

**CAPACITY REGION OF TWO INTERFERING MIMO MULTIPLE  
ACCESS CHANNELS WITH STRONG INTERFERENCE**

**A Thesis  
Submitted to the Graduate Faculty  
of the  
North Dakota State University  
of Agriculture and Applied Science**

**By**

**Noopur Modi**

**In Partial Fulfillment of the Requirements  
for the Degree of  
MASTER OF SCIENCE**

**Major Department:  
Electrical and Computer Engineering**

**June 2015**

**Fargo, North Dakota**

North Dakota State University  
Graduate School

---

**Title**

Capacity Region of Two Interfering MIMO Multiple Access Channels with Strong  
Interference

---

**By**

Noopur Modi

---

The Supervisory Committee certifies that this *disquisition* complies with North Dakota State  
University's regulations and meets the accepted standards for the degree of

**MASTER OF SCIENCE**

SUPERVISORY COMMITTEE:

Dr. Sanjay Karmakar

Chair

Dr. Rajesh Kavasseri

Dr. Sudrashan Srinivasan

Dr. Indranil Sengupta

Approved:

06/29/2015

Date

Dr. Scott C. Smith

Department Chair

## ABSTRACT

We consider a communication network, where two mutually interfering 2-user MIMO Multiple Access Channels (MAC) operate simultaneously via the same time-frequency space, and characterize the capacity region of this network when the channel matrices satisfy a strong interference condition. This interfering MAC (IMAC) with aforementioned channel matrices is called strongly ordered IMAC in this work. We characterize the capacity region first using the genie aided approach to find out several constraints that must be satisfied by any achievable rate tuple. Then we show that independent Gaussian coding at each transmitter and joint decoding of the messages at the receivers can achieve all the rate pairs that satisfy all the aforementioned constraints. In an IMAC, there are two types of tradeoffs between rates of communication; it is between the rates of users from different MACs and among rates of users belonging to the same MAC. Result shows two tradeoffs are homogeneous.

## ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to the College of Engineering, NDSU for letting me fulfill my dream of being a student here.

I would like to express my sincere gratitude to my advisor Prof. Sanjay Karmakar for the continuous support of my Master's study and research, for his patience, motivation, enthusiasm, and immense knowledge.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Rajesh Kavessari, Dr. Sudrashan Srinivasan, Dr. Indranil Sengupta, for being part of my thesis committee.

I would also like to thank my family for the support they provided me through my entire life.

## DEDICATION

To my parents and family in Fargo.

## TABLE OF CONTENTS

ABSTRACT .....	iii
ACKNOWLEDGMENTS .....	iv
DEDICATION.....	v
LIST OF FIGURES .....	viii
CHAPTER 1. INTRODUCTION .....	1
1.1. Entropy .....	4
1.2. Joint and Conditional Entropy .....	6
1.3. Mutual Information.....	8
1.4. Differential Entropy .....	9
1.4.1. Jensen's Inequality.....	11
1.4.2. Data Processing Inequality.....	12
1.5. Channel Capacity .....	12
1.5.1. Discrete Memoryless Channel Capacity .....	13
1.5.2. Gaussian Channel Capacity .....	17
1.5.3. Multiple Access Channel (MAC).....	19
CHAPTER 2. INTERFERENCE CHANNEL .....	23
2.1. Interfering Channels .....	23
2.2. Interfering MAC channels .....	29
2.3. Problem Statement .....	31
CHAPTER 3. MIMO MULTIPLE ACCESS CHANNEL .....	34
3.1. Channel Model.....	34

3.2. Capacity Region .....	51
3.2.1. Converse: Outer bounds to the achievable rate tuples.....	52
3.2.2. Achievability .....	62
CHAPTER 4. PRACTICAL APPLICATIONS OF OUR RESULTS .....	65
CHAPTER 5. CONCLUSION .....	69
REFERENCES .....	70
APPENDIX A.....	73
APPENDIX B.....	76

## LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	A discrete channel . . . . .	2
2	The binary-symmetric channel . . . . .	3
3	Additive white Gaussian noise channel with power constraint . . . . .	4
4	Entropy, conditional entropy, and mutual information . . . . .	9
5	Binary noiseless channel. . . . .	14
6	Multiple-access channel . . . . .	20
7	Achievable region of multiple-access channel for a fixed input distribution	21
8	Cellular structure for 2-user interference channel . . . . .	25
9	The $(M_1, N_1, M_2, N_2)$ MIMO IC . . . . .	27
10	Cellular structure for two mutually interfering MIMO MACs . . . . .	30
11	Cellular structure for two mutually interfering MIMO MACs with 4 users	31
12	Channel Model for IMAC with 4 users . . . . .	32
13	The IMAC with 4 MIMO users . . . . .	35
14	The comparison of graphs by two methods in Cell-1 . . . . .	68
15	The comparison of graphs by two methods in Cell-2 . . . . .	68



## CHAPTER 1. INTRODUCTION

The goal of any communication system is to transmit information from an information source to a destination via a communication channel. A communication engineer usually has very little control over these three components. The engineer's role is to design transmitters and receivers that send the source output over the channel to the destination with high fidelity. In this chapter, we study the other important component of a communication system; i.e., the communication channel. We also introduce the concept of coding for protection of messages against channel errors.

A communication channel is any medium over which information can be transmitted or in which information can be stored. Coaxial cable, ionospheric propagation, free space, fiber optic cables, and magnetic and optical disks are examples of communication channels. What is common among all these is that they accept signals at their inputs and deliver signals at their outputs at a later time (storage case) or at another location (transmission case). Therefore, each communication channel is characterized by a relation between its input and output. In this sense a communication channel is a *system*.

There are many factors that cause the output of a communication channel to be different from its input. Among these factors are attenuation, nonlinearities, bandwidth limitations, multipath propagation, and noise. All these factors contribute to a usually complex input/output relation in a communication channel. Due to the presence of fading and noise, the input/output relation in a communication channel is, generally, a stochastic relation.

Channels encountered in practice are generally waveform channels that accept (continuous-time) waveforms as their inputs and produce waveforms as their outputs. Because the bandwidth of any practical channel is limited, by using the sampling theorem, a waveform channel becomes equivalent to a discrete-time channel. In a

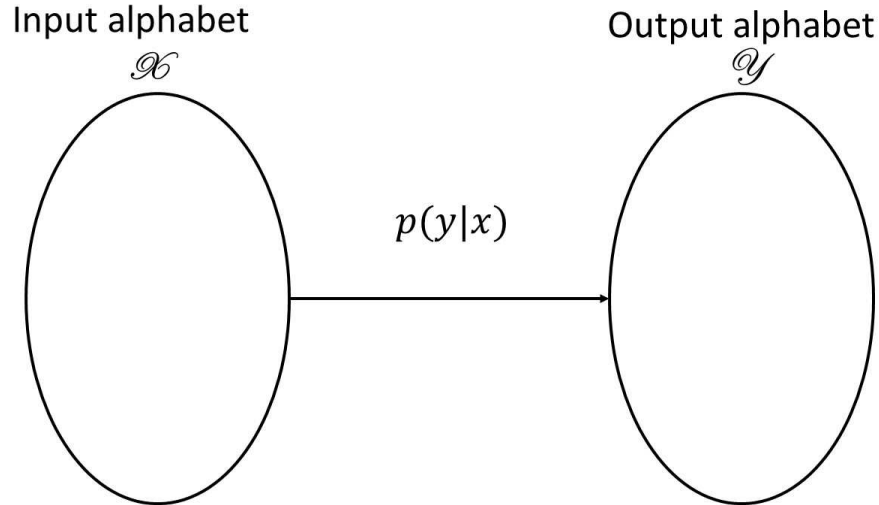


Figure 1. A discrete channel

discrete-time channel both input and output are discrete-time signals.

In a discrete-time channel, if the values that the input and output variables can take are finite, or countably infinite, the channel is called a *discrete channel*. An example of a discrete channel is a binary-input binary-output channel. In general, a discrete channel is defined by  $\mathcal{X}$ , the input alphabet,  $\mathcal{Y}$ , the output alphabet, and  $p(y|x)$  the conditional PMF of the output sequence given the input sequence. A schematic representation of a discrete channel is given in Figure 1. In general, the output  $y_i$  does not only depend on the input at the same time  $x_i$  but also on the previous inputs, or even previous and future inputs (in storage channels). Therefore, a channel can have *memory*. However, if a discrete channel does not have memory, it is called a *discrete-memoryless channel*, and for such a channel, for any  $y \in \mathcal{Y}^n$  and  $x \in \mathcal{X}^n$ , we have

$$p(y|x) = \prod_{i=1}^n p(y_i|x_i) \quad (1.1)$$

A special case of a discrete-memoryless channel is the *binary-symmetric channel*. Figure 2 shows a binary-symmetric channel. In a binary-symmetric channel,  $\epsilon = P(0|1) = P(1|0)$  is called the *crossover probability*.

