Kasper Fløe Trillingsgaard

Communications Schemes for Gaussian Two-way Relay Networks

Master of Science Thesis

Supervisors:
Petar Popovski
Torben Larsen

Aalborg University
Department of Electronic Systems

June 6th, 2013

Department of Electronic Systems
Fredrik Bajers Vej 7
DK-9220 Aalborg Ø
http://es.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Communication Schemes for Gaussian Two-way Relay Networks

Theme:
Wireless Communication

Project Period:
Spring Semester 2013

Project Group:
Group 1012

Participant(s):
Kasper Fløe Trillingsgaard

Supervisor(s):
Prof. Petar Popovski
Prof. Dr. Techn. Torben Larsen

Copies: 4

Page Numbers: 71

Date of Completion:
June 6th 2013

Abstract:

Relaying is a transmission technique that promises coverage extension and an increased spectral efficiency. The advent of two-way relaying based on network coding opened a large design space in wireless networks. Two related topics are addressed in this thesis: interactive two-way relaying and extension of the wireless network coding principle to larger networks.
The first part attempts to increase the achievable rates of the amplify-and-forward scheme for the traditional two-way relay channel using interaction. Two extended amplify-and-forward schemes are put forth, and the achievable rates are computed. Simulation results, however, show that no improvements are achieved using the proposed schemes. This once more confirms why there are so few interactive communication strategies in information theory that lead to nontrivial improvements.
The second part addresses coordinated relaying, which is a generalization of two-way relaying. In particular, the study focuses on the shared relay, in which two base stations wish to attain two-way communication to two terminals through one relay. The relay schemes, precode-and-forward and decode-and-forward, are extended to this relay network using the concepts of zero-forcing and alignment. The simulation results show that the decode-and-forward scheme performs well at low SNR while precode-and-forward is better at high SNR. Finally, it is illustrated how these concepts can scale to larger networks.

Institut for Elektroniske Systemer
Fredrik Bajers Vej 7
DK-9220 Aalborg Ø
http://es.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Titel:
Kommunikations Strategier for Gaussisk To-vejs Relæ Netværk

Tema:
Trådløs Kommunikation

Projektperiode:
Forårssemestret 2013

Projektgruppe:
Gruppe 1012

Deltager(e):
Kasper Fløe Trillingsgaard

Vejleder(e):
Prof. Petar Popovski
Prof. Dr. Techn. Torben Larsen

Oplagstal: 4

Sidetal: 71

Afleveringsdato:
6. juni 2013

Abstract:

Trådløse relæer forventes at få stor indflydelse i fremtidens trådløse netværk. Denne rapport omhandler to relaterede emner: interaktiv to-vejs kommunikation via et relæ og udvidelse af strategier for netværkskodning i relæ netværk til større netværk.
Første del af rapporten forsøger at øge de opnåelige transmissions rater for et tre-terminals relæ netvæk med to-vejs kommunikation ved brug af interaktive amplify-and-forward strategier. To interaktive metoder er introduceret og de opnåelige transmissions rater bliver beregnet. Ved brug af simulering vises dog at metoderne ikke opnår bedre transmissions rater end den tradionelle amplify-and-forward metode.
Anden del af rapporten omhandler koordinerede relæer, som er en generalisering af det simple tre-terminals relæ netværk. Der tages udgangspunkt i to mere avancerede relæ netværk, og der designes kommunikationsstrategier, som muliggøre effektiv kommunikation ved brug af principperne alignment og zero-forcing. To grundlæggende kommunikationsstrategier, kaldet precode-and-forward og decode-and-forward, udvides til disse mere avancerede relæ netværk. Herefter beregnes de opnåelige transmissions rater og sammenlignes. Simuleringer viser at precode-and-forward giver de bedste resultater ved høj signalstøjforhold, hvorimod decode-and-forward fungerer godt ved lave signalstøjforhold. Den endelige konklusion er at principperne, alignment og zero-forcing, skalere til større netværk.

Preface

This Master's Thesis documents a four months Master's project by 13gr1012 on the ELITE master education within Wireless Communication at Aalborg University in the spring of 2013.

The thesis concerns two aspects of two-way relaying; (1) increasing the achievable rates of the traditional two-way relay channel, and (2) coordinated relaying, which deals with larger relay networks.

Initially, the main motivation behind the project was solely to approach the outer bound for the traditional two-way relay channel, which has been an open problem for many years. The key difference between our and previous attempts, was the idea of using interactive communication strategies. Halfway through the project, without any positive results, we found that the problem was significantly more difficult than first anticipated. In particular, we found that the paper [1] even hypothesized that interaction can not extend the region. This problem is considered in Chapter 2 and 3, where the most developed of our interactive schemes are presented. The direction of the project was then changed towards relaying schemes for coordinated relaying which the remaining three chapters addresses.

The simulation code developed in the project is found on the CD attached to the thesis. A list of acronyms can be found in Appendix A and a list of symbols in Appendix B

Acknowledgments

I would like to thank my thesis supervisor Petar Popovski for his support during the many obstacles I faced throughout this project. Also, I would like to thank Elisabeth De Carvalho and Sabina Holm Skov for their valuable comments.

Kasper Fløe Trillingsgaard

Contents

CHAPTER 1

Introduction

Society is to higher and higher extend based on technology. The development within technology places heavy demands to the abilities of wireless communication; hence innovation and research within the field is necessary. Relaying is a central concept within wireless communication that allows two or more wireless nodes to communicate through another node, called a relay. In recent years, there has been an increasing interest in the concept of relaying in both industry and academia [2]. The applications of relaying ranges from satellite communication, in which two terminals far from each other wish to communicate through a relay, to ad hoc networks and cellular networks. One promising area in which relaying is expected to become beneficial is in cellular infrastructure. As the number of wireless devices increases, higher data rates are demanded and the frequency bands are to be located well above 2 GHz, the base station infrastructure needs to be increasingly more complex [2]. A straightforward but rather naive solution is to increase the density of base stations – at high deployment costs. The concept of relaying offers an alternative solution, in which relays are scattered within each cell area. The deployment cost of relays is significantly lower than of base stations since they are not connected to the wired backhaul network. If the density of the relays is sufficiently high, most terminals are closer to a relay than a base station. In that case, relaying promises range extension of existing base stations [3, 4] and higher spectral efficiency.

The scope of this thesis is limited to two-way half-duplex relaying on the Gaussian channel. Two-way relaying means that two or more terminals wish to exchange messages using a relay. All terminals are furthermore assumed to be half-duplex nodes implying that they can not simultaneously transmit and receive. In Fig. 1.1, three conceptual different approaches in which relaying can be accomplished are illustrated [5]. Suppose that the messages m_1 and m_2 of terminals 1 and 2 and destined to the terminal 2 and 1, respectively, are initially mapped to the codewords $\mathbf{x}_1$ and $\mathbf{x}_2$. The first strategy, also referred to as the *routing strategy*, corresponds to traditional point-to-point communication in which multiple transmissions at once are seen as interference. Using this strategy, the terminals may exchange messages within four rounds. The performance can be improved by the use of network coding [5]. Upon reception of each of the codewords $\mathbf{x}_1$ and $\mathbf{x}_2$, the relay forms a sum of the two messages, denoted by $\mathbf{x}_1 \oplus \mathbf{x}_2$ (e.g. modulo-2 sum for binary channels), which is transmitted to terminal 1 and 2. Each terminal can then cancel their own contribution from the sum to obtain the message from the other terminal. This is illustrated in the second strategy with only three rounds. However, as the relay does not need to decode each of the messages m_1 and m_2, a more efficient approach may be to form the sum of the messages at the physical layer [5] such that the relay obtains a noisy version

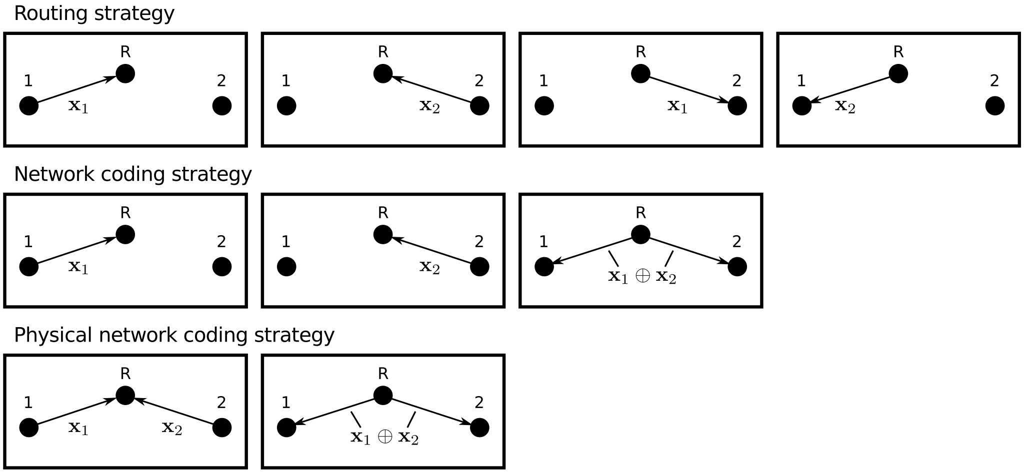


Figure 1.1: *The two-way relaying problem.*

of $\mathbf{x}_1 \oplus \mathbf{x}_2$ directly from the channel. Next, the relay broadcasts a codeword, representing the noisy sum, to the terminals. This leads to an approach in which the messages can be transmitted in only two rounds. This concept is termed *physical network coding* since the messages are combined at the physical layer. In this thesis, the strategies based on network coding and physical network coding are studied. These strategies, in general, consist of a multiple access (MA)-phase, where the terminals 1 and 2 transmit and the relay receives, and a broadcast (BC)-phase, where the relay transmits and the terminals receive.

The contributions of this thesis are two-fold. In the this first part, an attempt to extend the achievable rate region of the traditional two-way half-duplex relay channel, as depicted in Fig. 1.1, is investigated. The second part addresses coordinated relaying, and considers relaying schemes for more advanced relay networks with multiple antennas.

1.1. Approaching the Capacity of the Two-way Relay Channel

Since Van De Meulen in 1971 introduced the basic three-node two-way half-duplex relay channel with Gaussian noise in [6], there has been a massive amount of research within the field. The aim has been to reach the capacity region of the channel. An outer bound on the capacity region is easily derived as in [5], but no relaying scheme has yet managed to achieve it. During the years, several transmission schemes have been proposed, achieving rates increasingly close to the bound. The most fundamental schemes proposed are; decode-and-forward [7], amplify-and-forward [8, 9], compress-and-forward [10] and denoise-and-forward (also known as compute-and-forward) [5, 7] which all uses network coding or physical network coding.

In this work, a fundamentally different direction is taken. The key idea of the approach is to use the interaction which is created between the three nodes during multiple MA- and BC-phases. Interaction implies that the nodes can obtain feedback from other nodes that can be used to encode future messages. In point-to-point communication, it is well-known that feedback does not increase the capacity, but in multiuser systems and, in particular, two-way communication systems, feedback is known to increase the capacity region for several channels [11].

The organization of this part is as follows; Chapter 2 introduces the preliminaries

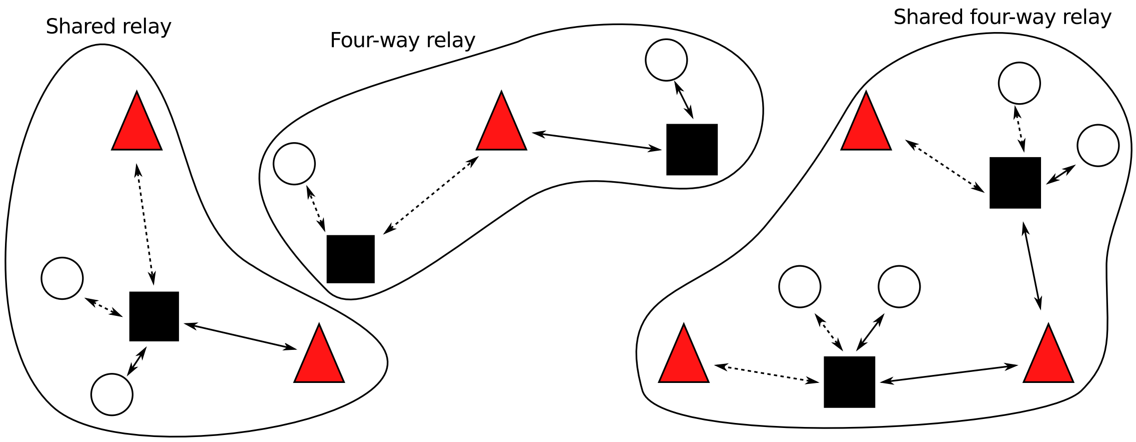


Figure 1.2: *A large relay network in which all terminals exchange messages with their base stations through relays. Red triangles denote base stations, black squares denote relays and circles denote terminals. Three different relay networks are identified; (1) the shared relay, (2) the four-way relay, and (3) the shared four-way relay.*

and describes an achievable scheme for multiple access channel with perfect feedback. This scheme is one of the main motivations of the relaying schemes to be presented. In Chapter 3, the two interactive schemes based on the amplify-and-forward scheme are proposed. The symmetric achievable rates are computed and the performance of the schemes are compared to the amplify-and-forward scheme.

1.2. Coordinated Relaying

Although the two-way relay channel has attracted much attention in the past years, in practical systems, the schemes for basic three-node relay channels are insufficient. Coordinated relaying [12] deals with larger relay networks in which techniques, similar to those for the three-node network, can be used to improve performance. The objective of the base stations is to serve a number of users in the most efficient way. Based on the two-way half-duplex relay channel, it is clear that a base station can exchange messages with all its users, without interference, in a Time Division Multiple Access (TDMA) or Frequency Division Multiple Access (FDMA) fashion. Coordinated relay is a broad term that covers rate-efficient schemes for larger relay networks. The two basic principles used for designing rate-efficient relaying schemes for larger networks are [12]:

1. *Aggregation of communication flows:* Instead of serving each of the terminals separately (through TDMA or FDMA), multiple terminals are served at once.
2. *Intentional cancellable interference:* The communication flows are allowed to interfere, but the relaying schemes are designed in such a way that interference can be cancelled upon reception.

As relay networks grow larger, more advanced sub networks can be identified. In Fig. 1.2, a large network of base stations, relays and terminals is depicted. Note that terminals communicating directly with the base stations are omitted for clarity. The goal of the base stations is to serve all the terminals. The figure identifies three sub networks which are more advanced than the three-node relay channel. The four-way relay is investigated in

[4] and with multiple antennas at the nodes in [3]. The shared relay is in general termed a multi-way relay in [13], and if the relay have multiple antennas, it is called a Multiple-Input Multiple-Output (MIMO) switch [14]. To the best of our knowledge, the shared four-way relay has not been investigated in literature.

One of the big advances in wireless communication is the use of multiple receive and transmit antennas for which the capacity was first found in [15]. MIMO channels in general increase reliability of communication and the achievable communication rates. Essentially, MIMO expands the time and frequency dimensions with a space dimension. For multiuser systems, this implies that a MIMO node can communicate with multiple nodes at once, at the same time and frequency, but in different spatial dimensions. In this thesis, the basic sub networks, the shared relay and shared four-way relay, depicted in Fig. 1.2, are studied with MIMO. In particular, schemes are designed using the following two principles:

1. *Zero-forcing:* Traditionally, if a node is broadcasting multiple messages to multiple nodes simultaneously, the receiving node obtains a superposition of the messages, and hence, successive interference cancellation have to be used to decode the messages [16]. Zero-forcing takes advantage of the fact that if the transmitting nodes have a sufficient number of antennas, it can transmit the messages for each node in different spatial dimensions such that all nodes receives an interference-free signal. This is principle is used for the four-way relay channel in [3].
2. *Alignment:* Suppose that two nodes wish to exchange messages through a relay with many antennas. Then, in general, the signals from the nodes are received in two different spatial subspaces at the relay. When two-way communication is desired, this is suboptimal since interference can be removed at the nodes in the end of the BC-phase. However, if the number of antennas at one of the nodes is increased, then the node can align its transmission such the signals are received in the same subspace at the relay, and hence decrease the number of spatial dimensions used at the relay by the two-way communication flow. This effectively enables the relay to serve more communication flows.

The part is organized as follows; the results of the MIMO channel, multiple access channel, the MIMO switch and the four-way relay channel are summarized in Chapter 4. In Chapter 5, a precode-and-forward [14] and two decode-and-forward relaying schemes are proposed for the shared relay channel based on alignment and zero-forcing. The achievable rates are derived and the performance is discussed. Finally, in Chapter 6, a decode-and-forward scheme is presented for the shared four-way relay and the achievable rates are compared to the shared relay.

Part I

Two-way Relaying with Interactive Communication

CHAPTER 2

State of the Art and Background

This chapter defines the two-way half-duplex relay channel, which is studied in several papers [5, 8, 7, 9, 10, 17]. Based on the system model, the Amplify-and-Forward (AF) scheme is described, which is used by the interactive relaying schemes discussed in the next chapter. The model considered in this chapter is simplified as much as possible in order to facilitate the analysis of the interactive schemes.

2.1. Two-way Relay Channel

The two-way half-duplex relay channel is depicted in Fig 2.1. The system model consists of two terminals, indexed 1 and 2, and one relay node. The time is divided into two phases; a MA-phase and a BC-phase, each of n channel uses. Terminal $j \in \{1, 2\}$ has a message $m_j \in \{1, \ldots, 2^{nR_j}\}$ destined to the other terminal, where R_j denotes the rate of the communication flow from terminal j to terminal $\bar{j}$[1]. The rates, R_1 and R_2, describes the number of bits that are communicated per channel use. In the MA-phase, the terminals simultaneously transmit their signals to the relay. Each terminal maps its message m_j to an n-dimensional vector, $\mathbf{x}_j \in \mathbb{C}^{n\times 1}$, using an encoding function $f_j^{(n)} : \{1, \ldots, 2^{nR_j}\} \to \mathbb{C}^{n\times 1}$, i.e. $\mathbf{x}_j = f_j^{(n)}(m_j)$. The encoding functions have to satisfy the power constraints at the terminals

$$\frac{1}{n}\mathbf{x}_j^{\mathrm{H}}\mathbf{x}_j \leq P_j \qquad \text{for } j \in \{1, 2\}. \tag{2.1}$$

The relay observes a noisy superposition of the signals from the terminals $\mathbf{y}_R \in \mathbb{C}^{n\times 1}$, i.e.

$$\mathbf{y}_R = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}_R, \tag{2.2}$$

and applies a mapping $g^{(n)} : \mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1}$ before broadcasting the signal back to the terminals. Consequently, the relay transmits $\mathbf{x}_R = g^{(n)}(\mathbf{y}_R)$. The relay mapping $g^{(n)}$ has to satisfy the power constraint at the relay

$$\frac{1}{n}\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R \leq P_R. \tag{2.3}$$

Each terminal then receives a noisy version of the signal from the relay $\mathbf{y}_j \in \mathbb{C}^{n\times 1}$, i.e.

$$\mathbf{y}_j = g^{(n)}(\mathbf{y}_R) + \mathbf{z}_j. \tag{2.4}$$

[1] $\bar{j}$ denotes the opposite terminal, i.e. $\bar{j} = 1$ if $j = 0$ and 0 otherwise.

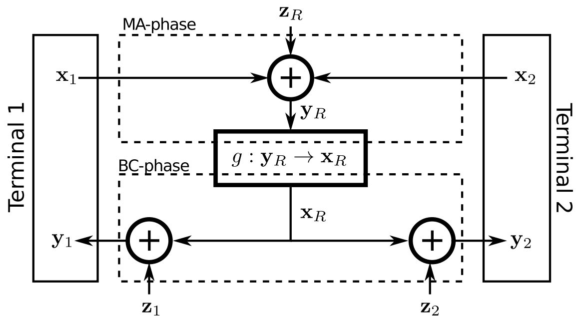


Figure 2.1: *The two-way relay channel.*

The noise signals added at the relay and at the terminals are identical and independently distributed (iid) zero-mean Gaussian variables with unit variance, i.e. $\mathbf{z}_R, \mathbf{z}_1, \mathbf{z}_2 \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_n)$, where $\overset{\text{iid}}{\sim}$ means "iid according to the distribution" and $\mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_n)$ denotes the distribution of zero-mean complex multivariate Gaussian random variables with covariance matrix $\mathbf{I}_n$. Finally, terminal j uses the decoding function $\hat{m}_{\bar{j}}^{(n)} : \mathbb{C}^{n\times 1} \to \{1, \ldots, 2^{nR_j}\}$ to reconstruct the message $m_{\bar{j}}$ from terminal $\bar{j}$.

The rates R_1 and R_2 are said to be *achievable* if and only if, a sequence of encoding functions $f_j^{(n)}$, decoding functions $\hat{m}_j^{(n)}$ and relay mappings $g^{(n)}$ exist, such that, for all $\varepsilon > 0$, there exists $n > 0$ such that

$$\Pr\left(\hat{m}_1^{(n)}(\mathbf{y}_1) \neq m_1 \text{ or } \hat{m}_2^{(n)}(\mathbf{y}_2) \neq m_2\right) < \varepsilon. \tag{2.5}$$

If the rates R_1 and R_2 are achievable, the tuple (R_1, R_2) is called an *achievable rate pair*. A set of achievable rate pairs (R_1, R_2) forms an *achievable rate region*, and if it can be shown that the achievable rate region contains all achievable rate pairs, the achievable rate region is also the *capacity region*, which in general can be shown to be convex using time-sharing arguments [16].

This definition implies that a relaying scheme is achievable if the probability of erronous reconstruction of the messages m_j with $j \in \{1, 2\}$ vanishes as n tends to infinity. It is important to note that finding the capacity region is a difficult task. The next section establishes a simple outer bound on the rate region using the capacity of the Additive White Gaussian Noise (AWGN) channel.

2.2. Outer Bound on the Capacity Region

From the system model, a simple outer bound for the capacity region of the two-way half duplex relay channel can be established. Outer bounds are of central importance in information theory in order to compare the performance of coding or relaying schemes to the optimum [16]. In particular, to show that a certain achievable rate region of a channel is also the capacity region it must the shown that the achievable rate region coincides with an outer bound. For the system model at hand, the rate R_1 can be bounded as

$$R_1 < I(\mathbf{x}_1; \mathbf{y}_2 | m_2) \overset{(a)}{\leq} \min(I(\mathbf{x}_1; \mathbf{y}_R | m_2), I(\mathbf{x}_R; \mathbf{y}_2 | m_2)), \tag{2.6}$$

where $I(\cdot;\cdot)$ and $I(\cdot;\cdot|\cdot)$ denotes the mutual information and the conditional mutual information, respectively, and (a) follows from the data processing inequality [16]. Intuitively, this means that terminal 1 can not communicate information at a higher rate than if both terminals and the relay knows the message from terminal 2. Considering the case in which terminal 2 does not transmit, it is clear that $I(\mathbf{x}_1; \mathbf{y}_R|m_2)$ and $I(\mathbf{x}_R; \mathbf{y}_2|m_2)$ are upper bounded by the capacity of the AWGN channel[2], and hence the following rate bounds are obtained

$$R_1 < \min\left(\log_2\left(1 + P_1\right), \log_2\left(1 + P_R\right)\right) \tag{2.7a}$$

$$R_2 < \min\left(\log_2\left(1 + P_2\right), \log_2\left(1 + P_R\right)\right). \tag{2.7b}$$

This outer bound has been extensively studied in literature [5, 7], but has not been achieved. Recent results, however, indicate that structured codes such as lattice codes needs be used to approach the bound, but the best known scheme only achieves the symmetric ($P_1 = P_2 = P$) rates [5]

$$R_1^{\text{lattice}} = R_2^{\text{lattice}} < \min\left(\log_2\left(\frac{1}{2} + P\right), \log_2\left(1 + P_R\right)\right). \tag{2.8}$$

In the next section, the simple relaying scheme amplify-and-forward is introduced. The performance of these schemes is, however, far from the lattice based schemes.

2.3. Amplify-and-Forward

The AF scheme is a simple scheme for two-way half-duplex relaying on the AWGN channel [8, 9]. The main idea is that the two terminals transmit their codewords simultaneously such that the relay receives a noisy sum of the codewords. The relay scales this noisy sum and relays it back to the terminals. Each node then subtracts its own contribution, resulting in an estimate of the other terminals codeword. This yields a very simple relaying scheme that performs well, with a constant gap to the outer bound in (2.7), for high SNR.

In the MA-phase, each terminal maps its message m_j to a codeword $\mathbf{x}_j \in \mathbf{C}^{n \times 1}$. The power of each codeword is limited by a power constraint P_j such that

$$\frac{1}{n}\mathbf{x}_j^{\mathrm{H}}\mathbf{x}_j \leq P_j. \tag{2.9}$$

The codewords are transmitted such that the relay receives a noisy superposition of the codewords

$$\mathbf{y}_R = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}_R. \tag{2.10}$$

where $\mathbf{z}_R \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_n)$. The signal $\mathbf{y}_j$ and $\mathbf{x}_j \in \mathbb{C}^{n \times 1}$ for $i \in \{1, 2, R\}$. At the relay, upon reception, $\mathbf{y}_R$ is scaled by an amplification factor $\gamma = \sqrt{\frac{P_R}{2P+1}}$ such that the power constraint $\gamma^2 \mathbf{y}_R^{\mathrm{H}}\mathbf{y} \leq nP_R$ is satisfied. In the BC-phase, the resulting signal is broadcast to the terminals, i.e. each terminal receive

$$\mathbf{y}_j = \sqrt{\frac{P_R}{2P+1}}\left(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}_R\right) + \mathbf{z}_j, \tag{2.11}$$

[2]The AWGN channel is given by $y = x + z$, where $\mathbb{E}[x] \leq P$ and $z \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$. The channel capacity is then given by $\log_2(1 + \frac{P}{\sigma^2})$ and the ratio $\frac{\mathbf{P}}{\sigma^2}$ is termed the Signal to Noise Ratio (SNR). The capacity is achieved using random codebooks with x distributed as $\mathcal{N}_{\mathbb{C}}(0, P)$ [16].

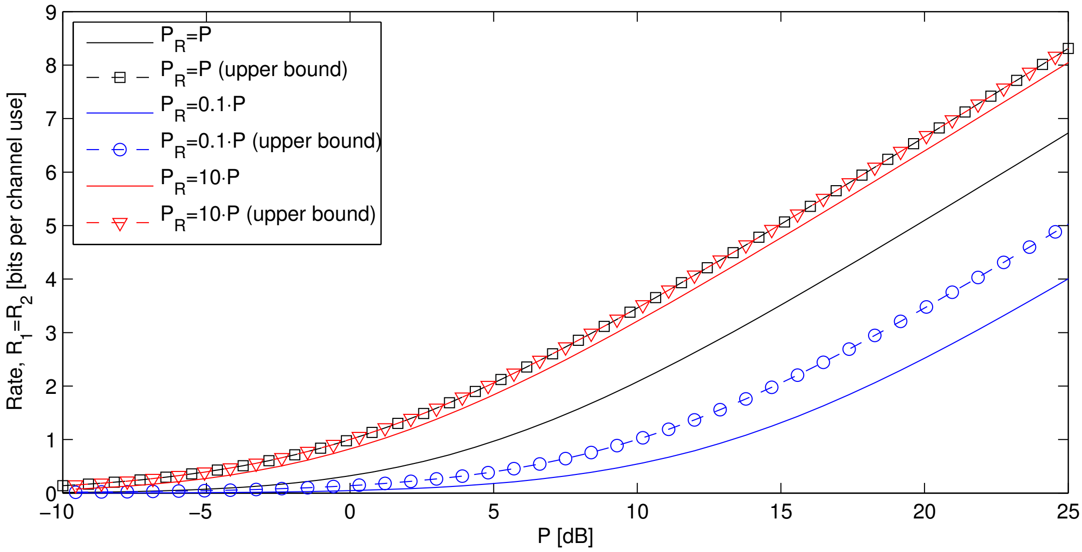


Figure 2.2: *The performance of* AF *compared to the outer bound derived in* (2.7) *with* $P = P_1 = P_2$.

where $\mathbf{z}_j \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_N)$ and $j \in \{1, 2\}$. Terminal j then cancels its own contribution from $\mathbf{y}_{\bar{j}}$. In that way, the AF scheme essentially creates a point-to-point AWGN channel from terminal j to terminal $\bar{j}$ with SNR given by

$$\mathrm{SNR}_j = \frac{\frac{P_R}{2P_j+1} P_j}{\frac{P_R}{2P_j+1} + 1} = \frac{P_R P_j}{P_R + 2P_j + 1}. \tag{2.12}$$

If the terminals use capacity achieving codes for the AWGN channel [16], the terminals can decode the codewords at the rates

$$R_j < \log_2 \left(1 + \frac{P_R P_j}{P_R + 2P_j + 1}\right). \tag{2.13}$$

In Fig. 2.2, the performance of the AF scheme is compared to the upper bound in (2.7). This figure shows that, at high SNR, the gap between the achievable rate and the upper bound is constant, and decreases as P_R becomes larger than P. The main problem with AF is that the noise from the MA-phase is amplified by the amplification factor as the signal.

The aim of the interactive relaying scheme is introduce multiple MA- and BC-phases such that feedback is allowed, and to investigate if this may increase the achievable rates of the AF scheme. The next section describes the Multiple Access Channel with Perfect Feedback (MACPF) that illustrates how feedback can be used to increase the capacity region.

2.4. Multiple Access Channel with Perfect Feedback

In this section, the Multiple Access Channel with Perfect Feedback is introduced. The system model is not directly used later, but the main goal of the section is to illustrate a scheme that does achieve an extended achievable rate region when feedback is facilitated. The system model is depicted in Fig. 2.3. The objective is determine the rates at which

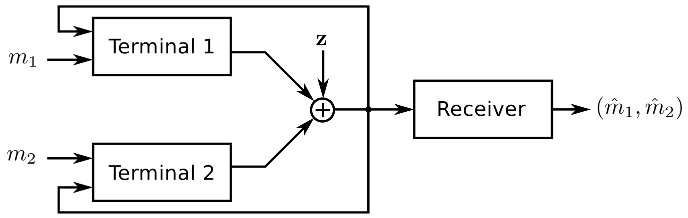


Figure 2.3: *The* MA *channel with perfect feedback [18].*

the two terminals can independently communicate to the receiver. Without feedback, it is widely known that the capacity region can be approached using successive interference cancellation. Using perfect feedback, however, the achievable rate region can be extended using a transmission scheme by Cover and Leung in [18] that exploits the fact that the terminals can transmit correlated information.

The time is divided into n channel uses and the two terminals have messages $m_1 \in \{1, \ldots, 2^{nR_1}\}$ and $m_2 \in \{1, \ldots, 2^{R_2}\}$ that are destined to the receiver. In channel use $i \in \{1, \ldots, n\}$, the terminals transmit signals $x_{1,i} \in \mathbb{R}$ and $x_{2,i} \in \mathbb{R}$, and the receiver obtains a noisy superposition of the signals, i.e. $y_i = x_{1,i} + x_{2,i} + z_i$, where z_i is a zero-mean Gaussian variable with unit variance. Before the next channel use, y_i is feed back to each of the terminals, such that the signals $x_{1,i+1}$ and $x_{2,i+1}$ may depend on $y_i, \ldots, y_1$. The total powers of the signals from the terminals are limited by the power constraints P_1 and P_2,

$$\frac{1}{n} \sum_{i=1}^{n} x_{j,i}^2 \leq P_j \qquad \text{for } j \in \{1, 2\}. \tag{2.14}$$

The objective is to reconstruct the messages m_1 and m_2 at the receiver with a vanishing error probability as the blocklength n tends to infinity.

If the feedback is not used by the terminals, the channel reduces to the traditional MA channel for which the capacity region is known to be (the achievable strategy is described for the MIMO MA channel in Section 4.2) [16]

$$R_j \leq \frac{1}{2} \log_2 (1 + P_j) \qquad \text{for } j \in \{1, 2\} \tag{2.15}$$

$$R_1 + R_2 \leq \frac{1}{2} \log_2 (1 + P_1 + P_2) \,. \tag{2.16}$$

In the following, the achievable region is improved by taking advantage of the feedback using a scheme first derived in [18]. The n channel uses are equally divided into $B+1$ blocks, indexed $0, \ldots, B$, such that each block consists of $n_B = \frac{n}{B+1}$ channel uses. Similarly, the messages m_j are divided into B messages $m_{j,b} \in \{1, \ldots, 2^{n_B R_j}\}$. The main idea is that in block $0 < b < B$, transmitter j uses a fraction α_j of its power for transmission of new information, i.e. the message $m_{j,b}$, whereas the remaining fraction $\overline{\alpha}_j = (1 - \alpha_j)$ of its power is used to resolve uncertainty about the messages, $m_{1,b-1}$ and $m_{2,b-1}$, in the previous block. In the first block with $b = 0$, the transmitter j only sends new information with power $\alpha_j P_j$. The rate of the new information R_j from terminal j is chosen such that terminal $\overline{j}$ can correctly decode the message from the feedback, in the end block b, using its own contribution, $x_{\overline{j},b}$, as side information. To allow this, the rate R_j have to satisfy

$$R_j \leq \frac{1}{2} \log_2 (1 + \alpha_j P_j) \,. \tag{2.17}$$

At the receiver, the new information can not be correctly decoded since it does not have any side information, but it does obtain $\frac{1}{2}n_B \log_2(1+\alpha_1 P_1+\alpha_2 P_2)$ out of $n_B(R_1+R_2)$ bits, and hence the clarifying information in the following block has to convey the remaining

$$n_B R_c \geq n_B(R_1 + R_2 - \frac{1}{2}\log_2(1 + \alpha_1 P_1 + \alpha_2 P_2)), \tag{2.18}$$

bits of information.

