7.20)
\end{aligned}
$$

Note that the expectations in (7.19) and (7.20) are taken with respect to $\mathbf{E}_{w,l}$. The total

transmission power of the BC system can be calculated using (7.18)-(7.20) and given by

$$
\begin{aligned}
&P_T^B \\
&= P_1^B + P_2^B + \sum_{l=2}^{L-1} P_{l+1}^B \\
&= tr\{\mathbf{UQU}^H\} + \sum_{l=1}^{L-1} \left(\prod_{m=1}^{l} c_m^2 \right) tr\left\{ \left(\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{D}_1^{(l)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{D}_1^{(l)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right) \mathbf{UQU}^H \right\} \\
&\quad + \sum_{l=1}^{L-1} tr\{c_l^2 \mathbf{F}_l \mathbf{F}_l^H\} + \sum_{l=2}^{L-1} \sum_{k=2}^{l} \left(\prod_{m=k-1}^{l} c_m^2 \right) tr\left\{ \mathbf{F}_l \mathbf{B}_{l+1}^{(k)} \mathbf{F}_l^H \right\}. \qquad (7.21)
\end{aligned}
$$

Then we write the SINR of each stream for the MAC and BC systems. SINR is defined as the ratio of the signal power to the summation of the interference power (interference from all other data streams), total noise power (propagated noise from previous hops plus the thermal noise at the destination node), and the residual power of the signal due to the channel estimation error. By using (7.7), the SINR of the ith data stream, $i = 1, \ldots, N_b$ at the destination node of the BC system is given by[2]

$$
\mathrm{SINR}_i^B = \frac{q_i \left(\prod_{m=1}^{L-1} c_m^2 \right) \left| \mathbf{v}_i^H \tilde{\mathbf{H}}_L \bigotimes_{l=L-1}^{1} \left(\mathbf{F}_l \tilde{\mathbf{H}}_l \right) \mathbf{u}_i \right|^2}{P_{I_i}^B}. \qquad (7.22)
$$

where $P_{I_i}^B$ is the total interference plus noise power of the ith stream, $i = 1, \ldots, N_b$, in the BC system, and is shown at the bottom of the following page.

The first term $R_{I_i}^B$ in $P_{I_i}^B$ is the residual interference power of the desired signal

[2]For the stochastic channel estimation error model adopted in Section 7.3, the channel estimation error is seen as noise [217, 232, 195].

in (7.7) stemming from the mismatch between the true and estimated CSI, and is given by

$$R_{I_i}^B = q_i \prod_{m=1}^{L-1} c_m^2 \mathbf{v}_i^H \Bigg[\sigma_{e,L}^2 tr\Big\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(i)} \mathbf{F}_{L-1}^H \Big\} \mathbf{R}_L + \tilde{\mathbf{H}}_L \Bigg(\sum_{k=1}^{L-2} \sigma_{e,k}^2 tr \Big\{ \mathbf{T}_k \mathbf{F}_{k-1} \mathbf{A}_{k-1}^{(i)} \mathbf{F}_{k-1}^H \Big\} \bigotimes_{m=L-1}^{k+1} \Big(\mathbf{F}_m \tilde{\mathbf{H}}_m \Big) \mathbf{F}_k \mathbf{R}_k \mathbf{F}_k^H \bigotimes_{m=k+1}^{L-1} \Big(\tilde{\mathbf{H}}_m^H \mathbf{F}_m^H \Big) + \sigma_{e,L-1}^2 tr \Big\{ \mathbf{T}_{L-1} \mathbf{F}_{L-2} \mathbf{A}_{L-2}^{(i)} \mathbf{F}_{L-2}^H \Big\} \mathbf{F}_{L-1} \mathbf{R}_{L-1} \mathbf{F}_{L-1}^H \Big) \tilde{\mathbf{H}}_L^H \Bigg] \mathbf{v}_i \tag{7.24}$$

where $\mathbf{A}_k^{(j)}$ are recursively defined as

$$\mathbf{A}_k^{(j)} = \begin{cases} \mathbf{F}_0^{-1} \mathbf{u}_j \mathbf{u}_j^H \mathbf{F}_0^{-H}, & \text{if } k = 0 \\ \tilde{\mathbf{H}}_1 \mathbf{u}_j \mathbf{u}_j^H \tilde{\mathbf{H}}_1^H + \sigma_{e,1}^2 tr\{\mathbf{T}_1 \mathbf{u}_j \mathbf{u}_j^H\} \mathbf{R}_1, & \text{if } k = 1 \\ \tilde{\mathbf{H}}_k \mathbf{F}_{k-1} \mathbf{A}_{k-1}^{(j)} \mathbf{F}_{k-1}^H \tilde{\mathbf{H}}_k^H + \sigma_{e,k}^2 tr \Big\{ \mathbf{T}_k \mathbf{F}_{k-1} \mathbf{A}_{k-1}^{(j)} \mathbf{F}_{k-1}^H \Big\} \mathbf{R}_k, & \text{if } k \geq 2. \end{cases} \tag{7.25}$$

Here we introduced the matrix $\mathbf{F}_0$ in (7.25) for the simplicity of presentation. In particular, $\mathbf{F}_0$ is an invertible matrix that is canceled by $\mathbf{F}_0^{-1}$ in $\mathbf{A}_0^{(i)}$ when $k = 1$ in the second term of (7.24).

By using (7.11), the SINR of the ith data stream, $i = 1, \ldots, N_b$ at the destination

---

$$P_{I_i}^B \tag{7.23}$$

$$= R_{I_i}^B + \underbrace{\sum_{j=1, j \neq i}^{N_b} q_j \left(\prod_{m=1}^{L-1} c_m^2 \right) \mathbf{v}_i^H \Big[\tilde{\mathbf{H}}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr \Big\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \Big\} \mathbf{R}_L \Big] \mathbf{v}_i}_{\text{interference power}}$$

$$+ \underbrace{\mathbf{v}_i^H \Bigg[\sum_{l=2}^{L} \Bigg(\prod_{m=l-1}^{L-1} c_m^2 \Bigg) \Big(\tilde{\mathbf{H}}_L \mathbf{F}_{L-1} \mathbf{B}_L^{(l)} \mathbf{F}_{L-1}^H \tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr \Big\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{B}_L^{(l)} \mathbf{F}_{L-1}^H \Big\} \mathbf{R}_L \Big) + \mathbf{I}_{N_L} \Bigg] \mathbf{v}_i}_{\text{propagated noise power + noise power at the destination}}$$

node of the MAC MIMO relay channel is given by

$$\mathrm{SINR}_i^M = \frac{p_i \left| \mathbf{u}_i^H \bigotimes_{l=1}^{L-1} \left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H \right) \tilde{\mathbf{H}}_L^H \mathbf{v}_i \right|^2}{P_{I_i}^M}. \tag{7.26}$$

where $P_{I_i}^M$ is the total interference plus noise power of the ith stream, $i = 1, \ldots, N_b$, in the MAC system, and is shown at the bottom of the current page. The first term $R_{I_i}^M$ in $P_{I_i}^M$ is the residual interference power of the desired signal in (7.11) stemming from the channel estimation error and is given by

$$\begin{aligned} R_{I_i}^M &= p_i \mathbf{u}_i^H \left[\sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{C}_{L-2}^{(i)} \mathbf{F}_1 \right\} \mathbf{T}_1 + \sum_{l=2}^{L} \sigma_{e,l}^2 tr\left\{ \mathbf{R}_l \mathbf{F}_l^H \mathbf{C}_{L-l-1}^{(i)} \mathbf{F}_l \right\} \right. \\ &\quad \left. \bigotimes_{k=1}^{l-1} \left(\tilde{\mathbf{H}}_k^H \mathbf{F}_k^H \right) \mathbf{T}_l \bigotimes_{k=l-1}^{1} \left(\mathbf{F}_k \tilde{\mathbf{H}}_k \right) \right] \mathbf{u}_i. \end{aligned} \tag{7.28}$$

where $\mathbf{C}_k^{(j)}$ are given by

$$\mathbf{C}_k^{(j)} = \begin{cases} \mathbf{F}_L^{-H} \mathbf{v}_j \mathbf{v}_j^H \mathbf{F}_L^{-1}, & \text{if } k = -1 \\ \tilde{\mathbf{H}}_L^H \mathbf{v}_j \mathbf{v}_j^H \tilde{\mathbf{H}}_L + \sigma_{e,L}^2 tr\{\mathbf{R}_L \mathbf{v}_j \mathbf{v}_j^H\} \mathbf{T}_L, & \text{if } k = 0 \\ \tilde{\mathbf{H}}_{L-k}^H \mathbf{F}_{L-k}^H \mathbf{C}_{k-1}^{(j)} \mathbf{F}_{L-k} \tilde{\mathbf{H}}_{L-k} + \sigma_{e,L-k}^2 tr\left\{ \mathbf{R}_{L-k} \mathbf{F}_{L-k}^H \mathbf{C}_{k-1}^{(j)} \mathbf{F}_{L-k} \right\} \mathbf{T}_{L-k}, & \text{if } k \geq 1. \end{cases}$$

$$\begin{aligned} P_{I_i}^M &= R_{I_i}^M + \underbrace{\mathbf{u}_i^H \sum_{j=1, j \neq i}^{N_b} p_j \left[\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right] \mathbf{u}_i}_{\text{interference power}} \\ &\quad + \underbrace{\mathbf{u}_i^H \left[\sum_{l=1}^{L-1} \left(\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{D}_1^{(l)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{D}_1^{(l)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right) + \mathbf{I}_{N_1} \right] \mathbf{u}_i}_{\text{propagated noise power + noise power at the destination}} \\ &\triangleq \mathbf{u}_i^H \mathbf{M}_i \mathbf{u}_i. \end{aligned} \tag{7.27}$$

Similar to $\mathbf{F}_0$ in (7.25), here we introduced the matrix $\mathbf{F}_L$ for the simplicity of presentation. In particular, $\mathbf{F}_L$ is an invertible matrix that is canceled by $\mathbf{F}_L^{-1}$ in $\mathbf{C}_{-1}^{(i)}$ when $l = L$ in the second term of (7.28).

7.5.2 Step 2

In this step, we rewrite the total transmission power of the BC system obtained in Step 1 based on the definition of the SINR duality.

In order to achieve identical SINRs at the MAC and BC systems, $\mathrm{SINR}_i^B = \mathrm{SINR}_i^M$, $i = 1, \ldots, N_b$ must be satisfied. Using (7.22) and (7.26), such SINR equality leads to $\prod_{m=1}^{L-1} c_m^2 q_i P_{I_i}^M = p_i P_{I_i}^B$. Summing this over all N_b streams, i.e., $\sum_{i=1}^{N_b} \mathrm{SINR}_i^B = \sum_{i=1}^{N_b} \mathrm{SINR}_i^M$, we have

$$\sum_{i=1}^{N_b} \left(\prod_{m=1}^{L-1} c_m^2 \right) q_i P_{I_i}^M = \sum_{i=1}^{N_b} p_i P_{I_i}^B. \tag{7.29}$$

By substituting (7.24) and (7.27) into (7.29), it can be seen that only the last terms (noise terms) in (7.24) and (7.27) remain and the other terms related to interference are canceled out, since[3]

$$\begin{aligned} &\sum_{i=1}^{N_b} \prod_{m=1}^{L-1} c_m^2 q_i \left(R_{I_i}^M + \sum_{j=1, j\neq i}^{N_b} p_j \mathbf{u}_i^H \left[\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right] \mathbf{u}_i \right) \\ = \quad &\sum_{i=1}^{N_b} p_i \left(R_{I_i}^B + \sum_{j=1, j\neq i}^{N_b} q_j \left(\prod_{m=1}^{L-1} c_m^2 \right) \mathbf{v}_i^H \left[\tilde{\mathbf{H}}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \tilde{\mathbf{H}}_L^H \right. \right. \\ &\qquad \left. \left. + \sigma_{e,L}^2 tr\left\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \right\} \mathbf{R}_L \right] \mathbf{v}_i \right). \end{aligned} \tag{7.30}$$

[3]The relation in (7.30) was used for single-hop, two-hop and multi-hop channels in [237], [215] and [216], respectively.

In other words, (7.29) can be written as

$$\sum_{i=1}^{N_b}\prod_{m=1}^{L-1}c_m^2 q_i\left(\mathbf{u}_i^H\left[\sum_{l=1}^{L-1}\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1+\sigma_{e,1}^2 tr\left\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\right\}\mathbf{T}_1\right)\right]\mathbf{u}_i+1\right)$$
$$=\sum_{i=1}^{N_b}p_i\left(\mathbf{v}_i^H\left[\sum_{l=2}^{L}\prod_{m=l-1}^{L-1}c_m^2\left(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H+\sigma_{e,L}^2 tr\left\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\right\}\mathbf{R}_L\right)\right]\mathbf{v}_i+1\right). \tag{7.31}$$

Substituting (7.31) back into (7.21), P_T^B can be written as

$$\begin{aligned}
P_T^B &= tr\{\mathbf{U}\mathbf{Q}\mathbf{U}^H\}+\sum_{l=1}^{L-1}\left(\prod_{m=1}^{l}c_m^2\right)tr\left\{\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1\right.\right.\\
&\qquad\left.\left.+\sigma_{e,1}^2 tr\left\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\right\}\mathbf{T}_1\right)\mathbf{U}\mathbf{Q}\mathbf{U}^H\right\}\\
&\quad+\sum_{l=2}^{L-1}\sum_{k=2}^{l}\left(\prod_{m=k-1}^{l}c_m^2\right)tr\left\{\mathbf{F}_l\mathbf{B}_{l+1}^{(k)}\mathbf{F}_l^H\right\}+\sum_{l=1}^{L-1}tr\{c_l^2\mathbf{F}_l\mathbf{F}_l^H\}-\left(\prod_{m=1}^{L-1}c_m^2\right)tr\{\mathbf{U}\mathbf{Q}\mathbf{U}^H\}\\
&\quad-\sum_{l=1}^{L-1}\left(\prod_{m=1}^{L-1}c_m^2\right)tr\left\{\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1+\sigma_{e,1}^2 tr\left\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\right\}\mathbf{T}_1\right)\mathbf{U}\mathbf{Q}\mathbf{U}^H\right\}\\
&\quad+\sum_{l=2}^{L}\left(\prod_{m=l-1}^{L-1}c_m^2\right)tr\left\{\left(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H\right.\right.\\
&\qquad\left.\left.+\sigma_{e,L}^2 tr\left\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\right\}\mathbf{R}_L\right)\mathbf{V}\mathbf{P}\mathbf{V}^H+\mathbf{P}\right\}.
\end{aligned}$$

where the first four terms are from (7.21) and the last three terms are from (7.31).

After some simple manipulations, P_T^B can be written as

$$\begin{aligned}
&P_T^B \qquad\qquad (7.32)\\
&=\sum_{l=2}^{L}\left(\prod_{m=l-1}^{L-1}c_m^2\right)tr\left\{\left(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H+\sigma_{e,L}^2 tr\left\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\right\}\mathbf{R}_L\right)\mathbf{V}\mathbf{P}\mathbf{V}^H+\mathbf{P}\right\}\\
&+\sum_{l=1}^{L-2}\left(1-\prod_{m=l+1}^{L-1}c_m^2\right)tr\left\{\left(\prod_{m=1}^{l}c_m^2\right)\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1+\sigma_{e,1}^2 tr\left\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\right\}\mathbf{T}_1\right)\mathbf{U}\mathbf{Q}\mathbf{U}^H\right\}\\
&+\left(1-\prod_{m=1}^{L-1}c_m^2\right)tr\{\mathbf{U}\mathbf{Q}\mathbf{U}^H\}+\sum_{l=1}^{L-1}tr\left\{c_l^2\mathbf{F}_l\mathbf{F}_l^H\right\}+\sum_{l=2}^{L-1}\sum_{k=2}^{l}\left(\prod_{m=k-1}^{l}c_m^2\right)tr\left\{\mathbf{F}_l\mathbf{B}_{l+1}^{(k)}\mathbf{F}_l^H\right\}.
\end{aligned}$$

For notational simplicity, for $l = 2, \ldots, L-1$, we denote

$$a_l \triangleq \sum_{k=2}^{l} \left(\prod_{m=k-1}^{l} c_m^2 \right) tr\Big\{ \mathbf{F}_l \mathbf{B}_{l+1}^{(k)} \mathbf{F}_l^H \Big\}. \tag{7.33}$$

Using (7.33) and with some manipulations, we have

$$\begin{aligned}
\sum_{l=2}^{L-1} a_l &= \sum_{l=2}^{L-1} \left(\prod_{m=l+1}^{L-1} c_m^2 \right) a_l + \sum_{l=2}^{L-1} \left(1 - \prod_{m=l+1}^{L-1} c_m^2 \right) a_l \\
&= \sum_{l=2}^{L-1} \sum_{k=2}^{l} \left(\prod_{m=k-1}^{L-1} c_m^2 \right) tr\left\{ \mathbf{F}_l \mathbf{B}_{l+1}^{(k)} \mathbf{F}_l^H \right\} + \sum_{l=2}^{L-1} \left(1 - \prod_{m=l+1}^{L-1} c_m^2 \right) a_l \\
&= \sum_{l=2}^{L-1} \sum_{k=l}^{L-1} \left(\prod_{m=l-1}^{L-1} c_m^2 \right) tr\left\{ \mathbf{F}_k \mathbf{B}_{k+1}^{(l)} \mathbf{F}_k^H \right\} + \sum_{l=2}^{L-1} \left(1 - \prod_{m=l+1}^{L-1} c_m^2 \right) a_l \\
&= \sum_{l=2}^{L-1} \left(\prod_{m=l}^{L} c_{m-1}^2 \right) \sum_{k=l}^{L-1} tr\left\{ \mathbf{F}_{l-1}^H \mathbf{D}_{l-1}^{(k)} \mathbf{F}_{l-1} \right\} + \sum_{l=2}^{L-1} \left(1 - \prod_{m=l+1}^{L-1} c_m^2 \right) a_l. \qquad (7.34)
\end{aligned}$$

where we have used the fact that $tr\Big\{ \mathbf{F}_k \mathbf{B}_{k+1}^{(l)} \mathbf{F}_k^H \Big\} = tr\Big\{ \mathbf{F}_{l-1}^H \mathbf{D}_{l-1}^{(k)} \mathbf{F}_{l-1} \Big\}$. We can also write

$$\begin{aligned}
\sum_{l=1}^{L-1} tr\{c_l^2 \mathbf{F}_l \mathbf{F}_l^H\} &= \sum_{l=1}^{L-1} tr\{c_l^2 \mathbf{F}_l \mathbf{F}_l^H\} + \underbrace{\sum_{l=2}^{L} \prod_{m=l}^{L} c_{m-1}^2 tr\{\mathbf{F}_{l-1}^H \mathbf{F}_{l-1}\} - \sum_{l=1}^{L-1} \left(\prod_{m=l}^{L-1} c_m^2 \right) tr\{\mathbf{F}_l \mathbf{F}_l^H\}}_{\text{equal to zero}} \\
&= \sum_{l=2}^{L} \prod_{m=l}^{L} c_{m-1}^2 tr\{\mathbf{F}_{l-1}^H \mathbf{F}_{l-1}\} + \sum_{l=1}^{L-1} \left(1 - \prod_{m=l+1}^{L-1} c_m^2 \right) tr\{c_l^2 \mathbf{F}_l \mathbf{F}_l^H\}. \qquad (7.35)
\end{aligned}$$

