On the Optimal Transmission Strategies for Sources without Channel State Information

by

Vahid Pourahmadi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

With the growth of multimedia services, it is essential to find new transmission schemes to support higher data rates in wireless networks. In this thesis, we study networks in which the Channel State Information (CSI) is only available at the destination. We focus on the analysis of three different network setups. For each case, we propose a transmission scheme which maximizes the average performance of the network.

The first scenario, which is studied in Chapter 2, is a multi-hop network in which the channel gain of each hop changes quasi-statically from one transmission block to the other. Our main motivation to study this network is the recent advances in deployment of relay nodes in wireless networks (e.g., LTE-A and IEEE 802.16j). In this setup, we assume that all nodes are equipped with a single antenna and the relay nodes are not capable of data buffering over multiple transmission blocks. The proposed transmission scheme is based on infinite-layer coding at all nodes (the source and all relays) in conjunction with the Decode-and-Forward (DF) relaying. The objective is to maximize the statistical average of the received rate per channel use at the destination. To find the optimal parameters of this code, we first formulate the problem for a two-hop scenario and describe the code design algorithm for this two-hop setting. The optimality of infinite-layer DF coding is also discussed for the case of two-hop networks. The result is then generalized to multi-hop scenarios. To show the superiority of the proposed scheme, we also evaluate the achievable average received rate of infinite-layer DF coding and compare it with the performance of previously known schemes.

The second scenario, studied in Chapter 3, is a single-hop network in which both nodes are equipped with multiple antennas, while the channel gain changes quasi-statically and the CSI is not available at the source. The main reason for selecting this network setup is to study the transmission of video signals (compressed using a scalable video coding

technique, e.g., SVC H.264/AVC) over a Multiple-Input Multiple-Output (MIMO) link. In this setup, although scalable video coding techniques compress the video signal into layers with different importance (for video reconstruction), the source cannot adapt the number of transmitted layers to the capacity of the channel (since it does not have the CSI in each time slot). An alternative approach is to always transmit all layers of the compressed video signal, but use unequal error protection for different layers. With this motivation, we focus on the design of multilayer codes for a MIMO link in which the destination is only able to perform successive decoding (not joint-decoding). In this chapter, we introduce a design rule for construction of multilayer codes for MIMO systems. We also propose a algorithm that uses this design rule to determine the parameters of the multilayer code. The performance analysis of the proposed scheme is also discussed in this chapter.

In the two previous scenarios, the ambiguity of the source regarding the channel state comes from the fact that the channel gains randomly change in each transmission block and there is no feedback to notify the source about the current state of the channel. Apart from these, there are some scenarios in which the channel state is unknown at the source, even though the channel gain is fixed and the source knows its value. The third scenario of this thesis presents an example of such network setups. More precisely, in Chapter 4, we study a multiple access network with K users and one Access Point (AP), where all nodes are equipped with multiple antennas. To access the network, each user independently decides whether to transmit in a time slot or not (no coordination between users). Considering a two-user random access network, we first derive the optimal value of network average Degrees of Freedom (DoF) (introduced in Section 4.1). Generalizing the result to multiuser networks, we propose an upper-bound for the network average DoF of a K-user random access network. This upper-bound is then analyzed for different network configurations to identify the network classes in which the proposed upper-bound is tight. It is also shown that simple single-stream data transmission achieves the upper-bound in most network settings. However, for some network configurations, we need to apply multi-stream data transmission in conjunction with interference alignment to reach the upper-bound. Some illustrative examples are also presented in this chapter.

Acknowledgements

First, I want to express my deep gratitude to my supervisor, Professor Amir K. Khandani for providing guidance throughout my PhD studies. He is a brilliant Professor with solid theoretical and practical knowledge. I have benefited tremendously from the opportunities that he provided for me and I am really thankful for all his helps.

I would like to thank the members of my dissertation committee, Professors Ravi Adve, Ali Ghodsi, Liang-Liang Xie, and Murat Uysal, who despite their busy schedules accepted to review my thesis and providing me with valuable suggestions. I also thankful to Professor Mohamed Oussama Damen for attending my PhD examination as the delegate of one of the committee members who was not able to attend.

I have been very fortunate to work with members of Coding and Signal Transmission (CST) Lab. In particular, I acknowledge Seyed Abolfazl Motahari for his support and encouragements. I would like to thank Behzad Nourani who patiently helped me in many different areas during the three years that we shared our office. I am also grateful to my friends, Hossin Bagheri, Afshar Ganjali, Masoud Ebrahimi, Majid Safari, Ali Zibaeenejad, Pooya Mahboubi, Jamshaid Abouei, Ali Ahmadzadeh, and Dr. Sharifkhani for their advice and support.

Finally, and most importantly, I thank my parents, for their endless love and support, without which I would not have succeeded. I also want to thank my dear sister, Zahra, for her continual inspiration and support. My special thanks and sincere appreciation are extended to the love of my life, my wife, Fathiyeh Faghih Khorasani, for all her love and understanding during the course of my PhD.

Dedication

To my parents:

Tahere Labafian *and* *Ahmad Pourahmadi*

and

To my beloved wife:

Fathiyeh Faghih Khorasani

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BS	Base Station
AP	Access Point
MAC	Multiple Access Channel
BC	Broadcast Channel
AWGN	Additive White Gaussian Noise
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
SISO	Single-Input Single-Output
CSI	Channel State Information
SNR	Signal-to-Noise Ratio
SINR	Signal-to-Interference-plus-Noise Ratio
DoF	Degrees Of Freedom
DMT	Diversity-Multiplexing Trade-off
pdf	Probability Density Function
cdf	Cumulative Density Function
SVC	Scalable Video Code
QoS	Quality of Service
MRC	Maximum Ratio Combining
HARQ	Hybrid Automatic Retransmission Request
LTE-A	Long Term Evolution - Advanced
DF	Decode and Forward
AF	Amplify and Forward
CF	Compress and Forward

Notation

Boldface Upper-Case Letters	Matrices
Boldface Lower-Case Letters	Vectors
$\mathbf{A}^{\dagger}$	Hermitian transpose of A
$\lvert\mathbf{A}\rvert$	Determinant of the matrix $\mathbf{A}$
$E[X]$	The expectation of the random variable X
x^{*}	The optimal solution to an optimization problem
$\mathcal{CN}(m, \sigma^2)$	Complex Gaussian distribution with mean m and variance σ^2

Chapter 1

Introduction

The rapid development and diffusion of multimedia applications motivate many researchers to study new and efficient schemes for data transmission. The importance of such studies is further enhanced when we aim to provide the multimedia services to mobile users via wireless channels where the channel gains are subject to change over time.

The studies related to the analysis of wireless networks with time-varying channels can be categorized based on their assumption regarding the amount of knowledge that the source and the destination have from the channel condition at each time slot. For instance, some studies assume both source and destination have access to the complete Channel State Information (CSI), while some other investigations assume that the destination knows the complete CSI, but only some partial CSI is available at the source.[1] Given the CSI at each time slot, the source can use this information to find the optimal transmission strategy.

The problem of determining the optimal transmission strategy becomes more challenging, if the source does not have information regarding the channel state of each time slot. In fact, there are many important practical applications, including TV broadcasting and satellite communications, in which there is no feedback channel and therefore the channel state information is only available at the destination and the source does not have access to the received Signal to Interference plus Noise Ratio (SINR). One common assumption in these cases is that although the source does not have the CSI in each time slot, it knows the

[1] One possible technique to estimate the channel state at the destination is that the destination measures the amplitude and phase of a pilot signal (transmitted by the source) and computes the channel gain by comparing these values with the initial amplitude and phase of the pilot signal at the transmitter side. Feeding-back this information to the source provides a knowledge of the channel state at the source.

statistics of the channel gain of the link. This is not an unrealistic assumption, since the environment that the source is going to be deployed is usually known a priori. Therefore, it is possible to first evaluate the statistics of the channel gain in that environment, and then use this information to design an efficient transmission scheme.

In this thesis, we investigate three important network configurations. To cover a broader range of applications, we study two different classes of networks. In the first class, we analyze networks in which the channel gains are changing quasi-statically from one transmission block to the other and since there is no feedback link, the source does not have access to the channel state information. In this class, we study a multi-hop network in Chapter 2 and a Multiple-Input Multiple-Output (MIMO) network in Chapter 3. In the second class, we analyze a network in which the channel gains remain fixed in different transmission blocks; however, the amount of interference on the desirable signal is changing from one transmission block to the other. Therefore, the source does not have access to the CSI (at the receiver) in each block. More precisely, in Chapter 4, we investigate a synchronous K-user random access network in which all users work in a similar frequency-band and the channel gains are fixed. However, due to the random access mechanism, the number of active users (and consequently the amount of interference) is changing in different transmission blocks. In this chapter, We aim to propose a transmission scheme that achieves the network average DoF (introduced in Section 4.1), where the users and the AP are equipped with M and N antennas, respectively.

In the following, we present a more detailed explanation of each chapter.

1.1 Transmission over Multi-hop Single-antenna Networks

As the first example of scenarios in which the channel state information is not available at the source, in Chapter 2, we study the transmission of data over a multi-hop network with quasi-static fading channels. It is assumed that CSI of each hop is only available at its corresponding receiving node and the relays are not capable of data buffering over multiple transmission blocks.

1.1.1 Related Work

The concept of relaying is first introduced by Van der Meulen in [1]. Following [1], Cover and El Gamal introduce two different coding strategies for a single-relay network with a direct link between source and destination [2]. In the first strategy, known as *Decode and Forward* (DF), the relay decodes the message and cooperates with source to send it in the next block. In the second strategy, known as *Compress and Froward* (CF) (also called *Quantize and Forward*), relay does not decode the message, instead, it compresses the received signal and forwards it to the destination in the next transmission block. Reference [2] proves the optimality of DF strategy in a single-relay network where the signal received at the destination is a degraded version of the signal received at the relay. Following [2], [3–6] investigate another relaying strategy called *Amplify and Forward* (AF). In AF relaying, the relay amplifies the received signal (without decoding it) and retransmits it to the destination.

Several studies analyze the performance of relaying in different network topologies [4–8]. Considering a single-relay network, references [4] and [5] present a single-letter expression for the maximum achievable rate of AF relaying using a simple linear scheme. Reference [6] shows that AF relaying achieves the capacity in Gaussian parallel single-antenna relay networks. The extension of [6] to the case of multiple-antenna Rayleigh fading is presented in [7] and [8].

Recent work also explores different transmission schemes for networks in which the CSI is only available at receiving nodes. The majority of these studies focus on Diversity-Multiplexing Trade-off (DMT) [9–14]. Other related studies focus on maximizing the statistical average of the rate per channel use at the destination [15–23]. In a pioneering work, Shamai finds the optimum strategy for a single-hop network where the CSI is available only at the receiver, and where the channel gain is quasi-static. The solution presented in [15] is based on replacing the receiver by a continuum of virtual receivers, each corresponding to a specific realization of the channel gain. Relying on an analogy with a degraded broadcast channel, reference [15] shows that an infinite-layer coding scheme with a proper power distribution maximizes the statistical average of the received rate at the destination.

There are several extensions of [15]. References [16–18] study the performance of different transmission strategies where some partial channel state information is available at the source. Reference [19] proposes the application of multilayer coding in a multicast network which has some QoS constraints. Reference [20] combines multilayer coding with a Hybrid Automatic Retransmission Request (HARQ) mechanism and shows that it results in high rates and low latency in a point-to-point link. References [21, 22] extend [15] to the case of MIMO channels.

Another important extension of [15] is studied in [24], where Steiner *et al.* present the two-hop extension of [15] in which they study a single-relay network (no direct link between source and destination) and the relay receive and transmit links are orthogonal to each other. In [24], it is assumed that the relay node is not capable of data buffering over multiple coding blocks and the CSI is only available at the receiver side of each link. The objective in [24] is to design a transmission scheme which maximizes the statistical average of the received rate per channel use at the destination. To this end, [24] studies different multilayer coding schemes in which the relay operates in different modes, including DF, AF, and CF. As discussed in [24], because of the high complexity of the infinite-layer DF coding scheme, only finite layer codes are considered in their proposed DF strategies. Reference [24] shows that, in the high SNR regime, AF relaying outperforms all other investigated transmission schemes. However, the authors note that, in spite of having the best performance among the compared strategies, AF relaying is not claimed to be optimal [24].

1.1.2 Motivations and Summary of the Main Contributions

Considering a two-hop network, results of [24] show the superiority of AF coding with respect to other schemes proposed in [24]. However, since the performance of infinite-layer DF coding scheme is not evaluated in [24], the question arises of whether or not the infinite-layer DF coding achieves a higher expected rate when compared to AF relaying.

In this chapter, we aim to answer this question and find the optimal performance of infinite-layer DF coding scheme for two-hop networks. The results are then used to find a transmission strategy for multi-hop networks. The main contributions of this chapter can be summarized as the following:

- The statistical average of the received rate per channel use at the final destination of a multi-hop network is formulated as an optimization problem.

- An algorithm is introduced to solve the introduced optimization problem and determine the optimum parameters of the proposed transmission scheme for two-hop networks.
- The optimality of infinite-layer DF coding is discussed for two-hop networks.
- The generalized design algorithm is proposed for construction of multilayer codes for multi-hop networks.

1.2 Transmission over Single-hop Multi-antenna Networks

With the growth of MIMO systems, it is essential to study the transmission of scalable video codes over MIMO systems. Similar to single antenna systems, multilayer coding is one of the main candidates for data transmission where unequal error protection is required for different parts of the transmitted message. In Chapter 3, we investigate the design of multilayer codes for single-hop MIMO links in which the channel state information is only available at the destination.

1.2.1 Related Work

Different video coding techniques, including MPEG-4 and H.264/AVC, have been recently introduced to provide efficient compression techniques for video content [25, 26]. Scalable Video Coding (SVC) is an extension of these coding schemes which compresses the video signal into different enhancement layers of temporal, spatial, or quality resolution. For instance, a SVC H.264/AVC stream consists of one base-layer and several enhancement layers [27, 28]. If a user is able to receive the base-layer it can reconstruct the video stream with a minimum required Quality of Service (QoS). Reception of each enhancement layer allows the user to improve the temporal, spatial, or resolution quality of the video stream.

One important application of scalable video codes is in scenarios where the aim is to transmit a video stream in a network with different link capacities. In such a network, all layers are transmitted only if the link has a capacity higher than the total rate of the compressed stream. Otherwise, the truncated version of the stream (with a lower rate) is transmitted through the link which allows the user to still decode the video (with a lower quality) [28, 29].

Focusing on wireless applications, scalable video codes can be used in a scenario in which there is one source broadcasting a video signal to several users, e.g., TV broadcasting and satellite communications. More precisely, in this application the source transmits a codeword and each user receives a noisy version of the codeword. The Signal to Noise Ratio (SNR) of the received signal at each user depends on the location of the user as well as its channel state in that transmission block. If we broadcast a message with a fixed rate, only a subset of users will be able to decode the message. Although this transmission scheme is useful in some applications, it is not appropriate for video transmission. This is due to the fact that there is no possibility of partial decoding of the transmitted data; therefore, even though the video is compressed using a scalable video coding standard, the destination can decode all or none of the layers.

An alternative scheme for layered data transmission is to use unequal error protection for different layers of the video stream [30–32]. For instance, it is possible to encode each layer of the video stream separately and then superimpose the resulting codes while we assign different coding redundancy and/or different powers to the layers which require different error protections. As an example, compared to the enhancement layers, we can assign a higher coding redundancy and/or higher power to the base-layer of a SVC H.264/AVC stream. By this, we can make sure that all users are able to decode at least the base-layer while some users can improve their video quality by decoding a number of enhancement layers. Multilayer coding is one method to implement such transmission schemes. In this chapter, we aim to present an algorithm to design multilayer codes which in turn leads to construction of more reliable codes for transmission of layered data such as SVC H.264/AVC. We rely on Information theoretic arguments to provide an upper-bound on the possible benefits, which in turns will provide insight into the practical construction of such codes.

To investigate the optimal design of multilayer codes in MIMO systems, [21] and [22] study the extension of [15] for MIMO systems where the CSI is only available at the destination and the objective is to maximize the average received rate at the destination. Shamai *et al.* in [21] show the analogy between multilayer coding for the MIMO setup and a general broadcast channel. Furthermore, they formulate the achievable average received rate at the destination as an optimization problem. This optimization problem can be used to determine the parameters of the multilayer code which maximizes the average received rate at the destination. Unfortunately, as stated in [21], the proposed optimization problem does not appear to have a straightforward solution. Therefore, [21] proposes some suboptimal multilayer transmission schemes and analyzes their performance. In a follow up study, [22] investigates the performance of some other suboptimal MIMO schemes including finite-layer coding strategies.

1.2.2 Motivations and Summary of the Main Contributions

In this chapter, we intend to continue the study of multilayer coding for MIMO systems in order to develop a transmission scheme which can be used for transmission of layered data such as SVC video streams. To this end, we follow the network structure of [21] and [22] where we have a single-user network and nodes are equipped with multiple antennas. The channel is assumed to be quasi-static and the CSI is only available at the destination. The objective is to determine the optimal design of a multilayer code which maximizes the average received rate at the destination. We limit our studies to the cases in which the destination is only able to perform successive decoding (not joint decoding). It should be noted that the proposed coding strategy might be not optimal for a destination which has the joint decoding capability, but this issue is not studied in this thesis. The main contributions of this chapter can be summarized as the following:

- It is shown that the design of the optimal multilayer code is not unique (if we consider single antenna systems).
- A design rule for construction of optimal multilayer codes for MIMO systems is introduced (as mentioned earlier, assuming successive decoding).
- An algorithm (based on the proposed design rule) is proposed to determine the parameters of the multilayer code.
- The performance of multilayer coding strategy is analyzed for networks with different fading statistics.
- A transmission scheme (based on multilayer coding) is presented to improve the performance of a typical cellular network which aims to broadcast layered data (e.g. a video signal coded using a scalable video coding standard) to all users.

1.3 Transmission over MIMO Random Access Networks

In this chapter, we study a synchronous K-user random access network in which the channel gains between the users and the Access Point (AP) are fixed (which are known by the users[2]) and users work in a similar frequency-band. There is no central controller in this network, and at the beginning of each time slot, all users independently decide (with probability ρ) whether or not to transmit in that time slot. In this network, the Signal to Interference plus Noise Ratio (SINR) of one user's signal at the AP depends on both the channel gain of the user-AP link, as well as the number of active users in that time slot (if more than one user become active in a time-slot, they interfere with each other). Therefore, although the channel gains are fixed and the source knows this value, it does not have access to the complete CSI (it does not know the number of active users in each time slot). In Chapter 4, we investigate the transmission over such random access networks, where nodes are equipped with multiple antennas.

1.3.1 Related Work

One of the main responsibilities of the medium access control layer in each network is to present an algorithm which coordinates how the network users should access the shared wireless medium. Some of these methods, including the Time and Frequency Division Multiple Access schemes, select a number of the network users in each time slot and allow them to send data in that time slot. These methods, therefore, require a central control unit which manages the users' transmission in each time slot. In addition, in these mechanisms, there should be a feedback channel to notify users regarding the bandwidth/time they should transmit their data. The other group of medium access mechanisms does not have this central brain and each user decides independently to whether or not to transmit in one time slot. These techniques are usually categorized in the group of random medium access mechanism. ALOHA [33], CSMA, and CSMA/CD [34] are some examples of random access mechanism. The main difficulty of the random medium access mechanism is that users do not have information about the number of active users in each time slot. Therefore,

[2]Since the channel gains are fixed there is no need to have a feedback channel to constantly report the channel gain to the source.

they cannot adapt their transmission rate to the number of active users. The goal of the random access mechanism is usually to maximize the average performance of the network.

After selecting the active users in each time slot (either by a central controller or using a random access mechanism), users and the AP can be modeled as a Multiple Access Channel (MAC). The capacity region of a K-user MAC is the set of all non-negative K-tuples $(R_1, R_2, \cdots, R_K)$ that satisfy [35]:

$$\sum_{j \in Q} R_j \leq \frac{1}{2} \log_2 \left(1 + \frac{\sum_{j \in Q} P_j}{N_0} \right), \quad \forall Q \subseteq \{1, 2, \cdots, K\}, \tag{1.1}$$

where P_j and N_0 are the jth user's transmission power and the noise spectral density, respectively. Different schemes have been proposed in order to achieve all points on the boundary of the capacity region. For instance, [36] proposes a scheme based on joint encoding/decodig for all K users. Reference [36] also presents a technique in which time sharing and user interference cancelation (successive decoding) can be applied to reach a particular point on the boundary. Centralized/ distributed rate splitting techniques are another methods of achieving the boundary of the MAC capacity region [35], [37]. Note that all of these schemes assume a fixed network setup, i.e., the active number of users and the channel gains between the users and the AP are constant during the whole communication.

Several studies also investigate the performance of non-static networks. For instance, [38] studies a K-user MAC while the channels are subjected to a quasi-static fading, i.e., the channel gains are only constant during one transmission block and they change independently from one transmission block to the other. In another pioneering work, Medard *et al.* analyze the achievable rate of a random access network and model the random access network as a Multiple Access Channel (with fixed channel gains), where in each time slot, each user (user i) become active with the probability of ρ_i [39]. A user's transmitted data in [39] consists of several layers of coded information with different coding rates. Therefore, if the complete message is not decodable (due to the interference), the AP is still able to decode parts of the transmitted data (according to the number of active users).

Following [39], Minero *et al.* in [40–42] propose a new upper-bound for the total expected rate of a random access channel where users and the AP are single antenna units. Applying multilayer coding techniques, [42] presents a transmission scheme which achieves within $\frac{\sqrt{3}}{2}$ bit of the maximum total expected rate of a symmetric two-user random access networks. Generalizing to a symmetric K-user random access network, the authors also propose an achievability scheme which is within one bit of the upper-bound of the maximum total expected rate. Interestingly, this scheme is very simple such that each user only needs to send one stream of data (does not need to apply multilayer coding).

1.3.2 Motivations and Summary of the Main Contributions

Reference [42] only discusses networks in which nodes are equipped *with a single antenna*. In this chapter, we extend the results of [42] to networks *with multiple antennas*. We focus on the symmetric K-user random access networks in which the users and the AP are equipped with M and N antennas, respectively. Our aim is to find the optimal network average Degrees of Freedom (DoF) (introduced in Section 4.1) where all users have the same probability of activation, ρ. The main contributions of this chapter can be summarized as the following:

- The optimal network average DoF is determined for all setups of symmetric two-user random access networks, i.e. different selections of M, N, and ρ.
- An upper-bound for the network average DoF of symmetric K-user random access networks is determined for different selections of M, N, and ρ.
- A simple single-stream transmission scheme is proposed that achieves the introduced upper-bound where: i) $N \leq M$, ii) $KM \leq N$, and iii) $M < N < KM$ and $\rho \geq \rho_{j_2-1}$ (ρ_{j_2-1} is defined in Section 4.4.1). Thus, we have the optimal network average DoF in these scenarios.
- A new transmission scheme based on multi-stream data transmission and interference alignment is proposed. The achievable network average DoF of this transmission scheme is higher compared to the single-stream data transmission, where $M < N < KM$ and $\rho < \rho_{j_2-1}$. It is also proved that this scheme achieves the corresponding upper-bound for some classes of random access networks. Thus, we have the optimality result for these network setups. It should be mentioned that there are some network configurations in which the proposed scheme cannot reach the upper-bound. Further investigation is required to close the gap between the upper and the lower bounds in these situations.

Chapter 2

Transmission over Multi-hop Single-antenna Networks

In this chapter[1], we study transmission over a multi-hop network in which information is transmitted from a source via a number of relays to a destination. It is assumed that channels are quasi-static fading with additive white Gaussian noise and that all nodes poses a equipped with a single antenna. The Channel State Information (CSI) of each hop is available only at the corresponding receiver and relays are not capable of data buffering over multiple transmission blocks. The objective is to maximize the statistical average of the received rate per channel use[2] at the destination. We start the analysis, in Section 2.2 and Section 2.3, by considering a two-hop network in which both the source and the relay use infinite-layer coding for data transmission and the relay is operating in decode and forward mode. This section also presents an algorithm to optimally distribute the available source and relay powers to different layers of their corresponding codes. Next, Section 2.4 shows how this transmission technique can be generalized to a multi-hop setting. Assuming Rayleigh fading, in Section 2.5, the performance of the proposed coding scheme is evaluated for a two-hop network and the results are compared with the performance of previously known strategies.

[1]The work in this chapter is partially reported in [43, 44].

[2]In this thesis, the statistical average of the received rate per channel use is referred to as *average received rate*.

2.1 Preliminaries

This section provides a review of key studies on the application of the multilayer codes in single-hop scenarios.

2.1.1 Single-hop Rate-Limited Broadcasting Strategy

Consider a single-hop network with one source and one destination where both nodes are equipped with a single antenna. Let integer i represent the time instant. In this network, we have:

$$y_i = hx_i + n_i, \tag{2.1}$$

where $\{x_i\}$ and $\{y_i\}$ denote samples of the transmitted and received signals, respectively. Furthermore, $\{n_i\}$ are samples of the independent and identically distributed (i.i.d.) additive white Gaussian noise which are zero-mean complex Gaussian, $\mathcal{CN}(0, 1)$. The source-destination channel gain changes quasi-statically and is denoted by h. It means that the value of h remains fixed during one transmission block and changes independently from block to block. It is also assumed that the source transmission rate in each block is limited to Γ_{in} and the CSI in each transmission block is only available at the destination; however, both the source and destination know the probability density function (pdf) of the channel fading power, i.e., $l = h^2$.

The aim is to design a transmission scheme such that it maximizes the average received rate at the destination. This problem is first studied in [15] where the authors proposed a scheme called the broadcast approach.

In the broadcast approach, the transmitted signal x is a multilayer code which consists of the superposition of continuum of layers of Gaussian coded information. According to the channel gain, in each block, the destination decodes up to a certain layer of the code. Note that the remaining code layers generate interference at the receiver. Defining $\rho(a, \Gamma_{in})da$ as the portion of the normalized source power which is assigned to the code layer a, the differential data rate of each code layer is computed in [21] as:

$$\begin{aligned}\zeta(a,\Gamma_{in})da &= \log\left(1+\frac{a\rho(a,\Gamma_{in})da}{1+aI(a,\Gamma_{in})}\right) \\ &\overset{(a)}{=} \frac{a\rho(a,\Gamma_{in})}{1+aI(a,\Gamma_{in})}da,\end{aligned} \tag{2.2}$$

where

$$I(a,\Gamma_{in}) = \int_a^\infty \rho(\gamma,\Gamma_{in})d\gamma, \tag{2.3}$$

and (a) follows from the assumption of infinitesimal rate assignment. Denoting the channel fading power $l=|h|^2$, [21] shows that in each channel state l, the destination receiving rate can be computed as:

$$R(l,\Gamma_{in}) = \int_0^l \zeta(a,\Gamma_{in})da, \tag{2.4}$$

Let $f_L(\cdot)$ and $F_L(\cdot)$ denote the *pdf* and the cumulative density function (*cdf*) of the channel fading power, respectively. The average received rate at the destination can be written as:

$$\begin{aligned}R_{avg} &= \int_0^\infty f_L(l)R(l,\Gamma_{in})dl \\ &= \int_0^\infty f_L(l)\int_0^l \frac{a\rho(a,\Gamma_{in})}{1+aI(a,\Gamma_{in})}dadl \\ &= \int_0^\infty \Big[1-F_L(a)\Big]\frac{a\rho(a,\Gamma_{in})}{1+aI(a,\Gamma_{in})}da.\end{aligned} \tag{2.5}$$

The objective is to find $\rho(\cdot,\Gamma_{in})$ such that the average received rate at the destination is maximized. Following [24], we have:

$$R^*_{avg} = \max_{\rho(\cdot,\Gamma_{in})} \quad \int_0^\infty f_L(l)\int_0^l \frac{a\rho(a,\Gamma_{in})}{1+aI(a,\Gamma_{in})}dadl \tag{2.6}$$

$$s.t. \quad \int_0^\infty \rho(l,\Gamma_{in})dl = P_1, \tag{2.7}$$

$$\int_0^\infty \frac{a\rho(a,\Gamma_{in})}{1+aI(a,\Gamma_{in})}da \le \Gamma_{in}. \tag{2.8}$$

where P_1 denotes the total source transmitted power.

As shown in [24], the optimal solution of the above optimization problem satisfies:

$$I^*(l, \Gamma_{in}) = \int_l^\infty \rho^*(\gamma, \Gamma_{in}) d\gamma = \begin{cases} P_1 & l < l^b \\ \frac{1 - F_L(l) + \lambda - l f_L(l)}{l^2 f_L(l)} & l^b \leq l \leq l^e \\ 0 & l^e < l \end{cases}, \quad (2.9)$$

where l^b, l^e are selected such that $I^*(l^b, \Gamma_{in}) = P_1$ and $I^*(l^e, \Gamma_{in}) = 0$, respectively. The parameter λ is computed such that the rate condition is satisfied.

Assuming a Rayleigh fading channel, i.e., $F_L(l) = 1 - e^{-l}$, we have [24]:

$$I^*(l, \Gamma_{in}) = \begin{cases} P_1 & l < l^b \\ \frac{\lambda}{l^2 e^{-l}} + \frac{1}{l^2} - \frac{1}{l} & l^b \leq l \leq l^e \\ 0 & l^e < l \end{cases}. \quad (2.10)$$

The values of λ, l^b, and l^e are computed by solving the following system of equations:

$$\begin{cases} \Gamma_{in} &= 2\log(l^e) - l^e - (2\log(l^b) - l^b), \\ I^*(l^b, \Gamma_{in}) &= P_1, \\ l^e &= 1 - W(-\lambda e), \end{cases} \quad (2.11)$$

where $W(x)$ is the Lambert W-function which is the inverse of the function we^w.

2.1.2 Single-Hop Broadcasting Strategy

Shamai *et al.* in [15] and [24] also consider the application of multilayer codes for single-hop networks with no rate limitation at the source. The formulation of the average received rate at the destination is similar to the previous case for $\Gamma_{in} = \infty$. For instance, (2.2)-(2.4) can be written as:

$$R(l, \infty) = \int_0^l \frac{a\rho(a, \infty)}{1 + aI(a, \infty)} da. \quad (2.12)$$

For simplicity of notation, henceforth we use $\tilde{R}(l) = R(l, \infty)$. Similarly, expression (2.6) simplifies to:

$$\begin{aligned} R^*_{avg} = \max_{\rho(\cdot,\infty)} \quad & \int_0^\infty f_L(l) \int_0^l \frac{a\rho(a,\infty)}{1 + aI(a,\infty)} da dl \\ s.t. \quad & \int_0^\infty \rho(l,\infty) dl = P_1. \end{aligned} \tag{2.13}$$

The optimal solution of (2.13) satisfies [15]:

$$I^*(l,\infty) = \int_l^\infty \rho^*(\gamma,\infty) d\gamma = \begin{cases} P_1 & l < l^b \\ \frac{1 - F_L(l) - l f_L(l)}{l^2 f_L(l)} & l^b \le l \le l^e \\ 0 & l^e < l \end{cases}, \tag{2.14}$$

where l^b and l^e satisfy $I^*(l^b,\infty) = P_1$ and $I^*(l^e,\infty) = 0$, respectively. The total rate transmitted over different layers of the optimal multilayer code is equal to:

$$R^*_T = \int_0^\infty \frac{a\rho^*(a,\infty)}{1 + aI^*(a,\infty)} da. \tag{2.15}$$

It is important to note that that in the case of $\Gamma_{in} \ge R^*_T$, the optimization problem in (2.6) is simplified to (2.13). Although this statement can be verified mathematically, the intuitive explanation is insightful. To explain, note that R^*_T in (2.15) shows the total rate transmitted using the optimal multilayer code when there is no rate constraint at the transmitter. This means even though the available rate at the source is more than R^*_T, the source needs to transmit only R^*_T bits of information in each block. Hence, if $\Gamma_{in} \ge R^*_T$, the rate constraint would not become active and the optimum solutions of (2.6) and (2.13) will be equal.

Example: For the case of Rayleigh fading where $F(l) = 1 - e^{-l}$, we have:

$$I^*(l,\infty) = \begin{cases} P_1 & l < l^b \\ \frac{1}{l} - \frac{1}{l^2} & l^b \le l \le l^e \\ 0 & l^e < l \end{cases}, \tag{2.16}$$

where $l^b = \frac{\sqrt{1+4P}-1}{2P}$ and $l^e = 1$. Furthermore, the maximum average received rate at the

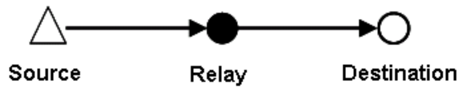

Figure 2.1: Two-hop network model

destination is:

$$R^*_{avg} = \int_{l^b}^{l^e} \left[\frac{2}{l} - 1\right] e^{-l} dl, \tag{2.17}$$

2.2 Network Model

Our network model, in this chapter, is the same as [24]. We first focus on a two-hop setting and then, in Section 2.4, generalize the results to multi-hop cases.

In a two-hop setting, the relay retransmits the data received from the source to the destination and the destination can only receive data via the relay. All nodes are single antenna and the two hops are orthogonal with the same bandwidth, Fig. 2.1. Therefore, we have:

$$\begin{aligned} \text{source to relay:} \quad \mathbf{y}_r &= h_1\mathbf{x}_s + \mathbf{n}_s, \\ \text{relay to destination:} \quad \mathbf{y}_d &= h_2\mathbf{x}_r + \mathbf{n}_r, \end{aligned} \tag{2.18}$$

where $\mathbf{x}_s$ and $\mathbf{x}_r$ denote the source and the relay transmitted vectors in one transmission block of length N. $\mathbf{y}_r$ and $\mathbf{y}_d$ represent the received vectors in one transmission block at the relay and at the destination, and h_1, h_2 are the source-relay and relay-destination channel gains. Elements of the noise vectors, $\mathbf{n}_s$ and $\mathbf{n}_r$, are zero-mean complex Gaussian i.i.d. with unit variance, $\mathcal{CN}(0, 1)$.

As in [24], in this work, it is assumed that:

1. The channel gains of both hops are quasi-static. In other words, channel gains remain fixed during one transmission block and change independently from block to block.
2. The source and the relay power constraints are expressed as $E\left[|x_s|^2\right] = P_1$ and $E\left[|x_r|^2\right] = P_2$, respectively, where x_s and x_r show an arbitrary element of $\mathbf{x}_s$ and $\mathbf{x}_r$.

3. The relay is not capable of storing data over multiple transmission blocks. Therefore, in each transmission, the relay can only retransmit the data that it has received from the source-relay link. This means if the relay transmits with rate $R < \Gamma_{in}$ (where Γ_{in} is the rate that the relay has received from the source-relay link) any remaining data cannot be stored at the relay and should be discarded.

4. The source has no information about either of the channel gains, the relay knows only the channel gain between itself and the source, and the destination knows both channel gains. [3]

5. The channel fading powers of the first and the second hops are defined as $l_1 = |h_1|^2$ and $l_2 = |h_2|^2$, respectively. It is assumed that the source and the relay both know the *pdf* of channel fading power for both links. $f_{L_1}(l_1)$ and $f_{L_2}(l_2)$ denote the *pdf* of the channel fading powers of the source-relay and the relay-destination links, respectively.