Since each terminal knows the message of the other terminal at the end of block b, both terminals know what information the receiver has about the messages $m_{1,b}$ and $m_{2,b}$, and hence they can cooperate in order to send the remaining $n_B R_c$ bits in block $b+1$. Cooperation is achieved by transmitting the same codewords of amplitude $\sqrt{\overline{\alpha}_j P_j}$ such that the power of the clarifying information becomes $(\sqrt{\overline{\alpha}_1 P_1} + \sqrt{\overline{\alpha}_2 P_2})^2 = \overline{\alpha}_1 P_1 + \overline{\alpha}_2 P_2 + 2\sqrt{\overline{\alpha}_1 \overline{\alpha}_2 P_1 P_2}$ after superposition on the channel. Note that since the terminals transmit the same signal, the power of the clarifying information is higher than if the terminals transmitted independent information. This effect is termed beamforming and is the key idea of the scheme. Consequently, the receiver can decode the clarifying message in block $b+1$, treating the new information as noise, at the rate

$$R_c \leq \frac{1}{2}\log_2\left(1 + \frac{\overline{\alpha}_1 P_1 + \overline{\alpha}_2 P_2 + 2\sqrt{\overline{\alpha}_1 \overline{\alpha}_2 P_1 P_2}}{1 + \alpha_1 P_1 + \alpha_2 P_2}\right). \tag{2.19}$$

After decoding the clarifying information in block $b+1$, the decoded codeword is subtracted such that only the noisy new information is left. In this way the scheme continuous until the last block B, where no new information is sent. Combining (2.18) and (2.19) yields

$$R_1 + R_2 \leq \frac{1}{2}\left[\log_2\left(1 + \frac{\overline{\alpha}_1 P_1 + \overline{\alpha}_2 P_2 + 2\sqrt{\overline{\alpha}_1 \overline{\alpha}_2 P_1 P_2}}{1 + \alpha_1 P_1 + \alpha_2 P_2}\right) + \log_2(1 + \alpha_1 P_1 + \alpha_2 P_2)\right] \tag{2.20}$$

$$\leq \frac{1}{2}\log_2\left(1 + P_1 + P_2 + 2\sqrt{\overline{\alpha}_1 \overline{\alpha}_2 P_1 P_2}\right). \tag{2.21}$$

The scheme thus achieves the rate region given by (2.17) and (2.21) for any $\alpha_1, \alpha_2 \in [0,1]$. It is readily seen that, when $\alpha_1 = \alpha_2 = 1$, the achievable region of the scheme reduces to the capacity region of the traditional MA channel in (2.16), but in general, the scheme achieves an extended achievable rate region.

Observe that the main idea of this scheme is that, in the end of each block, the terminals knows the message of the other terminal, such that the terminals have common information destined to the receiver, and hence they are able to cooperate and induce beamforming on the MA channel. In the next chapter, this key idea is used to develop interactive schemes for the AF scheme based on this idea.

CHAPTER 3

Amplify-and-Forward with Interaction

In this chapter, the AF scheme is generalized to include interaction. The main idea is to use multiple MA- and BC-phases and a block based coding scheme similar to that for the MACPF in Section 2.4. The signals from the terminals transmitted in each MA-phase are then allowed depend on the signals received in previous BC-phases. In this way, part of the signals of the BC-phase can be seen as feedback for the terminals.

In point-to-point communication, it is well-known that feedback from the receiver to transmitter does not affect the capacity of the system, although complexity of the capacity achieving codes may be significantly decreased [19]. In multiuser systems, however, the MACPF is a clear example where feedback indeed enlarges the capacity region. However, it has also been shown that, even with multiple terminals, the capacity region does not improve using interaction for several channels [20, 21]. Therefore, it is not obvious whether interaction may increase the capacity region of the two-way relay channel.

Here, the hypothesis is that feedback may enlarge the achievable rate region for AF due to the similarities to the MACPF. AF consists of a MA-phase in which both terminals transmit and a BC-phase in which a noisy version of the received codewords are broadcast to the terminals. This is similar to the MACPF except that the noisy superposition is feed back perfectly. However, the major difference lies in the fact that in the MACPF, the aim is to decode the messages at the relay, while for the two-way relay channel, the goal is for each terminal to decode the message of the other terminal.

In the schemes introduced in this chapter, a block coding scheme similar to the one for the MACPF is used. The time is divided into B blocks, each consisting of a MA- and a BC-phase of n channel uses. In contrast to the traditional relaying schemes with only one block, this division enables the terminals to use previously received information in the encoding process, and hence introduce interaction. This is done as following; in block $b \in \{2, \dots, B\}$, the n-dimensional spaces of the MA- and the BC-phases are divided into two orthogonal subspaces $S_{(n)}$ and $S_{(c)}$ such that (1) the subspace $S_{(n)}$ contains new information, and (2) the subspace $S_{(c)}$ contains clarifying information about the new information sent in the previous block. Upon reception at the terminals, each terminal estimates the message of the opposite terminal using the new information received through the subspace $S_{(n)}$. Next, this estimate is improved using the clarifying information in the subspace $S_{(c)}$. The hypothesis is that this model can improve the symmetric achievable rates, i.e. with $R_1 = R_2$, and, eventually, approach the outer bound discussed in Section 2.2.

This chapter proposes two different schemes to send clarifying information that uses the fact the clarifying information becomes correlated as for the MACPF, and hence part of the clarifying information is beamformed such that the relay receives higher power than if the signals were uncorrelated.

Signal-cancelling scheme: In block b, in the end of the BC-phase, suppose that terminal j estimates the signal of terminal $\bar{j}$ as $\mathbf{x}^{(\bar{j})} + \mathbf{z}^{(R)} + \mathbf{z}^{(j)}$ from the new information. During the MA-phase of block $b+1$, the signal cancelling scheme then transmits $-\mathbf{x}^{(j)} + \mathbf{x}^{(\bar{j})} + \mathbf{z}^{(R)} + \mathbf{z}^{(j)}$ as clarifying information such that the signal vectors $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are cancelled at the channel. The relay then receives $2\mathbf{z}^{(R)} + \mathbf{z}^{(1)} + \mathbf{z}^{(2)}$. The key idea of the is that the power of the relay noise $\mathbf{z}^{(R)}$ is beamformed such that it may be used to improve the estimate $\mathbf{x}^{(\bar{j})} + \mathbf{z}^{(R)} + \mathbf{z}^{(j)}$.

Noise-cancelling scheme: For the noise cancelling scheme, terminal 1 transmits $-(-\mathbf{x}^{(1)} + \mathbf{x}^{(2)} + \mathbf{z}^{(R)} + \mathbf{z}^{(1)})$ and terminal 2 transmits $\mathbf{x}^{(1)} - \mathbf{x}^{(2)} + \mathbf{z}^{(R)} + \mathbf{z}^{(2)}$ as clarifying information such that the noise vector $\mathbf{z}^{(R)}$ is cancelled at the relay. The relay then receives $2\mathbf{x}^{(1)} - 2\mathbf{x}^{(2)} - \mathbf{z}^{(1)} + \mathbf{z}^{(2)}$. The signal amplitudes of $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are hence doubled and the corresponding power is increased by a factor 4.

The following describes the system model more carefully and the transmission of new information in the subspace $S_{(n)}$.

3.1. System Model

This section introduces the system model with multiple blocks of MA- and BC-phases. The time is divided into B blocks, each consisting of a MA-phase and a BC-phase. Each of these phases are of n channel uses. Let $\mathbf{x}_b^{(i)} \in \mathbb{C}^{n\times 1}$ denote the signals transmitted from each terminal in block B such that the power constraints

$$\mathbf{x}_b^{(i),\mathrm{H}} \mathbf{x}_b^{(i)} \leq nP \tag{3.1}$$

are satisfied. The main idea is to use a certain number of the n-dimensions in the MA-phase to transmit new information and, the remaining dimensions, to transmit clarifying information. In the MA-phase, the signal dimensions are divided by a division constant $\alpha \in [0.5, 1]$ such that αn of the dimensions are used for transmission of new information and $\bar{\alpha} n = (1-\alpha)n$ of the dimensions are used for clarifying information. Define $\mathbf{U}_{(n)} \in \mathbb{R}^{n\times\alpha n}$, $\mathbf{U}_{(c)} \in \mathbb{R}^{n\times\bar{\alpha} n}$ as matrices such that $[\mathbf{U}_{(n)}, \mathbf{U}_{(c)}]$ is a unitary matrix, i.e.

$$\mathbf{U}_{(n)}^{\mathrm{H}} \mathbf{U}_{(n)} = \mathbf{I}_{\alpha n}, \quad \mathbf{U}_{(c)}^{\mathrm{H}} \mathbf{U}_{(c)} = \mathbf{I}_{\bar{\alpha} n}, \quad \mathbf{U}_{(c)}^{\mathrm{H}} \mathbf{U}_{(n)} = \mathbf{0} \text{ and } \mathbf{U}_{(n)}^{\mathrm{H}} \mathbf{U}_{(c)} = \mathbf{0}, \tag{3.2}$$

where $\mathbf{I}_{\alpha n}$ is the αn-by-αn identity matrix. These matrices defines the orthogonal subspaces $S_{(n)}$ and $S_{(c)}$ in which the new information and clarifying information are transmitted in. For division in time, $\mathbf{U}_{(n)}$ and $\mathbf{U}_{(c)}$ needs to be chosen such that $[\mathbf{U}_{(n)}, \mathbf{U}_{(c)}]$ is the identity matrix $\mathbf{I}_n$. To distribute the power between the two subspaces spanned by the matrices $\mathbf{U}_{(n)}$ and $\mathbf{U}_{(c)}$ in the MA-phase, the amplification factors $\eta_{(n)}$ and $\eta_{(c)}$ are used. The variable $\eta_{(n)}$ scales the new information in the subspace $S_{(n)}$ while $\eta_{(c)}$ scales the clarifying information transmitted in the subspace $S_{(c)}$. Similarly, at the relay two amplification factors $\gamma_{(n)}$ and $\gamma_{(c)}$ are used to scale the signals in the two subspaces.

With these definitions, the signals $\mathbf{x}_b^{(i)}$ can be decomposed into the signal conveying the new information, $\hat{\mathbf{x}}_b^{(i)} \in \mathbb{C}^{\alpha n \times 1}$, and the signal communicating the clarifying information, $\check{\mathbf{x}}_b^{(i)} \in \mathbb{C}^{\overline{\alpha} n \times 1}$, such that $\mathbf{x}_b^{(i)} = \mathbf{U}_{(n)}\hat{\mathbf{x}}_b^{(i)} + \mathbf{U}_{(c)}\check{\mathbf{x}}_b^{(i)}$. Due to this orthogonal division of the subspaces, each of the subspaces can be considered as being independent. For the remaining part of this chapter the accents ˆ and ˇ denote signals in the subspaces $S_{(n)}$ and $S_{(c)}$, respectively. Moreover, we define the matrices $\mathbf{C}_{(n)} \in \mathbb{C}^{(1-2\alpha)\times\alpha}$ and $\mathbf{C}_{(c)} \in \mathbb{C}^{\overline{\alpha}\times\alpha}$. The matrix $\mathbf{C}_{(c)}$ is later used as a compression matrix to compress an αn-dimensional signal, i.e. signals in the subspace $S_{(n)}$, into a $\overline{\alpha} n$-dimensional signal in the subspace $S_{(c)}$. The matrices are chosen such that $\begin{bmatrix} \mathbf{C}_{(n)} \\ \mathbf{C}_{(c)} \end{bmatrix}$ is an unitary matrix.

In the following, the communication within the subspace $S_{(n)}$ is described. In each block, each terminal transmit the new information

$$\hat{\mathbf{x}}_b^{(j)} = \eta_{(n)}\mathbf{w}_b^{(j)}, \tag{3.3}$$

where $\mathbf{w}_b^{(j)}$ is an equivalent, but unscaled version of the new information to be send, and satisfies the power constraint $\mathbf{w}_b^{(j),\mathrm{H}}\mathbf{w}_b^{(j)} \leq \alpha n P$. The relay then receives

$$\hat{\mathbf{y}}_b^{(R)} = \eta_{(n)}\left(\mathbf{w}_b^{(1)} + \mathbf{w}_b^{(2)}\right) + \hat{\mathbf{z}}_b^{(R)}, \tag{3.4}$$

where $\hat{\mathbf{z}}_b^{(R)} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_{\alpha n})$. The signal $\hat{\mathbf{y}}_b^{(R)}$ is scaled by the amplification factor $\gamma_{(n)}$ before being broadcast to the terminals, i.e.

$$\hat{\mathbf{x}}^{(R)} = \gamma_{(n)}\hat{\mathbf{y}}_b^{(R)}. \tag{3.5}$$

At the end of the BC-phase, each terminal receive

$$\hat{\mathbf{y}}_b'^{(j)} = \gamma_{(n)}\left(\eta_{(n)}\mathbf{w}_b^{(\bar{j})} + \hat{\mathbf{z}}_b^{(R)}\right) + \hat{\mathbf{z}}_b^{(j)} \tag{3.6}$$

after cancellation of their own contribution, where $\hat{\mathbf{z}}_b^{(j)} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_{\alpha n})$. Based on (3.6), the terminals may obtain an estimate of $\mathbf{w}_b^{(\bar{j})}$ by dividing $\hat{\mathbf{y}}_b'^{(j)}$ by $\gamma_{(n)}\eta_{(n)}$.

The following sections describes the communication in the clarifying subspace, $S_{(c)}$, for the two interactive schemes.

3.2. Noise Cancelling Scheme

In this section, the noise cancelling scheme is introduced based on the previously described system model. First, the signal model in the clarifying subspace is derived, and then the achievable rates are found. The scheme is introduced in the following.

First block ($b = 1$): Initially, the terminals transmit only new information, and hence the relay only receives noise in the subspace $S_{(c)}$

$$\check{\mathbf{y}}_1^{(R)} = \check{\mathbf{z}}_1^{(R)}. \tag{3.7}$$

where $\check{\mathbf{z}}_b^{(R)} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_{\overline{\alpha}n})$. The relay broadcasts this signal back to terminals in the BC-phase after multiplication by an amplification factor $\gamma_{(c)}$

$$\check{\mathbf{y}}_1^{(i)} = \check{\mathbf{x}}_1^{(R)} + \check{\mathbf{z}}_1^{(i)} = \gamma_{(c)}\check{\mathbf{z}}_1^{(R)} + \check{\mathbf{z}}_1^{(i)} \qquad \text{for } i \in \{1,2\} \tag{3.8}$$

where $\check{\mathbf{z}}_b^{(i)} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_{\overline{\alpha}n})$.

Block $b \in \{2, \dots, B\}$**:** In block b, the terminals transmit the clarifying information

$$\begin{aligned}
\check{\mathbf{x}}_b^{(i)} &= (-1)^i \eta_{(c)} \mathbf{C}_{(c)} (\hat{\mathbf{y}}_{b-1}^{(i)} - 2\gamma_{(n)}\eta_{(n)}\mathbf{w}_{b-1}^{(i)}) && (3.9)\\
&= (-1)^i \eta_{(c)} \mathbf{C}_{(c)} \left(\gamma_{(n)} \left(\eta_{(n)} \left\{-\mathbf{w}_{b-1}^{(i)} + \mathbf{w}_{b-1}^{(\bar{i})}\right\} + \hat{\mathbf{z}}_{b-1}^{(R)}\right) + \hat{\mathbf{z}}_{b-1}^{(i)}\right). && (3.10)
\end{aligned}$$

where $\eta_{(c)}$ is an amplification factor used to distribute the power between the subspaces $S_{(n)}$ and $S_{(c)}$. Note that terminal i is allowed to subtract $2\gamma_{(n)}\eta_{(n)}\mathbf{w}_{b-1}^{(i)}$ from $\hat{\mathbf{y}}_{b-1}^{(i)}$ since it was transmitted by itself in the previous block. Following this step, the resulting signal is compressed using the compression matrix $\mathbf{C}_{(c)}$. Due to the compression by the matrix $\mathbf{C}_{(c)}$, the clarifying signal essentially only contains information about $\overline{\alpha}n$ of the αn dimensions. At the end of the MA-phase, the relay receives

$$\begin{aligned}
\check{\mathbf{y}}_b^{(R)} &= \check{\mathbf{x}}_b^{(1)} + \check{\mathbf{x}}_b^{(2)} + \check{\mathbf{z}}_b^{(R)} && (3.11)\\
&= \eta_{(c)}\mathbf{C}_{(c)} \left[2\gamma_{(n)}\eta_{(n)} \left(\mathbf{w}_{b-1}^{(1)} - \mathbf{w}_{b-1}^{(2)}\right) - \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)} && (3.12)
\end{aligned}$$

The key idea is now that, since the signals $\check{\mathbf{x}}_b^{(1)}$ and $\check{\mathbf{x}}_b^{(2)}$, conveying the clarifying information, are correlated, i.e. the transmitted signals from the terminals both contain $\mathbf{w}_{b-1}^{(1)}$ and $\mathbf{w}_{b-1}^{(2)}$, the signal power is increased due to the beamforming effect. This effect is seen from the factor of 2 in (3.12). The relay scales the signal before broadcasting it back to the terminals

$$\begin{aligned}
\check{\mathbf{x}}_b^{(R)} &= \gamma_{(c)}\check{\mathbf{y}}_b^{(R)} && (3.13)\\
&= \gamma_{(c)} \left\{\eta_{(c)}\mathbf{C}_{(c)} \left[2\gamma_{(n)}\eta_{(n)} \left(\mathbf{w}_{b-1}^{(1)} - \mathbf{w}_{b-1}^{(2)}\right) - \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)}\right\}. && (3.14)
\end{aligned}$$

The terminals then receive the following clarifying information

$$\check{\mathbf{y}}_b^{(i)} = \gamma_{(c)} \left\{\eta_{(c)}\mathbf{C}_{(c)} \left[2\gamma_{(n)}\eta_{(n)} \left(\mathbf{w}_{b-1}^{(1)} - \mathbf{w}_{b-1}^{(2)}\right) - \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)}\right\} + \check{\mathbf{z}}_b^{(i)}. \tag{3.15}$$

After each terminal subtracts its own contribution, the terminals obtain

$$\begin{aligned}
\check{\mathbf{y}}_b'^{(1)} &= \gamma_{(c)} \left\{\eta_{(c)}\mathbf{C}_{(c)} \left[-2\gamma_{(n)}\eta_{(n)}\mathbf{w}_{b-1}^{(2)} - \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)}\right\} + \check{\mathbf{z}}_b^{(i)}. && (3.16)\\
\check{\mathbf{y}}_b'^{(2)} &= \gamma_{(c)} \left\{\eta_{(c)}\mathbf{C}_{(c)} \left[2\gamma_{(n)}\eta_{(n)}\mathbf{w}_{b-1}^{(1)} - \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)}\right\} + \check{\mathbf{z}}_b^{(i)}. && (3.17)
\end{aligned}$$

Estimates of the $\overline{\alpha}n$-dimensions, carried by the clarifying information, of $\mathbf{w}_{b-1}^{(j)}$ can then be obtained from (3.16)-(3.17) by left multiplying $\mathbf{C}_{(c)}^{\mathrm{H}}$ on $\check{\mathbf{y}}_b'^{(j)}$ [1]. In that way, the clarifying

[1] The matrix $\mathbf{C}_{(c)}^{\mathrm{H}}\mathbf{C}_{(c)}$ is essentially an orthogonal projection matrix that maps from the αn-dimensional space into to the αn-dimensional space spanned by $\mathbf{C}_{(c)}^{\mathrm{H}}$. Note that left multiplying $\mathbf{C}_{(c)}^{\mathrm{H}}\mathbf{C}_{(c)}$ onto a vector $\mathbf{v}$ with iid random entries leaves only $\frac{\overline{\alpha}}{\alpha}$ of the power.

information can be used to improve the estimate of $\mathbf{w}_{b-1}^{(j)}$ from the subspace $S_{(n)}$. Note that the difference between the estimates in the subspaces $S_{(c)}$, given by (3.16)-(3.17), and in the subspace $S_{(n)}$, given by (3.6), is that the desired signal is amplified by a factor of 2 and that only $\overline{\alpha}n$ of the αn signal dimensions of $\mathbf{w}_{b-1}^{(j)}$ are contained in the estimate $\check{\mathbf{y}}_{b}^{(j)}$.

With the signaling in both subspaces established, the next section considers conditions for the amplification factors.

3.2.1 Amplification factors

In the derivation of the scheme, there are four amplification factors, $\eta_{(n)}$, $\eta_{(c)}$, $\gamma_{(n)}$ and $\gamma_{(c)}$, that distributes the power among the subspaces $S_{(n)}$ and $S_{(c)}$ in MA- and BC-phase, respectively. The objective is to choose these parameters along with the division constant α in such a way that the rates achievable using the scheme are maximized.

To make the terminals and relay use the available power, the power of the signals in (3.3),(3.5), (3.10) and (3.14) are computed and equated to the power constraints, yielding the following following equations

$$\begin{aligned}
&\frac{1}{n}\hat{\mathbf{x}}_{b}^{(j),\mathrm{H}}\hat{\mathbf{x}}_{b}^{(j)} + \check{\mathbf{x}}_{b}^{(j),\mathrm{H}}\check{\mathbf{x}}_{b}^{(j)} \\
&\qquad = \eta_{(c)}^{2}\overline{\alpha}(\gamma_{(n)}^{2}[2\eta_{(n)}^{2}P+1]+1) + \eta_{(n)}^{2}\alpha P = P, && (3.18)\\
&\frac{1}{n}\hat{\mathbf{x}}_{b}^{(R),\mathrm{H}}\hat{\mathbf{x}}_{b}^{(R)} + \check{\mathbf{x}}_{b}^{(R),\mathrm{H}}\check{\mathbf{x}}_{b}^{(R)} \\
&\qquad = \gamma_{(c)}^{2}\overline{\alpha}\left(\eta_{(c)}^{2}\left[8\gamma_{(n)}^{2}\eta_{(n)}^{2}P+2\right]+1\right) + \gamma_{(n)}^{2}\alpha\left(2\eta_{(n)}^{2}P+1\right) = P_{R}. && (3.19)
\end{aligned}$$

As there are four variables and only two equations to be satisfied, $\eta_{(c)}^{2}$ and $\gamma_{(c)}^{2}$ are found in terms of $\eta_{(n)}^{2} \in \left[0, \frac{1}{\alpha}\right]$ and $\gamma_{(n)}^{2} \in \left[0, \frac{P_R}{\alpha(2\eta_{(n)}^{2}P+1)}\right]$. From (3.18)-(3.19), it is found that

$$\eta_{(c)}^{2} = \frac{P - \eta_{(n)}^{2}\alpha P}{\overline{\alpha}(\gamma_{(n)}^{2}[2\eta_{(n)}^{2}P+1]+1)} \tag{3.20a}$$

$$\gamma_{(c)}^{2} = \frac{P_R - \gamma_{(n)}^{2}\alpha\left(2\eta_{(n)}^{2}P+1\right)}{\overline{\alpha}\left(\eta_{(c)}^{2}\left[8\gamma_{(n)}^{2}\eta_{(n)}^{2}P+2\right]+1\right)}. \tag{3.20b}$$

Consequently, the rate maximization later only has to be performed over the parameters α, $\eta_{(n)}$ and $\gamma_{(n)}$.

3.2.2 Optimal reconstruction

At the terminal, it is desired to reconstruct the best possible estimate of the signal $\mathbf{w}_{b}^{(j)}$ for $j \in \{1,2\}$. In this section, the unbiased linear Minimum Mean Square Error (MMSE) estimator is used to reconstruct the signal $\mathbf{w}_{b}^{(1)}$ at terminal 2, and similar calculations can be carried out at terminal 1. From the subspace $S_{(n)}$, an estimate of $\mathbf{w}_{b}^{(1)}$ is given by (3.6). The objective is then to improve this estimate using the clarifying information in (3.17).

The new and clarifying information are combined using an unbiased estimater

$$\begin{aligned}\hat{\mathbf{w}}_b^{(1)} &= \frac{1}{\gamma_{(n)}\eta_{(n)}}\mathbf{C}_{(n)}^{\mathrm{H}}\mathbf{C}_{(n)}\hat{\mathbf{y}}_b^{\prime(2)} \\ &\quad + \lambda\frac{1}{\gamma_{(n)}\eta_{(n)}}\mathbf{C}_{(c)}^{\mathrm{H}}\mathbf{C}_{(c)}\hat{\mathbf{y}}_b^{\prime(2)} + (1-\lambda)\frac{1}{2\gamma_{(n)}\gamma_{(c)}\eta_{(n)}\eta_{(c)}}\mathbf{C}_{(c)}^{\mathrm{H}}\check{\mathbf{y}}_{b+1}^{\prime(2)} \qquad (3.21) \\ &= \mathbf{w}_b^{(1)} + \frac{1}{\eta_{(n)}}\Bigg[(\mathbf{C}_n^{\mathrm{H}}\mathbf{C}_n + \lambda\mathbf{C}_c^{\mathrm{H}}\mathbf{C}_c)\hat{\mathbf{z}}_b^{(R)} \\ &\qquad + \frac{1}{\gamma_{(n)}}\left(\mathbf{C}_n^{\mathrm{H}}\mathbf{C}_n + \frac{1+\lambda}{2}\mathbf{C}_c^{\mathrm{H}}\mathbf{C}_c\right)\hat{\mathbf{z}}_b^{(2)} \\ &\qquad - \frac{1-\lambda}{2\gamma_{(n)}}\mathbf{C}_c^{\mathrm{H}}\mathbf{C}_c\hat{\mathbf{z}}_b^{(1)} + \frac{1-\lambda}{2\gamma_{(n)}\eta_{(c)}}\mathbf{C}_c^{\mathrm{H}}\check{\mathbf{z}}_{b+1}^{(R)} \\ &\qquad + \frac{1-\lambda}{2\gamma_{(n)}\gamma_{(c)}\eta_{(c)}}\mathbf{C}_c^{\mathrm{H}}\check{\mathbf{z}}_{b+1}^{(2)}\Bigg], \qquad (3.22)\end{aligned}$$

where the matrices $\mathbf{C}_{(n)}^{\mathrm{H}}\mathbf{C}_{(n)} \in \mathbb{R}^{\alpha\times\alpha}$ and $\mathbf{C}_{(c)}^{\mathrm{H}}\mathbf{C}_{(c)} \in \mathbb{R}^{\alpha\times\alpha}$ denote orthogonal projection matrices. The parameter $\lambda \in [0,1]$ is a weighting parameter that determines the weight of the subspaces $S_{(n)}$ and $S_{(c)}$ in the final estimate. In order to obtain the MMSE estimate, the noise variance has to be minimized with respect to the parameter λ.

From (3.22), the noise variance of the estimate can be found as

$$\begin{aligned}&\sigma^2_{\mathrm{AF,NC}}(\lambda,\eta_{(n)},\gamma_{(n)},\alpha) \\ &\quad = \frac{1}{\eta_{(n)}^2\overline{\alpha}}\Bigg[\frac{2\alpha-1}{1-\alpha} + \lambda^2 + \frac{1}{\gamma_{(n)}^2}\left(2\alpha - 1 + \frac{1}{4}(1+\lambda)^2\right) \\ &\qquad + \frac{(1-\lambda)^2(1+\gamma_{(c)}^2+\gamma_{(c)}^2\eta_{(c)}^2)}{4\gamma_{(n)}^2\gamma_{(c)}^2\eta_{(c)}^2}\Bigg]. \qquad (3.23)\end{aligned}$$

The parameter λ is then chosen such that $\sigma^2_{\mathrm{AF,NC}}(\lambda,\eta_{(n)},\gamma_{(n)},\alpha)$ is minimized, i.e.

$$\sigma^2_{\mathrm{AF,NC}}(\eta_{(n)},\gamma_{(n)},\alpha) = \min_{\lambda\in[0,1]}\sigma^2_{\mathrm{AF,NC}}(\lambda,\eta_{(n)},\gamma_{(n)},\alpha). \qquad (3.24)$$

The minimization of (3.23) with respect to λ is easily performed since (3.23) is a parabola in terms of λ.

Achievable rates The aim of the this section is to maximize the achievable rate using the signal cancelling scheme. Within the nB channel uses, this scheme transmits $\alpha(B-1)$ Gaussian variables $\mathbf{w}_b^{(j)}$. Assuming that the number of blocks B tends to ∞, the achievable rates of each terminal can be expressed as

$$R_{\mathrm{AF,NC}}(\eta_{(n)},\gamma_{(n)},\alpha) = \alpha\log_2\left(1 + \frac{\alpha P}{\sigma^2_{\mathrm{AF,NC}}(\eta_{(n)},\gamma_{(n)},\alpha)}\right). \qquad (3.25)$$

This rate is then maximized with respect to the division constant α and the amplification factors $\eta_{(n)}$ and $\gamma_{(n)}$ (the amplification factors $\eta_{(c)}$ and $\gamma_{(c)}$ are given by (3.20))

$$R^{(\mathrm{opt})}_{\mathrm{AF,NC}} = \max_{\eta_{(n)},\gamma_{(n)},\alpha} R_{\mathrm{AF}}(\eta_{(n)},\gamma_{(n)},\alpha). \qquad (3.26)$$

Note that the problem is not convex. The problem is therefore solved using `fmincon()` in MATLAB with a large number of initial points. Results are shown in the numerical results in Section 3.4

3.3. Signal Cancelling Scheme

This section considers the signal cancelling scheme. As opposed to the noise cancelling AF scheme where the signal contributions in the clarifying signals are correlated, this scheme cancels the signal contributions on the channel and beamforms the relay noise. The approach for the derivation is similar to the noise cancelling scheme, and hence repeated derivations, including signaling in the first block, are omitted.

Block $b \in \{2, \ldots, B\}$: In block b, the signal cancelling scheme

$$\check{\mathbf{x}}_b^{(i)} = \eta_{(c)}\mathbf{C}_{(c)}(\hat{\mathbf{y}}_{b-1}^{(i)} - 2\gamma_{(n)}\eta_{(n)}\mathbf{w}_{b-1}^{(i)}) \tag{3.27}$$

$$= \eta_{(c)}\mathbf{C}_{(c)}\left(\gamma_{(n)}\left(\eta_{(n)}\left\{-\mathbf{w}_{b-1}^{(i)} + \mathbf{w}_{b-1}^{(\bar{i})}\right\} + \hat{\mathbf{z}}_{b-1}^{(R)}\right) + \hat{\mathbf{z}}_{b-1}^{(i)}\right). \tag{3.28}$$

where $\eta_{(c)}$ is a scaling variable used for power distribution among the subspaces $S_{(n)}$ and $S_{(c)}$. The relay then receives

$$\check{\mathbf{y}}_b^{(R)} = \eta_{(c)}\mathbf{C}_{(c)}\left[2\gamma_{(n)}\hat{\mathbf{z}}_{b-1}^{(R)} + \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)} \tag{3.29}$$

As opposed to the noise cancelling scheme, the key idea with this scheme is that the noise received at the relay in block $b-1$ is beamformed as clarifying information in block b. The relay scales the signal before broadcasting it back to terminals

$$\check{\mathbf{x}}_b^{(R)} = \gamma_{(c)}\left\{\eta_{(c)}\mathbf{C}_{(c)}\left[2\gamma_{(n)}\hat{\mathbf{z}}_{b-1}^{(R)} + \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)}\right\}. \tag{3.30}$$

Finally, the terminals receive the clarifying information

$$\check{\mathbf{y}}_b^{(i)} = \gamma_{(c)}\left\{\eta_{(c)}\mathbf{C}_{(c)}\left[2\gamma_{(n)}\hat{\mathbf{z}}_{b-1}^{(R)} + \hat{\mathbf{z}}_{b-1}^{(1)} + \hat{\mathbf{z}}_{b-1}^{(2)}\right] + \check{\mathbf{z}}_b^{(R)}\right\} + \check{\mathbf{z}}_b^{(i)}. \tag{3.31}$$

The signal obtained in (3.31) is essentially an estimate of the noise $\hat{\mathbf{z}}_{b-1}^{(R)}$. As for the noise cancelling scheme, estimates of $\mathbf{w}_{b-1}^{(j)}$ can then be obtained from (3.6) and improved using the clarifying information in (3.31) by cancelling the noise $\hat{\mathbf{z}}_{b-1}^{(R)}$.