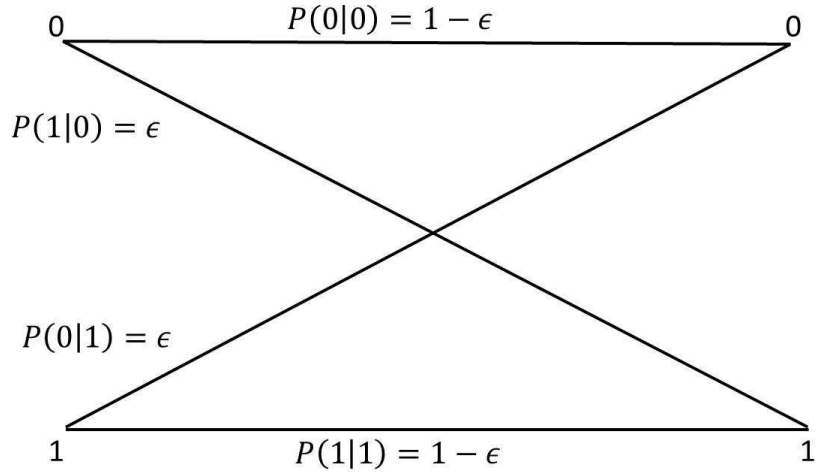


Figure 2. The binary-symmetric channel

The most important continuous alphabet channel is the discrete-time additive white Gaussian noise channel with an input power constraint. In this channel both  $\mathcal{X}$  and  $\mathcal{Y}$  are the set of real numbers, and the input-output relation is given by

$$Y = X + Z \tag{1.2}$$

where  $Z$  denotes the channel noise, which is assumed to be Gaussian with mean equal to 0 and variance equal to  $P_N$ . It is further assumed that inputs to this channel satisfy some power constraint. For example, for large  $n$ , input blocks of length  $n$  satisfy

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \leq P \tag{1.3}$$

where  $P$  is some fixed power constraint. The channel model is shown in Figure 3.

Of course, the main objective when transmitting information over any communication channel is *reliability*, which is measured by the probability of correct reception at the receiver. For that we need to derive the *Channel Capacity*, for that we need to understand other important fundamentals which we will see in next sections. In

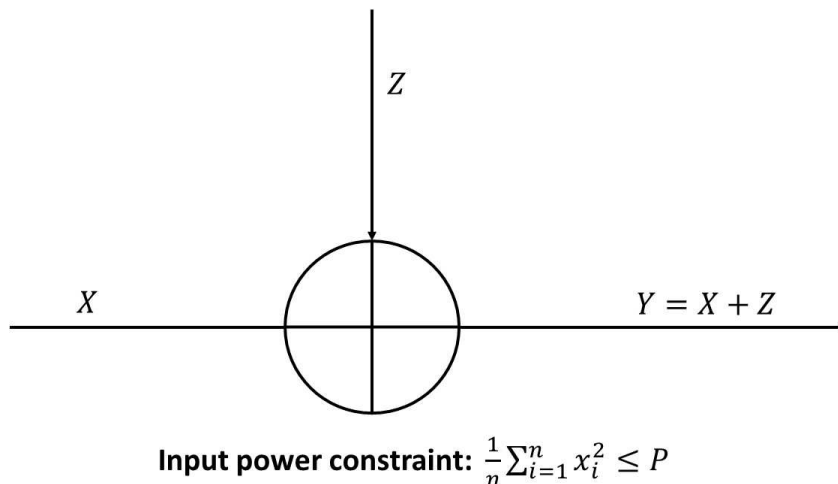


Figure 3. Additive white Gaussian noise channel with power constraint

section 1.5 we will get back to *channel capacity*.

### 1.1. Entropy

The development of the idea of entropy of random variables and processes by Claude Shannon provided the beginning of information theory and of the modern age of ergodic theory. We shall see that entropy and related information measures provide useful descriptions of the long term behavior of random processes and that this behavior is a key factor in developing the coding theorems of information theory.

Suppose now that we have  $n$  symbols  $\{a_1, a_2, \dots, a_n\}$ , and some source is providing us with a stream of these symbols. Suppose further that the source emits the symbols with probabilities  $\{p_1, p_2, \dots, p_n\}$ , respectively. For now, we also assume that the symbols are emitted independently (successive symbols do not depend in any way on past symbols).

The average amount of *information* we get from each symbol can be achieved from weighted average. If we observe the symbol  $a_i$ , we will be getting  $\log\left(\frac{1}{p_i}\right)$  *information* from that particular observation. In a long run (say  $N$ ) of observations, we will see (approximately)  $N \times p_i$  occurrences of symbol  $a_i$ . Thus, in the  $N$  (independent)

observations, we will get total information  $I$  of

$$I = \sum_{i=1}^n (N \times p_i) \times \log \left( \frac{1}{p_i} \right) \quad (1.4)$$

But then, the average information we get per symbol observed will be,

$$\frac{I}{N} = \sum_{i=1}^n p_i \times \log \left( \frac{1}{p_i} \right) \quad (1.5)$$

Note that  $\lim_{x \rightarrow 0} x \times \log \left( \frac{1}{x} \right) = 0$ , so we can, for our purposes, define  $p_i \times \log \left( \frac{1}{p_i} \right)$  to be 0 when  $p_i = 0$ .

This brings us to a fundamental definition. This definition is essentially due to Shannon in 1948, in the seminal papers in the field of information theory. As we have observed, we have defined information strictly in terms of the probabilities of events.

**Definition 1.** *The entropy of random variable  $X$  is a function of its PMF, if it is a discrete RV it is defined by:*

$$H(X) = \sum_{i=1}^N p_i \log \left( \frac{1}{p_i} \right). \quad (1.6)$$

*Another worth while way to think about this is in terms of expected value.*

$$H(X) = -E[\log P(x)] \quad (1.7)$$

In other words, the *entropy* of a probability distribution is just the expected value of the information of the distribution. Here, we use the logarithm to the base 2, which is well adapted to digital communication, and the entropy is then expressed in **bits**.

Intuitively, the entropy gives a measure of the uncertainty of the random variable. It is sometimes called the missing information: the larger the entropy, the less a priori information one has on the value of the random variable. This measure is, roughly speaking, the logarithm of the number of typical values that the variable can take. For example, A fair coin has two values with equal probability. Its entropy is 1 bit. A fair dice with  $M$  faces has entropy  $\log M$ .

## 1.2. Joint and Conditional Entropy

When dealing with two or more random variables, we can introduce joint and conditional entropies in exactly the same way that joint and conditional probabilities are introduced. These concepts are especially important when dealing with sources with memory.

**Definition 2.** *The joint entropy of two discrete random variables  $(X, Y)$  is defined by,*

$$H(X, Y) = - \sum_{x,y} p(x, y) \log p(x, y). \quad (1.8)$$

*For the case of  $n$  random variables  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ , we have,*

$$H(\mathbf{X}) = - \sum_{x_1, x_2, \dots, x_n} p(x_1, x_2, \dots, x_n) \log p(x_1, x_2, \dots, x_n). \quad (1.9)$$

As seen, the joint entropy is simply the entropy of a vector-valued random variable.

The conditional entropy of the random variable  $X$ , given the random variable  $Y$ , can be defined by noting that if  $Y = y$ , then the PMF of the random variables  $X$  will be  $p(x|y)$ , and the corresponding entropy is  $H(X|Y = y) = - \sum_x p(x|y) \log p(x|y)$ , which is intuitively the amount of uncertainty in  $X$  when one knows  $Y = y$ . The weighted average of the above quantities over all  $y$  is the uncertainty in  $X$  when  $Y$  is known. This quantity is known as the conditional entropy and defined as follows:

**Definition 3.** The conditional entropy of the random variable  $X$  given the random variable  $Y$  is defined by,

$$H(X|Y) = - \sum_{x,y} p(x,y) \log p(x|y) \quad (1.10)$$

In general, we have

$$H(X_n|X_1, \dots, X_{n-1}) = - \sum_{x_1, \dots, x_n} p(x_1, \dots, x_n) \log p(x_n|x_1, \dots, x_{n-1}) \quad (1.11)$$

Using chain rule for PMFs,  $p(x,y) = p(y)p(x|y)$ , we can show that  $H(X,Y) = H(Y) + H(X|Y)$ . Generalizing this result to the case of  $n$  random variables to show the following chain rule for entropies,

$$H(\mathbf{X}) = H(X_1) + H(X_2|X_1) + \dots + H(X_n|X_1, X_2, \dots, X_{n-1}) \quad (1.12)$$

If the random variable  $X_n$  denotes the output of a discrete (not necessarily memoryless) source at time  $n$ , then  $H(X_2|X_1)$  denotes the fresh information provided by source output  $X_2$  to someone who already knows the source output  $X_1$ . In the same way,  $H(X_n|X_1, X_2, \dots, X_{n-1})$  denotes the fresh information in  $X_n$  for an observer who has observed the sequence  $(X_1, X_2, \dots, X_{n-1})$ . The limit of the above conditional entropy as  $n$  tends to infinity is known as the *entropy rate* of the random process.

