SPECTRALLY-EFFICIENT PROTOCOLS FOR WIRELESS RELAY NETWORKS

by

Ramy M. Tannious

APPROVED BY SUPERVISORY COMMITTEE:

____

Dr. Aria Nosratinia, Chair

____

Dr. Naofal Al-Dhahir

____

Dr. John P. Fonseka

____

Dr. Hlaing Minn

To my mother and the loving memory of my father

SPECTRALLY-EFFICIENT PROTOCOLS FOR WIRELESS RELAY NETWORKS

by

RAMY M. TANNIOUS, B.S.E.E., M.S.E.E.

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

December 2008

PREFACE

This dissertation was produced in accordance with guidelines which permit the inclusion as part of the dissertation the text of an original paper or papers submitted for publication. The dissertation must still conform to all other requirements explained in the "Guide for the Preparation of Master's Theses and Doctoral Dissertations at The University of Texas at Dallas." It must include a comprehensive abstract, a full introduction and literature review and a final overall conclusion. Additional material (procedural and design data as well as descriptions of equipment) must be provided in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported.

It is acceptable for this dissertation to include as chapters authentic copies of papers already published, provided these meet type size, margin and legibility requirements. In such cases, connecting texts which provide logical bridges between different manuscripts are mandatory. Where the student is not the sole author of a manuscript, the student is required to make an explicit statement in the introductory material to that manuscript describing the student's contribution to the work and acknowledging the contribution of the other author(s). The signatures of the Supervising Committee which precede all other material in the dissertation attest to the accuracy of this statement.

ACKNOWLEDGEMENTS

Many people deserve my gratitude for supporting me during my graduate studies and the development of this dissertation. The words mentioned below are not enough to express my appreciation.

My advisor, Aria Nosratinia, deserves a special acknowledgment for several reasons. He taught me how to conduct the highest quality research. He supported me financially as well as provided guidance whenever I needed. I cannot forget the late nights we spent together attempting to make a deadline for submitting a technical paper. I am immensely thankful for his investment in me. I would also like to thank my dissertation supervisory committee of Naofal Al-Dhahir, John Fonseka, and Hlaing Minn for their time and their useful comments on my work.

I want to sincerely thank all my current and former colleagues at the Multimedia Communications Laboratory: Negar Bazargani, Mohamed Abou El-Seoud, Ali Tajer, Ahmadreza Hedayat, Mohammad Janani, Shahab Sanayei, Annie Le, Frank Namin, Vijay Suryavanshi, Kamtorn Ausavapattanakun, Ahmed Mehana, Yang Li, Thuy Nguyen, and Youngsam Yoon for their friendship. In addition, I would like to thank my Egyptian friends in the Dallas area. We have spent memorable times together, and I am fortunate to know each one of you.

Coming to study in The United States is not only about academic pursuits. It is also about the American experience. I appreciate the efforts of the International Students Services Office at UT Dallas for helping me live this experience. I would like to extend many thanks to Cristen Casey and Anne Caldwell. I was also fortunate to be a founding member of Global Village, a student organization that encompasses all international students. Thanks go to my "global villager" friends.

I cannot forget to thank the people who instilled in me many aspects of my personality during my formation days in Egypt. My deep appreciation goes to my teachers and mentors at Victoria College, Alexandria and Alexandria University.

Most importantly, I would like to thank my mom and my late dad. Their loving support, the sacrifices they made, and their constant encouragement throughout my life cannot be expressed in words. It was difficult to lose my dad during the development of this dissertation, but my dad's wisdom and the principles he instilled in me still guide my way. My love and thanks will be everlasting. I'm very grateful to my brother, Rafik as well. He has been instrumental in developing my character. Being the elder child, he has always been a role model. I am grateful for his advice and love.

November 2008

SPECTRALLY-EFFICIENT PROTOCOLS FOR WIRELESS RELAY NETWORKS

Publication No. ___________________

Ramy M. Tannious, Ph.D.
The University of Texas at Dallas, 2008

Supervising Professor: Dr. Aria Nosratinia

In this dissertation, several signaling protocols in relay networks are proposed and analyzed to improve the throughput and/or reliability of wireless networks.

In the first part of this dissertation, a new relay channel model is proposed and analyzed that supports a mixture of relayed and non-relayed messages. This model not only subsumes previous relay models, but also allows for an analysis of the tradeoff between relayed and private messages. The second part of this dissertation has the aim of improving the spectral efficiency of half-duplex relay systems, in multi-relay scenarios. This is accomplished by the design of methods and protocols employing opportunistic relay selection with limited feedback. The third and final part of this dissertation investigates the role of a multi-antenna relay in multi-user networks with cross-interference, also known as interference networks. Several strategies at the relay are investigated to manage the interference at the receivers and improve the overall quality of communication.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

Wireless communication is attractive for both the consumer as well as service providers. On one hand, it provides an ubiquitous mode of communications. On the other hand, it requires lower infrastructure costs and easier implementation compared to wire-line networks. However, communication over the wireless medium poses challenges for engineers and scientists. The wireless channel suffers from path loss, scattering, and shadowing effects. These effects are sometimes collectively called fading, and are detrimental to wireless communications. Furthermore, interference arises as a serious issue in a wireless networks due to the broadcast nature of wireless transmissions.

Fortunately, over time, new paradigms and approaches have been devised to combat the fading and the interference effects. One of the key recent directions in wireless communications has been to exploit, rather than suppress, the unique features of the wireless channel such as fading and interference. For example, the paradigm of opportunistic communication is based on transmitting at the peak of the fading channel. Multiple-input-multiple-output (MIMO) communications is enabled by the fading in rich scattering wireless environments. Likewise, the relaying in wireless networks takes advantage of the broadcast nature of wireless transmissions. These new developments give further hope that future wireless networks will provide higher data rates with improved reliability.

Relay networks are communication networks where one or more nodes assist the communication of a source-destination pair or a group of pairs. The simplest form of a relay network is the three node network known as the relay channel, which was introduced by van der Muelen in 1971 [1]. The seminal work of Cover and El Gamal further investigated the relay channel [2], establishing the capacity in several special cases. In the most general case, the capacity of the relay channel has remained an open problem for almost four decades.

Research in relay networks, with a few exceptions, remained dormant after the work of Cover and El Gamal, until the work of Sendonaris et. al [3] on cooperative diversity generated renewed interest in relaying and cooperation. Cooperative communication, which is a form of mutual relaying between the network nodes, enlarges the capacity region of a fading multiple-access channel, improves signals reliability, and extends the coverage area [3].

After the appearance of [3], a large body of research on relaying has been produced, and many aspects of relaying have been investigated. It is now generally accepted that the improvements made possible by relays can translate into tangible benefits for wireless communication systems, and that continued study of relaying remains of importance from both a theoretical as well as a practical point of view. This dissertation makes contributions to several key problems in this area.

1.2 Motivation and Objectives

The common theme of the research in this dissertation is to improve the rate of wireless relay networks. This can refer to the capacity region of a network, the spectral efficiency, or the degrees of freedom. These notions will be rigorously defined in the next chapter.

In the recent past, research in relays has focused on achieving better diversity via the additional signal paths provided by the relays. However, much less attention has been paid to achieving good spectral efficiency while maintaining (near) optimal spatial diversity in a relay network. The main challenge in this task is due to the repetition of information by the half-duplex relays, i.e., the nodes cannot transmit and receive at the same time/frequency.

This dissertation addresses the spectral efficiency in various relay network architectures. We start with the three node relay problem. Then, the multiple-relay network is considered, where a set of nodes are willing to help the communications of a source-destination pair. We then consider relaying in interference networks, where multiple source-destination pairs attempt to communicate together and they are assisted by a dedicated relay.

1.3 Contributions and Outline

Chapter 2 provides the mathematical background that is used in the following chapters. It collects some techniques and ideas used in multi-terminal information theory. It also defines some of the performance measures used in this dissertation.

Chapter 3 presents the first contribution of this dissertation. A new relay channel model is proposed and analyzed for the simple three node network. Simultaneous communications between the three nodes is proposed to maximize the spectral efficiency of the network. This leads to a channel model where relay, broadcast, and multiple-access components are featured. We derive inner and outer bounds on the capacity region for this channel. We further show that many of the existing capacity results for the three-node relay network can be recovered as special cases of the results obtained in this chapter.

Chapter 4 studies a multiple-relay network, where it is assumed that a group of nodes are willing to assist the source to communicate its message to the destination. We take an opportunistic approach for the relaying process via a relay selection framework. Two models are considered: where the direct link exists, and where it is absent. Two corresponding novel relay selection protocols are devised. Under a block-fading model, outage probability and diversity-multiplexing tradeoff are the performance limits of interest. The proposed protocols exhibit excellent performance at high spectral efficiencies compared to all existing protocols in the literature, while also requiring minimal overhead. Moreover, a novel concept of unequal error protection via a family of DMT curves is produced in our work.

The last contribution presented in this dissertation appears in Chapter 5. We study the inclusion of a MIMO relay to assist communications in an interference network. Thus, this chapter considers the simultaneous communication of several source-destination pairs. Several coding strategies are proposed and analyzed for this network model. Achievable bounds on the sum-rate capacity (network throughput) are obtained. Then, the achievable (lower bound) degrees of freedom (DOF) of this network for different signaling schemes is obtained. We then obtain an upper bound on the DOF and finally establish the exact DOF of the network under study.

CHAPTER 2

PRELIMINARIES

This chapter is dedicated to review some of the information-theoretic concepts and performance measures used throughout this dissertation. Some of the definitions and the proofs of theorems that are discussed below appear in [4].

2.1 Typicality Decoding

We use random codes in our analysis of various signaling schemes. Decoding of these codes at the receiver is based on the idea of typicality known as typicality decoding. Hence, we start this chapter by defining weak and strong typicality and their differences and then consider the definition of joint typicality.

Theorem 1 *Asymptotic Equipartition Property (AEP) If X_1, X_2,$\cdots X_n$ are i.i.d $\sim p(x)$, then*

$$-\frac{1}{n}\log p(X_1, X_2, \cdots, X_n) \rightarrow \quad H(X) \quad \textit{in probability} \tag{2.1}$$

where $H(X)$ is the entropy of X.

Definition 1 *(Weak Typicality) The typical set A_ϵ with respect to $p(x)$ is the set of sequences $x^n = (x_1, x_2, ...x_n) \in \mathcal{X}^n$ with the property*

$$2^{-n(H(X)+\epsilon)} \leq p(x_1, x_2, ...x_n) \leq 2^{-n(H(X)-\epsilon)} \tag{2.2}$$

As a consequence of the AEP, it can be shown that the typical set has the following properties. It has probability nearly 1, all elements belonging to the typical set are nearly equiprobable and the cardinality of the typical set is almost 2^{nH}.

There is another form of typicality known as strong typicality. Strong typicality put limits on the relative occurrence of the symbols within a sequence not on the probability of the sequence.

Definition 2 *(Strong Typicality) The strongly typical set A_ϵ with respect to $p(x)$ is the set of sequences $x^n = (x_1, x_2, ...x_n) \in \mathcal{X}^n$ with the property*

$$\left| \frac{1}{n} N(a; x^n) - p(a) \right| < \frac{\epsilon}{|\mathcal{X}|} \tag{2.3}$$

where $N(a; x^n)$ is the number of occurrences of the symbol a in the sequence x^n and $|\mathcal{X}|$ is the size (cardinality) of the alphabet $\mathcal{X}$.

Strong typicality implies weak typicality while the opposite is not always true.

Definition 3 *(Joint Typicality) The set A_ϵ of jointly typical sequences (x^n, y^n) with respect to $p(x, y)$ is the set of length n sequences with the properties*

$$\left| -\frac{1}{n} p(x^n) - H(X) \right| < \epsilon \tag{2.4}$$

$$\left| -\frac{1}{n} p(y^n) - H(Y) \right| < \epsilon \tag{2.5}$$

$$\left| -\frac{1}{n} p(x^n, y^n) - H(X, Y) \right| < \epsilon \tag{2.6}$$

where

$$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i) \tag{2.7}$$

It is easy to show that the probability of any randomly chosen typical pair hitting a jointly-typical pair is $\sim 2^{-nI(X;Y)}$. The notion of joint typicality is the basis of the typicality decoder. The typicality decoder is in general suboptimal but is simple to analyze and achieves all rates below capacity.

2.2 Some Network Information Theory Techniques

Most of the significant results in network information theory hinges upon the use of novel techniques in the construction of the signaling (coding) scheme. Most of the currently used techniques have been introduced during the 1970's. We review some of these tools that are used in our proofs throughout the dissertation.

2.2.1 Random Binning

The important idea of random binning is now introduced. The idea is that the encoder indexes at random all the sequences in the codebook (typical and atypical). The range of indexes is usually smaller then the total number of available sequences. Then, instead of sending the desired codeword itself, the encoder sends only the index of this sequence. The set of sequences which have the same index are said to form a bin. The number of bins determines the rate of the encoder. For decoding, the index of the bin is revealed to the decoder and it looks for one and only one typical sequence in that bin. This powerful idea can be extended to several sequences transmitted by same encoder as in the broadcast channel problem or transmitted by different sources as in distributed source-coding problems. Here, the sets of two sequences (or more) are divided into bins in such away that the pair of indexes specifies a product bin. The decoder in turn searches for a one and only one jointly-typical pair in this product bin.

2.2.2 Superposition Block Markov Encoding

Superposition block Markov encoding is an idea that was devised by Cover and Leung [5]. It gained a lot of attention afterwards as it was the key idea in achieving the capacity of the relay channel in some special cases. This idea is relevant when two or more transmitters cooperate in sending information to a receiver. Superposition block Markov encoding for two cooperating transmitters works as follows. Transmission occurs in blocks. At each block, the main transmitter sends high rate information that does not allow correct decoding by the destination while another node might be able to decode this information reliably. However, in the next block, both transmitters cooperate by sending "resolution information" that removes the uncertainty of the receiver about the data in the previous block. Moreover, in order to keep this cooperation possible in the next blocks, the main transmitter superimposes a fresh information about the message of the current block, hence the "superposition" part in the name of the technique. When the sizes of the two codebooks used by the source and the cooperating transmitter are the same, the encoding is called regular encoding, otherwise it is called irregular encoding.

2.2.3 Backward Decoding

Backward decoding is a decoding technique that was introduced by Willems [6]. It can simplify the decoding of a superposition block Markov encoded signal. All blocks of transmission are collected at the destination until the last block of transmission is completed. Decoding starts from the last block which contains only "resolution information". The decoder then moves to the previous blocks sequentially and decodes the information in a backward manner. Hence, the name "backward decoding". The obvious issues with this type of decoding is that it causes high decoding delay and requires large buffering.

2.3 Some Performance Measures for the Rate of Information

Moving on from ideas and coding techniques, several measures for performance limits used in this dissertation are now defined.

2.3.1 Shannon Capacity

Shannon capacity represents the maximum rate of transmission in a noisy channel with almost zero probability of error. Consider the following linear time invariant channel model with an input power constraint P:

$$y = hx + z \tag{2.8}$$

where h represents the random channel gain between the transmitter and receive and z is the thermal noise at receiver modeled as AWGN$\sim \mathcal{N}(0, N_o)$. Conditioned on knowing h, the channel follows a Gaussian model and its Shannon capacity is given by:

$$C(h, \rho) = \frac{1}{2} \log \left(1 + \rho\right) \tag{2.9}$$

achieved when $X \sim \mathcal{N}(0, P)$ and under infinite block length assumption, where $\rho = \frac{P}{N_o}$ is the signal-to-noise ratio (SNR). When h varies like in fading channels, if the codeword is long enough such that h reveals its ergodic nature within one codeword, then the Shannon (ergodic) capacity is obtained by taking the expectation of (2.9)

$$C_{ergodic}(\rho) = E[C(h, \rho)] \tag{2.10}$$

2.3.2 Outage Probability and Outage Capacity

When the codeword length spans only one channel realization whose value is unknown at the transmitter, there exist a non-zero probability of decoding error and outage occurs. In this case, the Shannon capacity in its strict sense is zero. One can instead define an outage probability P_{out} performance limit which at high SNR and long enough block length was proved to tightly lower bound the error probability [7]. For a given rate R, P_{out} is given by:

$$P_{out} = Pr[C(h, \rho) < R] \tag{2.11}$$

Also, one can define a corresponding outage capacity, C_ϵ. It is the rate below which the outage probability over any channel realization does not exceed ϵ. It is expressed by:

$$C_\epsilon = sup\big\{R : Pr[C(h, \rho) < R] \leq \epsilon\big\} \tag{2.12}$$

2.3.3 Diversity-Multiplexing Tradeoff

The outage probability described above provides insight into the tradeoff between error probability and transmission rate at a fixed SNR. However, formulating a closed-form outage expression is not possible in many cases. In order to circumvent this hurdle, Zheng and Tse elegantly formulated another performance measure known as the diversity-multiplexing tradeoff (DMT) [8]. The DMT provides a tradeoff between the reliability provided by a certain signaling scheme versus the rate expressed as a fraction of the AWGN channel capacity at high SNR. Communicating at fixed rate would yield the maximum reliability of a channel. However, at high SNR, any fixed rate R becomes vanishingly small with respect to the channel capacity. Thus, one would be sacrificing the spectral efficiency of the channel in order to attain the maximum diversity for the signal transmitted. Therefore, the key for the Zheng-Tse formulation is the notion of multiplexing gain r. It is expressed as:

$$r = \lim_{\rho \to \infty} \frac{R(\rho)}{log(\rho)} \tag{2.13}$$

The diversity gain d at a specific multiplexing gain r is defined as:

$$d(r) = \lim_{\rho \to \infty} -\frac{P_{out}(R(\rho))}{log(\rho)} \tag{2.14}$$

2.3.4 Degrees of Freedom

The degrees of freedom (DOF) is a characteristic of the channel. It is also known as the pre-log factor or the multiplexing gain. It is formally defined as:

$$DOF = \lim_{\rho \to \infty} \frac{C(\rho)}{\log(\rho)} \tag{2.15}$$

That is for a MIMO channel with m transmit antennas and n receive antennas, Telatar's seminal paper shows that such channel has a $DOF = min(m, n)$ [9].

CHAPTER 3

RELAY CHANNEL WITH PRIVATE MESSAGES

3.1 Introduction

The three-terminal relay channel was proposed by van der Meulen [1] and thoroughly studied by Cover and El Gamal [2]. The relay node in this model has no role aside from relaying, in particular, it is neither a source nor a sink of information.

When dedicated relays are unavailable, relaying must be done by network nodes that are also a source/sink of data. Thus, one is interested in the network performance limits when a relay must handle both relayed messages, as well as their own (private) messages. A representative channel model is shown in Fig. 3.1, which we call *relay channel with private messages* (RCPM). This is a network with both point-to-point as well as relayed links, a generalization of the traditional relay channel.

In the following we mention some of the literature that is most directly related to the present discussion. Liang and Veeravalli [10] studied a cooperative relay

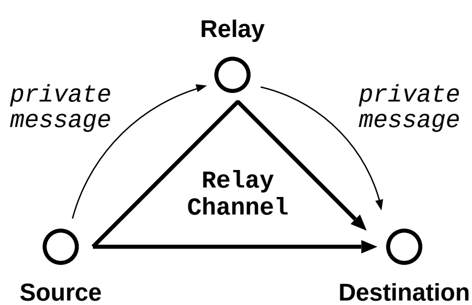

Figure 3.1. Relay channel with private messaging.

broadcast channel. The bounds in this work are further improved by Liang and Kramer [11]. Reznik et al. [12] address a similar problem with multiple relays. Dabora and Servetto [13] study a broadcast channel with a secure cooperative link between the receivers. Lai et al [14] study a half-duplex, fading, three-way channel.

The contributions of this chapter are described as follows. In our analysis of RCPM we use new combinations of coding strategies inspired by the MAC channel with generalized feedback, and Marton's approach to the broadcast channel. We derive achievable rates for the discrete memoryless and Gaussian RCPM based on decode-and-forward and compress-and-forward schemes, and outer bounds for the DMC case. The discrete memoryless and Gaussian RCPM generalize their counterparts in the original relay channel and relay-broadcast channels. Numerical results are provided that give insights into the trade-offs between private messaging and relayed messaging in this hybrid three-node network.

Throughout this chapter, the techniques of regular encoding and backward decoding defined in Chapter 2 are used. Backward decoding has been used previously for degraded relay channel in [15] and more recently for the general relay channel with partial decode-and-forward in [16]. As a by-product of our work, we demonstrate that backward decoding does not improve the achievable rate of a non-degraded relay channel employing compress-and-forward scheme.

3.2 Definitions and System Model

In this chapter, X, $\mathcal{X}$, and $||\mathcal{X}||$ denote a random variable, its range, and cardinality. $A_\epsilon^{(n)}(X)$ denotes the ϵ-typical set according to X, in the strong or weak sense. Deterministic scalars and vectors are shown by lower case and lower-case bold-face letters. We further define $X_t^i \triangleq (X_{t,1}, X_{t,2}..., X_{t,i})$, the capacity function

$\mathcal{C}(x) = \frac{1}{2}\log_2(1+x)$, and for usage in convex combinations, we define $\bar{x} = 1 - x$.

Definition 4 *A* relay channel with private messages *consists of a channel input alphabet* $\mathcal{X}_1$*, a relay input alphabet* $\mathcal{X}_2$*, two channel output alphabets* $\mathcal{Y}_2$ *and* $\mathcal{Y}_3$*, and a probability transition function* $p(y_2, y_3|x_1, x_2)$*, where* x_1*,* x_2 *denote source and relay inputs, respectively, while* y_2 *and* y_3 *denote the outputs at the relay and destination nodes, respectively.*

Definition 5 *A* $\left(\left(2^{nR_{12}}, 2^{nR_{23}}, 2^{nR_{13}}\right), n\right)$ *code for the relay channel with private messages consists of the following :*

- *Three sets of integers,* $\mathcal{W}_{12} = \left\{1, 2, ..., 2^{nR_{12}}\right\}$*,* $\mathcal{W}_{23} = \left\{1, 2, ..., 2^{nR_{23}}\right\}$*, and* $\mathcal{W}_{13} = \left\{1, 2, ..., 2^{nR_{13}}\right\}$*.*
- *An encoder,*

$$X_1 \ : \ \mathcal{W}_{12} \times \mathcal{W}_{13} \rightarrow \mathcal{X}_1^n$$

- *A set of relay functions* $\{f_i\}_{i=1}^n$*,*

$$x_{2,i} = f_i(y_{2,1}, ..., y_{2,i-1}, w_{23}), \qquad 1 \leq i \leq n$$

- *Two decoding functions,*

$$\begin{aligned} d_1 &: \mathcal{Y}_2^n \rightarrow \mathcal{W}_{12} \\ d_2 &: \mathcal{Y}_3^n \rightarrow \mathcal{W}_{13} \times \mathcal{W}_{23}. \end{aligned}$$

Figure 3.2 illustrates the encoding and decoding structure of different messages. The channels considered in this chapter are memoryless. Thus, the current outputs depend on the past only through present input symbols. The relay node is assumed to operate in full duplex mode and to be causal, i.e., its input is allowed to depend only on its past observations.