3.3.1 Scaling coefficients

As for the noise cancelling scheme, constraints for the amplification factors $\eta_{(n)}$, $\eta_{(c)}$, $\gamma_{(n)}$ and $\gamma_{(c)}$ are computed. This is done by computing the power of (3.3),(3.5), (3.28) and (3.30) and equating to zero, yielding the equations

$$\eta_{(c)}^2\overline{\alpha}(\gamma_{(n)}^2[2\eta_{(n)}^2 P + 1] + 1) + \eta_{(n)}^2\alpha P = P, \tag{3.32}$$

$$\gamma_{(c)}^2\overline{\alpha}\left(\eta_{(c)}^2\left[4\gamma_{(n)}^2 + 2\right] + 1\right) + \gamma_{(n)}^2\alpha\left(2\eta_{(n)}^2 P + 1\right) = P_R. \tag{3.33}$$

As there are four variables for only two equations, $\eta_{(c)}^2$ and $\gamma_{(c)}^2$ are found in terms of $\eta_{(n)}^2 \in \left[0, \frac{1}{\alpha}\right]$ and $\gamma_{(n)}^2 \in \left[0, \frac{P_R}{\alpha(2\eta_{(n)}^2 P+1)}\right]$. From (3.32)-(3.33), it is found that

$$\eta_{(c)}^2 = \frac{P - \eta_{(n)}^2\alpha P}{\overline{\alpha}(\gamma_{(n)}^2[2\eta_{(n)}^2 P + 1] + 1)} \tag{3.34}$$

$$\gamma_{(c)}^2 = \frac{P_R - \gamma_{(n)}^2\alpha\left(2\eta_{(n)}^2 P + 1\right)}{\overline{\alpha}\left(\eta_{(c)}^2\left[4\gamma_{(n)}^2 + 2\right] + 1\right)}. \tag{3.35}$$

3.3.2 Optimal reconstruction

An estimate of $\mathbf{w}_b^{(j)}$ can be constructed based (3.6) and (3.31). As for the noise cancelling scheme, an unbiased linear estimater is used to obtain the estimate

$$\hat{\mathbf{w}}_b^{(1)} = \frac{1}{\gamma_{(n)}\eta_{(n)}}\hat{\mathbf{y}}_b^{(2)} - (1-\lambda)\frac{1}{2\gamma_{(n)}\gamma_{(c)}\eta_{(c)}}\mathbf{C}_{(c)}^{\mathrm{H}}\check{\mathbf{y}}_{b+1}^{(2)} \tag{3.36}$$

$$\begin{aligned} &= \mathbf{w}_b^{(1)} + \frac{1}{\eta_{(n)}}\Big[(\mathbf{C}_n^{\mathrm{H}}\mathbf{C}_n + \lambda\mathbf{C}_c^{\mathrm{H}}\mathbf{C}_c)\hat{\mathbf{z}}_b^{(R)} \\ &\quad + \frac{1}{\gamma_{(n)}}\left(\mathbf{C}_n^{\mathrm{H}}\mathbf{C}_n + \frac{1+\lambda}{2}\mathbf{C}_c^{\mathrm{H}}\mathbf{C}_c\right)\hat{\mathbf{z}}_b^{(2)} \\ &\quad - \frac{1-\lambda}{2\gamma_{(n)}}\mathbf{C}_c^{\mathrm{H}}\mathbf{C}_c\hat{\mathbf{z}}_b^{(1)} - \frac{1-\lambda}{2\gamma_{(n)}\eta_{(c)}}\mathbf{C}_c^{\mathrm{H}}\check{\mathbf{z}}_{b+1}^{(R)} \\ &\quad - \frac{1-\lambda}{2\gamma_{(n)}\gamma_{(c)}\eta_{(c)}}\mathbf{C}_c^{\mathrm{H}}\check{\mathbf{z}}_{b+1}^{(2)}\Big] \end{aligned} \tag{3.37}$$

Surprisingly, this expression is essentially identical to the estimate for the noise cancelling scheme in (3.22) apart from signs differences. This implies that the MMSE estimate of $\mathbf{w}_b^{(j)}$ has a noise variance given by

$$\sigma^2_{\mathrm{AF,SC}}(\eta_{(n)}, \gamma_{(n)}, \alpha) = \min_{\lambda\in[0,1]} \sigma^2_{\mathrm{AF,NC}}(\lambda, \eta_{(n)}, \gamma_{(n)}, \alpha). \tag{3.38}$$

The difference from the noise cancelling schemes lies constraints on the amplification factors in (3.33). The rates achievable with this scheme are hence found by the same optimization problem as for the noise cancelling scheme in (3.26).

The next section solves the two optimization problems and compares the performance of the schemes to the AF scheme.

3.4. Numerical Results

In this section, the achievable symmetric rates $R = R_1 = R_2$ of the two schemes are found. The achievable rates are computed by solving the optimization problem in (3.26) with the amplification factors for each of the schemes. As the optimization problems have not been shown to be convex, the optimization over the three parameters is performed using `fmincon()` with 100 random feasible initial points. The achievable symmetric rates are computed in terms of the terminal power constraint P and the relay power constraint P_R, and the results are depicted in Fig. 3.1. It is seen that no improvement, compared to the AF scheme, is obtained since the achievable rates of the proposed schemes coincides with the AF scheme for all values of P and P_R.

To develop insights, Fig. 3.2 shows the achievable rates if the division constant α is fixed to a certain value in the interval $[0.5, 1]$. These results shows that the achievable rates strictly decreases as the division factor α decreases. When $P = P_R = 1$, the two schemes yields identical achievable rates. This is because the optimization problem yields the trivial solution in which $\eta_{(n)} = \gamma_{(n)} = 1$ for both schemes. This is not the case when $P = P_R = 100$, where the signal cancelling scheme is better for low α values.

Although the developed schemes do not show improvements, it does not mean that interaction can not improve the achievable rates of the two-way half-duplex relay channel.

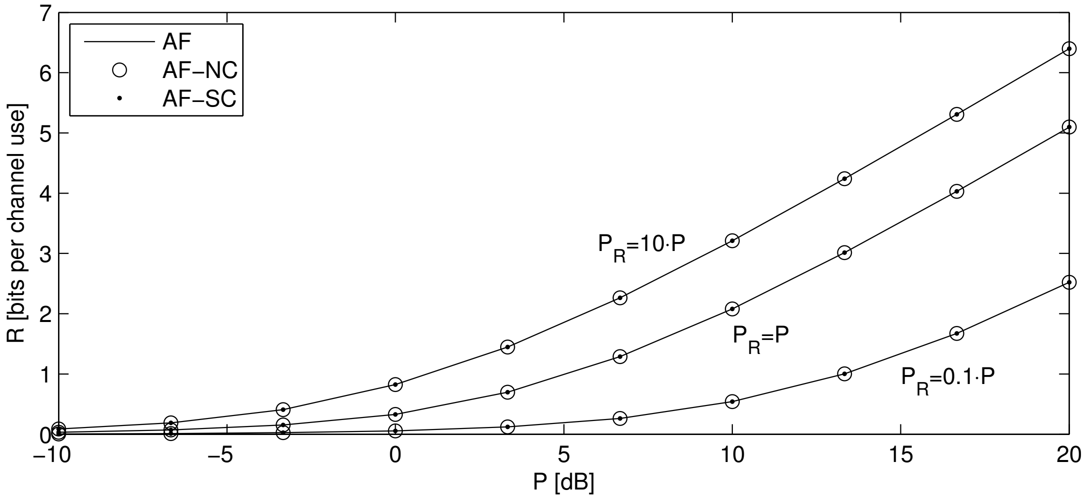
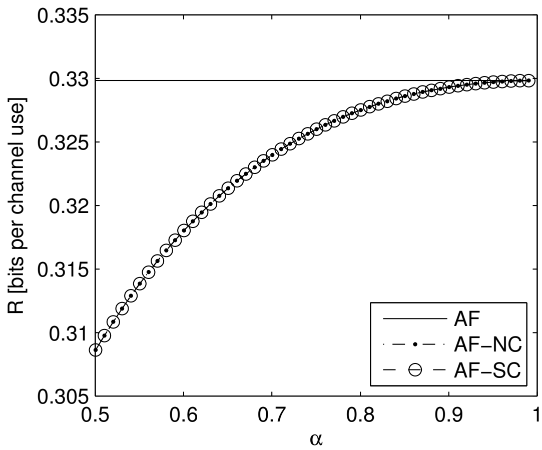


Figure 3.1: *Comparison of the proposed AF schemes.*

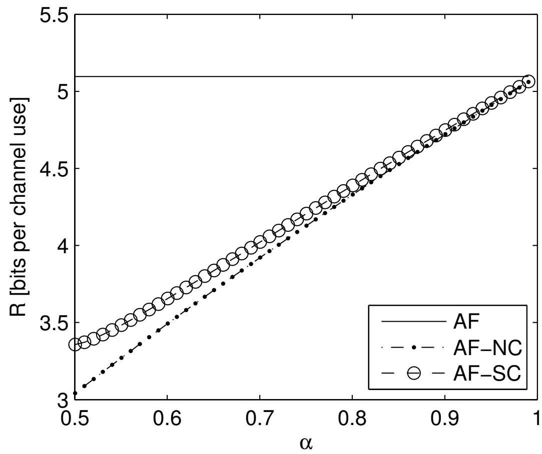
The main problem with the developed schemes seem to be that the noises added each time a signal is received at the terminals or the relay accumulates when it is send back and forth between the terminal and relay. This problem may be circumvented by using compress-and-forward or more advanced relaying schemes based on lattice coding [22], and hence further work may explore different relaying strategies with signaling similar to the proposed schemes.

A different path to follow is to take the contrary hypothesis, i.e. that feedback can not increase the achievable rates of two-relay relay channel for some relaying scheme. Then the approach the taken in [20, 21] can be used to derive upper bounds on the achievable rates with feedback. If it can be shown that the upper bounds coincides with the achievable rates, it can be included that interaction does not improve the scheme.

In fact, it is hypothesized in [1] that interaction can not improve the capacity regions if the following three conditions are satisfied: (1) self-interference can be cancelled, (2) there are no loops in the network, and (3) there are no "coherent gains" such no nodes in the network has common information destined to other terminals. The two-way half-duplex relay channel indeed satisfies these conditions.



(a) $P = P_R = 1$.



(b) $P = P_R = 100$.

Figure 3.2: *Achievable rates of the signal cancelling and noise cancelling schemes compared to the AF scheme in terms of fixed* $\alpha \in [0.5, 1]$.

Part II

Coordinated Relaying

CHAPTER 4

State of the Art and Background

This chapter contains the preliminaries of the part on coordinated relaying. In particular, the capacity results and the optimal power allocation strategy for MIMO channels are summarized. MIMO essentially extends the time and frequency dimensions with a spatial dimension. In that way, the throughput and reliability in wireless communication systems is increased. Besides these advantages, MIMO also allows a node to communicate with multiple other at the same time and frequency, but in different spatial dimensions. This idea is fundamental for the principles alignment and zero-forcing, which are introduced later. Based on the results of the MIMO channels several channels, which are used later, are introduced.

After introducing the MIMO channel, the fundamental results on the multiple access channel, the wireless MIMO switch, the two-way MIMO relay channel and the MIMO four-way relay channel are summarized.

4.1. MIMO Channel

The Multiple-Input Multiple-Output channel with M_1 transmitting and M_2 receiving antennas is depicted in Fig 4.1. In this section the capacity results and the optimal power allocation strategy are summarized. The derivation of the capacity of the MIMO channel first appeared in [15] by Teletar, and since then, MIMO channels have been subject to a massive amount of research. MIMO channels, as opposed to Single-Input Single-Output (SISO) channels, not only increases the throughput under the same power constraints, but also makes the channels more robust to fading and enables the possibility of creating multiple separate spatial channels. These properties are termed *transmit diversity* and *spatial multiplexing*, respectively [23].

The signal $\mathbf{x} \in \mathbb{C}^{M_1 \times 1}$ is emitted by the transmitter and is subject to the average power constraint P, i.e.

$$\mathbb{E}\left[\mathbf{x}^{\mathrm{H}}\mathbf{x}\right] = \operatorname{tr}\left(\mathbf{x}\mathbf{x}^{\mathrm{H}}\right) \leq P. \tag{4.1}$$

For the Gaussian SISO channel, Gaussian signaling is optimal [16]. Likewise multivariate Gaussian signaling, with covariance matrix denoted by the positive definite matrix $\mathbf{Q}_{\mathbf{x}} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^{\mathrm{H}}\right] \in \mathbb{C}^{M_1 \times M_1}$, is also optimal for the MIMO channel [15]. The signal is propagated through a linear channel modeled by the matrix $\mathbf{H} \in \mathbb{C}^{M_2 \times M_1}$, such that the receiver observes $\mathbf{y} \in \mathbb{C}^{M_2 \times 1}$, i.e.

$$\mathbf{y} = \mathbf{H}\mathbf{y} + \mathbf{z}, \tag{4.2}$$

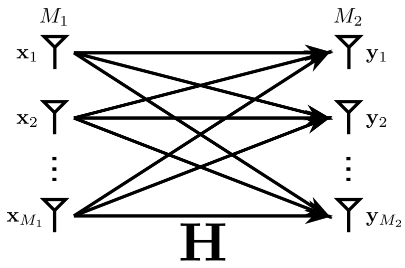


Figure 4.1: *The MIMO channel.*

where $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_\mathbf{z})$ with $\mathbf{C}_\mathbf{z} \in \mathbb{C}^{M_2 \times M_2}$ being a positive definite noise covariance matrix. The objective is to determine the maximum amount of information, denoted $I(\mathbf{x}, \mathbf{y})$, that can be transmitted from the transmitter to the receiver per channel use, in terms of the covariance matrix $\mathbf{Q}_\mathbf{x}$. An expression for the capacity is readily derived as in [15]

$$
\begin{aligned}
C_{\text{MIMO}} &= \max_{\mathbf{Q}_\mathbf{x} \succeq 0, \operatorname{tr}(\mathbf{Q}_\mathbf{x}) \leq P} I(\mathbf{x}; \mathbf{y}) && (4.3)\\
&= \max_{\mathbf{Q}_\mathbf{x} \succeq 0, \operatorname{tr}(\mathbf{Q}_\mathbf{x}) \leq P} H(\mathbf{y}) - H(\mathbf{z}) && (4.4)\\
&= \max_{\mathbf{Q}_\mathbf{x} \succeq 0, \operatorname{tr}(\mathbf{Q}_\mathbf{x}) \leq P} \log_2 \left| \pi e (\mathbf{I}_{M_2} + \mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^{\mathrm{H}}) \right| - \log_2 \left| \pi e \mathbf{C}_\mathbf{z} \right| && (4.5)\\
&= \max_{\mathbf{Q}_\mathbf{x} \succeq 0, \operatorname{tr}(\mathbf{Q}_\mathbf{x}) \leq P} \log_2 \left| \mathbf{I}_{M_2} + \mathbf{C}_\mathbf{z}^{-1}\mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^{\mathrm{H}} \right|, && (4.6)
\end{aligned}
$$

where $|\mathbf{A}|$ denotes the determinant of the matrix $\mathbf{A}$. The objective function of the maximization problem in (4.6) is concave and the set of covariance matrices $\mathbf{Q}_\mathbf{x}$ satisfying the power constraint is convex. Therefore, the expression in (4.6) can be computed using standard convex optimization tools. However, the solution is efficiently found using the *waterfilling* algorithm, which is explained next and also gives insight to the problem of optimal power allocation.

Waterfilling solution In the following, the waterfilling algorithm used to perform the maximization in (4.6) is derived based on [15]. To simplify the treatment in the following, assume that $\mathbf{C}_\mathbf{z}^{-1/2}\mathbf{H}$ is of full rank and that $M = M_1 = M_2$[1]. Moreover, denote the singular value decomposition (SVD) of $\mathbf{C}_\mathbf{z}^{-1/2}\mathbf{H}$ by $\mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{H}}$. By Sylvester's determinant theorem[2] and the property $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}_M$, the following manipulations on the objective function of (4.6) can then be performed

$$
\begin{aligned}
\log_2 \left| \mathbf{I}_M + \mathbf{C}_\mathbf{z}^{-1}\mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^{\mathrm{H}} \right| &= \log_2 \left| \mathbf{I}_M + \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{H}}\mathbf{Q}_\mathbf{x}\mathbf{V}\mathbf{S}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}} \right| && (4.7)\\
&= \log_2 \left| \mathbf{I}_M + \mathbf{S}^{\mathrm{H}}\mathbf{S}\mathbf{V}^{\mathrm{H}}\mathbf{Q}_\mathbf{x}\mathbf{V} \right|. && (4.8)
\end{aligned}
$$

[1] In [15], the waterfilling algorithm is derived for any M_1 and M_2.

[2] Sylvester's determinant theorem states that $|\mathbf{I}_p + \mathbf{A}\mathbf{B}| = |\mathbf{I}_q + \mathbf{B}\mathbf{A}|$ for $\mathbf{A} \in \mathbb{C}^{p \times q}$ and $\mathbf{B} \in \mathbb{C}^{q \times p}$.

Let $\tilde{\mathbf{Q}}_{\mathbf{x}}$ denote $\mathbf{V}^{\mathrm{H}}\mathbf{Q}_{\mathbf{x}}\mathbf{V}$ and s_i denote the i-th diagonal entry of $\mathbf{S}^{\mathrm{H}}\mathbf{S}$. Then (4.8) can be upper bounded by the Hardamard's inequality[3] as

$$\log_2 \left|\mathbf{I}_M + \mathbf{S}^{\mathrm{H}}\mathbf{S}\mathbf{V}^{\mathrm{H}}\mathbf{Q}_{\mathbf{x}}\mathbf{V}\right| \leq \log_2 \left(\prod_{i=1}^{M}(1 + s_i\tilde{\mathbf{Q}}_{\mathbf{x},ii})\right) \tag{4.9}$$

$$= \sum_{i=1}^{M} \log_2 \left(1 + s_i\tilde{\mathbf{Q}}_{\mathbf{x},ii}\right) \tag{4.10}$$

where $\tilde{\mathbf{Q}}_{\mathbf{x},ii}$ denotes the i-th diagonal entry of $\tilde{\mathbf{Q}}_{\mathbf{x}}$. Since $\mathbf{Q}_{\mathbf{x}}$ is required to be positive semidefinite, it is seen that $\tilde{\mathbf{Q}}_{\mathbf{x}}$ has to be chosen as diagonal matrix with positive entries such that the inequality in (4.9) is satisfied with equality. The optimization problem in (4.6) is then reduced to the problem of finding the diagonal entries $\tilde{\mathbf{Q}}_{\mathbf{x},ii}$ for $i \in \{1, \ldots, M\}$, i.e.

$$C_{\mathrm{MIMO}} = \max_{\tilde{\mathbf{Q}}_{\mathbf{x},ii} \geq 0, \mathrm{tr}\left(\tilde{\mathbf{Q}}_{\mathbf{x}}\right) \leq P} \sum_{i=1}^{M} \log_2 \left(1 + s_i\tilde{\mathbf{Q}}_{\mathbf{x},ii}\right). \tag{4.11}$$

This problem is readily solved using Lagrange multipliers. Let $\mu \geq 0$ and $\lambda_i \geq 0$ denote the Lagrange multipliers for the constraints $\mathrm{tr}\left(\tilde{\mathbf{Q}}_{\mathbf{x}}\right) \leq P$ and $\tilde{\mathbf{Q}}_{\mathbf{x},ii} \geq 0$. Then we have the stationarity equations

$$\frac{s_i}{1 + s_i\tilde{\mathbf{Q}}_{\mathbf{x},ii}} + \mu - \lambda_i = 0 \Leftrightarrow \mathbf{Q}_{\mathbf{x},ii} = \frac{1}{\mu - \lambda_i} - \frac{1}{s_i}, \tag{4.12}$$

for $i \in \{1, \ldots, M\}$ and the complementary conditions, $\tilde{\mathbf{Q}}_{\mathbf{x},ii}\lambda_i = 0$. Consequently, the solution of the problem is given

$$\tilde{\mathbf{Q}}_{\mathbf{x},ii} = \max\left(0, \frac{1}{\mu} - \frac{1}{s_i}\right), \tag{4.13}$$

and the corresponding Lagrange multipliers for the positivity of the diagonal entries are given by $\lambda_i = \max(0, \mu - s_i)$. The remaining Lagrange multiplier, μ, for the power constraint is found numerically. The value $\frac{1}{\mu}$ is termed the waterlevel and has to be chosen such that $\sum_{i=1}^{M} \tilde{\mathbf{Q}}_{\mathbf{x},ii} = P$.

Finally, the optimal covariance matrix of the input $\mathbf{x}$ to the MIMO channel is given by $\mathbf{V}\tilde{\mathbf{Q}}_{\mathbf{x}}\mathbf{V}^{\mathrm{H}}$. In the field of MIMO, the column vectors of the matrix $\mathbf{V}$ are called eigenmodes, and since they are right singular vectors of the channel $\mathbf{C}_{\mathbf{n}}^{-1/2}\mathbf{H}$, they essentially describes M orthogonal Gaussian SISO channels, each with the channel coefficient s_i. The idea, and the reason for the naming, of the waterfilling algorithm is that, as the waterlevel increases, the algorithm starts pouring power into the strong eigenmodes. That is, at low power, only the strongest eigenmodes are used, whereas most eigenmodes are used at high power. Achieving the capacity of the MIMO can therefore be achieved using traditional capacity achieving Gaussian codes, such as Low-Density Parity Check (LDPC) codes [24], for each of the eigenmodes [16].

[3] The Hardmard's inequality states that $\det \mathbf{A} \leq \prod_i \mathbf{A}_{ii}$, where $\mathbf{A}$ is a positive semidefinite matrix and $\mathbf{A}_{ii}$ denotes the i-th diagonal entry column of it. The inequality is satisfied with equality if and only if the matrix $\mathbf{A}$ is a diagonal matrix or at least on of the diagonal entries are zero.

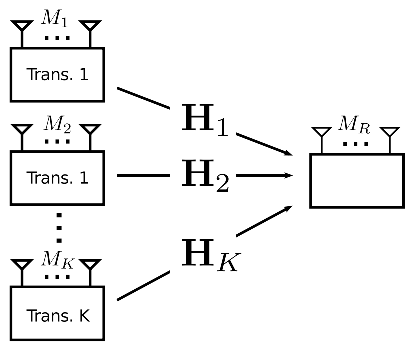


Figure 4.2: *The general MIMO MA channel.*

Degrees of Freedom At high SNR, the performance of the MIMO channel is determined by the multiplexing gain, measured in Degrees of Freedom (DoF) [23]. The DoF basically describes the number of interference free dimensions at high SNR, which corresponds to the pre-log factor of the capacity formula. Assuming that $\mathbf{C}_{\mathbf{n}}^{-1/2}\mathbf{H}$ is of full rank, the DoF for the MIMO channel may easily be found. At high SNR, i.e. when $P = \sum_{i=1}^{M_1} \tilde{\mathbf{Q}}_{\mathbf{x},ii}$ tends to ∞, all entries of $\widetilde{\mathbf{Q}}_{\mathbf{x},ii}$ are non-zero, and hence the capacity of the MIMO channel can be expressed as

$$C_{\text{MIMO,high}} = \min(M_1, M_2)\log_2(P) + o(\log_2(P)), \tag{4.14}$$

where $o(\log_2(P))$ collects all terms growing slower than $\log_2(P)$. From this equation, it is seen that the DoF is $\min(M_1, M_2)$, and hence limited by the number of antennas at either the transmitter or receiver.

4.2. Multiple-Access MIMO Channel

This section extends the results of the MIMO channel to the MA channel with K transmitters with M_i antennas and one receiver with M_R antennas. This setting is depicted in Fig 4.2. Transmitter $i \in \{1, \dots, K\}$ has a message $m_i \in \{1, \dots, 2^{nR_i}\}$ destined to the receiver. The aim is to characterize the achievable rates R_i over n channel uses. Each transmitter encodes its message into a codeword of n channel uses, $\mathbf{x}_i \in \mathbf{C}^{n\times 1}$, with the constraint that $\frac{1}{n}\mathbf{x}^{\mathrm{H}}\mathbf{x} \leq P_i$, where P_i denotes the average power allowed at transmitter i. Moreover, the codewords are Gaussian distributed with covariance matrices $\mathbf{Q}_i$. In the following, the achievable rate regions and the corresponding coding strategies are summarized.

Let $\mathbf{H}_i \in \mathbb{C}^{M_R\times M_i}$ for $i \in \{1, 2, \dots, K\}$ denote the channel matrices from transmitter i to the receiver, which are assumed to be of full rank. After transmission the receiver obtains a superposition of the codewords from each transmitter $\mathbf{y} \in \mathbb{C}^{m_R\times 1}$

$$\mathbf{y} = \sum_{i=1}^{K} \mathbf{H}_i\mathbf{x}_i + \mathbf{z}, \tag{4.15}$$

where $\mathbf{z} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{M_R})$. The capacity region for fixed covariance matrices, containing the achievable rates, is then given as [25]

$$\mathcal{C}_{\text{MAC}} = \bigcup_{\mathbf{Q}_i \succeq 0, \text{tr}(\mathbf{Q}_i) \leq P_i} \left\{ \sum_{i \in S} R_i < \log_2 \left| \mathbf{I}_{M_R} + \sum_{i \in S} \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^{\text{H}} \right|, \forall S \subseteq \{1, 2, \ldots, K\} \right\}. \tag{4.16}$$

In particular, when $K = 2$, the achievable rate region satisfies

$$R_1 < R_1^{(A)} = \log_2 \left| \mathbf{I}_{M_R} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} \right| \tag{4.17a}$$

$$R_2 < R_2^{(D)} = \log_2 \left| \mathbf{I}_{M_R} + \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\text{H}} \right| \tag{4.17b}$$

$$R_1 + R_2 < \log_2 \left| \mathbf{I}_{M_R} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} + \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\text{H}} \right|. \tag{4.17c}$$

For fixed $\mathbf{Q}_1$ and $\mathbf{Q}_2$, in general, the achievable rate region in (4.17) is a pentagon as depicted in Fig 4.3. Point A corresponds to the rate achievable if the receiver knows the codeword transmitted by transmitter 2. Thus to achieve point A, transmitter 2 emits a codeword $\mathbf{x}_2$ with covariance matrix $\mathbf{Q}_2$, which is known by the receiver (the rate of transmitter 2 should be 0). The receiver subtract the known codeword, $\mathbf{x}_2$, and decodes the message from transmitter 1 at the rate $R_1 < R_1^{(A)}$. To achieve point B, the receiver first decodes the message from transmitter 2, treating the signal from transmitter 1 as noise. The achievable rate of this operation is

$$R_2 < R_2^{(B)} = \log_2 \left| \mathbf{I}_{M_R} + (\mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} + \mathbf{I}_{m_R})^{-1} \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\text{H}} \right| \tag{4.18}$$

The receiver subtracts the decoded codeword $\mathbf{x}_2$ and the message from transmitter 1 is then decoded at rate

$$R_1 < R_1^{(A)} = \log_2 \left| \mathbf{I}_{M_R} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} \right|. \tag{4.19}$$

Finally, it can be verified that the sum rate inequality can approached arbitrary close

$$\begin{aligned} R_1 + R_2 &< R_1^{(A)} + R_2^{(B)} && (4.20) \\ &= \log_2 \left| \mathbf{I}_{M_R} + (\mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} + \mathbf{I}_{M_R})^{-1} \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\text{H}} \right| + \log_2 \left| \mathbf{I}_{m_R} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} \right| \\ &= \log_2 \left| (\mathbf{I}_{M_R} + (\mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} + \mathbf{I}_{M_R})^{-1} \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\text{H}})(\mathbf{I}_{m_R} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}}) \right| && (4.21) \\ &= \log_2 \left| \mathbf{I}_{M_R} + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\text{H}} + \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\text{H}} \right|. && (4.22) \end{aligned}$$

The point D and C are achieved in similar ways and all remaining points can be achieved using time-sharing.

Degrees of Freedom Using similar arguments for the MA channel as for the MIMO channel, DoF constraints can be found for the transmitters. Let $d_1, \ldots, d_K$ denote the DoF of the terminals $1, \ldots, K$. It can then be found that DoF of the terminals have to satisfy [26]

$$\sum_{i \in S} d_i \leq \min \left(\sum_{i \in S} M_i, M_R \right) \qquad \forall S \in \{1, \ldots, K\}. \tag{4.23}$$

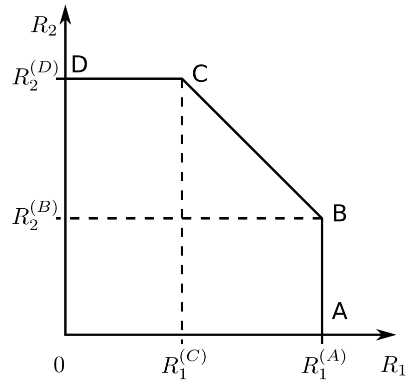


Figure 4.3: *The general MIMO MA channel.*

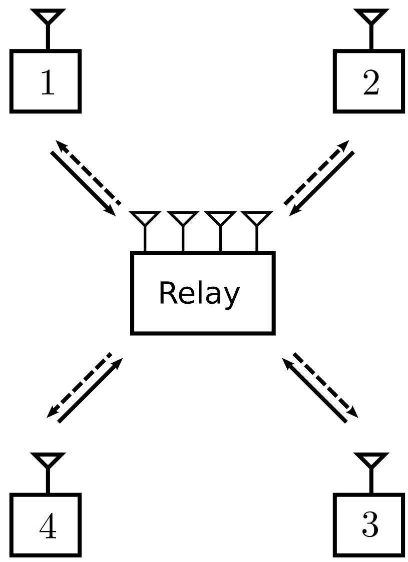


Figure 4.4: *The wireless MIMO switch with $K = 4$*

4.3. Wireless MIMO Switching

The wireless MIMO switch is depicted in Fig. 4.4, and consists of a K terminals, indexed $1, \ldots, K$, with one antenna and one relay with at least K antennas. For simplicity, assume that the relay has exactly K antennas. The message from each terminal is destined to another terminal in such a way that exactly one message is destined to each terminal. These communication flows can be represented through a permutation matrix $\mathbf{P}$[4]. Communication is achieved through a MA-phase and a BC-phase. This system model is considered in multiple works, e.g. [14, 27]. In [27], the main finding is that, due to the K antennas at the relay, the multiplexing gain is fully utilized, and hence one DoF is achieved for each communication flow.

In [14], the practical relaying scheme Precode-and-Forward, is considered. In the MA-phase, each terminal transmit their signal $x_i \in \mathbb{C}$ to the relay through a linear channel $\mathbf{H}_i \in \mathbb{C}^{K \times 1}$ for $i \in \{1, \ldots, K\}$. The relay receives a noisy superposition of the signals from

[4]A matrix with exactly one '1' on each row and column. That is, if there is a '1' in the p-th row and the q-th column, then the message from terminal q is addressed to terminal p.

the terminals, $\mathbf{y}_R = \underbrace{[\mathbf{H}_1, \ldots, \mathbf{H}_K]}_{\mathbf{H}} \underbrace{[x_1, \ldots, x_K]^{\mathrm{T}}}_{\mathbf{x}} + \mathbf{z}$, where $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_K)$. Upon reception at the relay, the received signal $\mathbf{y}_R$ is simply multiplied by a precoding matrix $\mathbf{G}$ before being broadcasted back to the terminals in the BC-phase. The terminals then receive

$$\mathbf{y} = [y_1, \ldots, y_K]^{\mathrm{T}} = \mathbf{H}^{\mathrm{T}}\mathbf{G}\mathbf{H}\mathbf{x} + \mathbf{H}^{\mathrm{T}}\mathbf{G}\mathbf{z} + \mathbf{w} \tag{4.24}$$

where $\mathbf{w} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_K)$. By choosing, the precoder matrix $\mathbf{G}$ such that the matrix $\mathbf{H}^{\mathrm{T}}\mathbf{G}\mathbf{H}$ is forced to be on the form

$$\mathbf{H}^{\mathrm{T}}\mathbf{G}\mathbf{H} = \mathbf{A}\mathbf{P} + \mathbf{B} \Leftrightarrow \mathbf{G} = (\mathbf{H}^{\mathrm{T}})^{-1}(\mathbf{P}\mathbf{A} + \mathbf{B})\mathbf{H}^{-1} \tag{4.25}$$

where $\mathbf{P}$ is the permutation matrix describing the communication flows and $\mathbf{A}, \mathbf{B}$ are diagonal matrices. Note that the diagonal entries of $\mathbf{A}$ represents the amplification of the desired signal and the diagonal entries of $\mathbf{B}$ represents the amplification of its own signal, leading to self-interference. In [14], the problem of minimizing the relay power consumption subject to SNR constraints on each communication flow is considered in two cases, (1) when $\mathbf{B}$ is the zero matrix, termed *zero-forcing relaying*, and (2) when the diagonal entries of $\mathbf{B}$ are allowed to take non-zero values, termed *network coded relaying*. For zero-forcing relaying, the optimal choice of $\mathbf{A}$ is found analytically for $K = 2$ and suboptimal solutions are found for $K > 2$, and for network coded relaying only suboptimal solutions are found.

4.4. Two-way MIMO Relay Channel

The two-way MIMO relay channel extends the traditional two-way half-duplex relay channel to multiple antennas at each of the terminals. The channel is studied in several works, e.g. [28, 29, 30, 31], and the most studied relaying schemes for this channel are Decode-and-Forward (DF) and Precode-and-Forward (PF) schemes.

