國 立 交 通 大 學

電子工程學系 電子研究所

博 士 論 文

分散式放大前送多輸入多輸出中繼站網路之設計

Design of Distributed Amplify-Forwarding MIMO Relay Networks

研 究 生：吳俊榮

指導教授：林大衛

中 華 民 國 一 百 年 七 月

分散式放大前送多輸入多輸出中繼站網路之設計

Design of Distributed Amplify-Forwarding MIMO Relay Networks

研 究 生：吳俊榮　　Student：Chun-Jung Wu

指導教授：林大衛　　Advisor：David W. Lin

國 立 交 通 大 學

電子工程學系 電子研究所

博 士 論 文

A Dissertation

Submitted to Department of Electronics Engineering and

Institute of Electronics

College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Electronics Engineering

July 2011

Hsinchu, Taiwan, Republic of China

中華民國一百年七月

分散式放大前送多輸入多輸出中繼站網路之設計

研究生：吳俊榮　　指導教授：林大衛博士

國立交通大學
電子工程學系　電子研究所

摘要

多進多出系統(MIMO)搭配中繼站之應用可提升系統涵蓋率或提高系統容量。採用分散式平行單天線中繼站網路則額外享有建置彈性、多集性與成本之優勢，然而亦面臨因結構上缺乏中繼站相互連通而衍生的設計議題。本論文針對分散式放大-前送(amplify-forwarding)中繼站型態，首先考慮提升中繼站容量之中繼前送增益(forwarding)設計。在設計中數學結構的限制下，目前並無實用有效之快速設計方式，因此我們採用矩陣低階更新(matrix low-rank updating)的方式，設計遞迴式最佳化解法。此解法雖可快速計算並提供提高系統容量之設計，對於分散式中繼站之特性研究並無實質貢獻，因此我們設計雜訊優勢模型以簡化系統，並據此提除對應之增益設計方法。此設計方法搭配中繼站部分啟用之模式，我們發現此設計方法可類比於單站多天線選擇系統(single-hop MIMO antenna selection)。因此我們可採用多天線選擇系統之已知特性，類比分析我們針對中繼站所提出之設計方法。由模擬結果可知，我們所設計之設計方法可有效提升系統容量，且印證類比於多天線選擇系統之容量分佈特性。

當中繼站的數量增加時，基於中繼站選擇之設計方法將面臨運算量大幅上升之問題。同樣還是由優勢雜訊模型出發，我們設計新的增益最佳化運算並避免中繼站選擇。此演算法更進一步簡化並轉換設計問題，並使用修正之解題空間(solution space)。應用於多用戶系統之低散逸波束設計亦包含於我們所設計之演算法。由

模擬結果可知，免除中繼站選擇之演算法，相較於使用中繼站選擇方法，其效能相若或更好。

在第三部分我們討論針對高度散佈(highly dispersive)通道之分頻正交多工系統(OFDM)之通道估計，以及其應用於使用 OFDM 之中繼站網路。對於 OFDM 通道估計問題，時域(time-domain)通道估計方法可利用有限的領航訊號(pilot symbols)提供精確之通道估計。然而此類方法須掌握通道多路徑之延遲位置(multipath delay positions)。已知針對延遲位置估計的方法需要特定格式之領航訊號，且在通過慢速衰落(slow fading)通道時可能會降低估計效能。我們所設計之延遲位置估計，採用追求最大程度相合(matching pursuit)之演算法，考慮相鄰 OFDM 訊框中領航訊號，且不限制領航訊號的位置。經由模擬結果可知，我們所設計之延遲位置估計方法，可大幅提升通道估計之準確率。另外我們討論採用 OFDM 之放大前送中繼站其對應之傳輸通道模型，經由分析得知可對應於一高度散佈之傳輸通道，因此適用於本文所提出之估計方法。

Design of Distributed Amplify-Forwarding MIMO Relay Networks

Student: Chun Jung Wu Advisor: Dr. David-W. Lin

Department of Electronics Engineering
& Institute of Electronics
National Chiao Tung University

Abstract

Relays can potentially enhance the transmission performance of multi-input multi-output (MIMO) systems. A parallel single-antenna relay network has additional advantages in flexibility, diversity, and cost, but also poses significant design problems because the absence of inter-antenna connections over different relays makes the underlying mathematical problems much more difficult to solve. In this thesis, we first consider the design of parallel amplify-and-forward relay networks. More specifically, we focus on the design of relay gains to maximize the system capacity. As no closed-form analytic solution can be found, we first develop an iterative algorithm based on low-rank updating to efficiently find a locally optimal solution. Since iterative optimization provides little insight into the analytical properties of the solution, we attempt analytical solutions based on asymptotic noise conditions and relay selection. It is discovered that the proposed algorithms could be modeled as single-hop MIMO antenna selection systems and could be analyzed in the similar way. We examine the resulting capacity outage diversity orders and confirm the analysis and equivalent modeling with simulation results.

As the algorithms based on relay selection may encounter expensive computation burden as number of relays increases, we develop algorithms based on noise dominant models but avoid computational intensive relay selection. The proposed algorithms are made possible by modified criterion and specifically constrained solution space. Low-leakage beamforming used for multiuser communication is applied in our algorithm. Numerical results demonstrate the proposed algorithms exhibit comparable or better performance than previous selection-based approaches.

In the third part of this thesis, we consider OFDM channel estimation with highly dispersive channel and application in distributed relay network. Time-domain channel estimation techniques have been proposed for OFDM systems for their ability to yield relatively accurate estimates with only a few pilots. Key information needed in such techniques is the multipath delays of the channel. Prior approaches to estimation of multipath delays require regular pilot structures and may not work in slow fading. We propose a group matching pursuit technique for channel estimation. The technique is an extension of the orthogonal matching pursuit technique. It employs the pilots in several OFDM symbols to estimate the multipath delays in a sequential manner, where the pilots can have an arbitrary structure. Simulation results show that the proposed algorithm has superior performance. We then demonstrate that distributed AF relay network with OFDM transmission could be modeled as single-hop OFDM system with equivalently high dispersive delay profile. Thus we apply and examine the proposed channel estimation schemes for the relay network of interest.

誌　謝

這份論文與漫長的修業得以完成，首先要感謝我的指導教授林大衛博士。林教授在這一路艱辛的過程中，提供我全方位的支持，在我的研究進度陷入困頓之時，還是不斷的鼓勵與協助。林教授長期的栽培與指導，對於我的研究與學位是最關鍵的因素，對於我個人則是永誌不忘的師恩。電子系桑梓賢教授亦提供諸多寶貴意見，並協助提升研究的品質與成果。在兩位老師的提攜之下，我才得以在專業的道路上徐緩學步前行。

我的家人與玉如在我成長與學業的過程中，以永恆無盡的包容與關懷，陪伴我走過幽暗，支持我奮力向前。親情的力量與偉大無以名狀，如果個人幸得尺寸之功，所有的成果皆榮歸於我的親人。

Contents

1 Introduction **1**

1.1 Research Background, Motivation and Issues 1

1.1.1 On the Capacity of Distributed AF MIMO Relay Network 1

1.1.2 On the OFDM Channel Estimation and Applications in Relay Networks 3

1.2 Outline of This Thesis 5

1.3 Contribution of This Thesis 6

2 Capacity of Relay Network and Alternating Optimization **8**

2.1 System Model and Capacity of Relay Networks 8

2.1.1 System Model and Power Limits 8

2.1.2 Noise Whitening and Capacity 10

2.2 Power Scaling and Upper bounds of Capacity 10

2.3 Iterative Alternating Optimization 13

2.3.1 Low-rank Updating for Capacity Computation 14

2.3.2 Optimization of α and β 17

2.3.3 Numerical Results 19

3 Capacity Improvement based on Noise-Dominant Models **21**

3.1 Preliminary: MIMO Antenna-Selection Systems 22

3.2 Designs for Destination-Noise Dominant Conditions 23

3.3 Designs for Relay-Noise Dominant Conditions 26

3.4 Numerical Results 28

4 Efficient Algorithms for Noise-Dominant Models **32**

4.1 Designs for Relay-Noise Dominant Conditions 33
4.1.1 Downlink Low-leakage Beamforming Design 33
4.1.2 Capacity Approximation and Relay Selection 34
4.1.3 Capacity Improvement based on Maximizing the Ratio of Norms . . 36
4.1.4 Algorithms Summary and Complexity Comparison 36
4.2 Designs for Destination-Noise Dominant Conditions 38
4.2.1 Capacity Approximation and Relay Selection 38
4.2.2 Capacity Improvement based on Partial Zero Forcing 38
4.2.3 Iterative Greedy Optimization 41
4.2.4 Algorithms Summary and Complexity Comparison 42
4.3 Numerical Results 44

5 Channel Estimation for Distributed Relay Networks with OFDM Transmission **47**
5.1 Matching Pursuit Algorithms for OFDM Channel Estimation 47
5.1.1 OFDM Transmission System 47
5.1.2 Time-Domain Approach to Channel Estimation 48
5.1.3 Estimation of Multipath Delays 48
5.2 Relay System Model with OFDM Transmission 52
5.3 Numerical Results 54

6 Conclusion **58**

Bibliography **60**

List of Figures

2.1 MIMO system with distributed relays. 9

2.2 Progression of average capacity with number of iterations. 20

2.3 Average of normalized capacity with respect to varying β values 20

3.1 Comparison of MIMO antenna selection and full MIMO systems in terms of CDF of capacity (top) and biased capacity (bottom). 24

3.2 Capacity CDF of distributed relay system designed under destination noise-dominant assumption. 29

3.3 Horizontally "biased" capacity CDF curves of distributed relay systems for diversity comparison. 30

3.4 Capacity CDF of distributed relay system designed under relay noise-dominant assumption. 30

3.5 Comparison of diversity orders of capacity CDFs of distributed relay systems of different sizes designed under relay noise-dominant assumption. . . 31

4.1 Downlink signal and interference flows for user 1 34

4.2 Downlink signal and leakage flows from user 1 35

4.3 Simulate under relay-noise dominating condition. 45

4.4 Simulate under destination-noise dominating condition. 45

4.5 Simulation with two noise terms having equal power level 46

5.1 In the pth iteration,projection and selection based on projection residual . 51

5.2 Relay operations for OFDM transmission 53

5.3 Normalized mean-square channel estimation errors of different channel estimation methods 55

5.4 Average symbol error rates (SERs) at data subcarriers with different channel estimation methods. 56

5.5 Average SERs at data subcarriers with different channel estimation methods when multipath delay are spaced apart. 56
5.6 Average SERs of OFDM distributed relay network at data subcarriers with different channel estimation methods when multipath delay are randomly spaced. 57

Chapter 1

Introduction

1.1 Research Background, Motivation and Issues

1.1.1 On the Capacity of Distributed AF MIMO Relay Network

Relays have been considered a useful means for coverage extension and capacity enhancement of wireless systems [33]. Among all conceivable relaying strategies, two have received the most attention: amplify-and-forward (AF) [30] and decode-and-forward (DF) [31]. In AF systems the relays amplify or beamform the received signals without further processing, while in DF systems they decode (or demodulate if there is no channel coding) the received signals and transmit the re-encoded (or remodulated) signals to the destination. Besides the forwarding strategies, an important subject in relay system design is the overall wireless system architecture. In this, due to the capacity advantage of multi-input multi-output (MIMO) transmission over single-input single-output (SISO) transmission, many have sought to incorporate some MIMO concepts one way or another. The present work is concerned with AF-based distributed relay networks, whose architecture will be described further later.

The simplest relay-aided transmission system consists of three nodes: source, relay (cooperator) and destination [44]. To facilitate MIMO transmission, an intuitive approach is to install multiple antennas on one or more of the nodes. For simplicity, consider the situation where the source and the destination have an equal number of antennas. A case with a single-antenna source (SAS) and a single relay (SR) equipped with multiple antennas (MAR) is considered in [48]. A natural extension to have a multiple-antenna source (MAS) and an SR-MAR to enable spatial multiplexing [25, 6]. In studies of MAS-SR-MAR systems, the multiple antennas on each terminal are usually assumed to be fully connected and may have arbitrary interconnection weights. In this case, known matrix theory can be used to decompose a MIMO transmission channel into parallel SISO links (via, for example, the singular value decomposition (SVD) or the QR decomposition).

Each spatially multiplexed signal stream can then be transported over one parallel link, and the matrix decomposition can be viewed as simultaneous beamforming for these streams. Typical performance measures, such as the signal-to-interference-plus-noise ratio (SINR) or the mean-square error (MSE) in received signal values, can be expressed in terms of the parameters of the decomposed channel. System optimization may then become essentially a problem of power allocation among the individual streams [25, 6].

On the other hand, use of multiple, parallel relays (PR) has also been considered by many researchers and shown to be potentially beneficial in various aspects [5, 21, 40, 22, 13, 9, 12, 24, 26, 11, 15, 47, 20, 8, 3]. For example, it is found that an increased number of relays can benefit the system capacity [5]. In fact, the parallel relays can function as virtual transmitter antennas and effect transmitter diversity either in the form of distributed space-time coding [21, 40] or in the form of distributed beamforming [22, 13, 9]. The corresponding diversity order has been examined in [22] and [12], respectively. Moreover, parallel relaying has also been studied in the contexts of sensor networks [24], two-way relaying [26], and secrecy communication [11]. However, despite the potential benefits, the fact that the relays are not connected but stand in parallel raises a cooperation problem which, if not dealt with, could severely limit the realizable benefit.

To see why, let L be the number of parallel relays and N_i $(1 \leq i \leq L)$ the number of antennas on relay i. Let M denote the number of antennas on the source terminal as well as that on the destination terminal. Consider first the simplest case where each terminal has only one antenna, i.e., SAS-PR-SAR where $M = N_i = 1\ \forall i$ [13, 9]. In this case, the relays effectively constitute a distributed beamformer for the single signal stream. Applying the same design philosophy to an MAS-PR-MAR system with $M > 1$ and $N_i > 1\ \forall i$, there can be $M_S = \min\{M, N_i\ \forall i\}$ concurrent signal streams. The beamforming techniques used in MAS-SR-MAR systems can be extended to this scenario with a twist [15, 47]. That is, the available antennas on the relays can be used to provide M_S parallel subchannels between the source and the destination. Systems operating in the above ways have been considered in some works [48, 13, 9, 15, 47, 25, 6]. In terms of capacity, however, such systems suffer from two consequences. First, the number of supported subchannels (i.e., the number of concurrent spatially multiplexed streams) does not grow with the relay number L, but is upper-bounded by M_S. Secondly, to increase the number of streams we need to ensure that all relays are equipped with sufficient antennas. Designs that can obviate the above limits are of interest and importance.

In this work we consider the design of MAS-PR-SAR systems (where $N_i = 1$ and $\sum N_i = L$) to support multiple signal streams. More specifically, we consider the design of AF relay forwarding gains for maximization of system capacity. Previously, Jin *et al.* [20] considered the case where the relays had equal gain and analyzed the statistics of the resulting ergodic capacity. Chen *et al.* [8] considered the minimization of transmission power subject to per-stream SINR targets. The problem is related to system capacity, but somewhat indirectly. Bae and Lee [3] proposed algorithms for capacity optimization under the condition that the product of the source-to-relay and the relay-to-destination

channel matrices was asymptotically diagonal in the limit of a large number of relays. But in sum, there is as yet no extensive work on the design of distributed parallel relay networks for capacity maximization. Actually, the relationship between number of relay terminals and system capacity also needs to be further clarified. The present work is motivated by these observations.

We consider two approaches to maximizing the capacity of a distributed relay network with presence of perfect channel state information (CSI). The first is algorithmic, as so far no closed-form solution to the problem exists. However, although algorithmic optimization can yield good results, it provides little insight into the analytical properties of the solutions. We thus also attempt an analytical approach. Because no closed-form solution can be obtained for the general situation, we consider two asymptotic situations which are more amenable to analysis. In one of them the relay noise dominates the overall noise in the received signal at the destination and in the other the destination terminal noise dominates. Alternatively, these two situations can also be viewed as providing two upper bounds to the system capacity.

Given the simplification lead by noise-dominant models, it is still a challenging task to optimize the relay network performance. Instead of considering all relays simultaneously, we discover that closed form optimization is possible if only partial of relays are active. In consequence, we develop algorithms for capacity improvement which work with relay selection. In addition, we then observe the proposed selection-based algorithms could be modeled and analyzed with equivalent single-hop MIMO antenna selection systems [16, 17]. Given the results and analysis of existing works in the area of MIMO antenna selection, we gain more in depth understanding about the application of proposed algorithms for distributed relay networks.

As the number of relays increases, it is found that both capacity and outage diversity order increases. However, the proposed algorithms based on relay selection may become intractable as the number of relay terminals is large. To counteract the situation, we design further simplifications by modified criterion and shrunken solution space. Though being suboptimal in nature, the proposed algorithms avoid computational intensive operation and could fully utilize the whole relay network. Simulations confirm the proposed approaches could result in slightly inferior or better performance than selection-based schemes, with considerably smaller computation cost as number of relays is large.

1.1.2 On the OFDM Channel Estimation and Applications in Relay Networks

Coherent demodulation of orthogonal frequency-division multiplexing (OFDM) signals critically depends on proper channel estimation. Since OFDM systems usually reserves some subcarriers as pilots, most channel estimation methods are pilot-aided. A common approach is to estimate the channel frequency response at pilot locations first, and

then "extend" the estimate to other subcarrier locations. One frequently considered way of "extension" is low-order polynomial interpolation, which can take the form of one-dimensional interpolation in the frequency domain (within the boundary of one OFDM symbol) or two-dimensional interpolation over frequency and time (across several OFDM symbols) [23], [7]. The performance of this sort of methods is limited by the pilot density and the channel characteristics. For example, if the channel has small coherence bandwidth (i.e., long delay spread) and low coherence time (e.g., due to fast motion) and the pilots are widely spaced in frequency, then they would have difficulty obtaining accurate channel estimates.

Another way of "extension" is based on exploiting the time-domain characteristics of the channel [29]. Since, in many cases, only a few multipaths contribute significantly to the channel response (in other words, the channel is "sparse"), the unknowns in time-domain estimation (which consist of the path coefficients of the significant multipaths if their delays are known) are usually much fewer than that in frequency-domain-based interpolation (which consist of the frequency response at all subcarriers). Hence the few pilots can be put to better use and result in more accurate channel estimates. This is especially the case when the pilots are very few and very widely spaced (as, for example, in the case of the IEEE 802.16-2004 OFDMA uplink [19]).

Evidently, a fundamental issue in time-domain channel estimation is to find the delays of the significant multipaths. In [46], an effective delay acquisition technique is developed, but the pilots need to be equally spaced. In [32], the MUSIC algorithm widely used for spectrum analysis is employed for channel estimation, but again assuming equally spaced pilots. The algorithm can be easily extended to deal with irregular pilot spacings, but the pilot locations in the multiple OFDM symbols used in channel estimation should be identical. To the best of our knowledge, there is no time-domain channel estimation technique proposed so far that makes use of arbitrarily organized pilots in multiple OFDM symbols in the presence of channel fading.

In this work, we propose an effective technique for time-domain sparse channel estimation based on the matching pursuit (MP) approach. MP algorithms have been used in audio and video signal processing to select bases of subspaces [27], [1]. We extend the prior MP method for multipath delay estimation by jointly considering a group of OFDM symbols; thus we term the proposed algorithm a group MP (GMP) algorithm. And we design the algorithm such that it can deal with arbitrary pilot assignment that may vary from OFDM symbol to OFDM symbol. In [43], [34] the similar MP based processing for multiple measurement vectors are discussed. Note that the *dictionary*, or the range of basis searching, in these works is unique. In the scenario of this contribution, however, there are multiple reference dictionaries due to arbitrary pilot assignment.

The proposed subspace-based approaches for OFDM channel estimation is effective to highly dispersive frequency-selective channels with limited resource of pilots, which is also an potential threat to distributed relays with OFDM transmission. For most works on relay systems, flat-fading channels are the presumed channel models. However in

practice frequency selective channels would pose substantial effects to relay networks We demonstrate that for AF distributed relays, the end-to-end OFDM transmission could be modeled as equivalent single-hop OFDM with a delay profile composed of summation of multiple convolution, and consequently exhibits severe channel dispersion. We would apply the proposed schemes for the relay transmission and examine the performance.

1.2 Outline of This Thesis

In Chap. 2, we first briefly describe the system model of amplify-forwarding distributed relay network used in this work, then present capacity (i.e. mutual information) of this model and the associated optimization problem. To simplify the optimization and gain some insights into system design, we discuss the effects of relay transmission power scaling which ultimately leads to an upper bound of capacity. As our first approach to capacity optimization, in the third part of this chapter we present the alternating optimization. Specifically, this approach applies successive greedy optimization, low-rank updating of matrix computation, and quadratic approximation. Finally some numerical results are shown to confirm the performance of proposed approach.

While alternating optimization based on low-rank updating discussed in Chap. 2 can yield good results, it provides little insight into the analytical properties of the solutions. We thus consider an analytical approach in Chap. 3. In particular, note that in AF systems the receiver noise arises from two sources: the relay noise $\boldsymbol{n}_R$ and the destination terminal noise $\boldsymbol{n}_D$. The design problem becomes mathematically more tractable when one of the two dominates in the overall receiver noise so that the other may be ignored. We term the simplification as noise-dominant models. Algorithms to design relay gains corresponding to the models are discussed in this chapter. We discover that the proposed algorithm could be linked to equivalent single-hop MIMO antenna selection system. Thus we would first briefly review the important results and findings on single-hop MIMO antenna selection systems. For analytical insights we show how to model equivalent MIMO antenna selection systems. Finally some numerical results are shown for performance evaluation and validating the links between equivalent MIMO antenna selection and proposed relay selection schemes.

We discuss system simplification and approximation of distributed relay network based on noise-dominant models in Chap. 3, wherein relay selection algorithms and associated relay gains designs are also presented. Though the algorithms based on relay selection is conceptually concise, the computation burden would growth exponentially as the number of relays increases. To circumvent the problem, in Chap. 4 we discuss algorithms based on noise-dominant models but avoid relay selection. For relay-noise dominant model we cast the problem into projections with two subspace and minimizing the ratio of projected vector norms. In Sec. 4.1.1 the idea and algorithm for multiuser low-leakage beamforming [35] are briefly reviewed. We would apply the algorithm to solve minimization of

norms in Sec. 4.1.3. As for destination-noise dominant model, in Sec. 4.2.2 we simplify the problem by making the end-to-end MIMO channel $\boldsymbol{H}$ in (2.3) an upper-triangular matrix so that the matrix determinant maximization in (2.4) could be approximated as product of diagonal terms of $\boldsymbol{H}$. Then in Sec. 4.2.3 we transform the design problem with specifically constrained solution space and propose iterative algorithm to reach local optimizer. Finally we summarize all the proposed algorithms (with and without relay selection), and compare the order of computation complexity for selection-based algorithms and proposed efficient designs. Numerical results of respective algorithms are shown and compared in Sec. 4.3.

In Chap. 5 we discuss subspace-based algorithms for OFDM transmission with highly dispersive delay profile and limited pilot resources. We propose MP based algorithms to reconstruct the delay subspace which enables efficient and accurate time-domain channel estimation. Then we present the channel modeling of distributed AF relay network with OFDM transmission. It turns out the relay network of interest actually suffer equivalent highly dispersive channel. Thus we propose the apply subspace-based algorithm to handle the channel estimation. Finally in Chap. 6 we provide some concluding remarks on this thesis work.

1.3 Contribution of This Thesis

We briefly outline the contribution of this work.

1. We propose iterative alternating optimization to improve the capacity of distributed AF MIMO relay network. The proposed algorithm is based on matrix low-rank updating and thus could efficiently generate optimized results.

2. Noise-dominant models is proposed to simplify system design on distributed AF MIMO relay network. We show the models serve as upper bounds of system capacity and lead to several efficient design algorithms.

3. Based on noise-dominant models and relay selection, we develop two corresponding relay gains designs. We also discover the performance of proposed algorithm could be modeled as equivalent MIMO antenna selection systems. Thus we could reply on the well-studied characteristics of MIMO antenna selection to approximate and analyze the proposed algorithms.

4. To circumvent the potential expensive computation burden of selection-based schemes, we further transform and simplify the design problem in order to consider and utilize all the relays simultaneously. For the relay noise dominant model, we develop algorithms based on low-leakage beamforming designs. As to destination noise dominant model, we propose to simplify the problem by matrix triangularization. An

algorithms with specifically constrained solution space is developed to iteratively improve the results.

