RICE UNIVERSITY

Bounds and Protocols for a Two-way Multiple Node Channel

by

Debashis Dash

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Science

APPROVED, THESIS COMMITTEE:

[Signatures]

Ashutosh Sabharwal, Chair
Assistant Professor of
Electrical and Computer Engineering

Behnaam Aazhang
J. S. Abercrombie Professor of
Electrical and Computer Engineering

Don H. Johnson
J. S. Abercrombie Professor of
Electrical and Computer Engineering
and of Statistics

Houston, Texas

May 2007
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 1441813
Copyright 2007 by ProQuest Information and Learning Company. All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346
At Thy Lotus Feet

गुरुर्ब्रह्मा गुरुविषणुः गुरुदेवो महेश्वरः।
गुरुरेव परंप्रह्व तस्मै श्रीगुरवे नमः॥३॥
-आदि शंकर (श्री गुरुस्तोत्रम्)

(Guru is the representation of Lord Brahma, Vishnu and Shiva. He creates and sustains knowledge and destroys the weeds of ignorance. I salute such a Guru.)

Dedicated to all my teachers who have shaped my life, especially my parents,

Mrs. Sudhanjali Dash and Dr. Jagannath Dash.
ABSTRACT

Bounds and Protocols for a Two-way Multiple Node Channel

by

Debashis Dash

Often in multiuser wireless systems like cellular networks, data is present in both directions (uplink and downlink). One way to formulate protocols for such networks, is to design the uplink and downlink independently. But such a design ignores the two-way nature of data. We show that by jointly designing the uplink and downlink, we get rate gains as long as there is data in both the directions. We make a distinction between data being two-way and the channel being two-way and show that for the case where only the channel is two-way, a disjoint design is optimal. The rate gains can be attributed to the inherent feedback in two-way schemes which enables cooperation in all topologies (including hidden node topologies). Furthermore, in near-far situations, the weak user gains appreciably from such cooperation. Finally we demonstrate that several well known inner bounds can be derived as special cases of our rate region.
Acknowledgements

I would like to gratefully acknowledge the enthusiastic supervision of my thesis adviser Prof. Ashutosh Sabharwal during this work. He has been lot more than just an advisor at every step. His technical and literary guidance helped me not only to formulate the content better but his suggestions were helpful in improving the presentation of my Master's defense immensely. His own hard-work, persistence and adherence to high standards have been a major driving force for my work. His constant support and assistance in addition to his non-acceptance of anything less than my best efforts always challenged me as well as encouraged me in my research. I would also like to thank Dr. Ahmad Khoshnevis who collaborated with me in the beginning part of my work and helped me get started on the problem.

I also heartily thank Prof. Behnaam Aazhang, Prof. Don Johnson and all the members of CMC Lab for their support, technical discussions and numerous suggestions and comments. In particular I would like to acknowledge the help of Chris Steger and Dr. Arnab Chakrabarti for many stimulating discussions and their help with refining my ideas.
Contents

Abstract ii

Acknowledgements iii

List of Tables vii

List of Figures viii

1 Introduction 1

2 Preliminaries 7

2.1 Superposition Coding 7

2.2 Block Markov Encoding 9

2.3 Half-Duplex Two-Way channel 10

2.4 Auxiliary random variables 11

2.5 Markov Chains 12

3 Main Results 13

3.1 Notation 13

3.2 Channel Models 13

3.2.1 Full-duplex Channel 14

3.2.2 Half-duplex Channel 15

3.3 Definitions 15
3.4 Proposed Coding Structure .................................. 17
3.5 Inner Bound ................................................... 19
3.6 Outer Bound .................................................. 20

4 Inner Bound ......................................................... 21
  4.1 Code generation ............................................. 21
  4.2 Encoding Scheme .......................................... 22
  4.3 Decoding Scheme .......................................... 22
  4.4 Probability of Error ...................................... 23
  4.5 Inner Bound for Half-duplex Channel .................. 26
    4.5.1 Test Channel ........................................... 26
    4.5.2 Encoding scheme ...................................... 28
    4.5.3 Decoding scheme ....................................... 31
    4.5.4 Probability of Error .................................. 32

5 Outer Bound ......................................................... 37
  5.1 Outer Bound for $R_3$ ..................................... 38
  5.2 Outer Bound for $R_1$ and $R_2$ ......................... 40
  5.3 Outer Bound for $R_1 + R_2$ ............................ 42

6 Special cases .......................................................... 45
  6.1 Gaussian case ............................................. 45
    6.1.1 Achievable Rate Region ................................ 48
  6.2 Symmetric Case .......................................... 50
  6.3 Asymmetric Case .......................................... 52
  6.4 Two-way Channel with One-way Data ....................... 56
  6.5 Reduction to Carleial rate region ...................... 59
  6.6 Reduction to Cover-Leung rate region ................... 60
List of Tables

4.1 Input alphabets and their sizes. ........................................... 28
4.2 Encoding functions for the half duplex case. ....................... 31
List of Figures

1.1 A typical cell in a cellular network where the mobile stations communicate through the base station (red). In most cases data is present in both directions of such communication links. ........................................ 2

1.2 Three user hidden node topology ............................................ 3

1.3 (a) Data in uplink and only feedback in the downlink and (b) No feedback (inner bound), genie-aided perfect feedback (outer bound) and MAC with finite feedback (two inner regions) with different amount of feedback power. ......................................................... 4

2.1 TDMA: Users orthogonally separated in time ................................ 8

2.2 FDMA: Users orthogonally separated in frequency ........................ 8

2.3 Superposition coding: codewords occupy the whole bandwidth and time .......................................................... 9

2.4 Markov chain formed by the channel blocks ................................. 10

2.5 Shannon's two-way channel ...................................................... 11

2.6 Half-duplex two-way channel .................................................. 11

3.1 Feedback based coding in multiuser two-way channels. For clarity of representation, the broadcast channel is \( p(y_1, y_2|x) = p(y_1|x)p(y_2|x) \). .............................................. 17

4.1 Half-duplex two-way network: Two states in the system. .............. 27

4.2 Usual superposition coding using binning. The cloud centers signify one of the message sets. ....................................................... 29
4.3 Superposition coding with binning. Note that in this case $p(z|u_{10}) = p(u_{20})$ and $p(z|u_{20}) = p(u_{10})$ by construction. 

5.1 Markov chain model of the multiuser feedback scenario. 

6.1 (a) Uplink and (b) Downlink for the half-duplex Gaussian channel 

6.2 Power distribution for Gaussian MAC-MC with half-duplex links 

6.3 Rate region without feedback, $\mathcal{R}_{nf}$ 

6.4 The 3-d rate region for the MAC + BC half-duplex two-way channel. 

6.5 MAC sum-rate versus BC common rate with and without feedback. 

User powers are $P_1 = P_2 = 50$, and gray node power $P = 50, 100, 1000$. 

$N = 1$ and $N_1 = N_2 = 1.5$. 

6.6 AWGN with pathloss: Near-far topology 

6.7 Percentage gain in the rates of the strong user (a) and weak user (b). 

6.8 Behaviour of Gaussian capacity at high and low SNRs. 

6.9 The strong user is able to decode the downlink data at a higher rate and hence cooperate more due to its better channel. This results in a better rate gain for the weaker user. 

6.10 Gaussian case with asymmetric power constraints 

6.11 No feedback (inner bound), genie-aided perfect feedback (outer bound) and MAC with finite feedback (two inner regions) with different amount of feedback power.
Chapter 1

Introduction

“One-wayness is obviously a serious limitation in talking about the most common kind of communication, which is two-way”

– Peter Elias on how Shannon might have come up with two-way channels in [23].

In a typical multiuser network, it is common to have data in both directions of a communication link. For example consider a cellular network as shown in Figure 1.1, where there are multiple mobile stations which communicate to the base station. These mobile stations can be devices like cell phones, pagers, smart phones, PDAs, laptops etc, using a number of different applications e.g. voice or web applications etc. Hence in many cases there is data to be sent from the base station to the mobile station in addition to data from the base stations. However, the current solution for above multiuser problems is to divide it into two independent problems (uplink and downlink). That is we design protocols and algorithms for the uplink separately from the downlink. The uplink resembles the information theoretic formulation of a Multiple Access Channel (MAC) [21, 2, 12, 8, 25, 5, 34, 17] and the downlink resembles a Broadcast Channel (BC) [7, 13, 3, 14, 11, 24]. Such independent design ignores the two-way nature of the data present in the network. One big reason for such disjoint design is its simplicity and ease of implementation. But if we can afford
to make our devices more computationally capable, we could use the two-way nature of the network to attain better throughput.

Figure 1.1: A typical cell in a cellular network where the mobile stations communicate through the base station (red). In most cases data is present in both directions of such communication links.

Two-way channels, first proposed and studied by Shannon in 1961 [32], is a natural model for many communication scenarios in which all parties may have data for each other. Shannon proposed both inner and outer bounds for the two-node two-way channel. Further, he also provided cases where the bounds match (deterministic channels like modulo-2 adding channel and non-deterministic channels like binary erasure-noiseless two-way channel and push-to-talk channel), and where they do not match (e.g. the well-studied example of binary multiplying channel [27, 28, 33, 22, 29, 30]). Though comparatively less widely studied than the one-way channel model [31], both the inner and outer bounds were improved subsequently [16, 35, 18]. The two-way formulation allows us to study multiuser networks with noisy feedback, and potentially when feedback shares its resources with the forward link. An interesting point to note is that while feedback is not needed to achieve the capacity of two-way Gaussian channel, we show that it is beneficial in multiple node Gaussian networks. The gain can be attributed to cooperative beamforming by the two nodes in uplink communication.
Figure 1.2: Three user hidden node topology

We focus on a three-node network shown in Figure 1.2 for all our analysis. We make a distinction between a two-way channel and two-way data. If the channel can support data on both directions then it is said to be a two-way channel, but if there are independent messages to be sent in both directions then it is said to have two-way data. For example even if there is no data in one of the directions, a two-way channel can be used to send only feedback. It is not always optimal to use joint encoding. For the case of a two-way channel which does not have data in one direction, we show that if feedback has to share resources with the forward link as in a conventional multiple access channel, the receiver (User 3 in Figure 1.2) has to use exponentially large power to offset the time loss. In essence, a two-way design is practically useless to implement in half-duplex networks, especially when there is no downlink data. To get an idea, consider Figure 1.3 which indicates that in order to do better than the well known MAC rate region, even if the downlink user employs enormous power it cannot even get close to the promised noiseless upper bound (more details in Chapter 6). But we claim that this result is merely because due to not using the downlink channel optimally by reserving it solely for feedback.

Perhaps due to limited attention given to two-way channels, its multi-node extension has received even less attention. Recent work on two-way networks [18, 20], studies both inner and outer bounds for arbitrary network topologies with two-way networks as a special case. The elegant formulation using code trees [18] clearly demonstrates the complexity of incorporating feedback information in code construction for generalized networks. However, the inner and outer bounds derived in [18] are not computable in many cases of interest.
Figure 1.3: (a) Data in uplink and only feedback in the downlink and (b) No feedback (inner bound), genie-aided perfect feedback (outer bound) and MAC with finite feedback (two inner regions) with different amount of feedback power.

Our key contribution is an inner bound using a novel binning strategy to form a two-way code using feedback which results in a larger rate region for the multiuser two-way channel. The inner bound is derived for both full-duplex and half-duplex channels. Our proposed coding strategy is inspired by the two-way coding scheme of Han [16] and generalizes our recent results in [10]. In fact, we show that the achievable rate region presented in [9] is a special case of the new coding structure.

The topology considered in this paper is inspired by two applications. First, it resembles the hidden node topology commonly considered in the design of contention resolution protocols, where the two senders do not have a channel between them. Due to the lack of a direct cooperative link, cooperative coding is not feasible. However, since the receivers are also senders, the feedback based on the received signals can be used to induce cooperation between uplink senders (Users 1 and 2). To some extent, the above conclusion was also independently derived for multiple access channels with noisy feedback [15]. Our results show that such cooperation can be induced for both full- and half-duplex networks, with and without downlink data (for the latter case an overhead is incurred in power). Second, for cellular networks, the
analysis shows that there is a sum-rate gain in jointly designing uplink and downlink communication. While all networks have decoupled uplink and downlink systems, a joint design can increase the system data rates.

We note that the usual min-cut max-flow bound [6, Chapter 14] does not apply to two-way networks. An extension of the bound for full-duplex two-way networks was considered in [19]. An improved outer bound was found for the conventional AWGN multiple access channel (where the third link does not have any data of its own to send) in [15]. In this thesis, we derive the outer bound from first principles, which demonstrates the challenge of deriving such an outer bound. The two-way channel implicitly allows sending feedback about the signal received in one mode (multiple access mode) while transmitting in the second mode (broadcast mode). Thus it leads to an infinite Markov chain (as shown in Figure 5.1).

Our proof uses the techniques proposed in [35], where the authors studied the conventional two-node two-way network. As proposed by Shannon [32], the outer bound typically uses an arbitrary input distribution. On the other hand, the inner bounds are obtained with independent inputs. Following the line of reasoning of [32, 35], our proposed outer bounds also use non-independent inputs to reflect how the feedback can induce “correlation” in the inputs.

Feedback is known to improve the capacity of multiuser channels and in some cases reduce the computational complexity. It is interesting to notice that most of the practical feedback protocols are based on “genie-aided” theoretical analysis, which either have a noiseless feedback link or some other practically difficult assumption like full-duplex radios [7, 12, 11, 5, 34, 25]. Hence, it becomes necessary to put practical constraints and see if the existing feedback protocols still give us gains. But our analysis of the two-way channel models also answer this additional question that the correct way to analyze a practical feedback scheme is by using a two-way channel model.
The rest of the thesis is organized as follows. In Chapter 2, we will discuss some basic principles which will be used repeatedly in the later chapters. In Chapter 3 we will discuss the channel model, assumptions, nomenclature and give the main results of the work. Chapter 4 contains the proof of the inner bound, where we derive special encoding and decoding functions to find the half-duplex special case of the inner bound. Chapter 5 contains the proof of the outer bound and Chapter 6 has some interesting Gaussian half-duplex cases. Then we consider a pathloss model which simulates a near-far situation and show that such cooperation through feedback is more useful for the weak user. For the case when both the uplink transmitters are equidistant from the receiver, we consider cases when they have symmetric or asymmetric power constraints. In the last part of this chapter we show that some of the previous known results are special cases of our rate region in Chapter 4. We conclude our findings in Chapter 7.
Chapter 2

Preliminaries

In this chapter we review some basic information theoretic ideas, which will be used throughout the thesis to derive our inner and outer bounds. We start with a common encoding method, known as superposition coding, used in multiuser communication systems.