2.3 Multilayer Coding Scheme for Two-hop Networks

Before discussing the details of the design of infinite-layer code for a two-hop network with DF relaying, in Section 2.3.1, we study a preliminary example. In this example, we study the design of the optimal multilayer code for a two-hop network in which the channel fading of the source-relay and the relay-destination have two states. The analysis presented in this example is much simpler compared to the case of a network with continuous channel fading powers.

Following Section 2.3.1, we formulate the average received rate at the destination of a two-hop network which uses infinite-layer DF coding. Then, Section 2.3.3 presents an algorithm to optimally determine the code parameters at the source and at the relay. Optimality analysis of this scheme is discussed at the end.

[3] These CSI assumptions are indeed practical, since the receiver of each hop can evaluate its immediate channel gain by measuring the pilot signal sent from the corresponding transmitter. In addition, the receiver can measure the equivalent channel (the source to the destination) if the relay forwards the pilot signal of the source towards the destination. Having the equivalent channel gain and the relay to destination channel gain, the destination can find the source to relay channel gain as well.

2.3.1 Optimal Transmission over a Two-State Two-hop Network

Consider a two-hop network in which the source-relay channel fading power is equal to $L_1(1)$ or $L_1(2)$ with probabilities $f_1(1)$ and $1-f_1(1)$, respectively, and $L_1(1) < L_1(2)$. Similarly, the relay-destination channel fading power is equal to $L_2(1)$ or $L_2(2)$ with probabilities $f_2(1)$ and $1-f_2(1)$, respectively, and $L_2(1) < L_2(2)$. The source and the relay power constraints are denoted by P_1 and P_2, respectively.

In this setting, the relay transmits a two-layer code. The rates associated with these layers are denoted by $R_2(1)$ and $R_2(2)$ which are transmitted with power $\rho_2(1)$ and $P_2 - \rho_2(1)$, respectively. Furthermore, the input rate of the relay is denoted by Γ_{in}.

The objective is to find $R_2(1)$, $R_2(2)$, and $\rho_2(1)$ to maximize the average received rate at the destination. Based on the multilayer structure of the code, these parameters should be set such that the destination can always decode the first layer (with rate $R_2(1)$) and can decode the second layer if the channel fading power is equal to $L_2(2)$. Considering a given Γ_{in} and $\rho_2(1)$, the problem of maximizing the average received rate at the destination can be be written as:

$$\begin{aligned} A^*(\rho_2(1), \Gamma_{in}) = \max_{R_2(1), R_2(2)} \quad & f_2(1)R_2(1) + (1-f_2(1))(R_2(1)+R_2(2)) \\ s.t. \quad & R_2(1) \leq \beta_1, \\ & R_2(2) \leq \beta_2, \\ & R_2(1)+R_2(2) \leq \Gamma_{in}, \\ & R_2(1), R_2(2) \geq 0, \end{aligned} \tag{2.19}$$

where $\beta_1 = \log\left(1 + \frac{\rho_2(1)L_2(1)}{1+(P_2-\rho_2(1))L_2(1)}\right)$ and $\beta_2 = \log\left(1+(P_2-\rho_2(1))L_2(2)\right)$. The optimum value of $A^*(\rho_2(1), \Gamma_{in})$ can be found using Fourier Motzkin elimination technique. The result is:

$$A^*(\rho_2(1), \Gamma_{in}) = \min\left\{ f_2(1)\beta_1 + (1-f_2(1))(\beta_1+\beta_2), f_2(1)\beta_1 + (1-f_2(1))\Gamma_{in}, \Gamma_{in} \right\}. \tag{2.20}$$

Table 2.1: The optimal values of $B^*(\Gamma_{in})$ for different values of the input rate

Case	$\rho_2(1)$	$R_2{}^*(1)$	$R_2{}^*(2)$	$B^*(\Gamma_{in})$
$0 \le \Gamma_{in} < \mathbb{TR}_1$	$\mathbb{TP}_1$	Γ_{in}	0	Γ_{in}
$\mathbb{TR}_1 \le \Gamma_{in} < \mathbb{TR}_2$	$\mathbb{TP}_2$	$\log\left(1+\frac{\mathbb{TP}_2 L_2(1)}{1+(P_2-\mathbb{TP}_2)L_2(1)}\right)$	$\Gamma_{in} - R_2{}^*(1)$	$R_2{}^*(1)+(1-f_2(1))R_2{}^*(2)$
$\mathbb{TR}_2 \le \Gamma_{in}$	$\mathbb{TP}_3$	$\log\left(1+\frac{\mathbb{TP}_3 L_2(1)}{1+(P_2-\mathbb{TP}_3)L_2(1)}\right)$	$\log(1+(P_2-\mathbb{TP}_3)L_2(2))$	$R_2{}^*(1)+(1-f_2(1))R_2{}^*(2)$

For a given Γ_{in}, the optimum value of $\rho_2(1)$ then can be found as:

$$
\begin{aligned}
B^*(\Gamma_{in}) = \max_{\rho_2(1)} \quad & A^*(\rho_2(1), \Gamma_{in}) \\
s.t. \quad & \rho_2(1) \le P_2, \\
& \rho_2(1) \ge 0.
\end{aligned}
\tag{2.21}
$$

The optimization problem in (2.21) can be solved for different values of Γ_{in}. The results are summarized in Table 2.1 where $\mathbb{TP}_1$ and $\mathbb{TP}_2$ are defined such that they satisfy $\log(1+\mathbb{TP}_2 L_2(1)) = \Gamma_{in}$ and $\log(1+P_2L_2(1)) + \log\left(\frac{1+(P_2-\mathbb{TP}_2)L_2(2)}{1+(P_2-\mathbb{TP}_2)L_2(1)}\right) = \Gamma_{in}$, respectively. Furthermore,

$$
\mathbb{TP}_3 = \left(P_2 - \left(\frac{L_2(2)(1-f_2(1)) - L_2(1)}{f_2(1)L_2(1)L_2(2)}\right)^+\right)^+, \tag{2.22}
$$

and $\mathbb{TR}_1$, $\mathbb{TR}_2$ are defined as:

$$
\begin{aligned}
\mathbb{TR}_1 &= \log(1+P_2L_2(1)), \\
\mathbb{TR}_2 &= \log(1+P_2L_2(1)) + \log\left(\frac{1+(P_2-\mathbb{TP}_3)L_2(2)}{1+(P_2-\mathbb{TP}_3)L_2(1)}\right).
\end{aligned}
\tag{2.23}
$$

Having obtained the optimal design of the relay (for each input rate), we now focus on the source-relay link. Similar to the relay part, the source transmits a two-layer code. The rates associated with these layers are denoted by $R_1(1)$ and $R_1(2)$ which are transmitted with power $\rho_1(1)$ and $P_1 - \rho_1(1)$, respectively. Using this structure, the input rate of the relay is $\Gamma_{in} = R_1(1)$ or $\Gamma_{in} = R_1(1) + R_1(2)$ if the channel fading power of the source-relay link is equal to $L_1(1)$ or $L_1(2)$, respectively.

We should now calculate $R_1(1)$, $R_1(2)$ and $\rho_1(1)$ to maximize the average received rate at the destination. For a fixed $\rho_1(1)$, we have:

$$\begin{aligned} D^*(\rho_1(1)) = \max_{R_1(1),R_1(2)} \quad & f_1(1)B^*(R_1(1)) + (1-f_1(1))B^*(R_1(1)+R_1(2)) \\ s.t. \quad & R_1(1) \leq \delta_1, \\ & R_1(2) \leq \delta_2, \\ & R_1(1), R_1(2) \geq 0, \end{aligned} \tag{2.24}$$

where $\delta_1 = \log\left(1 + \frac{\rho_1(1)L_1(1)}{1+(P_1-\rho_1(1))L_1(1)}\right)$ and $\delta_2 = \log(1+(P_1-\rho_1(1))L_1(2))$. Noting that $B^*(\Gamma_{in})$ is a non-decreasing function with respect to Γ_{in}, we conclude that the optimal solution of (2.24) satisfies $R_1(1) = \delta_1$ and $R_1(2) = \delta_2$, i.e.:

$$D^*(\rho_1(1)) = f_1(1)B^*(\delta_1) + (1-f_1(1))B^*(\delta_1+\delta_2). \tag{2.25}$$

Finally, the optimal value of $\rho_1(1)$ can be evaluated by:

$$\begin{aligned} R^* = \max_{\rho_1(1)} \quad & f_1(1)B^*(\delta_1) + (1-f_1(1))B^*(\delta_1+\delta_2) \\ s.t. \quad & \rho_1(1) \leq P_1, \\ & \rho_1(1) \geq 0, \end{aligned} \tag{2.26}$$

where $B^*(\delta_1)$ and $B^*(\delta_1+\delta_2)$ can be determined using Table 2.1 and using $\Gamma_{in} = \delta_1$ and $\Gamma_{in} = \delta_1+\delta_2$, respectively. Note that $\delta_1 = \log\left(1 + \frac{\rho_1(1)L_1(1)}{1+(P_1-\rho_1(1))L_1(1)}\right)$ and $\delta_2 = \log(1+(P_1-\rho_1(1))L_1(2))$. By obtaining the optimal value of $\rho_1(1)$, we have all the required parameters for designing the multilayer codes at both the source and the relay.

2.3.2 Infinite-layer DF Coding for Two-hop Networks

In this section, we continue the study of two-hop networks, but this time we consider a network in which both hops have continuous channel gains. Consequently, instead of finite-layer codes, the source and the relay use infinite-layer codes. In this setting, the relay still works in DF mode. A more precise description of each transmission phase is presented in the following.

1. To construct an infinite-layer code, the source selects a distribution function to distribute its power among its different layers. Since there is no rate-limitation at the source, this power distribution function is denoted by $\rho_1(\cdot, \infty)$. Note that $\rho_1(\cdot, \infty)$ should be selected such that it satisfies the source power constraint, i.e.: $\int_0^\infty \rho_1(a, \infty)da = P_1$.

   The relay is able to decode up to layer l_1 of the transmitted code, where l_1 represents the source-relay channel fading power. Thus, similar to (2.12), the relay input rate would be:

$$\tilde{R}_1(l_1) = \int_0^{l_1} \frac{a\rho_1(a, \infty)}{1 + aI_1(a, \infty)} da, \tag{2.27}$$

   where $I_1(a, \infty) = \int_a^\infty \rho_1(\gamma, \infty)d\gamma$.

2. In the second phase, the relay transmits the data to the destination. As noted in Section 2.2, the transmission rate of the relay cannot exceed what it has received from the source-relay link, i.e., $\tilde{R}_1(l_1)$. For each value of $\tilde{R}_1(l_1)$, the relay should choose a power distribution function for distributing the relay power among the code layers while also satisfying the power constraint P_2. Defining $\rho_2(\cdot, \tilde{R}_1(l_1))$ as the power distribution at the relay conditioned on the input rate $\tilde{R}_1(l_1)$, we have:

$$\text{Power constraint at the relay:} \forall \tilde{R}_1(l_1) : \int_0^\infty \rho_2(a, \tilde{R}_1(l_1))da = P_2, \tag{2.28}$$

$$\text{Rate constraint at the relay:} \forall \tilde{R}_1(l_1) : \int_0^\infty \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} da \leq \tilde{R}_1(l_1), \tag{2.29}$$

   where $I_2(a, \tilde{R}_1(l_1)) = \int_a^\infty \rho_2(\gamma, \tilde{R}_1(l_1))d\gamma$.

3. Transmitting the resulting infinite-layer code on the relay-destination link, the destination is able to decode up to layer l_2, where l_2 denotes the channel fading power of the second hop. Therefore, for each $\tilde{R}_1(l_1)$, the average received rate at the destination can be written as:

$$R_2(l_2, \tilde{R}_1(l_1)) = \int_0^{l_2} \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} da. \tag{2.30}$$

   Note that for successful decoding, the destination should know the power distribution strategy of the relay. This information can be obtained by evaluating the source-relay channel fading power at the destination.

The average received rate at the destination can be written as:

$$\begin{aligned} R_{avg} &= E_{l_1}\left[E_{l_2}\left[R_2\left(l_2, \tilde{R}_1(l_1)\right)\right]\right] \\ &= \int_0^\infty \int_0^\infty f_{L_1}(l_1) f_{L_2}(l_2) R_2(l_2, \tilde{R}_1(l_1)) dl_2 dl_1, \end{aligned} \tag{2.31}$$

where $f_{L_1}(l_1)$ and $f_{L_2}(l_2)$ denote the probability density functions of the channel fading powers of the source-relay and relay-destination links, respectively. Given (2.28)-(2.31), the final optimization problem is:

$$R^*_{avg} = \max_{\rho_1(\cdot,\infty),\ \rho_2(\cdot,\cdot)} \quad \int_0^\infty \int_0^\infty f_{L_1}(l_1) f_{L_2}(l_2) \int_0^{l_2} \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} da dl_2 dl_1 \tag{2.32}$$

$$s.t. \quad \int_0^\infty \rho_1(a, \infty) da = P_1, \tag{2.33}$$

$$\forall \tilde{R}_1(l_1) : \int_0^\infty \rho_2(a, \tilde{R}_1(l_1)) da = P_2, \tag{2.34}$$

$$\forall \tilde{R}_1(l_1) : \int_0^\infty \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} da \leq \tilde{R}_1(l_1). \tag{2.35}$$

Note that the above optimization problem is similar to the one derived in [24]. However, in [24], the relay rate constraint in (2.35) is stated as an equality. The inequality in (2.35) reflects the fact that the relay may not need to send all the information that it has received from the source-relay link. For instance, consider a scenario in which the rate received at the relay is higher than the corresponding R^*_T, where R^*_T is defined in (2.15). In such a scenario, the relay only uses R^*_T bits of the received information and ignores the rest.

2.3.3 Optimal Design of Infinite-layer DF Codes for Two-hop Networks

In this section, we present an algorithm to solve the optimization problem in (2.32)-(2.35). Note that unlike single-hop scenarios, (2.32) cannot be directly modeled as a constrained variation problem[4]. The reason is that $\tilde{R}_1(l_1)$ in (2.35) is not a constant value for different selections of $\rho_1(\cdot,\infty)$ and l_1. To resolve this issue, we use the following lemma.

Lemma 2.1. *The maximization problems in* (2.36) *and* (2.37) *are equivalent.*

$$\max_{g_1(\cdot),g_2(\cdot)} \int_a^b \int_c^d \mathcal{H}(x)\mathcal{K}(g_1(\cdot),g_2(\cdot),x,y)dydx, \quad g_1(\cdot)\geq 0, g_2(\cdot)\geq 0, \tag{2.36}$$

$$\max_{g_1(\cdot)} \int_a^b \mathcal{H}(x) \max_{g_2(\cdot)|(g_1(\cdot),x)} \int_c^d \mathcal{K}(g_1(\cdot),g_2(\cdot),x,y)dydx, \quad g_1(\cdot)\geq 0, g_2(\cdot)\geq 0, \tag{2.37}$$

where a,b,c,d are constant and $\mathcal{H}(\cdot)$ and $\mathcal{K}(\cdot)$ are known non-negative functions. The subscript $g_2(\cdot)|(g_1(\cdot),x)$ emphasizes that we should find the optimal $g_2(\cdot)$ once $g_1(\cdot)$ and x are determined by the outer maximization problem. Note that $\mathcal{H}(\cdot)$ depends only on the value of x, and $\mathcal{K}(\cdot)$ depends on the selection of $g_1(\cdot)$ and $g_2(\cdot)$ as well as the values of x and y.

Proof. Let us denote the solution of (2.36) by $(\hat{g}_1(\cdot),\hat{g}_2(\cdot))$ and the solution of (2.37) by $(\check{g}_1(\cdot),\check{g}_2(\cdot))$. Starting from (2.36), we can write:

$$\begin{aligned}
\max_{g_1(\cdot),g_2(\cdot)}\int_a^b\int_c^d \mathcal{H}(x)\mathcal{K}(g_1(\cdot),g_2(\cdot),x,y)dydx &= \int_a^b\int_c^d \mathcal{H}(x)\mathcal{K}(\hat{g}_1(\cdot),\hat{g}_2(\cdot),x,y)dydx \qquad (2.38)\\
&= \int_a^b \mathcal{H}(x)\int_c^d \mathcal{K}(\hat{g}_1(\cdot),\hat{g}_2(\cdot),x,y)dydx\\
&\leq \int_a^b \mathcal{H}(x) \max_{g_2(\cdot)|(\hat{g}_1(\cdot),x)} \int_c^d \mathcal{K}(\hat{g}_1(\cdot),g_2(\cdot),x,y)dydx\\
&\leq \max_{g_1(\cdot)} \int_a^b \mathcal{H}(x) \max_{g_2(\cdot)|(g_1(\cdot),x)} \int_c^d \mathcal{K}(g_1(\cdot),g_2(\cdot),x,y)dydx.
\end{aligned}$$

[4] A comprehensive discussion of variation problems can be found in [45].

On the other hand, if we start from from (2.37), we have:

$$\begin{aligned}\max_{g_1(\cdot)}\int_a^b \mathcal{H}(x)\max_{g_2(\cdot)|(g_1(\cdot),x)}\int_c^d \mathcal{K}(g_1(\cdot),g_2(\cdot),x,y)dydx &= \int_a^b \mathcal{H}(x)\int_c^d \mathcal{K}(\check{g}_1(\cdot),\check{g}_2(\cdot),x,y)dydx \qquad (2.39)\\ &= \int_a^b\int_c^d \mathcal{H}(x)\mathcal{K}(\check{g}_1(\cdot),\check{f}_2(\cdot),x,y)dydx\\ &\le \max_{g_1(\cdot),g_2(\cdot)}\int_a^b\int_c^d \mathcal{H}(x)\mathcal{K}(g_1(\cdot),g_2(\cdot),x,y)dydx.\end{aligned}$$

Combining (2.38) and (2.39), the lemma is proved. □

Using this lemma and noting that $f_{L_1}(l_1) \ge 0, \forall l_1$, we can express (2.32) as follows:

$$R^*_{avg} = \max_{\rho_1(\cdot,\infty)}\int_0^\infty f_{L_1}(l_1)\max_{\rho_2(\cdot,\cdot)|(\rho_1(\cdot,\infty),l_1)}\int_0^\infty f_{L_2}(l_2)\int_0^{l_2}\frac{a\rho_2(a,\tilde{R}_1(l_1))}{1+aI_2(a,\tilde{R}_1(l_1))}dadl_2dl_1, \qquad (2.40)$$

where the outer maximization problem is subject to:

$$\int_0^\infty \rho_1(a,\infty)da = P_1, \qquad (2.41)$$

and the constraints of the inner maximization problem are as follows:

$$\forall \tilde{R}_1(l_1):\int_0^\infty \rho_2(a,\tilde{R}_1(l_1))da = P_2, \qquad (2.42)$$

$$\forall \tilde{R}_1(l_1):\int_0^\infty \frac{a\rho_2(a,\tilde{R}_1(l_1))}{1+aI_2(a,\tilde{R}_1(l_1))}da \le \tilde{R}_1(l_1). \qquad (2.43)$$

Note that, in (2.40), the dependency of $\int_0^\infty f_{L_2}(l_2)\int_0^{l_2}\frac{a\rho_2(a,\tilde{R}_1(l_1))}{1+aI_2(a,\tilde{R}_1(l_1))}dadl_2$ on $\rho_1(\cdot,\infty)$ and l_1 is only through $\tilde{R}_1(l_1)$; therefore, R^*_{avg} can be written as:

$$R^*_{avg} = \max_{\rho_1(\cdot,\infty)}\int_0^\infty f_{L_1}(l_1)\max_{\rho_2(\cdot,\tilde{R}_1(l_1))}\int_0^\infty f_{L_2}(l_2)\int_0^{l_2}\frac{a\rho_2(a,\tilde{R}_1(l_1))}{1+aI_2(a,\tilde{R}_1(l_1))}dadl_2dl_1, \qquad (2.44)$$

where $\max\limits_{\rho_2(\cdot,\tilde{R}_1(l_1))}$ means that we should find the optimal power distribution function at the relay once the value of $\tilde{R}_1(l_1)$ is determined by the outer maximization problem.

Given (2.40)-(2.44), in the next two sections, we will discuss how this two-step maximization problem can be solved using Euler's equations.

2.3.3.1 Relay-Destination Link Maximization Problem

Assume that the rate decoded at the relay is equal to $\tilde{R}_1(l_1)$. Conditioned on $\tilde{R}_1(l_1)$, the objective is to find the optimum power distribution at the relay, namely $\rho_2^*(\cdot, \tilde{R}_1(l_1))$, such that the average received rate at the destination is maximized. In other words, $\rho_2^*(\cdot, \tilde{R}_1(l_1))$ is the solution to the following problem:

$$\begin{aligned} h^*(\tilde{R}_1(l_1)) \quad &= \quad \max_{\rho_2(\cdot,\tilde{R}_1(l_1))} \int_0^\infty f_{L_2}(l_2) \int_0^{l_2} \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} dadl_2 \\ & \qquad s.t. \quad \int_0^\infty \rho_2(a, \tilde{R}_1(l_1))da = P_2, \\ & \qquad\qquad \int_0^\infty \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} da \leq \tilde{R}_1(l_1). \end{aligned} \tag{2.45}$$

In (2.45), $\tilde{R}_1(l_1)$ is known (determined by the outer maximization problem). This problem is equivalent to the case of the rate-limited broadcasting strategy in (2.6). Therefore, the optimum solution is $\rho_2^*(l_2, \tilde{R}_1(l_1)) = -\frac{dI_2^*(l_2,\tilde{R}_1(l_1))}{dl_2}$, for:

$$I_2^*(l_2, \tilde{R}_1(l_1)) = \begin{cases} P_2 & l_2 < l_2^b \\ \frac{1-F_{L_2}(l_2)+\lambda-l_2 f_{L_2}(l_2)}{l_2^2 f_{L_2}(l_2)} & l_2^b \leq l_2 \leq l_2^e \\ 0 & l_2^e < l_2 \end{cases}, \tag{2.46}$$

where $F_{L_2}(l_2)$ is the *cdf* of the second-hop channel fading power and l_2^b, l_2^e are determined as a function of λ to satisfy $I_2^*(l_2^b, \tilde{R}_1(l_1)) = P_2$ and $I_2^*(l_2^e, \tilde{R}_1(l_1)) = 0$, respectively. Finally, λ is computed to satisfy:

$$\int_0^\infty \frac{a\rho_2(a, \tilde{R}_1(l_1))}{1 + aI_2(a, \tilde{R}_1(l_1))} da = \min(\tilde{R}_1(l_1), R_T^*), \tag{2.47}$$

where R_T^* is given in (2.15). Note that (2.47) states that the total transmission rate of the optimal multilayer code at the relay does not exceed R_T^*.

2.3.3.2 Source-Relay Link Maximization Problem

Knowing the optimum solution for the inner maximization problem in (2.45), i.e., $h^*(\tilde{R}_1(l_1))$, R^*_{avg} in (2.44) can be written as:

$$\begin{aligned} R^*_{avg} = \max_{\rho_1(\cdot,\infty)} \quad & \int_0^\infty f_{L_1}(l_1) h^*(\tilde{R}_1(l_1)) dl_1 \\ s.t. \quad & \int_0^\infty \rho_1(a,\infty) da = P_1, \end{aligned} \tag{2.48}$$

where:

$$\tilde{R}_1(l_1) = \int_0^{l_1} \frac{a\rho_1(a,\infty)}{1 + aI_1(a,\infty)} da, \tag{2.49}$$

and $I_1(a,\infty) = \int_a^\infty \rho_1(\gamma,\infty) d\gamma$. As (2.49) shows, $\tilde{R}_1(l_1)$ depends on l_1, $\rho_1(\cdot,\infty)$, and $I_1(\cdot,\infty)$. Furthermore, since $\rho_1(l_1,\infty) = -I'_1(l_1,\infty) = -\frac{dI_1(l_1,\infty)}{dl_1}$, the integrand of (2.48) can be expressed as a function of l_1, I_1, and I'_1. Hereafter, $G(l_1, I_1, I'_1)$ refers to the integrand of (2.48), i.e.:

$$G(l_1, I_1, I'_1) = f_{L_1}(l_1) h^*(\tilde{R}_1(l_1)). \tag{2.50}$$

Using this notation, equation (2.48) takes the form of a fixed end-point variation problem and the optimal $I_1(l_1,\infty)$ can be computed as the solution of the Euler's equation [46]:

$$\frac{\partial G}{\partial I_1} - \frac{d\frac{\partial G}{\partial I'_1}}{dl_1} = 0, \qquad l_1^b \le l_1 \le l_1^e, \tag{2.51}$$

where l_1^b and l_1^e are selected such that $I_1(l_1^b,\infty) = P_1$ and $I_1(l_1^e,\infty) = 0$, respectively, and $I_1(l_1,\infty) = P_1, \forall l_1 \le l_1^b$, and $I_1(l_1,\infty) = 0, \forall l_1 \ge l_1^e$. Furthermore, following from (2.50), we have:

$$\frac{\partial G}{\partial I_1} = f_{L_1}(l_1)h^{*'}(\tilde{R}_1(l_1)) \int_0^{l_1} \frac{a^2 I_1'(a,\infty)}{(1+aI_1(a,\infty))^2} da, \tag{2.52}$$

$$\frac{\partial G}{\partial I_1'} = f_{L_1}(l_1)h^{*'}(\tilde{R}_1(l_1)) \int_0^{l_1} \frac{-a}{1+aI_1(a,\infty)} da, \tag{2.53}$$

$$\begin{aligned}\frac{d\frac{\partial G}{\partial I_1'}}{dl} &= f_{L_1}'(l_1)h^{*'}(\tilde{R}_1(l_1)) \int_0^{l_1} \frac{-a}{1+aI_1(a,\infty)} da + f_{L_1}(l_1)h^{*'}(\tilde{R}_1(l_1)) \frac{-l_1}{1+l_1 I_1(l_1,\infty)} \\ &\quad + f_{L_1}(l_1)h^{*''}(\tilde{R}_1(l_1)) \frac{-l_1 I_1'(l_1,\infty)}{1+l_1 I_1(l_1,\infty)} \int_0^{l_1} \frac{-a}{1+aI_1(a,\infty)} da,\end{aligned} \tag{2.54}$$

where $h^{*'}(\cdot)$ and $h^{*''}(\cdot)$ denote the first and second order derivatives of $h^*(\cdot)$ with $h^*(\cdot)$ given in (2.45). [5] Substituting (2.52)-(2.54) in (2.51), we can simplify the Euler's equation and compute the optimum $I_1(l_1,\infty)$.

A numerical example for a two-hop network in which both hops have a Rayleigh fading distribution are provided in Section 2.5.

2.3.4 Optimality Analysis of Infinite-layer DF Coding

Considering the average received rate at the destination, we first review the optimality of multilayer transmission for single-hop networks. Next, the optimality result for the two-hop network will be presented.

2.3.4.1 Single-hop Network

A single-hop network consists of one source and one destination where the link between the two is quasi-static fading. Following work by Shamai in [15], this network can be modeled as a broadcast system where there is one source and κ, $\kappa \to \infty$, virtual destinations; see Fig. 2.2. In this equivalent model, the channel fading power between the source and each of the virtual destinations is fixed and corresponds to one of the realizations of the channel fading power denoted by $\{l_1, l_2, \cdots, l_\kappa\}$. It is also important to note that the average received rate at the destination in the original model (Fig. 2.2a) translates to the weighted sum of the rate received at the destinations in the broadcast model (Fig. 2.2b). The weight

[5] Since $h^*(\cdot)$ does not have a closed form solution, $h^{*'}(\cdot)$ and $h^{*''}(\cdot)$ should be determined numerically.

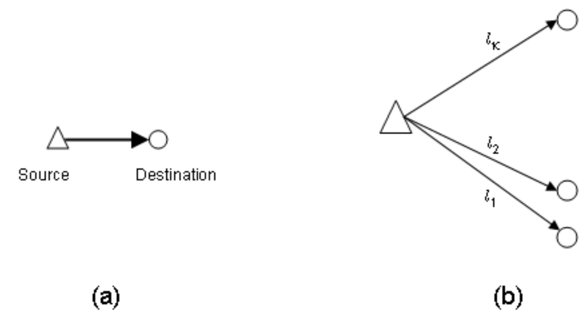

Figure 2.2: A single-user network and its equivalent broadcast network model

of the ith virtual destination is the probability that the channel fading power is equal to l_i.

Note that since all nodes are single antenna units, the resulting broadcasting network is degraded and its capacity region is known [47]. For example, the solid line in Fig. 2.3 shows the boundary of a typical capacity region of a two-user degraded broadcast channel. We also know that the multilayer coding (superposition coding) can achieve the boundary of the capacity region of the degraded broadcast channel [47].

The goal is to find a point of the capacity region which maximizes the weighed sum of the rates received at different virtual destinations. Clearly, since all weights and rates are non-negative, we expect that the optimal point is on the boundary of the capacity region. To determine this point, we can construct a family of κ dimensional planes with a given normal vector (determined based on the weights of the virtual destinations) and select the plane which is tangent to the boundary of the capacity region. For instance, the dashed line in Fig. 2.3 shows the tangent line corresponding to the weights of $\frac{2}{3}$ and $\frac{1}{3}$ for virtual destinations 1 and 2, respectively. The intersection of the dashed line and the boundary of the capacity region shows the point corresponding to the optimal design of the multilayer code.

An immediate consequence is that it is possible to find a multilayer code which maximizes the weighted sum rate in a single-hop network. This is possible because: i) the optimal point is on the boundary of the capacity region, and ii) all boundary points can be achieved using multi-layer coding. This proves that if we want to find the optimal transmission scheme of a single-hop network (with respect to the average received rate at the destination), it is sufficient to search over all possible multilayer codes.

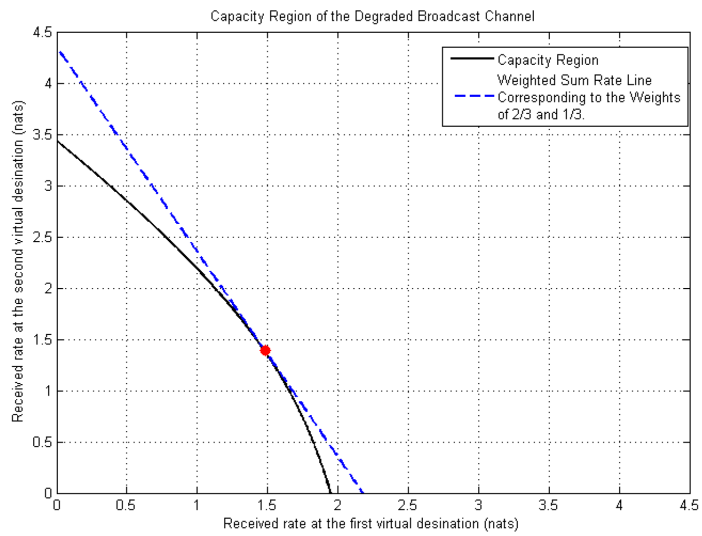

Figure 2.3: Capacity region of a single-user network

2.3.4.2 Rate Limited Single-hop Network

Here, we show that multilayer coding is the optimal transmission scheme (with respect to the average received rate at the destination) for rate-limited single-hop set-up[6] as well.

The equivalent broadcast channel model of this network is similar to the broadcast channel model of Section 2.3.4.1 (see Fig. 2.2b). The only difference is that, in this case, we should also consider the rate constraint at the transmitter of the broadcast channel. In other words, the capacity region of the rate-limited degraded broadcast network is the set of points which are inside the capacity region of the original degraded broadcast channel, and in addition, they are below the surface associated with the input rate constraint. To illustrate, the black solid line in Fig. 2.4 shows the boundary of a capacity region of a typical rate-limited two-user degraded broadcast channel. In this figure, the blue dotted line depicts the boundary of the capacity region of the original degraded broadcast channel (without rate limitation) and the red dashed line captures the source transmission rate limitation. Note that since all nodes are single antenna units, we still have a degraded broadcast setting.

[6]The structure of the rate-limited single-hop network is described in Section 2.1

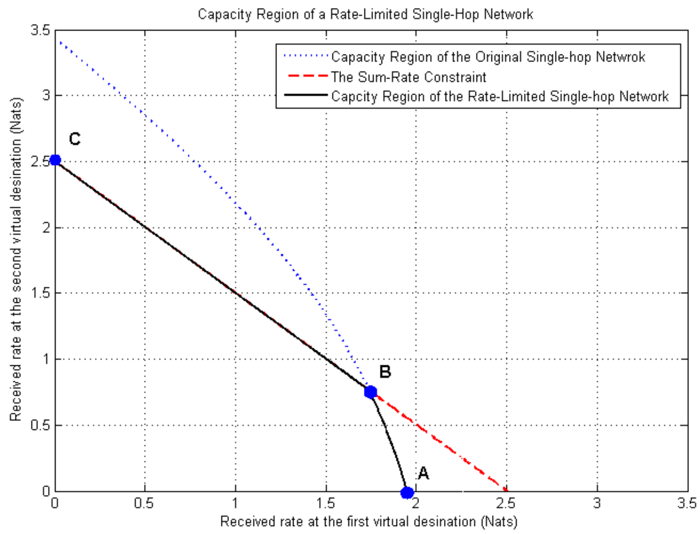

Figure 2.4: Capacity region of a rate-limited single-hop network

The goal here is to find a point in the capacity region which maximizes the weighted sum of the received rates at the virtual destinations. Similar to the previous part, due to the non-negativity of the weights and rates, the optimal point is on the boundary of the resulting capacity region (the solid line). Therefore, the optimal point is either over Arc $\widehat{AB}$ or point **C** (the end point of the line representing the input rate limitation). Arc $\widehat{AB}$ is a part of the original capacity region and all of its points can be achieved using multilayer codes. Furthermore, point **C** can be achieved using a single layer code which is again a special case of multilayer codes. This proves that a multilayer coding scheme is the optimal transmission scheme (with respect to the average received rate at the destination) in rate limited single-hop scenarios.

2.3.4.3 Optimality of Two-hop Infinite-layer DF Coding

First, note that according to the two-hop structure of the network, all the information received by the destination should be first passed through the relay. As a result, the signal received at the destination is a degraded version of what has been received at the relay. In other words, no information can be decoded by the destination unless it is also decodable at the relay. Because of this characteristic of the network, it can be concluded that decode and forward is the optimal relaying scheme in two-hop settings.