5. To perform OFDM channel estimation with highly dispersive delay profile but with few pilot or training symbols, we apply subspace-based algorithm to efficiently utilize limited pilot resource. Since the unknown subspace is critical to this approach, we apply OMP and invent GMP to iterative reconstruct the subspace. The proposed MP based algorithms could bypass some limitations of other algorithms and utilize the information brought by consecutive OFDM symbols. Also we discover that AF distributed relay network with OFDM transmission could be modeled as single-hop OFDM with equivalent highly dispersive channel. Thus the proposed subspace-based algorithms is suitable to handle the channel estimation for the relay network of interest.

Chapter 2

Capacity of Relay Network and Alternating Optimization

In this chapter, we first briefly describe the system model of amplify-forwarding distributed relay network used in this work, then present capacity (i.e. mutual information) of this model and the associated optimization problem. To simplify the optimization and gain some insights into system design, we discuss the effects of relay transmission power scaling which ultimately leads to an upper bound of capacity. As our first approach to capacity optimization, in the third part of this chapter we present the alternating optimization. Specifically, this approach applies successive greedy optimization, low-rank updating of matrix computation, and quadratic approximation. Finally some numerical results are shown to confirm the performance of proposed approach.

2.1 System Model and Capacity of Relay Networks

2.1.1 System Model and Power Limits

We consider a distributed MIMO relay system consisting of one source terminal, one destination terminal and L single-antenna relay terminals. Both source and destination are equipped with M antennas. To preserve the degree of freedom in the end-to-end transmission, it is reasonable to assume $L \geq M$. Fig. 2.1 demonstrates a conceptual representation for the system of interest.

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ ($\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^M$) denote the signal vector transmitted from the source and received at the destination, respectively. Let $\boldsymbol{r} = [\boldsymbol{r}(1)\ \boldsymbol{r}(2) \cdots \boldsymbol{r}(L)]^T$ represent the forwarding gain vector of relay network, wherein the ith relay performs amplify-forwarding by multiplying the received signal with the complex gain $\boldsymbol{r}(i)$ and then transmitting the result. Signals in the system pass through two MIMO channels, $\boldsymbol{F}$ and $\boldsymbol{G}^H$, to go from the source to the destination, where $\boldsymbol{F} \in \mathbb{C}^{M \times L}$ and $\boldsymbol{G}^H \in \mathbb{C}^{L \times M}$ denote channel between des-

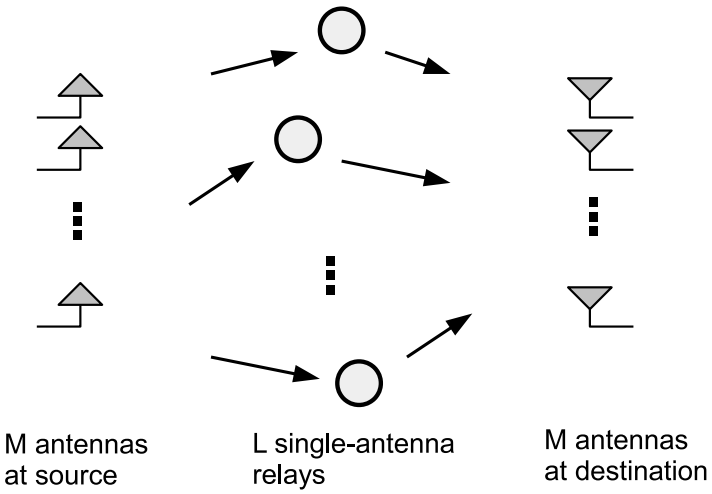

Fig. 2.1: MIMO system with distributed relays.

tination and relays and channel between relays and source, respectively. The superscript $()^H$ defines matrix Hermitian transpose. The received signals at the relays are assumed to be subject to additive complex circular white Gaussian noise $\boldsymbol{n}_R \sim CN(0, \sigma_R^2 \boldsymbol{I}_L)$, where $\boldsymbol{I}_L$ denotes an identity matrix of size $L \times L$. Likewise, the received signals at the destination are assumed to be subject to additive complex circular white Gaussian noise $\boldsymbol{n}_D \sim CN(0, \sigma_D^2 \boldsymbol{I}_M)$. Assume that there is no direct propagation path from the source to the destination. The received signal vector $\boldsymbol{y}$ at the destination can be described as

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{F}\,\mathrm{diag}(\boldsymbol{r})\boldsymbol{G}^H\boldsymbol{x} + \boldsymbol{F}\,\mathrm{diag}(\boldsymbol{r})\boldsymbol{n}_R + \boldsymbol{n}_D \\ &\triangleq \boldsymbol{F}\boldsymbol{R}\boldsymbol{G}^H\boldsymbol{x} + \boldsymbol{F}\boldsymbol{R}\boldsymbol{n}_R + \boldsymbol{n}_D \end{aligned} \tag{2.1}$$

where $\boldsymbol{R}$ is a diagonal matrix defined from $\boldsymbol{r}$, and $\mathrm{diag}(\boldsymbol{r})$ denotes a diagonal matrix whose diagonal terms are equal to $\boldsymbol{r}$.

To optimize the system capacity fairly, transmission power limits are imposed on the source and the relays. Specifically, we assume that the source transmits independent signal streams over its M antennas with equal power σ_x^2, subject to a total power limit of P_S. Further, the relays are subject to a total power limit of P_R. Mathematically, these constraints can be expressed as

$$\begin{aligned} P_S &\geq \mathrm{tr}(\mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^H\}) = M\sigma_x^2, \\ P_R &\geq \mathrm{tr}(\mathrm{E}\{(\boldsymbol{R}\boldsymbol{G}^H\boldsymbol{x} + \boldsymbol{R}\boldsymbol{n}_R)(\boldsymbol{R}\boldsymbol{G}^H\boldsymbol{x} + \boldsymbol{R}\boldsymbol{n}_R)^H\}) \\ &= \sum_{i=1}^{L}(\sigma_R^2 + \sigma_x^2|\boldsymbol{G}^{(i)}|^2)|\boldsymbol{r}(i)|^2 \triangleq \sum_{i=1}^{L} p(i)|\boldsymbol{r}(i)|^2, \end{aligned}$$

where $|\boldsymbol{G}^{(i)}|^2 = (\boldsymbol{G}^{(i)})^H\boldsymbol{G}^{(i)}$, with $\boldsymbol{G}^{(i)}$ representing the ith column of $\boldsymbol{G}$ and being associated with the channel vector between the source and the ith relay.

2.1.2 Noise Whitening and Capacity

Note that the noise vector $(\boldsymbol{FRn}_R+\boldsymbol{n}_D)$ in (2.1) as received by the destination is spatially correlated. To find the system capacity, a noise whitening filter $\boldsymbol{W}^{-1/2}$ at the destination is used, where $\boldsymbol{W}$ is given by

$$\begin{aligned}\boldsymbol{W} &= \mathrm{E}\{(\boldsymbol{FRn}_R+\boldsymbol{n}_D)(\boldsymbol{FRn}_R+\boldsymbol{n}_D)^H\} \\ &= \sigma_D^2(\boldsymbol{I}+\sigma^2(\boldsymbol{FR})(\boldsymbol{FR})^H),\end{aligned} \tag{2.2}$$

with $\sigma \triangleq \sigma_R/\sigma_D$. Let $\boldsymbol{H}$ denote the noiseless end-to-end equivalent MIMO channel, i.e., $\boldsymbol{H} \triangleq \sqrt{\sigma_x^2}\boldsymbol{FRG}^H$. Then we may derive the system capacity as a function of $\boldsymbol{R}$ as [41]

$$\begin{aligned}\log_2 \mathrm{C}(\boldsymbol{R}) &\triangleq \log_2 \det(\boldsymbol{I}_N+\boldsymbol{H}^H\boldsymbol{W}^{-1}\boldsymbol{H}) \\ &= \log_2 \det(\boldsymbol{W}+\boldsymbol{HH}^H)-\log_2 \det(\boldsymbol{W})\end{aligned} \tag{2.3}$$

where $\det(\cdot)$ denotes matrix determinant. The optimization problem can be stated as

$$\boldsymbol{R}_{opt} = \arg\max_{\boldsymbol{R}} \mathrm{C}(\boldsymbol{R}) \tag{2.4}$$

subject to

$$P_S \geq M\sigma_x^2,\ P_R \geq \sum_{i=1}^{L} p(i)|\boldsymbol{r}(i)|^2. \tag{2.5}$$

By observing (2.3) it is clear increasing source transmitting power or σ_x is always beneficial to capacity regardless of channel condition or the design of $\boldsymbol{R}$. Thus we may readily set $\sigma_x^2 = MP_S$. Further, the combination of σ_x^2 and $\boldsymbol{G}$ matrix in (2.3) could be replaced by a scaled version of $\boldsymbol{G}$ without lose any generality, so we ignore σ_x hereafter in this work.

2.2 Power Scaling and Upper bounds of Capacity

The inequality about P_S in (2.5) is discussed in last section. The remaining inequality power constraint naturally prompts one to think: is it possible to simplify the constraint by considering only the equality therein without impacting the optimality of the solution? Or, alternatively, given a certain $\boldsymbol{r}$ that satisfies (2.5) with inequality, will the system capacity be increased by scaling $\boldsymbol{r}$ to reach equality in (2.5)? Intuitively, the answer may seem to be a no-brainer as increasing the transmission power should be beneficial to the signal-to-noise ratio (SNR) and thus the capacity. However, because $\boldsymbol{R}$ affects both $\boldsymbol{H}$ and $\boldsymbol{W}$, it is not easy to intuitively determine in (2.3) the consequence of scaling $\boldsymbol{R}$ up. Thus a solid proof nevertheless requires some works. We state the result as a theorem and present the proof as follows.

Theorem 2.2.1 (Capacity scaling)**.** *When the (complex) relay gains* $\boldsymbol{R}$ *are scaled by* $s \in \mathbb{C}$ *with* $|s| > 1$, $\mathrm{C}(s\boldsymbol{R}) > \mathrm{C}(\boldsymbol{R})$.

Proof. First, it is clear in (2.3) that $\mathrm{C}(s\boldsymbol{R}) = \mathrm{C}(|s|\boldsymbol{R})$. Without loss of generality we assume $s \in \mathbb{R}^+$ (the set of positive real numbers) hereafter.

Consider a singular value decomposition of $\boldsymbol{FR}$ given by

$$\boldsymbol{FR} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^{\boldsymbol{H}} \tag{2.6}$$

where for convenience we let $\boldsymbol{\Lambda}$ be $M \times M$. Thus $\boldsymbol{U} \in \mathbb{C}^{M\times M}$ is the matrix of left singular vectors as usual, but the matrix of right singular vectors $\boldsymbol{V}$ becomes $L \times M$, that is, $\boldsymbol{V} \in \mathbb{C}^{L\times M}$. Further, let the singular values along the diagonal of $\boldsymbol{\Lambda}$ be arranged in descending numerical order. Let λ_i denote the ith diagonal element in $\boldsymbol{\Lambda}$. Substituting the above into (2.2) and (2.3), we get

$$\begin{aligned} \boldsymbol{W} &= \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^H, \\ \mathrm{C}(\boldsymbol{R}) &= \log\det\{\boldsymbol{I}_M + \boldsymbol{G}[(\boldsymbol{FR})^H\boldsymbol{W}^{-1}(\boldsymbol{FR})]\boldsymbol{G}^H\} \\ &= \log\det\{\boldsymbol{I}_M + \boldsymbol{G}[\boldsymbol{V}\overline{\boldsymbol{\Sigma}}\boldsymbol{V}^H]\boldsymbol{G}^H\}, \end{aligned} \tag{2.7}$$

where $\boldsymbol{\Sigma}$ and $\overline{\boldsymbol{\Sigma}}$ are diagonal matrices with their ith diagonal terms given by

$$\boldsymbol{\Sigma}(i,i) = \sigma_D^2 + (\sigma_R|\lambda_i|)^2, \tag{2.8}$$

$$\overline{\boldsymbol{\Sigma}}(i,i) = \frac{|\lambda_i|^2}{\sigma_D^2 + (\sigma_R|\lambda_i|)^2} = \frac{|\lambda_i|^2}{\boldsymbol{\Sigma}(i,i)}. \tag{2.9}$$

By scaling $\boldsymbol{R}$ to $s\boldsymbol{R}$ and using $\boldsymbol{W}_s$ to denote the resulting noise correlation matrix at the destination (in place of $\boldsymbol{W}$), we find

$$\begin{aligned} \boldsymbol{W}_s &= \sigma_D^2\boldsymbol{I} + (s\sigma_R)^2(\boldsymbol{FR})(\boldsymbol{FR})^H = \boldsymbol{U}\boldsymbol{\Sigma}_s\boldsymbol{U}^H, \\ \mathrm{C}(s\boldsymbol{R}) &= \log\det\{\boldsymbol{I}_M + \boldsymbol{G}[s^2(\boldsymbol{FR})^H\boldsymbol{W}_s^{-1}(\boldsymbol{FR})]\boldsymbol{G}^H\} \\ &= \log\det\{\boldsymbol{I}_M + \boldsymbol{G}[\boldsymbol{V}\overline{\boldsymbol{\Sigma}}_s\boldsymbol{V}^H]\boldsymbol{G}^H\}, \end{aligned} \tag{2.10}$$

where $\boldsymbol{\Sigma}_s$ and $\overline{\boldsymbol{\Sigma}}_s$ are diagonal matrices with their ith diagonal terms given by

$$\begin{aligned} \boldsymbol{\Sigma}_s(i,i) &= \sigma_D^2 + (s\sigma_R|\lambda_i|)^2, \\ \overline{\boldsymbol{\Sigma}}_s(i,i) &= \frac{s|\lambda_i|^2}{\sigma_D^2 + (s\sigma_R|\lambda_i|)^2} \triangleq a_i\overline{\boldsymbol{\Sigma}}(i,i), \end{aligned} \tag{2.11}$$

with a_i defined as

$$a_i = \frac{s^2[\sigma_D^2 + (\sigma_R|\lambda_i|)]^2}{\sigma_D^2 + s^2(\sigma_R^2|\lambda_i|)^2}.$$

Note that $a_i > 1$ and $a_i \le a_j$ for $i \le j$.

From (2.10) and (2.11), $\overline{\boldsymbol{\Sigma}}_s$ can be expressed as the sum of two diagonal matrices as

$$\overline{\boldsymbol{\Sigma}}_s = a_1\overline{\boldsymbol{\Sigma}} + \overline{\boldsymbol{\Sigma}}_\Delta,$$

where $\overline{\boldsymbol{\Sigma}}_\Delta$ is some nonnegative diagonal matrix. Then, based on the eigenvalue inequalities concerning the sum of two nonnegative-definite matrices [37, Sec. 6.4], we have

$$\begin{aligned} \mathrm{C}(s\boldsymbol{R}) &= \log\det\{\boldsymbol{I}_M + \boldsymbol{G}[\boldsymbol{V}\overline{\boldsymbol{\Sigma}_s}\boldsymbol{V}^H]\boldsymbol{G}^H\} \\ &\geq \log\det\{\boldsymbol{I}_M + a_1\boldsymbol{G}[\boldsymbol{V}\overline{\boldsymbol{\Sigma}}\boldsymbol{V}^H]\boldsymbol{G}^H\} \\ &> \log\det\{\boldsymbol{I}_M + \boldsymbol{G}[\boldsymbol{V}\overline{\boldsymbol{\Sigma}}\boldsymbol{V}^H]\boldsymbol{G}^H\} = \mathrm{C}(\boldsymbol{R}). \end{aligned}$$

■

Therefore, we confirm that scaling up of the relay gains can increase system capacity. Hence we may simplify the optimization constraint to

$$P_R = \sum_{i=1}^{L}(\sigma_R^2 + |\boldsymbol{G}^{(i)}|^2)|\boldsymbol{r}(i)|^2 \triangleq \sum_{i=1}^{L} p(i)|\boldsymbol{r}(i)|^2. \tag{2.12}$$

That is, the relays should transmit at the maximum allowed total power.

Next, one may wonder if the capacity could increase without bound if the total relay transmission power tends to infinity. Intuitively, the answer may appear to be another no-brainer because, from (2.1), the quality of the source-to-relay links should place a cap on the amount of information rate that the system can support no matter how much the relay transmission power can be. But again, a solid mathematical proof requires a few lines of reasoning. Again we state the result as a theorem and present the proof as follows.

Theorem 2.2.2 (Asymptotic capacity with high relay power)**.** *As* $|s| \to \infty$, $\mathrm{C}(s\boldsymbol{R})$ *is upper-bounded by*

$$\log\det[\boldsymbol{I}_M + \sigma_R^{-2}\boldsymbol{G}\boldsymbol{G}^H]$$

and it approaches the upper bound if and only if $\boldsymbol{G}$ *and* $\boldsymbol{F}\boldsymbol{R}$ *span the same row space.*

Proof. To complete the proof, we will need the following inequality of eigenvalues about matrix product [37, Sec. 6.6]. Assume $\boldsymbol{A}$ and $\boldsymbol{B}$ are $N \times N$ Hermitian non-negative definite matrices, and let $\rho_i(\boldsymbol{A})$ be the ith largest eigenvalue of $\boldsymbol{A}$. Then we have

$$\rho_i(\boldsymbol{A})\rho_N(\boldsymbol{B}) \leq \rho_i(\boldsymbol{A}\boldsymbol{B}) \leq \rho_i(\boldsymbol{A})\rho_1(\boldsymbol{B})$$

From (2.11), as $|s| \to \infty$ the significance of σ_D vanishes, so that $\overline{\boldsymbol{\Sigma}}_s(i,i) \lesssim \sigma_R^{-2}$ and

$$\mathrm{C}(s\boldsymbol{R}) \approx \log\det[\boldsymbol{I}_M + \sigma_R^{-2}\boldsymbol{G}\boldsymbol{V}\boldsymbol{V}^H\boldsymbol{G}^H] \tag{2.13}$$

where $\boldsymbol{V}$ is the matrix of right singular vectors of $\boldsymbol{FR}$ as given in (2.6). To obtain the result, the key is to grasp the eigenvalue structure of $\boldsymbol{GVV}^H\boldsymbol{G}^H$, or equivalently that of $\boldsymbol{G}^H\boldsymbol{GVV}^H$. For this, let $\rho_i(\boldsymbol{M})$ denote the ith largest eigenvalue of a matrix $\boldsymbol{M}$ that has real eigenvalues. We have

$$\begin{aligned}
&\rho_1(\boldsymbol{G}^H\boldsymbol{G}) \geq \rho_2(\boldsymbol{G}^H\boldsymbol{G}) \geq \cdots \geq \rho_M(\boldsymbol{G}^H\boldsymbol{G}) > 0,\\
&\rho_i(\boldsymbol{VV}^H) = 1,\ 1 \leq i \leq M,\\
&\rho_i(\boldsymbol{G}^H\boldsymbol{G}) = \rho_i(\boldsymbol{VV}^H) = 0,\ M+1 \leq i \leq L.
\end{aligned}$$

Therefore, based on the eigenvalue properties concerning matrix products [37, Sec. 6.6], we have

$$\begin{aligned}
\rho_i(\boldsymbol{G}^H\boldsymbol{GVV}^H) &\leq \rho_i(\boldsymbol{G}^H\boldsymbol{G})\rho_1(\boldsymbol{VV}^H)\\
&= \rho_i(\boldsymbol{G}^H\boldsymbol{G}).
\end{aligned} \tag{2.14}$$

The equality in the first line of the above equation holds if and only if $\boldsymbol{G}$ and $\boldsymbol{V}$ span the same row space, or equivalently, if and only if $\boldsymbol{G}$ and $\boldsymbol{FR}$ span the same row space. In conclusion, as $|s| \to \infty$,

$$\mathrm{C}(s\boldsymbol{R}) \leq \log\det[\boldsymbol{I}_M + \sigma_R^{-2}\boldsymbol{GG}^H], \tag{2.15}$$

where the equality holds if and only if $\boldsymbol{FR}$ and $\boldsymbol{G}$ span the same row space.

■

With Theorem 2.2.2, it is verified that $\mathrm{C}(\boldsymbol{R})$ is upper-bounded irrespective of the power level of the relays.

2.3 Iterative Alternating Optimization

In Sec. 1.1.1 we briefly review the state of the art of relay network designs and the challenges of distributed relay networks. In this section we take a closer look at the optimization problem in (2.4) and present the efficient suboptimal successive alternating algorithms based on low-rank updating and quadratic approximation.

Observing the criteria in (2.3), it is clear the computation for matrix determinant is critical for optimization. Given multiple-antenna relays, the forwarding gain matrix $\boldsymbol{R}$ is not diagonal, which mean we could rely on matrix decomposition (such as SVD or QR) to transform cascaded MIMO channels into parallel subchannels then simplify determinant computation [6]. Now to deal with diagonal $\boldsymbol{R}$ and non-convex problem, intuitively we have two options to design the optimization algorithms:

1. Brute-force or systematic global search: since (2.4) could not apply convex programming, we could always utilize systematic global optimization (*e.g.* genetic algorithm

[10]) or even brute-force search to find the optimal $\boldsymbol{R}$. However, this approach would cost huge computation when L is large.

2. Replace determinant computation with high-order multivariate polynomial function: the matrix determinant of a $N \times N$ matrix $\boldsymbol{H}$ is defined to be the scalar [28]

$$\det(\boldsymbol{H}) = \sum_{p} s(p)\boldsymbol{H}(1, p_1)\boldsymbol{H}(2, p_2) \cdots \boldsymbol{H}(N, p_N), \tag{2.16}$$

where the sum is taken over the $N!$ permutations $p = (p_1, p_2 \cdots p_N)$ of $(1, 2 \cdots N)$, $\boldsymbol{H}(i, j)$ denotes the item at ith row and jth column of $\boldsymbol{H}$, and the sign scalar $s(p) \in \{+1, -1\}$ is defined by permutation p. Following (2.3) and (2.16), we may transform the optimization cost function in (2.4) into a high-order multivariate polynomial of $\boldsymbol{r}(i)$ then perform derivative-based algorithms (*e.g.* Newton's method [10]) to approach local optimizer. Such algorithms would simplify the original multivariate problem as successive one-dimensional line searches and guide the searches to most favorable directions. However, even heading to favorable directions, line searches with repeated calculation of high-order multivariate polynomial function would still cost considerable computations and thus may not suit practical purpose.

Given the above two observations, we should design the optimization which not only simplifies the multivariate problem but does effectively suppress extensive computation. Alternating optimization [4] is one approach that meets the requirements. Instead of simultaneously considering all variables, by definition alternating optimization focus on one variable at one time and alternates between variables. By optimizing one variable (and fix the others) in one iteration and considering another variable in next round, alternating optimization again transform multivariate problem into successive one-dimensional optimization. Since we deal with cost function with single variable throughout line searches, we may apply low-rank updating methods [28, Sec. 6.2] to avoid extensive matrix computation. Further, we could approximate polynomial function based on low-rank updating as second-order function. Thus the approximated optimizer could be derived without line search.

2.3.1 Low-rank Updating for Capacity Computation

Some matrix computation, such as decomposition, inversion or determinant, require considerable computing resource. When the variation of a matrix comes from one or two vectors, it is possible to compute some quantities associated with the matrix efficiently via an updating procedure rather than via a full-blown computational procedure. Since the variations are frequently of rank one or two, the efficient updating methods are often termed *low-rank* updating. In what follows we describe some updating methods used in our work.

Assume square matrices $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{C}$ with updating equalities as follows

$$\boldsymbol{B} \triangleq \boldsymbol{A} + \boldsymbol{u}_1 \boldsymbol{v}_1^H, \boldsymbol{C} \triangleq \boldsymbol{B} + \boldsymbol{u}_2 \boldsymbol{v}_2^H$$

where $\boldsymbol{A}$ is assumed to be full-rank Hermitian matrix. Given $\det(\boldsymbol{A})$ and $\boldsymbol{A}^{-1}$, we have rank-one updating for efficient computation about $\boldsymbol{B}$ as

$$\boldsymbol{B}^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}_1 \boldsymbol{v}_1^H \boldsymbol{A}^{-1}}{1 + \boldsymbol{v}_1^H \boldsymbol{A}^{-1} \boldsymbol{u}_1}, \tag{2.17}$$

$$\det(\boldsymbol{B}) = \det(\boldsymbol{A})(1 + \boldsymbol{v}_1^H \boldsymbol{A}^{-1} \boldsymbol{u}_1). \tag{2.18}$$

Using (2.18) and (2.17) we could save a great deal of calculation and extend to rank-two updating for $\det(\boldsymbol{C})$ as

$$\begin{aligned}
\det(\boldsymbol{C}) &= \det(\boldsymbol{B})(1 + \boldsymbol{v}_2^H \boldsymbol{B}^{-1} \boldsymbol{u}_2) \\
&= \det(\boldsymbol{A})(1 + \boldsymbol{v}_1^H \boldsymbol{A}^{-1} \boldsymbol{u}_1) \left[1 + \boldsymbol{v}_2^H \left(\boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}_1 \boldsymbol{v}_1^H \boldsymbol{A}^{-1}}{1 + \boldsymbol{v}_1^H \boldsymbol{A}^{-1} \boldsymbol{u}_1}\right) \boldsymbol{u}_2\right] \\
&= \det(\boldsymbol{A})\{(1 + \boldsymbol{v}_1^H \boldsymbol{A}^{-1} \boldsymbol{u}_1)(1 + \boldsymbol{v}_2^H \boldsymbol{A}^{-1} \boldsymbol{u}_2) - \boldsymbol{v}_2^H \boldsymbol{A}^{-1} \boldsymbol{u}_1 \boldsymbol{v}_1^H \boldsymbol{A}^{-1} \boldsymbol{u}_2\}.
\end{aligned} \tag{2.19}$$

Now we describe how the above-mentioned low-rank updating is applied in our alternating optimization. In the proposed iterative relay gain adjustment, we adjust only one relay gain at a time. Specifically, in each iteration, we replace one relay gain by some value $\alpha \in \mathbb{C}$. The other relay gains are multiplied by a factor $\beta \in \mathbb{R}_+$ (where $\mathbb{R}_+$ stands for the set of positive real numbers) such that the power constraint (2.12) is satisfied. The factors α and β are chosen to maximize $C(\boldsymbol{R})$.

