

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Multihop Power Scheduling and MIMO Relay Channel Estimation

Permalink

<https://escholarship.org/uc/item/88d7f2gb>

Author

Kong, Ting

Publication Date

2011

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Multihop Power Scheduling and MIMO Relay Channel Estimation

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

Doctor of Philosophy

in

Electrical Engineering

by

Ting Kong

March 2011

Dissertation Committee:

Professor Yingbo Hua, Chairperson
Professor Ilya Dumer
Professor Albert Wang

Copyright by
Ting Kong
2011

The Dissertation of Ting Kong is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would never have been able to finish my dissertation without the guidance of my committee members, help from friends and support from my family. I would like to take the opportunity to thank these people who pleasantly involved themselves in helping me undertake this dissertation.

I would like to express my deepest gratitude to my advisor, Prof. Yingbo Hua, for his excellent guidance, caring and support. I have benefited tremendously from countless interactions and insightful discussions with him. His emphasis on the fundamental nature of techniques and problems has not only inspired many aspects of this dissertation, but also instilled me with a better understanding of research. He also deserves many, many thanks for carefully reading my manuscripts and patiently improving my English writing.

I thank the other members of my dissertation committee, Prof. Dumer and Prof. Wang, for all the time and effort they have invested in me. Thanks for their additional guidance and instruction through the years, and for always being there whenever I needed any help or advice.

I also want to thank the members of Laboratory of Signals, Systems and Networks in University of California, Riverside: Zheng Fang, Kezhu Hong, Yi Huang, Yuan Yu, Ben Fellows, Shengyang Xu, Xiang Dong, Haitao Liu, Qian Gao and Ali C. Cirik. Thanks for the helpful discussions and encouragement. Their friendship has made my Ph.D. experience both fun and rewarding.

I would like to thank my parents, my husband and my parents-in-law for their unconditional love, whole-hearted support and encouragement through the years.

Finally, I would like to acknowledge that this work was supported in part by the U.S. Army Research Office under the MURI Grant W911NF-04-1-0224.

To my parents.

ABSTRACT OF THE DISSERTATION

Multihop Power Scheduling and MIMO Relay Channel Estimation

by

Ting Kong

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2011
Professor Yingbo Hua, Chairperson

While single-hop wireless networks are commonly deployed today, multihop wireless networks are still in early stages of development. These networks have tremendous potential to be the technology of choice for providing ubiquitous Internet connectivity; minimizing the need for expensive wired infrastructure; and are relatively easy to deploy and maintain. However, there are still many fundamental challenges in wireless multihop relay networks. In this dissertation, we address two challenges that are of significant importance for wireless multihop relay networks.

The first challenge we address is multihop transmission and power scheduling. In a multihop relay network, there are two types of nodes: terminal nodes and router nodes. Terminal nodes transmit directly to the local router node with a single hop. Router node collects the data from its local terminal nodes and is responsible for transmitting these data to the access point (AP). As the router nodes are more sophisticated than the terminal nodes, we assume they can support complex signal processing techniques such as Dirty Paper Coding (DPC) and can support multiple antennas communications as well. In this

dissertation, our goal is to design a transmission scheme which can balance the power consumption in each router node so as to prolong the network life-time.

We first propose a DPC based multihop transmission scheme. An optimization problem of power scheduling and rate allocation to minimize a power related objective function or to maximize a rate related objective function is formulated. Specifically, we model the power balance problem by using min max power function as the objective function. Then, a general gradient projection method is proposed to solve the optimization problem for networks where both single antennas and multiple antennas can be equipped in each node. Some useful properties are explored to realize fast computation. Furthermore, an alternative subgroup method is also provided to reach a tradeoff between performance and complexity when the network size becomes large. Numerical results show that our proposed method achieves better power saving and balance performances compared with existing schemes.

The second challenge we address is the channel estimation and training design for multihop relay channels. We consider a two-hop amplify and forward (AF) Multiple-Input Multiple-Output (MIMO) channel first. To overcome the ambiguity problem in channel estimates, we propose an innovative channel estimation scheme. This scheme has two phases. In the first phase, the source transmits no signal while the relay transmits and the destination receives. In the second phase, the source transmits, the relay amplifies and forwards, and the destination receives. At the destination, the data received in the first phase are used to estimate the relay-to-destination channel, and the data received in the second phase are used to estimate the source-to-relay channel. The linear minimum mean

square error estimation (LMMSE) is used for channel estimation, which allows the use of prior knowledge of channel correlations. The algorithms for finding the optimal source training matrix used at the relay for the first phase, and the optimal source training matrix at the source and the optimal relay training matrix at the relay for the second phase, are developed. Power allocation along the diagonals of source and relay training matrices is solved by using an alternating algorithm with low complexity and fast convergence. The two-phase LMMSE based channel estimation method for two-hop AF MIMO relay channels can be extended for multihop AF MIMO relay channel estimation.

In summary, we discuss and provide solutions to two critical challenges in wireless multihop relay networks: multihop transmission and power scheduling; MIMO relay channel estimation and training design. Our work advance the state-of-the-art in wireless multihop relay networks, and bring us closer to realizing the vision of ubiquitous multihop relay networks.

Contents

List of Figures	xii
List of Tables	xiii
1 Introduction	1
1.1 Background	1
1.2 Challenges and Contributions	5
1.2.1 Multihop Transmission and Power Scheduling	5
1.2.2 MIMO Relay Channel Estimation and Training Design	6
1.3 Dissertation Scope and Outline	8
2 A Review of Transmission Strategies in Multihop Relay Networks	10
2.1 Introduction	10
2.2 System Model	13
2.2.1 System Model	13
2.2.2 Channel Model	16
2.2.3 System Parameters	17
2.3 Single Relay Transmission Schemes	19
2.4 Multiple Relays Two Stage Transmission Schemes	22
2.5 Multiple Relays Multiple Stages Transmission Schemes	30
2.6 Multiple Relays Multiple Stages Parallel Transmission Schemes	34
2.7 Multiple Antennas Transmission Schemes	35
3 Multihop Transmission and Power Scheduling	38
3.1 Introduction	38
3.2 System Model	41
3.3 Dirty Paper Coding based Multihop Transmission	45
3.4 Space Time Power Scheduling	47
3.4.1 Single Antenna At Each Node (SISO)	47
3.4.2 Multiple Antennas At Each Node (MIMO)	52
3.4.3 Low Complexity Algorithm	57

3.5	Numerical Results	59
3.6	Conclusion	63
4	MIMO Relay Channel Estimation and Training Design	65
4.1	Introduction	65
4.2	System Model	70
4.3	Channel Estimation and Training Design for Phase 1	73
4.4	Channel Estimation and Training Design for Phase 2	75
4.4.1	Channel estimation	76
4.4.2	Training Design Problem	77
4.4.3	Decomposition of Trainings	78
4.4.4	Optimal Unitary Components of the Trainings	80
4.4.5	Optimal Diagonal Components of the Trainings: General Case . . .	81
4.4.6	Optimal Diagonal Components of the Trainings: A Special Case . .	91
4.5	Numerical Results	95
4.6	Conclusion	97
5	Conclusion	103
	Bibliography	105
A	Proof of Optimal Trainings for Phase 1	115
B	Proof of Theorem 1	117
C	Proof of Theorem 2	120

List of Figures

2.1	Single relay model	14
2.2	Two stage multiple relays model	15
2.3	Multi-stage multiple relays model	15
2.4	Multi-stage multiple relays parallel model	16
3.1	Multihop relay network topology: (a). AP outside router node cluster (b). AP inside router node cluster	42
3.2	Transmission schemes: (a). Direct Access (b). Nearest Neighbor (c). Dirty Paper Coding	46
3.3	Illustration of dividing multihop network into subgroups	58
3.4	Comparison of P_k for different transmission schemes in SISO ($K=10, a=3$) .	60
3.5	Comparison of P_k for different transmission schemes in MIMO ($K=10, a=1$)	61
3.6	Comparison of the ratio of $\max P_k$ ($K = 3, 4, \dots, 13, a = 3$)	62
3.7	Comparison of $\max P_k$ versus different subgroup size ($a = 4, D = 9$, initialization=100)	63
3.8	Comparison of $\max P_k$ versus different subgroup size ($a = 4, D = 9$, initialization=1,2,3,4,5)	64
4.1	A two-hop AF MIMO relay system	71
4.2	Normalized MSE of \mathbf{H}_1 with optimal trainings and orthogonal trainings. . .	96
4.3	Normalized MSE of \mathbf{H}_1 where \mathbf{H}_1 is uncorrelated and \mathbf{H}_2 is correlated with the correlation factor ρ	97
4.4	Normalized MSE of \mathbf{H}_1 where \mathbf{H}_1 is correlated with the correlation factor ρ and \mathbf{H}_2 is uncorrelated.	98
4.5	Normalized MSE of \mathbf{H}_2 estimated in phase 1 and the normalized MSE of \mathbf{H}_1 estimated in phase 2.	99

List of Tables

2.1	An Overview of Multihop Network Transmission Schemes with Single Antenna	12
2.2	An Overview of Multihop Network Transmission Schemes with Multiple Antennas	13
2.3	A List of System Parameters for Multihop Network Transmission Schemes .	17
2.4	Single Relay Transmission Scheme Comparison	21
2.5	An Overview of Transmission Schemes with 2-stage Multiple Relays	26
2.6	Comparison of Different Transmission Schemes under GNAF Framework . .	27
2.7	Comparison of Beamforming Schemes in 2-Stage Single Antenna Transmission	28
2.8	Performance Comparison of Beamforming Schemes in 2-Stage Single Antenna Transmission	29
2.9	Multistage Transmission Schemes Comparison: Non-cooperative Schemes .	31
2.10	Multistage Transmission Schemes Comparison: Cooperative Schemes	32
2.11	Comparison of Multi-antenna based Transmission Schemes: Two-hop	36
2.12	Comparison of Multi-antenna based Transmission Schemes: Multihop	37
4.1	Comparison of Channel Estimation Methods for Two-hop AF MIMO Relay Channels	70
4.2	Two-layer Bisection Algorithm to Calculate $\lambda_F(j)$	100
4.3	Bisection Algorithm to Calculate $\{x_k^*\}$	101
4.4	Alternating Algorithm to Find Local Optimum $\{x_{lopt}^k, y_{lopt}^k\}$	102

Chapter 1

Introduction

1.1 Background

The rapid evolution of the mobile Internet technology has provided incentives for building efficient multihop wireless networks [1]. The standard task group developing IEEE 802.16j Mobile Multihop Relay (MMR) suggests extending the coverage of a base station by deploying several relay stations around the base station [2]. Multihop relay networks distinguish themselves from the existing infrastructure networks in that wireless transmission is no longer limited between the base station and users, but is utilized to relay information among users and relays as well. The potential advantages of such networks include extending the coverage, reducing the infrastructure costs and expediting deployment time.

Depending on the signal processing capabilities at relay nodes, relay node in multihop networks can be categorized as amplify-and-forward (AF) relay and decode-and-forward

(DF) relay. AF relay only amplifies and retransmits its received signal. Therefore, it introduces noises at the relay. DF relay receives signal, decodes and encodes the information before forwarding to the next relay.

Multihop relay networks can be classified by their application: mobile ad hoc networks (MANET), wireless mesh networks (WMN) and wireless sensor networks (WSN).

In MANET, there is no fixed infrastructure and no pre-determined organization of available links in the network. Individual nodes are responsible for dynamically discovering which other nodes they can directly communicate with. An important assumption is that not all the nodes can directly communicate with each other, so nodes are required to relay packets on behalf of other nodes in order to deliver data across the network. A significant feature of MANET is that rapid changes in connectivity and link characteristics are introduced due to node mobility and power control practices [1].

In WMNs, nodes are comprised of mesh routers and mesh clients. Each mesh router operates not only as a host but also as a router, forwarding packets on behalf of other mesh routers that may not be within direct wireless transmission range of their destinations. A WMN is dynamically self-organized and self-configured, with the mesh routers in the network automatically establishing and maintaining mesh connectivity among themselves (creating, in effect, an ad hoc network). This feature brings many advantages to WMNs such as low up-front cost, easy network maintenance, robustness, and reliable service coverage [3].

A WSN is composed of a large number of sensor nodes that are densely deployed either inside the phenomenon or very close to it. The position of sensor nodes need not

be predetermined. This allows random deployment in inaccessible terrains or disaster relief operations. A unique feature of sensor networks is the cooperative effort of sensor nodes. Sensor nodes are limited in power, computational capacities and memory [4].

The use of multiple antennas for wireless communication systems has gained overwhelming interest during the last decade - both in academia and industry. Multiple-input multiple-output (MIMO) systems can be utilized in order to accomplish a multiplexing gain, a diversity gain, or an antenna gain, thus enhancing the bit rate, the error performance, respectively [5]. MIMO can be sub-divided into three main categories, precoding, spatial multiplexing (SM) and diversity coding [5]-[7].

Precoding is multi-stream beamforming (BF), in the narrowest definition. In more general terms, it is considered to be all spatial processing that occurs at the transmitter. In (single-layer) beamforming, the same signal is emitted from each of the transmit antennas with appropriate phase (and sometimes gain) weighting such that the signal power is maximized at the receiver input. The benefits of beamforming are to increase the received signal gain, by making signals emitted from different antennas add up constructively, and to reduce the multipath fading effect [8].

In spatial multiplexing, a high rate signal is split into multiple lower rate streams and each stream is transmitted from a different transmit antenna in the same frequency channel. If these signals arrive at the receiver antenna array with sufficiently different spatial signatures, the receiver can separate these streams into (almost) parallel channels. Spatial multiplexing is a very powerful technique for increasing channel capacity at higher signal-to-noise ratios (SNR). The maximum number of spatial streams is limited by the

lesser in the number of antennas at the transmitter or receiver [8].

In diversity coding, a single stream (unlike multiple streams in spatial multiplexing) is transmitted, but the signal is coded using techniques called space time coding (STC). Two types of space time codes widely used are space time block code (STBC) and space time trellis code (STTC). The signal is emitted from each of the transmit antennas with full or near orthogonal coding. Diversity coding exploits the independent fading in the multiple antenna links to enhance signal diversity [8].

Employing MIMO in multihop relay networks have received much attention in recent years, e.g., see [9][19]. It is well established that relays can substantially improve the wireless coverage for users subject to limited power and spectral resources. MIMO relays can provide additional power and spectral savings by exploiting the spatial multiplexing and diversity of multiple antennas.

Although MIMO relays bring many unique advantages, they also present some fundamental challenges due to the additional relay nodes. In this dissertation, we will address two challenges in multihop MIMO relay networks: 1. multihop transmission and power scheduling; 2. MIMO relay channel estimation and training design. A detailed description of the challenges and a summary of our contributions are presented in next section.

1.2 Challenges and Contributions

1.2.1 Multihop Transmission and Power Scheduling

In a wireless multihop relay network, different nodes can play different roles. For example, in sensor networks, by introducing clustering hierarchy, nodes organize themselves into local clusters, with one node acting as the local cluster-head. Cluster-heads are generally more powerful than other sensor nodes and have more strict power control requirements. If a cluster-head dies, it will end the useful life-time of the sensor nodes belonging to the cluster [20][21].

We consider two types of nodes in a multihop relay network: terminal nodes and router nodes. Terminal nodes transmit directly to their local router node with a single hop. Router node collects the data from its local terminal nodes and is responsible for transmitting these data to the access point (AP). Similar to the hierarchy in sensor networks, router nodes provide backbone to relay data to AP and are subject to power consumption limitation. As the router nodes are more sophisticated than the terminal nodes, we assume they can support complex signal processing techniques such as DPC [22] and can support multiple antennas communications as well. In this dissertation, we investigate multihop transmission and power scheduling in such a multihop network. It is assumed that DPC is applied in each router node. Single antenna as well as multiple antennas might be equipped in router nodes. Our goal is to balance the power consumption in each router node so as to prolong the network life-time.

Some contributions of our work are as follows:

- A DPC based multihop transmission strategy is proposed.
- An optimization problem of power scheduling and rate allocation is formulated to minimize a power related objective function or to maximize a local rate demand related objective function. Specifically, we model the power balance problem by using min max power function as the objective function.
- A general gradient projection method is proposed to solve the optimization problem for networks where both single antennas and multiple antennas can be equipped in each node. Some useful properties are explored to realize fast computation.
- An alternative subgroup method is also provided to achieve a tradeoff between performance and complexity when the network size becomes large.
- Our proposed method achieves better power performances compared with other existing transmission schemes.

1.2.2 MIMO Relay Channel Estimation and Training Design

Channel estimation and training design for single-hop MIMO channel is well known [23]-[25]. However, for multihop relay channel, the study of both channel estimation algorithms and training design is limited.

In this dissertation, a two-hop AF MIMO relay system is considered. An AF relay is subject to limited signal processing functions and can not decode or estimate information. Therefore, an AF relay may not be able to complete the task of channel estimation by following a single-hop MIMO channel estimation approach. It is because of such a reason that

researchers have started to explore non-conventional MIMO channel estimation methods for MIMO relays.

In previous work [26][27], researchers estimate the source-to-relay channel matrix \mathbf{H}_1 and the relay-to-destination channel matrix \mathbf{H}_2 from the observed composite source-relay-destination channel matrix $\mathbf{H}_c = \mathbf{H}_2\mathbf{F}\mathbf{H}_1$ where \mathbf{F} is a known transformation matrix applied at the AF relay. However, a disadvantage of this approach is that there is always a scalar ambiguity for the estimates of \mathbf{H}_1 and \mathbf{H}_2 .

In this dissertation, aiming at addressing the ambiguity problem in MIMO relay channel estimates, we consider a different channel estimation scheme for the same type of two-hop MIMO relay system as discussed in [26] and [27]. The optimal source and relay training design is discussed.

Some of our contributions are as follows:

- A two-phase LMMSE estimation method for two-hop AF MIMO relay channel is proposed to minimize the channel estimation mean square error (MSE) subject to both energy constraints at the source and the relay. The method results in exact channel estimates without any ambiguity.
- Optimal structures of source and relay training matrices are derived by using convex optimization and majorization theory. Power allocation along the diagonals of source and relay training matrices is solved by using an alternating algorithm with low complexity and fast convergence.
- The two-phase LMMSE based channel estimation method for two-hop AF MIMO

relay channels can be extended for multihop AF MIMO relay channel estimation.

1.3 Dissertation Scope and Outline

The rest of this dissertation is organized as follows: in chapter 2, a review of transmission strategies in multihop relay networks is provided. Different transmission schemes are categorized and compared. Descriptions of the transmission mechanism and performances such as diversity, capacity are presented. Multihop transmission and power scheduling are discussed in chapter 3. A DPC based multihop transmission scheme is proposed. An optimization problem of power scheduling and rate allocation is formulated to minimize a power related objective function or to maximize a local rate demand related objective function. A general gradient projection method as well as a low-complexity method are proposed to solve the problem. Compared with existing transmission schemes, our proposed DPC based transmission scheme is more advantageous in power saving and power balance. In chapter 4, we discuss the channel estimation and training design for MIMO relay channels. A two-hop AF MIMO relay model is considered. A two-phase LMMSE estimation method is proposed to minimize the channel estimation MSE subject to both energy constraints at the source and the relay. Compared with previous channel estimation methods, our proposed algorithm results in exact channel estimates without any ambiguity. Optimal structures of source and relay training matrices are derived by using convex optimization and majorization theory. Power allocation along the diagonals of source and relay training matrices is solved by using an alternating algorithm with low complexity and fast convergence. Some conclusions are drawn in chapter 5.

Shown below is a list of my papers related to this dissertation.

- T. Kong and Y. Hua, “Optimal channel estimation and training design for MIMO relays,” *IEEE 42th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2010.
- T. Kong and Y. Hua, “Optimal training and relay matrices for LMMSE estimation of MIMO channels,” *IEEE Trans. on Signal Processing*, under review.
- T. Kong and Y. Hua, “On power scheduling for multihop wireless relays,” under preparation.

Chapter 2

A Review of Transmission Strategies in Multihop Relay Networks

2.1 Introduction

Transmission strategy is a research topic of significant importance in multihop relay networks. It not only reveals the performance properties such as throughput and diversity, but provides a basis for resource management such as routing, scheduling and power allocation. A large variety of schemes for multihop network transmission have been investigated in the past a few years. In this chapter, we attempt to systematize these research efforts, and provide an overview. Descriptions of the mechanisms and summaries of performances such as throughput, diversity and complexity are presented. However,

due to the consideration of article length, we omit detailed description and mathematical presentation of each scheme. Please refer to [28] for details.

The existing transmission strategies for multihop relay networks are categorized as follows:

- Single Antenna
 - Single Relay
 - Multiple Relays
 - * Two Stage Transmission
 - * Multi-stage Transmission
 - * Multi-stage Parallel Transmission
- Multiple Antennas
 - Two Hop Transmission
 - Multihop Transmission
 - Parallel Partition Transmission

Above methods share a common assumption that all relays are synchronized. Transmission schemes discussed in this chapter are shown in Table 2.1 and Table 2.2. The categorization of different transmission methods can be difficult because of the various criteria. However, aiming at revealing the fundamental differences, we categorize transmission schemes based on following criteria:

- Type of relay: AF, which introduces noise at relay; DF, which demands extra signal processing complexity.
- Number of buffers at relay: one buffer which can only processes instantaneous information; multiple buffers which can process the information received in a period of time .
- Cooperation among relays: no cooperation, e.g., repetition based transmission; with cooperation, e.g. distributed space time code (DSTC)
- Other criteria: for example, different system parameters such as CSI availability. These criteria are considered with lowest priority as they do not reveal the fundamental differences among the transmission strategies.

Table 2.1: An Overview of Multihop Network Transmission Schemes with Single Antenna

Single Antenna	Single Relay	repetition based		
		channel coding based		
		space time coding based		
		special cases		
	Multiple Relays	2 stage	repetition based	
			space time coding based	
			beamforming based	
		multi-stage	non-cooperative transmission	
			cooperative transmission	
		multi-stage parallel	Distributed Space Time Code (DSTC) AF	
Distributed Space Time Code (DSTC) DF				

For DSTC, in particular, either based on AF or DF, some criteria can be applied to evaluate the virtual space time codes include diversity, code rate and maximum likelihood (ML) decodability.

Table 2.2: An Overview of Multihop Network Transmission Schemes with Multiple Antennas

Multi-antennas	2 hop	convolutional code based (DF)
		beamforming (BF) and BLAST based (DF)
		DSTC based (AF)
	Multihop	serial MIMO (DF)
		scale and forward (AF)
		cascaded Orthogonal Space Time Block Code (OSTBC) AF
	Parallel Partition Transmission	independent parallel partition based (AF)
		flip and forward (AF)

2.2 System Model

2.2.1 System Model

Single Relay System Model

Single relay system model consists a source, a relay and a destination. There are two time slots (orthogonal channel) involved in the transmission. Depending on varying degrees of broadcasting and receive collision, four different cooperative transmission schemes are proposed as follows [29]:

1. The source communicates with the relay and destination during the first time slot. In the second time slot, both the relay and source communicate with the destination. This transmission scheme realizes maximum degrees of broadcasting and receive collision.
2. The source communicates with the relay and destination over the first time slot. In the second time slot, only the relay communicates with the destination terminal. This transmission scheme realizes a maximum degree of broadcasting and exhibits no receive collision. Fig.2.1 shows an example of this transmission scheme.

3. The third transmission scheme is identical to transmission scheme 1 apart from the fact that the destination chooses not to receive the direct source to destination signal during the first time slot. This scheme does not implement broadcasting but realizes receive collision.
4. The fourth transmission scheme is identical to transmission scheme 3 apart from the fact that the source does not transmit to destination in the second time slot. This scheme implements neither broadcasting nor receive collision.

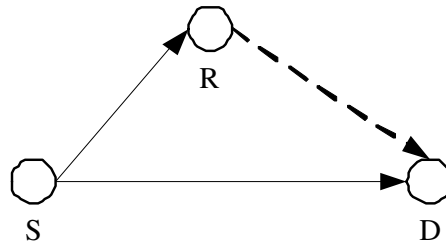


Figure 2.1: Single relay model

Multiple Relays Two Stage Transmission System Model

The multiple relays two stage model under consideration is illustrated in Fig.2.2. There are a source, a destination and N relay nodes. The transmission consists two stages: broadcasting stage and cooperation stage. Treating the group of relay nodes R_1, R_2, \dots, R_N as the same position of the relay node R in the single relay model, the four transmission schemes can be naturally extended to multiple relays case. Fig.2.2 is a natural extension of transmission scheme 4 in single relay transmission model.

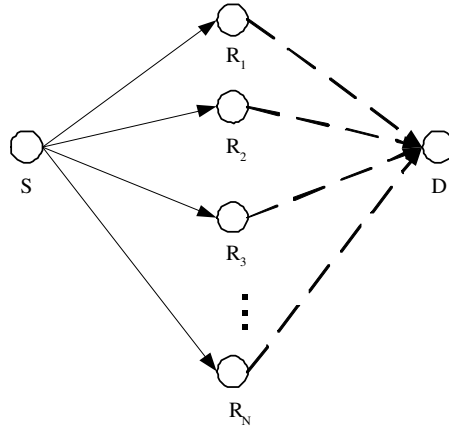


Figure 2.2: Two stage multiple relays model

Multiple Relays Multi-stages Transmission System Model

The multi-stage multiple relays model is illustrated in Fig.2.3. This model consists of a source node S and a destination node D . There are N intermediate relay nodes $R_i, i = 1, \dots, N$ being sequentially placed between S and D . Source node is indexed as 0 and destination node is indexed as $N + 1$. The intermediate relay nodes $R_i, i = 1, \dots, N$ are indexed as $1, 2, \dots, N$ sequentially. All the nodes in the network share a band of radio frequency of $B(\text{Hz})$.

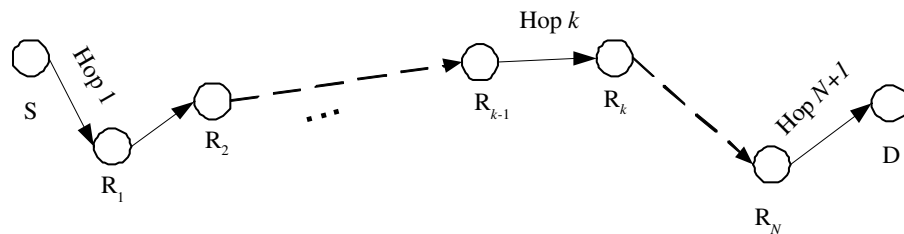


Figure 2.3: Multi-stage multiple relays model

Multiple Relays Multi-stage Parallel Transmission System Model

The multiple relays multi-stage parallel model is illustrated in Fig.2.4. There are a source and a destination as well as N clusters of relay nodes. N clusters are placed sequentially between source and destination with $n_i (i = 1, 2, \dots, N)$ relays in each cluster. This model is a natural extension of the multiple relays multi-stage transmission model with each relay node being replaced by a cluster of relay nodes. One particular constraint imposed is that an intermediate relay cluster only receives packets from its nearest previous cluster and transmits to its nearest following cluster.

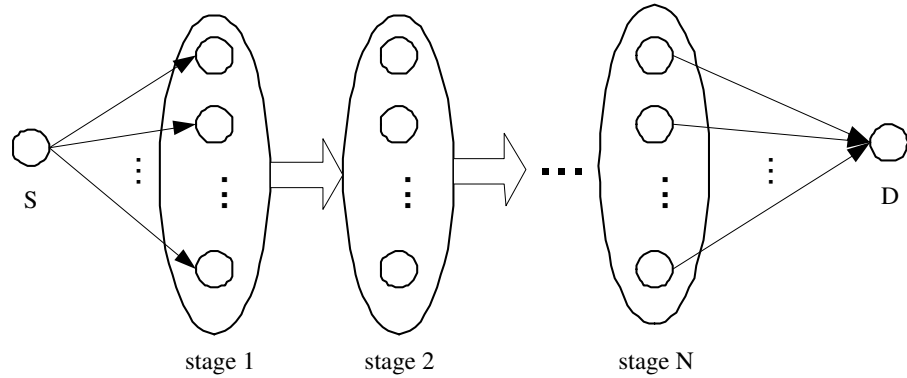


Figure 2.4: Multi-stage multiple relays parallel model

2.2.2 Channel Model

Path loss and microscopic (Rayleigh) fading are considered in the channel model. The basic node to node discrete-time complex baseband transmission model without node cooperation is

$$y_{i,j} = \left(\frac{1}{d_{i,j}} \right)^{\alpha/2} h_{i,j} s_i + n_j \quad (2.1)$$

where $y_{i,j}$ is the received signal at node j from node i . s_i is the transmitted signal from node i . $d_{i,j}$ is the distance between node i and node j . α is the path loss exponent. $h_{i,j}$ is the Rayleigh fading coefficient between node i and node j and is modelled as a complex Gaussian random variable with zero mean and unit variance. The noise at each node is assumed to be complex Gaussian distributed with variance σ_n^2 .

Furthermore, we assume the channel fading coefficients $h_{i,j}$ are independent across different hops. Different CSI availability (coherence/differential) and channel coherence time-symbol duration relation (frequency flat/selective channel) are both considered.

2.2.3 System Parameters

System parameters can be used to distinguish a transmission scheme. In the following table, we list some of the system parameters.

Table 2.3: A List of System Parameters for Multihop Network Transmission Schemes

Node Type	AF	DF	
Duplexity	full duplex	half duplex	
Node Selection	fixed Nodes	selected Nodes	incremental with feedback
CSI availability	full CSI at R and D	full CSI at D partial CSI at R	partial/no CSI at R and D
Partial CSI	second order statistics	quantized feedback	
Channel Type	frequency flat fading	frequency selective fading	
Power Constraint	R,S total power	R total power	R, S individual power
Modulation Type	coherent	differential	
Synchronization	block,carrier and symbol	block and carrier	carrier

Without special claim, the general scenario assumed in the following discussion is half duplex; fixed nodes; full CSI at destination, no CSI at relays for AF and full CSI at

relays for DF; no partial CSI; frequency flat fading channel; individual power constraint at source and relay; coherent modulation; block, carrier and symbol synchronization.