Substituting (7.34) and (7.35) back into (7.33) and after rearranging terms, P_T^B can be re-written as

$$P_T^B = \underbrace{\sum_{l=2}^{L-1}\prod_{m=l}^{L} c_{m-1}^2 tr\Big\{\Big(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr\Big\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\Big\}\mathbf{R}_L\Big)\mathbf{VPV}^H\Big\}}_{\text{part of first term in (7.33)}}$$

$$+\underbrace{c_{L-1}^2 tr\Big\{\Big(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{F}_{L-1}^H\}\mathbf{R}_L\Big)\mathbf{VPV}^H\Big\} + tr\{\mathbf{P}\}}_{\text{part of first term in (7.33), particularly when } l=L}$$

$$+\underbrace{\sum_{l=2}^{L-2}\Big(1-\prod_{m=l+1}^{L-1}c_m^2\Big)tr\Big\{\prod_{m=1}^{l}c_m^2\Big(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\Big\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\Big\}\mathbf{T}_1\Big)\mathbf{UQU}^H\Big\}}_{\text{part of second term in (7.33)}}$$

$$+\underbrace{\Big(1-\prod_{m=2}^{L-1}c_m^2\Big)tr\Big\{c_1^2\Big(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{F}_1\tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{F}_1\}\mathbf{T}_1\Big)\mathbf{UQU}^H\Big\}}_{\text{part of second term in (7.33), particularly when } l=1}$$

$$+\underbrace{\Big(1-\prod_{m=1}^{L-1}c_m^2\Big)tr\{\mathbf{UQU}^H\}}_{\text{third term in (7.33)}} + \underbrace{\sum_{l=2}^{L-1}\prod_{m=l}^{L}c_{m-1}^2 tr\{\mathbf{F}_{l-1}^H\mathbf{F}_{l-1}\} + c_{L-1}^2 tr\{\mathbf{F}_{L-1}^H\mathbf{F}_{L-1}\}}_{\text{first part of the fourth term in (7.33) given in (7.35)}}$$

$$+\underbrace{\sum_{l=2}^{L-2}\Big(1-\prod_{m=l+1}^{L-1}c_m^2\Big)tr\{c_l^2\mathbf{F}_l\mathbf{F}_l^H\} + \Big(1-\prod_{m=2}^{L-1}c_m^2\Big)tr\{c_1^2\mathbf{F}_1\mathbf{F}_1^H\}}_{\text{second part of the fourth term in (7.33) given in (7.35)}}$$

$$+\underbrace{\sum_{l=2}^{L-1}\prod_{m=l}^{L}c_{m-1}^2 tr\Big\{\sum_{k=l}^{L-1}\mathbf{F}_{l-1}^H\mathbf{D}_{l-1}^{(k)}\mathbf{F}_{l-1}\Big\} + \sum_{l=2}^{L-2}\Big(1-\prod_{m=l+1}^{L-1}c_m^2\Big)tr\Big\{\sum_{k=2}^{l}\prod_{m=k-1}^{l}c_m^2\mathbf{F}_l\mathbf{B}_{l+1}^{(k)}\mathbf{F}_l^H\Big\}}_{\text{last term in (7.33)}}$$

$$= \sum_{l=2}^{L-1}\prod_{m=l}^{L}c_{m-1}^2 tr\Big\{\Big(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr\Big\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{B}_L^{(l)}\mathbf{F}_{L-1}^H\Big\}\mathbf{R}_L\Big)\mathbf{VPV}^H$$
$$+\sum_{k=l}^{L-1}\mathbf{F}_{l-1}^H\mathbf{D}_{l-1}^{(k)}\mathbf{F}_{l-1} + \mathbf{F}_{l-1}^H\mathbf{F}_{l-1}\Big\}$$
$$+c_{L-1}^2 tr\Big\{\Big(\tilde{\mathbf{H}}_L\mathbf{F}_{L-1}\mathbf{F}_{L-1}^H\tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr\{\mathbf{T}_L\mathbf{F}_{L-1}\mathbf{F}_{L-1}^H\}\mathbf{R}_L\Big)\mathbf{VPV}^H + \mathbf{F}_{L-1}^H\mathbf{F}_{L-1}\Big\} + tr\{\mathbf{P}\}$$
$$+\sum_{l=2}^{L-2}\Big(1-\prod_{m=l+1}^{L-1}c_m^2\Big)tr\Big\{\prod_{m=1}^{l}c_m^2\Big(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\Big\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{D}_1^{(l)}\mathbf{F}_1\Big\}\mathbf{T}_1\Big)\mathbf{UQU}^H$$

$$+\sum_{k=2}^{l}\prod_{m=k-1}^{l} c_m^2\mathbf{F}_l\mathbf{B}_{l+1}^{(k)}\mathbf{F}_l^H + c_l^2\mathbf{F}_l\mathbf{F}_l^H\Bigg\} + \left(1-\prod_{m=1}^{L-1}c_m^2\right) tr\{\mathbf{U}\mathbf{Q}\mathbf{U}^H\}$$
$$+\left(1-\prod_{m=2}^{L-1}c_m^2\right) tr\left\{c_1^2\left(\tilde{\mathbf{H}}_1^H\mathbf{F}_1^H\mathbf{F}_1\tilde{\mathbf{H}}_1+\sigma_{e,1}^2 tr\{\mathbf{R}_1\mathbf{F}_1^H\mathbf{F}_1\}\mathbf{T}_1\right)\mathbf{U}\mathbf{Q}\mathbf{U}^H + c_1^2\mathbf{F}_1\mathbf{F}_1^H\right\} \quad (7.36)$$

7.5.3 Step 3

In this step, using the final expression of the total transmission power of the BC system in (7.36), we find the conditions on $\mathbf{P}$, $\mathbf{Q}$, and $c_l,\ l = 1,\ldots,L-1$, such that both the MAC and BC systems consume the same amount of total transmission power.

Using P_l^M in (7.15)-(7.17) and P_l^B in (7.18)-(7.20), $l = 1.\ldots,L$, (7.36) can be written as

$$\begin{aligned} P_T^B &= \sum_{l=2}^{L-1}\left(\prod_{m=l}^{L} c_{m-1}^2\right) P_{L+2-l}^M + c_{L-1}^2 P_2^M + P_1^M + \sum_{l=2}^{L-2}\left(1-\prod_{m=l+1}^{L-1}c_m^2\right) P_{l+1}^B \\ &\quad + \sum_{l=1}^{2}\left(1-\prod_{m=l}^{L-1}c_m^2\right) P_l^B \\ &= \sum_{l=1}^{L-1}\left(\prod_{m=l}^{L-1} c_m^2\right) P_{L+1-l}^M + P_1^M + \sum_{l=1}^{L-1}\left(1-\prod_{m=l}^{L-1}c_m^2\right) P_l^B. \end{aligned} \quad (7.37)$$

By adding and subtracting $\left(\sum_{l=1}^{L-1} P_{L+1-l}^M\right)$ to and from (7.37), which is the total relay transmission power for the MAC system, we obtain[4]

$$P_T^B - P_T^M = \sum_{l=1}^{L-1}\left(\prod_{m=l}^{L-1} c_m^2 - 1\right)\left(P_{L+1-l}^M - P_l^B\right). \quad (7.38)$$

Since the MAC and BC systems should consume the same amount of total transmission power, we need to find the conditions such that $P_T^B - P_T^M = 0$. Obviously, for any $L \geq 2$, the condition $P_T^B - P_T^M = 0$ is true if $\prod_{m=l}^{L-1} c_m^2 = 1$ for $l = 1,\ldots,L-1$, which is

[4] P_T^M on the left hand side of (7.38) is obtained by adding $\sum_{l=1}^{L-1} P_{L+1-l}^M$ and P_1^M in (7.37).

equivalent to $c_l = 1,\ l = 1, \ldots, L-1$. Thus, the first part of Theorem 1 (without transmission power constraint at each node) is proven. Moreover, the condition $P_T^B - P_T^M = 0$ also holds if $P_{L+1-l}^M = P_l^B,\ l = 1, \ldots, L-1$. Then we have $P_1^M = P_L^B$ due to the fair assumption $P_T^B = P_T^M$. Thus, we have $P_{L+1-l}^M = P_l^B,\ l = 1, \ldots, L$ and the second part of Theorem 1 (with transmission power constraint at individual nodes) is proven.

7.6 Appendix 7.B: Proof of Theorem 2

When the destination node of a MAC MIMO relay system employs SIC, the source symbols are detected successively with the last symbol detected first and the first symbol detected last, and thus the interference from the previously detected symbols is subtracted to detect the current symbol. Therefore, the SINR of the ith data stream, $i = 1, \ldots, N_b$ at the MAC MIMO relay system is written as

$$\mathrm{SINR}_i^M = \frac{p_i \left| \mathbf{u}_i^H \bigotimes_{l=1}^{L-1} \left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H \right) \tilde{\mathbf{H}}_L^H \mathbf{v}_i \right|^2}{P_{I_i}^M}. \tag{7.39}$$

where $P_{I_i}^M$ is the total interference plus noise power of the ith stream, $i = 1, \ldots, N_b$, in the MAC system, and can be expressed as

$$\begin{aligned} P_{I_i}^M &= R_{I_i}^M + \mathbf{u}_i^H \sum_{j=1}^{i-1} p_j \left[\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right] \mathbf{u}_i \\ &\quad + \mathbf{u}_i^H \left[\sum_{l=1}^{L-1} \left(\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{D}_1^{(l)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr\left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{D}_1^{(l)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right) + \mathbf{I}_{N_1} \right] \mathbf{u}_i. \end{aligned} \tag{7.40}$$

where $R_{I_i}^M$ is defined in (7.28).

When the source node of a BC MIMO relay system employs DPC, the source symbols are encoded successively with the first symbol encoded first and the last symbol

encoded last, and thus the interference from previously encoded symbols is subtracted to encode the current symbol. Therefore, the SINR of the ith data stream, $i = 1, \ldots, N_b$ at the BC MIMO relay system is written as

$$\mathrm{SINR}_i^B = \frac{q_i \left(\prod_{m=1}^{L-1} c_m^2 \right) \left| \mathbf{v}_i^H \tilde{\mathbf{H}}_L \bigotimes_{l=L-1}^{1} \left(\mathbf{F}_l \tilde{\mathbf{H}}_l \right) \mathbf{u}_i \right|^2}{P_{I_i}^B}. \tag{7.41}$$

where $P_{I_i}^B$ is the total interference plus noise power of the ith stream, $i = 1, \ldots, N_b$, in the BC system, and can be written as

$$\begin{aligned} & P_{I_i}^B \qquad (7.42) \\ &= \sum_{j=i+1}^{N_b} q_j \left(\prod_{m=1}^{L-1} c_m^2 \right) \mathbf{v}_i^H \left[\tilde{\mathbf{H}}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr \left\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \right\} \mathbf{R}_L \right] \mathbf{v}_i + R_{I_i}^B \\ &\quad + \mathbf{v}_i^H \left[\sum_{l=2}^{L} \left(\prod_{m=l-1}^{L-1} c_m^2 \right) \left(\tilde{\mathbf{H}}_L \mathbf{F}_{L-1} \mathbf{B}_L^{(l)} \mathbf{F}_{L-1}^H \tilde{\mathbf{H}}_L^H + \sigma_{e,L}^2 tr \left\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{B}_L^{(l)} \mathbf{F}_{L-1}^H \right\} \mathbf{R}_L \right) + \mathbf{I}_{N_L} \right] \mathbf{v}_i \end{aligned}$$

where $R_{I_i}^B$ is defined in (7.24).

Using (7.39) and (7.41), and the identity below (similar to (7.30))

$$\begin{aligned} & \sum_{i=1}^{N_b} \prod_{m=1}^{L-1} c_m^2 q_i \left(R_{I_i}^M + \sum_{j=1}^{i-1} p_j \mathbf{u}_i^H \left[\tilde{\mathbf{H}}_1^H \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \tilde{\mathbf{H}}_1 + \sigma_{e,1}^2 tr \left\{ \mathbf{R}_1 \mathbf{F}_1^H \mathbf{C}_{L-2}^{(j)} \mathbf{F}_1 \right\} \mathbf{T}_1 \right] \mathbf{u}_i \right) \\ = & \sum_{i=1}^{N_b} p_i \left(R_{I_i}^B + \sum_{j=i+1}^{N_b} q_j \left(\prod_{m=1}^{L-1} c_m^2 \right) \mathbf{v}_i^H \left[\tilde{\mathbf{H}}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \tilde{\mathbf{H}}_L^H \right. \right. \\ & \left. \left. + \sigma_{e,L}^2 tr \left\{ \mathbf{T}_L \mathbf{F}_{L-1} \mathbf{A}_{L-1}^{(j)} \mathbf{F}_{L-1}^H \right\} \mathbf{R}_L \right] \mathbf{v}_i \right). \end{aligned} \tag{7.43}$$

we obtain P_T^B as in (7.33) from $\sum_{i=1}^{N_b} \mathrm{SINR}_i^B = \sum_{i=1}^{N_b} \mathrm{SINR}_i^M$, and the steps in (7.33)-(7.38) are still the same. Thus Theorem 2 is proven.

7.7 Appendix 7.C: SINR-MSE Relation

To prove the MSE duality, we first rewrite SINR_i^M in (7.26) as

$$\begin{aligned}\mathrm{SINR}_i^M &= \frac{p_i \left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1}\left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H\right)\tilde{\mathbf{H}}_L^H \mathbf{v}_i\right|^2}{\mathbf{u}_i^H \mathbf{P}_i \mathbf{u}_i} \\ &\leq p_i \mathbf{v}_i^H \tilde{\mathbf{H}}_L \bigotimes_{l=L-1}^{1}\left(\mathbf{F}_l \tilde{\mathbf{H}}_l\right)\mathbf{P}_i^{-1}\bigotimes_{l=1}^{L-1}\left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H\right)\tilde{\mathbf{H}}_L^H \mathbf{v}_i.\end{aligned} \tag{7.44}$$

where $\mathbf{M}_i$ is defined in (7.27), and the inequality comes from Cauchy-Schwarz's inequality [238]. The upper bound is achieved by

$$\mathbf{u}_i^{\mathrm{SINR}} = \alpha_i \mathbf{M}_i^{-1} \bigotimes_{l=1}^{L-1}\left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H\right)\tilde{\mathbf{H}}_L^H \mathbf{v}_i. \tag{7.45}$$

Here $\alpha_i \neq 0$ is an arbitrary scalar.

Using (7.11), we can express the MSE of the ith data stream for the MAC channel as[5]

$$\begin{aligned}&\mathrm{MSE}_i^M \\ &= \mathbb{E}\left[|\hat{s}_i^M - s_i^M|^2\right] \\ &= \mathbb{E}\left[p_i \left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1}\left(\mathbf{H}_l^H \mathbf{F}_l^H\right)\mathbf{H}_L^H \mathbf{v}_i\right|^2\right] + \mathbf{u}_i^H \mathbf{u}_i + \sum_{j=1, j\neq i}^{N_b}\mathbb{E}\left[p_j \left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1}\left(\mathbf{H}_l^H \mathbf{F}_l^H\right)\mathbf{H}_L^H \mathbf{v}_j\right|^2\right] \\ &+ \sum_{k=1}^{L-1}\mathbb{E}\left[\mathbf{u}_i^H \bigotimes_{m=1}^{k}\left(\mathbf{H}_m^H \mathbf{F}_m^H\right)\bigotimes_{m=k}^{1}\left(\mathbf{F}_m \mathbf{H}_m\right)\mathbf{u}_i\right] - 2p_i^{\frac{1}{2}}\Re\left\{\mathbb{E}\left[\mathbf{u}_i^H \bigotimes_{l=1}^{L-1}\left(\mathbf{H}_l^H \mathbf{F}_l^H\right)\mathbf{H}_L^H \mathbf{v}_i\right]\right\} + 1.\end{aligned} \tag{7.46}$$

Using $\mathbf{M}_i$ defined in (7.27), (7.46) can be written as

$$\begin{aligned}\mathrm{MSE}_i^M &= p_i \left|\mathbf{u}_i^H \bigotimes_{l=1}^{L-1}\left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H\right)\tilde{\mathbf{H}}_L^H \mathbf{v}_i\right|^2 + \mathbf{u}_i^H \mathbf{M}_i \mathbf{u}_i \\ &\quad - 2p_i^{\frac{1}{2}}\Re\left\{\mathbf{u}_i^H \bigotimes_{l=1}^{L-1}\left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H\right)\tilde{\mathbf{H}}_L^H \mathbf{v}_i\right\} + 1.\end{aligned} \tag{7.47}$$

[5]Since we consider single-user downlink and uplink multi-hop MIMO relay system in our proof, the inter-user interference term from (11) can be eliminated. In other words, the inter-user interference becomes intra-user interference when we treat multiple users as a "super" node as in Section IV.

The optimal receive vector $\mathbf{u}_i$ minimizing (7.47) is the Wiener filter [239] and given by

$$\begin{aligned} \mathbf{u}_i^{\mathrm{MSE}} &= (\mathbf{M}_i + \tilde{\mathbf{v}}_i \tilde{\mathbf{v}}_i^H)^{-1} \tilde{\mathbf{v}}_i \\ &= \frac{\mathbf{M}_i^{-1} \tilde{\mathbf{v}}_i}{1 + \tilde{\mathbf{v}}_i^H \mathbf{M}_i^{-1} \tilde{\mathbf{v}}_i} \end{aligned} \tag{7.48}$$

where $\tilde{\mathbf{v}}_i \triangleq p_i^{\frac{1}{2}} \bigotimes_{l=1}^{L-1} \left(\tilde{\mathbf{H}}_l^H \mathbf{F}_l^H \right) \tilde{\mathbf{H}}_L^H \mathbf{v}_i$ and the matrix inversion lemma, which is given by $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B} \left(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1} \right)^{-1} \mathbf{DA}^{-1}$, is applied to obtain the second equation in (7.48).

Comparing (7.45) with (7.48), we find that $\mathbf{u}_i^{\mathrm{MSE}}$ in (7.48) also maximizes the SINR with $\alpha_i = p_i^{\frac{1}{2}} / (1 + \tilde{\mathbf{v}}_i^H \mathbf{M}_i^{-1} \tilde{\mathbf{v}}_i)$. By substituting (7.48) back into (7.47), the MSE of the ith data stream of the MAC system is given by

$$\begin{aligned} \mathrm{MSE}_i^M &= \frac{1}{1 + \tilde{\mathbf{v}}_i^H \mathbf{M}_i^{-1} \tilde{\mathbf{v}}_i} \\ &= \frac{1}{1 + \mathrm{SINR}_i^M}. \end{aligned} \tag{7.49}$$

Chapter 8

On Uplink-Downlink Sum-MSE Duality of Multi-hop MIMO Relay Channel

In this chapter, we establish the sum-MSE duality between uplink and downlink multi-hop AF MIMO relay channels, which is a generalization of several previously established results on sum-MSE duality. Unlike the previous methods that prove the duality by direct calculations of the MSEs for each stream, we introduce an interesting perspective to the relation of the uplink-downlink duality based on the Karush-Kuhn-Tucker (KKT) optimality conditions. Joint linear minimum sum-MSE transceiver optimization problems are formulated under the power constraints of the relays and user nodes for both uplink and downlink channels. Based on the KKT conditions associated with both optimization problems and by swapping the power constraints of the nodes in the downlink and uplink

channel, the uplink-downlink duality in sum-MSE is established. As a result, the sum-MSE in both uplink and downlink systems are the same and any achievable downlink system satisfying the KKT conditions can be transformed to the uplink system, vice versa.

8.1 Introduction

MSE is an important performance measure to approach the information-theoretic limits of Gaussian channels. In the uplink-downlink MSE duality, the MSEs remain the same during the conversion from uplink to downlink and vice versa. The stream-wise MSE duality was first derived for multiple-input single-output (MISO) systems with a sum power constraint [218] and then was extended to MIMO systems in [214, 219]. It was observed in [219] that under a total power constraint, any MSE point achievable in the uplink can also be achieved in the downlink. Stream-wise MSE duality was extended in [220] to sum-MSE and individual user-MSE dualities for MIMO systems.

All of the aforementioned MSE duality results are established by assuming that perfect channel state information (CSI) is available at all the nodes in the system. Motivated by this, the authors in [217] and [221] establish the MSE duality under imperfect CSI for single-hop MISO systems, which generalizes the MSE duality with perfect CSI in [218]. The MSE duality in [217] and [221] for MISO systems is extended to MIMO systems under imperfect CSI in [222]. None of [217, 221, 222] considers antenna correlation in their channel model. In [223], the sum-MSE uplink-downlink duality has been established for MIMO systems by considering imperfect CSI both at the base station and users, and with antenna correlation only at the base station. The duality result in [223] is established based on

the Karush-Kuhn-Tucker (KKT) optimality conditions for the uplink and downlink channel transceiver optimization problems. The sum-MSE duality result in [223] is extended in [224], where the authors have established three kinds of MSE duality (sum-MSE, individual user-MSE and stream-MSE) by incorporating the imperfect CSI and antenna correlation both at the base station and the users. The duality results in [224] can be seen as the extension of duality results in [220] to imperfect CSI case.