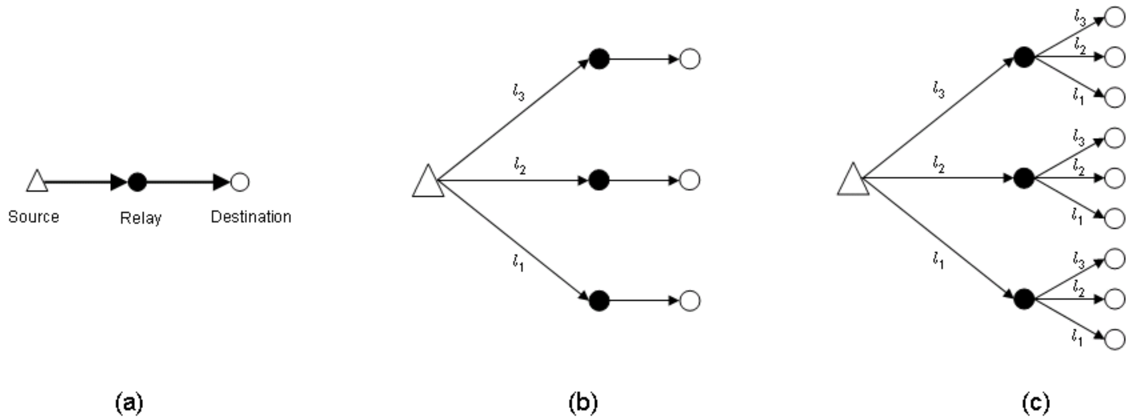

Figure 2.5: A two-hop network and its equivalent broadcast network model where $\kappa = 3$

To analyze the two-hop settings, consider a typical two-hop network depicted in Fig. 2.5a. Assume that the channel fading powers of each hop is randomly selected from a set of discrete values $\{l_1, l_2, \cdots, l_\kappa\}$.[7] The probability associated with each channel state is denoted by $\{\breve{f}_{L_1}(l_1), \breve{f}_{L_1}(l_2), \cdots, \breve{f}_{L_1}(l_\kappa)\}$ and $\{\breve{f}_{L_2}(l_1), \breve{f}_{L_2}(l_2), \cdots, \breve{f}_{L_2}(l_\kappa)\}$ for the first hop and for the second hop, respectively.

Similar to the single-hop network, we construct the equivalent broadcast channel of a two-hop setting. Starting from the source-relay link, we can model it with a broadcast channel consisting of one source and κ virtual relays in which the channel fading power between the source and each of the virtual relays are fixed and corresponds to one of the actual channel states. For illustration, Fig. 2.5b depicts this model for the case of $\kappa = 3$.

Considering the link between one of the virtual relays and its corresponding destination, we can model this link with another broadcast channel in which there are κ virtual destinations (corresponding to different states of the relay-destination link). Figure 2.5c shows an example for the case of $\kappa = 3$.

To find the optimal transmit strategy, we start from the relay-destination link and assume that the source uses an arbitrary transmission scheme such that the ith virtual relay decodes rate B_i where $i \in \{1, 2, \cdots, \kappa\}$.

Given rate B_i at the ith relay, the optimal relaying strategy is to maximize the weighted

[7]This represents a discrete approximation of the channel fading powers. Modeling the continuous channel fading power with discrete values is accurate only if $\kappa \to \infty$. Here, for simplicity, we assume that the channel fading power of both links are selected from the same set. However, the statements of optimality holds in the general case.

sum of the rate at the virtual destinations. To design the optimal code, we note that the transmission strategy of each of any given virtual relay only affects the rate of the virtual destinations which are associated to it. Therefore, the optimal coding scheme can be independently designed for each of the virtual relays, i.e., each virtual relay should transmit data such that it maximizes the weighted sum of the rates associated with its corresponding virtual destinations. Based on this observation, the problem of finding the optimal transmission scheme (conditioned on the value of B_i) for the ith virtual relay can be modeled as finding the optimal code for data transmission over a single-hop network with an input rate limitation. As presented in Section 2.1.1, the optimal transmission strategy in this setting is a multilayer code and the corresponding average received rate, i.e. $h^*(B_i)$, can be computed using (2.45) (or similarly (2.6)) where $\tilde{R}_1(l_1) = B_i$.

Given $h^*(B_i), \forall i \in \{1, 2, \cdots, \kappa\}$, the average received rate at the destination can be written as:

$$\overline{R} = \sum_{i=1}^{\kappa} \breve{f}_{L_1}(l_i) h^*(B_i), \tag{2.55}$$

where $\breve{f}_{L_1}(l_i)$ is the probability of the ith channel state in the source-relay link. Note that different transmission strategies at the source result in different sets of $\{B_i, i \in \{1, 2, \cdots, \kappa\}\}$ at the virtual relays, and consequently, due to (2.55), they result in different values for $\overline{R}$. Based on the equivalent broadcast model, the domain of the achievable sets of $\{B_i, i \in \{1, 2, \cdots, \kappa\}\}$ is the same as the capacity region of the underlying broadcast channel model corresponding to the source-relay link (the source and the virtual relays). Therefore, to find the optimal strategy, we should search over this capacity region and select set $\{B_i, i \in \{1, 2, \cdots, \kappa\}\}$ which maximizes $\overline{R}$. Furthermore, we know that since the equivalent broadcast channel is degraded, all points on the boundary of this capacity region can be achieved using a multilayer code [47]. Therefore, to prove the optimality of the multilayer coding for the source-relay link, it is enough to show that the optimal selection of set $\{B_i, i \in \{1, 2, \cdots, \kappa\}\}$ is a point on the boundary of the capacity region, i.e., we should show that $\frac{\partial \overline{R}}{\partial B_i}$ is positive $\forall i \in \{1, 2, ..., \kappa\}$ where:

$$\frac{\partial \overline{R}}{\partial B_i} = \breve{f}_{L_1}(l_i) h^{*\prime}(B_i). \tag{2.56}$$

By definition, $\breve{f}_{L_1}(l_i)$ is a non-negative value. Moreover, following Section 2.1, we can conclude that $h^{*\prime}(B_i)$ is non-negative (indeed, $h^{*\prime}(B_i) = 0$ for $B_i \geq R_T^*$ and $h^{*\prime}(B_i) > 0$, otherwise). Therefore, $\frac{\partial \overline{R}}{\partial B_i} \geq 0$ for $\forall i \in \{1, 2, ..., \kappa\}$. Consequently, $\overline{R}$ is maximized for a selection of set $\{B_i, i \in \{1, 2, \cdots, \kappa\}\}$ which is on the boundary of the capacity region of the broadcast channel. Thus, multilayer coding is the optimal transmission strategy for the source-relay link as well.

2.4 Infinite-layer DF Coding for Multi-hop Networks

In this section, we study the generalization of the two-hop infinite-layer DF coding strategy to the case of multi-hop scenarios. The following lemma is required in the sequel.

Lemma 2.2. *Consider the following optimization problem:*

$$\begin{aligned} \max_{\rho(\cdot,\Gamma_{in})} \quad & \int_0^\infty f(l) h(R(l,\Gamma_{in})) dl \\ s.t. \quad & \int_0^\infty \rho(a,\Gamma_{in}) da = P, \\ & \int_0^\infty \frac{a\rho(a,\Gamma_{in})}{1 + aI(a,\Gamma_{in})} da \leq \Gamma_{in}, \end{aligned} \tag{2.57}$$

where $I(a,\Gamma_{in}) = \int_\gamma^\infty \rho(a,\Gamma_{in}) d\gamma$, $f(\cdot) \geq 0$, $h(\cdot) \geq 0$, $h(\cdot)$ *is a non-decreasing function,* P *and* Γ_{in} *are given parameters, and:*

$$R(l,\Gamma_{in}) = \int_0^l \frac{a\rho(a,\Gamma_{in})}{1 + aI(a,\Gamma_{in})} da. \tag{2.58}$$

Then, the lemma states that the optimal $\rho(\cdot,\Gamma_{in})$ *which maximizes* (2.57) *can be evaluated as* $\rho^*(l,\Gamma_{in}) = -I^{*\prime}(l,\Gamma_{in}) = \frac{-dI^*(l,\Gamma_{in})}{dl}$ *where* $I^*(l,\Gamma_{in})$ *is the solution of:*

$$\begin{aligned} f(l)h'(R(l,\Gamma_{in})) & \int_0^l \frac{1}{(1 + aI^*(a,\Gamma_{in}))^2} da - f'(l)h'(R(l,\Gamma_{in})) \int_0^l \frac{-a}{1 + aI^*(a,\Gamma_{in})} da \\ - & \quad f(l)h''(R(l,\Gamma_{in})) \frac{-lI^{*\prime}(l,\Gamma_{in})}{1 + lI^*(l,\Gamma_{in})} \int_0^l \frac{-a}{1 + aI^*(a,\Gamma_{in})} da \\ + & \quad \frac{\lambda}{(1 + lI^*(l,\Gamma_{in}))^2} = 0, \qquad l^b \leq l \leq l^e, \end{aligned} \tag{2.59}$$

where l^b satisfies $I^*(l^b, \Gamma_{in}) = P$, l^e satisfies $I^*(l^e, \Gamma_{in}) = 0$, and λ is determined such that:

$$\int_0^\infty \frac{a\rho(a, \Gamma_{in})}{1 + aI(a, \Gamma_{in})} da = \min(\Gamma_{in}, R_T^*), \tag{2.60}$$

with R_T^* given in (2.15)*, and $I^*(l, \Gamma_{in}) = P, \forall l < l^b$ and $I^*(l, \Gamma_{in}) = 0, \forall l > l^e$.*

Proof. We first write (2.57) as a variation problem with a subsidiary constraint. For a given Γ_{in}, let $G_1(l, I, I') = f(l)h(R(l, \Gamma_{in}))$ and $G_2(l, I, I') = \frac{-lI'(l,\Gamma_{in})}{1+lI(l,\Gamma_{in})}$. The optimal $I(l, \Gamma_{in})$ satisfies:

$$\frac{\partial G_1}{\partial I} + \lambda \frac{\partial G_2}{\partial I} - \frac{d\frac{\partial G_1}{\partial I'}}{dl} - \lambda \frac{d\frac{\partial G_2}{\partial I'}}{dl} = 0, \quad l^b \leq l \leq l^e, \tag{2.61}$$

where, for each Γ_{in}, l^b and l^e are selected such that they satisfy $I(l^b, \Gamma_{in}) = P$ and $I(l^e, \Gamma_{in}) = 0$, respectively, and $I^*(l, \Gamma_{in}) = P, \forall l < l^b$ and $I^*(l, \Gamma_{in}) = 0, \forall l > l^e$. Furthermore, we have:

$$\frac{\partial G_1}{\partial I} = f(l)h^{*'}(R(l, \Gamma_{in})) \int_0^l \frac{a^2 I'(a, \infty)}{(1 + aI(a, \infty))^2} da, \tag{2.62}$$

$$\frac{d\frac{\partial G_1}{\partial I'}}{dl} = f'(l)h^{*'}(R(l, \Gamma_{in})) \int_0^l \frac{-a}{1 + aI(a, \infty)} da + f(l)h^{*'}(R(l, \Gamma_{in})) \frac{-l}{1 + lI(l, \infty)} \tag{2.63}$$

$$+ f(l)h^{*''}(R(l, \Gamma_{in})) \frac{-lI'(l, \infty)}{1 + lI(l, \infty)} \int_0^l \frac{-a}{1 + aI(a, \infty)} da,$$

$$\frac{\partial G_2}{\partial I} = \frac{l^2 I'(l, \Gamma_{in})}{(1 + lI(l, \Gamma_{in}))^2}, \tag{2.64}$$

$$\frac{d\frac{\partial G_2}{\partial I'}}{dl} = \frac{-1 + l^2 I'(l, \Gamma_{in})}{(1 + lI(l, \Gamma_{in}))^2}. \tag{2.65}$$

Substituting (2.62)-(2.65) into (2.61) completes the proof of Lemma 2.2. □

Figure 2.6 represents a typical multi-hop network setup. Similar to (2.32), the maximum achievable average received rate at the destination is the solution of the following

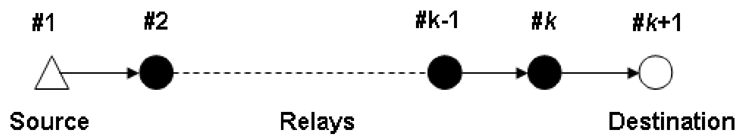

Figure 2.6: Multi-hop network model

optimization problem:

$$\begin{aligned} R_{avg} = & \max_{\substack{\rho_j(\cdot,\cdot) \\ \forall j\in\{1,2,\cdots,k\}}} \int_0^\infty f_{L_1}(l_1) \int_0^\infty f_{L_2}(l_2) \cdots \int_0^\infty f_{L_k}(l_k) R_k(l_k, \Gamma_{in}^k) dl_k \cdots dl_2 dl_1 \quad (2.66) \\ & s.t. \quad \int_0^\infty \rho_j(a, \Gamma_{in}^j) da = P_j, \qquad \forall j \in \{1,2,\cdots,k\}, \\ & \int_0^\infty \frac{a\rho_j(a, \Gamma_{in}^j)}{1 + aI_j(a, \Gamma_{in}^j)} da \leq \Gamma_{in}^j, \quad \forall j \in \{1,2,\cdots,k\}, \end{aligned}$$

where f_{L_j} represents the *pdf* of the channel fading power of the jth hop, i.e., the link between nodes j and $j+1$, $I_j(a, \Gamma_{in}^j) = \int_\gamma^\infty \rho_j(a, \Gamma_{in}^j) d\gamma$. Furthermore, $R_j(l_j, \Gamma_{in}^j)$ is the decodable rate at node $(j+1)$ where the channel fading power of the jth hop is in state l_j and the input rate of node j is Γ_{in}^j. We have:

$$R_j(l_j, \Gamma_{in}^j) = \int_0^{l_j} \frac{a\rho_j(a, \Gamma_{in}^j)}{1 + aI_j(a, \Gamma_{in}^j)} da. \quad (2.67)$$

Note that Γ_{in}^j (input rate of node j) for $j \in \{2, 3, \cdots, k\}$ can be evaluated as $\Gamma_{in}^j = R_{j-1}(l_{j-1}, \Gamma_{in}^{j-1})$ where $\Gamma_{in}^1 = \infty$ (there is no rate constraint at the source).

Similar to Lemma 2.1, (2.66) can be written as:

$$\max_{\rho_1(\cdot,\Gamma_{in}^1)} \int_0^\infty f_{L_1}(l_1) \max_{\rho_2(\cdot,\Gamma_{in}^2)} \int_0^\infty f_{L_2}(l_2) \cdots \max_{\rho_k(\cdot,\Gamma_{in}^k)} \int_0^\infty f_{L_k}(l_k) R_k(l_k, \Gamma_{in}^k) dl_k \cdots dl_2 dl_1. \quad (2.68)$$

The constraints of each maximization problem, say the jth one, is:

$$\int_0^\infty \rho_j(a, \Gamma_{in}^j) da = P_j, \quad (2.69)$$

$$\int_0^\infty \frac{a\rho_j(a, \Gamma_{in}^j)}{1 + aI_j(a, \Gamma_{in}^j)} da \leq \Gamma_{in}^j. \quad (2.70)$$

Based on (2.68), the maximum achievable average received rate at the destination can be found by determining the optimal design of the infinite-layer code corresponding to each node (conditioned on its input rate) starting from node $j = k$ and proceeding sequentially in a reverse order. This algorithm is summarized in the following:

1. Starting from the last hop, this step computes $\rho_k^*(\cdot, \Gamma_{in}^k)$ for all possible values of the corresponding input rate, i.e., Γ_{in}^k:

$$\begin{aligned} h_k^*(\Gamma_{in}^k) = & \max_{\rho_k(\cdot,\Gamma_{in}^k)} \int_0^\infty f_{L_k}(l_k) R_k(l_k, \Gamma_{in}^k) dl_k \\ & s.t. \quad \int_0^\infty \rho_k(a, \Gamma_{in}^k) da = P_k, \\ & \qquad \int_0^\infty \frac{a\rho_k(a, \Gamma_{in}^k)}{1 + aI_k(a, \Gamma_{in}^k)} da \leq \Gamma_{in}^k. \end{aligned} \tag{2.71}$$

   The resulting maximization problem has the form of (2.57) where $h(R_k(l_k, \Gamma_{in}^k)) = R_k(l_k, \Gamma_{in}^k)$. Therefore, optimal $\rho_k(\cdot, \Gamma_{in}^k)$, or equivalently $I_k^*(\cdot, \Gamma_{in}^k)$, can be found using Lemma 2.2 where (2.59) is satisfied with $h'(R_k(l_k, \Gamma_{in}^k)) = 1$, and $h''(R_k(l_k, \Gamma_{in}^k)) = 0$.

2. This step computes the optimal power distribution for node $k-1$, i.e., $\rho_{k-1}^*(\cdot, \Gamma_{in}^{k-1})$, such that:

$$\begin{aligned} h_{k-1}^*(\Gamma_{in}^{k-1}) &= \max_{\rho_{k-1}(\cdot,\Gamma_{in}^{k-1})} \int_0^\infty f_{L_{k-1}}(l_{k-1}) \max_{\rho_k(\cdot,\Gamma_{in}^k)} \int_0^\infty f_{L_k}(l_k) R_k(l_k, \Gamma_{in}^k) dl_k dl_{k-1} \\ &\overset{(a)}{=} \max_{\rho_{k-1}(\cdot,\Gamma_{in}^{k-1})} \int_0^\infty f_{L_{k-1}}(l_{k-1}) h_k^*(\Gamma_{in}^k) dl_{k-1} \\ &\overset{(b)}{=} \max_{\rho_{k-1}(\cdot,\Gamma_{in}^{k-1})} \int_0^\infty f_{L_{k-1}}(l_{k-1}) h_k^*(R_{k-1}(l_{k-1}, \Gamma_{in}^{k-1})) dl_{k-1} \\ & \quad s.t. \quad \int_0^\infty \rho_{k-1}(a, \Gamma_{in}^{k-1}) da = P_{k-1}, \\ & \qquad\quad \int_0^\infty \frac{a\rho_{k-1}(a, \Gamma_{in}^{k-1})}{1 + aI_{k-1}(a, \Gamma_{in}^{k-1})} da \leq \Gamma_{in}^{k-1}, \end{aligned} \tag{2.72}$$

   where (a) is derived by substituting the last maximization with its optimal solution, i.e., $h_k^*(\Gamma_{in}^k)$, and (b) is due to $\Gamma_{in}^k = R_{k-1}(l_{k-1}, \Gamma_{in}^{k-1})$. The resulting optimization problem, again, has the form of (2.57). Therefore, the optimal $\rho_{k-1}(\cdot, \Gamma_{in}^{k-1})$, or equivalently $I_{k-1}^*(\cdot, \Gamma_{in}^{k-1})$, can be found based on Lemma 2.2.

3. The parameters of the optimal codes for the remaining nodes can be computed similarly. For instance, for node j, we should first replace the last $k-j$ maximizations with $h^*_{j+1}(\Gamma_{in}^{j+1})$ and then use Lemma 2.2 to solve the resulting optimization problem and determine $\rho^*_j(\cdot, \Gamma_{in}^{j})$.

2.5 Numerical Results

This section presents some numerical results related to a two-hop network with *Rayleigh fading* in both hops. The performance of the proposed scheme is compared with the cut-set bound and with two other sub-optimal schemes proposed in [24].

Infinite-layer DF Coding: This scheme follows the design steps presented in Section 2.3.3.

Step I: Concerning the relay-destination link, (2.46) and (2.47) provide the optimum $\rho_2(l_2, \tilde{R}_1(l_1))$ for different relay input rates, namely $\tilde{R}_1(l_1)$, where due to the Rayleigh fading assumption we have $F_{L_2}(l_2) = 1 - e^{-l_2}$ and $f_{L_2}(l_2) = e^{-l_2}$. Having obtained the optimum power distribution of the relay (for each $\tilde{R}_1(l_1)$), we can use (2.45) to find the maximum average received rate at the destination. Figure 2.7 shows $h^*(\tilde{R}_1(l_1))$ for different values of $\tilde{R}_1(l_1)$ and P_2.

Step II: Concerning the source-relay link, the optimum $\rho_1(l_1, \infty)$ can be determined by finding $I_1(l_1, \infty)$ such that it satisfies (2.51). Note that due to (2.52)-(2.54) and the Rayleigh fading assumption, i.e., $f_{L_1}(l_1) = e^{-l_1}$, (2.51) can be simplified as:

$$h^{*'}(\tilde{R}_1(l_1)) \int_0^{l_1} \frac{1 - a - a^2 I_1(a,\infty)}{(1 + aI_1(a,\infty))^2} da - h^{*''}(\tilde{R}_1(l_1)) \frac{-l I_1'(l_1,\infty)}{1 + l I_1(l_1,\infty)} \int_0^{l_1} \frac{-a}{1 + aI_1(a,\infty)} da = 0, \tag{2.73}$$

where $\tilde{R}_1(l_1)$ is given in (2.49) and $h^{*'}(\tilde{R}_1(l_1))$, $h^{*''}(\tilde{R}_1(l_1))$ are the first and the second derivative of $h^*(\tilde{R}_1(l_1))$. To solve (2.73), we approximate the continuous function $I_1(l_1, \infty)$ with a discrete function $\breve{I}_1(m), m \in \{1, 2, ..., N\}$ where $\breve{I}_1(1) = P_1$ and $\breve{I}_1(N) = 0$. Note that $h^*(\tilde{R}_1(l_1))$ is already derived by the result of Step I. The optimal power distribution is then computed using numerical methods.

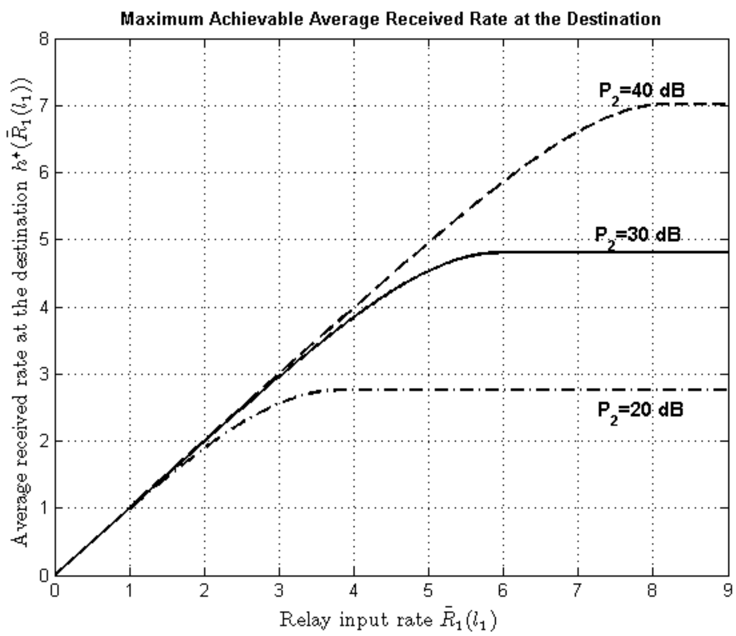

Figure 2.7: Maximum average received rate at the destination for the case of Rayleigh fading two-hop network

Figure 2.8 depicts $I_1^*(l_1, \infty)$ for $P_1 = P_2 = 20$dB assuming Rayleigh fading. Having $I_1^*(l_1, \infty)$, the amount of power associated with each layer of the code can be determined using $\rho_1^*(l_1, \infty) = -{I^*}'_1(l_1, \infty)$. The optimal $\rho_1^*(l_1, \infty)$ can then be used in (2.48) to find the maximum average received rate at the destination. The dashed lines with square marks in Fig. 2.9 and Fig. 2.10 represent the average received rate at the destination versus the relay power P_2 where $P_1 = 20$dB and $P_1 = 30$dB, respectively.

Broadcasting Cutset Bound: This bound simply indicates that the achievable average received rate at the destination of a two-hop network cannot exceed the achievable average rate of the associated single-hop links, i.e., the source-relay link and the relay-destination link [24]. This is independent of the relay structure and its relaying mode. Based on (2.13), the cutset bound can be written as[8]:

$$C_{cutset} = \min\left(\int_0^\infty f_{L_1}(l_1) \int_0^{l_1} \frac{a\rho_1^*(a,\infty)}{1 + aI_1^*(a,\infty)} da dl_1, \int_0^\infty f_{L_2}(l_2) \int_0^{l_2} \frac{a\rho_2^*(a,\infty)}{1 + aI_2^*(a,\infty)} da dl_2\right), \quad (2.74)$$

[8]See [24] for more details.

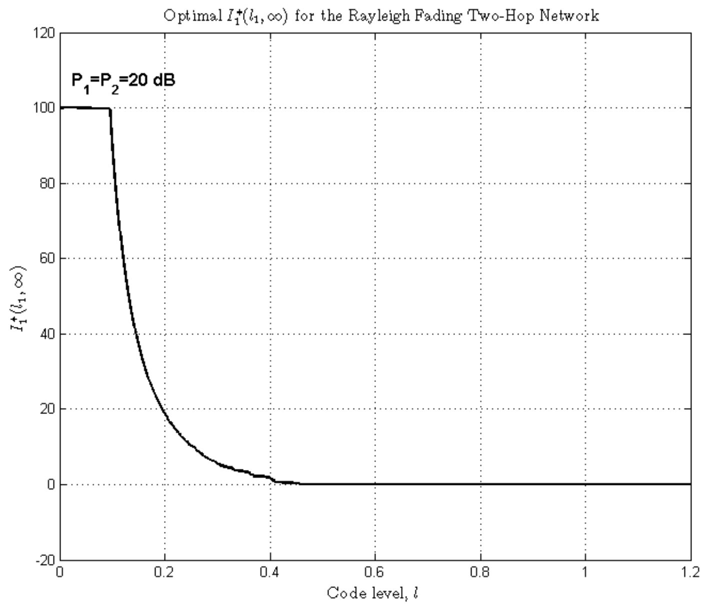

Figure 2.8: Optimal $I_1(l_1, \infty)$ for the case of Rayleigh fading two-hop network

where the first term shows the maximum achievable rate at the relay (considering the source-relay link) and the second term shows the maximum achievable rate at the destination (considering the relay-destination link). Furthermore, $I_1^*(a, \infty)$ and consequently $\rho_1^*(a, \infty)$ can be evaluated using (2.14) where $f_L(l)$ is substituted by $f_{L_1}(l_1)$. Similarly, $I_2^*(a, \infty)$ and consequently $\rho_2^*(a, \infty)$ can be evaluated using (2.14) where $f_L(l)$ is substituted by $f_{L_2}(l_2)$. The corresponding upper bounds versus different P_2 values are shown in Fig. 2.9 and Fig. 2.10 where $P_1 = 20$dB and $P_1 = 30$dB, respectively.

Infinite-layer AF Coding: In this scheme, the relay amplifies and forwards the received signal. To design the optimum infinite-layer code, first, the equivalent end-to-end channel should be found. In other words, the source-relay and the relay-destination channels combined with AF relaying can be modeled as a single channel with a new probability density function for the effective fading power. Once this new *pdf* is obtained, the optimum power distribution can be evaluated. The maximum achievable average received rate of this scheme is presented in [24]. Figure 2.9 and Fig. 2.10 depict the corresponding maximum average received rate at the destination where $P_1 = 20$dB and $P_1 = 30$dB, respectively.

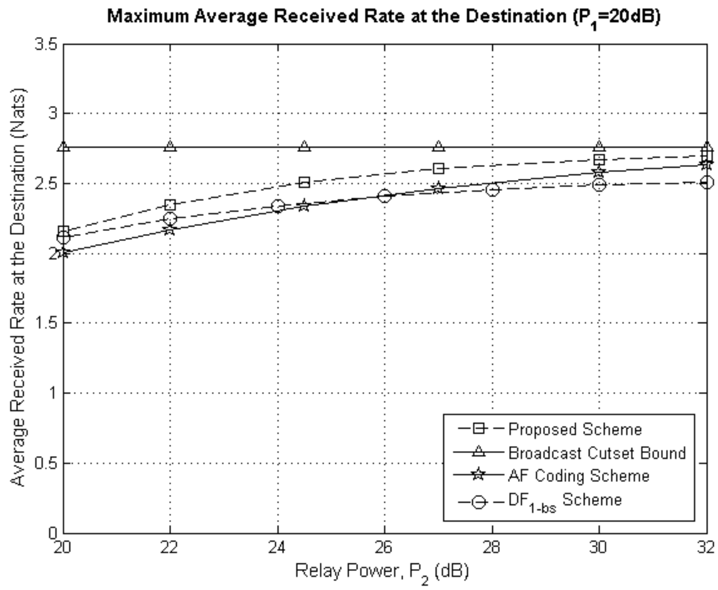

Figure 2.9: Average received rate at the destination P_1=20 dB, P_2=20 - 32 dB

Outage at the Source, Broadcasting at the Relay: This is another suboptimal strategy that has been studied in [24]. In this case, the source uses a one-layer code, known as the *outage approach*, and the relay uses an infinite-layer code. We refer to this scheme by $\mathbf{DF_{1-bs}}$.[9] The achievable average received rate of this scheme is computed in [24]. For the sake of comparison, this achievable average received rate is evaluated for different relay powers and presented in Fig. 2.9 and Fig. 2.10 where $P_1 = 20$dB and $P_1 = 30$dB, respectively.

As Fig. 2.9 and Fig. 2.10 show, in both network setups, the proposed infinite-layer DF coding strategy is strictly superior to the infinite-layer AF coding. Note that, prior to this work, the infinite-layer AF coding was the best known high SNR transmission strategy for two-hop networks [24]. As expected, Fig. 2.9 and Fig. 2.10 also show that the proposed scheme achieves a superior performance as compared to $\mathbf{DF_{1-bs}}$ ($\mathbf{DF_{1-bs}}$ is a special case of the proposed infinite-layer DF coding). Furthermore, as P_2 increases, the rate of the proposed scheme approaches the cutset bound which means that for high values of P_2, relays without buffering capability have small degradation compared to the ideal performance. In addition, Fig. 2.9 and Fig. 2.10 demonstrate that, as P_1 decreases,

[9]For consistency, subscript $\mathbf{1-bs}$ is used according to the notation used in [24]. This subscript represents the one-layer coding and the broadcasting scheme at the source and at the relay, respectively.

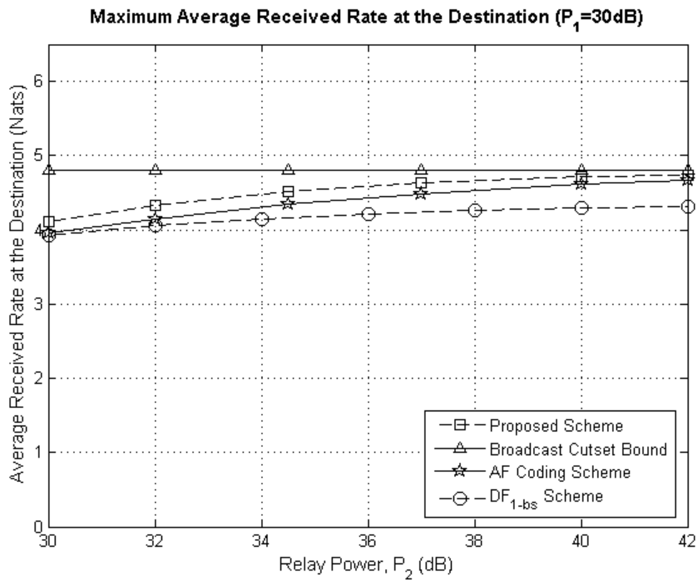

Figure 2.10: Average received rate at the destination, P_1=30 dB, P_2=30 - 42 dB

the performance of $\mathbf{DF_{1-bs}}$ approaches the optimal performance. This is because if the power of the relay is much smaller than the power of the source, the performance of the relay-destination link limits the performance of the system. Therefore, in most cases, even if the source transmits using a one-layer code, the relay is able to decode a rate of R_T^* (if the relay receives a rate more than R_T^*, the extra rate should be discarded).

2.6 Summary

In this chapter, we first examined a quasi-static two-hop network in which data is transmitted from a source via a relay to a destination. It is assumed that knowledge of the channel for each hop is not available at the corresponding transmitter, and the relay is not capable of data buffering over multiple transmission blocks. For this network setup, an infinite-layer coding scheme is proposed at the source, as well as at the relay. Furthermore, an optimal power distribution function is derived which is used to determine the amount of power that should be assigned to each code layer at the source and at the relay. We also discussed how these results can be generalized to multi-hop settings. Finally, we provided some numerical results that demonstrate the superior performance of the infinite-layer DF coding over the infinite-layer AF coding.

Chapter 3

Transmission over Single-hop Multi-antenna Networks

A number of recent source-coding standards compress the video signal into multiple layers such that a destination is able to reconstruct the video (with a lower quality) even if it has not received all layers. Implementation of such source coding techniques in wireless networks requires the application of coding mechanisms (such as multilayer coding) which allow unequal error protection for different layers of the coded video stream. In this chapter[1], we study the performance of multilayer coding for quasi-static fading channels where the source and the destination are equipped with multiple antennas and the CSI is only known at the destination. The objective is to maximize the average received rate at the destination.

The organization of this chapter is as follows: Section 3.2 proposes a design rule for construction of multilayer codes for MIMO systems. As a preliminary example, Section 3.3 shows how the proposed design rule can be used to construct the optimal multilayer code for single-antenna systems. Afterwards, Section 3.4 presents the code design algorithm for the case of MIMO systems. The performance of the designed code is then evaluated for different network setups and Section 3.5 shows how multilayer coding can improve the performance of wireless networks.

Notation: The input rate of the source in this scenario is infinity, i.e. $\Gamma_{in} = \infty$ and based on the notations of Section 2.1.2, we should use $R(\cdot, \infty)$, $\rho(\cdot, \infty)$, and $I(\cdot, \infty)$. However,

[1] The work in this chapter is partially reported in [23, 48].

with a slight misuse of notations, in this chapter, we denote $R(\cdot,\infty)$, $\rho(\cdot,\infty)$, and $I(\cdot,\infty)$ simply by $R(\cdot)$, $\rho(\cdot)$, and $I(\cdot)$, respectively.

3.1 Network Model

Consider a link where the source and the destination are equipped with M and N antennas, respectively. The channel is quasi-static (channel gain remains fixed during one transmission block) and CSI is only available at the destination (the source only knows the statistical behavior of the channel). It is also assumed that the destination is only able to perform successive decoding (joint-decoding is not possible). For this network, we have:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{3.1}$$

where $\mathbf{x}$ and $\mathbf{y}$ are $(M\times 1)$ input and $(N\times 1)$ output vectors, respectively. The additive noise vector $\mathbf{n}$ is an N-dimensional vector with zero-mean complex Gaussian i.i.d. distributed entries. The channel propagation matrix $\mathbf{H}$ is assumed to be a circularly symmetric $M\times N$ matrix with entries that have zero mean i.i.d. distributions. Furthermore, K denotes the rank of $\mathbf{H}\mathbf{H}^{\dagger}$ which is equal to $\min(M,N)$ since $\mathbf{H}$ is full rank.

The transmitted signal $\mathbf{x}$ is a multilayer code which consists of the superposition of a continuum layers of Gaussian coded information. Let $\mathbf{x}_l$ denote the stream that corresponds to the lth layer of the code. Similar to Section 2.1, we assume that the source assigns $\rho(l)dl$ portion of its normalized power to transmit $\mathbf{x}_l$. Furthermore, the incremental rate of code layer $\mathbf{x}_l$ is denoted by $\zeta(l)dl$ and $I(l)$ is defined as $I(l) = \int_l^{\infty}\rho(\gamma)d\gamma$.

In this work, unlike [21], we do not limit the study to a specific selection of $\zeta(l)$, i.e., (2.2). In contrast, we find the optimal multilayer code for an arbitrary $\zeta(l)$ and therefore the result provides a general rule for the design multilayer codes.

3.2 Multilayer Coding Scheme for Multiple Antenna Systems

In this section, we aim to extend the results of Section 2.1 for MIMO systems, and present an algorithm to determine families of optimal multilayer codes.