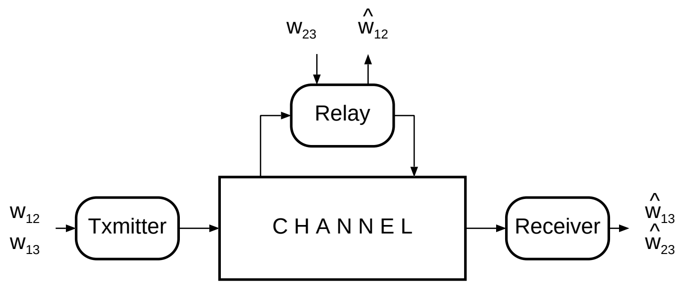

Figure 3.2. The encoding and decoding structure for relay with private messages.

Definition 6 *A relay channel with private messages is said to be degraded if its transition probability satisfies*

$$p(y_2, y_3|x_1, x_2) = p(y_2|x_1, x_2)p(y_3|y_2, x_2) \tag{3.1}$$

i.e., Y_3 is independent of X_1 conditioned on knowing Y_2 and X_2.

Definition 7 *An* AWGN *relay channel with private messages is a RCPM with a continuous input and output alphabets and independent, additive white Gaussian noise. The channel outputs of the relay and destination are given by:*

$$Y_2 = X_1 + Z_2 \tag{3.2}$$

$$Y_3 = X_1 + X_2 + Z_3 \tag{3.3}$$

where $Z_2 \sim \mathcal{N}(0, N_2)$ and $Z_3 \sim \mathcal{N}(0, N_3)$ are independent Gaussian noise.

Figure 3.3 illustrates this channel and the flow of information between the three nodes. The input sequences are subject to the power constraints $\mathcal{E}[\mathbf{x}_1^2] < P_1$ and $\mathcal{E}[\mathbf{x}_2^2] < P_2$.

Definition 8 *A degraded AWGN relay channel with private messages is a AWGN RCPM where the source and destination signals are independent given the relay input*

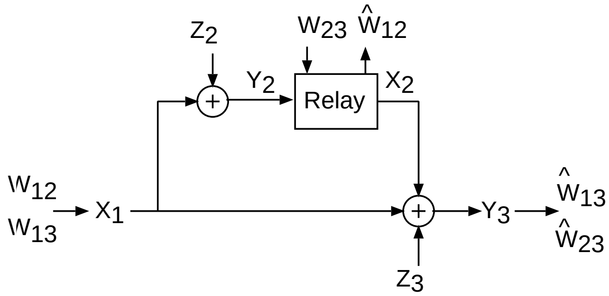

Figure 3.3. Gaussian relay channel with private messages.

and output, which is equivalent to saying that the relay knows everything that the destination knows. The channel outputs at the relay and destination are given by

$$Y_2 = X_1 + Z_2 \tag{3.4}$$

$$Y_3 = X_1 + X_2 + Z_2 + Z' \tag{3.5}$$

where Z_2 and Z' are independent zero mean Gaussian random variables with variances N_2 and $N_3 - N_2$, respectively, where $N_2 < N_3$.

Definition 9 *The average probability of error is defined as the probability that the decoded messages are different from the transmitted ones:*

$$P_e^{(n)} = P\left(\hat{W}_{12} \neq W_{12} \quad or \quad (\hat{W}_{13}, \hat{W}_{23}) \neq (W_{13}, W_{23})\right) \tag{3.6}$$

where $\hat{W}$ denotes an estimate of W. We assume that the source and relay nodes select their messages (W_{12}, W_{23}, W_{13}) independently and uniformly over $\mathcal{W}_{12} \times \mathcal{W}_{23} \times \mathcal{W}_{13}$. he probability of error at the relay and destination, respectively, is defined as:

$$P_{e,R}^{(n)} = P(\hat{W}_{12} \neq W_{12}) \tag{3.7}$$

$$P_{e,D}^{(n)} = P\left((\hat{W}_{13}, \hat{W}_{23}) \neq (W_{13}, W_{23})\right) \tag{3.8}$$

where each codeword contains n symbols. Note that by the union bound, we have:

$$\max\{P_{e,R}^{(n)}, P_{e,D}^{(n)}\} \leq P_e^{(n)} \leq P_{e,R}^{(n)} + P_{e,D}^{(n)} \tag{3.9}$$

Hence, if $P_e^{(n)} \to 0$ then both $P_{e,R}^{(n)}$ and $P_{e,D}^{(n)}$ go to zero.

Definition 10 *A rate triple (R_{12}, R_{23}, R_{13}) is said to be achievable for the relay channel with private messages if there exist a sequence of codes $\left((2^{nR_{12}}, 2^{nR_{23}}, 2^{nR_{13}}), n\right)$ with average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$.*

3.3 Achievable Rate Regions

In this section, we obtain achievable rate regions for RCPM when the relay node use the well known decode-and-forward and compress-and-forward schemes [2, Theorem 1, Theorem 6].

3.3.1 Relaying via Decode-and-Forward

Here, we assume the relay is able to fully decode both his private message and the message intended for the destination. The relay node then re-encodes the source node message W_{13} along with its message for the destination W_{23}.

Theorem 2 *The rates (R_{12}, R_{23}, R_{13}) are achievable for the discrete memoryless relay channel with private messages if*

$$R_{13} < \min\{I(U, V; Y_3), I(V; Y_2|U, X_2)\}, \tag{3.10}$$

$$R_{23} < I(X_2; Y_3|U, V), \tag{3.11}$$

$$R_{12} < I(X_1; Y_2|U, V, X_2) \tag{3.12}$$

for some joint distribution $p(u)p(v|u)p(x_1|u, v)p(x_2|u)p(y_2, y_3|x_1, x_2)$.

Proof:

The coding arguments use ideas from relay channels, broadcast channels and MAC channel with generalized feedback [2, 17, 6]. The source uses a three-level superposition block Markov encoding, while the relay uses superposition coding. Furthermore, we use the regular encoding/backward decoding techniques.

Consider a transmission period of B blocks, each of n symbols. We assume that n is sufficiently large to allow reliable decoding. The source and relay send sequences of $B-1$ messages $(W_{13}(b), W_{12}(b))$ and $W_{23}(b)$, respectively, over the channel in nB transmissions, where b denotes the block index, $b = 1, 2, ..., B-1$. The rate tuple $(R_{13}\frac{B-1}{B}, R_{12}\frac{B-1}{B}, R_{23}\frac{B-1}{B})$ approaches (R_{13}, R_{12}, R_{23}) as $B \to \infty$. In the following, we use random variables chosen according to an arbitrary probability distribution $p(u, v, x_1, x_2) = p(u)p(v|u)p(x_1|u, v)p(x_2|u)$.

<u>Random Codebook Construction</u>:

1. Generate $2^{nR_{13}}$ i.i.d. $\mathbf{u} = (u_1, u_2, ...u_n)$ sequences, each with distribution $p(\mathbf{u}) = \prod_{i=1}^{n} p(u_i)$. Label them $\mathbf{u}(w'_{13})$.

2. For each $\mathbf{u}(w'_{13})$ generate $2^{nR_{13}}$ i.i.d. $\mathbf{v}$ sequences, each with distribution $p(\mathbf{v}) = \prod_{i=1}^{n} p(v_i|u_i)$. Label them $\mathbf{v}(w'_{13}, w_{13})$.

3. For every pair $(\mathbf{u}(w'_{13}), \mathbf{v}(w'_{13}, w_{13}))$ generate $2^{nR_{12}}$ i.i.d $\mathbf{x}_1$ sequences, each with distribution
$$p(\mathbf{x_1}) = \prod_{i=1}^{n} p(x_{1,i}|u_i(w'_{13}), v_i(w'_{13}, w_{13}))$$
Label them $\mathbf{x}_1(w'_{13}, w_{13}, w_{12})$.

4. For each $\mathbf{u}(w'_{13})$ generate $2^{nR_{23}}$ i.i.d. $\mathbf{x}_2$ sequences, each with distribution $p(\mathbf{x}_2) = \prod_{i=1}^{n} p(x_{2,i}|u_i(w'_{13}))$. Label them $\mathbf{x}_2(w'_{13}, w_{23})$.

Encoding:

At Block b

1. The source sends $\mathbf{x}_1(w_{13,b-1}, w_{13,b}, w_{12,b})$, where $w_{13,b-1}$ was denoted above as w'_{13}.

2. Assuming the relay has estimated $w_{13,b-1}$ correctly from the previous block, it then sends $\mathbf{x}_2(w_{13,b-1}, w_{23,b})$.

So, the transmitted codeword pair is:

$$\begin{array}{lll} \mathbf{x}_1(1, w_{13,1}, w_{12,1}), & \mathbf{x}_2(1, w_{23,1}) & b = 1 \\ \mathbf{x}_1(w_{13,b-1}, w_{13,b}, w_{12,b}), & \mathbf{x}_2(w_{13,b-1}, w_{23,b}) & b = 2, ..., B-1 \\ \mathbf{x}_1(w_{13,B-1}, 1, 1), & \mathbf{x}_2(w_{13,B-1}, 1) & b = B. \end{array}$$

Decoding:

1. Assuming the relay has decoded $w_{13,b-1}$, it can decode $w_{13,b}$ by looking for a unique $\hat{w}_{13,b}$ such that $(\mathbf{u}(w_{13,b-1}), \mathbf{v}(w_{13,b-1}, \hat{w}_{13,b}), \mathbf{x}_2(w_{13,b-1}, w_{23,b}), \mathbf{y}_2(b))$ are jointly typical. This step can be made reliable if:

$$R_{13} < I(V; Y_2|U, X_2). \tag{3.13}$$

2. The relay decodes $w_{12,b}$ by looking for $\hat{w}_{12,b}$ such that $\mathbf{u}(w_{13,b-1})$, $\mathbf{v}(w_{13,b-1},$ $w_{13,b})$, $\mathbf{x}_1(w_{13,b-1}, w_{13,b}, \hat{w}_{12,b})$, $\mathbf{x}_2(w_{13,b-1}, \hat{w}_{23,b})$, and $\mathbf{y}_2(b)$ are jointly typical. The decoding is reliable if:

$$R_{12} < I(X_1; Y_2|U, V, X_2). \tag{3.14}$$

3. The destination waits until all blocks are received before it starts to decode. Suppose it has decoded $w_{13,b}$ in block $(b+1)$, then in block b, it looks for a unique

$\hat{w}_{13,b-1}$ such that $\mathbf{u}(\hat{w}_{13,b-1})$, $\mathbf{v}(\hat{w}_{13,b-1},\ w_{13,b})$, and $\mathbf{y}_3(b)$ are jointly typical. Upon successful decoding of $w_{13,b-1}$, the destination decodes $w_{23,b}$ by looking for a unique $\hat{w}_{23,b}$ such that $\mathbf{u}(w_{13,b-1})$, $\mathbf{v}(w_{13,b-1},\ w_{13,b})$, $\mathbf{x}_2(w_{13,b-1}, \hat{w}_{23,b})$, and $\mathbf{y}_3(b)$ are jointly typical.

This sequential decoding at the destination node clearly achieves the following rates:

$$R_{13} < I(U, V; Y_3). \tag{3.15}$$

$$R_{23} < I(X_2; Y_3|U, V). \tag{3.16}$$

The achievable rate region then follows directly from combining the previous set of equations. □

Remark 1 *The capacity of partially cooperative relay broadcast channel in [10] and also the capacity of the degraded relay channel [2] can be recovered from the above rate region. To see that, set $U = X_2$, $V = U$ and the region will reduce to that in [10, Theorem 3], also if we let $U = X_2$, $V = X_1$, we have the result of [2, Theorem 1].*

3.3.2 Relaying via Compress-and-Forward

Even when the relay node cannot fully decode the message it is supposed to relay, it can still render some help to the destination. If the channel between relay and destination had unlimited capacity, Y_2 could be transferred to the destination, however, this is often not the case. Therefore, the received signal of the relay is compressed into a new random variable, $\hat{Y}_2$, characterized by an index z —possibly via vector quantization— which is conveyed to the destination via X_2. Upon decoding z, the receiver uses $\hat{Y}_2(z)$ to resolve the uncertainty in Y_3 about the source's message. This strategy which was introduced in [2] has been commonly known as estimate-and-forward or compress-and-forward.

Theorem 3 *In the discrete memoryless relay channel with private messages, a set of private and relayed rates* (R_{12}, R_{23}, R_{13}) *is achievable if*

$$\begin{aligned} R_{13} &< I(U_1; \hat{Y}_2, Y_3 | V, X_2), & (3.17)\\ R_{12} &< I(U_2; Y_2 | X_2), & (3.18)\\ R_{13} + R_{12} &< I(U_1; \hat{Y}_2, Y_3 | V, X_2) + I(U_2; Y_2 | X_2) - I(U_1; U_2), & (3.19)\\ R_{23} &< I(V; Y_3), & (3.20) \end{aligned}$$

subject to

$$I(\hat{Y}_2; Y_2 | U_2, V, X_2) \leq I(X_2, \hat{Y}_2; Y_3 | V) \qquad (3.21)$$

where the random variables are drawn from any joint distribution

$$p(u_1, u_2)p(v)p(x_1|u_1, u_2)p(x_2|v)p(\hat{y_2}|y_2, x_2, v)p(y_2, y_3|x_1, x_2).$$

Proof:

The proof uses ideas developed for the general non-degraded relay channels, the non-degraded broadcast channels and source coding with side information at the decoder [2, 18, 19]. In particular, the source uses a Slepian-Wolf binning-type coding strategy developed by Marton for the general broadcast channels, while the relay uses superposition coding. We note that the coding used by the source corresponds to Marton's simplified region (in the Remarks of [18, Theorem 2]) where no condition on the independence of different random variables of the source's codeword is imposed and the random variable representing information decoded by both receivers is set to a constant. Again, the regular encoding/backward decoding technique is adopted. The relationship between the auxiliary random variables, channel inputs and channel outputs are depicted in Fig 3.4.

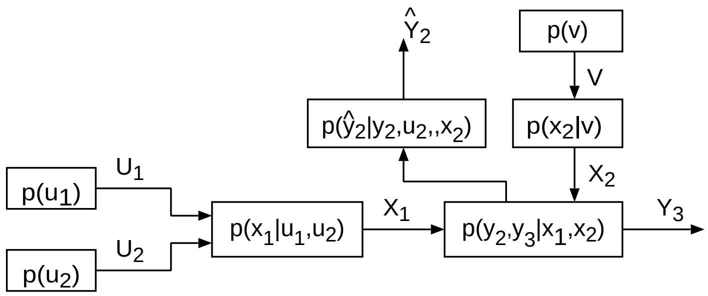

Figure 3.4. Relationship of auxiliary variables.

We consider a transmission over B blocks, each of n symbols. A sequence of $B-1$ messages $W_{13}(b)$, $W_{12}(b)$, and $W_{23}(b)$ will be sent over the channel in nB transmissions, where b denotes the block index, $b = 1, 2, ..., B-1$. The rate tuple $(R_{13}\frac{B-1}{B}, R_{12}\frac{B-1}{B}, R_{23}\frac{B-1}{B})$ approaches (R_{13}, R_{12}, R_{23}) as $B \to \infty$.

Let $A^{(n)}(U_1)$, $A^{(n)}(U_2)$, denote the set of sequences $\mathbf{u}_1$ and $\mathbf{u}_2$ that are strongly typical in U_1 and U_2, respectively, and $A^{(n)}(U_1, U_2)$ the set of strongly jointly typical sequences. Let $S^{(n)}(U_1)$ denote the set of all sequences $\mathbf{u}_1 \in A^{(n)}(U_1)$, such that $A^{(n)}(U_2|U_1)$ is nonempty. Thus, according to [20, Theorem 5.9], we have,

$$A^{(n)}(U_2|U_1) = \left\{\mathbf{u_2} \in A^{(n)}(U_2) : (\mathbf{u}_1, \mathbf{u}_2) \in A^{(n)}(U_1, U_2)\right\} \tag{3.22}$$

Similarly, define $S^{(n)}(U_2)$ for the sequence $\mathbf{u}_2$. Consider an arbitrary probability distribution $p(u_1, u_2, v, x_1, x_2, \hat{y}_2) = p(u_1, u_2)p(v)p(x_1|u_1, u_2)p(x_2|v)p(\hat{y}_2|y_2, x_2, v)$.

<u>Random Codebook Construction</u>:

1. Generate $2^{nR(U_1)}$ sequences $\mathbf{u}_1$ by drawing i.i.d. according to the probability

$$p(\mathbf{u}_1) = \begin{cases} \frac{1}{\|S^{(n)}(U_1)\|} & \mathbf{u}_1 \in S^{(n)}(U_1) \\ 0 & \text{otherwise.} \end{cases}$$

2. Generate $2^{nR(U_2)}$ sequences $\mathbf{u}_2$ by drawing i.i.d. according to the probability

$$p(\mathbf{u}_2) = \begin{cases} \frac{1}{\|S^{(n)}(U_2)\|} & \mathbf{u}_2 \in S^{(n)}(U_2) \\ 0 & \text{otherwise.} \end{cases}$$

3. Randomly assign $\mathbf{u}_1$'s into $2^{nR_{13}}$ bins and the $\mathbf{u}_2$'s into $2^{nR_{12}}$ bins.

4. For each product bin find a pair $(\mathbf{u}_1, \mathbf{u}_2)$ that belong in $A^{(n)}(U_1, U_2)$. For a sufficiently large n, random binning arguments ([4, 21]) guarantee that such a pair exist with high probability if

$$R_{13} + R_{12} < R(U_1) + R(U_2) - I(U_1; U_2) \tag{3.23}$$

5. For each product bin and its designated jointly typical pair $(\mathbf{u}_1, \mathbf{u}_2)$, generate $\mathbf{x}_1(u_1, u_2)$ according to $\prod_{i=1}^{n} p(x_{1,i}|u_{1,i}, u_{2,i})$. Label these $\mathbf{x}_1(w_{13}, w_{12})$.

6. Generate $2^{nR_{23}}$ sequences $\mathbf{v}$ by drawing i.i.d. according to the probability $p(\mathbf{v}) = \prod_{i=1}^{n} p(v_i)$. Label these $\mathbf{v}(w_{23})$.

7. For each $\mathbf{v}$ generate $2^{n\hat{R}}$ sequences $\mathbf{x}_2$ according to $\prod_{i=1}^{n} p(\mathbf{x}_{2,i}|v_i)$. Label these $\mathbf{x}_2(w_{23}, z')$.

8. For each $\mathbf{v}(w_{23})$ and $\mathbf{x}_2(w_{23}, z')$, generate $2^{n\hat{R}}$ sequences $\hat{\mathbf{y}}_2$ according to $\prod_{i=1}^{n} p(\hat{y}_{2,i}|x_{2,i}, v_i)$. Label these $\hat{\mathbf{y}}_2(w_{23}, z', z)$.

Encoding:

At Block b:

1. The source sends $\mathbf{x}_1(w_{13,b}, w_{12,b})$.

2. Assuming the relay has determined z_{b-1}, -denoted above as z'-, of the compressed signal $\hat{\mathbf{y}}_2$, it sends $\mathbf{x}_2(w_{23,b}, z_{b-1})$.

So, the transmitted codeword pair is given by:

$$\begin{array}{lll}
\mathbf{x}_1(w_{13,1}, w_{12,1}), & \mathbf{x}_2(w_{23,1}, 1) & b = 1 \\
\mathbf{x}_1(w_{13,b}, w_{12,b}), & \mathbf{x}_2(w_{23,b}, z_{b-1}) & b = 2, ..., B-1 \\
\mathbf{x}_1(1, 1), & \mathbf{x}_2(1, z_{B-1}) & b = B.
\end{array}$$

Decoding:

1. The relay decodes $w_{12,b}$ by looking for a unique $\hat{w}_{12,b}$ such that $(\mathbf{u}_2(\hat{w}_{12,b}),$ $\mathbf{x}_2(w_{23,b}, z_{b-1}), \mathbf{y}_2(b))$ are jointly typical. This can be made possible with small probability of error if

$$R_{12} < I(U_2; Y_2|X_2) \tag{3.24}$$

2. The relay can determine the index z_b of the hypothetical output $\hat{\mathbf{y}}_2(w_{23,b}, z_{b-1}, z_b)$ given it has determined z_{b-1} correctly, if $(\hat{\mathbf{y}}_2(w_{23,b}, z_{b-1}, \hat{z}_b), \mathbf{y}_2(b), \mathbf{x}_2(w_{23,b}, z_{b-1}),$ $\mathbf{v}(w_{23,b}))$ are jointly typical. Correct decision of z_b will occur with high probability if

$$\hat{R} > I(\hat{Y}_2; Y_2|U_2, X_2) \tag{3.25}$$

3. The destination waits until all blocks are received before it starts to decode. At block B, we let $(w_{12,B}, w_{13,B}, w_{23,B}) = (0, 0, 0)$ and consequently we let $z_B = 0$. Thus, we are left only with z_{B-1}, which can be decoded if the receiver finds a unique $\hat{z}_{B-1}$ such that: $(\hat{\mathbf{y}}_2(w_{23,B}, \hat{z}_{B-1}, z_B), \mathbf{y}_3(B)$ $, \mathbf{x}_2(w_{23,B}, \hat{z}_{B-1}), \mathbf{v}(w_{23,B}))$ are jointly typical. This step can be made reliable if

$$\hat{R} < I(X_2, \hat{Y}_2; Y_3|V) \tag{3.26}$$

4. Moving to block $B-1$ and for a general block b, the destination finds a unique $\hat{w}_{23,b}$ such that $(\mathbf{v}(\hat{w}_{23,b}), \mathbf{y}_3(b))$ are jointly typical. The decoding error can be made small if

$$R_{23} < I(V; Y_3) \tag{3.27}$$

5. Now assuming z_b, z_{b-1} and $w_{23,b}$ have been decoded correctly, then $w_{13,b}$ can be decoded at block b if the receiver finds a unique $\hat{w}_{13,b}$ such that: $(\mathbf{u}_1(\hat{w}_{13,b}),$

$\mathbf{v}(\hat{w}_{23,b}), \mathbf{x}_2(w_{23,b}, \hat{z}_{b-1}), \hat{\mathbf{y}}_2(w_{23,b}, z_{b-1}, z_b), \mathbf{y}_3(b))$ are jointly typical. The decoding error can be made small if

$$R_{13} < I(U_1; \hat{Y}_2, Y_3 | V, X_2) \tag{3.28}$$

Therefore, as long as n is chosen sufficiently large, $R(U_1) \geq I(U_1; \hat{Y}_2, Y_3|V, X_2)$ and $R(U_2) \geq I(U_2; Y_2|X_2)$, we can decode $w_{13,b}$ and $w_{12,b}$, respectively with an arbitrarily small probability of error.