In more detail, let $\boldsymbol{r} = [r(1), \ldots, r(L)]^T$ denote the gain vector of the relay network where superscript T denotes transpose. Let the ith term of $\boldsymbol{r}$, or $\boldsymbol{r}(i)$, be the relay to be optimized in certain iteration, and $\boldsymbol{r}_o$ be the same as $\boldsymbol{r}$ except with $\boldsymbol{r}(i)$ replaced by zero. Also let $\boldsymbol{r}_u$ denote the gain vector after the above-described update. Then

$$\boldsymbol{r}_o = (\boldsymbol{I}_L - \boldsymbol{S}_i)\boldsymbol{r}, \quad \boldsymbol{r}_u = \beta \boldsymbol{r}_o + \alpha \boldsymbol{S}_i \boldsymbol{1}, \tag{2.20}$$

where $\boldsymbol{S}_i$ denotes the "selection matrix" whose elements are all zero except for a 1 at the ith diagonal position and $\boldsymbol{1}$ represents an all-ones vector. Clearly, following (2.12) α is subject to the constraint

$$0 \le |\alpha| < \sqrt{P_R / p(i)}, \tag{2.21}$$

and for given α, we have

$$\beta = \sqrt{\frac{P_R - |\alpha|^2 p(i)}{\sum_i p(i) |\boldsymbol{r}_o(i)|^2}}. \tag{2.22}$$

Let $\boldsymbol{R}_o = \mathrm{diag}(\boldsymbol{r}_o)$ and $\boldsymbol{R}_u = \mathrm{diag}(\boldsymbol{r}_u)$. Then the noise-free equivalent end-to-end channel matrix after gain updating is given by

$$\boldsymbol{H}_u = \beta \boldsymbol{H}_o + \alpha \boldsymbol{F}^{(i)} (\boldsymbol{G}^{(i)})^H \tag{2.23}$$

where $\boldsymbol{H}_o = \boldsymbol{F}\boldsymbol{R}_o\boldsymbol{G}^H$ and $\boldsymbol{H}_u = \boldsymbol{F}\boldsymbol{R}_u\boldsymbol{G}^H$. And the autocorrelation matrix of the received noise vector at the destination becomes

$$\begin{aligned}\boldsymbol{W}_u &= \sigma_D^2[\boldsymbol{I}_M + (\sigma\beta)^2(\boldsymbol{F}\boldsymbol{R}_o)(\boldsymbol{F}\boldsymbol{R}_o)^H] + |\sigma_R\alpha|^2\boldsymbol{F}^{(i)}(\boldsymbol{F}^{(i)})^H \\ &\triangleq \sigma_D^2\boldsymbol{W}_{\boldsymbol{o}} + |\sigma_R\alpha|^2\boldsymbol{F}^{(i)}(\boldsymbol{F}^{(i)})^H. \end{aligned} \tag{2.24}$$

We see that $\boldsymbol{H}_u$ is different from $\beta\boldsymbol{H}_o$ by a rank-one matrix and that $\boldsymbol{W}_u$ is different from $\sigma_D^2\boldsymbol{W}_o$ also by a rank-one matrix. Further, we could express $\boldsymbol{W}_u + \boldsymbol{H}_u\boldsymbol{H}_u^H$ with rank-two updating as

$$\begin{aligned}\boldsymbol{W}_u + \boldsymbol{H}_u\boldsymbol{H}_u^H = (\sigma_D^2\boldsymbol{W}_o + \beta^2\boldsymbol{H}_o\boldsymbol{H}_o^H) + (\alpha^H\beta)(\boldsymbol{H}_o\boldsymbol{G}^{(i)})(\boldsymbol{F}^{(i)})^H \\ +\boldsymbol{F}^{(i)}\left[(\alpha\beta)(\boldsymbol{H}_o\boldsymbol{G}^{(i)})^H + (|\alpha\boldsymbol{G}^{(i)}|^2 + |\sigma_R\alpha|^2)(\boldsymbol{F}^{(i)})^H\right]. \end{aligned} \tag{2.25}$$

Given these matrix updating forms, we proceed the design of alternating optimization by decomposing the problem into three parts: 1) how to express $C(\boldsymbol{R}_u)$ in terms of α and β; 2) how to optimize the values of α and β; and 3) how to iterate. We address these subproblems in order below.

First, consider the term $\det(\boldsymbol{W}_u)$. From (2.24) it can readily be seen to be a polynomial in $|\alpha|^2$ and β^2. Applying the rank-one determinant update formula to it results in

$$\det(\boldsymbol{W}_u) = \sigma_D^{2M}\det(\boldsymbol{W}_o)(1 + |\sigma\alpha|^2(\boldsymbol{F}^{(i)})^H\boldsymbol{W}_o^{-1}\boldsymbol{F}^{(i)}). \tag{2.26}$$

Its dependence on $|\alpha|^2$ and β^2 can be expressed more concretely in terms of an eigenvalue decomposition of $(\boldsymbol{F}\boldsymbol{R}_o)(\boldsymbol{F}\boldsymbol{R}_o)^H$:

$$(\boldsymbol{F}\boldsymbol{R}_o)(\boldsymbol{F}\boldsymbol{R}_o)^H = \boldsymbol{V}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^H. \tag{2.27}$$

Then, letting $e_1(i)$ denote the ith eigenvalue (i.e., the ith diagonal element of $\boldsymbol{\Sigma}_1$), we have

$$\det(\boldsymbol{W}_u) = \sigma_D^{2M}\prod_{i=1}^{M}\{1 + (\sigma\beta)^2e_1(i)\}\{1 + |\sigma\alpha|^2(\boldsymbol{F}^{(i)})^H\boldsymbol{V}_1[\boldsymbol{I}_M + (\sigma\beta)^2\boldsymbol{\Sigma}_1]^{-1}\boldsymbol{V}^H\boldsymbol{F}^{(i)}\}, \tag{2.28}$$

where a leading product in (2.18) is canceled by the common denominator of the braced quantity.

Next, consider the term $\det(\boldsymbol{W}_u + \boldsymbol{H}_u\boldsymbol{H}_u^H)$, for which we make use of the rank-2 update formula for matrix determinants. Employing (2.19) with the following identifications of variables:

$$\begin{aligned}&\boldsymbol{C} \leftrightarrow \boldsymbol{W}_u + \boldsymbol{H}_u\boldsymbol{H}_u^H, \\ &\boldsymbol{A} \leftrightarrow \sigma_D^2[\boldsymbol{I}_M + (\sigma\beta)^2\boldsymbol{F}\boldsymbol{R}_o(\boldsymbol{F}\boldsymbol{R}_o)^H] + \beta^2\boldsymbol{H}_o\boldsymbol{H}_o^H, \\ &\boldsymbol{u}_1 \leftrightarrow \boldsymbol{F}^{(i)}, \boldsymbol{v}_1 \leftrightarrow (\alpha^H\beta)(\boldsymbol{H}_o\boldsymbol{G}^{(i)}) + (|\alpha\boldsymbol{G}^{(i)}|^2 + |\sigma_R\alpha|^2)\boldsymbol{F}^{(i)}, \\ &\boldsymbol{u}_2 \leftrightarrow \boldsymbol{H}_o\boldsymbol{G}^{(i)}, \boldsymbol{v}_2 \leftrightarrow (\alpha\beta)\boldsymbol{F}^{(i)}, \end{aligned}$$

we get

$$\begin{aligned}&\det(\boldsymbol{W}_u + \boldsymbol{H}_u\boldsymbol{H}_u^H)\\&= \det(\boldsymbol{A})\{(1+\boldsymbol{v}_1^H\boldsymbol{A}^{-1}\boldsymbol{u}_1)(1+\boldsymbol{v}_2^H\boldsymbol{A}^{-1}\boldsymbol{u}_2) - \boldsymbol{v}_2^H\boldsymbol{A}^{-1}\boldsymbol{u}_1\boldsymbol{v}_1^H\boldsymbol{A}^{-1}\boldsymbol{u}_2\}.\end{aligned} \tag{2.29}$$

As in the case of $\det(\boldsymbol{W}_u)$, its polynomial functional dependence on α and β can be brought out more concretely with an eigenvalue decomposition of a constituent factor of $\boldsymbol{A}$:

$$\boldsymbol{F}\boldsymbol{R}_o(\boldsymbol{F}\boldsymbol{R}_o)^H + \sigma_R^{-2}\boldsymbol{H}_o\boldsymbol{H}_o^H = \boldsymbol{V}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^H. \tag{2.30}$$

Then, letting $e_2(i)$ denote the ith eigenvalue, we have

$$\begin{aligned}&\det(\boldsymbol{A}) = \sigma_D^{2M}\prod_{i=1}^{M}[1+(\sigma\beta)^2 e_2(i)]\\&\det(\boldsymbol{W}_u + \boldsymbol{H}_u\boldsymbol{H}_u^H)\\&= \det(\boldsymbol{A})[(1+l_1\boldsymbol{p}_1^H\boldsymbol{p}_1 + l_2\boldsymbol{p}_2^H\boldsymbol{p}_1)(1+l_2^H\boldsymbol{p}_1^H\boldsymbol{p}_2) - (l_2^H\boldsymbol{p}_1^H\boldsymbol{p}_1)(l_1\boldsymbol{p}_1^H\boldsymbol{p}_2 + l_2\boldsymbol{p}_2^H\boldsymbol{p}_2)]\\&= \sigma_D^{2M}\prod_{i=1}^{M}[1+(\sigma\beta)^2 e_2(i)] \times [1 + l_1|\boldsymbol{p}_1|^2 + 2\Re(l_2\boldsymbol{p}_2^H\boldsymbol{p}_1) + |l_2|^2(|\boldsymbol{p}_1^H\boldsymbol{p}_2|^2 - |\boldsymbol{p}_1|^2|\boldsymbol{p}_2|^2)],\end{aligned} \tag{2.31}$$

where $\Re(\cdot)$ denotes the real part of a quantity and we have made the following definitions to simplify the notation:

$$l_1 \triangleq (|\alpha\boldsymbol{G}^{(i)}|^2 + |\sigma_R\alpha|^2)\sigma_D^{-2}, \quad l_2 \triangleq \alpha\beta^2\sigma_D^{-2}, \tag{2.32}$$

$$\boldsymbol{p}_1 \triangleq [\boldsymbol{I}_M + (\sigma\beta)^2\boldsymbol{\Sigma}_2]^{-\frac{1}{2}}\boldsymbol{V}_2^H\boldsymbol{F}^{(i)}, \quad \boldsymbol{p}_2 \triangleq [\boldsymbol{I}_M + (\sigma\beta)^2\boldsymbol{\Sigma}_2]^{-\frac{1}{2}}\boldsymbol{V}_2^H(\boldsymbol{H}_o\boldsymbol{G}^{(i)}). \tag{2.33}$$

In summary, $C(\boldsymbol{R}_u)$ can be expressed as the difference between (2.31) and (2.28). We now turn to the problem of finding α and β that maximize it.

2.3.2 Optimization of α and β

To start, note that none of the terms constituting $\det(\boldsymbol{W}_u)$ and $\det(\boldsymbol{W}_u +^2 \boldsymbol{H}_u\boldsymbol{H}_u^H)$ depend on the phase of α except $\Re(l_2\boldsymbol{p}_2^H\boldsymbol{p}_1)$ that appears in (2.31). As a result, for any given $|\alpha|$ and β, $C(\boldsymbol{R}_u)$ can be maximized by choosing the phase of α such that $\Re(l_2\boldsymbol{p}_2^H\boldsymbol{p}_1)$ is maximized. This can be achieved by letting $\angle\alpha = \angle\boldsymbol{p}_1^H\boldsymbol{p}_2$, so that $\Re(l_2\boldsymbol{p}_2^H\boldsymbol{p}_1) = |l_2\boldsymbol{p}_2^H\boldsymbol{p}_1|$. The problem thus reduces to one of finding the best $|\alpha|$ and β. But since there is a one-to-one relation between $|\alpha|$ and β (see (2.22)), we only need to solve for β. After some straightforward algebra based on (2.3), (2.28), (2.31), we can show that the optimal β is one that maximizes the following function:

$$q(\beta) \triangleq \frac{1 + l_1|\boldsymbol{p}_1|^2 + 2|l_2\boldsymbol{p}_2^H\boldsymbol{p}_1| + |l_2|^2(|\boldsymbol{p}_1^H\boldsymbol{p}_2|^2 - |\boldsymbol{p}_1|^2|\boldsymbol{p}_2|^2)}{1 + |\sigma\alpha|^2(\boldsymbol{F}^{(i)})^H\boldsymbol{V}_1[\boldsymbol{I}_M + (\sigma\beta)^2\boldsymbol{\Sigma}_1]^{-1}\boldsymbol{V}^H\boldsymbol{F}^{(i)}}\prod_{i=1}^{M}\frac{1+(\sigma\beta)^2 e_2(i)}{1+(\sigma\beta)^2 e_1(i)}, \tag{2.34}$$

where $|\alpha|$, l_1, l_2, $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are all functions of β.

Due to the complicated nature of (2.34), there is in general no closed-form solution for the optimal β. We need to resort to a search technique, and such techniques are innumerable. A simplest one is non-iterative line search, in which one examines a sufficiently dense subset of all admissible values of β to find the one maximizing $q(\beta)$. From (2.21) and (2.22), the set of admissible values of β are given by

$$0 < \beta \leq \sqrt{\frac{P_R}{\sum_{i=1}^{L} p(i)|r_o(i)|^2}}. \tag{2.35}$$

A second method is to iteratively update a trial solution to β by solving a low-order polynomial approximation to $q(\beta)$ in each iteration. For example, one may, in each iteration, use a quadratic approximation obtained by taking the second-order Taylor series expansions of $q(\beta)$ around some β value and take the β value that maximizes the quadratic approximation as the updated trial solution. If this value should fall outside the admissible range given in (2.35), we may replace it by the nearest boundary value of the range. In addition, if the resulting $q(\beta)$ value should decrease in some iteration, then we may stop the iteration and revert to an earlier solution.

In fact, to find the optimal β one need not work with $q(\beta)$ directly. Any monotone increasing function of $q(\beta)$ can be used in its stead. For example, since, from (2.34), $q(\beta)$ is a product of multiple factors, it may be easier to consider maximizing a logarithm of $q(\beta)$ than $q(\beta)$ itself, for then products become sums. This approach is taken in our implementation of the quadratic approximation method. Moreover, in implementing the quadratic approximation method we have chosen to take the Taylor series expansion at $\beta = 1$. The reason is that, since we adjust one relay gain at a time, the overall optimization process belong to the category of alternating optimization which is guaranteed to converge to a local optimum [4]. Upon convergence, the values in $\boldsymbol{r}_u$ will change little from one iteration to the next. In other words, the optimal β values will approach unity upon convergence of the overall algorithm. Hence a series expansion around $\beta = 1$ should provide a good approximation to the performance surface in the later stages of algorithm progression and benefit its convergence behavior there. In summary, in our implementation of the quadratic approximation method we seek to maximize $q_l(\beta) \triangleq \log q(\beta)$. And for it we define

$$q_a(\beta) \triangleq q_l(1) + q_l'(1)(\beta - 1) + \frac{q_l''(1)}{2}(\beta - 1)^2, \tag{2.36}$$

where $q_l'(1)$ and $q_l''(1)$ are the first and second derivatives of $q_l(\beta)$ evaluated at $\beta = 1$. The solution to the equation $q_a'(\beta) = 0$ is then taken to be the current trial solution for β.

We summaries the proposed successive optimization algorithm as follows. It finds a locally optimal solution.

1. Select some relay i for gain adjustment, where i can be chosen in round-robin fashion, for example.

2. Perform eigenvalue decomposition of $(\boldsymbol{FR}_o)(\boldsymbol{FR}_o)^H$ and $\boldsymbol{FR}_o(\boldsymbol{FR}_o)^H+\sigma_R^{-2}\boldsymbol{H}_o\boldsymbol{H}_o^H$ as in (2.27) and (2.30) to find $\boldsymbol{\Sigma}_i$ and $\boldsymbol{V}_i$, $i=1,2$.

3. Solve for the β that maximizes $q(\beta)$ as given in (2.34) by a search method, such as the line search or the quadratic approximation method described in the last subsection. Obtain the corresponding $|\alpha|$ using (2.22) and let $\angle\alpha=\angle\boldsymbol{p}_1^H\boldsymbol{p}_2$.

4. Update the relay gains by setting the gain of the ith relay to α and multiplying the gains of all other relays by β.

5. Exit if some stopping criteria are satisfied, or go to step 1 otherwise.

2.3.3 Numerical Results

In presenting the simulation results, we arbitrarily let $M=4$, $P_R=10$, and $\sigma_R^2=\sigma_D^2=0.1$. And we consider two relay network sizes: $L=6$ and $L=12$. The channel matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are generated by letting all their elements be independent and identically distributed (i.i.d.) complex Gaussian random variables. The relays are initialized to an identical gain that satisfies the power constraint (2.12). Thus their initial performance also serves as a benchmark to compare algorithm results with.

Fig. 2.2 illustrates the progression of average capacity with number of iterations under two methods of solving for the relay gains adjustment factor β: line search and quadratic approximation, where the former has a much higher computational complexity than the latter. The results show that line search performs better than quadratic approximation, but both show a qualitatively similar convergence behavior and the final results after convergence are quite close. Also note that the converged capacity, for both approaches, are much better than the initial capacity, especially when L is large. Such significant improvement of capacity suggests that the proposed algorithms could contribute substantial performance upgrading when compared to simple identical gain design.

In Fig. 2.3 we show the average behavior of capacity variation when performing line search of optimal β value in step 3 of procedures described in Sec. 2.3.2. Note the capacity curve is first normalized with maximal $q(\beta)$ in step 3, then is averaged over random channel conditions as set for Fig. 2.2. With this simulation, we could confirm that the optimal solution to line search is around $\beta=1$ in average. Thus applying second-order Taylor Series expansion at $\beta=1$ not only lead to efficient suboptimal $q(\beta)$ solution, but also provides a good approximation to actual $q(\beta)$ behavior around optimal β.

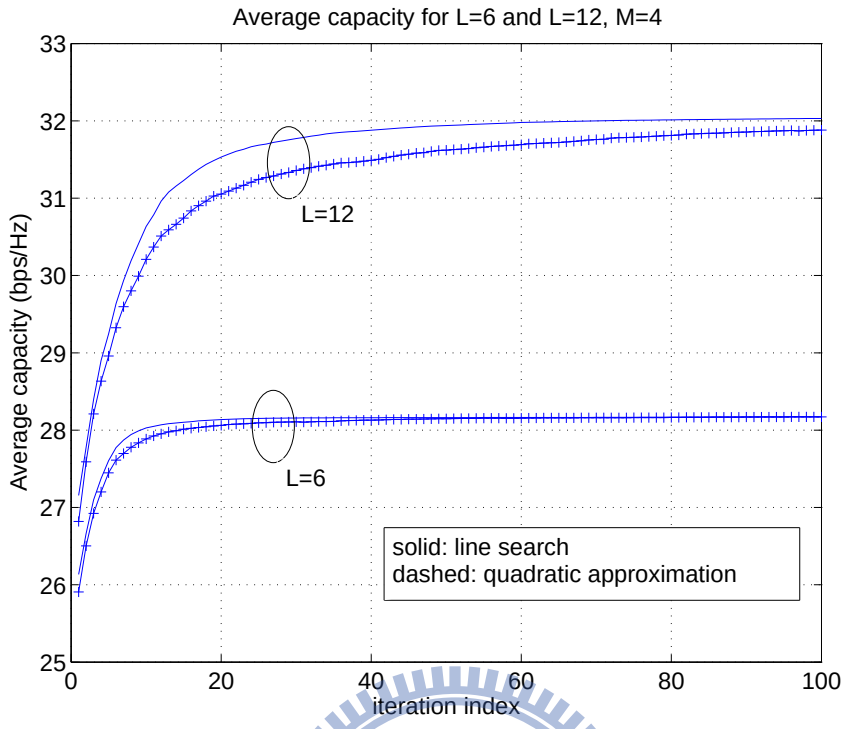

Fig. 2.2: Progression of average capacity with number of iterations.

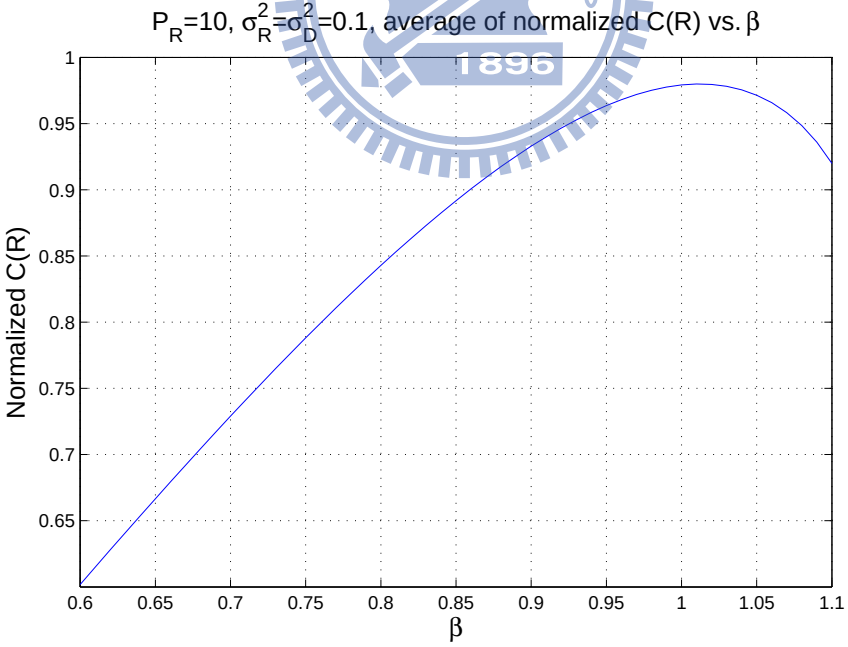

Fig. 2.3: Average of normalized capacity with respect to varying β values

Chapter 3

Capacity Improvement based on Noise-Dominant Models

While alternating optimization based on low-rank updating discussed in Chap. 2 can yield good results, it provides little insight into the analytical properties of the solutions. We thus consider an analytical approach in this chapter. Since no closed-form solution can be obtained for the general situation, we consider several simplified situations which are more amenable to analysis. In particular, note that in AF systems the receiver noise arises from two sources: the relay noise $\boldsymbol{n}_R$ and the destination terminal noise $\boldsymbol{n}_D$. The design problem becomes mathematically more tractable when one of the two dominates in the overall receiver noise so that the other may be ignored. The results obtained from ignoring one noise source may be viewed as upper bounds on system capacity or as asymptotic performance of the system. For convenience, we term the two simplified conditions the relay noise-dominant condition and the destination noise-dominant condition, respectively.

Interestingly, closed-form analytical solutions are not available for arbitrary L even in these simplified conditions. But such solutions can be found if L is restricted to some specific values depending on M. It thus prompts a (suboptimal) relay selection approach wherein a judiciously selected subset of the relays is used to participate in signal transmission and the subset size is such that an analytical solution exists. This approach also helps us to study the resulting capacity outage diversity and compare it to that of single-hop MIMO systems with or without antenna selection [17, 14].

Since MIMO antenna selection systems lay the foundation for the analytical interpretation of the proposed relay selection schemes, In this chapter we would first briefly review the important results of the studies on single-hop MIMO antenna selection systems. Then we present noise-dominant models for relay network design and propose corresponding algorithms for individual models. For analytical insights we design equivalent MIMO antenna selection systems. Finally some numerical results are shown for performance evaluation and validating the links between equivalent MIMO antenna selection and proposed relay selection schemes.