3.2.1 Successive Decoding Requirement at the Destination

As mentioned in Section 3.1, in this work, we assume that the destination has only the capability of successive decoding and is not able to perform joint decoding on the received signal. It means that the destination cannot decode layer ℓ_1 before it decodes all code layers l where $l < \ell_1$. In other words, the code layer ℓ_1, i.e. $\mathbf{x}_{\ell_1}$, can be decoded by the destination if for all $l \leq \ell_1$, the mutual information between the received signal $\mathbf{y}$ and $\mathbf{x}_l$ is larger than the differential rate of $\mathbf{x}_l$, i.e.,

$$\zeta(l)dl \leq \mathcal{I}(\mathbf{y}; \mathbf{x}_l|\mathbf{x}_a, \quad \forall a \leq l), \forall l \leq \ell_1, \tag{3.2}$$

where $\mathcal{I}(\mathbf{y}; \mathbf{x}_l|\mathbf{x}_a, \forall a \leq l)$ represents the mutual information between $\mathbf{y}$ and $\mathbf{x}_l$ where all $\mathbf{x}_a, a < l$ have been already decoded.

3.2.2 Formulating the Average Received Rate at the Destination (R_{avg})

The average received rate at the destination, in MIMO cases, can be formulated by either of the following two expressions in (3.3) and (3.4):

$$R_{avg} = \int_{\mathbf{H}} f_{\mathbb{H}}(\mathbf{H}) R(\mathbf{H}) d\mathbf{H}, \tag{3.3}$$

where $f_{\mathbb{H}}(\mathbf{H})$ denotes the *pdf* of the channel propagation matrix, and $R(\mathbf{H})$ represents the decodable rate at state $\mathbf{H}$. As shown in [21], the calculation of (3.3) for Single-Input Single-Output (SISO) systems is straightforward. The reason is that in SISO cases, matrix $\mathbf{H}_i$ should be substituted by a complex number which simplifies the integration of (3.3).

Unfortunately, this simplification is not applicable in MIMO scenarios. To overcome this mathematical complexity, we evaluate R_{avg} by integrating $\mathcal{P}(l)\zeta(l)dl$ over all layers of the code, i.e.:

$$R_{avg} = \int_0^\infty \mathcal{P}(l)\zeta(l)dl, \tag{3.4}$$

where $\mathcal{P}(l)$ represents the probability that the channel is in a state that the lth code layer can be decoded, and $\zeta(l)dl$ shows the differential rate associated to the code layer l.

To simplify (3.4), we start by finding the probability that the destination can decode a certain code layer, namely $\mathcal{P}(\ell_1)$. Due to the successive decodability scheme, this probability is equivalent to the probability of:

$$\zeta(l)dl \le \mathcal{I}(\mathbf{y}; \mathbf{x}_l|\mathbf{x}_a, \forall a \le l), \forall l \le \ell_1. \tag{3.5}$$

To compute $\mathcal{P}(\ell_1)$, we assume that $\zeta(l)$ is selected such that if for a particular channel realization, we have $\zeta(\ell_1)dl \le \mathcal{I}(\mathbf{y}; \mathbf{x}_{\ell_1}|\mathbf{x}_a, \forall a \le \ell_1)$, we also have $\zeta(l)dl \le \mathcal{I}(\mathbf{y}; \mathbf{x}_l|\mathbf{x}_a, \forall a \le l), \forall l \le \ell_1$. In the sequel, we refer to this property as the *successive decodability requirement*.

To show the decodability of layer ℓ_1 (assuming that $\zeta(l)$ satisfies the *successive decodability requirement*), it is enough to verify (3.5) for only ℓ_1 (not for all $l < \ell_1$). Therefore, $\mathcal{P}(\ell_1)$ can be written as:

$$\begin{aligned} \mathcal{P}(\ell_1) &= \mathbf{Pr}\{\zeta(l)dl \le \mathcal{I}(\mathbf{y}; \mathbf{x}_l|\mathbf{x}_a, \forall a \le l), \forall l < \ell_1\} \\ &= \mathbf{Pr}\{\zeta(\ell_1)dl \le \mathcal{I}(\mathbf{y}; \mathbf{x}_{\ell_1}|\mathbf{x}_a, \forall a \le \ell_1)\}. \end{aligned} \tag{3.6}$$

Equation (3.6) can be further simplified noting that $\mathbf{H}$ is a circularly symmetric $M \times N$ matrix where its entries have zero mean i.i.d. distributions. For such a channel, [49] shows that equal power distribution between antenna elements maximizes the mutual information between the source and the destination. Thus, $\mathcal{I}(\mathbf{y};\mathbf{x}_{\ell_1}|\mathbf{x}_a,\forall a\le \ell_1)$ can be written as:

$$\begin{aligned}\mathcal{I}(\mathbf{y};\mathbf{x}_{\ell_1}|\mathbf{x}_a,\forall a\le \ell_1) =&\mathcal{I}(\mathbf{y};\mathbf{x}_{l\ge\ell_1}|\mathbf{x}_a,\forall a<\ell_1)-\mathcal{I}(\mathbf{y};\mathbf{x}_{l>\ell_1}|\mathbf{x}_a,\forall a\le \ell_1)\\ \overset{(a)}{=}&\log\det(\mathbf{I}+\frac{I(\ell_1)}{M}\mathbf{H}\mathbf{H}^\dagger)-\log\det(\mathbf{I}+\frac{I(\ell_1+dl)}{M}\mathbf{H}\mathbf{H}^\dagger)\\ =&-dI(l)\frac{d}{dI(l)}\log\det\left(\mathbf{I}+\frac{I(l)}{M}\mathbf{H}\mathbf{H}^\dagger\right)\Big|_{l=\ell_1}\\ \overset{(b)}{=}&-dI(l)\mathrm{Tr}\left(\left[\mathbf{I}+\frac{I(l)}{M}\mathbf{H}\mathbf{H}^\dagger\right]^{-1}\times\frac{\partial(\mathbf{I}+\frac{I(l)}{M}\mathbf{H}\mathbf{H}^\dagger)}{\partial I(l)}\right)\Big|_{l=\ell_1},\end{aligned}\tag{3.7}$$

where (a) can be justified using the results of [49], and (b) is based on $\frac{d\det(\mathbf{G})}{da}=\det(\mathbf{G})\times \mathrm{Tr}\left(\mathbf{G}^{-1}\frac{d\mathbf{G}}{da}\right)$, where $\mathbf{G}$ is an invertible matrix and its elements are independent of each other [50]. Equation (3.7) can be further simplified as:

$$\begin{aligned}\mathcal{I}(\mathbf{y};\mathbf{x}_{\ell_1}|\mathbf{x}_a,\forall a\le \ell_1)=&-dI(l)\mathrm{Tr}\left(\left[\mathbf{I}+\frac{I(l)}{M}\mathbf{H}\mathbf{H}^\dagger\right]^{-1}\frac{\mathbf{H}\mathbf{H}^\dagger}{M}\right)\Big|_{l=\ell_1}\\ =&-dI(l)\mathrm{Tr}\left(\left[\mathbf{I}+\frac{I(l)}{M}\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger\right]^{-1}\frac{\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger}{M}\right)\Big|_{l=\ell_1}\\ \overset{(a)}{=}&-dI(l)\mathrm{Tr}\left(\left[\mathbf{V}\left(\mathbf{I}+\frac{I(l)}{M}\boldsymbol{\Lambda}\right)\mathbf{V}^\dagger\right]^{-1}\frac{\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger}{M}\right)\Big|_{l=\ell_1}\\ =&-dI(l)\mathrm{Tr}\left(\mathbf{V}\left[\mathbf{I}+\frac{I(l)}{M}\boldsymbol{\Lambda}\right]^{-1}\mathbf{V}^\dagger\frac{\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger}{M}\right)\Big|_{l=\ell_1}\\ \overset{(b)}{=}&-dI(l)\mathrm{Tr}\left(\left[\mathbf{I}+\frac{I(l)}{M}\boldsymbol{\Lambda}\right]^{-1}\frac{\boldsymbol{\Lambda}}{M}\right)\Big|_{l=\ell_1}\\ \overset{(c)}{=}&-dI(l)\sum_{i=1}^{K}\frac{\frac{\lambda_i}{M}}{1+\frac{I(l)\lambda_i}{M}}\Big|_{l=\ell_1}\\ =&-I'(\ell_1)\sum_{i=1}^{K}\frac{\lambda_i}{M+I(\ell_1)\lambda_i}\,dl,\end{aligned}\tag{3.8}$$

where $\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger$ represents the eigenvalue decomposition of matrix $\mathbf{H}\mathbf{H}^\dagger$ where $\lambda_i, i\in\{1,\cdots,K\}$ represent the non-zero eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ and $\mathbf{V}$ is the unitary matrix of corresponding eigenvectors. In (3.8), (a) and (b) are due to $\mathbf{I}=\mathbf{V}\mathbf{V}^\dagger$ and $\mathrm{Tr}(\mathbf{AB})=\mathrm{Tr}(\mathbf{BA})$ identities, respectively. Furthermore, (c) is justified using the fact that both $\left(\mathbf{I}+\frac{I(l)}{M}\mathbf{\Lambda}\right)^{-1}$ and $\frac{\mathbf{\Lambda}}{M}$ are diagonal matrices and we know that the product of two diagonal matrices is diagonal as well. Substituting (3.8) in (3.6), we have:

$$\mathcal{P}(\ell_1) = \mathbf{Pr}\left\{\zeta(\ell_1)dl \leq -I'(\ell_1)\sum_{i=1}^{K}\frac{\lambda_i}{M+I(\ell_1)\lambda_i}\,dl\right\}. \tag{3.9}$$

To simplify the notation, we define random variable U^{α} as:

$$U^{\alpha} = \sum_{i=1}^{K}\frac{\lambda_i}{M+\alpha\lambda_i}, \tag{3.10}$$

where α is a constant. Based on this notation and (3.9), we have:

$$\mathcal{P}(\ell_1) = \mathbf{Pr}\{\zeta(\ell_1) \leq -I'(\ell_1)\times U^{I(\ell_1)}\}. \tag{3.11}$$

It is noted that depending on the *pdf* of matrix $\mathbf{H}$ entries, the joint *pdf* of the λ_is', i.e., $f(\lambda_1,\lambda_2,\cdots,\lambda_K)$, can be determined. For instance, [51] presents $f(\lambda_1,\lambda_2,\cdots,\lambda_K)$ for the case where the matrix elements have zero-mean complex Gaussian distribution, $\mathcal{CN}(0,1)$. Knowing $f(\lambda_1,\lambda_2,\cdots,\lambda_K)$, it is possible to compute the *pdf* and *cdf* of $U^{I(\ell_1)}$ and consequently $\mathcal{P}(\ell_1)$. The terms $f_{U^{\alpha}}(u)$ and $F_{U^{\alpha}}(u)$ denote the *pdf* and *cdf* of U^{α}, respectively.

Given (3.11), the probability of successful decoding of layer ℓ_1, where $l^b \leq \ell_1 \leq l^e$, can be written as:

$$\begin{aligned}\mathcal{P}(\ell_1) &= \mathbf{Pr}\left\{U^{I(\ell_1)} \geq \frac{\zeta(\ell_1)}{-I'(\ell_1)}\right\} \\ &= 1-\mathbf{Pr}\left\{U^{I(\ell_1)} \leq \frac{\zeta(\ell_1)}{-I'(\ell_1)}\right\} \\ &= 1-F_{U^{I(\ell_1)}}\left(\frac{\zeta(\ell_1)}{-I'(\ell_1)}\right),\end{aligned} \tag{3.12}$$

where $F_{U^{I(\ell_1)}}$ represents the *cdf* of the random variable $U^{I(\ell_1)}$ and l^b, l^e are selected such that $I(l) = P, \forall l \leq l^b$ and $I(l) = 0, \forall l^e \leq l$, respectively. Note that, in (3.12), $I'(\cdot)$ is always a negative value. Substituting (3.12) in (3.4), we have the following theorem.

Theorem 3.1. *Consider a single-user MIMO system in which the source uses multilayer coding and the destination is able to perform successive decoding. In this network, the average received rate at the destination can be evaluated as:*

$$R_{avg} = \int_{l^b}^{l^e} \left[1 - F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)\right] \zeta(l) dl. \tag{3.13}$$

where l^b, l^e are selected such that $I(l) = P, \forall l \leq l^b$ and $I(l) = 0, \forall l^e \leq l$, respectively. Note that this theorem is valid only if $\zeta(l)$ is selected such that it satisfies the successive decodability requirement.

As (3.13) shows, R_{avg} depends on l, $I(\cdot)$, $I'(\cdot)$, and $\zeta()$; therefore, our goal is to find $I(\cdot)$ and $\zeta(\cdot)$ such that they maximize (3.13). This maximization problem is discussed in the next section.

3.2.3 Maximizing the Average Received Rate at the Destination (R_{avg})

Hereafter $G(l, I, I', \zeta)$ refers to the integrand of (3.13), i.e.,

$$G(l, I, I', \zeta) = Q(l, I, I', \zeta)\zeta(l), \tag{3.14}$$

where $Q(l, I, I', \zeta) = 1 - F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)$. Using this notation and (3.13), R^*_{avg} can be written as:

$$R^*_{avg} = \max_{I(\cdot),\zeta(\cdot)} \int_{l^b}^{l^e} G(l, I, I', \zeta) dl, \tag{3.15}$$

where l^b, l^e are selected such that $I(l) = P, \forall l \leq l^b$ and $I(l) = 0, \forall l^e \leq l$, respectively. Note that this problem has the form of a variation problem with two functions $\zeta(\cdot)$ and $I(\cdot)$. To find the optimal solution of (3.15), we use the following theorem [45].

Theorem 3.2. *[45], A necessary condition for the curves:*

$$y_i = y_i(x), \qquad \forall i \in \{1, \cdots, n\}, \tag{3.16}$$

to be an extremum of:

$$\int_b^a F(x, y_1, \cdots, y_n, y_1', \cdots, y_n')dx, \tag{3.17}$$

is that for all $i \in \{1, \cdots, n\}$, $y_i(x)$ *satisfies the following Euler equations for all* $a \leq x \leq b$*:*

$$F_{y_i} - \frac{d}{dx}F_{y_i'} = 0. \tag{3.18}$$

where the subscripts denote the partial derivatives with respect to the corresponding arguments.

Using Theorem 3.2, the optimal solution of the maximization problem in (3.15) can be determined as:

$$\frac{\partial G}{\partial I} - \frac{d}{dl}(\frac{\partial G}{\partial I'}) = 0, \tag{3.19}$$

$$\frac{\partial G}{\partial \zeta} - \frac{d}{dl}(\frac{\partial G}{\partial \zeta'}) = 0. \tag{3.20}$$

where $l^b \leq l \leq l^e$ and l^b, l^e are selected such that $I(l) = P, \forall l \leq l^b$ and $I(l) = 0, \forall l^e \leq l$, respectively.

In the following, we show that equations (3.19) and (3.20) are not independent, i.e., if G satisfies (3.20) it also satisfies (3.19). To prove, we first simplify (3.20) for $l^b \leq l \leq l^e$ where:

$$\begin{aligned} \frac{\partial G}{\partial \zeta} &= \frac{\partial Q(l, I, I', \zeta)}{\partial \zeta(l)} \times \zeta + Q(l, I, I', \zeta) \times 1, \\ \frac{\partial G}{\partial \zeta'} &= 0. \end{aligned} \tag{3.21}$$

Therefore, to satisfy (3.20), $\zeta(l)$ and $Q(l, I, I', \zeta)$ should be selected such that:

$$-\frac{\partial Q(l, I, I', \zeta)}{\partial \zeta} \times \zeta(l) = Q(l, I, I', \zeta), \forall l^b \leq l \leq l^e. \tag{3.22}$$

We also have the following lemma.

Lemma 3.1. *Consider $Z(x,y)$ defined as:*

$$Z(x,y) = \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}, \tag{3.23}$$

where u is a random variable and x, y are two given numbers. The lemma states that we have the following identity:

$$\frac{\partial Z(x,y)}{\partial y} = \frac{-x}{y} \times \frac{\partial Z(x,y)}{\partial x}. \tag{3.24}$$

Proof. We have:

$$Z(x,y) = \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}. \tag{3.25}$$

Therefore:

$$\begin{aligned} \frac{\partial Z(x,y)}{\partial x} &= \lim_{dx \to 0} \frac{\mathbf{Pr}\left\{u \le \frac{-x-dx}{y}\right\} - \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}}{dx} \\ &= \lim_{dx \to 0} \frac{\mathbf{Pr}\left\{u \le \frac{-x}{y} + \frac{-dx}{y}\right\} - \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}}{dx}, \end{aligned} \tag{3.26}$$

and:

$$\begin{aligned} \frac{\partial Z(x,y)}{\partial y} &= \lim_{dy \to 0} \frac{\mathbf{Pr}\left\{u \le \frac{-x}{y+dy}\right\} - \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}}{dy} \\ &\overset{(a)}{=} \lim_{dy \to 0} \frac{\mathbf{Pr}\left\{u \le \frac{-x}{y} + \frac{xdy}{y^2}\right\} - \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}}{dx}, \end{aligned} \tag{3.27}$$

where in (a) we used the fact that $\frac{xdy}{y(y+dy)} = \frac{xdy}{y^2}$ as dy goes to zero. Defining $dh = \frac{-x}{y}dy$, (3.28) can be written as:

$$\begin{aligned} \frac{\partial Z(x,y)}{\partial y} &= \lim_{dy \to 0} \frac{\mathbf{Pr}\left\{u \le \frac{-x}{y} + \frac{-dh}{y}\right\} - \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}}{\frac{-y}{x}dh} \\ &= \frac{-x}{y} \lim_{dh \to 0} \frac{\mathbf{Pr}\left\{u \le \frac{-x}{y} + \frac{-dh}{y}\right\} - \mathbf{Pr}\left\{u \le \frac{-x}{y}\right\}}{dh}. \end{aligned} \tag{3.28}$$

Thus, the lemma is proved since:

$$\frac{\partial Z(x,y)}{\partial y} = \frac{-x}{y}\frac{\partial Z(x,y)}{\partial x}. \tag{3.29}$$

□

Considering (3.19) and (3.14), $\forall l^b \leq l \leq l^e$ we have:

$$\frac{\partial G}{\partial I} = \frac{\partial Q(l,I,I',\zeta)}{\partial I} \times \zeta(l), \tag{3.30}$$

$$\frac{\partial G}{\partial I'} = \frac{\partial Q(l,I,I',\zeta)}{\partial I'} \times \zeta(l). \tag{3.31}$$

Furthermore, using Lemma 3.1, $\frac{\partial Q(l,I,I',\zeta)}{\partial I'}$ can be simplified as:

$$\begin{aligned}\frac{\partial Q(l,I,I',\zeta)}{\partial I'} &= -\frac{\partial F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)}{\partial I'} \\ &= -\frac{-\zeta(l)}{I'(l)} \times \frac{\partial F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)}{\partial \zeta} \\ &= -\frac{\zeta(l)}{I'(l)} \times \frac{\partial Q(l,I,I',\zeta)}{\partial \zeta} \\ &\overset{(a)}{=} \frac{Q(l,I,I',\zeta)}{I'(l)},\end{aligned} \tag{3.32}$$

where (a) follows from the identity imposed by the Euler condition of (3.22).

Given (3.32) and considering $l^b \leq l \leq l^e$, (3.30) and (3.31) can be written as:

$$\frac{\partial G(l,I,I',\zeta)}{\partial I} = \frac{\partial Q(l,I,I',\zeta)}{\partial I} \times \zeta(l), \tag{3.33}$$

$$\begin{aligned}\frac{\partial G(l,I,I',\zeta)}{\partial I'} &= \frac{\partial Q(l,I,I',\zeta)}{\partial I'} \times \zeta(l) \\ &= \frac{\zeta(l)}{I'(l)} \times Q(l,I,I',\zeta).\end{aligned} \tag{3.34}$$

Furthermore, we have:

$$
\begin{aligned}
\frac{d}{dl}\frac{\partial G(l,I,I',\zeta)}{\partial I'} &= \frac{d}{dl}\frac{\zeta(l)}{I'(l)} \times Q(l,I,I',\zeta) + \frac{\zeta(l)}{I'(l)} \times \frac{d}{dl}Q(l,I,I',\zeta) \\
&= \frac{\zeta'(l)}{I'(l)}Q(l,I,I',\zeta) + \frac{-\zeta(l)I''(l)}{I'^2(l)}Q(l,I,I',\zeta) \\
&\quad + \frac{\zeta(l)}{I'(l)}\left[\frac{\partial Q(l,I,I',\zeta)}{\partial I} \times I'(l) + \frac{\partial Q(l,I,I',\zeta)}{\partial I'} \times I''(l) \right. \\
&\qquad \left. + \frac{\partial Q(l,I,I',\zeta)}{\partial \zeta} \times \zeta'(l) + \frac{\partial Q(l,I,I',\zeta)}{\partial l}\right] \\
&\overset{(a)}{=} \frac{\zeta'(l)}{I'(l)}Q(l,I,I',\zeta) + \frac{-\zeta(l)I''(l)}{I'^2(l)}Q(l,I,I',\zeta) \\
&\quad + \frac{\partial Q(l,I,I',\zeta)}{\partial I}\zeta(l) + \frac{Q(l,I,I',\zeta)\zeta(l)I''(l)}{I'^2(l)} - \frac{Q(l,I,I',\zeta)\zeta'(l)}{I'(l)} \\
&= \frac{\partial Q(l,I,I',\zeta)}{\partial I}\zeta(l),
\end{aligned}
\tag{3.35}
$$

where (a) follows from (3.22), (3.32), and the fact that $\frac{\partial Q(l,I,I',\zeta)}{\partial l} = 0$.

Substituting (3.33) and (3.36) into the Euler equation in (3.19), $\frac{\partial G}{\partial I}$ and $\frac{d}{dl}(\frac{\partial G}{\partial I'})$ cancel out each other and therefore (3.19) is always satisfied. In other words, as long as $I(\cdot)$ and $\zeta(\cdot)$ satisfy (3.20), they satisfy (3.19) as well. Based on this observation, the optimal multilayer design rule can be summarized as the following theorem.

Theorem 3.3. *Consider a single-user MIMO system in which the source uses multilayer coding and the destination is able to perform successive decoding. In this network, the average received rate at the destination is maximized if $I(\cdot)$ and $\zeta(\cdot)$ are related as:*

$$
\frac{\partial F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)}{\partial \zeta} \times \zeta(l) = 1 - F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right), \forall l^b \le l \le l^e,
\tag{3.36}
$$

where l^b, l^e are selected such that $I(l) = P, \forall l \le l^b$ and $I(l) = 0, \forall l^e \le l$, respectively. Note that this theorem is valid only if $\zeta(l)$ satisfies the successive decodability requirement*.*

3.3 The Design Rule Sanity Check

Theorem 3.3 states a design rule for constructing the optimal multilayer code for a general single-user MIMO system with successive decoder. It is important to note that unlike the previously existing designs of multilayer codes, the proposed scheme is valid for arbitrary selections of $\zeta(l)$ (as long as it satisfies the successive decodability requirement).

To illustrate the application of Theorem 3.3, we study a SISO network (in which the channel gains are Rayleigh fading) and use (3.36) to determine the optimal multilayer code for two different examples.

We start by simplifying $F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)$ for a Rayleigh fading SISO setup. Note that in single-antenna scenarios, $\mathbf{HH}^\dagger$ is equal to the fading power, i.e., $\lambda = |h|^2$. Therefore, for all $l^b \le l \le l^e$, we have:

$$\begin{aligned} F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right) &= \mathbf{Pr}\left\{\frac{\lambda}{1+\alpha\lambda} \le \frac{\zeta(l)}{-I'(l)}\right\} \\ &= \mathbf{Pr}\left\{-I'(l)\lambda - I(l)\zeta(l)\lambda < \zeta(l)\right\} \\ &= \mathbf{Pr}\left\{\lambda < \frac{\zeta(l)}{-I'(l) - I(l)\zeta(l)}\right\} \\ &= \mathbf{Pr}\left\{|h|^2 < \frac{\zeta(l)}{-I'(l) - I(l)\zeta(l)}\right\} \\ &\overset{(a)}{=} 1 - e^{\frac{\zeta(l)}{I'(l)+I(l)\zeta(l)}}, \end{aligned} \tag{3.37}$$

where l^b, l^e are selected such that $I(l) = P, \forall l \le l^b$ and $I(l) = 0, \forall l^e \le l$, respectively. In (3.37), (a) follows from the fact that the power of a Rayleigh fading channel has an exponential distribution, i.e., $\mathbf{Pr}\left\{|h|^2 < a\right\} = 1 - e^{-a}$. Therefore, R_{avg} in (3.13) can be written as:

$$R_{avg} = \int_{l^b}^{l^e} \left[1 - F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)\right]\zeta(l)dl \tag{3.38}$$

$$= \int_{l^b}^{l^e} e^{\frac{\zeta(l)}{I'(l)+I(l)\zeta(l)}}\zeta(l)dl. \tag{3.39}$$

Using Theorem 3.3, R_{avg} would be maximized if $I(l)$ and $\zeta(l)$ satisfy (3.36), i.e.:

$$\frac{\partial\left[1-e^{\frac{\zeta(l)}{I'(l)+I(l)\zeta(l)}}\right]}{\partial\zeta}\times\zeta(l)=e^{\frac{\zeta(l)}{I'(l)+I(l)\zeta(l)}},\forall l^b\leq l\leq l^e. \tag{3.40}$$

Equation (3.40) can be further simplified to:

$$-I'(l)\zeta(l)=\left[I'(l)+I(l)\zeta(l)\right]^2,\forall l^b\leq l\leq l^e. \tag{3.41}$$

Meaning that the average received rate at the destination of a single-antenna Rayleigh fading network, i.e., (3.39), becomes maximized if its associated $I(l)$ and $\zeta(l)$ satisfy (3.41).

We now use this result in the following two cases.

- **<u>Case I:</u>**

  First assume $\zeta(l)=\frac{-lI'(l)dl}{1+lI(l)}$. Note that, the optimal multilayer code for the case of $\zeta(l)=\frac{-lI'(l)dl}{1+lI(l)}$ is originally computed in [21] using a deferent approach.

  Based on (3.41), the optimal power distribution function can be evaluated by determining $I_1(l)$ such that:

$$-I_1{}'(l)\times\frac{-lI_1'(l)dl}{1+lI_1(l)}=\left[I_1'(l)+I_1(l)\frac{-lI_1'(l)dl}{1+lI_1(l)}\right]^2,\quad l^b<l<l^e, \tag{3.42}$$

  where l^b, l^e are selected such that $I(l)=P,\forall l\leq l^b$ and $I(l)=0,\forall l^e\leq l$, respectively. Equation (3.42) can be further simplified to:

$$l(1+lI_1(l))=1\Rightarrow I_1(l)=\frac{1}{l^2}-\frac{1}{l}.\quad l^b<l<l^e, \tag{3.43}$$

  Substituting $\zeta(l)$ and $I_1(l)$ in (3.39), the maximum average received rate will be:

$$R^*_{avg}=\int_{l^b}^{l^e}\left[\frac{2}{l}-1\right]e^{-l}dl, \tag{3.44}$$

  where $l^b=\frac{\sqrt{1+4P}-1}{2P}$ and $l^e=1$. As the last step we should show that $\zeta(l)$ satisfies the *successive decodability requirement* which is apparent for the case of $\zeta(l)=\frac{-lI'(l)dl}{1+lI(l)}$ (Assuming $\ell_1<|h|^2$, it is obvious that for all $\ell_2:\ell_2<\ell_1$, we have $\ell_2<|h|^2$).

As expected equations (3.43) and (3.44) are identical to the results reported in [21]. This verifies the accuracy of the proposed scheme.

- **<u>Case II:</u>**

As the second example, assume $\zeta(l) = -lI'(l)$. Similar to *Case I* The optimal $I_2(l)$ for $s^b \leq l \leq s^e$ can be evaluated as:

$$I_2{}'(l) \times \; lI_2'(l) = [I_2'(l) + lI_2(l)I_2'(l)]^2 \Rightarrow lI_2{}'(l) = I_2{}'(l)^2[1 - lI_2(l)]^2, \qquad (3.45)$$

where s^b, s^e are selected such that $I(l) = P, \forall l \leq s^b$ and $I(l) = 0, \forall s^e \leq l$, respectively. Equation (3.45) can be further simplified to:

$$I_2(l) = \frac{1}{l} - \frac{1}{\sqrt{l}}, \quad s^b < l < s^e, \qquad (3.46)$$

where $s^b = \sqrt{\frac{\sqrt{1+4P}-1}{2P}}$ and $s^e = 1$. Having $I_2(l)$, (3.39) can be written as:

$$\begin{aligned} R^*_{avg} &= \int_{s^b}^{s^e} -l\left[1 - F_U^{I_2(l)}(l)\right] I_2'(l)dl \qquad (3.47) \\ &= \int_{s^b}^{s^e} -e^{-\sqrt{l}}\left[\frac{-1}{l} + \frac{1}{2\sqrt{l}}\right] dl \\ &\overset{(a)}{=} \int_{(s^b)^2}^{(s^e)^2} \left[\frac{2}{\hat{l}} - 1\right] e^{-\hat{l}} d\hat{l}, \end{aligned}$$

where (a) is deduced by changing the integral variable from l to $\hat{l} = \sqrt{l}$. To finalize the design, we should check the *successive decodability requirement*. To this end, assuming:

$$\ell_1 \leq \frac{\lambda}{1 + I_2(\ell_1)\lambda}, \qquad (3.48)$$

we should show that $\forall \ell_2, \ell_2 < \ell_1$ we have:

$$\ell_2 \leq \frac{\lambda}{1 + I_2(\ell_2)\lambda}. \qquad (3.49)$$

The proof is straight forward. Since $I_2(s) = \frac{1}{s} - \frac{1}{\sqrt{s}}$, (3.48) and (3.49) can be simplified to $\sqrt{\ell_1} \leq \lambda$ and $\sqrt{\ell_2} \leq \lambda$, respectively. Therefore, if we know $\sqrt{\ell_1} \leq \lambda$, it is adequate to show that for all ℓ_2 less than ℓ_1, $\sqrt{\ell_2} \leq \lambda$ is valid. Obviously, the last inequality is accurate since $\ell_2 < \ell_1 \Rightarrow \sqrt{\ell_2} < \sqrt{\ell_1}$; therefore, the *successive decodability requirement* holds.

Note that R^*_{avg} in (3.47) is similar to R^*_{avg} in (3.44) for *Case I*. In other words, although changing $\zeta(l)$ resulted in different optimal power distributions, i.e. $I_1(l)$ and $I_2(l)$, the values of R^*_{avg} are the same in both cases. This result also shows that the optimal design of multi-layer code is not unique for single-antenna networks.

3.4 Multilayer Code Design Algorithm for MIMO Systems

Considering a single antenna network, Section 3.3 presents two designs of multilayer codes for two assumptions of $\zeta(l)$. Generalizing to MIMO scenarios, this section shows how Theorem 3.3 can be applied to construct the optimal multilayer code for a MIMO system where $\zeta(l) = \zeta_2(l) = -lI'(l)$.

We use Theorem 3.3 and first simplify $F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)$ based on $\zeta(l) = \zeta_2(l)$, i.e.,

$$\begin{aligned} F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right) &= \mathbf{Pr}\left\{\sum_{i=1}^{K} \frac{\lambda_i}{M + I(l)\lambda_i} \leq \frac{\zeta(l)}{-I'(l)}\right\} \\ &\overset{(a)}{=} \mathbf{Pr}\left\{\sum_{i=1}^{K} \frac{\lambda_i}{M + I(l)\lambda_i} \leq l\right\} \\ &= F_{U^{I(l)}}(l), \end{aligned} \tag{3.50}$$

where $l^b \leq l \leq l^e$ and l^b, l^e are selected such that $I(l) = P, \forall l \leq l^b$ and $I(l) = 0, \forall l^e \leq l$, respectively. We also have:

$$\begin{aligned} \frac{\partial F_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right)}{\partial \zeta} &= \frac{1}{-I'(l)} f_{U^{I(l)}}\left(\frac{\zeta(l)}{-I'(l)}\right) \\ &= \frac{-f_{U^{I(l)}}(l)}{I'(l)}, \end{aligned} \tag{3.51}$$

where $f_{U^{\alpha}}(l)$ represents the *pdf* of the random variable U^{α}.

Substituting (3.50) and (3.51) into (3.36), the optimal $I(l)$ should be selected such that it satisfies:

$$l f_{U^{I(l)}}(l) = 1 - F_{U^{I(l)}}(l), \forall l^b \leq l \leq l^e, \tag{3.52}$$

where l^b, l^e are selected such that $I(l) = P, \forall l \leq l^b$ and $I(l) = 0, \forall l^e \leq l$, respectively. Note that $I(l)$ is valid only if we are able to show that the selected $\zeta(l)$ satisfies the *successive decodability requirement*.

MultiLayer Code Design Algorithm:

The following algorithm summarizes the steps of the algorithm that we have used for finding the optimal $I(l)$ which satisfies (3.52). Note that $I(l)$ is a continuous function; however, we have quantized it into Q levels and solved (3.52) for these discrete points. Note that, the result becomes more accurate as $Q \to \infty$.

1. Uniformly quantize $[0, P]$ to Q levels, i.e., $\mathbb{P} = \{\alpha_1, \cdots, \alpha_Q\}$, where $\alpha_1 = P$ and $\alpha_Q = 0$. Furthermore, set $q = 1$.

2. Set $I(l) = \alpha_q$.

3. Based on the fading characteristic of the channel (and consequently channel eigenvalues, i.e., λ_i's), find the *pdf* and *cdf* of $U^{\alpha_q} = \sum_{i=1}^{K} \frac{\lambda_i}{M+\alpha_q \lambda_i}$ [2].

4. Find l_q^* such that $1 - F_{U^{\alpha_q}}(l_q^*) = l_q^* f_{U^{\alpha_q}}(l_q^*)$.

5. If $q < Q$; set $q = q + 1$ and return to step 2.

6. Set $I^*(l) = P, \forall l < \ell_1^*$ and $I^*(l) = 0, \forall l > l_Q^*$.

7. Having the optimal $I(l)$, the rate of each code layer is determined by $\zeta(l) = \zeta_2(l)$. The transmitted signal ($\mathbf{x}$) would be the superposition of all these code layers.

8. Verify that the *successive decodability requirement* is satisfied for the selected $\zeta(l)$.

[2] As will be discussed in Section 3.5.1, a closed form solution for the *pdf* and *cdf* of U^{α_q} is not required and it is sufficient if we are able to find a numerical estimation of them.

3.5 Numerical Results and Discussion

To analyze the performance of the proposed transmission scheme, in this section, we investigate the performance of multilayer codes in three different network setups. For the purpose of comparison, we also compute the ergodic capacity (C_{erg}) and the outage rate (R_{outage}) as defined in [21]. A brief explanation of C_{erg} and R_{outage} is presented in the following.