2.3 Single Relay Transmission Schemes

Single relay model shown in Fig.2.1 is the building block for more complicated multihop relay transmission models. The basic principle of transmission scheme design for single relay is to achieve spatial diversity and/or temporal diversity through source-relay cooperation. The research of transmission strategy for single relay model was initiated by A. Sendonaris [30]-[32] with the idea of user cooperation. A. Sendonaris considers two sources which have information s_1, s_2 respectively to transmit to the same destination. Each source receives an attenuated and noisy version of the other source's transmitted signal and uses that, in conjunction with its own data, to construct its transmit signal. The destination receives a noisy version of the sum of the attenuated signals of both sources. It is indicated that user cooperation is beneficial and can result in a higher data rate and a decreased sensitivity to channel variations. Later, it is generalized to realize temporal and spatial diversity through the cooperation of source and relay. Generally, there are three ways to achieve diversity: repetition based cooperation, channel coding based cooperation and space time coding based cooperation. In [30]-[43], the relay is considered as a user so there is no extra resource consumption compared with end-to-end transmission. In the following, we will mainly present the transmission schemes from the relay cooperation perspective. However, some additional information are also provided by considering relay as a user from user cooperation perspective.

From end-to-end transmission (non-cooperative) to single relay assisted transmission (cooperative), the main goal is to introduce diversity. Different schemes are discussed to provide diversity through different mechanics like repetition, channel coding and space

time codes. As more orthogonal channel might be introduced by taking advantage of the relay, our concern is the diversity multiplexing tradeoff (DMT). Assume B bits can be transmitted from source to destination in T seconds in an end-to-end transmission, then the benchmark information rate is $\frac{B}{T}$.

We assume there are two orthogonal channels. In the first orthogonal channel, source broadcasts to relay and destination. In the second orthogonal channel, relay transmits to destination. Transmission schemes in single relay case are compared in Table 2.4. It shows that Nonorthogonal Amplify and Forward (NAF) and Dynamic Decode and Forward (DDF) are two special cases of repetition based cooperation. The main difference of both methods from the repetition method is they assume that a codeword consists of l' consecutive symbol intervals. Source transmits in every symbol interval while relay waits l' intervals and starts to transmit. In Table 2.4, $d(r)$ is the diversity-multiplexing trade off and r is the multiplexing gain. Please refer to [28] for detailed descriptions of each transmission scheme listed in Table 2.4.

Table 2.4: Single Relay Transmission Scheme Comparison

Name	Characteristics	Diversity and Capacity	Performance
Repetition Based Transmission			
fixed relaying AF/DF [35]	always use S-R-D	AF: $d(r) = 2(1 - 2r)$, $0 \leq r \leq 0.5$ full diversity DF: $d(r) = 1 - 2r$, $0 \leq r \leq 0.5$ no full diversity	DF: limited by S-R AF > direct in low rate
selection relaying AF/DF [35]	$ h_{S,R} > h_t$: S-R-D otherwise: S-D	DF: $d(r) = 2(1 - 2r)$ full diversity	selection DF > direct in low rate
incremental relaying [35]	1 bit feedback from D R relay when S-D fails	$d(r) = 1 - r, 0 \leq r \leq 1$ full diversity	> selection relay
NAF [36]	wait l' time slots to transmit	$d(r) \leq (1 - r) + (1 - 2r)^+$ (equality: $l' = 0.5l$)	DMT: upper bound of single relay AF
DDF [36]	R transmits until $l' C_{S,R} > lR$ R: data rate at S	$d(r) = \begin{cases} 2(1 - r), & 0 \leq r \leq 0.5 \\ \frac{1-r}{r}, & 0.5 \leq r \leq 1 \end{cases}$	DMT optimal for single relay ($0 \leq r \leq 0.5$)
Channel Coding Based Transmission			
no ST diversity [41]	R decode N_1	full diversity rate > repetition	entitle to unfairness
with ST diversity [43]	relay both N_2 in time slot 2	the same	> above in fast fading
Space Time Code Based Transmission			
AF based [29]	R relays 2nd row S sends 1st row	full diversity	Code design criteria: rank, determinant
DF based [46]	R, S as ST code	protocol 1,3 full diversity	
Special Transmission			
differential [47]	diff. mod at S	full diversity	PDF of SNR and BER
frequency selective [49]	D-TR STBC D-SC STBC D-OFDM STBC	maximum diversity: $\min\{L_{s,r}, L_{r,d}\} + L_{s,d} + 2$ L: the channel length	Pairwise Error Probability analysis

2.4 Multiple Relays Two Stage Transmission Schemes

For the multiple relays two-stage transmission, there are two phases: broadcast phase and cooperation phase. Depending on the active links in each phase, we have four versions of basic transmission scenarios as follows:

1. In broadcast phase, source broadcasts to relays and destination. In cooperation phase, both source and relays communicate to destination.
2. In broadcast phase, source broadcasts to relays and destination. In cooperation phase, only relays communicate to destination.
3. In broadcast phase, source communicates to relay only. In cooperation phase, source and relays communicate to destination.
4. In broadcast phase, source communicates to relay and in cooperation phase, relays communicate to destination.

The transmission schemes are categorized as Repetition Based Transmission and Beamforming (BF) Based Transmission. An overview of the schemes are shown in Table 2.5 where GNAF is short for general non-orthogonal amplify and forward. The transmission schemes under the framework of GNAF is summarized in Table 2.6 with specific parameter values and performances. The transmission schemes under the framework of beamforming are shown in Table 2.7 and Table 2.8. The system parameters, CSI availability, the objective functions and constraints of beamforming design, the close form of beamforming and performances are compared. In Table 2.7, P_i stands for the power constraint for relay i , P_R

stands for overall power constraint for relays and P_{SR} stands for overall power constraint for source and relays. Please refer to [28] for detailed description of each scheme.

\mathbf{A}_i , \mathbf{B}_i , T_1 and T_2 are the system parameters of the GNAF transmission. In order to clarify above parameters, we summarize the GNAF framework as follows:

In our general framework, the power and channel use allocation are included. And we will consider the most comprehensive case in which source broadcasts to relays and destination in broadcast phase, both source and relays communicate to destination in cooperation phase. The broadcast phase comprises of T_1 channel uses and the cooperation phase comprises of T_2 channel uses. The transmitted vector from the source is a $T_1 \times 1$ complex vector \mathbf{s} satisfying $E(\mathbf{s}^H \mathbf{s}) = 1$. The quantities α_1, α_2 and α_3 are the power allocation factors satisfying $\alpha_1 + \alpha_2 + \alpha_3 = T_1 + T_2$ so that P represents the total average power spent by the source and the relays together. There are two steps:

- Step 1: in the broadcast phase, the received signal at the i th relay and the destination is

$$\mathbf{y}_i = \sqrt{\alpha_1 P} h_{0,i} \mathbf{s} + \mathbf{n}_i \quad i = 1, 2, \dots, N$$

$$\mathbf{y}_{N+1}(1) = \sqrt{\alpha_1 P} h_{0,N+1} \mathbf{s} + \mathbf{n}_{N+1}(1)$$

There are complex matrices \mathbf{A}_i and \mathbf{B}_i at the i th relay. \mathbf{A}_i and \mathbf{B}_i are subject to Frobenius norm constraint $\|\mathbf{A}_i\|^2 + \|\mathbf{B}_i\|^2 \leq 1$. The transmitted signal at relay i is

$$\mathbf{x}_i = \frac{1}{\sqrt{\alpha_1 P + 1}} (\mathbf{A}_i \mathbf{y}_i + \mathbf{B}_i \mathbf{y}_i^*) \quad i = 1, 2, \dots, N$$

- Step 2: in the cooperation phase, the received signal at the destination is

$$\begin{aligned}
\mathbf{y}_{N+1}(2) &= \sqrt{\alpha_2 P} h_{0,N+1} (\mathbf{A}_0 \mathbf{s} + \mathbf{B}_0 \mathbf{s}^*) + \sum_{i=1}^N \sqrt{\alpha_3 P} h_{i,N+1} \mathbf{x}_i + \mathbf{n}_{N+1}(2) \\
&= \sqrt{\alpha_2 P} h_{0,N+1} (\mathbf{A}_0 \mathbf{s} + \mathbf{B}_0 \mathbf{s}^*) + \sum_{i=1}^N \frac{\sqrt{\alpha_3 P} \sqrt{\alpha_1 P} h_{i,N+1}}{\sqrt{\alpha_1 P + 1}} (\mathbf{A}_i h_{0,i} \mathbf{s} + \mathbf{B}_i h_{0,i}^* \mathbf{s}^*) \\
&\quad + \sum_{i=1}^N \frac{\sqrt{\alpha_3 P} h_{i,N+1}}{\sqrt{\alpha_1 P + 1}} (\mathbf{A}_i \mathbf{n}_i + \mathbf{B}_i \mathbf{n}_i^*) + \mathbf{n}_{N+1}(2)
\end{aligned}$$

Stacking the received signal at destination at the first phase and the second phase together, we have

$$\mathbf{y}_{N+1} = \begin{bmatrix} \mathbf{y}_{N+1}(1) \\ \mathbf{y}_{N+1}(2) \end{bmatrix} = \sqrt{\frac{\alpha_3 \alpha_1 P^2}{\alpha_1 P + 1}} \mathbf{S} \mathbf{H} + \mathbf{W}$$

where

$$\begin{aligned}
\mathbf{S} &= \begin{bmatrix} \sqrt{\frac{\alpha_1 P + 1}{\alpha_3 P}} \mathbf{s} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \sqrt{\frac{(\alpha_1 P + 1) \alpha_2}{\alpha_1 \alpha_3 P}} \mathbf{A}_0 \mathbf{s} & \mathbf{A}_1 \mathbf{s} & \cdots & \mathbf{A}_N \mathbf{s} & \sqrt{\frac{(\alpha_1 P + 1) \alpha_2}{\alpha_1 \alpha_3 P}} \mathbf{B}_0 \mathbf{s}^* & \mathbf{B}_1 \mathbf{s}^* & \cdots & \mathbf{B}_N \mathbf{s}^* \end{bmatrix} \\
\mathbf{H}^H &= \begin{bmatrix} h_{0,N+1} & h_{1,N+1} h_{0,1} & \cdots & h_{N,N+1} h_{0,N} & h_{0,N+1} & \cdots & h_{N,N+1} h_{0,N}^* \end{bmatrix} \\
\mathbf{W} &= \begin{bmatrix} \mathbf{n}_{N+1}(1) \\ \sum_{i=1}^N \frac{\sqrt{\alpha_3 P} h_{i,N+1}}{\sqrt{\alpha_1 P + 1}} (\mathbf{A}_i \mathbf{n}_i + \mathbf{B}_i \mathbf{n}_i^*) + \mathbf{n}_{N+1}(2) \end{bmatrix}
\end{aligned}$$

f_i in Table 2.7 and Table 2.8 is the i th beamformer in beamforming based transmission. The framework of beamforming transmission is summarized as follows:

- In the first stage, the transmitter broadcasts to the relay and the destination. The received signals are respectively as follows:

$$y_i(1) = h_{0,i} s + n_i(1) \tag{2.2}$$

$$y_{N+1}(1) = h_{0,N+1}s + n_{N+1}(1) \quad (2.3)$$

- The received signal at the destination at the second stage is

$$\begin{aligned}
y_{N+1}(2) &= \sum_{i=1}^N \frac{f_i}{\sqrt{1+|h_{0,i}|^2}} h_{i,N+1}(h_{0,i}s + n_i(1)) + n_{N+1}(2) \quad (2.4) \\
&= \begin{bmatrix} \frac{h_{0,1}h_{1,N+1}}{\sqrt{1+|h_{0,1}|^2}} & \dots & \frac{h_{0,N}h_{N,N+1}}{\sqrt{1+|h_{0,N}|^2}} \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_N \end{bmatrix} s \\
&\quad + \begin{bmatrix} \frac{n_1(1)h_{1,N+1}}{\sqrt{1+|h_{0,1}|^2}} & \dots & \frac{n_N(1)h_{N,N+1}}{\sqrt{1+|h_{0,N}|^2}} \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_N \end{bmatrix} + n_{N+1}(2) \\
&= \mathbf{h}^T \mathbf{f} s + \mathbf{n}^T(1) \mathbf{f} + n_{N+1}(2)
\end{aligned}$$

where f_i is the beamformer at the i th relay. Combined with (2.3), the received signal at the destination can be written as

$$\begin{bmatrix} y_{N+1}(1) \\ y_{N+1}(2) \end{bmatrix} = \begin{bmatrix} h_{0,N+1} \\ \mathbf{h}^T \mathbf{f} \end{bmatrix} s + \begin{bmatrix} n_{N+1}(1) \\ \mathbf{n}^T(1) \mathbf{f} + n_{N+1}(2) \end{bmatrix} \quad (2.5)$$

Table 2.5: An Overview of Transmission Schemes with 2-stage Multiple Relays

Repetition Based	general repetition[46][34][72]		
Space Time Code Based	AF	GNAF [44][45]	distributed dispersion code [54]
			real/complex orthogonal, quasi-orthogonal code [59]
			single symbol ML decodable code [60]
			2-group ML decodable code [62]
			multi-group ML decodable code [63]
	Hurwitz random matrix based [64][50][65]		
	DSTC with low-rate feedback [66]		
	distributed differential ST code [67][68]		
	DF	space time code based [46][34]	
		randomized space time code [69][70]	
distributed non-coherent STC [71]			
Beamforming Based	centralized BF and power [73][52]		
	distributed BF [75][76][77][81]		
	BF with 2nd order statistics [78]		
	BF with quantized feedback [79]		
	BF in frequency selective channel [80]		
	joint BF and decoder design [82]		

Table 2.6: Comparison of Different Transmission Schemes under GNAF Framework

T_1, T_2	$\mathbf{A}_i, \mathbf{B}_i$	Diversity/Capacity	Performance	
GNAF				
T_1, T_2	$\mathbf{A}_i, \mathbf{B}_i$	DMT lower bound: $d(r) \geq \max [1 - r, (N + 1) (1 - \frac{r(N+1)}{N})^+]$	delay/coding gain efficient non-exponential ML complexity	
Distributed Dispersion Code/OAF [54]				
$T_1 = T_2$	$\mathbf{A}_i, \mathbf{B}_i = \mathbf{0}$	DMT: $\min(T, N) (1 - \frac{\log \log P}{\log P})$ full diversity ($T > N$)	optimal power allocation	
Real Orthogonal Code [59]				
T_1, T_2	$\mathbf{A}_i^T \mathbf{A}_i = \kappa \mathbf{I},$ $\mathbf{A}_i^T \mathbf{A}_j = -\mathbf{A}_j^T \mathbf{A}_i,$ $\mathbf{B}_i = \mathbf{0}$	full diversity	linear decoding complexity; error rate < random code; diversity > selection DF;	
Complex Orthogonal Code [59]				
T_1, T_2	$\mathbf{G}_i^T \mathbf{G}_i = \kappa \mathbf{I},$ $\mathbf{G}_i^T \mathbf{G}_j = -\mathbf{G}_j^T \mathbf{G}_i,$ $\mathbf{G}_i = \mathbf{A}_i \text{ or } \mathbf{B}_i$	-		
Quasi-orthogonal Code [59]				
T_1, T_2	$\mathbf{G}_i^T \mathbf{G}_i = \kappa \mathbf{I}(\text{unitary}),$ $\mathbf{G}_i = \mathbf{A}_i \text{ and } \mathbf{B}_i$	-		
Single-Symbol ML Decodable Code [60]				
$T_1 = T_2$	1. \mathbf{A}_i and \mathbf{B}_i no nonzero entries at the same position. 2. $\mathbf{A}_i, \mathbf{B}_i$ and $\mathbf{A}_i + \mathbf{B}_i$ are column-monomial 3. $\mathbf{A}_{k_1} \mathbf{R}^{-1} \mathbf{A}_{k_2}^H$ $= \mathbf{B}_{k_1} \mathbf{R}^{-1} \mathbf{B}_{k_2}^H$ $= \mathbf{0}_{T_2 \times T_2},$ $1 \leq k_1 \neq k_2 \leq N - 1.$ 4. $\mathbf{A}_{k_1} \mathbf{R}^{-1} \mathbf{A}_{k_2}^H$ $+ \mathbf{B}_{k_1}^* \mathbf{R}^{-1} \mathbf{B}_{k_2}^H$ $= \text{diag}[D_{1,k}, \dots, D_{T_2,k}]$ for $1 \leq k \leq N - 1.$	full diversity upper bound of data rate	single symbol ML decodability; rate twice higher repetition based	
Two Group ML Decodable Code [62]				
$T_1 = T_2$	unitary, satisfy doubling construction rule	full diversity	2 group ML decodable	
Multi-group ML Decodable Code [63]				
$T_1 = T_2$	$\mathbf{A}_i^H \mathbf{R}^{-1} \mathbf{A}_j + \mathbf{A}_j^H \mathbf{R}^{-1} \mathbf{A}_i = \mathbf{0}$	full diversity	multigroup ML decodable	

Table 2.7: Comparison of Beamforming Schemes in 2-Stage Single Antenna Transmission

System Parameter	CSI	Objective Function Constraints	BF design
Centralized BF [73][52]			
$h_{0,N+1} = 0$ $f_0 = 0$	R: $h_{0,i}, h_{i,N+1}$; D: full CSI	Max SNR $\leq P_i$	$f_i = \alpha_i e^{j\theta_i}$, optimize α_i $\theta_i = -[\arg(h_{0,i}) + \arg(h_{i,N+1})]$
$f_0 = 0$	full CSI at scheduler	Max E(SNR) $\leq P_R$	f_i : largest eigenvector
Distributed BF [75][76][77][81]			
$P_0 = 1$ $f_0 = 0$	R: $h_{0,i}, h_{i,N+1}$; D: full CSI	Max E(SNR) $\leq P_R$	$f_i = \frac{\lambda h_{0,i}^* h_{i,N+1}^* \ h_{0,i}\ ^2}{\sqrt{1+\ h_{0,i}\ ^2}} \sqrt{P_R}$ $\lambda = \frac{1}{\sqrt{\sum_{i=1}^N \frac{ h_{0,i} ^2 h_{i,N+1} ^2}{1+ h_{s,i} ^2}}}$ λ : feedback from D
$h_{0,N+1} = 0$ $f_0 = 0$	R: $h_{0,i}, h_{i,N+1}$; D: full CSI	Max SNR $\leq P_i$	D feedback common info to R f_i found locally
BF with Second Order Statistics [78]			
$h_{0,N+1} = 0$ $f_0 = 0$	2nd order channel statistics	Max E{SNR} $\leq P_i$	f_i solved by SDP
BF with Quantized Feedback [79]			
$h_{0,N+1} = 0$ $f_0 = 0$	R: $h_{0,i} + Bbits$; D full CSI	Min BER $\leq P_i$	quantize \mathbf{h} to index m map to \mathbf{f}_m feedback to R
BF in Frequency Selective Channel [80]			
$h_{0,N+1} = 0$ $f_0 = 0$	D: full CSI	Min P_R $SINR \geq \gamma$	$\mathbf{f} = \beta \mathbf{D}^{-\frac{1}{2}} P\{\mathbf{Q}\}$
Joint BF and Decoder [82]			
$f_0 = 0$	3 CSI cases	Max SNR $\leq P_{S+R}$	decoder: MRC; BF: joint/selection

Table 2.8: Performance Comparison of Beamforming Schemes in 2-Stage Single Antenna Transmission

Name	Reference	Diversity	Performance Analysis
centralized BF	[73][52]	full diversity N	-
	[75]	full diversity $N + 1$	-
distributed BF	[75][76][77][81]	full diversity $N + 1$ DMT: $d(r) = (N + 1)(1 - 2r)$	outage probability analysis
	[81]	full diversity N	-
BF with 2nd order statistics	[78]	-	$O(\frac{1}{\log N})$ approximate to nonconvex max SNR
BF with QF	[79]	full diversity when $B \geq \log N$	-
BF freq. sele.	[80]	-	-
Joint BF/decoder	[82]	full diversity $(N+1)$ for full CSI	-

2.5 Multiple Relays Multiple Stages Transmission Schemes

Multiple relays multiple stages transmission refers to the relaying strategy where the packets generated in source are transmitted through several stages until they reach the destination. According to the differences in basic transmission mechanism, the transmission schemes can be categorized into two:

- Non-cooperative transmission: there is no cooperation between source and relays and no cooperation among relays during the transmission process.
- Cooperative transmission: source and relay or relays cooperate to transmit packets from source to destination.

Table 2.9 and Table 2.10 show a summary of multiple relays multiple stages transmission schemes where S-SAF stands for sequential slotted amplify and forward (SAF).

For some relatively simple transmission schemes (such as single hop transmission), we will try to present some mathematical analysis of the system performance such as end-to-end throughput and end-to-end delay. We make following assumptions for the convenience of analysis. For the tractability of the problem and the ease to analyze, we assume the relays are placed equal-distantly from source to destination with a adjacent distance of Δd . Although this network is obviously a simplification and rarely encountered in practice. It allows for tractable analysis and for the establishment of important insights into the performance of general multihop networks [83].

There are several sources of latency in a communication system [84]. In delay analysis for linear multihop network, we only consider the signal processing delay caused

Table 2.9: Multistage Transmission Schemes Comparison: Non-cooperative Schemes

Characteristics	Diversity Capacity	Performance
Single Hop Transmission [83][84]		
S to D directly	power-bandwidth tradeoff	> multihop in bandwidth limited regime
Multihop Transmission: without Spatial Reuse [85][84][86][87]		
forward to nearest node	multihop diversity (Non-)Ergodic Power -bandwidth efficiency	$N_{opt} = \operatorname{argmin} P$ $s.t. r \geq \frac{R}{B}$ (w/o delay constr) [84] $N_{opt} = \operatorname{argmin} P_{out}$ [83] outage probability
Multihop Transmission: with Rate Adaptive [85][83][88]		
parallel transmission	multihop diversity	power-bandwidth tradeoff
CSI known (at Tx)	multihop diversity	$N_{opt} = \operatorname{argmin} P$ $s.t. r \geq \frac{R}{B}$ (non) ergodic power -bandwidth tradeoff
Multihop Transmission: with Spatial Reuse and Rate Adaptive [88]		
parallel transmission	-	Power-bandwidth tradeoff
Multinode Cooperative Diversity Based Transmission [91][96]		
TDMA, each node broadcasts; receiver combine copies	full diversity = cooperation nodes	SER expression for N-node with MPSK or QAM

Table 2.10: Multistage Transmission Schemes Comparison: Cooperative Schemes

Characteristics	Diversity Capacity	Performance
NAF [36]		
s1: S broadcasts; s2: R sends last rx symbol, S sends new symbols	DMT: $d(r) = (1-r) + N(1-2r)^+$	Outperform LW-STC [34]
DDF [36]		
Rs starts to transmit when $l' C_{SR} > lR$	DMT: $d(r) = (N+1)(1-r)$, $0 \leq r \leq \frac{1}{N+1}$ $1 + \frac{N(1-2r)}{1-r}$, $\frac{1}{N+1} \leq r \leq 0.5$ $\frac{1-r}{r}$, $0.5 \leq r \leq 1$	achieve best DMT when $0 \leq r \leq \frac{1}{N+1}$.
SAF [90]		
N Rs M slots S,R sends in each slot	DMT upper bound: $d(r) = (1-r)^+ + N(1 - \frac{M}{M-1}r)$	finite M , SAF DMT < MISO upper bound for $r > \frac{M-1}{M}$; SAF < Non-cooperative
S-SAF [90]		
1 relay per slot R sends last rx	high multiplexing: exploit diversity Rs isolated: upper DMT bound	2 Rs 3 slots case reaches DMT Ubound; $r \leq \frac{2}{3}$: > NAF best 2 relay scheme
Recursive Backward IC Based Transmission (full/half duplex)[83]		
$B+1$ slots send B bits	rate grows: $O(\log N)$ $B \rightarrow \infty$: no rate loss	> multihop in higher rate; delay grows exponentially with N
Coherent Multistage Relay with IC (DF) [94][95]		
block Markov encoding list decoding stripping	achievable rate	achieve the best known rate for Gaussian multiple relay channel

by encoding and decoding at the transmitter and receiver, respectively. We assume the time needed for one time encoding and decoding is T_p . The type of delay we considered is end-to-end delay T_{end} . It is defined as the period of time between a symbol is encoded at the source and is decoded at the destination.

2.6 Multiple Relays Multiple Stages Parallel Transmission Schemes

Multiple relays multiple stages parallel transmission refers to the transmission scenario where there are multiple clusters. Different clusters are placed sequentially between the source and the destination. Messages are received from adjacent previous cluster and only pass on to adjacent following cluster. There are multiple relays in a cluster. If there are only one node in each cluster, then the transmission scenario reduces to multiple relays multiple stages transmission.

There are mainly two transmission schemes falling in this category:

- Multistage Distributed Space Time Codes (DSTC) based on AF protocol
- Multistage Distributed Space Time Codes (DSTC) based on DF protocol

The characteristic of Multistage DSTC protocol is as follows:

- A MIMO channel is modelled between two clusters.
- A broadcast channel is modelled from source to the first cluster.
- A multi-access channel is modelled from the last cluster to destination.
- Within one cluster, AF or DF based DSTC can be applied.

For the purpose of concise, we will omit the description of each transmission scheme here. Please refer to [28] for more details.

2.7 Multiple Antennas Transmission Schemes

When a node (either source, destination and/or relay) is equipped with more than one antenna, the transmission schemes are categorized as multiple antennas transmission schemes. Transmission schemes falling in this category are divided into two parts:

- Two-hop Transmission: the layer between source and destination can be either a single multi-antenna relay or a group of single-antenna relays
- Multihop Transmission: multiple layers are placed from source to destination. Each layer can be a single multi-antenna relay or a group of single-antenna relays

Table 2.12 shows a comparison of the major transmission schemes in multi-antenna category. The characteristics, diversity and other performances are considered. It is assumed that the source has M_s antennas, the relay has M_r antennas and the destination has M_d antennas.

Table 2.11: Comparison of Multi-antenna based Transmission Schemes: Two-hop

Characteristics	Diversity/ Capacity	Performances
Convolutional Code Based Transmission (DF)[100][101][102]		
relay gets N_2 by decoding N_1 STC encoded	full diversity	code design criteria ST cooperation reduces error rate of all nodes
BF and BLAST Based Transmission (DF)[103]		
CSI at Tx; no coordination among relays (single antenna relay)	achieve the same capacity as MIMO	full degrees of freedom
Distributed Space Time Code (AF)[?][105]		
unitary matrix multiplication (single antenna relay)	DMT: $d(r) =$ $(\min\{M_s, M_d\} - r)$ $*(N - r)$ $r \leq \min\{M_s, M_d, N\}$; $M_s \neq M_d$: max diversity	Pairwise Error Probability (PEP) analysis: $M_s = M_d$: $\left(\frac{1}{P}\right)^{\min\{M_s, M_d\}N}$; $M_s \neq M_d$: $\left(\frac{\log^{1/M_s} P}{P}\right)^{M_s N}$;
Threshold Selection Combining (SC)/MRC (DF) [106]		
multi-antenna at each relay SC/MRC received signal forward/BF signal to D D: MRC with delayed signal	-	few relays with many antennas is not significantly worse than many relays with few antennas; SC based is not significantly worse than MRC based

Table 2.12: Comparison of Multi-antenna based Transmission Schemes: Multihop

Characteristics	Diversity/ Capacity	Performances
Serial MIMO Transmission (DF) [107]		
multi-antenna relay as a layer; from layer to layer as MIMO channel	optimal in DMT	-
Scale and Forward Transmission (AF) [103][107]		
each relay in each layer simply scale and forward (single antenna relay)	full degree freedom in high SNR; lower bound of DMT	tradeoff of network size, rate and diversity in the outage formulation
Cascaded OSTBC (AF) [108]		
OSTBC transmitted from source; single antenna relay; \mathbf{A}_i at relay	maximal diversity	single symbol decodability
Parallel Partition Transmission		
Independent Parallel Partition Based Transmission (AF) [107]		
divides a layer into supernodes; find independent parallel AF paths; multi-antennas relay	achieves diversity upper bound	-
Flip and Forward Transmission (AF) [107]		
use flip matrix in each supernode; mismatch between two channels removed; multi-antennas relay	maximal diversity	maximal multiplexing gain

Chapter 3

Multihop Transmission and Power Scheduling

3.1 Introduction

Multihop relay transmission is being considered to improve coverage and increase throughput for the next generation of wireless networks [1]. Different nodes play different roles in a multihop relay networks. For example, in sensor networks, by introducing clustering hierarchy, nodes organize themselves into local clusters, with one node acting as the local cluster-head. Cluster-heads are generally more powerful than other sensor nodes and have more strict power control requirements. If a cluster-head dies, it will end the useful life-time of the sensor nodes belonging to the cluster [20][21].

Dirty Paper Coding (DPC) is introduced by Costa in [22]. The basic premise of DPC is that if interference to a given user is known in advance, the encoding strategy can

exploit the structure of the interference such that the capacity is the same as if there was no interference at all. It is believed that the integration of DPC may lead to advanced multihop networks with high spectral efficiency [110]. And some research work investigating DPC in multihop relay networks are as follows. [111] proposed cooperative transmit diversity using one-dimensional superposition modulation which is an instance of DPC and evaluated its performance with computer simulations. [112] analyzed the superposition modulated cooperative transmission with outage probability and expands to two-dimensional modulation. The author also applied a constellation rotation technique to this system in order to achieve maximum coding gain.

In this chapter, we consider a multihop network where there are two types of nodes: terminal nodes and router nodes. Terminal nodes transmit directly to the local router node with a single hop. Router node collects the data from its local terminal nodes and is responsible for transmitting these data to access point (AP). Similar to the hierarchy in sensor networks, router nodes provide infrastructure/backbone to relay data to AP and are subject to power consumption limitation. As the router nodes are more sophisticated than the terminal nodes, we assume they can support complex signal processing techniques such as DPC and can support multiple antennas communications as well. In this chapter, we investigate the power scheduling in such a multihop networks where DPC is applied in each router node. Single antenna as well as multiple antennas might be equipped in router nodes. Our goal is to balance the power consumption in each router node so as to prolong the network life-time.

Our work is related to [20][21] in using the similar multihop relay network model

with additional DPC based signal processing capability in each router node. The differences between our work and previous DPC based multihop networks [111] [112] is that the previous work investigate DPC based diversity techniques while ours focus on space time power scheduling of a DPC based multihop transmission scheme without taking advantage of diversity.