Recently, the uplink-downlink sum-MSE duality for single-hop systems [223] has been extended to two-hop AF MIMO relay systems, where all the nodes in the system are equipped with multiple antennas [225]. By considering the antenna correlation and channel estimation error at all nodes, the sum-MSE uplink-downlink duality is established by analyzing the KKT conditions of both uplink and downlink sum-MSE transceiver optimization problems. In [240], the stream-MSE duality is established for multi-hop MIMO relay channels under the imperfect CSI, which generalizes all of the previously published stream-MSE duality results.

8.1.1 Contributions of This Work

1. Unlike the methods in [214, 217, 218, 221, 240] that show the duality by direct calculations of the MSEs of each stream of all users, we use KKT conditions associated with the transceiver optimization problems of the uplink and downlink channels for the duality proof, and thus provide an interesting perspective to the relation of the uplink-downlink duality.

2. We show that uplink and downlink multi-hop AF MIMO relay channels share the

same achievable sum-MSE region under the transmission power constraints at the relays and the user nodes. Our proof generalizes the results in [223] and [225], which also use KKT conditions to prove the sum-MSE duality for single-hop and two-hop MIMO channels, respectively.

3. Unlike [240] that shows the sum-MSE duality for multi-hop AF MIMO relay systems under the assumption that receivers employ linear minimum MSE (MMSE) receivers, the sum-MSE duality result in this chapter is applicable to any linear receiver.

We show that the sum-MSE duality for multi-hop MIMO AF relay system can be achieved based on the KKT conditions of the sum-MSE transceiver optimization problems for both the uplink and downlink channels, and by swapping the power constraints at the lth node of the downlink system and the $(L+1-l)$th node of the uplink system, $l = 1, \ldots, L$, where L is the number of hops of the relay network. As a direct application of this sum-MSE duality, the complicated downlink multiuser MIMO relay system optimization problem can be carried out efficiently by focusing on an equivalent uplink multiuser MIMO relay system.

The following notations are used in this chapter. Matrices and vectors are denoted as bold capital and lowercase letters, respectively. For matrices, $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate transpose, respectively. $\mathbb{E}[\cdot]$ stands for the statistical expectation; $\mathbf{I}_N$ denotes an $N \times N$ identity matrix; $\mathrm{tr}(\cdot)$ stands for matrix trace. For matrices $\mathbf{A}_i$, $\bigotimes_{i=l}^{k}(\mathbf{A}_i) \triangleq \mathbf{A}_l \ldots \mathbf{A}_k$. $\prod_{i=l}^{k}(\mathbf{A}_i) \triangleq \mathbf{A}_l \ldots \mathbf{A}_k$ for $l \leq k$ and equal to identity matrix for $l > k$.

8.2 System Model

We consider a wireless communication system with K users, $L-1$ ($L \geq 2$) relay nodes, and one base station (BS) node, where each node is equipped with multiple antennas. The number of antennas at the lth relay node of the uplink system (users to BS) is $N_l,\ l = 1, \ldots, L-1$ and the BS is equipped with N_L antennas. We assume that the signal transmitted by the lth node can only be received by the $(l+1)$th node due to the propagation path-loss. Thus, source signals travel through L hops before reaching their destination. The ith user transmits (receives) M_i independent data streams using M_i antennas. Thus, the total number of independent data streams from all users is $N_0 = \sum_{i=1}^{K} M_i$. To be able to support N_0 data streams in each transmission, there is $N_0 \leq \min(N_1, N_2, \ldots, N_L)$. Each relay node works in half-duplex mode and employs a linear AF (non-regenerative) relay matrix to amplify and forward its received signals.

8.2.1 Uplink MIMO Relay System

For the uplink MIMO relay system shown in Fig. 8.1, the ith user linearly precodes the symbol vector $\mathbf{s}_i^{UL} \in C^{M_i \times 1}$ by the source precoding matrix $\mathbf{B}_i \in C^{M_i \times M_i}$ and the precoded signal vector $\mathbf{u}_i = \mathbf{B}_i \mathbf{s}_i^{UL}$ is transmitted from the ith user to the first relay node. We assume independent unit-power transmit symbols, i.e., $\mathbb{E}\left[\mathbf{s}_i^{UL} \left(\mathbf{s}_i^{UL}\right)^H\right] = \mathbf{I}_{M_i}$. The received signal at the first relay node is given by

$$\mathbf{y}_1^{UL} = \sum_{i=1}^{K} \mathbf{G}_i \mathbf{B}_i \mathbf{s}_i^{UL} + \mathbf{v}_1 \tag{8.1}$$

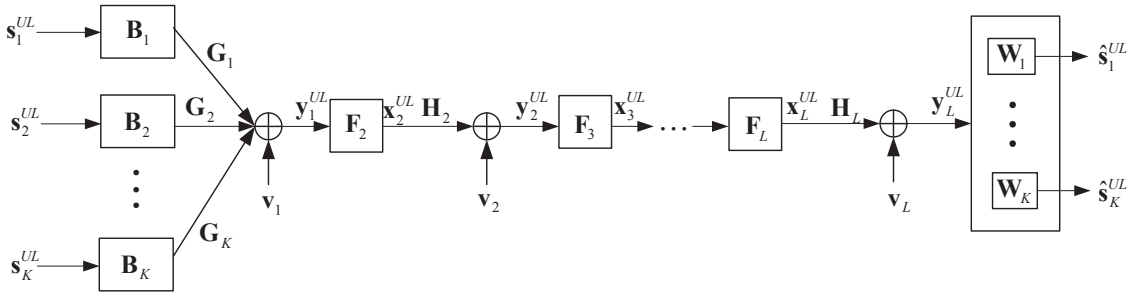

Figure 8.1: Uplink multi-hop AF MIMO relay system.

where $\mathbf{G}_i \in C^{N_1 \times M_i}$, $i = 1, \ldots, K$, is the MIMO channel matrix between the first relay node and the ith user and $\mathbf{v}_1$ is the $N_1 \times 1$ independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vector at the first relay node.

The lth relay node, $l = 1, \ldots, L-1$, employs $\mathbf{F}_{l+1} \in C^{N_l \times N_l}$ to amplify and forward the received signals. The transmitted signal vector from the lth relay node is given by

$$\mathbf{x}_{l+1}^{UL} = \mathbf{F}_{l+1}\mathbf{y}_l^{UL}, \qquad l = 1, \ldots, L-1 \tag{8.2}$$

where $\mathbf{y}_l^{UL} \in C^{N_l \times 1}$ is the signal vector received at the lth relay node, $l = 1, \ldots, L-1$. From (8.1) and (8.2), the received signal vector at the relay nodes, $l = 1, \ldots, L-1$, and the received signal vector at the BS ($l = L$) can be written as

$$\mathbf{y}_l^{UL} = \mathbf{A}_l \sum_{i=1}^{K} \mathbf{G}_i \mathbf{B}_i \mathbf{s}_i^{UL} + \bar{\mathbf{v}}_l, \qquad l = 1, \ldots, L \tag{8.3}$$

where $\mathbf{A}_l$ is the equivalent MIMO channel matrix from the first relay node to the lth relay

node and $\bar{\mathbf{v}}_l$ is the equivalent noise vector given by

$$\mathbf{A}_l = \begin{cases} \bigotimes_{i=l}^{2} (\mathbf{H}_i \mathbf{F}_i), & l = 2, \ldots, L \\ \mathbf{I}_{N_1}, & l = 1 \end{cases} \tag{8.4}$$

$$\bar{\mathbf{v}}_l = \begin{cases} \sum_{j=2}^{l} \left(\bigotimes_{i=l}^{j} (\mathbf{H}_i \mathbf{F}_i) \, \mathbf{v}_{j-1} \right) + \mathbf{v}_l, & l = 2, \ldots, L \\ \mathbf{v}_1, & l = 1 \end{cases}. \tag{8.5}$$

Here $\mathbf{H}_l \in C^{N_l \times N_{l-1}}$, $l = 2, \ldots, L$, is the MIMO channel matrix of the lth hop, and $\mathbf{v}_l$ is the i.i.d. AWGN vector at the $(l+1)$-th node of the uplink system, $l = 1, \ldots, L$. We assume that all noises are complex circularly symmetric with zero mean and unit variance.

From (8.5), the covariance matrix of $\bar{\mathbf{v}}_l$ can be written as,

$$\mathbf{C}_l = \mathbb{E}\left[\bar{\mathbf{v}}_l \bar{\mathbf{v}}_l^H\right] = \begin{cases} \sum_{j=2}^{l} \left(\bigotimes_{i=l}^{j} (\mathbf{H}_i \mathbf{F}_i) \bigotimes_{i=j}^{l} \left(\mathbf{F}_i^H \mathbf{H}_i^H \right) \right) + \mathbf{I}_{N_l} & l = 2, \ldots, L \\ \mathbf{I}_{N_1} & l = 1 \end{cases}. \tag{8.6}$$

With a linear receiver at the BS, the estimated signal vector is given by

$$\begin{aligned} \hat{\mathbf{s}}_j^{UL} &= \mathbf{W}_j \mathbf{y}_L^{UL} \\ &= \mathbf{W}_j \left[\mathbf{A}_L \sum_{i=1}^{K} \mathbf{G}_i \mathbf{B}_i \mathbf{s}_i^{UL} + \bar{\mathbf{v}}_L \right], \qquad j = 1, \ldots, K \end{aligned} \tag{8.7}$$

where $\mathbf{W}_j$ is the $M_j \times N_L$ weight matrix of the linear receiver. From (8.3) and (8.7), the MSE matrix of the jth user can be written as

$$\begin{aligned} \mathbf{E}_j^{UL} &= \mathbb{E}\left[\left(\mathbf{s}_j^{UL} - \hat{\mathbf{s}}_j^{UL} \right) \left(\mathbf{s}_j^{UL} - \hat{\mathbf{s}}_j^{UL} \right)^H \right] \\ &= \mathbf{I}_{M_j} - \mathbf{W}_j \mathbf{A}_L \mathbf{G}_j \mathbf{B}_j - \mathbf{B}_j^H \mathbf{G}_j^H \mathbf{A}_L^H \mathbf{W}_j^H \\ &\quad + \mathbf{W}_j \left[\mathbf{A}_L \mathbf{A}^{UL} \mathbf{A}_L^H + \mathbf{C}_L \right] \mathbf{W}_j^H, \quad j = 1, \ldots, K \end{aligned} \tag{8.8}$$

where $\mathbf{A}^{UL} = \sum_{i=1}^{K} \mathbf{G}_i \mathbf{B}_i \mathbf{B}_i^H \mathbf{G}_i^H$.

The transmission power consumed at the lth relay node is

$$\begin{aligned} \mathrm{tr}\left(\mathbb{E}\left[\mathbf{x}_{l+1}^{UL}\left(\mathbf{x}_{l+1}^{UL}\right)^H\right]\right) &= \mathrm{tr}\left(\mathbf{F}_{l+1}\mathbb{E}\left[\mathbf{y}_l^{UL}\left(\mathbf{y}_l^{UL}\right)^H\right]\mathbf{F}_{l+1}^H\right) \\ &= \mathrm{tr}\left(\mathbf{F}_{l+1}\left(\mathbf{A}_l\sum_{i=1}^{K}\mathbf{G}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{G}_i^H\mathbf{A}_l^H + \mathbf{C}_l\right)\mathbf{F}_{l+1}^H\right) \\ &= \mathrm{tr}\left(\mathbf{F}_{l+1}\left(\mathbf{A}_l\mathbf{A}^{UL}\mathbf{A}_l^H + \mathbf{C}_l\right)\mathbf{F}_{l+1}^H\right),\ l = 1,\ldots,L-1. \quad (8.9) \end{aligned}$$

With the optimization variables $\{\mathbf{F}_l\}_{l=2}^{L}, \{\mathbf{B}_j\}_{j=1}^{K}, \{\mathbf{W}_j\}_{j=1}^{K}$, the uplink transceiver optimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{F}_l,\mathbf{B}_j,\mathbf{W}_j} \quad & \sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{E}_j^{UL}\right) && (8.10) \\ \text{s.t.} \quad & \sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{B}_j\mathbf{B}_j^H\right) \le P_1^{UL} && (8.11) \\ \text{s.t.} \quad & \mathrm{tr}\left(\mathbf{F}_l\left(\mathbf{A}_{l-1}\mathbf{A}^{UL}\mathbf{A}_{l-1}^H + \mathbf{C}_{l-1}\right)\mathbf{F}_l^H\right) \le P_l^{UL}, \quad l = 2,\ldots,L && (8.12) \end{aligned}$$

where (8.11) and (8.12) are the total transmit power at the users and transmission power constraints at each relay node, respectively, and P_l^{UL}, $l = 1,\ldots,L$, are the corresponding power budget.

8.2.2 Downlink MIMO Relay System

In the downlink communication channel shown in Fig. 8.2, the base station linearly precodes the symbol vector $\mathbf{s}_i^{DL} \in C^{M_i \times 1}$ destined for user i using the matrix $\mathbf{T}_i \in C^{N_L \times M_i}$. We assume independent unit-power transmit symbols, i.e., $\mathbb{E}\left[\mathbf{s}_i^{DL}\left(\mathbf{s}_i^{DL}\right)^H\right] = \mathbf{I}_{M_i}$. The base station transmits the $N_L \times 1$ linearly precoded symbol vector $\sum_{i=1}^{K}\mathbf{T}_i\mathbf{s}_i^{DL}$ and the $N_{L-1} \times 1$

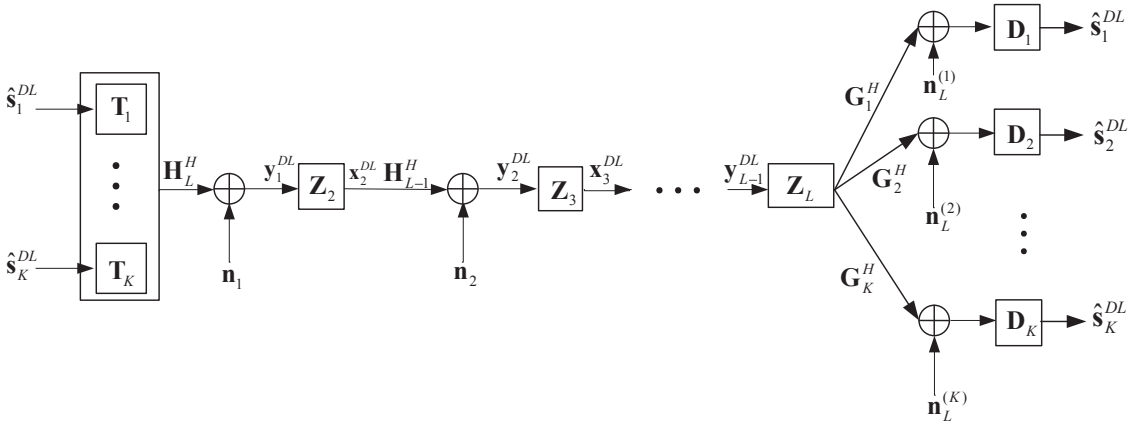

Figure 8.2: Downlink multi-hop AF MIMO relay system.

signal vector received at the first relay node of the downlink system can be written as

$$\mathbf{y}_1^{DL} = \mathbf{H}_L^H \sum_{i=1}^{K} \mathbf{T}_i \mathbf{s}_i^{DL} + \mathbf{n}_1 \tag{8.13}$$

where $\mathbf{n}_1 \in C^{N_{L-1} \times 1}$ is the i.i.d. AWGN vector at the first relay node.

The lth relay node in the downlink system, $l = 1, \ldots, L-1$ employs $\mathbf{Z}_{l+1} \in C^{N_{L-l} \times N_{L-l}}$ to amplify and forward the received signals. The transmitted signal vector from the lth relay node is given by

$$\mathbf{x}_{l+1}^{DL} = \mathbf{Z}_{l+1} \mathbf{y}_l^{DL}, \qquad l = 1, \ldots, L-1 \tag{8.14}$$

where $\mathbf{y}_l^{DL} \in C^{N_{L-l} \times 1}$, $l = 1, \ldots, L-1$, is the signal vector received at the lth relay node at the downlink system and can be written as

$$\mathbf{y}_l^{DL} = \mathbf{K}_l \mathbf{H}_L^H \sum_{i=1}^{K} \mathbf{T}_i \mathbf{s}_i^{DL} + \bar{\mathbf{n}}_l, \qquad l = 1, \ldots, L-1. \tag{8.15}$$

Here $\mathbf{K}_l$ is the equivalent MIMO channel matrix from the first relay node to the lth relay

node in the downlink channel and $\bar{\mathbf{n}}_l$ is the equivalent noise vector given by

$$\mathbf{K}_l = \begin{cases} \bigotimes_{m=L-l+1}^{L-1} \left(\mathbf{H}_m^H \mathbf{Z}_{L-m+1}\right), & l = 2, \ldots, L-1 \\ \mathbf{I}_{N_{L-1}}, & l = 1 \end{cases} \tag{8.16}$$

$$\bar{\mathbf{n}}_l = \begin{cases} \sum_{k=1}^{l-1} \bigotimes_{m=L-l+1}^{L-k} \left(\mathbf{H}_m^H \mathbf{Z}_{L-m+1}\right) \mathbf{n}_k + \mathbf{n}_l, & l = 2, \ldots, L-1 \\ \mathbf{n}_1, & l = 1 \end{cases} \tag{8.17}$$

where $\mathbf{n}_l$ is the i.i.d. AWGN vector at the lth relay node of the downlink system, $l = 1, \ldots, L-1$. We assume that all noises are complex circularly symmetric with zero mean and unit variance. The received signal vector at the ith user is written as

$$\begin{aligned} \mathbf{y}_L^{(i)} &= \mathbf{G}_i^H \mathbf{Z}_L \mathbf{y}_{L-1}^{DL} + \mathbf{n}_L^{(i)} \\ &= \mathbf{G}_i^H \mathbf{Z}_L \mathbf{K}_{L-1} \mathbf{H}_L^H \sum_{i=1}^{K} \mathbf{T}_i \mathbf{s}_i^{DL} + \bar{\mathbf{n}}_L^{(i)}, \qquad i = 1, \ldots, K \end{aligned} \tag{8.18}$$

where $\bar{\mathbf{n}}_L^{(i)} = \mathbf{G}_i^H \mathbf{Z}_L \bar{\mathbf{n}}_{L-1} + \mathbf{n}_L^{(i)}$ is the equivalent noise vector at the ith user.