**Definition 4.** The entropy rate of a stationary discrete-time random process is defined by

$$H = \lim_{n \rightarrow \infty} H(X_n|X_1, X_2, \dots, X_{n-1}) \quad (1.13)$$

Stationarity ensures the existence of the limit, and it can be proved that an alternative

definition of the entropy rate for sources with memory is given by

$$H = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n) \quad (1.14)$$

Entropy rate plays the role of entropy for sources with memory. It is basically a measure of the uncertainty per output symbol of the source.

### 1.3. Mutual Information

For discrete random variables,  $H(X|Y)$  denotes the entropy (or uncertainty) of the random variable  $X$  after random variable  $Y$  is known. Therefore, if the starting entropy of the random variable  $X$  is  $H(X)$ , then  $H(X) - H(X|Y)$  denotes the amount of uncertainty of  $X$  that has been removed by revealing random variable  $Y$ . In other words,  $H(X) - H(X|Y)$  is the amount of information provided by the random variable  $Y$  about random variable  $X$ . This quantity plays an important role in both source and channel coding and is called the mutual information between two random variables.

**Definition 5.** *The mutual information between two discrete random variables  $X$  and  $Y$  is denoted by  $I(X; Y)$  and defined by*

$$I(X; Y) = H(X) - H(X|Y). \quad (1.15)$$

Mutual information has certain properties that are explored in problems and summarized here.

1.  $I(X; Y) \geq 0$  with equality if and only if  $X$  and  $Y$  are independent.
2.  $I(X; Y) \leq \min(H(X), H(Y))$ .
3.  $I(X; Y) = \sum_{x,y} p(x, y) \log \frac{p(x,y)}{p(x)p(y)}$ .
4.  $I(X; Y) = H(X) + H(Y) - H(X, Y)$ .



5.  $I(X;Y|Z)$  is the conditional mutual information and defined by  $I(X;Y|Z) = H(X|Z) - H(X|Y, Z)$ .
6.  $I(X;Y|Z) = \sum_z p(z)I(X;Y|Z = z)$ .
7.  $I(XY; Z) = I(X; Z) + I(Y; Z|X)$ . This is the chain rule for mutual information.
8. In general,  $I(X_1, \dots, X_n; Y) = I(X_1; Y) + I(X_2; Y|X_1) + \dots + I(X_n; Y|X_1, \dots, X_{n-1})$ .

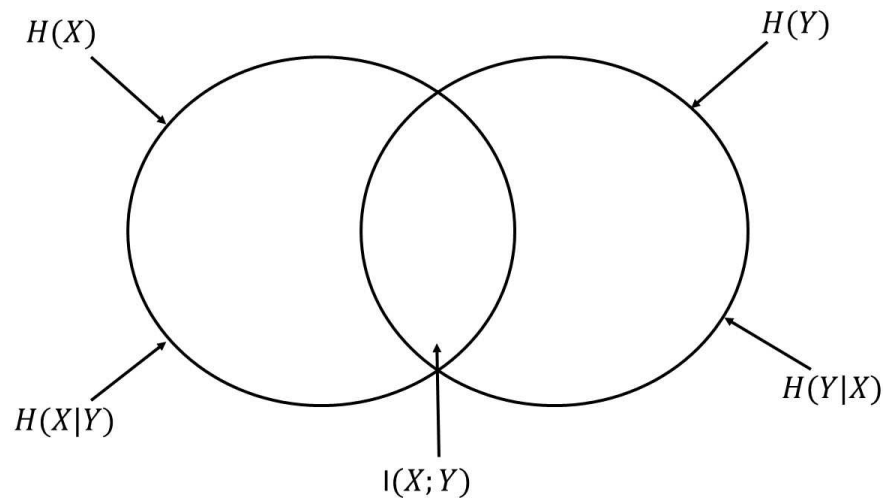


Figure 4. Entropy, conditional entropy, and mutual information

Figure 4 represents the relation among entropy, conditional entropy and mutual information quantities.

#### 1.4. Differential Entropy

So far we have defined entropy and mutual information for discrete sources. If we are dealing with a discrete-time, continuous-alphabet source whose outputs are real numbers, nothing exists that has the intuitive meaning of entropy. In the continuous case, another quantity that resembles entropy, called differential entropy, is defined. However, it does not have the intuitive meaning of entropy. In fact, to reconstruct the output of a continuous source reliably, an infinite number of bits/source output are

required because any output of the source is a real number and the binary expansion of a real number has infinitely many bits.

**Definition 6.** *The differential entropy of a continuous random variable  $X$  with PDF  $f_X(x)$  is denoted by  $h(X)$  and defined by*

$$h(X) = - \int_{-\infty}^{\infty} f_X(x) \log f_X(x) dx. \quad (1.16)$$

where  $0 \log 0 = 0$ .

The differential entropy of a Gaussian random variable with mean zero and variance  $\sigma^2$  can be derived as follows:

The PDF is  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$ . Therefore, using natural logarithms we find the differential entropy in nats,

$$h(X) = - \int_{-\infty}^{\infty} \ln \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right) f(x) dx - \int_{-\infty}^{\infty} \ln \left( e^{-\frac{x^2}{2\sigma^2}} \right) f(x) dx \quad (1.17)$$

$$= \ln \left( \sqrt{2\pi\sigma^2} \right) + \frac{\sigma^2}{2\sigma^2} \quad (1.18)$$

$$= \frac{1}{2} \ln(2\pi e\sigma^2) \text{ nats} \quad (1.19)$$

where we have used  $\int_{-\infty}^{\infty} f(x) dx = 1$  and  $\int_{-\infty}^{\infty} x^2 f(x) dx = \sigma^2$ . Changing the base of logarithms to 2, we have

$$h(X) = \frac{1}{2} \log(2\pi e\sigma^2) \text{ bits} \quad (1.20)$$

Extensions of the definition of differential entropy to joint differential variables and conditional differential entropy are straightforward. For two random variables,

we have

$$h(X, Y) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log f(x, y) dx dy \quad (1.21)$$

and

$$h(X|Y) = h(X, Y) - h(Y) \quad (1.22)$$

The mutual information between two continuous random variables  $X$  and  $Y$  is defined similarly to the discrete case as

$$I(X; Y) = h(Y) - h(Y|X) = h(X) - h(X|Y) \quad (1.23)$$

Although differential entropy does not have the intuitive interpretation of discrete entropy, it can be shown that the mutual information of continuous random variables has basically the same interpretation as the mutual information of discrete random variables; i.e., the information provided by one random variable about the other random variable.

#### 1.4.1. Jensen's Inequality

This inequality is one of the most widely used in mathematics and one that underlies many of the basic results in information theory. It is called Jensen's Inequality.

**Theorem 1.** *If  $f$  is a convex function and  $X$  is a random variable,*

$$E[f(X)] \geq f(E[X]) \quad (1.24)$$

*Moreover, if  $f$  is strictly convex, the equality in equation (1.24) implies that  $X = E[X]$  with probability 1 (i.e.,  $X$  is a constant).*

The inequality can be extended to continuous distributions by continuity arguments. Using these results we can prove some of the properties of entropy and

relative entropy. The following theorem is of fundamental importance.

### 1.4.2. Data Processing Inequality

The data-processing inequality can be used to show that no clever manipulation of the data can improve the inferences that can be made from the data.

**Definition 7.** *Random variables  $X, Y, Z$  are said to form a Markov chain in that order (denoted by  $X \rightarrow Y \rightarrow Z$ ) if the conditional distribution of  $Z$  depends only on  $Y$  and is conditionally independent of  $X$ . Specifically,  $X, Y$ , and  $Z$  form a Markov chain  $X \rightarrow Y \rightarrow Z$  if the joint probability mass function can be written as,*

$$p(x, y, z) = p(x)p(y|x)p(z|y) \tag{1.25}$$

We can now prove an important and useful theorem demonstrating that no processing of  $Y$ , deterministic or random, can increase the information that  $Y$  contains about  $X$ .

**Theorem 2.** *If  $X \rightarrow Y \rightarrow Z$ , then  $I(X; Y) \geq I(X; Z)$ .*

We have equality if and only if  $I(X; Y|Z) = 0$  (i.e.,  $X \rightarrow Z \rightarrow Y$  forms a Markov chain). Similarly, one can prove that  $I(Y; Z) \geq I(X; Z)$ .

The supporting results to this theorem are as follows:

1. In particular, if  $Z = g(Y)$ , we have  $I(X; Y) \geq I(X; g(Y))$ .
2. If  $X \rightarrow Y \rightarrow Z$ , then  $I(X; Y|Z) \leq I(X; Y)$ .