Some contributions of our work are as follows:

- A DPC based transmission strategy is proposed for multihop networks.
- An optimization problem of power scheduling and rate allocation is formulated to minimize a power related objective function or to maximize a local rate demand related objective function. Specifically, we model the power balance problem by using min max power function as the objective function.
- A general gradient projection method is proposed to solve the optimization problem for networks where both single antennas and multiple antennas can be equipped in each node. Some useful properties are explored to realize fast computation.
- An alternative subgroup method is also provided to achieve a tradeoff between performance and complexity when the network size becomes large.
- Our proposed method achieves better performances compared with existing schemes.

The rest of the chapter is organized as follows. The system model and problem formulation is given in Section 3.2. In Section 3.3 we introduce the DPC based transmission strategy. Gradient projection method as well as a low-complexity subgroup method are

provided in Section 3.4. Numerical simulation results are presented and analyzed in Section 3.5. Finally, some concluding remarks are given in Section 3.6.

3.2 System Model

In this section, we investigate a multihop network composed of terminal nodes and router nodes. The network is divided into many local clusters. Within each cluster, terminal nodes transmit directly to the local router with a single hop. Routers are organized as a backbone, responsible for transmitting the collected information to the AP. Terminal nodes are assumed to have limited signal processing capabilities while router nodes can support advanced signal processing techniques such as DPC and multiple antennas. Depending on the relative positions of router nodes and AP, there are two types of network topologies: (a) AP outside router node cluster, (b). AP inside router node cluster. These two types of topology are shown in Fig. 3.1. For the first case which is shown in Fig. 3.1 (a), it is natural that the far away router nodes use nearer router nodes to relay information to AP. And we can number the router nodes depending on its relative distance to AP in ascending order as $1, 2, \dots, K - 1$ and AP is numbered as K . So there are totally K nodes in the multihop relay network only considering router nodes and AP. For case 2 which is shown in Fig. 3.1 (b), it is more complicated to number the router nodes. However, we can always divide the network into several subnetworks by using angular separations. Within each subnetwork, the AP is outside the corresponding router nodes cluster and we can use the method in case 1 to number the router nodes in each subnetwork. Without loss of generality, we use the topology of case 1 in the following discussions. The information collected at router node k

is assumed to be $r_k(\text{bits/s/Hz}), k = 1, 2, \dots, K$. As we focus on the power scheduling for router nodes, in the following we refer router nodes and AP as nodes for simplicity.

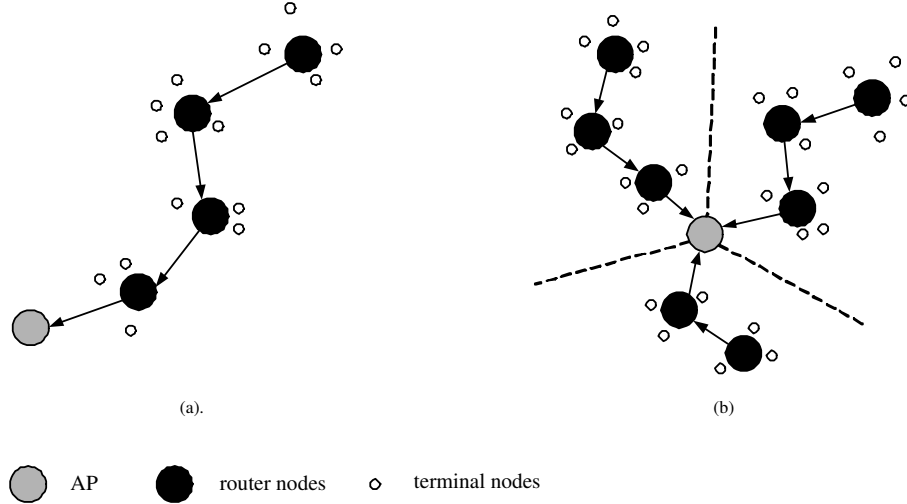


Figure 3.1: Multihop relay network topology: (a). AP outside router node cluster (b). AP inside router node cluster

We assume the whole network share a bandwidth of W Hz and no frequency reuse is applied within the network. Following [35], we assume that the nodes can not transmit and receive at the same time.

We model a frequency flat fading channel between any two nodes. Channel coefficient remain unchanged during the transmission interval. Path loss and small scale fading are both considered. We consider two scenarios in the following: single antenna equipped at each node and multiple antennas equipped at each node.

For node with single antenna,

$$h_{k,j} = \frac{1}{d_{k,j}^{\alpha/2}} w_{k,j} \quad (3.1)$$

where $h_{k,j}$ is the channel coefficient from node k to node j . $d_{k,j}$ is the distance between

node k and node j . α is the path loss exponent and $w_{k,j}$ is a complex Gaussian random variable with zero mean and unit variance, modelling the Rayleigh fading between node k and node j .

For node with M antennas,

$$\mathbf{H}_{k,j} = \frac{1}{d_{k,j}^{\alpha/2}} \mathbf{W}_{k,j} \quad (3.2)$$

where $\mathbf{H}_{k,j}$ is the channel response matrix between node j and node k . $\mathbf{W}_{j,k}$ consists of i.i.d. random complex Gaussian elements and each of them conforms to the distribution $CN(0, 1)$.

Define $S_{in}(k)$ as the set of the nodes which directly transmit information to node k and $S_{out}(k)$ as the set of the nodes which directly receive information from node k . The power consumed at transmit nodes i ($i = 1, 2, \dots, K-1$) is P_k . And the power for the link from node i to node j is $P_{i,j}$ where $P_i = \sum_{j \in S_{out}(i)} P_{i,j}$. The rate (bits/s/Hz) for the link from node i to node j is $R_{i,j}$. The noise at each node j is modelled as a Gaussian variable with zero mean and variance σ_n^2 . Noises in different nodes are independent from each other.

A flow conservation constraint for the multihop relay network is shown as follows:

$$r_k \leq \sum_{j \in S_{out}(k)} R_{k,j} - \sum_{i \in S_{in}(k)} R_{i,k} \quad k = 1, 2, \dots, K-1 \quad (3.3)$$

For every link, the power $p_{i,j}$ and data rate $R_{i,j}$ are related by the Shannon Capacity

$$R_{i,j} = \log_2(1 + SINR_{i,j}) \quad (3.4)$$

where $SINR_{i,j}$ is the signal to interference and noise ratio for the link from node i to node j . Our interest is to investigate the interaction between the power vector $\mathbf{P} =$

$[P_1, P_2, \dots, P_{K-1}]^T$ and the information rate vector demanded to transmit $\gamma = [r_1, r_2, \dots, r_{K-1}]^T$ while satisfying the flow conservation constraint.

Generally speaking, two types of optimization problem can be considered. One is aiming at minimizing power related cost function and the other is aiming at maximizing user rate related function which are shown as follows respectively:

$$\begin{aligned} & \min_{\{P_{i,j}, R_{i,j}, \forall i,j\}} J(\mathbf{P}) & (3.5) \\ \text{s.t. } & r_k \leq \sum_{j \in \mathcal{S}_{out}(k)} R_{k,j} - \sum_{i \in \mathcal{S}_{in}(k)} R_{i,k} \quad k = 1, 2, \dots, K-1 \\ & \mathbf{P} \geq \mathbf{0} \end{aligned}$$

$$\begin{aligned} & \max_{\{P_{i,j}, R_{i,j}, \forall i,j\}} J(\gamma) & (3.6) \\ \text{s.t. } & r_k \leq \sum_{j \in \mathcal{S}_{out}(k)} R_{k,j} - \sum_{i \in \mathcal{S}_{in}(k)} R_{i,k} \quad k = 1, 2, \dots, K-1 \\ & 0 \leq P_k \leq p_k, \quad k = 1, 2, \dots, K-1 \\ & r_k \geq 0, \quad k = 1, 2, \dots, K-1 \end{aligned}$$

where p_k is the peak power constraint for node k . Both problems can be useful depending on applications. Note that in each problem, the variables are $R_{i,j}$ and $P_{i,j}$ ($i = 1, 2, \dots, K-1; j = i+1, i+2, \dots, K$).

One example of the power related function is

$$J(\mathbf{P}) = \left(\sum_{k=1}^{K-1} c_k P_k^p \right)^{\frac{1}{p}} \quad (3.7)$$

where c_k are weights and p is a parameter. When $p = 1$ and $c_k = 1, k = 1, 2, \dots, K-1$,

the cost function reduces to minimizing the sum of power and

$$J(\mathbf{P}) = \sum_{k=1}^{K-1} P_k \quad (3.8)$$

From the definition of norm, we know $\|\mathbf{x}\|_p \approx \|\mathbf{x}\|_\infty$ when $p \rightarrow \infty$. When $c_k = 1, k = 1, 2, \dots, K-1$, p is large, then the cost function becomes

$$J(\mathbf{P}) = \left(\sum_{k=1}^{K-1} P_k^p \right)^{\frac{1}{p}} \approx \max(P_1, P_2, \dots, P_{K-1}) \quad (3.9)$$

In the following, we will use equation (3.9) as the objective function to solve the optimization problem in (3.5). By minimizing the maximum power of the router node set, we can enforce a policy to realize power balance.

3.3 Dirty Paper Coding based Multihop Transmission

In this section, we propose a DPC based transmission strategy for multihop relay networks. In DPC based transmission, for a network with $K-1$ ordered router nodes and one AP (indexed as node K), we assign $K-1$ orthogonal channels for $K-1$ router nodes. In orthogonal channel 1, node 1 divides its own data into $K-1$ parts and send them to node 2, node 3, \dots , and node K (respectively) at the rates $R_{1,2}, R_{1,3}, \dots, R_{1,K}$, (respectively) using DPC. In orthogonal channel 2, node 2 combines its own data with that from node 1 and then divides them into $K-2$ parts and then send these parts to node 3, node 4, \dots , and node K (respectively) at the rates $R_{2,3}, R_{2,4}, \dots, R_{2,K}$, (respectively) using DPC. This process continues until node $K-1$ sends to node K its own data combined with those from node 1, node 2, \dots , and node $K-2$. We can use DPC in such a way that the data stream meant for node j from node k is cancelled at node l where $k < l < j$. In Fig. 3.2,

DPC based transmission as well as two conventional transmission schemes (Direct Access and Nearest Neighbor) are illustrated. Direct Access is shown in Fig. 3.2 (a) where each router node transmits to the AP directly with a single hop. Nearest Neighbor is shown in Fig. 3.2 (b) where one router node only transmits to the router node which is nearest to itself.

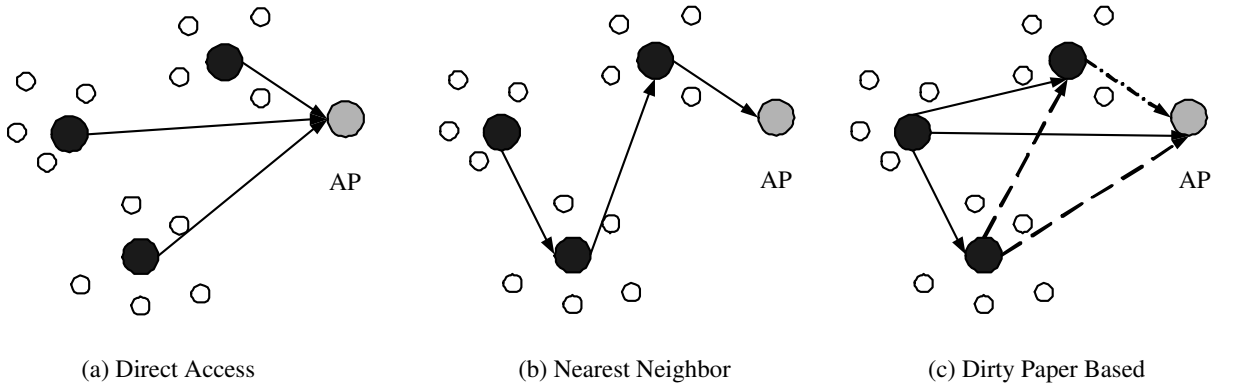


Figure 3.2: Transmission schemes: (a). Direct Access (b). Nearest Neighbor (c). Dirty Paper Coding

In dirty paper coding based transmission, flow conservation constraint in (3.3) can be written as

$$r_k \leq \sum_{j=k+1}^K R_{k,j} - \sum_{i=1}^{k-1} R_{i,k} \quad k = 1, 2, \dots, K-1 \quad (3.10)$$

In the following section, we will provide solutions for the optimization problem shown in (3.5) by applying DPC based transmission schemes both in SISO case and MIMO case.

3.4 Space Time Power Scheduling

In the following, we will provide solutions for the optimization problems under DPC based transmission scheme in both SISO and MIMO cases. Without loss of generality, we will discuss the optimization problem defined in (3.5) with (3.9) as cost function and (3.10) as flow conservation constraint. Assume the time used for transmission for the whole network is normalized to 1, so for DPC based transmission, every node occupies $\frac{1}{K-1}$ of the total time.

3.4.1 Single Antenna At Each Node (SISO)

In DPC, when node k transmits to downstream nodes $i, (k < i \leq K)$, one of the downstream nodes j only interferes with its upstream nodes $l, k < l < j$, the rate from node k to node j is

$$R_{k,j} = \log_2 \left[1 + P_{k,j} \left(\frac{\sigma_n^2}{|h_{k,j}|^2} + \sum_{l=k+1}^{j-1} P_{k,l} \right)^{-1} \right] \quad (3.11)$$

The total power consumed by node k over the whole channel is $P_k = \sum_{j=k+1}^K P_{k,j}$

where

$$P_{k,j} = \left(\frac{\sigma_n^2}{|h_{k,j}|^2} + \sum_{m=k+1}^{j-1} P_{k,m} \right) (2^{R_{k,j}} - 1) \quad (3.12)$$

where $P_{k,k} = 0$.

Stacking the flow conservation constraint in (3.10) for $k = 1, 2, \dots, K-1$ into a vector form, it can be written as:

$$\gamma_K \leq \mathbf{A}_K \mathbf{Y}_K \quad (3.13)$$

where $\gamma_K = (r_1 \ r_2 \ \cdots \ r_{K-1})^T$ and $\mathbf{y}_K = \mathbf{x} = (\mathbf{x}_1^T \ \mathbf{x}_2^T \ \cdots \ \mathbf{x}_{K-1}^T)^T$. We define

$$\mathbf{x}_1 = (R_{1,2} \ R_{1,3} \ \cdots \ R_{1,K-1} \ R_{1,K})^T \quad (3.14)$$

$$\mathbf{x}_2 = (R_{2,3} \ R_{2,4} \ \cdots \ R_{2,K})^T$$

\vdots

$$x_{K-1} = R_{K-1,K}$$

The constraint matrix \mathbf{A}_K is a $(K-1) \times \frac{K(K-1)}{2}$ matrix depending on the number of nodes in the network. Some examples are as follows:

$$\mathbf{A}_2 = [1] \quad (3.15)$$

$$\mathbf{A}_3 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad (3.16)$$

$$\mathbf{A}_4 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{bmatrix} \quad (3.17)$$

\vdots

$$\mathbf{A}_K = \left[\begin{array}{cccc|c} 1 & \cdots & \cdots & 1 & 0 \\ \hline -1 & & & 0 & \\ & \ddots & & \vdots & \mathbf{A}_{K-1} \\ & & -1 & 0 & \end{array} \right] \quad (3.18)$$

And \mathbf{A}_K has following properties

$$\mathbf{A}_K \mathbf{A}_K^T = \begin{bmatrix} K-1 & -1 & \cdots & -1 \\ -1 & K-1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ -1 & \cdots & -1 & K-1 \end{bmatrix} \quad (3.19)$$

$$(\mathbf{A}_K \mathbf{A}_K^T)^{-1} = \frac{1}{K} \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2 \end{bmatrix} \quad (3.20)$$

As \mathbf{P} is a function of \mathbf{y}_K and $\mathbf{y}_K \geq \mathbf{0}$ guarantees $\mathbf{P} \geq \mathbf{0}$, the optimization problem in (3.5) with (3.10) as constraint can be written as

$$\min_{\gamma_K \leq \mathbf{A}_K \mathbf{y}_K, \mathbf{y}_K \geq \mathbf{0}} J(\mathbf{y}_K) \quad (3.21)$$

Optimization with Reduced Dimension

In order to simplify (3.21), we need following theorems.

Theorem 1: For any \mathbf{y}_K , if the equality $\gamma_K \leq \mathbf{A}_K \mathbf{y}_K$ does not hold, there is another $\tilde{\mathbf{y}}_K$ such that $\gamma_K \leq \mathbf{A}_K \tilde{\mathbf{y}}_K$ and $J(\tilde{\mathbf{y}}_K) \leq J(\mathbf{y}_K)$.

Proof: If $\gamma_K \leq \mathbf{A}_K \mathbf{y}_K$, then without of generality, there is k_0 such that for $k < k_0$, $r_k = \sum_{m=k+1}^K R_{k,m} - \sum_{i=1}^{k-1} R_{i,k}$; $r_{k_0} < \sum_{m=k_0+1}^K R_{k_0,m} - \sum_{i=1}^{k_0-1} R_{i,k_0}$; and for $k > k_0$, $r_k \leq \sum_{m=k+1}^K R_{k,m} - \sum_{i=1}^{k-1} R_{i,k}$. Now, if we reduce any non-zero $R_{k_0,m}$ where $m > k_0$, then the resulting cost is also reduced.

Corollary: If $\mathbf{y}_K^* = \arg \min_{\gamma_K \leq \mathbf{A}_K \mathbf{y}_K, \mathbf{y}_K \geq \mathbf{0}} J(\mathbf{y}_K)$, then $\gamma_K = \mathbf{A}_K \mathbf{y}_K^*$.

Now we can write

$$\min_{\gamma_K \leq \mathbf{A}_K \mathbf{y}_K, \mathbf{y}_K \geq 0} J(\mathbf{y}_K) = \min_{\gamma_K = \mathbf{A}_K \mathbf{y}_K, \mathbf{y}_K \geq 0} J(\mathbf{y}_K) \quad (3.22)$$

Define

$$\mathbf{S}_3^T = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \quad (3.23)$$

$$\mathbf{S}_4^T = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & 1 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \end{bmatrix} \quad (3.24)$$

\vdots

$$\mathbf{S}_k^T = \left[\begin{array}{c|c|c} \mathbf{0} & & \mathbf{S}_{k-1}^T \\ \hline 1 & & \\ \vdots & -\mathbf{I} & \mathbf{I} | \mathbf{0} \\ 1 & & \end{array} \right] \quad (3.25)$$

where \mathbf{S}_k^T is a $\frac{(k-1)(k-2)}{2} \times \frac{k(k-1)}{2}$ matrix.

Theorem 2: For $k \geq 3$, $\text{null}(\mathbf{A}_k) = \text{range}(\mathbf{S}_k)$, i.e., $\mathbf{A}_k \mathbf{S}_k = \mathbf{0}$ and $\text{rank}(\mathbf{A}_k) + \text{rank}(\mathbf{S}_k) = \frac{k(k-1)}{2}$

Proof: It is easy to verify.

The solution space of $\gamma_K = \mathbf{A}_K \mathbf{y}_K$ can now be expressed as

$$\mathbf{y}_K = \mathbf{A}_K^T (\mathbf{A}_K \mathbf{A}_K^T)^{-1} \gamma_K + \mathbf{S}_K \mathbf{a}_K \doteq t(\mathbf{a}_K) \quad (3.26)$$

Then (3.22) can be written as

$$\min_{\gamma_K = \mathbf{A}_K \mathbf{y}_K, \mathbf{y}_K \geq 0} J(\mathbf{y}_K) = \min_{t(\mathbf{a}_K) \geq 0} J(\mathbf{a}_K) \quad (3.27)$$

This is an optimization problem with a reduced dimension.

Gradient Projection Method

In the following, we use gradient projection method to solve the optimization problem in (3.27). The gradient is given by

$$\frac{\partial J}{\partial \mathbf{a}} = \frac{\partial \mathbf{t}^T}{\partial \mathbf{a}} \frac{\partial J}{\partial \mathbf{t}} = \mathbf{S}_K^T \frac{\partial J(\mathbf{t})}{\partial \mathbf{t}} \Big|_{\mathbf{t}=\mathbf{t}(\mathbf{a})} \quad (3.28)$$

where $\frac{\partial J(\mathbf{t})}{\partial \mathbf{t}} \Big|_{\mathbf{t}=\mathbf{t}(\mathbf{a})} = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{t}(\mathbf{a})} = \mathbf{g}(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{t}(\mathbf{a})}$ and we have dropped the subscript K from \mathbf{a}_K for convenience. Provided that $t(\mathbf{a}) > \mathbf{0}$, we can apply the unconstrained gradient search with the Armijo rule for the step size. Namely, at iteration k ,

$$\mathbf{a}^{(k+1)} = \mathbf{a}^{(k)} - \beta^m \frac{\partial J}{\partial \mathbf{a}} \Big|_{\mathbf{a}=\mathbf{a}^{(k)}} \quad (3.29)$$

where m is the smallest non-negative integer such that

$$J(t(\mathbf{a}^{(k)})) - J(t(\mathbf{a}^{(k+1)})) \geq \sigma \beta^m \left\| \frac{\partial J}{\partial \mathbf{a}} \Big|_{\mathbf{a}=\mathbf{a}^{(k)}} \right\|^2 \quad (3.30)$$

where $0 < \sigma < 1$. In order to maintain $t(\mathbf{a}) > \mathbf{0}$ during the iteration process, we choose the smallest non-negative integer m satisfying both (3.30) and $t(\mathbf{a}^{(k+1)}) > \mathbf{0}$.

Recall the cost function in (3.9), the gradient of the cost with respect to $R_{l,m}$, $m \geq l + 1$, is

$$\frac{\partial J}{\partial R_{l,m}} = \left(\sum_{k=1}^{K-1} c_k P_k^p \right)^{\frac{1-p}{p}} \sum_{k=1}^{K-1} c_k P_k^{p-1} \sum_{j=k+1}^K \frac{\partial P_{k,j}}{\partial R_{l,m}} \quad (3.31)$$

Note that $R_{l,m} = 0$ and $P_{l,m} = 0$ for $m \leq l$.

For $l \neq k$,

$$\frac{\partial P_{k,j}}{\partial R_{l,m}} = 0 \quad (3.32)$$

Hence,

$$\frac{\partial J}{\partial R_{l,m}} = \left(\sum_{k=1}^{K-1} c_k P_k^p \right)^{\frac{1-p}{p}} c_l P_l^{p-1} \sum_{j=l+1}^K \frac{\partial P_{l,j}}{\partial R_{l,m}} \quad (3.33)$$

For $m \geq j + 1$,

$$\frac{\partial P_{l,j}}{\partial R_{l,m}} = 0 \quad (3.34)$$

It reflects that the “rate” from node l to node m does not affect the “power” from node l to node j , if $j \leq m$.

Hence, for $m \geq l + 1$ (by default for $R_{l,m}$)

$$\frac{\partial J}{\partial R_{l,m}} = \left(\sum_{k=1}^{K-1} c_k P_k^p \right)^{\frac{1-p}{p}} c_l P_l^{p-1} \sum_{j=m}^K \frac{\partial P_{l,j}}{\partial R_{l,m}} \quad (3.35)$$

For $j = m$,

$$\frac{\partial P_{l,m}}{\partial R_{l,m}} = 2^{R_{l,m}} \ln 2 \left(\frac{\sigma_n^2}{|h_{l,m}|^2} + \sum_{\hat{m}=l}^{m-1} P_{l,\hat{m}} \right) \quad (3.36)$$

For $m < j \leq K$

$$\frac{\partial P_{l,j}}{\partial R_{l,m}} = (2^{R_{l,j}} - 1) \sum_{\hat{m}=l}^{j-1} \frac{\partial P_{l,\hat{m}}}{\partial R_{l,m}} \quad (3.37)$$

which should be used together with the previous expression to recursively compute $\frac{\partial P_{l,j}}{\partial R_{l,m}}$.

That is, for each l , choose $j = l + 1, l + 2, \dots, K$ in increasing order. For each j , compute $\frac{\partial P_{l,j}}{\partial R_{l,m}}$ for all $m = l + 1, l + 2, \dots, j$. With the knowledge of (3.31), $\mathbf{g}(\mathbf{x})$ is known. And the gradient $\frac{\partial J}{\partial \mathbf{a}}$ can be calculated accordingly.

In short, gradient projection method can be used to solve the reduced dimension optimization problem in SISO case.

3.4.2 Multiple Antennas At Each Node (MIMO)

When each node in wireless multihop relay networks is equipped with M antennas, we define $\mathbf{P}_{k,j}$ as the covariance matrix of the signal component $\mathbf{s}_{k,j}$ transmitted from node k to node j , where $k = 1, 2, \dots, K - 1$ and $j = j + 1, j + 2, \dots, K$. Then the data rate for

$\mathbf{s}_{k,j}$ is

$$R_{k,j} = \log_2 |\mathbf{I} + \mathbf{P}_{k,j} \mathbf{G}_{k,j}| \quad (3.38)$$

where

$$\mathbf{G}_{k,j} = \mathbf{H}_{k,j}^H \left[\sigma_n^2 \mathbf{I} + \mathbf{H}_{k,j} \left(\sum_{m=k+1}^{j-1} \mathbf{P}_{k,m} \right) \mathbf{H}_{k,j}^H \right]^{-1} \mathbf{H}_{k,j} \quad (3.39)$$

The power consumed by $\mathbf{s}_{k,j}$ is $p_{k,j} = \text{tr}(\mathbf{P}_{k,j})$. The power consumed by node k is

$$P_k = \sum_{j=k+1}^K p_{k,j} = \sum_{j=k+1}^K \text{tr}(\mathbf{P}_{k,j}) \quad (3.40)$$

When $\mathbf{P}_{k,m}$, $m = k+1, \dots, j-1$, are given, we have $\mathbf{G}_{k,j}$ and its eigen-decomposition is denoted by

$$\mathbf{G}_{k,j} = \mathbf{Q}_{k,j} \mathbf{\Lambda}_{k,j} \mathbf{Q}_{k,j}^H \quad (3.41)$$

where $\mathbf{\Lambda}_{k,j} = \text{diag}(\lambda_{k,j}(1), \lambda_{k,j}(2), \dots, \lambda_{k,j}(n_k))$ consists of the eigenvalues in descending order, and n_k is the number of transmit antennas from node k . Then the optimal $\mathbf{P}_{k,j}$ is given by

$$\mathbf{P}_{k,j} = \mathbf{Q}_{k,j} \mathbf{F}_{k,j} \mathbf{Q}_{k,j}^H \quad (3.42)$$

where $\mathbf{F}_{k,j} = \text{diag}(f_{k,j}(1), f_{k,j}(2), \dots, f_{k,j}(n_j))$.

According to waterfilling,

$$f_{k,j}(i) = \left(v_{k,j} - \frac{1}{\lambda_{k,j}(i)} \right)^+ \quad (3.43)$$

and

$$\sum_{i=1}^{n_k} \left(v_{k,j} - \frac{1}{\lambda_{k,j}(i)} \right)^+ = p_{k,j} \quad (3.44)$$

where $(y)^+ = \max(y, 0)$. For each given $p_{k,j}$, the number of non-zero diagonal elements in $\mathbf{F}_{k,j}$ is easy to compute and is denoted by $n_{k,j}$.

Optimization with Reduced Dimension

Similarly to SISO, we define

$$\mathbf{x} = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \cdots \ \mathbf{x}_{K-1}^T] \quad (3.45)$$

where $\mathbf{x}_j = [R_{j,j+1} \ R_{j,j+2} \ \cdots \ R_{j,K}]^T$.

$$\mathbf{p} = [\mathbf{p}_1^T \ \mathbf{p}_2^T \ \cdots \ \mathbf{p}_{K-1}^T] \quad (3.46)$$

where $\mathbf{p}_j = [p_{j,j+1} \ p_{j,j+2} \ \cdots \ p_{j,K}]^T$. Also define $\mathbf{r} = [r_1 \ r_2 \ \cdots \ r_{K-1}]^T$.

Using the introduced matrix \mathbf{S} (4.26) and vector \mathbf{a} (3.26), the reduced dimension optimization problem (3.27) can be equivalently written as

$$\min_{\mathbf{x}=\mathbf{A}+\mathbf{r}+\mathbf{S}\mathbf{a} \geq 0, \mathbf{p} \leq \bar{\mathbf{p}}} J(\mathbf{p}) \quad (3.47)$$

Gradient Projection Method

The cost is generally non-linear and non-quadratic and even non-convex. But we can use the gradient projection (GP) method and many initializations to search for the best solution.

For (3.47), the GP method requires

$$\frac{\partial J}{\partial \mathbf{a}} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{a}} \frac{\partial J}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{a}} \frac{\partial \mathbf{p}^T}{\partial \mathbf{x}} \frac{\partial J}{\partial \mathbf{p}} = \mathbf{S}^T \frac{\partial \mathbf{p}^T}{\partial \mathbf{x}} \frac{\partial J}{\partial \mathbf{p}} \quad (3.48)$$

where $\frac{\partial J}{\partial \mathbf{p}}$ is easy to find in general once J is defined, but $\frac{\partial \mathbf{p}^T}{\partial \mathbf{x}}$ is not easy to find.

However, there is a one-to-one mapping between \mathbf{p} and \mathbf{x} . In an abstract form, we can write $\mathbf{x} = f(\mathbf{p})$, but the inverse of this function is not available in a closed form. To

compute $\frac{\partial \mathbf{p}^T}{\partial \mathbf{x}}$, we can apply the gradient operation on both sides of $\mathbf{x} = f(\mathbf{p})$:

$$\frac{\partial \mathbf{x}^T}{\partial \mathbf{x}} = \frac{\partial \mathbf{p}^T}{\partial \mathbf{x}} \frac{\partial f^T(\mathbf{p})}{\partial \mathbf{p}} \quad (3.49)$$

where the left side equals the identity matrix, and hence:

$$\frac{\partial \mathbf{p}^T}{\partial \mathbf{x}} = \left(\frac{\partial f^T(\mathbf{p})}{\partial \mathbf{p}} \right)^{-1} \quad (3.50)$$

Each entry of $\frac{\partial f^T(\mathbf{p})}{\partial \mathbf{p}}$ is determined by $\frac{\partial R_{j,k}}{\partial p_{m,l}}$ for some j, k, m and l .