From (8.17), the covariance matrix of the equivalent noise vector $\bar{\mathbf{n}}_l$ at the lth relay node, $l = 2, \ldots, L-1$ and the covariance matrix of the equivalent noise vector $\bar{\mathbf{n}}_L^{(i)}$ at the ith user can be written as

$$\begin{aligned} \mathbf{C}_l^{DL} &= \mathbb{E}\left[\bar{\mathbf{n}}_l \bar{\mathbf{n}}_l^H\right] \qquad (8.19) \\ &= \sum_{k=1}^{l-1} \left(\bigotimes_{m=L-l+1}^{L-k} \left(\mathbf{H}_m^H \mathbf{Z}_{L-m+1}\right) \bigotimes_{m=L-k}^{L-l+1} \left(\mathbf{Z}_{L-m+1}^H \mathbf{H}_m\right) \right) + \mathbf{I}_{N_{L-l}}, \; l = 2, \ldots, L-1 \\ \mathbf{C}_L^{(i)} &= \mathbb{E}\left[\bar{\mathbf{n}}_L^{(i)} \left(\bar{\mathbf{n}}_L^{(i)}\right)^H\right] \\ &= \mathbf{G}_i^H \mathbf{Z}_L \mathbf{C}_{L-1}^{DL} \mathbf{Z}_L^H \mathbf{G}_i + \mathbf{I}_{M_i}. \qquad (8.20) \end{aligned}$$

A linear receiver matrix $\mathbf{D}_j \in C^{M_j \times M_j}$ is applied at the jth user to estimate the

symbol vector $\mathbf{s}_j^{DL}$. The estimated symbol vector $\hat{\mathbf{s}}_j^{DL}$ is expressed as

$$\begin{aligned}\hat{\mathbf{s}}_j^{DL} &= \mathbf{D}_j \mathbf{y}_L^{(j)} \\ &= \mathbf{D}_j \mathbf{G}_j^H \mathbf{Z}_L \mathbf{K}_{L-1} \mathbf{H}_L^H \sum_{i=1}^{K} \mathbf{T}_i \mathbf{s}_i^{DL} + \mathbf{D}_j \bar{\mathbf{n}}_L^{(j)}, \quad j = 1, \ldots, K. \end{aligned} \tag{8.21}$$

From (8.21), the MSE matrix of the jth user can be written as

$$\begin{aligned}\mathbf{E}_j^{DL} &= \mathbb{E}\left[\left(\mathbf{s}_j^{DL} - \hat{\mathbf{s}}_j^{DL}\right)\left(\mathbf{s}_j^{DL} - \hat{\mathbf{s}}_j^{DL}\right)^H\right] \\ &= \mathbf{I}_{M_j} - \mathbf{D}_j \mathbf{G}_j^H \mathbf{Z}_L \mathbf{K}_{L-1} \mathbf{H}_L^H \mathbf{T}_j - \mathbf{T}_j^H \mathbf{H}_L \mathbf{K}_{L-1}^H \mathbf{Z}_L^H \mathbf{G}_j \mathbf{D}_j^H \\ &\quad + \mathbf{D}_j \left[\mathbf{G}_j^H \mathbf{Z}_L \mathbf{K}_{L-1} \mathbf{A}^{DL} \mathbf{K}_{L-1}^H \mathbf{Z}_L^H \mathbf{G}_j + \mathbf{C}_L^{(j)}\right] \mathbf{D}_j^H, \qquad j = 1, \ldots, K \end{aligned} \tag{8.22}$$

where $\mathbf{A}^{DL} = \mathbf{H}_L^H \sum_{i=1}^{K} \mathbf{T}_i \mathbf{T}_i^H \mathbf{H}_L$.

The transmission power consumed at the lth relay node is

$$\begin{aligned}\mathrm{tr}\left(\mathbb{E}\left[\mathbf{x}_{l+1}^{DL}\left(\mathbf{x}_{l+1}^{DL}\right)^H\right]\right) &= \mathrm{tr}\left(\mathbf{Z}_{l+1}\mathbb{E}\left[\mathbf{y}_l^{DL}\left(\mathbf{y}_l^{DL}\right)^H\right]\mathbf{Z}_{l+1}^H\right) \\ &= \mathrm{tr}\left(\mathbf{Z}_{l+1}\left(\mathbf{K}_l \mathbf{H}_L^H \sum_{i=1}^{K} \mathbf{T}_i \mathbf{T}_i^H \mathbf{H}_L \mathbf{K}_l^H + \mathbf{C}_l^{DL}\right)\mathbf{Z}_{l+1}^H\right) \\ &= \mathrm{tr}\left(\mathbf{Z}_{l+1}\left(\mathbf{K}_l \mathbf{A}^{DL} \mathbf{K}_l^H + \mathbf{C}_l^{DL}\right)\mathbf{Z}_{l+1}^H\right), \; l = 1, \ldots, L-1 \end{aligned} \tag{8.23}$$

With the optimization variables $\{\mathbf{Z}_l\}_{l=2}^{L}, \{\mathbf{T}_j\}_{j=1}^{K}, \{\mathbf{D}_j\}_{j=1}^{K}$, the downlink transceiver optimization problem can be formulated as

$$\min_{\mathbf{Z}_l, \mathbf{T}_j, \mathbf{D}_j} \quad \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{E}_j^{DL}\right) \tag{8.24}$$

$$\text{s.t.} \quad \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{T}_j \mathbf{T}_j^H\right) \le P_1^{DL} \tag{8.25}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{Z}_l\left(\mathbf{K}_{l-1}\mathbf{A}^{DL}\mathbf{K}_{l-1}^H + \mathbf{C}_{l-1}^{DL}\right)\mathbf{Z}_l^H\right) \le P_l^{DL}, \quad l = 2, \ldots, L \tag{8.26}$$

where (8.25) and (8.26) are the total transmit power at the users and transmission power

constraints at each relay node, respectively, and P_l^{DL}, $l = 1, \ldots, L$, are the corresponding power budget.

8.3 Uplink-Downlink Duality

The minimum sum-MSE transceiver optimization problems of uplink and downlink systems are both non-convex, but the objective and constraint functions are continuously differentiable. Thus the uplink-downlink duality can be established based on their KKT optimality conditions [241].

8.3.1 The KKT Conditions of the Uplink Problem

The Lagrangian function of the problem (8.10)-(8.12) is written as

$$\begin{aligned}\mathcal{L}^{UL} &= \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{E}_j^{UL}\right) + \lambda_1 \left(\sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{B}_j \mathbf{B}_j^H\right) - P_1^{UL} \right) \\ &\quad + \sum_{l=2}^{L} \lambda_l \left(\mathrm{tr}\left(\mathbf{F}_l \left(\mathbf{A}_{l-1} \mathbf{A}^{UL} \mathbf{A}_{l-1}^H + \mathbf{C}_{l-1} \right) \mathbf{F}_l^H \right) - P_l^{UL} \right) \end{aligned} \tag{8.27}$$

where λ_1 and λ_l, $l = 2, \ldots, L$, are the Lagrange multipliers associated with the total power constraint at the users and power constraint at the $(l-1)$-th relay node, respectively. The gradient condition associated with the Lagrangian function (8.27) is given by

$$\mathbf{G}_k^H \mathbf{A}_L^H \mathbf{W}_k^H = \left(\lambda_1 \mathbf{I}_{M_k} + \sum_{l=2}^{L} \lambda_l \mathbf{G}_k^H \mathbf{A}_{l-1}^H \mathbf{F}_l^H \mathbf{F}_l \mathbf{A}_{l-1} \mathbf{G}_k + \sum_{j=1}^{K} \mathbf{G}_k^H \mathbf{A}_L^H \mathbf{W}_j^H \mathbf{W}_j \mathbf{A}_L \mathbf{G}_k \right) \mathbf{B}_k \tag{8.28}$$

$$\sum_{j=1}^{K}\mathbf{H}_k^H\prod_{m=k+1}^{L}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\mathbf{W}_j^H\mathbf{B}_j^H\mathbf{G}_j^H\prod_{m=2}^{k-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)=$$
$$\sum_{j=1}^{K}\mathbf{H}_k^H\prod_{m=k+1}^{L}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\mathbf{W}_j^H\mathbf{W}_j\mathbf{A}_L\mathbf{A}^{UL}\prod_{m=2}^{k-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)$$
$$+\sum_{j=1}^{K}\sum_{m=2}^{k}\mathbf{H}_k^H\prod_{l=k+1}^{L}\left(\mathbf{F}_l^H\mathbf{H}_l^H\right)\mathbf{W}_j^H\mathbf{W}_j\bigotimes_{l=L}^{m}\left(\mathbf{H}_l\mathbf{F}_l\right)\prod_{l=m}^{k-1}\left(\mathbf{F}_l^H\mathbf{H}_l^H\right)$$
$$+\lambda_k\mathbf{F}_k\left(\mathbf{A}_{k-1}\mathbf{A}^{UL}\mathbf{A}_{k-1}^H+\mathbf{C}_{k-1}\right)$$
$$+\sum_{l=k+1}^{L}\lambda_l\Bigg(\mathbf{H}_k^H\prod_{m=k+1}^{l-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\mathbf{F}_l^H\mathbf{F}_l\mathbf{A}_{l-1}\mathbf{A}^{UL}\prod_{m=2}^{k-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)$$
$$+\sum_{j=2}^{k}\mathbf{H}_k^H\prod_{m=k+1}^{l-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\mathbf{F}_l^H\mathbf{F}_l\bigotimes_{i=l-1}^{j}\left(\mathbf{H}_i\mathbf{F}_i\right)\prod_{m=j}^{k-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\Bigg) \tag{8.29}$$

$$\mathbf{B}_k^H\mathbf{G}_k^H\mathbf{A}_L^H=\mathbf{W}_k\left(\mathbf{A}_L\mathbf{A}^{UL}\mathbf{A}_L^H+\mathbf{C}_L\right) \tag{8.30}$$

where we have used the identities from [242] that $\frac{\partial\mathrm{tr}\left(\mathbf{A}\mathbf{Z}^H\right)}{\partial\Re\mathbf{Z}}=\mathbf{A}$, $\frac{\partial\mathrm{tr}(\mathbf{B}\mathbf{Z})}{\partial\Re\mathbf{Z}}=\mathbf{B}^T$, $i\frac{\partial\mathrm{tr}\left(\mathbf{A}\mathbf{Z}^H\right)}{\partial\Im\mathbf{Z}}=\mathbf{A}$, $i\frac{\partial\mathrm{tr}(\mathbf{B}\mathbf{Z})}{\partial\Im\mathbf{Z}}=-\mathbf{B}^T$ and $\frac{df(z)}{dz^*}=\frac{1}{2}\left[\frac{\partial f(z)}{\partial\Re z}+i\frac{\partial f(z)}{\partial\Im z}\right]$. Here $i=\sqrt{-1}$. The other KKT conditions associated with the problem (8.10)-(8.12) are given below

$$\lambda_1\left(\sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{B}_j\mathbf{B}_j^H\right)-P_1^{UL}\right)=0 \tag{8.31}$$
$$\lambda_l\left(\mathrm{tr}\left(\mathbf{F}_l\left(\mathbf{A}_{l-1}\mathbf{A}^{UL}\mathbf{A}_{l-1}^H+\mathbf{C}_{l-1}\right)\mathbf{F}_l^H\right)-P_l^{UL}\right)=0,\quad l=2,\ldots,L \tag{8.32}$$
$$\lambda_1\geq 0,\quad\sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{B}_j\mathbf{B}_j^H\right)\leq P_1^{UL} \tag{8.33}$$
$$\lambda_l\geq 0,\quad\mathrm{tr}\left(\mathbf{F}_l\left(\mathbf{A}_{l-1}\mathbf{A}^{UL}\mathbf{A}_{l-1}^H+\mathbf{C}_{l-1}\right)\mathbf{F}_l^H\right)\leq P_l^{UL},\quad l=2,\ldots,L. \tag{8.34}$$

Lemma 1. For any solutions satisfying the KKT conditions (8.28)-(8.34), the Lagrange

multipliers are

$$\lambda_L = \frac{\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k^H \mathbf{W}_k\right)}{P_L^{UL}} \tag{8.35}$$

$$\lambda_{L-1} = \frac{\mathrm{tr}\left(\mathbf{F}_L^H \left(\mathbf{H}_L^H \sum_{j=1}^{K} \mathbf{W}_j^H \mathbf{W}_j \mathbf{H}_L + \lambda_L \mathbf{I}_{N_{L-1}}\right) \mathbf{F}_L\right)}{P_{L-1}^{UL}} \tag{8.36}$$

$$\begin{aligned} \lambda_l = \frac{1}{P_l^{UL}} \mathrm{tr}\Bigg(& \mathbf{F}_{l+1}^H \Bigg(\bigotimes_{m=l+1}^{L-1} \mathbf{H}_m^H \mathbf{F}_{m+1}^H \mathbf{H}_L^H \sum_{j=1}^{K} \mathbf{W}_j^H \mathbf{W}_j \mathbf{H}_L \bigotimes_{m=L-1}^{l+1} \mathbf{F}_{m+1} \mathbf{H}_m \\ & + \sum_{k=l+2}^{L} \lambda_k \bigotimes_{m=l+1}^{k-1} \mathbf{H}_m^H \mathbf{F}_{m+1}^H \bigotimes_{m=k-1}^{l+1} \mathbf{F}_{m+1} \mathbf{H}_m + \lambda_{l+1} \mathbf{I}_{N_l} \Bigg) \mathbf{F}_{l+1} \Bigg), \; l = 1, \ldots, L-2 \end{aligned} \tag{8.37}$$

Proof. See Appendix 8.A. □

8.3.2 The KKT Conditions of the Downlink Problem

The Lagrangian function of the problem (8.24)-(8.26) is written as

$$\begin{aligned} \mathcal{L}^{DL} = & \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{E}_j^{DL}\right) + \alpha_1 \left(\sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{T}_j \mathbf{T}_j^H\right) - P_1^{DL} \right) \\ & + \sum_{l=2}^{L} \alpha_l \left(\mathrm{tr}\left(\mathbf{Z}_l \left(\mathbf{K}_{l-1} \mathbf{A}^{DL} \mathbf{K}_{l-1}^H + \mathbf{C}_{l-1}^{DL} \right) \mathbf{Z}_l^H \right) - P_l^{DL} \right) \end{aligned} \tag{8.38}$$

where α_1 and $\alpha_l,\ l = 2, \ldots, L$, are the Lagrange multipliers associated with the total power constraint at the users and power constraint at the $(l-1)$-th relay node, respectively. The

KKT conditions associated with the problem (8.24)-(8.26) are given by

$$\mathbf{H}_L\mathbf{K}_{L-1}^H\mathbf{Z}_L^H\mathbf{G}_k\mathbf{D}_k^H = \left(\alpha_1\mathbf{I}_{M_k} + \sum_{l=2}^{L}\alpha_l\mathbf{H}_L\mathbf{K}_{l-1}^H\mathbf{Z}_l^H\mathbf{Z}_l\mathbf{K}_{l-1}\mathbf{H}_L^H + \sum_{j=1}^{K}\mathbf{H}_L\mathbf{K}_{L-1}^H\mathbf{Z}_L^H\mathbf{G}_j\mathbf{D}_j^H\mathbf{D}_j\mathbf{G}_j^H\mathbf{Z}_L\mathbf{K}_{L-1}\mathbf{H}_L^H\right)\mathbf{T}_k \tag{8.39}$$

$$\begin{aligned}\sum_{j=1}^{K}\mathbf{X}_k^{(L)}\mathbf{G}_j\mathbf{D}_j^H\mathbf{T}_j^H\mathbf{H}_L\mathbf{Y}_k^{(1)} &= \sum_{j=1}^{K}\mathbf{X}_k^{(L)}\mathbf{G}_j\mathbf{D}_j^H\mathbf{D}_j\mathbf{G}_j^H\mathbf{Z}_L\mathbf{K}_{L-1}\mathbf{A}^{DL}\mathbf{Y}_k^{(1)}\\ &+ \sum_{j=1}^{K}\mathbf{X}_k^{(L)}\mathbf{G}_j\mathbf{D}_j\mathbf{D}_j^H\mathbf{G}_j^H\mathbf{Z}_L\left(\sum_{c=1}^{k-1}\prod_{m=2}^{L-c}(\mathbf{H}_m^H\mathbf{Z}_{L-m+1})\mathbf{Y}_k^{(c)}\right)+\alpha_k\mathbf{Z}_k\left(\mathbf{K}_{k-1}\mathbf{A}^{DL}\mathbf{Y}_k^{(1)}+\mathbf{C}_{k-1}^{DL}\right)\\ &+ \sum_{l=k+1}^{L}\alpha_l\mathbf{X}_k^{(l)}\mathbf{Z}_l\left(\mathbf{K}_{l-1}\mathbf{A}^{DL}\mathbf{Y}_k^{(1)}+\sum_{n=1}^{k-1}\bigotimes_{m=L-l+2}^{L-n}\mathbf{H}_m^H\mathbf{Z}_{L-m+1}\mathbf{Y}_k^{(n)}\right)\end{aligned} \tag{8.40}$$

$$\mathbf{T}_k^H\mathbf{H}_L\mathbf{K}_{L-1}^H\mathbf{Z}_L^H\mathbf{G}_k = \mathbf{D}_k\left(\mathbf{G}_k^H\mathbf{Z}_L\mathbf{K}_{L-1}\mathbf{A}^{DL}\mathbf{K}_{L-1}^H\mathbf{Z}_L^H\mathbf{G}_k + \mathbf{C}_L^{(k)}\right) \tag{8.41}$$

$$\alpha_1\left(\sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{T}_j\mathbf{T}_j^H\right) - P_1^{DL}\right) = 0 \tag{8.42}$$

$$\alpha_l\left(\mathrm{tr}\left(\mathbf{Z}_l\left(\mathbf{Z}_{l-1}\mathbf{A}^{DL}\mathbf{Z}_{l-1}^H + \mathbf{C}_{l-1}^{DL}\right)\mathbf{Z}_l^H\right) - P_l^{DL}\right) = 0, \quad l = 2,\ldots,L \tag{8.43}$$

$$\alpha_1 \geq 0, \quad \sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{T}_j\mathbf{T}_j^H\right) \leq P_1^{DL} \tag{8.44}$$

$$\alpha_l \geq 0, \quad \mathrm{tr}\left(\mathbf{Z}_l\left(\mathbf{Z}_{l-1}\mathbf{A}^{DL}\mathbf{Z}_{l-1}^H + \mathbf{C}_{l-1}^{DL}\right)\mathbf{Z}_l^H\right) \leq P_l^{DL}, \quad l = 2,\ldots,L. \tag{8.45}$$

In (8.40), $\mathbf{X}_k^{(c)}$ and $\mathbf{Y}_k^{(c)}$ are defined as

$$\mathbf{X}_k^{(c)} = \begin{cases}\bigotimes_{m=L-k+1}^{L-c+2}\left(\mathbf{H}_m\mathbf{Z}_{L-m+2}^H\right), & \text{otherwise}\\ \mathbf{I}_{N_1}, & k = c\end{cases} \tag{8.46}$$

and

$$\mathbf{Y}_k^{(c)} = \begin{cases}\bigotimes_{m=L-c}^{L-k+2}\left(\mathbf{Z}_{L-m+1}^H\mathbf{H}_m\right), & \text{otherwise}\\ \mathbf{I}_{N_{L-k+1}}, & k = c+1\end{cases}. \tag{8.47}$$

Lemma 2. For any solutions satisfying the KKT conditions (8.39)-(8.45), the Lagrange multipliers are

$$\alpha_L = \frac{\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{D}_k \mathbf{D}_k^H\right)}{P_L^{DL}} \tag{8.48}$$

$$\alpha_{L-1} = \frac{\mathrm{tr}\left(\mathbf{Z}_L^H \left(\sum_{i=1}^{K} \mathbf{G}_i \mathbf{D}_i^H \mathbf{D}_i \mathbf{G}_i^H + \alpha_L \mathbf{I}_{M_i}\right) \mathbf{Z}_L\right)}{P_{L-1}^{DL}} \tag{8.49}$$

$$\begin{aligned} \alpha_{L-l+1} = \frac{1}{P_{L-l+1}^{DL}} \mathrm{tr} \Bigg(& \mathbf{Z}_{L-l+2}^H \Bigg(\bigotimes_{i=l-1}^{2} \mathbf{H}_i \mathbf{Z}_{L-i+2}^H \sum_{i=1}^{K} \mathbf{G}_i \mathbf{D}_i^H \mathbf{D}_i \mathbf{G}_i^H \bigotimes_{i=2}^{l-1} \mathbf{Z}_{L-i+2} \mathbf{H}_i^H \\ & + \sum_{j=2}^{l-1} \alpha_{L-j+2} \left[\bigotimes_{i=l-1}^{j} \mathbf{H}_i \mathbf{Z}_{L-i+2}^H \bigotimes_{i=j}^{l-1} \mathbf{Z}_{L-i+2} \mathbf{H}_i^H \right] \\ & + \alpha_{L-l+2} \mathbf{I}_{N_{l-1}} \Bigg) \mathbf{Z}_{L-l+2} \Bigg), \quad l = 3, \ldots, L. \end{aligned} \tag{8.50}$$

Proof. Similar to the proof of Lemma 1, Lemma 2 can also be proved easily, which we will not repeat. □

8.3.3 Sum-MSE Uplink-Downlink Duality

Theorem 1. Assume $\{\mathbf{F}_l\}_{l=2}^{L}, \{\mathbf{B}_j\}_{j=1}^{K}, \{\mathbf{W}_j\}_{j=1}^{K}$ denote a set of uplink transceiver that satisfies the uplink KKT conditions (8.28)-(8.34). Let $\mathbf{T}_j = \sqrt{1/\lambda_L}\mathbf{W}_j^H$, $\mathbf{D}_j = \sqrt{\lambda_1}\mathbf{B}_j^H$, $\mathbf{Z}_l = \sqrt{\lambda_{L-l+2}/\lambda_{L-l+1}}\mathbf{F}_{L-l+2}^H$, $l = 2, \ldots, L$. Then, when the power constraint of the lth node of the downlink channel is swapped with the power constraint of the $(L-l+1)$-th node of the uplink channel, i.e., $P_l^{DL} = P_{L-l+1}^{UL}$, $l = 1, \ldots, L$, sum-MSE achieved by $\{\mathbf{F}_l\}_{l=2}^{L}, \{\mathbf{B}_j\}_{j=1}^{K}, \{\mathbf{W}_j\}_{j=1}^{K}$ can also be achieved by downlink transceiver $\{\mathbf{Z}_l\}_{l=2}^{L}, \{\mathbf{T}_j\}_{j=1}^{K}$, $\{\mathbf{D}_j\}_{j=1}^{K}$, which satisfies the downlink KKT conditions (8.39)-(8.45). Conversely, assume $\{\mathbf{Z}_l\}_{l=2}^{L}, \{\mathbf{T}_j\}_{j=1}^{K}, \{\mathbf{D}_j\}_{j=1}^{K}$ denote a set of downlink transceiver that satisfies the KKT conditions (8.39)-(8.45). Let $\mathbf{B}_j = \sqrt{1/\alpha_L}\mathbf{D}_j^H$, $\mathbf{W}_j = \sqrt{\alpha_1}\mathbf{T}_j^H$ and $\mathbf{F}_{L-l+2} = \sqrt{\alpha_l/\alpha_{l-1}}\mathbf{Z}_l^H$, $l =$

$2, \ldots, L$. Then, when the power constraint of the lth node of the uplink channel is swapped with the power constraint of the $(L-l+1)$-th node of the downlink channel, i.e., $P_l^{UL} = P_{L-l+1}^{DL},\ l = 1, \ldots, L$, the sum-MSE achieved by $\{\mathbf{Z}_l\}_{l=2}^L, \{\mathbf{T}_j\}_{j=1}^K, \{\mathbf{D}_j\}_{j=1}^K$ can also be achieved by the uplink transceiver $\{\mathbf{F}_l\}_{l=2}^L, \{\mathbf{B}_j\}_{j=1}^K, \{\mathbf{W}_j\}_{j=1}^K$, which satisfies the uplink KKT conditions (8.28)-(8.34).