## 1.5. Channel Capacity

In this chapter we characterize the channel capacity. The characterization of the channel capacity (the logarithm of the number of distinguishable signals) as the maximum mutual information is the central and most famous success of information theory.

### 1.5.1. Discrete Memoryless Channel Capacity

We define a discrete channel to be a system consisting of an input alphabet  $X$  and output alphabet  $Y$  and a probability transition matrix  $p(y|x)$  that expresses the probability of observing the output symbol  $y$  given that we send the symbol  $x$ . The channel is said to be memoryless if the probability distribution of the output depends only on the input at that time and is conditionally independent of previous channel inputs or outputs. We will first understand the channel capacity of a discrete memoryless channel:

**Definition 8.** *We define the information channel capacity of a discrete memoryless channel as*

$$C = \max_{p(x)} I(X; Y) \quad (1.26)$$

where the maximum is taken over all possible input distributions  $p(x)$ .

This definition follows for continuous channel with only difference is that now the maximization is carried out over the all possible input PDFs instead of PMFs. Next, we consider few examples of channel capacity for some simplest channels.

Now let us see the example of a Noiseless Binary channel. Suppose that we have a channel where the binary input is reproduced exactly at the output. In this case, any transmitted bit is received without error. Intuitively, one error-free bit can be transmitted per use of the channel, and the capacity should be 1 bit.

We have the rate upper-bound

$$R \leq \max I(X; Y) \quad (1.27)$$

$$= \max [H(X) - H(X|Y)] \quad (1.28)$$

$$= \max [H(X)] \quad (1.29)$$

$$\leq 1 \quad \text{bit/channel use,} \quad (1.30)$$

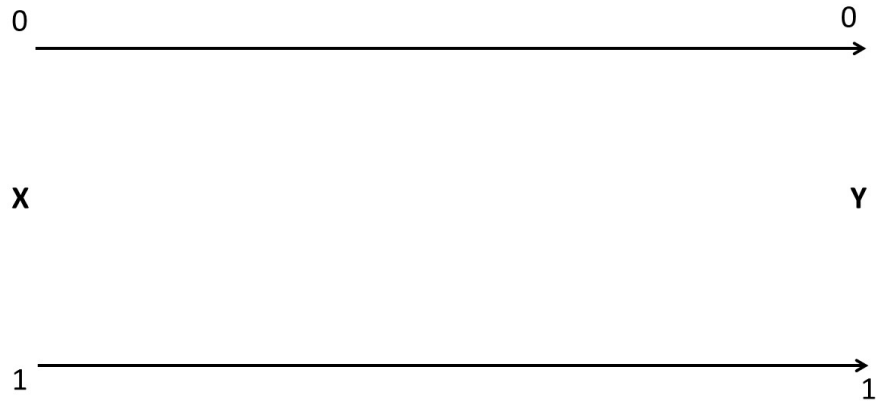


Figure 5. Binary noiseless channel.

because for a binary RV  $X$ ,  $H(X) \leq 1$ .

Now, let us consider the achievable rate for a particular input  $X^*$  with  $p_{X^*}(x) = (\frac{1}{2}, \frac{1}{2})$ . The achieved rate  $R^*$  for that input is

$$R^* = I(X^*; Y) \tag{1.31}$$

$$= H(Y) - H(Y|X^*) \tag{1.32}$$

$$= H(X^*) \tag{1.33}$$

Since  $H(X^*) = 1$  bit, the achievable rate

$$R^* = H(X^*) = 1 \text{ bit/channel use.} \tag{1.34}$$

The upper-bound in (1.30) matches with the achievable rate (1.34). Hence, we can say the capacity of the noiseless binary channel is 1 bit/per channel use. Now consider a *Binary Symmetric Channel* as shown in Figure 2. This is a binary channel in which the input symbols are complemented with probability  $\epsilon$ . This is the simplest model of a channel with errors, yet it captures most of the complexity of the general

problem.

When an error occurs, a 0 is received as a 1, and vice versa. The bits received do not reveal where the errors have occurred. In a sense, all the bits received are unreliable. Later we show that we can still use such a communication channel to send information at a nonzero rate with an arbitrarily small probability of error.

We bound the mutual information by

$$I(X : Y) = H(Y) - H(Y|X) \tag{1.35}$$

$$= H(Y) - \sum p(x)H(Y|X = x) \tag{1.36}$$

$$= H(Y) - \sum p(x)H(p) \tag{1.37}$$

$$= H(Y) - H(p) \tag{1.38}$$

$$\leq 1 - H(p) \tag{1.39}$$

where the last inequality follows because  $Y$  is a binary random variable. Equality is achieved when the input distribution is uniform. Hence, the information capacity of a binary symmetric channel with parameter  $p$  is

$$C = 1 - H(p) \text{ bits} \tag{1.40}$$

If we have a message  $W$ , drawn from the index set  $\{1, 2, \dots, M\}$ , results in the signal  $X^n(W)$ , which is received as a random sequence  $Y^n$ . The receiver then guesses the index  $W$  by an appropriate decoding rule  $\hat{W} = g(Y^n)$ . The receiver makes an error if  $\hat{W}$  is not the same as the index  $W$  that was transmitted. We now define these ideas formally.

**Definition 9.** A discrete channel, denoted by  $(X, p(y|x), Y)$ , consists of two finite sets  $\mathcal{X}$  and  $\mathcal{Y}$  and a collection of probability mass functions  $p(y|x)$ , one for each  $x \in \mathcal{X}$ , such that for every  $x$  and  $y$ ,  $p(y|x) \geq 0$ , and for every  $x$ ,  $\sum_y p(y|x) = 1$ , with the interpretation that  $X$  is the input and  $Y$  is the output of the channel.

**Definition 10.** The  $n$ th extension of the discrete memoryless channel (DMC) is the channel  $(\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n)$ , where,

$$p(y_k|x^k, y^{k-1}) = p(y_k|x_k), \quad k = 1, 2, \dots, n. \quad (1.41)$$

When we refer to the discrete memoryless channel, we mean the discrete memoryless channel without feedback unless we state explicitly otherwise.

Now we will define the *conditional probability of error*.

**Definition 11.** Let,

$$\lambda_i = Pr(g(Y^n) \neq i | X^n = x^n(i)) = \sum_{y^n} p(y^n|x^n(i)) I(g(y^n) \neq i) \quad (1.42)$$

be the conditional probability of error given that index  $i$  was sent, where  $I(\cdot)$  is the indicator function.

Now we will define the *rate* and the *achievable rate* and then finally the *capacity*.

**Definition 12.** The rate  $R$  of an  $(M, n)$  code is,

$$R = \frac{\log M}{n} \text{ bits per transmission} \quad (1.43)$$

**Definition 13.** A rate  $R$  is said to be *achievable* if there exists a sequence of  $(\lceil 2^{nR} \rceil, n)$  codes such that the maximal probability of error  $\lambda^{(n)}$  tends to 0 as  $n \rightarrow \infty$ .



Later we write  $(2^{nR}, n)$  codes to mean  $(\lceil 2^{nR} \rceil, n)$  codes. This will simplify the notation.

**Definition 14.** *The capacity of a channel is the supremum of all achievable rates.*

Thus, rates less than capacity yield arbitrarily small probability of error for sufficiently large block lengths.

Next we will see the Gaussian Channel Capacity. The results of Gaussian Channel Capacity is very useful for this thesis work.

### 1.5.2. Gaussian Channel Capacity

The most important continuous alphabet channel is the Gaussian channel depicted in Figure 3. This is a time-discrete channel with output  $Y_i$  at time  $i$ , where  $Y_i$  is the sum of the input  $X_i$  and the noise  $Z_i$ . The noise  $Z_i$  is drawn i.i.d. from a Gaussian distribution with variance  $N$ . Thus,

$$Y_i = X_i + Z_i, \quad Z_i \sim \mathcal{N}(0, N) \quad (1.44)$$

The noise  $Z_i$  is assumed to be independent of the signal  $X_i$ . This channel is a model for some common communication channels, such as wired and wireless telephone channels and satellite links.

The most common limitation on the input is an energy or power constraint. We assume an average power constraint. For any codeword  $(x_1, x_2, \dots, x_n)$  transmitted over the channel, we require that,

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \leq P \quad (1.45)$$

We now define the (information) capacity of the channel as the maximum of the mutual information between the input and output over all distributions on the input that satisfy the power constraint.