2.1 Superposition Coding

The simplest way to separate users (or messages) when the resources are constrained, is by orthogonal allocation. Users (or messages) can be divided in time (TDMA) or in frequency (FDMA) or both. As shown in Figure 2.1, the messages $W_1$ and $W_2$ are independently mapped into channel input $X_1$ and $X_2$ respectively and they are sent over the channel in orthogonal time bands. Similar is the situation for FDMA as shown in Figure 2.2. Due to its orthogonal nature, the decoding scheme is very simple. However, this is not an optimal scheme if one can allow a more complicated multiuser decoding scheme. Superposition coding is known to perform better than time and frequency (orthogonal) division of resources. As shown in Figure 2.3, now the channel input $X$ is a function of both $W_1$ and $W_2$, where the cloud centers denote the message from the first user and the around each cloud center the alphabet points
code for the message from the second user, forming a larger code $X$, which now can take up the whole time bandwidth space on the channel. But due to its structure it is more complicated to decode it than any orthogonal scheme.

![Figure 2.1: TDMA: Users orthogonally separated in time](image)

![Figure 2.2: FDMA: Users orthogonally separated in frequency](image)

Superposition coding was first used in [7] and was shown to be optimal and perform better than any orthogonal scheme [4]. The idea behind such a coding scheme is to allow both users to use the channel simultaneously rather than dividing it between users or messages orthogonally. Superposition coding was introduced in the context of the broadcast channel, so the authors also called it cooperative
broadcasting. For Gaussian channels, the Gaussian codewords can just be added with a proper power scaling. For a more general channel, first typical sequences are selected for the input message layers and then a typical channel sequence is selected which is jointly typical with all the message layer sequences. Decoding of a superposition code is more complex than an equivalent orthogonal code. The best way is to first decode the message with the best signal to noise ratio (SINR), considering all other messages as noise and then subtract (for Gaussian channels) the decoded message. Then the next message is decoded following the same procedure till the last message has been decoded.

![Figure 2.3: Superposition coding: codewords occupy the whole bandwidth and time](image)

**2.2 Block Markov Encoding**

The idea of block Markov superposition codes was introduced in [8]. Instead of superimposing multiple user’s data, the authors superimposed messages (independent and incremental) from different blocks and used noiseless feedback to attain a larger rate region for a multiple access channel.
Superposition is used to encode messages (of more than one block) at different layers that can be decoded once enough information is acquired about that layer. These message layers give rise to different auxiliary random variables which increase the resolution at which we can manage the rate division between the different users depending on the noise variance at the receiver but it increases the complexity of the decoder.

As the coding is done over blocks as a whole, as compared to some schemes which adaptively modify each channel symbol, the blocks form a Markov chain as shown in Figure 2.4.

The block structure of the coding scheme enables to take off the part of the message that could be decoded after each block, which can reduce the amount of data to be sent as feedback.

2.3 Half-Duplex Two-Way channel

The two-way channel was first introduced by Shannon in 1961 [32]. A two-way channel model utilizes the previous channel outputs in addition to the message to form the channel input. As shown in Figure 2.5, the channel inputs also depend on the channel outputs for the same user in addition to the message. Such a channel model can be easily extended for feedback schemes.

The two-way channel initially introduced by Shannon and those usually consid-
Figure 2.5: Shannon’s two-way channel

Figure 2.6: Half-duplex two-way channel

2.4 Auxiliary random variables

The usual way of writing rate regions is done in terms of the channel input and output random variables. But superposition schemes typically, gives rise to random variables other than channel inputs and outputs due to the layered code structure. These additional random variables are called auxiliary variables. The auxiliary random variables help in encoding the messages into the layered code to form the channel
inputs. To keep the rate region computable, one of the major concerns of any superposition code is to bound the cardinality of such auxiliary variables [26]. However, in many multiuser cases, this is still an open problem [5, 26, 16, 35].

2.5 Markov Chains

Definition 2.1 Three random variables $X, Y$ and $Z$ form a Markov chain (written as $X \rightarrow Y \rightarrow Z$) if the joint density function has the following form:

$$p(x, y, z) = p(x)p(y|x)p(z|y)$$

Some well known information theoretic properties of Markov chains that are used to prove outer bounds are:

- Conditioning on functions of random variables $H(A|B, C) = H(A|B, C, D)$ for $D = f(B, C)$
  
  Typical use: $H(Y_n|W, Y^{n-1}) = H(Y_n|W, Y^{n-1}, X_n)$ for $X_n = f(W, Y^{n-1})$.

- Conditioning reduces entropy $H(A|B) < H(A)$
  
  Typical use: $H(Y_n|W, X_n, Y^{n-1}) < H(Y_n|X_n, Y^{n-1})$.

- Markov Chain property $H(A|B, C) = H(A|B)$ for a Markov chain $C \rightarrow B \rightarrow A$

  Typical use: $H(Y_n|W, X_1, X_2, Y^{n-1}) = H(Y_n|X_1, X_2, Y^{n-1})$.

- Memoryless channel property $H(Y_n|X^n, Y^{n-1}) = H(Y_n|X_n)$ if $p(y^n|x^n) = \prod_i p(y_i|x_i)$.

  Another example: $H(Y_n|X_1^n, X_2^n, Y^{n-1}) = H(Y_n|X_1, X_2)$.

- Data processing theorem The most common way to write the data processing theorem is $I(X; Y) \geq I(X; Z)$. While finding outer bounds, the following form: $H(Z|X) \geq H(Z|Y)$ is also useful.
Chapter 3

Main Results

Before presenting our main results, we define the notation, the channel model and
the performance criteria.

3.1 Notation

We define the following notation for representing all the random variables. Through-
out this paper, random variables are represented by capitals, (e.g. $U_1$, $X_1$, $W_1$, etc.),
their alphabet sets by the corresponding script symbols, (e.g. $\mathcal{U}_1$, $\mathcal{X}_1$, $\mathcal{W}_1$, etc.),
and any particular instance of those random variables by small case symbols, (e.g.
$u_1$, $x_1$, $w_1$, etc.). In general, $A^{b}_{u,c}$ will denote random variable $A$ for the $c^{th}$ channel use
of block $b$ of user $u$. Boldface variables will be used to denote vectors. For example,
$X_1 = (X_1^1, X_1^2, \ldots, X_1^B)$ is the vector input to the channel for user 1 for $B$ blocks.
Each of the block inputs is a vector ($n$-tuple) itself (i.e. $X_i^1 = (X_{i,1}, X_{i,2}, \ldots, X_{i,n})$).

3.2 Channel Models

Now we describe the channel models considered in this work. We will consider
two types of channels depending on certain assumptions on the transmitter-receiver
CHAPTER 3. MAIN RESULTS

capabilities. If the transmitters and receivers can send and receive at the same time, they are said to be in a full-duplex mode. Usually, by operating in a full-duplex mode, we get a gain in the rate region. This gain is similar to what we had seen in superposition coding as compared to TDMA or FDMA in Chapter 2. On the other hand, if the transmitters and receivers can only send or receive and not do both at the same time, they are said to be in a half-duplex mode. The formal mathematical definitions of these channels are given below:

3.2.1 Full-duplex Channel

Three variables are associated with each terminal: $w_i$ to represent the message, $x_i$ to represent the channel input and $y_i$ to represent the channel output at terminal $i$. Then the two-way network is described by the channel probability distribution $p(y_1, y_2, y_3|x_1, x_2, x_3)$. In many cases of interest, the two-way channel can be decomposed such that

$$p(y_1, y_2, y_3|x_1, x_2, x_3) = p(y_3|x_1, x_2, x_3)p(y_1, y_2|x_1, x_2, x_3).$$

In the case of Gaussian channels, the channel input-output relation can be described as follows:

$$y_1 = x_1 + x_2 + x_3 + n_1,$$

$$y_2 = x_1 + x_2 + x_3 + n_2,$$

$$y_3 = x_1 + x_2 + x_3 + n_3,$$

where $n_1$, $n_2$ and $n_3$ are independent zero mean Gaussians with variances $N_1, N_2$ and $N_3$ respectively. We further assume that each source operates under an average power constraint of $P_i, i = 1, 2, 3$. Note that the inputs $x_1$ and $x_2$ do not appear in
outputs $y_1$ and $y_2$, since we have assumed that there is no direct channel between Users 1 and 2.

### 3.2.2 Half-duplex Channel

Since the users are assumed to be half-duplex, the network time-shares between two possible states: multiple access channel (MAC) mode for $\eta$ fraction of time and broadcast channel (BC) mode for $(1 - \eta)$ fraction of time, where $0 \leq \eta \leq 1$ denotes the time-sharing variable. In this case, the channel probability law can be written as

$$p(y_1, y_2, y_3|x_1, x_2, x_3) = \begin{cases} p(y_3|x_1, x_2) & \text{MAC} \\ p(y_1, y_2|x_3) & \text{BC}. \end{cases} \quad (3.4)$$

In the case of Gaussian channels, the input-output relation is given by

$$\begin{align*}
y_1 &= x_3 + n_1, \\
y_2 &= x_3 + n_2, \\
y_3 &= x_1 + x_2 + n_3,
\end{align*} \quad (3.5) \quad (3.6) \quad (3.7)$$

where $n_1, n_2$ and $n_3$ are independent zero mean Gaussians with variances $N_1, N_2$ and $N_3$ respectively. We further assume that each source operates under an average power constraint of $P_i$, $i = 1, 2, 3$ in their respective transmission modes.

### 3.3 Definitions

We define the alphabet sets from which the input messages $m_1, m_2, m_3$ are chosen ($m_1 \in \mathcal{M}_1, m_2 \in \mathcal{M}_2, m_3 \in \mathcal{M}_3$) as, $\mathcal{M}_1 = \{1, 2, \ldots, M_1\}, \mathcal{M}_2 = \{1, 2, \ldots, M_2\}, \mathcal{M}_3 = \{1, 2, \ldots, M_3\}$. $P_{e1}$ and $P_{e2}$ are the error probabilities of decoding error at user 3 de-
defined by,

\[ P_{e1} = \frac{1}{M_1 M_2 M_3} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \sum_{m_3=1}^{M_3} Pr(\hat{m}_1 \neq m_1 | m_1, m_2, m_3 \text{ was sent}) \]

\[ P_{e2} = \frac{1}{M_1 M_2 M_3} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \sum_{m_3=1}^{M_3} Pr(\hat{m}_2 \neq m_2 | m_1, m_2, m_3 \text{ was sent}) \]

Similarly, \( P_{e3}^1 \) and \( P_{e3}^2 \) are the error probabilities of decoding error at User 1 and User 2 (in decoding what User 3 sent to both of them) defined by

\[ P_{e3}^1 = \frac{1}{M_1 M_2 M_3} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \sum_{m_3=1}^{M_3} Pr(\hat{m}_3^1 \neq m_1 | m_1, m_2, m_3 \text{ was sent}) \]

\[ P_{e3}^2 = \frac{1}{M_1 M_2 M_3} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \sum_{m_3=1}^{M_3} Pr(\hat{m}_3^2 \neq m_2 | m_1, m_2, m_3 \text{ was sent}) \]

**Definition 3.1 (Achievable Region)** The rate tuple \((R_1, R_2, R_3)\) is said to be achievable if there exists a \((M_1, M_2, M_3, n)\) code with,

\[ R_1 \leq \frac{\log M_1}{n}, R_2 \leq \frac{\log M_2}{n}, R_3 \leq \frac{\log M_3}{n} \]

such that, \( P_{e1} \to 0, P_{e2} \to 0, P_{e3}^1 \to 0 \) and \( P_{e3}^2 \to 0 \), as \( n \to \infty \).

**Definition 3.2 (Capacity Region)** The closure of all achievable rate tuples is called the capacity region of the two-way multiple access channel and denoted by \( C \).
3.4 Proposed Coding Structure

Consider the simplest multiuser case with three users as shown in Figure 3.1. Users have data for each other; in our case, User 1 and 2 have messages for User 3 and User 3 has a common message for Users 1 and 2.

Figure 3.1: Feedback based coding in multiuser two-way channels. For clarity of representation, the broadcast channel is \( p(y_1, y_2|x) = p(y_1|x)p(y_2|x) \).

Figure 3.1 shows the coding structure at each user. Since each user has data to send to other users, each user has an encoder to encode their transmissions and decode the received transmissions. In decoupled systems, the encoder and decoder are decoupled. However, in our proposed scheme, the input to decoder is also an input to encoder and vice versa, thereby implicitly allowing feedback from receivers to transmitters.

The coding scheme can be understood to operate in three phases. In the first
phase, both Users 1 and 2 split their messages, $M_1$ at rate $R_1$ and $M_2$ at rate $R_2$, into two parts. The rate pair $(R_1, R_2)$ is outside the capacity region of multiple access region with no feedback [6]. Both Users 1 and 2 use superposition coding to encode their two-part messages. At the receiver of User 3, only the inner layer of superposition code can be decoded, and a portion of message from both Users 1 and 2 is left undecoded (similar to [5]).

In the second phase, User 3 strips away the portion of information it has decoded from the received signal. In essence, it compresses the received message to only that portion of the message that could not be decoded. It then jointly encodes this codeword with the downlink message $M_3$ at rate $R$ and transmits it in the broadcast channel (BC) mode. Since the users know the portion of their own message which could not be decoded, they use this side information to decode both $M_3$ and the undecoded message from the other user.

Finally, in phase three, both Users 1 and 2 know a part of undecoded message sent by each other from previous transmissions. They superimpose that on top of their new messages leading to a three-layer superposition code. The receiver at User 3, uses the previous reception and the current reception to decode the undecodable portion from last instance and some new information (the inner two layers). The third layer is undecodable. Thus, the system goes through second and third phases for $b$ consecutive blocks.