Proof. See Appendix 8.B. □

Theorem 1 shows that if a transceiver design satisfying the KKT conditions associated with uplink optimization problem achieves a certain sum-MSE, it can also be achieved by a transceiver design satisfying the KKT conditions associated with downlink optimization problem, and vice versa. Therefore, the downlink transceiver optimization problems can be solved through an equivalent uplink problem, and vice versa. Since the optimization problems (8.10)-(8.12) and (8.24)-(8.26) are non-convex, the KKT conditions are necessary for local (global) minimums. By Theorem 1, the local minimum of the uplink transceiver optimization problem corresponds to the same local minimum in the downlink transceiver optimization problem, so it is expected that the global minimum of the uplink and downlink transceiver optimization problems must be the same. Furthermore, according to the proof of Theorem 1, at each local optimum when the minimum sum-MSE is achieved, each user's individual MSEs in both links are also identical.

8.4 Numerical Examples

In this section, we justify the sum-MSE duality theorem for multi-hop MIMO relay systems through numerical simulations. We simulate a flat Rayleigh fading environment where all channel matrices have entries with zero mean. The variance of entries in $\mathbf{G}_i$ is $1/M_i, i = 1, \cdots, K$, and the variance of entries in $\mathbf{H}_l$ is $1/N_l, l = 2, \cdots, L$. All noises are complex circularly symmetric with zero mean and unit variance.

All simulation results are averaged over 1000 independent channel realizations. We use the iterative algorithm in [243] to design optimal uplink transceivers $\{\mathbf{F}_l\}_{l=2}^{L}, \{\mathbf{B}_j\}_{j=1}^{K}$, $\{\mathbf{W}_j\}_{j=1}^{K}$ and use the proposed duality theorem to obtain the dual downlink transceivers $\{\mathbf{Z}_l\}_{l=2}^{L}, \{\mathbf{T}_j\}_{j=1}^{K}, \{\mathbf{D}_j\}_{j=1}^{K}$. For all examples, we set $P_L^{UL} = P_1^{DL} = 20$dB and assume that $P_l^{DL} = P_{L-l+1}^{UL} = P, l = 2, \cdots, L$ for simplicity.

We simulate five-hop multiuser MIMO relay systems in our examples. Since there are many parameters on the system setup for multi-hop relays, for simplicity, we consider relay systems where all users have the same number of antennas (i.e., $M_i = M, i = 1, \cdots, K$) and all relay nodes and the destination node in the uplink have the same number of antennas (i.e., $N_l = N, l = 1, \cdots, L$). The extension to systems where different nodes have different number of antennas is straightforward. Fig. 8.3 shows the MSE performance of the uplink and downlink systems versus P with $K = 3$, $M = 2$, and $N = 10$. The BER performance of both systems with QPSK constellations are illustrated in Fig. 8.4 versus P. It can be clearly seen from Figs. 8.3 and 8.4 that the curves overlap, indicating that both the uplink and downlink systems achieve the same sum-MSE and BER. Similar results can also be obtained for other transceiver design approaches.

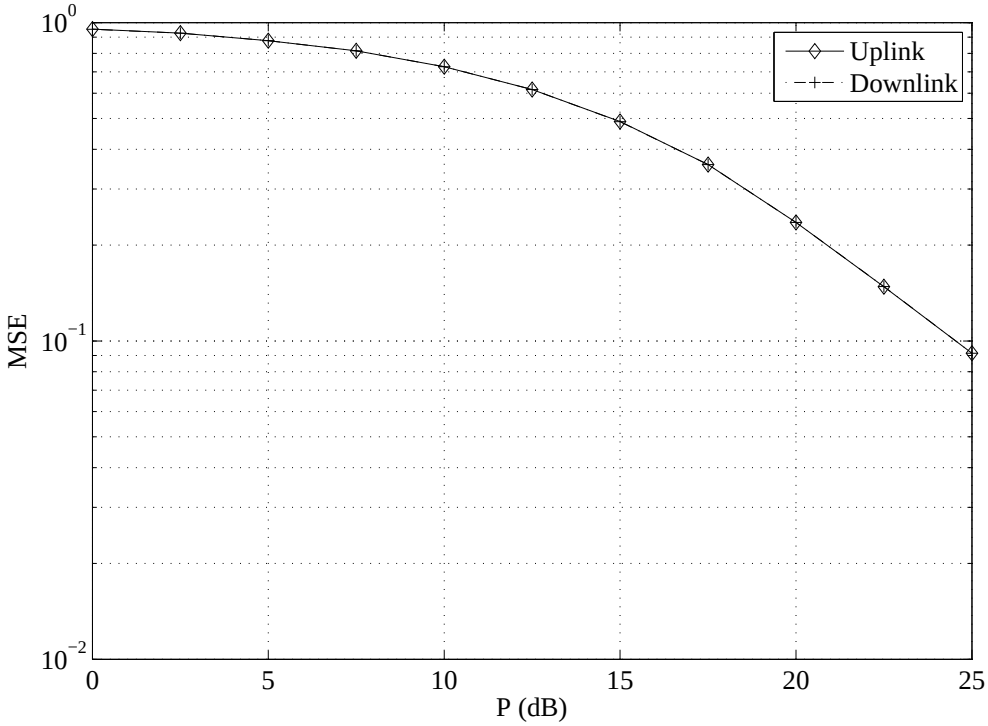

Figure 8.3: Example 1: MSE versus P. $K = 3$, $M = 2$, $N = 10$, $P_L^{UL} = P_1^{DL} = 20$dB.

8.5 Appendix 8.A: Proof of Lemma 1

Proof. Multiplying both sides of (8.29) by $\mathbf{F}_k^H$ and taking the trace, we get

$$\begin{aligned}\sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{A}_L^H\mathbf{W}_j^H\mathbf{B}_j^H\mathbf{G}_j^H\right) &= \sum_{j=1}^{K}\mathrm{tr}\left(\mathbf{W}_j\mathbf{A}_L\mathbf{A}^{UL}\mathbf{A}_L^H\mathbf{W}_j^H\right)\\ &+\sum_{j=1}^{K}\sum_{m=2}^{k}\mathrm{tr}\left(\bigotimes_{l=m}^{L}\left(\mathbf{F}_l^H\mathbf{H}_l^H\right)\mathbf{W}_j^H\mathbf{W}_j\bigotimes_{l=L}^{m}\left(\mathbf{H}_l\mathbf{F}_l\right)\right)\\ &+\lambda_k\mathrm{tr}\left(\mathbf{F}_k\left(\mathbf{A}_{k-1}\mathbf{A}^{UL}\mathbf{A}_{k-1}^H+\mathbf{C}_{k-1}\right)\mathbf{F}_k^H\right)\\ &+\sum_{l=k+1}^{L}\lambda_l\mathrm{tr}\left(\bigotimes_{m=2}^{l-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\mathbf{F}_l^H\mathbf{F}_l\mathbf{A}_{l-1}\mathbf{A}^{UL}\right)\\ &+\sum_{l=k+1}^{L}\sum_{j=2}^{k}\lambda_l\mathrm{tr}\left(\bigotimes_{m=j}^{l-1}\left(\mathbf{F}_m^H\mathbf{H}_m^H\right)\mathbf{F}_l^H\mathbf{F}_l\bigotimes_{i=l-1}^{j}\left(\mathbf{H}_i\mathbf{F}_i\right)\right)\end{aligned}\quad(8.51)$$

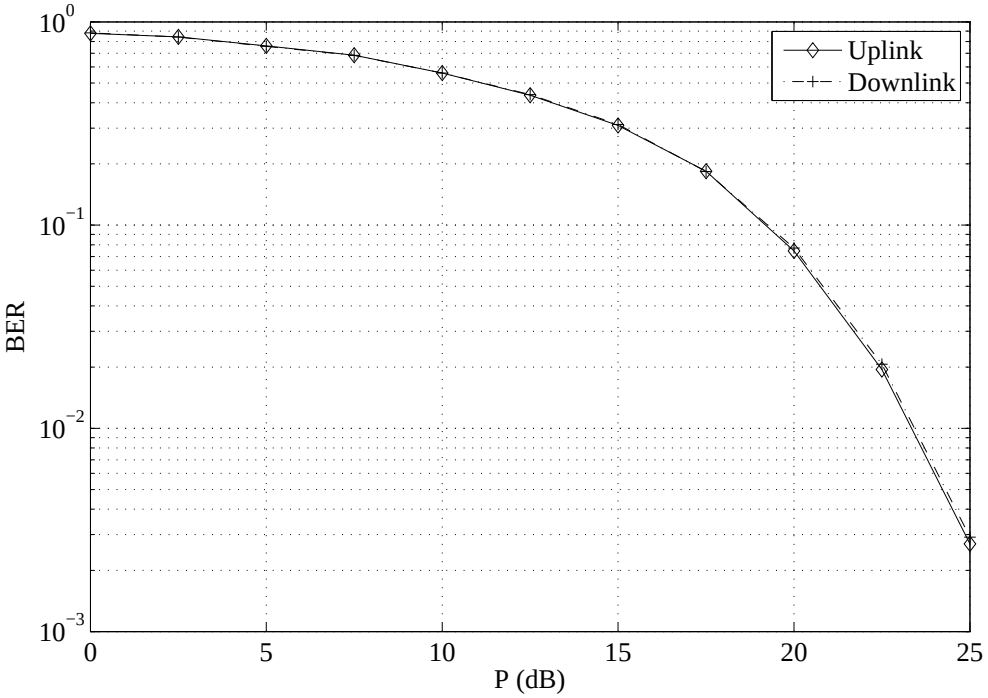

Figure 8.4: Example 2: BER versus P. $K = 3$, $M = 2$, $N = 10$, $P_L^{UL} = P_1^{DL} = 20$dB.

After some straightforward steps and using the KKT conditions in (8.33) and (8.34), (8.51) can be written as

$$\begin{aligned}
&\sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{B}_j^H \mathbf{G}_j^H \mathbf{A}_L^H \mathbf{W}_j^H\right) = \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{A}_L \mathbf{A}^{UL} \mathbf{A}_L^H \mathbf{W}_j^H\right) + \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{C}_L \mathbf{W}_j^H\right) \\
&- \sum_{j=1}^{K} \sum_{m=k+1}^{L} \mathrm{tr}\left(\mathbf{W}_j \bigotimes_{l=L}^{m} \left(\mathbf{H}_l \mathbf{F}_l\right) \bigotimes_{l=m}^{L} \left(\mathbf{F}_l^H \mathbf{H}_l^H\right) \mathbf{W}_j^H\right) - \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{W}_j^H\right) \\
&+ \sum_{l=k}^{L} \lambda_l P_l^{UL} - \sum_{l=k+1}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \left(\sum_{j=k+1}^{l-1} \bigotimes_{i=l-1}^{j} \left(\mathbf{H}_i \mathbf{F}_i\right) \bigotimes_{i=j}^{l-1} \left(\mathbf{F}_i^H \mathbf{H}_i^H\right)\right) \mathbf{F}_l^H\right) \\
&- \sum_{l=k+1}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \mathbf{F}_l^H\right). \qquad (8.52)
\end{aligned}$$

By multiplying both sides of (8.30) with $\mathbf{W}_k^H$, taking the trace operation, and

summing over all k, we have

$$\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{B}_k^H \mathbf{G}_k^H \mathbf{A}_L^H \mathbf{W}_k^H\right) = \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{A}_L \mathbf{A}^{UL} \mathbf{A}_L^H \mathbf{W}_k^H\right) + \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{C}_L \mathbf{W}_k^H\right) \quad (8.53)$$

By substituting (8.53) back into (8.52), we obtain

$$\begin{aligned}\sum_{l=k}^{L} \lambda_l P_l^{UL} &= \sum_{j=1}^{K} \sum_{m=k+1}^{L} \mathrm{tr}\left(\mathbf{W}_j \bigotimes_{l=L}^{m} \left(\mathbf{H}_l \mathbf{F}_l\right) \bigotimes_{l=m}^{L} \left(\mathbf{F}_l^H \mathbf{H}_l^H\right) \mathbf{W}_j^H\right) + \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{W}_j^H\right) \\ &\quad + \sum_{l=k+1}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \left(\sum_{j=k+1}^{l-1} \bigotimes_{i=l-1}^{j} \left(\mathbf{H}_i \mathbf{F}_i\right) \bigotimes_{i=j}^{l-1} \left(\mathbf{F}_i^H \mathbf{H}_i^H\right)\right) \mathbf{F}_l^H\right) \\ &\quad + \sum_{l=k+1}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \mathbf{F}_l^H\right). \end{aligned} \quad (8.54)$$

For each $\mathbf{F}_k,\ k = 2, \ldots, L$, we can find the corresponding $\lambda_k,\ k = 2, \ldots, L$ using (8.54). Starting from $k = L$, (8.54) comes down to

$$\lambda_L P_L^{UL} = \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{W}_j^H\right) \quad (8.55)$$

which is equal to (8.35). For $k = L - 1$, (8.54) comes down to

$$\begin{aligned}\lambda_{L-1} P_{L-1}^{UL} + \lambda_L P_L^{UL} &= \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{H}_L \mathbf{F}_L \mathbf{F}_L^H \mathbf{H}_L^H \mathbf{W}_j^H\right) + \lambda_L \mathrm{tr}\left(\mathbf{F}_L \mathbf{F}_L^H\right) \\ &\quad + \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{W}_j^H\right). \end{aligned} \quad (8.56)$$

After substituting (8.55) into (8.56), we obtain

$$\lambda_{L-1} P_{L-1}^{UL} = \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{H}_L \mathbf{F}_L \mathbf{F}_L^H \mathbf{H}_L^H \mathbf{W}_j^H\right) + \lambda_L \mathrm{tr}\left(\mathbf{F}_L \mathbf{F}_L^H\right) \quad (8.57)$$

which is equal to (8.36). For $k = L-2$, (8.54) comes down to

$$\begin{aligned}
&\lambda_{L-2}P_{L-2}^{UL} + \lambda_{L-1}P_{L-1}^{UL} + \lambda_L P_L^{UL} \\
&= \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \bigotimes_{l=L}^{L-1} \mathbf{H}_l \mathbf{F}_l \bigotimes_{l=L-1}^{L} \mathbf{F}_l^H \mathbf{H}_l^H \mathbf{W}_j^H \right) + \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{W}_j^H \right) \\
&+ \sum_{l=L-1}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \left(\sum_{j=L-1}^{l-1} \bigotimes_{i=l-1}^{j} \mathbf{H}_i \mathbf{F}_i \bigotimes_{i=j}^{l-1} \mathbf{F}_i^H \mathbf{H}_i^H \right) \mathbf{F}_l^H \right) + \sum_{l=L-1}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \mathbf{F}_l^H \right) \quad (8.58)
\end{aligned}$$

After substituting (8.55) and (8.57) into (8.58), we have

$$\begin{aligned}
\lambda_{L-2}P_{L-2}^{UL} &= \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \bigotimes_{l=L}^{L-1} \mathbf{H}_l \mathbf{F}_l \bigotimes_{l=L-1}^{L} \mathbf{F}_l^H \mathbf{H}_l^H \mathbf{W}_j^H \right) \\
&\quad + \lambda_L \mathrm{tr}\left(\mathbf{F}_L \mathbf{H}_{L-1} \mathbf{F}_{L-1} \mathbf{F}_{L-1}^H \mathbf{H}_{L-1}^H \mathbf{F}_L^H \right) + \lambda_{L-1} \mathrm{tr}\left(\mathbf{F}_{L-1} \mathbf{F}_{L-1}^H \right) \quad (8.59)
\end{aligned}$$

which is equal to (8.37) for $l = L-2$. In a similar fashion, we can also calculate λ_k, for $k = L-3, \ldots, 2$.