The information capacity of the Gaussian channel with power constraint  $P$  is

$$I(X;Y) = h(Y) - h(Y|X) \quad (1.46)$$

$$= h(Y) - h(X + Z|X) \quad (1.47)$$

$$= h(Y) - h(Z|X) \quad (1.48)$$

$$= h(Y) - h(Z) \quad (1.49)$$

$$\leq \frac{1}{2} \log 2\pi e(P + N) - \frac{1}{2} \log 2\pi eN \quad (1.50)$$

$$= \frac{1}{2} \log \left( 1 + \frac{P}{N} \right) \quad (1.51)$$

since  $E[Y^2] = P + N$  and the Gaussian is the maximum entropy distribution for a given variance. So,

$$C = \frac{1}{2} \log \left( 1 + \frac{P}{N} \right), \text{ bits per channel use} \quad (1.52)$$

**Definition 15.** An  $(M, n)$  code for the Gaussian channel with power constraint  $P$  consists of the following:

1. An index set  $\{1, 2, \dots, M\}$ .
2. An encoding function  $x : \{1, 2, \dots, M\} \rightarrow \mathcal{X}^n$ , which maps an input index into a sequence that is  $n$  elements long,  $x^n(1), x^n(2), \dots, x^n(M)$ , such that the average power constraint is satisfied:

$$\sum_{i=1}^n (x_i^n(w))^2 \leq nP \quad (1.53)$$

for  $w = 1, 2, \dots, M$ .

3. A decoding function  $g : \mathcal{Y}^n \rightarrow \{1, 2, \dots, M\}$ .

**Definition 16.** A rate  $R$  is said to be achievable for a Gaussian channel with a power constraint  $P$  if there exists a sequence of  $(2^{nR}, n)$  codes with codewords satisfying the power constraint such that the maximal probability of error  $\lambda^{(n)} \rightarrow 0$ . The **capacity** of the channel is the supremum of the achievable rates.

**Theorem 3.** The capacity of a Gaussian channel with power constraint  $P$  and noise variance  $N$  is

$$C = \frac{1}{2} \log \left( 1 + \frac{P}{N} \right), \text{ bits per transmission} \quad (1.54)$$

### 1.5.3. Multiple Access Channel (MAC)

Multiple Access Channel is the channel in which two (or more) senders send information to a common receiver. The channel is illustrated in Figure 6. Common example of this channel is a satellite receiver with many independent ground stations, or a set of cell phones communicating with a base station. We see that the senders must contend not only with the receiver noise but with interference from each other as well.

**Definition 17.** A discrete memoryless multiple-access channel consists of three alphabets,  $\mathcal{X}_1$ ,  $\mathcal{X}_2$ , and  $\mathcal{Y}$ , and a probability transition matrix  $p(y|x_1, x_2)$ .

**Definition 18.** A  $((2^{nR_1}, 2^{nR_2}), n)$  code for the multiple-access channel consists of two sets of integers  $\mathcal{W}_1 = \{1, 2, \dots, 2^{nR_1}\}$  and  $\mathcal{W}_2 = \{1, 2, \dots, 2^{nR_2}\}$ , called the message sets, two encoding functions,

$$X_1 : \mathcal{W}_1 \rightarrow \mathcal{X}_1^n \quad (1.55)$$

and

$$X_2 : \mathcal{W}_2 \rightarrow \mathcal{X}_2^n \quad (1.56)$$

and a decoding function,

$$g : \mathcal{Y} \rightarrow \mathcal{W}_1 \times \mathcal{W}_2. \quad (1.57)$$

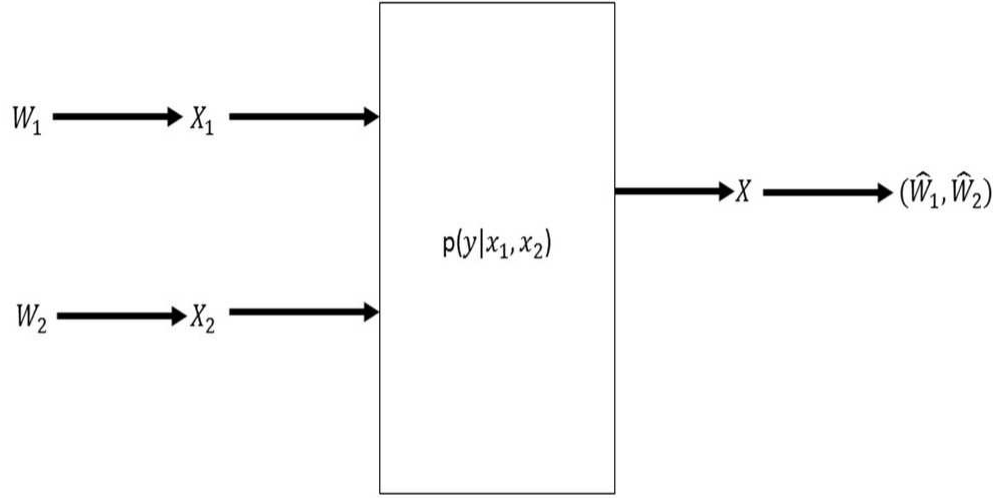


Figure 6. Multiple-access channel

There are two senders and one receiver for this channel. Sender 1 chooses an index  $W_1$  uniformly from the set  $\{1, 2, \dots, 2^{nR_1}\}$  and sends the corresponding codeword over the channel. Sender 2 does likewise. Assuming that the distribution of messages over the product set  $\mathcal{W}_1 \times \mathcal{W}_2$  is uniform (i.e., the messages are independent and equally likely), we define the average probability of error for the  $((2^{nR_1}, 2^{nR_2}), n)$  code as follows:

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{(w_1, w_2) \in \mathcal{W}_1 \times \mathcal{W}_2} \Pr \{g(Y^n) \neq (w_1, w_2) | (w_1, w_2)\}. \quad (1.58)$$

**Definition 19.** A rate pair  $(R_1, R_2)$  is said to be achievable for the multiple access channel if there exists a sequence of  $((2^{nR_1}, 2^{nR_2}), n)$  codes with  $P_e^{(n)} \rightarrow 0$ .

**Definition 20.** The capacity region of the multiple-access channel is the closure of the set of achievable  $(R_1, R_2)$  rate pairs.

The next theorem states the channel capacity or the capacity region of a Multiple access channel.

**Theorem 4.** *The capacity of a multiple-access channel  $(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y})$  is the closure of the convex hull of all  $(R_1, R_2)$  satisfying*

$$R_1 < I(X_1; Y|X_2), \quad (1.59)$$

$$R_2 < I(X_2; Y|X_1), \quad (1.60)$$

$$R_1 + R_2 < I(X_1, X_2; Y) \quad (1.61)$$

for some product distribution  $p_1(x_1)p_2(x_2)$  on  $\mathcal{X}_1 \times \mathcal{X}_2$ .

For a particular  $p_1(x_1)p_2(x_2)$ , the region is illustrated in Figure 7.

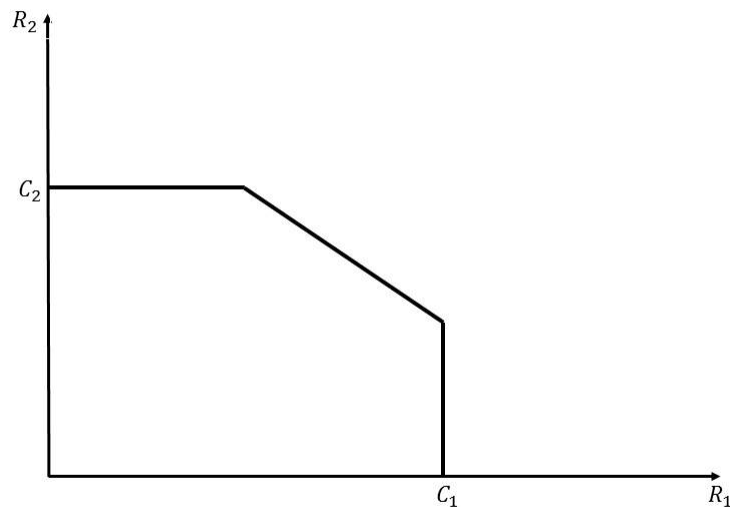


Figure 7. Achievable region of multiple-access channel for a fixed input distribution

All the definitions and data in this chapter is reproduced and inspired from [1] and [2]. The detailed proofs of the theorems and the definitions are mentioned in [1] and [2].

The knowledge of this chapter is very important to support and to understand the forth coming topics in this thesis. The interference channel is the very important topic for this thesis and hence it is discussed in detail in next section. Also the problem statement is discussed in the next section.

## CHAPTER 2. INTERFERENCE CHANNEL

Interference is the most common problem faced in Wireless Communication. The effects of interference ranges from being a minor annoyance to making the wireless system unusable. For example, while talking on phone we experience disturbance due to interference. The formal definition would be: An Interference Channel (IC) models the situation where a number ( $N$ ) of independent transmitters try to communicate their separate information to  $N$  different receivers via a common channel. There is a strict one-to-one correspondence between transmitters and receivers. Consequently, the transmission of information from each source to its corresponding receiver interferes with the communication between the other transmitters and their receivers. The interference channel can therefore be viewed as being comprised of  $N$  principal links and  $N(N - 1)$  interference links. An example of an IC is when far end cross talk occurs between two twisted pair cables in the same binder in a DSL. For wireless communication this is more frequent phenomenon, for example, in radio communications, since the electromagnetic spectrum is a limited resource, frequency bands are often simultaneously used by several radio links that are not completely isolated.