3.1 Preliminary: MIMO Antenna-Selection Systems

Diversity techniques based on antenna selection or maximal ratio combining are mature and popular approach to utilize multiple antennas for performance improvement. When only single stream is transmitted with N_t Tx antennas and N_r Rx antennas, the diversity order of $N_t N_r$ could be achieved [2]. However, this principle could not be applied directly to spatial multiplexing transmission with multiple concurrent streams. To this end, MIMO antenna selection systems are developed [16] [17] and shown to be an efficient way to provide additional diversity for spatial multiplexing systems.

Consider a point-to-point MIMO system with M transmitter antennas and N receiver antennas, where $N \geq M$. Let $\boldsymbol{H} \in \mathbb{C}^{N \times M}$ be the channel matrix and let the transmitted signal-to-received noise power ratio (transmit-to-receive SNR) ρ^2. Then the system capacity is given by

$$\log_2 C(\boldsymbol{H}) = \log_2 \det(\boldsymbol{I}_M + \rho^2 \boldsymbol{H}^H \boldsymbol{H}). \tag{3.1}$$

For a flat-fading $\boldsymbol{H}$, a statistical lower bound is [14]

$$\log_2 C(\boldsymbol{H}) \geq \sum_{i=1}^{M} \log_2(1 + \rho^2 \gamma^2_{N-i+1}) \tag{3.2}$$

where γ^2_{N-i+1} denotes a gamma-distributed random variable with $N - i + 1$ degrees of freedom. This lower bound indicates that the capacity of an $M \times N$ MIMO system is statistically equivalent to or better than that of a system composed of M parallel independent single-input multi-output (SIMO) subsystems wherein the ith subsystem performs maximal-ratio combining (MRC) on $N - M + i$ receiver antennas. In other words, the overall capacity outage diversity of an $M \times N$ MIMO system is bounded between $N - M + 1$ and N.

Consider a system where the receiver selects M out of its N antennas for use in signal detection. Let $\boldsymbol{H}_S$ be the $M \times M$ channel matrix of the resulting MIMO channel. This matrix contains the M rows in $\boldsymbol{H}$ that correspond to the selected receiver antennas. There are $\binom{N}{M}$ possible antenna choices. Let $(M, N; M)_S$ denote a system wherein the antennas are chosen to maximize the capacity. Then the system capacity can be described as

$$\log_2 C_S = \max_{\boldsymbol{H}_S} \log_2 \det(\boldsymbol{I} + \rho^2 \boldsymbol{H}_S^H \boldsymbol{H}_S). \tag{3.3}$$

It is shown in [17] that the capacity of such a system is again statistically equivalent to or better than a MIMO system composed of M parallel independent SIMO subsystems wherein the ith subsystem performs antenna selection by choosing one out of $N - i + 1$ receiver antennas.

In a nutshell, both the full system $(M, N; N)_S$ and the receiver antenna selection system $(M, N; M)_S$ can be statistically modeled as a set of parallel SIMO transmissions and

thus share a similar capacity outage diversity order. We show the cumulative distribution function (CDF) curves of capacity for both systems. Given $M = 3$, in Fig. 3.1 the top part presents CDF curves of both systems with $N = 5$ and $N = 12$, respectively. We could observe that increasing N would improve capacity, and full systems outperform antenna selection systems. To gain more insights into the capacity distribution, in the bottom part of Fig. 3.1 we show biased CDF where curves are overlapped and shifted horizontally with the 50% points collocated. With the biased CDF it is clear that antenna selection systems share similar capacity distribution characteristic with corresponding full systems.

3.2 Designs for Destination-Noise Dominant Conditions

Following Sec. 2.1.2, the noise in system could be represented as

$$\begin{aligned} \boldsymbol{W} &= \mathrm{E}\{(\boldsymbol{F}\boldsymbol{R}\boldsymbol{n}_R + \boldsymbol{n}_D)(\boldsymbol{F}\boldsymbol{R}\boldsymbol{n}_R + \boldsymbol{n}_D)^H\} \\ &= \sigma_D^2(\boldsymbol{I} + \sigma^2(\boldsymbol{F}\boldsymbol{R})(\boldsymbol{F}\boldsymbol{R})^H). \end{aligned} \tag{3.4}$$

Thus, when the destination noise dominates in the overall noise we have $\boldsymbol{W} \gtrsim \sigma_D^2\boldsymbol{I}$. Then we have capacity approximation as

$$\begin{aligned} \mathrm{C}(\boldsymbol{R}) &= \log_2 \det(\boldsymbol{I} + \boldsymbol{H}^H\boldsymbol{W}^{-1}\boldsymbol{H}) \\ &\lesssim \det[\boldsymbol{I} + \sigma_D^{-2}\boldsymbol{H}^H\boldsymbol{H}]. \end{aligned} \tag{3.5}$$

Hence the capacity is approximately that of an $M \times M$ single-hop point-to-point MIMO system with channel matrix $\boldsymbol{H}$ and transmitted signal-to-received noise power ratio (transmit-to-receive SNR) $1/\sigma_D^2$. However, even in this rather simplified condition, no general solution is available to the optimization problem (2.4) for arbitrary $L > M$. But an analytical solution can be obtained for $L = M$. Thus we consider a relay selection approach wherein M relays are selected to perform the relaying. Let the total of $\binom{L}{M}$ selections be indexed from 1 to $\binom{L}{M}$. For the kth selection define the corresponding optimization target based on (3.5) as

$$C_D(k, \boldsymbol{R}_D) \triangleq \det[\boldsymbol{I}_M + \sigma_D^{-2}(\boldsymbol{F}_k\boldsymbol{R}_D\boldsymbol{G}_k^H)(\boldsymbol{F}_k\boldsymbol{R}_D\boldsymbol{G}_k^H)^H] \tag{3.6}$$

where $\boldsymbol{F}_k \in \mathbb{C}^{M\times M}$ and $\boldsymbol{G}_k \in \mathbb{C}^{M\times M}$, respectively, denote the submatrices of $\boldsymbol{F}$ and $\boldsymbol{G}$ constructed by collecting the columns corresponding to the active relays in the kth selection, and $\boldsymbol{R}_D$ denotes the diagonal matrix of relay gains of the active relays. Let $r_D(i)$ be the ith diagonal term in $\boldsymbol{R}_D$. In high SNR,

$$\begin{aligned} C_D(k, \boldsymbol{R}_D) &\lesssim \det[\sigma_D^{-2}(\boldsymbol{F}_k\boldsymbol{R}_D\boldsymbol{G}_k^H)(\boldsymbol{F}_k\boldsymbol{R}_D\boldsymbol{G}_k^H)^H] \\ &= \sigma_D^{-2M}\det(\boldsymbol{F}_k\boldsymbol{F}_k^H)\det(\boldsymbol{G}_k\boldsymbol{G}_k^H)\prod_i |r_D(i)|^2. \end{aligned} \tag{3.7}$$

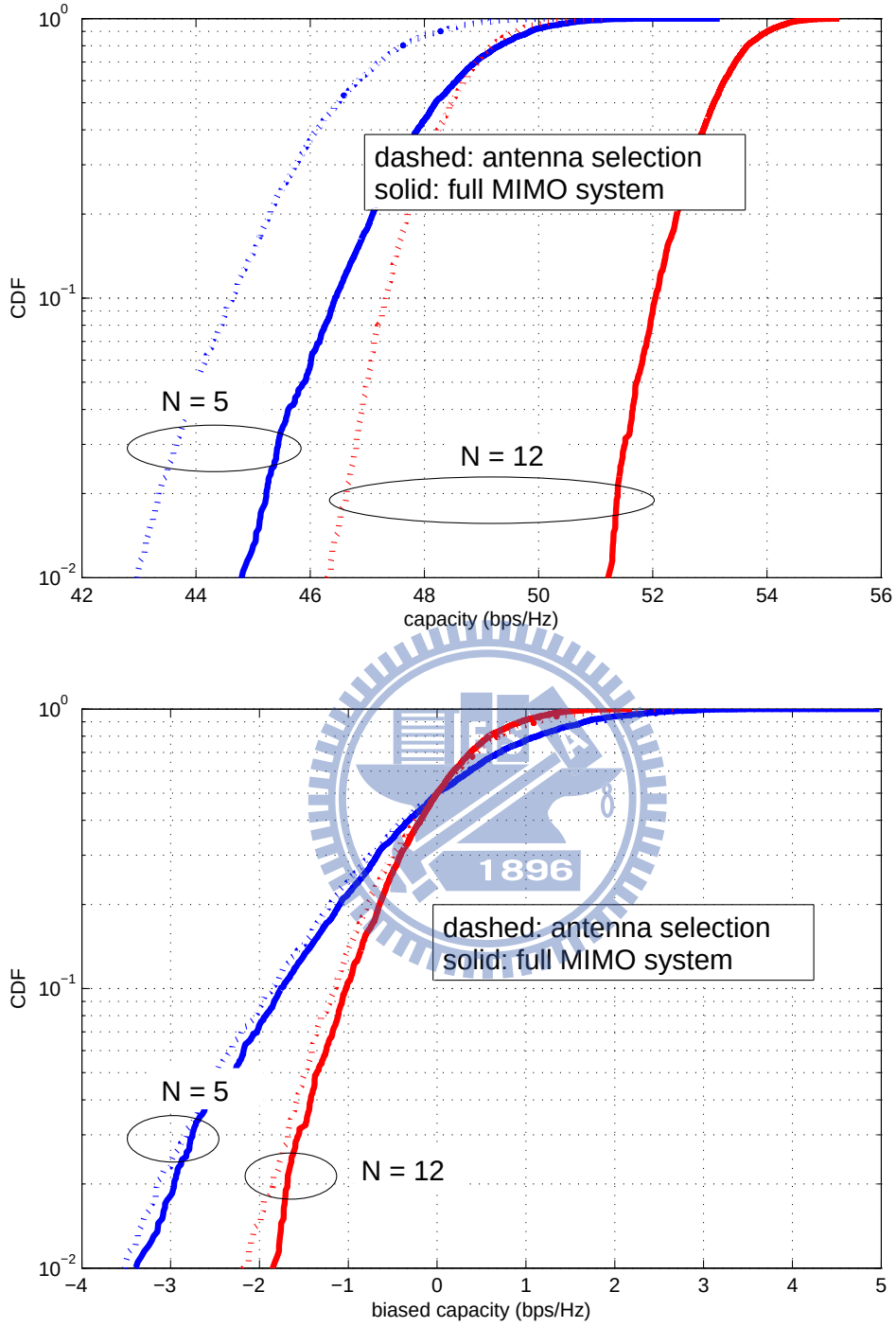

Fig. 3.1: Comparison of MIMO antenna selection and full MIMO systems in terms of CDF of capacity (top) and biased capacity (bottom).

To maximize $C_D(k, \boldsymbol{R}_D)$ subject to the power constraint (2.12) given $\boldsymbol{F}_k$ and $\boldsymbol{G}_k$, we equivalently find the optimum $\boldsymbol{r}_D$ such that

$$\boldsymbol{r}_D = \arg\max_{\boldsymbol{r}_D} \prod_i |r_D(i)| \tag{3.8}$$

subject to

$$\sum_{i=1}^{M} (\sigma_R^2 + \|\boldsymbol{G}_k^{(i)}\|^2)|r_D(i)|^2 = P_R, \tag{3.9}$$

where $\boldsymbol{r}_D = [r_D(1), \ldots, r_D(M)]^T$ and $\boldsymbol{G}_k^{(i)}$ denotes the ith column of $\boldsymbol{G}_k$. Employing the Lagrange multiplier technique leads to the optimum relay power allocation as

$$|r_D(i)| = \sqrt{\frac{P_R}{M(\sigma_R^2 + \|\boldsymbol{G}_k^{(i)}\|^2)}}. \tag{3.10}$$

Denote the resulting $C_D(k, \boldsymbol{R}_D)$ by $C_{DO}(k)$. The final solution is then given by the optimal selection

$$\bar{k} \triangleq \arg\max_k C_{DO}(k) \tag{3.11}$$

together with its corresponding optimum relay power allocation.

To analyze its performance, substitute (3.10) into (3.7) and assume $\|\boldsymbol{G}_k^{(i)}\|^2 \gg \sigma_R^2$ (i.e., consider the high SNR limit). Then we get an upper bound for any $C_{DO}^{(k)}(\boldsymbol{R}_D)$ as

$$C_{DO}(k) \lesssim \sigma_D^{-2M} \det(\boldsymbol{F}_k \boldsymbol{F}_k^H) \frac{\det(\boldsymbol{G}_k \boldsymbol{G}_k^H)}{\prod_i \|\boldsymbol{G}_k^{(i)}\|^2} (\frac{P_R}{M})^M. \tag{3.12}$$

A simpler upper bound can be obtained by considering a QR decomposition of $\boldsymbol{G}_k^H$ as $\boldsymbol{G}_k^H = \boldsymbol{Q}\boldsymbol{T}$, where $\boldsymbol{Q}$ is a unitary matrix and $\boldsymbol{T}$ is an upper triangular matrix. Denote the ith column of $\boldsymbol{T}$ by $\boldsymbol{T}^{(i)}$ and the ith diagonal term of $\boldsymbol{T}$ by $\boldsymbol{T}(i,i)$. Then $\|\boldsymbol{T}^{(i)}\|^2 = \|\boldsymbol{G}_k^{(i)}\|^2$ because $\boldsymbol{T}^{(i)}$ and $\boldsymbol{G}_k^{(i)}$ are related by a unitary transform $\boldsymbol{Q}$, and $|\boldsymbol{T}(i,i)|^2 \le \|\boldsymbol{T}^{(i)}\|^2$. Consequently,

$$\frac{\det(\boldsymbol{G}_k \boldsymbol{G}_k^H)}{\prod_i \|\boldsymbol{G}_k^{(i)}\|^2} = \frac{|\det(\boldsymbol{Q})|^2 |\det(\boldsymbol{T})|^2}{\prod_i \|\boldsymbol{G}_k^{(i)}\|^2} = \prod_{i=1}^{M} \frac{|\boldsymbol{T}(i,i)|^2}{\|\boldsymbol{G}_k^{(i)}\|^2} \le \prod_{i=1}^{M} \frac{\|\boldsymbol{T}^{(i)}\|^2}{\|\boldsymbol{G}_k^{(i)}\|^2} = 1, \tag{3.13}$$

where equality holds only when $\boldsymbol{G}_k$ has orthogonal columns. Substituting into (3.12) yields the desired upper bound

$$C_{DO}(k) < \det(\boldsymbol{F}_k \boldsymbol{F}_k^H)(\frac{P_R}{\sigma_D^2 M})^M. \tag{3.14}$$

Thus we obtain an upper bound C_{DU} on the capacity measure for the suboptimal solution as

$$C_{DO}(\bar{k}) < \det(\boldsymbol{F}_{\bar{k}} \boldsymbol{F}_{\bar{k}}^H)(\frac{P_R}{\sigma_D^2 M})^M \le \max_k \det(\boldsymbol{F}_k \boldsymbol{F}_k^H)(\frac{P_R}{\sigma_D^2 M})^M \triangleq C_{DU}. \tag{3.15}$$

Note that $\log_2 C_{DU}$ is actually the asymptotic capacity of an $(M, L; M)_S$ system at transmit-to-receive SNR $P_R^M/(\sigma_D^2 M)^M$, where $(X, Y; Z)_S$ denotes an $X \times Y$ single-hop point-to-point MIMO system wherein the receiver selects, out of the total Y received antenna signals, the Z that yields the maximum capacity for receiver processing. (See the review in Sec. 3.1)

By (3.12) we may also obtain a lower bound for $C_{DO}(\bar{k})$: Letting

$$\underline{k} = \arg\max_k \frac{\det(\boldsymbol{G}_k \boldsymbol{G}_k^H)}{\prod_i \|\boldsymbol{G}_k^{(i)}\|^2}, \tag{3.16}$$

we have the lower bound C_{DL} as

$$C_{DL} \triangleq C_{DO}(\underline{k}) \leq \max_k C_{DO}(k). \tag{3.17}$$

When $L \gg M$, it becomes more likely to find a set of M nearly orthogonal columns in $\boldsymbol{G}$. In this case, we will have $\det(\boldsymbol{G}_{\underline{k}} \boldsymbol{G}_{\underline{k}}^H)/\prod_i \|\boldsymbol{G}_{\underline{k}}^{(i)}\|^2 \lesssim 1$ and thus

$$C_{DL} = C_{DO}(\underline{k}) \lesssim \det(\boldsymbol{F}_{\underline{k}} \boldsymbol{F}_{\underline{k}}^H)(\frac{P_R}{\sigma_D^2 M})^M. \tag{3.18}$$

Now since $\underline{k}$ is a selection based on $\boldsymbol{G}$ without taking $\boldsymbol{F}$ into consideration and since from (3.18) $\log_2 C_{DL}$ resembles the form of the capacity of an $M \times M$ single-hop point-to-point MIMO system with channel matrix $\boldsymbol{F}_{\underline{k}}$ at transmit-to-receive SNR $P_R^M/(\sigma_D^2 M)^M$, we can view $\log_2 C_{DL}$ as the capacity of an $(M, M; M)_S$ system.

Therefore, from (3.15) and (3.17) we conclude that in the destination-noise dominant condition, the performance of relay selection with optimal power allocation is asymptotically upper-bounded by that of the $(M, L; M)_S$ MIMO antenna selection system and lower-bounded by that of $(M, M; M)_S$. The capacity outage diversity order is thus similarly bounded by that of these two systems. Simulation results in Sec. 3.4 will show that, although the above derivation has been carried out mostly assuming asymptotic conditions, relay selection systems operating in practical conditions exhibit some similar performance characteristics.

3.3 Designs for Relay-Noise Dominant Conditions

We now turn to the relay noise-dominant situation. Again, no closed-form general solution can be found for arbitrary values of L and M, but a solution can be found if they are related in a specific way. We thus again propose a relay selection scheme.

To start, let N out of the L relays be selected to participate in the relaying, where $N \geq M$ but is otherwise undetermined for the moment. Altogether there are $\binom{L}{N}$ selections. For the jth selection we let $\boldsymbol{F}_j \in \mathbb{C}^{M \times N}$ and $\boldsymbol{G}_j \in \mathbb{C}^{M \times N}$ denote the corresponding channel

submatrices of $\boldsymbol{F}$ and $\boldsymbol{G}$, respectively. From the derivation up to (2.9) in Theorem. 2.2.1 we may infer that, in the relay noise-dominant situation,

$$C_R(j, \boldsymbol{R}_R) \lesssim \det(\boldsymbol{I}_M + \sigma_R^{-2} \boldsymbol{G}_j \boldsymbol{V}_j \boldsymbol{V}_j^H \boldsymbol{G}_j^H) \tag{3.19}$$

where $\boldsymbol{R}_R$ is the diagonal matrix of relay gains, $\boldsymbol{V}_j \in \mathbb{C}^{N \times M}$ is the matrix of right singular vectors of $\boldsymbol{F}_j \boldsymbol{R}_R$ with its jth column corresponding to the jth largest singular value of $\boldsymbol{F}_j \boldsymbol{R}_R$. Comparing with the situation addressed in Theorem 2.2.1 (in particular, see (2.13)) we find that relay noise-dominant systems behave similarly to systems with very high relay transmission power. Hence by Theorem 2.2.2, $C_R^{(j)}(\boldsymbol{R}_R)$ is upper-bounded as

$$C_R(j, \boldsymbol{R}_R) \leq \det(\boldsymbol{I}_M + \sigma_R^{-2} \boldsymbol{G}_j \boldsymbol{G}_j^H), \tag{3.20}$$

where equality holds and $C_R(j, \boldsymbol{R}_R)$ is maximized if the rows of $\boldsymbol{F}_j \boldsymbol{R}_R$ span the same space as that of $\boldsymbol{G}_j$. Note that, contrary to the destination noise-dominant case, in the present case the total relay transmission power does not affect the performance at all, only the row space of $\boldsymbol{F}_j \boldsymbol{R}_R$ matters. And the relay network should try to align the row space of $\boldsymbol{F}_j \boldsymbol{R}_R$ with that of $\boldsymbol{G}_j$, which is a beamforming problem.

To proceed, let $\boldsymbol{F}_j^{<i>}$ denote the ith row of $\boldsymbol{F}_j$. Let $\boldsymbol{O}_j \in \mathbb{C}^{N \times (N-M)}$ be a matrix of basis vectors for the orthogonal complement of the row space of $\boldsymbol{G}_j$; that is, $\boldsymbol{O}_j$ is such that $\boldsymbol{G}_j \boldsymbol{O}_j = \boldsymbol{0}$ where $\boldsymbol{0}$ denotes a zero matrix. Also, let $\boldsymbol{\Phi}_{ij} \triangleq \operatorname{diag}(\boldsymbol{F}_j^{<i>}) \boldsymbol{O}_j$. Immediately we have

$$\boldsymbol{r}_R^T \boldsymbol{\Phi}_{ij} = \boldsymbol{r}_R^T \operatorname{diag}(\boldsymbol{F}_j^{<i>}) \boldsymbol{O}_j = (\boldsymbol{F}_j^{<i>})^H \boldsymbol{R}_R \boldsymbol{O}_j \tag{3.21}$$

where $\boldsymbol{r}_R \in \mathbb{C}^N$ is the vector formed of the diagonal elements of $\boldsymbol{R}_R$. To make the row space of $\boldsymbol{F}_j \boldsymbol{R}_R$ equal to that of $\boldsymbol{G}_j$, we may equivalently find $\boldsymbol{r}_R$ such that $\boldsymbol{r}_R^T \boldsymbol{\Phi}_{ij} = \boldsymbol{0}\ \forall i$. For this, define

$$\boldsymbol{\Phi}_j \triangleq [\boldsymbol{\Phi}_{1j}\ \boldsymbol{\Phi}_{2j}\ \cdots\ \boldsymbol{\Phi}_{Mj}] \in \mathbb{C}^{N \times N(N-M)}. \tag{3.22}$$

Then the optimal solution or beamformer $\boldsymbol{r}_R$ should be such that $\boldsymbol{r}_R^T \boldsymbol{\Phi}_j = \boldsymbol{0}$. The existence of such a solution would require $\boldsymbol{\Phi}_j$ to have a non-empty null column space. Therefore let $M(N-M) < N$. Combined with the earlier assumption that $N \geq M$, the only choice is $N = M + 1$ for any $M \geq 2$.

In conclusion, the final solution is given by the selection

$$\bar{j} = \arg \max_{1 \leq j \leq \binom{L}{M+1}} C_R(j, \boldsymbol{R}_R) \tag{3.23}$$

where for each j, $\boldsymbol{R}_R$ is given by $\operatorname{diag}(\boldsymbol{r}_R)$ with $\boldsymbol{r}_R$ being the solution to the equation $\boldsymbol{r}_R^T \boldsymbol{\Phi}_j = \boldsymbol{0}$ and "normalized" such that $\sum_{i=1}^{M+1} (\sigma_R^2 + \|\boldsymbol{G}_j^{(i)}\|^2) |\boldsymbol{r}_R(i)|^2 = P_R$ (where $\boldsymbol{G}_j^{(i)}$ is the ith column of $\boldsymbol{G}_j$ and $\boldsymbol{r}_R(i)$ is the ith element of $\boldsymbol{r}_R$).

Regarding its performance, from (3.20) we see that the resulting capacity measure is approximately given by C_{RO} as follows:

$$\max_j C_R(j, \boldsymbol{R}_R) \approx \max_j \det(\boldsymbol{I}_M + \sigma_R^{-2} \boldsymbol{G}_j \boldsymbol{G}_j^H) \triangleq C_{RO}, \tag{3.24}$$

where the middle expression indicates that $\log_2 C_{RO}$ should behave similarly to an $(M, L; M+1)_S$ MIMO antenna selection system with transmit-to-receive SNR σ_x^2/σ_R^2. We will not develop upper and lower bounds to the capacity performance as in the destination noise-dominant case because (3.24) is already a good approximation.

3.4 Numerical Results

Next, we consider the methods derived under the two dominant noise assumptions. This serves two main purposes. First, their performance is compared with two benchmarks, namely, that of equal-gain allocation and that obtained with alternating iterative algorithm (discussed in Sec. 2.3). And secondly, their capacity outage behavior is observed. It is suggested in Sec. 3.2 and 3.3 that more relays would represent larger capacity outage diversity, which is expected to be presented in simulation. In this we also look at how close the bounds C_{DU} and C_{DL} and the approximation C_{RO} are to the actual results. To demonstrate the capacity outage diversity, we present the cumulative distribution function (CDF) of capacity (with respect to capacity).