By using (8.54), we can calculate $\lambda_k,\ k = 2, \ldots, K$, but not λ_1. To calculate λ_1, we left-multiply both sides of (8.28) with $\mathbf{B}_k^H$, take the trace of both sides and sum over k. We also right-multiply both sides of (8.30) with $\mathbf{W}_k^H$, take the trace of both sides and sum over k to make both equations equal. Then we obtain

$$\begin{aligned}
&\sum_{k=1}^{K} \lambda_1 \mathrm{tr}\left(\mathbf{B}_k \mathbf{B}_k^H \right) + \sum_{k=1}^{K} \sum_{l=2}^{L} \lambda_l \mathrm{tr}\left(\mathbf{B}_k^H \mathbf{G}_k^H \mathbf{A}_{l-1}^H \mathbf{F}_l^H \mathbf{F}_l \mathbf{A}_{l-1} \mathbf{G}_k \mathbf{B}_k \right) \\
&+ \sum_{k=1}^{K} \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{B}_k^H \mathbf{G}_k^H \mathbf{A}_L^H \mathbf{W}_j^H \mathbf{W}_j \mathbf{A}_L \mathbf{G}_k \mathbf{B}_k \right) \\
&= \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{A}_L \mathbf{A}^{UL} \mathbf{A}_L^H \mathbf{W}_k^H \right) + \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{C}_L \mathbf{W}_k^H \right). \quad (8.60)
\end{aligned}$$

By using $\mathbf{A}^{UL}$ defined in (8.8), (8.60) can be written as

$$\begin{aligned}
&\lambda_1 \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{B}_k \mathbf{B}_k^H\right) + \sum_{l=2}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \mathbf{A}_{l-1} \mathbf{A}^{UL} \mathbf{A}_{l-1}^H \mathbf{F}_l^H\right) \\
&+ \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{A}_L \mathbf{A}^{UL} \mathbf{A}_L^H \mathbf{W}_j^H\right) \\
&= \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{A}_L \mathbf{A}^{UL} \mathbf{A}_L^H \mathbf{W}_k^H\right) + \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{C}_L \mathbf{W}_k^H\right) \qquad (8.61)
\end{aligned}$$

After canceling the common terms in (8.61) and using the KKT conditions in (8.33) and (8.34), (8.61) can be written as

$$\lambda_1 P_1^{UL} + \sum_{l=2}^{L} \lambda_l \left(P_l^{UL} - \mathrm{tr}\left(\mathbf{F}_l \mathbf{C}_{l-1} \mathbf{F}_l^H\right)\right) = \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{C}_L \mathbf{W}_k^H\right) \qquad (8.62)$$

which is equivalent to

$$\lambda_1 P_1^{UL} + \sum_{l=2}^{L} \lambda_l P_l^{UL} = \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_k \mathbf{C}_L \mathbf{W}_k^H\right) + \sum_{l=2}^{L} \lambda_l \mathrm{tr}\left(\mathbf{F}_l \mathbf{C}_{l-1} \mathbf{F}_l^H\right). \qquad (8.63)$$

After substituting $\sum_{l=k}^{L} \lambda_l P_l^{UL}$ in (8.54) for $k = 2$ into (8.63), we can find λ_1 after tedious but straightforward steps, which is equal to (8.37) for $l = 1$. □

8.6 Appendix 8.B: Proof of Theorem 1

Proof. The theorem includes two parts. We first consider the forward part, i.e., the uplink to downlink transformation. Assume that we are given $\{\mathbf{F}_l\}_{l=2}^{L}, \{\mathbf{B}_j\}_{j=1}^{K}, \{\mathbf{W}_j\}_{j=1}^{K}$, a set of the uplink transceiver that satisfies the KKT conditions (8.28)-(8.34). Then, sum-MSE for

the uplink channel can be written as

$$\begin{aligned}
&\sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{E}_j^{UL}\right) \qquad (8.64)\\
&\overset{(a)}{=} \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{I}_{M_j}\right) - \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{A}_L \mathbf{G}_j \mathbf{B}_j\right)\\
&\overset{(b)}{=} \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{I}_{M_j}\right) - \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{A}_L \mathbf{G}_j \left(\lambda_1 \mathbf{I}_{M_j} + \sum_{l=2}^{L} \lambda_l \mathbf{D}_l^{UL} + \mathbf{E}^{UL}\right)^{-1} \mathbf{G}_j^H \mathbf{A}_L^H \mathbf{W}_j^H\right)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{D}_l^{UL} &= \mathbf{G}_j^H \mathbf{A}_{l-1}^H \mathbf{F}_l^H \mathbf{F}_l \mathbf{A}_{l-1} \mathbf{G}_j \qquad (8.65)\\
\mathbf{E}^{UL} &= \sum_{k=1}^{K} \mathbf{G}_j^H \mathbf{A}_L^H \mathbf{W}_k^H \mathbf{W}_k \mathbf{A}_L \mathbf{G}_j. \qquad (8.66)
\end{aligned}$$

In (8.64), (a) comes from multiplying (8.30) with $\mathbf{W}_j^H$ on the right and substituting the equation back into (8.8), and (b) comes from (8.28).

Sum-MSE of the downlink channel can be written as

$$\begin{aligned}
&\sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{E}_j^{DL}\right)\\
&\overset{(c)}{=} \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{I}_{M_j}\right) - \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{T}_j^H \mathbf{H}_L \mathbf{K}_{L-1}^H \mathbf{Z}_L^H \mathbf{G}_j \mathbf{D}_j^H\right)\\
&\overset{(d)}{=} \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{I}_{M_j}\right) - \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{T}_j^H \mathbf{H}_L \mathbf{K}_{L-1}^H \mathbf{Z}_L^H \mathbf{G}_j \left(\mathbf{G}_j^H \mathbf{Z}_L \mathbf{K}_{L-1} \mathbf{A}^{DL} \mathbf{K}_{L-1}^H \mathbf{Z}_L^H \mathbf{G}_j + \mathbf{C}_L^{(j)}\right)^{-1}\right.\\
&\qquad\qquad \left. \times\, \mathbf{G}_j^H \mathbf{Z}_L \mathbf{K}_{L-1} \mathbf{H}_L^H \mathbf{T}_j\right) \qquad (8.67)
\end{aligned}$$

where (c) is obtained by multiplying the conjugate transpose of (8.41) on the left with $\mathbf{D}_j$ and substituting the equation back into (8.22), and (d) comes from (8.41).

In the downlink channel, let

$$\mathbf{T}_j = \eta_1 \mathbf{W}_j^H,\ \mathbf{D}_j = \frac{1}{\eta_L} \mathbf{B}_j^H,\ \mathbf{Z}_l = \frac{\eta_l}{\eta_{l-1}} \left(\mathbf{F}_{L-l+2}\right)^H,\ l = 2, \ldots, L. \qquad (8.68)$$

where $\eta_l,\ l = 1, \ldots, L \in \Re^+$. By substituting (8.68) into (8.67), we have $\mathbf{H}_L \mathbf{K}_{L-1}^H \mathbf{F}_2 = \frac{\eta_{L-1}}{\eta_1} \mathbf{A}_L$ and the MSE of the downlink channel is given by

$$\begin{aligned}
&\sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{E}_j^{DL}\right) \\
&= \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{I}_{M_j}\right) - \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j \mathbf{A}_L \mathbf{G}_j \left(\sum_{k=1}^{K} \mathbf{G}_j^H \mathbf{A}_L^H \mathbf{W}_k^H \mathbf{W}_k \mathbf{A}_L \mathbf{G}_j \right.\right. \\
&+ \left.\left. \sum_{k=1}^{L-2} \frac{1}{\eta_k^2} \mathbf{G}_j^H \mathbf{A}_{L-k}^H \mathbf{F}_{L-k+1}^H \mathbf{F}_{L-k+1} \mathbf{A}_{L-k} \mathbf{G}_j + \frac{1}{\eta_{L-1}^2} \mathbf{G}_j^H \mathbf{F}_2^H \mathbf{F}_2 \mathbf{G}_j + \frac{1}{\eta_L^2} \mathbf{I}_{M_j} \right)^{-1} \right. \\
&\quad \left. \times \mathbf{G}_j^H \mathbf{A}_L^H \mathbf{W}_j^H \right). \qquad (8.69)
\end{aligned}$$

On the other hand, the total power constraints for the downlink (8.25)-(8.26) can be written as

$$\eta_1^2 \sum_{j=1}^{K} \mathrm{tr}\left(\mathbf{W}_j^H \mathbf{W}_j\right) \le P_1^{DL} \qquad (8.70)$$

$$\mathrm{tr}\left(\eta_2^2 \mathbf{F}_L^H \Big[\mathbf{H}_L^H \sum_{j=1}^{K} \mathbf{W}_j^H \mathbf{W}_j \mathbf{H}_L + \frac{1}{\eta_1^2} \mathbf{I}_{N_{L-1}} \Big] \mathbf{F}_L \right) \le P_2^{DL}, \qquad (8.71)$$

$$\begin{aligned}
&\mathrm{tr}\Bigg(\eta_l^2 \mathbf{F}_{L-l+2}^H \Big[\bigotimes_{m=L-l+2}^{L-1} \mathbf{H}_m^H \mathbf{F}_{m+1}^H \mathbf{H}_L^H \sum_{j=1}^{K} \mathbf{W}_j^H \mathbf{W}_j \mathbf{H}_L \bigotimes_{m=L-1}^{L-l+2} \mathbf{F}_{m+1} \mathbf{H}_m \qquad (8.72) \\
&+ \sum_{k=1}^{l-2} \frac{1}{\eta_k^2} \bigotimes_{m=L-l+2}^{L-k} \mathbf{H}_m^H \mathbf{F}_{m+1}^H \bigotimes_{m=L-k}^{L-l+2} \mathbf{F}_{m+1} \mathbf{H}_m + \frac{1}{\eta_{l-1}^2} \mathbf{I}_{N_{L-l+1}} \Big] \mathbf{F}_{L-l+2} \Bigg) \le P_l^{DL},\ l = 3, \ldots, L
\end{aligned}$$

Comparing (8.70)-(8.72) with (8.35)-(8.37) in Lemma 1, if we choose $\eta_1 = \sqrt{1/\lambda_L}$, $\eta_2 = \sqrt{1/\lambda_{L-1}}, \ldots, \eta_L = \sqrt{1/\lambda_1}$, particularly, $\eta_l = \sqrt{1/\lambda_{L-l+1}},\ l = 1, \ldots, L$, and assign $P_l^{DL} = P_{L-l+1}^{UL},\ l = 1, \ldots, L$, then the sum-MSE of the downlink (8.69) and uplink channel (8.64) will be the same while (8.70)-(8.72) will be satisfied with equality. Moreover, let $\alpha_l = \lambda_{L-l+1}$, the KKT conditions of the downlink channel (8.39)-(8.45) are all satisfied. Therefore, from the set $\{\mathbf{F}_l\}_{l=2}^L, \{\mathbf{B}_j\}_{j=1}^K, \{\mathbf{W}_j\}_{j=1}^K$ satisfying the uplink KKT conditions,

we can find a set $\{\mathbf{Z}_l\}_{l=2}^{L}, \{\mathbf{T}_j\}_{j=1}^{K}, \{\mathbf{D}_j\}_{j=1}^{K}$ which satisfies the downlink KKT conditions and has the identical sum-MSE. This concludes the forward part.

The converse part, i.e., the downlink to uplink transformation, can also be proven using the same arguments. Details are omitted to avoid repetition. □

Chapter 9

Conclusion

9.1 Summary of Contributions

In the first part of this thesis, joint transceiver optimization for MIMO FD systems with linear transmit and receive processing is studied. We consider different performance measures: Ergodic mutual information, weighted sum-rate, sum MSE, and min-max MSE subject to power constraints. In the second part of this thesis, the SINR, MSE, capacity uplink-downlink duality for multi-hop MIMO AF relay systems has been established. The topics covered are:

- In Chapter 3, we have studied the ergodic mutual information maximization of two FD MIMO radio systems (bi-directional system and relay system) that suffer from a (digitally manageable residual) self-interference under a fast fading channel model. The source covariance matrices are treated as a function of time and/or frequency within any given time/frequency band so that both spatial and temporal freedoms of

the source covariance matrices can be exploited. Since the globally optimal solution is difficult to obtain due to the non-convex nature of the problem, a gradient projection algorithm is developed to optimize the power allocation vectors at two respective nodes with the knowledge of statistical CSI useful for the transmitters. In addition to an exact closed-form ergodic mutual information expression, we introduced a much simpler asymptotic closed-form ergodic mutual information expression, which is shown to be an accurate approximation and in turn simplifies the computation of the power allocation vectors. It is shown through numerical simulations that the ergodic mutual information increases with the number of antennas, decreases as the channel estimation error and/or the transmitter distortion increases. Moreover, it is demonstrated that at a high self-interference power level, the optimal power schedule reduces to the HD mode, and at a low self-interference power level, the optimal power schedule switches to the FD mode.

- In Chapter 4, we have addressed the transmit filter design for WSR maximization problem in FD MIMO bi-directional systems that suffer from self-interference under the imperfect CSI knowledge and limited dynamic ranges at the transmitters and receivers. Both sum-power constraint and individual power constraint were considered. Since the globally optimal solution is difficult to obtain due to the non-convex nature of the problems, an alternating iterative algorithm to find a local WSR optimum was proposed based on the relationship between WSR and WMMSE problems. The source covariance matrices are treated as a function of time and/or frequency within any given time/frequency band so that both spatial and temporal freedoms of the

source covariance matrices can be exploited. It is shown through numerical simulations that for the individual power constrained problem, at a high self-interference power level, the optimal power schedule reduces to the HD mode, and at a low self-interference power level, the optimal power schedule switches to the FD mode. For the sum-power constrained problem, the sum-rate of FD scheme is always more than the sum-rate of HD scheme, and using two time slots and one time slot transmission give the same performance.

- Most of the works on FD systems have studied the maximization of the achievable rate and to the best of our knowledge, MSE based transceiver designs have not been studied. MSE is an important performance measure to approach the information-theoretic limits of Gaussian channels, and has been widely considered as an optimization metric in precoding design in the literature. In Chapter 5, the effects of residual self-interference, due to the imperfect CSI and limited DR at the transmitters and receivers, on the sum-MSE and Min-Max MSE transceiver design problems for FD MIMO bi-directional system is studied. Since the transceiver design problems are non-convex, an iterative alternating algorithm is proposed that compute the transmit precoding or receive filtering matrices in an alternating fashion while keeping the other one fixed. It is shown through numerical simulations that MSE at each node increases as the channel estimation error and the power of the transmitter/receiver impairments increases. Moreover, in Min-Max MSE transceiver design the nodes achieve the same MSE, which is fair, and sum-MSE transceiver design achieves the minimum total MSE over two FD nodes.

- In Chapter 7, we have investigated the MAC-BC SINR (MSE, capacity) duality in a multi-hop AF MIMO relay system under imperfect CSI and antenna correlation at each hop, which is a generalization of several previously established results. We proved that identical stream-wise SINR (MSE, capacity) in the MAC and BC systems can be achieved by two approaches. Firstly, under the same total network transmission power constraint, the relay nodes of the BC system employ the Hermitian transposed MAC system relay amplifying matrices. Secondly, under the individual transmission power constraint at each node of the system, the relays of the BC system use the scaled and Hermitian transposed MAC system relay amplifying matrices, where the scaling factors are obtained by swapping the power constraints of the nodes of the MAC system. Moreover, we proved that the two approaches developed above are also valid for relay systems with nonlinear transceivers at the source and destination nodes.

- In Chapter 8, we have established the uplink-downlink sum-MSE duality in a multi-hop AF MIMO relay system, which is a generalization of several previously established sum-MSE duality results. Unlike the previous methods that require the direct calculation of MSEs for each stream, the proof of the duality in this paper is based on the KKT conditions of the uplink and downlink minimum sum-MSE transceiver optimization problems. By analyzing the KKT conditions of the these optimization problems, it is shown that both the uplink and the downlink systems share the same achievable sum-MSE region when the power constraint at the lth node of the downlink system is switched with the power constraint at the $(L+1-l)$-th node of the uplink

system, $l = 1, \ldots, L$, where L is the number of hops of the relay network. Therefore, the downlink transceiver optimization problems satisfying the KKT conditions can be carried out by solving an equivalent uplink problem, and vice versa.

9.2 Future Work

1. This thesis has discussed only a simple model, i.e., two FD nodes exchanging information simultaneously. It might be good approach to extend this system into a cellular network level, where the UL and DL are done at the same time and same frequency through a FD base station. Various optimization problems related to UL and DL links, and scheduling algorithms can be proposed.

2. In addition to FD systems, cognitive radio networks is also a promising technology in future wireless communication systems to enhance spectrum efficiency. We can combine FD systems with cognitive radio networks. Particularly, FD secondary users can exchange information simultaneously with each other, while they should provide protection (limited interference) to primary users. Various optimization problems can be formulated under this scenario.

3. The Quality-of-Service (QoS) issue of FD bi-directional scheme where the total transmitted power is minimized subject to minimum rate constraints of each node can be considered. The penalty method can be applied to develop a very efficient optimization algorithm to minimize the total system power. This algorithm requires a central scheduler. However, a centralized planning would require the exchange of a huge

amount of data among the nodes. This would induce an excessive signaling traffic. It is then of special interest to devise decentralized mechanisms which are able to adapt resource allocation dynamically in order to limit interference adequately.

4. Uplink-downlink duality results have been established under a total power constraint of the system and/or individual power constraints at each node of the system. So, it will be interesting to extend the uplink-downlink duality results of this thesis to generalized multiple linear constraints, that can handle the per-antenna power constraints, interference constraints, etc.

Bibliography

[1] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band Full-duplex Wireless: Challenges and Opportunities [arxiv]".

[2] H. Ju, E. Oh, and D. Hong, "Catching resource-devouring worms in next-generation wireless relay systems: Two-way relay and full-duplex relay," *IEEE Comm. Magazine*, vol. 47, no. 9, pp. 58-65, Sept. 2009.

[3] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond" *IEEE Communications Magazine*, Feb 2014.

[4] Quellan Inc. Qhx220 narrowband noise canceller ic. http://www.quellan.com/products/qhx220 ic.php.

[5] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, H. V. Balan, and G. Dejean, "Rethinking indoor wireless: Low power, low frequency, full-duplex," in *IEEE Wireless Mesh Networks*, pages 1-6, 2010.

[6] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. Asilomar Conf. Signals, Syst. Computers*, pp. 1558-1562, 2010.

[7] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti. "Achieving single channel, full duplex wireless communication," in *Proc. Int. Conf. Mobile Comput. Netw.*, pages 1-12, 2010.

[8] M. Jain, J. I. Choi, T. Kim, D. Bharadia, K. Srinivasan, S. Seth, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless,' in *Proc. Int. Conf. Mobile Comput. Netw*, 2011.

[9] J.-H. Lee, "Self-Interference cancelation using phase rotation in full-duplex wireless," *IEEE Trans. on Vehicular Technology*, vol. 62, no. 9, pp. 4421-4429, Nov. 2013.

[10] Z. Zhan, G. Villemaud, and J-M. Gorce, "Design and evaluation of a wideband Full-Duplex OFDM system based on AASIC," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 68-72, Sept. 2013.

[11] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full duplex wireless communications", in *Proc. of Hotnets*, Nov 2011.

[12] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "MIDU: Enabling MIMO full duplex", in *Proc. of the ACM Mobicom*, 2012.

[13] D. Korpi, S. Venkatasubramanian, T. Riihonen, L. Anttila, S. Tretyakov, M. Valkama, and R. Wichman, "Advanced self-interference cancellation and multiantenna techniques for full-duplex radios," *47th Annual Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, Pacific Grove, California, USA, November 2013.

[14] P. Meerasri, P. Uthansakul, and M. Uthansakul, "Self-Interference cancellation based mutual coupling model for full-duplex single-channel MIMO systems," *International Journal of Antennas and Propagations*, vol. 2013, pp. 1-10, 2013.

[15] J. G. McMichael, and K. E. Kolodziej, "Optimal tuning of analog self-interference cancellers for full-duplex wireless communication," *Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 246-251, Oct. 2012.

[16] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. on Wireless Communications*, vol. 11, no. 12, pp. 4296-4307, Dec. 2012.

[17] A. Balatsoukas-Stimming, P. Belanovic, K. Alexandris and A. Burg, "On self-interference suppression methods for low-complexity full-duplex MIMO", *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 3-6, 2013.

[18] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of full-duplex: Design and real-time implementation", Online at arXiv, 2011.

[19] Y. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 5992-6010, December 2013.

[20] S. H. Li and R. D. Murch, "Full-duplex wireless communication using transmitter output based echo cancellation," in *Proc. 2011 IEEE Global Commun. Conf.*, pp. 1-5.

[21] M. Knox, "Single antenna full duplex communications using a common carrier," in *13th Annual Wireless and Microwave Technology Conference*, Apr. 2012.

[22] E. Everett, M. Duarte, C. Dick. A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," *IEEE Asilomar*, pp. 2002-2006, Nov. 2011.

[23] S. Hong, J. Mehlman, and S. Katti, "Picasso: Flexible RF and spectrum slicing", in *SIGCOMM*, 2012.

[24] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM'13*, Aug. 2013.

[25] J. I. Choi, S. Hong, M. Jain, S. Katti, P. Levis, J. Mehlman, "Beyond full duplex wireless," *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp. 40-44, Nov. 2012.