The study of the channel similar to the Interference Channel from information theoretic standpoint was initiated by [3] and the research in this field was pursued further, in spite of over four decades of research, the capacity region of general IC is still unsolved. A lot of research on this topic has focused on the two user-two receiver case.

### 2.1. Interfering Channels

The interfering channels can be classified according to their environment. The interfering channel can be a Discrete Memoryless Channel (DMC) or it can be a Gaussian Channel as discussed in [4]. The capacity calculation i.e. the achievability

and the converse depend on the type of the channel. Also there are different types of interference i.e. Very strong interference, Strong interference, Aligned Strong Interference, Weak interference, etc. Let us consider a 2-user interference channel, which is considered the simplest channel model in which multiple transmit-receive pairs communicate over a common noisy channel. This model was first mentioned in [3] and was studied in a series of works in [5], [6], [7], [8], [9], [10], [11] that considered certain special cases of IC where the capacity regions of the so-called *very strong* IC, the *strong* IC and certain classes of *degraded* and *deterministic* ICs, respectively, were established.

Since most modern wireless communication systems feature multiple antennas at some or all terminals it is of interest to study the 2-user Gaussian MIMO IC. Now as the capacity of the general case of the Interference Channel is not known, the capacity region of its sub classes are approached to solve. Intuitively, the *weak interference* channel is when the direct link channel is better than the cross link channel. The *strong interference* channel is when the cross link channel is better than the direct link channel and the *very strong interference* channel is where the cross link is more dominant than the direct link channel. We are accustomed to thinking of *strong interference* as a more detrimental effect than *weak interference*. From the information theoretic point of view there are number of works that have shown that *strong interference* is less harmful than *weak interference* and *very strong* interference is as good as no interference at all [5]. The *aligned-strong* interference is when the direct and the cross link channel matrices satisfy a matrix equation i.e. the direct link's channel matrix is a matrix multiple of the cross link's channel, where the matrix multiple satisfies some particular constraint. In general, the problem of characterizing the exact capacity of a MIMO IC even for small and special classes can be challenging.



In this work, we characterize capacity region for class of strong IC's. Now if we consider a simple case of two user MIMO, then it will be 2 user MIMO Gaussian IC with class of *strong* ICs. The result for this is derived in paper [12]. Now if we consider a practical situation of a cellular structure where there are two cells Cell-1 and Cell-2 nearby. Cell-1 has one user, i.e  $T_{x1}$  with multiple antennas and one Base Station, ie.  $R_{x1}$ . Similarly Cell-2 has one user,  $T_{x2}$  with multiple antennas and one Base Station, ie.  $R_{x2}$ . The cellular structure described is as shown in Figure 8.

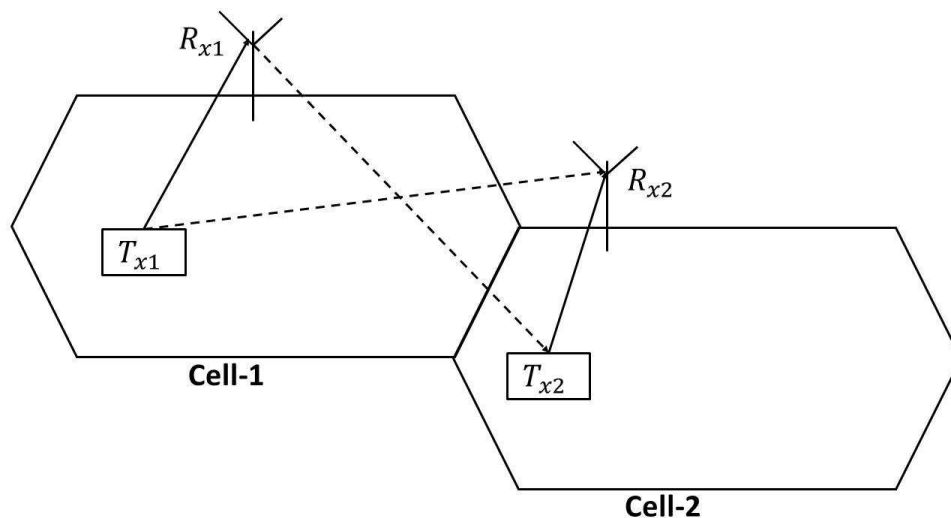


Figure 8. Cellular structure for 2-user interference channel

Here let us understand the real problem and what solution is given by [12]. We know that, Cell-1 has one transmitter and one receiver. In the absence of Cell-2, it will be a simple SISO communication, hence receiver 1 will receive a strong signal. Now if transmitter 2 is also communicating in Cell-2, then interference will come in the picture. Now if the transmitter 1 becomes selfless and communicates, with less power so that there will be less interference at receiver 2, then receiver 1 will not be able to receive the signal properly and hence will face interference from transmitter 2 signal. And if the transmitter 1 becomes selfish and transmits with high power then receiver 2 will face interference. Hence there has to be a tradeoff of

rates at which both the transmitters can communicate simultaneously by providing least interference. This rate region is the desired capacity region derived in [12].

Now, we reproduce the result of [12] here; the capacity of the 2-user strong MIMO IC  $\mathcal{C}_{IC}$  was defined as follows:

The capacity region  $\mathcal{C}_{IC}$  of an  $(M_1, N_1, M_2, N_2)$  *strong in partial order* IC, where the channel matrices satisfy the following condition:

$$H_{ii}^\dagger H_{ii} \preceq H_{ij}^\dagger H_{ij}, \forall i \neq j \in \{1, 2\}, \quad (2.1)$$

is given by the set of rate tuples satisfying the following constraints:

$$R_1 \leq \log |(I_{N_1} + H_{11} K_{x1} H_{11}^\dagger)|; \quad (2.2)$$

$$R_2 \leq \log |(I_{N_2} + H_{22} K_{x2} H_{22}^\dagger)|; \quad (2.3)$$

$$R_1 + R_2 \leq \log |(I_{N_2} + H_{12} K_{x1} H_{12}^\dagger + H_{22} K_{x2} H_{22}^\dagger)|; \quad (2.4)$$

$$R_1 + R_2 \leq \log |(I_{N_1} + H_{21} K_{x2} H_{21}^\dagger + H_{11} K_{x1} H_{11}^\dagger)| \quad (2.5)$$

To understand the physical interpretation and this result, it is first important to know the channel model. The 2-user IC as shown in Figure 9, where user  $i$  ( $T_{xi}$ ) has  $M_i$  antennas and receiver  $i$  ( $R_{xi}$ ) has  $N_i$  antennas, respectively for  $i = 1, 2$  is considered. Such a MIMO IC will be referred to as a  $(M_1, N_1, M_2, N_2)$  MIMO IC in the sequel.  $H_{ij} \in \mathbb{C}^{N_j \times M_i}$  models the channel matrix between  $T_{xi}$  and  $R_{xj}$ , which remain fixed for the entire duration of communication.  $\mathcal{H} = \{H_{11}, H_{12}, H_{21}, H_{22}\}$ . At time  $t$ ,  $T_{xi}$  chose a vector  $X_{it} \in \mathbb{C}^{M_i \times 1}$  and sends  $\sqrt{P_i} X_{it}$  into the channel, where for the input signals we assume the following power constraint:

$$\sum_{t=1}^n \frac{1}{n} \mathbb{E} \left( X_{it} X_{it}^\dagger \right) \preceq K_{xi}, \forall i = 1, 2. \quad (2.6)$$

The input and output alphabets are the set of real numbers, and the outputs are linear combinations of the inputs, plus the Gaussian noise. The received signals at time  $t$  can be written as:

$$Y_{1t} = H_{11}X_{1t} + H_{21}X_{2t} + Z_{1t}; Y_{1t} \in \mathbb{C}^{N_1 \times 1}, \quad (2.7)$$

$$Y_{2t} = H_{22}X_{2t} + H_{12}X_{1t} + Z_{2t}; Y_{2t} \in \mathbb{C}^{N_2 \times 1}, \quad (2.8)$$

where  $Z_{it} \in \mathbb{C}^{N_i \times 1}$  are i.i.d as  $\mathcal{CN}(0, I_{N_i})$  across  $i$  and  $t$ .