The system model consists of two terminals, equipped with M antennas, and the objective is to exchanging messages. There is no direct link between the terminals, and communication is solely done through a relay with M antennas. The channel matrices from terminal i, with $i \in \{1, 2\}$, to the relay are denoted $\mathbf{H}_i$ and reciprocal channels are assumed in the BC phase, i.e. $\mathbf{H}_i^{\mathrm{T}}$ from the relay to terminal i. The time is divided into n channel uses, and the division constant τ divides these channel uses into a MA- and a BC-phase of $n_{\mathrm{MA}} = \lfloor \tau n \rfloor$ and $n_{\mathrm{BC}} = n - n_{\mathrm{MA}}$ channel uses, respectively. In the MA-phase, terminal i transmits the signal $\mathbf{x}_i \in \mathbb{C}^{M \times n_{\mathrm{MA}}}$, which satisfies the power constraint $\frac{1}{n_{\mathrm{MA}}}\mathbf{x}_i^{\mathrm{H}}\mathbf{x}_i \leq P_i$. The relay then receives

$$\mathbf{y}_R = \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 + \mathbf{z} \tag{4.26}$$

where $\mathbf{z} \overset{\mathrm{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_M)$. The relay performs a mapping $g : \mathbb{C}^{M \times n_{\mathrm{MA}}} \to \mathbb{C}^{M \times n_{\mathrm{BC}}}$, from $\mathbf{y}_R$ into $\mathbf{x}_R \in \mathbb{C}^{M \times n_{\mathrm{BC}}}$, satisfying the power constraint $\frac{1}{n_{\mathrm{BC}}}\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R \leq P_R$. The relay broadcasts this signals to the terminals in the BC-phase such that the terminals receive

$$\mathbf{y}_i = \mathbf{H}_i^{\mathrm{T}} g(\mathbf{y}_R) + \mathbf{w}_i \tag{4.27}$$

where $\mathbf{w}_i \overset{\mathrm{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_M)$. The objective is then to choose the mapping g and the covariance matrices of $\mathbf{x}_i$ such that a certain performance metric is optimized. In the following the DF and the PF schemes are briefly summarized.

Decode-and-forward (DF) DF decodes the codewords from each terminal in the end of the MA-phase. In the BC-phase, the relay reencodes the information into a codeword $\mathbf{x}_R \in \mathbb{C}^{M \times n_{\mathrm{BC}}}$, such that, upon reception at the terminals, each terminal can reconstruct the message from the opposite terminal using its contribution to the reencoded signal as side information. In the MA-phase, to decode the codewords from the terminals, the MA-phase forms a MA channel, and hence the codewords can be decoded at rates satisfying

$$R_i \leq \tau \log_2 \left|\mathbf{I}_M + \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^{\mathrm{H}}\right| \qquad \text{for } i \in \{1, 2\} \tag{4.28}$$

$$R_1 + R_2 \leq \tau \log_2 \left|\mathbf{I}_M + \mathbf{H}_1 \mathbf{Q}_1 \mathbf{H}_1^{\mathrm{H}} + \mathbf{H}_2 \mathbf{Q}_2 \mathbf{H}_2^{\mathrm{H}}\right|, \tag{4.29}$$

where $\mathbf{Q}_i$ denotes the covariance matrices of $\mathbf{x}_i$. In the BC-phase, the channel is known as a broadcast channel with side information at the receiver and the channel is studied in [28], where the following rate constraints are found

$$R_i \leq (1-\tau) \log_2 \left|\mathbf{I}_M + \mathbf{H}_i^{\mathrm{T}} \mathbf{Q}_R \mathbf{H}_i^*\right| \qquad \text{for } i \in \{1, 2\}, \tag{4.30}$$

where $\mathbf{Q}_R$ denotes the covariance matrix of the reencoded signal $\mathbf{x}_R$. A performance metric can then be optimized with respect to $\mathbf{Q}_i$, $\mathbf{Q}_R$ and τ in order to obtain the most suitable solution for a given problem.

Finally, let d_i denote the DoF attainable by each user, then d_1 and d_2 must satisfy $d_1 + d_2 \leq \frac{1}{2}M$, and at high SNR the performance is hence limited by the MA-phase in which the relay has to decode the codewords from both terminals. The fraction $\frac{1}{2}$ originates from the fact that $\max(\tau, 1-\tau)$ is $\frac{1}{2}$.

Precode-and-Forward (PF) The precode-and-forward scheme extends the AF scheme to MIMO and is studied in [30, 31, 32]. Upon reception in the end of the MA-phase, the relay left multiplies a relay precode matrix, $\mathbf{G}_R \in \mathbb{C}^{M \times M}$ onto the received signal $\mathbf{y}_R$ to obtain $\mathbf{x}_R$. This implies that $n_{\mathrm{MA}} = n_{\mathrm{BC}}$ and hence $\tau = \frac{1}{2}$. The terminals then receive

$$\mathbf{y}_i = \mathbf{H}_i^{\mathrm{T}} \mathbf{G}_R (\mathbf{H}_1 \mathbf{x}_1 + \mathbf{H}_2 \mathbf{x}_2 + \mathbf{z}) + \mathbf{w}_i. \tag{4.31}$$

Each of the terminals may cancel their own contribution, and in that way the scheme essentially creates a point-to-point MIMO channels in each direction between the two terminals. The achievable rate of the communication flow from i to $\bar{i}$ is hence given by

$$R_i \leq \frac{1}{2} \log_2 \left|\mathbf{I}_M + \mathbf{H}_{\bar{i}}^{\mathrm{T}} \mathbf{G}_R \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^{\mathrm{H}} \mathbf{G}_R^{\mathrm{H}} \mathbf{H}_{\bar{i}}^*\right|, \tag{4.32}$$

where $\frac{1}{2}$ in front originates from the fact the communication is performed in two phases of $\frac{1}{2}n$ channel uses. A performance metric can then be optimized over the covariance matrices $\mathbf{Q}_i$ and the relay precode matrix $\mathbf{G}_R$. This problem is in general non-convex [32], and hence difficult to solve. Various relaxation are performed in literature in order to obtain a solution. In [32], alternate optimization between the covariance matrices and the relay precode matrix is used, i.e. optimization alternates between optimizing with respect to $\mathbf{Q}_i$ with $\mathbf{G}_R$ fixed and with respect $\mathbf{G}_R$ with $\mathbf{Q}_i$ fixed.

The obvious advantage of PF is that the relay mapping is a per-channel use mapping, and hence significantly less complex than DF. Moreover, the behavior of PF at high SNR is also better, since the scheme achieves $\frac{1}{2}M$ DoF.

Figure 4.5: *The MIMO Four-way relay channel. The solid lines denote communication in the MA-phase and the dashed lines denote communication in the BC-phase.*

4.5. Four-way MIMO Relay Channel

In this section, the results for the Four-way MIMO relay channel, one of the three sub networks identified in Chapter 1.2, are summarized. This channel is considered in [3], where a DF relaying scheme is developed. The system model is depicted in Fig. 4.5, and consists of two terminals with two-way communication, two relays and one base station. The aim of the base station is to serve the two terminals, through the relays, assuming that there are no direct links. In the MA-phase of n_{MA} channel uses, the base station and the terminals transmit simultaneously such that the relay is can decode the codewords from the base station and the terminals. In order to avoid interference at the relay, the base station uses zero-forcing. Let $\mathbf{x}_B \in \mathbb{C}^{2M \times n_{\mathrm{MA}}}$ denote the signal transmitted by the base station during the MA-phase, and let $\mathbf{u}_i \in \mathbb{C}^{M \times n_{\mathrm{MA}}}$ be the signal carrying the message to terminal i. To avoid relay $\bar{i}$ to receive a superposition of the signals $\mathbf{u}_{B_1}$ and $\mathbf{u}_{B_2}$, the base station uses a precode matrices $\mathbf{G}_{B_i}$ such that

$$\mathbf{x}_B = \mathbf{G}_{B_1}\mathbf{u}_1 + \mathbf{G}_{B_2}\mathbf{u}_2. \tag{4.33}$$

The precode matrices $\mathbf{G}_{B_i}$ are chosen such that $\mathbf{H}_{\bar{i}\bar{i}}\mathbf{G}_{B_i} = \mathbf{0}$, i.e. any signals, $\mathbf{u}_{B_i}$, multiplied onto the precode matrix $\mathbf{G}_{B_i}$ are cancelled at relay $\bar{i}$. This property is called zero-forcing. Similarly, the terminals transmit $\mathbf{x}_i \in \mathbb{C}^{M \times n_{\mathrm{MA}}}$, with $i \in \{1, 2\}$ such that relay i receives

$$\begin{aligned} \mathbf{y}_{R_i} &= \mathbf{G}_{B_1}\mathbf{u}_1 + \mathbf{G}_{B_2}\mathbf{u}_2 + \mathbf{x}_i + \mathbf{z} && (4.34)\\ &= \mathbf{G}_{B_i}\mathbf{u}_i + \mathbf{x}_i + \mathbf{z}, && (4.35) \end{aligned}$$

where $\mathbf{z} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_M)$ and $\mathbf{G}_{B_{\bar{i}}}\mathbf{u}_{\bar{i}}$ is cancelled by zero-forcing. Consequently, as DF for the two-way MIMO relay channel, the rates have to satisfy the following constraints in the MA-phase

$$\begin{aligned} R_i &\leq \tau \log_2 \left|\mathbf{I}_M + \mathbf{H}_{i\bar{i}}\mathbf{Q}_i\mathbf{H}_{i\bar{i}}^{\mathrm{H}}\right| && (4.36a)\\ R_{B_i} &\leq \tau \log_2 \left|\mathbf{I}_M + \mathbf{H}_{ii}\mathbf{G}_{B_i}\mathbf{Q}_i\mathbf{G}_{B_i}^{\mathrm{H}}\mathbf{H}_{ii}^{\mathrm{H}}\right| && (4.36b)\\ R_i + R_{B_i} &\leq \tau \log_2 \left|\mathbf{I}_M + \mathbf{H}_{i\bar{i}}\mathbf{Q}_i\mathbf{H}_{i\bar{i}}^{\mathrm{H}} + \mathbf{H}_{ii}\mathbf{Q}_i\mathbf{H}_{ii}^{\mathrm{H}}\right| && (4.36c) \end{aligned}$$

where $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ are the covariance matrices of $\mathbf{x}_i$ and $\mathbf{u}_{B_i}$, respectively. In the BC-phase of n_{BC} channel uses, after decoding each of the codewords from the MA-phase, the messages are reencoded into codewords $\mathbf{x}_{R_i} \in \mathbb{C}^{M \times n_{\mathrm{BC}}}$ which are broadcast to the terminals. In [3], it is found that the rates have to satisfy the following rate constraint

$$\begin{aligned} R_i &\leq (1-\tau) \log_2 \left|\mathbf{I}_M + \mathbf{H}_{ii}^{*}\mathbf{Q}_{R_i}\mathbf{H}_{ii}^{\mathrm{T}}\right| && (4.37a)\\ R_{B_i} &\leq (1-\tau) \log_2 \left|\mathbf{I}_M + \mathbf{H}_{i\bar{i}}^{*}\mathbf{Q}_{R_i}\mathbf{H}_{i\bar{i}}^{\mathrm{T}}\right| && (4.37b)\\ R_i + R_{B_i} &\leq (1-\tau) \log_2 \left|\mathbf{I}_M + \mathbf{H}_{ii}^{*}\mathbf{Q}_{R_i}\mathbf{H}_{ii}^{\mathrm{T}} + \mathbf{H}_{i\bar{i}}^{*}\mathbf{Q}_{R_i}\mathbf{H}_{i\bar{i}}^{\mathrm{T}}\right|, && (4.37c) \end{aligned}$$

where $\mathbf{Q}_{R_i}$ are the covariance matrices of $\mathbf{x}_{R_i}$. Essentially, the broadcast phase creates a broadcast channel with side information, but since the base station receives a superposition of the signals from both relays, it is also a MA channel. However, in [3], the rates satisfying (4.37) are shown to be achievable using successive interference cancellation at the base station with arguments similar to those for the MA channel considered in Section 4.2.

A performance metric can then be optimized with respect to the covariance matrices $\mathbf{Q}_i$, $\mathbf{Q}_{B_i}$ and $\mathbf{Q}_{R_i}$ subject to the rate constraint in (4.36) and (4.37) and power constraints.

Using the principles developed, the next chapter considers more advanced relay networks.

CHAPTER 5

Shared Relay Channel

In this chapter, the shared relay, one of the identified sub networks in Section 1.2, is addressed. The system model is depicted in Fig. 5.1, and consists of two base stations, B_1 and B_2, and two terminals, T_1 and T_2, that are paired such that B_i and T_i, with $i \in \{1, 2\}$, wish to attain two-way communication. All nodes are assumed to be half-duplex, there are no direct links between the base stations and the terminals. The relay and the base stations are each equipped with $2M$ antennas and the terminals have M antennas. In this network, the rates of each of the communication flows are denoted by R_i and R_{B_i}, i.e. the rate of the communication flow from terminal i to base station and from base station i to terminal i, respectively. The objective is to find the best achievable rates subject to a certain performance metric. In this thesis, the Weighted Minimum Rate (WMR) performance metric is used. The metric is defined as

$$\mathrm{WMR}(R_{B_1}, R_1, R_{B_2}, R_2) = \min(w_1 R_1, w_{B_1} R_{B_1}, w_2 R_2, w_{B_2}), \tag{5.1}$$

where $(w_1, w_{B_1}, w_2, w_{B_2}) \in \mathbb{R}_+^4$ is the weight tuple. This metric is often used in communications systems because it ensures that all nodes are obtain a certain fraction of the sum of all the rates [3, 14].

The objective of the shared relay considered in this chapter, has the same overall objective as the wireless MIMO switch, which is summarized in Section 4.3. However, the main difference is that the shared relay has another antenna configuration such that the number of antennas at the relay is reduced at the relay at the cost of a higher number of antennas at the base stations. To accomplish that, the idea of alignment is used.

Two different relaying strategies using a MA-phase and a BC-phase are investigated.

- *Precode-and-forward*: Since the base stations have more antennas than the terminals, each base station may spatially align their signal to the corresponding terminal in such a way that the base station and the corresponding terminal transmit in the same subspace. At the relay, the received signal is multiplied by a precoding matrix before broadcasting the received signal back to the nodes. Note that this strategy is similar to the amplify-and-forward scheme described in Section 2.3 with the important difference that the relay may distribute its power among the communication flows and spatial dimensions in an optimal way.
- *Decode-and-forward*: In the MA-phase the transmitted messages from the base stations and the terminals are completely decoded at the relay. Achievable rates for errorfree decoding of all messages are found using standard results for the general

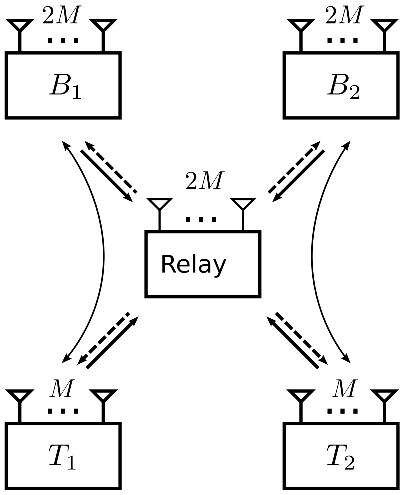


Figure 5.1: *Base stations B_1 and B_2 wish to perform two-way communication with terminal T_1 and T_2, respectively, through a relay. The base stations are equipped with $2M$ antennas, the relay is equipped with $2M$ antennas and the terminals have M antennas. The solid arrows depict communication in the MA-phase and dashed lines depict communication in the BC-phase.*

MA MIMO channel. In the BC-phase the relay broadcasts reencoded versions of the received messages back to each base station-terminal pair using spatial alignment.

In the remaining part of this chapter, the relaying schemes are analyzed and the achievable rates are derived. Based on these results, the WMR is maximized and the schemes are compared.

5.1. System Model

Two base stations, B_1 and B_2 with $2M$ antennas, and two terminals, T_1 and T_2 with M antennas, wish to exchange messages, in pairs, using a single relay with $2M$ antennas and n channel uses. That is, base station i, with $i \in \{1, 2\}$, have the message $m_{B_i} \in \{1, \ldots, 2^{nR_{B_i}}\}$ that is destined to terminal i, while terminal i have a message $m_i \in \{1, \ldots, 2^{nR_i}\}$ which is destined to base station i. The system model is depicted in Fig 5.1. The time is divided into a MA-phase of $n_{\mathrm{MA}} = \lfloor \tau n \rfloor$ channel uses and a BC-phase of $n_{\mathrm{BC}} = n - n_{\mathrm{BC}}$ channel uses, where $\tau \in [0, 1]$.

Initially, the messages m_{B_i} and m_i are mapped into codewords $\mathbf{x}_{B_i}^{n_{\mathrm{MA}}} \in \mathbb{C}^{2M \times n_{\mathrm{MA}}}$ and $\mathbf{x}_i^{n_{\mathrm{MA}}} \in \mathbb{C}^{2M \times n_{\mathrm{MA}}}$ using the encoding functions $f_{B_i}^{(n)} : \{1, \ldots, 2^{nR_{B_i}}\} \to \mathbb{C}^{M \times n_{\mathrm{MA}}}$ and $f_i^{(n)} : \{1, \ldots, 2^{nR_i}\} \to \mathbb{C}^{M \times n_{\mathrm{MA}}}$, i.e. $\mathbf{x}_{B_i}^n = f_{B_i}^{(n)}(m_{B_i})$ and $\mathbf{x}_i^n = f_i^{(n)}(m_i)$. These mappings are subject to the following power constraints

$$\frac{1}{n}\mathrm{tr}\left(\mathbf{x}_{B_i}^{n_{\mathrm{MA}}}\mathbf{x}_{B_i}^{n_{\mathrm{MA}},\mathrm{H}}\right) \leq P_{B_i} \quad \text{and} \quad \frac{1}{n}\mathrm{tr}\left(\mathbf{x}_i^{n_{\mathrm{MA}}}\mathbf{x}_i^{n_{\mathrm{MA}},\mathrm{H}}\right) \leq P_i. \tag{5.2}$$

For instance, with $M = 1$, $n = 6$ and $\tau = \frac{1}{2}$, the terminals transmit vectors of dimension 1×3, the base stations transmit matrices of dimension 2×3 during the MA-phase. In the BC-phase, the relay transmits a matrix of dimension 2×3.

In the MA-phase, the channels from the nodes to the relay are denoted $\mathbf{H}_{B_1}, \mathbf{H}_{B_2} \in \mathbb{C}^{2M\times 2M}$ for the base stations and $\mathbf{H}_1, \mathbf{H}_2 \in \mathbb{C}^{2M\times M}$ for the terminals. Reciprocal channels and full Channel State Information (CSI) are assumed, i.e. the channel matrices from the relay to the nodes in the BC-phase, known by all nodes, are given as $\mathbf{H}_i^{\mathrm{T}}$ and $\mathbf{H}_{B_i}^{\mathrm{T}}$. All channels are assumed to be independent Rayleigh fading channels such that the entries of the channel matrices are i.i.d zero mean Gaussian random variables unit variance. Note that the matrices $\mathbf{H}_{B_1}$, $\mathbf{H}_{B_2}$ and $[\mathbf{H}_1, \mathbf{H}_2]$ are of full rank almost surely since the entries of the matrices are random complex numbers. In this treatment, it is assumed that the channel matrices are of full rank. To simplify notation, the channel matrix $\mathbf{H}$ is introduced as

$$\mathbf{H} = [\mathbf{H}_{B_1}, \mathbf{H}_1, \mathbf{H}_{B_2}, \mathbf{H}_2] \in \mathbb{C}^{2M\times 6M}. \tag{5.3}$$

With these assumptions, the relay receives $\mathbf{y}_R^{n_{\mathrm{BC}}} \in \mathbb{C}^{2M\times n}$ in the MA-phase

$$\mathbf{y}_R^{n_{\mathrm{MA}}} = \mathbf{H} \begin{bmatrix} \mathbf{x}_{B_1}^{n_{\mathrm{MA}}} \\ \mathbf{x}_1^{n_{\mathrm{MA}}} \\ \mathbf{x}_{B_2}^{n_{\mathrm{MA}}} \\ \mathbf{x}_2^{n_{\mathrm{MA}}} \end{bmatrix} + \mathbf{z}^{n_{\mathrm{MA}}}. \tag{5.4}$$

where $\mathbf{z}^{n_{\mathrm{MA}}} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{2M})$. The relay applies a mapping $g^{(n)} : \mathbb{C}^{2M\times n_{\mathrm{MA}}} \to \mathbb{C}^{2M\times n_{\mathrm{BC}}}$ to the received signal $\mathbf{y}_R^n$ and broadcasts the resulting signal back to the base stations and terminals. The relay mapping $g^{(n)}$ has to satisfy the relay power constraint

$$\frac{1}{n}\mathrm{tr}\left(\mathbf{x}_R^{n_{\mathrm{MA}}}\mathbf{x}_R^{n_{\mathrm{MA}},\mathrm{H}}\right) \leq P_R. \tag{5.5}$$

Terminal i and base station i, with $i \in \{1,2\}$, then receives $\mathbf{y}_i^{n_{\mathrm{BC}}} \in \mathbb{C}^{M\times n_{\mathrm{BC}}}$ and $\mathbf{y}_{B_i}^{n_{\mathrm{BC}}} \in \mathbb{C}^{2M\times n_{\mathrm{BC}}}$, respectively, i.e.

$$\mathbf{y}_i^{n_{\mathrm{BC}}} = \mathbf{H}_i^{\mathrm{T}} g(\mathbf{y}_R^{n_{\mathrm{BC}}}) + \mathbf{w}_i^{n_{\mathrm{BC}}} \tag{5.6}$$

$$\mathbf{y}_{B_i}^{n_{\mathrm{BC}}} = \mathbf{H}_{B_i}^{\mathrm{T}} g(\mathbf{y}_R^{n_{\mathrm{BC}}}) + \mathbf{w}_{B_i}^{n_{\mathrm{BC}}} \tag{5.7}$$

where $\mathbf{w}_i^{n_{\mathrm{BC}}} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_M)$ and $\mathbf{w}_{B_i}^{n_{\mathrm{BC}}} \overset{\text{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{2M})$. Finally, using using decoding functions $\hat{m}_i^{(n)} : \mathbf{C}^{2M\times n_{\mathrm{BC}}} \to \{1, \ldots, 2^{nR_i}\}$, the base stations decode the signals $\mathbf{y}_{B_i}^n$ in order to obtain the messages m_i from the terminals. Likewise, the terminals use the decoding functions $\hat{m}_{B_i}^{(n)} : \mathbf{C}^{M\times n_{\mathrm{BC}}} \to \{1, \ldots, 2^{nR_i}\}$ to decode the received signals $\mathbf{y}_i^n$ in order to reconstruct the messages m_{B_i}. In this way, the achievable rates are defined as for the traditional two-way relay channel in Section 2.1.

For PF, the mapping $g^{(n)}$ is simply a per-channel use linear transformation, while for the DF schemes, the mapping is complicated, involving both decoding and reencoding at the relay.

The main concept of the schemes considered in this chapter is alignment. This idea is illustrated through an example depicted in Fig. 5.2. The terminals have channel matrices that are linear independent. Each of the base stations have channel matrices that are invertible, and hence the base stations can transmit in a subspace in such a way that the signals of each communication flow coincides at the relay. In that way, since the relay knows the channel matrices, it can separate the communication flows.

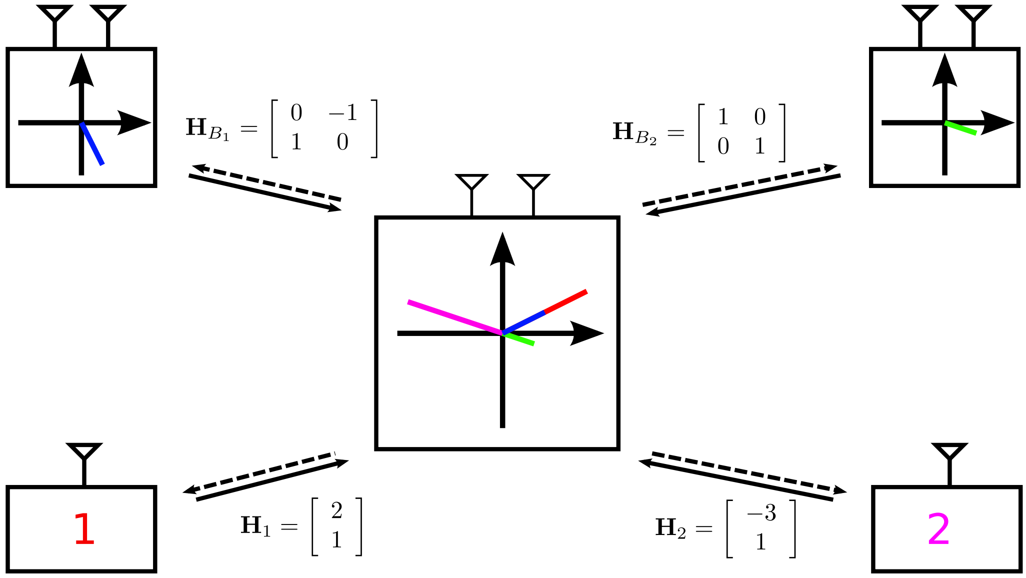


Figure 5.2: *Concept of alignment. Each node is transmitting a signal in the MA-phase. The colors illustrate the contribution at the relay. Since the terminals only have a single antenna, they are only able to transmit a scalar (note that the values and magnitudes of the vectors are not to scale). The base stations can transmit a two-dimensional vector due to their two antennas. However, to align the signal spaces at the relay of each base station to the corresponding terminal, the base stations can only transmit in a one-dimensional subspace. The vector spaces spanned by the channel matrices* $\mathbf{H}_i$ *are termed alignment spaces.*

5.2. Precode-and-Forward Relaying

For the two-way MIMO relay channel and the wireless MIMO switch discussed in Section 4.3 and 4.4, the PF scheme has proved to be very efficient. In this section, PF along with zero-forcing and alignment is used to enable efficient two-way communication for the shared relay channel. Having more antennas at the base stations and the relay than at the terminals enable the base stations to align their signals to the terminals signals in the MA-phase, such that each base station-terminal pair may communicate in different subspaces (not necessarily orthogonal) that are perfectly separable at the relay. Since the base station-terminal pairs are separated, network coded relaying can employed.

Throughout this section, $\tau = \frac{1}{2}$ and only one channel use in the MA- and BC-phase is considered, i.e. $n = 2$ and $n_{\mathrm{MA}} = n_{\mathrm{BC}} = 1$. This is sufficient since PF operates in a channel use-by-channel use fashion. For notational convenience, the superscripts, describing the number of channel uses are dropped such that $\mathbf{x}_i \in \mathbb{C}^{M\times 1}$ and $\mathbf{x}_{B_i} \in \mathbb{C}^{2M\times 1}$, with $i \in \{1,2\}$, are used instead of $\mathbf{x}_i^{n_{\mathrm{MA}}} \in \mathbb{C}^{M\times n_{\mathrm{MA}}}$ and $\mathbf{x}_{B_i}^{n_{\mathrm{MA}}} \in \mathbb{C}^{2M\times n_{\mathrm{MA}}}$, respectively.

The remaining part of this section describes the precoding techniques used at the base stations, terminals and relay.

Base station-terminal precoding In the MA-phase, the base stations transmit their signals through alignment matrices $\mathbf{G}_{B_1}, \mathbf{G}_{B_2} \in \mathbb{C}^{2M\times M}$, introduced shortly, such that

$$\mathbf{x}_{B_1} = \mathbf{G}_{B_1}\mathbf{u}_{B_1} \qquad \text{and} \qquad \mathbf{x}_{B_2} = \mathbf{G}_{B_2}\mathbf{u}_{B_2} \tag{5.8}$$

where the signals $\mathbf{u}_{B_i} \in \mathbb{C}^{M\times 1}$, for $i \in \{1,2\}$, denote the signals within the alignment subspace and carries the messages m_{B_i}. For example, at the base stations, due to the $2M$ antennas, communication from the channel to the relay effectively has $2M$ dimensions, but due to spatial alignment of the signals to the terminals is only of dimension M. The signals $\mathbf{u}_i \in \mathbb{C}^{M\times 1}$, carrying the messages m_i, are similarly defined such that the terminals transmit $\mathbf{x}_i = \mathbf{u}_i$. The matrices $\mathbf{Q}_{B_i}$ and $\mathbf{Q}_i$ denote the positive semidefinite covariance matrices of $\mathbf{u}_{B_i}$ and $\mathbf{u}_i$ that are used to allocate the power at the nodes in an optimal way in order to take advantage of the multiplexing gain that MIMO offers.

The linear relation between $\mathbf{x}_j$ and $\mathbf{u}_j$, with $j \in \{1,2,B_1,B_2\}$ is described by the alignment matrix $\mathbf{G} \in \mathbb{C}^{6M\times 4M}$

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{B_1} & 0 & 0 & 0 \\ 0 & \mathbf{I}_M & 0 & 0 \\ 0 & 0 & \mathbf{G}_{B_2} & 0 \\ 0 & 0 & 0 & \mathbf{I}_M \end{bmatrix}. \tag{5.9}$$

With this notation the transmitted signals are collected in signals $\mathbf{x} \in \mathbb{C}^{6M\times 1}$ and $\mathbf{u} \in \mathbb{C}^{4M\times 1}$ such that

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{B_1} \\ \mathbf{x}_1 \\ \mathbf{x}_{B_2} \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{G} \begin{bmatrix} \mathbf{u}_{B_1} \\ \mathbf{u}_1 \\ \mathbf{u}_{B_2} \\ \mathbf{u}_2 \end{bmatrix} = \mathbf{G}\mathbf{u}. \tag{5.10}$$

Similarly, the covariance matrices are collected in a matrix $\mathbf{Q} \in \mathbb{C}^{4M\times 4M}$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{B_1} & 0 & 0 & 0 \\ 0 & \mathbf{Q}_1 & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{B_2} & 0 \\ 0 & 0 & 0 & \mathbf{Q}_2 \end{bmatrix}, \tag{5.11}$$

such that the covariance matrix of $\mathbf{u}$ is given by $\mathbf{Q}$. The alignment matrices $\mathbf{G}_{B_i}$ are chosen such that the signals from the base stations $\mathbf{x}_{B_1}$ and $\mathbf{x}_{B_2}$ are aligned with the channel matrices of the corresponding terminals, i.e.

$$\mathrm{span}(\mathbf{H}_{B_i}\mathbf{G}_{B_i}) = \mathrm{span}(\mathbf{H}_i) \qquad \text{for } i \in \{1,2\}. \tag{5.12}$$

This implies that base station i and terminal i transmit in the same subspace, termed the alignment subspace of pair i. It is readily that this is successfully achieved when

$$\mathbf{G}_{B_i} = \mathbf{H}_{B_i}^{-1}\mathbf{H}_i, \tag{5.13}$$

for $i \in \{1,2\}$. Note that the alignment matrices $\mathbf{G}_{B_i}$ are not unique. However, the specific choice does not matter since the covariance matrices $\mathbf{Q}_{B_i}$ are also freely chosen subject to the power constraints in (5.2).

Using this alignment scheme, the power constraints in (5.2) become

$$\mathrm{tr}\left(\mathbf{x}_{B_i}\mathbf{x}_{B_i}^{\mathrm{H}}\right) = \mathrm{tr}\left(\mathbf{H}_{B_i}^{-1}\mathbf{H}_i\mathbf{Q}_{B_i}\mathbf{H}_i^{\mathrm{H}}\mathbf{H}_{B_i}^{-\mathrm{H}}\right) \le P_{B_i}, \tag{5.14}$$

and

$$\mathrm{tr}\left(\mathbf{x}_i\mathbf{x}_i^{\mathrm{H}}\right) = \mathrm{tr}\left(\mathbf{Q}_i\right) \le P_i, \tag{5.15}$$

for the base stations and terminals, respectively.