[26] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. Ramakrishnan, C. Rice, and N. Shankaranarayanan, "Design and characterization of a full-duplex multi-antenna system for WiFi networks," *arXiv preprint arXiv:1210.1639*, 2012.

[27] V. Aggarwal, N. K. Shankaranarayanan, "Performance of a random-access wireless network with a mix of full and half-duplex stations," *IEEE Communications Letters*, vol. 17, no. 11, pp. 2200-2203, November 2013.

[28] D. Bharadia, and S. Katti, "Full-duplex MIMO radios," NSDI 2014.

[29] B. Chen, V. Yenamandra and K. Srinivasan, "FlexRadio: Fully flexible radios," *NSDI*, 2014.

[30] Y. Hua, P. Liang, Y. Ma, A. Cirik and Q. Gao, "A method for broadband full-duplex MIMO radio," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 793-796, Dec 2012.

[31] B. Kaufman, J. Lilleberg, B. Aazhang, "An analog baseband approach for designing full-duplex radios," *47th Annual Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, Pacific Grove, California, USA, November 2013.

[32] M. A. Khojastepour, S. Rangarajan, "Wideband digital cancellation for full-duplex communications," *Forty Sixth Asilomar Conference on Signals, Systems and Computers* pp. 1300-1304, Nov. 2012.

[33] J. Krier, and I. F. Akyildiz, "Active self-interference cancellation of passband signals using gradient descent," in *Proc. of Personal, Indoor and Mobile Radio Communications (PIMRC)*, London, UK, September 2013.

[34] N. Li, W. Zhu, and H. Han, "Digital interference cancellation in single channel, full duplex wireless communication," *Wireless Communications, Networking and Mobile Computing (WiCOM)*, pp. 1-4, Sept. 2012.

[35] K. Junghwan, K. Shamaileh, S. Adusumilli, and V. Rao, "Digital interference cancellation for multimedia transmission in full duplex communication link," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1-5, June 2013.

[36] M. Bernhardt, F. H. Gregorio, J. E. Cousseau "A robust wireless OFDM echo cancellation system," XV Reunin de Trabajo en Procesamiento de la Informacin y Control, RPIC 2013, Universidad Nacional de Ro Negro, San Carlos de Bariloche, Sep. 2013.

[37] E. Ahmed, A. Eltawil, and A. Sabharwal, "Rate gain region and design trade-offs for full-duplex wireless communications," accepted for *IEEE Trans. on Wireless Communications*, 2013.

[38] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancellation in wireless full-duplex," December 2012, submitted to IEEE Transactions on Vehicular Technology. [Online]. Available: arXiv:1212.5462v1

[39] V. Syrjala, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," submitted and under review in *IEEE Transactions on Wireless Communications*, to be published, 2013. [Online].

[40] E. Ahmed, A. Eltawil, and A. Sabharwal, "Self-interference cancellation with phase noise induced ICI suppression for full-duplex systems," 2013. [Online]. Available: http://arxiv.org/abs/1307.4149

[41] E. Ahmed and A. M. Eltawil," On phase noise suppression in full-duplex systems", submitted to *IEEE Trans. on Wireless Communications*, January, 2014.

[42] T. Riihonen, P. Mathecken, and R. Wichman, "Effect of oscillator phase noise and processing delay in full-duplex OFDM repeaters," *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp. 1947-1951, Nov. 2012.

[43] Y. Hua, Y. Ma, P. Liang, and A. C. Cirik, "Breaking the barrier of transmission noise in full-duplex radio," *IEEE Military Communications Conference, MILCOM 2013*, pp. 1558-1563, Nov. 2013.

[44] D. W. Bliss, T. M. Hancock, and P. Schniter, "Hardware phenomenological effects on cochannel full-duplex MIMO relay performance," *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp. 34-39, Nov. 2012.

[45] E. Ahmed, A. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, 2013.

[46] L. Anttila, D. Korpi, V. Syrjala, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*,2013.

[47] D. Korpi, T. Riihonen, V. Syrjala, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," submitted and under review in IEEE Transactions on Wireless Communications, 2013.

[48] T. Riihonen, and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp. 45-49, Nov. 2012.

[49] Z. Zhan, G. Villemaud and J.-M. Gorce, "Analysis and reduction of the impact of thermal noise on the full-duplex OFDM radio," *IEEE Radio and Wireless Symposium*, 2014.

[50] K. Haneda, E. Kahra, S. Wyne, C. Icheln, and P. Vainikainen, "Measurement of loopback interference channels for outdoor-to-indoor full-duplex radio relays," in *Proc. 4th European Conference on Antennas and Propagation (EuCAP)*, pp. 1 - 5, Barcelona, Spain, April 2010.

[51] E. Everett, A. Sahai, A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes", accepted *IEEE Trans. on Wireless Communication*, January 2013.

[52] A. Sahai, G. Patel, and A. Sabharwal, "Asynchronous full-duplex wireless," *Fourth International Conference on Communication Systems and Networks (COMSNETS)*, pp. 1-9, Jan. 2012.

[53] K. M. Nasr, J. P. Cosmas, M. Bard, and J. Gledhill, "Performance of an echo canceller and channel estimator for on-channel repeaters in DVBT/H networks," *IEEE Trans. Broadcast.*, vol. 53, pp. 609-618, Sep. 2007.

[54] H. Hamazumi, K. Imamura, N. Iai, K. Shibuya, and M. Sasaki, "A study of a loop interference canceller for the relay stations in an SFN for digital terrestrial broadcasting," in *Proc. IEEE Global Commun. Conf*, pp. 167-171, 2000.

[55] T. Riihonen, S. Werner, and R. Wichman, "Optimized gain control for single-frequency relaying with loop interference," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2801-2806, June 2009.

[56] T. Riihonen, S. Werner, and R. Wichman, "Transmit power optimization for multi-antenna decode-and-forward relays with loopback self-interference from full-duplex operation," in *45th Annual Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, pp. 1408-1412, Nov. 2011.

[57] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983-5993, Dec. 2011.

[58] T. Riihonen, S. Werner, and R. Wichman, "Residual self-interference in full-duplex MIMO relays after null-space projection and cancellation," in *Proc. 44th Annual Asilomar Conference on Signals, Systems, and Computers*, November 2010.

[59] B. Chun, E. Jeong, J. Joung, Y. Oh, and Y. H. Lee, "Pre-nulling for self-interference suppression in full-duplex relays," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit Conference (APSIPA ASC)*, Sapporo, Oct. 2009.

[60] B. Chun and Y. H. Lee, "A spatial self-interference nullification method for full duplex amplify-and-forward MIMO relays," in *Proc. IEEE Wireless Communications and Networking Conference*, April 2010.

[61] T. Riihonen, A. Balakrishnan, K. Haneda, S. Wyne, S. Werner, and R. Wichman, "Optimal eigenbeamforming for suppressing self interference in full-duplex MIMO relays", *45th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, Maryland, March 2011.

[62] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex MIMO relays," in *Proc. Asilomar Conf. Signals Syst. Comput.*, (Pacific Grove, CA), pp. 658-662, Oct. 2010.

[63] B. Chun and H. Park; , "A spatial-domain joint-nulling method of self-interference in full-duplex relays," *IEEE Communications Letters*, vol. 16, no. 4, pp.436-438, April 2012.

[64] D. Senaratne and C. Tellambura, "Beamforming for space division duplexing," in *Proc. IEEE Int. Conf. Commun.*, pp. 1-5, Japan, Jun. 2011,

[65] E. A. Rodriguez, R. L. Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Adaptive self-interference cancellation in wideband full-duplex decode-and-forward MIMO relays," *IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 370-374, June 2013.

[66] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and J. Lilleberg, "Full-Duplex in large-scale wireless systems", in *47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013.

[67] T. Riihonen, M. Vehkapera, and R. Wichman,"Large-system analysis of rate regions in bidirectional full-duplex MIMO link: Suppression versus cancellation," in *Proc. Conf. Inform. Sciences Syst.*, Mar. 2013.

[68] M. Vehkapera, T. Riihonen, and R. Wichman, "Asymptotic analysis of full-duplex bidirectional MIMO link with transmitter noise," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 1265-1270, Sept. 2013.

[69] D. W. Bliss, P. A. Parker, and A. R. Margetts, "Simultaneous transmission and reception for improved wireless network performance," in *Proc. IEEE 14th Workshop on Statistical Signal Processing*, pp. 26-29, Aug. 2007.

[70] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Trans. Wireless Communications*, vol. 8, pp. 3933-3938, Aug. 2009.

[71] P. Larsson and M. Prytz, "MIMO on-frequency repeater with self-interference cancellation and mitigation," in *Proc. IEEE 69th Vehicular Technology Conference*, April 2009.

[72] D. Choi and D. Park, "Effective self interference cancellation in full-duplex relay systems," *Electron. Lett.*, vol. 48, pp. 129-130, Jan. 2012.

[73] C. Kim, E.-R. Jeong, Y. Sung, and Y. H. Lee, "Asymmetric complex signaling for full-duplex decode-and-forward relay channels," *ICT Convergence (ICTC)*, pp. 28-29, Oct. 2012.

[74] Y. Hua, "An overview of beamforming and power allocation for MIMO relays," in *Proc. IEEE Military Commun. Conf.*, (San Jose, CA), pp. 375-380, Nov. 2010.

[75] R. Lopez-Valcarce, E. Antonio-Rodriguez, T. Riihonen, S. Werner, and R. Wichman, "Autocorrelation-based adaptation rule for feedback for feedback equalization in wide-band full-duplex amplify-and-forward MIMO relays," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4968-4972, May 2013.

[76] R. Lopez-Valcarce, E. Antonio-Rodriguez, C. Mosquera, F. Perez-Gonzalez, "An adaptive feedback canceller for full-duplex relays based on spectrum shaping," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1566-1577, September 2012.

[77] E. A. Rodriguez and R. L. Valcarce, "Adaptive self-interference suppression for full-duplex relays with multiple receive antennas," in *Proc. 13th IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, 2012.

[78] E. A. Rodriguez and R. L. Valcarce, "Cancelling self-interference in full-duplex relays without angle-of-arrival information," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4731-4735, May 2013.

[79] H. Sakai, T. Oka, and K. Hayashi, "Simple adaptive filter method for cancellation of coupling wave in OFDM signals at SFN relay station," in Proc. *EUSIPCO 2006*, Sept. 2006.

[80] K. Hayashi, Y. Fujishima, M. Kaneko, H. Sakai, R. Kudo, and T. Murakami, "Self-interference canceller for full-duplex radio relay station using virtual coupling wave paths," *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, pp. 1-5, Dec. 2012.

[81] K. Hayashi, M. Kaneko, M. Noguchi, and H. Sakai, "A single frequency full-duplex radio relay station for frequency domain equalization systems," *IEEE/CIC International Conference on Communications in China (ICCC), 2013*, pp. 33-38, Aug. 2013.

[82] S. O. Al-Jazzar, T. Al-Naffouri, "Relay self interference minimization using tapped filter," *8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*, pp. 316-319, May 2013.

[83] J. Sangiamwong, T. Asai, J. Hagiwara, Y. Okumura, and T. Ohya, "Joint multi-filter design for full-duplex MU-MIMO relaying," in *Proc. IEEE 69th Veh. Technol. Conf.*, Apr. 2009.

[84] J.-H. Lee and O.-S. Shin, "Precoding and power allocation for full-duplex MIMO relays," *IEICE Trans. Commun.*, vol. E94-B, no. 8, pp. 2316-2327, Aug. 2011.

[85] C.-H. Lee, J.-H. Lee, Y.-W. Kwak, Y.-H. Kim, and S.-C. Kim, "The realization of full duplex relay and sum rate analysis in multiuser MIMO relay channel," in *Proc. IEEE Veh. Technol. Conf.*, Ottawa, ON, Canada, Sep. 2010, pp. 1-5.

[86] J.-H. Lee and O.-S. Shin, "Full-duplex relay based on block diagonalization in multiple-input multiple-output relay systems," *IET Commun.*, vol. 4, no. 15, pp. 1817-1826, Oct. 2010.

[87] J.-H. Lee and O.-S. Shin, "Full-Duplex relay based on distributed beamforming in multiuser MIMO systems," *IEEE Transactions on Vehicular Technology*, vol.62, no. 4, pp. 1855-1860, May 2013.

[88] J.-H. Lee and O.-S. Shin, "Distributed beamforming approach to full-duplex relay in multiuser MIMO transmission," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Paris, France, Apr. 2012, pp. 278-282.

[89] T. Taniguchi, Y. Karasawa, "Design and analysis of MIMO multiuser system using full-duplex multiple relay nodes," *Wireless Days (WD)*, pp. 1-8, Nov. 2012.

[90] J.-H. Lee and O.-S. Shin, "Full-duplex relay based on zero-forcing beamforming," *IEICE Trans. Commun.*, vol. E94-B, pp. 978-985, April 2011.

[91] Y.-W. Kwak, J.-H. Lee, Y.-H. Kim, and S.-C. Kim, "Precoder design and capacity analysis for multi-antenna full-duplex relay," *IEICE Trans. Commun.*, vol. E95-B, no. 7, pp. 2446-2450, July 2012.

[92] C.-H. Lee, J.-H. Lee, O.-S. Shin, and S.-C. Kim, "Sum rate analysis of multiantenna multiuser relay channel," *IET Communications*, vol. 4, no. 17, pp. 2032-2040, 2010.

[93] D. Nguyen, L.-N. Tran, P. Pirinen, M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Transactions on Signal Processing,* vol. 61, no. 16, pp. 4038-4050, Aug.15, 2013.

[94] D. Kim, S. Park, S, H. Ju, and D. Hong, "Transmission capacity of full-duplex based two-way ad-hoc networks with ARQ protocol," *IEEE Trans. on Vehicular Technology*, 2014.

[95] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," *IEEE Information Theory Workshop (ITW)*, pp. 1-5, Sept. 2013.

[96] P. Li, S. Guo, and W. Zhuang, "Optimal transmission scheduling of cooperative communications with a full-duplex relay," *IEEE Trans. on Parallel and Distributed Systems*, Aug. 2013.

[97] W. Cheng, X. Zhang, H. Zhang, "Full-Duplex wireless communications for cognitive radio networks," Available: http://arxiv.org/abs/1105.0034.

[98] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Simultaneous transmit and sense for cognitive radios using full-duplex: A first study," to appear: *2012 IEEE International Symposium on Antennas and Propagation and USNC-URSI National Radio Science Meeting*.

[99] W. Cheng, X. Zhang, and H.'Zhang, "Imperfect full duplex spectrum sensing in cognitive radio networks," in *Proc. 17th ACM MOBICOM, The Third Workshop on Cognitive Wireless Networking(CoRoNet) Workshop*, Las Vegas, Nevada, USA, Sep. 2011.

[100] W. Afifi, M. Krunz, " Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *Proc. of the IEEE INFOCOM 2013 Main-Conference*, 2013.

[101] H. L. Kim, S. Lim, H. Wang, and D. Hong, "Optimal power allocation and outage analysis for cognitive full duplex relay systems," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 10, pp. 3754-3765, Oct. 2012.

[102] G. Zheng, I. Krikidis, and B. Ottersten, "Full-Duplex cooperative cognitive radio with transmit imperfections," *IEEE Trans. on Wireless Communications*, vol. 12, no. 5, pp. 2498-2511, May 2013.

[103] T. Riihonen, S. Werner, R. Wichman, and E. Zacarias B., "On the feasibility of full-duplex relaying in the presence of loop interference," in *Proc. 10th IEEE Workshop on Signal Processing Advances in Wireless Communications*, June 2009.

[104] T. Riihonen, S. Werner, and R. Wichman, "Comparison of full-duplex and half-duplex modes with a fixed amplify-and-forward relay," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1-5, Sep. 2009.

[105] T. Riihonen, S. Werner, and R. Wichman, "Rate-interference trade-off between duplex modes in decode-and-forward relaying," in *Proc. IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 690-695, Istanbul, Turkey, Sep. 2010.

[106] Y. Y. Kang and J. H. Cho, "Capacity of MIMO wireless channel with full-duplex amplify-and-forward relay," in *Proc. IEEE 20th Int. Symp. Pers., Indoor and Mobile Radio Commun.*, Sep. 2009.

[107] Y. Y. Kang, B. J. Kwak, and J. C. Cho, "An optimal self-interference cancellation for full-duplex amplify-and-forward relay," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 13217-1221, Sept. 2013.

[108] A. Lo and P. Guan, "Performance of in-band full-duplex amplify-and-forward and decode-and-forward relays with spatial diversity for next-generation wireless broadband," in *Proc. 2011 Intl. Conf. Inf. Netw*, pp. 290-294, Jan 2011.

[109] B. P. Day, D. W. Bliss, A. R. Margetts, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702-3713, July 2012.

[110] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30 , no. 8, pp. 1541-1553, Sept. 2012.

[111] A. C. Cirik, Y. Rong, and Y. Hua, "Ergodic mutual information of full-duplex MIMO radios with residual self-interference," presented at *2013 Asilomar Conf. Signals, Syst. Computers*, Nov. 2013.

[112] A. C. Cirik, R. Wang, and Y. Hua, "Weighted-sum-rate maximization for bi-directional full-duplex MIMO systems," presented at *2013 Asilomar Conf. Signals, Syst. Computers*, Nov. 2013.

[113] T. M. Kim, H. J. Yang, A. J. Paulraj, "Distributed sum-rate optimization for full-duplex MIMO system under limited dynamic range," *IEEE Signal Processing Letters*, vol. 20, no. 6, pp. 555-558, June 2013.

[114] J. Zhang, O. Taghizadeh, J. Luo, and M. Haardt, "Full duplex wireless communications with partial interference cancellation," in *Proc. of the 46th Asilomar Conference on Signals, Systems, and Computers, (Pacific Grove, CA)*, Nov. 2012.

[115] J. Zhang, O. Taghizadeh, and M. Haardt, "Joint source and relay precoding design for one-way full-duplex MIMO relaying systems," in *Proc. of Tenth International Symposium on Wireless Communication Systems (ISWCS 2013)*, pp. 1-5, Aug. 2013.

[116] J. Zhang, O. Taghizadeh, and M. Haardt, "Transmit strategies for full-duplex point-to-point systems with residual self-interference," in *Proc. International ITG Workshop on Smart Antennas (WSA 2013)*, Mar. 2013.

[117] J. Zhang, O. Taghizadeh, and M. Haardt, "Robust transmit beamforming design for full-duplex point-to-point MIMO systems," *Tenth International Symposium on Wireless Communication Systems (ISWCS 2013)*, pp. 1-5, Aug. 2013.

[118] S. P. Herath, T. Le-Ngoc, "Sum-rate performance and impact of self-interference cancellation on full-duplex wireless systems," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 881-885, Sept. 2013.

[119] X. Cheng, B. Yu, X. Cheng, L. Yang, "Two-Way full-duplex amplify-and-forward relaying," *IEEE Military Communications Conference, MILCOM*, pp. 1-6, Nov. 2013.

[120] H. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. Smith, "Low-Complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. on Wireless Communications*, 2014.

[121] K. Lee, H. M. Kwon, M. Jo, H. Park and Y. H. Lee, "MMSE-based optimal design of full-duplex relay system," *IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1-5, Sept. 2012.

[122] K. Lee, H. M. Kwon; E. M. Sawan, H Park, and Y. H. Lee, "Half-duplex and full-duplex distributed relay systems under power constraints," *51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 684-689, 2-4 Oct. 2013.

[123] W. Cheng, X. Zhang, and H. Zhang, "QoS driven power allocation over full-duplex wireless links," in *ICC*, pp. 5286-5290, June, 2012.

[124] W. Cheng, X. Zhang, and H. Zhang, "Full/half duplex based resource allocations for statistical quality of service provisioning in wireless relay networks," in *Proc. of the IEEE INFOCOM'12 Conf.*, pp. 864-872, Mar. 2012.

[125] W. Cheng, X. Zhang, and H. Zhang, "Optimal dynamic power control for full-duplex bidirectional-channel based wireless networks," in *Proc. of INFOCOM*, pp. 3120-3128, April 2013.