Now we are ready for the two main results: inner and outer bounds for the half-duplex 3 user channel model with hidden node topology.
3.5 Inner Bound

Theorem 3.1 All the rate points defined by \((R_1, R_2, R_3) \in \mathcal{R}\) which satisfy the following equations:

\[
R_1 \leq I(\tilde{U}_1; \tilde{U}_2, X_3, Y_3, \tilde{U}_3, W_3)
\]
\[
R_2 \leq I(\tilde{U}_2; \tilde{U}_1, X_3, Y_3, \tilde{U}_3, W_3)
\]
\[
R_3 \leq \min[\{I(\tilde{U}_3; X_1, Y_1, \tilde{U}_1, W_1), I(\tilde{U}_3; X_2, Y_2, \tilde{U}_2, W_2)\}]
\]
\[
R_1 + R_2 \leq I(\tilde{U}_1, \tilde{U}_2; X_3, Y_3, \tilde{U}_3, W_3)
\]

are achievable (i.e. \(\mathcal{R} \subseteq \mathcal{C}\)) where \(X_1, X_2, X_3\) are the channel inputs, \(Y_1, Y_2, Y_3\) are the channel outputs and \(U_1, U_2, U_3, \tilde{U}_1, \tilde{U}_2, \tilde{U}_3, W_1, W_2, W_3\) are auxiliary random variables whose joint distribution is of the form,

\[
P(u_1, u_2, u_3, \tilde{u}_1, \tilde{u}_2, \tilde{u}_3, w_1, w_2, w_3, x_1, x_2, x_3, y_1, y_2, y_3)
\]
\[
= P_1(u_1)P_2(u_2)P_3(u_3)P(\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, w_1, w_2, w_3)
\]
\[
\cdot P(x_1|u_1, \tilde{u}_1, w_1)P(x_2|u_2, \tilde{u}_2, w_2)
\]
\[
\cdot P(x_3|u_3, \tilde{u}_3, w_3)P(y_1, y_2, y_3|x_1, x_2, x_3)
\]

We prove the above inner bound and find the half-duplex special case in Chapter 4. The proof is based on typical decoding arguments and superposition coding based on Carleial’s inner bound for a full duplex MAC with feedback [5]. We will further investigate the inner bound in Chapter 6 with applications to several interesting special cases.
3.6 Outer Bound

**Theorem 3.2** The capacity region $\mathcal{C}$ of the three-node half-duplex two-way network is a subset of the following region

$$\mathcal{C}' \equiv \{(R_1, R_2, R_3) :$$

$$R_1 \leq I(X_1; Y_2|X_2, Z_2, W_3, Q) + I(X_1; Y_3|X_2, Z_2, Z_3, Q)$$

$$R_2 \leq I(X_2; Y_1|X_1, Z_1, W_3, Q) + I(X_2; Y_3|X_1, Z_1, Z_3, Q)$$

$$R_3 \leq \min[I(X_3; Y_1|X_1, Z_1, Q), I(X_3; Y_2|X_2, Z_2, Q)]$$

$$R_1 + R_2 \leq I(X_1, X_2; Y_3|Z_3, Q)\},$$

where $X_1, X_2, X_3, Y_1, Y_2, Y_3, Z_1, Z_2, Z_3, Q$ are random variables whose joint distribution is of the form,

$$p(q)p(z_1, z_2, z_3|q)p(x_1|z_1, q)p(x_2|z_2, q)p(x_3|z_3, q)$$

$$\quad \cdot p(y_3|x_1, x_2, q)p(y_1|x_3, q)p(y_2|x_3, q),$$

where $Q$ is the time sharing variable.

We will prove the outer bound in Chapter 5 and find some special cases like its half duplex extension and no feedback case. The proof of the outer bound is based on Han's outer bound for a two user two-way channel [16], and uses the Markov structure of the channel input, output and auxiliary random variables.
Chapter 4

Inner Bound

In the first part of this chapter we will prove Theorem 3.1. Later, we will also consider the special cases of half-duplex two-way channels. The proof of the inner bound will be shown in four steps. The first step involves generation of the codebooks. The second step includes the encoding scheme from the codebook. Then we will give the details of the decoding scheme in the third step, which is used to calculated the probability of error in the fourth step. The probability of error calculation finally gives rise to the rate equations that govern the achievability rate region.

4.1 Code generation

Generate $M_1 = |\mathcal{M}_1|$ independent identically distributed (i.i.d.) random vectors $\mathbf{U}_1$ according to $P_1(\mathbf{u}_1) = \prod_{i=1}^n P_1(u_{1,i})$. Then randomly assign each random vector to one of the messages from the input alphabet of User 1($\mathcal{M}_1$). Similarly, we generate $M_2 = |\mathcal{M}_2|$ and $M_3 = |\mathcal{M}_3|$ i.i.d. random vectors $\mathbf{U}_2$ and $\mathbf{U}_3$ according to $P_2(\mathbf{u}_2) = \prod_{i=1}^n P_2(u_{2,i})$ and $P_3(\mathbf{u}_3) = \prod_{i=1}^n P_3(u_{3,i})$ respectively. Then randomly assign each of these vectors to each member of the input alphabets of User 2 and 3 respectively.


4.2 Encoding Scheme

Suppose at block \( b - 1 \), the \( n \)-tuples \( x_1^{b-1}, x_2^{b-1} \) and \( x_3^{b-1} \) is sent by users 1, 2 and 3 respectively and \( y_1^{b-1}, y_2^{b-1} \) and \( y_3^{b-1} \) is received at respective nodes. In the \( b^{th} \) block we define the “feedback” \( w_1^b, w_2^b \) and \( w_3^b \) as

\[
\begin{align*}
w_1^b &= x_1^{b-1}, y_1^{b-1}, \\
w_2^b &= x_2^{b-1}, y_2^{b-1}, \\
w_3^b &= x_3^{b-1}, y_3^{b-1},
\end{align*}
\]

and set the “incremental” part as:

\[
\begin{align*}
\tilde{u}_1^b &= u_1^{b-1}(m_1^{b-1}), \\
\tilde{u}_2^b &= u_2^{b-1}(m_2^{b-1}), \\
\tilde{u}_3^b &= u_3^{b-1}(m_3^{b-1}).
\end{align*}
\]

The encoding functions use the channel output and the information about the previous blocks \( (1, 2, \ldots, b - 1) \) that the user already has, in addition to the new message for the present block to construct the channel inputs for block \( b \)

\[
(x_1^b = f_1(u_1^b(m_1^b), \tilde{u}_1^b, w_1^b))
\]

\[
x_2^b = f_2(u_2^b(m_2^b), \tilde{u}_2^b, w_2^b)
\]

\[
x_3^b = f_3(u_3^b(m_3^b), \tilde{u}_3^b, w_3^b))
\]

4.3 Decoding Scheme

At the end of block \( b \), \( x_1^b \) and \( y_1^b \) are known to User 1, \( x_2^b \) and \( y_2^b \) are known to User 2 and \( x_3^b \) and \( y_3^b \) are known to User 3. The decoding procedure for the three users
are as follows: User 1 decodes the message sent by User 3 during block $b$ as $\hat{m}_3^b \in \mathcal{K}$ such that,

$$ (u_3^b(\hat{m}_3^b), x_2^b, y_2^b, \tilde{u}_3^b, w_2^b) \in T_e(\tilde{U}_3, X_1, Y_1, \tilde{U}_1, W_1). $$

(4.1)

Note that the information about $y_1^{b-1}$ is hidden in $W_1$ and the incremental information about $m_3^b$ is in $y_1^b$. Similarly, User 2 decodes the message sent by User 3 during block $b$ as $\hat{m}_3^b \in \mathcal{K}$ such that,

$$ (u_3^b(\hat{m}_3^b), x_2^b, y_2^b, \tilde{u}_2^b, w_2^b) \in T_e(\tilde{U}_3, X_2, Y_2, \tilde{U}_2, W_2). $$

(4.2)

User 3 decodes the messages sent by Users 1 and 2 during block $b-1$ as $\hat{m}_1^{b-1} \in \mathcal{I}$ and $\hat{m}_2^{b-1} \in \mathcal{J}$ such that,

$$ (u_1^{b-1}(\hat{m}_1^{b-1}), u_2^{b-1}(\hat{m}_2^{b-1}), x_3^b, y_3^b, \tilde{u}_3^b, w_3^b) \in T_e(\tilde{U}_1, \tilde{U}_2, X_3, Y_3, \tilde{U}_3, W_3). $$

4.4 Probability of Error

Due to the symmetry in the random generation of the codes, any message sent by the users for a given block yields the same probability of error. So, without loss of generality, we can assume that $(m_1^{b-1}, m_2^{b-1}, m_3^b) = (1, 1, 1)$ was sent. Consider decoding at User 1. For block $b$, define $E_k = (\hat{m}_3^b = k | m_3^b = 1)$ as (i.e. the event that User 1 decodes the message sent by User 3 during block $b$ as $k$ when User 3 sent the message $m_3^b = 1$). An error event happens if either the correct codeword is not found to be jointly typical with the channel output or an incorrect code word is found to be jointly typical. Therefore,

$$ P_{e_1}^b = Pr(\tilde{E}_1 \cup \bigcup_{k \neq 1} E_k | (1, 1, 1) \text{was sent}) $$

$$ \leq P(\tilde{E}_1) + \sum_{k \neq 1} P(E_k), $$
where \( P \) is the conditional probability given that \((1, 1, 1)\) was sent (as defined in Cover and Thomas pg 394). By using the properties of jointly typical sequences we can show that, \( P(\tilde{E}_1) \to 0 \) and we can bound the second term as follows

\[
P(\tilde{E}_k) = P_r((u_{3 \ell}^{b-1}(k), x_1^b, y_1^b, \tilde{u}_1^b, w_1^b) \\
\in T_c((\tilde{U}_3, X_1, Y_1, \tilde{U}_1, W_1)|(1, 1, 1) \text{was sent}) \\
= \sum_{(\tilde{u}_3^{\ell}(k), x_1^b, y_1^b, \tilde{u}_1^b, w_1^b) \in T_c} P(\tilde{u}_3^b) P(x_1^b, y_1^b, \tilde{u}_1^b, w_1^b) \\
\leq |T_c|2^{-n(H(\tilde{U}_3)-\epsilon)}2^{-n(H(X_1, Y_1, \tilde{U}_1, W_1)-\epsilon)} \\
\leq 2^{n(H(\tilde{U}_3, X_1, Y_1, \tilde{U}_1, W_1)-\epsilon)}2^{-n(H(\tilde{U}_3))-\epsilon})2^{-n(H(X_1, Y_1, \tilde{U}_1, W_1)-\epsilon)} \\
= 2^{-n(I(\tilde{U}_3; X_1, Y_1, \tilde{U}_1, W_1)-3\epsilon)}. 
\]

Hence, it follows that

\[
P_{c1}^b \leq P(\tilde{E}_1) + \sum_{k \neq 1} 2^{-n(I(\tilde{U}_3; X_1, Y_1, \tilde{U}_1, W_1)-3\epsilon)} \\
= P(\tilde{E}_1) + (K - 1)2^{-n(I(\tilde{U}_3; X_1, Y_1, \tilde{U}_1, W_1)-3\epsilon)} \\
\leq P(\tilde{E}_1) + 2^{nR_3}2^{-n(I(\tilde{U}_3; X_1, Y_1, \tilde{U}_1, W_1)-3\epsilon)}. 
\]

So, to make the probability of error to be arbitrarily small,

\[
R_3 \leq I(\tilde{U}_3; X_1, Y_1, \tilde{U}_1, W_1). \tag{4.3}
\]

Following exactly the same steps for User 2, it can be shown that

\[
R_3 \leq I(\tilde{U}_3; X_2, Y_2, \tilde{U}_2, W_2). \tag{4.4}
\]
Now consider the decoding at User 3. For block $b$, define $E_{ij} = ((\hat{m}^b_i, \hat{m}^b_j) = (i, j)|(m^b_1, m^b_2) = (1, 1))$ as (i.e. the event that User 3 decodes the message sent by User 1 and User 2 during block $b$ as $i$ and $j$ respectively when $(m^b_1, m^b_2) = (1, 1)$ was sent). An error event occurs if either the correct codewords are not found to be jointly typical with the channel output or an incorrect code word is found to be jointly typical. Therefore,

$$P_{e3}^b = P(\bar{E}_{11} \cup \bigcup_{i,j \neq (1,1)} E_k | (1, 1, 1)\text{was sent})$$

$$\leq P(\bar{E}_{11}) + \sum_{i \neq 1, j = 1} P(E_{11}) + \sum_{i = 1, j \neq 1} P(E_{1j}) + \sum_{i \neq 1, j \neq 1} P(E_{ij}).$$

By using the properties of jointly typical sequences we can show that, $P(\bar{E}_{1}) \to 0$ and we can bound the second term as follows

$$P(E_{11}) = Pr((u^b_1^{-1}(i), u^b_2^{-1}(1), x^b_3, y^b_3, \tilde{u}^b_3, w^b_3)$$

$$\in T_c((\bar{U}_1, \bar{U}_2, X_3, Y_3, \bar{U}_3, W_3)|(1, 1, 1)\text{was sent})$$

$$= \sum_{(\tilde{u}^b_1, \tilde{u}^b_2, x^b_3, y^b_3, \tilde{u}^b_3, w^b_3) \in T_c} P(\tilde{u}_1)P(u^b_2, x^b_3, y^b_3, \tilde{u}^b_3, w^b_3)$$

$$\leq |T_c|2^{-n(H(\bar{U}_1))-\epsilon}2^{-n(H(\bar{U}_2,X_3,Y_3,\bar{U}_3,W_3))-\epsilon}$$

$$\leq 2^n(H(\bar{U}_1,\bar{U}_2,X_3,Y_3,\bar{U}_3,W_3)-\epsilon)2^{-n(H(\bar{U}_1))-\epsilon}2^{-n(H(\bar{U}_2,X_3,Y_3,\bar{U}_3,W_3))-\epsilon}$$

$$= 2^{-n(H(\bar{U}_1))-\epsilon}2^n(H(\bar{U}_1|\bar{U}_2,X_3,Y_3,\bar{U}_3,W_3)-2\epsilon)$$

$$= 2^{-n(I(\bar{U}_1;\bar{U}_2,X_3,Y_3,\bar{U}_3,W_3))-3\epsilon}.$$

Hence, it follows that

$$R_1 \leq I(\bar{U}_1; \bar{U}_2, X_3, Y_3, \bar{U}_3, W_3). \quad (4.5)$$
CHAPTER 4. INNER BOUND

Following the same steps we can show that to make the probability of error arbitrarily small,

\[ R_2 \leq I(\bar{U}_2; \bar{U}_1, X_3, Y_3, \bar{U}_3, W_3) \] (4.6)

\[ R_1 + R_2 \leq I(\bar{U}_1, \bar{U}_2; X_3, Y_3, \bar{U}_3, W_3). \] (4.7)

This concludes the proof of Theorem 3.1.  