Precoding at the relay In the MA-phase, the relay receives $\mathbf{y}_R \in \mathbb{C}^{2M\times 1}$

$$\begin{aligned}\mathbf{y}_R &= \mathbf{H}_{B_1}\mathbf{G}_{B_1}\mathbf{u}_{B_1} + \mathbf{H}_{B_2}\mathbf{G}_{B_2}\mathbf{u}_{B_2} \\ &\quad + \mathbf{H}_1\mathbf{G}_1\mathbf{u}_1 + \mathbf{H}_2\mathbf{G}_2\mathbf{u}_2 + \mathbf{w} && (5.16)\\ &= [\mathbf{H}_{B_1}, \mathbf{H}_1, \mathbf{H}_{B_2}, \mathbf{H}_2]\,\mathbf{G}\mathbf{u} + \mathbf{z} && (5.17)\\ &= \underbrace{[\mathbf{H}_1, \mathbf{H}_2]}_{\tilde{\mathbf{H}}}\underbrace{\begin{bmatrix}\mathbf{I}_M & \mathbf{I}_M & 0 & 0\\ 0 & 0 & \mathbf{I}_M & \mathbf{I}_M\end{bmatrix}}_{\tilde{\mathbf{I}}_M}\mathbf{u} + \mathbf{z}, && (5.18)\end{aligned}$$

where the matrix $\tilde{\mathbf{H}} = [\mathbf{H}_1, \mathbf{H}_2] \in \mathbb{C}^{2M\times 2M}$ is termed the *effective channel matrix.* Note that, due to the alignment of the two base station-terminal pairs, the signal spaces can be completely separated by left-multiplication of $\tilde{\mathbf{H}}^{-1}$ on the received signal $\mathbf{y}_R$. Upon reception, the relay precodes the received signal $\mathbf{y}_R$ by a precoding matrix $\mathbf{G}_R \in \mathbb{C}^{2M\times 2M}$ before broadcast in the BC-phase, and hence the nodes receive

$$\left[\mathbf{y}_{B_1}^{\mathrm{T}}, \mathbf{y}_1^{\mathrm{T}}, \mathbf{y}_{B_2}^{\mathrm{T}}, \mathbf{y}_2^{\mathrm{T}}\right]^{\mathrm{T}} = \mathbf{y} = \mathbf{H}^{\mathrm{T}}\mathbf{G}_R\tilde{\mathbf{H}}\tilde{\mathbf{I}}_M\mathbf{u} + \mathbf{H}^{\mathrm{T}}\mathbf{G}_R\mathbf{z} + \mathbf{w}, \qquad (5.19)$$

where $\mathbf{w} = \left[\mathbf{w}_{B_1}^{\mathrm{T}}, \mathbf{w}_1^{\mathrm{T}}, \mathbf{w}_{B_2}^{\mathrm{T}}, \mathbf{w}_2^{\mathrm{T}}\right]^{\mathrm{T}}$. Note that base station receive $\mathbf{y}_{B_i}$ which is of $2M$ dimensions. In order to extract only the spatial dimensions from the alignment subspace, where the signal of terminal i, $\mathbf{u}_i$, resides, base station i use the receive matrix $\mathbf{G}_{B_i}^{\mathrm{T}}$, matching the alignment matrix $\mathbf{G}_{B_i}^{\mathrm{T}}$. The resulting signals, after left-multiplication by the receive matrices, are denoted $\tilde{\mathbf{y}}_{B_i} \in \mathbb{C}^{M\times 1}$, i.e. $\tilde{\mathbf{y}}_{B_i} = \mathbf{G}_{B_i}^{\mathrm{T}}\mathbf{y}_{B_i}$. By similarly defining $\tilde{\mathbf{y}}_i = \mathbf{y}_i \in \mathbb{C}^{M\times 1}$, (5.19) can be expressed as

$$\begin{aligned}[\tilde{\mathbf{y}}_{B_1}^{\mathrm{T}}, \tilde{\mathbf{y}}_1^{\mathrm{T}}, \tilde{\mathbf{y}}_{B_2}^{\mathrm{T}}, \tilde{\mathbf{y}}_2^{\mathrm{T}}]^{\mathrm{T}} &= \tilde{\mathbf{y}} && (5.20)\\ &= \mathbf{G}^{\mathrm{T}}(\mathbf{H}^{\mathrm{T}}\mathbf{G}_R\tilde{\mathbf{H}}\tilde{\mathbf{I}}_M\mathbf{u} + \mathbf{H}^{\mathrm{T}}\mathbf{G}_R\mathbf{z} + \mathbf{w}) && (5.21)\\ &= \tilde{\mathbf{I}}_M^{\mathrm{T}}\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{G}_R\tilde{\mathbf{H}}\tilde{\mathbf{I}}_M\mathbf{u} + \tilde{\mathbf{I}}_M^{\mathrm{T}}\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{G}_R\mathbf{z} + \mathbf{G}^{\mathrm{T}}\mathbf{w}. && (5.22)\end{aligned}$$

The relay precoding matrix $\mathbf{G}_R$ is designed as a zero-forcing precoder as in [14], that is, each base station-terminal pair do not receive any interference from the other pair. Also, alignment enables the relay to treat the two communication flow pairs separately. To this end, $\mathbf{G}_R$ needs to be chosen such that $\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{G}_R\tilde{\mathbf{H}}$ has a block diagonal structure, i.e. it has to be on the following form

$$\begin{aligned}\mathbf{G}_R &= \tilde{\mathbf{H}}^{-\mathrm{T}}\begin{bmatrix}\mathbf{A}_1 & 0\\ 0 & \mathbf{A}_2\end{bmatrix}\tilde{\mathbf{H}}^{-1}, && (5.23)\\ &= \tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{A}\tilde{\mathbf{H}}^{-1}, && (5.24)\end{aligned}$$

where $\mathbf{A}_i \in \mathbb{C}^{M\times M}$ are amplification matrices and are analogous to the amplification factors for the AF scheme. The amplification matrices can be chosen arbitrarily as long as the power constraint in (5.5) is satisfied. Substituting (5.24) into (5.22) yields

$$\begin{aligned}\tilde{\mathbf{y}} &= \tilde{\mathbf{I}}_M^{\mathrm{T}}\mathbf{A}\tilde{\mathbf{I}}_M\mathbf{u} + \tilde{\mathbf{I}}_M^{\mathrm{T}}\mathbf{A}\tilde{\mathbf{H}}^{-1}\mathbf{z} + \mathbf{G}^{\mathrm{T}}\mathbf{w} && (5.25)\\ &= \underbrace{\begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_1 & & \\ \mathbf{A}_1 & \mathbf{A}_1 & & \\ & & \mathbf{A}_2 & \mathbf{A}_2\\ & & \mathbf{A}_2 & \mathbf{A}_2\end{bmatrix}}_{\tilde{\mathbf{A}}}\mathbf{u} + \begin{bmatrix}\mathbf{A}_1 & \\ \mathbf{A}_1 & \\ & \mathbf{A}_2\\ & \mathbf{A}_2\end{bmatrix}\tilde{\mathbf{H}}^{-1}\mathbf{z} + \mathbf{G}^{\mathrm{T}}\mathbf{w}. && (5.26)\end{aligned}$$

The diagonal matrices of $\tilde{\mathbf{A}}$ represent self-interference and the off-diagonal matrices describes the signal of interest. Note that there is no more than one off-diagonal matrix for each row, which is due to alignment at the base stations and zero-forcing at the relay. Since each node knows the signal it transmitted in the MA-phase, self-interference may be cancelled, and hence, after cancellation of self-interference, the nodes receive

$$\tilde{\mathbf{y}}'_i = \mathbf{A}_i \mathbf{u}_{B_i} + \mathbf{A}_i \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \mathbf{z} + \mathbf{w}_i \tag{5.27a}$$

$$\tilde{\mathbf{y}}'_{B_i} = \mathbf{A}_i \mathbf{u}_i + \mathbf{A}_i \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \mathbf{z} + \mathbf{H}_i^{\mathrm{T}} \mathbf{H}_{B_i}^{-\mathrm{T}} \mathbf{w}_i, \tag{5.27b}$$

where $\mathbf{B}_1 = [\mathbf{I}_M, \mathbf{0}]$ and $\mathbf{B}_2 = [\mathbf{0}, \mathbf{I}_M]$ are defined to select the upper and lower half of the matrix $\tilde{\mathbf{H}}^{-1}$. In this representation, it is seen that two-way communication is successfully achieved. Observe that each communication flow is a MIMO channel with $\frac{1}{2}M$ DoF, which is also the maximum achievable. This is because each communication flow is limited by the channels between the relay and the terminals with only M antennas. Note that the fraction $\frac{1}{2}$ originates from the fact that computation of rates are done over both the MA- and BC-phase.

Achievable rates From the received signals in (5.27), the achievable rates are derived given the signal covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ along with the relay amplification matrices $\mathbf{A}_i$. The role of the signal covariance matrices are to allocate the available power among the strongest spatial dimensions in order to maximize the capacity of each of the MIMO channels. The interpretation of the amplification matrices $\mathbf{A}_i$ is similar to the amplifying constant for the traditional AF scheme; larger values in the amplification matrices increase the amplification of the signal of interest in the MA-phase, but also amplifies the noise. Compared to AF, the use of amplification matrices also allows the possibility to amplify specific eigenmodes more than others. For example, suppose the base stations chooses to use most of its power to transmit in only few of its eigenmodes, i.e. it only uses of subset of the spatial dimensions available, then the amplification matrix can be chosen such that only these eigenmodes are amplified, and hence the noise in the remaining eigenchannels is suppressed.

The channels induced using the PF scheme, given in (5.27), are traditional MIMO channel described in Section 4.1, and hence the achievable rate region is characterized by

$$\begin{cases} R_{B_i} & \leq \frac{1}{2} \log_2 \left| \mathbf{I} + \mathbf{C}_{\mathbf{n},i}^{-1} \mathbf{A}_i \mathbf{Q}_{B_i} \mathbf{A}_i^{\mathrm{H}} \right| \\ R_i & \leq \frac{1}{2} \log_2 \left| \mathbf{I} + \mathbf{C}_{\mathbf{n},B_i}^{-1} \mathbf{A}_i \mathbf{Q}_i \mathbf{A}_i^{\mathrm{H}} \right| \end{cases}, \tag{5.28}$$

where the fraction $\frac{1}{2}$ stems from the fact that communication is done in a MA- and a BC-phase. The matrices $\mathbf{C}_{\mathbf{n},i}$ and $\mathbf{C}_{\mathbf{n},B_i}$ denote the noise covariance matrices of $\tilde{\mathbf{y}}'_i$ and $\tilde{\mathbf{y}}'_{B_i}$ in (5.27), respectively, and are given as

$$\mathbf{C}_{\mathbf{n},i} = \mathbf{A}_i \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\mathrm{H}} \mathbf{B}_i^{\mathrm{H}} \mathbf{A}_i^{\mathrm{H}} + \mathbf{I}_M \tag{5.29}$$

$$\mathbf{C}_{\mathbf{n},B_i} = \mathbf{A}_i \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\mathrm{H}} \mathbf{B}_i^{\mathrm{H}} \mathbf{A}_i^{\mathrm{H}} + \mathbf{H}_i^{\mathrm{T}} \mathbf{H}_{B_i}^{-\mathrm{T}} \mathbf{H}_{B_i}^{-*} \mathbf{H}_i^{*}. \tag{5.30}$$

The aim is to choose the amplification matrices $\mathbf{A}_i$ and covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ such that the achievable rates are optimized for each communication flow according the metric WMR defined in (5.1). The general optimization problem, maximizing the WMR can then be formulated as

$$\begin{aligned}
\underset{\substack{\mathbf{A}_1,\mathbf{A}_2,\mathbf{Q}_i,\mathbf{Q}_{B_i}\\ R_1,R_{B_1},R_2,R_{B_2}}}{\text{maximize}} \quad & \min(w_1 R_1, w_2 R_2, w_{B_1} R_{B_1}, w_{B_2} R_{B_2}) && (5.31\text{a})\\
\text{subject to} \quad & (5.28), \mathbb{E}\left[\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R\right] \le P_R && (5.31\text{b})\\
& \operatorname{tr}(\mathbf{Q}_i) \le P_i, \mathbf{Q}_i \succeq 0 && (5.31\text{c})\\
& \operatorname{tr}\left(\mathbf{H}_{B_i}^{-1}\mathbf{H}_i\mathbf{Q}_{B_i}\mathbf{H}_i^{\mathrm{H}}\mathbf{H}_{B_i}^{-\mathrm{H}}\right) \le P_{B_i}, \mathbf{Q}_{B_i} \succeq 0 && (5.31\text{d})
\end{aligned}$$

This problem is in general non-convex and difficult to solve since the matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ are involved in the expressions of $\mathbf{C}_{\mathbf{n},i}$ and $\mathbf{C}_{\mathbf{n},B_i}$ and the relay power consumption $\mathbb{E}\left[\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R\right]$.

Even for the two-way MIMO relay channel, described in Section 4.4, joint optimization over the covariance matrices and the relay precode matrix has proved to be a difficult problem [32]. In [32], the optimizing problem is relaxed by alternately optimizing the covariance matrices and relay precode matrices. The optimal covariance matrices are then found using standard convex optimization tools, and a suboptimal relay precode matrix is found using a gradient descent method. In the following, similar relaxations are introduced for the problem in (5.31).

5.2.1 Suboptimal Power Allocation Scheme

As joint optimization over $\mathbf{Q}_i$, $\mathbf{Q}_{B_i}$ and $\mathbf{A}_i$ is non-convex, a simplified scheme is proposed that enables computation of suboptimal covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ and amplification matrices $\mathbf{A}_i$ through convex optimization. In this thesis, the following relaxations are done.

1. Alternate optimization, as in [32], is used to avoid joint optimization with respect to covariance matrices and amplification matrices. In that way, the optimization problem is solved by alternately optimizing with respect to the covariance matrices and the amplification matrices.

2. In order to make the optimization problem with respect to the amplification matrices tractable, the following highly simplified amplification matrices are used

$$\mathbf{A}_i = a_i \mathbf{I}_M \qquad \text{for } i \in \{1,2\}, \tag{5.32}$$

   where the amplification constants $a_i \in \mathbb{C}$, for $i \in \{1,2\}$, are used to distribute the power at the relay in an optimal manner.

Alternate optimization ensures that a good solutions is found [32] and the simplified amplification matrices makes the optimization problem with respect to the amplification constants a_i analytically tractable. For notational convenience, let the amplification vector $\mathbf{a}$ denote $[a_1, a_2]^{\mathrm{T}}$. It is important to note that certain performance degradations, due to the use alternate optimization and the simplified choice of the amplification matrices $\mathbf{A}_i$, are expected when $M \ge 2$.

In the following, the relay power constraint in (5.5) is expressed in terms of the covariance matrices and the amplification matrices, which is needed for the optimization problems. Then the necessary optimization problems are formulated and an analytical solution for the optimal amplification vector $\mathbf{a}$ with fixed $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ is found. Similarly, a waterfilling-type algorithm is developed to optimize with respect to $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ with $\mathbf{a}$ fixed.

Power constraint at the relay With the simplified choice of amplification matrices $\mathbf{A}_i$, the relay power constraint $\mathbb{E}\left[\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R\right] \leq P_R$ can be simplified in terms of the covariance matrices $\mathbf{Q}_i$, $\mathbf{Q}_{B_i}$ and the amplification vector $\mathbf{a}$. Since the optimization is performed alternately, the relay power consumption $\mathbb{E}\left[\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R\right]$ is first derived in a simple form in terms of the amplification vector $\mathbf{a}$, and subsequently in terms of the covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$.

Using the definition of the trace and its properties, the following manipulations can be performed

$$\begin{aligned}
\mathbb{E}\left[\mathbf{x}_R^{\mathrm{H}}\mathbf{x}_R\right] &= \mathbb{E}\left[\mathbf{u}^{\mathrm{H}}\tilde{\mathbf{I}}_M^{\mathrm{H}}\tilde{\mathbf{H}}^{\mathrm{H}}\mathbf{G}_R^{\mathrm{H}}\mathbf{G}_R\tilde{\mathbf{H}}\tilde{\mathbf{I}}_M\mathbf{u} + \mathbf{z}^{\mathrm{H}}\mathbf{G}_R^{\mathrm{H}}\mathbf{G}_R\mathbf{z}\right] && (5.33)\\
&= \operatorname{tr}\left(\mathbf{G}_R\tilde{\mathbf{H}}\tilde{\mathbf{I}}_M\mathbf{Q}\tilde{\mathbf{I}}_M^{\mathrm{H}}\tilde{\mathbf{H}}^{\mathrm{H}}\mathbf{G}_R^{\mathrm{H}} + \mathbf{G}_R\mathbf{G}_R^{\mathrm{H}}\right) && (5.34)\\
&= \operatorname{tr}\left(\mathbf{G}_R\left(\tilde{\mathbf{H}}\tilde{\mathbf{Q}}\tilde{\mathbf{H}}^{\mathrm{H}} + \mathbf{I}\right)\mathbf{G}_R^{\mathrm{H}}\right) && (5.35)\\
&= \operatorname{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{A}\left(\tilde{\mathbf{Q}} + \tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\right)\mathbf{A}^{\mathrm{H}}(\tilde{\mathbf{H}}^{*})^{-1}\right) && (5.36)\\
&= \operatorname{tr}\left(\mathbf{A}\mathbf{K}\mathbf{A}^{\mathrm{H}}(\tilde{\mathbf{H}}^{*})^{-1}\tilde{\mathbf{H}}^{-\mathrm{T}}\right) && (5.37)\\
&= \tilde{\mathbf{a}}^{\mathrm{H}}\tilde{\mathbf{S}}\tilde{\mathbf{a}} \leq P_R && (5.38)
\end{aligned}$$

where $\tilde{\mathbf{a}}$ is the column matrix of the diagonal entries of $\mathbf{A}$, $\tilde{\mathbf{Q}}$ denotes $\tilde{\mathbf{I}}_M\mathbf{Q}\tilde{\mathbf{I}}_M^{\mathrm{H}}$, $\mathbf{K}$ denotes $\tilde{\mathbf{Q}} + \tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}$ and the entries of $\tilde{\mathbf{S}} \in \mathbb{C}^{2M\times 2M}$ are given as [14]

$$\tilde{\mathbf{S}}_{ij} = \mathbf{K}_{ji}[(\tilde{\mathbf{H}}^{*})^{-1}\tilde{\mathbf{H}}^{-\mathrm{T}}]_{ij}. \tag{5.39}$$

The quadratic form (5.38) can be further simplified to

$$\mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a} \leq P_R \tag{5.40}$$

where $\mathbf{a} = [a_1, a_2]^{\mathrm{T}}$, $\mathbf{S} = \begin{bmatrix} \mathbf{1}_M^{\mathrm{H}}\mathbf{B}_1\tilde{\mathbf{S}}\mathbf{B}_1^{\mathrm{H}}\mathbf{1}_M & \mathbf{1}_M^{\mathrm{H}}\mathbf{B}_1\tilde{\mathbf{S}}\mathbf{B}_2^{\mathrm{H}}\mathbf{1}_M \\ \mathbf{1}_M^{\mathrm{H}}\mathbf{B}_2\tilde{\mathbf{S}}\mathbf{B}_1^{\mathrm{H}}\mathbf{1}_M & \mathbf{1}_M^{\mathrm{H}}\mathbf{B}_2\tilde{\mathbf{S}}\mathbf{B}_2^{\mathrm{H}}\mathbf{1}_M \end{bmatrix} \in \mathbb{C}^{2\times 2}$ and $\mathbf{1}_M$ is the column matrix in $\mathbb{R}^{M\times 1}$ with all entries equal to '1'. Note that $\tilde{\mathbf{a}} = \mathbf{a}$ in the special case where $M = 1$. It is later seen that the optimization problem in terms of the amplification vector $\mathbf{a}$ is particularly simple because the power constraint at the relay can be represented as the simple quadratic form in (5.40).

Alternatively, to express the relay power consumption in terms of the covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$, first note that $\tilde{\mathbf{Q}} = \sum_{i=1}^{2}\mathbf{B}_i^{\mathrm{H}}(\mathbf{Q}_i + \mathbf{Q}_{B_i})\mathbf{B}_i$. Using this identity, (5.36) can be equivalently expressed as

$$\begin{aligned}
\operatorname{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{A}\tilde{\mathbf{Q}}\mathbf{A}^{\mathrm{H}}(\tilde{\mathbf{H}}^{*})^{-1}\right) &= \operatorname{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{A}\left(\sum_{i=1}^{2}\mathbf{B}_i^{\mathrm{H}}\mathbf{B}_i\tilde{\mathbf{Q}}\mathbf{B}_i^{\mathrm{H}}\mathbf{B}_i\right)\mathbf{A}^{\mathrm{H}}(\tilde{\mathbf{H}}^{*})^{-1}\right) && (5.41)\\
&= \sum_{i=1}^{2}\operatorname{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{A}\mathbf{B}_i^{\mathrm{H}}(\mathbf{Q}_{B_i} + \mathbf{Q}_i)\mathbf{B}_i\mathbf{A}^{\mathrm{H}}(\tilde{\mathbf{H}}^{*})^{-1}\right) && (5.42)\\
&\leq P_R - \operatorname{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{A}\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{A}^{\mathrm{H}}(\tilde{\mathbf{H}}^{*})^{-1}\right). && (5.43)
\end{aligned}$$

From these representations of the relay power constraint, the next sections study the optimization problem with respect to the covariance matrices with $\mathbf{a}$ fixed, and the optimization problem with respect to $\mathbf{a}$ with the covariance matrices fixed.

Optimization with respect to $\mathbf{Q}_i$ and $\mathbf{Q}_{\mathbf{B}_\mathrm{i}}$

The optimization with respect to the covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ with fixed $\mathbf{a}$ is formulated as

$$\begin{aligned}
&\underset{\substack{\mathbf{Q}_i\succeq 0,\mathbf{Q}_{B_i}\succeq 0\\ R_1,R_{B_1},R_2,R_{B_2}}}{\text{maximize}} && \min(w_1R_1, w_2R_2, w_{B_1}R_{B_1}, w_{B_2}R_{B_2}) && (5.44a)\\
&\text{subject to} && (5.28), (5.42)-(5.43), \operatorname{tr}(\mathbf{Q}_i) \le P_i, && (5.44b)\\
& && \operatorname{tr}\left(\mathbf{H}_{B_i}^{-1}\mathbf{H}_i\mathbf{Q}_{B_i}\mathbf{H}_i^{\mathrm{H}}\mathbf{H}_{B_i}^{-\mathrm{H}}\right) \le P_{B_i}. && (5.44c)
\end{aligned}$$

The problem in (5.44) is convex and may be solved using standard convex optimization tools e.g. CVX or SeDuMi. However, as for power allocation on the traditional MIMO channel, the problem is efficiently solved using an iterative waterfilling-type algorithm which is developed in the following.

First note that the problem in (5.44) may be generalized to the following optimization problem with appropriate substitutions

$$\begin{aligned}
&\underset{\mathbf{Q}_1\succeq 0,\ldots,\mathbf{Q}_L\succeq 0}{\text{maximize}} && \min_{l\in\{1,\ldots,L\}} w_l \log_2\left|\mathbf{I}+\mathbf{H}_l\mathbf{Q}_l\mathbf{H}_l^{\mathrm{H}}\right| && (5.45a)\\
&\text{subject to} && \operatorname{tr}(\mathbf{Q}_l\mathbf{T}_l) \le P_l && (5.45b)\\
& && \sum_{l=1}^{L}\operatorname{tr}(\mathbf{Q}_l\mathbf{T}_{c,l}) \le P_c, && (5.45c)
\end{aligned}$$

where $\mathbf{H}_l \in \mathbb{C}^{M\times M}$ is of full rank, and $\mathbf{T}_l$ and $\mathbf{T}_{c,l}$ are positive definite matrices in the vector space $\mathbb{C}^{M\times M}$.

Remark 5.1. *The problem in* (5.45) *in general represents a system model in which L MIMO transmitters, that does not interfere and have individual and common constraints on the covariance matrices, wish to transmit to L different receivers with respect to the WMR metric. An example of a system in which this optimization problem is of relevance is in a wireless device with multiple transmitters at different frequency bands (to avoid co-channel interference) under individual and common power constraints.*

This problem is solved in an iterative manner using Algorithm 1. The following lemma guarantees that the algorithm converges to the global optimum.

Lemma 5.1. *The algorithm in Algorithm 1 converges to the solution of the optimization problem in* (5.45).

Proof. The problem in (5.45) is readily seen to be convex, and hence a global optimum to the problem exists.

In iteration i, the channel index $l_{\min}^{(i)}$ of the channel with the smallest weighted rate in the objective function in (5.45a), termed the bottleneck channel, is found, i.e.

$$l_{\min}^{(i)} = \arg\min_l w_l \log_2\left|\mathbf{I}+\mathbf{H}_l\mathbf{Q}_l^{(i)}\mathbf{H}_l^{\mathrm{H}}\right|. \tag{5.46}$$

If there are more than one solution to the optimization problem in (5.46), $l_{\min}^{(i)}$ is chosen to be one of them at random. Define $t_{\min}^{(i)} = w_{\min}\log_2\left|\mathbf{I}+\mathbf{H}_{l_{\min}}\mathbf{Q}_{l_{\min}}^{(i)}\mathbf{H}_{l_{\min}}^{\mathrm{H}}\right|$, which denotes

Algorithm 1 Solves the optimization problem in (5.45).

input: Channel matrix $\mathbf{H}_l$, power constraints $\mathbf{T}_j$ and $\mathbf{T}_{c,l}$ along with P_l and P_c for $l \in \{1, \ldots, L\}$.
output: $\mathbf{Q}_i$
initialize: $\mathbf{Q}_i = \mathbf{0}$ for $i \in \{1, \ldots, L\}$.
repeat
  $l_{\min}^{(i)} \leftarrow \arg\min_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i)} \mathbf{H}_l^{\mathrm{H}}\right|$.
  $t_{\min}^{(i)} \leftarrow \log_2 \left|\mathbf{I} + \mathbf{H}_{l_{\min}} \mathbf{Q}_{l_{\min}}^{(i)} \mathbf{H}_{l_{\min}}^{\mathrm{H}}\right|$.
  Find $\delta_l^{(i)}$ satisfying $t_{\min}^{(i)} = \log_2 \left|\mathbf{I} + (1-\delta_l^{(i)}) \mathbf{H}_l \mathbf{Q}_l^{(i)} \mathbf{H}_l^{\mathrm{H}}\right|$ for $l \neq l_{\min}$.
  $\mathbf{Q}_l^{(i+1)} \leftarrow \left(1 - \frac{\delta_l}{2}\right) \mathbf{Q}_l^{(i)}$ for $l \neq l_{\min}$.
  Solve the optimization problem in (5.49) to obtain $\mathbf{Q}_{l_{\min}}^{(i)}$ using Algorithm 2.
until convergence

the achievable rate of the bottleneck channel, i.e. the $l_{\min}$-th channel. Moreover, $t_l^{(i)}$ is defined as $w_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i)} \mathbf{H}_l^{\mathrm{H}}\right|$ and denotes the achievable rate of the l-th channel.

Let $\mathcal{L}_{\min}^{(i)}$ denote the set of all bottleneck channels to the optimization problem, i.e. $\mathcal{L}_{\min}^{(i)} = \left\{l \ : \ l_{\min}^{(i)} = t_l^{(i)}\right\}$, and let $|\mathcal{L}_{\min}^{(i)}|$ denote the cardinality of the set.

The constants $\delta_l^{(i)} \in [0,1]$ are then defined by the solutions of the equations

$$t_{\min}^{(i)} = w_l \log_2 \left|\mathbf{I} + (1-\delta_l^{(i)}) \mathbf{H}_l \mathbf{Q}_l^{(i)} \mathbf{H}_l^{\mathrm{H}}\right| \tag{5.47}$$

for $l \notin \mathcal{L}_{\min}^{(i)}$. Since $\log_2 |\mathbf{I} + a\mathbf{X}|$ is monotonically non-decreasing in a, a solution to (5.47) indeed exists for each $l \notin \mathcal{L}_{\min}^{(i)}$. The following updates are then performed for $l \neq l_{\min}^{(i)}$ using the constants $\delta_l^{(i)}$

$$\mathbf{Q}_l^{(i+1)} = \begin{cases} \left(1 - \frac{\delta_l^{(i)}}{2}\right) \mathbf{Q}_l^{(i)} & \text{for } l \notin \mathcal{L}_{\min}^{(i)} \\ \mathbf{Q}_l^{(i)} & \text{for } l \in \mathcal{L}_{\min}^{(i)} \text{ and } l \neq l_{\min} \end{cases} \tag{5.48}$$

This update, essentially ensures that the common power constraint is not satisfied with equality, such that the common power constraint in (5.45c) is strict. This implies that rate of the $l_{\min}$-th channel may be increased. Note that, $w_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i+1)} \mathbf{H}_l^{\mathrm{H}}\right| > t_{\min}^{(i)}$ for $l \notin \mathcal{L}_{\min}^{(i)}$, i.e. $l_{\min}$-th channel, is still a bottleneck after the update. Finally, the following optimization problem is solved to obtain $\mathbf{Q}_{l_{\min}}^{(i+1)}$ and $t_{l_{\min}}^{(i+1)}$.

$$\underset{\mathbf{Q}_{l_{\min}}^{(i+1)} \succeq 0}{\text{maximize}} \quad t_{l_{\min}}^{(i+1)} = \log_2 \left|\mathbf{I} + \mathbf{H}_{l_{\min}} \mathbf{Q}_{l_{\min}}^{(i+1)} \mathbf{H}_{l_{\min}}^{\mathrm{H}}\right| \tag{5.49a}$$

$$\text{subject to} \quad \operatorname{tr}\left(\mathbf{Q}_{l_{\min}}^{(i+1)} \mathbf{T}_{l_{\min}}\right) \leq P_{l_{\min}} \tag{5.49b}$$

$$\operatorname{tr}\left(\mathbf{Q}_{l_{\min}}^{(i+1)} \mathbf{T}_{c,l_{\min}}\right) \leq P_c - \sum_{l=1, l \neq l_{\min}}^{L} \operatorname{tr}\left(\mathbf{Q}_l^{(i+1)} \mathbf{T}_{c,l}\right) = \zeta_c, \tag{5.49c}$$

where $\zeta_c = P_c - \sum_{l=1, l\neq l_{\min}}^{L} \mathrm{tr}\left(\mathbf{Q}_l^{(i+1)}\mathbf{T}_{c,l}\right)$ is introduced for notational convenience. Due to the power reduction step in (5.48), ζ_c is strictly larger than zero if $|\mathcal{L}_{\min}^{(i)}| < L$. The following exhaustive list of cases is then considered.

1. $|\mathcal{L}_{min}^{(i)}| = 1$ *and* $t_{l_{min}}^{(i+1)} > t_{l_{min}}^{(i)}$: Then the following inequality is satisfied

$$t_{l_{\min}}^{(i)} = \min_l w_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i)} \mathbf{H}_l^{\mathrm{H}}\right| < \min_l w_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i+1)} \mathbf{H}_l^{\mathrm{H}}\right| = t_{\min}^{(i+1)}, \quad (5.50)$$

   and hence the objective function of (5.45) is successfully increased towards the optimum from iteration i to $i+1$.

2. $1 < |\mathcal{L}_{min}^{(i)}| \leq L$ *and* $t_{l_{min}}^{(i+1)} > t_{l_{min}}^{(i)}$: Since $1 < |\mathcal{L}_{\min}^{(i)}|$, we have

$$t_{l_{\min}}^{(i)} = \min_l w_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i+1)} \mathbf{H}_l^{\mathrm{H}}\right| = t_{\min}^{(i+1)}. \quad (5.51)$$

   However, $t_{l_{\min}}^{(i+1)} > t_{l_{\min}}^{(i)}$ implies that the number of bottleneck channels in iteration $(i+1)$ is decreased by one, i.e. $|\mathcal{L}_{\min}^{(i+1)}| = |\mathcal{L}_{\min}^{(i)}| - 1$.

3. $|\mathcal{L}_{min}^{(i)}| \leq L$ *and* $t_{l_{min}}^{(i+1)} = t_{l_{min}}^{(i)}$: Since the inequality constraint (5.49c) is strict (due to the power reduction step in (5.48)), the problem is independent of the remaining covariance matrices $\mathbf{Q}_l$ for $l \neq l_{\min}$, and hence $t_{l_{\min}}^{(i)} = t_{l_{\min}}^{(i+1)}$ is an global optimum of the optimization problem in (5.45).

Since either the objective function $\min_l w_l \log_2 \left|\mathbf{I} + \mathbf{H}_l \mathbf{Q}_l^{(i)} \mathbf{H}_l^{\mathrm{H}}\right|$ is increased or the cardinality of the set of bottleneck channels decreased, eventually, the algorithm converges to the optimum of (5.45). □

Note that, a unique optimum can not be guaranteed. To see this, observe that if P_c in (5.45c) tends to infinity, the optimization problem can be solved simply by finding the capacity of each of the L channels subject to the individual power constraint. The solution of the optimization problem in (5.45) is then given by the capacity of the weakest channel. However, the covariance matrices of the remaining channels may be chosen arbitrarily as long as their resulting rates are higher than the capacity of weakest channel.

As part of Algorithm 1, the following optimization problem is repeatedly solved in each iteration

$$\begin{aligned}
&\underset{\mathbf{Q}\succeq 0}{\text{maximize}} && \log_2 \left|\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^{\mathrm{H}}\right| && (5.52a)\\
&\text{subject to} && \mathrm{tr}\left(\mathbf{Q}\mathbf{T}\right) \leq P && (5.52b)\\
& && \mathrm{tr}\left(\mathbf{Q}\mathbf{T}_c\right) \leq \zeta_c. && (5.52c)
\end{aligned}$$

where $\mathbf{H} \in \mathbb{C}^{M\times M}$ is full rank and $\mathbf{T}$ and $\mathbf{T}_c$ are positive definite matrices in the vector space $\mathbb{C}^{M\times M}$. In the following lemma, it is shown that this optimization problem can be efficiently solved using the traditional waterfilling algorithm.

Lemma 5.2. *Algorithm 2 converges to the solution of the optimization problem in* (5.52).