The achievable rate region of Theorem 3 then follows directly from combining equations (3.28), (3.24), (3.23) and (3.27), while the constraint given by (3.21) follows from combining (3.25) and (3.26). $\square$

Remark 2 *The symbols* $\mathbf{y}_2$ *are here compressed to* $\hat{\mathbf{y}}_2$ *after peeling off the component of* X_1 *intended to the relay node which is represented by* U_2*. Hence, we condition on knowing* U_2 *in (3.25). A similar situation arises when using compress-and-forward for multiple relays [22].*

We now briefly explain why strong typicality condition is needed in the CF scheme. First, we state the following Lemma due to Berger [23].

Lemma 1 *If* X *and* Y *are jointly typical,* Y *and* Z *are also jointly typical, both in the strong sense, then if* $X \to Y \to Z$*, all three sequences are jointly typical in the strong sense.*

Note that the above lemma does not hold for weakly typical sequences. In the CF scheme used, the relay does not know X_1 fully and moreover it does not know Y_3, the channel output at the destination. However, by the Markov chain $(X_1, Y_3) \to (X_2, Y_2) \to \hat{Y}_2$ and by Berger's Lemma, typicality between all sequences is guaranteed.

Another issue in which strong and weakly typical sequences differ is the random binning procedure. Recall that we limited the sequences U_1 and U_2 chosen for binning to be drawn from $S^{(n)}(U_1)$ and $S^{(n)}(U_2)$ and not from $A^{(n)}(U_1)$ and $A^{(n)}(U_2)$, respectively. The reason is that we try to guarantee the existence of at least one typical U_2 for any typical U_1 such that the pair (U_1, U_2) is jointly typical. This is not guaranteed by independent (random) drawing of each sequence as in weak typicality.

Remark 3 *The achievability result for the general RCPM given by Theorem 3 generalizes the achievability result of the general relay channel in [2, Theorem 6]. This is can be shown by setting $U_1 = X_1$, $U_2 = 0$ and $V = 0$. The constraint on the relay observation compression rate at the end Theorem 3 also corresponds to the constraint provided in [2, Theorem 6]. To see that, (3.21) becomes:*

$$\begin{aligned} I(\hat{Y}_2; Y_2|X_2) &\leq I(X_2, \hat{Y}_2; Y_3), \\ &= I(X_2; Y_3) + I(\hat{Y}_2; Y_3|X_2) \end{aligned} \tag{3.29}$$

But we have the Markov relation $Y_3 \rightarrow X_2, Y_2 \rightarrow \hat{Y}_2$, therefore,

$$I(\hat{Y}_2; Y_2, Y_3|X_2) \leq I(X_2; Y_3) + I(\hat{Y}_2; Y_3|X_2) \tag{3.30}$$

Further simplification leads to

$$I(\hat{Y}_2; Y_2|X_2, Y_3) \leq I(X_2; Y_3) \tag{3.31}$$

which is the constraint given in [2, Theorem 6].

Hence, we see that backward decoding does not improve the achievable rate of a relay channel employing compress-and-forward scheme. This is in contrast to the results of [16] when the relay uses the partial decode-forward of [2, Theorem 7].

Recall that in the strategy of Remark 2, the relay makes an observation, removes the part intended for relay, then compresses and transmits X_2 to destination.

Alternatively, X_2 may represents the compressed version of Y_2 together with W_{23}. In this approach, the relay does not peel off any component from its observation. The destination decodes X_2 (and hence Z and W_{23}), reconstructs $\hat{Y}_2$, and decodes W_{12} with the help of both $\hat{Y}_2$ and Y_3. It then decodes W_{13} knowing both X_2 and U_2. It can be shown, in a manner similar to Theorem 3, that the following rate region is achievable:

$$
\begin{aligned}
R_{13} &< I(U_1; \hat{Y}_2, Y_3 | U_2, V, X_2), && (3.32)\\
R_{12} &< \min\{I(U_2; Y_2 | X_2), I(U_2; \hat{Y}_2, Y_3 | X_2)\}, && (3.33)\\
R_{13} + R_{12} &< I(U_1; \hat{Y}_2, Y_3 | U_2, V, X_2) - I(U_1; U_2) \\
&\quad + \min\{I(U_2; Y_2 | X_2), I(U_2; \hat{Y}_2, Y_3 | X_2)\}, && (3.34)\\
R_{23} &< I(V; Y_3), && (3.35)
\end{aligned}
$$

subject to

$$I(\hat{Y}_2; Y_2 | V, X_2) \leq I(X_2, \hat{Y}_2; Y_3 | V) \tag{3.36}$$

where the random variables are drawn from any joint distribution

$$p(u_1, u_2)p(v)p(x_1|u_1, u_2)p(x_2|v)p(\hat{y}_2|y_2, x_2, v)p(y_2, y_3|x_1, x_2).$$

3.4 Upper Bounds on the Capacity Region

3.4.1 Upper Bound via Cut-set Theorem

The following upper rate bounds are obtained using the max-flow-min-cut theorem [4, Theorem 15.10.1] with different choices of subsets.

(a) A cut through source-relay and source-destination links gives:

$$R_{12} + R_{13} < I(X_1; Y_2, Y_3 | X_2) \tag{3.37}$$

(b) A cut through source-destination and relay-destination links gives:

$$R_{13} + R_{23} < I(X_1, X_2; Y_3) \tag{3.38}$$

(c) A cut through source-relay and relay-destination links gives:

$$R_{12} < I(X_1; Y_2|X_2) \tag{3.39}$$

$$R_{23} < I(X_2; Y_3|X_1) \tag{3.40}$$

3.4.2 Upper Bounds via Auxiliary Random Variables

The above cut-set bounds have two shortcomings. First, they ignore the components of the codewords represented by the auxiliary random variables. Second, there is no explicit individual bound on R_{13}. By appropriate definition of auxiliary random variables, a new set of bounds is obtained, including an upper bound on R_{13}, and a tighter bound on R_{12} in the degraded case.

Given any $\left(\left(2^{nR_{12}}, 2^{nR_{23}}, 2^{nR_{13}}\right), n\right)$ code for RCPM, the probability mass function on the joint ensemble space $W_{13} \times W_{13} \times W_{23} \times \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}_2^n \times \mathcal{Y}_3^n$ is given by:

$$\begin{aligned} p(w_{12}, w_{13}, w_{23}, \mathbf{x_1}, \mathbf{x_2}, \mathbf{y_2}, \mathbf{y_3}) = p(w_{12})p(w_{13})p(w_{23})p(\mathbf{x_1}|w_{12}, w_{13}) \\ \times \prod_{i=1}^{n} p(x_{2,i}|w_{23}, y_2^{i-1})p(y_{2,i}y_{3,i}|x_{1,i}x_{2,i}) \end{aligned} \tag{3.41}$$

Now, based on Fano's inequality, we have

$$H(W_{12}|Y_2^n) \leq nR_{12}P_{e,R}^{(n)} + 1 = n\delta_{R,n} \tag{3.42}$$

$$H(W_{13}, W_{23}|Y_3^n) \leq n(R_{13} + R_{23})P_{e,D}^{(n)} + 1 = n\delta_{D,n} \tag{3.43}$$

where $\delta_{R,n}$, $\delta_{D,n} \to 0$ as $P_e^{(n)} \to 0$.

R_{13} can be upper bounded as follows:

$$\begin{aligned}
nR_{13} &= H(W_{13}),\\
&= I(W_{13};Y_3^n) + H(W_{13}|Y_3^n),\\
&\overset{(a)}{\leq} I(W_{13};Y_3^n) + n\delta_{D,n},\\
&= \sum_{i=1}^{n} I(W_{13};Y_{3,i}|Y_3^{i-1}) + n\delta_{D,n},\\
&= \sum_{i=1}^{n} H(Y_{3,i}|Y_3^{i-1}) - H(Y_{3,i}|Y_3^{i-1},W_{13}) + n\delta_{D,n},\\
&\leq \sum_{i=1}^{n} H(Y_{3,i}) - H(Y_{3,i}|Y_2^{i-1},Y_3^{i-1},W_{13}) + n\delta_{D,n},\\
&= \sum_{i=1}^{n} H(Y_{3,i}) - H(Y_{3,i}|U_i,V_i) + n\delta_{D,n},\\
&= \sum_{i=1}^{n} I(U_i,V_i;Y_{3,i}) + n\delta_{D,n}
\end{aligned} \tag{3.44}$$

where (a) follows from Fano's inequality since $H(W_{13}|Y_3^n) \leq H(W_{13},W_{23}|Y_3^n)$ and we define $U_i = (Y_2^{i-1}, Y_3^{i-1})$ and $V_1 = V_2 = ... = V_n = W_{13}$.

And R_{12} can be upper bounded as follows:

$$\begin{aligned}
nR_{12} &= H(W_{12}),\\
&= I(W_{12};Y_2^n) + H(W_{12}|Y_2^n),\\
&\leq I(W_{12};Y_2^n,Y_3^n,W_{13},W_{23}) + n\delta_{R,n},\\
&\overset{(b)}{=} I(W_{12};Y_2^n,Y_3^n|W_{13},W_{23}) + n\delta_{R,n},\\
&= \sum_{i=1}^{n} H(Y_{2,i},Y_{3,i}|Y_2^{i-1},Y_3^{i-1},W_{13},W_{23}) - H(Y_{2,i}Y_{3,i}|Y_2^{i-1}Y_3^{i-1},W_{12},W_{13},W_{23})\\
&\qquad + n\delta_{R,n},\\
&\overset{(c)}{=} \sum_{i=1}^{n} H(Y_{2,i},Y_{3,i}|Y_2^{i-1},Y_3^{i-1},W_{13},W_{23},X_{2,i})\\
&\qquad - H(Y_{2,i}Y_{3,i}|Y_2^{i-1}Y_3^{i-1},W_{12},W_{13},W_{23},X_{2,i}) + n\delta_{R,n},
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{n} H(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, W_{13}, W_{23}, X_{2,i}) \\
&\qquad - H(Y_{2,i}Y_{3,i}|Y_2^{i-1}Y_3^{i-1}, W_{12}, W_{13}, W_{23}, X_{1,i}, X_{2,i}) + n\delta_{R,n}, \\
&\overset{(d)}{\leq} \sum_{i=1}^{n} H(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, W_{13}, X_{2,i}) - H(Y_{2,i}Y_{3,i}|X_{1,i}, X_{2,i}) + n\delta_{R,n}, \\
&\leq \sum_{i=1}^{n} H(Y_{2,i}, Y_{3,i}|Y_2^{i-1}, Y_3^{i-1}, W_{13}, X_{2,i}) - H(Y_{2,i}Y_{3,i}|Y_2^{i-1}Y_3^{i-1}, W_{13}, X_{1,i}, X_{2,i}) \\
&\qquad + n\delta_{R,n}, \\
&= \sum_{i=1}^{n} I(X_{1,i}; Y_{2,i}Y_{3,i}|U_i, V_i, X_{2,i}) + n\delta_{R,n} \qquad (3.45)
\end{aligned}
$$

(b) follows from independence of W_{13}, W_{13} and W_{23}, (c) is due to the fact that $X_{2,i}$ is a function of Y_2^{i-1} and W_{23} and (d) is justified because of removing W_{23} from the conditioning of the first term and by noting that $(W_{12}, W_{13}, W_{23}, Y_2^{i-1}, Y_3^{i-1}) \rightarrow (X_{1,i}, X_{2,i}) \rightarrow (Y_{2,i}, Y_{3,i})$ form a Markov chain from the memoryless property of the channel.

In case of degraded RCPM, $X_{1,i} \rightarrow (U_i, V_i, X_{2,i}, Y_{2,i}) \rightarrow Y_{3,i}$ form a Markov chain and so

$$\sum_{i=1}^{n} I(X_{1,i}; Y_{3,i}|U_i, V_i, X_{2,i}, Y_{2,i}) = 0 \qquad (3.46)$$

Hence R_{12} is upper bounded by

$$nR_{12} \leq \sum_{i=1}^{n} I(X_{1,i}; Y_{2,i}|U_i, V_i, X_{2,i}) + n\delta_{R,n} \qquad (3.47)$$

Finally, the single letter characterization for the two previous bounds can be obtained by introducing a time-sharing random variable Q which is uniformly distributed over the n symbols and independent of W_{12}, W_{13}, W_{23}, X_1, X_2, Y_2 and Y_3. Next, define $U = (Q, U_Q)$, $V = V_Q$, $X_1 = X_Q$, $X_2 = X_{2,Q}$, $Y_2 = Y_{2,Q}$, and $Y_3 = Y_{3,Q}$. Following steps similar to those in [4, Ch. 15.3.4] the bounds can be reduced to the single letter form given by:

$$R_{13} < I(U, V; Y_3), \qquad (3.48)$$

$$R_{12} < I(X_1; Y_2|U, V, X_2) \tag{3.49}$$

It can be easily shown that for the degraded RCPM, the bound on R_{12} is tighter than the cut-set bound of (3.39) and in fact comparing (3.12) and (3.49), the bound is tight.

Remark 4 *The auxiliary random variables U and V are not independent, therefore their cardinality cannot be bounded using existing methods (e.g. [24]). This problem has been previously noted in [13, Comment 4.1].*

3.5 Application to Gaussian Channels

Characterizing the capacity region in Gaussian channels is of high interest as Gaussian channel provide approximation to realistic wireless channel models. Moreover, the DMC bounds developed in the previous sections are incomputable due to the problem of bounding the cardinality of the auxiliary random variables. Nevertheless, by applying previous results in Gaussian RCPM channel we can obtain some numerical results and assess the behavior of various rates in our three-node network. For simplicity, the following results assume the input distributions to be Gaussian. Although, Gaussian inputs may not be optimal, however, searching over all possible distributions is a tedious task. All rate values in this section are expressed in (bps/Hz).

3.5.1 Decode-and-forward

We consider a degraded Gaussian RCPM of Definition 8, for which the following result applies:

Corollary 1 *An achievable rate region for the AWGN relay channel with private* *messages is the convex hull of the rates* (R_{12}, R_{23}, R_{13}) *satisfying:*

$$R_{13} < \min\left\{\mathcal{C}\left(\frac{\bar{\alpha}P_1 + \bar{\gamma}P_2 + 2\sqrt{\bar{\alpha}\bar{\beta}\bar{\gamma}P_1P_2}}{\alpha P_1 + \gamma P_2 + N_3}\right), \mathcal{C}\left(\frac{\beta\bar{\alpha}P_1}{\alpha P_1 + N_2}\right)\right\}, \tag{3.50}$$

$$R_{23} < \mathcal{C}\left(\frac{\gamma P_2}{\alpha P_1 + N_3}\right), \tag{3.51}$$

$$R_{12} < \mathcal{C}\left(\frac{\alpha P_1}{N_2}\right) \tag{3.52}$$

for some α, β *and* $\gamma \in [0, 1]$,

Here α indicates the fraction of the source's power allocated to convey its private message to the relay, β controls the power allocated to the source-to-destination message in the previous block and in the current block and finally, γ indicates the part of the power allocated to convey the private message from the relay to the destination.

Proof:

This rate region is achieved via the coding strategy described in Theorem 2 and then evaluating mutual informations using the following independent distributions:

- $U \sim \mathcal{N}(0, P_u)$, $V' \sim \mathcal{N}(0, \beta\bar{\alpha}P_1)$.
- $X_1' \sim \mathcal{N}(0, \alpha P_1)$, $X_2' \sim \mathcal{N}(0, \gamma P_2)$.

Furthermore, we let $V = \sqrt{\frac{\bar{\alpha}\bar{\beta}P_1}{P_u}}U + V'$, $X_1 = V + X_1'$ and $X_2 = \sqrt{\frac{\bar{\gamma}P_2}{P_u}}U + X_2'$.

□

Remark 5 *One would recover the decode-and-forward capacity in a degraded Gaussian relay channel defined in [2, Theorem 5] by setting* $\alpha = 0$ *and* $\gamma = 0$.

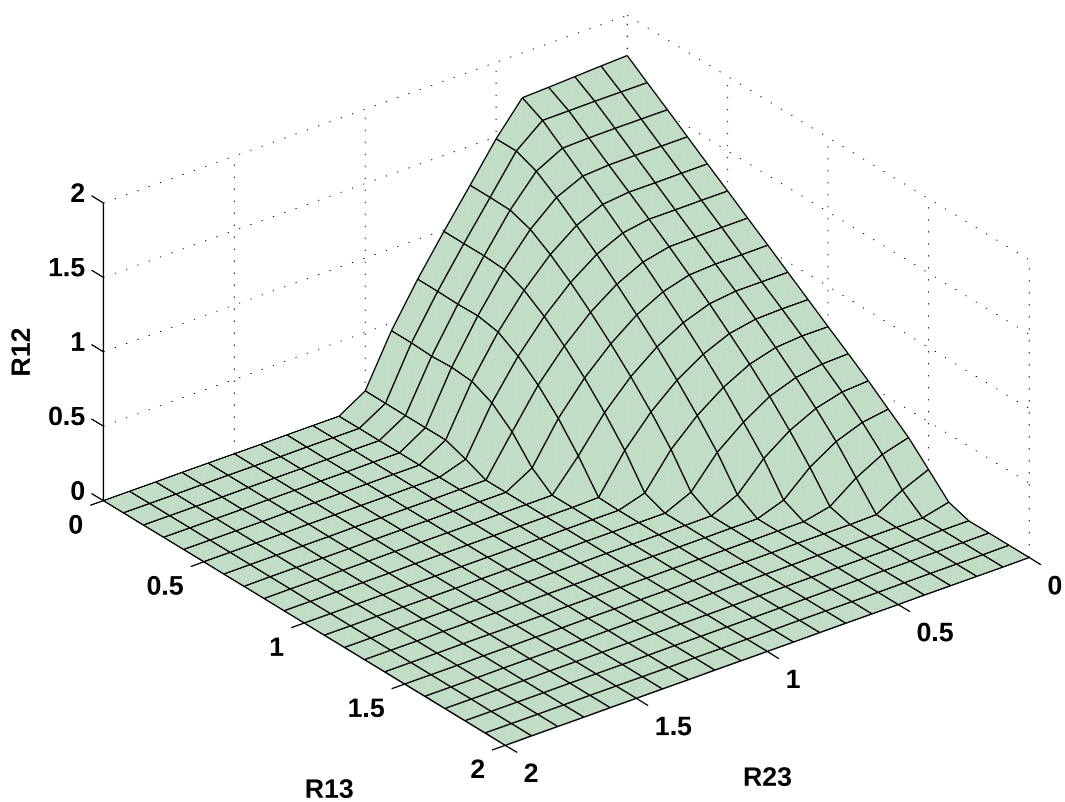

Figure 3.5. An Achievable rate region of the degraded Gaussian relay channel with private messages.