For the destination noise-dominant case, Fig. 3.2 shows some cumulative distribution function (CDF) curves of the obtained capacity at $M = 3$, $L = 6$, $\sigma_D = 0.1$ and $P_R = 1$. Not surprisingly, the iterative algorithm performs better than the suboptimal relay-selection solutions, and the equal-gain allocation performs worse. The C_{DU} curve is rather close to the iterative algorithm results at the same σ_D/σ_R ratio. As to the relay-selection solutions, we see that the capacity performance drops as σ_R increases (which worsens the SNR). But even though the destination noise becomes less dominant with increasing σ_R, the capacity outage diversity order (indicated by the slope of the curve) remains similar and similar to that of C_{DU}.

Next, we consider how the diversity order varies with number of relays (L). Fig. 3.3 shows some results with all system parameters the same as above except for a fixed $\sigma_R = 10^{-2}$ and a variable L. As the purpose is to examine the diversity order behavior but not the actual capacity, we "bias" the CDF curves horizontally to make their 50% points co-located at zero capacity. The curves verify that the proposed relay selection method indeed yields a similar diversity order to C_{DU}, and the diversity order increases (i.e., the CDF curve steepens) with number of relays. To compare with the diversity behavior of C_{DL}, we also show a curve for a $(3, 3; 3)_S$ system.

Now consider the design based on the relay noise-dominant assumption. Fig. 3.4 shows some results. Again, the iterative algorithm performs better and the equal-gain worse. And we verify that the diversity order behavior at $\sigma_R/\sigma_D = 10$ is similar to that of a $(3, 6; 4)_S$ system (the behavior of C_{RO}). As expected, capacity drops as σ_D increases (which lowers the SNR and also makes the relay noise less dominant). But the great difference with Fig. 3.2 is the reduction in diversity order (i.e., reduction in steepness of CDF curve) with reduced relay noise-dominance.

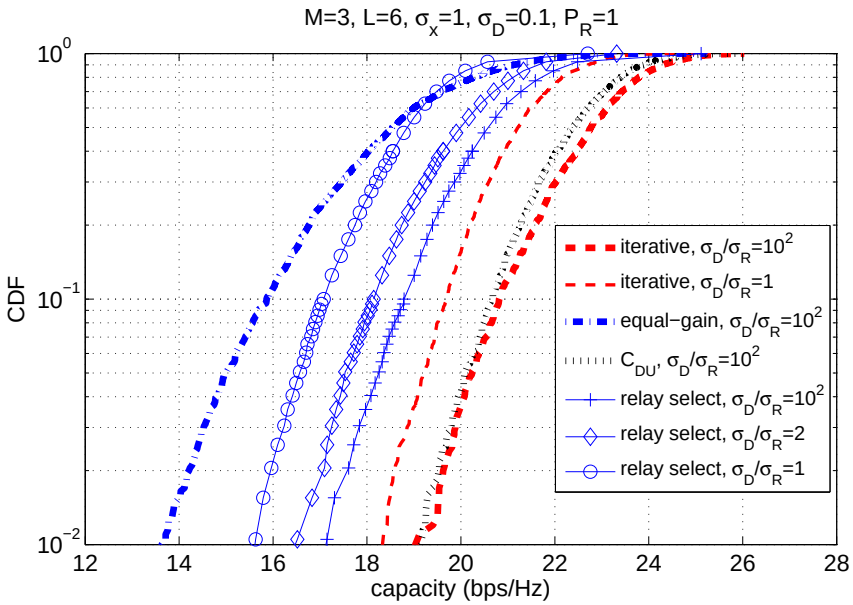

Fig. 3.2: Capacity CDF of distributed relay system designed under destination noise-dominant assumption.

In Fig. 3.5, we compare the capacity CDFs of distributed relay networks of different sizes, all designed with the relay selection method for the relay noise-dominant condition, with the $(M, L; M+1)_S$ MIMO antenna selection systems. We see that the performance of the latter tightly upper-bounds the corresponding distributed relay systems.

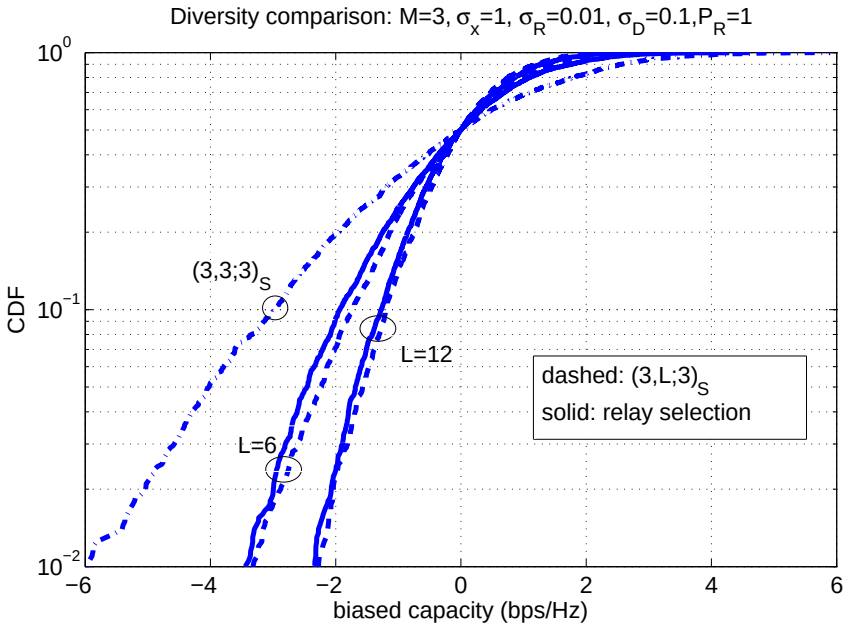

Fig. 3.3: Horizontally "biased" capacity CDF curves of distributed relay systems for diversity comparison.

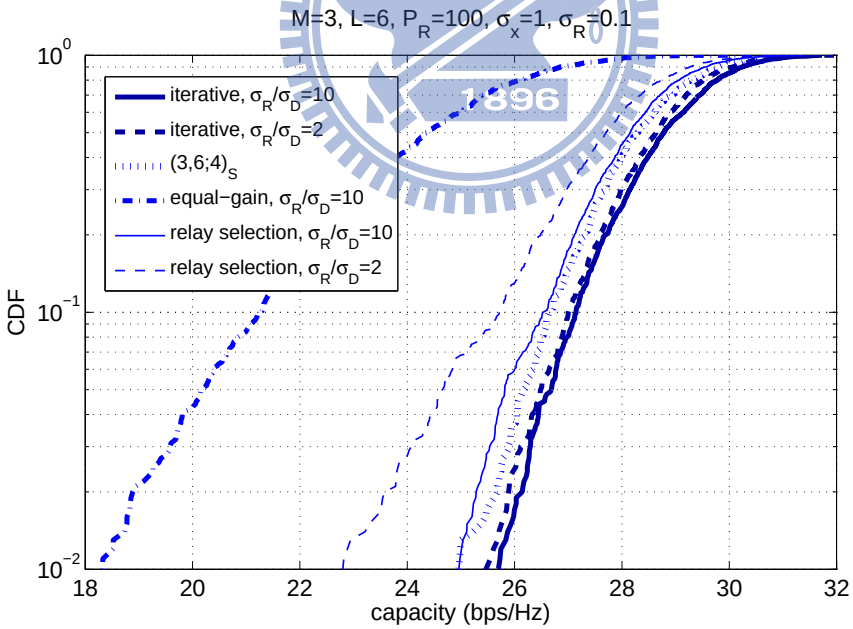

Fig. 3.4: Capacity CDF of distributed relay system designed under relay noise-dominant assumption.

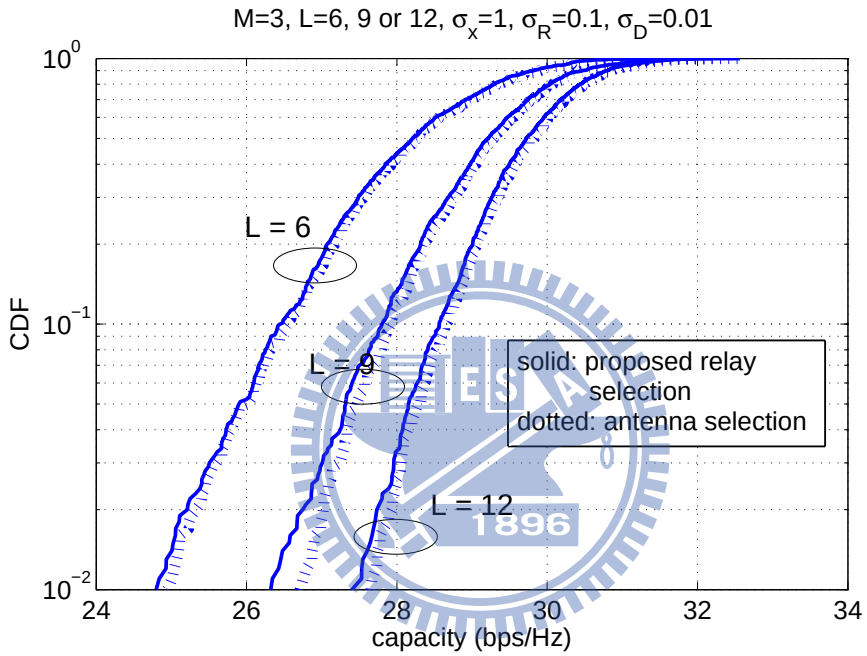

Fig. 3.5: Comparison of diversity orders of capacity CDFs of distributed relay systems of different sizes designed under relay noise-dominant assumption.

Chapter 4

Efficient Algorithms for Noise-Dominant Models

We discuss system simplification and approximation of distributed relay network based on noise-dominant models in Chap. 3, wherein relay selection algorithms and associated relay gains designs are also presented. Though the algorithms based on relay selection is conceptually concise, the computation burden would growth exponentially as the number of relays increases. To circumvent the problem, in this chapter we discuss algorithms based on noise-dominant models but avoid relay selection. As the foundation of proposed efficient designs, we briefly review relay selection algorithms in Sec. 4.1.2 and Sec. 4.2.1.

For relay-noise dominant model we start from the equivalent beamforming problem discussed in Sec.3.3. Since we consider L relay gains simultaneously and optimal beamformer is not available, we cast the problem into projections onto two subspace and minimizing the ratio of projected vector norms. In Sec.4.1.1 the idea and algorithm for multiuser low-leakage beamforming [35] are briefly reviewed. We would apply the algorithm to solve minimization of norms in Sec. 4.1.3.

As for destination-noise dominant model, in Sec. 4.2.2 we simplify the problem by making the end-to-end MIMO channel $\boldsymbol{H}$ in (2.3) an upper-triangular matrix so that the matrix determinant maximization in (2.4) could be approximated as product of diagonal terms of $\boldsymbol{H}$. In other words, we zero-force some terms of $\boldsymbol{H}$ then focus on a product maximization problem. In Sec. 4.2.3 we transform the design problem with specifically constrained solution space and propose iterative algorithm to reach local optimizer.

It is the potentially expensive computation of selection-based algorithms that motivates us to develop efficient designs for noise-dominant models. Thus in Sec. 4.1.4 and 4.2.4 we would summarize all the proposed algorithms (with and without relay selection), and compare the order of computation complexity for selection-based algorithms and proposed efficient designs. Finally the respective performance would be shown and compared in Sec.4.3.

4.1 Designs for Relay-Noise Dominant Conditions

4.1.1 Downlink Low-leakage Beamforming Design

Assume base station is equipped with N transmission antennas and serves K users. For the kth user, base station send signal scalar s_k via dedicated transmission beamforming vector $\boldsymbol{r}_k \in \mathbb{Z}^N$. Thus the transmitted signal vector $\boldsymbol{x}$ could be described as

$$\boldsymbol{x} = \sum_{k=1}^{K} \boldsymbol{r}_k s_k, \tag{4.1}$$

where $\boldsymbol{r}_k$ and s_k is normalized such that $\mathrm{E}\{|s_k|^2\} = 1$, $\|\boldsymbol{r}_k\|^2 = 1$. Assume the number of receiver antennas at the kth user is M_k, and the MIMO channel matrix between base station and the kth user is $\boldsymbol{F}_k \in \mathbb{Z}^{M_k \times N}$. Then the reception vector of the ith user, $\boldsymbol{y}_i$, could be modeled as

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{F}_i \boldsymbol{x} + \boldsymbol{v}_k \\ &= \boldsymbol{F}_i \boldsymbol{r}_i s_i + \sum_{\forall k \neq i} \boldsymbol{F}_i \boldsymbol{r}_k s_k + \boldsymbol{v}_k, \end{aligned} \tag{4.2}$$

where $\boldsymbol{v}_k \sim \mathcal{CN}(0, \sigma_k^2 \boldsymbol{I})$ denotes the reception noise of the ith user. Note the first term in the RHS of (4.2) represents desired downlink signal for the ith user, while the second terms represents unwanted interference.

Following (4.2) the SINR (signal to interference plus noise ratio) at the input of the ith user could be stated as

$$\mathrm{SINR}_i = \frac{\|\boldsymbol{F}_i \boldsymbol{r}_i\|^2}{M_i \sigma_i^2 + \sum_{\forall k \neq i} \|\boldsymbol{F}_i \boldsymbol{r}_k\|^2}. \tag{4.3}$$

Note the beamformers meant for other users, $\boldsymbol{r}_k (\forall k \neq i)$, account for the interference to the ith user. It is conceptually depicted in Fig. 4.1 with $i = 1$. SINR is a performance index of fundamental importance and frequently used as design criteria. However, it is challenging to optimize K beamformers simultaneously with K SINR index. Another potential approach is to arrange beamformers so that the interference is completely nulled out at transmitter. It is intuitive and simple, but would require larger size of transmitter antennas to ensure that $\boldsymbol{F}_k \boldsymbol{r}_i = \boldsymbol{0}$ $\forall i, k \neq i$. Thus base station may need additional antennas to serve more users.

To accommodate the above-mentioned problems of SINR-based beamforming design, another relaxed but viable criterion is proposed in [35]. Instead of considering interfering terms $\boldsymbol{F}_i \boldsymbol{r}_k$ $\forall i \neq k$, Sadek *et al.* proposed applying SLNR (signal to leakage plus noise ratio) measure which emphasis the *leakage* signals from $\boldsymbol{r}_i$ to users $j (\forall k \neq i)$. The SLNR

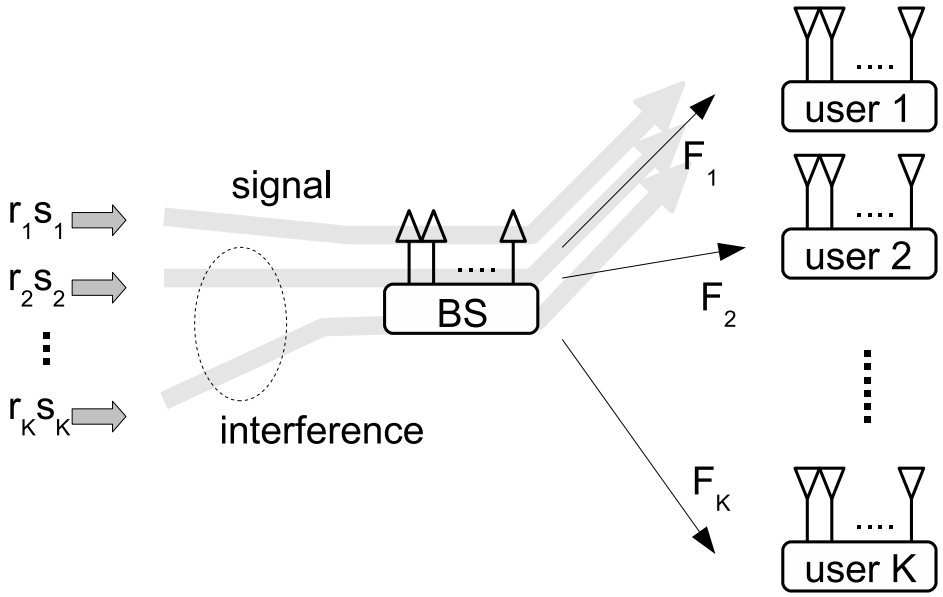

Fig. 4.1: Downlink signal and interference flows for user 1

measure for the ith user could be stated as

$$\text{SLNR}_i = \frac{\|\boldsymbol{F}_i\boldsymbol{r}_i\|^2}{M_i\sigma_i^2 + \sum_{\forall k\neq i}\|\boldsymbol{F}_k\boldsymbol{r}_i\|^2} \triangleq \frac{\|\boldsymbol{F}_i\boldsymbol{r}_i\|^2}{M_i\sigma_i^2 + \|\boldsymbol{\Phi}_i\boldsymbol{r}_i\|^2} \tag{4.4}$$

and depicted in Fig. 4.2 assuming $i = 1$. $\boldsymbol{\Phi}_i$ is defined to be a matrix composed of $\boldsymbol{F}_k(\forall k \neq i)$ and is used to measure the leakage from $\boldsymbol{r}_i$ It is clear that SLNR_i depends only on $\boldsymbol{r}_i$ and is easier to be optimized. The optimal solution of $\boldsymbol{r}_i$ that maximize SLNR_i is shown to be [35] the scaled version of eigenvector of $(M_i\sigma_i^2\boldsymbol{I} + \boldsymbol{\Phi}_i^H\boldsymbol{\Phi}_i)^{-1}\boldsymbol{F}_i^H\boldsymbol{F}_i$ corresponding to the maximal eigenvalue. Note that maximizing SLNR_i is equivalent to design $\boldsymbol{r}_i$ with two criteria, minimizing $\|\boldsymbol{\Phi}_i\boldsymbol{r}_i\|$ and maximizing $\|\boldsymbol{F}_i\boldsymbol{r}_i\|$, exercised simultaneously. In later sections we would again apply the composite criteria for relay network design.

4.1.2 Capacity Approximation and Relay Selection

For the case of relay-noise dominating mode, the significance of $\boldsymbol{I}_M$ in (2.2) vanishes. Following Sec. 3.3 we derive the upper bound and approximation of capacity measure as

$$C(\boldsymbol{R}) \lesssim \det(\boldsymbol{I}_M + \frac{\sigma_x^2}{\sigma_R^2}\boldsymbol{G}\boldsymbol{V}\boldsymbol{V}^H\boldsymbol{G}^H) \tag{4.5}$$

$$\leq \det(\boldsymbol{I}_M + \frac{\sigma_x^2}{\sigma_R^2}\boldsymbol{G}\boldsymbol{G}^H), \tag{4.6}$$

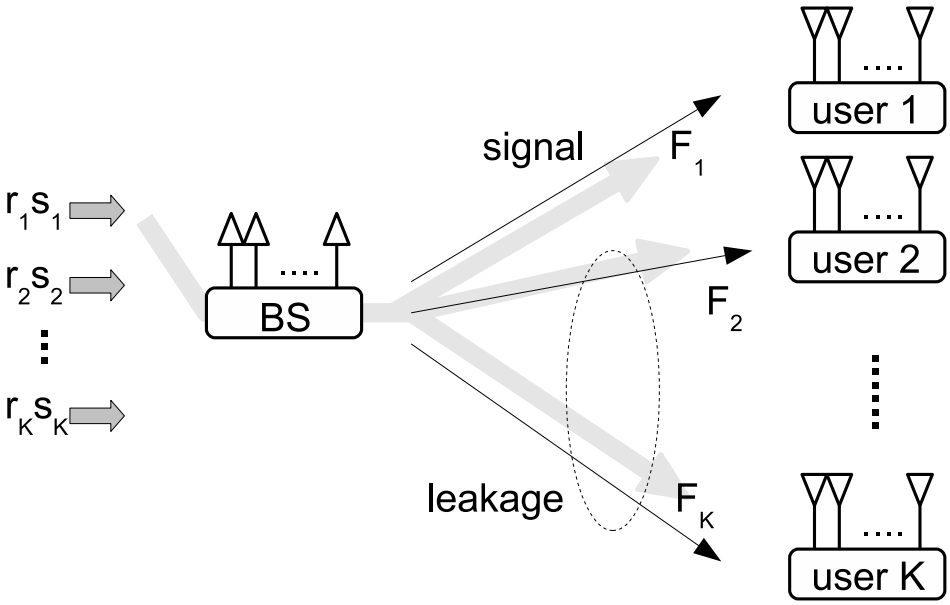

Fig. 4.2: Downlink signal and leakage flows from user 1

where $\boldsymbol{V} \in \mathbb{C}^{L\times M}$ is the matrix composed of right singular vectors of $\boldsymbol{FR}$ corresponding to the M largest singular values, and the equality in (4.6) holds when $\boldsymbol{V}$ spans the range of $\boldsymbol{G}^H$. To explain how and when this condition is satisfied, we first define following matrices: let $\boldsymbol{V}_{GR} \in \mathbb{C}^{L\times M}$ be the orthogonal basis spanning range of $\boldsymbol{G}^H$, and let $\boldsymbol{V}_{GN}$ be the orthogonal complement of $\boldsymbol{V}_{GR}$. Also we define $\boldsymbol{\Phi}_i$ and $\boldsymbol{\Phi}$ such that

$$\boldsymbol{\Phi}_i \triangleq \operatorname{diag}(\boldsymbol{F}^{<i>})\boldsymbol{V}_{GN} \tag{4.7}$$

$$\boldsymbol{\Phi} \triangleq [\boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \cdots \boldsymbol{\Phi}_M] \in \mathbb{C}^{L\times M(L-M)}, \tag{4.8}$$

where $\boldsymbol{F}^{<i>}$ stands for the ith row vector of $\boldsymbol{F}$.

Since we require $\boldsymbol{V}$ spans the range of $\boldsymbol{G}^H$, the row vectors of $\boldsymbol{FR}$ should not fall in the null space of $\boldsymbol{G}^H$. That is, $(\boldsymbol{FR})^{<i>}\boldsymbol{V}_{GN} = \boldsymbol{0}\ \forall i$, which could be rewritten as

$$\begin{aligned}(\boldsymbol{FR})^{<i>}\boldsymbol{V}_{GN} &= \boldsymbol{r}^T \operatorname{diag}(\boldsymbol{F}^{<i>})\boldsymbol{V}_{GN} \\ &= \boldsymbol{r}^T\boldsymbol{\Phi}_i = \boldsymbol{0}\ \forall i,\end{aligned} \tag{4.9}$$

where $\boldsymbol{0}$ means the matrix with appropriate size and all zero terms. Equivalently we should find the solution to $\boldsymbol{r}^T\boldsymbol{\Phi} = \boldsymbol{0}$, and scale $\boldsymbol{r}$ such that power limitation (2.5) is satisfied. A non-trivial solution requires non-empty null space of $\boldsymbol{\Phi}^T$, which means the size of relay networks should satisfies $L = M + 1$ (see Sec. 3.3 for detail).

When it comes to larger relay networks ($L > M + 1$), one simple extension based on (4.9) is to select $M + 1$ relays out of L. Since we may guarantee asymptotic capacity reaching the upper bound in (4.6) for $M + 1$ relays, the selection with best value in (4.6) would be used to activate the selected relays and deactivate otherwise.

4.1.3 Capacity Improvement based on Maximizing the Ratio of Norms

For larger size of relay network, the exhaustive $\binom{L}{M+1}$ selections would require considerable computation as L increases, thus we may resort to another approach to design $\boldsymbol{R}$ efficiently. Recall in Sec. 4.1.2 the fundamental is to design $\boldsymbol{R}$ such that $\boldsymbol{G}^H$ and $(\boldsymbol{FR})^H$ have identical range space. When $L \neq M+1$ we could still apply the principle and design the algorithm with composite criteria.

On one hand, though in (4.9) we could not find $\boldsymbol{r}$ as the solution to $\boldsymbol{r}^T\boldsymbol{\Phi} = \boldsymbol{0}$ when $L \neq M+1$, we may try to minimize $\|\boldsymbol{r}^T\boldsymbol{\Phi}\|$ as an approximated criterion based on (4.9). On the other hand, note that

$$\|(\boldsymbol{FR})^{<i>}\boldsymbol{V}_{GR}\| \leq \|(\boldsymbol{FR})^{<i>}\| \ \forall i$$

and the equality holds when $\boldsymbol{G}^H$ and $(\boldsymbol{FR})^H$ have identical range space, which indicates that we should try to maximize $\|(\boldsymbol{FR})^{<i>}\boldsymbol{V}_{GR}\|$. To do so we define $\boldsymbol{\Psi}_i$ and $\boldsymbol{\Psi}$ as

$$\boldsymbol{\Psi}_i \triangleq \text{diag}(\boldsymbol{F}^{<i>})\boldsymbol{V}_{GR} \tag{4.10}$$

$$\boldsymbol{\Psi} \triangleq [\boldsymbol{\Psi}_1\boldsymbol{\Psi}_2\cdots\boldsymbol{\Psi}_M] \in \mathbb{C}^{L\times M^2}. \tag{4.11}$$

Then maximizing $\|(\boldsymbol{FR})^{<i>}\boldsymbol{V}_{GR}\|$ is equivalent to maximizing $\|\boldsymbol{r}^T\boldsymbol{\Psi}\|$.