To compute $\frac{\partial R_{j,k}}{\partial p_{m,l}}$, we first note that

$$\frac{\partial R_{j,k}}{\partial p_{m,l}} = 0 \quad \text{if } j \neq m \quad (3.51)$$

and

$$\frac{\partial R_{j,k}}{\partial p_{j,l}} = 0 \quad \text{if } k < l \quad (3.52)$$

If $j = m$ and $k = l$, as $\partial \ln |\mathbf{x}| = \text{tr}[\mathbf{x}^{-1} \partial \mathbf{x}]$, we have

$$\begin{aligned} \frac{\partial R_{j,k}}{\partial p_{j,k}} &= \frac{\partial}{\partial p_{j,k}} \log_2 |\mathbf{I} + \mathbf{P}_{j,k} \mathbf{G}_{j,k}| \\ &= \frac{\log_2 e}{\partial p_{j,k}} \text{tr} \left[(\mathbf{I} + \mathbf{P}_{j,k} \mathbf{G}_{j,k})^{-1} \partial \mathbf{P}_{j,k} \mathbf{G}_{j,k} \right] \\ &= \frac{\log_2 e}{\partial p_{j,k}} \text{tr} \left[(\mathbf{I} + \mathbf{F}_{j,k} \Lambda_{j,k})^{-1} \partial \mathbf{F}_{j,k} \Lambda_{j,k} \right] \end{aligned} \quad (3.53)$$

Since $p_{j,k} = \text{tr}(\mathbf{F}_{j,k}) = \sum_{i=1}^{n_j} \left(v_{j,k} - \frac{1}{\lambda_{j,k}(i)} \right)^+$, we have $\partial p_{j,k} = n_{j,k} \partial v_{j,k}$. Then, we have

$$\begin{aligned} \partial \mathbf{F}_{j,k} &= \text{diag} \left(\mathbf{I}_{n_{j,k} \times n_{j,k}} \partial v_{j,k} \mathbf{0}_{(n_j - n_{j,k}) \times (n_j - n_{j,k})} \right) \\ &= \frac{1}{n_{j,k}} \text{diag} \left(\mathbf{I}_{n_{j,k} \times n_{j,k}} \mathbf{0}_{(n_j - n_{j,k}) \times (n_j - n_{j,k})} \right) \partial p_{j,k} \end{aligned} \quad (3.54)$$

Therefore,

$$\frac{\partial R_{j,k}}{\partial p_{j,k}} = \frac{\log_2 e}{n_{j,k}} \sum_{i=1}^{n_{j,k}} \frac{\lambda_{j,k}(i)}{1 + f_{j,k} \lambda_{j,k}(i)} \quad (3.55)$$

$$\begin{aligned} &= \frac{\log_2 e}{n_{j,k}} \sum_{i=1}^{n_{j,k}} \frac{1}{v_{j,k}} \\ &= \frac{\log_2 e}{v_{j,k}} \end{aligned} \quad (3.56)$$

If $k > l$, we have

$$\begin{aligned} &\frac{\partial R_{j,k}}{\partial p_{j,l}} \quad (3.57) \\ &= \frac{\partial}{\partial p_{j,l}} \log_2 |\mathbf{I} + \mathbf{P}_{j,k} \mathbf{G}_{j,k}| \\ &= \frac{\log_2 e}{\partial p_{j,l}} \text{tr} \left[(\mathbf{I} + \mathbf{P}_{j,k} \mathbf{G}_{j,k})^{-1} \mathbf{P}_{j,k} \partial \mathbf{G}_{j,k} \right] \\ &= \frac{\log_2 e}{\partial p_{j,l}} \text{tr} \left[(\mathbf{I} + \mathbf{P}_{j,k} \mathbf{G}_{j,k})^{-1} \mathbf{P}_{j,k} \mathbf{H}_{j,k} \partial \left(\mathbf{I} + \mathbf{H}_{j,k} \sum_{m=j+1}^{k-1} \mathbf{P}_{j,m} \mathbf{H}_{j,m}^H \right)^{-1} \mathbf{H}_{j,k}^H \right] \\ &= -\frac{\log_2 e}{\partial p_{j,l}} \text{tr} \left[(\mathbf{I} + \mathbf{P}_{j,k} \mathbf{G}_{j,k})^{-1} \mathbf{P}_{j,k} \mathbf{H}_{j,k} \left(\mathbf{I} + \mathbf{H}_{j,k} \sum_{m=j+1}^{k-1} \mathbf{P}_{j,m} \mathbf{H}_{j,m}^H \right)^{-1} \mathbf{H}_{j,k} (\partial \mathbf{P}_{j,k}) \mathbf{H}_{j,k}^H \right. \\ &\quad \left. \left(\mathbf{I} + \mathbf{H}_{j,k} \sum_{m=j+1}^{k-1} \mathbf{P}_{j,m} \mathbf{H}_{j,m}^H \right)^{-1} \mathbf{H}_{j,k}^H \right] \end{aligned}$$

The dimension of $\frac{\partial f^T(p)}{\partial p}$ is $\frac{K(K-1)}{2} \times \frac{K(K-1)}{2}$. But $\frac{\partial f^T(p)}{\partial p}$ is highly structured. For example,

$$\frac{\partial f^T(p)}{\partial p} \Big|_{K=3} = \begin{bmatrix} \frac{1}{v_{1,2} \ln 2} & \frac{\partial r_{1,3}}{\partial p_{1,2}} & 0 \\ 0 & \frac{1}{v_{1,3} \ln 2} & 0 \\ 0 & 0 & \frac{1}{v_{2,3} \ln 2} \end{bmatrix} \quad (3.58)$$

In order to compute $\frac{\partial J}{\partial \mathbf{a}} = \mathbf{S}^T \frac{\partial \mathbf{p}^T}{\partial \mathbf{x}} \frac{\partial J}{\partial \mathbf{p}}$, we need to compute $\mathbf{t} = \frac{\partial \mathbf{p}^T}{\partial \mathbf{x}} \frac{\partial J}{\partial \mathbf{p}} = \left(\frac{\partial f^T(\mathbf{p})}{\partial \mathbf{p}} \right)^{-1} \frac{\partial J}{\partial \mathbf{p}}$.

We can define

$$\mathbf{t} = \begin{bmatrix} \mathbf{t}_{K-1} \\ \mathbf{t}_{K-2} \\ \vdots \\ \mathbf{t}_1 \end{bmatrix}, \quad \frac{\partial J_1}{\partial \mathbf{p}} = \begin{bmatrix} \mathbf{a}_{K-1} \\ \mathbf{a}_{K-2} \\ \vdots \\ \mathbf{a}_1 \end{bmatrix} \quad (3.59)$$

$$\frac{\partial f^T(\mathbf{p})}{\partial \mathbf{p}} = \begin{bmatrix} \mathbf{B}_{K-1} & 0 & \cdots & 0 \\ 0 & \mathbf{B}_{K-2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{B}_1 \end{bmatrix} \quad (3.60)$$

where \mathbf{t}_k has the dimension $k \times 1$, \mathbf{a}_k has the dimension $k \times 1$, and \mathbf{B}_k is upper triangular and has the dimension $k \times k$. Then we have

$$\mathbf{B}_k \mathbf{t}_k = \mathbf{a}_k, \quad k = 1, 2, \dots, K-1 \quad (3.61)$$

since \mathbf{B}_k is upper triangular, $\mathbf{t}_k = \mathbf{B}_k^{-1} \mathbf{a}_k$ can be easily computed by back-substitution. That is, first compute the last element of \mathbf{t}_k , then the second last element, and so on.

3.4.3 Low Complexity Algorithm

It is also noted that when the dimension of rate distribution vector is large, the gradient projection method becomes less efficient and the computation complexity is high. Hence, in this section, we propose a solution to achieve a better tradeoff between performance and complexity.

The basic idea is to decompose a large multihop relay network into several subgroups. The number of nodes in a subgroup is denoted by the subgroup size s . Fig.3.3 shows

how a large multihop network is decomposed. In order to realize power balance, we should try to pick up the nodes for a subgroup as evenly as possible from the original network. In Fig.3.3, the original networks consisting 7 nodes. And it can be decomposed to either 3 subgroups with $s = 3$ or 2 subgroups with $s = 4$. It is worth noticing that when $s = 2$, it is Direct Access. Therefore, our space time power scheduling scheme can be applied to each subgroup with the dimension of rate distribution vector greatly reduced.

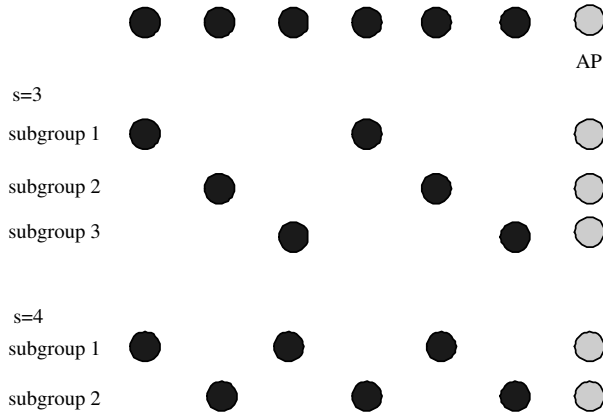


Figure 3.3: Illustration of dividing multihop network into subgroups

The size of the subgroup s is a crucial parameter to decompose a large multihop network. Given a fixed size of original multihop network, the possible choices of s is discrete and should satisfy $\frac{K-1}{s-1} = integer, s > 0, s \in integer$. Some numerical results in the next section show that power performance varies with different choices of subgroup size. Given a fixed distance, one interesting choice of the subgroup size is $s = 3$ for which the optimization problem in (3.5) can produce a deterministic solution. By decomposing the large multihop network into subgroups with group size $s = 3$ and use the deterministic solution, the computation complexity is greatly reduced. Therefore, subgroup based low complexity

method is a good candidate for suboptimal solution.

3.5 Numerical Results

In this section, we present some numerical results through simulations to provide further insights. The fading factors are fixed $w_{k,j} = 1$ and \mathbf{W} is a matrix with all element being 1. The number of antennas in each node is $M = 2$ for MIMO case. Local traffic demands are assumed to be $r_k = 1, k = 1, 2, \dots, K - 1$. $p = 50$ is used in the objective function. In the gradient projection method, some typical choices of the parameters are $\beta = 0.5, \sigma = 0.1$.

- Comparison of power performances of different transmission schemes

We first compare our proposed gradient projection method with Direct Access and Nearest Neighbor methods. In the following simulation, we assume the distance between two adjacent nodes is $\Delta d = 1$. Fig.3.4 and Fig.3.5 compare maximal P_k of Direct Access, Nearest Neighbor and DPC methods in SISO and MIMO cases, respectively. In SISO case, the power loss exponent $a = 3$ while in MIMO case the power loss exponent $a = 1$. Large a in MIMO case will result in huge power consumption P_1 as the power increases exponentially with the increase of distance. Therefore, we choose a to be a small number in MIMO case for the convenience of comparing the three transmission schemes. From both figures, we can observe that DPC method balances the power consumption for each node while meeting the traffic demand. In Direct Access, far away node is in disadvantage for the long distance to AP. In Nearest

Neighbor, nearest node to AP is in disadvantage for relaying all the previous data to AP. With DPC, we have additional freedom to optimize over power scheduling and rate allocation to achieve a better performance. It is also worth noticing that under different parameters such as path loss exponent, the performance of Direct Access and Nearest Neighbor can be quite different.

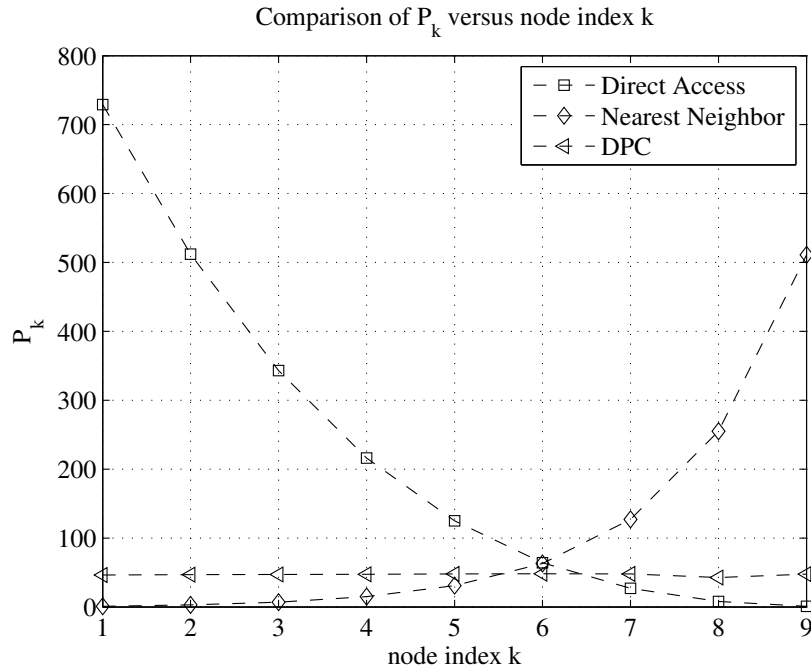


Figure 3.4: Comparison of P_k for different transmission schemes in SISO ($K=10, a=3$)

In order to see the performance variation with regard to network size K , we define two ratios: $\frac{MaxPower_{DA}}{MaxPower_{DPC}}$ and $\frac{MaxPower_{NN}}{MaxPower_{DPC}}$ where $MaxPower_{DA}$ stands for the maximal power of P_k with respect to k for Direct Access method and $MaxPower_{NN}$ stands for the maximal power of P_k with respect to k for Nearest Neighbor method. Fig. 3.6 shows both the ratio variations with the increase of the network size K . It is observed

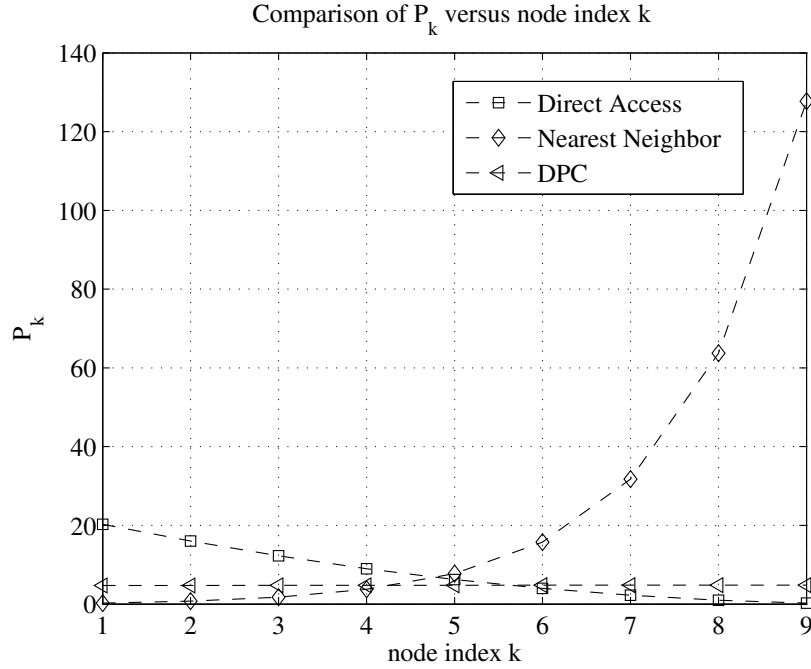


Figure 3.5: Comparison of P_k for different transmission schemes in MIMO ($K=10, a=1$)

that both ratios are far above 1 which means that the DPC method always consumes less maximal power. One interesting observation is that the ratio $\frac{MaxPower_{DA}}{MaxPower_{DPC}}$ first increases and then slowly decreases. The peak appears around $K = 9$. $\frac{MaxPower_{NN}}{MaxPower_{DPC}}$ increases with the increase of network size K . Because with more nodes in the network, the power consumption at node $K - 1$ grows larger. Initialization of DPC method is chosen to be 100 for $K \leq 11$, 200 for $K = 12$, 300 for $K = 13$.

- Subgroup based low complexity algorithm

As discussed in previous section, the optimal subgroup size s can be investigated by simulation given a fixed distance. Here, instead of fixing the distance between two adjacent nodes, we fix the distance between the first node and AP. In the following

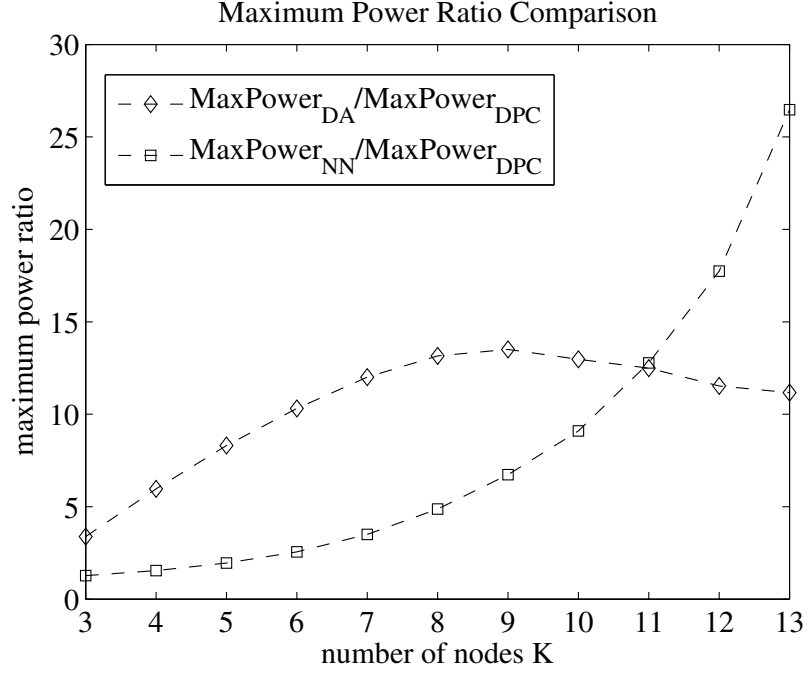


Figure 3.6: Comparison of the ratio of max P_k ($K = 3, 4, \dots, 13, a = 3$)

simulations, we choose the distance to be $D = 9$. And the number of nodes in the network $K = 20$. we will show the maximal power per node with different subgroup sizes. Fig.3.7 and Fig.3.8 show the power performances when the initialization=100 and intializaiton=1,2,3,4,5, respectively. From both figures, we can observe that when $s = 2$, it is Direct Access. From $s = 2$ to $s = 3$, the power drops most significantly. The power differences between $s = 3$ and other subgroup are not very large. Therefore, to achieve a tradeoff between performance and complexity, $s = 3$ seems to be a good choice. Plus, when $s = 3$, the optimization problem in (3.5) has deterministic solution which can be solved by bisection method, for example. That's why in all figures, no matter what the initialization situation is, the value for $s = 3$ is always the same. Thus

the deterministic solution greatly reduces the computation complexity introduced by gradient projection.

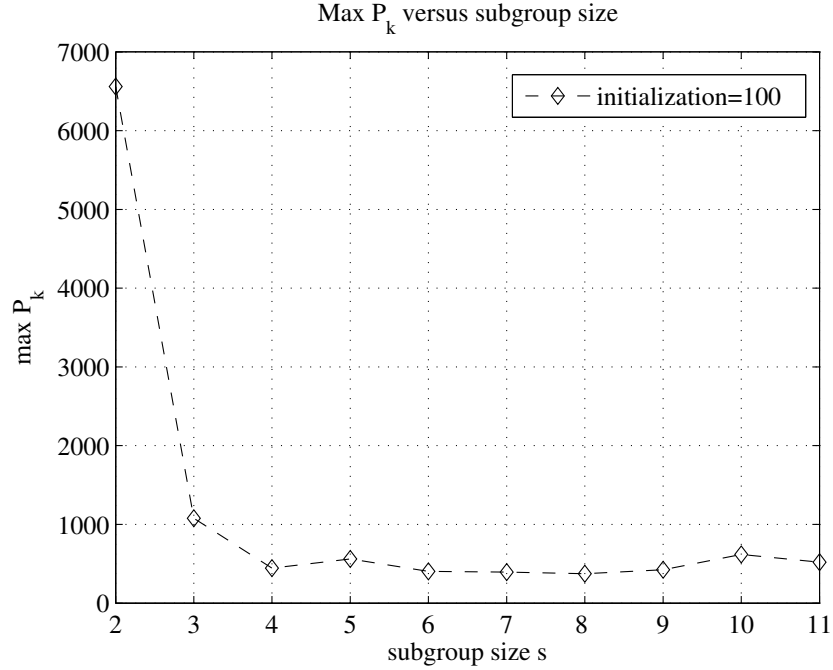


Figure 3.7: Comparison of $\max P_k$ versus different subgroup size ($a = 4, D = 9$, initialization=100)

3.6 Conclusion

We proposed a DPC based transmission strategy for multihop relay networks. By taking advantage of DPC, we have additional freedom to optimize over power scheduling and rate allocation to realize the power balance goal. A general gradient projection method is proposed to solve the optimization problem for networks where both single antennas and multiple antennas can be equipped in each node. Some useful properties are explored to realize fast computation. An alternative subgroup method is also provided to reach a good

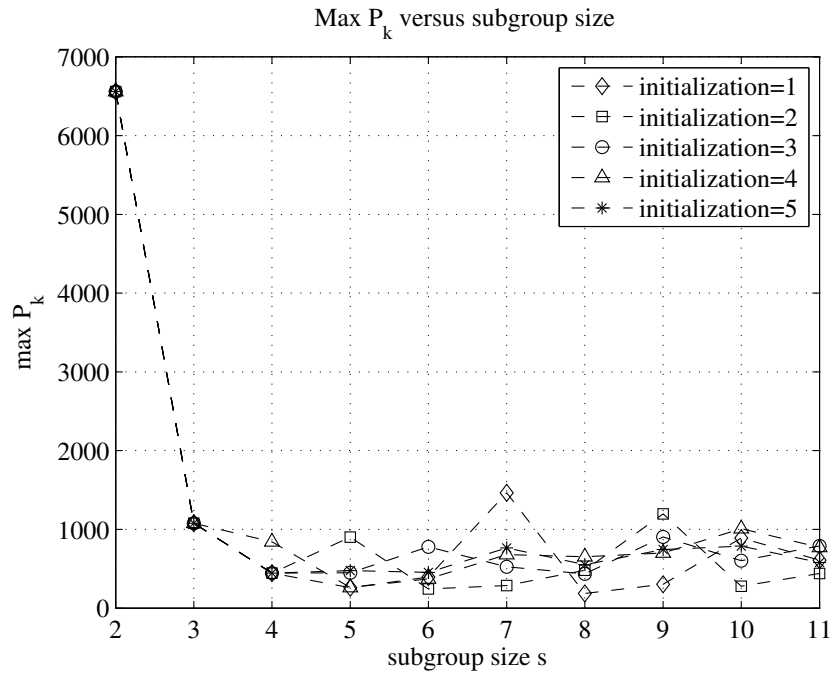


Figure 3.8: Comparison of $\max P_k$ versus different subgroup size ($a = 4, D = 9$, initialization=1,2,3,4,5)

tradeoff between performance and complexity when the network size becomes large. Our proposed method achieves performance advantage compared with existing schemes.

Chapter 4

MIMO Relay Channel Estimation and Training Design

4.1 Introduction

Relay-based cooperative communications have drawn great interest recently [30][35]. When multiple antennas are deployed at one or more nodes of the relay system, we also call such relay system a MIMO relay system. The achievable rate and capacity upper bound of a MIMO relay system have been studied in [9]. The diversity-multiplexing tradeoff of multiantenna cooperative systems has been studied in [113].

Depending on the signal processing capabilities at relay nodes, MIMO cooperative relay systems can be categorized as amplify-and-forward (AF) relays and decode-and-forward (DF) relays. For AF MIMO relays, relay node only amplifies and retransmits its received signal. Compared with DF relays, the complexity of AF relays is much lower. And

this advantage is especially important when all nodes are equipped with multiple antennas, since decoding multiple data streams involves much more computational efforts than decoding a single data stream.

Many studies of AF MIMO relays can be found in [9]-[19] and the references therein. As shown in these works, the knowledge of the channel matrices between nodes is important for optimized system performances. For example, it is shown in [18] that if the channel matrices between adjacent nodes in a multihop AF MIMO relay system are known, then an optimal design of the source covariance matrix and the transformation matrices at all relays should meet a diagonalization condition. In [12] and [14], the knowledge of channel matrices are used to determine the optimal relay matrix and power allocation for two-hop AF MIMO relays to maximize the mutual information between source and destination. [11] extends the work in [12] to the joint design of source matrix and relay matrix. In [114], capacity bounds and power allocation are derived with the assumption the channel matrices are known.

Despite the importance of channel matrices information for AF MIMO relays, there are limited discussions on channel estimation of two-hop AF MIMO relay systems. To the best of our knowledge, channel estimation for single-hop MIMO channel has been discussed in [23]-[25]. The single-hop MIMO channel estimation methods can be applied to MIMO relays if every pair of adjacent MIMO nodes can be treated as a pair of transmitter and receiver. However, AF MIMO relays are subject to limited signal processing functions and can not decode or estimate information. Therefore, AF MIMO relays may not be able to complete the task of channel estimation by following a single-hop MIMO channel

estimation approach. It is because of such a reason that researchers have started to explore channel estimation methods for MIMO relays. [26] and [27] addressed channel estimation for two-hop AF MIMO relays. In [26] and [27], source-relay channel matrix \mathbf{H}_1 and the relay-to-destination matrix \mathbf{H}_2 are estimated from the composite source-relay-destination matrix $\mathbf{H}_C = \mathbf{H}_2\mathbf{F}\mathbf{H}_1$ where \mathbf{F} is the relay transformation matrix. [26] proposes a least square (LS) based channel estimation method to obtain \mathbf{H}_1 and \mathbf{H}_2 respectively from $\mathbf{H}_2\mathbf{F}\mathbf{H}_1$. The method consists of a sequence of LS problems and requires a set of n_R different relay matrices $\mathbf{F}_i, i = 1, 2, \dots, n_R$ where n_R is the number of antennas at the relay. [27] studied sufficient and necessary conditions on \mathbf{F} to ensure a successful estimation of \mathbf{H}_1 and \mathbf{H}_2 from \mathbf{H}_c . The advantage of using \mathbf{H}_c to estimate \mathbf{H}_1 and \mathbf{H}_2 is that for channel estimation, the relay node does not need to do anything different from that for data transmission, and the destination node performs all the tasks needed for estimation of \mathbf{H}_1 and \mathbf{H}_2 . But a disadvantage of the above approach is that there is always a scalar ambiguity for the estimates of \mathbf{H}_1 and \mathbf{H}_2 , i.e., $\hat{\mathbf{H}}_1 = \alpha\mathbf{H}_1, \hat{\mathbf{H}}_2 = \frac{1}{\alpha}\mathbf{H}_2$. Furthermore, the choices of the training and relay matrices to minimize the MSE of channel estimation has not been optimized in above literatures yet.

In this chapter, we proposed an LMMSE based channel estimation method for two-hop AF MIMO relay channels in which the channel estimates are not subject to any ambiguity.

The proposed channel estimation algorithm includes two phases. In the first phase, the source node transmits no signal, but the relay transmits a training matrix \mathbf{S}_R and the destination estimates \mathbf{H}_2 . In the second phase, the source transmits a training matrix \mathbf{S}_S , the relay amplifies and forwards with the transformation matrix \mathbf{F} , and the destination

estimates \mathbf{H}_1 with the prior knowledge of \mathbf{H}_2 . In this scheme, there is no ambiguity in the channel estimates (i.e., the exact \mathbf{H}_1 and \mathbf{H}_2 are always found in the absence of noise). This scheme requires the relay to generate and transmit signals in the first phase. But no other advanced operation is required at the relay.

For phase 1, the channel estimation problem is the same as for a single-hop MIMO channel. The channel estimation method we used in phase 1 is essentially a special case of the one discussed in [25]. [25] discussed LMMSE channel estimation and training design for single-hop MIMO channel under Rician fading with both noise and interferences at the receiver. In [25], the received signal at the receiver is $\mathbf{Y} = \mathbf{H}\mathbf{P} + \mathbf{N}$. By considering Kronecker-structure system, the received signal becomes $vec(\mathbf{Y}) = (\mathbf{P}^T \otimes \mathbf{I})vec(\mathbf{H}) + vec(\mathbf{N})$ where vec forms a vector by stacking up all columns of the matrix and \otimes denotes the Kronecker product. [25] assumes that $vec(\mathbf{H}) \in \mathcal{CN}(vec(\bar{\mathbf{H}}), \mathbf{R})$ and $vec(\mathbf{N}) \in \mathcal{CN}(vec(\bar{\mathbf{N}}), \mathbf{S})$ where $\bar{\mathbf{H}}$ and $\bar{\mathbf{N}}$ are the mean of \mathbf{H} and \mathbf{N} while \mathbf{R} and \mathbf{S} are covariance matrix of $vec(\mathbf{H})$ and $vec(\mathbf{N})$ respectively. In our case, we consider correlated Rayleigh fading channel with noise only in which $\bar{\mathbf{H}} = \mathbf{0}$, $\bar{\mathbf{N}} = \mathbf{0}$ and $\mathbf{S} = \sigma_n^2 \mathbf{I}$. Although our solution for single-hop channel estimation falls into the framework of [25], we developed it independently and the solution can adapt to different optimization objectives. Furthermore, we provide another perspective to derive optimal training structure by using KKT conditions while [25] uses majorization theory to find the structure.

For phase 2, we assume the Kronecker channel correlation model for \mathbf{H}_1 , i.e., $\mathbf{H}_1 = \mathbf{C}_{r_1}^{\frac{1}{2}} \mathbf{W}_1 \mathbf{C}_{t_1}^{\frac{T}{2}}$ or equivalently $\mathbf{h}_1 = (\mathbf{C}_{t_1}^{\frac{1}{2}} \otimes \mathbf{C}_{r_1}^{\frac{1}{2}}) \mathbf{w}_1$ with $\mathbf{h}_1 = vec(\mathbf{H}_1)$ and $\mathbf{w}_1 = vec(\mathbf{W}_1)$. All elements in \mathbf{W}_1 are modelled as i.i.d. random variables with zero mean and unit variance.