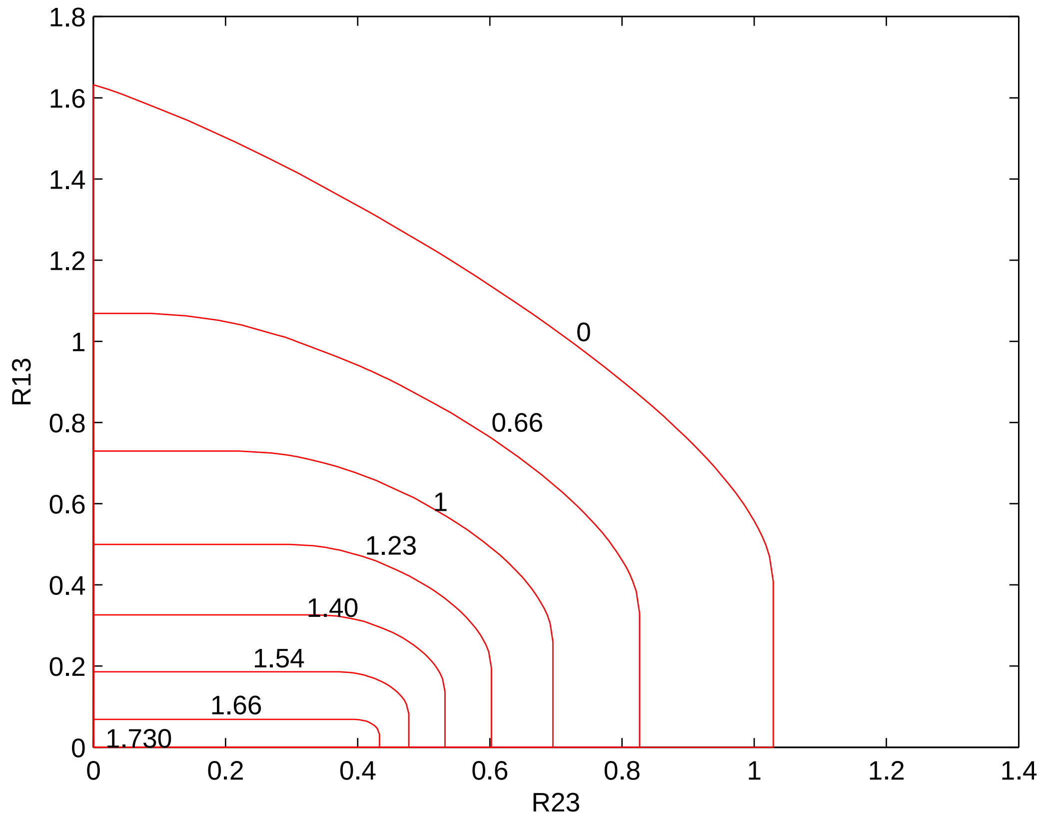

Figure 3.6. Cross sections of achievable rate region of the degraded Gaussian relay channel with private messages, parameterized by the private rate R_{12}

Example 1 *Consider the degraded Gaussian RCPM with* $\frac{P_1}{N_2} = 10dB$, $\frac{P_2}{N_3} = 5dB$. *Figure 3.5 shows an achievable rate region of this Gaussian RCPM. Figure 3.6 shows contour plots of the achievable rate region in the* R_{13}-R_{23} *plane to demonstrate the trade-off between the relayed rate and one of the private rates.*

3.5.2 Compress-and-forward

We consider a general (non-degraded) Gaussian RCPM. In addition to the channel outputs relations given in section 3.2, the random variable representing the compressed channel output of the relay is chosen as:

$$\hat{Y}_2 = U_1 + Z_2 + Z_c \tag{3.53}$$

where Z_c refers to the compression noise whose variance N_c is determined by the constraint given in Theorem 3 . Then, the following result applies:

Corollary 2 *An achievable rate region for the AWGN relay channel with private messages is given by:*

$$R_{13} < \mathcal{C}\left(\frac{P_{u_1}P_{u_2}(1-\rho^2) + (P_{u_1} + 2\rho\sqrt{P_{u_1}P_{u_2}})(N_2+N_c) + P_{u_1}N_3}{P_{u_2}(N_2+N_c) + N_3(N_2+N_c)}\right), \tag{3.54}$$

$$R_{12} < \mathcal{C}\left(\frac{P_{u_2} + 2\rho\sqrt{P_{u_1}P_{u_2}}}{P_{u_1} + N_2}\right), \tag{3.55}$$

$$R_{13} + R_{12} < \mathcal{C}\left(\frac{(P_{u_1} + 2\rho\sqrt{P_{u_1}P_{u_2}})(N_2+N_3+N_c)}{P_{u_2}(N_2+N_3+N_c) + N_3(N_2+N_c)}\right) + \mathcal{C}\left(\frac{P_{u_2} + 2\rho\sqrt{P_{u_1}P_{u_2}}}{P_{u_1} + N_2}\right) + \mathcal{C}\left(-\rho^2\right), \tag{3.56}$$

$$R_{23} < \mathcal{C}\left(\frac{\gamma P_2}{P_1 + \bar{\gamma}P_2 + N_3}\right) \tag{3.57}$$

subject to

$$N_c \geq \left(\frac{P_{u_1}P_{u_2}(1-\rho^2) + P_{u_1}N_3 + N_2(P_1+N_3)}{\bar{\gamma}P_2}\right) \tag{3.58}$$

for some α *and* $\gamma \in [0,1]$ *and* $0 < \rho < 1$.

$$P_{u_1} \leq \left(\sqrt{P_1(1-\alpha(1-\rho^2))} - \rho\sqrt{\alpha P_1}\right)^2 \tag{3.59}$$

$$P_{u_2} \leq \alpha P_1 \tag{3.60}$$

We shall see that the variable ρ denotes the correlation coefficient between the two Gaussian codebooks U_1 and U_2 used by the source, α is the fraction of source power dedicated to its private message, and γ is the fraction of relay power dedicated to the private message of the relay.

Proof:

We attempt to find attractive compress-and-forward achievable rates, according to Equations (3.17)-(3.20), by finding suitable X_1, X_2, U_1, U_2, and V, and their distributions. The extension of compress-and-forward DMC to the Gaussian channel faces two main problems:

- The proof of achievability requires strong typicality. In general, strong typicality does not apply to continuous random variables. However, for Gaussian input distributions the problem can be solved, as explained in [22, Remark 30].
- Every time we wish to extend DMC results to Gaussian channels, we need to find optimal codebooks which are often left unspecified in the DMC developments. In some early cases, e.g. dirty paper coding of Costa [25], the optimum codebook was correctly guessed, and the guess was verified by tightness against upper bounds. In many cases, such as ours, it is not feasible to verify optimality of the guesses. We use jointly Gaussian U_1 and U_2 with correlation ρ as our guess, where ρ will be optimized.

The codebook is therefore:

- $U_1 \sim \mathcal{N}(0, P_{u_1})$, $U_2 \sim \mathcal{N}(0, P_{u_2})$ with correlation coefficient ρ

- $V \sim \mathcal{N}(0, \gamma P_2)$ and $X_2' \sim \mathcal{N}(0, \bar{\gamma} P_2)$

The transmit signals are $X_1 = U_1 + U_2$ and $X_2 = V + X_2'$. The achievability of Equations (3.54)-(3.57) follows from this choice of codebooks (the details of proof appear in the Appendix). The power constraint of X_1 gives rise to power constraints (3.59) and (3.60) for U_1 and U_2, respectively. Equation (3.58) gives N_c, the compression noise variance, in terms of other variables in the system. □

Remark 6 *One would recover the compress-and-forward rate in a non-degraded Gaussian relay channel provided in [22] by setting $\alpha = 0$ and $\gamma = 0$.*

Note that the above achievable rate was obtained by successive nulling and cancellation in a particular order, namely, the private messages are decoded first. Other achievable rates can be obtained by different orderings of nulling and canceling the codeword components representing Z and W_{12} at the relay, and W_{13} and W_{23} at the destination. This leads to a total of four possibilities. Each of the decoding orders gives one achievable rate region, and naturally the overall achievable rate region is the convex hull of the four (the remaining rate regions can be similarly derived as Corollary 2) .

We briefly comment on the various orderings possible for nulling and canceling. Consider the achievable region in Corollary 2 where the private messages are decoded first and peeled off, following by the decoding of relayed message at rate R_{13}. It is known that for successive decoding to work well, one must start with the dominant signal in the superposition. However, at least at some operating points we may have very low rates for private messages, which translates into low power allocated for these messages. Therefore the remainder of the power will be available for relaying. This means that the private messages do not always correspond to the dominant signal.

For example, consider the case where R_{23} is small, that is, the signal corresponding to the private message from relay to destination has small power. If we insist on decoding the private message first, it will limit the power and hence the rate associated with R_{13} below the levels possible in the system, and hence is very suboptimal. It is then reasonable to proceed with decoding as follows.

At the destination, we start by considering the signal component of R_{23} as "noise" and decode the relayed signal at rate R_{13}. This will allow us to know (part of) the signal transmitted by the source, which can now be peeled off. Note that the transmission by the relay cannot be peeled off in general, because the system is not degraded. Now, the private message from the relay is decoded.

As yet another example, consider that R_{12} is much smaller than the other rates in the system. Therefore, if the relay wishes to peel off W_{12} from its input, this will severely limit the amount of information that can be relayed in a decode-and-forward protocol, since this assumes R_{12} is the dominant signal at the relay input. Once again, by reversing the order of nulling and cancelling whenever appropriate, one may be able to obtain better achievable rates.

3.6 Appendix: Derivation of (3.54)-(3.58)

Given the input-output relationship between different random variables in Gaussian channels, the rate region represented by (3.54)-(3.57) is obtained as follows

$$\begin{aligned} R_{13} &\leq I(U_1; \hat{Y}_2, Y_3 | V, X_2) \\ &= h(\hat{Y}_2, Y_3 | V, X_2) - h(\hat{Y}_2, Y_3 | U_1, V, X_2) \end{aligned} \tag{3.61}$$

Now,

$$
\begin{aligned}
h(\hat{Y}_2, Y_3|V, X_2) &= h(U_1 + Z_2 + Z_c, X_1 + X_2 + Z_3|V, X_2) \\
&= h(U_1 + Z_2 + Z_c, U_1 + U_2 + Z_3) \\
&= \frac{1}{2}\log(2\pi e)^2 \begin{vmatrix} P_{u_1} + N_2 + N_c & P_{u_1} + \rho\sqrt{P_{u_1}P_{u_2}} \\ P_{u_1} + \rho\sqrt{P_{u_1}P_{u_2}} & P_1 + N_3 \end{vmatrix} \\
&= \frac{1}{2}\log(2\pi e)^2 \Big(P_{u_1}P_{u_2}(1-\rho^2) + P_1(N_2 + N_c) + P_{u_1}N_3 \\
&\qquad + N_3(N_2 + N_c) \Big) \qquad (3.62)
\end{aligned}
$$

while

$$
\begin{aligned}
h(\hat{Y}_2, Y_3|U_1, V, X_2) &= h(Z_2 + Z_c, U_2 + Z_3) \\
&= \frac{1}{2}\log(2\pi e)^2 \begin{vmatrix} N_2 + N_c & 0 \\ 0 & P_{u_2} + N_3 \end{vmatrix} \\
&= \frac{1}{2}\log(2\pi e)^2 \Big((P_{u_2} + N_3)(N_2 + N_c) \Big) \qquad (3.63)
\end{aligned}
$$

After straight forward simplifications we get,

$$
R_{13} < \mathcal{C}\left(\frac{P_{u_1}P_{u_2}(1-\rho^2) + (P_{u_1} + 2\rho\sqrt{P_{u_1}P_{u_2}})(N_2 + N_c) + P_{u_1}N_3}{P_{u_2}(N_2 + N_c) + N_3(N_2 + N_c)} \right) \qquad (3.64)
$$

Next,

$$
\begin{aligned}
R_{12} &\le I(U_2; Y_2|X_2) \\
&= h(Y_2|X_2) - h(Y_2|U_2, X_2) \\
&= \frac{1}{2}\log(2\pi e)(P_1 + N_2) - \frac{1}{2}\log(2\pi e)(P_{u_1} + N_2) \qquad (3.65)
\end{aligned}
$$

simplifying we get,

$$
R_{12} < \mathcal{C}\left(\frac{P_{u_2} + 2\rho\sqrt{P_{u_1}P_{u_2}}}{P_{u_1} + N_2} \right) \qquad (3.66)
$$

To compute $I(U_1; U_2)$, we have

$$
I(U_1; U_2) = h(U_1) + h(U_2) - h(U_1, U_2),
$$

$$= \frac{1}{2}\log(2\pi e)(P_{u_1}) + \frac{1}{2}\log(2\pi e)(P_{u_2})$$
$$- \frac{1}{2}\log(2\pi e)^2 \begin{vmatrix} P_{u_1} & \rho\sqrt{P_{u_1}P_{u_2}} \\ \rho\sqrt{P_{u_1}P_{u_2}} & P_{u_2} \end{vmatrix} \tag{3.67}$$

Simplifying, we get

$$I(U_1;U_2) = -\mathcal{C}\big(-\rho^2\big) \tag{3.68}$$

Combining (3.64), (3.66) and (3.68), we have

$$R_{13} + R_{12} < \mathcal{C}\left(\frac{(P_{u_1} + 2\rho\sqrt{P_{u_1}P_{u_2}})(N_2 + N_3 + N_c)}{P_{u_2}(N_2 + N_3 + N_c) + N_3(N_2 + N_c)}\right) + \mathcal{C}\left(\frac{P_{u_2} + 2\rho\sqrt{P_{u_1}P_{u_2}}}{P_{u_1} + N_2}\right)$$
$$+ \mathcal{C}\left(-\rho^2\right) \tag{3.69}$$

For R_{23}, we have

$$R_{23} \leq I(V;Y_3),$$
$$= h(Y_3) - h(Y_3|V),$$
$$= \mathcal{C}\left(\frac{\gamma P_2}{P_1 + \bar{\gamma}P_2 + N_3}\right) \tag{3.70}$$

We are left with computing the constraint on the compression noise variance N_c. From (3.25),

$$\hat{R} > I(\hat{Y}_2;Y_2|U_2,V,X_2)$$
$$= h(\hat{Y}_2|U_2,V,X_2) - h(\hat{Y}_2|U_2,V,X_2,Y_2)$$
$$= \frac{1}{2}\log(2\pi e)(P_{u_1} + N_2 + N_c) - \frac{1}{2}\log(2\pi e)(N_c) \tag{3.71}$$

Hence,

$$\hat{R} > \mathcal{C}\left(\frac{P_{u_1} + N_2}{N_c}\right) \tag{3.72}$$

On the other hand, we have from (3.26)

$$\hat{R} < I(X_2,\hat{Y}_2;Y_3|V)$$
$$= I(X_2;Y_3|V) + I(\hat{Y}_2;Y_3|V,X_2) \tag{3.73}$$

Now,

$$I(X_2;Y_3|V) = \mathcal{C}\left(\frac{\bar{\gamma}P_2}{P_1+N_3}\right) \tag{3.74}$$

Moreover, we have

$$\begin{aligned} I(\hat{Y}_2;Y_3|V,X_2) &= h(\hat{Y}_2|V,X_2) - h(\hat{Y}_2|V,X_2,Y_3) \\ &= h(\hat{Y}_2) - h(\hat{Y}_2|Y_3') \end{aligned} \tag{3.75}$$

where we defined

$$Y_3' = Y_3 - X_2 = U_1 + U_2 + Z_3$$

Now,

$$h(\hat{Y}_2) = \frac{1}{2}\log(2\pi e)(P_{u_1} + N_2 + N_c) \tag{3.76}$$

$$\begin{aligned} h(\hat{Y}_2|Y_3') &= h(\hat{Y}_2, Y_3') - h(Y_3') \\ &= \frac{1}{2}\log(2\pi e)^2 \begin{vmatrix} P_{u_1} + N_2 + N_c & P_{u_1} + \rho\sqrt{P_{u_1}P_{u_2}} \\ P_{u_1} + \rho\sqrt{P_{u_1}P_{u_2}} & P_1 + N_3 \end{vmatrix} - \frac{1}{2}\log(2\pi e)(P_1 + N_3) \end{aligned} \tag{3.77}$$

Simplifying we get,

$$h(\hat{Y}_2|Y_3') = \frac{1}{2}\log(2\pi e)\left(\frac{P_{u_1}P_{u_2}(1-\rho^2) + P_1(N_2+N_c) + P_{u_1}N_3 + N_3(N_2+N_c)}{P_1+N_3}\right) \tag{3.78}$$

Combining (3.76) and (3.78),

$$I(\hat{Y}_2;Y_3|V,X_2) = \mathcal{C}\left(\frac{P_1P_{u_1} - P_{u_1}P_{u_2}(1-\rho^2)}{P_{u_1}P_{u_2}(1-\rho^2) + P_1(N_2+N_c) + P_{u_1}N_3 + N_3(N_2+N_c)}\right) \tag{3.79}$$

Combining (3.74) and (3.79),

$$\hat{R} = \mathcal{C}\left(\frac{\bar{\gamma}P_2}{P_1+N_3}\right) + \mathcal{C}\left(\frac{P_1P_{u_1} - P_{u_1}P_{u_2}(1-\rho^2)}{P_{u_1}P_{u_2}(1-\rho^2) + P_1(N_2+N_c) + P_{u_1}N_3 + N_3(N_2+N_c)}\right) \tag{3.80}$$

Finally, from (3.72) and (3.80), we have

$$\frac{P_{u_1} + N_2 + N_c}{N_c} \leq \left(\frac{\bar{\gamma}P_2 + P_1 + N_3}{P_1 + N_3}\right) \times \left(\frac{(P_1 + N_3)(P_{u_1} + N_2 + N_c)}{P_{u_1}P_{u_2}(1-\rho^2) + P_1(N_2 + N_c) + P_{u_1}N_3 + N_3(N_2 + N_c)}\right) \tag{3.81}$$

Solving (3.81) for N_c leads to:

$$N_c \geq \left(\frac{P_{u_1}P_{u_2}(1-\rho^2) + P_{u_1}N_3 + N_2(P_1 + N_3)}{\bar{\gamma}P_2}\right) \tag{3.82}$$

CHAPTER 4

SPECTRALLY-EFFICIENT RELAY SELECTION WITH LIMITED FEEDBACK

4.1 Introduction

Relays can improve the performance of a wireless system via a number of mechanisms, such as increased spatial diversity or beamforming effects (whenever available). But for half-duplex relays, some time must be set aside for listening to the source, during which the relay must be silent. These silent times lead to a loss of spectral efficiency (also known as the multiplexing loss).

In this chapter, we address the issue of multiplexing loss in relay networks. As our main tool, we use variations on relay selection, which has nice properties but requires an exchange of channel state information between the nodes. We aim to recover the multiplexing loss using relay selection, under the constraint of very limited feedback (on the order of merely bits/relay).

Relay selection has been recently proposed to overcome some shortcomings of the existing relaying approaches in networks with multiple relays. Relay selection simplifies signaling, avoids complex synchronization schemes, and with careful design can preserve the spatial diversity provided by the total number of relays available in the network [26]. However, the selection process requires an overhead. This overhead grows with the number of relays in the network. Moreover, in practice, the control channel that often conveys the feedback information is of very limited rate [27]. Hence, one is motivated to devise relay selection methods with limited feedback.

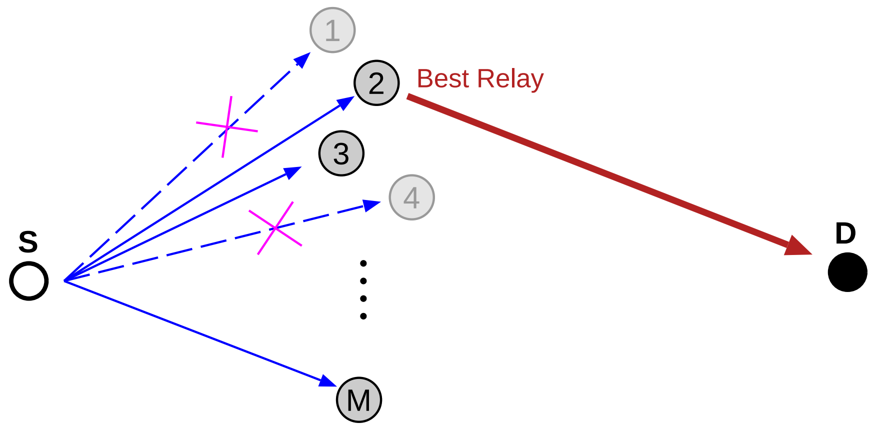

Figure 4.1. Wireless network with relay selection. The best successful relay transmits to destination.

The setting of this chapter includes a source, a destination, and multiple relays, all of them single-antenna nodes in a fading wireless environment (see Figure 4.1). The relays employ a regenerative, decode-and-forward (DF) strategy. We consider two topologies: Either (a) the destination has a viable link to all transmitters, so a direct link from the source to destination exists, or (b) a direct link does not exist between the source and destination, so they can only communicate through the relays. The topologies include a very general inter-relay interference, which is not assumed to be either strong or weak. In fact, the general relay-on-relay interference is a key challenge and interesting facet of this work. Throughout this chapter, we define the best relay as the one with the highest instantaneous channel gain to the destination.

In the scenario where a direct source-destination link exists, one may use feedback not only to select the best relay, but also to select *no* relay when none is needed [28]. Thus, feedback can be used to gain both diversity as well as multiplexing. Motivated by this argument, we present an *Incremental Transmission Relay Selection (ITRS)* protocol, which employs a Type-I hybrid ARQ with packet combining at the destination, and includes a limited-feedback handshake for relay selection. We perform a detailed outage and DMT analysis of this method. ITRS meets the multiple-input-single-output (MISO) DMT bound, and provides a distinct improvement on a number of existing comparable methods.

In the scenario where a direct source-destination link does not exist, we employ overlapping (non-orthogonal) transmissions from the source to the relays and from the selected relay to the destination. This strategy leads to a *Multi-Hop Relay Selection (MHRS)* protocol. We analyze the performance of this method when the relays employ either successive cancellation, joint decoding of the incoming messages, or hybrid adaptive strategies. This system provides an embedded set of DMT curves that can be used for unequal-error protection (UEP), a very attractive feature for practical systems. Even the minimal DMT of this system is shown to improve on comparable existing methods.

We now outline some past works on multiple relay networks that use the DF relaying scheme. To improve the spectral efficiency of relaying, Laneman and Wornell [29] propose distributed space-time codes (DSTC), which requires synchronization between the nodes. Azarian et al. [30] propose dynamic decode and forward (DDF) for multiple relays. The basic ideas of DDF are very nice, but unfortunately DDF does not scale with increasing number of relays in the high-rate regime. Bletsas et al. [26] propose an opportunistic relaying scheme that achieves the DMT of DSTC without the synchronization requirement, but requires transmit and receive-side channel-state information at the relays. Recently, Tajer and Nosratinia [31] show that it is possible to achieve the same DMT with very little information exchange.

Relay selection has also generated a sizable literature. The work in [32] shows the outage-optimality of relay selection under aggregate power constraint, which borrows much from the earlier work in [26]. In the multi-source, multi-destination scenarios, only a few works exist. Nosratinia and Hunter [33] demonstrate relay selection techniques that can capture maximum diversity in the number of cooperating nodes, while each node only knows its own receive channel state. Lin et. al. [34] presents relay selection criteria in the presence of node locations. Beres and Adve [35] considers various levels of centralization and compares selection with DSTC under instantaneous channel knowledge. There are also several works on relay selection adopt the amplify-and-forward (AF) scheme [36], [37], [38], [39], [40], [41] whose details are beyond the scope of this study.

There are some works with similarities to our ITRS protocol: Zhao and Valenti [28] were the first to consider hybrid-ARQ in relays, but they select relays based on average channel gains, resulting in coding gain and second order diversity but not a diversity order that is equal to the number of available relays in the network. Lo et al. [42] propose a decentralized, limited-feedback, HARQ-based relay selection, and concentrate on BER and throughput studies.

Recently, Yang and Belfiore [43] present a sequential AF technique where, like our MHRS protocol, the relays transmit in succession. The results of [43] on AF networks cannot be directly compared with the present work, which is on DF networks. Furthermore, the achievable DMT of [43] is not known except for special cases where relays are isolated, or when two-slot transmission is used.

To summarize, the contributions of this chapter are as follows: Relay selection methods are devised under very limited feedback and very general inter-relay interference conditions, for the purpose of recovering multiplexing gain in half-duplex DF relay networks. For two topologies with and without a direct link, we propose two protocols, named ITRS and MHRS, which are analyzed in detail and their DMT is provided or bounded. The MHRS protocol gives rise to an embedded set of DMT curves that can be used for unequal error protection. Our protocols improve over existing methods for half-duplex DF relays, including DSTC, DDF, and opportunistic relaying.

4.2 System and Channel Models

The system model consists of a destination node, a source, and M half-duplex relays (see Figure 4.1). The channel gains between any two nodes is described by a flat, quasi-static block Rayleigh fading model. We also consider the case where the source-destination link is non-existent, thus the communication must take place in a two-hop fashion through the relays, creating a bigger challenge for spectral efficiency. The analysis is general and avoids any special assumptions, such as isolated relays or strong inter-relay interference.