Now we need to handle norm minimization and maximization simultaneously. The design of $\boldsymbol{r}$ could be realized by maximizing the ratio of norms as

$$\boldsymbol{r}_{opt} = \arg\max_{\boldsymbol{r}} \frac{\|\boldsymbol{r}^T\boldsymbol{\Psi}\|}{\|\boldsymbol{r}^T\boldsymbol{\Phi}\|}. \tag{4.12}$$

Recall we discuss low-leakage beamforming technique in Sec.4.1.1. The criterion in (4.12) resembles that in (4.4). Thus we derive $\boldsymbol{r}_{opt}$ such that it should be proportional to the eigenvector of $(\boldsymbol{\Phi}\boldsymbol{\Phi}^H)^{-1}\boldsymbol{\Psi}\boldsymbol{\Psi}^H$ corresponding to the maximum eigenvalue, and is scaled according to (2.5). Note the $\boldsymbol{\Phi}$ in (4.12) is wide matrix, so $\boldsymbol{\Phi}\boldsymbol{\Phi}^H$ is invertible (given $\boldsymbol{\Phi}$ is full rank).

4.1.4 Algorithms Summary and Complexity Comparison

For better understanding and comparison, in what follows we summarize and itemize the two algorithms (with and without relay selection) for relay-noise . Moreover, among the steps of algorithms we highlight some key procedures that are computationally demanding or repeatedly executed. Based on the assessment of computation burden for these steps, we may obtain a rough yet useful comparison of complexity of the two algorithms.

For the algorithm based on relay selection in Sec. 4.1.2, the procedures could be summarized as follows:

1. Set $\binom{L}{M+1}$ combinations to represent all possible selection of choosing $M+1$ relays out of L.

2. For the ith combination ($1 \leq i \leq \binom{L}{M+1}$), we define $\boldsymbol{F}_i$ and $\boldsymbol{G}_i$ as the submatrix of $\boldsymbol{F}$ and $\boldsymbol{G}$ respectively and is corresponding to the selected $M+1$ relays for the ith selection. Also we denote $\boldsymbol{r}_i$ as the corresponding $M+1$ relay gains.

3. For all the combination, we compute the individual capacity approximation as (4.6).

4. Find the combination with largest capacity and set relay gains for the selected relays by following steps. The remaining (unselected) gains are set as zero.

5. Assume the index of selected combination is k. Define matrix $\boldsymbol{V}_{GN}$ be the orthogonal basis spanning the null space of $\boldsymbol{G}_k$ (*i.e.* $\boldsymbol{G}_k\boldsymbol{V}_{GN} = \boldsymbol{0}$). Set matrix $\boldsymbol{\Phi}$ based on (4.7) and (4.8).

6. Set $\boldsymbol{r}_k$ as the solution of $\boldsymbol{r}_k^T\boldsymbol{\Phi} = 0$ then scale $\boldsymbol{r}_k$ such that the power limitation (2.5) is satisfied.

In step 3 the capacity approximation for each combination requires calculating matrix determinant and matrix multiplication. We may apply QR matrix decomposition based on Gram-Schmidt process to obtain matrix determinant, and consume computation of order $\mathcal{O}((M+1)^3) = \mathcal{O}(M^3)$ [42]. The matrix multiplication requires $M^2(2M+1)$ flops (complex Floating-Point number Operations Per Second), thus again is of order $\mathcal{O}(M^3)$. Since the computation of matrix determinant and multiplication is repeated for all the combination, the computational complexity associated with step 3) is about $\mathcal{O}(L^{M+1}M^3)$.

Now we summarize the procedures of the algorithm based on maximizing the ratio of norms in Sec. 4.1.3 as follows:

1. Define matrix $\boldsymbol{V}_{GN}$ be the orthogonal basis spanning the null space of $\boldsymbol{G}$ (*i.e.* $\boldsymbol{G}\boldsymbol{V}_{GN} = \boldsymbol{0}$), and $\boldsymbol{V}_{GR}$ be the orthogonal complement of $\boldsymbol{V}_{GN}$.

2. Set matrix $\boldsymbol{\Phi}$ based on (4.7) and (4.8). Also set matrix $\boldsymbol{\Psi}$ in a similar way based on (4.10) and (4.11).

3. Following (4.12) we set $\boldsymbol{r}$ as the eigenvector of $(\boldsymbol{\Phi}\boldsymbol{\Phi}^H)^{-1}\boldsymbol{\Psi}\boldsymbol{\Psi}^H$ corresponding to its maximum eigenvalue, and scale the resulting $\boldsymbol{r}$ such that the power limitation (2.5) is satisfied.

In step 1 we perform QR decomposition for $\boldsymbol{G}$ which costs about $2L^3$ flops [42]. Among the three matrix multiplications in step 3), the one with largest matrix size, $\boldsymbol{\Phi}\boldsymbol{\Phi}^H$, takes $L^2(2ML - 2M^2 - 1)$ flops. Thus for large size of relay the matrix multiplications are of order $\mathcal{O}(ML^3)$. Also found in 3 is an inversion for $L \times L$ Hermitian matrix, which would require $L^3 + L^2 + L$ flops [18]. Based on these assessment we conclude that the algorithm computation is of order $\mathcal{O}(ML^3)$.

4.2 Designs for Destination-Noise Dominant Conditions

4.2.1 Capacity Approximation and Relay Selection

For the case of destination-noise dominating mode, in (2.2) and (2.3) the matrix $\boldsymbol{W}$ could be approximated as $\sigma_D^2 \boldsymbol{I}$. So we have

$$C(\boldsymbol{r}) \lesssim \det(\boldsymbol{I}_M + \frac{\sigma_x^2}{\sigma_D^2} \boldsymbol{H}^H \boldsymbol{H}). \tag{4.13}$$

Clearly $C(\boldsymbol{r})$ is now closely related to $|\det(\boldsymbol{H})|$, especially when system works with high SNR. Therefore it is reasonable to use $|\det(\boldsymbol{H})|$ as alternative cost function to be maximized during system design. In other words, (2.4) could be replaced as

$$\begin{aligned} \boldsymbol{r}_{opt} &= \arg\max_{\boldsymbol{r}} |\det(\boldsymbol{H})| \\ \text{subject to } &\sum_{\forall i} p(i)|\boldsymbol{r}(i)|^2 \le P_R. \end{aligned} \tag{4.14}$$

For the case $L = M$ and full-rank square $\boldsymbol{F}$ and $\boldsymbol{G}$, (4.14) turns out to be

$$\boldsymbol{r}_{opt} = \arg\max_{\boldsymbol{r}} |\det(\boldsymbol{F})\det(\boldsymbol{G})| \prod_{\forall i} |\boldsymbol{r}(i)| \tag{4.15}$$

with optimal solution as

$$|\boldsymbol{r}_{opt}(i)| = \sqrt{\frac{P_R}{M p(i)}}. \tag{4.16}$$

For the case $L > M$, we may select M relays out of L such that (4.16) is applied for each selection and choose the combination with largest $|\det(\boldsymbol{H})|$.

4.2.2 Capacity Improvement based on Partial Zero Forcing

When $L > M$, $|\det(\boldsymbol{H})|$ could not be decoupled as (4.15) therefore the solution in (4.16) is no longer applicable. Selection-based approaches may simply become too expensive as the combinations grows exponentially when L increases. Therefore another efficient algorithm is required if we would like to handle and make use of all relays simultaneously.

Denote $\boldsymbol{H}(i,j)$ as the element at the ith row and jth column of $\boldsymbol{H}$. The determinant of $\boldsymbol{H}$ is defined to be

$$\det(\boldsymbol{H}) \triangleq \sum_{\forall \boldsymbol{u}} a(\boldsymbol{u}) \prod_{i=1}^{M} \boldsymbol{H}(i, \boldsymbol{u}(i)), \tag{4.17}$$

where the summation is taken over the $M!$ permutations $\boldsymbol{u} = [\boldsymbol{u}(1)\boldsymbol{u}(2)\cdots\boldsymbol{u}(M)]$ of $(1, 2\cdots M)$ [28], and the scalar $a(\boldsymbol{u}) \in \{1, -1\}$ is a function of $\boldsymbol{u}$. Directly pursuing the optimization of (4.17) is complicate and impracticable. To approach a viable algorithm we first simplify (4.17) by designing $\boldsymbol{R}$ such that partial terms of $\boldsymbol{H}$ are zero forced and $\boldsymbol{H}$ becomes a lower-triangular matrix. Equivalently we would have

$$\boldsymbol{H}(i,j) = 0, 1 \leq i \leq M, i < j \leq M. \tag{4.18}$$

$$\det(\boldsymbol{H}) = \prod_{i=1}^{M} \boldsymbol{H}(i,i). \tag{4.19}$$

Note the matrix determinant computation is now greatly simplified as product of diagonal terms.

Note that $\boldsymbol{H}(i,j)$ could be expressed as $\boldsymbol{H}(i,j) = (\boldsymbol{F}^{<i>} \odot \overline{\boldsymbol{G}}^{<j>})\boldsymbol{r}$, where $\odot$ represents element-wise product, and $\overline{\boldsymbol{G}}^{<j>}$ denotes the conjugate of $\boldsymbol{G}^{<j>}$. Combining (4.18) and (4.19) we modify (4.14) as

$$\boldsymbol{r}_{opt} = \arg\max_{\boldsymbol{r}} |\prod_{\forall i} \boldsymbol{H}_R^{<i>}\boldsymbol{r}| \tag{4.20}$$

subject to

$$\boldsymbol{H}_N\boldsymbol{r} = \boldsymbol{0}, \tag{4.21}$$

$$\sum_{\forall i} p(i)|\boldsymbol{r}(i)|^2 \leq P_R, \tag{4.22}$$

where $\boldsymbol{H}_R$ is defined such that its the ith row $\boldsymbol{H}_R^{<i>}$ is equal to $\boldsymbol{F}^{<i>} \odot \overline{\boldsymbol{G}}^{<i>}$. That is, $\boldsymbol{H}_R\boldsymbol{r} = \text{diag}(\boldsymbol{H})$. $\boldsymbol{H}_N \in \mathbb{C}^{\frac{M(M-1)}{2} \times L}$ is defined in a similar way such that $\boldsymbol{H}_N\boldsymbol{r}$ corresponds to the upper-half elements (besides diagonal terms) of $\boldsymbol{H}$. To ensure non-empty solution domain in (4.21) we require $L > \frac{M(M-1)}{2}$, which would be true for large relay systems.

To transform (4.20), (4.21) and (4.22) into more tractable forms, we first introduce a diagonal matrix $\boldsymbol{S}$ whose ith diagonal term is equal to $p(i)^{-0.5}$, and define $\boldsymbol{r}_S \triangleq \boldsymbol{S}^{-1}\boldsymbol{r}$. Following (4.21) we know $\boldsymbol{H}_N\boldsymbol{S}\boldsymbol{r}_S$ should be $\boldsymbol{0}$. For this we define $\boldsymbol{V}_N \in \mathbb{C}^{L \times (L - \frac{M(M-1)}{2})}$ as the orthogonal basis of null space of $\boldsymbol{H}_N\boldsymbol{S}$, and define $\boldsymbol{r}_N$ such that $\boldsymbol{r}_N \triangleq \boldsymbol{V}_N^H\boldsymbol{r}_S$. Now the optimization problem becomes

$$\boldsymbol{r}_{opt} = \boldsymbol{S}\boldsymbol{V}_N \arg\max_{\boldsymbol{r}_N} |\prod_{\forall i} (\boldsymbol{H}_R\boldsymbol{S}\boldsymbol{V}_N)^{<i>}\boldsymbol{r}_N| \tag{4.23}$$

$$\text{subject to } \sum_{\forall i} |\boldsymbol{r}_N(i)|^2 = P_R. \tag{4.24}$$

Note we set equality constraint in (4.24) since scaling $\boldsymbol{r}$ (or $\boldsymbol{r}_N$) with a positive real number would always be beneficial in (4.23).

For convenience of algorithm presentation, we assume $\boldsymbol{H}_R\boldsymbol{S}\boldsymbol{V}_N \in \mathbb{C}^{M\times(L-\frac{M(M-1)}{2})}$ be a full-rank wide matrix , while the proposed algorithm could be easily extended when $\boldsymbol{H}_R\boldsymbol{S}\boldsymbol{V}_N$ is tall matrix. We do SVD (singular value decomposition) such that $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{V}_R^H = \boldsymbol{H}_R\boldsymbol{S}\boldsymbol{V}_N$, where $\boldsymbol{A}$ is unitary, $\boldsymbol{\Sigma}$ is diagonal matrix, and $\boldsymbol{V}_R \in \mathbb{C}^{(L-\frac{M(M-1)}{2})\times M}$. Then we arrange another transforming by defining $\boldsymbol{\epsilon} \triangleq \boldsymbol{V}_R^H\boldsymbol{r}_N$, and obtain the optimization problem as

$$\boldsymbol{r}_{opt} = \boldsymbol{S}\boldsymbol{V}_N\boldsymbol{V}_R \arg\max_{\boldsymbol{\epsilon}} |\prod_{\forall i}(\boldsymbol{A}\boldsymbol{\Sigma})^{<i>}\boldsymbol{\epsilon}| \tag{4.25}$$

$$\text{subject to } \boldsymbol{\epsilon}^H\boldsymbol{\epsilon} = P_R, \tag{4.26}$$

Now we could focus on $\max \prod_i(\boldsymbol{A}\boldsymbol{\Sigma}^{<i>}\boldsymbol{\epsilon})$. To best of our knowledge (4.25) and (4.26) could not be handled by closed-form solution or convex programming, thus we design iterative algorithm for an suboptimal solution based on the following observations:

1. Define $\boldsymbol{\eta} \triangleq \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{\epsilon}$ and $c_\eta \triangleq \|\boldsymbol{\eta}\|$. Given c_η being a known and fixed value, by Jensen's inequality the upper bound of $|\prod_i \boldsymbol{\eta}(i)|$ would be $(c_\eta/\sqrt{M})^M$, where the bound is reached when $|\boldsymbol{\eta}(i)| = c_\eta/\sqrt{M}\ \forall i$. In other words, given a fixed $\|\boldsymbol{\eta}\|$, we would conclude that $|\prod_i \boldsymbol{\eta}(i)|$ is maximized when all elements in $\boldsymbol{\eta}$ have identical absolute value.

2. Assume the diagonal terms of $\boldsymbol{\Sigma}$ are listed in descending order. Since $\boldsymbol{A}$ is unitary, we know $\|\boldsymbol{\eta}\| = \|\boldsymbol{\Sigma}\boldsymbol{\epsilon}\|$. Given $\|\boldsymbol{\epsilon}\|^2 = P_R$ as limited by (4.26), it is obvious that $\boldsymbol{\Sigma}(M,M)\sqrt{P_R} \le \|\boldsymbol{\eta}\| \le \boldsymbol{\Sigma}(1,1)\sqrt{P_R}$, where the upper and lower bound are reached when all elements in $\boldsymbol{\epsilon}$ are zeros except $|\boldsymbol{\epsilon}(1)| = \sqrt{P_R}$ and $|\boldsymbol{\epsilon}(M)| = \sqrt{P_R}$, respectively.

Following the observations, we set two principles for optimization:

1. Based on observation 1, when the transformed $\boldsymbol{\eta}$ shows the unique form that all its elements has identical absolute value, the product value $|\prod_i \boldsymbol{\eta}(i)|$ reaches maximum in terms of c_η. Consequently we would force all elements in $\boldsymbol{\eta}$ to have identical absolute value. That is, $\boldsymbol{\eta}$ could be described as $\boldsymbol{\eta} = (c_\eta/\sqrt{M})\boldsymbol{\theta}$, where $\boldsymbol{\theta} \in \mathbb{C}^M$ and $|\boldsymbol{\theta}(i)| = 1\ \forall i$.

2. Based on observation 2, the 2-norm of $\boldsymbol{\epsilon}$ is fixed to P_R while $\|\boldsymbol{\eta}\|^2$ is varying. Since $\|\boldsymbol{\eta}\|^2$ affects the achievable upper bound of $|\prod_i \boldsymbol{\eta}(i)|$, we should consider maximizing $\|\boldsymbol{\eta}\|^2$.

Combining (4.25), (4.26) and the principles, we arrive

$$|\prod_{\forall i}(\boldsymbol{A\Sigma})^{<i>}\boldsymbol{\epsilon}| = |\prod_{\forall i}\boldsymbol{\eta}(i)| = (c_\eta/\sqrt{M})^M$$
$$\|\boldsymbol{\epsilon}\| = \sqrt{P_R} = (\frac{c_\eta}{\sqrt{M}})\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\|.$$

Clearly maximizing in (4.25) is now equivalent to minimizing $\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\|$. The new forms of optimization problems could be stated as

$$\boldsymbol{r}_{opt} = \boldsymbol{S}\boldsymbol{V}_N\boldsymbol{V}_R\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\arg\min_{\boldsymbol{\theta}}\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\| \tag{4.27}$$
$$\text{subject to } \boldsymbol{\theta}\in\mathbb{C}^M, |\boldsymbol{\theta}(i)| = 1\ \forall i \tag{4.28}$$

Note the solution to (4.27) is suboptimal to (4.25), because the optimizer for (4.25) may not result in $\boldsymbol{\eta}$ with the form $(c_\eta/\sqrt{M})\boldsymbol{\theta}$ and thus would be ruled out of the domain of (4.28). Since the actual optimizer is not available, we have no way to assess the the performance degradation of suboptimality. With simulation results, however, it is shown to be an efficient yet powerful approach.

Also note that though we apply lower-triangular matrix to arrange the zero forcing and approximate computation of matrix determinant, according to (4.17) it is possible use other zero forcing arrangements (for example, do row permutation of $\boldsymbol{H}$ and/or set upper-triangular zero forcing). However, selecting among the potential arrangements is beyond the scope of this work. What we focus here is to design efficient approach given zero forcing arrangement is fixed.

4.2.3 Iterative Greedy Optimization

To handle the optimization for (4.27) we consider iterative greedy algorithm which gradually improve the cost function in (4.27). The idea is to select one element in $\boldsymbol{\theta}$, say $\boldsymbol{\theta}(k)$, as varying variable and optimize it in each iteration, then pick another $\boldsymbol{\theta}(j)$ $j\neq k$ in next iteration. To clearly separate and express the varying and invariant parts, we define

$$\boldsymbol{u} \triangleq (\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1})^{(k)}, \tag{4.29}$$
$$\boldsymbol{u}_o \triangleq \sum_{\forall\tilde{k}\neq k}(\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1})^{(\tilde{k})}\boldsymbol{\theta}(\tilde{k}), \tag{4.30}$$
$$\boldsymbol{\theta}_\Delta\in\mathbb{C}^M, \boldsymbol{\theta}_\Delta(i) \triangleq \exp(j\angle\boldsymbol{u}(i) - j\angle\boldsymbol{u}_o(i)), \tag{4.31}$$

where $\angle \boldsymbol{u}(i)$ denotes the phase of $\boldsymbol{u}(i)$. With these definitions we rewrite the square of cost function in (4.27) as

$$\begin{aligned}
&\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\|^2 = \|\boldsymbol{\theta}(k)\boldsymbol{u} + \boldsymbol{u}_o\|^2 \\
=&\|\boldsymbol{\theta}(k)|\boldsymbol{u}| \odot \boldsymbol{\theta}_\Delta + |\boldsymbol{u}_o|\ \|^2 \\
=&\sum_{i=1}^{M} |\boldsymbol{u}(i)|^2 + |\boldsymbol{u}_o(i)|^2 + 2|\boldsymbol{u}(i)\boldsymbol{u}_o(i)|\cos(\angle\boldsymbol{\theta}_\Delta(i) + \angle\boldsymbol{\theta}(k)).
\end{aligned} \tag{4.32}$$

Differentiate (4.32) with respect to $\angle\boldsymbol{\theta}(k)$ yields

$$\begin{aligned}
\frac{\partial\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\|^2}{\partial\angle\boldsymbol{\theta}(k)} =& -2\sum_{i=1}^{M} |\boldsymbol{u}(i)\boldsymbol{u}_o(i)|\sin(\angle\boldsymbol{\theta}_\Delta(i) + \angle\boldsymbol{\theta}(k)) \\
=& -2\sin\angle\boldsymbol{\theta}(k)\sum_{i=1}^{M} |\boldsymbol{u}(i)\boldsymbol{u}_o(i)|\cos\angle\boldsymbol{\theta}_\Delta(i) \\
&-2\cos\angle\boldsymbol{\theta}(k)\sum_{i=1}^{M} |\boldsymbol{u}(i)\boldsymbol{u}_o(i)|\sin\angle\boldsymbol{\theta}_\Delta(i).
\end{aligned} \tag{4.33}$$

To find optimized $\angle\boldsymbol{\theta}(k)$ such that (4.32) is minimized, we derive the root of (4.33). The solution would be

$$\angle\boldsymbol{\theta}(k) = \arctan\left(\frac{-\sum_{i=1}^{M} |\boldsymbol{u}(i)\boldsymbol{u}_o(i)|\sin\angle\boldsymbol{\theta}_\Delta(i)}{\sum_{i=1}^{M} |\boldsymbol{u}(i)\boldsymbol{u}_o(i)|\cos\angle\boldsymbol{\theta}_\Delta(i)}\right) \tag{4.34}$$

Note due to the nature of trigonometric functions we would find two solutions (roots) in the range $0 \le \angle\boldsymbol{\theta}(k) \le 2\pi$, and the one corresponding to smaller $\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\|$ is the desired $\boldsymbol{\theta}(k)$ optimizer. As described earlier, in each iteration we pick one $\boldsymbol{\theta}(k)$ as varying variable and repeat (4.32)-(4.34) closed-forms to optimize $\|\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\theta}\|$.

4.2.4 Algorithms Summary and Complexity Comparison

Similar to what we presented in Sec. 4.1.4, now we summarize the procedures and compare the complexity of two algorithms for destination-noise dominating model. First, the algorithm based on relay selection in Sec. 4.2.1 could be described as follows:

1. Set $\binom{L}{M}$ combinations to represent all possible selection of choosing M relays out of L.

2. For the ith combination ($1 \leq i \leq \binom{L}{M}$), we define $\boldsymbol{F}_i$ and $\boldsymbol{G}_i$ as the submatrix of $\boldsymbol{F}$ and $\boldsymbol{G}$ respectively and is corresponding to the selected M relays for the ith selection. Also we denote $\boldsymbol{r}_i$ as the corresponding M relay gains.

3. For the ith combination, we compute the individual relay gains $\boldsymbol{r}_i$ following (4.16) then calculate the capacity with (4.15).

4. Repeat step 3 for all the combinations then selected the one with largest capacity and set relay gains accordingly.

The computationally demanding step 3 requires twice matrix determinant computation for $M \times M$ matrices, and would be repeated for $\binom{L}{M}$ combinations. Thus the computation burden is of order $\mathcal{O}(L^M M^3)$.

For the algorithm discussed in Sec. 4.2.2 and 4.2.3, the procedures could be summarized as follows:

1. Define $\boldsymbol{H}_R$, $\boldsymbol{H}_N$, $\boldsymbol{S}$, $\boldsymbol{V}_N$, $\boldsymbol{V}_R$, $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$ as denoted in Sec. 4.2.2.

2. Define $\boldsymbol{\theta}$ by randomly choosing all the terms from $[0, 2\pi]$

3. Set iteration index $i = 1$

4. Set index $k =$ mod $(i, M) + 1$, where mod (i, M) means the value of i modulo M.

5. With index k, define $\boldsymbol{u}$, $\boldsymbol{u}_o$ and $\boldsymbol{\theta}_\Delta$ based on (4.29), (4.30) and (4.31).

6. Calculate $\boldsymbol{\theta}(k)$ based on (4.34).

7. Increase i by one and go back to step 4, or stop if certain criteria are satisfied.

8. Using the final result of $\boldsymbol{\theta}$, we set $\boldsymbol{r}$ based on (4.27).

The iterative operations between step 4 and 7 do not require complicate computation, and typically would converge within 10 iterations. Thus we may ignore this part for complexity assessment. In step 8 there are multiple matrix multiplication and inversion. Note $\boldsymbol{A}$ is unitary, and $\boldsymbol{S}, \boldsymbol{\Sigma}$ are diagonal matrices. Thus the corresponding matrix computation is simple. The maximal matrix size for the remaining matrix multiplication in step 8 is L, so the overall complexity of step 8 is of order $\mathcal{O}(L^3)$. In step 1 we need SVD computation for a $M \times (L - \frac{M(M-1)}{2})$ matrix , which require [42, p. 234]

$$\{L - \frac{M(M-1)}{2}\}^2(2M-1) + 6\{L - \frac{M(M-1)}{2}\}^3 \tag{4.35}$$

flops and is of order $\mathcal{O}(L^3)$. Thus we could conclude that the order of overall computation is of $\mathcal{O}(L^3)$.