Ergodic Capacity, C_{erg}**:** To compute the ergodic capacity, instead of quasi-static fading, we assume that we have a fast fading channel, i.e., the code block length is much longer than the dynamics of the channel. The optimal transmission scheme for such a network is to design a single-layer code with a rate equal to the channel ergodic capacity which can be computed as [51]:

$$C_{erg} = \mathbb{E}\left\{\log\det\left(1 + \frac{P\mathbf{H}\mathbf{H}^{\dagger}}{M}\right)\right\}. \quad (3.53)$$

The reason is that each transmitted codeword experiences a large number of channel realizations. Thus, the destination can decode this message if the rate of the coded data is less than the expectation of the channel capacities for different channel realizations. Note that since in our network model the channel gain remains constant during one transmission block, C_{erg} would be an upper bound for the maximum achievable average received rate at the destination.

Outage Rate, R_{outage}: Considering a quasi-static fading channel, in this strategy, the source transmits a single-layer code with rate R for all channel conditions. The transmitted data can be decoded if the channel capacity (corresponding to the channel state of that block) is more than R. Otherwise, the channel would be in outage and no information can be extracted. For each value of R, the average received rate at the destination would be: $R_o(R) = R \times \mathbf{Pr}\left\{R < \log\det\left(1 + \mathbf{H}\mathbf{H}^{\dagger}P/M\right)\right\}$ (refer to [21] for details). In this scheme, the design of the optimal code is to find a rate $R = R^*$ which maximizes $R_o(R)$. The term R_{outage} denotes this maximum achievable rate:

$$R_{outage} = \max_{R} R \times \mathbf{Pr}\left\{R < \log\det\left(1 + \mathbf{H}\mathbf{H}^{\dagger}P/M\right)\right\}. \quad (3.54)$$

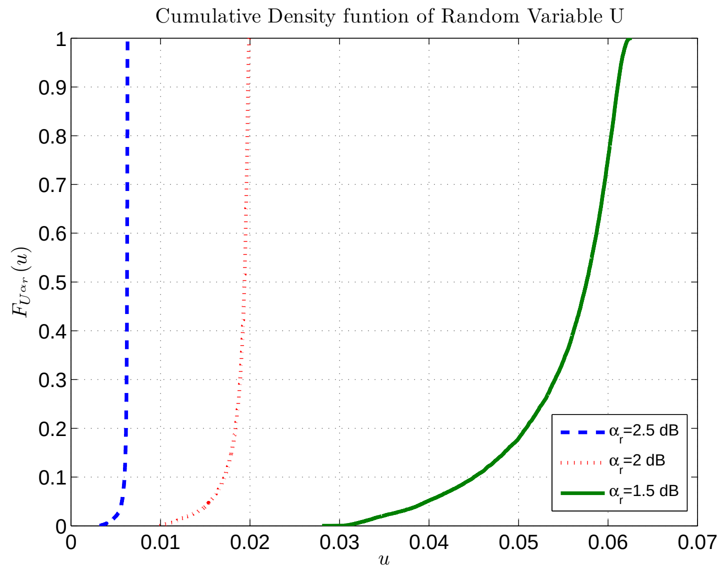

Figure 3.1: Cumulative density function of U where $\alpha_r = 2.5, 2,$ and 1.5 dB.

3.5.1 Setup One

Network Setup: Consider a network with one source and one destination, each equipped with two antennas. The link between the source and the destination is considered as a quasi-static Rayleigh fading, i.e., entries of the channel matrix $\mathbf{H}$ are zero-mean complex Gaussian, $\mathcal{CN}(0, 1)$ random variables, and the CSI is only available at the destination[3]. Note that this setup is the same as [21] and [22]. Therefore, we can compare the achievable rate of the proposed scheme with the achievable rates presented in [21] and [22].

Code Design: To obtain the optimal multilayer code design, we first construct set $\mathbb{P}$ as described in the first step of the design algorithm presented in section 3.4. In the next step, for each value of $\alpha_r \in \mathbb{P}$, we empirically estimate the *pdf* and *cdf* of U^{α_r}.[4] For illustration, the three curves in Fig. 3.1 depict $F_{U^{\alpha_r}}(l)$ where $\alpha_r = 2.5, 2,$ and 1.5 dB.

Given the *cdf* and *pdf* of $U^{\alpha_r}, \forall \alpha_r \in \mathbb{P}$, the optimal $I(l)$ can be determined by solving

[3]In this study we assume that the destination is only able to perform successive decoding.

[4]It is noted that having the joint distribution of λ_is' for the case of Rayleigh fading [51], $f_{U^{\alpha_r}}(l)$ and $F_{U^{\alpha_r}}(l)$ could be determined analytically. However, these functions will not have simple mathematical forms; hence, in step 4 of the design algorithm we still need to solve (3.52) numerically and we cannot use the closed form expression for $f_{U^{\alpha_r}}(l)$ and $F_{U^{\alpha_r}}(l)$. This is the reason that a numerical estimation of $f_{U^{\alpha_r}}(l)$ and $F_{U^{\alpha_r}}(l)$ is sufficient for finding the close-to-optimal $I(l)$.

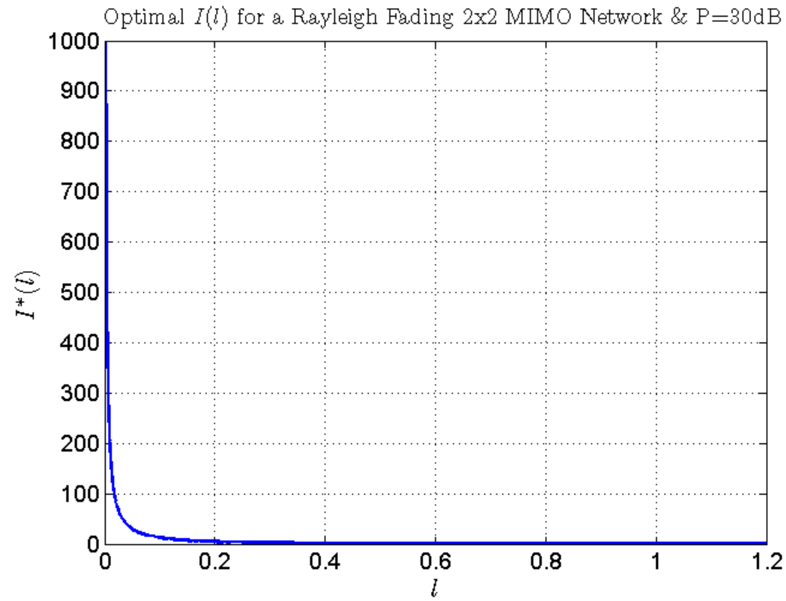

Figure 3.2: $I^*(l)$ for a Rayleigh fading 2x2 MIMO system where P=30dB.

(3.52) as described in step 4 of the proposed design algorithm. Figure 3.2 shows the resulting $I^*(l)$ for the understudy network where the source power is P = 30dB. Note that since we do not have the closed form expression for $I^*(l)$, the *successive decodability requirement* is verified numerically in this scenario.

Finally, we use Theorem 3.1 to find the corresponding average received rate at the destination. The solid line in Fig. 3.3 shows the achievable rate of the proposed scheme where P varies from 0dB to 70dB. For comparison, we also compute R_{outage} and C_{erg} for this network setup. The results are depicted in Fig. 3.3.

Discussion: As Fig. 3.3 shows, multilayer coding, namely R_{bs}, achieves higher data rates compared to the outage approach, i.e., R_{outage}. This is due to the possibility of partial data decoding where multilayer codes are used.

Another important observation is related to the efficiency of the multilayer codes in MIMO system with Rayleigh fading. As Fig. 3.3 suggests multilayer codes are not as beneficial as they appear in single-antenna cases (refer to Fig. 3 in [21] for the performance of multilayer coding in single-antenna setups). This effect is also reported by the authors of [21] and [22]. As noted in [21] and [22], the *hardening effect* is probably the main reason for this behavior. It means that the existence of multiple antennas and the fact

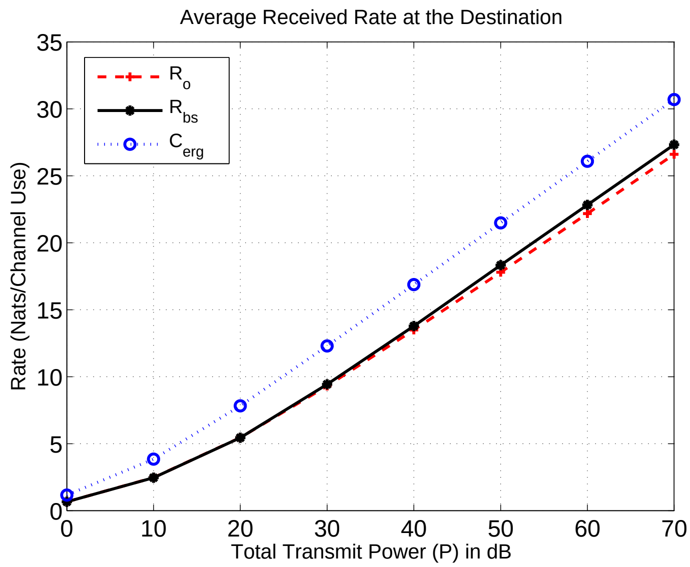

Figure 3.3: Average received rate at the destination for different transmission schemes (setup one).

that each of the entries of the propagation matrix is drawn independently from a complex normal distribution eliminates the large variation of the effective channel fading for different channel realizations. This characteristic reduces the efficiency of the multilayer coding and leads to a a small improvement as compared to the outage approach (single-layer coding).

3.5.2 Setup Two

Network Setup: The network setup is similar to *setup one*. The only difference is that instead of a pure Rayleigh fading, the source-destination link is modeled as a link which has both quasi-static Rayleigh fading and shadowing. More precisely, the source-destination propagation matrix $\hat{\mathbf{H}}$ can be written as [52]:

$$\hat{\mathbf{H}} = \sqrt{z}\mathbf{H}, \tag{3.55}$$

where $\mathbf{H}$ is a Rayleigh fading matrix similar to *setup one* and z is a positive random variable with log-normal distribution. To model practical environments, the mean and variance of the log-normal random variable are set to 0dB and 8dB, respectively [53].

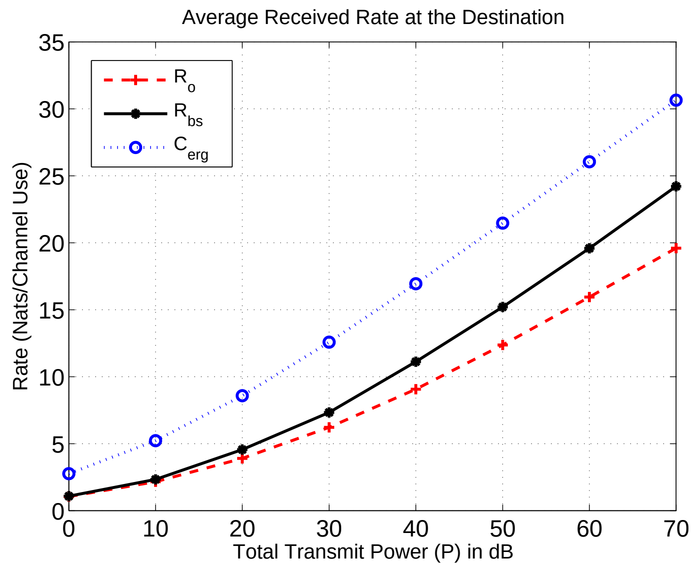

Figure 3.4: Average received rate at the destination for different transmission schemes (setup two).

Code Design: We follow the steps of *setup one* with the exception that we also consider the effect of the log-normal shadowing in evaluating the *cdf* and *pdf* of U^{α_r}, i.e., $F_{U^{\alpha_r}}(l)$ and $f_{U^{\alpha_r}}(l)$. The results are depicted in Fig. 3.4 where the achievable average received rate of the proposed scheme is shown by a solid line, namely R_{bs}. R_{outage} and C_{erg} are also plotted for comparison.

Discussion: Similar to *setup one*, the superiority of the proposed approach compared to the outage approach can be verified, Fig. 3.4. However, unlike *setup one*, multilayer coding considerably enhances the performance of the outage approach and the *hardening effect* is not significant as in *setup one*. The reason is that in *setup one* each channel state depends on the values of four independent random variables (corresponding to each transmit-receive antenna pairs). The independence between these values leads to the hardening effect, i.e., there would not be a very large deviation in the total effect of all four variables. On the other hand, in *setup two*, after fixing the Rayleigh fading matrix, i.e., $\mathbf{H}$, the whole matrix is multiplied by a shadowing factor, which is square root of a log-normal random variable, i.e., z. Thus, if z becomes small it affects all the transmitter-receiver antenna gains simultaneously and hence reduces the channel capacity considerably. The opposite effect (large increment in capacity) can be seen where the shadowing factor takes

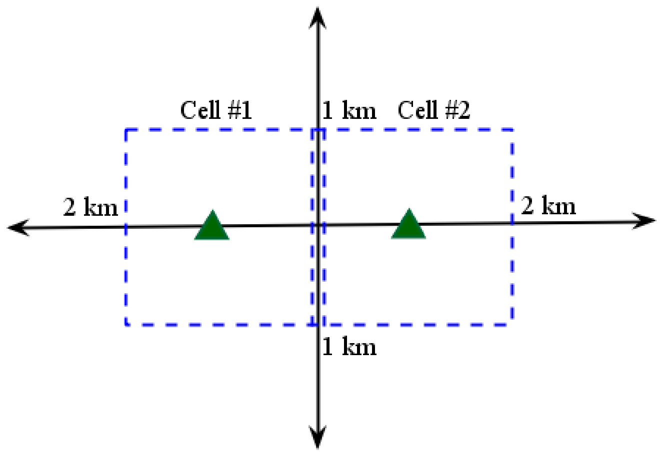

Figure 3.5: Network topology

a large value. This large deviation in the network capacity is the reason that the multilayer code is more beneficial in *setup two*.

This result points out that although multilayer coding is not very beneficial in some network setups (e.g., quasi-static Rayleigh fading), it might significantly improve the network performance in some other channel distributions (e.g. quasi-static Rayleigh fading with shadowing).

3.5.3 Setup Three

Network Setup: Consider a network of size 4km×2km which is partitioned into two cells.[5] The Base Station (BS) associated with each cell is located at the middle of each cell, Fig. 3.5.

The aim is to broadcast a TV channel (e.g., a popular TV show) to all users. It is assumed that users are uniformly distributed in the network. The channel between each user and each BS is modeled as Rayleigh fading with log-normal shadowing. The distance between the user and each BS is used to estimate the path loss effect. In this study, the path loss exponent is set to 2.6 [54], and all nodes are equipped with a single antenna.

To improve the performance of video broadcasting in a multiuser network, we should

[5]This structure can be considered as two adjacent cells of a cellular network which covers the whole region. To simplify the analysis, here, we only consider two of the cells in this network.

first compress the video signal using a Scalable Video Coding (SVC) standard, e.g., SVC H.264/AVC. As briefly discussed in Section 1.2, scalable video code streams consist of one base-layer and several enhancement layers. One way to use this feature in a wireless network is to transmit a SVC signal using a multilayer code. Based on multilayer transmission scheme, each user (according to its channel condition) can decode a number of code layers and partially regenerate the transmitted video.

Code Design: In the following, we discuss two transmission schemes for the network. The first scheme is based on the common implementation of cellular networks while in the second scheme we use the idea of BS cooperation for transmission of multilayer codes.

A. **Macro-Diversity Scheme**: In this scheme we follow the common design of a cellular network and partition the available bandwidth into two equal parts, e.g., $[0, W/2)$ and $[W/2, W)$. Each of these frequency ranges is assigned to one of the BSs. For instance, the right BS transmits in the range of $[0, W/2)$ and the left one transmits one $[W/2, W)$. Both BSs use multilayer coding to broadcast the SVC coded video signal to the users. Based on this scheme, a typical user (for instance in the right cell), receives two signals:
   a) The multilayer code transmitted by its own BS in the frequency range of $[0, W/2)$.
   b) The multilayer code transmitted by the BS of the adjacent cell in the frequency range of $[W/2, W)$.
   One simple decoding strategy is that the user disregards the signal received from the BS of the adjacent cell and decodes the data only based on the signal received in the frequency range of $[0, W/2)$. This decoding scheme can be improved, considering the fact that the other BS is also transmitting exactly the same message. Therefore, the user can receive the signal from both BSs and then perform Maximum Ratio Combining (MRC).

   To design the multilayer code for this case, distribution of the channel gain is selected corresponding to the distribution of the effective SNR (after performing MRC) at a typical network user.[6] Having this equivalent model, the optimal design of the multilayer code can be found based on the results of Section 2.1. Furthermore, the average received rate of a typical user can be evaluated using (2.13) and (2.14)

[6] The effective SNR depends on the distance of the user from each of the BSs, the shadowing factor, and the power of the Rayleigh fading for each link. It is assumed that users are uniformly distributed in each cell.

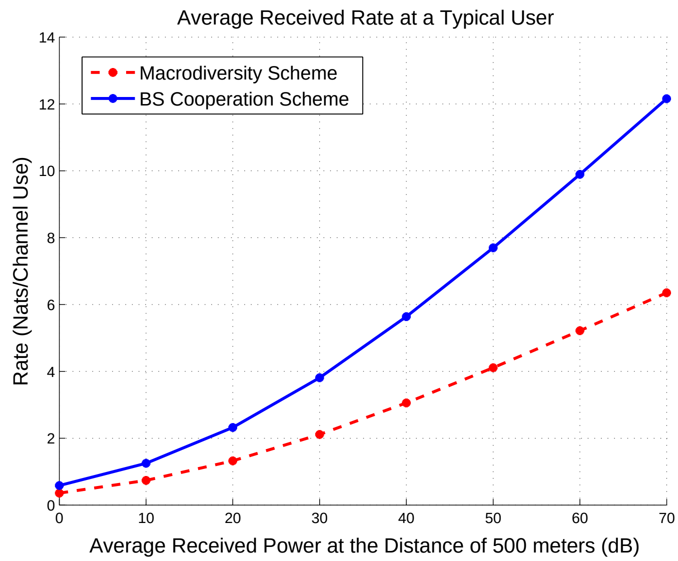

Figure 3.6: Average received rate at the destination for different transmission schemes (setup three).

where $F(l)$ and $f(l)$ are substituted by the *cdf* and *pdf* of the equivalent channel gain, respectively. The achievable rate of this scheme is shown with dashed line in Fig. (3.6). In this figure, the x-axis represents the average power that a typical user (located at the middle of a cell) receives from its associated BS.

B. **BS Cooperation Diversity Scheme**: In this example, both BSs intend to transmit the same message to the users, and users are able to receive the signal transmitted by both BSs (with different attenuations). In addition, since both BSs have access to the whole message, no extra communication is require if BSs want to cooperatively transmit the data.

   Based on this observation, in this scheme, instead of two separate BSs, the two BSs are considered as two antennas of a single source which are not co-located (antennas are distributed over the network). Therefore, the problem of broadcasting a TV signal over two separate cells of a network translates to broadcasting of a TV signal from a source which is equipped with multiple antennas. The main advantage of this method is that since both BSs are transmitting cooperatively, we do not need to partition the bandwidth between them and the whole bandwidth can be used for transmission of the multilayer code. We should also mention that, in this MIMO scheme, each BS transmits independent codewords which means that the antennas do not need cooperation and thus they do not need to be co-located .

The optimal multilayer code can be determined by following the design algorithm proposed in Section 3.4. The solid line in Fig. 3.6 depicts the performance of this transmission scheme.

Discussion: Comparing the results of the two schemes, we can see that the second transmission strategy considerably improves the network performance. Intuitively, by cooperative data transmission (over the two BSs), the transmission rate can be increased because of doubling the available bandwidth and in addition we get the multiple antenna gain. This result implies that in a cellular network in which the BSs are broadcast the same message to the users, instead of partitioning the available bandwidth between different cells, the bandwidth should be shared by all BSs which broadcast the message cooperatively toward the destination.

3.6 Summary

In this work, we investigated the optimal design of multilayer coding, in order to develop a scheme that can be used for transmission of layered coded data (such as SVC H.264/AVC streams). The objective is to maximize the average received rate at the destination and we studied a quasi-static fading MIMO system where the CSI is only known at the destination and the destination relies on successive decoding. For such network settings, we proposed a design rule for constructing a multilayer code where nodes are equipped with multiple antennas. We also suggested an algorithm to determine the optimal power distribution function of the multilayer code. Furthermore, the maximum average received rate at the destination was evaluated for different network setups and studied the gain that we can get by transmission of multilayer codes.

Chapter 4

Transmission over MIMO Random Access Networks

In this chapter[1], we study a symmetric multiple access network with K users and one Access Point (AP). It is assumed that users and the AP are equipped with M and N antennas, respectively. To access the network, each user independently decides whether to transmit in a time slot or not (no coordination between users). We begin the study by first considering a symmetric two-user random access network in Section 4.3 and determining the optimal value of network average Degrees of Freedom (DoF) for this two-user network. Generalizing the results to symmetric K-user random access networks, in Section 4.4, we propose an upper-bound of the network average DoF. This upper-bound is then analyzed for different network configurations and we determine the network classes in which the proposed upper-bound is tight. It is also shown that in most network settings the upper-bound can be achieved using simple single-stream data transmission. However, for some network configurations we need to apply multi-stream data transmission in conjunction with interference alignment to reach the upper-bound. Some illustrative examples are also presented in this chapter.

[1]The work in this chapter is partially reported in [55, 56].

4.1 Network Model

Consider a network with K synchronous users and one Access Point (AP) where the users and the AP are equipped with M and N antennas, respectively. The maximum transmit power of each node is denoted by P. $\mathbf{H}_j$ is an $M \times N$ matrix representing channel propagation matrix between user j and the AP. The entries of $\mathbf{H}_j$ are fixed and drawn independently from a complex Gaussian distribution with zero mean and unit variance, $\mathcal{CN}(0,1)$. The input-output relationship is:

$$\mathbf{y} = \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 + \cdots + \mathbf{H}_K\mathbf{x}_K + \mathbf{n}, \tag{4.1}$$

where $\mathbf{x}_j$ and $\mathbf{y}$ represent the transmitted signal from user j and the received signal at the AP, respectively.

At the beginning of each time slot, each user independently decides, with probability ρ, whether to transmit or not in that time slot. The set of active users (in each time slot) is known at the AP, but users do not know the status of other users.

Similar to the classic packet collision model in the networking literature, in [42], a transmission is considered successful if the set of transmission rates of active users corresponds to a point inside of the capacity region of that network state. The average network throughput is then computed based on the achievable rate in each network state and their corresponding probability of occurrence [42]. Unfortunately, computation of the network average throughput becomes complicated in MIMO cases. Therefore, we investigate the behavior of the system in high SNR regimes. In particular, we study the growth rate of the average throughput (average DoF of the network), i.e.:

$$\begin{aligned} \overline{D} &= \lim_{SNR\to\infty} \frac{\sum_{i=1}^{2^K-1} \Pr[\mathcal{S}_i] R[\mathcal{S}_i]}{\log SNR} \\ &= \sum_{i=1}^{2^K-1} \Pr[\mathcal{S}_i] \lim_{SNR\to\infty} \frac{R[\mathcal{S}_i]}{\log SNR} \\ &= \sum_{i=1}^{2^K-1} \Pr[\mathcal{S}_i] D[\mathcal{S}_i], \end{aligned} \tag{4.2}$$

where $\mathcal{S}_i \subseteq \{1, 2, \cdots, K\}, i \in \{1, 2, \cdots, 2^K - 1\}$ shows the set of users which are active in state i. Furthermore, $\Pr[\mathcal{S}_i]$ depicts the probability that the network is in state $\mathcal{S}_i$. Note that since we have K nodes in the network, there are $2^K - 1$ different states of active users. $R[\mathcal{S}_i]$ and $D[\mathcal{S}_i]$ represent the achievable throughput and the achievable DoF of the network operating in state $\mathcal{S}_i$, respectively.

4.2 Problem Formulation

4.2.1 Virtual AP Representation

To evaluate the DoF of different network states, we use the idea of the virtual AP proposed in [42]. In this method, the actual AP is substituted by a number of virtual APs each of which corresponds to one of the network states, namely $\mathrm{AP}[\mathcal{S}], \mathcal{S} \subseteq \{1, 2, \cdots, K\}$. Thus, there are $2^K - 1$ virtual APs and each user is connected to $2^{K-1} - 1$ virtual APs.

With this representation, we can consider the network with K users (becomes active with probability ρ), as a network in which all users are always active, but we have a number of APs each of them is connected to a subset of users. The problem of finding a scheme that maximizes *the network average DoF* is also translated to finding a strategy which maximizes *the weighted sum of the DoF of the received data at different APs*. The weight of each AP is the probability that the network operates in such a network condition. Considering ρ as the probability of activation of a node in a time slot, the weight of each $\mathrm{AP}[\mathcal{S}]$ is equal to:

$$\mathrm{P}_{|\mathcal{S}|} = \Pr[\mathcal{S}] = \rho^{|\mathcal{S}|}(1-\rho)^{K-|\mathcal{S}|}, \tag{4.3}$$

where $|\mathcal{S}|$ is the cardinality of the set of active users $\mathcal{S}$. As can be seen, this probability only depends on the number of active users and not the actual active users.

4.2.2 User Signaling

Based on the virtual AP representation, we now focus on a network with K users and $2^K - 1$ destinations, where each user broadcasts data to $2^{K-1} - 1$ destinations.

Let us define $\mathcal{W}_j$ (with DoF of d_j) as the data transmitted by user j and $\mathcal{W}_j[\mathcal{S}]$ (where $\mathcal{S} \subseteq \{1, 2, \cdots, K\}$) as the portion of user j's transmitted data which can be decoded at

AP[$\mathcal{S}$]. The DoF of $\mathcal{W}_j[\mathcal{S}]$ is denoted by $d_j[\mathcal{S}]$. The DoF of the total decodable data at state $\mathcal{S}$, i.e., $D[\mathcal{S}]$, can be written as:

$$D[\mathcal{S}] = \sum_{\forall j \in \mathcal{S}} d_j[\mathcal{S}]. \tag{4.4}$$

As discussed in Section 4.2.1, the network average DoF is equal to the weighted sum of the degrees of freedom for each of the virtual APs. The weight of each AP is the probability that network operates in that state, see (4.3). Equation (4.2), therefore, can be written as:

$$\begin{aligned} \overline{D} &= \sum_{\forall \mathcal{S}:\mathcal{S}\subseteq\{1,\cdots,K\}} \Pr[\mathcal{S}]D[\mathcal{S}] \\ &= \sum_{\forall \mathcal{S}:\mathcal{S}\subseteq\{1,\cdots,K\}} \mathrm{P}_{|\mathcal{S}|}D[\mathcal{S}]. \end{aligned} \tag{4.5}$$

Note that the goal is to determine a transmission scheme which maximizes the value of $\overline{D}$.

Knowing the virtual AP representation of the network, we are now ready to study the problem of determining the best transmission scheme which maximizes the network average DoF. In the following we need the following lemma. For the proof the reader is referred to [42].

Lemma 4.1. *Consider $\mathcal{S}, \mathcal{S}' \subseteq \{1, 2, \cdots, K\}$. If $\mathcal{S} \subseteq \mathcal{S}'$, $d_j[\mathcal{S}] \geq d_j[\mathcal{S}']$.*

4.3 Two-user MIMO Random Access Networks

Before going to the general K-user network, in this section we first study the behavior of a network with two users and one AP. Later, in Section 4.4, we extend these results to the general case of multiuser networks. The main result of the two-user random access network is presented in the following theorem.

Theorem 4.1. *Consider a two-user random access network where each user and the AP have M and N antennas, respectively. The node activation probability is denoted by ρ. The optimal network average DoF for this network is presented in Table 4.1.*

Table 4.1: Optimal network average DoF: Two-user random access network

	$\rho \geq \frac{1}{2}$	$\rho < \frac{1}{2}$
$N \leq M$	$N\rho$	$2N\rho(1-\rho)$
$M < N < 2M$	$N\rho$	$2M\rho(1-\rho) + 2\Delta\rho^2$
$2M \leq N$	$2M\rho$	

$\Delta = (N - M)$

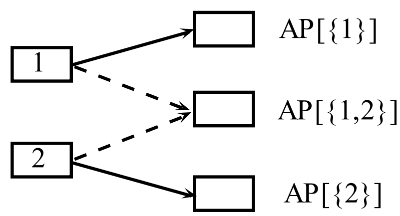

Figure 4.1: Virtual AP representation of two-user random access network.

4.3.1 Converse of Theorem 4.1

Figure 4.1 shows the virtual AP representation of a two-user MIMO random access network. The DoF of the decoded data at each virtual AP can be computed by (4.4):

$$\begin{aligned} D[\{1\}] &= d_1[\{1\}], \\ D[\{2\}] &= d_2[\{2\}], \\ D[\{1,2\}] &= d_1[\{1,2\}] + d_2[\{1,2\}]. \end{aligned} \tag{4.6}$$

One trivial upper-bound for the achievable DoF *in each state* is for the case that we assume full cooperation between the active users. Therefore, an upper-bound for the DoF in each state is the minimum of the number of AP's antenna and the total number of antennas available at the active users (in that state), i.e.:

$$\begin{aligned} D[\{1\}] &= d_1[\{1\}] \leq \min(M, N), \\ D[\{2\}] &= d_2[\{2\}] \leq \min(M, N), \\ D[\{1,2\}] &= d_1[\{1,2\}] + d_2[\{1,2\}] \leq \min(2M, N). \end{aligned} \tag{4.7}$$

Therefore, the basic upper-bound of the average network DoF will be:

$$\begin{aligned} \max_{\mathcal{A}} \quad & \overline{D}_u = \mathrm{P}_1 d_1[\{1\}] + \mathrm{P}_1 d_2[\{2\}] + \mathrm{P}_2 \Big(d_1[\{1,2\}] + d_2[\{1,2\}] \Big) \\ s.t. \quad & d_1[\{1\}] \leq \min(M, N) \\ & d_2[\{2\}] \leq \min(M, N) \\ & d_1[\{1,2\}] + d_2[\{1,2\}] \leq \min(2M, N) \\ & d_1[\{1,2\}] \leq d_1[\{1\}] \\ & d_2[\{1,2\}] \leq d_2[\{2\}]. \end{aligned} \tag{4.8}$$

where $\mathcal{A} = \{d_1[\{1\}], d_2[\{2\}], d_1[\{1,2\}], d_2[\{1,2\}]\}$ and the last two inequalities are due to Lemma 4.1.

This upper-bound is not tight in general. To improve the upper-bound we can make use of providing some side information for different network states through a genie. The same procedure is applied for improving the upper-bound in [42]. In the following, we intuitively describe this technique for a two-user network. For the detailed proof of this method refer to the proof of Theorem 1 in [42].

The virtual AP at state $\mathcal{S} = \{1,2\}$ is able to decode a part of data transmitted from each of the users denoted by $\mathcal{W}_1[\{1,2\}]$ and $\mathcal{W}_2[\{1,2\}]$. Now, assume a genie provides the undecoded part of user 2 transmitted signal, i.e., $\{\mathcal{W}_2 \backslash \mathcal{W}_2[\{1,2\}]\}$ for AP$[\{1,2\}]$. Thus, AP$[\{1,2\}]$ can decode $\mathcal{W}_2$ completely (it has already $\mathcal{W}_2[\{1,2\}]$ part). Having $\mathcal{W}_2$, the contribution of user 2 on the received signal can be canceled. Therefore, it can be assumed that the network is working in state $\mathcal{S} = \{1\}$. In this network state (user 1 is the only active user), there is no interferer at the AP and therefore the AP is able to decode all transmitted data from user 1, $\mathcal{W}_1$ (of course, rate of $\mathcal{W}_1$ should be less than or equal to the capacity of the point-to-point link between user 1 and the AP). So, if we provide $\{\mathcal{W}_2 \backslash \mathcal{W}_2[\{1,2\}]\}$ for AP$[\{1,2\}]$, it should be able to decode $\{\mathcal{W}_1[\{1\}], \mathcal{W}_2[\{1,2\}]\}$. Note that the total DoF of these streams still needs to follow the same DoF upper-bound. Therefore, we have:

$$d_1[\{1\}] + d_2[\{1,2\}] \leq \min(2M, N). \tag{4.9}$$

In a same way, if a genie reveals the message part of $\{\mathcal{W}_1 \backslash \mathcal{W}_1[\{1,2\}]\}$ to the AP$[\{1,2\}]$, we will have:

$$d_1[\{1,2\}] + d_2[\{2\}] \leq \min(2M, N). \tag{4.10}$$

Providing the side information for AP[{1}] and AP[{2}] will not generate new conditions. Therefore, the new upper-bound for the average network DoF will be:

$$\begin{aligned}
\max_{\mathcal{A}} \quad & \overline{D}_u = \mathrm{P}_1 d_1[\{1\}] + \mathrm{P}_1 d_2[\{2\}] + \mathrm{P}_2 \Big(d_1[\{1,2\}] + d_2[\{1,2\}] \Big) \\
s.t. \quad & d_1[\{1\}] \leq \min(M, N) \\
& d_2[\{2\}] \leq \min(M, N) \\
& d_1[\{1\}] + d_2[\{1,2\}] \leq \min(2M, N) \\
& d_2[\{2\}] + d_1[\{1,2\}] \leq \min(2M, N) \\
& d_1[\{1,2\}] \leq d_1[\{1\}] \\
& d_2[\{1,2\}] \leq d_2[\{2\}].
\end{aligned} \tag{4.11}$$

where $\mathcal{A} = \{d_1[\{1\}], d_2[\{2\}], d_1[\{1,2\}], d_2[\{1,2\}]\}$.

In order to simplify the upper-bound, we first prove the following lemma.

Lemma 4.2. *Consider a two-user MIMO random access network. The upper-bound optimization problem in* (4.11) *has at least one solution which is symmetrical for all users.*

Proof: The optimization problem in (4.11) has the form of a linear program problem, i.e., it has a linear objective function with respect to the DoF of messages and the constraints define a convex region.

Now, assume $\Upsilon = \{d_1[\{1\}] = a, d_1[\{1,2\}] = b, d_2[\{2\}] = c, d_2[\{1,2\}] = d\}$ is one optimal non-symmetrical solution for the problem. Due to the symmetry of the network, we can change the role of user 1 and user 2. Therefore, $\Upsilon' = \{d_1[\{1\}] = c, d_1[\{1,2\}] = d, d_2[\{2\}] = a, d_2[\{1,2\}] = b\}$ is also an optimal solution for the problem. Now consider a new scheme called $\Upsilon'' = \frac{\Upsilon + \Upsilon'}{2}$. Note that the DoF of each stream is now equal for both users (they have the same transmission scheme). Due to the convexity of the objective function, we know that the achievable DoF of Υ'' is at least equal to Υ, therefore, it is optimal. Moreover, the convexity of the search domain ensures that Υ'' is still a feasible point. Thus, there exists at least one optimal solution which is invariant under user permutation.

Applying Lemma 4.2 to the two-user random access network, we can conclude that

$$\begin{aligned} \alpha_1 &= d_1[\{1\}] \quad = d_2[\{2\}], \\ \alpha_2 &= d_1[\{1,2\}] = d_2[\{1,2\}]. \end{aligned} \tag{4.12}$$

Combining (4.3), (4.11), and (4.12), we get the following theorem.