4.5 Inner Bound for Half-duplex Channel

The full-duplex encoding scheme can be used to derive a half-duplex case by finding special encoding and decoding functions. In this section, we will consider a channel that is time duplexed between a multiple access (MAC) phase and a multicast (MC) phase 4.1. During the MAC phase, User 1 and User 2 send their codewords to User 3. We also refer to this phase as the uplink. This takes up \( \eta \) fraction of the total time. During the MC phase, User 3, sends a common codeword (message) to Users 1 and 2. This consumes \( 1 - \eta \) fraction of the total time. We call this phase MC to distinguish between the case when User 3 has independent messages for Users 1 and 2 (which we call a broadcast phase or BC phase). We also refer to this phase as the downlink.

4.5.1 Test Channel

The half-duplex two-way channel that was explained above can be written as:

\[ p(y_1, y_2, y_3|x_1, x_2, x_3) = p_{MAC}(y_3|x_1, x_2)p_{MC}(y_1, y_2|x_3) \]

As explained in the full-duplex case, the channel inputs are formed by auxiliary random variables which are derived from different layers of messages. If we consider
Figure 4.1: *Half-duplex two-way network: Two states in the system.*

a new channel whose inputs are these auxiliary random variables \( (u_{10}, \tilde{u}_{10}, u_{11}, u_{20}, \tilde{u}_{20}, u_{22}, u_3 \text{ and } w_3) \), then such a test channel (see [5]) can be represented as:

\[
p^*(y_1, y_2, y_3 | u_1, \tilde{u}_1, u_2, \tilde{u}_2, u_3, \tilde{w}_3) = \sum_{x_1, x_2, x_3} p_{MAC}(y_3 | x_1, x_2) \gamma_1(x_1 | u_1, \tilde{u}_1, \tilde{u}_2) \gamma_2(x_2 | u_2, \tilde{u}_1, \tilde{u}_2) \\
\cdot p_{BC}(y_1, y_2 | x_3) \gamma_3(x_3 | u_3, \tilde{w}_3)
\]

where,

\[
p(u_1, u_2, u_3, \tilde{u}_1, \tilde{u}_2, \tilde{w}_3) = p(u_1)p(u_2)p(u_3)p(\tilde{u}_1 | u_1, y_1)p(\tilde{u}_2 | u_2, y_2)p(\tilde{w}_3 | y_3).
\]

Note the independent new information and the dependent incremental and feedback information.
4.5.2 Encoding scheme

We will denote the alphabets for the channel inputs and outputs in script letters. The coding scheme divides the input into blocks each consisting of \( n \) channel uses. For each block, the channel input is formed, in general from three parts:

- New information independent of messages of any previous block (we will denote this by \( u \), and its alphabet by \( \mathcal{U} \)).

- Feedback information obtained from the previous channel output (we will denote this by \( \hat{u} \), and its alphabet by \( \mathcal{W} \)).

- Incremental information about the previous channel input obtained using the previous feedback signal (we will denote this by \( \hat{u} \), and its alphabet by \( \mathcal{W} \)).

Let us define the input alphabets and the index sets first. All the sequences shown in the table above are typical sequences independently generated from their respective probability distributions. A major contribution of our two-way code is the way the feedback codebook is formed. So, next we explain how the feedback alphabet \( \mathcal{W}_3 \) is generated from \((U_{10}, U_{20})\). After the sequences \( u_{10} \) and \( u_{20} \) are generated with respect to the distributions \( p(u_{10}) \) and \( p(u_{20}) \) respectively, \( \max(I, J) \) number of sequences \( w_3(u_{10}, u_{20}) \) are generated such that, \( p(w_3|u_{10}) = p(u_{20}) \) and \( p(w_3|u_{20}) = p(u_{10}) \). This implies that if a receiver has access to \( w_3 \) and \( u_{10} \), it can decode \( u_{20} \) and similarly if
it has access to \( w_3 \) and \( u_{20} \), it can decode \( u_{10} \). In order to show that we can generate such sequences, we review the usual procedure to perform superposition coding.

Let us consider the case when a superposition code \( Z \) is formed from \( U_{10} \) and \( U_{20} \). A simple example alphabet set is shown in Figure 4.2. The usual way to produce such a code is to produce bins which are mapped to one of the alphabets (\( U_{20} \) in our example) and then generate individual codes in each bin mapped to the other alphabet (\( U_{10} \) in our example). But note that in such a case, the bin index can be mapped to the message index of one of the constituting alphabets. To understand this properly we have indexed the superimposed codebook with two indexes in Figure 4.2 (b). Note that the second index is repeated in each bin. This is what causes the redundancy in the information about user 2 and the bin index.

![Figure 4.2: Usual superposition coding using binning. The cloud centers signify one of the message sets.](image)

In order to put in the property required by us, we reshuffle the bin as shown in Figure 4.3. Now, the way \( \bar{W}_3 \) is generated is by producing typical sequences mapped
Figure 4.3: Superposition coding with binning. Note that in this case $p(z|u_{10}) = p(u_{20})$ and $p(z|u_{20}) = p(u_{10})$ by construction.

to the bins in $Z$. Note that as no index can be repeated in a bin, we can have
$\min(I, K)$ points in any bin and hence the number of bins turn out to be $\max(I, K)$
which is the alphabet size of $\hat{W}_3$.

In general, the channel outputs are generated as: $x_i^b = f_i^b(u_i, \tilde{u}_i, \tilde{w}_i)$, where ($i = 1, 2$ or 3). Here, the encoding functions $f_1$ and $f_2$ perform superposition coding while
$f_3$ is used to generate a broadcast code. Note that due to our MAC-MC configuration,
Users 1 and 2 do not send any feedback information and User 3 does not send any
incremental information.

Hence the encoding functions are as follows:

User 1:  $x_1^b = f_1^b(u_1^b(m_1^b), \tilde{u}_1^b)$  $f_1^b : \tilde{U}_1^n \times \tilde{U}_1^n \rightarrow \tilde{X}_1^n$

User 2:  $x_2^b = f_2^b(u_2^b(m_2^b), \tilde{u}_2^b)$  $f_2^b : \tilde{U}_2^n \times \tilde{U}_2^n \rightarrow \tilde{X}_2^n$

User 3:  $x_3^b = f_3^b(u_3^b(m), \tilde{w}_3^b)$  $f_3^b : \tilde{U}_3^n \times \tilde{W}_3^n \rightarrow \tilde{X}_3^n$
<table>
<thead>
<tr>
<th>Encoded msgs:</th>
<th>$u_1^b = (u_{10}^b(i), u_{11}^b(j))$</th>
<th>$u_1^b \in \mathcal{U}<em>1, u</em>{10}^b \in \mathcal{U}<em>{10}, u</em>{11}^b \in \mathcal{U}_{11}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$u_2^b = (u_{20}^b(k), u_{22}^b(l))$</td>
<td>$u_2^b \in \mathcal{U}<em>2, u</em>{20}^b \in \mathcal{U}<em>{20}, u</em>{22}^b \in \mathcal{U}_{22}$</td>
</tr>
<tr>
<td></td>
<td>$u_3^b = u_3^b(m)$</td>
<td>$u_3^b \in \mathcal{U}_3, m \in \mathcal{M}$</td>
</tr>
<tr>
<td>Incremental msgs:</td>
<td>$\tilde{u}<em>{10}^b(i') = (y</em>{b-1}^b, u_{10}^b(i'))$</td>
<td>$u_{b-1}^b(i')$</td>
</tr>
<tr>
<td></td>
<td>$\tilde{u}<em>{20}^b(k') = (y</em>{b-1}^b, u_{20}^b(k'))$</td>
<td>$u_{b-1}^b(k')$</td>
</tr>
<tr>
<td>Feedback msg:</td>
<td>$\tilde{w}_3^b = \tilde{w}_3^b(n)$</td>
<td>$\tilde{w}_3^b \in \tilde{W}_3, n \in \mathcal{N}$</td>
</tr>
</tbody>
</table>

Table 4.2: Encoding functions for the half duplex case.

At the beginning of block $b$, Users 1 and 2 form the incremental information $\tilde{u}_1^b$ and $\tilde{u}_2^b$ based on $y_{b-1}^b$, $u_1^b$ and $u_2^b$, $u_3^b$ respectively. It forms $x_1$ and $x_2$ as mentioned above and sends through the channel using it for $\eta$ fraction of the time. User 3 receives $y_3^b$ and forms $x_3$ as given above. The decoding part is explained in the next section.

The incremental information and the feedback information do not assume any decoding. All the decoding is lumped into the encoding functions $f_1, f_2$ and $f_3$.

### 4.5.3 Decoding scheme

Consider block $b, 1 \leq b \leq B$, which is divided into a MAC phase and a MC phase. During the MAC phase of the previous block, User 3 received $y_{3}^{b-1}$, and decoded $(\delta'', \delta', k'', \delta')$. During the current block $b$, User 3 decodes $(\delta', \delta', k', \delta)$. Similarly, during the MC phase of the previous block, User 1 and User 2 received $y_{1}^{b-1}, y_{2}^{b-1}$ and User 1 was able to decode $m', k'$ and User 2 was able to decode $m', i'$. During the current block’s MC phase, User 1 decodes $(\tilde{m}, \tilde{k})$ and User 2 decodes $(\tilde{m}, \tilde{i})$. As User 3 uses a broadcast code to encode the feedback codebook and his own independent message, Users 1 and 2 use a broadcast code decoding with the knowledge of the structure of the feedback code in the following way:

User 1 already knows $\tilde{u}_{10}$ perfectly. First it decodes $u_3$ and then decodes $u_{20}$. In
other words, first it finds $\hat{m} \in \mathcal{M}$ such that

$$(u^b_3(\hat{m}), y^b_1, \bar{u}^b_{10}(i), \bar{u}^b_{20}) \in A^a_e(U_3, Y_1, \bar{U}_{10}, \bar{U}_{20}),$$

and then it finds $\hat{k}' \in \mathcal{K}$ such that

$$(u^b_3(m), y^b_2, \bar{u}^b_{10}(i), \bar{u}^b_{20}(\hat{k})) \in A^a_e(U_3, Y_2, \bar{U}_{10}, \bar{U}_{20}).$$

Similarly for User 2, it first finds $\hat{m} \in \mathcal{M}$ and then $\hat{i} \in \mathcal{I}$ and such that

$$(u^b_3(\hat{m}), y^b_2, \bar{u}^b_{10}(\hat{i}), \bar{u}^b_{20}(\hat{k})) \in A^a_e(U_3, Y_2, \bar{U}_{10}, \bar{U}_{20}),$$

User 3 first finds $\hat{i} \in \mathcal{I}, \hat{j} \in \mathcal{J}, \hat{k}' \in \mathcal{K}$ and $\hat{l} \in \mathcal{L}$ such that

$$(\bar{u}^b_{10}(\hat{i}'), u^b_{11}(\hat{j}), \bar{u}^b_{20}(\hat{k}'), u^b_{22}(\hat{l}), \bar{u}^b_{3}(y^b_3) \in A^a_e(\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3, Y_3)$$

$$(\bar{u}^b_{10}(-1)(\hat{i}''), u^b_{11}(-1)(\hat{j}''), \bar{u}^b_{20}(-1)(\hat{k}''), u^b_{22}(-1)(\hat{l}''), \bar{u}^b_{3}(-1)(y^b_3) \in A^a_e(\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3, Y_3).$$

Then it forms the feedback information by decoding the coarse index $\hat{n} \in \mathcal{N}$

$$(u^b_{10}(\hat{i}'), u^b_{11}(\hat{j}), \bar{u}^b_{20}(k'), u^b_{22}(l), \bar{u}^b_{3}(\hat{n}), y^b_3) \in A^a_e(\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3, Y_3).$$

### 4.5.4 Probability of Error

Assume that the indices used to generate the channel inputs were $(i' = 1, i = 1, j = 1, k' = 1, k = 1, l = 1, m = 1)$, during block $b$.

Users 1 and 2 decode a broadcast code formed using $U_3$ and $\bar{W}_3$. But from the
encoding scheme we know that \( W_3 \) itself uses a superposition code formed using \( U_{10} \) and \( U_{20} \) with reshuffled binning as explained in the Section 4.5.2. To determine the probability of error, define the following error events

\[
E_{m}^1 = \{(U_3(m), Y_1) \in A_e^n(U_3, U_{20}, Y_1|U_{10})\}
\]
\[
E_{km}^1 = \{(U_{20}(k), U_3(m), Y_1) \in A_e^n(U_3, U_{20}, Y_1|U_{10})\}
\]
\[
E_{m}^2 = \{(U_3(m), Y_2) \in A_e^n(U_3, U_{10}, Y_1|U_{20})\}
\]
\[
E_{km}^2 = \{(U_{10}(i), U_3(m), Y_2) \in A_e^n(U_3, U_{10}, Y_2|U_{20})\}
\]

Notice that the probability that the correct message sequence fails to satisfy Equations 4.8 and 4.9 (i.e. \( P(E_{1c}^1) \)) is less than some arbitrarily small positive number \( 2^{-n\delta} \) (see [5] Lemma 1). Hence the probability of error is bounded by (an index of 0 means an error)

\[
P_{e1} < 2^{-n\delta} + \sum_{m \neq 1} E_{m}^1 + \sum_{k \neq 1} E_{km}^1 < 2^{-n\delta} + (M - 1)E_0^1 + (K - 1)E_{10}^1 < 2^{-n\delta} + 2^{nR_3}E_0^1 + 2^{nR_{20}}E_{10}^1 < 2^{-n\delta} + 2^{nR_3}2^{-n[I(U_3;Y_1|U_{10})-\epsilon]} + 2^{nR_{20}}2^{-n[I(U_{20};Y_1|U_{10};U_3)-\epsilon]},
\]

and similarly for User 2

\[
P_{e2} < 2^{-n\delta} + 2^{nR_{20}}2^{-n[I(U_{20};Y_1|U_{10};U_3)-\epsilon]} + 2^{nR_3}2^{-n[I(U_3;Y_1|U_{10})-\epsilon]}.
\]