Algorithm 2 Waterfilling-type Algorithm for Computation of the Achievable Rate with Two Power Constraints (using bisection)

input: Channel matrix $\mathbf{H}$, power constraints $\mathbf{T}$ and $\mathbf{T}_c$ along with P and ζ_c.
output: $\mathbf{Q}$
initialize: $\lambda_{\min} = 0$ and $\lambda_{\max} = 1$.
repeat
  $\lambda \leftarrow \frac{\lambda_{\min}+\lambda_{\max}}{2}$.
  − Compute the Cholesky factorization $\mathbf{R}_\lambda \mathbf{R}_\lambda^{\mathrm{H}} = (\lambda \mathbf{T} + (1-\lambda)\mathbf{T}_c)^{-1}$.
  − Perform traditional waterfilling with the channel matrix $(\mathbf{H}\mathbf{R}_\lambda)$ and the power constraint $\lambda P + (1-\lambda)P_c$ to obtain the covariance matrix $\tilde{\mathbf{Q}}_\lambda$.
  − Compute the covariance matrix $\mathbf{Q}_\lambda = \mathbf{R}_\lambda^{-1}\tilde{\mathbf{Q}}\mathbf{R}_\lambda^{-1,\mathrm{H}}$.
  if $\operatorname{tr}(\mathbf{Q}_\lambda \mathbf{T}) \leq P$ **then**
    $\lambda_{\max} \leftarrow \lambda$.
  else
    $\lambda_{\min} \leftarrow \lambda$.
  end if
until convergence
$\mathbf{Q} \leftarrow \mathbf{Q}_\lambda$

Proof. First note that the optimization problem is convex since the objective function is concave and the constraints are linear. Slater's condition is satisfied since a strictly feasible $\mathbf{Q}$ can always be found for nonzero P and ζ_c [33], and hence the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient conditions for a global optimal solution. Therefore, the solutions of primal and the dual problems coincides.

The optimization problem in (5.52) can be rewritten as

$$\begin{aligned}
&\max_{\substack{\mathbf{Q}\succeq 0 \\ \operatorname{tr}(\mathbf{QT})\leq P, \operatorname{tr}(\mathbf{QT}_c)\leq \zeta_c}} \log_2 \left|\mathbf{I} + \mathbf{HQH}^{\mathrm{H}}\right| \\
&\overset{(a)}{=} \max_{\mathbf{Q}\succeq 0} \min_{\mu\geq 0, \lambda\in[0,1]} \log_2 \left|\mathbf{I} + \mathbf{HQH}^{\mathrm{H}}\right| \\
&\qquad + \mu\lambda[P - \operatorname{tr}(\mathbf{QT})) + \mu(1-\lambda)(\zeta_c - \operatorname{tr}(\mathbf{QT}_c)] \qquad (5.53)\\
&\overset{(b)}{=} \max_{\mathbf{Q}\succeq 0} \min_{\mu\geq 0, \lambda\in[0,1]} \log_2 \left|\mathbf{I} + \mathbf{HQH}^{\mathrm{H}}\right| \\
&\qquad + \mu[\lambda P + (1-\lambda)\zeta_c - \operatorname{tr}(\mathbf{Q}(\lambda\mathbf{T} + (1-\lambda))\mathbf{T}_c)] \qquad (5.54)\\
&\overset{(c)}{=} \min_{\lambda\in[0,1]} \max_{\mathbf{Q}\succeq 0} \min_{\mu\geq 0} \log_2 \left|\mathbf{I} + \mathbf{HQH}^{\mathrm{H}}\right| \\
&\qquad + \mu\left[\lambda P + (1-\lambda)\zeta_c - \operatorname{tr}\left(\mathbf{Q}\left(\lambda\mathbf{T} + (1-\lambda)\,\mathbf{T}_c\right)\right]\right., \qquad (5.55)
\end{aligned}$$

where (a) follows from the fact that the inner minimization is unbounded below if $\mathbf{Q}$ does not satisfy the constraint, (b) follows from the linearity of the trace operator and (c) follows from duality. Now, define $\mathbf{R}_\lambda\mathbf{R}_\lambda^{\mathrm{H}}$ as the Cholesky factorization of $(\lambda\mathbf{T}+(1-\lambda)\mathbf{T}_c)^{-1}$, where $\mathbf{R}_\lambda \in \mathbb{C}^{M\times M}$ is positive definite, and let $\tilde{\mathbf{Q}}$ denote $\mathbf{R}_\lambda^{-1}\mathbf{Q}\mathbf{R}_\lambda^{-\mathrm{H}}$. Then, by substitution, (5.55) can be rewritten as following

$$\min_{\lambda\in[0,1]} \max_{\tilde{\mathbf{Q}}\succeq 0} \min_{\mu\geq 0} \log_2 \left|\mathbf{I} + \mathbf{H}\mathbf{R}_\lambda\tilde{\mathbf{Q}}\mathbf{R}_\lambda^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\right| + \mu\left[\lambda P + (1-\lambda)\zeta_c - \operatorname{tr}\left(\tilde{\mathbf{Q}}\right)\right]. \qquad (5.56)$$

Note that the inner maximization problem is essentially identical to the waterfilling problem with channel $\mathbf{HR}_\lambda$ and power constraint $\lambda P + (1-\lambda)\zeta_c$, which was discussed in

Section 4.1. As for the waterfilling algorithm, μ corresponds to the Langrange multiplier of the power constraint, and hence $\frac{1}{\mu}$ denotes the waterlevel.

The inner maximization problem is hence solved using the waterfilling algorithm, and the outer minimization is readily solved using a one-dimensional line-search, such as bisection, within the interval $[0, 1]$ [33]. Since the waterfilling algorithm and the bisection methods are known converge, Algorithm 2 converges to the solution of the optimization problem in (5.52). $\square$

Optimization with respect to a

The optimization problem maximizing the WMR with respect to $\mathbf{a}$ with the covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ fixed is formulated as following using the achievable rates in (5.28) and the relay power constraint in (5.40)

$$\max_{\substack{\mathbf{a}\\R_1,R_{B_1}\\R_2,R_{B_2}}} \quad \min(w_1R_1, w_2R_2, w_{B_1}R_{B_1}, w_{B_2}R_{B_2}) \tag{5.57a}$$

$$\text{s.t.} \quad \mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a} \leq P_R, \tag{5.57b}$$

$$R_{B_i} \leq \frac{1}{2}\log_2\left|\mathbf{I} + |a_i|^2(|a_i|^2\mathbf{B}_i\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{B}_i^{\mathrm{H}} + \mathbf{I}_M)^{-1}\mathbf{Q}_{B_i}\right|, \tag{5.57c}$$

$$R_i \leq \frac{1}{2}\log_2\left|\mathbf{I} + |a_i|^2(|a_i|^2\mathbf{B}_i\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{B}_i^{\mathrm{H}} + \mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\mathbf{H}_{B_i}^{-*}\mathbf{H}_i^{*})^{-1}\mathbf{Q}_i\right|. \tag{5.57d}$$

To solve this problem, the optimization problem is reversed such that the relay power consumption given by $\mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a}$ is minimized subject to a minimum WMR and (5.57c)-(5.57d), i.e.

$$\max_{\substack{\mathbf{a}\\R_1,R_{B_1}\\R_2,R_{B_2}}} \quad \mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a} \tag{5.58a}$$

$$\text{s.t.} \quad \min(w_1R_1, w_2R_2, w_{B_1}R_{B_1}, w_{B_2}R_{B_2}) \geq t \tag{5.58b}$$

$$R_{B_i} \leq \frac{1}{2}\log_2\left|\mathbf{I} + |a_i|^2(|a_i|^2\mathbf{B}_i\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{B}_i^{\mathrm{H}} + \mathbf{I}_M)^{-1}\mathbf{Q}_{B_i}\right|, \tag{5.58c}$$

$$R_i \leq \frac{1}{2}\log_2\left|\mathbf{I} + |a_i|^2(|a_i|^2\mathbf{B}_i\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{B}_i^{\mathrm{H}} + \mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\mathbf{H}_{B_i}^{-*}\mathbf{H}_i^{*})^{-1}\mathbf{Q}_i\right|, \tag{5.58d}$$

where $t \geq 0$ is the required minimum WMR. The solution to the optimization problem in (5.57) can then be found by repeatedly solving the optimization in (5.58) for different values of t until $\mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a} = P_R$. This is done using a one-dimensional line search algorithm.

Now, define the $\zeta_i \in \mathbb{R}_+$, with $i \in \{1, 2\}$, as the solutions of the equations

$$\begin{aligned} t = \frac{1}{2}\min\Big(&\log\left|\mathbf{I} + \zeta_i^2(\zeta_i^2\mathbf{B}_i\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{B}_i^{\mathrm{H}} + \mathbf{I}_M)^{-1}\mathbf{Q}_{B_i}\right|, \\ &\log\left|\mathbf{I} + \zeta_i^2(\zeta_i^2\mathbf{B}_i\tilde{\mathbf{H}}^{-1}\tilde{\mathbf{H}}^{-\mathrm{H}}\mathbf{B}_i^{\mathrm{H}} + \mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\mathbf{H}_{B_i}^{-*}\mathbf{H}_i^{*})^{-1}\mathbf{Q}_i\right|\Big). \end{aligned} \tag{5.59}$$

Note that solutions exist, since the achievable minimum rate in (5.59) for pair i is monotonically non-decreasing in ζ_i. The optimization problem in (5.58) can then be further reduced to

$$\underset{\mathbf{a}}{\text{minimize}} \quad \mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a} \tag{5.60a}$$

$$\text{subject to} \quad |a_i| \geq \zeta_i, \tag{5.60b}$$

which is a Quadratically Constrained Quadratic Program. The following proposition gives a simple analytical solution to the optimization problem in (5.60).

Proposition 5.1. *The solution to the optimization problem in* (5.60) *are given as*

$$\mathbf{a} = \begin{cases} \left[\zeta_1, -\frac{|s_{12}|}{s_{12}}\zeta_2\right]^{\mathrm{T}} & \textit{if } \frac{\zeta_2}{\zeta_1} \in \left[\frac{|s_{12}|}{s_{22}}, \frac{s_{11}}{|s_{12}|}\right] \\ \left[\frac{s_{12}}{s_{11}}\zeta_2, -\zeta_2\right]^{\mathrm{T}} & \textit{if } \frac{\zeta_2}{\zeta_1} \geq \frac{s_{11}}{|s_{12}|} \\ \left[\zeta_1, -\frac{s_{12}^*}{s_{22}}\zeta_1\right]^{\mathrm{T}} & \textit{if } \frac{\zeta_2}{\zeta_1} \leq \frac{|s_{12}|}{s_{22}} \end{cases} \tag{5.61}$$

Proof. Denote the entry of $\mathbf{S}$ at the p-th row and the q-th column as s_{pq}. Note that $\mathbf{S} \in \mathbb{C}^{2\times 2}$ is Hermitian and hence $s_{12} = s_{12}^*$. The optimization problem is then solved using the KKT conditions which are necessary conditions for a global minimum [33],

$$\left(\mathbf{S} + \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\right)\mathbf{a} = \mathbf{0} \tag{5.62a}$$

$$\lambda_i \leq 0 \qquad i \in \{1,2\} \tag{5.62b}$$

$$|a_i| \geq \zeta_i \qquad i \in \{1,2\} \tag{5.62c}$$

$$\lambda_i\left(|a_i| - \zeta_i\right) = 0 \qquad i \in \{1,2\}. \tag{5.62d}$$

Notice that, in order to satisfy (5.62a), the columns of the matrix $\mathbf{S} + \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ needs to be linearly dependent, which is the case if and only if

$$\left|\mathbf{S} + \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\right| = (s_{11} + \lambda_1)(s_{22} + \lambda_2) - |s_{12}|^2 = 0 \tag{5.63}$$

Any pair (λ_1, λ_2) satisfying (5.63) and (5.62b) are hence parametrized as

$$(\lambda_1, \lambda_2) = \left(-s_{11} + |s_{12}|\alpha, -s_{22} + |s_{12}|\frac{1}{\alpha}\right) \tag{5.64}$$

with $\alpha \in \left[\frac{|s_{12}|}{s_{22}}, \frac{s_{11}}{|s_{12}|}\right]$. In order to satisfy (5.62a), given α, the values of $\mathbf{a}$ needs to satisfy $\frac{a_2}{a_1} = -\frac{|s_{12}|\alpha}{s_{12}}$. There are now have three cases.

1. For $\alpha = \frac{|s_{12}|}{s_{22}}$: In this case $\lambda_1 = -s_{11} + \frac{|s_{12}|^2}{s_{22}}$ and $\lambda_2 = 0$. This implies that $|a_1| = \zeta_1$ must be satisfied, and hence

$$\mathbf{a} = \left[\zeta_1, -\frac{s_{12}^*}{s_{22}}\zeta_1\right]^{\mathrm{T}} \tag{5.65}$$

satisfies the KKT conditions in (5.62). This solution holds if and only if $\frac{\zeta_2}{\zeta_1} \leq \frac{|s_{12}|}{s_{22}}$.

2. For $\alpha = \frac{s_{11}}{|s_{12}|}$: In this case $\lambda_1 = 0$ and $\lambda_2 = -s_{22} + \frac{|s_{12}|^2}{s_{11}}$. This implies that $|a_2| = \zeta_2$ must be satisfied, and hence

$$\mathbf{a} = \left[-\frac{s_{12}}{s_{11}}\zeta_2, \zeta_2\right]^{\mathrm{T}}. \tag{5.66}$$

This solution holds if and only if $\frac{\zeta_2}{\zeta_1} \geq \frac{s_{11}}{|s_{12}|}$

Algorithm 3 Algorithm for computation of covariance matrices $\mathbf{Q}_i$, $\mathbf{Q}_{B_i}$ and amplification vector $\mathbf{a}$ for PF

output: $\mathbf{Q}_{B_i}$, $\mathbf{Q}_i$ and $\mathbf{a}$ with $i \in \{1,2\}$.
initialize: $\mathbf{Q}_{B_i} = \mathbf{Q}_i = \mathbf{0}$ and $\mathbf{a} = 10^{-3}[1,1]^{\mathrm{T}}$.
repeat
  Solve the optimization problem in (5.44) using Algorithm 1 to update $\mathbf{Q}_{B_i}$ and $\mathbf{Q}_i$.
  repeat
    Choose $t > 0$ using a line search algorithm.
    Compute ζ_i, with $i \in \{1,2\}$, from the equations in (5.59) based on t.
    Compute the optimal $\mathbf{a}$ of the optimization problem in (5.60) using Proposition 5.1.
  until $|\mathbf{a}^{\mathrm{H}}\mathbf{S}\mathbf{a} - P_R| \leq \epsilon$
until convergence
$\mathbf{Q} \leftarrow \mathbf{Q}_\lambda$

3. For $\alpha \in \left(\frac{|s_{12}|}{s_{22}}, \frac{s_{11}}{|s_{12}|}\right)$: In this case $|a_1| = \zeta_1$ and $|a_2| = \zeta_2$, and thus

$$\mathbf{a} = \left[\zeta_1, -\frac{|s_{12}|}{s_{12}}\zeta_2\right]^{\mathrm{T}}. \tag{5.67}$$

□

Using Algorithm 1 and Proposition 5.1, alternate optimization between the covariance matrices and the amplification vector $\mathbf{a}$ can be performed. The algorithm is described in Algorithm 3, and the algorithm converges since each of the separate optimization problems converges. In the practical implementation, $\epsilon = 10^{-8}$ and $\mathbf{a}$ is initialized to have a small magnitude, i.e. $\mathbf{a} = 10^{-3}[1,1]^{\mathrm{T}}$. It has been found that this choice makes convergence faster.

Numerical results for this scheme are presented in Section 5.4. The next section deals with the decode-and-forward schemes.

5.3. Decode-and-Forward Relaying

In this section, the achievable rate region using two different DF schemes are investigated investigated, and optimization problems maximizing the WMR are formulated. As opposed to the PF scheme, DF completely decodes the messages from the base stations and the terminals at the relay.

In the MA-phase, two different strategies, termed Decode-and-Forward with Aligned MA-phase (DF-A) and Decode-and-Forward with Unaligned MA-phase (DF-U), are considered. DF-A yields an achievable region which is simply computed and it keeps the communication flows in to different subspaces both in the MA-phase and in the BC-phase. In contrast, DF-U yields an achievable rate region which is more complex to compute and alignment is not used during the MA-phase.

The achievable rate regions of DF-A and DF-U are described in the following.

Pair-aligned transmissions in MA-phase The base stations align their signals to the terminals in the same manner as the precode-and-forward scheme in Section 5.2 In that way, the relay can separate the two communication flows. As for the precode-and-forward scheme the relay receives

$$\mathbf{y}_R = \underbrace{[\mathbf{H}_1, \mathbf{H}_2]}_{\tilde{\mathbf{H}}} \underbrace{\begin{bmatrix} \mathbf{I}_M & \mathbf{I}_M & 0 & 0 \\ 0 & 0 & \mathbf{I}_M & \mathbf{I}_M \end{bmatrix}}_{\tilde{\mathbf{I}}} \mathbf{u} + \mathbf{z}. \tag{5.68}$$

Consequently, the relay may invert the channel by left-multiplying $\tilde{\mathbf{H}}$ onto $\mathbf{y}_R$ to separate the signal pairs, and hence obtain

$$\mathbf{y}_{R_i} = \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \mathbf{y}_R = \mathbf{u}_{B_i} + \mathbf{u}_i + \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \mathbf{z} \tag{5.69}$$

The noise covariance matrix of this expression is then given by $\mathbf{C}_{\mathbf{n},\mathbf{y}_{R_i}} = \mathbf{B}_i \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\mathrm{H}} \mathbf{B}_i^{\mathrm{H}}$. Note that the coefficients in front of the signals are 1, which is due to the channel inversion. The transformation on the covariance matrix of the noise compensates for this. After separation, the relay can decode the codewords $\mathbf{u}_i$ and $\mathbf{u}_{B_i}$ at rates R_i and R_{B_i}. The rates at which the codewords can be decoded correctly, with a vanishing error probability when n tends to ∞, are given by the capacity region of the MIMO MA channel described in Section 4.2

$$\begin{aligned} &\mathcal{C}_{\mathrm{DF,MAC,aligned}} \\ &= \left\{ \begin{array}{l} R_i \le \tau \log_2 \left| \mathbf{I}_M + (\mathbf{B}_i \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\mathrm{H}} \mathbf{B}_i^{\mathrm{H}})^{-1} \mathbf{Q}_i \right| \\ R_i \le \tau \log_2 \left| \mathbf{I}_M + (\mathbf{B}_i \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\mathrm{H}} \mathbf{B}_i^{\mathrm{H}})^{-1} \mathbf{Q}_{B_i} \right| \\ R_i + R_{B_i} \le \tau \log_2 \left| \mathbf{I}_M + (\mathbf{B}_i \tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\mathrm{H}} \mathbf{B}_i^{\mathrm{H}})^{-1} (\mathbf{Q}_i + \mathbf{Q}_{B_i}) \right| \end{array} \right\}, \end{aligned} \tag{5.70}$$

under the power constraints

$$\mathrm{tr}\left(\mathbf{Q}_i\right) \le P_i \qquad \text{and} \qquad \mathrm{tr}\left(\mathbf{H}_{B_i}^{-1} \mathbf{H}_i \mathbf{Q}_{B_i} \mathbf{H}_i^{\mathrm{H}} \mathbf{H}_{B_i}^{-\mathrm{H}}\right) \le P_{B_i}. \tag{5.71}$$

Unaligned transmissions in MA-phase While pair-aligned transmissions in the MA-phase yields a simple achievable rate region, i.e. (5.70), alignment is not necessary since the messages from the nodes have to be completely decoded at the relay. The MA-phase may simply be seen as a MA channel with four nodes, i.e. the two base stations and the two terminals. However, the resulting capacity region with no alignment becomes considerable more complicated and involves more rate constraints. This also implies that decoding is more complex.

The base stations and terminals transmit codewords $\mathbf{x}_i \in \mathbb{C}^{M \times n}$ and $\mathbf{x}_{B_i} \in \mathbb{C}^{2M \times n}$, and corresponding covariance matrices of the transmitted codewords are denoted by $\mathbf{Q}_i \in \mathbb{C}^{M \times M}$ and $\mathbf{Q}_{B_i} \in \mathbb{C}^{2M \times 2M}$, respectively. In order to satisfy the power constraints, $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$ are required to satisfy the constraints

$$\mathrm{tr}\left(\mathbf{Q}_i\right) \le P_i \qquad \text{and} \qquad \mathrm{tr}\left(\mathbf{Q}_{B_i}\right) \le P_{B_i}. \tag{5.72}$$

According to the results on the MIMO MA channel summarized in Section 4.2, the codewords from the nodes can be decoded at all rates contained in using successive interference cancellation

$$\begin{aligned} &\mathcal{C}_{\mathrm{DF,MAC,unaligned}} \\ &= \left\{ \sum_{j \in S} R_j \le \tau \log_2 \left| \mathbf{I}_{2M} + \sum_{j \in S} \mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^{\mathrm{H}} \right|, \forall S \subseteq \{1, 2, B_1, B_2\} \right\}. \end{aligned} \tag{5.73}$$

Note that the number of rate constraints in (5.73) is $\binom{4}{1}+\binom{4}{2}+\binom{4}{3}+\binom{4}{4}=15$ as opposed to 3 in (5.70).

BC-phase In the BC-phase, the relay has successfully decoded the messages from each node using pair-aligned or unaligned transmissions in the MA-phase. In the BC-phase, alignment is employed. Alignment has the advantage that it forms a broadcast channel with side information for each communication flow. The broadcast channel with side information was also used for DF for the two-way MIMO relay channel used described in Section 4.4. Each of the signal pairs $(\mathbf{u}_{B_1},\mathbf{u}_1)$ and $(\mathbf{u}_{B_2},\mathbf{u}_2)$ are then reencoded to the codewords $\mathbf{u}_{R_1}\in\mathbb{C}^{M\times n}$ and $\mathbf{u}_{R_2}\in\mathbb{C}^{M\times n}$ with covariance matrices denoted by $\mathbf{Q}_{R_1}$ and $\mathbf{Q}_{R_2}$, respectively. For notational convenience, let $\mathbf{u}_R=[\mathbf{u}_{R_1}^{\mathrm{T}},\mathbf{u}_{R_2}^{\mathrm{T}}]^{\mathrm{T}}$. Note that the goal is to choose $\mathbf{u}_{R_i}$ such that, terminal i can decode the message m_{B_i} given a noisy version of $\mathbf{u}_{R_i}$ and its own message m_i as side information, and likewise the communication flow in the opposite direction. Finally, the codewords $\mathbf{u}_{R_1}$ and $\mathbf{u}_{R_2}$ are transmitted to the nodes using zero-forcing precoding.

In order to perform zero-forcing, the relay precodes $\mathbf{u}_R$ using the precoding matrix $\mathbf{G}_R\in\mathbb{C}^{2M\times 2M}$ such that $\mathbf{x}_R=\mathbf{G}_R\mathbf{u}_R$. The covariance matrix of $\mathbf{x}_R$ is then given as $\mathrm{tr}\left(\mathbf{G}_R\mathbf{u}_R(\mathbf{G}_R\mathbf{u}_R)^{\mathrm{H}}\right)=\mathbf{Q}_R$, and the relay have to satisfy the power constraints $\mathrm{tr}\left(\mathbf{Q}_R\right)\leq P_R$. As in the PF scheme, the precode matrix $\mathbf{G}_R$ is thus given by

$$\mathbf{G}_R=\tilde{\mathbf{H}}^{-\mathrm{T}} \tag{5.74}$$

Using this precode matrix, the covariance matrix $\mathbf{Q}_R$ is expressed as

$$\begin{aligned}\mathbf{Q}_R&=\tilde{\mathbf{H}}^{-\mathrm{T}}\begin{bmatrix}\mathbf{Q}_{R_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_{R_2}\end{bmatrix}\tilde{\mathbf{H}}^{-*} &(5.75)\\ &=\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{B}_1^{\mathrm{T}}\mathbf{Q}_{R_1}\mathbf{B}_1\tilde{\mathbf{H}}^{-*}+\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{B}_2^{\mathrm{T}}\mathbf{Q}_{R_2}\mathbf{B}_2\tilde{\mathbf{H}}^{-*}, &(5.76)\end{aligned}$$

and finally, the power constraint at the relay is given by

$$\mathrm{tr}\left(\mathbf{Q}_R\right)=\mathrm{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{B}_1^{\mathrm{T}}\mathbf{Q}_{R_1}\mathbf{B}_1\tilde{\mathbf{H}}^{-*}\right)+\mathrm{tr}\left(\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{B}_2^{\mathrm{T}}\mathbf{Q}_{R_2}\mathbf{B}_2\tilde{\mathbf{H}}^{-*}\right)\leq P_R. \tag{5.77}$$

With the precoding matrix in (5.74), each of the terminals receive

$$\tilde{\mathbf{y}}_i=\mathbf{H}_i^{\mathrm{T}}\mathbf{x}_R+\mathbf{z}_i=\mathbf{u}_{R_i}+\mathbf{z}_i. \tag{5.78}$$

Note that terminal i only receives the codeword associated with the base station-terminal pair i due to the zero forcing precoding matrix. Since the base stations have $2M$ antennas, the desired signal pair at base station i can be obtained using the receive matrix $\mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}$, in order to extract only the spatial dimensions in which the signal $\mathbf{x}_{R_i}$ reside. That is

$$\begin{aligned}\tilde{\mathbf{y}}_{B_i}&=\mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\left(\mathbf{H}_{B_i}^{\mathrm{T}}\tilde{\mathbf{H}}^{-\mathrm{T}}\mathbf{x}_R+\mathbf{z}_{B_i}\right) &(5.79)\\ &=\mathbf{x}_{R_i}+\mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\mathbf{z}_{B_i}. &(5.80)\end{aligned}$$

The noise covariance matrix of this signal is given as $\mathbf{C}_{\mathbf{n},\tilde{\mathbf{y}}_{B_i}}=\mathbf{H}_i^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\mathbf{H}_{B_i}^{-*}\mathbf{H}_i^{*}$

From the discussion of the broadcast channel with side information in Section 4.4, the achievable rate region for the BC-phase is given by [34]

$$\mathcal{C}_{\mathrm{DF},BC}=\left\{\begin{array}{c}R_i\leq(1-\tau)\log_2\left|\mathbf{I}+\mathbf{C}_{\mathbf{n},\tilde{\mathbf{y}}_{B_i}}^{-1}\mathbf{Q}_{R_i}\right|\\ R_{B_i}\leq(1-\tau)\log_2\left|\mathbf{I}+\mathbf{Q}_{R_i}\right|\end{array}\right\} \tag{5.81}$$

Degrees of Freedom The DoF of each communication flow are denoted d_1, d_2, d_{B_1} and d_{B_2}, and limited by both the MA- and BC-phase. Constraints on the DoF are stated in the following.

- *DF-A*:The DoF are only limited by the MA-phase, and hence the DoF are given as for the MA channel

$$d_i + d_{B_i} \le \frac{1}{2}M. \tag{5.82}$$

- *DF-U*: The DoF are both limited by the MA- and BC-phase, i.e.

$$d_1 + d_2 + d_{B_1} + d_{B_2} \le M, \tag{5.83}$$

$$d_i \le \frac{1}{2}M \quad \text{and} \quad d_{B_i} \le \frac{1}{2}M, \tag{5.84}$$

  where the constraint in (5.83) originates from the MA channel with four terminals and (5.84) stems from the BC channel with side information at the receiver.

Both schemes achieves M DoF in total, which is half of the total DoF of PF. This degradation is solely due to the fact that the messages from each node have to be decoded at the relay. Interestingly, DF-U is seen be scale better to asymmetrical DoF situations, i.e. if one of the base station-terminal pair are weighted higher than the other pair. This is due to the fact that $d_i + d_{B_i}$ may be larger than $\frac{1}{2}M$ as opposed to DF-A.

5.3.1 Problem formulation

From the achievable rate regions in (5.70), (5.73) and (5.81) this section maximizes the WMR for DF-A and DF-U.

The problem of maximizing the WMR for DF-A is given by

$$\max_{\substack{\mathbf{Q}_{R_i}\mathbf{Q}_i,\mathbf{Q}_{B_i},\tau \\ R_1,R_{B_1},R_2,R_{B_2}}} \quad \min(w_1R_1, w_2R_2, w_{B_1}R_{B_1}, w_{B_2}R_{B_2}) \tag{5.85a}$$

$$\text{s.t.} \quad (R_1, R_2, R_{B_1}, R_{B_2}) \in \mathcal{C}_{\text{DF,MAC,aligned}} \cap \mathcal{C}_{\text{DF,BC}}, \tag{5.85b}$$

$$(5.77), \mathbf{Q}_i \succeq 0, \mathbf{Q}_{B_i} \succeq 0, \mathbf{Q}_{R_i} \succeq 0, \tag{5.85c}$$

$$\operatorname{tr}(\mathbf{Q}_i) \le P_i, \operatorname{tr}\left(\mathbf{H}_{B_i}^{-1}\mathbf{H}_i\mathbf{Q}_{B_i}\mathbf{H}_i^{\mathrm{H}}\mathbf{H}_{B_i}^{\mathrm{H}}\right) \le P_{B_i}. \tag{5.85d}$$

The problem of maximizing the WMR for DF-U, is simply formulated by exchanging $\mathcal{C}_{\text{DF,MAC,aligned}}$ by $\mathcal{C}_{\text{DF,MAC,unaligned}}$.

These problems are in general non-convex, and hence the problems can not be solved using standard convex optimization algorithms. Following [3], the optimization problem in (5.85) is relaxed using alternate optimization. The optimization is performed in two steps

1. The optimization problem in (5.85) is solved with respect to $\mathbf{Q}_i$, $\mathbf{Q}_{B_i}$ and $\mathbf{Q}_{R_i}$ while keeping τ fixed. In that case, the problem reduces to a convex optimization problem.

2. The optimization problem in (5.85) is solved with respect to τ while keeping $\mathbf{Q}_i$, $\mathbf{Q}_{B_i}$ and $\mathbf{Q}_{R_i}$ fixed. The problem is then reduced to a linear programming problem.

To find a local solution to (5.85), the two optimization problems are solved alternately until convergence[1]. Each of these problems are solved using CVX in MATLAB. In the practical implementation, τ is initialized to 0.5.

5.4. Numerical Results

In this section, the proposed relaying schemes are evaluated by simulation. The objective of the simulations are to evaluate the following:

– *Achievable WMR in terms of M and SNR*: As the number of antennas is increased, the DoFs, and hence the WMR, of each the schemes are expected to increase linearly. Similarly, as the SNR increases, the WMR is expected to increase. Since PF achieves the maximum number of DoF, PF is expected to perform well at high SNR while DF-A and DF-U are expected perform well at low SNR. Moreover, DF-U is expected to perform better than DF-A, since it is not restricted to the alignment subspaces in the MA-phase.

– *Achievable regions with fixed uplink-downlink ratio*: For fixed P_j with $j \in \{1, 2, B_1, B_2, R\}$, and ratios $\frac{w_1}{w_{B_1}}$ and $w_1 = w_2$ and $w_{B_1} = w_{B_2}$, the achievable rate regions can be found by simulation, i.e. the points $(\min(R_1, R_{B_1}), \min(R_2, R_{B_2}))$ that can be achieved for the different schemes. This shows how well the schemes adapt to asymmetric weights.