For relay selection, we assume the existence of a low-rate, reliable feedback from the destination to the relays (and possibly from destination to the source). Aside from this, no transmit-side channel state information (CSI) is assumed. The nodes have access to perfect receive CSI.

We assume that the input codewords are obtained from a random Gaussian codebook. The length of a codeword is asymptotically large but spans one coherence interval of the channel. Source and relay nodes each transmit under an average power constraint P. The receive noises are normally distributed $\sim \mathcal{N}(0, \sigma^2)$. The average receive SNR at each receiver is denoted ρ, i.e., $\rho = \frac{P}{\sigma^2}$. The system has M relays, indexed $m = 1, \ldots, M$. The channels between source and relays ($h_{s,m}$), relays and destination ($h_{m,d}$) and the inter-relay channels ($h_{m,m'}$) are zero-mean independent, circularly symmetric complex Gaussian random variables whose variances are $\lambda_{s,m}$, $\lambda_{m,d}$, $\lambda_{m,m'}$, respectively. The magnitude square of channel coefficients, also known as effective channel gain, are denoted $g_{s,m}$, $g_{m,d}$, $g_{m,m'}$ and follow exponential distributions. Whenever it exists, the source-destination channel is described with $h_{s,d}$ and follows similar statistics as the other links in the system. For simplicity of exposition, throughout the chapter we assume that the source-relay channels have identical distributions, and the same holds for relay-destination and inter-relay channels, respectively. However, the DMT results do not depend on this assumption and continue to hold even if channels have non-identical (but finite) variance.

The performance of protocols is measured by outage [44], and the diversity-multiplexing tradeoff [8]. A channel is said to achieve multiplexing gain r and diversity gain d if there exists a sequence of codes $C(\rho)$ with rate $R(\rho)$ and resulting outage probability $P_{out}(\rho)$ such that:

$$\lim_{\rho\to\infty} \frac{R(\rho)}{\log(\rho)} = r \qquad\qquad \lim_{\rho\to\infty} \frac{\log P_{out}(\rho)}{\log(\rho)} = -d \tag{4.1}$$

In the following developments, we say $f(\rho)$ is *exponentially equal* to ρ^v , denoted by $f(\rho) \doteq \rho^v$, if

$$\lim_{\rho\to\infty} \frac{\log(f(\rho))}{\log(\rho)} = v \tag{4.2}$$

4.3 Incremental Transmission Relay Selection

This section presents a protocol for a multi-relay network with limited feedback, called *Incremental Transmission Relay Selection (ITRS)*. The network consists of a source, M relays, and a destination, where the destination has a fading link to the source as well as the relays (see Section 4.2). In this protocol, the limited feedback has dual use: it selects the best relay, thus improving diversity, and also enables retransmission (HARQ), thus improving spectral efficiency. The broad outline of the protocol is as follows: A packet is broadcast by the source. If the destination cannot decode, a limited-feedback handshake is performed that identifies the best available node (among source and relays), which will retransmit the packet. The ITRS protocol is described in detail in Figure 4.2. Note that the channel gains are assumed to remain fixed during steps 3-5.

1. The source transmits a packet.
2. If the destination correctly decodes the message, it broadcasts an ACK and system returns to Step 1. Otherwise destination broadcasts a NACK.
3. Upon receiving the NACK, the relays that successfully decoded the packet will declare their status via a one-bit packet (RTS - Request to Send) to the destination. The RTS packet includes a pilot.
4. The destination estimates channel gains, picks the best transmitter from among successful relays and the source, and broadcasts the index of the best node.
5. The best node will retransmit the packet. The destination combines its two received packets and decodes. If unsuccessful, destination is in outage.

Figure 4.2. The Incremental Transmission with Relay Selection (ITRS) protocol

The ITRS protocol uses a maximum of one retransmission. Further retransmissions would reduce (and eventually eliminate) outage, but also incur further delay. We study the case of one retransmission, which incurs modest delay and yet captures the biggest part of the gains available through retransmissions.

The ITRS protocol uses type-I H-ARQ with packet combining, i.e., relays use the same codebook as the source. Type-II H-ARQ, where the relays use non-identical codebooks, has better mutual information but also increases complexity. The two methods achieve the same DMT.

The ITRS protocol includes the source in the competition for the re-transmission, thus improving the diversity as well as throughput, as seen in the sequel.

The protocols presented in this chapter require feedback, whose transmission in turn requires a channel and a protocol. Feedback often goes through a control channel that exists in many wireless standards. The medium access layer for these channels can be either contention-based or slotted. In the former, all relays contend in sending their RTS to the destination, in which case the relay address (ID) must be attached to the RTS packet. In a time-slotted system, on the other hand, each relay transmits an RTS in its designated mini-slot only. This avoids collision between relays, but some mini-slots may go unused depending on the number of available relays, therefore usage of channel resources may be inefficient.

4.3.1 Outage Probability and Effective Rate

During the first transmission of a packet by the source, the received signals at the relays and the destination are given by:

$$\mathbf{y}_m = h_{s,m}\,\mathbf{x}_s + \mathbf{z}_m \qquad m = 1, \ldots, M \tag{4.3}$$

$$\mathbf{y}_d = h_{s,d}\,\mathbf{x}_s + \mathbf{z}_d \tag{4.4}$$

During a re-transmission, the received signal at the destination is given by

$$\mathbf{y}_d = h_{m^*,d}\,\mathbf{x}_{m^*} + \mathbf{z}_d \tag{4.5}$$

Where m^* denotes the index of the selected relay. We emphasize again that for the retransmission, the best relay is chosen from among all the nodes (including the source) that have possession of the packet data at that time.

During the original packet transmission, the mutual information across the source-destination channel is:

$$I_D = \log(1 + \rho g_{s,d}) \tag{4.6}$$

If a retransmission occurs, the combination of the two transmissions forms an equivalent channel between the source and the destination, whose mutual information is:

$$I^*_{itrs} = \frac{1}{2}\log\left[1 + \rho(g_{s,d} + g_{m^*,d})\right] \tag{4.7}$$

Denote the set of all nodes (in addition to the source) that have decoded the message of the source with $D(s)$. Using the law of total probability, the outage probability can be expressed as:

$$P_{out} = \sum_{t=1}^{M+1} Pr\Big\{ I^*_{itrs} < \frac{R}{2} \,\Big|\, I_D < R\,,\, |D(s)| = t \Big\} Pr\{I_D < R\}\ Pr\{|D(s)| = t\},$$
$$= \sum_{t}^{M+1} Pr\Big\{ I^*_{itrs} < \frac{R}{2} \,\Big|\, |D(s)| = t \Big\}\ Pr\{|D(s)| = t\} \tag{4.8}$$

The outage probability in (4.8) is computed for a rate R in case of successful source transmission and for a rate $\frac{R}{2}$ in case of incremental transmission due to information repetition.

The probability that exactly t nodes (including the source) know the message is given by [29],

$$Pr\{|D(s)| = t\} = \binom{M}{t-1}\ \exp\Big(-\frac{2^R-1}{\lambda_{s,m}\rho}\Big)^{t-1}\Big[1-\exp\Big(-\frac{2^R-1}{\lambda_{s,m}\rho}\Big)\Big]^{M-t+1} \tag{4.9}$$

By substituting (4.9) in (4.8) and obtaining the CDF of I^*_{itrs} one can find a closed form expression for the overall outage probability ($M \geq 1$):

$$P_{out,ITRS} = \sum_{t=1}^{t=M+1} F_W(\gamma) \times \binom{M}{t-1} \exp\Big(-\frac{\gamma}{\lambda_{s,m}}\Big)^{t-1}\Big(1-\exp\Big(-\frac{\gamma}{\lambda_{s,m}}\Big)\Big)^{M-t+1} \tag{4.10}$$

Where

$$F_W(\gamma) = \Big[t\sum_{k=1}^{t-1}\binom{t-1}{k}\frac{(-1)^k}{k}\Big(1+\frac{\exp(-\mu(k+1)\gamma)-1}{(k+1)} - \exp(-\mu\gamma)\Big)\Big] + \ t\Big(1-(\mu\gamma+1)\exp(-\mu\gamma)\Big) \tag{4.11}$$

$\gamma = \frac{2^R-1}{\rho}$ and for simplicity we let $\lambda_{s,d} = \lambda_{m^*,d} = \frac{1}{\mu}$. The details of the analysis are carried out in the Appendix.

Figure 4.3 depicts the outage probability of several relaying schemes for a network with two relays. The benchmark for direct transmission is a HARQ scheme with two rounds of transmission for which the following outage expression is developed:

$$P_{out,HARQ} = \Gamma(2, \mu\gamma) \tag{4.12}$$

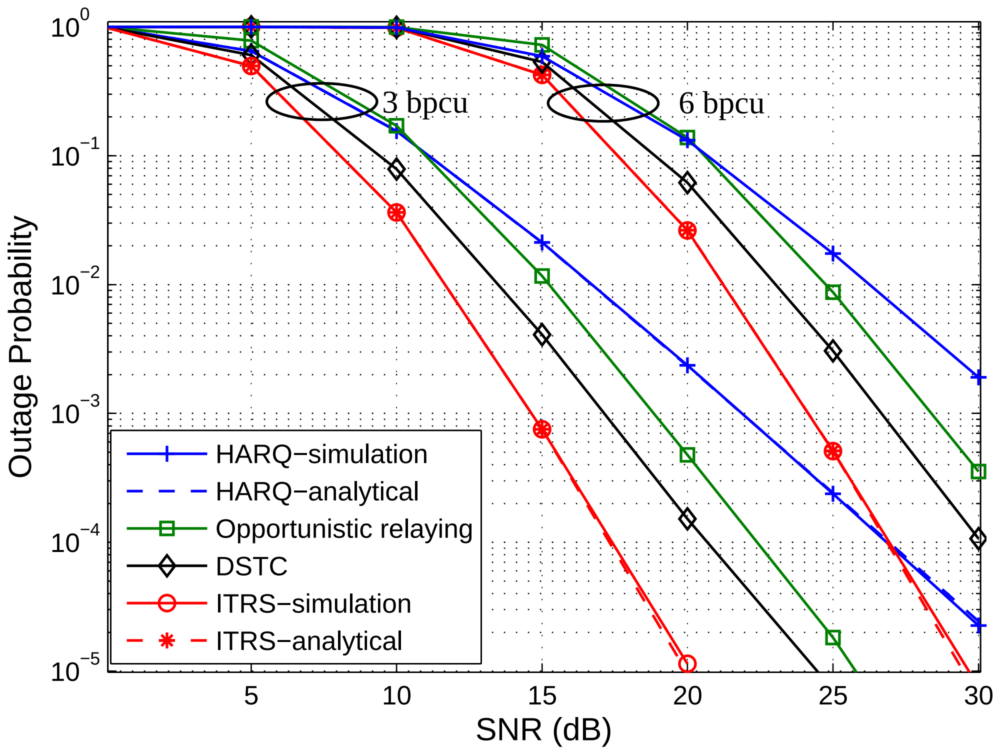

Figure 4.3. Outage performance of ITRS compared with distributed space-time coding, opportunistic relaying and HARQ non-cooperative transmission.

where $\Gamma(.)$ is the incomplete gamma function. ITRS performs better than the distributed space-time coding and opportunistic relaying schemes. Note the almost perfect match between the simulation results and the analytical expressions developed for HARQ and ITRS protocols.

We now calculate the throughput η, also known as effective rate or expected rate, for the ITRS protocol. This value has two contributing terms: for packets that are received in one try, or two tries, as shown below:

$$\eta = R\exp\left(-\frac{2^R-1}{\rho\lambda_{s,d}}\right) + \frac{R}{2}\left[\left(1-\exp\left(-\frac{2^R-1}{\rho\lambda_{s,d}}\right)\right)(1-P_{out})\right] \tag{4.13}$$

The first term is the average rate from the direct link and it occurs with the associated success probability. The second term is the average rate from HARQ with relay selection. Therefore, the rate is reduced to half since two blocks are used to transmit the same information. This second round of transmission is successful under the following two conditions:

1. The first round transmission failed.

2. The second round transmission with relay selection is successful.

We note that a somewhat similar notion of expected spectral efficiency was developed in [45] for a single-relay Amplify and Forward (AF) incremental relaying. The mapping $R \rightarrow \eta$ is highly nonlinear and one may choose R to maximize the throughput η.

The ITRS protocol requires $1 + \frac{\log(M+1)}{M+1}[1 - \exp(-\frac{2^R-1}{\rho\lambda_{s,d}})]$ bits of overhead per transmitting node. First, the destination broadcasts one bit of ACK/NACK. With probability $1 - \exp(-\frac{2^R-1}{\rho\lambda_{s,d}})$, the response is a NACK. The available relays and the source will respond with one-bit (known as Request To Send, or RTS). Finally, the destination will broadcast the index of the best node via $\log(M+1)$ bits. Asymptotically, this overhead is one bit per node per packet.

The above overhead analysis only counts the information bits in the feedback/control channels. It does not include the extra overhead that must be included in practice, for example a preamble. We also note that although we strive to design protocols with minimal overhead, this overhead will not affect the DMT results. In the high SNR regime, any constant overhead will diminish with respect to the channel capacity.

Remark 7 *If the source is excluded from the competition for relaying the expected rate will be given by*

$$\eta = R\exp\left(-\frac{2^R-1}{\rho\lambda_{s,d}}\right) + \frac{R}{2}\left\{\left[1-\exp\big(-\frac{2^R-1}{\rho\lambda_{s,d}}\big)\right]\right.$$

$$\left[1-\left(1-\exp\big(-\frac{2^R-1}{\rho\lambda_{m*,d}}\big)\right)^M\right](1-P_{out})\Big\} \tag{4.14}$$

This expected rate will approach (4.13) for large number of relays and high ρ.

4.3.2 DMT Analysis

In the high-SNR regime the performance of ITRS is described as follows, where we denote $(\cdot)^+ = \max\{\cdot, 0\}$.

Theorem 4 *The ITRS protocol achieves the following diversity-multiplexing tradeoff:*

$$d_{ITRS}(r) = (M+2)(1-r)^+ \tag{4.15}$$

which is equivalent to the optimal DMT of a system with one source node and M relay nodes [30, 8].

Proof: See the Appendix. □

Corollary 3 *ITRS with independent codebooks (type-II ARQ) achieves the same DMT.*

Proof: Since the identical codebooks achieve the MISO DMT bound, and independent codebooks will do no worse, then type-II ARQ will also achieve the same upper bound. □

The DMT of ITRS protocol with eight relays is shown in Figure 4.4. Also shown are other DF-based protocols, including the DDF of Azarian et al. [30, Theorem 6], and the DSTC of Laneman-Wornell [29], which has DMT equivalent to Bletsas et al. [26]. For fairness, we have compared our algorithm with a slight enhancement of DSTC by allowing its source to participate in the second phase of transmission.

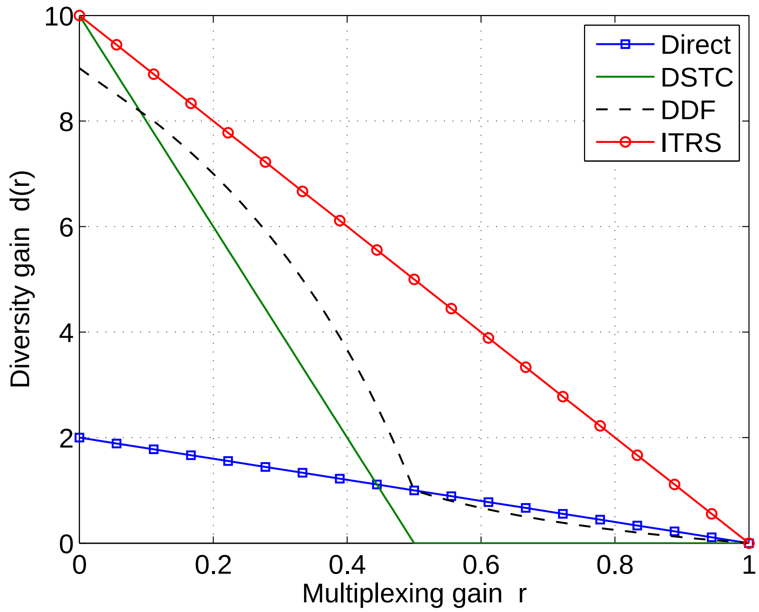

Figure 4.4. Diversity-multiplexing tradeoff of ITRS compared with distributed space-time coding, dynamic decode-and-forward and HARQ non-cooperative transmission. There are eight relays and the source-destination link exists.

For the non-cooperative benchmark, the DMT of HARQ signaling is shown, where a maximum diversity order of two is possible via packet combining [46, Corollary 3]. We see that ITRS has improved performance over previous protocols across all r, while requiring only limited feedback.

Protocol analysis corroborates the merits of allowing the source to compete for transmission in the relaying phase, which results in higher effective rate and diversity order $M+2$ (since $M+1$ nodes act as distributed antennas in the second phase).

Remark 8 *Consider the case where the destination node is limited to a type-I HARQ without diversity combining. Then the ITRS protocol still works, and achieves a slightly diminished maximum diversity order of* $M+1$*. Thus, ITRS can also be used in networks with very simple nodes without packet combining capabilities, e.g., wireless sensor networks.*

Remark 9 *When SNR is low, retransmissions are frequent. If, furthermore, relays are not abundant, the source may be called upon to re-transmit frequently, which is a strain on its power resources. Under these conditions, one may use a variation of ITRS, where the source will re-transmit* only *if all relays have failed to decode. This results in a slightly diminished maximal diversity of $M+1$, while extending the lifetime of the network.*

4.4 Two-Hop Relay Selection

When a direct path between the source and destination is unavailable, the relays must repeat the signal in a two-hop fashion. But it has been well-known that repeating the source's transmission limits the spectral efficiency in relay networks. The work in this section shows that in the presence of multiple relays, one may recover a good part of the rate loss with appropriate protocol design. We present a Multi-Hop Relay Selection (MHRS) protocol with attractive spectral efficiency, using non-orthogonal decode-and-forward signaling. The basic operation of the algorithm is as follows: in each time interval, the source transmits a new packet for the benefit of the relays. Simultaneously, the "best" relay re-transmits a packet for the benefit of the destination, interfering with the reception of other relays. All relays attempt to decode in the presence of interference, to be able to participate in the next round of transmission. The details of the MHRS protocol is described in Figure 4.5. It is assumed the channel remains constant within steps 2-4.

A sample timing diagram of the MHRS protocol is shown in Figure 4.6. The reception status of the relays is shown with a check or a cross. A check mark means successful decoding while a cross means failed decoding. Notice that whenever a relay transmits, due to the half-duplex constraint, it cannot receive. Therefore, in

1. The source transmits alone in the first time slot. Then, in each time slot:
2. Relays that successfully decode the source packet, declare their status to the destination via a one-bit RTS packet (which includes a pilot).
3. The destination picks the best relay and broadcasts its index.
4. The best relay retransmits its decoded packet, which the destination attempts to decode. At the same time, the source transmits a new packet.
5. The source packet and relayed packet combine at other relays. Relays attempt to decode new source packet in the presence of interference. Then continue to Step 2.

Figure 4.5. The Multi-Hop with Relay Selection (MHRS) protocol.

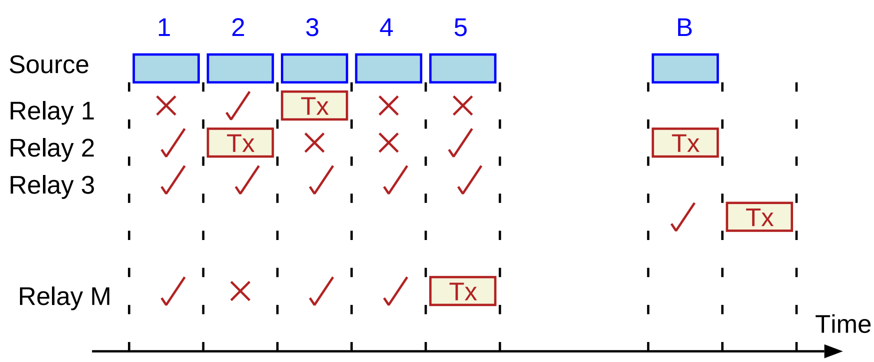

Figure 4.6. Frame structure of MHRS protocol.

the following time interval, it is operating at a disadvantage since it cannot peel-off the interference signal from the source transmission. Thus, in this sample path many relays are shown to fail in decoding immediately after transmission. In each time interval, the best available relay is called upon to transmit to the destination. Note that at the very end, the source is silent while a relay communicates the final packet with the destination. The overhead for control in the MHRS protocol is $1+\frac{\log(M-1)-1}{M}$ bits per node per packet (as mentioned in the previous section, this overhead does not include preamble and pilots of packets transmitted through the control channel).

Remark 10 *If no relay has successfully decoded the sources message, i.e. an empty decoding set, the source will notice the absence of the RTS (Request to Send) signal from all the relays, and can resend the packet. This is easy to implement, but in general not straight forward to analyze. The full analysis of this extension can be a subject of future work.*

In the following, we calculate a DMT upper bound and then present two decoding protocols at the relays, each with their own achievable DMT. The first decoding protocol is based on successive cancellation at the relays. The key result in successive cancellation is that, after each transmission by a relay, due to interference it cannot recover its own decoding diversity, thus it cannot contribute to the overall diversity any longer. It follows that across time, a family of DMT curves are produced with varying diversity. Every $B+1$ blocks, the maximal diversity is restored when the source transmits alone and relays are silent.

The interesting outcome of the family of DMT curves is that it allows variable error-protection. The overall data can be divided into several groups with varying error sensitivity. The most sensitive data is transmitted early, and enjoy the best

DMT, while other packets with lower sensitivity are transmitted later. To our knowledge, this is the first formal introduction of a variable-error protection scheme in relay networks, in the DMT sense.

If the relays have enough computational power, they may be able to jointly decode the two interfering signals. We show that a hybrid strategy, incorporating both successive cancellation and joint decoding, in part meets the DMT upper bound, and is superior to successive cancellation.

4.4.1 DMT Upper bound

We upper bound the DMT of our system by considering a hypothetical system where the individual relays are replaced with one MIMO relay. This will result in a system that operates as follows: during each interval, the best antenna for the relay-destination channel is used for relaying, while the other antennas listen to the source to receive the next frame. Since the new system is equivalent to perfect information exchange between relays, its performance upper bounds the performance of our system.