Note that User 3 uses the well-known typical set decoding of the different message layers before it decodes the feedback layer. Define the following error events \( (E_{ijkl}F_{ik}) \)

\[
E_{ijkl}^3 = \{(\hat{u}_{10}^b(\hat{v}), u_{11}^b(\hat{j}), \bar{u}_{20}^b(\hat{k}), u_{22}^b(\hat{\ell}), \bar{w}_3^b, y_3^b) \in A_e^n(\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3, Y_3)\},
\]
\[ F^3_{tk} = \{ (u_{10}^{b-1}(i''), u_{10}^{b-1}(i'), u_{11}^{b-1}(j'), u_{20}^{b-1}(k''), u_{20}^{b-1}(k'), u_{22}^{b-1}(l'), w_3^{b-1}, y_8^{b-1}) \in A_k^n(\bar{U}_{10}, U_{10}, U_{11}, \bar{U}_{20}, U_{20}, U_{22}, \bar{W}_3, Y_3) \}. \]

Using these error events, the probability of error for User 3 can be bounded as:

\[
P_{e3} < 2^{-n\delta} + (I - 1)E^3_{1000}F^3_{10} + (J - 1)E^3_{0100} + (K - 1)E^3_{0010}F^3_{01} + (L - 1)E^3_{0001} \\
+ (I - 1)(J - 1)E^3_{1000}F^3_{10} + (I - 1)(K - 1)E^3_{1010}F^3_{11} + (I - 1)(L - 1)E^3_{1001}F^3_{10} \\
+ (J - 1)(K - 1)E^3_{0110}F^3_{01} + (J - 1)(L - 1)E^3_{0101}F^3_{01} + (K - 1)(L - 1)E^3_{0011}F^3_{01} \\
+ (I - 1)(J - 1)(K - 1)E^3_{1110}F^3_{11} + (I - 1)(J - 1)(L - 1)E^3_{1101}F^3_{11} \\
+ (I - 1)(K - 1)(L - 1)E^3_{1011}F^3_{11} + (I - 1)(K - 1)(L - 1)E^3_{1111}F^3_{11} \\
< 2^{-n\delta} + 2^n[R_{10} - I(\bar{U}_{10}; Y_3|\bar{U}_{20}, U_{11}, U_{22}, \bar{W}_3) - I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 6\epsilon] \\
+ 2^n[R_{11} - I(U_{11}; Y_3|\bar{U}_{10}, \bar{U}_{20}, U_{22}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 3\epsilon] \\
+ 2^n[R_{20} - I(\bar{U}_{20}; Y_3|\bar{U}_{10}, U_{11}, U_{22}, \bar{W}_3) - I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 6\epsilon] \\
+ 2^n[R_{22} - I(U_{22}; Y_3|\bar{U}_{10}, \bar{U}_{20}, U_{11}, \bar{W}_3) + 3\epsilon] \\
+ 2^n[R_{10} + R_{11} - I(\bar{U}_{10}, U_{11}; Y_3|\bar{U}_{20}, U_{22}, \bar{W}_3) - I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 8\epsilon] \\
+ 2^n[R_{10} + R_{20} - I(\bar{U}_{10}, U_{20}; Y_3|U_{11}, U_{22}, \bar{W}_3) - I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 8\epsilon] \\
+ 2^n[R_{10} + R_{22} - I(U_{10}, U_{22}; Y_3|U_{11}, \bar{U}_{20}, \bar{W}_3) - I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 8\epsilon] \\
+ 2^n[R_{11} + R_{20} - I(U_{11}, U_{20}; Y_3|\bar{U}_{10}, U_{22}, \bar{W}_3) - I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 8\epsilon] \\
+ 2^n[R_{11} + R_{22} - I(U_{11}, U_{22}; Y_3|\bar{U}_{10}, \bar{U}_{20}, \bar{W}_3) + 3\epsilon] \\
+ 2^n[R_{20} + R_{22} - I(U_{20}, U_{22}; Y_3|\bar{U}_{10}, U_{11}, \bar{W}_3) - I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 8\epsilon] \\
+ 2^n[R_{10} + R_{11} + R_{20} - I(\bar{U}_{10}, U_{11}, U_{20}; Y_3|U_{22}, \bar{W}_3) - I(U_{10}; U_{20}; Y_3|\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 6\epsilon] \\
+ 2^n[R_{10} + R_{11} + R_{22} - I(\bar{U}_{10}, U_{11}, \bar{U}_{20}; Y_3|\bar{U}_{20}, \bar{W}_3) - I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 6\epsilon] \\
+ 2^n[R_{10} + R_{20} + R_{22} - I(U_{10}, U_{20}, U_{22}; Y_3|U_{11}, \bar{W}_3) - I(U_{10}; U_{20}; Y_3|\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 6\epsilon] \\
+ 2^n[(R_{11} + R_{20} + R_{22}) - I(U_{11}, U_{20}, U_{22}; Y_3|\bar{U}_{10}, \bar{W}_3) - I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22}, \bar{W}_3) + 6\epsilon] \]
CHAPTER 4. INNER BOUND

\[ + 2^n((R_{10} + R_{11} + R_{20} + R_{22}) - I(\tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}; Y_3 | \tilde{W}_3) - I(U_{10}, U_{20}; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3) + 6\epsilon) \]

The last decoding operation gives rise to the rate equation,

\[
\max(R_{10}, R_{20}) < I(\tilde{W}_3; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, U_3).
\] (4.10)

Hence, to make the probabilities of error arbitrarily small (as \(n \to \infty\)), the rates are constrained by the following equations which gives the achievability rate region for the half-duplex case:

\[
R_{10} < \eta \min[I(U_{10}; Y_2 | U_{20}, U_3), I(\tilde{W}_3; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, U_3)
\]

\[
I(\tilde{U}_{10}; Y_3 | \tilde{U}_{20}, U_{11}, U_{22}) + I(U_{10}; Y_3 | U_{20}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{11} < \eta I(U_{11}; Y_3 | \tilde{U}_{10}, \tilde{U}_{20}, U_{10}, U_{20}, U_{22})
\]

\[
R_{20} < \eta \min[I(U_{20}; Y_1 | U_{10}, U_3), I(\tilde{W}_3; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, U_3)
\]

\[
I(\tilde{U}_{20}; Y_3 | \tilde{U}_{10}, U_{11}, U_{22}) + I(U_{20}; Y_3 | U_{10}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{22} < \eta I(U_{22}; Y_3 | U_{10}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22})
\]

\[
R_{10} + R_{11} < \eta[I(\tilde{U}_{10}, U_{11}; Y_3 | \tilde{U}_{20}, U_{22}) + I(U_{10}; Y_3 | U_{20}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{10} + R_{20} < \eta[I(\tilde{U}_{10}, U_{20}; Y_3 | U_{11}, U_{22}) + I(U_{10}, U_{20}; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{10} + R_{22} < \eta[I(\tilde{U}_{10}, U_{22}; Y_3 | U_{11}, \tilde{U}_{20}) + I(U_{10}; Y_3 | U_{20}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{11} + R_{20} < \eta[I(U_{11}, \tilde{U}_{20}; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}) + I(U_{20}; Y_3 | U_{10}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{11} + R_{22} < \eta I(U_{11}, U_{22}; Y_3 | \tilde{U}_{10}, \tilde{U}_{20})
\]

\[
R_{20} + R_{22} < \eta[I(\tilde{U}_{20}, U_{22}; Y_3 | \tilde{U}_{10}, U_{11}) + I(U_{20}; Y_3 | U_{10}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{10} + R_{11} + R_{20} < \eta[I(\tilde{U}_{10}, U_{11}, \tilde{U}_{20}; Y_3 | U_{22}) + I(U_{10}, U_{20}; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{10} + R_{11} + R_{22} < \eta[I(\tilde{U}_{10}, U_{11}, U_{22}; Y_3 | \tilde{U}_{20}) + I(U_{10}; Y_3 | U_{20}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{10} + R_{20} + R_{22} < \eta[I(\tilde{U}_{10}, \tilde{U}_{20}, U_{22}; Y_3 | U_{11}) + I(U_{10}, U_{20}; Y_3 | \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]

\[
R_{11} + R_{20} + R_{22} < \eta[I(U_{11}, \tilde{U}_{20}, U_{22}; Y_3 | \tilde{U}_{10}) + I(U_{20}; Y_3 | U_{10}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)]
\]
\[ R_{19} + R_{11} + R_{20} + R_{22} < \eta[I(\tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}; Y_3) \\
+ I(U_{10}, U_{20}; Y_3|\tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22}, \tilde{W}_3)] \]

\[ R_3 < (1 - \eta) \min[I(U_3; Y_1|U_{10}, U_{20}), I(U_3; Y_2|U_{10}, U_{20})]. \]

The rate region that we obtained appears to be similar to the rate region for a four sender MAC [1] and MAC with generalized feedback [5]. One thing to notice is the reduction in the signal to interference ratio for Users 1 and 2 in the downlink in our coding scheme due to the processing done by User 3. We show in Chapter 6 that this inner bound can be reduced to that in [5] which has been shown to be a generalized version of a four sender MAC. It is difficult to derive any intuition from these rate equations but the gains will be clear when we analyze the encoding scheme for a Gaussian channel and consider some interesting special cases in Chapter 6.
Chapter 5

Outer Bound

This chapter contains the proof of the outer bound given in Theorem 3.2. We restate the naming conventions in more detail for the random variables for clarity. We will denote channel inputs by \( X \), channel outputs by \( Y \) and the messages by \( W \) with proper subscripts and superscripts as described below. The encoding will use \( B \) blocks of \( n \) channel uses each. We introduce the following notation for all variables. First, \( A_{u,b} \) will denote the \( b \) block of User \( u \); note the block is of length \( n \) symbols. The vector of \( b \) consecutive blocks will be denoted by \( A^b_u = (A_{u,1}, \ldots, A_{u,b-1}) \). Finally, the vector with all blocks \( b = 1, \ldots, B \) will be denoted by \( \mathbf{A} = (A_{u,1}, \ldots, A_{u,B}) \). Note that boldface variables denote vectors for \( B \) blocks each of length \( n \) drawn from the same distribution. For example, \( \mathbf{X}_1 = (X_{1,1}, X_{1,2}, \ldots, X_{1,B}) \) is the vector input to the channel for User 1 for \( B \) blocks and \( (X_{1,i} = (X_{1,i}^1, X_{1,i}^2, \ldots, X_{1,i}^n)) \) is the channel input for User 1 for the \( i^{th} \) block which is a vector \( (n \text{-tuple}) \) itself each independently drawn from the same distribution \( (X_{1,i}^k \sim X_1) \). Finally, we introduce a random variable \( Z_{u,b} = (X_{u,b-1}, Y_{u,b-1}) \).

The model in Figure 3.1 can be mathematically defined by a Markov chain as shown in Figure 5.1.

Note that \( X \)'s are deterministic functions of \( W \)'s and \( Y \)'s (i.e. \( X_{1,b} = f_1(W_1, Y_{1,b-1}) \),

37
$X_{2,b} = f_2(W_2, Y_{2,b-1})$ and $X_{3,b} = g(W_3, Y_{3,b})$ for some deterministic encoding functions $f_1, f_2, g$.

We prove each of the inequalities in the following subsections. To avoid notational overload, we will *suppress* conditioning on the time-sharing variable $q$ in the proof, but the reader should modify the appropriate steps to account for time-sharing.

### 5.1 Outer Bound for $R_3$

We start by bounding $R_3$ first as follows.

$$nBR_3 = H(W_3|W_1, W_2)$$

$$= H(W_3|W_1, W_2, Y_1) + I(W_3; Y_1|W_1, W_2)$$  \hspace{1cm} (5.1)

$$= H(W_3|W_1, W_2, Y_2) + I(W_3; Y_2|W_1, W_2).$$  \hspace{1cm} (5.2)

The first term can be upper bounded by an arbitrarily small positive number using Fano’s theorem (similarly for all other bounds given below). The mutual information (second term in Equation 5.1) can be upper bounded as

$$I(W_3; Y_1|W_1, W_2)$$

$$= H(Y_1|W_1, W_2) - H(Y_1|W_1, W_2, W_3)$$

$$= \sum_{b=1}^{B} [H(Y_{1,b}|W_1, W_2, Y_{1,b-1}) - H(Y_{1,b}|W_1, W_2, W_3, Y_{1,b-1})]$$
\[ \sum_{b=1}^{B} \{ H(Y_1, W_1, W_2, Y_b^{-1}, X_1) - H(Y_1, W_1, W_2, X_1, Y_b^{-1}, Y_3) \} \]

\[ \sum_{b=1}^{B} \{ H(Y_1, W_1, W_2, Y_b, Y_b^{-1}, Z_1) - H(Y_1, W_1, W_2, X_1, Y_b^{-1}, Y_b, X_3, Z_1) \} \]

\[ \sum_{b=1}^{B} \{ H(Y_1, X_1, Z_{1,b}) - H(Y_1, X_1, X_{3,b}, Z_{1,b}) \} \]

\[ = \sum_{b=1}^{B} I(Y_1; X_{3,b}|X_1, Z_1) \]

\[ \leq n BI(X_3; Y_1|X_1, Z_1), \]

where

(a) Using \( X_{1,b} = f_1(W_1, Y_{1,b-1}) \) in both the terms and the inequality is due to introducing \( Y_3^b \) in the second term.

(b) Using \( Z_{1,b} = (X_{1,b-1}, Y_{1,b-1}) \) in both the terms and in the second term using the fact that \( X_{3,b} = g(W_3, Y_{3,b}) \).

(c) For the first term, we just remove a few variables and hence get an upper bound and for the second term, \( Y_1,b \) being the channel output is independent of all other terms given \( X_{3,b} \) and \( X_{1,b} \) as it uses feedback.