4.3 Numerical Results

To evaluate the performance of algorithms described in Sec. 4.1 and Sec. 4.2, we simulate random channels and perform capacity optimization by the two selection-based methods and proposed efficient designs. Selection-based method for relay-noise dominating and destination-noise dominating are marked as 'relay dom selec' and 'desti dom selec', respectively. The approach maximizing ratio of norms in Sec. 4.1.3 is marked as 'relay dom MRN', while the algorithm based on lower-triangular matrix zero forcing in Sec. 4.2.2 is marked as 'desti dom ZF'. For all the simulations we set $M = 3$. The resulting capacity measures are used to generate cumulative distribution function (CDF) curves so that we could compare the capacity distribution of various system configurations and algorithms. Each curve shows the distribution of 10^3 channel realizations. For benchmarking purpose we also simulate a simple relay gains design which set all gains with identical value and scale relay the transmission power according to (2.5).

Since handling relay network with larger size is one of our motivations to develop efficient designs, two relay sizes are simulated so that we could examine if any worth noting difference shown between the sizes, where solid and dash lines are for $L = 9$ and $L = 18$, respectively.

In Fig. 4.3 we simulate the relay system with relay-noise dominating condition by setting $P_R = 100, \sigma_R = 0.1$ and $\sigma_D = 0.01$. Not surprisingly the algorithms developed for relay-noise dominating, as described in Sec. 4.1, performs better than those for destination-noise dominating. For $L = 18$ the algorithm maximizing ratio of norms shows slight performance degradation compared to selection-based method, but would save considerable computation. Hence the proposed algorithm would be beneficial for large relay systems given noise modeling fit presumed condition.

Next we consider relaying under destination-noise dominating condition and set $P_R = 1, \sigma_R = 0.01$ and $\sigma_D = 0.1$. Again algorithms with mismatched model perform worse than those with correct noise model. Note with larger relay size , the approach based on lower-triangle matrix zero forcing shows better results than selection-based method, which suggests that for large relay system the proposed algorithm not only works efficiently but also demonstrates powerful performance by collaborating the whole relay network, while selection-based method could only utilize a small portion of relays thus results in inferior capacity.

Finally in Fig. 4.5, a particular noise condition, set by $P_R = 10, \sigma_R = 0.1$ and $\sigma_D = 0.01$, is chosen to examine if the algorithms presented in Sec. 4.1 and Sec. 4.2 still work properly when noise model is neither relay-noise or destination-noise dominant. Note that noise dominant condition is affected not only by ratio of σ_R and σ_D, but also P_R and L. Thus for generic noise condition (none noise dominates) it would not be easy to fully investigate and conclude the superiority between algorithms. But for this particular noise condition we could observe that in general the proposed algorithms results in better

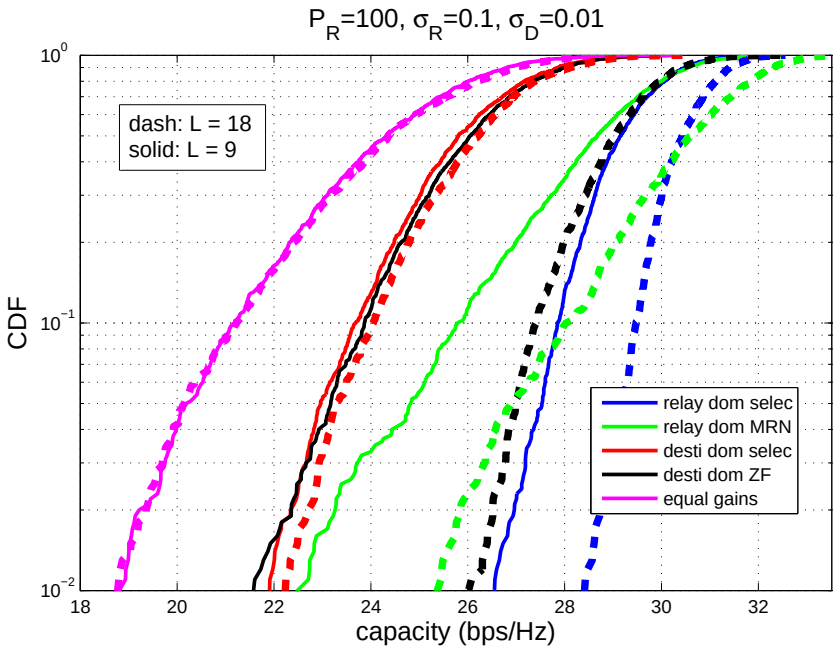

Fig. 4.3: Simulate under relay-noise dominating condition.

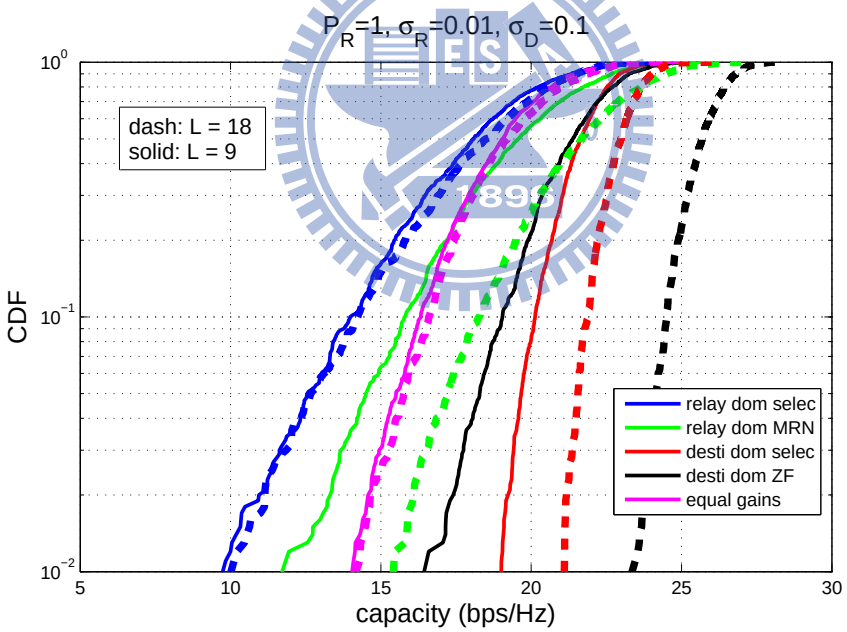

Fig. 4.4: Simulate under destination-noise dominating condition.

system capacity than equal gains design. It is shown in this simulation that the proposed algorithms performs robustly against model mismatch.

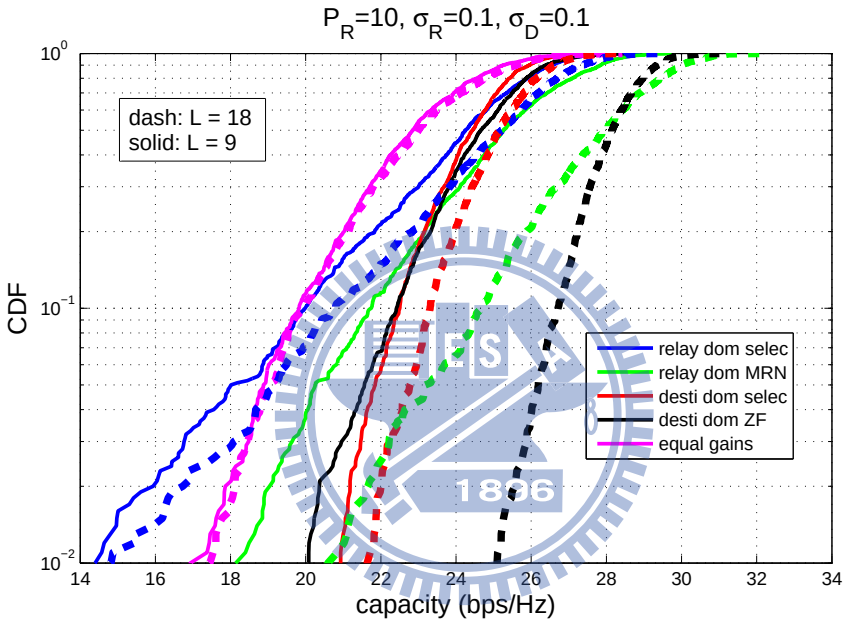

Fig. 4.5: Simulation with two noise terms having equal power level

Chapter 5

Channel Estimation for Distributed Relay Networks with OFDM Transmission

5.1 Matching Pursuit Algorithms for OFDM Channel Estimation

5.1.1 OFDM Transmission System

Assume that the coherence time of the fading wireless channel is much larger than the OFDM symbol duration. Let the size of the discrete Fourier transform (DFT) used in OFDM transmission be N and let the OFDM symbol duration be T. The transmission mechanism associated with each OFDM symbol can be described in terms of matrix-vector notations as

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X}\boldsymbol{W}\boldsymbol{h} + \boldsymbol{n} \\ &= \boldsymbol{X}\boldsymbol{g} + \boldsymbol{n}, \end{aligned} \tag{5.1}$$

where $\boldsymbol{X} = \mathrm{diag}(x(0), x(1), \ldots, x(N-1))$ is the diagonal matrix composed of the transmitted data, W is the Fourier transform matrix, $\boldsymbol{h}$ is the channel impulse response vector and $\boldsymbol{g}$ is the N-vector of the corresponding frequency response vector, $\boldsymbol{n}$ is the N-vector of additive noise samples (assumed white Gaussian), and $\boldsymbol{y}$ is the N-vector of received signal in the frequency domain (i.e., after DFT). The structures of $\boldsymbol{h}$ and $\boldsymbol{W}$ are as follows.

Let the multipath channel have L paths with delays given by τ_l, $l = 0, \ldots, L-1$, where $0 \le \tau_l \le \tau_{\max}$ for some maximum possible path delay $\tau_{\max}$ and each τ_l may be nonsample-spaced (that is, it need not be an integer multiple of the OFDM sample spacing). Vector $\boldsymbol{h}$ has L elements, which are the complex gains of the multipaths. Matrix W has dimension

$N \times L$, with its lth column given by $1/\sqrt{N}[1, e^{-j2\pi\tau_l/T}, \ldots, e^{-j2\pi(N-1)\tau_l/T}]^T$ where $(\cdot)^T$ denotes matrix transpose. Note that the lth column of W is parametrized by the path delay τ_l. The range space of W, or that of any matrix structured similarly to W, has been called a *delay subspace* [39].

Assume there are D pilot subcarriers in each OFDM symbol and assume $D \geq L$. Let S be the $D \times N$ *selection matrix* that selects the pilot locations of an N-vector. For example, $\underline{\boldsymbol{y}} \triangleq S\boldsymbol{y}$ is the vector of received pilots and $\underline{\boldsymbol{g}} \triangleq S\boldsymbol{g}$ is the vector of channel frequency response at the pilot locations. Then for the pilot locations, we have

$$\begin{aligned} \underline{\boldsymbol{y}} &= \underline{\boldsymbol{X}\boldsymbol{W}}\boldsymbol{h} + \underline{\boldsymbol{n}} \\ &= \underline{\boldsymbol{X}\boldsymbol{g}} + \underline{\boldsymbol{n}} \end{aligned} \tag{5.2}$$

where

$$\underline{\boldsymbol{X}} \triangleq \boldsymbol{S}\boldsymbol{X}\boldsymbol{S}^T, \quad \underline{\boldsymbol{W}} \triangleq \underline{\boldsymbol{S}\boldsymbol{W}}. \tag{5.3}$$

5.1.2 Time-Domain Approach to Channel Estimation

Given the pilot data $\underline{\boldsymbol{X}}$ and the received pilot vector $\boldsymbol{y}$, one way of time-domain channel estimation is to first derive the least-square (LS) estimates of $\underline{\boldsymbol{g}}$ and $\boldsymbol{h}$, which are given by [46]

$$\hat{\underline{\boldsymbol{g}}} = \underline{\boldsymbol{X}}^{-1}\underline{\boldsymbol{y}} \tag{5.4}$$

and

$$\hat{\boldsymbol{h}} = (\underline{\boldsymbol{W}}^H\underline{\boldsymbol{W}})^{-1}\underline{\boldsymbol{W}}^H\hat{\underline{\boldsymbol{g}}} \triangleq \underline{\boldsymbol{W}}^\dagger\hat{\underline{\boldsymbol{g}}}, \tag{5.5}$$

respectively, where superscript H denotes Hermitian transpose. Then the estimated channel frequency response is given by

$$\begin{aligned} \hat{\boldsymbol{g}} &= \boldsymbol{W}\hat{\boldsymbol{h}} \\ &= \boldsymbol{W}\underline{\boldsymbol{W}}^\dagger\underline{\boldsymbol{X}}^{-1}\underline{\boldsymbol{y}}. \end{aligned} \tag{5.6}$$

With know pilot locations and pilot values, in order to complete the computation described in the right-hand side of (5.6), the only information that need to be estimated is the delay subspace, or equivalently, the set of path delays $\{\tau_l\}$. To this subject we now turn in the next section.

5.1.3 Estimation of Multipath Delays

Consider a group of L_g successive OFDM symbols and let them be indexed $j = 0, \ldots, L_g - 1$. Assume that, within the time span of these L_g symbols (i.e., L_gT), the complex multipath gains may vary due to fading, but the path delays remain the same. This assumption is appropriate because the path delays usually change much more slowly than

the path gains [39]. In our earlier notations, W stays constant over this period but $\boldsymbol{h}$ may change. For convenience, we attach an index to $\boldsymbol{h}$ and let $\boldsymbol{h}^j$ denote the channel response in the jth OFDM symbol period in the group. Likewise, we also use superscript j to index other quantities that may change with symbols, such as $\underline{\hat{\boldsymbol{g}}}^j$ and $\boldsymbol{S}^j$.

Let there be Q candidate delay values between 0 and $\tau_{\max}$ from which we will identify L for the delay subspace. One reasonable choice of these Q values is $\tau_{\max}k/Q$, $k = 0, \ldots, Q-1$. We can define an $N \times Q$ *dictionary matrix* as $\boldsymbol{V} = [\boldsymbol{v}_0, \ldots, \boldsymbol{v}_{Q-1}]$ where its kth column is given by

$$\boldsymbol{v}_k = [1, e^{-j2\pi\tau_{\max}\frac{kT}{Q}}, \ldots, e^{-j2\pi(N-1)\tau_{\max}\frac{kT}{Q}}]^T. \tag{5.7}$$

Define $\underline{\boldsymbol{V}}^j \triangleq \boldsymbol{S}^j\boldsymbol{V}$.

Estimation Based on the MUSIC Algorithm

The MUltiple SIgnal Classification algorithm (MUSIC) [36] algorithm has been widely used in array signal processing for direction of signal arrival (DOA) estimation. With the assumption of uncorrelated sources, MUSIC algorithm generally is capable of high resolution identification. We mentioned that the MUSIC has been proposed for use in multipath delay estimation for OFDM transmission, with the assumption that the pilot locations be fixed and equal-spaced [32]. Below we outline the algorithm without giving all the details. It is written in a form applicable to the case with fixed but not necessarily equal-spaced pilots.

The fundamental idea of the MUSIC technique is to first find the null (noise) subspace based on the received signal and then project all candidate basis vectors of the delay subspace (i.e., columns of $\underline{\boldsymbol{V}}^j$ or $\boldsymbol{V}$) into the null subspace. Since the actual basis vectors of the delay subspace (which correspond to signal) do not lie in the null subspace, the reciprocals of the projections should show peak at these basis vectors. From this we can identify the delay subspace. Procedure-wise, the steps are as follows:

1. For each OFDM symbol group, collect the L_g estimated channel frequency response vectors $\underline{\hat{\boldsymbol{g}}}^j$ for pilot locations. Solve for the projection matrix $\boldsymbol{P}_N$ of the noise subspace with rank $D - L$.

2. Project all the columns in $\underline{\boldsymbol{V}}$ with $\boldsymbol{P}_N$. Find the L columns with the smallest projection magnitudes. These L columns define the desired delays.

3. Follow the procedure in Sec. 5.1.2 to complete the channel estimation.

Note in the second step we omit the index j for $\underline{\boldsymbol{V}}$ because $\underline{\boldsymbol{V}}^j$ is identical for all j. Indeed, having fixed pilot locations is a requirement of the MUSIC technique. Besides the limitation of fixed pilot locations, a property of the MUSIC technique is that, if some

path coefficients do not change significantly over the L_g OFDM symbols, then there may be a rank-deficiency problem. The result is that these paths may not be identified and resolved properly. This property appears quite undesirable, because it seems to imply the unpalatable conclusion that, in order to achieve good multipath delay estimation, we should make the OFDM symbol period a significant fraction of the channel coherence time. In the area of direction-of-arrivals estimation, this effect has been known as the problem of correlated signal sources, and it may heavily degrade the estimation performance even in high SNR [38]. A technique called spatial smoothing [32], [38] can solve the problem, but the remedy itself also requires equal-spaced pilots. Moreover, it would divide the pilots into several groups, which is an unaffordable solution when the pilots are very few.

Estimation Based on Orthogonal Matching Pursuit Employing One Single OFDM Symbol

In preparation for the description of the proposed GMP technique, we describe how conventional orthogonal matching pursuit (OMP) can be applied to multipath delay estimation with a single OFDM symbol [45].

Ideally, to choose L delays out of Q candidate values, we should try all $\binom{Q}{L}$ possible combinations. For each combination, $\hat{\underline{\boldsymbol{g}}}$ may be projected into the corresponding delay subspace. We then choose the combination with the largest projection magnitude as the estimation result. But the above exhaustive search approach is obviously impractical even with a moderate number of candidate delays Q. One suboptimal but much more efficient approach is the OMP technique [1], which employs a kind of greedy search method to determine the chosen candidates in a sequential fashion.

In applying OMP to multipath delay estimation for OFDM, we determine one path delay at a time. At each iteration, say iteration p, let $\boldsymbol{U}_p$ be the matrix containing the columns from $\underline{\boldsymbol{V}}$ that define the (partial) delay subspace found so far. We project $\hat{\underline{\boldsymbol{g}}}$ to the subspace and find the residual. Then from all columns of $\underline{\boldsymbol{V}}$ that have not entered or covered by $\boldsymbol{U}_p$, we choose the one that has the maximum inner product with the residual and add it to $\boldsymbol{U}_p$. At this, we go to the next iteration until the required number of paths is found. The concept of above-mentioned orthogonal matching pursuit is shown in Fig. 5.1.

Mathematically, let d_p be the index of the column from $\underline{\boldsymbol{V}}$ that is chosen in iteration p. Let $\boldsymbol{k}_p$ denote this vector, that is, $\boldsymbol{k}_p = \underline{\boldsymbol{v}}_{d_p}$. Let P_{U_p} denote the matrix that, when premultiplied to a vector, projects the vector onto the range space of $\boldsymbol{U}_p$. And let $\boldsymbol{r}_p$ be the residual after the pth iteration. Then the OMP algorithm, in iteration p, works as follows:

$$d_p = \arg\max_i |\boldsymbol{r}_{p-1}^H \underline{\boldsymbol{v}}_i|, \quad 0 \le i \le Q-1, \tag{5.8}$$

$$\boldsymbol{k}_p = \underline{\boldsymbol{v}}_{d_p}, \quad \boldsymbol{U}_p = [\boldsymbol{U}_{p-1}, \boldsymbol{k}_p], \tag{5.9}$$

$$\boldsymbol{r}_p = (\boldsymbol{I} - \boldsymbol{P}_{U_p})\hat{\underline{\boldsymbol{g}}}, \tag{5.10}$$

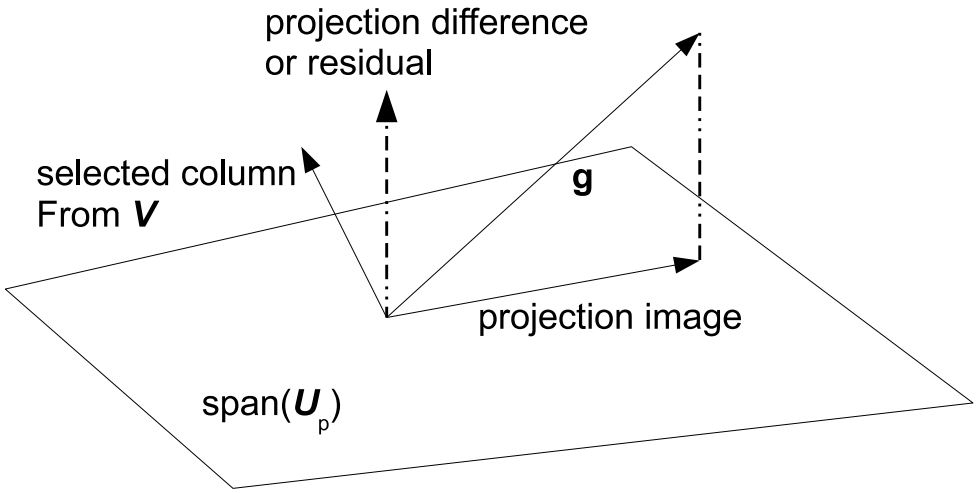

Fig. 5.1: In the pth iteration,projection and selection based on projection residual

where $\boldsymbol{r}_0 \triangleq \hat{\underline{\boldsymbol{g}}}$, $\boldsymbol{U}_0 = \varnothing$, and $\boldsymbol{U}_{L-1}$ gives the desired estimate of $\underline{\boldsymbol{W}}$ (See Sec.5.1.2).

The Group Matching Pursuit Algorithm for Multipath Delay Estimation

Now we turn to the proposed GMP algorithm for estimation of multipath delays based on observation of one OFDM symbol group. While the multipath delays and the delay subspace characterized by W are (assumed to be) fixed within one OFDM symbol group, the changing pilot locations result in different $\underline{\boldsymbol{W}}^j$ and different $\underline{\boldsymbol{V}}^j$. One approach, based on OMP, to address this condition is to perform L_g OMP operations, one for each OFDM symbol, and combine the results. But how the results can be combined poses a problem, because the estimated delays may be different for different OFDM symbols.

The idea of GMP is to make use of the whole set of $\hat{\underline{\boldsymbol{g}}}^j$, $j = 0, \ldots, L_g - 1$, and obtain a jointly optimal delay estimation in some sense. This results in the following steps for iteration p of the algorithm:

$$d_p = \arg\max_i \sum_{j=0}^{L_g-1} |(\boldsymbol{r}_{p-1}^j)^H \underline{\boldsymbol{v}}_i^j|, \quad 0 \le i \le Q-1, \tag{5.11}$$

$$\boldsymbol{k}_p^j = \underline{\boldsymbol{v}}_{d_p}^j, \quad \boldsymbol{U}_p^j = [\boldsymbol{U}_{p-1}^j, \boldsymbol{k}_p^j], \quad 0 \le j \le L_g - 1, \tag{5.12}$$

$$\boldsymbol{r}_p^j = (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{U}_p^j})\hat{\underline{\boldsymbol{g}}}, \quad 0 \le j \le L_g - 1. \tag{5.13}$$

(See the beginning paragraph of Sec. 5.1.3 for the meaning of the superscript j.) As in OMP, $\boldsymbol{U}_p^j$, $j = 0, \ldots, L_g - 1$, give the the desired estimates of $\underline{\boldsymbol{W}}^j$, $j = 0, \ldots, L_g - 1$, that define the delay subspace and can be used as described in Sec. 5.1.2 to obtain a channel estimate for each OFDM symbol in the group. Note that the channel estimates may vary

for different symbols, because the channel is subject to fading, but the delay subspace is the same.

On the Number of Path Delays to Estimate

Throughout the chapter, we have assumed that the number of path delays to be estimated is known. This information can be obtained through other means of channel analysis [46] or empirical data. Even if the number of estimated delays is different from the actual number, in many cases it should not be critical. For example, if we have estimated less path delays than the actual but have captured the most significant paths, then the loss may be acceptable. Conversely, if we have estimated several more path delays than the actual, the resulting enhancement in noise may have little implication as long as its correlation with the actual delay subspace is small [45]. In any case, the number of multipath delays that can be estimated with the proposed technique is upper bounded by D, for otherwise we would have an under-determined set of equations for $\hat{\boldsymbol{h}}$ (see, e.g., (5.5)).