Using the above model, we have the following result:

Theorem 5 *The DMT of the multi-hop with relay selection (MHRS) protocol is upper bounded by:*

$$d^*(r) = (M-1)\left(1 - \frac{B+1}{B}r\right)^+ \tag{4.16}$$

Proof: According to [47, Lemma 1], the DMT of a channel with a single MIMO relay is bounded by the minimum of the source-relay and relay-destination DMT bounds. The source-relay DMT bound is the well-known SIMO bound

$$d^*_{SR}(r) = (M-1)(1-r)^+ \tag{4.17}$$

The relay-destination DMT is bounded by a MISO DMT with single-antenna selection out of $M-1$ available antennas, which has been recently reported in [48, Theorem 4.1].

$$d^*_{RD}(r) = (M-1)(1-r)^+ \tag{4.18}$$

The proof is completed by taking the minimum of the previous two bounds and taking into account the rate loss due to the causality of the relay. The optimal DMT is given by,

$$d^*(r) = (M-1)\left(1 - \frac{B+1}{B}r\right)^+ \tag{4.19}$$

□

4.4.2 Successive Cancellation DMT and Variable-Error Protection

Based on the protocol description, the received signals at the intermediate nodes and the destination are respectively given by:

$$\mathbf{y}_m = h_{s,m}\,\mathbf{x}_s + h_{m^*,m}\,\mathbf{x}_{m^*} + \mathbf{z}_m \tag{4.20}$$

$$\mathbf{y}_d = h_{m^*,d}\mathbf{x}_{m^*} + \mathbf{z}_d \tag{4.21}$$

Where $m = 1, \ldots, M$ and $m \neq m^*$. The transmission of the packets occurs in cycles. The source packets in each cycle are indexed by $b = 1, \ldots, B$. At the end of the cycle, the source stays silent for one period so that the last packet can be cleared to the destination. Then the entire process starts again (see Figure 4.6).

The mutual information of the channel between the "best" relay node and the destination is given by:

$$I^*_{mhrs} = \log\left(1 + \rho \max_{m \in D(s)} g_m\right) \tag{4.22}$$

The outage probability can be expressed as:

$$Pr\{I^*_{mhrs} < R\} = \sum_t Pr\{|D(s)| = t\} Pr\{I^*_{mhrs} < R \;\big|\; |D(s)| = t\} \tag{4.23}$$

Thus, the diversity-multiplexing tradeoff is governed by two probabilities: the relay decoding probability, and the outage probability conditioned on a decoding set $D(s)$.

Recall that from the viewpoint of a relay, there are two packets arriving simultaneously over the air. For example, when the source is transmitting packet b, another relay is transmitting packet $b-1$. We aim to calculate the probability that a given relay fails to decode the source packet b, denoted $P(O_b)$. Conditioned on having decoded packet $b-1$, the probability of relay outage for packet b is:

$$P(O_b|\bar{O}_{b-1}) \doteq \rho^{-(1-r)^+}$$

However, if the previous packet $b-1$ cannot be peeled off, the interference has the same order of magnitude as the signal, and thus

$$P(O_b|O_{b-1}) \doteq \rho^0$$

Now, we derive a equation using the law of total probability

$$\begin{aligned} P(O_b) &= P(O_b|\bar{O}_{b-1})P(\bar{O}_{b-1}) + P(O_b|O_{b-1})P(O_{b-1}) \\ &\doteq \rho^{-(1-r)^+}[1 - P(O_{b-1})] + \rho^0 P(O_{b-1}) \\ &= \rho^{-(1-r)^+} + P(O_{b-1})[1 - \rho^{-(1-r)^+}] \\ &\doteq \rho^{-(1-r)^+} + P(O_{b-1}) \qquad \text{for } \ r < 1 \end{aligned} \tag{4.24}$$

During transmission of the first packet, the relays listen to the source signal without interference, so $P(O_1) \doteq \rho^{-(1-r)^+}$. Then according to the above recursion, each relay will continue to decode with $P(O_b) \doteq \rho^{-(1-r)^+}$ until it is called upon to transmit. During a transmission interval by a relay node, it cannot listen to the source signal, so in the next interval, it will have to decode the source signal without knowledge of the interference. This task has outage probability proportional to ρ^0.

From this point onwards, the recursion shows that the relay will continue to experience outage proportional to ρ^0. In other words, the loss of diversity propagates in time. The diversity of all relays is restored at the end of the transmission cycle.

Now consider $Pr\{|D(s)| = t\}$. To have exactly t decoding relays, $M-t$ relays must be in outage. For the first packet, all relays decode without interference on i.i.d. channels, therefore

$$Pr\{|D(s)| = t\} \doteq \rho^{-(M-t)(1-r)^+} \qquad \text{for } \ b = 1$$

Subsequently, one of the relays is chosen to relay packet $b = 1$. This relay will lose its diversity for all subsequent packets, until the end of the cycle. Each relay that transmits will then stay out of the decoding set in successive blocks with probability proportional to ρ^0. Effectively, as we go through the packets, the number of available relays is reduced one-by-one.[1] Thus, for packet b, the probability that there are t relays ready to transmit is:

$$Pr\{|D(s)| = t\} \doteq \rho^{-(M-b+1-t)^+(1-r)^+} \tag{4.25}$$

Now we look at the destination outage conditioned on the decoding set.

$$\begin{aligned} Pr\left\{ I^*_{mhrs} < R \,\middle|\, |D(s)| = t \right\} &= \left(1 - \exp\left(-\frac{2^R - 1}{\rho \lambda_{m^*,d}}\right)\right)^t, \\ &= \left(1 - \exp\left(-\frac{\rho^{r-1}}{\lambda_{m^*,d}}\right)\right)^t \\ &\doteq \rho^{-t(1-r)^+} \end{aligned} \tag{4.26}$$

Substituting (4.25) and (4.26) in (4.23), we get:

$$Pr\{I^*_{mhrs}(b) < R\} \doteq \rho^{-(M-b+1)^+(1-r)^+} \tag{4.27}$$

---

[1]Please note that it is possible for the relays to return to the decoding pool, as is shown in Figure 4.6, but not with probability-1 asymptotically with SNR.

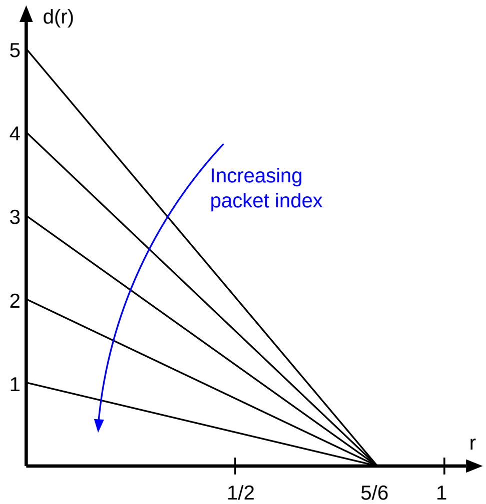

Figure 4.7. Family of DMT's for various data positions in the transmission cycle of a 5-relay MHRS network, resulting in a variable error protection system.

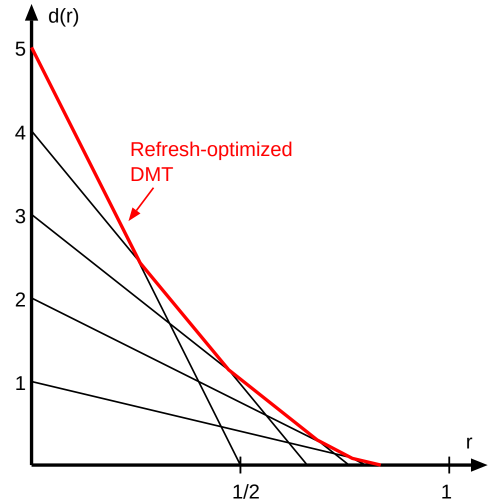

Figure 4.8. The DMT of the overall MHRS with five relays, where the refresh cycle is optimized separately for each r.

Finally, we have to take into account a fractional rate loss, because as seen in Figure 4.6, overall B blocks are transmitted in $B+1$ time intervals, due to causality requirement of the relays. Therefore, we must make the adjustment $r \to \frac{B+1}{B}r$. The final result can be described as follows:

Theorem 6 *For the MHRS protocol, the following diversity-multiplexing tradeoff is achievable for the packet b, where $b \in \{1, \ldots, B\}$.*

$$d(r,b) = (M-b+1)^{+}\left(1 - \frac{B+1}{B}\,r\right)^{+} \tag{4.28}$$

The variable-error protection strategy is an attractive feature of this system that allows a tailoring of transmission to the application requirements. Figure 4.7 shows the family of DMT's obtained in a MHRS protocol with five relays and $B = 5$.

In some applications, we may not be interested in a multiplicity of DMT's, thus the diversity across different packets b is dominated by the smallest diversity gain, i.e.,

$$d_{SC}(r) = \min_b d(r,b) = (M-B+1)^+ \left(1 - \frac{B+1}{B} r\right)^+$$

Note that in this expression, $B+1$ is a *refresh cycle* of the system, i.e., the period after which the source will transmit alone and will reset all the interferences at the relays. For an overall DMT above, since the two terms $M-B+1$ and $(B+1)/B$ move in opposite directions, one may optimize B for each multiplexing gain r so that the best diversity is obtained. This will lead to an overall DMT curve as shown in Figure 4.8.

4.4.3 MHRS Protocol with Hybrid Joint Decoding

In the previous section, we observed that successive cancellation in the MHRS protocol leads to error propagation and a gradual loss of diversity with increasing packet index. This loss arises from the reduced ability of the relays, after their own transmission, to correctly estimate and subtract the interference caused by other relays.

For better performance, we can employ a more powerful decoding technique at the relay. Whenever possible, the relays will decode by successive cancellation, but whenever that is not possible, the relays attempt an optimal joint decoding of the two arriving signals. Compare this with the method of Section 4.4.2, where the unavailable interfering signals were treated as noise. The more powerful method, denoted *MHRS with hybrid joint decoding*, improves the DMT of the MHRS protocol, and in fact meets the DMT upper bound up to a certain multiplexing gain, as we shall see in the sequel.

To calculate the DMT of MHRS with hybrid joint decoding, we use certain recent results on the so-called Z-channel. It is not difficult to see that our system

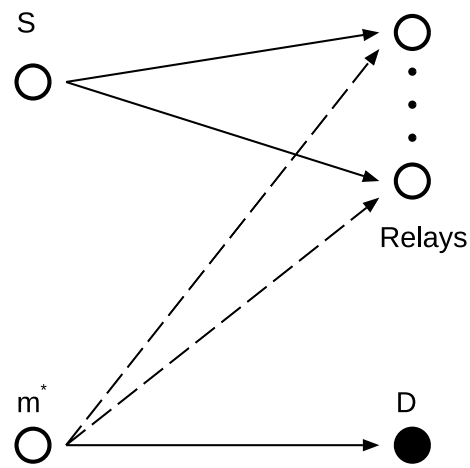

Figure 4.9. The multiple Z channels arising from the MHRS channel model.

model is a special case of the Z-channel, since the source is heard only by the relays, while the best relay in each interval is heard by both the destination *and* relays. (See Figure 4.9).

Recently, the DMT of the Z channel under general decoding was reported in [49]. Specializing the result of [49] to our channel model gives in the following relay outage diversity for a single-block decoding:

$$d_Z(r) = (\min\{(1-r), 2(1-2r)\})^+ \tag{4.29}$$

Since hybrid MHRS, for each packet detection, chooses the better of successive cancellation or joint decoding, the hybrid method must perform strictly better than either of its components. We will use this fact to find a lower bound to the DMT of the hybrid method.

The DMT of the non-hybrid method that always attempts successive cancellation was calculated in Section 4.4.2. In the following Lemma, we calculate the DMT of the non-adaptive method that always uses joint decoding.

Lemma 2 *The DMT of MHRS protocol under joint decoding is lower bounded by*

$$d_{JD}(r) = \Bigg(\min \Big\{ (M-1)(1 - \frac{B+1}{B} r), \\ \min_{t=0,\ldots,M-1} (2M-2-t) - \frac{B+1}{B} r(4M-4-3t) \Big\} \Bigg)^+ \quad (4.30)$$

Proof: We first consider an individual frame. When all relays (excluding the selected one) use joint decoding for detecting the message of the source and using the DMT of the Z-channel given in (4.29), we can write

$$Pr\{|D(s)| = t\} \doteq \Bigg(1 - \rho^{-(\min\{(1-r),2(1-2r)\})^+}\Bigg) \times \rho^{-(M-1-t)(\min\{(1-r),2(1-2r)\})^+} \\ \doteq \rho^{-(M-1-t)(\min\{(1-r),2(1-2r)\})^+} \quad (4.31)$$

Now, from (4.26), we have

$$Pr\Bigg\{ I^*_{mhrs} < R \,\Bigg|\, |D(s)| \Bigg\} \doteq \rho^{-t(1-r)^+} \quad (4.32)$$

Therefore, the asymptotic outage probability is expressed as

$$P_{out} \doteq \sum_t \rho^{-\Big(t(1-r)+(M-1-t)(\min\{(1-r),2(1-2r)\})\Big)^+} \quad (4.33)$$

Hence, the diversity order is given by

$$d_{JD}(r) = \Bigg(\min_t \Bigg\{ \min\Big[(M-1)(1-r), t(1-r) + 2(M-1-t)(1-2r)\Big] \Bigg\} \Bigg)^+ \\ = \Bigg(\min \Bigg\{ (M-1)(r-1) \min_{t=0,\ldots,M-1} (2M-2-t) - r(4M-4-3t) \Bigg\} \Bigg)^+ \quad (4.34)$$

We now consider the overall rate loss due to sending B frames in $B+1$ time intervals, therefore we must make the substitution $r \to \frac{B+1}{B} r$. This completes the proof of the Lemma. $\square$

At low spectral efficiencies, the above expression shows a distinct improvement over successive cancellation.

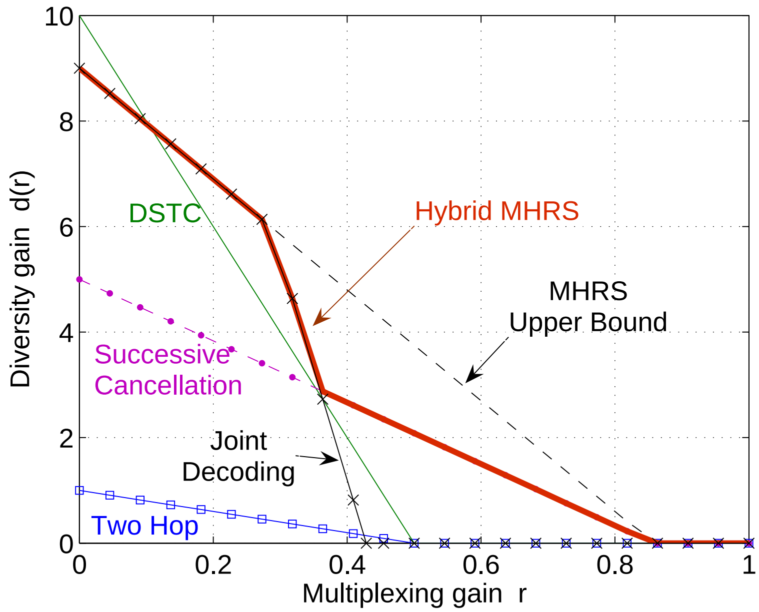

Figure 4.10. Diversity-multiplexing tradeoff of different protocols in a two-hop network with ten relays.

The overall DMT of the hybrid joint decoding method is bounded by the following.

Theorem 7 *The DMT of hybrid joint decoding is bounded below by:*

$$\max\left\{(M-B+1)^{+}\left(1-\frac{B+1}{B}r\right)^{+}, \min\left\{(M-1)(1-\frac{B+1}{B}r)^{+}, \right.\right.$$
$$\left.\left.\min_{t=0,\ldots,M-1}[(2M-2-t)-\frac{B+1}{B}r(4M-4-3t)]^{+}\right\}\right\} \tag{4.35}$$

Proof: The DMT of the adaptive (hybrid) protocol $d(r)$ is bounded below by the two DMT's belonging to the pure successive cancellation and joint decoding methods $d_{SC}(r), d_{JD}(r)$. It immediately follows that $d(r) \geq \max\{d_{SC}(r), d_{JD}(r)\}$. □

Figure 4.10 compares the diversity-multiplexing tradeoff of several DF-based protocols in a two-hop relay network with $M = 10$ and $B = 6$. Also, shown the MHRS protocol upper bound derived in Section 4.4.1. The hybrid MHRS protocol attains better DMT performance, across a large range of spectral efficiencies, compared with distributed space-time codes and opportunistic relaying[2].

Remark 11 *The hybrid MHRS protocol, as analyzed above, used a successive cancellation component whose minimal DMT was included in the analysis. The reader may recall, however, that the successive cancellation MHRS produces a family of unequal error protection DMT's. Thus, the hybrid strategy can also inherit the embedded DMT property of the successive cancellation. Since the hybrid DMT is influenced by the successive cancellation DMT in the high-rate regime, it follows that the embeddedness of the hybrid DMT is observable at high multiplexing rates. In the low multiplexing rates, all frames will experience the same diversity. To summarize, the embedded hybrid MHRS has the following achievable DMT, where b is the index of the packet.*

$$\max\left\{(M-b+1)^+\left(1-\frac{B+1}{B}r\right)^+, \min\left\{(M-1)(1-\frac{B+1}{B}r)^+,\right.\right.$$
$$\left.\left.\min_{t=0,\ldots,M-1}[(2M-2-t)-\frac{B+1}{B}r(4M-4-3t)]^+\right\}\right\} \tag{4.36}$$

4.5 Appendix

4.5.1 ITRS Outage Analysis

Computing the outage probability of ITRS hinges upon a closed-form expression for

$$Pr\left\{\frac{1}{2}\log\left[1+\rho(g_{s,d}+g_{m^*,d})\right] \leq \frac{R}{2}\ \middle|\ D(s)\right\} \tag{4.37}$$

whose calculation is the main goal of this appendix.

---

[2]Distributed space-time codes and opportunistic relaying have the same DMT, for compactness, only one of them is marked in Figure 4.10.

Lemma 3 *Consider an exponential random variable U with mean λ_u, and a random variable V which is maximum of a group of t i.i.d. exponential random variables, each with mean λ_g. The PDF of $W = U + V$ is given by*

$$f_W(w) = t \exp\Big(-\frac{w}{\lambda_u}\Big) \sum_{k=0}^{t-1} \binom{t-1}{k} \frac{(-1)^k}{(k+1)\lambda_u - \lambda_g} \Big(1 - \exp\big[-(\frac{k+1}{\lambda_g} - \frac{1}{\lambda_u})w\big]\Big) \tag{4.38}$$

Proof:

$$f_U(u) = \frac{1}{\lambda_u} \exp\Big(-\frac{u}{\lambda_u}\Big) \tag{4.39}$$

The CDF of a maximum of i.i.d. exponential random variables is given by:

$$F_V(v) = \Big(1 - \exp\Big(-\frac{v}{\lambda_g}\Big)\Big)^t \tag{4.40}$$

Differentiating with respect to v, we get

$$f_V(v) = \frac{t}{\lambda_g} \exp\Big(-\frac{v}{\lambda_g}\Big)\Big(1 - \exp\Big(-\frac{v}{\lambda_g}\Big)\Big)^{t-1} \tag{4.41}$$

Using the convolution integral, the PDF of W is given by:

$$f_W(w) = \frac{t}{\lambda_g \lambda_u} \exp\Big(-\frac{w}{\lambda_u}\Big) \int_0^w \exp\Big(-\frac{v}{\lambda_{eq}}\Big)\Big(1 - \exp\Big(-\frac{v}{\lambda_g}\Big)\Big)^{t-1} dv \tag{4.42}$$

where we have defined:

$$\frac{1}{\lambda_{eq}} \triangleq \frac{1}{\lambda_g} - \frac{1}{\lambda_u} \tag{4.43}$$

Using the binomial expansion, performing the integration and simplifying, one obtains,

$$f_W(w) = t \exp\Big(-\frac{w}{\lambda_u}\Big) \sum_{k=0}^{t-1} \binom{t-1}{k} \frac{(-1)^k}{(k+1)\lambda_u - \lambda_g} \Big(1 - \exp\big[-(\frac{k+1}{\lambda_g} - \frac{1}{\lambda_u})w\big]\Big) \tag{4.44}$$

□

Now, let $U = g_{s,d}$, $V = g_{m^*,d}$ and $W = U + V$. Then, the PDF of W is given by:

$$f_W(w) = t \exp\Big(-\frac{w}{\lambda_{s,d}}\Big) \sum_{k=0}^{t-1} \binom{t-1}{k} \frac{(-1)^k}{(k+1)\lambda_{s,d} - \lambda_{m,d}} \Big(1 - \exp\big[-(\frac{k+1}{\lambda_{m,d}} - \frac{1}{\lambda_{s,d}})w\big]\Big) \tag{4.45}$$

The CDF of this expression can be obtained via integration;

$$F_W(\gamma) = t \sum_{k=0}^{t-1} \binom{t-1}{k} \frac{(-1)^k}{(k+1)\lambda_{s,d} - \lambda_{m,d}} \times \int_0^{\gamma} \exp\Big(-\frac{w}{\lambda_{s,d}}\Big)\Big(1 - \exp\big[-(\frac{k+1}{\lambda_{m,d}} - \frac{1}{\lambda_{s,d}})w\big]\Big)\, dw \tag{4.46}$$

where as previously defined, $\gamma = \frac{2^R - 1}{\rho}$. Also, to simplify the calculations, we let $\lambda_{s,d} = \lambda_{m^*,d} = \frac{1}{\mu}$. After integration by parts and collecting terms, whose details are omitted for brevity, we obtain:

$$F_W(\gamma) = \Big[t \sum_{k=1}^{t-1} \binom{t-1}{k} \frac{(-1)^k}{k} \Big(1 + \frac{\exp(-\mu(k+1)\gamma) - 1}{(k+1)} - \exp(-\mu\gamma)\Big)\Big] + t\Big(1 - (\mu\gamma + 1)\exp(-\mu\gamma)\Big) \tag{4.47}$$

Finally, the overall outage probability for ITRS protocol is calculated by

$$P_{out,ITRS} = \sum_{D(s)} F_W(\gamma) \times \binom{M}{t-1} \exp\Big(-\frac{\gamma}{\lambda_{s,m}}\Big)^{t-1} \Big(1 - \exp\Big(-\frac{\gamma}{\lambda_{s,m}}\Big)\Big)^{M-t+1} \tag{4.48}$$

4.5.2 ITRS DMT Analysis

The effective rate of the ITRS protocol is η, as defined in (4.13), therefore the multiplexing gain must be defined with respect to η. However, one can equivalently use the nominal per-packet transmission rate R for the DMT calculations, since an inspection of (4.13) shows they are asymptotically equivalent:

$$\lim_{\rho\to\infty} \eta = R \tag{4.49}$$

Hence we can proceed and calculate the multiplexing gain based on R as defined in (4.1).