The matrix \mathbf{C}_{r_1} is known as the receive correlation matrix of \mathbf{H}_1 , and \mathbf{C}_{t_1} the transmit correlation matrix of \mathbf{H}_1 . We will develop an algorithm for computing the optimal pair of \mathbf{S}_S and \mathbf{F} , which minimizes a cost of the covariance matrix of the LMMSE channel estimation errors in \mathbf{H}_1 . This problem is non-convex. But we will apply the majorization theory to determine the optimal structures of \mathbf{S}_S and \mathbf{F} so that the remaining problem is much simpler to solve numerically.

It is clear that \mathbf{H}_1 can also be estimated by the same technique as in phase 1 if the relay can perform advanced computations. The scheme shown above for phase 2 provides an alternative approach, which reduces the burden on the relay.

To sum up, the contributions of this chapter are as follows:

- A two-phase LMMSE estimation method for two-hop AF MIMO relay channel is proposed to minimize the channel estimation MSE subject to both power energy constraints at the source and the relay. The method results in exact channel estimates without any ambiguity.
- Optimal structures of training and relay matrices are derived by using convex optimization and majorization theory. Power allocation along the diagonal of training and relay matrices is solved by using an alternating algorithm with low complexity and fast convergence.
- The two-phase LMMSE based channel estimation method for two-hop AF MIMO relay channels can be extended to multihop AF MIMO relay channel estimation.

A comparison of our proposed LMMSE channel estimation method with the meth-

ods in [26] and [27] is shown in Table 4.1. In the table, “Y” stands for “Yes” and “N” stands for “No”. From the following table, we see that [26] and [27] both require substantial system resources. For example, [26] requires $2n_R$ time slots while [27] requires at least 6 time slots to complete channel estimation.

Table 4.1: Comparison of Channel Estimation Methods for Two-hop AF MIMO Relay Channels

criteria	proposed method	LS [26]	Interim [27]
required time slots	4	$2n_R$	6
power energy constraint at relay	Y	N	N
training matrix design	Y	N	N
relay matrix requirement	any	diagonal	unitary
estimation ambiguity	N	Y	Y

The rest of the chapter is organized as follows. The system model is given in Section II. In Section III we present a general LMMSE based single-hop channel estimation method. And the details of the two-phase LMMSE channel estimation method and training design are discussed in IV. Some numerical results are demonstrated in section V. Finally, some concluding remarks are given in VI.

4.2 System Model

A system model of two-hop MIMO relay system is shown in Fig.4.1. There are three nodes in the system: source node S , relay node R and destination node D , all of which are half-duplex and equipped with multiple antennas.

Here, we consider a single narrow-band subcarrier for which all channels are flat fading. The number of antennas at S , R and D are n_S , n_R and n_D , respectively. The MIMO channel between S and R is \mathbf{H}_1 and the one between R and D is \mathbf{H}_2 . In general, \mathbf{H}_1 and

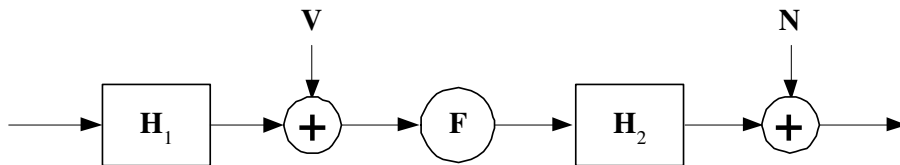


Figure 4.1: A two-hop AF MIMO relay system

\mathbf{H}_2 can be correlated and are modeled as

$$\mathbf{H}_i = \mathbf{C}_{r_i}^{\frac{1}{2}} \mathbf{W}_i \mathbf{C}_{t_i}^{\frac{T}{2}}, \quad i = 1, 2 \quad (4.1)$$

where $\mathbf{C}_{t_i} = \mathbf{C}_{t_i}^{\frac{1}{2}} \mathbf{C}_{t_i}^{\frac{H}{2}}$, $\mathbf{C}_{r_i} = \mathbf{C}_{r_i}^{\frac{1}{2}} \mathbf{C}_{r_i}^{\frac{H}{2}}$, $i = 1, 2$. All elements in \mathbf{W}_i are modeled as i.i.d. random variables with zero mean and unit variance. The matrix \mathbf{C}_{r_i} is known as the receive correlation matrix of \mathbf{H}_i , and \mathbf{C}_{t_i} the transmit correlation matrix of \mathbf{H}_i .

\mathbf{F} is the $n_R \times n_R$ relay transformation matrix at R . \mathbf{S}_S is $n_S \times L$ training matrix at S and \mathbf{S}_R is the $n_R \times L$ training matrix at R . The noise at both relay node and destination node are assumed to be complex Gaussian, zero-mean with unit variance. The noise matrix at relay and at the destination are $\mathbf{V} \in \mathbb{C}^{n_R \times L}$ and $\mathbf{N} \in \mathbb{C}^{n_D \times L}$ respectively.

Our channel estimation method includes two phases:

- Phase 1: the source node S transmits no signal, the relay R transmits a training matrix \mathbf{S}_R and the destination D estimates \mathbf{H}_2 .
- Phase 2: the source S transmits a training matrix \mathbf{S}_S , the relay R amplifies and forwards with the transformation matrix \mathbf{F} , and the destination D estimates \mathbf{H}_1 with the prior knowledge of \mathbf{H}_2 .

It is assumed that \mathbf{H}_1 , \mathbf{H}_2 keep constant during the two-phase estimation process. In the

first phase, the received signal at the destination is

$$\mathbf{Y}_D^{(1)} = \mathbf{H}_2 \mathbf{S}_R + \mathbf{N}^{(1)} \quad (4.2)$$

In the second phase, the received signal at the relay is

$$\mathbf{Y}_R = \mathbf{H}_1 \mathbf{S}_S + \mathbf{V} \quad (4.3)$$

Either concurrently in full-duplex mode or with L-slot delay in half-duplex mode, the relay transmits $\mathbf{X}_R = \mathbf{F} \mathbf{Y}_R$. The received signal at the destination is

$$\mathbf{Y}_D^{(2)} = \mathbf{H}_2 \mathbf{F} \mathbf{H}_1 \mathbf{S}_S + \mathbf{H}_2 \mathbf{F} \mathbf{V} + \mathbf{N}^{(2)} \quad (4.4)$$

Power constraints at both source and relay are considered. The maximum power consumed at the source and the relay are respectively P_S and P_R . In the first phase, the energy constraint at the relay is

$$\text{tr}(\mathbf{S}_R \mathbf{S}_R^H) \leq P_R \quad (4.5)$$

In the second phase, the energy constraint at the source is

$$\text{tr}(\mathbf{S}_S \mathbf{S}_S^H) \leq P_S \quad (4.6)$$

and an averaged energy constraint is imposed on relay

$$E \{ \text{tr} [\mathbf{X}_R \mathbf{X}_R^H] \} \leq P_R \quad (4.7)$$

where \mathbf{X}_R is the transmitted signal from relay and E is the expectation over the distribution of \mathbf{H}_1 since \mathbf{H}_1 is unknown. The value of P_R in (4.7) can be different from that in (4.5).

4.3 Channel Estimation and Training Design for Phase 1

In the following section, we will briefly introduce our channel estimation algorithm and training design for phase 1. Although the solution is a special case of [25], the optimal training matrix structure is developed independently from the perspective of solving a convex optimization problem.

In phase 1, the relay transmits $\mathbf{S}_R \in C^{m_R \times L}$ over L time slots subject to the total energy constraint in (4.5).

Recall the fact $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$ where \otimes denotes the Kronecker product [115]. We will also use frequently $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ and $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$. With $\mathbf{H}_2 = \mathbf{C}_{r_2}^{\frac{1}{2}} \mathbf{W}_2 \mathbf{C}_{t_2}^{\frac{T}{2}}$, we apply Kronecker product in [115] to (4.2) and have

$$\mathbf{y}_D^1 = \left(\mathbf{S}_R^T \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{1}{2}} \right) \mathbf{w}_2 + \mathbf{n}^{(1)} \quad (4.8)$$

where $\mathbf{y}_D^{(1)} = \text{vec}(\mathbf{Y}_D^{(1)})$, $\mathbf{w}_2 = \text{vec}(\mathbf{w}_2)$ and $\mathbf{n}^{(1)} = \text{vec}(\mathbf{N}^{(1)})$. We consider the LMMSE estimator of \mathbf{w}_2 from $\mathbf{y}_D^{(1)}$, i.e., $\hat{\mathbf{w}}_2 = \mathbf{T} \mathbf{y}_D^{(1)}$ where \mathbf{T} is such that the following is minimized:

$$J_2 = E \left\{ \text{tr} \left[\mathbf{C}_0 (\mathbf{w}_2 - \hat{\mathbf{w}}_2) (\mathbf{w}_2 - \hat{\mathbf{w}}_2)^H \right] \right\} \quad (4.9)$$

We will be interested in the following two choices of \mathbf{C}_0 :

$$\text{If } \mathbf{C}_0 = \mathbf{I}, J_2 = E(\|\mathbf{w}_2 - \hat{\mathbf{w}}_2\|^2).$$

$$\text{If } \mathbf{C}_0 = \mathbf{C}_{t_2}^{\frac{H}{2}} \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{H}{2}} \mathbf{C}_{r_2}^{\frac{1}{2}}, J_2 = E(\|\mathbf{h}_2 - \hat{\mathbf{h}}_2\|^2).$$

It is useful to note that as long as \mathbf{C}_0 is positive definite, \mathbf{C}_0 does not affect the optimal \mathbf{T} which is given by

$$\mathbf{T} = \mathbf{R}_{\mathbf{w}_2 \mathbf{y}_D^{(1)H}} \mathbf{R}_{\mathbf{y}_D^{(1)} \mathbf{y}_D^{(1)H}}^{-1} \quad (4.10)$$

where $\mathbf{R}_{\mathbf{w}_2 \mathbf{y}_D^{(1)H}} = E \left[\mathbf{w}_2 \mathbf{y}_D^{(1)H} \right]$ and $\mathbf{R}_{\mathbf{y}_D^{(1)} \mathbf{y}_D^{(1)H}} = E \left[\mathbf{y}_D^{(1)} \mathbf{y}_D^{(1)H} \right]$.

By substituting $\hat{\mathbf{w}}_2 = \mathbf{T} \mathbf{y}_D^{(1)}$, the covariance matrix of the estimation error $\delta \mathbf{h}_2 = \mathbf{h}_2 - \hat{\mathbf{h}}_2$ is

$$\begin{aligned}
\mathbf{R}_{\delta \mathbf{h}_2, \delta \mathbf{h}_2} &= E \left\{ \mathbf{C}_0 (\mathbf{w}_2 - \hat{\mathbf{w}}_2) (\mathbf{w}_2 - \hat{\mathbf{w}}_2)^H \right\} \\
&= \mathbf{C}_0 \left(\mathbf{I} - \mathbf{R}_{\mathbf{w}_2 \mathbf{y}_D^{(1)}} \mathbf{R}_{\mathbf{y}_D^{(1)} \mathbf{y}_D^{(1)H}}^{-1} \mathbf{R}_{\mathbf{y}_D^{(1)} \mathbf{w}_2}^H \right) \\
&= \mathbf{C}_0 \left[\mathbf{I} + \mathbf{C}_{t_2}^{\frac{H}{2}} \mathbf{S}_R^* \otimes \mathbf{C}_{r_2}^{\frac{H}{2}} \left(\mathbf{I} + \mathbf{S}_R^T \mathbf{C}_{t_2} \mathbf{S}_R^* \otimes \mathbf{C}_{r_2} \right)^{-1} \mathbf{S}_R^T \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{1}{2}} \right] \\
&= \mathbf{C}_0 \left(\mathbf{I} + \mathbf{C}_{t_2}^{\frac{H}{2}} \mathbf{S}_R^* \mathbf{S}_R^T \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{H}{2}} \mathbf{C}_{r_2}^{\frac{1}{2}} \right)^{-1} \\
&= \mathbf{C}_0 \left(\mathbf{I} + \mathbf{C}_{t_2}^{\frac{H}{2}} \mathbf{C}_S^R \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{H}{2}} \mathbf{C}_{r_2}^{\frac{1}{2}} \right)^{-1}
\end{aligned} \tag{4.11}$$

where $\mathbf{C}_S^R = \mathbf{S}_R^* \mathbf{S}_R^T$ and we apply $(\mathbf{I} + \mathbf{A}\mathbf{B})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I} + \mathbf{B}\mathbf{A})^{-1}\mathbf{B}$ to derive the last two steps.

It is easy to verify that with the optimal \mathbf{T} ,

$$J_2 = J'_2 \doteq \text{tr} \left(\mathbf{C}_0 \left(\mathbf{I} + \mathbf{C}_{t_2}^{\frac{H}{2}} \mathbf{C}_{S_R} \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{H}{2}} \mathbf{C}_{r_2}^{\frac{1}{2}} \right)^{-1} \right) \tag{4.12}$$

If \mathbf{T} is arbitrary, $J_2 \geq J'_2$.

Once \mathbf{C}_{S_R} is fixed, J'_2 is invariant to \mathbf{S}_R . So, to find the optimal relay training matrix \mathbf{S}_R for phase 1, it suffices to find \mathbf{C}_{S_R} by solving the following problem:

$$\begin{aligned}
\min_{\mathbf{C}_{S_R} \geq \mathbf{0}} \quad & J'_2 \\
s.t. \quad & \text{tr}(\mathbf{C}_{S_R}) \leq P_R
\end{aligned} \tag{4.13}$$

where $\mathbf{C}_{S_R} \geq \mathbf{0}$ is the positive semidefinite constraint on \mathbf{C}_{S_R} , and P_R is the power bound at the relay.

The problem (4.13) is convex. In Appendix A, we apply the generalized KKT conditions [117] to arrive at the following optimal solution: $\mathbf{C}_{S_R} = \mathbf{U}_{t_2} \mathbf{C} \mathbf{U}_{t_2}^H$ where \mathbf{U}_{t_2} is the unitary eigenvector matrix of \mathbf{C}_{t_2} and \mathbf{C} is a diagonal matrix with its diagonal elements $c(j)$, $j = 1, 2, \dots, n_R$, either equal to zero or given by the positive solution of the following equations:

$$\lambda_1(j) \lambda_{t_2}(j) \sum_{i=1}^{n_R} \frac{\lambda_2(i) \lambda_{r_2}(i)}{[1 + \lambda_{r_2}(i) \lambda_{t_2}(j) c(j)]^2} = \mu, \quad j = 1, 2, \dots, n_R \quad (4.14)$$

where $\mu > 0$ is such that $\text{tr}(\mathbf{C}) = P_R$. Recall $\mathbf{C}_0 = \mathbf{C}_1 \otimes \mathbf{C}_2$. Here, $\lambda_2(i)$ and $\lambda_{r_2}(i)$ are the i th largest eigenvalue of \mathbf{C}_2 and \mathbf{C}_{r_2} . $\lambda_1(j)$ and $\lambda_{t_2}(j)$ are the j th largest eigenvalue of \mathbf{C}_1 and \mathbf{C}_{t_2} .

For any given $\mu > 0$, for each j , $c(j)$ can be easily found by the bi-section search [119] since the left side expression of (4.14) is a monotonically decreasing function of $c(j)$. Consequently, $\text{tr}(\mathbf{C}_{S_R}) = \text{tr}(\mathbf{C}) = \sum_{j=1}^{n_R} c(j)$ is a monotonically decreasing function of μ , and hence the optimal μ can be found by an outer-loop bi-section search.

The above solution for \mathbf{C}_{S_R} is similar to one in [25] although a different method of derivation was used in [25].

4.4 Channel Estimation and Training Design for Phase 2

In the following section, we will discuss the channel estimation and training design for two-hop channel estimation phase 2. An optimization problem to minimize the MSE of channel estimation is formulated. Optimal source training and relay matrices structures are derived using majorization theory. With the derived structures, the power allocation along

the diagonal of training and relay matrices is greatly simplified. Different power allocation solutions are proposed for different channel correlation situations.

4.4.1 Channel estimation

In phase 2, the source transmits $\mathbf{S}_S \in C^{m_S \times L}$ over L time slots subject to the energy constraint (4.6). Applying the vec operator to \mathbf{X}_R , we have $\mathbf{x}_R = vec(\mathbf{X}_R) = (\mathbf{S}_S^T \otimes \mathbf{F})vec(\mathbf{H}_1) + (\mathbf{I} \otimes \mathbf{F})vec(\mathbf{V})$. Also recall $\mathbf{h}_1 = vec(\mathbf{H}_1) = (\mathbf{C}_{t_1}^{1/2} \otimes \mathbf{C}_{r_1}^{1/2})vec(\mathbf{W}_1)$. Then, it is easy to verify that (4.7) becomes

$$tr(\mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H + \mathbf{I} \otimes \mathbf{F} \mathbf{F}^H) \leq P_R \quad (4.15)$$

Applying vec operator to (4.4), the vector form of $\mathbf{Y}_D^{(2)}$ is

$$\mathbf{y}_D^{(2)} = (\mathbf{S}_S^T \otimes \mathbf{H}_2 \mathbf{F}) \mathbf{h}_1 + (\mathbf{I} \otimes \mathbf{H}_2 \mathbf{F}) \mathbf{v} + \mathbf{n}^{(2)} \quad (4.16)$$

Unlike the discussion in Section 4.3, we now only focus on the mean squared errors of $\hat{\mathbf{h}}_1 = vec(\hat{\mathbf{H}}_1)$. The choice of the mean squared errors of $\hat{\mathbf{w}}_1 = vec(\hat{\mathbf{W}}_1)$ would make the optimal training design more difficult, which will not be further mentioned. Namely, we define the cost

$$J_1 \doteq E\{tr[(\mathbf{h}_1 - \hat{\mathbf{h}}_1)(\mathbf{h}_1 - \hat{\mathbf{h}}_1)^H]\} \quad (4.17)$$

without any other weighting.

The LMMSE estimation of \mathbf{h}_1 is given by $\hat{\mathbf{h}}_1 = \mathbf{R}_{\mathbf{h}_1, \mathbf{y}_D^{(2)}} \mathbf{R}_{\mathbf{y}_D^{(2)}, \mathbf{y}_D^{(2)}}^{-1} \mathbf{y}_D^{(2)}$. The covariance matrix of the estimation error $\delta \mathbf{h}_1 = \mathbf{h}_1 - \hat{\mathbf{h}}_1$ is well known as

$$\mathbf{R}_{\delta \mathbf{h}_1, \delta \mathbf{h}_1} = \mathbf{R}_{\mathbf{h}_1, \mathbf{h}_1} - \mathbf{R}_{\mathbf{h}_1, \mathbf{y}_D^{(2)}} \mathbf{R}_{\mathbf{y}_D^{(2)}, \mathbf{y}_D^{(2)}}^{-1} \mathbf{R}_{\mathbf{h}_1, \mathbf{y}_D^{(2)}}^H \quad (4.18)$$

where $\mathbf{R}_{\mathbf{h}_1, \mathbf{h}_1} = E \{ \mathbf{h}_1 \mathbf{h}_1^H \} = \mathbf{C}_{t_1} \otimes \mathbf{C}_{r_1}$. Therefore

$$\mathbf{R}_{\mathbf{h}_1 \mathbf{y}_D^{(2)}} = E \left[\mathbf{h}_1 \mathbf{y}_D^{(2)H} \right] = (\mathbf{C}_{t_1} \otimes \mathbf{C}_{r_1}) (\mathbf{S}_S^T \otimes \mathbf{H}_2 \mathbf{F})^H = \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H \quad (4.19)$$

and

$$\mathbf{R}_{\mathbf{y}_D^{(2)} \mathbf{y}_D^{(2)}} = E \left[\mathbf{y}_D^{(2)} \mathbf{y}_D^{(2)H} \right] = \mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I} \otimes \mathbf{H}_2 \mathbf{F} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I} \quad (4.20)$$

Therefore, with the LMMSE of \mathbf{h}_1 and (4.19), (4.20), we have

$$\begin{aligned} J_1 &= J_1' & (4.21) \\ &\doteq tr \left\{ \mathbf{R}_{\delta \mathbf{h}_1, \delta \mathbf{h}_1^H} \right\} \\ &= tr [\mathbf{C}_{t_1} \otimes \mathbf{C}_{r_1}] - tr \left[(\mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H) \right. \\ &\quad \left. \cdot [\mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I} \otimes \mathbf{H}_2 \mathbf{F} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I}]^{-1} (\mathbf{S}_S^T \mathbf{C}_{t_1}^H \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1}^H) \right] \end{aligned}$$

In other words, with any other linear estimator of \mathbf{h}_1 , $J_1 \geq J_1'$.

4.4.2 Training Design Problem

The optimal training and relay matrices design problem for phase 2 is formulated as

$$\begin{aligned} &\min_{\mathbf{S}_S, \mathbf{F}} J_1' & (4.22) \\ &s.t. \quad tr(\mathbf{S}_S \mathbf{S}_S^H) \leq P_S \\ &\quad \quad tr(\mathbf{R}_{\mathbf{x}_R, \mathbf{x}_R}) \leq P_R \end{aligned}$$

where $\mathbf{R}_{\mathbf{x}_R, \mathbf{x}_R} = E \{ \mathbf{x}_R \mathbf{x}_R^H \} = \mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H + \mathbf{I} \otimes \mathbf{F} \mathbf{F}^H$ as in (4.15).

With (4.15) and (4.21), the optimization problem in (4.22) can be written as

$$\begin{aligned}
& \max_{\mathbf{S}_S, \mathbf{F}} \quad tr \left\{ [\mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I} \otimes \mathbf{H}_2 \mathbf{F} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I}]^{-1} \right. \\
& \quad \left. (\mathbf{S}_S^T \mathbf{C}_{t_1}^H \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1}^H \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H) \right\} \\
s.t. \quad & tr \{ \mathbf{S}_S^T \mathbf{S}_S^* \} \leq P_S \tag{4.23} \\
& tr \{ \mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H + \mathbf{I} \otimes \mathbf{F} \mathbf{F}^H \} \leq P_R
\end{aligned}$$

4.4.3 Decomposition of Trainings

In this section, we show a decomposition of Trainings into two sets of components: unitary components and diagonal components.

Denote the eigenvalue decompositions (EVD) of $\mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^*$ and $\mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H$, respectively, as

$$\mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* = \mathbf{U}_S \mathbf{\Lambda}_S \mathbf{U}_S^H \tag{4.24}$$

$$\mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H = \mathbf{U}_F \mathbf{\Lambda}_F \mathbf{U}_F^H \tag{4.25}$$

where the \mathbf{U} matrices are the unitary eigenvector matrices and the $\mathbf{\Lambda}$ matrices are the diagonal eigenvalue matrices with descending diagonal elements.

Also let $\mathbf{C}_{t_1} = \mathbf{U}_{t_1} \mathbf{\Lambda}_{t_1} \mathbf{U}_{t_1}^H$ and $\mathbf{C}_{r_1} = \mathbf{U}_{r_1} \mathbf{\Lambda}_{r_1} \mathbf{U}_{r_1}^H$ be the EVDs of \mathbf{C}_{t_1} and \mathbf{C}_{r_1} , respectively, with descending eigenvalues. Define $\mathbf{C}_{t_1}^{1/2} = \mathbf{U}_{t_1} \mathbf{\Lambda}_{t_1}^{1/2}$ and $\mathbf{C}_{r_1}^{1/2} = \mathbf{U}_{r_1} \mathbf{\Lambda}_{r_1}^{1/2}$.

Then, we can write

$$\mathbf{S}_S^T \mathbf{C}_{t_1}^{1/2} = \mathbf{U}_S \mathbf{\Lambda}_S^{1/2} \mathbf{Q}_S \tag{4.26}$$

$$\mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1}^{1/2} = \mathbf{U}_F \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \tag{4.27}$$

where \mathbf{Q}_S and \mathbf{Q}_F are unitary matrices.

It is important to note here that if \mathbf{C}_{t_1} , \mathbf{C}_{r_1} and \mathbf{H}_2 are nonsingular (which is assumed unless specified otherwise), the training matrices \mathbf{S}_S and \mathbf{F} are uniquely determined by the unitary components: \mathbf{U}_S , \mathbf{U}_F , \mathbf{Q}_S , \mathbf{Q}_F and the diagonal components: $\mathbf{\Lambda}_S$, $\mathbf{\Lambda}_F$. Namely, $\mathbf{S}_S^T = \mathbf{U}_S \mathbf{\Lambda}_S^{1/2} \mathbf{Q}_S \mathbf{C}_{t_1}^{-1/2}$ and $\mathbf{F} = \mathbf{H}_2^{-1} \mathbf{U}_F \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{C}_{r_1}^{-1/2}$.

It then follows from (4.17) that

$$\begin{aligned}
& J'_1 - tr(\mathbf{C}_{t_1} \otimes \mathbf{C}_{r_1}) \\
&= -tr \left\{ [\mathbf{S}_S^T \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I} \otimes \mathbf{H}_2 \mathbf{F} \mathbf{F}^H \mathbf{H}_2^H + \mathbf{I}]^{-1} \right. \\
&\quad \left. (\mathbf{S}_S^T \mathbf{C}_{t_1}^H \mathbf{C}_{t_1} \mathbf{S}_S^* \otimes \mathbf{H}_2 \mathbf{F} \mathbf{C}_{r_1}^H \mathbf{C}_{r_1} \mathbf{F}^H \mathbf{H}_2^H) \right\} \\
&= -tr \left\{ \left[\mathbf{U}_S \mathbf{\Lambda}_S \mathbf{U}_S^H \otimes \mathbf{U}_F \mathbf{\Lambda}_F \mathbf{U}_F^H + \mathbf{U}_S \mathbf{U}_S^H \otimes \mathbf{U}_F \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{C}_{r_1}^{-1/2} \mathbf{C}_{r_1}^{-H/2} \mathbf{Q}_F^H \mathbf{\Lambda}_F^{H/2} \mathbf{U}_F^H + \mathbf{I} \right]^{-1} \right. \\
&\quad \left. \left(\mathbf{U}_S \mathbf{\Lambda}_S^{1/2} \mathbf{Q}_S \mathbf{C}_{t_1}^{-1/2} \mathbf{C}_{t_1}^H \mathbf{C}_{t_1} \mathbf{C}_{t_1}^{-H/2} \mathbf{Q}_S^H \mathbf{\Lambda}_S^{H/2} \mathbf{U}_S^H \otimes \mathbf{U}_F \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{C}_{r_1}^{-1/2} \mathbf{C}_{r_1}^H \mathbf{C}_{r_1} \mathbf{C}_{r_1}^{-H/2} \mathbf{Q}_F^H \mathbf{\Lambda}_F^{H/2} \mathbf{U}_F^H \right) \right\} \\
&= -tr \left\{ \left[\mathbf{U}_S \mathbf{\Lambda}_S \mathbf{U}_S^H \otimes \mathbf{U}_F \mathbf{\Lambda}_F \mathbf{U}_F^H + \mathbf{U}_S \mathbf{U}_S^H \otimes \mathbf{U}_F \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{\Lambda}_{r_1}^{-1} \mathbf{Q}_F^H \mathbf{\Lambda}_F^{H/2} \mathbf{U}_F^H + \mathbf{I} \right]^{-1} \right. \\
&\quad \left. \left(\mathbf{U}_S \mathbf{\Lambda}_S^{1/2} \mathbf{Q}_S \mathbf{\Lambda}_{t_1} \mathbf{Q}_S^H \mathbf{\Lambda}_S^{H/2} \mathbf{U}_S^H \otimes \mathbf{U}_F \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{\Lambda}_{r_1} \mathbf{Q}_F^H \mathbf{\Lambda}_F^{H/2} \mathbf{U}_F^H \right) \right\} \\
&= -tr \left\{ \left(\mathbf{\Lambda}_S \otimes \mathbf{\Lambda}_F + \mathbf{I} \otimes \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{\Lambda}_{r_1}^{-1} \mathbf{Q}_F^H \mathbf{\Lambda}_F^{H/2} + \mathbf{I} \right)^{-1} \right. \\
&\quad \left. \left(\mathbf{\Lambda}_S^{1/2} \mathbf{Q}_S \mathbf{\Lambda}_{t_1} \mathbf{Q}_S^H \mathbf{\Lambda}_S^{H/2} \otimes \mathbf{\Lambda}_F^{1/2} \mathbf{Q}_F \mathbf{\Lambda}_{r_1} \mathbf{Q}_F^H \mathbf{\Lambda}_F^{H/2} \right) \right\} \tag{4.28}
\end{aligned}$$

We can see from the above equation that the cost J'_1 is invariant to \mathbf{U}_S and \mathbf{U}_F but depends on $\mathbf{\Lambda}_S$, $\mathbf{\Lambda}_F$, \mathbf{Q}_S and \mathbf{Q}_F .

It is easy to verify that the energy constraint at the source can now be written as

$$tr \left\{ \mathbf{\Lambda}_S \mathbf{Q}_S \mathbf{\Lambda}_{t_1}^{-1} \mathbf{Q}_S^H \right\} \leq P_S \tag{4.29}$$

which depends on $\mathbf{\Lambda}_S$ and \mathbf{Q}_S , and is invariant to all other components of the trainings.

To simplify the energy constraint at the relay, we denote the singular value decomposition (SVD): $\mathbf{H}_2 = \mathbf{U}_{H_2} \boldsymbol{\Sigma}_{H_2} \mathbf{V}_{H_2}^H$, with descending singular values, where \mathbf{U}_{H_2} and \mathbf{V}_{H_2} are (square) unitary singular vector matrices. Note that we will use $\boldsymbol{\Sigma}_{H_2}^2 = \boldsymbol{\Sigma}_{H_2} \boldsymbol{\Sigma}_{H_2}^H$ in the case where \mathbf{H}_2 is non-square. Then, using (4.26), (4.27) and $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$, one can verify that the relay energy constraint (4.7) can be rewritten as

$$\begin{aligned} & \text{tr} \{ \boldsymbol{\Lambda}_S \} \text{tr} \left\{ \boldsymbol{\Sigma}_{H_2}^{-2} \mathbf{U}_{H_2}^H \mathbf{U}_F \boldsymbol{\Lambda}_F \mathbf{U}_F^H \mathbf{U}_{H_2} \right\} \\ & + L \cdot \text{tr} \left\{ \boldsymbol{\Sigma}_{H_2}^{-2} \mathbf{U}_{H_2}^H \mathbf{U}_F \boldsymbol{\Lambda}_F^{\frac{1}{2}} \mathbf{Q}_F \boldsymbol{\Lambda}_{r_1}^{-1} \mathbf{Q}_F^H \boldsymbol{\Lambda}_F^{\frac{H}{2}} \mathbf{U}_F^H \mathbf{U}_{H_2} \right\} \leq P_R \end{aligned} \quad (4.30)$$

which depends on $\boldsymbol{\Lambda}_S$, $\boldsymbol{\Lambda}_F$, \mathbf{U}_F and \mathbf{Q}_F , and is invariant to \mathbf{U}_S and \mathbf{Q}_S .