Theorem 4.2. *Consider a two-user random access network where each user and the AP have M and N antennas, respectively. One upper-bound for the average network DoF is:*

$$\begin{aligned} \max_{\alpha_1,\alpha_2} \quad & \overline{D}_u = 2\rho(1-\rho)\alpha_1 + 2\rho^2\alpha_2 \\ s.t. \quad & \alpha_1 \leq \min(M,N) \\ & \alpha_1 + \alpha_2 \leq \min(2M,N) \\ & \alpha_2 \leq \alpha_1 \\ & 0 \leq \alpha_2, \end{aligned} \tag{4.13}$$

where ρ is the probability of the node activation.

To compute this upper-bound, we define $\beta_2 = \alpha_2$ and $\beta_1 = \alpha_1 - \alpha_2$. The above optimization problem can then be written as the following set of inequalities:

$$\begin{aligned} \overline{D}_u &\leq 2\rho(1-\rho)\beta_1 + 2\rho\beta_2 \\ \beta_1 + \beta_2 &\leq \min(M,N) \\ \beta_1 + 2\beta_2 &\leq \min(2M,N) \\ \beta_1 &\geq 0 \\ \beta_2 &\geq 0. \end{aligned} \tag{4.14}$$

We use the Fourier Motzkin elimination technique to compute $\overline{D}_u$ in terms of N, M, and ρ (see Section 4.6 for more details). The results are similar to that presented in Table 4.1. This proves that Table 4.1 is the upper-bound of the network achievable DoF, $\overline{D}$.

4.3.2 Achievability Scheme and Signal Alignment

Based on (4.1), the channel input-output relation is:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 + \mathbf{n} \\ &= \begin{bmatrix}\mathbf{H}_1 & \mathbf{H}_2\end{bmatrix}\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix} + \mathbf{n} \\ &= \mathbf{H}\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix} + \mathbf{n}, \end{aligned} \tag{4.15}$$

where $\mathbf{H}$ is a $2M \times N$ matrix representing the combined channel between the users and the AP. Since the channel gains are drawn independently from a random variable, $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}$ are full rank almost surely. In other words, the column spaces of $\mathbf{H}_1$ and $\mathbf{H}_2$ is $\min(M, N)$-dimensional subspace in the N-dimensional space of the received vectors. The rank of $\mathbf{H}$ is equal to $\min(2M, N)$, i.e., we can determine $\min(2M, N)$ independent $N \times 1$ vectors in the space of the received vectors.

The **core idea** of the achievability scheme is that user j partitions its data into $\min(M, N)$ sub-messages and transmits each of them along one optimally selected spatial direction. To find these optimal directions, in general, we first decide on $\min(2M, N)$ independent directions at the space of the received vectors. Depending on the values of M and N, we select some of these directions and *assign* them to the users. If one direction is *assigned* to a user, the other user will not transmit any stream along that direction. There are also some settings in which we let both users transmit along one direction, i.e., this direction is *shared* between both users.

The data sent along a direction which is *assigned* to user 1 or user 2 (if its DoF is less than one) is always decodable regardless of the state of the other user. However, the data sent along a *shared* direction can be decoded if i) only one of the users is active in that network state, or ii) both users are active but each transmits with half of its available DoF. Otherwise, the transmitted streams along that direction cannot be extracted. The optimal design of the code, therefore, relies on determining these spatial directions such that they minimize the dimension of the interference [57, 58]. In the following we present one optimal selection of the transmit directions. These directions will be used in the achievability scheme proposed in Theorem 4.3.

a) $N \leq M$: First select N *independent* vectors in the space of the received signal:

$$\mathbf{v} = \{\mathbf{v}_1, \mathbf{v}_2, , \cdots, \mathbf{v}_N\}.$$

Then, we define $\mathbf{v}_i^j$ as an $M \times 1$ vector that if user j transmits data along that spatial direction, its data will be received at the AP along the direction of $\mathbf{v}_i$, i.e.,

$$\mathbf{v}_i = \mathbf{H}_j \mathbf{v}_i^j, \tag{4.16}$$

where $i \in \{1, 2, \cdots, N\}$ and $j \in \{1, 2\}$. In fact, we select N directions at the AP and *share* all of them between the two users. Note that in (4.16) we have N equations and M variables where $M \geq N$. Thus, there is at least one solution for each $\mathbf{v}_i^j$.

b) $N \geq 2M$: Due to the independence between entries of the channel matrices, the ranks of the column spaces of $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}$ are equal to M, M and $2M$, respectively. This implies that the subspace of the intersection of the column spaces of $\mathbf{H}_1$ and $\mathbf{H}_2$ has rank zero with probability one. Therefore, it is possible to find $2M$ *independent* vectors such that the first M vectors, $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_M\}$, span the column space of $\mathbf{H}_1$ and the remaining vectors, $\{\mathbf{v}_{M+1} \cdots, \mathbf{v}_{2M}\}$, span the column space of $\mathbf{H}_2$. We, then, define $\mathbf{v}_i^j$ (where $i \in \{1, 2, \cdots, M\}$) as a direction along which user j transmits its ith data stream such that:

$$\begin{aligned} \mathbf{v}_i &= \mathbf{H}_1 \mathbf{v}_i^1, \\ \mathbf{v}_{M+i} &= \mathbf{H}_2 \mathbf{v}_i^2. \end{aligned} \tag{4.17}$$

Intuitively, we select $2M$ directions, M of them are *assigned* to user 1 and the other M directions are *assigned* to user 2.

c) $M \leq N \leq 2M$: In such settings, the ranks of $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}$ are M, M, and N, respectively. Furthermore, since $N < 2M$, the intersection of the column spaces of $\mathbf{H}_1$ and $\mathbf{H}_2$ is of rank of $2M - N$ with probability one. This means that it is possible to find N *independent* $N \times 1$ vectors, such that:

1) $2M - N$ directions reside in the intersection of the column spaces of $\mathbf{H}_1$ and $\mathbf{H}_2$:

$$p1 = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{2M-N}\}, \tag{4.18}$$

2) $N-M$ directions reside inside the column space of $\mathbf{H}_1$, but not $\mathbf{H}_2$:

$$p2 = \{\mathbf{v}_{2M-N+1}, \mathbf{v}_{2M-N+2}, \cdots, \mathbf{v}_M\}, \tag{4.19}$$

3) and $N-M$ directions reside inside the column space of $\mathbf{H}_2$, but not $\mathbf{H}_1$:

$$p3 = \{\mathbf{v}_{M+1}, \mathbf{v}_{M+2}, \cdots, \mathbf{v}_N\}. \tag{4.20}$$

$p1, p2$, and $p3$ are the proper receiving directions at the destination. We should find the corresponding optimal transmit directions. For directions in ρ_1, we use:

$$\begin{aligned} \mathbf{v}_i &= \mathbf{H}_1 \mathbf{v}_i^1, \\ \mathbf{v}_i &= \mathbf{H}_2 \mathbf{v}_i^2, \end{aligned} \tag{4.21}$$

where $i \in \{1, 2, \cdots, 2M-N\}$. It means that user 1 and user 2 transmit their data along $\mathbf{v}_i^1$ and $\mathbf{v}_i^2$ which are received aligned with each other at the destination. In other words, the directions in set ρ_1 are *shared* between both users.

Next, we select $N-M$ directions of $\mathbf{v}_{2M-N+i}^1$ and $\mathbf{v}_{2M-N+i}^2$ (where $i \in \{1, 2, \cdots, N-M\}$) such that the transmitted data from user 1 and user 2 are received along directions of ρ_2 and ρ_3, respectively, i.e.:

$$\begin{aligned} \mathbf{v}_{2M-N+i} &= \mathbf{H}_1 \mathbf{v}_{2M-N+i}^1, \\ \mathbf{v}_{M+i} &= \mathbf{H}_2 \mathbf{v}_{2M-N+i}^2. \end{aligned} \tag{4.22}$$

As will be shown, this technique enables us to decode some parts of the transmitted data if both users are active. The achievability scheme can be summarized as the following.

Theorem 4.3. *Consider a two-user random access network where each user and the AP have M and N antennas, respectively. The transmission scheme presented in Table 4.2 achieves the optimal value of the network average DoF. In this table, $\mathbf{x}_j$ represents the data transmitted by user j. The i th stream of $\mathbf{x}_j$ is denoted by s_i^j and the DoF of s_i^j is shown by $d_{s_i^j}$.*

Proof. We should verify that the proposed scheme achieves the upper-bound for all network setups. Here, we present the proof of each scenarios.

Table 4.2: Achievability scheme: Two-user random access network

	Signal Design	$\rho \geq \frac{1}{2}$	$\rho < \frac{1}{2}$
$N \leq M$	$\mathbf{x}_1 = \sum_{i=1}^{N} s_i^1 \mathbf{v}_i^1, \mathbf{x}_2 = \sum_{i=1}^{N} s_i^2 \mathbf{v}_i^2$	$i \in \{1, 2, \cdots, N\}:$ $d_{s_i^1} = d_{s_i^2} = \frac{1}{2}$	$i \in \{1, 2, \cdots, N\}:$ $d_{s_i^1} = d_{s_i^2} = 1$
$M < N < 2M$	$\mathbf{x}_1 = \sum_{i=1}^{2M-N} s_i^1 \mathbf{v}_i^1 + \sum_{i=2M-N+1}^{M} s_i^1 \mathbf{v}_i^1$ $\mathbf{x}_2 = \sum_{i=1}^{2M-N} s_i^2 \mathbf{v}_i^2 + \sum_{i=2M-N+1}^{M} s_i^2 \mathbf{v}_i^2$	$i \in \{1, 2, \cdots, 2M-N\}:$ $d_{s_i^1} = d_{s_i^2} = \frac{1}{2},$ $i \in \{2M-N+1, \cdots, M\}:$ $d_{s_i^1} = d_{s_i^2} = 1,$	$i \in \{1, 2, \cdots, 2M-N\}:$ $d_{s_i^1} = d_{s_i^2} = 1,$ $i \in \{2M-N+1, \cdots, M\}:$ $d_{s_i^1} = d_{s_i^2} = 1,$
$2M \leq N$	$\mathbf{x}_1 = \sum_{i=1}^{M} s_i^1 \mathbf{v}_i^1, \mathbf{x}_2 = \sum_{i=1}^{M} s_i^2 \mathbf{v}_i^2$	$d_{s_i^1} = d_{s_i^2} = 1, i \in \{1, 2, \cdots, M\}$	

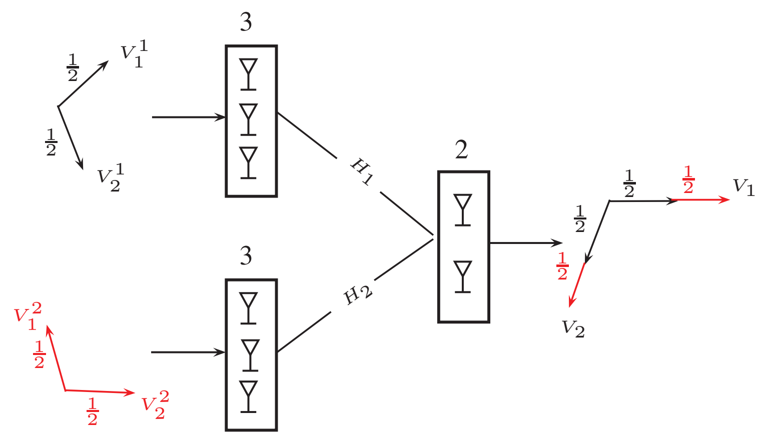

Figure 4.2: Signal structure: Two-user random access network where $M = 3$ and $N = 2$ ($N \leq M$), and $\rho \geq \frac{1}{2}$.

a) Consider a network in which $N \leq M$ and $\rho \geq \frac{1}{2}$. From Table 4.2, the achievability scheme is that each user partitions its data to N streams and sends them along $\mathbf{v}_i^j$ as defined in (4.16). Based on (4.16), $\mathbf{v}_i^1$ and $\mathbf{v}_i^2$ are received aligned with each other and along the direction of $\mathbf{v}_i$ (all N receiving directions are shared between both users). Noting that the DoF of each stream is $\frac{1}{2}$, we conclude that both streams can be decoded at high SNR regardless of the number of active users. Thus, from (4.5) the network achievable DoF will be:

$$\begin{aligned} \overline{D}_l &= \rho(1-\rho)\left[\sum_{i=1}^{N} d_{s_i^1} + \sum_{i=1}^{N} d_{s_i^2}\right] + \rho^2 \sum_{i=1}^{N}\left[d_{s_i^1} + d_{s_i^2}\right] \\ &= 2\rho(1-\rho)\frac{N}{2} + \rho^2(2 \times \frac{N}{2}) = N\rho. \end{aligned} \tag{4.23}$$

We can verify that this achievable network average DoF is the same as the upper-bound presented in Table 4.1. For illustration, Fig. 4.2 shows an example of this scheme where $M = 3$ and $N = 2$.

b) Now, consider a network in which $N \leq M$ and $\rho < \frac{1}{2}$. The achievability scheme is similar to part **a**, except that in this case the DoF of each stream is 1. With this selection, and since all direction are shared between both users we can not decode any data when both users are active. However, we achieve the DoF of N when only one user is active.

Therefore, similar to (4.23) we have:

$$\overline{D}_l = \rho(1-\rho)\left[\sum_{i=1}^{N} d_{s_i^1} + \sum_{i=1}^{N} d_{s_i^2}\right] + \rho^2 \sum_{i=1}^{N} \times 0 \qquad (4.24)$$
$$= 2\rho(1-\rho)N.$$

This result is the same as what presented in Table 4.1.

c) The next scenario is for a case in which $2M \leq N$ (either $\rho \geq \frac{1}{2}$ or $\rho < \frac{1}{2}$). In these situations, user 1 and user 2 partition their data into M streams and transmit them along the direction of $\mathbf{v}_i^1$ and $\mathbf{v}_i^2$ (where $i \in \{1, 2, \cdots, M\}$) as selected in (4.17), respectively. Based on (4.17), $\mathbf{v}_i^1$ and $\mathbf{v}_i^2$ will be separately received along $\mathbf{v}_1, \cdots, \mathbf{v}_M$ and $\mathbf{v}_{M+1}, \cdots, \mathbf{v}_{2M}$, respectively (they do not interfere with each other). This property, in addition to the fact that the DoF of each stream is 1 ensure us that the data can be decoded (in high SNR) regardless of the number of active users. The network average DoF is thus:

$$\overline{D}_l = \rho(1-\rho)\left[\sum_{i=1}^{M} d_{s_i^1} + \sum_{i=1}^{M} d_{s_i^2}\right] + \rho^2 \sum_{i=1}^{M} \left[d_{s_i^1} + d_{s_i^2}\right] \qquad (4.25)$$
$$= 2\rho(1-\rho)M + \rho^2(2M) = 2M\rho,$$

which is the same as the upper-bound introduced in Table 4.1. Figure 4.3 demonstrates an example of $M = 2$ and $N = 5$.

d) In this case we study is a setup in which $M < N < 2M$ and $\rho < \frac{1}{2}$. The idea is to align the transmitted data such that if both users become active, only some parts of the data interfere with each other and we can still decode some portion of the transmitted messages. To this end, we select N receiving directions as presented in $p1$, $p2$ and $p3$, (4.18)-(4.20).

The signal transmitted by user 1 and user 2, each consists of M streams where $2M - N$ of them are transmitted along $\mathbf{v}_i^1$ and $\mathbf{v}_i^2$ (as defined in (4.21)), respectively. $\mathbf{v}_i^1$ and $\mathbf{v}_i^2$ both get to the AP aligned with each other and along the direction of $\mathbf{v}_i$ (from set $\rho1$). The remaining $N - M$ streams of user 1 and user 2 are transmitted along the directions defined in (4.22). These streams are get to the AP along their corresponding directions from set $\rho2$ and $\rho3$ (without interference). Figure 4.4 shows an example in which $M = 3$ and $N = 5$.

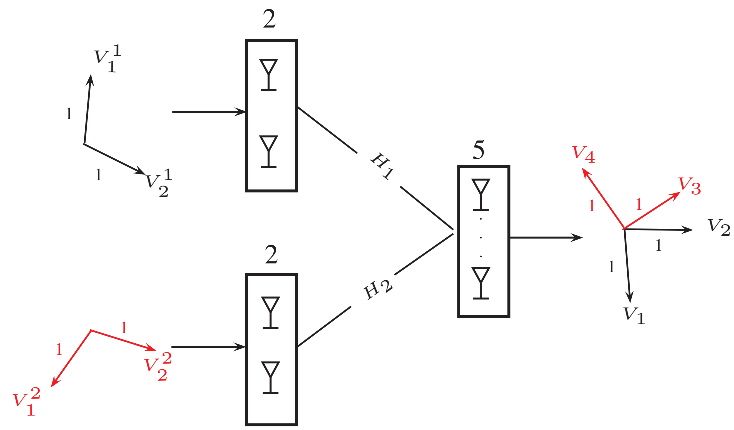

Figure 4.3: Signal structure: Two-user random access network where $M = 2$ and $N = 5$ ($2M \leq N$).

Note that the DoF of all streams are one; therefore, the data sent along a direction can be decoded if only one stream is transmitted along that direction. Thus, if only user 1 (or user 2) transmits, we can decode all M streams of sets $p1$ and $p2$ (or $p3$). If both users transmit, $2(N-M)$ messages transmitted along $p2$ and $p3$ can be successfully decoded (no interference along these directions). However, the streams received along the directions of set $p1$ will be lost since these directions are occupied with two transmitters, each of them sends with the DoF of 1. The achievable network average DoF, therefore, can be evaluated as:

$$\begin{aligned} \overline{D}_l &= \rho(1-\rho)\left[\sum_{i=1}^{M} d_{s_i^1} + \sum_{i=1}^{M} d_{s_i^2}\right] + \rho^2\left[\sum_{i=2M-N+1}^{M}\left[d_{s_i^1} + d_{s_i^2}\right]\right] \\ &= \rho(1-\rho)\left[M+M\right] + \rho^2\left[N-M+N-M\right] \\ &= 2\rho(1-\rho)M + 2(N-M)\rho^2 \\ &= 2\rho(1-\rho)M + 2\rho^2\Delta, \end{aligned} \tag{4.26}$$

where $\Delta = N - M$. This is the same as the upper-bound presented in Table 4.1.

e) The last case is when $M < N < 2M$ and $\rho \geq \frac{1}{2}$. The transmission scheme of this case is similar to part **d** where the transmit DoF over the shared directions is $\frac{1}{2}$ (instead of 1 in the previous case). By this assumption, we are able to decode the messages sent along the shared directions even when both users are active; however, the achievable DoF

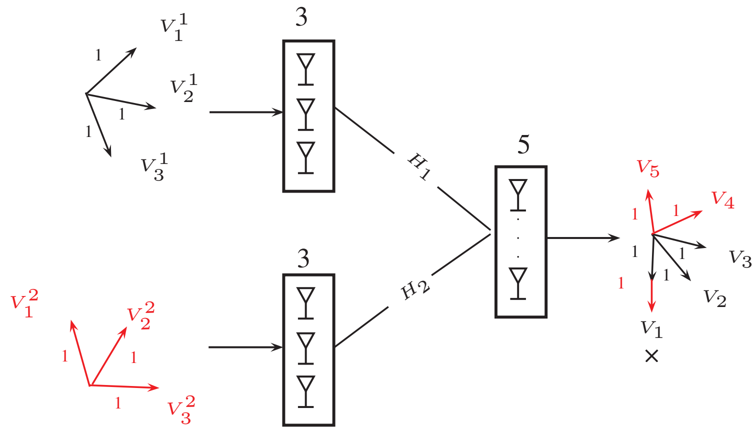

Figure 4.4: Signal structure: Two-user random access network where $M = 3$ and $N = 5$ $(M < N < 2M)$, and $\rho < \frac{1}{2}$.

is reduced when only one user is active.

$$\overline{D}_l = \rho(1-\rho)\left[\sum_{i=1}^{M} d_{s_i^1} + \sum_{i=1}^{M} d_{s_i^2}\right] + \rho^2\left[\sum_{i=1}^{2M-N}\left[d_{s_i^1} + d_{s_i^2}\right] + \sum_{i=2M-N+1}^{M}\left[d_{s_i^1} + d_{s_i^2}\right]\right] \tag{4.27}$$
$$= 2\rho(1-\rho)\left[N - M + \frac{2M-N}{2}\right] + \rho^2\left[(2M-N)\times(\frac{1}{2}+\frac{1}{2}) + (N-M)\times(1+1)\right]$$
$$= 2N\rho$$

This result is also the same as the upper-bound presented in Table 4.1 and Theorem 4.3 is proved. □

Figure 4.5 depicts the optimal network average DoF, for different values of the node activation probability. The results are plotted for different cases, where $M = 3, 4$ and $N = 5, 6$.

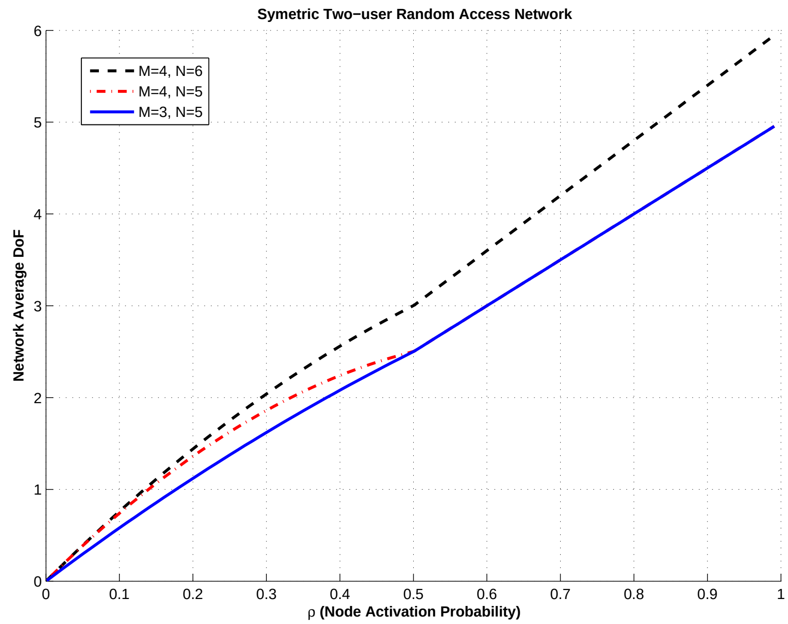

Figure 4.5: Network average DoF: Two-user MIMO random access network.

4.4 K-user MIMO Random Access Networks

In this section we extend the study to multiuser networks. Similar to the two-user network, we will use the virtual AP representation of the network, i.e.:

a) There are $2^K - 1$ virtual APs and each user sends data to $2^{K-1} - 1$ virtual APs.

b) Each user transmits a message, $\mathcal{W}_j$, with the DoF of d_j.

c) Assuming that the network is in state $\mathcal{S}$, the decodable portion of user j transmitted data is denoted by $\mathcal{W}_j[\mathcal{S}]$ with the DoF of $d_j[\mathcal{S}]$. The total DoF of the decoded signal at AP$[\mathcal{S}]$ is therefore can be evaluated by:

$$D[\mathcal{S}] = \sum_{\forall j \in \mathcal{S}} d_j[\mathcal{S}]. \tag{4.28}$$

Note that based on Lemma 4.1 all $d_j[\mathcal{S}]$ have the property that, if $\mathcal{S} \subseteq \mathcal{S}'$, then:

$$d_j[\mathcal{S}'] \leq d_j[\mathcal{S}]. \tag{4.29}$$

One upper-bound for the achievable DoF at each network state is the minimum of the number of AP's antennas and the total number of antennas available at all active users, i.e., $\forall \mathcal{S} \subseteq \{1, 2, \cdots, K\}$:

$$D[\mathcal{S}] = \sum_{\forall j \in \mathcal{S}} d_j[\mathcal{S}] \leq \min(|\mathcal{S}|M, N). \tag{4.30}$$

d) The goal is to determine a transmission scheme which maximizes the network average DoF which can be evaluated as:

$$\overline{D} = \sum_{\forall \mathcal{S}: \mathcal{S} \subseteq \{1, \cdots, K\}} \mathrm{P}_{|\mathcal{S}|} D[\mathcal{S}]. \tag{4.31}$$

4.4.1 Main Results

In the following we present the main results of this section. The detailed discussion will follow up later.

Theorem 4.4. *Consider a symmetric K-user MIMO random access network in which each user and the AP are equipped with M and N antennas, respectively. The node activation probability is denoted by ρ. One upper-bound for the network average DoF is:*

$$\begin{aligned} \max_{\mathcal{A}} \quad & \overline{D}_u = \sum_{j=1}^{K} \sum_{k=1}^{j} k \Gamma_k \beta_j \\ s.t. \quad & a1. \quad \sum_{k=1}^{K} \beta_k \leq \min(M, N) \\ & a2. \quad \sum_{k=1}^{K} k\beta_k \leq \min(KM, N) \\ & b. \quad \beta_k \geq 0; \quad k \in \{1, 2, \cdots, K\}, \end{aligned} \tag{4.32}$$

where $\mathcal{A} = \{\beta_j, \forall j \in \{1, 2, \cdots, K\}\}$ and Γ_k is the probability that k users are active in the network, i.e.,

$$\Gamma_k = \binom{K}{k} \rho^k (1-\rho)^{K-k}. \tag{4.33}$$

The optimization problem in Theorem 4.4 is indeed a Linear Program (LP) problem with K non-negative variables and $K+2$ constraints. The **important property of** (4.32) is that except the positivity constraints, there are only two other inequalities among the boundary conditions. Therefore, (4.32) has at least one optimal solution in which at most two variables take non-zero values. Note that, the solution of (4.32) depends on the values of K, M, N, and ρ. To determine these optimal solutions (in different settings), we have used the simplex method. The following corollary summarizes the results.

Corollary 4.1. *The optimization problem of* (4.32) *has at least one solution in which all* β_j *except at most two of them are equal to zero. The exact solution of different network settings (different combinations of* $\rho, K, M,$ *and* N*) is presented in Table 4.3.*

In Table 4.3, ρ_j, $j \in \{1, 2, \cdots, K\}$ *are defined as the solution of:*

$$B_{K-j,j}(1-\rho_j) - j \begin{pmatrix} K-1 \\ j \end{pmatrix} \rho_j^j (1-\rho_j)^{K-1-j} = 0, \tag{4.34}$$

where $B_{a,b}(x)$ *is the incomplete beta function defined as:*

$$B_{a,b}(x) = \frac{\int_0^x u^{a-1}(1-u)^{b-1}du}{\int_0^1 u^{a-1}(1-u)^{b-1}du}. \tag{4.35}$$

Note that the value of ℓ *in Table 4.3 is a function of the node activation probability,* ρ.

Furthermore, the definition of ρ_j in (4.34) can be used to prove the following lemma.

Lemma 4.3. $\arg\max_j \frac{\sum_{k=1}^{j} k\Gamma_k}{j}$ *is equal to* ℓ *if*

$$\rho_{\ell-1} < \rho \leq \rho_\ell,$$

where ρ_j *is the solution of equation* (4.34)

Table 4.3: Network average DoF upper-bound: K-user random access network

$N \leq M$ $\ell : \rho_{\ell-1} < \rho \leq \rho_\ell$	$N \geq KM$	$M < N < KM$ $j_1 : \rho_{j_1-1} < \rho \leq \rho_{j_1}, j_2 = \lfloor \frac{N}{M} \rfloor + 1$	
		$\rho_{j_2-1} \leq \rho : \ell = j_1$	$\rho < \rho_{j_2-1} : \ell = j_2$
$\forall j \neq \ell,\ \beta_j = 0,$ $\beta_\ell = \frac{N}{\ell},$	$\forall j \neq K,\ \beta_j = 0,$ $\beta_K = M,$	$\forall j \neq \ell,\ \beta_j = 0,$ $\beta_\ell = \frac{N}{\ell},$	$\forall j \notin \{\ell-1, \ell\},\ \beta_j = 0,$ $\beta_{\ell-1} = \ell M - N, \beta_\ell = N - (\ell-1)M,$
$\overline{D}_u = \left(\sum_{k=1}^{\ell} k\Gamma_k\right) \frac{N}{\ell}$	$\overline{D}_u = KM\rho$	$\overline{D}_u = \left(\sum_{k=1}^{\ell} k\Gamma_k\right) \frac{N}{\ell}$	$\overline{D}_u = \left(\sum_{k=1}^{\ell-1} k\Gamma_k\right)(\ell M - N) + \left(\sum_{k=1}^{\ell} k\Gamma_k\right)(N - (\ell-1)M)$

Note: ℓ is a function of of the node activation probability, ρ.

The optimality of the proposed upper-bound (existence of an achievability scheme) depends on the values of M, N, and K. The next theorem summarizes the achievability results.

Theorem 4.5. *Consider a symmetric K-user MIMO random access network in which each user and the AP are equipped with M and N antennas, respectively. The network average DoF upper-bound presented in Theorem 4.4 is achievable if any of of the following conditions holds:*

I) $N \leq M$,

II) $KM \leq N$,

III) $M < N < KM$ *and* $\rho \geq \rho_{\lfloor \frac{N}{M} \rfloor}$ *where* ρ_j *is defined in* (4.34).

IV) The upper-bound is also achievable for some special network setups (discussed in section 4.4.3) where $M < N < KM$ *and* $\rho < \rho_{\lfloor \frac{N}{M} \rfloor}$.

Note that the optimality of the upper-bound is not proved for the general case of $M < N < KM$ *and* $\rho < \rho_{\lfloor \frac{N}{M} \rfloor}$.

The idea behind the code design is similar to what discussed for the two-user random access network. Each user partitions its data to sub-messages and transmits them over appropriate directions. These directions are determined such that if more than one user are transmitting, the AP can still decode some of the sub-messages from each user.

4.4.2 Proof of Theorem 4.4 (The Genie Aided Upper-bound)

The immediate upper-bound for the average network DoF can be derived by finding the maximum value of (4.31) inside the region defined by (4.29) and (4.30), i.e.,

$$\begin{aligned} \max_{\mathcal{A}} \quad & \overline{D}_u = \sum_{\forall \mathcal{S}: \mathcal{S} \subseteq \{1, \cdots, K\}} \mathrm{P}_{|\mathcal{S}|} D[\mathcal{S}] \\ s.t. \quad a. \quad & D[\mathcal{S}] = \sum_{\forall j \in \mathcal{S}} d_j[\mathcal{S}] \leq \min(|\mathcal{S}|M, N) \\ b. \quad & \mathcal{S} \subseteq \mathcal{S}' : d_j[\mathcal{S}'] \leq d_j[\mathcal{S}] \\ & \forall \mathcal{S}, \mathcal{S}' \subseteq \{1, 2, \cdots, K\}. \end{aligned} \tag{4.36}$$

where $\mathcal{A} = \{d_j[\mathcal{S}], \forall j \in \{1, \cdots, K\}, \forall \mathcal{S} \subseteq \{1, 2, \cdots, K\}\}$.

In order to get a tighter upper-bound, similar to the two-user scenario, we use the idea of providing some side information to the virtual APs [42].

Based on the virtual AP representation we know that the received signal at AP[$\mathcal{S}$] is:

$$\mathcal{G}[\mathcal{S}] = \bigcup_{\forall j \in \mathcal{S}} \mathcal{W}_j. \tag{4.37}$$

First we select one of the users in $\mathcal{S}$, say user $\hat{1}$. The part of user $\hat{1}$ message decoded at AP[$\mathcal{S}$] is denoted by $\mathcal{W}_{\hat{1}}[\mathcal{S}]$ and its DoF is equal to $d_{\hat{1}}[\mathcal{S}]$.

Let $\mathcal{V}_{\hat{1}}[\mathcal{S}]$ be the portion of $\hat{1}$ transmitted signal which is not decoded by AP[$\mathcal{S}$], i.e.,

$$\mathcal{V}_{\hat{1}}[\mathcal{S}] = \{\mathcal{W}_{\hat{1}} \backslash \mathcal{W}_{\hat{1}}[\mathcal{S}]\}. \tag{4.38}$$

Now assume a genie provides $\mathcal{V}_{\hat{1}}[\mathcal{S}]$ for AP[$\mathcal{S}$]. Therefore, AP[$\mathcal{S}$] has access to all data streams transmitted by user $\hat{1}$. Thus, it can remove the contribution of user $\hat{1}$ signal on the received signal, i.e.:

$$\mathcal{G}[\mathcal{S}_1] = \bigcup_{\forall j \in \mathcal{S}_1} \mathcal{W}_j, \tag{4.39}$$

where $\mathcal{S}_1 = \{\mathcal{S} \backslash \hat{1}\}$. $\mathcal{G}[\mathcal{S}_1]$ represents the signal that the AP would receive if it works in state $\mathcal{S}_1$ (equivalently the signal that received at AP[$\mathcal{S}_1$]). Furthermore, from the virtual AP representation we know that AP[$\mathcal{S}_1$] can decode $\mathcal{W}_j[\mathcal{S}_1]$ for all the remaining users, i.e., $j \in \{\mathcal{S} \backslash \hat{1}\}$. Note that, since $\mathcal{S}_1 \subset \mathcal{S}$, from Lemma 4.1 we have $d_j[\mathcal{S}_1] \geq d_j[\mathcal{S}]$, for $j \in \mathcal{S}_1$ (the DoF of decodable signal is increased when the effect of user $\hat{1}$ is eliminated). Thus:

$$\sum_{\forall j \in \mathcal{S}_1} d_j[\mathcal{S}_1] \geq \sum_{\forall j \in \mathcal{S}_1} d_j[\mathcal{S}]. \tag{4.40}$$

Therefore, to confine the search domain, instead of each constraint of:

$$d_{\hat{1}}[\mathcal{S}] + \sum_{\forall j \in \mathcal{S}_1} d_j[\mathcal{S}] \leq \min(|\mathcal{S}|M, N), \tag{4.41}$$

we use:

$$d_{\hat{1}}[\mathcal{S}] + \sum_{\forall j \in \mathcal{S}_1} d_j[\mathcal{S}_1] \leq \min(|\mathcal{S}|M, N). \tag{4.42}$$

We now select another active user, user $\hat{2}$. AP$[\mathcal{S}_1]$ can decode a part of user $\hat{2}$ transmitted data denoted by $\mathcal{W}_{\hat{2}}[\mathcal{S}_1]$ with the DoF of $d_{\hat{2}}[\mathcal{S}_1]$. Similar to the previous case we define:

$$\mathcal{V}_{\hat{2}}[\mathcal{S}_1] = \{\mathcal{W}_{\hat{2}} \backslash \mathcal{W}_{\hat{2}}[\mathcal{S}_1]\}, \tag{4.43}$$

as the uncoded part of user $\hat{2}$ transmitted data. Providing $\mathcal{V}_{\hat{2}}[\mathcal{S}_1]$ for AP$[\mathcal{S}_1]$, the effect of user $\hat{2}$ can be eliminated as well. Thus, we have:

$$\mathcal{G}[\mathcal{S}_2] = \bigcup_{\forall j \in \mathcal{S}_2} \mathcal{W}_j, \tag{4.44}$$

where $\mathcal{S}_2 = \{\mathcal{S} \backslash \{\hat{1}, \hat{2}\}\}$. Therefore, the modified upper-bound is:

$$d_{\hat{1}}[\mathcal{S}] + d_{\hat{2}}[\mathcal{S}_1] + \sum_{\forall j \in \mathcal{S}_2} d_j[\mathcal{S}_2] \leq \min(|\mathcal{S}|M, N). \tag{4.45}$$

Continuing this procedure for the remaining active nodes of $\mathcal{S}$, it can be proved that the basic DoF constraint of:

$$\sum_{\forall j \in \mathcal{S}} d_j[\mathcal{S}] \leq \min(|\mathcal{S}|M, N), \tag{4.46}$$

can be substituted by:

$$d_{\hat{1}}[\mathcal{S}] + d_{\hat{2}}[\mathcal{S}_1] + \cdots + d_{\hat{q}}[\mathcal{S}_{q-1}] \leq \min(qM, N), \tag{4.47}$$

where $q = |\mathcal{S}|$.