(d) Using the DMC property for the n-channel uses in each block. Hence, \( R_3 \leq I(X_3; Y_1|X_1, Z_1) \) and similarly from equation 5.2, \( R_3 \leq I(X_3; Y_2|X_2, Z_2) \). So,

\[ R_3 \leq \min[I(X_3; Y_1|X_1, Z_1), I(X_3; Y_2|X_2, Z_2)]. \]  \hspace{1cm} (5.3)
5.2 Outer Bound for $R_1$ and $R_2$

Note that

$$nBR_1 = H(W_1|W_2, W_3)$$

$$= H(W_1|W_2, W_3, Y_3) + I(W_1; Y_3|W_2, W_3).$$

The second term can be upper bounded as

$$I(W_1; Y_3|W_2, W_3) \leq I(W_1; Y_2, Y_3|W_2, W_3)$$

$$= I(W_1; Y_2|W_2, W_3) + I(W_1; Y_3|W_2, W_3, Y_2).$$  (5.4)

The first term can be upper bounded very similar to the proof of Section 5.1,

$$I(W_1; Y_2|W_2, W_3)$$

$$= H(Y_2|W_2, W_3) - H(Y_2|W_1, W_2, W_3)$$

$$= \sum_{b=1}^{B} [H(Y_{2,b}|W_2, W_3, Y_{2,b}^{-1}) - H(Y_{2,b}|W_1, W_2, W_3, Y_{2,b}^{-1})]$$

$$\leq \sum_{b=1}^{B} [H(Y_{2,b}|W_2, W_3, X_{2,b}^{-1}, Y_{2,b}^{b-1}, X_2^b) - H(Y_{2,b}|W_1, W_2, W_3, X_{2,b}^{b-1}, Y_{2,b}^{b-1}, X_2^b)]$$

$$\leq \sum_{b=1}^{B} [H(Y_{2,b}|W_2, W_3, X_{2,b}^{-1}, Y_{2,b}^{b-1}, Z_2^b) - H(Y_{2,b}|W_1, W_2, W_3, X_{2,b}^{b-1}, Y_{2,b}^{b-1}, X_2^b, Z_2^b)]$$

$$\leq \sum_{b=1}^{B} I(Y_{2,b}; X_{1,b}, X_{2,b}, Z_{2,b}, W_3)$$

$$\leq \sum_{b=1}^{B} \sum_{i=1}^{n} I(Y_{2,b}^i; X_{1,b}^i, X_{2,b}^i, Z_{2,b}^i, W_3)$$

$$\leq nBI(X_1; Y_2|X_2, Z_2, W_3),$$
where

(a) Using $X_{2,b} = f_2(W_2, Y_{2,b-1})$ in both the terms and the inequality is due to introducing $Y_1^{b-1}$ in the second term.

(b) Using $Z_{2,b} = (X_{2,b-1}, Y_{2,b-1})$ in both the terms and in the second term using the fact that $X_{3,b} = g(W_3, Y_{3,b})$.

(c) For the first term, we just remove a few variables and hence get an upper bound and for the second term, $Y_{2,b}$ being the channel output is independent of all other terms given $X_{1,b}$, $X_{2,b}$ and $W_3$.

(d) Using the DMC property for the $n$-channel uses in each block.

Now consider the second term in (5.4) as follows

\[
I(W_1; Y_3 | W_2, W_3, Y_2) \\
= H(Y_3 | W_2, W_3, Y_2) - H(Y_3 | W_1, W_2, W_3, Y_2) \\
\leq \sum_{b=1}^{B} [H(Y_{3,b} | W_2, W_3, Y_2, X_2^{b-1}, W_3, Y_3^{b-1}, X_3^{b-1})] \\
- H(Y_{3,b} | W_1, W_2, W_3, X_3^{b-1}, Y_3^{b-1}, Y_1^{b-1}, Y_2^{b-1})] \\
= \sum_{b=1}^{B} [H(Y_{3,b} | W_2, W_3, X_2^{b-1}, Y_2^{b-1}, Z_2, X_3^{b-1}, Y_3^{b-1}, Z_3)] \\
- H(Y_{3,b} | W_1, W_2, W_3, X_3^{b-1}, X_1^{b-1}, Y_1^{b-1}, Y_2^{b-1}, X_1^{b-1}, X_2^{b-1}, Z_2^{b}, Z_3^{b})] \\
\leq \sum_{b=1}^{B} I(X_{1,b}; Y_{3,b} | X_{2,b}, Z_{2,b}, Z_{3,b}) \\
= \sum_{b=1}^{B} n I(X_1^i; Y_3^i | X_2^i, Z_2^i, Z_3^i) \\
\leq n B I(X_1; Y_3 | X_2, Z_2, Z_3),
\]
where
(a) Using \( X_{3,b-1} = g(W_3, Y_{3,b-1}) \) in both the terms and the inequality is due to introducing \( Y_1^{b-1} \) and \( Y_2^{b-1} \) in the second term.
(b) Using \( Z_3^b = (X_3^{b-1}, Y_3^{b-1}) \) in both the terms and in the second term using the fact that \( X_{1,b-1} = f_1(W_1, Y_{1,b-1}) \) and \( X_{2,b-1} = f_2(W_2, Y_{2,b-1}) \).
(c) For the first term, we just remove a few variables and hence get an upper bound and for the second term, \( Y_{3,b} \) being the channel output is independent of all other terms given \( X_{1,b} \) and \( X_{2,b} \).
(d) Using the DMC property for the n-channel uses in each block. Hence,

\[
R_1 \leq I(X_1; Y_2|X_2, Z_2, W_3) + I(X_1; Y_3|X_2, Z_2, Z_3),
\]

(5.5)

and similarly by symmetry,

\[
R_2 \leq I(X_2; Y_1|X_1, Z_1, W_3) + I(X_2; Y_3|X_1, Z_1, Z_3).
\]

(5.6)

### 5.3 Outer Bound for \( R_1 + R_2 \)

Finally, consider the sum bound

\[
nB(R_1 + R_2) = H(W_1, W_2|W_3)
\]

\[
= H(W_1, W_2|W_3, Y_3) + I(W_1, W_2; Y_3|W_3).
\]

The second term can be bounded as

\[
I(W_1, W_2; Y_3|W_3)
\]

\[
= H(Y_3|W_3) - H(Y_3|W_1, W_2, W_3)
\]
\[= \sum_{b=1}^{B} [H(Y_{3,b}|W_3, Y_3^{b-1}) - H(Y_{3,b}|W_1, W_2, W_3, Y_3^{b-1})] \]

\[\leq \sum_{b=1}^{B} [H(Y_{3,b}|W_3, Y_3^{b-1}, X_3^{b-1}) - H(Y_{3,b}|W_1, W_2, W_3, X_3^{b-1}, Y_3^{b-1}, Y_1^{b-1}, Y_2^{b-1})] \]

\[\leq \sum_{b=1}^{B} [H(Y_{3,b}|W_3, X_3^{b-1}, Y_3^{b-1}, Z_3^b) - H(Y_{3,b}|W_1, W_2, W_3, X_3^{b-1}, Y_1^{b-1}, Y_2^{b-1}, Y_3^{b-1}, X_1^b, X_2^b, Z_3^b)] \]

\[\leq \sum_{b=1}^{B} [H(Y_{3,b}|Z_3^b) - H(Y_{3,b}|X_1, X_2, Z_3^b)] \]

\[= \sum_{b=1}^{B} I(X_1, X_2, Y_3, Z_3^b) \]

\[\leq \sum_{b=1}^{B} \sum_{i=1}^{n} I(X_{1,b}^i, X_{2,b}^i, Y_{3,b}^i, Z_{3,b}^i) \]

\[\leq nBI(X_1, X_2; Y_3|X_3, Z_3), \]

where

(a) Using \(X_{3,b-1} = g(W_3, Y_{3,b-1})\) in both the terms and the inequality is due to introducing \(Y_1^{b-1}\) and \(Y_2^{b-1}\) in the second term.

(b) Using \(Z_3^b = (X_3^{b-1}, Y_3^{b-1})\) in both the terms and in the second term using the fact that \(X_{1,b-1} = f_1(W_1, Y_{1,b-1})\) and \(X_{2,b-1} = f_2(W_2, Y_{2,b-1})\).

(c) For the first term, we just remove a few variables and hence get an upper bound and for the second term, \(Y_{3,b}\) being the channel output is independent of all other terms given \(X_{1,b}\) and \(X_{2,b}\).

(d) Using the DMC property for the n-channel uses in each block. Hence,

\[R_1 + R_2 \leq I(X_1, X_2; Y_3|Z_3). \quad (5.7)\]

This concludes the proof. \[\blacksquare\]

To find a half duplex outer bound, the three users time share between two states
(namely MAC and MC). This can be achieved by choosing, $Q \in \{q_{mc}, q_{mac}\}$ such that $p(Q = q_{mac}) = \eta$ and $p(Q = q_{mc}) = 1 - \eta$. So, $|Q| = 2$ in this case. Also, if there is no feedback, there is no need to use the auxiliary random variables $Z_i$ and hence the outer bound simplifies to the known capacity region of the multiple-access channel.

The applicability of Theorem 3.2 is reduced by our inability to prove an upper bound on the alphabet sizes $Z_1$, $Z_2$ and $Z_3$ (alphabets for the auxiliary random variables $Z_1$, $Z_2$ and $Z_3$). A future improvement can be to find special cases like shown in [35] where such an upper bound is possible to complete.
Chapter 6

Special cases

6.1 Gaussian case

One of the most commonly analyzed special case of a memoryless channel is a Gaussian channel. When the channel is Gaussian, Gaussian distributed codewords have the best performance. Hence we consider Gaussian distributed codewords to find the rate region. The system model is shown in Figure 6.1. The achievable rate region for this channel is similar to the discrete memoryless channel we considered before. There are some small distinctions though. We no longer need a time sharing random

![Figure 6.1](image)

Figure 6.1: (a) Uplink and (b) Downlink for the half-duplex Gaussian channel
variable $Q$ to make the rate region convex. The rate region is made convex by taking a convex hull over all possible values of the power distribution variables as shown in Figure 6.2. These variables ($\alpha, \beta$ and $\lambda$) distribute the power allocated to the users among the different message layers of the coding scheme. The code generation part has two more subtle differences from the DMC case. Firstly, instead of looking for typical sequences, we draw i.i.d. Gaussian sequences from the respective distributions. Secondly, the $Z$ sequence (the feedback codeword) as is formed by just adding the two Gaussian components $U_{10}$ and $U_{20}$ with the proper power distribution values. Users 1 and 2 generate Gaussian channel inputs from their messages in the following way (according to the power constraints $E[X_1^2] = P_1$ and $E[X_2^2] = P_2$)

$$x_1 = \sqrt{\alpha_1 \beta_1 P_1} \ u_{10} + \sqrt{\alpha_1 \lambda_1 P_1} \ \tilde{u}_{10} + \sqrt{\alpha_1 \lambda_2 P_1} \ \tilde{u}_{20} + \sqrt{\alpha_1 \beta_1 P_1} \ u_{11}$$

$$x_2 = \sqrt{\alpha_2 \beta_2 P_2} \ u_{20} + \sqrt{\alpha_2 \lambda_1 P_2} \ \tilde{u}_{10} + \sqrt{\alpha_2 \lambda_2 P_2} \ \tilde{u}_{20} + \sqrt{\alpha_2 \beta_2 P_2} \ u_{22}.$$ 

During the MAC mode, User 3 receives the channel output formed by $x_1$ and $x_2$ as $y_3$

$$y_3 = x_1 + x_2 + z_3. \quad (6.1)$$

It decodes a part of the message ($u_{11}, u_{22}, \tilde{u}_{10}, \tilde{u}_{20}, u'_{10}, u'_{20}$) and subtracts the decoded...
information from the channel output. So it is left with $y'_3$

$$y'_3 = \sqrt{\alpha_1\bar{\beta}_1 P_1} \ u_{10} + \sqrt{\alpha_2\bar{\beta}_2 P_2} \ u_{20} + z_3. \quad (6.2)$$

The power constraint of User 3 is given as $E[X_i^2] = P_3$. Now it can decode $Z$ which gives rise to a sum rate constraint which is not tight as in the DMC case hence it does not show up in the final equations. User 3 allocates a fraction of its available power to send feedback data $P_f = \gamma P_3$ and the rest is used for its own data,

$$x_3 = \sqrt{(1 - \gamma) P_3} \ u_3 + \sqrt{\frac{\gamma P_3}{P_f^*}} z. \quad (6.3)$$

where $P_f^* = N_3 + \alpha_1\bar{\beta}_1 P_1 + \alpha_2\bar{\beta}_2 P_2$. This is sent out by User 3 during the MC phase and Users 1 and 2 receive $y_1 = x_3 + z_1$ and $y_2 = x_3 + z_2$ respectively. So, the channel outputs can be expanded as

$$y_1 = \sqrt{(1 - \gamma) P_3} \ u_3 + \sqrt{\frac{\gamma P_3}{P_f^*}} \left( \sqrt{\alpha_1\bar{\beta}_1 P_1} \ u_{10} + \sqrt{\alpha_2\bar{\beta}_2 P_2} \ u_{20} \right) + z_1$$

$$y_2 = \sqrt{(1 - \gamma) P_3} \ u_3 + \sqrt{\frac{\gamma P_3}{P_f^*}} \left( \sqrt{\alpha_1\bar{\beta}_1 P_1} \ u_{10} + \sqrt{\alpha_2\bar{\beta}_2 P_2} \ u_{20} \right) + z_2$$

$$y_3 = \sqrt{\alpha_1\bar{\beta}_1 P_1} \ u_{10} + \sqrt{\alpha_1\lambda_1 P_1} \ \tilde{u}_{10} + \sqrt{\alpha_1\lambda_2 P_1} \ \tilde{u}_{20} + \sqrt{\alpha_1\bar{\beta}_1 P_1} \ u_{11}$$

$$+ \sqrt{\alpha_2\bar{\beta}_2 P_2} \ u_{20} + \sqrt{\alpha_2\lambda_1 P_2} \ \tilde{u}_{10} + \sqrt{\alpha_2\lambda_2 P_2} \ \tilde{u}_{20} + \sqrt{\alpha_2\bar{\beta}_2 P_2} \ u_{22} + z_3. \quad (6.4)$$

The decoding scheme is also very similar to the DMC case. For users 1 and 2, they subtract our their own part of the message and decode the remaining part like a Gaussian broadcast code. First they decode User 3’s message and then they decode the other user’s message. Hence the first users decoding depends on

$$y_1 - \sqrt{\frac{\gamma P_3\alpha_1\bar{\beta}_1 P_1}{P_f^*}} \ u_{10} = \sqrt{(1 - \gamma) P_3} \ u_3 + \sqrt{\frac{\gamma P_3\alpha_2\bar{\beta}_2 P_2}{P_f^*}} \ u_{20} + z_1,$$
and similarly, the second user’s decoding depends on

\[ y_2 = \sqrt{\frac{\gamma P_3 \alpha_1 \beta_2 P_2}{P_f^*}} u_{20} = \sqrt{(1 - \gamma) P_3} u_3 + \sqrt{\frac{\gamma P_3 \alpha_1 \beta_1 P_1}{P_f^*}} u_{10} + z_2. \]