These cases are simulated using a simulation framework developed in MATLAB that implements the three relaying schemes PF, DF-A and DF-U. Besides these three schemes, a simple Amplify-and-Forward scheme is implemented as comparison to PF. The AF scheme operates in the same way as PF, with the simplification that the relay precode matrix $\mathbf{G}_R$ is chosen as a scaled identity matrix

$$\mathbf{G}_R = \gamma \mathbf{I}_{2M} \Leftrightarrow \mathbf{A}_i = \gamma \mathbf{I}_M, \tag{5.86}$$

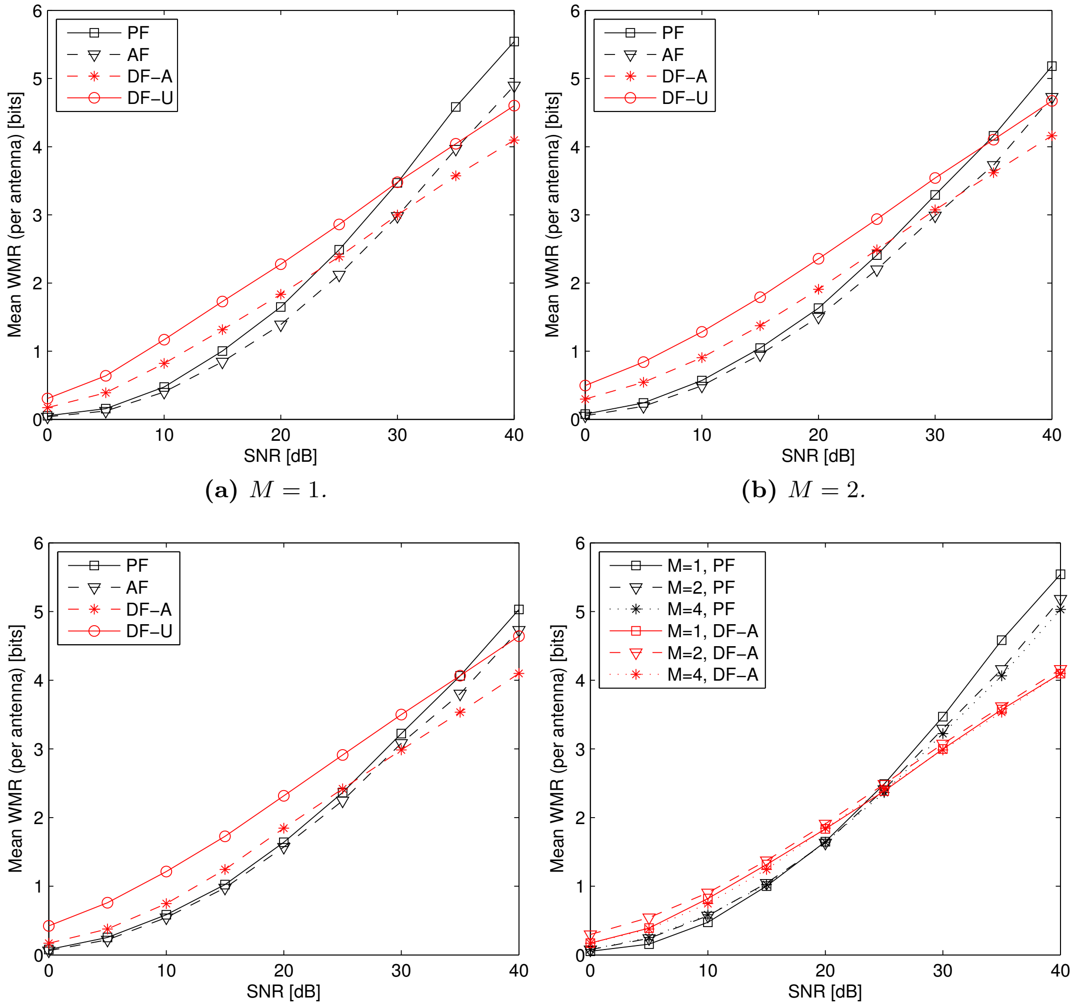
where γ is chosen such that the relay power consumption is used. Optimization is then only performed over the covariance matrices $\mathbf{Q}_i$ and $\mathbf{Q}_{B_i}$. The AF scheme, as opposed to the other relaying schemes, lacks the ability to distribute the power of the relay to each communication flow in an optimal way. This leads to hypothesis that it performs significantly worse than PF, but achieves better WMR than DF-A and DF-U at very high SNR.

All simulations are performed by averaging over 300 realizations of channel matrices. The entries of the channel matrices, as previously explained, chosen from a complex zero-mean Gaussian distribution with unit variance. All algorithms are stopped when

$$\frac{\mathrm{WMR}^{(k+1)} - \mathrm{WMR}^{(i)}}{\mathrm{WMR}^{(k+1)}} \leq 10^{-5}. \tag{5.87}$$

where $\mathrm{WMR}^{(k)}$ denotes the WMR in the k-th iteration of the respective scheme. Finally, when the SNR is given in decibels, it is computed as

$$\mathrm{SNR}\ [\mathrm{dB}] = 10 \log_{10} \mathrm{SNR}. \tag{5.88}$$



(a) $M = 1$.

(b) $M = 2$.

(c) $M = 4$.

(d) *WMR of PF and DF plotted in terms of* M *and SNR.*

Figure 5.3: *Performance of the developed schemes with* $P_1 = P_2 = P_{B_1} = P_{B_2} = SNR$ *and* $P_R = 4$ *SNR.*

Achievable WMR in terms of SNR and M To compare the developed relaying schemes, the WMR, normalized to the number of antennas, has been simulated in terms of SNR and number of antennas M and powers at the nodes. The results are depicted on Fig. 5.3.

It is seen that DF-A and DF-U performs well at low SNR, while the performance of PF increases faster due to the higher number of DoF, and hence performs better at high SNR. This is explained by the fact that in DF-A and DF-U, the signals are completely decoded at the relay after the MA-phase, whereas for PF, the signals are allowed to add up in their respective alignment subspaces, at the relay, since each terminal are able to cancel their contribution upon reception. Moreover, from Fig. 5.3d, observe that at high SNR, the WMR per antenna is almost the same for $M \in \{1, 2, 4\}$, and hence the WMR approximately

[1] Note that alternate optimization indeed converges since each step decreases the WMR, and the optimization problems in each step are bounded above.

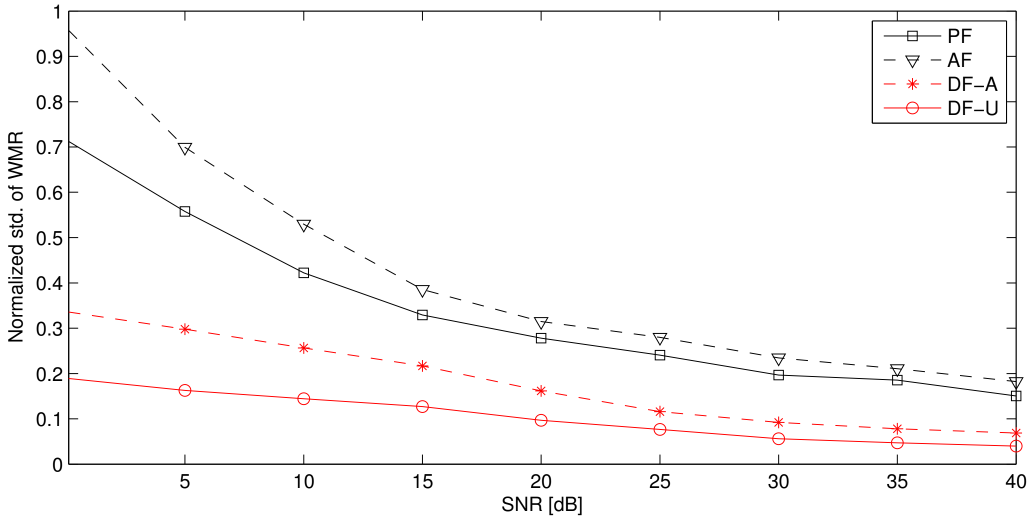


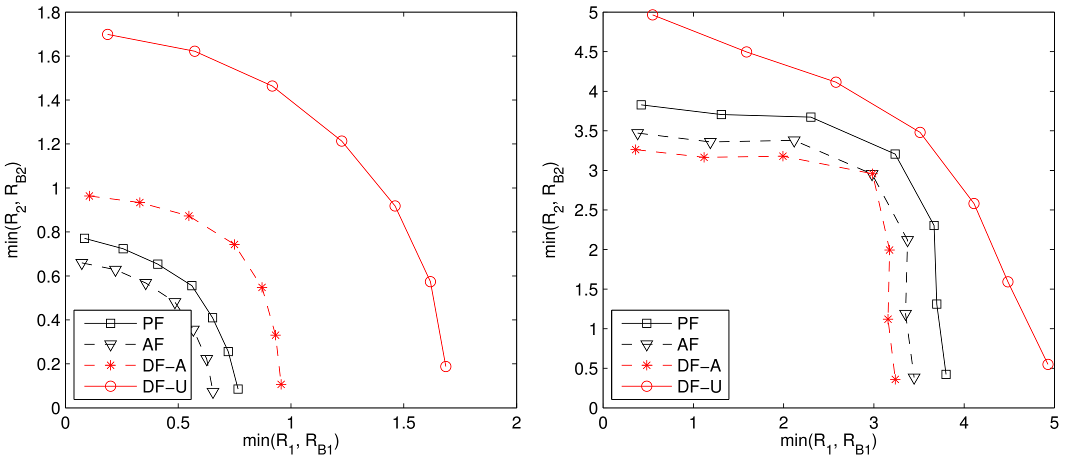
Figure 5.4: *Standard deviation in percentage of the* WMR *with the same power settings as in Fig. 5.3 and with* $M = 2$.

increases linearly with M.

The improvement of PF compared to AF is limited. Clearly, both schemes features the same number of DoF, and the performance gap between the two schemes decreases as the number of antennas increases. This may be due to the fact that the waterfilling algorithm has more parameters to optimize over. For example, if $M = 1$, all the covariance matrices are scalars, and hence the waterfilling algorithm can only pour energy into one eigenmode. Therefore, distributing the relay power in the optimal way naturally yields a significant performance increase. However, for high M, the waterfilling algorithm can pour power into many eigenmodes, which may decrease the need for optimal power allocation at the relay.

The standard deviation of the WMR normalized with respect to the WMR is depicted in Fig 5.4, and is interpreted as a measure of diversity. If the standard deviation is high, the WMR highly depends on the specific channel realization, whereas a low standard deviation means that the scheme is able to adapt to the channel well. It is seen that AF and PF have standard deviation that are significantly higher than DF-A and DF-U. This is likely due to the simplified choice of the relay precode matrix that prevents the relay from reallocating the power of each communication flow in an optimal way. Moreover, DF-U, which is not restricted by alignment in the MA-phase, achieves the lowest standard deviation of all the schemes.

Achievable regions The achievable rate regions simulated are depicted in Fig. 5.5. As previously seen, PF performs well at high SNR while DF-A and DF-U perform well at low SNR. Moreover, at high SNR, it is seen that DF-U adapts well to asymmetric weights compared to DF-A, PF and AF. This is explained by the fact that DF-U is the only relaying scheme that is able to adjust the DoF of each communication flows to the weight tuple. This result is also seen from Fig. 5.6 that shows the WMR in terms of SNR with the communication flows weighted asymmetrically. The figure shows that the slope of the WMR is stepper for DF-U than DF-A, and hence DF-U achieves higher DoF. This indicates that DF-U in fact uses its DoF in a more efficient way than DF-A with asymmetric weights.



(a) $P_1 = P_2 = P_{B_1} = P_{B_2} = 10\,\text{dB}$ *and* $P_R = 16\,\text{dB}$.

(b) $P_1 = P_2 = P_{B_1} = P_{B_2} = 30\,\text{dB}$ *and* $P_R = 36\,\text{dB}$.

Figure 5.5: *Achievable rate regions with* $M = 2$ *(per antenna).*

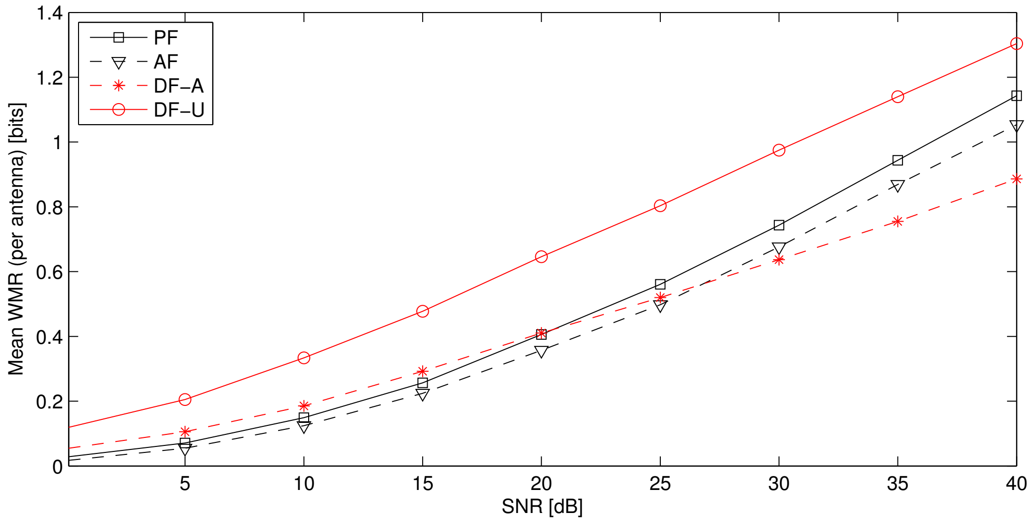


Figure 5.6: *with asymmetric weights* $w_1 = w_{B_1} = 0.2$ *and* $w_2 = w_{B_2} = 1.8$*, and the same power settings as in Fig. 5.3 and with* $M = 2$.

Validation considerations This paragraph briefly describes considerations on validations of the results.

- All equations (5.8) through (5.43) have been verified through simulation. For instance, all power consumptions and covariance matrices, i.e. at the base stations, terminals and the relay, have been found numerically by averaging over a high number of signal realizations and these results are compared expressions for the corresponding power consumptions.

- The solutions of the optimization problem in (5.44) obtained using Algorithm 1 have been verified using CVX.

- The solutions of the optimization problem in (5.60) which is obtained from the Proposition 5.1 have been verified using `fmincon()` in MATLAB. Further, the KKT conditions are automatically verified during the simulations when optimizing with respect to $\tilde{\mathbf{a}}$.

In general, verification of the achievable rate regions are difficult since verification requires an achievable scheme. Such achievable schemes involve designing capacity approaching codes such as LDPC codes.

The next chapter extends the shared relay channel using the results from the four-way relay.

CHAPTER 6

Shared Four-way Relay Channel

With the results for the shared relay and the four-way relay, the aim of this chapter is to demonstrate that the concepts of zero-forcing and alignment scales to larger networks of nodes. This chapter proposes a Decode-and-Forward scheme, which is an extension of DF-A described in Section 5.3, for the shared four-relay channel depicted in Fig. 6.1. The network is essentially a combination of two shared relay channels that are combined such that one of the base stations is common for both networks. The common base station is thus serving two terminals. The common base station then uses a combination of zero-forcing and alignment to align to both terminals. In order to accomplish this, the number of antennas at the common base station have to be $4M$.

The next section describes the system model and in Section 6.2, the scheme and the corresponding achievable rate region is derived. Finally, the scheme is compared to DF-A for the shared relay.

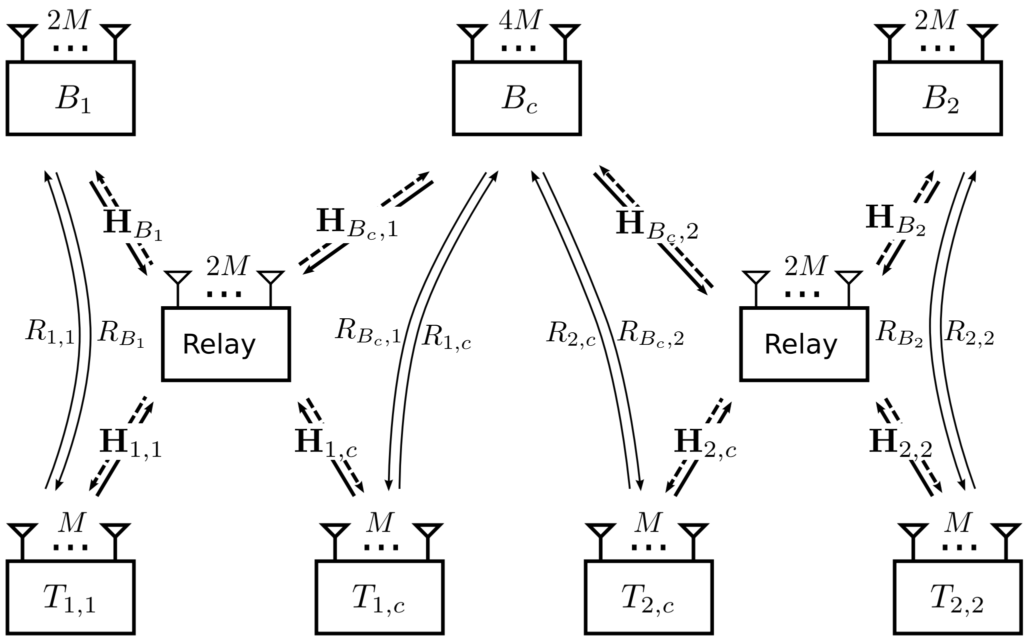


Figure 6.1:

6.1. System model

The shared four-way relay channel, depicted in Fig. 6.1, consists of the four terminals, $T_{1,1}$, $T_{1,c}$, $T_{2,2}$ and $T_{2,c}$, the three base stations, B_1, B_2 and B_c and two relays R_1 and R_2. The terminals are all equipped with M antennas and the base stations, B_1 and B_2, and the relay, R_1 and R_2, are equipped with $2M$ antennas to enable alignment. The common station has $4M$ antennas to enable both alignment and zero-forcing.

The time is divided into a total of n channel uses. The first $n_{\mathrm{MA}} = \lfloor n\tau \rfloor$ channel uses are devoted to the MA-phase while the remaining $n_{\mathrm{BC}} = n - n_{\mathrm{MA}}$ are for the BC-phase. The following list describes the messages and the encoding functions to map the messages to codewords.

- $T_{i,i}$ have messages $m_{i,i} \in \{1, \ldots, 2^{nR_{i,i}}\}$ destined to B_i with $i \in \{1, 2\}$. The messages are mapped to codewords $\mathbf{x}_{i,i}^{n_{\mathrm{MA}}} \in \mathbb{C}^{M \times n_{\mathrm{MA}}}$ using the encoding function $f_{i,i}^{(n)} : \{1, \ldots, 2^{nR_{i,i}}\} \to \mathbb{C}^{M \times n_{\mathrm{MA}}}$. codewords
- $T_{i,c}$ have messages $m_i \in \{1, \ldots, 2^{nR_{i,c}}\}$ destined to B_c. The messages are mapped to codewords $\mathbf{x}_{i,c}^{n_{\mathrm{MA}}} \in \mathbb{C}^{M \times n_{\mathrm{MA}}}$ using the encoding function $f_{i,c}^{(n)} : \{1, \ldots, 2^{nR_{i,c}}\} \to \mathbb{C}^{M \times n_{\mathrm{MA}}}$.
- B_i have messages $m_{B_i} \in \{1, \ldots, 2^{nR_{B_i}}\}$ destined to $T_{i,i}$. The messages are mapped to codewords $\mathbf{x}_{B_i}^{n_{\mathrm{MA}}} \in \mathbb{C}^{2M \times n_{\mathrm{MA}}}$ using the encoding function $f_{B_i}^{(n)} : \{1, \ldots, 2^{nR_{B_i}}\} \to \mathbb{C}^{2M \times n_{\mathrm{MA}}}$.
- B_c have messages $m_{B_c,i} \in \{1, \ldots, 2^{nR_{B_c,i}}\}$ destined to $T_{i,c}$ for $i \in \{1, 2\}$. The messages are mapped to a single codeword $\mathbf{x}_{B_c}^{n_{\mathrm{MA}}} \in \mathbb{C}^{4M \times n_{\mathrm{MA}}}$ using the encoding function $f_{B_c}^{(n)} : \{1, \ldots, 2^{nR_{B_c,1}}\} \times \{1, \ldots, , 2^{nR_{B_c,2}}\} \to \mathbb{C}^{4M \times n_{\mathrm{MA}}}$. That is $f_{B_c}^{(n)}$ is a function of both $m_{B_c,1}$ and $m_{B_c,2}$.

These codewords are chosen from capacity achieving point-to-point codes for the Gaussian MIMO channel and are subject to the power constraints

$$\mathbb{E}\left[\mathbf{x}_{i,i}^{n_{\mathrm{MA}},\mathrm{H}}\mathbf{x}_{i,i}^{n_{\mathrm{MA}}}\right] \leq P_{i,i}, \quad \mathbb{E}\left[\mathbf{x}_{i,c}^{n_{\mathrm{MA}},\mathrm{H}}\mathbf{x}_{i,c}^{n_{\mathrm{MA}}}\right] \leq P_{i,c} \tag{6.1a}$$

$$\mathbb{E}\left[\mathbf{x}_{B_i}^{n_{\mathrm{MA}},\mathrm{H}}\mathbf{x}_{B_i}^{n_{\mathrm{MA}}}\right] \leq P_{B_i}, \quad \mathbb{E}\left[\mathbf{x}_{B_c}^{n_{\mathrm{MA}},\mathrm{H}}\mathbf{x}_{B_c}^{n_{\mathrm{MA}}}\right] \leq P_{B_c}. \tag{6.1b}$$

Thus there are four two-way communications flows between the base stations and the terminals with the rates denoted by $R_{i,i}$, $R_{i,c}$, R_{B_i} and $R_{B_c,i}$.

The relay i, for $i \in \{1, 2\}$, receives noisy superposition of the codewords from the base stations B_i, B_c and the terminals $T_{i,i}$ and $T_{i,c}$ through the channels $\mathbf{H}_{B_i} \in \mathbb{C}^{2M \times 2M}$, $\mathbf{H}_{B_c,i} \in \mathbb{C}^{2M \times 4M}$, $\mathbf{H}_{i,i} \in \mathbb{C}^{2M \times M}$ and $\mathbf{H}_{i,c} \in \mathbb{C}^{2M \times M}$. As for the shared relay, the entries of these channel matrices are given by iid complex Gassian random variables with unit variance, and the channel matrices are assumed to be of full rank. To simplify notation, the channel matrix $\mathbf{H}_i$ is introduced

$$\mathbf{H}_i = \left[\mathbf{H}_{B_i}, \mathbf{H}_{i,i}, \mathbf{H}_{B_c,i}, \mathbf{H}_{1,c}\right]. \tag{6.2}$$

With this notation, the relay obtains $\mathbf{y}_{R_i}^{n_{\mathrm{MA}}} \in \mathbb{C}^{2M \times n_{\mathrm{MA}}}$ in the MA-phase

$$\mathbf{y}_{R_i}^{n_{\mathrm{MA}}} = \mathbf{H}_i \begin{bmatrix} \mathbf{x}_{B_i}^{n_{\mathrm{MA}}} \\ \mathbf{x}_{i,i}^{n_{\mathrm{MA}}} \\ \mathbf{x}_{B_c}^{n_{\mathrm{MA}}} \\ \mathbf{x}_{i,c}^{n_{\mathrm{MA}}} \end{bmatrix} + \mathbf{z}_i^{n_{\mathrm{MA}}} \tag{6.3}$$

for $i \in \{1, 2\}$, where $\mathbf{z}_i^{n_{\mathrm{MA}}} \overset{\mathrm{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{2M})$. Using the relay mapping $g_i^{(n)} : \mathbb{C}^{2M \times n_{\mathrm{MA}}} \to \mathbb{C}^{4M \times n_{\mathrm{MA}}}$, relay i maps $\mathbf{y}_{R_i}^{n_{\mathrm{MA}}}$ into $\mathbf{x}_{R_i}^{n_{\mathrm{BC}}}$, which is subject to the relay power constraints

$$\mathbb{E}\left[\mathbf{x}_{R_i}^{n_{\mathrm{BC}},\mathrm{H}} \mathbf{x}_{i,i}^{n_{\mathrm{BC}}}\right] \leq P_{R_i}. \tag{6.4}$$

The signals $\mathbf{x}_{R_i}$ are then broadcast to the base stations and terminals through the reciprocal channels, and hence each node receives

$$\mathbf{y}_{i,i}^{n_{\mathrm{BC}}} = \mathbf{H}_{i,i}^{\mathrm{T}} \mathbf{x}_{R_i}^{n_{\mathrm{BC}}} + \mathbf{w}_{i,i}^{n_{\mathrm{BC}}} \tag{6.5}$$

$$\mathbf{y}_{i,c}^{n_{\mathrm{BC}}} = \mathbf{H}_{i,c}^{\mathrm{T}} \mathbf{x}_{R_i}^{n_{\mathrm{BC}}} + \mathbf{w}_{i,c}^{n_{\mathrm{BC}}} \tag{6.6}$$

$$\mathbf{y}_{B_i}^{n_{\mathrm{BC}}} = \mathbf{H}_{B_i}^{\mathrm{T}} \mathbf{x}_{R_i}^{n_{\mathrm{BC}}} + \mathbf{w}_{B_i}^{n_{\mathrm{BC}}} \tag{6.7}$$

$$\mathbf{y}_{B_c}^{n_{\mathrm{BC}}} = \mathbf{H}_{B_c,1}^{\mathrm{T}} \mathbf{x}_{R_i}^{n_{\mathrm{BC}}} + \mathbf{H}_{B_c,2}^{\mathrm{T}} \mathbf{x}_{R_2}^{n_{\mathrm{BC}}} + \mathbf{w}_{B_c}^{n_{\mathrm{BC}}}, \tag{6.8}$$

where $\mathbf{w}_{i,i}^{n_{\mathrm{BC}}}, \mathbf{w}_{i,c}^{n_{\mathrm{BC}}} \overset{\mathrm{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_M)$, $\mathbf{w}_{B_i}^{n_{\mathrm{BC}}} \overset{\mathrm{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{2M})$ and $\mathbf{w}_{B_c}^{n_{\mathrm{BC}}} \overset{\mathrm{iid}}{\sim} \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{4M})$. Finally, the decoding functions $\hat{m}_{i,i}^{(n)}$, $\hat{m}_{i,,c}^{(n)}$, $\hat{m}_{B_i}^{(n)}$ and $\hat{m}_{B_c,i}^{(n)}$ are used to reconstruct the desired messages. The objective is to choose the encoding functions, relay mapping and decoding functions such that the achievable WMR is maximized and the probability of errornous reconstruction of the messages tends to zero as n tends infinity.

In the remaining part of this chapter, the superscripts are omitted for notational convenience.

6.2. Decode-and-Forward Scheme

In this section, a DF scheme based on zero-forcing and alignment is developed. The main idea of the scheme is to demonstrate that these concepts scales to larger networks of nodes and relays. As for the shared relay, each base stations aligns their signal to the corresponding terminal, creating an interference-free communication flows between the base station and corresponding terminal. The vector pairs $(\mathbf{u}_{i,i}, \mathbf{u}_{B_i})$ and $(\mathbf{u}_{i,c}, \mathbf{u}_{B_c,i})$ for $i \in \{1, 2\}$ denote the signals to transmitted by the corresponding nodes in each of the four communication flows before alignment. The vectors are contained in the space $\mathbb{C}^{M \times n_{\mathrm{BC}}}$, and their covariance matrices are denoted by $\mathbf{Q}_{i,i}$, $\mathbf{Q}_{B_i}$, $\mathbf{Q}_{i,c}$ and $\mathbf{Q}_{B_c,i}$ from the vector space $\mathbb{C}^{M \times M}$. In the BC-phase, the messages are reencoded at the relays, and the relay ensures that each communication flow is interference-free using zero-forcing precoding . The signals before zero-forcing are denoted by $\mathbf{u}_{R_i,i}$ and $\mathbf{u}_{R_i,c}$, with $i \in \{1, 2\}$, and the corresponding covariance matrices are given by $\mathbf{Q}_{i,i}$ and $\mathbf{Q}_{B_i}$.

Since they terminals do not use alingment, $\mathbf{x}_{i,i} = \mathbf{u}_{i,i}$ and $\mathbf{x}_{i,c} = \mathbf{u}_{i,c}$. In the following, the precoding schemes at the base stations are described.

Precoding at the base stations The precoding scheme for the base stations, B_1 and B_2, is the same as for the DF-A for the shared relay, i.e.

$$\mathbf{x}_{B_i} = \mathbf{H}_{B_i}^{-1}\mathbf{H}_{i,i}\mathbf{u}_{B_i}, \tag{6.9}$$

where $\mathbf{u}_{B_i}$ is the signal destined to the terminal $T_{i,i}$.

For the base station B_c, let $\mathbf{u}_{B_c,i}$ denote the signal destined to terminal $T_{i,c}$, and let $\mathbf{G}_{B_c,i} \in \mathbb{C}^{4M\times M}$ be the alignment matrix for the corresponding terminal. The common base station then transmits

$$\mathbf{x}_{B_c} = \mathbf{G}_{B_c,1}\mathbf{u}_{B_c,1} + \mathbf{G}_{B_c,2}\mathbf{u}_{B_c,2}. \tag{6.10}$$

For this scheme, it is desirable to choose $\mathbf{G}_{B_c,i}$ such that zero-forcing and alignment is achieved, and hence the following conditions have to be satisfied

$$\operatorname{span}(\mathbf{H}_{B_c,i}\mathbf{G}_{B_c,i}) = \operatorname{span}(\mathbf{H}_{i,c}) \tag{6.11}$$

$$\mathbf{H}_{B_c,\bar{i}}\mathbf{G}_{B_c,i} = \mathbf{0}. \tag{6.12}$$

Alignment is achieved through the constraint in (6.11) that restricts the subspaces that base station B_c and terminal $T_{i,c}$ are transmitting in to coincide at relay i. Zero-forcing is achieved through the constraint in (6.12) that ensures that the signal, $\mathbf{u}_{B_c,i}$, sent through the precoding matrix $\mathbf{G}_{B_c,i} \in \mathbb{C}^{4M\times M}$ is cancelled at relay $\bar{i}$. Note that the only requirements that have to be fulfilled to satisfy (6.11)-(6.12) are that $\mathbf{G}_{B_c,i}$ is in the nullspace of $\mathbf{H}_{B_c,\bar{i}}$ and that $\mathbf{G}_{B_c,i}$ is in the column space of $\mathbf{H}_{B_c,i}$. As for the shared relay, the matrices $\mathbf{G}_{B_c,i}$ are not uniquely determined, but the specific choice does not affect the achievable rates since the covariance matrices can be freely chosen. A simple way to find an alignment matrix is then to solve the following system for $\mathbf{G}_{B_c,i}$

$$\begin{bmatrix} \mathbf{H}_{B_c,i} \\ \mathbf{H}_{B_c,\bar{i}} \end{bmatrix} \mathbf{G}_{B_c,i} = \begin{bmatrix} \mathbf{H}_{i,c} \\ \mathbf{0} \end{bmatrix} \Leftrightarrow \mathbf{G}_{B_c,i} = \begin{bmatrix} \mathbf{H}_{B_c,i} \\ \mathbf{H}_{B_c,\bar{i}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{H}_{i,c} \\ \mathbf{0} \end{bmatrix}. \tag{6.13}$$

The concatenated alignment matrix, $\mathbf{G}_{B_c} \in \mathbb{C}^{4M\times 2M}$, is then defined as

$$\mathbf{G}_{B_c} = [\mathbf{G}_{B_c,1}, \mathbf{G}_{B_c,2}] = \begin{bmatrix} \mathbf{H}_{B_c,1} \\ \mathbf{H}_{B_c,2} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{H}_{1,c} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{2,c} \end{bmatrix}. \tag{6.14}$$

Using these alignment matrices, the i-th relay receives a superposition of the signals from the terminals $T_{i,i}$ and $T_{i,c}$, and base stations B_i and B_c,

$$\mathbf{y}_{R_i} = \mathbf{H}_{B_i}\mathbf{G}_{B_i}\mathbf{u}_{B_i} + \mathbf{H}_{B_c,i}\mathbf{G}_{B_c}\begin{bmatrix} \mathbf{u}_{B_c,1} \\ \mathbf{u}_{B_c,2} \end{bmatrix} + \mathbf{H}_{i,1}\mathbf{u}_i + \mathbf{H}_{i,c}\mathbf{u}_{i,c} + \mathbf{z}_i \tag{6.15}$$

$$= \tilde{\mathbf{H}}_i\tilde{\mathbf{I}}\mathbf{u}_i + \mathbf{z}_i. \tag{6.16}$$

where $\tilde{\mathbf{H}}_i = [\mathbf{H}_{i,1}, \mathbf{H}_{i,c}]$ and $\mathbf{u}_i = [\mathbf{u}_{B_1}^{\mathrm{T}}, \mathbf{u}_1^{\mathrm{T}}, \mathbf{u}_{B_2}^{\mathrm{T}}, \mathbf{u}_2^{\mathrm{T}}]^{\mathrm{T}}$. As for the shared relay, the communication flows can be separated by inverting the effective channel $\tilde{\mathbf{H}}_i$ for $i \in \{1,2\}$. In that way, the relay i can obtain noisy versions of each communication flow

$$\mathbf{y}_{R_i,i} = \mathbf{u}_{B_i} + \mathbf{u}_{i,i} + \mathbf{B}_1\tilde{\mathbf{H}}_i^{-1}\mathbf{z}_i \tag{6.17a}$$

$$\mathbf{y}_{R_i,c} = \mathbf{u}_{B_c,i} + \mathbf{u}_{i,c} + \mathbf{B}_2\tilde{\mathbf{H}}_i^{-1}\mathbf{z}_i. \tag{6.17b}$$

To express the achievable rate region of the MA-phase, the noise covariance matrices of these signals are given computed as $\mathbf{C}_{\mathbf{n},i,i} = \mathbf{B}_1\tilde{\mathbf{H}}_i^{-1}\tilde{\mathbf{H}}_i^{-\mathrm{H}}\mathbf{B}_1^{\mathrm{H}}$ and $\mathbf{C}_{\mathbf{n},i,c} = \mathbf{B}_2\tilde{\mathbf{H}}_i^{-1}\tilde{\mathbf{H}}_i^{-\mathrm{H}}\mathbf{B}_2^{\mathrm{H}}$.