In the following two sub-sections, we will show how to identify the optimal unitary components and the optimal diagonal components, respectively.

4.4.4 Optimal Unitary Components of the Trainings

Among the training components, we have the unitary (matrix) components \mathbf{U}_S , \mathbf{Q}_S , \mathbf{U}_F and \mathbf{Q}_F , and the diagonal (matrix) components $\boldsymbol{\Lambda}_S$ and $\boldsymbol{\Lambda}_F$. The optimality of the choices of the unitary components are given by following two theorems.

Theorem 1: For any \mathbf{C}_{r_1} and \mathbf{C}_{t_1} , the solution to the problem (4.22) is such that $\mathbf{Q}_S = \mathbf{I}$ and \mathbf{U}_S is arbitrary unitary.

Proof: See Appendix B.

Theorem 2: If $\mathbf{C}_{r_1} = \alpha \mathbf{I}$, the solution to (4.22) is such that $\mathbf{U}_F = \mathbf{U}_{H_2}$ and \mathbf{Q}_F is arbitrary unitary.

Proof: See Appendix C.

Namely, if $\mathbf{C}_{r_1} = \alpha \mathbf{I}$, then we can choose $\mathbf{U}_S = \mathbf{I}$ and $\mathbf{Q}_F = \mathbf{I}$ as optimal, and the optimal \mathbf{S}_S and \mathbf{F} have the following structures:

$$\mathbf{S}_S^T = \Lambda_S^{1/2} \mathbf{C}_{t_1}^{-1/2} \quad (4.31)$$

$$\mathbf{F} = \frac{1}{\sqrt{\alpha}} \mathbf{V}_{H_2} \Sigma_{H_2}^{-1} \Lambda_F^{1/2} \quad (4.32)$$

If $\mathbf{C}_{r_1} \neq \alpha \mathbf{I}$, finding the optimal \mathbf{Q}_F (which also affects the optimal \mathbf{U}_F) is a difficult problem.

In the following two sections, we will propose algorithms to find Λ_S and Λ_F for the general case as well as for the special case where $\mathbf{C}_{t_1} = \beta \mathbf{I}$, $\mathbf{C}_{r_1} = \alpha \mathbf{I}$.

4.4.5 Optimal Diagonal Components of the Trainings: General Case

In this section, we apply $\mathbf{Q}_S = \mathbf{I}$, $\mathbf{Q}_F = \mathbf{I}$, $\mathbf{U}_F = \mathbf{U}_{H_2}$ and $\mathbf{U}_S = \mathbf{I}$ to develop efficient algorithms for finding the optimal Λ_S and Λ_F . The above choice of the unitary components is optimal if \mathbf{C}_{r_1} is proportional to the identity matrix, and has no known optimality property otherwise.

Then, the training design problem (4.22) becomes

$$\begin{aligned} \min_{\Lambda_S \geq 0, \Lambda_F \geq 0} & \quad -tr \left\{ \left(\Lambda_S \otimes \Lambda_F + \mathbf{I} \otimes \Lambda_F^{1/2} \Lambda_{r_1}^{-1} \Lambda_F^{H/2} + \mathbf{I} \right)^{-1} \left(\Lambda_S^{1/2} \Lambda_{t_1} \Lambda_S^{H/2} \otimes \Lambda_F^{1/2} \Lambda_{r_1} \Lambda_F^{H/2} \right) \right\} \\ \text{s.t.} & \quad tr \{ \Lambda_S \Lambda_{t_1}^{-1} \} \leq P_S \\ & \quad tr \{ \Lambda_S \} tr \{ \Sigma_{H_2}^{-2} \Lambda_F \} + L tr \left\{ \Sigma_{H_2}^{-2} \Lambda_F^{1/2} \Lambda_{r_1}^{-1} \Lambda_F^{H/2} \right\} \leq P_R \end{aligned} \quad (4.33)$$

Denote $\lambda_S(i)$, $\lambda_F(i)$, $\lambda_{t_1}(i)$, $\lambda_{r_1}(i)$ and $\sigma_{H_2}(i)$ as the i th diagonal element of Λ_S , Λ_F , Λ_{t_1} , Λ_{r_1}

and Σ_{H_2} , respectively. The problem (4.33) can be further written as

$$\begin{aligned}
& \min_{\{\lambda_S(i) \geq 0, \forall i\}, \{\lambda_F(j) \geq 0, \forall j\}} - \sum_{i=1}^L \sum_{j=1}^{n_D} \frac{\lambda_S(i) \lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j)}{\lambda_S(i) \lambda_F(j) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1} \quad (4.34) \\
s.t. \quad & \sum_{i=1}^L \lambda_{t_1}^{-1}(i) \lambda_S(i) \leq P_S \\
& \left(\sum_{i=1}^L \lambda_S(i) \right) \left(\sum_{j=1}^{n_D} \sigma_{H_2}^{-2}(j) \lambda_F(j) \right) + L \sum_{j=1}^{n_D} \sigma_{H_2}^{-2}(j) \lambda_{r_1}^{-1}(j) \lambda_F(j) \leq P_R
\end{aligned}$$

Let $\bar{n}_S = \text{rank}(\mathbf{C}_{t_1})$ and $\bar{n}_F = \min(\text{rank}(\mathbf{C}_{r_1}), \text{rank}(\mathbf{H}_2))$. Obviously, for $i > \bar{n}_S$, $\lambda_{t_1}^{-1}(i) = \infty$, and for $j > \bar{n}_F$, $\sigma_{H_2}^{-2}(j) \lambda_{r_1}^{-1}(j) = \infty$. From the energy constraints in (4.34), it is easy to see that the solution to (4.34) must be such that $\lambda_S(i) = 0$ for $L \geq i > \bar{n}_S$ and $\lambda_F(j) = 0$ for $n_D \geq j > \bar{n}_F$. So, we can replace L and n_D in (4.34) by \bar{n}_S and \bar{n}_F , respectively. Also note that for the problem (4.34), we do not require the non-singularity condition on \mathbf{C}_{t_1} , \mathbf{C}_{r_1} and \mathbf{H}_2 . It also means that the inverse in (4.31) and (4.32) should be replaced by pseudoinverse in the case where \mathbf{H}_2 and/or \mathbf{C}_{t_1} is singular.

The problem (4.34) is non-convex. But if we fix $\lambda_F(j)$ for all j , the optimization over $\lambda_S(i)$ for all i is convex. Similarly, if we fix $\lambda_S(i)$ for all i , the optimization over $\lambda_F(j)$ for all j is also convex. By alternating between the two sub-optimizations, we can find a local optimal solution to (4.34).

The algorithms of the two sub-optimizations are shown next.

1. Optimizing $\{\lambda_F(j)\}$ with fixed $\{\lambda_S(i)\}$

We use a group of $\lambda_S(i), i = 1, \dots, \bar{n}_S$ which satisfies $\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} = P_S$. Therefore,

the original problem in (4.34) is transformed as

$$\begin{aligned} \min_{\{\lambda_F(j)\}} & - \sum_{i=1}^{\bar{n}_S} \sum_{j=1}^{\bar{n}_F} \frac{\lambda_S(i) \lambda_{t_1}(i) \lambda_{r_1}(j) \lambda_F(j)}{[\lambda_S(i) + \lambda_{r_1}^{-1}(j)] \lambda_F(j) + 1} \\ \text{s.t.} & \sum_{i=1}^{\bar{n}_S} \lambda_S(i) \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} \leq P_R \end{aligned} \quad (4.35)$$

In the following, we will use Lagrange multiplier method to solve above problem.

The Lagrange function is

$$\begin{aligned} L^F = & - \sum_{i=1}^{\bar{n}_S} \sum_{j=1}^{\bar{n}_F} \frac{\lambda_S(i) \lambda_{t_1}(i) \lambda_{r_1}(j) \lambda_F(j)}{[\lambda_S(i) + \lambda_{r_1}^{-1}(j)] \lambda_F(j) + 1} \\ & + \mu \left[\sum_{i=1}^{\bar{n}_S} \lambda_S(i) \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} - P_R \right] \end{aligned} \quad (4.36)$$

The derivative of L^F with respect to $\lambda_F(j)$ is

$$\frac{\partial L^F}{\partial \lambda_F(j)} = - \sum_{i=1}^{\bar{n}_S} \frac{\lambda_{t_1}(i) \lambda_{r_1}(j) \lambda_S(i)}{\{[\lambda_S(i) + \lambda_{r_1}^{-1}(j)] \lambda_F(j) + 1\}^2} + \mu \left[\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\sigma_{H_2}^2(j)} + \frac{L}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} \right] \quad (4.37)$$

Therefore, according to KKT conditions [117], the solution $\lambda_F(j), j = 1, \dots, \bar{n}_F$ need to satisfy following equations:

$$\sum_{i=1}^{\bar{n}_S} \frac{\lambda_{t_1}(i) \lambda_{r_1}(j) \lambda_S(i)}{\{[\lambda_S(i) + \lambda_{r_1}^{-1}(j)] \lambda_F(j) + 1\}^2} = \mu \left[\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\sigma_{H_2}^2(j)} + \frac{L}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} \right] \quad (4.38)$$

$$f(\boldsymbol{\lambda}_F) = \sum_{i=1}^{\bar{n}_S} \lambda_S(i) \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} = P_R \quad (4.39)$$

where $\boldsymbol{\lambda}_F = [\lambda_F(1), \dots, \lambda_F(\bar{n}_F)]^T$.

Let the left hand side of (4.38)

$$g_j(\lambda_F(j)) = \sum_{i=1}^{\bar{n}_S} \frac{\lambda_{t_1}(i) \lambda_{r_1}(j) \lambda_S(i)}{\{[\lambda_S(i) + \lambda_{r_1}^{-1}(j)] \lambda_F(j) + 1\}^2} \quad (4.40)$$

It can be observed that $g_j(\cdot)$ is a monotonically decreasing function with respect to $\lambda_F(j)$.

Therefore, μ is monotonically decreasing with the increase of $\lambda_F(j)$.

So for each set of the newly found $\lambda_S(i)$ for all i , we check the energy constraint. If the constraint is violated, we increase μ . Or otherwise, we reduce μ . This is because $g_j(\lambda_F(j))$ is a decreasing function of $\lambda_S(i)$. The search for the optimal μ can follow the bi-section method [119].

In order to apply bi-section method, we need the range of both $\lambda_F(j)$ and μ :

The range of $\lambda_F(j)$ is

$$0 \leq \lambda_F(j) \leq \frac{P_R}{\sum_{i=1}^{\bar{n}_S} \lambda_S(i) \sum_{j=1}^{\bar{n}_F} \sigma_{H_2}^{-2}(j) + L \sum_{j=1}^{\bar{n}_F} \sigma_{H_2}^{-2}(j) \lambda_{r_1}^{-1}(j)} = \lambda_F^{max} \quad (4.41)$$

The range of μ is

$$0 \leq \mu \leq \max_j \frac{\lambda_{r_1}(j)}{\sigma_{H_2}^{-2}(j) \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sigma_{H_2}^{-2}(j) \lambda_{r_1}^{-1}(j)} \sum_{i=1}^{\bar{n}_S} \lambda_{t_1}(i) \lambda_S(i) = \mu^{max} \quad (4.42)$$

The two-layer bi-section algorithm used to find $\lambda_F(j)$ satisfying (4.38) and (4.39) is shown in Table 4.2.

2. Optimizing $\{\lambda_S(i)\}$ with fixed $\{\lambda_F(j)\}$:

We fix a group of $\lambda_F(j), j = 1, 2, \dots, \bar{n}_F$, then optimization problem in (4.33) can be transformed as

$$\begin{aligned} \min_{\{\lambda_S(i)\}} & - \sum_{i=1}^{\bar{n}_S} \sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j) \lambda_S(i)}{\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1} \\ \text{s.t.} & \sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} \leq P_S \\ & \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} \leq P_R \end{aligned} \quad (4.43)$$

It can be observed that the above optimization problem has two energy constraints.

Therefore the Lagrange function is

$$L^S = - \sum_{i=1}^{\bar{n}_S} \sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j) \lambda_S(i)}{\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1} + \mu_1 \left[\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} - P_S \right] \quad (4.44)$$

$$+ \mu_2 \left[\sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} - P_R \right]$$

The derivative of L^S with respect to $\lambda_S(i)$ is

$$\frac{\partial L^S}{\partial \lambda_S(i)} = - \sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j) [\lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]}{[\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]^2} + \mu_1 \lambda_{t_1}^{-1}(i) + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.45)$$

According to KKT conditions, there are three possibilities:

- Possibility 1:

$$\sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j) [\lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]}{[\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]^2} = \frac{\mu_1}{\lambda_{t_1}(i)} + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.46)$$

$$\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} < P_S$$

$$\sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} = P_R$$

$$\mu_1 = 0$$

$$\mu_2 > 0$$

- Possibility 2:

$$\begin{aligned}
\sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j) [\lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]}{[\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]^2} &= \frac{\mu_1}{\lambda_{t_1}(i)} + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.47) \\
\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} &= P_S \\
\sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} &< P_R \\
\mu_1 &> 0 \\
\mu_2 &= 0
\end{aligned}$$

- Possibility 3:

$$\begin{aligned}
\sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i) \lambda_F(j) \lambda_{r_1}(j) [\lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]}{[\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]^2} &= \frac{\mu_1}{\lambda_{t_1}(i)} + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.48) \\
\sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} &= P_S \\
\sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j) \lambda_{r_1}(j)} &= P_R \\
\mu_1 &> 0 \\
\mu_2 &> 0
\end{aligned}$$

As the optimization problem in (4.43) is convex. There exists a solution that satisfies one of the three possibilities. We can try the first two possibilities by using two layers of 1-D bi-section search, which is similar to the algorithm shown in Table 4.2. If for the first two possibilities no solution is found, we then try Possibility 3. The latter involves a 2-D search of μ_1 and μ_2 . But for each given pair of the values of μ_1 and μ_2 , we do the same 1-D bi-section search for $\lambda_S(i)$ for each i .

Let

$$h_i(\lambda_S(i)) = \sum_{j=1}^{\bar{n}_F} \frac{\lambda_{t_1}(i)\lambda_F(j)\lambda_{r_1}(j)[\lambda_F(j)\lambda_{r_1}^{-1}(j) + 1]}{[\lambda_F(j)\lambda_S(i) + \lambda_F(j)\lambda_{r_1}^{-1}(j) + 1]^2}, \quad i = 1, 2, \dots, \bar{n}_S \quad (4.49)$$

$$f_a(\boldsymbol{\lambda}_S) = \sum_{i=1}^{\bar{n}_S} \frac{\lambda_S(i)}{\lambda_{t_1}(i)} \quad (4.50)$$

$$f_b(\boldsymbol{\lambda}_S) = \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)\lambda_{r_1}(j)} \quad (4.51)$$

where $\boldsymbol{\lambda}_S = [\lambda_S(1), \lambda_S(2), \dots, \lambda_S(\bar{n}_S)]^T$.

- 1-D Bi-section search for Possibility 1 and Possibility 2

Take Possibility 1 for example, it is expressed as follows:

$$h_i(\lambda_S(i)) = \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.52)$$

$$f_a(\boldsymbol{\lambda}_S) < P_S$$

$$f_b(\boldsymbol{\lambda}_S) = P_R$$

As the range of μ_2 is

$$0 \leq \mu_2 \leq \max_i \lambda_{t_1}(i) \frac{\sum_{j=1}^{\bar{n}_F} \frac{\lambda_{r_1}(j)\lambda_F(j)}{\lambda_{r_1}^{-1}(j)\lambda_F(j)+1}}{\sum_{j=1}^{\bar{n}_F} \sigma_{H_2}^{-2}\lambda_F(j)} \quad (4.53)$$

And the range of $\lambda_S(i)$ is

$$0 \leq \lambda_S(i) \leq \frac{P_R - L \sum_{j=1}^{\bar{n}_F} \sigma_{H_2}^{-2}(j)\lambda_{r_1}^{-1}\lambda_F(j)}{\sum_{j=1}^{\bar{n}_F} \sigma_{H_2}^{-2}(j)\lambda_F(j)} \quad (4.54)$$

Therefore a two-layer bi-section algorithm can be used to find $\lambda_S(i), i = 1, 2, \dots, \bar{n}_S$.

For Possibility 2, $\lambda_S(i)$ can be found in a similar manner.

- 2-D search combined with 1-D bi-section for Possibility 3

Possibility 3 can be written as

$$h_i(\lambda_S(i)) = \frac{\mu_1}{\lambda_{t_1}(i)} + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.55)$$

$$f_a(\boldsymbol{\lambda}_S) = P_S$$

$$f_b(\boldsymbol{\lambda}_S) = P_R$$

The algorithm we proposed to solve (4.55) involves two steps:

- step 1: 2-D search for μ_1 and μ_2
- step 2: 1-D bi-section search for each given pair of μ_1 and μ_2

Step 1:

In the 2-D search for μ_1 and μ_2 , we first identify the ranges of μ_1 and μ_2 . Secondly, we choose a resolution and divide the 2-D area into small grids. Finally, we start with a pair of μ_1 and μ_2 within the grid and apply gradient method to find a better pair of μ_1 and μ_2 resulting in a local optimum. It is worth noticing that if the resolution is small enough, a global optimum is guaranteed.

The ranges of μ_1 and μ_2 are respectively

$$0 < \mu_1 \leq \lambda_{t_1}^2(i) \sum_{j=1}^{\bar{n}_F} \frac{\lambda_{r_1}(j)\lambda_F(j)}{\lambda_{r_1}^{-1}(j)\lambda_F(j) + 1} = \mu_1^{max} \quad (4.56)$$

$$0 < \mu_2 \leq \left[\lambda_{t_1}(i) \sum_{j=1}^{\bar{n}_F} \frac{\lambda_{r_1}(j)\lambda_F(j)}{\lambda_{r_1}^{-1}(j)\lambda_F(j) + 1} - \mu_1 \lambda_{t_1}^{-1}(i) \right] \cdot \left[\sum_{j=1}^{\bar{n}_F} \sigma_{H_2}^{-2}(j)\lambda_F(j) \right]^{-1} = \mu_2^{max} \quad (4.57)$$

For the gradient method, let the cost function be

$$J(\boldsymbol{\mu}_S) = [f_a(\boldsymbol{\lambda}_S) - P_S]^2 + [f_b(\boldsymbol{\lambda}_S) - P_R]^2 \quad (4.58)$$

Then the conditions are transformed to a unconstrained optimization problem

$$\min_{\boldsymbol{\mu}_S} J(\boldsymbol{\mu}_S) = [f_a(\boldsymbol{\lambda}_S) - P_S]^2 + [f_b(\boldsymbol{\lambda}_S) - P_R]^2 \quad (4.59)$$

where $\boldsymbol{\mu}_S = [\mu_1, \mu_2]^T$. $\boldsymbol{\lambda}_S$, μ_1 and μ_2 are related by $h_i(\lambda_S(i)) = \frac{\mu_1}{\lambda_{t_1}(i)} + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)}$.

The gradient is

$$\frac{\partial J(\boldsymbol{\mu}_S)}{\partial \boldsymbol{\mu}_S} = [f_a(\boldsymbol{\lambda}_S) - P_S]^2 \frac{\partial f_a(\boldsymbol{\lambda}_S)}{\partial \boldsymbol{\mu}_S} + [f_b(\boldsymbol{\lambda}_S) - P_R]^2 \frac{\partial f_b(\boldsymbol{\lambda}_S)}{\partial \boldsymbol{\mu}_S} \quad (4.60)$$

1. $\frac{\partial f_a(\boldsymbol{\lambda}_S)}{\partial \boldsymbol{\mu}_S}$:

As

$$\frac{\partial f_a(\boldsymbol{\lambda}_S)}{\partial \boldsymbol{\mu}_S} = \sum_{i=1}^{\bar{n}_S} \lambda_{t_1}^{-1} \frac{\partial \lambda_S(i)}{\partial \boldsymbol{\mu}_S} = \sum_{i=1}^{\bar{n}_S} \lambda_{t_1}^{-1} \begin{bmatrix} \frac{\partial \lambda_S(i)}{\partial \mu_1} \\ \frac{\partial \lambda_S(i)}{\partial \mu_2} \end{bmatrix} \quad (4.61)$$

Recall that

$$h_i(\lambda_S(i)) = \frac{\mu_1}{\lambda_{t_1}(i)} + \mu_2 \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \quad (4.62)$$

Take derivative of μ_1 and μ_2 on both sides of the equation respectively, we have

$$\frac{\partial h_i(\lambda_S(i))}{\partial \lambda_S(i)} \frac{\partial \lambda_S(i)}{\partial \mu_1} = \lambda_{t_1}^{-1} \implies \frac{\partial \lambda_S(i)}{\partial \mu_1} = \lambda_{t_1}^{-1} \left[\frac{\partial h_i(\lambda_S(i))}{\partial \lambda_S(i)} \right]^{-1} \quad (4.63)$$

$$\frac{\partial h_i(\lambda_S(i))}{\partial \lambda_S(i)} \frac{\partial \lambda_S(i)}{\partial \mu_2} = \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2} \implies \frac{\partial \lambda_S(i)}{\partial \mu_2} = \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2} \left[\frac{\partial h_i(\lambda_S(i))}{\partial \lambda_S(i)} \right]^{-1} \quad (4.64)$$

where $\frac{\partial h_i(\lambda_S(i))}{\partial \lambda_S(i)}$ is

$$\frac{\partial h_i(\lambda_S(i))}{\partial \lambda_S(i)} = -2\lambda_{t_1}(i) \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F^2(j) \lambda_{r_1}(j) [\lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]}{[\lambda_F(j) \lambda_S(i) + \lambda_F(j) \lambda_{r_1}^{-1}(j) + 1]^3} \quad (4.65)$$

2. $\frac{\partial f_b(\boldsymbol{\lambda}_S)}{\partial \boldsymbol{\mu}_S}$:

$$\frac{\partial f_b(\boldsymbol{\lambda}_S)}{\partial \boldsymbol{\mu}_S} = \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \frac{\partial \lambda_S(i)}{\partial \boldsymbol{\mu}_S} = \sum_{j=1}^{\bar{n}_F} \frac{\lambda_F(j)}{\sigma_{H_2}^2(j)} \sum_{i=1}^{\bar{n}_S} \begin{bmatrix} \frac{\partial \lambda_S(i)}{\partial \mu_1} \\ \frac{\partial \lambda_S(i)}{\partial \mu_2} \end{bmatrix} \quad (4.66)$$

With (4.60) and the initialization from grid method, $\boldsymbol{\mu}_S^k$ is updated as

$$\boldsymbol{\mu}_S^{k+1} = \boldsymbol{\mu}_S^k - \gamma^{m_k} \frac{\partial J(\boldsymbol{\mu}_S)}{\partial \boldsymbol{\mu}_S} \quad (4.67)$$

where m_k is given by Armijo rules. m_k is the minimal nonnegative integer that satisfies the following inequality

$$J(\boldsymbol{\mu}_S^{k+1}) - J(\boldsymbol{\mu}_S^k) \leq -\sigma \gamma^{m_k} \left(\frac{\partial J(\boldsymbol{\mu}_S)}{\partial \boldsymbol{\mu}_S} \right)^H \frac{\partial J(\boldsymbol{\mu}_S)}{\partial \boldsymbol{\mu}_S} \quad (4.68)$$

where σ and γ are constants. According to [119], σ is often chosen close to zero, for example, $\sigma \in [10^{-5}, 10^{-1}]$. A proper choice of γ is usually from 0.1 to 0.5. The convergence criterion of the gradient algorithm can be chosen as $\max \text{abs}(\boldsymbol{\mu}_S^{k+1} - \boldsymbol{\mu}_S^k) \leq \epsilon$ where $\max \text{abs}(\cdot)$ denotes the maximal absolute value of each element of a vector, and ϵ is a positive constant close to 0.

As the unconstrained optimization problem (4.59) is non-convex, we choose a proper resolution to sample μ_1 and μ_2 . Then we take each pair of μ_1 and μ_2 sample as initial point for gradient iteration. Then we find a local optimum to minimize the objective function. For all pairs of the initializations, we find the one results in the minimal objective function. If the resultant value of cost function is small enough (close to zero), then $\boldsymbol{\lambda}_S = [\lambda_S(1), \dots, \lambda_S(\bar{n}_S)]^T$ is found. If not, we can always reduce the resolution and repeat the process again.

Step 2:

For a given pair of μ_1 and μ_2 , 1-D bi-section method can be used to find a group of $\lambda_S(i), i = 1, 2 \dots, \bar{n}_S$.

4.4.6 Optimal Diagonal Components of the Trainings: A Special Case

In this section, we will propose an algorithm to find the energy allocation for training and relay matrices when $\mathbf{C}_{t_1} = \beta \mathbf{I}, \mathbf{C}_{r_1} = \alpha \mathbf{I}$. The motivation behind this is that when $\mathbf{C}_{r_1} = \alpha \mathbf{I}$, the optimal structures of training and relay matrices are guaranteed. And when $\mathbf{C}_{t_1} = \beta \mathbf{I}$, either the two constraints in (4.43) are equivalent or only one of them is active (where equality holds for optimal solution). It will be shown next that the algorithm to find the power allocation can be greatly simplified when $\mathbf{C}_{t_1} = \beta \mathbf{I}$.

When $\mathbf{C}_{t_1} = \beta \mathbf{I}, \mathbf{C}_{r_1} = \alpha \mathbf{I}$, the optimal structures of training and relay matrices in Theorem 2 are respectively

$$\mathbf{S}_{S,opt}^T = \frac{1}{\sqrt{\beta}} \Lambda_S^{1/2} = \Lambda_1^{1/2}, \quad \mathbf{F}_{opt} = \mathbf{V}_{H_2} \Sigma_{H_2}^{-1} \Lambda_F^{1/2} = \mathbf{V}_{H_2} \Lambda_2^{1/2} \quad (4.69)$$

Instead of finding $\lambda_S(i)$ and $\lambda_F(j)$, we find $\lambda_1(i) = \frac{1}{\beta} \lambda_S(i)$ and $\lambda_2(j) = \sigma_{H_2}^{-2}(j) \lambda_F(j)$.

With the optimal training and relay matrices structures suggested in Theorem 2, the optimization problem in (4.22) can be written as

$$\begin{aligned} \min_{\Lambda_1, \Lambda_2} \quad & -tr \left\{ \left[\left(\beta \Lambda_1 + \frac{1}{\alpha} \mathbf{I} \right) \otimes \Sigma_{H_2} \Lambda_2 \Sigma_{H_2}^H + \mathbf{I} \right]^{-1} \left(\beta^2 \Lambda_1 \otimes \alpha \Sigma_{H_2} \Lambda_2 \Sigma_{H_2}^H \right) \right\} \quad (4.70) \\ s.t. \quad & tr \{ \Lambda_1 \} \leq P_S \\ & tr \left\{ \left(\beta \Lambda_1 + \frac{1}{\alpha} \mathbf{I} \right) \otimes \Lambda_2 \right\} \leq P_R \end{aligned}$$

Let $\Lambda_1 \otimes \mathbf{I} = \bar{\Lambda}_1$, $\mathbf{I} \otimes \Lambda_2 = \bar{\Lambda}_2$ and $\mathbf{I} \otimes \Sigma_{H_2} = \bar{\Sigma}_{H_2}$ the optimization problem transforms to

$$\begin{aligned} \min_{\bar{\Lambda}_1, \bar{\Lambda}_2} \quad & -tr \left\{ \left(\beta \bar{\Lambda}_1 \bar{\Sigma}_{H_2} \bar{\Lambda}_2 \bar{\Sigma}_{H_2}^H + \frac{1}{\alpha} \bar{\Sigma}_{H_2} \bar{\Lambda}_2 \bar{\Sigma}_{H_2}^H + \mathbf{I} \right)^{-1} \bar{\Lambda}_1 \bar{\Sigma}_{H_2} \bar{\Lambda}_2 \bar{\Sigma}_{H_2}^H \right\} \\ \text{s.t.} \quad & tr \{ \bar{\Lambda}_1 \} \leq \bar{n}_F P_S \\ & tr \left\{ \left(\beta \bar{\Lambda}_1 + \frac{1}{\alpha} \mathbf{I} \right) \bar{\Lambda}_2 \right\} \leq P_R \end{aligned} \quad (4.71)$$

Let $\bar{\lambda}_1(k)$, $\bar{\lambda}_2(k)$ and $\bar{\sigma}_{H_2}^2(k)$, $i = k, \dots, \bar{n}_S \bar{n}_F$ be the i th diagonal element of $\bar{\Lambda}_1$, $\bar{\Lambda}_2$ and $|\bar{\Sigma}_{H_2}|^2$ respectively. Therefore, the optimization problem becomes

$$\begin{aligned} \min_{\{\bar{\lambda}_1(k)\}, \{\bar{\lambda}_2(k)\}} \quad & - \sum_{k=1}^{\bar{n}_S \bar{n}_F} \frac{\bar{\sigma}_{H_2}^2(k) \bar{\lambda}_1(k) \bar{\lambda}_2(k)}{\beta \bar{\sigma}_{H_2}^2(k) \bar{\lambda}_1(k) \bar{\lambda}_2(k) + \frac{1}{\alpha} \bar{\sigma}_{H_2}^2(k) \bar{\lambda}_2(k) + 1} \\ \text{s.t.} \quad & \sum_{k=1}^{\bar{n}_S \bar{n}_F} \bar{\lambda}_1(k) \leq \bar{n}_F P_S \\ & \sum_{k=1}^{\bar{n}_S \bar{n}_F} \left(\beta \bar{\lambda}_1(k) + \frac{1}{\alpha} \right) \bar{\lambda}_2(k) \leq P_R \end{aligned} \quad (4.72)$$

To simplify the optimization problem above, we let

$$b_k = \bar{\sigma}_{H_2}^2(k) \quad (4.73)$$

$$x_k = \bar{\lambda}_1(k) \quad (4.74)$$

$$y_k = \bar{\lambda}_2(k) \left(\beta \bar{\lambda}_1(k) + \frac{1}{\alpha} \right) \quad (4.75)$$

Therefore, above optimization problem can be written in the form

$$\begin{aligned} \min_{\{x_k\}, \{y_k\}} \quad & \sum_{k=1}^K - \frac{x_k}{\frac{1}{\alpha} + \beta x_k} \frac{b_k y_k}{1 + b_k y_k} \\ \text{s.t.} \quad & \sum_{k=1}^K x_k \leq \bar{n}_F P_S \\ & \sum_{k=1}^K y_k \leq P_R \end{aligned} \quad (4.76)$$

Optimization problem (4.76) is nonconvex with respect to $x_k, y_k, k = 1, 2, \dots, K$. However, it is conditional convex if we fix either $x_k, k = 1, 2, \dots, K$ or $y_k, k = 1, 2, \dots, K$. Hence we can use an alternating algorithm to obtain a local optimal solution. Specifically, we fix $y_k, k = 1, 2, \dots, K$, and let $q_k = \frac{b_k y_k}{1 + b_k y_k}, k = 1, 2, \dots, K$ with y_k satisfying $\sum_{k=1}^K y_k = P_R$. Then we solve the convex problem

$$\begin{aligned} \min_{\{x_k\}} \quad & \sum_{k=1}^K -q_k \frac{x_k}{\frac{1}{\alpha} + \beta x_k} \\ \text{s.t.} \quad & \sum_{k=1}^K x_k \leq \bar{n}_F P_S \end{aligned} \quad (4.77)$$

The above convex problem can be solved using Lagrange multiplier method and has a water-filling type solution.