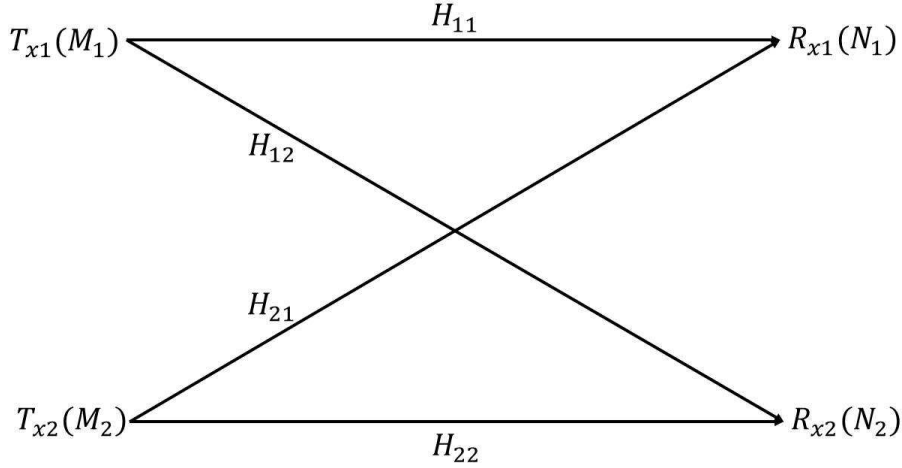


Figure 9. The  $(M_1, N_1, M_2, N_2)$  MIMO IC

Now the Equation (2.1) is the condition which the channel matrices should satisfy in order to be *Strong IC in partial order*. For example, for  $i = 1$  and  $j = 2$  we are considering the communication between the  $T_{x1}$  i.e. user-1 and  $R_{x2}$  which is BS-2. In this case the condition will be:

$$H_{11}^\dagger H_{11} \preceq H_{12}^\dagger H_{12}, \quad (2.9)$$

Hence we can say that the gram matrix of the direct link is less than the gram matrix

of interference link. Intuitively, we can say that the interference link is strong. It is easier to visualize this using Figure 9.

Now, Equations (2.2), (2.3), (2.4) and (2.5) define the capacity region. Capacity region is defined in a two-dimensional coordinate system where these bounds form a region where the communication between both the transmitters to their respective receivers is possible, i.e. the probability of error at the receiver tends to zero at both the receivers. Let us understand the physical interpretation of each equation. In equation (2.2), we have rate bound of  $R_1$ , it means it is the upper bound to the rate of communication between both the transmitters and receiver 1, where receiver 2 is considered mute. Similarly  $R_2$  is the upper bound to the rate of communication between both the transmitters and receiver 2, where receiver 1 is considered mute. In Equation (2.3) and (2.4) the sum rate bound of  $R_1 + R_2$  i.e. both the transmitters are simultaneously communicating to both the receivers. The difference in equation (2.3) and (2.4) is, in (2.3) it is assumed that genie provides the side information ( $X_2^n$ ) to  $R_{x1}$  and in (2.4) it is assumed that genie provides the side information ( $X_1^n$ ) to  $R_{x2}$ . The minimum of these two sum rate bounds is the actual  $R_1 + R_2$  sum rate bound. The capacity region of an IC is defined as the closure of the set of rate pairs  $(R_1, R_2)$  for which both receivers can decode their own messages with arbitrarily small positive error probability. Hence, intersection of all these rate bounds of  $(R_1, R_2)$  characterize the capacity region for 2-user strong MIMO IC. Each rate equation will form a plane in two-dimensional coordinate system and intersection of all these planes will characterize the capacity region.

The above work was focusing on the *strong interference* case, since *strong interference* is the major focus in this thesis we discussed it in detail. But there has been a noticeable advance in other class of ICs too. Recent results include the capacity regions of new and/or more general classes of channels than for which capacity was

previously known, e.g., the sum capacity of the so-called *noisy* interference channels ie., *low interference* regime, was found in [13], [14], [15]. The common feature of this line of work is that it focuses on subset of channel parameters but seeks to solve the challenging problem of obtaining the exact capacity of the SISO channel. In [16], the capacity region of a class of *very strong* SIMO ICs was characterized, this paper also demonstrates that, multiple transmit single receive antenna (MISO) Gaussian interference channels are much harder to characterize. The class of *aligned-strong* ICs are discussed in [17].

Our concentration will now be on Interfering Multiple Access Channels (IMAC). IMAC address the more typical problem in the practical scenerio. In the next section we will study IMAC in detail.

## 2.2. Interfering MAC channels

In the previous section we studied the so called *2-user interference channel* (IC) model is the simplest configuration to analyze and subsequently characterize the fundamental tradeoff that exists among the communication rates of several transmit-receive pairs operating via the same time frequency space. For instance, consider the uplink of two adjacent cells in the cellular network. The IC models the scenario when each cell contains exactly one user which intend to communicate to its corresponding base station and in the process causes interference to the other communication link. The intensive research on IC over the past decade has revealed numerous interesting and sometime surprising results which promises significant improvement of the overall performance of the network.

In practice however, the cellular region served by a base station rarely contains only one user and therefore in practice a scenario where both the adjacent cells have multiple users and all users in a particular cell is trying to communicate to their corresponding base station simultaneously and in the process interfering with the

communication of the adjacent cell is more typical. The communication within a single cell can be modelled by the well known *Multiple Access Channel* (MAC) and as a result we have two mutually interfering MAC (IMAC). In contrast to IC, the IMAC model not only captures the tradeoff between the communication rates of users belonging to different cells, but also simultaneously captures the tradeoff among the communication rates of users belonging to the same cell. Figure 10 shows the above mentioned scenario.

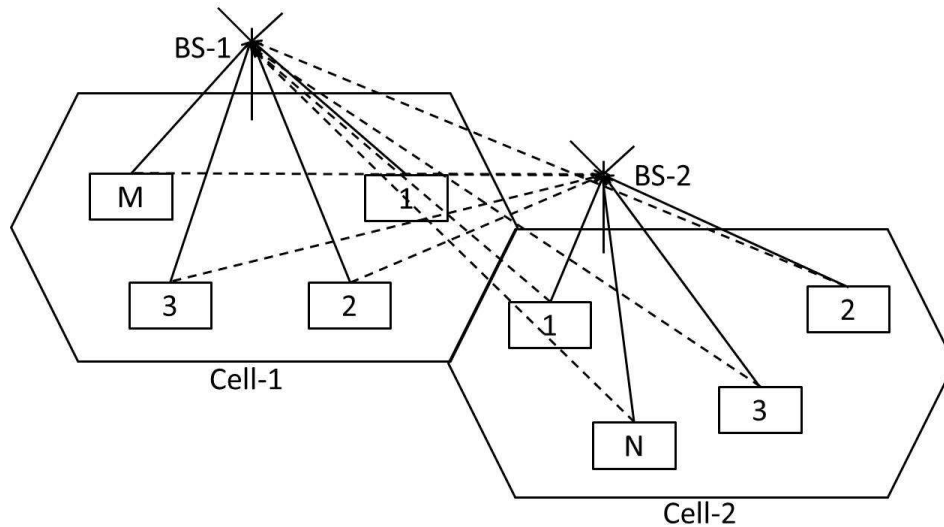


Figure 10. Cellular structure for two mutually interfering MIMO MACs

In Figure 10, we consider two mutually interfering MIMO MACs. MAC-1 is formed between  $M$  users and BS-1 and MAC-2 is formed between  $N$  users and BS-2. The dotted lines represents the interference channel and the solid line represents the direct channel. Now while characterizing the rate region there will be tradeoff between the rate of users from different MAC and tradeoff among rates of users belonging to the same MAC.

While there are plethora of research articles on ICs [18], [19], research efforts on the IMAC is almost non-existing. To the best of our knowledge, the only previous result on IMAC was reported in [20], where the authors consider an IMAC with

single antennal (SISO) nodes and finds capacity region where the interfering nodes have mixed *strong* [6] and *very strong* [5] interference, or both the interferers have stronger than the so called *strong* interference. They also derive lower and upper bounds to the sum-capacity of the SISO IMAC for *weak* interfering links. In next section, we will define the problem that is addressed in this thesis.

### 2.3. Problem Statement

This thesis is inspired from our paper [21]. In this work, we consider an IMAC with multiple antennas (MIMO) at all nodes. In particular we consider an IMAC with two mutually interfering MIMO MACs as shown in Figure 11, with arbitrary number of antennas at each transmitter and  $N$  antennas at each receiver.