Note that there are $q!$ ways to select users $\hat{1}, \hat{2}, \cdots, \hat{q}$ (different sorting order of set $\mathcal{S}$). Therefore, the basic condition of (4.46) for state $\mathcal{S}$ generates $q!$ conditions of the form (4.47).

Theorem 4.6. *Consider a K-user MIMO random access network. One upper-bound of the network average DoF is equal to:*

$$\max_{\mathcal{A}} \quad \overline{D} = \sum_{\forall \mathcal{S}:\mathcal{S}\subset\{1,\cdots,K\}} P_{|\mathcal{S}|} \sum_{\forall j \in \mathcal{S}} d_j[\mathcal{S}],$$

where $\mathcal{A} = \{d_j[\mathcal{S}], \forall j \in \{1,\cdots,K\}, \forall \mathcal{S} \subseteq \{1,2,\cdots,K\}\}$ and $d_j[\mathcal{S}]$'s are located inside the boundary **B1** *defined by:*

$$\begin{aligned} &a. \quad d_{\hat{1}}[\mathcal{S}] + d_{\hat{2}}[\mathcal{S}_1] + \cdots + d_{\hat{q}}[\mathcal{S}_{q-1}] \leq \min(|\mathcal{S}|M, N) \\ &b. \quad \mathcal{S} \subseteq \mathcal{S}' : d_j[\mathcal{S}'] \leq d_j[\mathcal{S}], \end{aligned} \tag{4.48}$$

for all $\mathcal{S}, \mathcal{S}' \subseteq \{1,\cdots,K\}$ and for all sorting order of $\mathcal{S}$ represented by $\hat{\mathcal{S}} = \{\hat{1}, \hat{2}, \cdots, \widehat{|\mathcal{S}|}\}$. In equation (4.48)*, $q = |\mathcal{S}|$ and $\mathcal{S}_j = \{\mathcal{S} \backslash \{\hat{1}, \hat{2}, \cdots, \hat{j}\}\}$.*

To solve this problem, we first simplify the constraints specifying **B1**. We need the following lemma.

Lemma 4.4. *The upper-bound optimization problem defined in* (4.48) *has at least one symmetric solution. This symmetric solution has the following properties:*

1) $d_j[\mathcal{Q}]$ is equal for all user inside a set $\mathcal{Q}$, i.e., $d[\mathcal{Q}] = d_j[\mathcal{Q}], \forall j \in \mathcal{Q}$.

2) $d[\mathcal{Q}_1]$ is equal to $d[\mathcal{Q}_2]$ as long as $q = |\mathcal{Q}_1| = |\mathcal{Q}_2|$, i.e., $\alpha_q = d[\mathcal{Q}_1] = d[\mathcal{Q}_2]$.

Proof: The proof is similar to the proof of Lemma 4.2 (for the two-user scenario), and is based on the linearity of the objective function (i.e., (4.48)), the convexity of **B1**, and the fact that if $|\mathcal{Q}_1| = |\mathcal{Q}_2|$, the coefficients of $d_j[\mathcal{Q}_1]$ and $d_j[\mathcal{Q}_2]$ in (4.48) are equal.

B1.a Boundary Constraints

As described in (4.48), $\mathcal{S} \subseteq \{1,\cdots,K\}$; thus, there are $2^K - 1$ different selection for $\mathcal{S}$. Furthermore, for each $\mathcal{S}$, there are $|\mathcal{S}|!$ different sorting order. Therefore, the constraints of **B1.a** are in fact $\sum_{k=1}^{2^K-1} k!$ inequalities. Here, we show that instead of $\sum_{k=1}^{2^K-1} k!$ inequalities, it is possible to determine only K inequalities which construct the same region of **B1.a**.

a) We start by considering one group of constraints in **B1.a** related to one particular $\mathcal{S}$ with different sorting orders, i.e.:

$$d_{\hat{1}}[\mathcal{S}] + d_{\hat{2}}[\mathcal{S}_1] + \cdots + d_{\hat{q}}[\mathcal{S}_{q-1}] \leq \min(|\mathcal{S}|M, N), \tag{4.49}$$

where $q = |\mathcal{S}|$, $\mathcal{S}_j = \{\mathcal{S}\backslash\{\hat{1}, \hat{2}, \cdots, \hat{j}\}\}$, and $\hat{\mathcal{S}} = \{\hat{1}, \hat{2}, \cdots, \hat{q}\}$ is one desirable sorting order of $\mathcal{S}$. There are $q!$ inequities in this group.

According to the first symmetry property of Lemma 4.4, if $\mathcal{Q} \subseteq \{1, \cdots, K\}$, then $d_j[\mathcal{Q}]$ is the same for all user j inside of $\mathcal{Q}$, i.e., $\forall j \in \mathcal{Q}$. Thus, we can suppress the user subscript from equation (4.49):

$$d[\mathcal{S}] + d[\mathcal{S}_1] + \cdots + d[\mathcal{S}_{q-2}] + d[\mathcal{S}_{q-1}] \leq \min(|\mathcal{S}|M, N). \tag{4.50}$$

In other words, different sorting of set $\mathcal{S}$ does not generate independent constraints; therefore, all $|\mathcal{S}|!$ inequalities associated for different sorting order of the set $\mathcal{S}$ can be substituted by one constraint of the form (4.50).

b) Now, from the second symmetry property we know that $d[\mathcal{S}]$ is equal for all $\mathcal{S}$'s which have the same cardinality, i.e., the value of $d[\mathcal{S}]$ only depends on the cardinality of $\mathcal{S}$ (not the actual active users of $\mathcal{S}$). Therefore, (4.50) is the same for all subset $\mathcal{S}$'s which have the same cardinality. Thus, we can write (4.50) as the following:

$$\alpha_{|\mathcal{S}|} + \alpha_{|\mathcal{S}_1|} + \cdots + \alpha_{|\mathcal{S}_{q-2}|} + \alpha_{|\mathcal{S}_{q-1}|} \leq \min(|\mathcal{S}|M, N), \tag{4.51}$$

Furthermore, since $|\mathcal{S}_i| = |\mathcal{S}| - i$, (4.51) can be written as:

$$\sum_{k=1}^{|\mathcal{S}|} \alpha_k \leq \min(|\mathcal{S}|M, N). \tag{4.52}$$

Based on this observation, we can replace all constraints related to all $\mathcal{S}$'s with the same cardinality with one constraint of the form (4.52). We also know that $|\mathcal{S}|$ can take a value between one and K. Therefore, all $2^K - 1$ constraints can be replaced by K inequalities of the form (4.52). The following lemma summarizes the above discussion.

Lemma 4.5. *The set of inequalities which define the boundary condition of* **B1.a** *can be substituted by the following K inequalities:*

$$\sum_{k=1}^{q} \alpha_k \leq \min(qM, N), \tag{4.53}$$

where $q \in \{1, 2, \cdots, K\}$.

B1.b Boundary Constraints

Consider $\mathcal{S}, \mathcal{S}' \subseteq \{1, 2, \cdots, K\}$ such that $\mathcal{S} \subseteq \mathcal{S}'$. Due to **B1.b** constraint, $d_j[\mathcal{S}']$ and $d_j[\mathcal{S}]$ should satisfy:

$$d_j[\mathcal{S}'] \leq d_j[\mathcal{S}] \tag{4.54}$$

Based on Lemma 4.4, $d_j[\mathcal{Q}]$, $\mathcal{Q} \subseteq \{1, 2, \cdots, K\}$ is invariant to the selection of the user j and in addition is the same for all sets $\mathcal{Q}$ with the same cardinality; thus, we can rewrite (4.54) as:

$$\alpha_{|\mathcal{S}'|} \leq \alpha_{|\mathcal{S}|}. \tag{4.55}$$

Furthermore, since $\mathcal{S} \subseteq \mathcal{S}'$, we have $|\mathcal{S}| \leq |\mathcal{S}'|$. Having (4.55) for all sets of $\mathcal{S}$ and $\mathcal{S}'$, we can write **B1.b** constraints as:

$$\alpha_{q'} \leq \alpha_q, \qquad \forall q', q \in \{1, 2, \cdots, K\}, q \leq q', \tag{4.56}$$

or equivalently,

$$\alpha_K \leq \alpha_{K-1} \leq \cdots \leq \alpha_2 \leq \alpha_1. \tag{4.57}$$

The Objective Function

The objective function of the optimization problem in (4.48) can be written as the following:

$$\begin{aligned}\overline{D}_u &= \sum_{\forall \mathcal{S}:\mathcal{S}\subseteq\{1,\cdots,K\}} \mathrm{P}_{|\mathcal{S}|} \sum_{\forall j\in\mathcal{S}} d_j[\mathcal{S}] \\ &\overset{(a)}{=} \sum_{\forall \mathcal{S}:\mathcal{S}\subseteq\{1,\cdots,K\}} \mathrm{P}_{|\mathcal{S}|} |\mathcal{S}| d[\mathcal{S}] \\ &\overset{(b)}{=} \sum_{k=1}^{K} k\Gamma_k \alpha_k,\end{aligned} \tag{4.58}$$

where Γ_k is define by (4.33) and shows the probability that k users are active in the network. In equation (4.58), (a) and (b) are justified based on the first and second symmetry properties of Lemma 4.4, respectively.

Based on Lemma 4.5 and equations (4.57) and (4.58), the next theorem summarizes the new outer-bound of the network average DoF.

Theorem 4.7. *Consider a symmetric K-user MIMO random access network in which each user and the AP are equipped with M and N antennas, respectively. One upper-bound of the network average DoF can be determined as:*

$$\begin{aligned}\max_{\mathcal{A}} \quad & \overline{D}_u = \sum_{k=1}^{K} k\Gamma_k \alpha_k \\ s.t. \quad & a. \quad \sum_{k=1}^{q} \alpha_k \le \min(qM, N), \quad \forall q \in \{1, 2, \cdots, K\} \\ & b. \quad 0 \le \alpha_K \le \alpha_{K-1} \le \cdots \le \alpha_2 \le \alpha_1,\end{aligned} \tag{4.59}$$

where $\mathcal{A} = \{\alpha_j, \forall j \in \{1, 2, \cdots, K\}\}$ and Γ_k is defined in (4.33).

Do We Need All Boundary Constraints

In this section we want to see whether any of the boundary constraints of the maximization problem of (4.59) is redundant or not. We need the following lemma.

Lemma 4.6. *For all $q \in \{1, 2, 3, \cdots, K\}$:*

a) $\sum_{k=1}^{q} \alpha_k \le q\alpha_1$,

b) $\sum_{k=1}^{q} \alpha_k \le \sum_{k=1}^{K} \alpha_k$.

Proof: The proof is straight forward and can be verified by noting that $0 \leq \alpha_K \leq \alpha_{K-1} \leq \cdots \leq \alpha_2 \leq \alpha_1$.

Lets write the boundary conditions of (4.59) for three cases of $q = 1$, $q = j$ where $j \in \{2, \cdots, K-1\}$, and $q = K$:

$$
\begin{aligned}
CO_1 : \alpha_1 &\leq \min(M, N), \\
CO_2 : \sum_{k=1}^{j} \alpha_k &\leq \min(jM, N), \; j \in \{2, \cdots, K-1\}, \\
CO_3 : \sum_{k=1}^{K} \alpha_k &\leq \min(N, KM).
\end{aligned}
\tag{4.60}
$$

Now assume that q such that $qM \leq N$; therefore, CO_1 and CO_2 reduce to:

$$
\begin{aligned}
CO_1 : \alpha_1 &\leq M, \\
CO_2 : \sum_{k=1}^{q} \alpha_k &\leq qM.
\end{aligned}
\tag{4.61}
$$

Multiplying the first inequality by q, we get:

$$
q\alpha_1 \leq qM. \tag{4.62}
$$

Furthermore, based on the Lemma 4.6-a and non-negativity of α_k's, we know that:

$$
\sum_{k=1}^{q} \alpha_k \leq q\alpha_1, \tag{4.63}
$$

Combining equations (4.62) and (4.63), we can conclude that as long as CO_1 is satisfied, all other boundary conditions of q in which $qM \leq N$ are also satisfied, and therefore they are redundant.

Now assume q such that $N < qM$, then, CO_2 and CO_3 can be written as:

$$\begin{aligned} CO_2 : \sum_{k=1}^{q} \alpha_k &\leq N, \\ CO_3 : \sum_{k=1}^{K} \alpha_k &\leq N. \end{aligned} \tag{4.64}$$

The inequality in Lemma 4.6-b shows that CO_2 is redundant with respect to CO_3 (where we have $N < qM$).

As a summary, we see that CO_2 constraints for all $j \in \{2, 3, \cdots, K-1\}$ are redundant with respect to either CO_1 or CO_3, depending on the values of jM and N. Therefore, we can remove all constraints $1 < q < K$ and only keep the first and the last boundary constraints. Therefore, the optimization problem of (4.59) can be written as:

$$\begin{aligned} \max_{\mathcal{A}} \quad & \overline{D}_u = \sum_{k=1}^{K} k\Gamma_k \alpha_k \\ s.t. \quad & a1. \quad \alpha_1 \leq \min(N, M) \\ & a2. \quad \sum_{k=1}^{K} \alpha_k \leq \min(N, KM) \\ & b. \quad 0 \leq \alpha_K \leq \alpha_{K-1} \leq \cdots \leq \alpha_2 \leq \alpha_1, \end{aligned} \tag{4.65}$$

where $\mathcal{A} = \{\alpha_j, \forall j \in \{1, \cdots, K\}, j \in \{1, 2, \cdots, K\}\}$ and Γ_k is defined in (4.33).

To further simplify the upper-bound, let's define $\beta_K = \alpha_K$ and $\beta_i = \alpha_i - \alpha_{i+1}$ where $i \in \{1, 2, \cdots, K-1\}$. We then rewrite the optimization problem of (4.65) based on these new variables. This leads us to the final upper-bound optimization problem which is presented in (4.32) and therefore Theorem 4.4 is proved.

4.4.3 Proof of Theorem 4.5 (The Achievability Scheme)

As briefly explained in Section 4.4.1, the achievability scheme depends on the network parameters. In this section, we discuss each situation in more details and bring an example for each scenario. We need the following lemma in the following.

Lemma 4.7. *[59], Consider a k-user multiple access network in which each user and the AP equipped with M and N antennas, respectively. There exist a coding scheme (based on random coding) in which the decodable DoF of each user is equal to* $\min(M, \frac{N}{k})$.

1) Single-Stream Codes

a) Case $N \leq M$:

For such network settings, Table 4.3 presents the network average upper-bound in which all β_k's are set to zero except one of them in which $\beta_\ell = \frac{N}{\ell}$ where ℓ is determined such that $\rho_{\ell-1} < \rho \leq \rho_\ell$ and ρ is the probability of node activation.

Given this upper-bound and considering the definition of β_j, it is implied that a scheme achieves the upper-bound if we propose a coding scheme such that all users transmit with the DoF of $\frac{N}{\ell}$ and the AP can decode the users transmitted data as long as the number of active users in a time slot is less than ℓ. To construct such a code, we use the idea of random coding. Assuming a long block length of $t \geq KM + N + 1$, each user (for instance user j) generates a Gaussian codebook $\mathcal{C}^j$ which has $SNR^{\delta_j \times t}$ codewords where $\delta_j = \frac{N}{\ell}$ (representing the DoF of the code) and ℓ is determined such that $\rho_{\ell-1} < \rho \leq \rho_\ell$. Each codeword is an $M \times t$ matrix which its elements are drawn from a zero mean Gaussian random variable with unit variance.

By this selection, while the number of active users k, is less than ℓ, we have $\sum_{j=1}^{k} \delta_j \leq N$. Furthermore, since in this scenario $N \leq M$, the DoF of each user, i.e., $\frac{N}{\ell}$, is also less than the number of antennas available at each user, i.e., M. Having these two properties, Lemma 4.7 guarantees that if there are k (where $k \leq \ell$) active users, the AP is able to decode all data streams, i.e.:

$$d_j[\mathcal{Q}] = \delta_j = \frac{N}{\ell}, \ \forall j \in \mathcal{Q} \text{ where } |\mathcal{Q}| \leq \ell \tag{4.66}$$

Substituting (4.66) in (4.31), the achievable network average DoF can be computed as:

$$\begin{aligned}
\overline{D}_l &= \sum_{\forall \mathcal{Q}: \mathcal{Q} \subseteq \{1, \cdots, K\}} \mathrm{P}_{|\mathcal{Q}|} D[\mathcal{Q}] \\
&= \sum_{\forall \mathcal{Q}: \mathcal{Q} \subseteq \{1, \cdots, K\}} \mathrm{P}_{|\mathcal{Q}|} \sum_{j \in \mathcal{Q}} d_j[\mathcal{Q}] \\
&= \sum_{k=1}^{K} \sum_{\substack{\forall \mathcal{Q}: |\mathcal{Q}|=k, \\ \mathcal{Q} \subseteq \{1, \cdots, K\}}} \mathrm{P}_{|\mathcal{Q}|} \sum_{j \in \mathcal{Q}} d_j[\mathcal{Q}] \\
&= \sum_{k=1}^{\ell} \sum_{\substack{\forall \mathcal{Q}: |\mathcal{Q}|=k, \\ \mathcal{Q} \subseteq \{1, \cdots, K\}}} k \mathrm{P}_k \frac{N}{\ell} \\
&= \left(\sum_{k=1}^{\ell} k \Gamma_k \right) \frac{N}{\ell}
\end{aligned} \tag{4.67}$$

where Γ_k is defined in (4.33) and shows the probability that k users are active in the network. As can be verified the achievable network average DoF is equal to the upper-bound presented in Table 4.3. Therefore, we have the optimal network average DoF in this scenario.

To illustrate, Fig. 4.6 shows the network average DoF for different network setups. Note that, to achieve the upper-bound, the DoF of each user should be modified according to the value of ρ. For instance, consider the case in which $K = 5, M = 3$, and $N = 2$. For this setup, as long as $0.669 < \rho \leq 1$ (or equivalently $\ell = 5$) the optimal scheme is that each user transmits a single-stream code with the DoF of $\frac{2}{5}$. This data can always be decoded regardless of the number of active users. Then, if $0.4860 < \rho \leq 0.669$ (or equivalently $\ell = 4$) the DoF of each user should be equal to $\frac{2}{4}$. In this case, the transmitted data can be decode, if less than five users are active. If five users become active, we loos all the transmitted data. In a same way, if $0.333 < \rho \leq 0.4860$ (or $\ell = 3$), $0.2 < \rho \leq 0.333$ (or $\ell = 2$), and $0 \leq \rho \leq 0.2$ (or $\ell = 1$), each user should transmit a single-stream data with the DoF of $\frac{2}{3}$, $\frac{2}{2}$, and $\frac{2}{1}$, respectively.

b) Case $N \geq KM$:

The achievability scheme of this case is similar to the previous case, i.e., each user sends one stream of data with a DoF of M, i.e, $\delta_j = M$. Assuming all users are active, the DoF

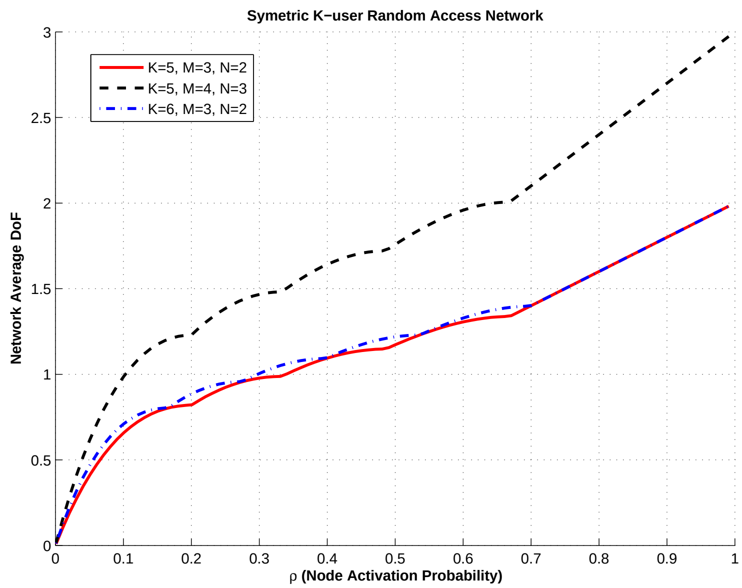

Figure 4.6: Network average DoF: K-user random access network where $N \leq M$.

of the received data is equal to $\sum_{j=1}^{K} \delta_j = KM$ which is less than N due to the $N \geq KM$ assumption in this case. Clearly, the DoF of the received data remain less than N when fewer users are active. Therefore, from Lemma 4.7 we conclude that the transmitted data from all users can be decoded in all network states and the DoF of each user data is equal to M. Similar to (4.67), the achievable network average DoF can be computed as:

$$\begin{aligned} \overline{D}_l &= \sum_{k=1}^{K} \sum_{\substack{\forall \mathcal{Q}:|\mathcal{Q}|=k, \\ \mathcal{Q}\subseteq\{1,\cdots,K\}}} \mathrm{P}_{|\mathcal{Q}|} \sum_{j\in\mathcal{Q}} d_j[\mathcal{Q}] \qquad (4.68) \\ &= \sum_{k=1}^{K} \sum_{\substack{\forall \mathcal{Q}:|\mathcal{Q}|=k, \\ \mathcal{Q}\subseteq\{1,\cdots,K\}}} k\mathrm{P}_k M \\ &= \left(\sum_{k=1}^{K} k\Gamma_k \right) M = KM\rho \end{aligned}$$

where Γ_k is defined in (4.33). Comparing this achievable network average DoF with the result of Table 4.3 proves that we have the optimality result in this case as well.

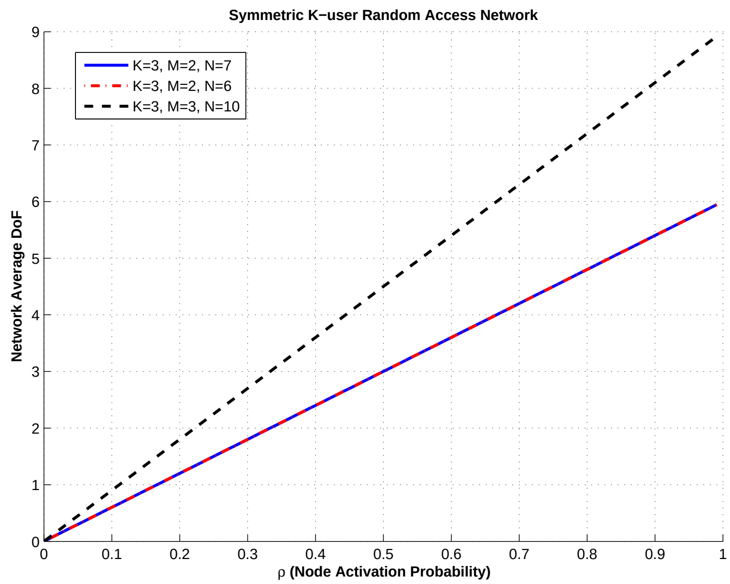

Figure 4.7: Network Average DoF: K-user random access network where $KM \leq N$.

As an example, Fig. 4.7 shows the network average DoF for different network settings. In each scenario, the optimal scheme is that each user sends a single-stream with the DoF of M. The figure verifies that the optimal transmission remains unchanged for all range of ρ. Another observation is that the optimal average network DoF is not increased if we increase N beyond $2M$ (i.e., from $N = 6$ to $N = 7$). It is due to the fact that there is no additional degrees of freedom at the transmitter; therefore, the extra dimension at the receiver cannot be exploited.

c) Case $M < N < KM$:

Two situations occur for the upper-bound of the network average DoF.

First, as long as $\rho \geq \rho_{j_2-1}$, there is only one non-zero β_k. The achievability scheme is therefore similar to the network of $N \leq M$. In other words, each user (for instance user j), construct a codebook ($\mathcal{C}^j$) which encodes a single-stream of data with the DoF of $\delta_j = \frac{N}{\ell}$ where ℓ is determined such that $\rho_{\ell-1} < \rho \leq \rho_\ell$. This selection ensures the decodability of the transmitted data as long as not more than ℓ users are active in the network. Thus, similar to (4.67), we have:

$$\overline{D}_l = \sum_{k=1}^{\ell} \sum_{\substack{\forall \mathcal{Q}:|\mathcal{Q}|=k, \\ \mathcal{Q}\subseteq\{1,\cdots,K\}}} k\mathrm{P}_k \frac{N}{\ell} = \left(\sum_{k=1}^{\ell} k\Gamma_k\right) \frac{N}{\ell} \quad (4.69)$$

where Γ_k is defined in (4.33). Comparing this result with the upper-bound of Table 4.3 shows that the single-stream code is optimal in the case that $\rho \geq \rho_{j_2-1}$. However, as will be shown, the single-stream coding cannot achieve the upper-bound where $\rho < \rho_{j_2-1}$. In the next part we discuss these situations in more details.

2) <u>Multi-Stream Codes:</u> $M < N < KM$ **and** $\rho < \rho_{j_2-1}$

The upper-bound of the network average DoF, in the scenarios where $M < N < KM$ and $\rho < \rho_{j_2-1}$, is presented in Table 4.3 in which we set $\beta_{\ell-1} = \ell M - N$, $\beta_\ell = N-(\ell-1)M$, and $\beta_j = 0, \forall j \neq \ell$ and $(\ell-1)$. This result implies that the upper-bound can be achieved if we can construct a code such that the decodable DoF of each user is equal to M if up to $\ell - 1$ users become active. Furthermore, where there is ℓ active users, we should be able to partially decode the transmitted data and get the DoF of $N - (\ell - 1)M$.

The **main difference** between the upper-bound of these scenarios and the previous case is that, here, the upper-bound is achieved when two β_j are not equal to zero while in the previous cases there is only one non-zero β_j. It is the reason that in these situations, although single-stream codes are still a lower-bound of the achievable network average DoF, they are no longer the optimal scheme and connot reach the upper-bound. The existence of the optimal code for arbitrarily selection of K, M, and N is an open problem. Here, we aim to present some examples of networks for which we are able to construct a coding scheme which achieves the upper-bound of the network average DoF. Therefore, we have the optimal network average DoF in these special network setups.

As will be discussed, the proposed code consists of multiple streams of data. It is the reason that we use the term *multi-stream code* for this transmission scheme. The transmit direction of each code is selected such that all of the streams from all users can be decoded if less than ℓ users are active. Furthermore, if ℓ users are active, the AP still can decode parts of the transmitted streams from each users [57,58]. Simultaneous activation of more than ℓ users will corrupt all streams and nothing can be decoded at the AP.

As noted, we are always able to construct the single-stream code (a code which can be either received completely or we loos all transmitted data) and consider it as a lower-bound for the network average DoF. In the single-stream code, each user (for instance user j), selects a code-book ($\mathcal{C}^j$) with DoF of:

$$\delta_j = \begin{cases} M & \text{if } \frac{\sum_{k=1}^{\ell-1} k\Gamma_k}{\ell\Gamma_\ell} \leq \frac{N}{M\ell - N} \\ \frac{N}{\ell} & \text{if } \frac{\sum_{k=1}^{\ell-1} k\Gamma_k}{\ell\Gamma_\ell} > \frac{N}{M\ell - N} \end{cases} . \tag{4.70}$$

The achievable average DoF, therefore, will be:

$$\overline{D}_l = \begin{cases} \sum_{k=1}^{\ell-1} k\Gamma_k M & \text{if } \frac{\sum_{k=1}^{\ell-1} k\Gamma_k}{\ell\Gamma_\ell} \leq \frac{N}{M\ell - N} \\ \sum_{k=1}^{\ell} k\Gamma_k \frac{N}{\ell} & \text{if } \frac{\sum_{k=1}^{\ell-1} k\Gamma_k}{\ell\Gamma_\ell} > \frac{N}{M\ell - N} \end{cases} . \tag{4.71}$$

In the sequel, the performance of the single-stream coding approach is also presented when we analyze different network setup.

4.4.4 Examples

In this section, we bring three examples. In the first two examples we discuss two network setups for which we are able to construct an achievability scheme based on a multi-stream data transmission. Thus, we have the optimality results in these situations. As we mentioned there are some scenarios in which the multi-stream data transmission is not able to achieve the upper-bound. The third example of this section presents one of these scenarios. Note that, based on Section 4.4.3.1-c, in all setups (including the last one) we have the optimal network average DoF if ρ is larger than a threshold, namely ρ_{j_2-1}.

Example 1: As the first example, consider a K-user random access network where $M = K - 1$ and $N = 2K - 3$. Therefore, $j_2 = \lfloor \frac{N}{M} \rfloor + 1 = 2$ and $\rho_{j_2-1} = \frac{1}{K}$.

In this network, $M < N < KM$. Therefore, we know that as long as $\rho \geq \frac{1}{K}$ single-

Table 4.4: Notations used for the intersections of column spaces

$\mathbf{a} = \mathfrak{I}(\mathbf{H}_1, \mathbf{H}_2)$	$\Rightarrow$	$\mathbf{a} = \mathbf{H}_1\mathbf{a}_1 = \mathbf{H}_2\mathbf{a}_2$
$\mathbf{b} = \mathfrak{I}(\mathbf{H}_1, \mathbf{H}_3)$	$\Rightarrow$	$\mathbf{b} = \mathbf{H}_1\mathbf{b}_1 = \mathbf{H}_3\mathbf{b}_3$
$\mathbf{c} = \mathfrak{I}(\mathbf{H}_1, \mathbf{H}_4)$	$\Rightarrow$	$\mathbf{c} = \mathbf{H}_1\mathbf{c}_1 = \mathbf{H}_4\mathbf{c}_4$
$\mathbf{d} = \mathfrak{I}(\mathbf{H}_2, \mathbf{H}_3)$	$\Rightarrow$	$\mathbf{d} = \mathbf{H}_2\mathbf{d}_2 = \mathbf{H}_3\mathbf{d}_3$
$\mathbf{e} = \mathfrak{I}(\mathbf{H}_2, \mathbf{H}_4)$	$\Rightarrow$	$\mathbf{e} = \mathbf{H}_2\mathbf{e}_2 = \mathbf{H}_4\mathbf{e}_4$
$\mathbf{f} = \mathfrak{I}(\mathbf{H}_3, \mathbf{H}_4)$	$\Rightarrow$	$\mathbf{f} = \mathbf{H}_3\mathbf{f}_3 = \mathbf{H}_4\mathbf{f}_4$

stream code, Section 4.4.3.1-c, achieves the optimal network average DoF. In these cases, we first find ℓ such that $\rho_{\ell-1} \leq \rho \leq \rho_\ell$. Then, we construct a single-stream code with the DoF of $\frac{N}{\ell}$ for each user.

Now we focus on the settings that $\rho < \frac{1}{K}$. Based on Table 4.3, we know that the upper-bound is achievable if we are able to design a code such that we can get the DoF of $K-1$ where only one user transmits. Furthermore, the DoF of $2K-4$ should be decodable where each two users become active. To design such a code, each user constructs multiple streams of data and transmits them over predetermined spatial directions. These directions are selected such that at least some of them do not overlap with each other (when more than one user become active); therefore, we can decode some parts of the data.

To simplify the explanation, we assume $K = 4$, i.e., we have a four-user random access network and users and the AP are equipped with 3 and 5 antennas, respectively. We focus on the cases where $\rho < \frac{1}{4}$. Table 4.3 suggests that we should design a code such that we can achieve the DoF of 3 and 2 (per user) when one and two users are active, respectively.

Let $\mathbf{H}_j,\ j \in \{1,2,3,4\}$ denotes the 5×3 channel propagation matrix between user j and the AP. The order of the column spaces of all $\mathbf{H}_j,\ j \in \{1,2,3,4\}$ is 3 and the intersection of the column spaces between each two $\mathbf{H}_j$ is one (the space of the received signal is of the order of 5). The intersections of each two channel matrices are summarized in Table 4.4. In this table, $\mathfrak{I}(\cdot,\cdot)$ is a function with two matrices as the inputs, and its output is the intersection of the two input matrices. For instance, $\mathbf{b} = \mathfrak{I}(\mathbf{H}_1, \mathbf{H}_3)$ means that the intersection of the column spaces of $\mathbf{H}_1$ and $\mathbf{H}_3$ is denoted by $\mathbf{b}$. Furthermore, the second

Table 4.5: Signal structure: Four-user random access network where $M = 3$ and $N = 5$

	s_1	s_2	s_3
User 1	$\mathbf{a}_1$	$\mathbf{b}_1$	$\mathbf{c}_1$
User 2	$\mathbf{a}_2$	$\mathbf{d}_2$	$\mathbf{e}_2$
User 3	$\mathbf{b}_3$	$\mathbf{d}_3$	$\mathbf{f}_3$
User 4	$\mathbf{c}_4$	$\mathbf{e}_4$	$\mathbf{f}_4$

expression in each row of Table 4.4 shows the direction along which a user should transmit such that the data gets to the AP along the direction of the corresponding intersection. For instance, $\mathbf{b}_1$ and $\mathbf{b}_3$ denote the directions that user 1 and user 3 should transmit such that their data gets to the AP along direction $\mathbf{b}$.

To transmit data, each user partitions its data into three streams (s_1, s_2, and s_3) each with the DoF of one. These streams are then transmitted over three predetermined directions. Table 4.5 summarizes the code structure of each user in this setup. For instance, Table 4.5 shows that user 3 transmits data over three directions of $\mathbf{b}_3$, $\mathbf{d}_3$, and $\mathbf{e}_3$. Therefore, s_1, s_2, and s_3 will be received along the directions of $\mathbf{b}$, $\mathbf{d}$, and $\mathbf{e}$, respectively.

We claim that with this structure we can achieve the DoF of 3 if one user is active and also we can decode the DoF of 2 (per user) in cases that two users are active. For instance assume that user 1 is active. Based on Table 4.5, it transmits three messages along $\mathbf{a}_1$, $\mathbf{b}_1$, and $\mathbf{c}_1$ spatial directions. These messages are received along directions $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ (see Fig. 4.8). Sine there is no interferer, all streams can be decoded and we get the DoF of 3.