Rate constraints in the MA-phase Upon reception of the signal from each node, using successive interference cancellation, the relays decode the messages from the base stations and the terminals based on separated received signals in (6.17). The corresponding achievable are then given by the achievable rates of the MA channel discussed in Section 4.2 (one MA channel for each relay).

$$\mathcal{C}_{\mathrm{DF,MAC}}^{\mathrm{S\text{-}FW}} = \left\{ \begin{array}{l} R_{i,i} \leq \tau \log_2 \left| \mathbf{I}_M + \mathbf{C}_{\mathbf{n},i,i}^{-1} \mathbf{Q}_{i,i} \right| \\ R_{B_i} \leq \tau \log_2 \left| \mathbf{I}_M + \mathbf{C}_{\mathbf{n},i,i}^{-1} \mathbf{Q}_{B_i} \right| \\ R_{i,i} + R_{B_i} \leq \tau \log_2 \left| \mathbf{I}_M + \mathbf{C}_{\mathbf{n},i,i}^{-1} (\mathbf{Q}_{i,i} + \mathbf{Q}_{B_i}) \right| \\ R_{i,c} \leq \tau \log_2 \left| \mathbf{I}_M + \mathbf{C}_{\mathbf{n},i,c}^{-1} \mathbf{Q}_{i,c} \right| \\ R_{B_c,i} \leq \tau \log_2 \left| \mathbf{I}_M + \mathbf{C}_{\mathbf{n},i,c}^{-1} \mathbf{Q}_{B_c,i} \right| \\ R_{i,c} + R_{B_c,i} \leq \tau \log_2 \left| \mathbf{I}_M + \mathbf{C}_{\mathbf{n},i,c}^{-1} (\mathbf{Q}_{i,c} + \mathbf{Q}_{B_c,i}) \right| \end{array} \right\}, \tag{6.18}$$

under the power constraints

$$\operatorname{tr}(\mathbf{Q}_i) \leq P_i, \qquad \operatorname{tr}\left(\mathbf{H}_{B_i}^{-1} \mathbf{H}_i \mathbf{Q}_{B_i} \mathbf{H}_i^{\mathrm{H}} \mathbf{H}_{B_i}^{-\mathrm{H}}\right) \leq P_{B_i}, \tag{6.19}$$

and

$$\operatorname{tr}\left(\mathbf{x}_{B_c} \mathbf{x}_{B_c}^{\mathrm{H}}\right) = \mathbf{G}_{B_c,1} \mathbf{Q}_{B_c,1} \mathbf{G}_{B_c,1}^{\mathrm{H}} + \mathbf{G}_{B_c,2} \mathbf{Q}_{B_c,2} \mathbf{G}_{B_c,2}^{\mathrm{H}} \leq P_{B_c}. \tag{6.20}$$

Broadcast-phase After decoding the messages from each node, each signal pairs $(\mathbf{u}_{B_i}, \mathbf{u}_{i,i})$ and $(\mathbf{u}_{B_c,i}, \mathbf{u}_{c,i})$, with $i \in \{1, 2\}$, are reencoded to the codewords $\mathbf{u}_{R_i,i} \in \mathbb{C}^{M\times 1}$ and $\mathbf{u}_{R_i,c} \in \mathbb{C}^{M\times 1}$ using capacity achieving codes for the Gaussian MIMO channel. The covariance matrices of these codewords are denoted $\mathbf{Q}_{R_i,i}$, $\mathbf{Q}_{R_i,c}$. Moreover, define $\mathbf{u}_{R_i} = [(\mathbf{u}_{R_i,i})^{\mathrm{T}}, (\mathbf{u}_{R_i,c})^{\mathrm{T}}]^{\mathrm{T}}$ and denote the corresponding covariance matrix by $\mathbf{Q}_{R_i} = \begin{bmatrix} \mathbf{Q}_{R_i,i} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{R_i,c} \end{bmatrix}$. Relay i then transmits the codeword $\mathbf{u}_{R_i}$ through the precoding matrix $\mathbf{G}_{R_i} \in \mathbb{C}^{2M\times 2M}$, and the transmissions during the BC-phase, at the relays, have to satisfy the power constraints $\operatorname{tr}\left(\mathbf{G}_{R_i} \mathbf{Q}_{R_i} \mathbf{G}_{R_i}^{\mathrm{H}}\right) \leq P_{R_i}$.

At the relays, the relay precode matrices $\mathbf{G}_{R_i}$ are chosen as for the DF-A and DF-U for the shared relay. That is, the relay transmits

$$\mathbf{x}_{R_i} = \mathbf{G}_{R_i} \mathbf{u}_{R_i} = \tilde{\mathbf{H}}_i^{-\mathrm{T}} \mathbf{u}_{R_i}. \tag{6.21}$$

Consequently, with zero-forcing precoding as in (6.21), the terminals receive

$$\mathbf{y}_{i,i} = \mathbf{H}_{i,i}^{\mathrm{T}} \mathbf{x}_{R_i} + \mathbf{z}_{i,i} = \mathbf{H}_{i,i}^{\mathrm{T}} \tilde{\mathbf{H}}_i^{-\mathrm{T}} \mathbf{u}_{R_i} + \mathbf{z}_{i,i} = \mathbf{u}_{R_i,i} + \mathbf{z}_{i,i} \tag{6.22}$$

$$\mathbf{y}_{i,c} = \mathbf{H}_{i,c}^{\mathrm{T}} \mathbf{x}_{R_i} + \mathbf{z}_{i,c} = \mathbf{H}_{i,c}^{\mathrm{T}} \tilde{\mathbf{H}}_i^{-\mathrm{T}} \mathbf{u}_{R_i} + \mathbf{z}_{i,c} = \mathbf{u}_{R_i,c} + \mathbf{z}_{i,c}. \tag{6.23}$$

Note that terminal i only receives the codeword associated with the base station-terminal pair i due to the zero forcing precoding matrix.

Since the base stations B_1 and B_2 have $2M$ antennas, the desired communication flow at base station i can be obtained as

$$\mathbf{y}_{B_i} = \mathbf{H}_{i,i}^{\mathrm{T}} \mathbf{H}_{B_i}^{-\mathrm{T}} \left(\mathbf{H}_{B_i}^{\mathrm{T}} \tilde{\mathbf{H}}_i^{-\mathrm{T}} \mathbf{x}_{R_i} + \mathbf{z}_{B_i} \right) \tag{6.24}$$

$$= \mathbf{u}_{R_i,i} + \mathbf{H}_{i,i}^{\mathrm{T}} \mathbf{H}_{B_i}^{-\mathrm{T}} \mathbf{z}_{B_i}. \tag{6.25}$$

The noise covariance matrix of this signal is given by $\mathbf{C}_{\mathbf{n},B_i} = \mathbf{H}_{i,i}^{\mathrm{T}}\mathbf{H}_{B_i}^{-\mathrm{T}}\mathbf{H}_{B_i}^{-*}\mathbf{H}_{i,i}^{*}$.

At the common base station B_c, a superposition of the signals from the relays is received. To only extract the spatial dimensions used for communication with $T_{1,c}$ and $T_{2,c}$, $\mathbf{G}_{B_c}^{\mathrm{T}}$ multiplied on the received signal to match the alignment

$$\tilde{\mathbf{y}}_{B_c} = \mathbf{G}_{B_c}^{\mathrm{T}}(\mathbf{H}_{B_1,c}^{\mathrm{T}}\tilde{\mathbf{H}}_1^{-\mathrm{T}}\mathbf{x}_{R_1} + \mathbf{H}_{B_2,c}^{\mathrm{T}}\tilde{\mathbf{H}}_2^{-\mathrm{T}}\mathbf{x}_{R_2} + \mathbf{z}_{B_c}) \tag{6.26}$$

$$= \begin{bmatrix} \mathbf{x}_{R_1,c} \\ \mathbf{x}_{R_2,c} \end{bmatrix} + \mathbf{G}_{B_c}^{\mathrm{T}}\mathbf{z}_{B_c} \tag{6.27}$$

The noise covariance matrix of this signal is thus given by $\mathbf{C}_{\mathbf{n},B_c} = \mathbf{G}_{B_c}^{\mathrm{T}}\mathbf{G}_{B_c}^{*}$.

Finally, the power constraints in (6.4) can be expressed in terms of the relay precode matrix

$$\begin{aligned}
&\mathrm{tr}\left(\mathbf{G}_{R_i}\mathbf{Q}_R\mathbf{G}_{R_i}^{\mathrm{H}}\right) \\
&\quad = \mathrm{tr}\left(\tilde{\mathbf{H}}_i^{-\mathrm{T}}\mathbf{B}_1^{\mathrm{T}}\mathbf{Q}_{R_i,i}\mathbf{B}_1\tilde{\mathbf{H}}_i^{-*}\right) + \mathrm{tr}\left(\tilde{\mathbf{H}}_i^{-\mathrm{T}}\mathbf{B}_2^{\mathrm{T}}\mathbf{Q}_{R_i,c}\mathbf{B}_2\tilde{\mathbf{H}}_i^{-*}\right) \le P_{R_i}.
\end{aligned} \tag{6.28}$$

Rate constraint in the BC-phase The rates for the communication flows between the base stations B_i and terminals $T_{i,i}$, with $i \in \{1, 2\}$, are identical to those for DF-A, i.e. [34, 35]

$$\mathcal{C}_{\mathrm{DF,BC}}^{\text{S-FW-1}} = \left\{ \begin{array}{c} R_{i,i} \le (1-\tau)\log_2\left|\mathbf{I} + \mathbf{C}_{\mathbf{n},B_i}^{-1}\mathbf{Q}_{R_i,i}\right| \\ R_{B_i} \le (1-\tau)\log_2\left|\mathbf{I} + \mathbf{Q}_{R_i,i}\right| \end{array} \right\} \tag{6.29}$$

However, for the common base station, the channel is essentially a combined MA channel and broadcast channel with side information. The MA channel is introduced because the common base station receives a superposition of the signals from both relays, while the channels seen from the relays are broadcast channels with side information. This corresponds to the four-way MIMO relay channel described in Section 4.5, where it was shown that the following achievable rate region can be achieved using successive interference cancellation and time-sharing [3]

$$\begin{aligned}
&\mathcal{C}_{\mathrm{DF,BC}}^{\text{S-FW-2}} \\
&= \left\{ \begin{array}{rl}
R_{i,c} & \le (1-\tau)\log_2\left|\mathbf{I} + \mathbf{C}_{\mathbf{n},B_c}^{-1}\mathbf{B}_i^{\mathrm{H}}\mathbf{Q}_{R_i,c}\mathbf{B}_i\right| \\
R_{B_c,i} & \le (1-\tau)\log_2\left|\mathbf{I} + \mathbf{Q}_{R_i,c}\right| \\
R_{1,c} + R_{2,c} & \le (1-\tau)\log_2\left|\mathbf{I} + \mathbf{C}_{\mathbf{n},B_c}^{-1}\mathbf{B}_1^{\mathrm{H}}\mathbf{Q}_{R_1,c}\mathbf{B}_1 + \mathbf{C}_{\mathbf{n},B_c}^{-1}\mathbf{B}_2^{\mathrm{H}}\mathbf{Q}_{R_2,c}\mathbf{B}_2\right|
\end{array} \right. .
\end{aligned} \tag{6.30}$$

6.2.1 Problem formulation

Based on the achievable rate regions for the MA- and BC-phase in given by $\mathcal{C}_{\mathrm{DF,MA}}^{\text{S-FW}}$ and $\mathcal{C}_{\mathrm{DF,BC}}^{\text{S-FW-1}} \cap \mathcal{C}_{\mathrm{DF,BC}}^{\text{S-FW-2}}$, respectively. As for the shared relay, WMR, with the achievable rates weighted by the tuple $(w_{1,1}, w_{2,2}, w_{1,c}, w_{2,c}, w_{B_1}, w_{B_2}, w_{B_c,1}, w_{B_c,2})$, is maximized over the covariance matrices and the division constant τ between the MA- and BC-phase. The general optimization problem can be formulated as

$$\max_{\substack{\mathbf{Q}_{R_i,i},\mathbf{Q}_{R_i,c},\mathbf{Q}_{i,i},\mathbf{Q}_{i,c}\\ \mathbf{Q}_{B_i},\mathbf{Q}_{B_c,i}\\ \tau,R_{i,i},R_{i,c},R_{B_i},R_{B_c,i}}} \quad \min(w_1 R_1, w_2 R_2, w_{B_1} R_{B_1}, w_{B_2} R_{B_2}) \tag{6.31a}$$

$$\text{s.t.} \quad (R_{i,i}, R_{i,c}, R_{B_i}, R_{B_c,i}) \in \mathcal{C}^{\text{S-FW}}_{\text{DF,MAC}} \cap \mathcal{C}^{\text{S-FW-1}}_{\text{DF,BC}} \cap \mathcal{C}^{\text{S-FW-2}}_{\text{DF,BC}}, \tag{6.31b}$$

$$(6.28), \mathbf{Q}_{i,i}, \mathbf{Q}_{i,c}, \mathbf{Q}_{B_i}, \mathbf{Q}_{B_c,i}, \mathbf{Q}_{R_i,i}, \mathbf{Q}_{R_i,c} \succeq 0 \tag{6.31c}$$

$$\operatorname{tr}(\mathbf{Q}_{i,i}) \le P_{i,i}, \operatorname{tr}(\mathbf{Q}_{i,c}) \le P_{i,c} \tag{6.31d}$$

$$\operatorname{tr}(\mathbf{Q}_{B_i}) \le P_{B_i} \operatorname{tr}(\mathbf{Q}_{B_c,i}) \le P_{B_c,i}. \tag{6.31e}$$

This problem is not convex. As for the shared relay and as in [3], the problem is relaxed using alternate optimization in order to obtain a local minimum [3]. That is, the optimization is solved by alternating between the following optimization problems

1. Optimize (6.31) with respect to the covariance matrices $\mathbf{Q}_{R_i,i}, \mathbf{Q}_{R_i,c}, \mathbf{Q}_{i,i}, \mathbf{Q}_{i,c}, \mathbf{Q}_{B_i}, \mathbf{Q}_{B_c,i}$ with τ fixed.
2. Optimize (6.31) with respect to τ with the covariance matrices $\mathbf{Q}_{R_i,i}, \mathbf{Q}_{R_i,c}, \mathbf{Q}_{i,i}, \mathbf{Q}_{i,c}, \mathbf{Q}_{B_i}, \mathbf{Q}_{B_c,i}$ fixed.

Each of these problems are convex, and are hence solved using CVX in MATLAB. In the practical implementation, τ is initially set to $\frac{1}{2}$ and the optimization start with step 1.

6.3. Numerical Results

To assess the performance of the shared four-way relay, the WMR has been simulated in terms of SNR, the number of antennas M and SNR. The main objective of the chapter is to demonstrate that the idea of alignment and zero-forcing indeed scales to larger networks, and in this case the shared four-way relay channel, without significant performance degradations. To evaluate how well the scheme scales, the results are compared to the shared relay with the relaying scheme DF-A. In order to generate comparable results, one unit of power is given to each of the nodes per node it communicates with, i.e. the relays each get 4 units of power since they serve four nodes, the common base station is given 2 units and the remaining nodes are given 1 unit. For this scenario, the performance of the shared relay and the shared four-way relay are expected to be similar.

The simulations are performed in the same manner as in Section 5.4. The results are shown in Fig. 6.2. It is seen that DF-A for the shared relay achieves slightly higher WMR than the shared four-way relay. One contribution to the higher WMR at high SNR for the shared relay is explained by the lower probability of having a bad channel matrix in the relay network.

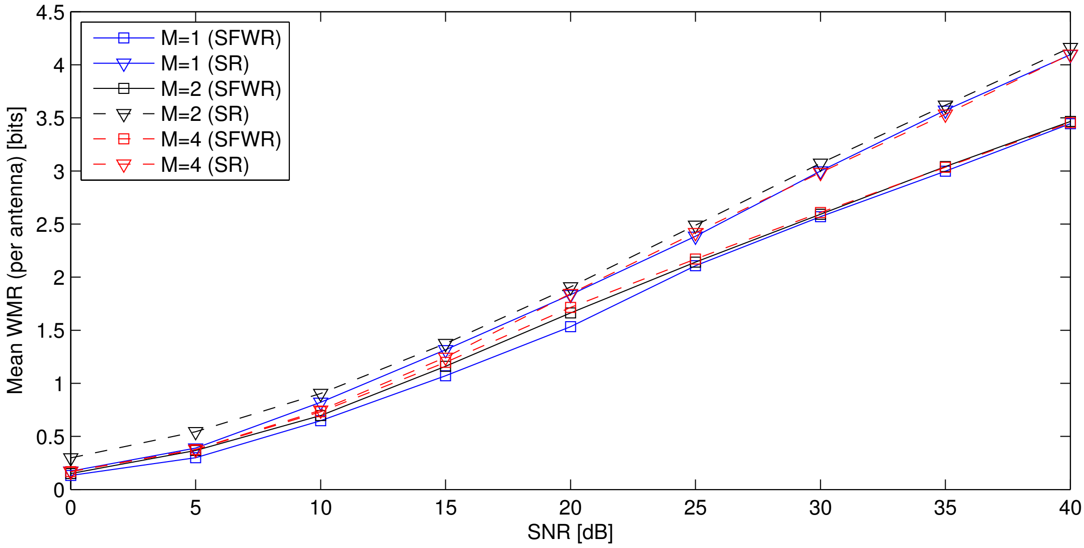


Figure 6.2: *WMR comparison between the DF scheme for the shared four-way relay (SFWR) and DF-A for the shared relay (SR). All weights are set to 1, and the powers for the shared four-way relay are set as* $P_{1,1} = P_{2,2} = P_{1,c} = P_{2,c} = P_{B_1} = P_{B_2} = \frac{1}{2}P_{B_c} = \frac{1}{4}P_{R_1} = \frac{1}{4}P_{R_2} = SNR$ *and for the shared relay,* $P_1 = P_2 = P_{B_1} = P_{B_2} = \frac{1}{4}P_R = SNR$.

CHAPTER 7

Discussion

The previous two chapters proposed a number of relaying schemes for the shared relay and the shared four-way relay. This chapter briefly discusses the schemes and results.

The main objective of the chapters have been show that alignment and zero-forcing can be used to enable efficient two-way communication in relay networks more complex than the traditional two-way relay channel. The vision of this idea is that general relay networks can be decomposed into a number of smaller relay networks as in the example in Section 1.2, which are then solved separately. In this thesis, the Weighted Minimum Rate fairness metric has been used to evaluate to the communication schemes. Another common fairness metric is the weighted sum rate, which does not ensure that each communication flow is given a certain fraction of the total performance. As opposed to the WMR, the weighted sum rate better evaluates the average performance of a scheme since a single weak channel matrix does not have the same effect on the metric. Further work may explore the performance using this metric.

In this treatment, the base stations and the relay have exactly $2M$ antennas, twice the number of the terminals. In practice, other combinations of antennas might be used. For example, if the base stations and the relay have fewer than twice the antennas of the terminals, a block diagonalization approach as in [38] may used for alignment. Such approach leads to a lower number of DoF since exactly M of the spatial dimensions at the base stations need to be devoted for alignment. Similarly, if the base stations and relay have more than $2M$ antennas, performance is increased, but the performance is still limited by number of DoF between relay and the terminals, and hence only the transmit diversity between the base stations and the relay is increased.

Another important assumption is that each node have full channel state information which is used to coordinate their transmissions. It is clear that this requirement is fundamental for the functionality of the proposed schemes. Lack of channel state information clearly implies that alignment can not be achieved and self-interference can not be cancelled. However, the assumption is widely used and is satisfied for the Rayleigh block fading channels, where the channels can be periodically estimated.

For the shared relay channel, the presented schemes have clear advantages and disadvantages. As was shown, PF achieves the maximum number of DoF, and hence performs well at high SNR. Similarly, at low SNR, DF-A and DF-U achieved significant better results and may therefore be more applicable. Besides the achievable rates, the complexity of the schemes is also of interest. PF essentially operates in a channel use-by-channel use fashion, and the relay mapping is hence implemented efficiently. This also implies that PF need the same number of channel uses in the MA and BC-phases, which may be suboptimal. Moreover, encoding and decoding of the signals is only done at the base stations and the terminals, and hence codebooks used by the terminals and base stations should not be implemented in the relay. In contrast, DF-A and DF-U requires relay mappings that are significantly more complex. At the relay, in the MA-phase, the received signals first have to be decoded using a combination of capacity achieving codes and successive interference cancellation. Likewise, reenconding of the signals using capacity achieving codes is required in the BC-phase.

The DF-A has the advantage over DF-U that successive interference cancellation only have to done among two nodes as opposed to all four nodes. Although the achievable rate regions for DF-U are easily found, successive interference cancellation among more nodes comes at the cost of more significant performance degradations due to unideal impairments in practical systems [36, 37]. Consequently, the performance degradations are expected to be more significant for DF-U.

Finally, in Chapter 6, a DF scheme was designed for the shared four-way relay and it was shown that the performance of the DF scheme was only slightly worse than DF-A for the shared relay channel. This indicates that the ideas of alignment indeed scales to larger networks. Furthermore, in this chapter, only the DF scheme using alignment was developed, However, due to simplicity introduced by alignment and zero-forcing, PF and an unaligned DF scheme may clearly be extended to the shared four-way relay channel.

Part III

Closing

CHAPTER 8

Conclusion

In this thesis, two-way relaying schemes that are based on a MA- and a BC-phase have been studied. The main objective of the thesis is two-fold; (1) to investigate the basic scheme, Amplify-and-Forward, to determine whether higher achievable rates can be achieved using interaction, and (2) to extend the basic schemes like AF and DF to more advanced relay networks using alignment and zero-forcing.

In the first part, AF for the traditional two-way relay channel is investigated. Two interactive extensions of AF, that are based on the fact that the terminals transmit correlated information, are introduced. The achievable symmetric rates are derived for the schemes and are compared to the traditional AF scheme. However, simulations over a range power values at the terminals and the relay shows that no performance improvements are obtained. This result however does not definitively imply that interaction can not be used to improve relaying schemes. A valid direction to take as future work, is to assume that the rates can not be improved and attempt to follow the approach in [1] to find an upper bound of the achievable rates, and show that these coincides with the achievable rates of a known scheme.

The second part of the thesis, concerns coordinated relaying. The key idea of this concept is that each node in the network coordinate their transmissions and cooperate in order to achieve the highest achievable rates. In this thesis, emphasis is put on the two techniques: zero-forcing and alignment. These techniques are used to design efficient relaying schemes for the shared relay channel and the shared four-way relay channel. Essentially, when these techniques are combined each two-way communication flow between the nodes can be separated and processed independently at the relays.

For the shared relay, the three relaying schemes PF, DF-A and DF-U are put forth. For PF, the WMR is optimized with respect to the transmit covariance matrices and the precode matrix at the relay. In order to perform this optimization, the optimization problem is relaxed by the use alternate optimization and a simplified relay precode matrix. The number of achievable DoF for the scheme is further shown to be the maximum, i.e. $\frac{1}{2}M$ per communication flow, and hence the PF operate optimally at very high SNR. Similarly, the WMR using DF-A and DF-U is optimized using alternate optimization between the transmit covariance matrices and the division constant dividing the channel uses between the MA and BC-phase. Finally, by simulation it is found that PF performs well at high SNR, while DF-A and DF-U perform better at low SNR. The shared four-way relay serves as an example of a more advanced network, in which two relays, three base stations and four terminals are involved. Using zero-forcing and alignment, a DF scheme is proposed, and by simulation, it is shown that the performance is only slightly degraded compared to the shared relay. Based on these results, it is concluded that zero-forcing and alignment indeed does scale to larger networks such as the shared four-way relay.

Appendices

APPENDIX A

Acronyms

This appendix contains a list of abbreviations and terms used throughout the report. The list is made to help the reader gain a better understanding of the report content, and should therefore be carefully read before reading the report itself.

AF Amplify-and-Forward

AWGN Additive White Gaussian Noise

BC broadcast

CSI Channel State Information

dB decibels

DF Decode-and-Forward

DF-A Decode-and-Forward with Aligned MA-phase

DF-U Decode-and-Forward with Unaligned MA-phase

DoF Degrees of Freedom

EVD eigenvalue decomposition

FDMA Frequency Division Multiple Access

iid identical and independently distributed

KKT Karush-Kuhn-Tucker

LDPC Low-Density Parity Check

MA multiple access

MACPF Multiple Access Channel with Perfect Feedback

MIMO Multiple-Input Multiple-Output

MMSE Minimum Mean Square Error

PF Precode-and-Forward

QCQP Quadratically Constrained Quadratic Program

QoS Quality of Service

SISO Single-Input Single-Output

SNR	Signal to Noise Ratio
SVD	singular value decomposition
TDMA	Time Division Multiple Access
WMR	Weighted Minimum Rate

APPENDIX B

List of Symbols

B.1 Two-way Relaying with Interactive Communication

$\bar{j}$	$\bar{j} = 1$ if $j = 0$ and 1 otherwise.
$I(x; y)$	Denotes the mutual information between the random variables x and y distributed according to the joint distribution $p(x, y)$.
$I(x; y\|z)$	Denotes the conditional mutual information between the random variables x and y given z distributed according to the conditional joint distribution $p(x, y\|z)$.
$\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \mathbf{C})$	denotes the multivariate Gaussian distribution with mean μ and covariance matrix $\mathbf{C}$.
$\overset{\text{iid}}{\sim}$	Means "iid according to"
$f^{(n)}$	Encoding function; encodes messages to codewords.
$g^{(n)}$	Relay mapping; maps received signal at the relay $\mathbf{y}_R$ to the signal to be transmitted $\mathbf{x}_R$.
$\hat{m}^{(n)}$	Decoding function; maps received signal at a terminal to the desired message.
$\hat{\ }$	Denotes a signal in the subspace $S_{(n)}$.
$\check{\ }$	Denotes a signal in the subspace $S_{(c)}$.
$\mathbf{U}_{(n)}$ and $\mathbf{U}_{(c)}$	Denote the matrices spanning the orthogonal subspaces $S_{(n)}$ and $S_{(c)}$. Moreover, $[\mathbf{U}_{(n)}, \mathbf{U}_{(c)}]$ is unitary.
$\mathbf{C}_{(c)}$	Benotes a compression matrix of dimension $\overline{\alpha}n \times \alpha n$.

B.2 Coordinated Relaying

Note that <> is used as placeholder for another symbol..

$\bar{j}$	$\bar{j} = 1$ if $j = 0$ and 1 otherwise.
$I(x; y)$	Denotes the mutual information between the random variables x and y distributed according to the joint distribution $p(x, y)$.
$I(x; y\|z)$	Denotes the conditional mutual information between the random variables x and y given z distributed according to the conditional joint distribution $p(x, y\|z)$.
$\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \mathbf{C})$	Denotes the multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$.
$\overset{\text{iid}}{\sim}$	Means "iid according to"
$\|\mathbf{A}\|$	Denotes the determinant of the matrix $\mathbf{A}$.
$\mathrm{tr}\,(\mathbf{A})$	The trace operater; computes the sum of the diagonal matrices of $\mathbf{A}$.
$\mathbf{I}_n$	Denotes the identity matrix of dimension n.
$\mathbf{B}_i \in \mathbb{R}^{M \times 2M}$	Given by $\mathbf{B}_1 = [\mathbf{I}_M, \mathbf{0}]$ and $\mathbf{B}_2 = [\mathbf{0}, \mathbf{I}_M]$
$f^{(n)}_{<>}$	Encoding function; encodes messages to codewords.
$g^{(n)}_{<>}$	Relay mapping; maps received signal at the relay $\mathbf{y}_R$ to the signal to be transmitted $\mathbf{x}_R$.
$\hat{m}^{(n)}_{<>}$	Decoding function; maps received signal at a terminal to the desired message.
$\mathbf{H}_{<>}$	Denotes channel matrices
$\tilde{\mathbf{H}}_{<>}$	Denote effective channel matrices, i.e. for the shared relay $\tilde{\mathbf{H}} = [\mathbf{H}_1, \mathbf{H}_2]$
$\mathbf{C}_{<>}$	Denotes noise covariance matrices.
$\mathbf{Q}_{<>}$	Denote covariance matrices of signals and codewords (i.e. without noise).
$\mathbf{A}_1$ and $\mathbf{A}_2$	Denote the amplification matrices used for PF.
$\mathbf{A}$	Denotes the matrix $\begin{bmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{A}_2 \end{bmatrix}$.
$\tilde{\mathbf{I}}$	Denote the matrix $\begin{bmatrix} \mathbf{I}_M & \mathbf{I}_M & 0 & 0 \\ 0 & 0 & \mathbf{I}_M & \mathbf{I}_M \end{bmatrix}$.
$\mathbf{G}_{<>}$	Denotes precode matrices at the base stations or the relay.
τ	Division constant dividing the n channel uses between the MA- and BC-phase used for the DF schemes.

Bibliography

[1] Z. Cheng and N. Devroye, "Two-way networks: when adaptation is useless," 2012, available on arXiv.

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

[3] H. Liu, F. Sun, E. D. Carvelho, P. Popovski, H. Thomsen, and Y. Zhao, "Mimo four-way relaying," in *IEEE Workshop on Signal Processing Advances in Wireless Communications*, 2013, pp. 1–6.

[4] H. Liu, P. Popovski, E. de Carvalho, Y. Zhao, F. Sun, and C. D. T. Thai, "Transmission schemes for four-way relaying in wireless cellular systems," *IEEE Communication Letters*, 2012.

[5] B. Nazer and M. Gastpar, "Reliable physical layer network coding," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 438–460, 2011.

[6] E. C. V. D. Meulen, "Three-terminal communication channels," *Applied Probability Trust*, 1971.

[7] P. Popovski and T. Koike-Akino, "Coded bidirectional relaying in wireless networks," in *Advances in Wireless Communications.* Springer, 2009.

[8] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *IEEE International Conference on Communications*, june 2007, pp. 707 –712.

[9] ——, "Bi-directional amplification of throughput in a wireless multi-hop network," in *IEEE 63rd Vehicular Technology Conference*, vol. 2, may 2006, pp. 588 –593.

[10] C. Schnurr, S. Stanczak, and T. J. Oechtering, "Achievable rates for the restricted half-duplex two-way relay channel under a partial-decode-and-forward protocol," in *IEEE Information Theory Workshop*, may 2008, pp. 134 –138.

[11] A. E. Gamal and Y.-H. Kim, *Network Information Theory.* Cambridge, 2011.

[12] C. D. T. Thai and P. Popovski, "Coordinated direct and relay transmission with interference cancelation in wireless systems," *IEEE Communications Letters*, vol. 15, no. 4, pp. 416–418, 2011.

[13] D. Gunduz, A. Yener, A. Goldsmith, and H. V. Poor, "The multi-way relay channel," in *IEEE International Symposium on Information Theory*, 2009, pp. 339–343.

[14] F. Wang, S. Chang, and D. Guo, "Wireless mimo switching with zero-forcing relaying and network-coded relaying," 2012, available on arXiv.

[15] E. Teletar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, nov 1999.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, 2006.

[17] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, 2011.

[18] T. Cover and C. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Transactions on Information Theory*, vol. 27, no. 3, pp. 292 – 298, may 1981.

[19] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–i: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 172 – 182, apr 1966.

[20] Z. Cheng and N. Devroye, "On the capacity of multi-user two-way linear deterministic channels," in *IEEE International Symposium on Information Theory*, 2012, pp. 2092–2096.

[21] S. Changho, W. I-Hsiang, and D. Tse, "Two-way interference channels," in *IEEE International Symposium on Information Theory*, 2012, pp. 2801–2805.

[22] U. Erez and R. Zamir, "Achieving 1/2 log (1+snr) on the awgn channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293 – 2314, oct. 2004.

[23] A. Lozano and N. Jindal, "Transmit diversity vs. spatial multiplexing in modern mimo systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 186–197, 2010.

[24] T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge University Press, 2008.

[25] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of mimo channels," pp. 684–702, 2003.

[26] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1859–1874, 2004.

[27] H. Ghozlan, Y. Mohasseb, and H. E.-G. G. Kramer, "The mimo wireless switch: Relaying can increase the multiplexing gain," in *IEEE International Symposium on Information Theory*, 2009, pp. 1448–1452.

[28] R. F. Wyrembelski, T. J. Oechtering, I. Bjelakovic, C. Schnurr, and H. Boche, "Capacity of gaussian mimo bidirectional broadcast channels," in *IEEE International Symposium on Information Theory*, 2008, pp. 584–588.

[29] W. Rui and T. Meixia, "Joint source and relay precoding designs for mimo two-way relay systems," in *IEEE International Conference on Communications*, 2011, pp. 1–5.

[30] R. Yue, "Joint source and relay optimization for two-way linear non-regenerative mimo relay communications," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6533–6546, 2012.

[31] T. Meixia and W. Rui, "Linear precoding for multi-pair two-way mimo relay systems with max-min fairness," *Signal Processing, IEEE Transactions on*, vol. 60, no. 10, pp. 5361–5370, 2012.

[32] X. Shengyang and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way mimo relay system," *IEEE Transactions on Wireless Communications*, vol. 10, no. 5, pp. 1645–1655, 2011.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[34] Z. Jian, M. Kuhn, A. Wittneben, and G. Bauch, "Optimum time-division in mimo two-way decode-and-forward relaying systems," in *Asilomar Conference on Signals, Systems and Computers*, 2008, pp. 1494–1500.

[35] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *IEEE Information Theory Workshop*, 2007, pp. 313–318.

[36] S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi, "Successive interference cancellation: a back-of-the-envelope perspective," in *Proceedings of the 9th Workshop on Hot Topics in Networks*, ser. Hotnets-IX. New York, NY, USA: ACM, 2010, pp. 17:1–17:6. [Online]. Available: http://doi.acm.org/10.1145/1868447.1868464

[37] X. Xu and N. Goertz, "Practical successive interference cancellation in the binary-input gaussian multiple-access channel," in *International ITG Conference on Source and Channel Coding*, 2008, pp. 1–6.

[38] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, 2004.