[126] P. C. Weeraddana, M. Codreanu, M. Latva-aho, and A. Ephremides, "On the effect of self-interference cancelation in multihop wireless networks," *EURASIP J. Wireless Commun. Netw.*, pp. 1-10, Oct. 2010.

[127] V. Aggarwal, M. Duarte, A. Sabharwal, and N. K. Shankaranarayanan, "Full- or half-duplex? A capacity analysis with bounded radio resources," in *IEEE Information Theory Workshop*, Sept. 2012.

[128] N. Shende, O. Gurbuz, and E. Erkip, "Half-duplex or full-duplex relaying: A capacity analysis under self-interference," *47th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, March 2013.

[129] S. Barghi, A. Khojastepour, K. Sundaresan, and S. Rangarajan, "Characterizing the throughput gain of single cell MIMO wireless systems with full duplex radios," in *WiOpt*, pp. 68-74, 2012.

[130] M. Vehkapera, M. A. Girnyk, T. Riihonen, R. Wichman, L. K. Rasmussen, "On achievable rate regions at large-system limit in full-duplex wireless local access," *First International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 7-11, July 2013.

[131] D. Kim, H. Ju, S. Park, and D. Hong, "Effects of channel estimation error on full-duplex two-way networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, November 2013.

[132] E. Everett, D. Dash, C. Dick, and A. Sabharwal, "Self-interference cancellation in multi-hop full-duplex networks via structured signaling," in *49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1619-1626, Sept. 2011.

[133] A. Thangaraj, R. K. Ganti, S. Bhashyam, "Self-interference cancellation models for full-duplex wireless communications," *International Conference on Signal Processing and Communications (SPCOM)*, pp. 1-5, July 2012.

[134] K. Yamamoto, K. Haneda, H. Murata, and S. Yoshida, "Optimal transmission scheduling for hybrid of full- and half-duplex relaying," *IEEE Commun. Lett.*, vol. 15, pp. 305-307, March 2011.

[135] M. Miyagoshi, K. Yamamoto, K. Haneda, H. Murata, S. Yoshida, "Multi-user transmission scheduling for a hybrid of full- and half-duplex relaying," in *Proc 8th International Conference on Information, Communications and Signal Processing (ICICS)*, pp. 1-5, Dec. 2011.

[136] D. Ng, E. Lo, and R. Schober, "Dynamic resource allocation in MIMO-OFDMA systems with full-duplex and hybrid relaying," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1291-1304, May 2012.

[137] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3074-3085, Sep. 2011.

[138] I. Harjula, R. Wichman, K. Pajukoski, E. Lahetkangas, E. Tiirola, and O. Tirkkonen, "Full duplex relaying for local area," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 2684-2688, Sept. 2013.

[139] Q. Meng, W. Feng, G. Zheng, S. Chatzinotas, and B. Ottersten, "Fixed full duplex relaying for wireless broadband communication," *International Conference on Wireless Communications & Signal Processing (WCSP)*, pp. 1-5, Oct. 2012.

[140] H. Ju, S. Lim, D. Kim; H. V. Poor, D. Hong, "Full duplexity in beamforming-based multi-hop relay networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1554-1565, September 2012.

[141] D. S. Michalopoulos, J. Schlenker, J. Cheng, and R. Schober, "Error rate analysis of full-duplex relaying," in *Proc. Intl. Waveform Diversity and Design Conf.*, pp. 165-168, 2010.

[142] T. Kwon, S. Lim, S. Choi, and D. Hong, "Optimal duplex mode for DF relay in terms of the outage probability," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 3628-3634, Sep. 2010.

[143] S. Tedik, and G. Kurt, "Practical full-duplex physical layer network coding", accepted in *VTC 2014*.

[144] T. K. Baranwal, D. S. Michalopoulos, R. Schober, "Outage analysis of multi-hop full duplex relaying," *IEEE Communications Letters*, vol. 17, no.1, pp. 63-66, January 2013.

[145] P. K. Sharma, P. Garg, "Outage analysis of full duplex decode and forward relaying over Nakagami-m channels," *National Conference on Communications (NCC)*, pp. 1-5, Feb. 2013.

[146] H. Alves, D. da Costa, R. Souza, and M. Latva-aho, "Performance of block-markov full duplex relaying with self interference in Nakagami-m fading," accepted in *IEEE Wireless Communications Letters*, vol. 2, no. 3, pp. 311-314, June 2013.

[147] T. M. Kim and A. Paulraj, "Outage probability of amplify-and-forward cooperation with full duplex relay," in *Proc. of IEEE Wireless Commun. Netw. Conf.*, pp. 75-79, 2012.

[148] M. Khafagy, A. Ismail, M.-S. Alouini, and S. Aissa, "On the outage performance of full-duplex selective decode-and-forward relaying," *IEEE Communications Letters*, vol. 17, no. 6, pp. 1180-1183, June 2013.

[149] M. Khafagy, A. Ismail, M.-S. Alouini, and S. Aissa, "Energy efficient cooperative protocols for full-duplex relay channels," 2014.

[150] H. Alves, G. Fraidenraich, R. D. Souza, M. Bennis, and M. Latva-aho, "Performance analysis of full duplex and selective and incremental half duplex relaying schemes," in *Proc. of IEEE Wireless Commun. Netw. Conf.*, pp. 781-785, 2012.

[151] H. Alves, D. Benevides da Costa, R. D. Souza, and M. Latva-aho, "On the performance of two-way half-duplex and one-way full-duplex relaying," *IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 56-60, June 2013.

[152] R. Hu, C. Hu, J. Jiang, X. Xie, and L. Song, "Full-duplex mode in amplify-and-forward relay channels : Outage probability and ergodic capacity," *International Journal of Antennas and Propagation*, January 2014.

[153] I. Krikidis, H. A. Suraweera, and C. Yuen, "Amplify-and-Forward with full-duplex relay selection," in *Proc. of IEEE Intl. Conf. Commun.*, pp. 1-6, 2012.

[154] Y. Liu, X.-G. Xia, and H. Zhang, "Distributed space-time coding for full-duplex asynchronous cooperative communications," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2680-2688, July 2012.

[155] Y. Liu, X.-G. Xia, and H. Zhang, "Distributed linear convolutional space-time coding for two-relay full-duplex asynchronous cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6406-6417, December 2013.

[156] I. Krikidis, H. A. Suraweera, S. Yang, K. Berberidis, "Full-duplex relaying over block fading channel: A diversity perspective," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4524-4535, December 2012.

[157] I. Krikidis, H. A. Suraweera, "Full-Duplex cooperative diversity with Alamouti space-time code," *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 519-522, October 2013.

[158] H. Jin, and V. C. M. Leung, "Full-duplex transmissions in fiber-connected distributed relay antenna systems," *IEEE Global Communications Conference (GLOBECOM)*, pp. 4284-4289, Dec. 2012.

[159] H. Jin, and V. C. M. Leung, "Performance analysis of full-duplex relaying employing fiber-connected distributed antennas," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 146-160, Jan. 2014.

[160] X. Rui, J. Hou, and L. Zhou, "On the performance of full-duplex relaying with relay selection," *Electron. Lett.*, vol. 46, pp. 1-2, Dec. 2010.

[161] X. Rui, J. Hou, and L. Zhou, "Performance analysis of full-duplex relaying with partial relay selection," *Wireless Pers. Commun*, March, 2013.

[162] X. Rui, J. Hou, and L. Zhou, "On the capacity of full-duplex relaying with partial relay selection," *Wireless Pers. Commun*, March, 2014.

[163] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 4381-4393, Dec. 2012.

[164] B. Yu, L. Yang, X. Cheng, and R. Cao, "Relay location optimization for full-duplex decode-and-forward relaying," *IEEE Military Communications Conference, MILCOM 2013*, pp. 13-18, Nov. 2013

[165] Y. Sung, J. Ahn, B. V. Nguyen and K. Kim, "Loop-interference suppression strategies using antenna selection in full-duplex MIMO relays," in *Proc. Intl. Symp. Intelligent Signal Process. and Commun. Syst. (ISPACS 2011)*, Chiangmai, Thailand, pp. 1-4, Dec. 2011.

[166] H. A. Suraweera, I. Krikidis and C. Yuen, "Antenna selection in the full-duplex multi-antenna relay channel," in *Proc. IEEE International Conference on Communications (ICC 2013)*, Budapest, Hungary, June 2013.

[167] M. Zhou, H. Cui, L. Song, and B. Jiao, "Transmit-Receive antenna pair Selection in full duplex systems," *IEEE Wireless Communications Letters*, 2014.

[168] T. Kwon, Y. Kim, and D. Hong, "Comparison of FDR and HDR under adaptive modulation with finite-length queues," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 2, pp. 838-843, Feb. 2012.

[169] A. Mukherjee and A. L. Swindlehurst, "A full-duplex active eavesdropper in MIMO wiretap channels: construction and countermeasures," in *Proc. of Asilomar Conf. on Signals, Syst., and Computers*, Nov. 2011.

[170] G. Zheng, I. Krikidis, L. Jiangyuan, A. P. Petropulu, and B. Ottersten, "Improving physical layer secrecy using full-duplex jamming receivers," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4962-4974, Oct. 2013.

[171] G. Zheng, E. A. Jorswieck, and B. Ottersten, "Cooperative communications against jamming with half-duplex and full-duplex relaying," *77th IEEE Vehicular Technology Conference (VTC Spring)*, pp. 1-5, June 2013.

[172] S. Vishwakarma, and A. Chockalingam, "Sum secrecy rate in full-duplex wiretap channel with imperfect CSI", 2014.

[173] I. Krikidis, G. Zheng and B. Ottersten, "Harvest-Use cooperative networks with half/full-duplex relaying," *IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, April 2013.

[174] H. Alves, M. Bennis, R. D. Souza, and M. Latva-aho, "Enhanced performance of heterogeneous networks through full-duplex relaying", *Special Issue on Femtocells in 4G Systems - EURASIP Journal on Wireless Communications and Networking*, 2012.

[175] J. Bai, and A. Sabharwal, "Distributed full-duplex via wireless side channels: bounds and protocols", *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4162-4173, August 2013.

[176] J. Bai, C. Dick, A. Sabharwal, "Vector bin-and-cancel for MIMO distributed full-duplex," submitted to "IEEE Transactions on Information Theory", Jan 2014.

[177] J. Bai, C. Dick, A. Sabharwal, "K-user symmetric MIMO distributed full-duplex network via wireless side-channels," *51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1144-1151, Oct. 2013.

[178] S. Goyal, P. Liu, S. Hua and S. S. Panwar, "Analyzing a full-duplex cellular system", to appear in the *Proc. of Conference on Information Sciences and Systems (CISS)*, 2013.

[179] X. Fang, D. Yang, and G. Xue, "Distributed algorithms for multipath routing in full-duplex wireless networks," in *Proc. of the IEEE MASS'11 Conf.*, pp. 102-111, Oct. 2011.

[180] X. Fang, D. Yang, and G. Xue, "Pathbook: Cross-layer optimization for full-duplex wireless networks," in *Elsevier*, vol. 57, pp. 1895-1912, June 2013.

[181] N. Singh, D. Gunawardena, A. Proutiere, B. Radunovic, H. V. Balan, and P. B. Key, "Efficient and fair MAC for wireless networks with self-interference cancellation," in *WiOpt*, 2011, pp. 94-101.

[182] J. Y. Kim, O. Mashayekhi, H. Qu, M. Kazadiieva, and P. Levis, "Janus: A novel MAC protocol for full duplex radio," CSTR July, 2013.

[183] K. Sanghoon, W. E. Stark, "On the performance of full duplex wireless networks," *47th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, March 2013.

[184] K. Tamaki, H. A. Raptino, Y. Sugiyama, M. Bandai, S. Saruwatari, and T. Watanabe, "Full duplex media access control for wireless multi-hop networks,"*IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 1-5, June 2013.

[185] D. Ramirez, and B. Aazhang, "Optimal routing and power allocation for wireless networks with imperfect full-duplex nodes," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4692-4704, September 2013.

[186] K. Kato, and M. Bandai, "Routing protocol for directional full-duplex wireless," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 3239-3243, 8-11 Sept. 2013.

[187] M. Fukumoto and M. Bandai, "MIMO full-duplex wireless: Node architecture and medium access control protocol," in *Proc. of International Conference on Mobile Computing and Ubiquitous Networking (ICMU'14)*, pp. 76-77, Singapore, Jan. 2014.

[188] Y. Cai, F. R. Yu, J. Li, Y. Zhou, and L. Lamont, "Medium access control for unmanned aerial vehicle (UAV) ad-hoc networks with full-duplex radios and multipacket reception capability," *IEEE Trans. on Vehicular Technology*, vol. 62, no. 1, pp. 390-394, Jan. 2013.

[189] S. Chen, M.A. Beach and J.P. McGeehan, "Division-free duplex for wireless applications," Electronics Letters, Vol. 34, No. 2, 1998.

[190] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, pp. 3117-3128, Dec. 2002.

[191] Y. Rong and Y. Hua, "Optimal power schedule for distributed MIMO links," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2896-2900, Aug. 2008.

[192] T. Yoo and A. J. Goldsmith, "Capacity and optimal power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203-2214, May 2006.

[193] H. Suzuki, T. V. A. Tran, I. B. Collings, G. Daniels, and M. Hedley, "Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 867-876, Aug. 2008.

[194] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual Tx-RF impairments", in *International ITG Workshop on Smart Antennas*, pp. 189-196, Feb. 2010.

[195] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, pp. 951-963, Apr. 2003.

[196] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585-595, Nov. 1999.

[197] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2781-2788, Oct. 2003.

[198] I. Abramowitz and M. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Washington, DC: U.S. Government Printing Office, 1972.

[199] Y. Rong, Y. Hua, A. Swami, and A. L. Swindlehurst, "Space-time power schedule for distributed MIMO links without instantaneous channel state information at the transmitting nodes," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 686-700, Feb. 2008.

[200] D. P. Bertsekas, *Nonlinear Programming*, 2nd. edition. Belmont, MA: Athena Scientific, 1995.

[201] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641-657. Mar. 1999.

[202] S. Shamai (Shitz) and S. Verdu, "The effect of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, May 2001.

[203] Y. Rong and Y. Hua, "Space-time power scheduling of MIMO links – Fairness and QoS considerations," *IEEE J. Selected Topics Signal Process.*, vol. 2, no. 2, pp. 171-180, Apr. 2008.

[204] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29-43, Jan. 2005.

[205] S. S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, "Weighted sum rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4792-4799, Dec. 2008.

[206] J. Shin and J. Moon, "Weighted-Sum-Rate-Maximizing Linear Transceiver Filters for the K-User MIMO Interference Channel," *IEEE Trans. Communications*, vol. 60, no. 10, pp. 2776-2783, Oct. 2012.

[207] M. Joham, K. Kusume, M. Gzara, W. Utschick, and J. Nossek, "Transmit Wiener filter for the downlink of TDDDS-CDMA systems," in *Proc. IEEE Seventh International Symposium on Spread Spectrum Techniques and Applications 2002*, vol. 1, pp. 9-13, Sept. 2002.

[208] H. Shen, B. Li, M. Tao, and X. Wang, "MSE-based transceiver designs for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480-3489, Nov. 2010.

[209] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second order cone programming," Linear Algebra APP., 1998.

[210] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB tool for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11-12, pp. 625-653, 1999.

[211] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912-1921, Aug. 2003.

[212] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 18-28, Jan. 2004.

[213] D. Tse and P. Viswanath, "On the capacity of the multiple antenna broadcast channel," in *Proc. Multiantenna Channels: Capacity, Coding and Signal Process., DIMACS Workshop Amer. Math. Soc. 2003*, Oct. 2002, vol. 62, pp. 87-105.

[214] A. Khachan, A. J. Tenenbaum, and R. Adve, "Joint transmitter-receiver optimization for downlink multiuser MIMO communications," *IEEE Trans. Wireless Commun.*, 2006, submitted for publication.

[215] 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.

[216] Y. Rong and M. R. A. Khandaker, "On uplink-downlink duality of multihop MIMO relay channel," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1923-1931, Jun. 2011.

[217] M. Ding and S. D. Blostein, "Uplink-downlink duality in normalized MSE or SINR under imperfect channel knowledge," in *Proc. IEEE Global Telecommun. Conf.*, Washington, DC, Nov. 26-30, pp. 3786-3790, 2007.

[218] S. Shi and M. Schubert, "MMSE transmit optimization for multi-user multi-antenna systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, Philadelphia, PA, Mar. 18-23, vol. 3, pp. 409-412, 2005.

[219] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization," *IEEE Trans. Signal Process.*, vol. 55, pp. 5436-5446, Nov. 2007.

[220] R. Hunger, M. Joham, and W. Utschick, "On the MSE-duality of the broadcast channel and the multiple access channel," *IEEE. Trans. Signal Process.*, vol. 57, pp. 698-713, Feb. 2009.

[221] M. B. Shenouda and T. Davidson, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1015-1024, Aug. 2008.

[222] P. Ubaidulla and A. Chockalingam, "Robust joint precoder/receive filter designs for multiuser MIMO downlink," in *Proc. 10th IEEE Workshop Signal Process. Advances Wireless Commun.*, pp. 136-140, Jun. 2009.

[223] M. Ding and S. D. Blostein, "Relation between joint optimizations for multiuser MIMO uplink and downlink with imperfect CSI," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, Las Vegas, NV, pp. 3149-3152, Apr. 2008.

[224] T. E. Bogale, B. K. Chalise, and L. Vandendorpe, "Robust transceiver optimization for downlink multiuser MIMO systems", *IEEE Trans. Signal Process.*, vol. 59, pp. 446-453, Jan. 2011.

[225] J. Liu and Z. Qiu, "Sum MSE uplink-downlink duality of multiuser amplify-and-forward MIMO relay systems," *IEEE Veh. Technol. Conf.*, San Francisco, CA, Sep. 5-8, 2011.

[226] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2658-2668, Oct. 2003.

[227] 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.

[228] S. Han, S. Ahn, E. Oh, and D. Hong, "Effect of channel estimation error on BER performance in cooperative transmission," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2083-2088, May 2009.

[229] O. Amin, S. S. Ikki, and M. Uysal, "On the performance analysis of multirelay cooperative diversity systems with channel estimation errors," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 2050-2059, June 2011.

[230] A. S. Ibrahim and K. J. R. Liu, "Mitigating channel estimation error via cooperative communications," in *Proc. IEEE ICC*, 2009.

[231] S. S. Ikki and S. Aïssa, "Impact of imperfect channel estimation and co-channel interference on dual-hop relaying systems," *IEEE Commun. Lett.*, vol. 16, pp. 324-327, Mar. 2012.

[232] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Process.*, vol. 57, pp. 1141-1150, Mar. 2009.

[233] T. Yoo, E. Yoon, and A. Goldsmith, "MIMO capacity with channel uncertainty: Does feedback help?" in *Proc. IEEE Global Telecommun. Conf.*, Dallas, TX, vol. 1, pp. 96-100, Nov. 29-Dec. 3, 2004.

[234] Y. Rong, "Simplified algorithms for optimizing multiuser multi-hop MIMO relay systems," *IEEE Trans. Commun.*, vol. 59, pp. 2896-2904, Oct. 2011.

[235] Y. Rong, "Optimal linear nonregenerative multihop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2268-2279, Jul. 2010.

[236] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837-4851, Dec. 2009.

[237] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2064-2078, Apr. 2012.

[238] D. Palomar, J. Cioffi, and M. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381-2401, Sep. 2003.

[239] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[240] A. C. Cirik, Y. Rong, Y. Ma, and Y. Hua, "On MAC-BC duality of multihop MIMO relay channel with imperfect channel knowledge," *IEEE Trans. Wireless Commun*, submitted.

[241] M. Ding, "Multiple-input multiple-out wireless system designs with imperfect channel knowledge," Ph.D. thesis, Queen's Univ., Kingston, Ontario, Canada, 2008.

[242] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiations: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740-2746, Jun. 2007.

[243] M. R. A. Khandaker and Y. Rong, "Joint transceiver optimization for multiuser MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 60, pp. 5977-5986, Nov. 2012.