Now assume that another user, for instant user 2, starts transmission. It transmits along $\mathbf{a}_2$, $\mathbf{d}_2$, and $\mathbf{e}_2$ directions. Therefore, at the AP we will have the following signals:

$$\begin{aligned} &\text{user } 1(s_2) : \text{direction } \mathbf{b}, \quad \text{user } 1(s_3) : \text{direction } \mathbf{c}, \\ &\text{user } 2(s_2) : \text{direction } \mathbf{d}, \quad \text{user } 2(s_3) : \text{direction } \mathbf{e}, \\ &\qquad \text{user } 1(s_1) + \text{user } 2(s_1) : \text{direction } \mathbf{a} \end{aligned} \tag{4.72}$$

This signal structure is also pictorially depicted in Fig. 4.8. As can be seen, we have five direction that over four of them ($\mathbf{b}, \mathbf{c}, \mathbf{d}$, and $\mathbf{e}$) only one stream is transmitted. Since the DoF of each stream is one, we can decode these parts and get the DoF of 4 (DoF of 2 per

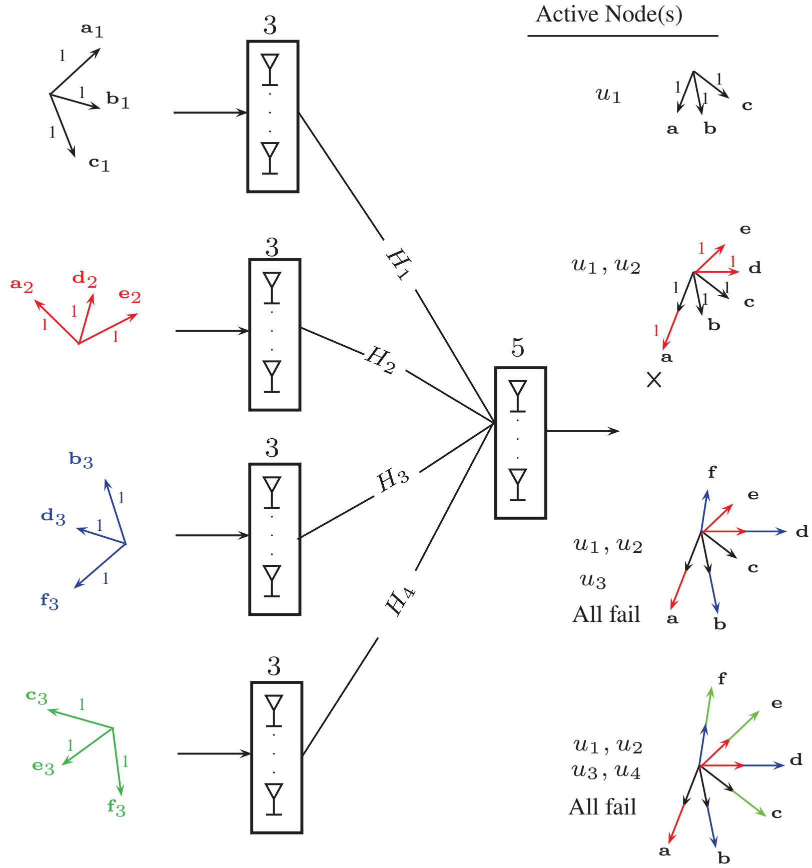

Figure 4.8: Signal structure: Four-user random access network where $M = 3$ and $N = 5$. The last two cases are not-decodable since there are more than 5 received directions at the AP.

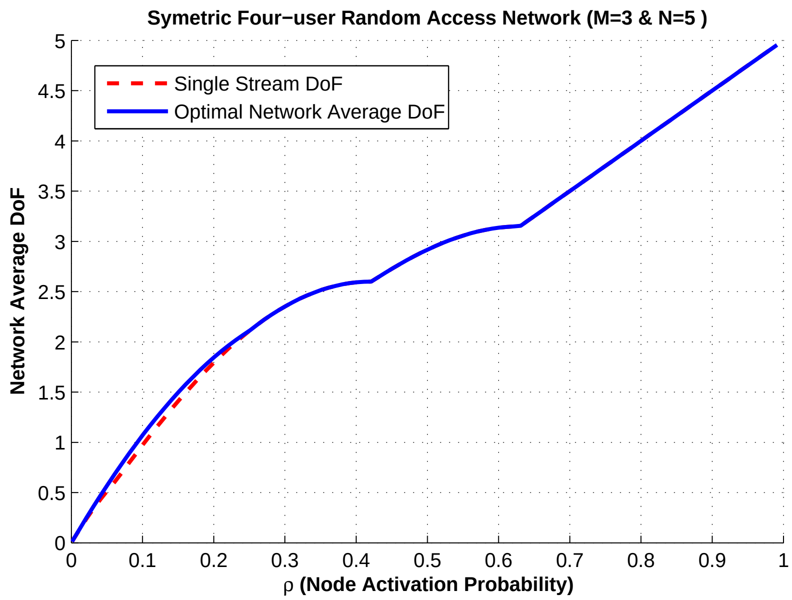

Figure 4.9: Network average DoF: Four-user random access network where $M = 3$ and $N = 5$.

user). Note that, we cannot decode the messages which are sent along direction **a**. It is due to the fact that in this direction we have two signal (s_1 of user 1 and s_1 of user 2) each has the DoF of 1.

The above reasoning is not limited to users 1 and 2 and applies to all cases that there are two active users in the network, i.e., regardless of actual active users, we can get the DoF of 3 and 2 (per user) if the number of active users is one and two, respectively. This code, therefore, achieves the upper-bound in the range of $\rho < \frac{1}{4}$ and shows the optimality of the results in Theorem 4.4. Note that, this strategy can be extended to design codes for the general cases where the number of users is not equal to 4.

Figure 4.9 shows the optimal network average DoF for this 4-user network. For the range of $\rho > 0.25$, the single-stream data transmission is optimal, i.e., where $0.63 < \rho \leq 1$, $0.422 < \rho \leq 0.63$, and $0.25 < \rho \leq 0.422$ each user transmits a single-stream with the DoF of $\frac{5}{4}$, $\frac{5}{3}$, $\frac{5}{2}$, respectively. In case that $\rho < 0.25$, the optimal scheme is to use the introduced multi-stream technique which aligns the signals of different users. For comparison, we have also plotted the achievable network average DoF if we use the single-stream code where $\rho < 0.25$. The gap between the two schemes is evident in the figure.

Table 4.6: Signal structure: Four-user random access network where $M = 4$ and $N = 7$

	s_1	s_2	s_3	s_4
User 1	$\mathbf{a}_1$	$\mathbf{b}_1$	$\mathbf{c}_1$	$\mathbf{g}$
User 2	$\mathbf{a}_2$	$\mathbf{d}_2$	$\mathbf{e}_2$	$\mathbf{h}$
User 3	$\mathbf{b}_3$	$\mathbf{d}_3$	$\mathbf{f}_3$	$\mathbf{i}$
User 4	$\mathbf{c}_4$	$\mathbf{e}_4$	$\mathbf{f}_4$	$\mathbf{j}$

Example 2: Another example of networks in which we can prove the optimality of the upper-bound is a K-user network where $M = K$ and $N = 2K-1$. Thus, $j_2 = \lfloor \frac{N}{M} \rfloor + 1 = 2$ and $\rho_{j_2-1} = \frac{1}{K}$.

Based on Table 4.3, a single-stream code, Section 4.4.3.1-c, achieves the optimal network average DoF if $\rho \geq \frac{1}{K}$.

Furthermore, Table 4.3 suggests that the upper-bound can be achieved for $\rho < \frac{1}{K}$ if we can design a code such that the AP can decode the DoF of K and $K-1$ (per user) where the number of active users is one and two, respectively. To simplify the explanation, here, we consider a 4-user network ($K = 4$) where $M = 4$ and $N = 7$. In this setup, $\mathbf{H}_j$, $j \in \{1, 2, 3, 4\}$ are matrices of size 7×4. Therefore, the order of the intersection of the column space ofs each two $\mathbf{H}_j$ is still 1. The notation that we used to refer to these intersections is the same as Example 1 (Table 4.4).

To transmit data, in this scenario, each user partitions its data into four streams (s_1, s_2, s_3 and s_4) each with the DoF of one. The transmit direction of each stream is summarized in Table 4.5. In Table 4.6, directions $\mathbf{g}, \mathbf{h}, \mathbf{i}$, and $\mathbf{j}$ are selected such that they are independent to the other vectors transmitted by their corresponding user. For instance, $\mathbf{g}$ is determined such that it is independent of $\mathbf{a}_1, \mathbf{b}_1$, and $\mathbf{c}_1$.

To prove the optimality, we should show that the achievable DoF of this code is 4 if one user is active and 3 (per user) in cases that two users are active. Figure 4.10 shows this situation. In this figure, $\mathbf{g}' = H_1\mathbf{g}$ and $\mathbf{h}' = H_2\mathbf{h}$. For instance, we assume user 1 is the only active user. In this case there is no interferer in the network and therefore, all four transmitted streams are received along four independent directions at the AP (they can be decode successfully). The achievable DoF is therefore equal to 4. Next, we study the case with two active users in the network, without loss of generality, we assume user 1 and

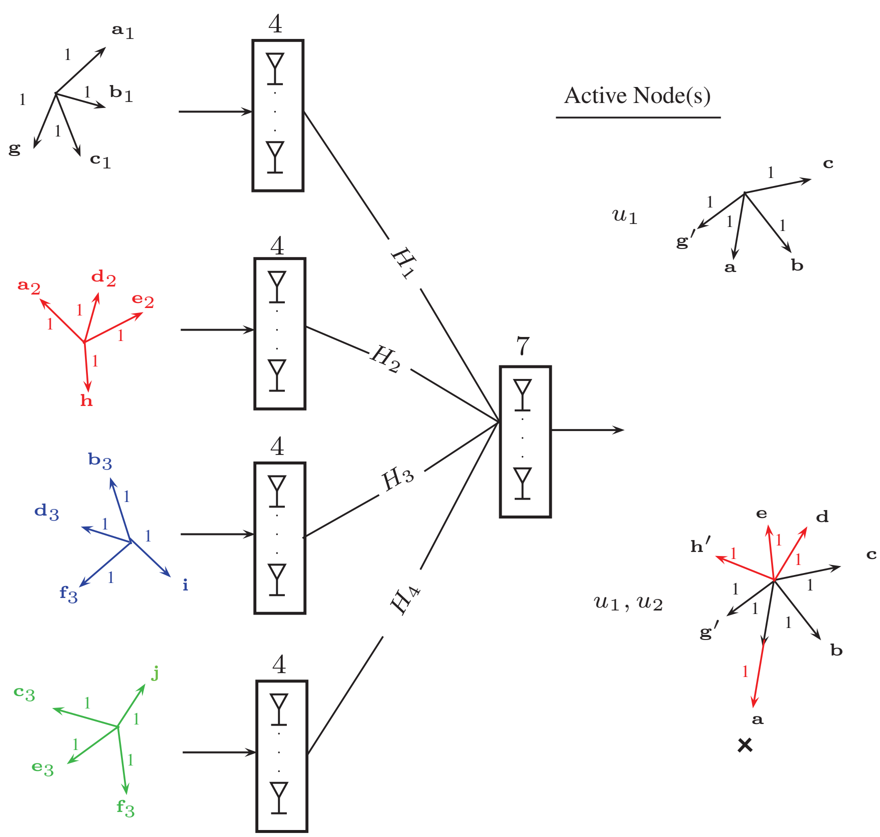

Figure 4.10: Signal Structure: Four-user random access network where $M = 4$ and $N = 7$.

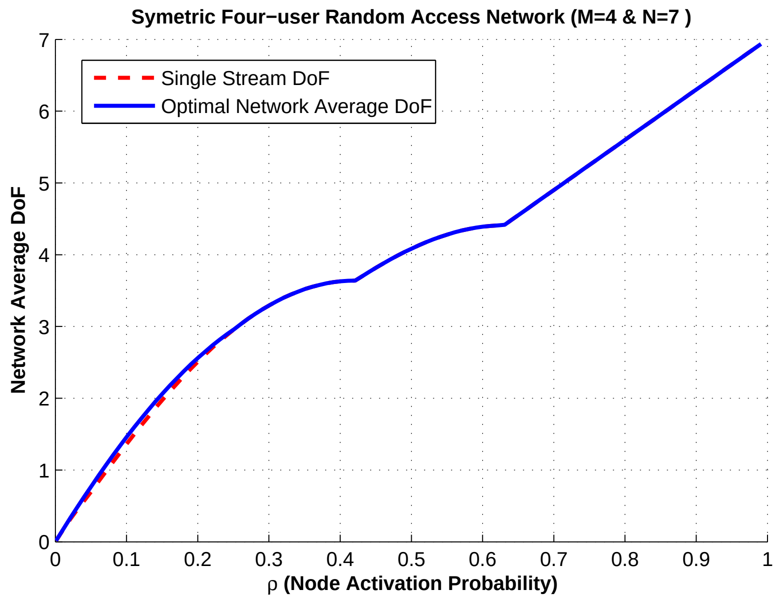

Figure 4.11: Network average DoF: Four-user random access network where $M = 4$ and $N = 7$.

user 2 are active. Looking into the structure of the received signal (depicted in Fig. 4.10), we see that (except direction a) there is only one data stream per each spatial direction. Thus, all of them could be decoded and we get the DoF of 6 (3 per user). Note that the data sent along a_1 and a_2 are received aligned with each other when they get to the AP and thus we cannot decode them.

Extending this results to any one or two active users, we see that for $\rho < \frac{1}{4}$ the DoF of 4 and 3 (per user) is achievable, respectively. This proves that the upper-bound of Theorem 4.4 is indeed optimal.

Figure 4.11 shows the optimal network average DoF for this 4-user network. Similar to Example 1, where ρ is higher than a threshold, i.e., $0.25 \leq \rho$, the single-stream code is optimal and where $\rho < 0.25$ we should use the proposed alignment technique. Figure 4.11 presents the achievable DoF of the sub-optimal single-stream code as well.

Note that these cases are only two examples that we mentioned to show that the upper-bound of Theorem 4.4 is tight in some scenarios. However, there are also some network setups that the introduced vector alignment technique cannot reach the upper-bound. In

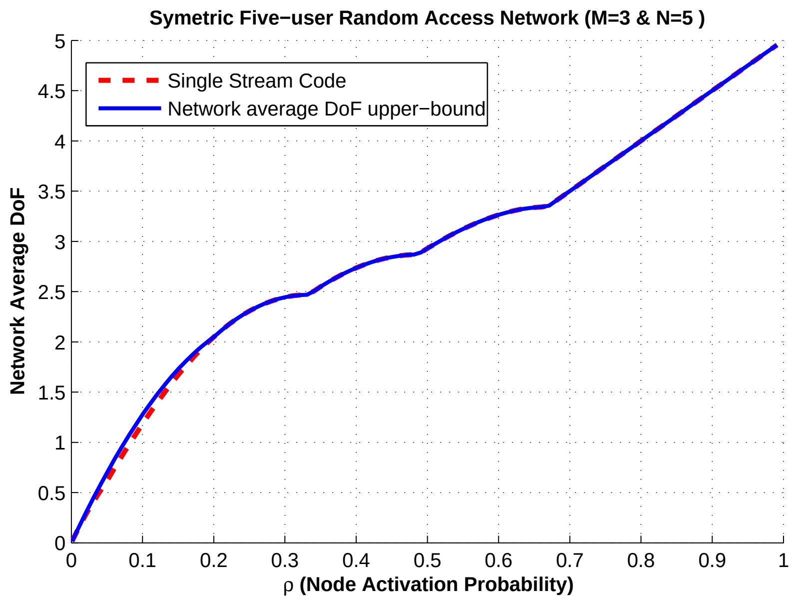

Figure 4.12: Network average DoF: Five-user random access network where $M = 3$ and $N = 5$.

these scenarios we can use the single-stream code as a lower-bound of the achievable network average DoF. For instance, in Example 3 we analyze one of these network conditions.

Example 3: Consider a network with five users where each user and the AP are equipped with 3 and 5 antennas, respectively.

For this network, from (4.34), we have $\rho_{j_2-1} = \frac{1}{5}$. Therefore, as long as $\rho \geq \frac{1}{5}$, a single-stream coding technique based on Section 4.4.3.1-c can achieve the optimal network average DoF. However, if $\rho < \frac{1}{5}$ there is a gap between the achievable rates of the single-stream code and the upper-bound of Theorem 4.4. The vector alignment scheme which we applied for the previous two examples fails to fill this gap as well. Further investigation is required to find out the optimal network average DoF in these situations.

The results are presented in Fig. 4.12 where we have the optimum results only for $\rho \geq 0.2$. The gap between the lower and the upper-bound (where $\rho < \frac{1}{5}$) is shown in the figure as well.

4.5 Summary

We studied a random multiple access network with K users (each equipped with M antennas) and one Access Point (with N antenna). There is no central controller for the network and, at the beginning of each a time slot, users (independently and with probability ρ) start transmission of data. In this chapter, we proposed an upper-bound of the network average DoF for different network setups. In addition, we presented an achievability scheme which reaches the upper-bound of the network average DoF for all symmetric two-user random access networks with different values of M and N. The optimality result also holds for symmetric K-user random access networks where i) $N \leq M$, or ii) $KM \leq N$, or iii) $M < N < KM$ and ρ is larger than a threshold. The existence of the achievability scheme is also proved for some specific network examples in which ρ is smaller than the threshold and $M < N < KM$. This optimal strategy is based on the idea of multi-stream data transmission and vector interference alignment.

4.6 Fourier-Motzkin Elimination Technique for Solving Equation (4.14)

As mentioned in the last paragraph of 4.3.1, we can use Fourier-Motzkin elimination technique to show that $\overline{D}_u$ in (4.14) is equivalent to the results of Table 4.1. To show the steps of the proof, in this section we find the value of $\overline{D}_u$ for the cases that $N \leq M$ and $\rho \leq \frac{1}{2}$ or $\rho > \frac{1}{2}$. The other cases can be proved similarly. For convenience, lets rewrite (4.14) considering that $N \leq M$.

$$\begin{aligned} \overline{D}_u &\leq 2\rho(1-\rho)\beta_1 + 2\rho\beta_2 && (4.73)\\ \beta_1 + \beta_2 &\leq N && (4.74)\\ \beta_1 + 2\beta_2 &\leq N && (4.75)\\ \beta_1 &\geq 0 && (4.76)\\ \beta_2 &\geq 0. && (4.77) \end{aligned}$$

First, we combine the above inequalities such that we can cancel β_1:

$$(4.73) + 2\rho(1-\rho) \times (4.74) \tag{4.78}$$
$$(4.73) + 2\rho(1-\rho) \times (4.75) \tag{4.79}$$
$$(4.76) + (4.74) \tag{4.80}$$
$$(4.76) + (4.75) \tag{4.81}$$
$$(4.77) \tag{4.82}$$

Two cases can happen:

1) Considering $\rho \leq \frac{1}{2}$, we have:

$$\overline{D}_u \leq 2\rho^2\beta_2 + 2\rho(1-\rho)N \tag{4.83}$$
$$\overline{D}_u + 2\rho(1-2\rho)\beta_2 \leq 2\rho(1-\rho)N \tag{4.84}$$
$$\beta_2 \leq N \tag{4.85}$$
$$2\beta_2 \leq N \tag{4.86}$$
$$\beta_2 \geq 0 \tag{4.87}$$

It can be seen that (4.85) is redundant with respect to (4.86). Next, we combine inequities (4.83)-(4.87) such that we can cancel β_2.

$$(1-2\rho) \times (4.83) + \rho \times (4.84) \tag{4.88}$$
$$(4.83) + \rho^2 \times (4.86) \tag{4.89}$$
$$2\rho(1-2\rho) \times (4.87) + (4.84) \tag{4.90}$$
$$2 \times (4.87) + (4.86) \tag{4.91}$$

Therefore, we have:

$$\overline{D}_u \leq 2\rho(1-\rho)N \tag{4.92}$$
$$\overline{D}_u \leq \rho(1-2\rho)N \tag{4.93}$$
$$\overline{D}_u \leq 2\rho(1-\rho)N \tag{4.94}$$
$$0 \leq N \tag{4.95}$$

It can be shown that (4.93) is redundant with respect to (4.92). Therefore, if $N \leq M$ and $\rho \leq \frac{1}{2}$, we have $\overline{D}_u = 2\rho(1-\rho)N$ which is equal to what is presented in Table 4.1.

2) Considering $\rho > \frac{1}{2}$, we have:

$$\overline{D}_u \leq 2\rho^2\beta_2 + 2\rho(1-\rho)N \tag{4.96}$$

$$\overline{D}_u \leq 2\rho(2\rho-1)\beta_2 + 2\rho(1-\rho)N \tag{4.97}$$

$$\beta_2 \leq N \tag{4.98}$$

$$2\beta_2 \leq N \tag{4.99}$$

$$\beta_2 \geq 0 \tag{4.100}$$

It can be shown that (4.96) and (4.98) are redundant with respect to (4.97) and (4.99), respectively. Therefore, we can eliminate them. Next, we combine inequities (4.96)-(4.100) such that they cancel out β_2.

$$(4.97) + \rho(2\rho-1) \times (4.99) \tag{4.101}$$

$$2 \times (4.100) + (4.99) \tag{4.102}$$

Therefore, we have:

$$\overline{D}_u \leq \rho N \tag{4.103}$$

$$0 \leq N \tag{4.104}$$

Therefore, if $N \leq M$ and $\rho > \frac{1}{2}$, we have $\overline{D}_u = \rho N$ which is equal to what is presented in Table 4.1.

Chapter 5

Concluding Remarks

5.1 Summary of Contributions

In this dissertation, we have studied three network setups in which the channel state information is only available at the destination.

In Chapter 2, we have considered a multi-hop network in which the channel gains of each hop changes quasi-statistically from one transmission block to the other. It is assumed that the knowledge of the channel gain of each hop is available only at its corresponding receiver, and relays are not capable of data buffering over multiple transmission blocks. Considering a single-antenna case, we have proposed a scheme based on multilayer data transmission in conjunction with decode and forward relaying. For this transmission scheme, we have formulated the statistical average of the received rate per channel use at the final destination as an optimization problem. We have then introduced an algorithm to solve the introduced optimization problem and determine the optimal parameters of the multilayer codes of each node. The optimality of the proposed scheme for the two-hop networks is also discussed in this chapter.

In Chapter 3, instead of multi-hop settings, we have studied a single-hop network (with a quasi-statistic fading channel) in which both source and destination are equipped with multiple antennas. Similar to Chapter 2, we have assumed that the CSI is not available at the source and the objective is to design a multilayer code such that it maximizes the statistical average of the received rate per channel use at the destination. This study is limited to the scenarios in which the destination is only able to perform successive decoding. In this chapter, first, we have proposed a general design rule for construction of the optimal multilayer codes for multiple-input multiple-output systems. The chapter also presents an algorithm that uses the proposed design rule to determine the parameters of the multilayer code. The superiority of the proposed scheme compared to the previously known scheme is also shown in this chapter.

Finally, in Chapter 4, we have analyzed a K-user random access network in which users and the AP are equipped with M and N antennas, respectively. In this network, although the channel gains are fixed and known at the users, the user still does not know the complete channel state. It is due to the fact that in each time slot, there could be more than one active user that produce interference over each other. The objective is to find a way to combat with the ambiguity regarding the channel state and propose a scheme that achieves the optimal value of the network average DoF (introduced in Section 4.1). This chapter also proposes a transmission scheme that maximizes the network average DoF for all settings of a two-user random access network, i.e., for all values of M, N, and ρ where ρ denotes the probability of node activation in a time-slot. Generalizing the result to a K-user random access network, we have presented an upper-bound for the network average DoF. Furthermore, we have proved that this upper-bound is tight, if either of these conditions holds: 1) $N \leq M$, or 2) $KM \leq N$, or 3) $M < N < KM$, and ρ is larger than a threshold. The optimality of the proposed upper-bound is also proved for some specific network examples in which ρ is smaller than the threshold and $M < N < KM$. The achievability schemes of these cases are based on the idea of multi-stream data transmission and vector interference alignment.

5.2 Future Research Directions

This dissertation can be extended in different directions. Some of them are briefly described in the following.

- In Chapter 2, we have studied a multi-hop network in which all relays have a maximum transmission power constraint. To extend the results, we can relax this constraint to an average power constraint. Based on the average power constraint, a relay is able to transmit with lower power in cases that its input rate (the amount of data that the relay is decoded from the multilayer code of the previous hop) is low, and instead use higher transmit power in cases that the input rate is higher, i.e., the relay saves power in some transmission blocks and uses it in some other transmission blocks. The optimization problem in this case is the extended version of the optimization problem introduce in (2.5)-(2.8), where we should also find the optimum transmit power of each relay (condition on its input rate), while the average transmit power is constant. In [60], we present some initial studies on this subject.

- In Chapter 3, we have discussed the optimal design of multilayer codes for single-hop MIMO networks. Inspired by the results of Chapter 2, this work can be extended to the optimal design of the multilayer codes, when they are transmitted over multi-hop MIMO networks.

- As pointed out in Chapter 4, the optimal network average DoF of a K-user MIMO random access networks is derived except for the cases that $M < N < KM$ and ρ is smaller than threshold ρ_{j_2-1}, where ρ denotes the user activation probability, and M, N represent the number of antennas available at each user and the AP, respectively. Threshold ρ_{j_2-1} is also defined in Corollary 4.1. In this chapter, we have proposed a transmission scheme that achieves the network average DoF even in networks where ρ is smaller than threshold ρ_{j_2-1} and $M < N < KM$. However, as shown in Example 3 of Section 4.4.4, the proposed scheme is not applicable for all network settings. Therefore, an important extension of the work is to close the gap between the upper-bound and lower-bound of the network average DoF by either finding a new achievability scheme or by refining the upper-bound in these cases.

- In the study of random access network in Chapter 4, the AP is either able to decode a message or it discards all messages that are collided with each other. One interesting idea is that instead of discarding the whole collided messages, the AP stores the received signal (without decoding it) and tries to decode it using the information that it receives at a later time. This is the same notion that is used in network coding literature. The technique is that even though we might not be able to decode a message independently, we might be able to decode it when we receive some information in the future, or this undecodable message might be useful for decoding a message that we will receive at a later time. Transmission schemes that use this idea potentially have this capability to improve the throughput of the network. Therefore, this can be a noteworthy extension of the results of Chapter 4.

References

[1] E.C. Van Der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[2] T. Cover and A.E.L. Gamal, "Capacity theorems for the relay channel," *Information Theory, IEEE Transactions on*, vol. 25, no. 5, pp. 572–584, 1979.

[3] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Information Theory, 2000. Proceedings. IEEE International Symposium on*. IEEE, 2000, p. 22.

[4] S. Zahedi, M. Mohseni, and A. El Gamal, "On the capacity of AWGN relay channels with linear relaying functions," in *IEEE International Symposium on Information Theory*, 2004, pp. 399–399.

[5] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1545–1561, 2006.

[6] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *Information Theory, IEEE Transactions on*, vol. 51, no. 3, pp. 765–779, 2005.

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

[8] S.O. Gharan, A. Bayesteh, and A.K. Khandani, "Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay network," *Information Theory, IEEE Transactions on*, vol. 57, no. 4, pp. 2070–2082, 2011.

[9] J.N. Laneman, D.N.C. Tse, and G.W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *Information Theory, IEEE Transactions on*, vol. 50, no. 12, pp. 3062–3080, 2004.

[10] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4152–4172, 2005.

[11] S. Yang and J.C. Belfiore, "Towards the optimal amplify-and-forward cooperative diversity scheme," *Information Theory, IEEE Transactions on*, vol. 53, no. 9, pp. 3114–3126, 2007.

[12] K. Sreeram, S. Birenjith, and P.V. Kumar, "Multi-hop cooperative wireless networks: diversity multiplexing tradeoff and optimal code design," *Arxiv preprint arXiv:0802.1888*, 2008.

[13] S.O. Gharan, A. Bayesteh, and A.K. Khandani, "Diversity-multiplexing tradeoff in multi-antenna multi-relay networks: Improvements and some optimality results," *Arxiv preprint arXiv:0912.3978*, 2009.

[14] S. Pawar, A.S. Avestimehr, and D.N.C. Tse, "Diversity-multiplexing tradeoff of the half-duplex relay channel," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*. IEEE, 2008, pp. 27–33.

[15] S. Shamai, "A broadcast strategy for the Gaussian slowly fading channel," in *Information Theory. 1997. Proceedings., 1997 IEEE International Symposium on*. IEEE, 1997, p. 150.

[16] A. Steiner and S. Shamai, "Achievable rates with imperfect transmitter side information using a broadcast transmission strategy," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 3, pp. 1043–1051, 2008.

[17] T.T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *Communications, IEEE Transactions on*, vol. 55, no. 4, pp. 820–829, 2007.

[18] S. Ekbatani, F. Etemadi, and H. Jafarkhani, "Transmission over slowly fading channels using unreliable quantized feedback," in *Data Compression Conference, 2007. DCC'07.* IEEE, 2007, pp. 353–362.

[19] S.R. Mirghaderi, A. Bayesteh, and A.K. Khandani, "On the capacity of wireless multicast networks," *Arxiv preprint arXiv:0805.4248*, 2008.

[20] A. Steiner and S. Shamai, "Multi-layer broadcast Hybrid-ARQ strategies," in *Communications, 2008 IEEE International Zurich Seminar on.* IEEE, pp. 148–151.

[21] S. Shamai and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *Information Theory, IEEE Transactions on*, vol. 49, no. 10, pp. 2617–2635, 2003.

[22] A. Steiner and S. Shamai, "Multi-layer broadcasting over a block fading MIMO channel," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 11, pp. 3937–3945, 2007.

[23] V. Pourahmadi, A.S. Motahari, and A.K. Khandani, "Infinite-layer codes for single-user slowly fading MIMO channels," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on.* IEEE, 2009, pp. 1525–1529.

[24] A. Steiner and S. Shamai, "Single-user broadcasting protocols over a two-hop relay fading channel," *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 4821–4838, 2006.

[25] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, 2003.

[26] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, no. 3, pp. 301–317, 2001.

[27] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1103–1120, 2007.

[28] H. Zhang, Y. Zheng, M.A. Khojastepour, and S. Rangarajan, "Scalable video streaming over fading wireless channels," in *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE.* IEEE, 2009, pp. 1–6.

[29] T. Schierl, T. Stockhammer, and T. Wiegand, "Mobile video transmission using scalable video coding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1204–1217, 2007.

[30] S. Jaeckel and V. Jungnickel, "Scalable video transmission in multiantenna broadcast systems," in *Wireless Conference, 2008. EW 2008. 14th European.* IEEE, 2008, pp. 1–6.

[31] R. D'Haenens, J. Doggen, D. Bakker, and T. Dams, "Transmitting scalable video with unequal error protection over 802.11 b/g," in *IEEE International Conference on Wireless & Mobile Computing, Networking & Communication.* IEEE, 2008, pp. 638–643.

[32] C. Atici and M.O. Sunay, "High data-rate video broadcasting over 3G wireless systems," *Broadcasting, IEEE Transactions on*, vol. 53, no. 1, pp. 212–223, 2007.

[33] N. Abramson, "THE ALOHA SYSTEM: another alternative for computer communications," in *Proceedings of the November 17-19, 1970, fall joint computer conference.* ACM, 1970, pp. 281–285.

[34] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: part I–carrier sense multiple-access modes and their throughput-delay characteristics," *Communications, IEEE Transactions on*, vol. 23, no. 12, pp. 1400–1416, 1975.

[35] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the gaussian multiple-access channel," *Information Theory, IEEE Transactions on*, vol. 42, no. 2, pp. 364–375, 1996.

[36] T.M. Cover, J.A. Thomas, J. Wiley, et al., *Elements of information theory*, vol. 6, Wiley Online Library, 1991.

[37] J. Cao and E.M. Yeh, "Asymptotically optimal multiple-access communication via distributed rate splitting," *Information Theory, IEEE Transactions on*, vol. 53, no. 1, pp. 304–319, 2007.

[38] S. Shamai, "A broadcast approach for the multiple-access slow fading channel," in *Information Theory, 2000. Proceedings. IEEE International Symposium on*. IEEE, 2000, p. 128.

[39] M. Médard, J. Huang, A.J. Goldsmith, S.P. Meyn, and T.P. Coleman, "Capacity of time-slotted ALOHA packetized multiple-access systems over the awgn channel," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 2, pp. 486–499, 2004.

[40] P. Minero and D.N.C. Tse, "A broadcast approach to multiple access with random states," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*. IEEE, pp. 2566–2570.

[41] P. Minero, D.N.C. Tse, and M. Franceschetti, "A broadcast approach to random access," in *Information Theory Workshop, 2009. ITW 2009. IEEE*. IEEE, pp. 615–619.

[42] P. Minero, M. Franceschetti, and D.N.C. Tse, "Random access: An information-theoretic perspective," *Arxiv preprint arXiv:0912.3264*, 2009.

[43] V. Pourahmadi, A. Bayesteh, and A.K. Khandani, "Multilevel coding strategy for two-hop single-user networks," in *Communications, 2008 24th Biennial Symposium on*. IEEE, 2008, pp. 115–119.

[44] V. Pourahmadi, A. Bayesteh, and A.K. Khandani, "Multilayer coding over multi-hop single-user networks," *Submitted to Information Theory, IEEE Transactions on*, 2009.

[45] I. Geldfand and S. Fomin, *Calculus of Variations*, Printice-Hall Inc., 1963.

[46] Gilbert Strang, "Calculus of Variations," *MIT OpenCourseware*, Spring 2006.

[47] P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise (corresp.)," *Information Theory, IEEE Transactions on*, vol. 20, no. 2, pp. 279–280, 1974.

[48] V. Pourahmadi, A. Motahari, and A.K. Khandani, "Multilayer codes for broadcasting over quasi-static fading MIMO networks," *To be shortly submitted to Wirless Communaction, IEEE Transactions on*, 2011.

[49] A.M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 3, pp. 662–671, 2006.

[50] K.B. Petersen and M.S. Pedersen, "The matrix cookbook," *Technical University of Denmark*, pp. 7–15, 2008.

[51] E. Telatar, "Capacity of multi-antenna gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[52] B. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in the sum rate," in *in Proc. 40th Annual Allerton Conf. Communications, Control and Computing*. Citeseer, 2002.

[53] S. Toumpis and A.J. Goldsmith, "Capacity regions for wireless Ad-hoc networks," *Wireless Communications, IEEE Transactions on*, vol. 2, no. 4, pp. 736–748, 2003.

[54] X. Zhao, T. Rautiainen, K. Kalliola, and P. Vainikainen, "Path-loss models for urban microcells at 5.3GHz," *Antennas and Wireless Propagation Letters, IEEE*, vol. 5, no. 1, pp. 152–154, 2006.

[55] V. Pourahmadi, S.A. Motahari, and A.K. Khandani, "Degrees of freedom of two-user MIMO networks with random medium access control mechanism," in *Information Theory, 2011. ISIT 2011. IEEE International Symposium on*, 2011, pp. 1886–1890.

[56] V. Pourahmadi, A. Motahari, and A.K. Khandani, "Degrees of freedom for a class of MIMO networks with random medium access controller," *To be shortly submitted to Telecommunications, European Transactions on*, 2011.

[57] M.A. Maddah-Ali, A.S. Motahari, and A.K. Khandani, "Communication over mimo x channels: Interference alignment, decomposition, and performance analysis," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3457–3470, 2008.

[58] A.S. Motahari, S.O. Gharan, M.A. Maddah-Ali, and A.K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *Arxiv preprint arXiv:0908.2282*, 2009.

[59] L. Zheng and D.N.C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *Information Theory, IEEE Transactions on*, vol. 49, no. 5, pp. 1073–1096, 2003.

[60] V. Pourahmadi, "On the optimal design of wireless relay network," Tech. Rep., (Ph.D. Thesis Proposal), University of Wateerloo, [Online] Available: http://shannon2.uwaterloo.ca/ vpourahm/Publications.html, April 2008.