There are two ways to obtain the diversity order in terms of the multiplexing gain. One can either find an upper bound on the outage expression in (4.8) in a manner similar to [26, Theorem 3], or use the closed-form outage expression that is developed in this chapter, in asymptotic SNR form. The former approach is easier and we briefly mention the required steps.

First, from (4.9), at high SNR

$$Pr\Big\{|D(s)| = t\Big\} \doteq \rho^{(r-1)(M-t+1)} \left(\frac{1}{\lambda_{s,m}}\right)^{M-t+1} \tag{4.50}$$

Now, from (4.7), at high SNR

$$\begin{aligned} Pr\left\{ I^*_{itrs} < \frac{r\log\rho}{2} \middle| D(s) \right\} &= Pr\left\{ \log(1+\rho(g_{s,d}+g_{m^*,d})) \leq r\log\rho \;\middle|\; D(s) \right\}, \\ &\leq Pr\Big\{g_{s,d} \leq \rho^{r-1} \;\Big|\; D(s)\Big\}\; Pr\Big\{g_{m^*,d} \leq \rho^{r-1} \;\Big|\; D(s)\Big\}, \\ &\doteq \rho^{r-1}\rho^{t(r-1)}, \\ &= \rho^{(t+1)(r-1)} \end{aligned} \tag{4.51}$$

where we have used the results of Lemmas 2 and 3 of [26]. Combining (4.50) and (4.51), the diversity order of the ITRS protocol is given by

$$d_{ITRS}(r) = (M+2)\,(1-r)^{+} \tag{4.52}$$

CHAPTER 5

RELAY-ASSISTED INTERFERENCE NETWORKS

5.1 Introduction

In addition to historical importance in network information theory, a better understanding of the interference channel [50] is becoming practically important as well, since many current wireless communication systems are interference-limited. Examples include ad-hoc networks with peer-to-peer communications that lack infrastructure and hence transmission coordination, interference between adjacent networks in wireless LAN systems, as well as cognitive networks, where primary and secondary users transmit in the same frequency band.

The capacity of the interference channel in the most general case remains unknown, thus a number of partial approaches for investigating the interference channel have been pursued. One of the tools for understanding the behavior of multi-terminal networks is the *degrees of freedom* (DOF), also known as the *multiplexing gain* or the *pre-log factor*, which characterizes the scaling behavior of a network throughput at high signal-to-noise ratios (SNR). We formally define the degrees of freedom as follows:

$$DOF = \lim_{P \to \infty} \frac{C_s}{\log(\frac{P}{\sigma^2})} \tag{5.1}$$

where P is the power constraint at each source node, σ^2 is the noise variance at a destination and C_s is the network sum-rate capacity.

As an example, the maximum degrees of freedom of a two-user (single-antenna) Gaussian interference channel is equal to one [51]. In this chapter we investigate the effect of having a dedicated MIMO relay shared by several source-destination pairs on the degrees of freedom of such network. The main issue is whether with simple single-user decoding at the destinations, exploiting direct links is of a benefit.

A brief outline of related work is as follows. The new advances in network information theory have allowed characterizing the degrees of freedom of many networks. It is well known that the MIMO MAC and MIMO BC have full degrees of freedom [52, 53]. Since the point-to-point MIMO channel has full degrees of freedom, the latter result is equivalent to perfect cooperation for transmitters (MIMO MAC) and receivers (MIMO BC) at high SNR. Recently, through the interesting idea of *interference alignment*, new results have been obtained that characterize the degrees of freedom in interference networks. The idea of interference alignment is to pre-code the transmitted symbols of each user into multiple dimensions (can be time or frequency) such that at the desired receiver the interference signals are aligned in some dimensions. Hence, other dimensions are left interference-free. In a K user time-varying interference network, $\frac{K}{2}$ degrees of freedom are achieved almost surely [54]. However, for fixed channel coefficients and single antenna nodes, the Host-Madsen-Nosratinia conjecture of a DOF equals to one remains unsolved [51]. Thus, the gap remains untouched between the DOF upper bound of $\frac{K}{2}$ and the achievable DOF of 1.

The first attempt to study the effect of relaying on the degrees of freedom of the interference network was performed in [51] and [55]. A rather negative result was obtained. Cooperation over fading links between the sources, between the destinations, or both, cannot improve the degrees of freedom of an interference network. On the other hand, if perfect cooperation between sources (destinations) is assumed, the network can mimic a MIMO system with antennas co-located at the transmitting (receiving) side as we mentioned previously. In [56], the link between the sources and the link between the destinations are assumed to have a constant fading coefficient (static AWGN) while the links between the sources and the destinations have phase fading. It is shown that with full channel state information at all nodes, cooperation can help in increasing the throughput of a two-user interference channel close to rates achieved by a 2×2 MIMO system.

The Gaussian interference channel with a *dedicated* relay was explicitly introduced and its capacity studied by Şahin and Erkip [57, 58]. Recent works related to this area are [59, 60]. Other works tangentially related to this area include [61, 62] that despite apparent similarities are different in their essential features, due to a two-hop amplify-and-forward model.

The addition of a MIMO relay to an interference channel (direct links exist) gives rise to a network model that we denote *the interference MIMO relay channel (IMRC)*. In this chapter we obtain the achievable sum-rate, and consequently, the achievable degrees of freedom of the Gaussian IMRC with the source and destination nodes having one antenna each. Towards this end, we devise new combinations of coding strategies that are inspired by the coding schemes used in relay channels, as well as MIMO MAC and MIMO broadcast channels. We start the analysis for the two-user case and generalize it for multiple-user case. We show that in a K-user network with a MIMO relay, one can achieve exactly $\frac{K}{2}$ degrees of freedom same as a two-hop strategy. It is assumed that the relay has global channel state information, but *other nodes have only their own channel-state information.* We also study upper bounds on the degrees of freedom by specializing the recently developed upper bounds in [63]. Our study establishes the fact that the interference MIMO relay channel has $\frac{K}{2}$ degrees of freedom.

We then take the investigation one step further to consider the effect of the availability of abundant power at the relay. This is motivated by real-world scenarios

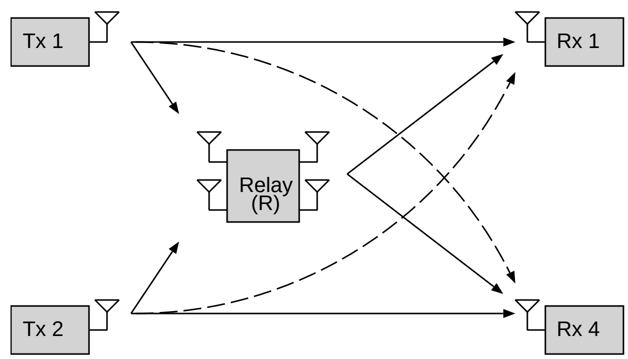

Figure 5.1. The two-User Interference MIMO Relay Channel

where a single relay tower, with easy access to power, is assisting many mobiles. We wish to understand whether additional power at the relay is helpful, and if so, what is the minimum amount of power needed to impart maximum degrees of freedom to the network. We find that if the relay has power proportional to $\mathcal{O}(P^2)$, it can impart the maximum K degrees of freedom to a K-user network whose users have power $\mathcal{O}(P)$, regardless of the number of users (c.f. of our definition of DOF in (5.1)).

5.2 System Model

Throughout the chapter, lower-case and upper-case boldface letters denote vectors and matrices, respectively. $\det(\mathbf{X})$ denotes the determinant of matrix $\mathbf{X}$ while $\mathbf{X}^\dagger$ and $\mathbf{X}^*$ denote the transpose and hermitian of $\mathbf{X}$, respectively. The norm of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\|$. $\log(.)$ stands for the base-2 logarithm, hence all rates are expressed in bits/channel use.

The *interference MIMO relay channel* (IMRC) is depicted in Fig. 5.1. Nodes 1 and 2 want to communicate independent messages W_1 and W_2 to their respective receivers, possibly with help from the relay (node R). The relay is assumed to be equipped with M antennas, where $M \geq 2$ while all other nodes have one antenna each. All links are subject to flat fading which remains constant during the transmission period. The channels from the sources to their corresponding destinations, from the sources to the relay and from the relay to destinations are denoted by the letters, f, g and h, respectively. A subscript ab is used to index the transmitting and receiving nodes, a and b, respectively.

The input-output relation of a Gaussian IMRC is given by:

$$y_1 = f_{11}x_1 + \mathbf{h}_{R1}^{\dagger}\mathbf{x}_R + f_{21}x_2 + z_1 \tag{5.2}$$

$$y_2 = f_{12}x_1 + \mathbf{h}_{R2}^{\dagger}\mathbf{x}_R + f_{22}x_2 + z_2 \tag{5.3}$$

$$y_R = \mathbf{g}_{1R}x_1 + \mathbf{g}_{2R}x_2 + z_R \tag{5.4}$$

where y_1, y_2 and y_R are the channel outputs at receivers 1, 2 and the relay, x_1, x_2 and $\mathbf{x}_R$ are the transmitted signals. The variables z_1, z_2 and z_R denote zero-mean, unit-variance additive white Gaussian noises at the receivers. We assume individual block power constraints on the transmitting nodes. Nodes 1 and 2 have equal transmit power constraint of P, i.e.

$$\sum_{i=1}^{n} \|x_k(i)\|^2 \leq nP_k, \qquad k = 1, 2 \tag{5.5}$$

where i is the symbol index within a block of n symbols.

We assume that the relay node has block power constraint of P_R, which may be different from P and will be specified in each instance in the sequel. The relay uses a decode-and-forward scheme [2] that includes linear pre-coding, in a manner to be explained shortly. The channel state information (CSI) knowledge assumptions are as follows. Transmitters 1 and 2 each have perfect knowledge about their own transmit-side CSI while receivers 1 and 2 have perfect knowledge of their receive-side CSI. Global channel knowledge is assumed at the relay. The relay is assumed to operate in full-duplex mode, i.e., it can receive and transmit at the same time. Throughout the chapter, we assume the input alphabets to be Gaussian. The average probability of error is defined as follows:

$$P_e^{(n)} = Pr\big[\{\hat{W}_1 \neq W_1\} \cup \{\hat{W}_2 \neq W_2\}\big] \tag{5.6}$$

where, $\hat{W}$ denotes an estimate of W. The rate of transmission from node k is $R_k = \frac{\log Q_k}{n}$, where Q_k is the size of the message transmitted by the $k'th$ node. A rate pair (R_1, R_2) is said to be achievable for the interference MIMO relay channel if there exist a sequence of codes $\big((2^{nR_1}, 2^{nR_2}), n\big)$ with average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$.

5.3 Coding strategies and Achievable Rates

Our approach to the problem is to use the relay in a way that minimizes the interference at the receivers. However, this task is highly nontrivial because the causality of the relay prohibits straight-forward interference cancelation. Therefore, sophisticated coding and power control strategies are needed to possibly manage the interference at the receivers.

Consider a transmission period of B blocks, each of n symbols. We assume that n is sufficiently large to allow reliable decoding. The sources and relay send sequences of $B-1$ messages ($W_1(b)$ and $W_2(b)$) over the channel in nB transmissions, where b denotes the block index, $b = 1, 2, ..., B-1$. The rate pair $(R_1\frac{B-1}{B}, R_2\frac{B-1}{B})$ approaches (R_1, R_2) as $B \to \infty$.

5.3.1 Encoding at the Sources

The source uses the super-position block Markov encoding technique devised in [2]. In particular at any block b,

$$X_1^{(b)} = U_1 + U_1' \tag{5.7}$$

$$X_2^{(b)} = U_2 + U_2' \tag{5.8}$$

where U_1 and U_1' are i.i.d Gaussian codebooks encoding the messages of the current and the previous blocks with powers $\chi(P)$ and $\psi(P)$, respectively, according to the power constraint

$$\chi(P) + \psi(P) = P \tag{5.9}$$

Similar definitions hold for the signal components transmitted by node 2, U_2 and U_2'.

5.3.2 Decoding and Re-encoding at the Relay

We use a space division multiple-access (SDMA) approach to communicate between nodes 1, 2 and the MIMO relay. Therefore, both sources transmit simultaneously and the MIMO relay attempts decoding both signals. At the end block b, given that the relay decoded both messages $W_1(b-1)$ and $W_2(b-2)$ correctly, it can decode the messages $W_1(b)$ and $W_2(b)$ of both users while achieving a $DOF = 2$. This can be achieved by a zero-forcing strategy, as long as the relay has no fewer antennas as the number of transmit nodes, and is made possible by the independence of the users' channels to the relay that is a result of spatial separation. The sum-rate constraint for correct decoding at the relay is given by [64, section 10.1]:

$$R_1 + R_2 \leq \log\det\big(\mathbf{I}_2 + \mathbf{G}\mathbf{K}_x\mathbf{G}^*\big) \tag{5.10}$$

where $\mathbf{G} = [\mathbf{g}_{1R}\ \mathbf{g}_{2R}]$, $\mathbf{K}_x = diag\big(\chi(P), \chi(P)\big)$, and $\mathbf{I}_2$ is the 2×2 identity matrix.

We now describe the encoding process at the relay. Ideally, we would have liked the relay to cancel the whole interference seen by each receiver. However, due to causality, the relay can only battle the interference arising from signals that it has already decoded. Thus, even if everything else can be accomplished perfectly, we can expect that not all of the interferences will be canceled. The question is, if interferences cannot be fully removed, then how to manage the remaining interference so that a good result may be obtained in terms of the degrees of freedom. We will address this problem via power allocation policies at the sources and at the relay, as will be shown in the sequel.

The channel from the relay to both destinations is similar to a Gaussian MIMO broadcast channel whose capacity region has been recently determined [65]. To help in canceling the interference, the relay uses a modified zero-forcing beamforming (ZF-BF) strategy [66]. ZF-BF achieves the maximum degrees of freedom of the sum-rate capacity of a Gaussian MIMO BC, although it is in general suboptimal compared to the capacity-achieving dirty-paper coding (DPC) strategy. The relay constructs and transmits the following signal:

$$\mathbf{x}_R^{(b)} = u_1'\mathbf{t}_1 + u_2'\mathbf{t}_2 \tag{5.11}$$

where $\mathbf{t}_1$ and $\mathbf{t}_2$ are 2×1 complex unitary beamforming vectors. Proper selection of beamforming vectors (magnitudes and phases) allows partial suppression of interference at the receivers. Simultaneously, t_1 and t_2 can be selected to allow beamforming (coherent combination) of the relay signal with the cooperative components of the signals transmitted by the sources at the current block. For simplicity, we assume the relay divides its power P_R equally between the two codebooks U_1' and U_2'.

5.3.3 Decoding at the Destinations

Given the structure of the signal formed by the relay, we re-write (5.2) and (5.3) as follows:

$$y_1^{(b)} = f_{11}u_1 + (f_{11} + \mathbf{h}_{R1}^\dagger \mathbf{t}_1)u_1' + (f_{21} + \mathbf{h}_{R1}^\dagger \mathbf{t}_2)u_2' + f_{21}u_2 + z_1 \tag{5.12}$$

$$y_2^{(b)} = f_{12}u_1 + (f_{12} + \mathbf{h}_{R2}^\dagger \mathbf{t}_1)u_1' + (f_{22} + \mathbf{h}_{R2}^\dagger \mathbf{t}_2)u_2' + f_{22}u_2 + z_2 \tag{5.13}$$

Therefore, the relay selects $\mathbf{t}_1$ and $\mathbf{t}_2$ such that $\mathbf{h}_{R2}^{\dagger}\mathbf{t}_1 = -f_{12}$ and $\mathbf{h}_{R1}^{\dagger}\mathbf{t}_2 = -f_{21}$. This will cancel part of the interference seen by each receiver, thus the received signals are modified to:

$$y_1^{(b)} = f_{11}u_1 + (f_{11} + \mathbf{h}_{R1}^{\dagger}\mathbf{t}_1)u_1' + f_{21}u_2 + z_1 \tag{5.14}$$

$$y_2^{(b)} = f_{12}u_1 + (f_{22} + \mathbf{h}_{R2}^{\dagger}\mathbf{t}_2)u_2' + f_{22}u_2 + z_2 \tag{5.15}$$

Receivers 1 and 2 can use Willems's backward decoding to decode their intended signals [6]. Backward decoding imposes decoding delays, however, it simplifies the analysis compared to list decoding or window decoding [22]. Backward decoding starts from block B. The receivers have interference-free channels to decode $u_1^{(B-1)}$ and $u_2^{(B-1)}$. In block $B-1$, they pre-subtract the components of $u_1^{(B-1)}$ and $u_2^{(B-1)}$ before attempting to decode $u_1^{(B-2)}$ and $u_2^{(B-2)}$. Therefore, at any block b the received signals can be further reduced to:

$$y_1^{(b)} = (f_{11} + \mathbf{h}_{R1}^{\dagger}\mathbf{t}_1)u_1' + f_{21}u_2 + z_1 \tag{5.16}$$

$$y_2^{(b)} = f_{12}u_1 + (f_{22} + \mathbf{h}_{R2}^{\dagger}\mathbf{t}_2)u_2' + z_2 \tag{5.17}$$

To simplify the analysis, one can convert the above channel into *standard form* without changing the capacity region [50].

$$y_{1,s}^{(b)} = u_{1,s}' + \sqrt{\alpha}\, u_{2,s} + z_{1,s} \tag{5.18}$$

$$y_{2,s}^{(b)} = \sqrt{\beta}\, u_{1,s} + u_{2,s}' + z_{2,s} \tag{5.19}$$

The subscript s indicates that the variables are in standard form. The relations with the original channel are as follows:

$$u_{1,s}' = (f_{11} + \mathbf{h}_{R1}^{\dagger}\mathbf{t}_1)u_1' \tag{5.20}$$

$$u_{2,s}' = (f_{22} + \mathbf{h}_{R2}^{\dagger}\mathbf{t}_2)u_2' \tag{5.21}$$

and,

$$\sqrt{\alpha} = \frac{f_{21}}{(f_{22} + \mathbf{h}_{R2}^{\dagger}\mathbf{t}_2)} \tag{5.22}$$

$$\sqrt{\beta} = \frac{f_{12}}{(f_{11} + \mathbf{h}_{R1}^{\dagger}\mathbf{t}_1)} \tag{5.23}$$

The power constraints on each of $u'_{1,s}$ and $u'_{2,s}$ are given by:

$$\psi_{1,s}(P) = \left\| f_{11}\sqrt{\psi(P)} + \mathbf{h}_{R1}^{\dagger}\mathbf{t}_1\sqrt{\frac{P_R}{2}} \right\|^2 \tag{5.24}$$

$$\psi_{2,s}(P) = \left\| f_{22}\sqrt{\psi(P)} + \mathbf{h}_{R2}^{\dagger}\mathbf{t}_2\sqrt{\frac{P_R}{2}} \right\|^2 \tag{5.25}$$

Note that the amplitudes of part of the signal components from the sources and the relay combine at the destination due to the beamforming effect. Based on the standard form and the corresponding power constraints, one can see that receivers 1 and 2 can decode their respective messages W_1 and W_2 reliably if:

$$R_1 \leq log\left(1 + \frac{\psi_{1,s}(P)}{\|\alpha\|\chi(P) + 1}\right) \tag{5.26}$$

$$R_2 \leq log\left(1 + \frac{\psi_{2,s}(P)}{\|\beta\|\chi(P) + 1}\right) \tag{5.27}$$

We proceed to specify two power allocation strategies and explore the corresponding achievable degrees of freedom.

- **Power Policy (a)**

  We let $\chi(P) = \psi(P) = \frac{P}{2}$ and $P_R = P$. According to this power allocation, the multi-access part of the channel according to (5.10) achieves $DOF = 2$. However, according to (5.26) and (5.27), the signal and interference have the same power order and hence a $DOF = 0$ is achieved. Therefore, the degrees of freedom of the network in this case is zero. Clearly this is not a desirable solution.

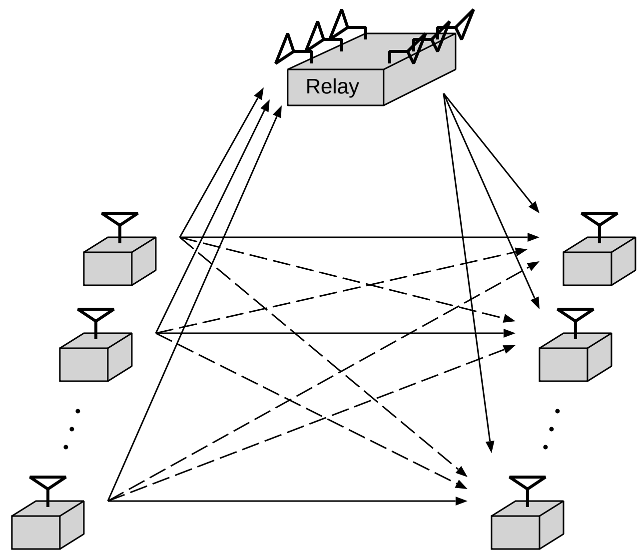

Figure 5.2. Relay for K-user interference network.

- **Power Policy (b)**

  Now we explore an asymmetric power allocation policy characterized by $\chi(P) = \sqrt{P}$, $\psi(P) = P - \sqrt{P}$ and $P_R = P$. Therefore, the cooperative information - also known as the resolution information- has a higher power order than the information of the current block of transmission. It is clear that $DOF = 1$ is achieved on the multi-access side of the channel. On the other hand, each of (5.26) and (5.27) provides a pre-log factor of $\frac{1}{2}$ leading to a sum-rate having $DOF = 1$ for the direct link with relaying. Therefore, an overall $DOF = 1$ is achieved.