Such decoding gives rise to the following rate equations

\[ R_{10} < (1 - \eta)C \left( \frac{\alpha_1 \beta_1 P_1 \gamma P_3}{P_f^* N_2} \right) \]  \hspace{1cm} (6.5a)

\[ R_{20} < (1 - \eta)C \left( \frac{\alpha_2 \beta_2 P_2 \gamma P_3}{P_f^* N_1} \right) \]  \hspace{1cm} (6.5b)

\[ R_3 < (1 - \eta)C \left( \frac{(1 - \gamma) P_3}{\gamma P_3 + \max(N_1, N_2)} \right). \]  \hspace{1cm} (6.5c)

User 3 in addition to decoding the layered messages, decodes the lumped feedback codebook as \( \tilde{w}_3 \). Hence, its decoding depends on

\[ \tilde{w}_3 + (\sqrt{\alpha_1 \lambda_1 P_1} + \sqrt{\lambda_2 \lambda_2 P_2}) \tilde{u}_{10} + (\sqrt{\alpha_1 \lambda_2 P_1} + \sqrt{\lambda_2 \beta_2 P_2}) \tilde{u}_{20} + \sqrt{\alpha_1 \beta_1 P_1} u_{11} \]
\[ + \sqrt{\alpha_2 \beta_2 P_2} u_{22} + z_3 \]

\[ = \tilde{w}_3 + \tilde{w}_3' + \sqrt{\alpha_1 \beta_1 P_1} u_{11} + \sqrt{\alpha_2 \beta_2 P_2} u_{22} + z_3. \]

### 6.1.1 Achievable Rate Region

Using the above, we can calculate the right hand side terms of the inner bound given in last part of Chapter 4. The rate region region achieved by the coding scheme proposed above is defined as follows; for simplicity denote \( C(x) = \frac{1}{2} \log(1 + x) \). \( \mathcal{R}_f = (R_1 = (R_{10} + R_{11}), R_2 = (R_{20}, R_{22}), R_3) \) is the set of all non-negative \( (R_{10}, R_{11}, R_{20}, R_{22}, R_3) \) satisfying the following inequalities:

\[ R_{10} < \min \left[ \eta C(\alpha_1 \beta_1 P_1), (1 - \eta)C \left( \frac{\alpha_1 \beta_1 P_1 \gamma P_3}{P_f^* N_2} \right) \right] \]
\[ R_{20} < \min \left[ \eta C(\alpha_2 \beta_2 P_2), (1 - \eta) C\left( \frac{\alpha_2 \beta_2 P_2 \gamma P_3}{P_f^* N_2} \right) \right] \]

\[ R_{11} < \eta C\left( \frac{\alpha_1 \beta_1 P_1}{N^*} \right) \]

\[ R_{22} < \eta C\left( \frac{\alpha_2 \beta_2 P_2}{N^*} \right) \]

\[ R_{10} + R_{20} < \eta C\left( \frac{P^*}{N^*} \right) + \eta C(\alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2) \]

\[ R_{10} + R_{11} < \eta C\left( \frac{\lambda_1 P^* + \alpha_1 \beta_1 P_1}{N^*} \right) + \eta C(\alpha_1 \beta_1 P_1) \]

\[ R_{10} + R_{22} < \eta C\left( \frac{\lambda_1 P^* + \alpha_2 \beta_2 P_2}{N^*} \right) + \eta C(\alpha_2 \beta_2 P_2) \]

\[ R_{20} + R_{11} < \eta C\left( \frac{\lambda_2 P^* + \alpha_1 \beta_1 P_1}{N^*} \right) + \eta C(\alpha_2 \beta_2 P_2) \]

\[ R_{20} + R_{22} < \eta C\left( \frac{\lambda_2 P^* + \alpha_2 \beta_2 P_2}{N^*} \right) + \eta C(\alpha_2 \beta_2 P_2) \]

\[ R_{11} + R_{22} < \eta C\left( \frac{\alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2}{N^*} \right) \]

\[ R_{10} + R_{20} + R_{11} < \eta C\left( \frac{P^* + \alpha_1 \beta_1 P_1}{N^*} \right) + \eta C(\alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2) \]

\[ R_{10} + R_{20} + R_{22} < \eta C\left( \frac{P^* + \alpha_2 \beta_2 P_2}{N^*} \right) + \eta C(\alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2) \]

\[ R_{10} + R_{11} + R_{22} < \eta C\left( \frac{\lambda_1 P^* + \alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2}{N^*} \right) + \eta C(\alpha_1 \beta_1 P_1) \]

\[ R_{20} + R_{11} + R_{22} < \eta C\left( \frac{\lambda_2 P^* + \alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2}{N^*} \right) + \eta C(\alpha_2 \beta_2 P_2) \]

\[ R_{10} + R_{11} + R_{20} + R_{22} < \eta C\left( \frac{P^* + \alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2}{N^*} \right) + \eta C(\alpha_1 \beta_1 P_1 + \alpha_2 \beta_2 P_2) \]

\[ R_{10} + R_3 < (1 - \eta) C\left( \frac{\alpha_1 \beta_1 P_1 \gamma P_3}{P_f^* N_2} + \frac{(1 - \gamma) P_3}{N_2} \right) \]
Figure 6.3: Rate region without feedback, $\mathcal{R}_{nf}$

\[
R_{20} + R_3 < (1 - \eta)C \left( \frac{\alpha_2 \beta_2 P_2 \gamma P_3}{P_f N_1} + \frac{(1 - \gamma)P_3}{N_1} \right)
\]
\[
R_3 < (1 - \eta)C \left( \frac{(1 - \gamma)P_3}{\gamma P_3 + \max(N_1, N_2)} \right)
\]

where, $P^* = ((\bar{\alpha}_1 P_1)^{\frac{1}{2}} + (\bar{\alpha}_2 P_2)^{\frac{1}{2}})^2$, $P_f^* = 1 + \alpha_1 \bar{\beta}_1 P_1 + \alpha_2 \bar{\beta}_2 P_2$
and $N^* = 1 + \alpha_1 \bar{\beta}_1 P_1 + \alpha_2 \bar{\beta}_2 P_2$. Here, $\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1$ are power distribution parameters and each is freely chosen between zero and unity, $N_3 = 1$, and $\lambda_2 = 1 - \lambda_1$.

### 6.2 Symmetric Case

Note that the rate region without any feedback from receivers is obtained by simple time-sharing between the MAC and MC mode. In this case the system uses spatially independent inputs, $p(x_1, x_2, x) = p(x_1)p(x_2)p(x)$, resulting in a region given by

\[
\mathcal{R}_{nf} = \{(\eta \mathcal{R}_{MAC}, (1 - \eta) \mathcal{R}_{BC})\},
\]  
(6.6)
where $\mathcal{R}_{MAC}$ and $\mathcal{R}_{BC}$ are the non-feedback capacity regions of AWGN multiple-access and broadcast channels as shown in Figure 6.2. In the terminology of two-way channels, the above choice of codebooks is also known as restricted codes [32], which do not send feedback to the receivers. The above rate region $\mathcal{R}_{nf}$ is the capacity achieved by current systems which decouple uplink and downlink in multiuser channels, either via time-division or frequency-division duplexing. Figure 6.4 shows the 3D rate region plot defined by $\mathcal{R}_f$.

![Figure 6.4: The 3-d rate region for the MAC + BC half-duplex two-way channel.](image)

The gain in the rate region is obtained by making the messages of the users correlated through feedback. Hence although the messages at the three users are independent, the channel inputs are not independent. It is easy to see that using the power constraints, the power distribution and the independence of the messages of the different users, we can show that

$$E[X_1X_2] = \sqrt{\alpha_1\bar{\alpha}_1\bar{\beta}_1\lambda_1 P_1 + \sqrt{\alpha_2\bar{\alpha}_2\bar{\beta}_2\lambda_2 P_2}},$$
where $\alpha_1 \tilde{b}_1$ and $\alpha_2 \tilde{b}_2$ are implicit functions of the feedback power.

We define the net symmetric sum rate gain as

$$G_s = \eta(R_{sum} - R_{sum}^{nf}) - (1 - \eta)(R_3^{nf} - R_3),$$

where $\gamma_{opt}$ is found by maximising $G_s$ over all the power distribution coefficients ($R_{sum} = R_1 + R_2$). Figure 6.5 shows the MAC sum rate ($R_{sum}$) for $R_1 = R_2$ and $N_1 = N_2$ as a function of broadcast rate $R_3$ with and without feedback. The objective is to optimize the fraction of power used for feedback to maximize the overall throughput of the network while keeping track of all the resources spent. For each $\eta$, we optimize over $\gamma$ and the power distribution variables give us different points in the region.

The gain increases as the total system power increases, especially the power available at User 3. Since the gain in capacity is logarithmic in power for Gaussian channels, higher power at the central receiver implies that it can use more power towards feedback without losing significant data rate. And more power for feedback implies that the feedback information for the MAC mode increases, which increases the amount of information the two white users cooperate on. Thus, there is a larger MAC sum rate gain compared to loss in BC rate due to sharing power with feedback.

### 6.3 Asymmetric Case

If we consider only channels with additive Gaussian noise, then the users can have asymmetric properties if they have either asymmetric channels or asymmetric power constraints. Consider a channel that has a multiplicative path-loss factor in addition to having a additive Gaussian noise (see Figure 6.6). The relationship between the channel input and output random variables change a little bit from what we saw for
Figure 6.5: MAC sum-rate versus BC common rate with and without feedback. User powers are $P_1 = P_2 = 50$, and gray node power $P = 50, 100, 1000$. $N = 1$ and $N_1 = N_2 = 1.5$.

Figure 6.6: AWGN with pathloss: Near-far topology

the Gaussian case as follows

$$Y_1 = d_1^{-\alpha/2} X_3 + Z_1,$$

$$Y_2 = d_2^{-\alpha/2} X_3 + Z_2,$$

$$Y_3 = d_1^{-\alpha/2} X_1 + d_2^{-\alpha/2} X_2 + Z_3.$$

Hence in the Gaussian rate equations, we need to replace $P_1$ by $d_1^{-\alpha} P_1$, $P_2$ by $d_2^{-\alpha} P_2$, $N_1$ by $N_1 d_1^\alpha$ and $N_2$ by $N_2 d_2^\alpha$. For the simulations, we consider $\alpha = 2$.

We fix the distance of User 2 from User 3 and keep increasing the $d_1$ to see the
Figure 6.7: Percentage gain in the rates of the strong user (a) and weak user (b).

effect of the near far case. Figure 6.7 (a) shows the gain obtained by the strong user and Figure 6.7 (b) shows how much the weak user gains from this cooperation. To compare our scheme with the noiseless Gaussian capacity, the Ozarow rate region [25] is shown in a dotted line.

It is evident that cooperation helps the weak user more than the strong user. This is due to the behaviour of Gaussian capacity at different SNRs. As shown in Figure 6.8, at low SNRs, the capacity grows linearly but at high SNRs, the capacity logarithmically. Therefore, for the weak user, any increase in effective SNR increases its rate linearly, whereas the increase in the strong user’s rates happen logarithmically.

The higher gains for the weak user can also be explained due to the unequal maximum rates the respective channels can support as shown in Figure 6.9. The amount of data each channel can support is represented by the width of the rectangles. The random variables \( U_{10}, U_{20}, \bar{W}_3 \) are as explained in Chapter 4. The black rectangle represents the independent message sent by User 3 (green), the blue rectangle represents the amount of data User 1 (red) can decode about User 2 (blue). As User 2’s link is stronger, it is able to decode more about the User 1 from the downlink channel and hence is able to cooperate more in the next block. Hence, this results
Figure 6.8: Behaviour of Gaussian capacity at high and low SNRs.

Figure 6.9: The strong user is able to decode the downlink data at a higher rate and hence cooperate more due to its better channel. This results in a better rate gain for the weaker user.
in a higher rate gain for the weak user. As we mentioned in the beginning of the section, asymmetry can also be brought about in asymmetric power constraints even if the channel is just Gaussian without any pathloss and the receivers are identical (so that the noise variances are identical in the uplink). A schematic is shown in Figure 6.3, where the size of the nodes reflect its power constraint.

6.4 Two-way Channel with One-way Data

In this section, we show why the two-way formulation is essential to build practical feedback systems. We make a distinction between the channel being two-way and data being two-way. We say that the channel is two-way if it can support messages to be sent in both directions. To demonstrate the claim, we will aim for a multiple access system with feedback from receiver but with no data in MC mode. Our objective will be to design a system which can outperform MAC capacity region without feedback with a system with feedback whose feedback resources are accounted for. To do the same, we will use $R_f$ with $R = 0$ and ask what power at User 3 will allow us to outperform a non-feedback system.

Figure 6.11 shows the answer to above objective. For $P_1 = P_2 = 6$, required feedback power at User 3 is of the order of $10^{10}$ to outperform non-feedback system! This exponentially large power is easy to justify as follows. When feedback channel is operated, it takes away temporal degrees of freedom from the forward channel, reducing rates in the pre-log factor. However, cooperative gain due to feedback
improves rates inside the log (in the SNR). Thus, for these cooperative gains to outperform loss due to feedback channel, the feedback power has to be increased significantly such that \((1 - \eta) \log(1 + \frac{P_f}{N_0}) \approx 1\). This immediately implies required feedback power is exponentially large; the following theorem proves it for the equal rate pair \((R_1 = R_2)\).

\[
P = 6
\]

![Figure 6.11: No feedback (inner bound), genie-aided perfect feedback (outer bound) and MAC with finite feedback (two inner regions) with different amount of feedback power.](image)

**Theorem 6.1 (Min feedback power)** The rate region of the MAC with feedback is strictly greater than a MAC without feedback for all receiver powers \(P_f\) such that \(P_f \geq P_f^{\text{min}}\), where

\[
P_f^{\text{min}} = \frac{(1 + 4\alpha \beta P)(1 + N_1)(1 + P)^{\frac{1}{2(1-\eta)}}}{\alpha \beta P(1 + \frac{2\alpha \beta P}{1+2\alpha \beta P})^{\frac{1}{2(1-\eta)}}}
\] (6.8)
CHAPTER 6. SPECIAL CASES

**Proof 6.1** For equal rate pair \( R_1 = R_2, \alpha_1 = \alpha_2 = \alpha, \beta_1 = \beta_2 = \beta, \lambda_1 = \lambda_2 = 0.5, P_1 = P_2 = P \) and additionally, we assume that \( N_1 = N_2 = N \). Denote \( r_1 = R_{10} = R_{20}, \ r_2 = R_{11} = R_{22} \) and \( r_1 + r_2 = R_{10} + R_{11} = R_{20} + R_{22} = R \). Also normalize \( N_0 = 1 \).