The Lagrange function is

$$L_x = \sum_{k=1}^K -q_k \frac{x_k}{\frac{1}{\alpha} + \beta x_k} + \mu_x \left(\sum_{k=1}^K x_k - \bar{n}_F P_S \right) \quad (4.78)$$

The derivative of L_x with respect to x_k is

$$\frac{\partial L_x}{\partial x_k} = -\frac{\frac{1}{\alpha} q_k}{(1 + x_k)^2} + \mu_x \quad (4.79)$$

Let $\frac{\partial L_x}{\partial x_k} = 0$, the water-filling type solution is

$$x_k^* = \left[\sqrt{\frac{\frac{1}{\alpha} q_k}{\mu_x}} - 1 \right]^+ \quad (4.80)$$

where $[x]^+ = \max(x, 0)$ and μ_x is the Lagrange multiplier satisfying $\sum_{k=1}^K x_k = \bar{n}_F P_S$.

The process of solving (4.77) is shown in the Table 4.3.

Similarly, we fix $x_k, k = 1, 2, \dots, K$, and let $p_k = \frac{x_k}{\frac{1}{\alpha} + \beta x_k}, k = 1, 2, \dots, K$ with x_k

satisfying $\sum_{k=1}^K x_k = \bar{n}_F P_S$. Then we solve the convex problem

$$\begin{aligned} \min_{\{y_k\}} \sum_{k=1}^K -p_k \frac{b_k y_k}{1 + b_k y_k} \\ \text{s.t. } \sum_{k=1}^K y_k \leq P_R \end{aligned} \quad (4.81)$$

The above convex problem can be solved using Lagrangian multiplier method and has a water-filling type solution.

$$y_k^* = \frac{1}{b_k} \left[\sqrt{\frac{b_k p_k}{\mu_y}} - 1 \right]^+ \quad (4.82)$$

where μ_y is the Lagrangian multiplier satisfying $\sum_{k=1}^K y_k = P_R$.

Since the conditional update of x_k or y_k by fixing the other, may either decrease or maintain but can not increase the MSE, monotonic convergence of $\{x_k, y_k\}$ follows directly from this observation. It is worth noticing that once a local optimum is reached, the updating process will terminate. Therefore, the alternating algorithm will achieve locally optimal solution. However, the alternating algorithm is subject to low complexity and fast convergence. The alternating algorithm to obtain local minimum is shown in Table 4.4.

Once a local optimum is obtained, the diagonal of $\bar{\Lambda}_1$ and $\bar{\Lambda}_2$ are given by

$$\bar{\lambda}_1(k) = x_k \quad (4.83)$$

and

$$\bar{\lambda}_2(k) = \frac{y_k}{\frac{1}{\alpha} + \beta x_k} \quad (4.84)$$

An example of this special case is the uncorrelated MIMO channel. When $\alpha = 1, \beta = 1, \mathbf{C}_{t_1} = \mathbf{C}_{r_1} = \mathbf{I}, \mathbf{H}_1$ is an uncorrelated MIMO channel. Then two-layered bi-section algorithm shown in Table 4.4 can be used to find a local optimal solution.

4.5 Numerical Results

In this section, we present some numerical examples to illustrate the performance of our proposed algorithm. We assume $P_S = P_R = P$ and $n_S = n_R = n_D = L = N$. We define a correlation matrix \mathbf{C}_ρ with $[\mathbf{C}_\rho]_{i,j} = \rho^{|i-j|}$ where ρ is the normalized correlation coefficient with magnitude $|\rho| < 1$ [120]. We also define the normalized MSE of \mathbf{H}_1 and \mathbf{H}_2 as $\frac{E_{\mathbf{H}_2}[J'_1]}{N^2}$ and $\frac{J'_2}{N^2}$, respectively. The average over \mathbf{H}_2 is computed by using 100 realizations of \mathbf{H}_2 .

Fig. 4.2 compares the normalized MSE of \mathbf{H}_1 between “optimal source and relay trainings” and “orthogonal source and relay trainings”, where $\mathbf{H}_1 = \mathbf{W}_1$, $\mathbf{H}_2 = \mathbf{W}_2$ and $N = 4$. For orthogonal trainings, we use $\mathbf{S}_S = \sqrt{\frac{P_S}{n_S}}\mathbf{I}$ and $\mathbf{F} = \sqrt{\frac{P_R}{(P_S+n_S)n_R}}\mathbf{I}$. Here, both \mathbf{S}_S and \mathbf{F} have orthogonal columns. As expected, the optimal trainings yield a better accuracy than the orthogonal trainings.

From (4.21), we can see that the channel correlation of both \mathbf{H}_1 and \mathbf{H}_2 (\mathbf{C}_{t_1} , $\mathbf{C}_{r_1}, \mathbf{C}_{t_2}$, \mathbf{C}_{r_2}) will impact the MSE of \mathbf{H}_1 . In the following, Fig. 4.3 and Fig. 4.4 will illustrate how the channel correlation of \mathbf{H}_1 and \mathbf{H}_2 impact the MSE of \mathbf{H}_1 differently. Generally speaking, MSE of \mathbf{H}_1 achieves better performance when \mathbf{H}_1 is strongly correlated and \mathbf{H}_2 is weakly correlated.

Fig. 4.3 illustrates the normalized MSE of \mathbf{H}_1 with the optimal source and relay trainings, where $\mathbf{H}_1 = \mathbf{W}_1$ and $\mathbf{H}_2 = \mathbf{C}_\rho^{1/2}\mathbf{W}_2\mathbf{C}_\rho^{1/2}$ with $\rho = 0.2$ (weak correlation) and $\rho = 0.8$ (strong correlation). We can see from Fig. 4.3, as \mathbf{H}_2 becomes strongly correlated, the MSE of \mathbf{H}_1 degrades, with the difference in performance being more apparent in high energy constraint region. We can also observe that with the increase of N , the gap between

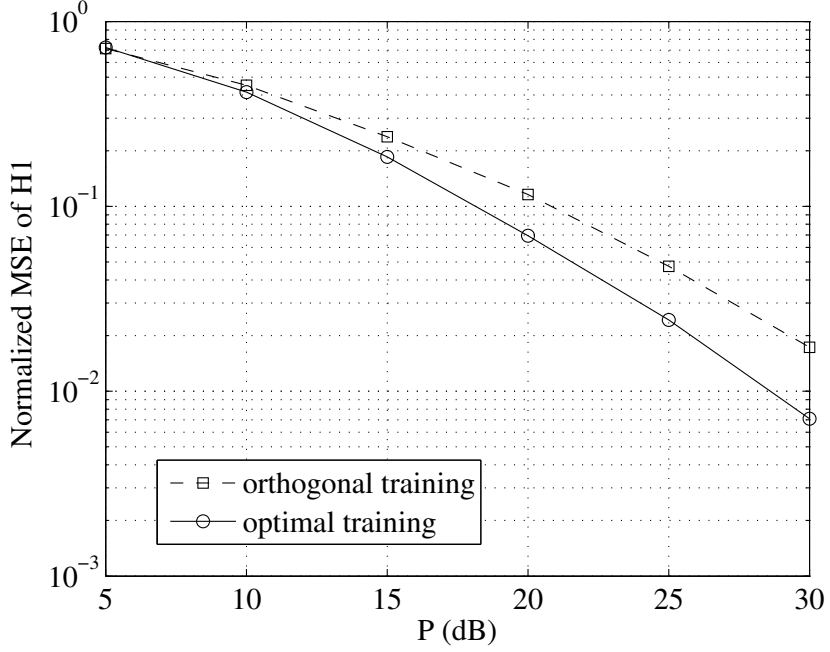


Figure 4.2: Normalized MSE of \mathbf{H}_1 with optimal trainings and orthogonal trainings.

the performances of weakly correlated \mathbf{H}_2 and strongly correlated \mathbf{H}_2 becomes larger.

Fig. 4.4 illustrates the normalized MSE of \mathbf{H}_1 with the optimal source and relay trainings, where $\mathbf{H}_2 = \mathbf{W}_2$ and $\mathbf{H}_1 = \mathbf{C}_\rho^{1/2} \mathbf{W}_1 \mathbf{C}_\rho^{1/2}$ with $\rho = 0.2$ (weak correlation) and $\rho = 0.8$ (strong correlation). Different from Fig. 4.3, Fig. 4.4 shows that with \mathbf{H}_1 getting strongly correlated, the MSE of \mathbf{H}_1 improves, with the performance gap more obvious in low energy constraint region. We can also observe that with the increase of N , the gap between the performances of weakly correlated \mathbf{H}_1 and strongly correlated \mathbf{H}_1 becomes larger.

Fig.4.5 compares the normalized MSE of \mathbf{H}_1 and that of \mathbf{H}_2 , where $\mathbf{H}_1 = \mathbf{W}_1$ and $\mathbf{H}_2 = \mathbf{W}_2$. We see that the estimation accuracy of \mathbf{H}_2 is much higher than that of \mathbf{H}_1 , which is expected. Recall that the estimation of \mathbf{H}_2 in phase 1 is based on a single-hop link

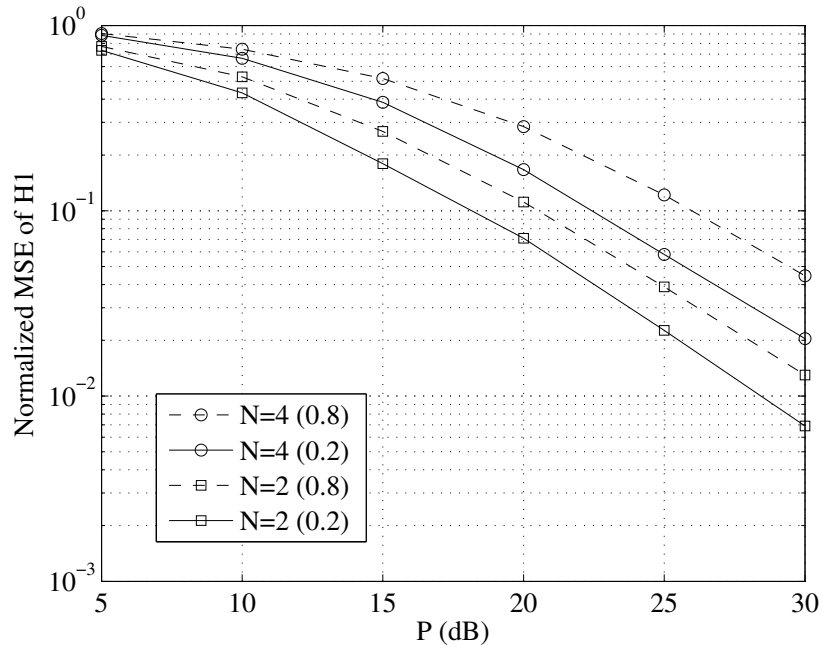


Figure 4.3: Normalized MSE of \mathbf{H}_1 where \mathbf{H}_1 is uncorrelated and \mathbf{H}_2 is correlated with the correlation factor ρ .

while the estimation of \mathbf{H}_1 in phase 2 is based on a two-hop relay system where the relay only does "amplify and forward". The high accuracy of \mathbf{H}_2 is in fact important for the estimation of \mathbf{H}_1 in phase 2 where \mathbf{H}_2 is assumed to be known. Also note that the method shown in this paper does not have the ambiguity problem suffered by those in [26] and [27].

4.6 Conclusion

In this paper, we have proposed a two-phase LMMSE-based channel estimation method for a two-hop AF MIMO relay system. In phase 1, the relay-to-destination channel is estimated for which the relay sends out a source training matrix. In phase 2, the source-to-relay channel is estimated for which the source sends out a source training matrix and

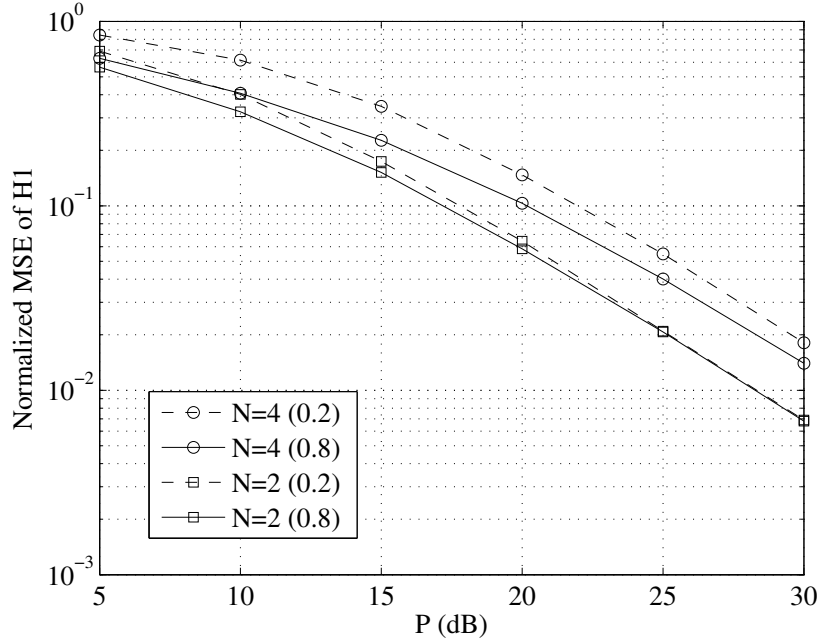


Figure 4.4: Normalized MSE of \mathbf{H}_1 where \mathbf{H}_1 is correlated with the correlation factor ρ and \mathbf{H}_2 is uncorrelated.

the relay applies a relay training matrix. For phase 1, an optimal design of the source training has been presented, the result of which is similar to one in [25] while our approach based on generalized KKT conditions provides a complementary perspective. For phase 2, an optimal joint design of the source and relay trainings has been developed, which is a much harder problem than in phase 1. The two-phase channel estimation scheme shown in this paper does not have the ambiguity problem suffered by the schemes in [26] and [27].

The two-phase scheme can be extended to an M -phase scheme for an M -hop AF relay system. If all nodes are indexed sequentially with the source node being node 0 and the destination node being node M , then in phase m the channel matrix between node $M - m$ and node $M - m + 1$ is estimated for which node $M - m$ sends out a source training

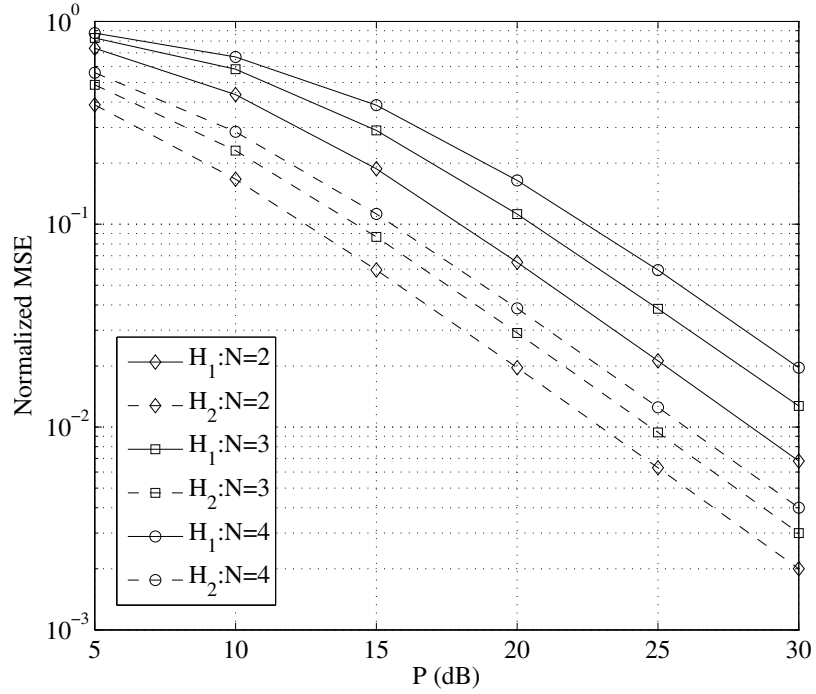


Figure 4.5: Normalized MSE of \mathbf{H}_2 estimated in phase 1 and the normalized MSE of \mathbf{H}_1 estimated in phase 2.

matrix, all other down-stream nodes (except node M) sends out relay training matrices and the channel matrices between the adjacent down-stream nodes can be assumed to be known. But a problem with such a scheme is that the estimation errors for the down-stream channels will accumulate and affect the estimation of their upper-stream channels. In practice, such a scheme can be useful only if SNR for each link is sufficiently high.

Table 4.2: Two-layer Bisection Algorithm to Calculate $\lambda_F(j)$

1. initialization:
 μ : $\mu^{min} = 0, \mu^{max}$
 λ_F :
 $\lambda_F^{min}(j) = 0, \lambda_F^{max}(j) = \lambda_F^{max}, j = 1, \dots, \bar{n}_F$

2. two loop bi-section:
outer loop:
 $\mu^{mid} = \frac{1}{2}(\mu^{min} + \mu^{max})$
while ($\mu^{mid} \neq \mu^{min}$ and $\mu^{mid} \neq \mu^{max}$)
 for $j = 1 : \bar{n}_F$
 inner loop:
 $\lambda_F^{mid}(j) = \frac{1}{2}[\lambda_F^{min}(j) + \lambda_F^{max}(j)]$
while ($\lambda_F^{mid}(j) \neq \lambda_F^{min}(j)$ and $\lambda_F^{mid}(j) \neq \lambda_F^{max}(j)$)
 if $g_j(\lambda_F^{mid}(j)) < \mu^{mid} \left[\sigma_{H_2}^{-2}(j) \sum_{i=1}^{\bar{n}_S} \lambda_S(i) + L\sigma_{H_2}^{-2}(j)\lambda_{r_1}^{-1}(j) \right]$
 $\lambda_F^{max}(j) = \lambda_F^{mid}(j)$
 else
 $\lambda_F^{min}(j) = \lambda_F^{mid}(j)$
 endif
 $\lambda_F^{mid}(j) = \frac{1}{2}[\lambda_F^{min}(j) + \lambda_F^{max}(j)]$
 endwhile
 $\lambda_F^*(j) = \lambda_F^{mid}(j)$
 endfor
if $f(\lambda_F^*) < P_R$
 $\mu^{max} = \mu^{mid}$
else
 $\mu^{min} = \mu^{mid}$
 endif
 $\mu^{mid} = \frac{1}{2}(\mu^{min} + \mu^{max})$
endwhile

3. $\lambda_F^{opt}(j) = \lambda_F^*(j), j = 1, \dots, \bar{n}_F$

Table 4.3: Bisection Algorithm to Calculate $\{x_k^*\}$

1. initialization:
 $\mu_x: \mu_x^{min} = 0, \mu_x^{max} = \max_k \frac{1}{\alpha} q_k$

2. loop:
 $\mu_x^{mid} = (\mu_x^{max} + \mu_x^{min})/2$
while $(\mu_x^{mid} \neq \mu_x^{min} \text{ and } \mu_x^{mid} \neq \mu_x^{max})$
 $x_k = \left[\sqrt{\frac{\frac{1}{\alpha} q_k}{\mu_x^{mid}} - 1} \right]^+$, for $k = 1, 2, \dots, K$
if $\sum_{k=1}^K x_k < \bar{n}_F \bar{P}_s$
 $\mu_x^{max} = \mu_x^{mid}$
else
 $\mu_x^{min} = \mu_x^{mid}$
endif
 $\mu_x^{mid} = (\mu_x^{max} + \mu_x^{min})/2$
endwhile

3. $x_k^* = x_k$

Table 4.4: Alternating Algorithm to Find Local Optimum $\{x_{lopt}^k, y_{lopt}^k\}$

1. initialization:
 $y_k^0 \rightarrow \sum_{k=1}^K y_k^0 = P_R$
 $q_k^0 = \frac{b_k y_k^0}{1 + b_k y_k^0}$
 flag=1, $i = 0$

2. Loop:
 while flag
 $i = i + 1$
 $x_k^i = \left[\sqrt{\frac{\frac{1}{\alpha} q_k^{i-1}}{\mu_x^i} - 1} \right]^+$ using bi-section algorithm
 $p_k^i = \frac{x_k^i}{\frac{1}{\alpha} + \beta x_k^i}$
 $MSE_1 = f_{MSE}(x_k^i, y_k^{i-1})$
 $y_k^i = \frac{1}{b_k} \left[\sqrt{\frac{b_k p_k^i}{\mu_y^i} - 1} \right]^+$ using bi-section algorithm
 $q_k^i = \frac{b_k y_k^i}{1 + b_k y_k^i}$
 $MSE_2 = f_{MSE}(x_k^i, y_k^i)$
 if $|MSE_1 - MSE_2| < \epsilon$
 flag=0
 endif
 endwhile

3. $x_{lopt}^k = x_k^i, y_{lopt}^k = y_k^i$

Chapter 5

Conclusion

In this dissertation, we have first reviewed the state-of-art of multihop relay transmission. Descriptions of the transmission mechanism have been presented and performances such as diversity, capacity and DMT have been compared. Then we have addressed two critical challenges in multihop relay networks.

The first challenge we have addressed is multihop transmission and power scheduling. We have proposed a DPC based multihop transmission strategy . By taking advantage of DPC, we have additional freedom to optimize over power scheduling and rate allocation to realize the goal of power saving and power balance. A general gradient projection method has been proposed to solve the optimization problem for networks where both single antennas and multiple antennas can be equipped in each node. Some useful properties have been explored to realize fast computation. An alternative subgroup method has also been provided to reach a tradeoff between performance and complexity when the network size becomes large. Our proposed method has achieved better power saving and power balance

performances compared with existing schemes.

The second challenge we have addressed is MIMO relay channel estimation and training design. We have proposed a two-phase LMMSE-based channel estimation method for a two-hop AF MIMO relay system. In phase 1, the relay-to-destination channel is estimated for which the relay sends out a source training matrix. In phase 2, the source-to-relay channel is estimated for which the source sends out a source training matrix and the relay applies a relay training matrix. For phase 1, an optimal design of the source training has been presented, the result of which is similar to one in [25] while our approach based on generalized KKT conditions provides a complementary perspective. For phase 2, an optimal joint design of the source and relay training matrices has been developed, which is a much harder problem than in phase 1. The two-phase channel estimation does not have the ambiguity problem suffered by the schemes in [26] and [27]. The two-phase scheme can be extended to an M -phase scheme for an M -hop AF relay system.

The research work presented in this dissertation has advanced the state-of-the-art in wireless multihop relay networks and brought us closer to realizing the vision of ubiquitous multihop wireless networks.

Bibliography

- [1] R. Ramanathan and J. Redi, "A Brief Overview of Ad Hoc Networks: Challenges and Directions," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 20-22, May 2002.
- [2] IEEE 802.16j: Air interface for fixed and mobile broadband wireless access systems. 2006.
- [3] I. Akyildiz, X. Wang and W. Wang, "Wireless mesh network: a survey," *Computer Networks*, pp. 445-487, Jan., 2005.
- [4] I. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 102-114, Aug 2002.
- [5] J. Mietzner, R. Schober, L. Lampe, W.H. Gerstacker and P.A. Hoeher, "Multiple antennas technique for wireless communications - a comprehensive literature survey," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 2, pp. 87-105, second quarter, 2009.
- [6] D. Gesbert, M. Shafi, Shiu Da-shan P. J. Smith and A. Naguib, "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems", *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 281-302, April 2003.
- [7] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-time Wireless Communications* Cambridge: Cambridge University Press, 2003.
- [8] MIMO Wikipedia, <http://en.wikipedia.org/wiki/MIMO>.
- [9] B. Wang, J. Zhang and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trnas. Inf. Theory*, vol.51, pp.29-43, Jan. 2005.
- [10] H. Bolcskei, R. U. Nabar, O. Oyman and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. on Wireless Communications*, vol. 5, no. 6, pp. 1433-1444, June 2006.
- [11] Z. Fang, Y. Hua and J. C. Koshy, "Joint source and relay optimization for non-regenerative MIMO relay," in *Proc. IEEE Workshop Sensor Array Multi-Channel Signal Processing*, Waltham, WA, July 2006, pp. 239-243.

- [12] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398-1407, Apr. 2007.
- [13] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. on Wireless Communications*, vol. 6, no. 5, pp. 1774-1786, May 2007.
- [14] O. Munoz, J. Vidal, and A. Agustin, "Linear transceiver design in non-regenerative relays with channel state information," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2593-2604, Jun. 2007.
- [15] C.-B. Chae, T. W. Tang, R. W. Heath and S.-Y. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 2, pp. 727-738, Feb. 2008.
- [16] A. S. Behbahani, R. Merched and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. on Signal Processing*, vol. 56, no. 10, part 2, pp. 5062-5073, Oct. 2008.
- [17] Y. Rong, X. Tang and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Transactions on Signal Processing*, Vol. 57, No. 12, pp. 4837-4852, Dec 2009.
- [18] Y. Rong and Y. Hua, "Optimality of diagonalization of multihop MIMO relays," *IEEE Transactions on Wireless Communications*, vol. 8, No. 12, pp. 6068-6077, Dec. 2009.
- [19] Y. Yu and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users", *IEEE Transactions on Signal Processing*, vol. 58, No. 5, pp. 2823-2835, May 2010.
- [20] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, Dec 2004, pp. 366-79.
- [21] R. Rajagopalan and P. Varshney, "Data aggregation techniques in sensor networks: A survey," *IEEE Comm. Surveys and Tutorials*, vol. 8, no. 4, pp. 48-63, 2006.
- [22] M. Costa, "Writing on dirty paper," *IEEE Transaction on Information Theory*, vol. 29, pp. 439-441, May 1983.
- [23] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 884-893, Mar. 2006.
- [24] M. Biguesh, S. Gazor and M. H. Shariat, "Optimal training sequence for MIMO wireless systems in colored environments," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 1807-1820, Aug. 2009.

- [25] E. Bjornson, B. Ottersten, “A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance,” *IEEE Trans. Signal Processing*, vol. 58, no. 3, pp. 1807-1820, Mar. 2010.
- [26] P. Lioliou and M. Viberg, “Least square based channel estimation for MIMO relays,” *IEEE WSA* 2008.
- [27] J. Ma, P. Orlik, J. Zhang and G. Y. Li, “Pilot matrix design for interim channel estimation in two hop MIMO AF relay systems,” *ICC* 2009.
- [28] Ting Kong, “A Survey of transmission Schemes for Multi-hop relay Networks,” *technical report*, 2009.
- [29] R. U. Nabar, H. Bolcskei, and F.W. Kneubuhler, “Fading relay channels: Performance limits and space-time signal design,” *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099-1109, Aug. 2004.
- [30] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity-part I: system description,” *IEEE Trans. Commun.*, vol. 51, pp. 1927-1938, Nov. 2003.
- [31] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity-part II: implementation aspects and performance analysis,” *IEEE Trans. Commun.*, vol. 51, pp. 1939-1948, Nov. 2003.
- [32] A. Nosratinia, T. Hunter, and A. Hedayat, “Cooperative communication in wireless networks,” *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68-73, October 2004.
- [33] J. N. Laneman and G.W.Wornell, “Energy-efficient antenna sharing and relaying for wireless networks,” in *Proc. IEEE WCNC*, 2000, pp. 7?2.
- [34] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, “An efficient protocol for realizing cooperative diversity in wireless networks,” in *Proc. IEEE Int. Symp. Information Theory*, Washington, DC, June 2001.
- [35] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062-3080, Dec 2004.
- [36] K. Azarian, H. E. Gamal, and P. Schniter, “On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 4152-4172, Dec. 2005.
- [37] N. Prasad, M. K. Varanasi, “Diversity and Multiplexing Tradeoff Bounds for Cooperative Diversity Protocols,” *Proc. International Symposium on Information Theory*, June 2004, Page:268.
- [38] T. E. Hunter and A. Nosratinia, “Cooperation diversity through coding,” in *Proc. Int. Symp. Inform. Theory*, Laussane, Switzerland, 2002, p. 220.