The results of this section is summarized in the following theorem:

Theorem 8 *Under the power allocation policies considered above, the Interference MIMO Relay Channel (IMRC) can achieve degrees of freedom of zero and one.*

5.3.4 The K-user Interference Network with MIMO Relay

The previous coding strategies can easily be extended to an interference network where K users transmit simultaneously and a MIMO relay having $M \geq K$ antennas helps all K nodes in their transmission. Details are omitted for brevity and the result is summarized as follows:

Corollary 4 *A K-user Interference MIMO relay network achieves zero degrees of freedom under power policy (a) and* $\frac{K}{2}$ *degrees of freedom under power policy (b).*

Corollary 4 leads to the following insight. Compared to a time-sharing strategy that achieves $DOF = 1$, a MIMO relay increases the degrees of freedom to $\frac{K}{2}$. This can be achieved by exploiting the direct links under power policy (b) or more simply using a two-hop strategy (a MIMO MAC followed by a MIMO BC).

While we consider the case of full-duplex relay, one can devise similar signaling strategies for the half-duplex case. However, the block-Markov coding is not required. A brief description of a possible coding scheme is given as follows. The sources transmit all the time. However, they divide each block of their transmission into two halves. Each source node transmits the same message in the two halves using i.i.d. Gaussian codebooks. During the second half, the relay transmits and manages the interference as discussed above in the full-duplex case. At the destinations, the received signals at the first and second halves form two Gaussian parallel channels, the first sees interference while the other is interference-free. It can be easily shown that the maximum degrees-of-freedom of this scheme is $\frac{K}{2}$. Hence, again a two-hop strategy suffices to achieve a $\frac{K}{2}$ degrees-of-freedom. However, exploiting the direct links provides an increase in the throughput compared to two-hop communications for all signal-to-noise-ratios.

5.4 Upper bounds on the Degrees of Freedom

In a recent work, an elegant way to find upper bounds on fully connected interference and X networks with relays and feedback has been presented in [63]. The upper bound on a $S \times R \times D$ fully connected network can be specialized to the network we study in this chapter, where S, R, and D refer to the number of sources, relays and destinations in the network. A fully connected network means that there is a message from every source to every destination that needs to be communicated. For completeness, we will first state the main result on the upper bound on the degrees of freedom of the $S \times R \times D$ network.

Theorem 9 *[63] If $\mathcal{D}$ represents the degrees of freedom region of the $S \times R \times D$ node X network, then the total degrees of freedoms can be upper bounded as follows:*

$$\max_{[(d_{i,j})] \in \mathcal{D}} \sum_{j=1}^{S} \sum_{i=S+R+1}^{S+R+D} d_{i,j} \leq \frac{SD}{S+D-1}$$

Note that in [63], the authors derives upper bound not only on the degrees of freedom of the $S \times R \times D$ but on the whole degrees of freedom *region*. The interested reader can refer to [63] for further details on the proof techniques.

Considering the K user interference network, the following corollary from [63] gives the exact degrees of freedom of this network.

Corollary 5 *[63] "Consider a fully connected K user interference network with R relays, where all the channel coefficients are time-varying/frequency-selective with values drawn randomly from a continuous distribution with support bounded below by a non-zero constant. Let all nodes be full-duplex allowing noisy transmitter/receiver cooperation. Also, let the source and relay nodes receive perfect feedback from all nodes. Then the interference network has $\frac{K}{2}$ degrees of freedom."*

The bounds in the previous theorem and the corollary are applicable to the MIMO relay. Since the proof of the converse assumes full cooperation between the distributed R relay nodes (see observation 3 in [63]). Now, the achievable degrees of freedom for the interference MIMO relay channel was shown to be $\frac{K}{2}$. Moreover, feedback and time/frequency selectivity of the channel do not reduce the degrees of freedom of the channel. Then one concludes the following.

Corollary 6 *The degrees of freedom of the interference MIMO relay network is given by* $\frac{K}{2}$*.*

Therefore, the ineffectiveness of exploiting the direct link to increase the degrees of freedom is not an artifact of the coding strategy used. It is a fundamental result.

5.5 Abundant Power at the Relay

Given that the relay can perform a pre-coding strategy that does not cause interference at the respective receivers at high SNR. One is interested to investigate a case where the relay can play a role in increasing the degrees of freedom of the interference network. Consider the effect of abundant power at the relay, specifically, assume $\chi(P) = \psi(P) = \frac{P}{2}$ while at the relay we have $P_R = P^2$ (or in general $\mathcal{O}(P^2)$). In this case, the network will achieve its maximum possible degrees of freedom of two, thanks to the pre-coding strategy employed by the relay, which allows the power of the relay not to cause interference at any node. Note that our definition of the degrees of freedom in (5.1) concentrates on the power of information-bearing nodes, thus allowing us to study the effect of abundant excess power at the relay for this special case.

Therefore, K degrees of freedom can be achieved using a relay that enjoys power proportional to P^2, but the power is independent on K. Thus, one relay station with enough antennas and abundant power can significantly boost the network throughput in the high SNR regime.

In the following, we shall refer to the power policy discussed in this section as **Power Policy (c)**.

5.6 Numerical Results

We corroborate the analysis by the following numerical example of an interference MIMO relay channel. The following setup is considered:

- Two-user channel and the relay has two antennas, i.e., $K = M = 2$.
- The noise variance at all nodes σ^2=1.
- The magnitude of channel coefficients are selected as: $h_{13} = 1.2$, $h_{14} = 0.5$, $h_{23} = 0.5$, $h_{24} = 1.2$, $\mathbf{h}_{15}^{\dagger} = (0.6 \;\; 1.2)$, $\mathbf{h}_{25}^{\dagger} = (1 \;\; 0.5)$, $\mathbf{h}_{53}^{\dagger} = (0.5 \;\; 1)$ and $\mathbf{h}_{54}^{\dagger} = (1 \; 2)$.

According to the selected channel gains, we have from (5.22), $\alpha = 0.1455$ and $\beta = 0.2157$. Note that the original interference channel provides standard form channel gains of $a = b = 0.1736$ (see *e.g.* [67]). Thus, both with and without the MIMO relay, the interference is considered weak/moderate. This is the case where the capacity region of the interference channel is unknown and where a form of relaying will be of greater impact on the capacity [55].

Figure 5.3 depicts the sum-rate of five different schemes. Curve (1) is the case where no relay is present and the two sources have ideal cooperation leading to a

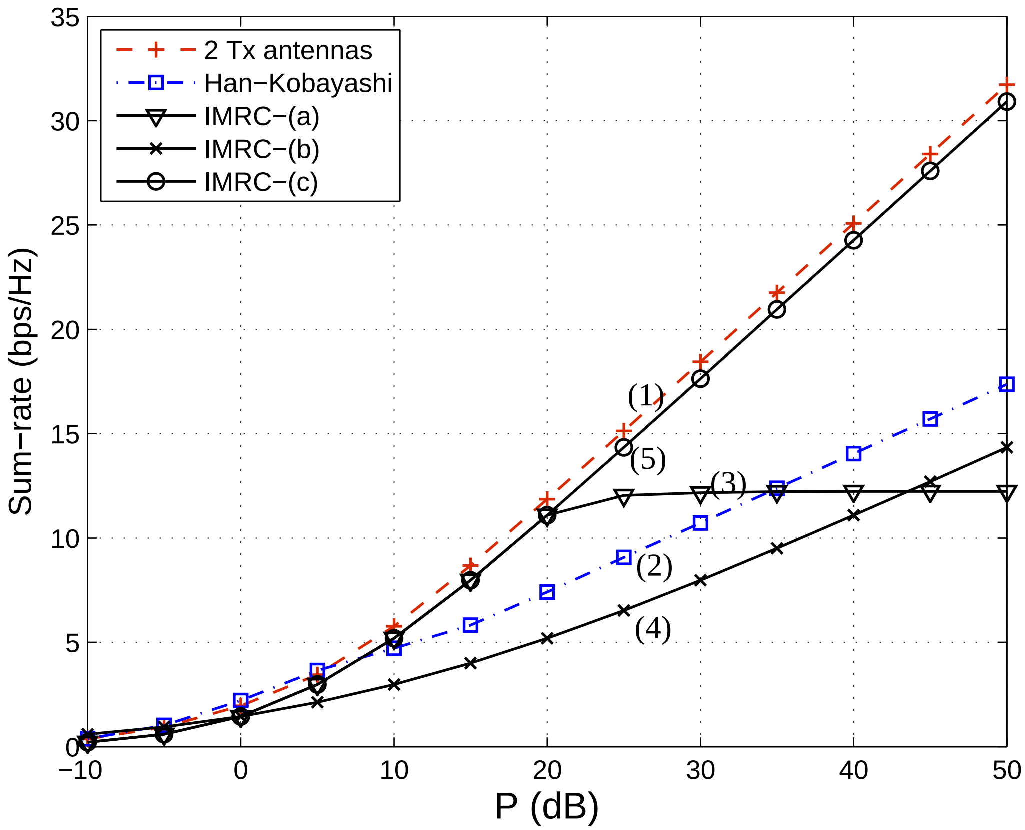

Figure 5.3. Network throughput versus transmit power of each source P: 2-user example.

MIMO BC model. The optimal power allocation and hence the sum-rate capacity are computed according to *Algorithm 2* of [68]. Curve (2) depicts the best known achievable sum-rate for the interference channel (with no relay present) using the Han-Kobayashi coding scheme. This scheme involves rate splitting, joint decoding at the receivers and moreover it includes a time-sharing random variable that switches between time-division transmission and simultaneous transmission. The cardinality of the time-sharing parameter is set to two and furthermore the power allocation of the rate-splitting scheme is optimized. This corresponds to curve 4 of [69]. Curves 3, 4 and 5 are the computable sum-rate of the interference MIMO relay channel under power policies (a), (b) and (c), respectively. The slopes of the these three curves verify that degrees of freedom of zero, one and two are indeed achievable. Also notice that although the power policy (a) provides $DOF = 0$ but the relay enhances the network throughput significantly at medium signal-to-noise-ratios. Throughput analysis at finite SNR is outside the scope of this chapter. We finally emphasize here that we use independent decoding at the nodes and we do not fully optimize the power allocation strategies at the sources and the relay.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

This chapter summarizes the contributions of this dissertation and provides some possible avenues for future directions in the area of relay networks. The findings of this dissertation also appear in [70, 71, 72, 73, 74, 75].

In Chapter 3 we present a new relay channel model for the three node network called the relay channel with private messages (RCPM). In addition to the traditional communication from source to destination (assisted by the relay), the source has a private message for the relay, and the relay has a private message for the destination. Achievable rate regions as well as outer bounds on the capacity region are obtained for the discrete memoryless relay channel with private messages. The Gaussian versions of this channel are also studied and achievable rate regions are characterized. Numerical results are provided that give insights into the trade-offs between private messaging and relayed messaging in this hybrid three-node network. We show that many of the previous results for the original relay and relay-broadcast channels can be recovered as special cases of the results presented in this chapter.

Chapter 4 addresses the spectral efficiency loss or the multiplexing loss that occurs in relay networks due to causality of relays and the half-duplex constraint. We devise spectrally-efficient relay selection techniques with limited feedback in decode-and-forward (DF) relay networks. When a direct link exists between the source and destination, we propose an Incremental Transmission Relay Selection (ITRS) protocol that leverage a H-ARQ mechanism. In the absence of a direct link, we propose a Multi-

Hop Relay Selection (MHRS) and several efficient non-orthogonal DF protocols are produced and analyzed. We use the powerful diversity-multiplexing tradeoff analysis tool to examine the performance of the algorithms proposed. The ITRS achieves the MISO DMT bound, in addition we derive an outage expression that is valid at any SNR. We develop upper and lower bounds for the diversity-multiplexing tradeoff of the MHRS. We show that the developed lower bound meets the upper bound over a portion of the multiplexing gains. The proposed protocols improve over existing methods for half-duplex DF relay systems in block-fading channels.

In Chapter 5 we characterize the high-SNR sum-rate behavior of the interference channel in the presence of a dedicated MIMO relay. The relay is used to manage the interference at the receivers. Using a number of hybrid encoding strategies and power allocation policies, we obtain non-asymptotic achievable sum-rates, subsequently leading to achievable degrees of freedom. The results are generalized from a two-user to a K-user network. The achievable degrees of freedom are tight against a recently developed upper bound. Our main result is that only $\frac{K}{2}$ degrees of freedom are achievable in an interference channel with MIMO relay, assuming global channel knowledge at the relay *but not at other nodes*. Thus, appropriate signaling in a two-hop scenario captures the degrees of freedom gains without the need for the direct links. We also investigate the case where the relay (unlike other nodes) has access to abundant power, showing that when the sources have power P and the relay is allowed power proportional to $\mathcal{O}(P^2)$, the full K degrees of freedom are available to the network.

To summarize, the major contributions of this dissertation are:

- Proposing a bandwidth efficient three-node channel model, known as the relay channel with private messages, and analyzing its capacity.

- Proposing relay selection algorithms that achieve the best known diversity-multiplexing tradeoff for the class of decode-and-forward relays in a multiple-relay network.

- Proposing the use of a MIMO relay to manage the interference in interference networks. We report the first case of achieving the full degrees-of-freedom (pre-log sum-capacity factor) of an interference network.

We forsee some extensions to the research topics proposed in this dissertation. The development of tight outer bounds for the Gaussian relay channel with private messages remains an open problem. Also, one can include a common message from the source to relay and destination to the channel model and perform a capacity analysis. An interesting direction is to consider the fading relay channel with private messages and study power allocation policies at the source and relay. Another promising direction is to use the relay channel with private messages as a building block of larger networks and study the rate scaling laws for these network architectures.

Considering the relay selection framework, few extensions are in order. We employ random coding arguments throughout this dissertation. A natural direction which has a strong practical impact, is to design channel codes for the relay selection protocols developed in this dissertation. Another extension is to consider the case of MIMO nodes and perform a DMT analysis. Design issues related to higher layers can also be explored, including selection of the number of transmission blocks for the multi-hop relay selection algorithm and designing signaling protocols for exchanging the feedback information over a control channel.

In the relay-assisted interference network, our analysis concentrates on the high SNR behavior of the network throughput. Many parameters can be further optimized for non-asymptotic SNRs. More complex coding/decoding techniques can also be employed, for example, a modified Han-Kobayashi scheme (in the presence of the MIMO relay) that combines rate-splitting, time-sharing (TDM), relaying and joint decoding at the receivers. Also, one can study the half-duplex version of this network under Gaussian and block-fading channel models.

BIBLIOGRAPHY

[1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

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

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley-Interscience, second edition, 2006.

[5] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inform. Theory*, vol. 27, no. 3, pp. 292–298, May 1981.

[6] F. M. J. Willems, *Informationtheoretical Results for the Discrete Memoryless Multiple Access Channel*, Ph.D. thesis, Katholieke Univ. Leuven, Leuven, Belgium, October 1982.

[7] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[8] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[9] I. E. Telatar, "Capacity of multiantenna Gaussian channels," *Eur. Trans. Commun.*, vol. 10, no. 6, pp. 585–595, 1999.

[10] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 3, pp. 900–928, 2007.

[11] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct 2007.

[12] A. Reznik, S. R. Kulkarni, and S. Verdu, “Broadcast-relay channel: capacity region bounds,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2005, pp. 820–824.

[13] R. Dabora and S. D. Servetto, “Broadcast channels with cooperating decoders,” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, 2006.

[14] L. Lai, K. Liu, and H. El Gamal, “The three node wireless network: Achievable rates and cooperation strategies,” *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 805–828, 2006.

[15] C.-M. Zing, F. Kuhlmann, and A. Buzo, “Achivability proof of some multiuser channel coding theorems using backward decoding,” *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1160–1165, 1989.

[16] H.-F. Chong, M. Motani, and K. H. Carg, “Generalized backward decoding strategies for the relay channel,” *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 394–401, 2007.

[17] T. M. Cover, “Comments on broadcast channels,” *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2524–2530, 1998.

[18] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[19] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[20] Raymond W. Yeung, *A First Course in Information Theory*, Springer-Verlag New York LLC, 2002.

[21] A. A. El Gamal and E. C. van der Meulen, “A proof of Marton’s coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 27, pp. 120–122, 1981.

[22] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[23] T. Berger, “Multiterminal source coding,” *Lecture notes presented at the 1977 CISM Summer School, Udine, Italy*, July 1977.

[24] B. E. Hajek and M. B. Pursley, "Evaluation of an acheivable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 36–46, 1979.

[25] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[26] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[27] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, Oct. 2004.

[28] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE J. Select. Areas Commun.*, vol. 23, no. 1, pp. 7–18, Jan. 2005.

[29] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

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

[31] A. Tajer and A. Nosratinia, "Opportunistic cooperation via relay selection with minimal information exchange," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[32] A. Bletsas, H. Shin, and M. Z. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Transactions on Wireless Communications*, vol. 6, no. 9, pp. 3450–3460, Sep. 2007.

[33] A. Nosratinia and T. E. Hunter, "Grouping and partner selection in cooperative wireless networks," *IEEE J. Select. Areas Commun.*, vol. 54, no. 4, pp. 760–378, Feb. 2006.

[34] Z. Lin, E. Erkip, and A. Stefanov, "Cooperative regions and partner choice in coded cooperative systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1323–1334, July 2006.

[35] E. Beres and R. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 118–127, Jan 2008.

[36] A. Bletsas, H. Shin, and M. Z. Win, "Outage optimality of opportunistic amplify-and-forward relaying," *IEEE Communication Letters*, vol. 11, no. 3, pp. 261–263, March 2007.

[37] Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay netwroks: Optimal power allocation versus selection," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 3114–3123, Aug. 2007.

[38] Y. Zhao, R. Adve, and T. J. Lim, "Symbol error rate of selection amplify-and-forward relay systems," *IEEE Communication Letters*, vol. 10, no. 11, pp. 757–759, Nov. 2006.

[39] I. Krikidis, J. Thompson, S. McLaughlin, and N. Goertz, "Amplify-and-forward with partial relay selection," *IEEE Communication Letters*, vol. 12, no. 4, pp. 235–237, April 2008.

[40] J. Cai, X. Shen, J. W. Mark, and A. S. Alfa, "Semi-distributed user relaying algorithm for amplify-and-forward wireless relay netwroks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1348–1357, April 2008.

[41] D. S. Michalopoulos, G. K. Karagiannidis, T. A. Tsiftsis, and R. K. Mallik, "Distributed transmit antenna selection (DTAS) under performance or energy consumption constraints," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1168–1173, April 2008.

[42] C. K. Lo, R. W. Heath, and S. Vishwanath, "Opportunistic relay selection with limited feedback," in *Proc. IEEE Vehicular Technology Conference (VTC)*, April 2007, pp. 135 – 139.

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

[44] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.

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

[46] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: Diversity-multiplexing-delay tradeoff," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3601–3621, Aug. 2006.

[47] M. Yuksel and E. Erkip, "Cooperative wireless systems: A diversity multiplexing tradeoff perspective," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3371 – 3393, Oct 2007.

[48] Y. Jiang and M. K. Varanasi, "Diversity-multiplexing gain tradeoff of MIMO systems with antenna selection," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[49] M. Ansari, A. Bayesteh, and A. K. Khandani, "Diversity-multiplexing tradeoff in Z-channel," in *Proc. 10th Canadian Workshop on Information Theory (CWIT)*, June 2007, pp. 21–24.

[50] A. B. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[51] Anders Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[52] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1859–1874, Sept. 2004.

[53] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2895–2909, Oct. 2003.

[54] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3425–3441, August 2008.

[55] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005, pp. 2310–2314.

[56] C. T. K. Ng, N. Jindal, A. J. Goldsmith, and U. Mitra, "Capacity gain from two-transmitter and two-receiver cooperation," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3822 – 3827, Oct 2007.

[57] O. Şahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel," in *Proc. IEEE GLOBECOM*, Nov. 2007, pp. 1627–1631.

[58] O. Şahin and E. Erkip, "On achievable rates for interference relay channel with interference cancelation," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Nov. 2007, pp. 805–809.

[59] I. Maric, R. Dabora, and A. Goldsmith, "On the capacity of the interference channel with a relay," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, July 2008, pp. 554–558.

[60] S. Sridharan, S. Vishwanath, S. A. Jafar, and S. Shamai, "On the capacity of cognitive relay assisted Gaussian interference channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, July 2008, pp. 549–553.

[61] V. I. Morgenshtern and H. Boelcskei, "Crystallization in large wireless networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3319 – 3349, Oct 2007.

[62] C. Rao and B. Hassibi, "High diversity scheme for wireless networks based on interference cancellation," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, June 2007, pp. 1256–1260.

[63] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, co-operation and full duplex operation," *IEEE Trans. Inform. Theory*, 2008, submitted for publication.

[64] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, New York, NY, USA, 2005.

[65] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[66] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[67] Gerhard Kramer, "Review of rate regions for interference channels," in *International Zurich Seminar on Communications*, Feb. 2006, pp. 162–165.

[68] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[69] Igal Sason, "On achievable rate regions for the Gaussian interfernce channel," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 1345–1356, June 2004.

[70] R. Tannious and A. Nosratinia, "Cooperation in wideband channels," in *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, June 2005, pp. 955–959.

[71] R. Tannious and A. Nosratinia, "Relay channel with private messages," in *Proc. IEEE GLOBECOM*, Nov. 2006, pp. 1–5.

[72] R. Tannious and A. Nosratinia, "Relay channel with private messages," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3777–3785, Oct 2007.

[73] R. Tannious and A. Nosratinia, "Spectrally efficient relay selection protocols in wireless networks," in *Proc. IEEE ICASSP*, June 2008, pp. 3225–3228.

[74] R. Tannious and A. Nosratinia, "Spectrally efficient relay selection with limited feedback," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1419–1428, Oct 2008.

[75] R. Tannious and A. Nosratinia, "The interference channel with MIMO relay: Degrees of freedom," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, July 2008, pp. 1908–1912.

VITA

Ramy Medhat Abdallah Tannious was born in Alexandria, Egypt, in 1977. He received the degrees of Bachelor of Science (with honors) in 1999 and Master of Science in 2003, both in Electrical Engineering from Alexandria University, Alexandria, Egypt. In January 2005, he started his Ph.D. studies in Electrical Engineering at the University of Texas at Dallas. From 2000 to 2001, he served in the Signal Corps of the Egyptian Armed Forces. From 2002 to 2003, he was with SIDPEC Petrochemicals, Alexandria, Egypt as an Instrumentation and Control Engineer. During 2004, he was a Research and Teaching Assistant at the University of South Florida. In the summer of 2008, he was an intern at Qualcomm Inc., San Diego, working on baseband algorithms design for the next generation of wireless systems.