Under such conditions, \( P^* = 4\alpha P, P_f^* = 1 + 2\alpha\beta P \) and \( N^* = 1 + 2\alpha\beta P \).

By using the fact that \( \frac{1}{2}C(2x) < C(x) < 2C\left(\frac{x}{2}\right) \) the rate region inequalities reduce to the following:

\[
r_1 < \min \left\{ (1 - \eta)C \left( \frac{\alpha \beta PP_f}{(1 + 2\alpha \beta P)N_1} \right), \eta C(\alpha \beta P), \frac{\eta}{2}C \left( \frac{4\alpha P}{1 + 2\alpha \beta P} \right) + \frac{\eta}{2}C(2\alpha \beta P) \right\}
\]

(6.9)

\[
r_2 < \frac{\eta}{2}C \left( \frac{2\alpha \beta P}{1 + 2\alpha \beta P} \right)
\]

(6.10)

\[
R + r_1 < \eta C \left( \frac{4\alpha P + \alpha \beta P}{1 + 2\alpha \beta P} \right) + \eta C(2\alpha \beta P)
\]

(6.11)

\[
R + r_2 < \eta C \left( \frac{2\alpha P + 2\alpha \beta P}{1 + 2\alpha \beta P} \right) + \eta C(\alpha \beta P)
\]

(6.12)

\[
R < \frac{\eta}{2}C \left( \frac{4\alpha P + 2\alpha \beta P}{1 + 2\alpha \beta P} \right) + \frac{\eta}{2}C(2\alpha \beta P)
\]

(6.13)

Let us refer to the right hand side of the above equations as \( k_1, k_2, k_3, k_4 \) and \( k_5 \) respectively. Note that the feedback power is not present in any of the equations except the first one \( (k_1) \). And as the convex hull is found by optimizing over all possible values of the power distribution variables and the time sharing variable \( \eta \), if \( P_f \) is very small to have any effect on the rate region (i.e. there is no \( P_f \) in \( k_1 \)), then \( \eta_{opt} = 1 \) and the rate region comes out to be same as the MAC region without feedback. When \( P_f \) increases above a certain \( P_{fmin} \), but is still small enough so that \( k_1 \) is a function of \( P_f \), the total sum rate is above the MAC sum rate. The only way \( P_f \) defines the maximum sum rate is when the first min term \( (k_1) \) is a function of \( P_f \) and \( \min(k_1 + k_2, k_3) = k_1 + k_2 \). At \( P_f = P_{fmin} \), the sum rate with feedback just
equals the MAC sum rate. This is easy to calculate. The MAC sum rate is given as:
\[ \frac{1}{2} C(P) . \] Therefore

\[
(1 - \eta)C \left( \frac{\alpha \beta P P_{f}^{\min}}{(1 + 4\alpha \beta P)(1 + N_1)} \right) + \eta C \left( \frac{2\alpha \beta P}{1 + 2\alpha \beta P} \right) = \frac{C(P)}{2} \tag{6.14}
\]

\[ \Rightarrow P_{f}^{\min} = \frac{N_1(1 + 2\alpha \beta P)(1 + P)^{\frac{2}{\alpha \beta P(1 + 2\alpha \beta P)^{2/3}}}}{\alpha \beta P(1 + 2\alpha \beta P)^{2/3}} \tag{6.15} \]

### 6.5 Reduction to Carleial rate region

In the Carleial generalized feedback scheme, the radios are assumed to be full-duplex ($\eta = 1$) and hence can hear other user's transmissions. Hence there is no compression at the receiver ($X_3 = Y_3$) and there is no data in the downlink (There is no $W_3$ or $U_3$ i.e. $R_3 = 0$).

\[
R_{10} < \min[I(U_{10}; Y_2|U_{20}, U_{11}, U_{22}),
I(\bar{U}_{10}; Y_3|\bar{U}_{20}) + I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{11} < I(U_{11}; Y_3|\bar{U}_{10}, \bar{U}_{20}, U_{10}, U_{20}, U_{22})
\]

\[
R_{20} < \min[I(U_{20}; Y_1|U_{10}, U_{11}, U_{22}),
I(\bar{U}_{20}; Y_3|\bar{U}_{10}) + I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{22} < I(U_{22}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})
\]

\[
R_{10} + R_{11} < \eta[I(\bar{U}_{10}, U_{11}; Y_3|\bar{U}_{20}, U_{22}) + I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{10} + R_{20} < [I(\bar{U}_{10}, \bar{U}_{20}; Y_3|U_{11}, U_{22}) + I(U_{10}, U_{20}; Y_3|\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{10} + R_{22} < [I(\bar{U}_{10}, U_{22}; Y_3|U_{11}, \bar{U}_{20}) + I(U_{10}; Y_3|U_{20}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{11} + R_{20} < [I(U_{11}, \bar{U}_{20}; Y_3|\bar{U}_{10}, U_{22}) + I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{11} + R_{22} < I(U_{11}, U_{22}; Y_3|\bar{U}_{10}, \bar{U}_{20})
\]

\[
R_{20} + R_{22} < [I(\bar{U}_{20}, U_{22}; Y_3|\bar{U}_{10}, U_{11}) + I(U_{20}; Y_3|U_{10}, \bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]

\[
R_{10} + R_{11} + R_{20} < [I(\bar{U}_{10}, U_{11}, \bar{U}_{20}; Y_3|U_{22}) + I(U_{10}, U_{20}; Y_3|\bar{U}_{10}, U_{11}, \bar{U}_{20}, U_{22})]
\]
\[ R_{10} + R_{11} + R_{22} < [I(\tilde{U}_{10}, U_{11}, U_{22}; Y_3|\tilde{U}_{20}) + I(U_{10}; Y_3|U_{20}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22})] \]
\[ R_{10} + R_{20} + R_{22} < [I(\tilde{U}_{10}, \tilde{U}_{20}, U_{22}; Y_3|U_{11}) + I(U_{10}, U_{20}; Y_3|\tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22})] \]
\[ R_{11} + R_{20} + R_{22} < [I(U_{11}, \tilde{U}_{20}, U_{22}; Y_3|\tilde{U}_{10}) + I(U_{20}; Y_3|U_{10}, \tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22})] \]
\[ R_{10} + R_{11} + R_{20} + R_{22} < [I(\tilde{U}_{10}, U_{10}, U_{11}, \tilde{U}_{20}, U_{20}, U_{22}; Y_3). \]

which is indeed the Carleial rate region with the following substitutions: \( U_1 = U_{10}, U_2 = U_{20}, V_1 = \tilde{U}_{10}, V_2 = U_{22}, W_1 = \tilde{U}_{10} \) and \( W_2 = \tilde{U}_{20} \). We also use the fact that \([I(\tilde{U}_{10}, U_{10}, U_{11}, \tilde{U}_{20}, U_{20}, U_{22}; Y_3) + I(U_{10}, U_{20}; Y_3|\tilde{U}_{10}, U_{11}, \tilde{U}_{20}, U_{22})] = [I(\tilde{U}_{10}, U_{10}, U_{11}, \tilde{U}_{20}, U_{20}, U_{22}; Y_3) \) in the last rate equation above.

### 6.6 Reduction to Cover-Leung rate region

In this section, we will show that the Carleial rate region can be reduced to the Cover-Leung rate region when we make certain substitutions based on the assumptions of both the schemes. Since, in the Cover-Leung scheme, the feedback channel is noiseless, we put \( Y_1 = Y_2 = Y \). Also, this makes a part of the alphabet common for both the users. Hence we put \( W_1 = W_2 = W, U_1 = U_2 = U \) and \( R_{10} = R_{20} = R_0 \). Let us now, write down the Carleial rate region with the above substitutions

\[ R_{10} < I(U; Y|Q, U, V_2) \]
\[ R_{10} < I(U; Y|Q, U, V_1) \]
\[ R_{10} < I(U; Y|Q, U, V_1, V_2) + I(U; Y|Q, U) \]
\[ R_{10} < I(U; Y|Q, U, V_1, V_2) + I(U; Y|Q, U) \]
\[ R_{11} < I(V_1; Y|Q, U, V_2) \]
\[ R_{22} < I(V_2; Y|Q, U, V_1) \]
\[ 2R_{10} < I(U; Y|Q, U, V_1, V_2) + I(U; Y|Q) \]
\[ R_{10} + R_{11} < I(U, V_1; Y|Q, U, V_2) + I(U; Y|Q, U) \]
\[ R_{10} + R_{22} < I(U, V_2; Y|Q, U, V_1) + I(U; Y|Q, U) \]
\[ R_{10} + R_{11} < I(U, V_1; Y|Q, U, V_2) + I(U; Y|Q, U) \]
\[ R_{10} + R_{22} < I(U, V_2; Y|Q, U, V_1) + I(U; Y|Q, U) \]
\[ R_{11} + R_{22} < I(V_1, V_2; Y|Q, U) \]
\[ 2R_{10} + R_{11} < I(U, V_1; Y|Q, U, V_2) + I(U; Y|Q) \]
\[ 2R_{10} + R_{22} < I(U, V_2; Y|Q, U, V_1) + I(U; Y|Q) \]
\[ R_{10} + R_{11} + R_{22} < I(U, V_1, V_2; Y|Q, U) + I(U; Y|Q, U) \]
\[ 2R_{10} + R_{22} < I(U, V_1, V_2; Y|Q, U) + I(U; Y|Q, U) \]
\[ 2R_{10} + R_{11} + R_{22} < I(U, V_1, V_2; Y|Q). \]

First remove all the equations that involve trivial bounds. The equations that involve the downlink channel now will vanish as feedback is noiseless in Cover-Leung scheme. Also note that as there is no time sharing in the Cover-Leung scheme, \( Q \) takes a deterministic value with probability 1 due to which we can take it off the equation. The time sharing between the codebooks is built into the random variable \( U \). We also use the fact that their scheme uses deterministic encoding functions: \( X_1 = f_1(U, V_1) \) and \( X_2 = f_2(U, V_2) \). Finally, we substitute \( R_1 = R_0 + R_{11} \) and \( R_2 = R_0 + R_{22} \) to get

\[ R_{11} < I(V_1; Y|X_2) \]
\[ R_{22} < I(V_2; Y|X_1) \]
\[ 2R_{10} < I(U; Y) \]
\[ R_1 < I(X_1; Y|U, V_2) = I(X_1; Y|U, X_2) \]
\[ R_2 < I(X_2; Y|U, V_1) = I(X_2; Y|U, X_1) \]
\[ R_{11} + R_{22} < I(V_1, V_2; Y|U) \]
\[ R_{10} + R_{11} + R_{22} < I(X_1, X_2; Y|U) \]
\[ R_1 + R_2 < I(X_1, X_2; Y). \]

Note that the \(2R_{10} + R_{11}\) and \(2R_0 + R_{22}\) bounds were removed as they were larger than the addition of the singleton bounds. Also as \(I(V_1, V_2; Y|U) = I(V_1; Y|X_2) + I(V_2; Y|X_1)\) (since \(V_1\) and \(V_2\) are independent of each other), the \(R_{11} + R_{22}\) equation is redundant. The next step is to show that the following equations dominate the rest of the equations

\[ R_1 < I(X_1; Y|U, X_2) \quad (6.16) \]
\[ R_2 < I(X_2; Y|U, X_1) \quad (6.17) \]
\[ R_1 + R_2 < I(X_1, X_2; Y), \quad (6.18) \]

\[ 2R_0 < I(U; Y) \quad (6.19) \]
\[ R_{11} < I(V_1; Y|X_2) \quad (6.20) \]
\[ R_{22} < I(V_2; Y|X_1) \quad (6.21) \]
\[ R_0 + R_{11} + R_{22} < I(X_1, X_2; Y|U). \quad (6.22) \]

(These equations constrain the way \(R_1\) and \(R_2\) can be decomposed. As long as they have no effect on the bounds of \(R_1\), \(R_2\) and \(R_1 + R_2\), they do not affect on the final rate region.) It is easy to see that

\[ \frac{1}{2}\text{RHS of eqn}(6.19) + \text{RHS of eqn}(6.20) > \text{RHS of eqn}(6.16) \]

since \(I(U, V_1; Y|U, V_2) = I(V_1; Y|U, V_2) + I(U; Y|U, V_1, V_2) = I(V_1; Y|U, V_2)\)

and similarly,

\[ \frac{1}{2}\text{RHS of eqn}(6.19) + \text{RHS of eqn}(6.21) > \text{RHS of eqn}(6.17) \]

For the sum rate, note that
RHS of eqn(6.19) + RHS of eqn(6.20) + RHS of eqn(6.21) = RHS of eqn(6.18)

Finally we can see that

\( \frac{1}{2} \text{RHS of eqn(6.19)} + \text{RHS of eqn(6.22)} > \text{RHS of eqn(6.18)} \)

since, \( I(U, V_1, V_2; Y|U) = I(V_1, V_2; Y|U) = I(V_2; Y|U) + I(V_1; Y|U, V_2) \).
Chapter 7

Conclusions

We considered the problem of a multi-node system when there are data to be sent in both the directions, and showed that a joint design of uplink and downlink coding schemes expands the rate region obtained by independent code design. A two-way analysis accounts for all the resources spent in multi-user systems which use feedback. We give an inner and outer bound for the capacity region for our joint coding scheme. For the Gaussian channel, we show that our coding scheme can enable users to cooperate with each other in hidden node topologies as well. We also show an interesting near-far situation where cooperation induced through feedback helps the weak user significantly. If there are no data in one direction, forcing it to use feedback to induce a two-way channel requires exponentially more resources as compared to a time-sharing disjoint design without using feedback. The major finding of this thesis is that if the communication system has data in both directions, then designing a joint coding scheme performs better and the right way to analyze it uses a two-way channel model.

Room for improvement still remains in certain parts of the scheme. We used messages spanning two code blocks in our encoding structure, which can be generalized to more blocks and can bring the rates closer to the capacity region, although
it will increase the complexity of the decoding scheme. A similar but fundamentally different approach to general networks has been considered in [18]. Also, as in similar multi-user superposition coding schemes, the cardinality of the auxiliary random variables is our scheme is not bounded which hinders its computability.
Bibliography