- [39] -, "Coded cooperation under slow fading, fast fading, and power control," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, 2002.
- [40] -, "Performance analysis of coded cooperation diversity," in *Proc. IEEE International Conference on Communications (ICC)*, Anchorage, AK, May 2003.
- [41] -, "Diversity through Coded Cooperation," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 283-289, Feb. 2006.
- [42] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1470-1476, Sep. 2004.
- [43] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Trans. Signal Processing*, vol. 53, pp. 362- 371, Feb. 2004.
- [44] G. S. Rajan and B. S. Rajan, "A non-orthogonal distributed space-time coded protocol, part I: signal model and design criterion," in *Proc. IEEE Inform. Theory. Workshop*, Chengdu, China, Oct 2006.
- [45] G. S. Rajan and B. S. Rajan, "A non-orthogonal distributed space-time coded protocol, part II: code construction and DMG tradeoff," in *Proc. IEEE Inform. Theory. Workshop*, Chengdu, China, Oct 2006.
- [46] J. Laneman and G. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415-2425, Oct. 2003.
- [47] Q. Zhao and H. Li, "Performance of Differential Modulation with Wireless Relays in Rayleigh Fading Channels," *IEEE Communication Letters*, vol. 9, no. 4, pp. 343-345, Apr. 2005.
- [48] Q. Zhao and H. Li, "Differential BPSK modulation for wireless relay networks," in *Proc. 38th Annual Conf. on Info Sci. and Systems*, Mar. 17-19, 2004.
- [49] H. Mheidat, M. Uysal and N. Al-Dhahir, "Equalization techniques for distributed space-time block codes with amplify-and-forward relaying," *IEEE Tran. on Signal Processing*, vol. 55, no. 5, 1839-1852, MAY 2007.
- [50] Y. Hua, Y. Chang and Y. Mei, "A networking perspective of mobile parallel relays," *IEEE Workshop on Digital Signal Processing*, pp. 249-253, Taos Ski Valley, NM, 1-4 Aug 2004.
- [51] Y. Chang and Y. Hua, "Application of space-time linear block codes to parallel wireless relays in mobile ad hoc networks," *Proc of Asilomar Annual Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2003.

- [52] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *Proc of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, NV, Apr. 2007.
- [53] G. Susinder Rajan and B. Sundar Rajan, "Distributed Space-Time Codes for Cooperative Networks with Partial CSI," in *Proc. IEEE WNCN2007*, 2007.
- [54] Y. Jing and B. Hassibi, "Distributed Space-Time Coding in Wireless Relay Networks," *IEEE Trans. on Wireless Communications*, vol. 5, no. 12, Dec. 2006, pp. 3524-3536.
- [55] F. Oggier and B. Hassibi, "An Algebraic Family of Distributed Space-Time Codes for Wireless Relay Networks," in *Proc IEEE International Symposium on Information Theory*, July 2006, pp. 538-541.
- [56] P. Elia, F. Oggier and P. Vijay Kumar, "Asymptotically Optimal Cooperative Wireless Networks with Reduced Signaling Complexity," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 258-267, Feb. 2007.
- [57] Petros Elia, K. Vinodh, M. Anand, P. Vijay Kumar, "D-MG Tradeoff and Optimal Codes for a Class of AF and DF Cooperative Communication Protocols," *IEEE Trans. Inform. Theory*, submitted. Available in Arxiv cs.IT/0611156.
- [58] T. Kiran and B.S. Rajan, "Partially coherent distributed space time codes with differential encoder and decoder," ISIT-2006, pp. 547-551, Seattle, July 9-14, 2006
- [59] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Trans. Inform. Theory*, vol. 53, pp. 4106-4118, Nov. 2007.
- [60] Z. Yi and I-Min Kim, "Single-Symbol ML Decodable Distributed STBCs for Cooperative Networks," *IEEE Trans. Info. Theory*, Vol. 53, No. 8, pp. 2977-2985, Aug 2007.
- [61] Md. Z. A. Khan and B. S. Rajan, "Single-symbol maximum likelihood decodable linear STBCs," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2062-2091, May 2006.
- [62] T. Kiran and B. Sundar Rajan, "Distributed Space-Time Codes with Reduced Decoding Complexity," *Proc. IEEE Intl. Symp. Inform. Theory*, Seattle, July 9-14, 2006, pp. 542-546.
- [63] G. S. Rajan and B. S. Rajan, "Multi-group ML decodable collocated and distributed space-time block codes," *IEEE Trans. Inform. Theory*, accepted for publication. Available online: http://aps.arxiv.org/PS_cache/arxiv/pdf/0712/0712.2384v2.pdf
- [64] Y. Hua, Y. Mei, and Y. Chang, "Wireless antennas making wireless communications perform like wireline communications", in IEEE AP-S Topical Conf. Wireless Communication Technology, Honolulu, HI, Oct. 2003, pp. 47-53.
- [65] Y. Hua, Y. Mei and Y. Chang, "Parallel wireless mobile relays with space-time modulation", *IEEE Workshop on Statistical Signal Processing*, St. Louis, MO, Sep. 2003.

- [66] J. M. Paredes, B. H. Khalaj and A. B. Gershman, "Cooperative Transmission for Wireless Relay Networks Using Low-Rate Feedback," submitted April 2009. Available online at http://arxiv.org/PS_cache/arxiv/pdf/0904/0904.1369v1.pdf
- [67] Y. Jing and H. Jafarkhani, "Distributed differential space-time coding for wireless relay networks," *IEEE Trans. On. Communications*, vol. 56, No. 7, pp. 1092-1100, Jul. 2008.
- [68] T. Kiran and B. S. Rajan, "Partially-coherent distributed space-time codes with differential encoder and decoder," *IEEE JSAC on. Communications*, vol. 25, no. 2, pp. 426-433, Feb. 2007.
- [69] B. Sirkeci-Mergen and A. Scaglione, "Randomized space-time coding for distributed cooperative communication," *IEEE Trans. Signal Processing*, vol. 55, no. 10, pp. 5003-5017, Oct. 2007.
- [70] S. Yiu, R. Schober, and L. Lampe, "Distributed space-time block coding," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1195-1206, Jul. 2006.
- [71] T. Wang, Y. Yao, and G. B. Giannakis, "Non-coherent distributed space-time processing for multiuser cooperative transmissions," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3339-3343, Dec. 2006.
- [72] Y. Zhao, R. Adve, and T. Lim, "Improving amplify-and-forward relay networks: Optimal power allocation versus selection," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3114-C3123, Aug. 2007.
- [73] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," submitted on Apr. 2008, available online <http://arxiv.org/abs/0804.1117>
- [74] P.Larsson, "Large-scale cooperative relaying network with optimal combining under aggregate relay power constraint" in *Proc of Future Telecommunications Conference*, 2003.
- [75] Z. Ding, W. H. Chin, and K. K. Leung, "Distributed beamforming and power allocation for cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1817-C1822, May 2008.
- [76] I. Hammerstrom, M. Kuhn, and A. Wittneben, "Impact of relay gain allocation on the performance of cooperative diversity networks," in *Proc. VTC 2004 Fall*, vol. 3, pp. 1815-1819, Sept. 2004.
- [77] A. F. Dana and B. Hassibi, "On the power efficiency of sensory and ad hoc wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, pp. 2890-2914, July 2006.
- [78] V. Havaray-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Network beamforming based on second-order statistics of the channel state information," in *Proc. ICASSP2008*, Las Vegas, USA, Apr. 2008, pp. 2605-2608.

- [79] E. Koyuncu, Y. Jing, and H. Jafarkhani, "Distributed beamforming in wireless relay networks with quantized feedback," *IEEE J. Select. Areas Commun.*, vol. 26, pp. 1429-1439, Oct. 2008.
- [80] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming for relay networks in frequency selective fading channels," in *Proc. ICASSP-09*, Taipei, Taiwan, April 2009.
- [81] G. Zheng, K. K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation," *IEEE Signal Processing Letters*, vol. 16, no. 4, pp. 257-260, April 2009.
- [82] Z. Yi and I. Kim, "Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 447-458, Feb. 2007.
- [83] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello and T. E. Fuja, "Bandwidth- and power-efficient routing in linear wireless networks", *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2624-2633, Jun. 2006.
- [84] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello and T. E. Fuja, "On the optimum number of hops in linear wireless networks", in *Proc. IEEE Information Theory Workshop*, pp. 165-169, 2004.
- [85] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388-404, Mar. 2000.
- [86] O. Oyman and S. Sandhu, "A Shannon-Theoretic Perspective on Fading Multihop Networks", in *Proc. IEEE CISS*, pp. 525-530, 2006.
- [87] I. Stanojev, O. Simeone, Y. Bar-Ness and M. Cho, "On the optimal number of hops in linear wireless ad hoc networks with hybrid ARQ", in *Proc. 6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops*, pp. 369-374, April 2008.
- [88] O. Oyman and J. N. Laneman, "Multi-hop Diversity in Wideband OFDM Systems: The Impact of Spatial Reuse and Frequency Selectivity", in *Proc. IEEE International Symposium on Spread-Spectrum Systems and Applications (ISSSTA '08)*, Bologna, Italy, August 2008.
- [89] O. Oyman and S. Sandhu, "Non-ergodic power-bandwidth tradeoff in linear multihop networks", in *Proc. IEEE ISIT*, pp. 1514-1518, Seattle, Jul. 2006.
- [90] S. Yang and J.-C. Belfiore, "Towards the Optimal Amplify-and-Forward Cooperative Diversity Scheme," *IEEE Transactions on Information Theory*, vol. 53, Issue 9, pp 3114-3126, Sept. 2007.
- [91] J. Boyer, D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Trans. Commun.*, vol. 52, pp. 1820-1830, Oct 2004.

- [92] K. Sreeram, S. Birenjith and P. Vijay Kumar, "Multi-hop Cooperative Wireless Networks: Diversity Multiplexing Tradeoff and Optimal Code Design," submitted Feb. 2008. (http://arxiv.org/PS_cache/arxiv/pdf/0802/0802.1888v1.pdf)
- [93] K. Sreeram, S. Birenjith, and P. Vijay Kumar, "Multi-hop cooperative wireless networks: Diversity multiplexing tradeoff and optimal code design," in *ITA Workshop*, 27 Jan.-1 Feb. 2008. U.C. San Diego, 2008 available on <http://ita.ucsd.edu/workshop/08>.
- [94] L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation", *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748-767, May 2004.
- [95] Z. Dawy, "Power allocation in wireless multihop networks with application to virtual antenna arrays", in *Proc. 15th International Symposium on Personal, Indoor, and Mobile Radio Communications*, pp. 1682-1688, Sep. 2004.
- [96] A. K. Sadek, W. Su and K. J. Liu, "Multinode cooperative communications in wireless networks," *IEEE Trans. Signal Processing*, vol. 55, no. 1, pp. 341-355, Jan. 2007.
- [97] P. Gupta and P. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1877-1894, Aug. 2003
- [98] F. Oggier and B. Hassibi, "Code Design for Multihop Wireless Relay Networks," accepted for publication in *EURASIP Journal on Advances in Signal Processing*. Available Online: <http://www.hindawi.com/journals/asp/raa.457307.html>, Nov. 2007.
- [99] M. Dohler, A. Gkelias and H. Aghvami, "A resource allocation strategy for distributed MIMO multi-hop communication systems," *IEEE Communication Letters*, vol. 8, no. 2, pp. 99-101, Feb 2004.
- [100] A. Stefanov and E. Erkip, "Cooperative space-time coding for wireless networks," *IEEE Trans. on Communications*, vol. 53, No.11, pp. 1804-1809, Nov. 2005.
- [101] A. Stefanov and E. Erkip, "Cooperative space-time coding for wireless networks," in *Proc. IEEE Information Theory Workshop*, La Sorbonne, Paris, France, April 2003.
- [102] H. El Gamal, A. R. Hammons and A. Stefanov, "Space-time overlays for convolutionally coded systems," *IEEE Trans. Commun.* vol. 51. no.9, pp 1603-1612, Sep.2003.
- [103] S. Borade, L. Zheng, and R. Gallager, "Amplify-and-forward in wireless relay networks: Rate, diversity, and network size," *IEEE Trans. On Inform. Theory*, vol. 53, pp. 3302-3318, Oct 2007.
- [104] C. Rao and B. Hassibi, "Diversity-Multiplexing Gain Trade-Off of a MIMO System with Relays," *IEEE Information Theory Workshop*, Norway, July 2007.
- [105] Y. Jing and B. Hassibi, "Diversity analysis of distributed spacetime codes in relay networks with multiple transmit/receive antennas," *EURASIP J. Advances in Signal Processing*, vol. 2008, article ID 254573, 17 pages, 2008 (doi:10.1155/2008/254573).

- [106] A. Adinoyi and H. Yanikomeroglu, "Cooperative Relaying in Multi-Antenna Fixed Relay Networks," *IEEE Trans. on Wireless Communications*, vol.6, no. 2, pp. 533-544, Feb 2007.
- [107] S. Yang and J.-C. Belfiore, "Diversity of MIMO Multihop Relay Channels," submitted to *IEEE Trans. on Inform. Theory*, Available Online: <http://arxiv.org/abs/0708.0386>, Aug. 2007.
- [108] R. Vaze and R. W. Heath Jr., "Maximizing reliability in multi-hop wireless networks with cascaded space-time codes," *Proceedings of Information Theory and Applications Workshop*, UCSD, Feb. 2008.
- [109] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multi-hop wireless networks," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '03)*, vol. 1, pp. 702-711, 2003.
- [110] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Visvanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, "Relay-based Deployment Concepts for Wireless and Mobile Broadband Cellular Radio." *IEEE Communications Magazine*, September 2004.
- [111] E. G. Larsson and B. R. Vojcic, "Cooperative transmit diversity based on superposition modulation," *IEEE Communications Letters*, vol. 9, no. 9, pp. 778-780, Sep. 2005.
- [112] K. Ishii, "Superposition modulated cooperative diversity for halfduplex scenario", *IEEE Journal of Communications*, vol. 2, 20-27, Dec., 2007.
- [113] M. Yuksel and E. Erkip, "Multi-antenna cooperative wireless systems: a diversity multiplexing tradeoff perspective," *IEEE Trans. Inf. Theory*, vol. 53. pp. 3371-3393, Oct. 2007.
- [114] A. H. Madsen, J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Info Theory*, vol.51, Jun. 2005.
- [115] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits and Systems*, vol.25, no. 9, pp. 772-781, Sep. 1978.
- [116] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [117] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [118] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.
- [119] D. P. Bertsekas, *Nonlinear Programming*, 2nd. edition. Belmont, MA: Athena Scientific, 1995.

- [120] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Letters*, Sept. 2001.

Appendix A

Proof of Optimal Trainings for Phase 1

Denote the eigenvalue decompositions (EVD) of \mathbf{C}_{t_2} and \mathbf{C}_{r_2} as $\mathbf{C}_{t_2} = \mathbf{U}_{t_2} \mathbf{\Lambda}_{t_2} \mathbf{U}_{t_2}^H$ and $\mathbf{C}_{r_2} = \mathbf{U}_{r_2} \mathbf{\Lambda}_{r_2} \mathbf{U}_{r_2}^H$ with descending eigenvalues. We can then write $\mathbf{C}_{t_2}^{\frac{1}{2}} = \mathbf{U}_{t_2} \mathbf{\Lambda}_{t_2}^{1/2}$ and $\mathbf{C}_{r_2}^{\frac{1}{2}} = \mathbf{U}_{r_2} \mathbf{\Lambda}_{r_2}^{1/2}$. We also write $\mathbf{C}_0 = \mathbf{C}_1 \otimes \mathbf{C}_2 = \mathbf{\Lambda}_1 \otimes \mathbf{\Lambda}_2$. If $\mathbf{C}_0 = \mathbf{I}$, both $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are the identity matrices. If $\mathbf{C}_0 = \mathbf{C}_{t_2}^{\frac{H}{2}} \mathbf{C}_{t_2}^{\frac{1}{2}} \otimes \mathbf{C}_{r_2}^{\frac{H}{2}} \mathbf{C}_{r_2}^{\frac{1}{2}}$, we have equivalently $\mathbf{\Lambda}_1 = \mathbf{\Lambda}_{t_2}$ and $\mathbf{\Lambda}_2 = \mathbf{\Lambda}_{r_2}$.

It then follows from (4.12) that

$$\begin{aligned}
 L &\doteq J'_2 + \mu \cdot \text{tr}(\mathbf{C}_{S_R}) \\
 &= \text{tr} \left\{ (\mathbf{\Lambda}_1 \otimes \mathbf{\Lambda}_2) \left[\mathbf{I} + \mathbf{\Lambda}_{t_2}^{1/2} \mathbf{U}_{t_2}^H \mathbf{C}_{S_R} \mathbf{U}_{t_2} \mathbf{\Lambda}_{t_2}^{1/2} \otimes \mathbf{\Lambda}_{r_2} \right]^{-1} \right\} + \mu \cdot \text{tr}(\mathbf{C}_{S_R}) \\
 &= \sum_{i=1}^{n_D} \lambda_2(i) \text{tr} \left\{ \mathbf{\Lambda}_1 \left(\mathbf{I} + \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} \mathbf{C} \mathbf{\Lambda}_{t_2}^{1/2} \right)^{-1} \right\} + \mu \cdot \text{tr}(\mathbf{C})
 \end{aligned} \tag{A.1}$$

where $\mathbf{C} = \mathbf{U}_{t_2}^H \mathbf{C}_{S_R} \mathbf{U}_{t_2}$, which has not yet been shown to be diagonal.

It follows from the generalized KKT conditions that the solution to (4.13) satisfies the sufficient and necessary conditions: $\frac{\partial L}{\partial \mathbf{C}^H} \geq \mathbf{0}$, $\mathbf{C} \geq 0$, $\mu > 0$ and $\text{tr}(\mathbf{C}) = P_R$, which is easy to prove by using (5.95) in [117].

To derive $\frac{\partial L}{\partial \mathbf{C}^H}$, we will use $\partial(\mathbf{A}\mathbf{X}\mathbf{B}) = \mathbf{A}\partial\mathbf{X}\mathbf{B}$, $\partial(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}\partial\mathbf{X}\mathbf{X}^{-1}$ and $\text{tr}(\mathbf{A}\mathbf{B}) = \text{tr}(\mathbf{B}\mathbf{A})$. Then, we have

$$\begin{aligned} \partial L = & - \sum_{i=1}^{n_D} \lambda_2(i) \text{tr} \left\{ \mathbf{\Lambda}_{t_2}^{1/2} \left(\mathbf{I} + \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} \mathbf{C} \mathbf{\Lambda}_{t_2}^{1/2} \right)^{-1} \mathbf{\Lambda}_1 \left(\mathbf{I} + \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} \mathbf{C} \mathbf{\Lambda}_{t_2}^{1/2} \right)^{-1} \right. \\ & \left. \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} \partial \mathbf{C} \right\} + \mu \cdot \text{tr}(\partial \mathbf{C}) \end{aligned} \quad (\text{A.2})$$

Recall that if $\partial L = \text{tr}(\mathbf{A}\partial\mathbf{X})$, then $\frac{\partial L}{\partial \mathbf{X}^H} = \mathbf{A}$ [19]. Therefore,

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{C}^H} = & - \sum_{i=1}^{n_D} \lambda_2(i) \mathbf{\Lambda}_{t_2}^{1/2} \left(\mathbf{I} + \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} \mathbf{C} \mathbf{\Lambda}_{t_2}^{1/2} \right)^{-1} \\ & \mathbf{\Lambda}_1 \left(\mathbf{I} + \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} \mathbf{C} \mathbf{\Lambda}_{t_2}^{1/2} \right)^{-1} \lambda_{r_2}(i) \mathbf{\Lambda}_{t_2}^{1/2} + \mu \mathbf{I} \end{aligned} \quad (\text{A.3})$$

It is easy to observe from (A.3) that for any $\mu > 0$, there is always a diagonal $\mathbf{C} \geq 0$ such that $\frac{\partial L}{\partial \mathbf{C}^H} \geq 0$. Therefore, the optimal \mathbf{C} is diagonal.

Let $\mathbf{C} = \text{diag}(c(1), \dots, c(n_R)) \geq 0$. It follows from (A.3) that for all $j = 1, 2, \dots, n_R$,

$$\left(\frac{\partial L}{\partial \mathbf{C}^H} \right)_{j,j} = \mu - \lambda_1(j) \lambda_{t_2}(j) \sum_{i=1}^{n_D} \frac{\lambda_2(i) \lambda_{r_2}(i)}{[1 + \lambda_{r_2}(i) \lambda_{t_2}(j) c(j)]^2} \geq 0 \quad (\text{A.4})$$

where $\mu > 0$ is such that $\text{tr}(\mathbf{C}) = P_R$.

Appendix B

Proof of Theorem 1

From (4.28), (4.29), and (4.30), it is obvious that the cost and the constraints in the problem (4.22) is invariant to \mathbf{U}_S and hence any unitary \mathbf{U}_S is optimal.

To prove the optimality of the choice of $\mathbf{Q}_S = \mathbf{I}$, we need the following definitions and lemmas from [118].

DEFINITION 1 [20, 1.A.1]: Consider any two real-valued $N \times 1$ vectors \mathbf{x} , \mathbf{y} , and let $x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[N]}$ and $y_{[1]} \geq y_{[2]} \geq \dots \geq y_{[N]}$ denote the elements of \mathbf{x} and \mathbf{y} , respectively, sorted in decreasing order. Then \mathbf{x} is said to be majorized by \mathbf{y} , denoted as $\mathbf{x} \prec \mathbf{y}$, if $\sum_{i=1}^n x_{[i]} \leq \sum_{i=1}^n y_{[i]}$, $n = 1, 2, \dots, N-1$ and $\sum_{i=1}^N x_{[i]} = \sum_{i=1}^N y_{[i]}$.

DEFINITION 2 [20, 1.A.2]: Using the same notations as in DEFINITION 1, \mathbf{x} is said to be weakly majorized by \mathbf{y} , denoted as $\mathbf{x} \prec_w \mathbf{y}$, if $\sum_{i=1}^n x_{[i]} \leq \sum_{i=1}^n y_{[i]}$, $n = 1, 2, \dots, N$.

LEMMA 1 [20, 9.H.1.h]: For two $N \times N$ positive semidefinite Hermitian matrices \mathbf{A} and \mathbf{B} with eigenvalues $\lambda_{a,i}$ and $\lambda_{b,i}$, $i = 1, \dots, N$, arranged in the descending order

respectively, it follows that $tr(\mathbf{A}\mathbf{B}) \geq \sum_{i=1}^N \lambda_{a,i} \lambda_{b,N+1-i}$.

LEMMA 2 [20, 9.B.1]: For a Hermitian matrix \mathbf{A} with the vector $\mathbf{d}[\mathbf{A}]$ of its main diagonal elements (in descending order for convenience) and the vector $\boldsymbol{\lambda}[\mathbf{A}]$ of its eigenvalues (in descending order for convenience), it follows that $\mathbf{d}[\mathbf{A}] \prec \boldsymbol{\lambda}[\mathbf{A}]$.

LEMMA 3 [20, 9.H.2]: For m $N \times N$ complex matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$, let $\mathbf{B} = \mathbf{A}_1 \mathbf{A}_2 \dots \mathbf{A}_m$, then $\sigma_b \prec_w \sigma_{a,1} \odot \sigma_{a,2} \odot \dots \odot \sigma_{a,m}$, where σ_b and $\sigma_{a,i}, i = 1, \dots, m$, denote $N \times 1$ vectors containing the singular values of \mathbf{B} and \mathbf{A}_i arranged in the same order, respectively, and \odot denotes the Schur (element-wise) product of two vectors.

LEMMA 4 [20, 3.A.8]: For a real-valued function f , $\mathbf{x} \prec_w \mathbf{y}$ implies $f(\mathbf{x}) \leq f(\mathbf{y})$ if and only if f is increasing with respect to each variable and Schur-convex.

Recall that among the two constraints in the problem (4.22), only the first, or equivalently (4.29), depends on \mathbf{Q}_S . From LEMMA 1, we have

$$tr[\boldsymbol{\Lambda}_S \mathbf{Q}_S \boldsymbol{\Lambda}_{t_1}^{-1} \mathbf{Q}_S^H] = tr[\boldsymbol{\Lambda}_S (\mathbf{Q}_S \boldsymbol{\Lambda}_{t_1}^{-1} \mathbf{Q}_S^H)] \geq tr[\boldsymbol{\Lambda}_S \boldsymbol{\Lambda}_{t_1}^{-1}] \quad (\text{B.1})$$

where the equality holds when $\mathbf{Q}_S = \mathbf{I}$. Namely, for any given $\boldsymbol{\Lambda}_S$, the source consumes the least amount of power when $\mathbf{Q}_S = \mathbf{I}$.

For the cost J'_1 in (4.22), let us define

$$\begin{aligned} \mathbf{X} &= \left(\boldsymbol{\Lambda}_S \otimes \boldsymbol{\Lambda}_F + \mathbf{I} \otimes \boldsymbol{\Lambda}_F^{\frac{1}{2}} \mathbf{Q}_F \boldsymbol{\Lambda}_{r_1}^{-1} \mathbf{Q}_F^H \boldsymbol{\Lambda}_F^{\frac{1}{2}} + \mathbf{I} \right)^{-1} (\mathbf{I} \otimes \boldsymbol{\Lambda}_F^{\frac{1}{2}} \mathbf{Q}_F \boldsymbol{\Lambda}_{r_1} \mathbf{Q}_F^H \boldsymbol{\Lambda}_F^{\frac{1}{2}}) \\ \mathbf{Y} &= \boldsymbol{\Lambda}_S^{\frac{1}{2}} \mathbf{Q}_S \boldsymbol{\Lambda}_{t_1} \mathbf{Q}_S^H \boldsymbol{\Lambda}_S^{\frac{1}{2}} \otimes \mathbf{I} \end{aligned}$$

Then, from (4.28),

$$J'_1 - tr[\mathbf{C}_{t_1} \otimes \mathbf{C}_{r_1}] = -tr[\mathbf{X}\mathbf{Y}] \quad (\text{B.2})$$

where \mathbf{X} depends on \mathbf{Q}_F , and \mathbf{Y} depends on \mathbf{Q}_S .

It follows from LEMMA 2 and LEMMA 3 that

$$\mathbf{d}[\mathbf{XY}] \prec \boldsymbol{\lambda}[\mathbf{XY}] \prec_w \boldsymbol{\lambda}[\mathbf{X}] \odot \boldsymbol{\lambda}[\mathbf{Y}] \quad (\text{B.3})$$

It is known [118] that $\boldsymbol{\lambda}[\mathbf{A}] \prec_w \boldsymbol{\lambda}[\mathbf{A}']$ implies $\boldsymbol{\lambda}[\mathbf{A} \otimes \mathbf{I}] \prec_w \boldsymbol{\lambda}[\mathbf{A}' \otimes \mathbf{I}]$, and $\boldsymbol{\lambda}[\mathbf{B}] \prec_w \boldsymbol{\lambda}[\mathbf{B}']$ implies $\boldsymbol{\lambda}[\mathbf{A}] \odot \boldsymbol{\lambda}[\mathbf{B}] \prec_w \boldsymbol{\lambda}[\mathbf{A}] \odot \boldsymbol{\lambda}[\mathbf{B}']$. It then follows that $\boldsymbol{\lambda}[\mathbf{Y}] \prec_w \boldsymbol{\lambda}[\mathbf{Y}']$ where $\mathbf{Y}' = \boldsymbol{\Lambda}_S^{\frac{1}{2}} \boldsymbol{\Lambda}_{t_1} \boldsymbol{\Lambda}_S^{\frac{1}{2}} \otimes \mathbf{I}$ which is \mathbf{Y} when $\mathbf{Q}_S = \mathbf{I}$. Furthermore,

$$\mathbf{d}[\mathbf{XY}] \prec \boldsymbol{\lambda}[\mathbf{XY}] \prec_w \boldsymbol{\lambda}[\mathbf{X}] \odot \boldsymbol{\lambda}[\mathbf{Y}'] \quad (\text{B.4})$$

Since $tr(\cdot)$ is increasing and Schur-convex function of $\mathbf{d}[\mathbf{XY}]$, from LEMMA 4, we have

$$tr(\mathbf{XY}) \leq tr(\boldsymbol{\Lambda}[\mathbf{X}] \boldsymbol{\Lambda}[\mathbf{Y}']) \quad (\text{B.5})$$

where $\boldsymbol{\Lambda}[\mathbf{X}]$ and $\boldsymbol{\Lambda}[\mathbf{Y}']$ are diagonal matrices with the elements of $\boldsymbol{\lambda}[\mathbf{X}]$ and $\boldsymbol{\lambda}[\mathbf{Y}']$ being their diagonal values, respectively. From (B.2) and (B.5), J'_1 is minimized when $\mathbf{Q}_S = \mathbf{I}$.

The above discussions show that when $\mathbf{Q}_S = \mathbf{I}$, the source consumes the least amount of power and J'_1 is minimized. Therefore, we reach the conclusion that $\mathbf{Q}_S = \mathbf{I}$ is optimal.

Appendix C

Proof of Theorem 2

It is easy to observe from (4.28), (4.29) and (4.30) that when $\mathbf{C}_{r_1} = \alpha\mathbf{I}$ or equivalently $\mathbf{\Lambda}_{r_1} = \alpha\mathbf{I}$, the cost function and both constraints in (4.22) are invariant to \mathbf{Q}_F . Therefore, any unitary \mathbf{Q}_F is optimal.

We know from (4.28) and (4.29) that the cost and the first constraint of (4.22) are invariant to \mathbf{U}_F . Using $\mathbf{C}_{r_1} = \alpha\mathbf{I}$ and LEMMA 1, the left hand side of (4.30) can be written as

$$\begin{aligned} & tr\{\mathbf{\Lambda}_S\} tr\left\{\mathbf{\Sigma}_{H_2}^{-2}\mathbf{U}_{H_2}^H\mathbf{U}_F\mathbf{\Lambda}_F\mathbf{U}_F^H\mathbf{U}_{H_2}\right\} + \frac{L}{\alpha} \cdot tr\left\{\mathbf{\Sigma}_{H_2}^{-2}\mathbf{U}_{H_2}^H\mathbf{U}_F\mathbf{\Lambda}_F\mathbf{U}_F^H\mathbf{U}_{H_2}\right\} \\ & \geq tr(\mathbf{\Lambda}_S)tr(\mathbf{\Sigma}_{H_2}^{-2}\mathbf{\Lambda}_F) + \frac{L}{\alpha}tr(\mathbf{\Sigma}_{H_2}^{-2}\mathbf{\Lambda}_F) \end{aligned} \tag{C.1}$$

where the lower bound is achieved when $\mathbf{U}_F = \mathbf{U}_{H_2}$.