5.2 Relay System Model with OFDM Transmission

Now we discuss how the end-to-end distributed relay forwarding could be modeled as conventional OFDM transmission. We are interested in amplify-forwarding distributed relay network with OFDM modulation throughout end-to-end transmission. Assume relay network is composed of L single-antenna relays. Both source and destination terminal are equipped with single-antenna. Note the discussion in this chapter could be intuitively extended to the cases of multi-antenna source or destination. Again we apply two-phase relay transmission as assumed in Chap. 2. In the first slot, source transmits one OFDM symbol. Each relay performs respective time synchronization to determine correct symbol boundary, then remove cyclic prefix (CP) and acquire a complete OFDM symbol with N subcarriers. To realize amplify-forwarding, the ith ($1 \leq i \leq L$) relay amplify received OFDM symbol by r_i, then add a new CP and transmit. Note that the procedure of removing CP then add another seems redundant. Actually, he received CP samples contain IBI (inter-block interference) signals. If these CP samples remain untouched and are forwarded to destination, the end-to-end transmission would supper two stages of IBI corruption. Thus removing the received CP samples would equivalently remove the IBI from source and therefore would reduce the overall IBI.

We assume a global synchronization for relays transmission is available so that relays could transmit at the same time. Upon receiving the forwarded OFDM symbol, destination works as a conventional OFDM receiver to process received OFDM symbols. The relays reception and forwarding process is depicted in Fig. 5.2.

To model the end-to-end OFDM transmission, we start by focusing on single relay.

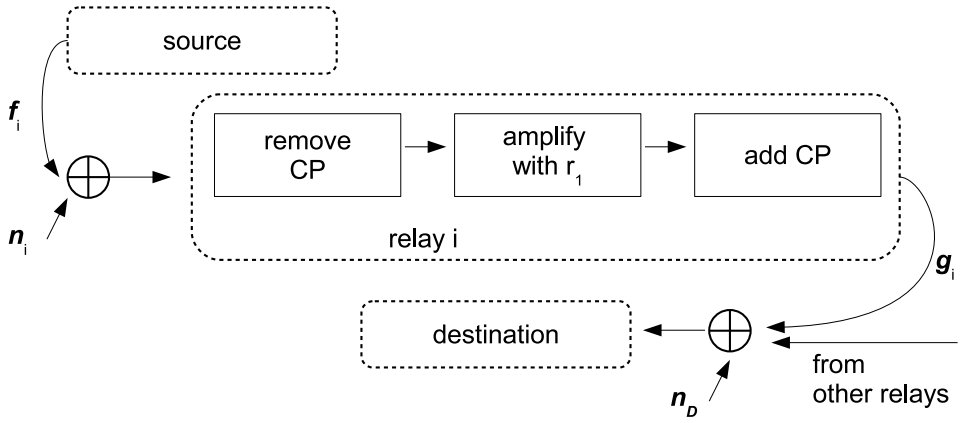

Fig. 5.2: Relay operations for OFDM transmission

For the ith relay, its frequency-domain reception could be stated as

$$\begin{aligned} \boldsymbol{z}_i &= \boldsymbol{X}\boldsymbol{W}\widetilde{\boldsymbol{f}}_i + \boldsymbol{n}_i = \boldsymbol{X}\boldsymbol{f}_i + \boldsymbol{n}_i \\ &= \boldsymbol{F}_i\boldsymbol{x} + \boldsymbol{n}_i \end{aligned} \tag{5.14}$$

where $\boldsymbol{z}_i \in \mathbb{Z}^N$, $\boldsymbol{X}$ denotes a diagonal matrix whose diagonal terms are equal to $\boldsymbol{x}$ and composed of signals from source, $\boldsymbol{n}_i$ is the additive noise, $\boldsymbol{W}$ the Fourier transform matrix, $\widetilde{\boldsymbol{f}}_i$ and $\boldsymbol{f}_i$ represent the impulse response and frequency response of channel between source and ith relay, respectively. $\boldsymbol{F}_i$ is defined to be a diagonal matrix whose diagonal terms are equal to $\boldsymbol{f}_i$. Note that though relays do not perform DFT or IDFT throughout the signal reception and forwarding, we use $\boldsymbol{z}_i$ as frequency-domain representation of received signal for sake of concise and convenient system modeling.

At destination, amplified $\boldsymbol{z}_i$ travels through another multipath channel and is summed up with OFDM signals from other relays. The frequency-domain reception of destination is described as

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{n}_D + \sum_{i=1}^{L} r_i \,\mathrm{diag}(\boldsymbol{z}_i)\boldsymbol{W}\widetilde{\boldsymbol{g}}_i = \boldsymbol{n}_D + \sum_{i=1}^{L} r_i \,\mathrm{diag}(\boldsymbol{z}_i)\boldsymbol{g}_i \\ &= \boldsymbol{n}_D + \sum_{i=1}^{L} r_i \boldsymbol{G}_i\boldsymbol{F}_i\boldsymbol{x} + \sum_{i=1}^{L} r_i \boldsymbol{G}_i\boldsymbol{n}_i \\ &\triangleq \boldsymbol{n}_D + \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}_R, \end{aligned} \tag{5.15}$$

where $\boldsymbol{n}_R$ is the overall forwarding noise, $\widetilde{\boldsymbol{g}}_i$ and $\boldsymbol{g}_i$ represent the impulse response and frequency response of channel between destination and ith relay, respectively. $\boldsymbol{G}_i$ is defined to be a diagonal matrix whose diagonal terms are equal to $\boldsymbol{g}_i$, and diagonal matrix $\boldsymbol{H} \triangleq \sum_{\forall i} \triangleq \boldsymbol{G}_i\boldsymbol{F}_i$ means frequency response of the equivalent end-to-end channel. Let $\boldsymbol{h}$ be the vector composed of diagonal terms of $\boldsymbol{H}$. Observing the subcarrier-wise product

in the definition of $\boldsymbol{H}$, we may conclude that

$$\boldsymbol{h} = \sum_{i=1}^{L} \boldsymbol{f}_i \odot \boldsymbol{g}_i = \boldsymbol{W} \sum_{i=1}^{L} \widetilde{\boldsymbol{f}}_i \otimes \widetilde{\boldsymbol{g}}_i, \tag{5.16}$$

where $\odot$ denotes element-wise product, $\otimes$ stands for circular convolution. Observing (5.16) we understand that the equivalent channel impulse response of end-to-end channel could be modeled as follows: for each relay, do circular convolution for the two-stage channel $\widetilde{\boldsymbol{f}}_i$ and $\widetilde{\boldsymbol{g}}_i$, then sum up the L copies of convolution results. Based on this model, it is clear that the equivalent channel is of widely dispersive channel impulse response.

5.3 Numerical Results

Let the DFT size in OFDM be 256, with 12 subcarriers assigned for pilots. The pilot locations are randomly determined. Let QPSK be employed for each data subcarrier. Consider transmission over a 4-path channel. The path coefficients vary randomly from one OFDM symbol to another, each following a complex Gaussian distribution. Besides the first path delay $\tau_0 = 0$, other path delays are uniformly distributed in the range $[0, \tau_{\max})$, but stay constant during the OFDM symbol group used in GMP channel estimation. We let $\tau_{\max} = 25$ and $L_g = 10$.

Two MP-based approaches are simulated: OMP and GMP. As mentioned previously, to the best of our knowledge there does not exist prior techniques suitable for subspace-based OFDM channel estimation under arbitrary pilot assignments that may vary from symbol to symbol. Thus we cannot compare with eigen-decomposition based schemes such as that in [46] or [32]. However, we simulate channel estimation methods based on linear interpolation and spline interpolation, for a comparison.

Fig. 5.3 shows the mean-square channel estimation errors of different approaches, and Fig. 5.4 the average symbol error rates for each simulated scheme. In the figures, the labels "GMP+MS" and "OMP+SS" mean "GMP approach for multi-symbol estimation" and "OMP for single-symbol estimation," respectively. While interpolation-based methods suffer from scarcity of pilots and are not able to estimate the shape of channel frequency responses accurately, MP-based methods can use the limited resource (pilots) efficiently and result in clearly superior estimation. The proposed GMP algorithm enjoys the greatest "diversity gain" from multi-symbol processing and thus has the better performance among all.

Fig. 5.5 shows the average symbol error rates when the four path delays are fixed at [0, 3, 6, 9]. The simulation demonstrates even better performance for GMP than that in Fig. 5.4. This is because subspace-based algorithms for OFDM channel estimation has a resolution limitation depending on the pilot ratio [32]. When some paths are close together, as occasionally happened in the simulation resulting in Fig. 5.4, MP-based

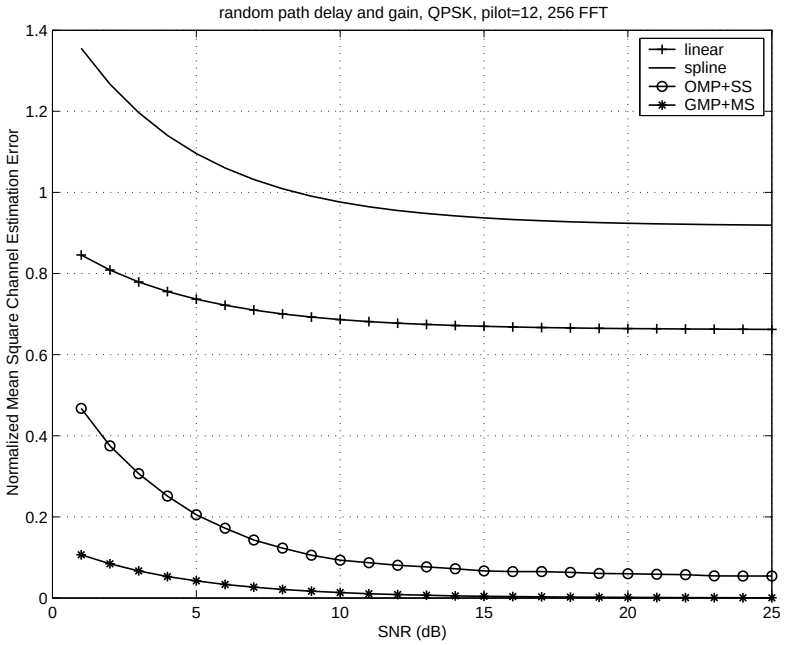

Fig. 5.3: Normalized mean-square channel estimation errors of different channel estimation methods

schemes may have difficulty telling them apart. But this is certainly not the case with the simulation resulting in Fig. 5.5, for the paths are well separated.

In Fig. 5.6 we simulate and examine channel estimations for OFDM distributed relay networks. We set the number of relay terminals as 10. It is assumed there are S_t multipath delay taps randomly spaced between $[0, \tau_{\max}]$. In OFDM transmission with 256-point FFT/IFFT we assign S_p pilot subcarriers and estimate S_e (the column size of $\underline{\boldsymbol{W}}$ in (5.5)) tap gains. In Fig. 5.6 we simulate two sets of configurations. For solid lines we examine the performance with fewer pilots and smaller delay range, thus we set $S_t = 3, \tau_{\max} = 4, S_e = 8$ and $S_p = 10$. For dotted lines we try wider delay range with more pilots by setting $S_t = 4, \tau_{\max} = 8, S_e = 18$ and $S_p = 20$. To compare the performance of various approach we realize channel estimation based on linear interpolation, time-domain least-square method (discussed in Sec. 5.1.2) and OMP-based algorithm. It is clear that OMP-based outperforms the others in both configurations.

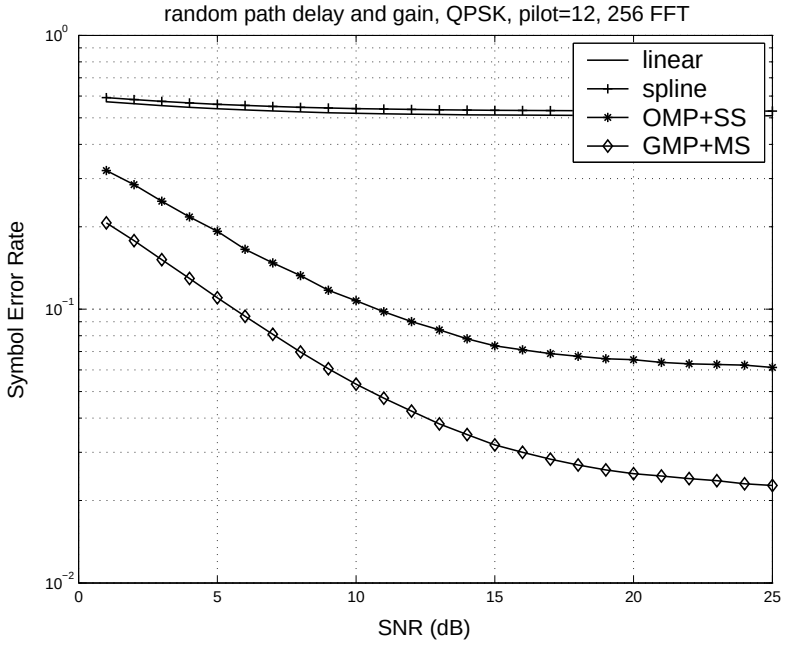

Fig. 5.4: Average symbol error rates (SERs) at data subcarriers with different channel estimation methods.

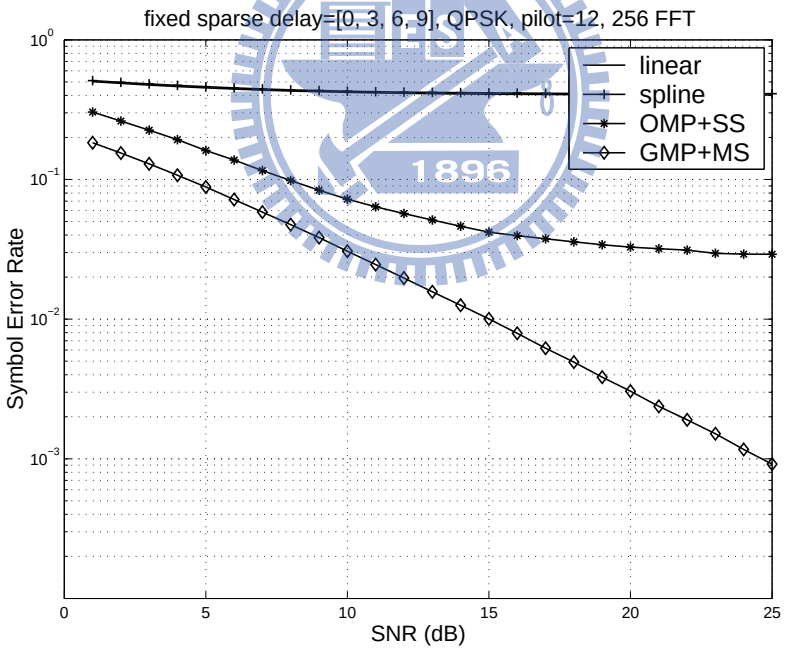

Fig. 5.5: Average SERs at data subcarriers with different channel estimation methods when multipath delay are spaced apart.

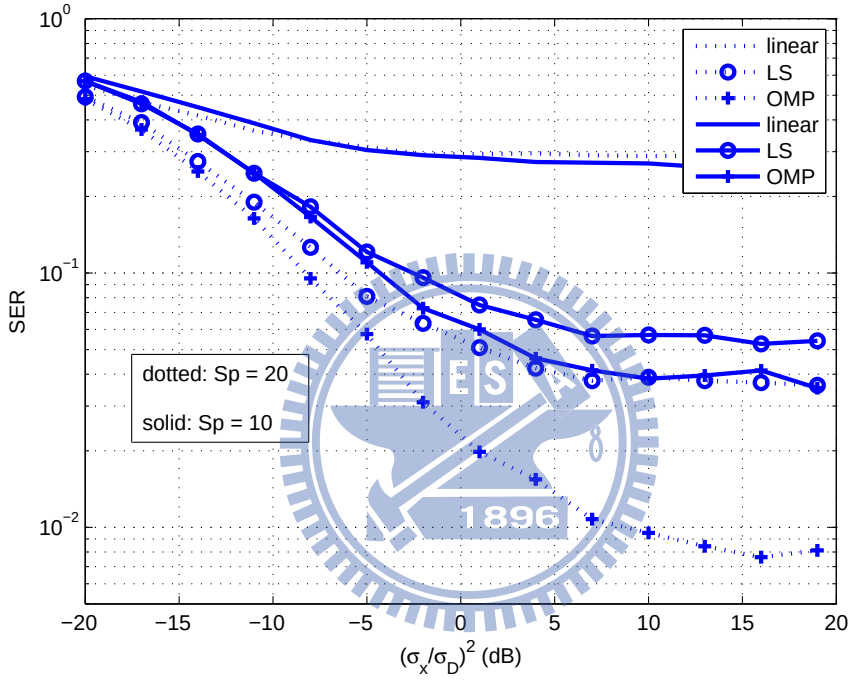

Fig. 5.6: Average SERs of OFDM distributed relay network at data subcarriers with different channel estimation methods when multipath delay are randomly spaced.

Chapter 6

Conclusion

In this thesis we studied the design of distributed AF MIMO relay networks. Specifically, we focused on the capacity improvement and present noise-dominant models to simplify design problem. As to channel estimation we applied proposed subspace-based algorithm to handle equivalent dispersive channel.

In Chap. 2 and Chap. 3, We considered the design of distributed amplify-and-forward relay networks for two-hop MIMO transmission. More specifically, we considered the determination of relay gains for maximization of system capacity. As no closed-form analytical solution could be found for the problem, we considered two alternative approaches. One approach was algorithmic, for which we derived an efficient iterative algorithm. Since the algorithmic solution gave little insight into the analytical properties of the solution, we also took an analytical approach, assuming some asymptotic noise conditions. The analytical approach resulted in several relay selection-type of solutions and facilitated an analysis of the diversity behavior of the solutions. It turned out that their capacity diversity performance behaved similarly to some single-hop point-to-point MIMO antenna selection systems previously analyzed by other researchers. Some simulation results were presented. The results showed that, not surprisingly, the iterative algorithm did yield better designs than the relay selection methods, but at the cost of a substantially higher computational complexity. More significantly, they also confirmed our outage diversity analysis and verified that increasing the number of relays could enhance the outage diversity performance.

In Chap.4 we considered the design of distributed amplify-and-forward relay networks for two-hop MIMO transmission. With the help of noise-dominating models we simplify the originally intractable capacity optimization problem. Further simplification is realized by designing modified criterion of maximal ratio of norms, and shrinking the solution space with zero-forcing. Since no relay selection is required, we could control the computation cost even with large size of relay network. The performance of proposed algorithms is verified by simulations and is proven to be comparable or better than selection-based algorithm.

Time-domain channel estimation techniques can obtain relatively accurate channel estimates for OFDM transmission with relatively few pilot subcarriers. But it requires knowledge of the multipath delays. In Chap. 5 We proposed a group matching pursuit technique for multipath delay estimation. Unlike previous techniques, the proposed technique allows arbitrary pilot structures that may vary from one OFDM symbol to the next. Simulation results showed that the proposed algorithm has superior estimation performance. We also presented the channel modeling of distributed AF relay network with OFDM transmission. It turns out the relay network of interest actually suffer equivalent highly dispersive channel. Thus we examined applying subspace-based algorithm to handle the channel estimation.

Bibliography

[1] J. Adler, B. D. Rao, and K. Kreutz-Delgado, "Comparison of basis selection methods," in *Asilomar Conf. Signals Systems Computers*, 1996, pp. 252–257.

[2] J. Andersen, "Antenna arrays in mobile communications: Gain, diversity, and channel capacity," *IEEE Antennas Propagat. Mag.*, vol. 42, pp. 12–16, 2000.

[3] Y. Bae and J. Lee, "Power allocation for MIMO systems with multiple non-regenerative single-antenna relays," in *Proc. IEEE Veh. Technol. Conf. (VTC-Spring)*, 2010.

[4] J. Bezdek and R. Hathaway, "Some notes on alternating optimization," in *Advances Soft Computing*, 2002, pp. 187–195.

[5] H. Bölcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, 2006.

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

[7] M. X. Chang and Y. T. Su, "2D regression channel estimation for equalizing OFDM signals," in *Proc. IEEE Veh. Technol. Conf.*, 2000.

[8] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Distributed peer-to-peer beamforming for multiuser relay networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, 2009.

[9] ——, "Filter-and-forward distributed beamforming for relay networks in frequency selective fading channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, 2009.

[10] E. Chong and S. Żak, *An introduction to optimization*. Wiley-interscience, 2008.

[11] L. Dong, Z. Han, A. Petropulu, and H. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, 2010.

[12] M. Elkashlan, P. L. Yeoh, R. H. Y. Louie, and I. B. Collings, "On the exact and asymptotic SER of receive diversity with multiple amplify-and-forward relays," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4602–4608, 2010.

[13] P. Fertl, A. Hottinen, and G. Matz, "A multiplicative weight perturbation scheme for distributed beamforming in wireless relay networks with 1-bit feedback," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, 2009.

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

[15] Y. Fu, L. Yang, and W. P. Zhu, "A nearly optimal amplify-and-forward relaying scheme for two-hop MIMO multi-relay networks," *IEEE Commun. Lett.*, vol. 14, no. 3, pp. 229–231, 2010.

[16] A. Gorokhov, D. Gore, and A. Paulraj, "Performance bounds for antenna selection in MIMO systems," in *Conf. Rec., IEEE Int. Conf. Commun.*, vol. 5, 2003.

[17] ——, "Receive antenna selection for MIMO flat-fading channels: theory and algorithms," *IEEE Inf. Theory*, vol. 49, no. 10, pp. 2687–2696, Oct. 2003.

[18] R. Hunger, "Floating point operations in matrix-vector calculus," Munich University of Technology, Tech. Rep., 2007.

[19] *IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access systems*, IEEE Std. 802.16-2004, Oct. 2004.

[20] S. Jin, M. R. McKay, C. Zhong, and K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, 2010.

[21] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, 2006.

[22] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, vol. 3, 2007.

[23] S. G. Kang, Y. M. Ha, and E. K. Joo, "A comparative investi-gation on channel estimation algorithms for OFDM in mobile communications," *IEEE Trans. Broadcasting*, vol. 49, no. 2, pp. 142–149, 2008.

[24] R. Krishna, Z. Xiong, and S. Lambotharan, "A cooperative MMSE relay strategy for wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 15, pp. 549–552, 2008.

[25] C. Li, L. Yang, and W. P. Zhu, "Joint power allocation based on link reliability for MIMO systems assisted by relay," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, 2009.

[26] C. Li, L. Yang, and W. Zhu, "Two-way MIMO relay precoder design with channel state information," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3358–3363, 2010.

[27] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[28] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000.

[29] H. Minn and V. K. Bhargava, "An investigation into time-domain approach for OFDM channel estimation," *IEEE Trans. Broadcast.*, vol. 46, no. 4, pp. 240–248, Dec. 2000.

[30] O. Muñoz, J. Vidal, and A. Agustín, "Non-regenerative MIMO relaying with channel state information," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, vol. 3, 2005, pp. 361–364.

[31] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading relay channels: performance limits and space–time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, p. 1099, 2004.

[32] M. Oziewicz, "On application of MUSIC algorithm to time delay estimation in OFDM channels," *IEEE Trans. Broadcast.*, vol. 51, no. 2, pp. 249–255, Jun. 2005.

[33] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 0, pp. 80–89, Sep. 2004.

[34] B. D. Rao, K. Engan, and S. Cotter, "Diversity measure minimization based method for comupting sparse solutions to linear inverse problems with multiple measurement vectors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP),*, 2004.

[35] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, p. 1711–1721, 2007.

[36] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectral Estimation Workshop*, 1979, pp. 234–258.

[37] G. A. F. Seber, *A matrix handbook for statisticians.* Wiley-Interscience, 2007.

[38] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation fo coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, pp. 806–811, 1985.

[39] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 315–324, Jan. 2004.

[40] S. Sun and Y. Jing, "Channel training and estimation in distributed space-time coded relay networks with multiple transmit/receive antennas," in *Proc. IEEE Wireless Commun. Networking Conf.*, 2010.

[41] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[42] L. Trefethen and D. Bau, *Numerical Linear Algebra.* Philadelphia, PA: SIAM, 1997, no. 50.

[43] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP),*, May 2005.

[44] E. Van Der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, no. 1, pp. 120–154, 1971.

[45] C. J. Wu and D. W. Lin, "Sparse channel estimation for OFDM transmission based on representative subspace fitting," in *Proc. IEEE Veh. Technol. Conf.*, May 2005.

[46] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–478, Mar. 2001.

[47] E. Yilmaz and M. O. Sunay, "Effects of Transmit Beamforming on the Capacity of Multi-Hop MIMO Relay Channels," in *Proc. IEEE Personal Indoor Mobile Radio Commun.*, Sep. 2007, pp. 1–5.

[48] E. Zeng, S. Zhu, X. Liao, Z. Zhong, and Z. Feng, "On the performance of amplify-and-forward relay systems with limited feedback beamforming," *IEICE Trans. Commun.*, vol. 91, no. 6, pp. 2053–2057, 2008.