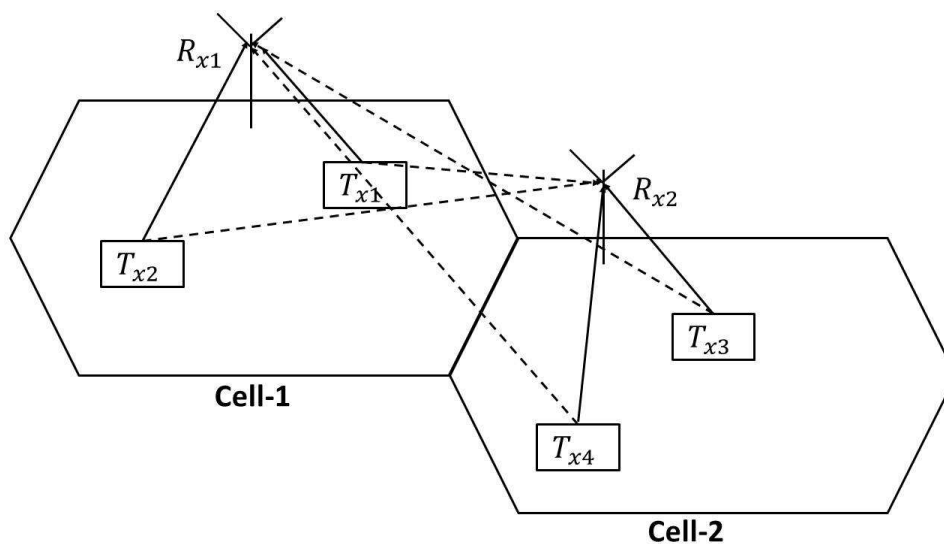


Figure 11. Cellular structure for two mutually interfering MIMO MACs with 4 users

We know that, to characterize the capacity region of two users IC we had to deal with the tradeoff between the rate of users from different MAC. But in the case of four users IMAC we have to deal with two different tradeoffs:

- Tradeoff between the rate of users from different MAC.
- Tradeoff among rates of users belonging to same MAC.

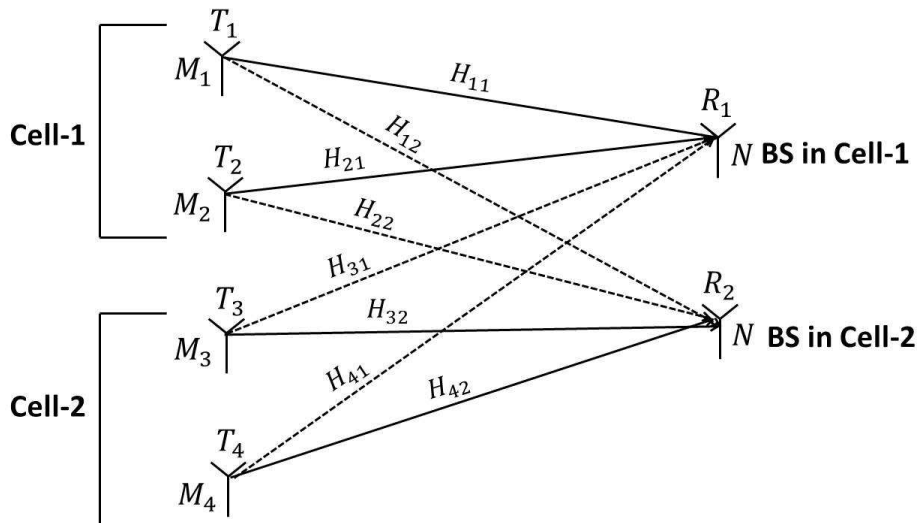


Figure 12. Channel Model for IMAC with 4 users

As shown in Figure 12,  $T_k$  is the user or transmitter with  $M_k$  antennas for  $k = \{1, 2, 3, 4\}$  and  $R_j$  is the Base Station or the receiver with  $N$  antennas for  $j = \{1, 2\}$ . Hence four users form two different MAC channels with BS-1 and BS-2. The two MACs for interfering MAC called IMAC. In this work we find the capacity region of IMAC for special class of channel matrices called Strongly ordered IMAC as explained in section 3.1. The non-trivial nature of this problem is to deal with the two communication problems simultaneously. Say for instance user 1 i.e.  $T_1$  is communicating with BS 1 i.e.  $R_1$  in Cell- 1, then it will experience interference from  $T_3$  and  $T_4$  from Cell-2 and it will also face MAC channel problem from  $T_2$ . Hence it is non trivial to derive the rate for which all the transmitters can communicate with the corresponding receiver at an optimum rate. The optimum rate would be the rate of communication at which all the transmitters can communicate such that both the receivers can decode their own messages with arbitrarily small positive error probability.



Our results will eventually prove that while characterizing the capacity region these two tradeoff will turn out to be unique, i.e. Interference channel problem and the MAC channel problem are the same. The capacity region of an IMAC is defined as the closure of the set of rate pairs  $(R_1, R_2, R_3, R_4)$  for which both the receivers can decode their messages as probability of error tends to zero. The capacity region for four user channel model will be intersection of planes in four dimensional co-ordinate system.

In next section we will describe the channel model in more detail and prove some preliminary results which will be followed by the derivation of main result i.e the capacity region.

## CHAPTER 3. MIMO MULTIPLE ACCESS CHANNEL

In this chapter, we characterize the entire capacity region of this channel assuming a relatively less restrictive *strong interference* condition. As a special case when both the MACs have only one user the IMAC turns into a 2-user IC and we can retrieve the previous result on 2-user MIMO IC derived in [17] and [22]. The different sections in this chapter address the problem defined in section 2.3. In section 3.1 the channel model considered is described and some preliminary results are proved, which is followed by the main result, i.e., the capacity region, in section 3.2. The converse is proved in subsection 3.2.1 and achievability is proved in subsection 3.2.2.

### 3.1. Channel Model

Here there's a communication network with two mutually interfering 2-user Multiple access channels (IMAC), as shown in Fig.13, where the  $k$ -th transmitter have  $M_k$  antennas and both the receivers have  $N$  antennas, respectively. The receivers and the transmitters are denoted by  $Rx_1$ ,  $Rx_2$  and  $Tx_k$  with  $k \in \Sigma = \{1, 2, 3, 4\}$ , respectively. Let  $H_{kj}$  represents the channel matrix from transmitter  $k$  to receiver  $j$ , where  $k \in \Sigma$  and  $j \in \{1, 2\}$ . Moreover, these channel matrices are also assumed to be *time-invariant*, i.e., the channel coefficients does not change with time. The received signal at  $Rx_j$  at the  $t$ -th channel use can be expressed as

$$z_{jt} = \sum_{k=1}^4 H_{kj} x_{kt} + u_{jt}, \quad (3.1)$$

where  $x_{kt} \in \mathbb{C}^{M_k \times 1}$  is the transmit vector from  $T_k$  at time  $t$ ,  $u_{jt} \sim \mathcal{CN}(0, I_N)$ <sup>1</sup> are independent across  $j \in \{1, 2\}$  and time. The transmitted signals from each

---

<sup>1</sup>We denote the distribution of a Gaussian vector with zero mean and covariance matrix  $Q$  by  $\mathcal{CN}(0, Q)$ .

transmitter satisfies the following average covariance constraint

$$\frac{1}{n} \sum_{t=1}^n \text{Cov}(x_{kt}) \preceq Q_k, \quad 1 \leq k \leq 4. \quad (3.2)$$

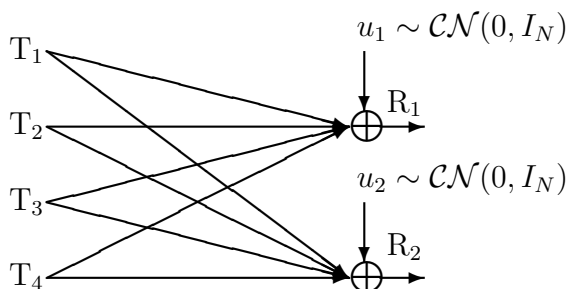


Figure 13. The IMAC with 4 MIMO users

On an IMAC,  $\text{Tx}_k$  wants to send a message  $\mathcal{W}_k$  at a rate of  $R_k$ , for  $k \in \Sigma$ . Among these messages  $\text{Rx}_j$  is interested only in  $\{\mathcal{W}_k : (2j - 1) \leq k \leq (2j)\}$ , for  $j \in \{1, 2\}$ . Suppose, given a message set  $w_1, \dots, w_4$ , where  $w_k \in \{1, \dots, 2^{nR_k}\}$ ,  $\text{Tx}_k$  chooses a codeword  $x_k^n \in \mathcal{C}_k(n)$  and sends it through the channel, where  $\mathcal{C}_k(n)$  is the codebook of the  $k$ -th transmitter containing  $nM_k$ -length codewords which satisfy the average covariance constraint in equation (3.2). Also assume that, the receivers  $\text{Rx}_1$  and  $\text{Rx}_2$  recover  $\hat{w}_1, \hat{w}_2$  and  $\hat{w}_3, \hat{w}_4$  from their respective received signals. Then the probability of detection error can be denoted as  $P_e(n) = \max\{\text{Pr}_1((\hat{w}_1, \hat{w}_2) \neq (w_1, w_2)), \text{Pr}_2((\hat{w}_3, \hat{w}_4) \neq (w_3, w_4))\}$ . Clearly, since  $(w_3, w_4)$  and  $(w_1, w_2)$  are not necessary at  $\text{Rx}_1$  and  $\text{Rx}_2$ , respectively, they do not appear in the error computation at the corresponding receivers. A rate tuple  $(R_1, R_2, R_3, R_4)$  - hereafter denoted as  $\bar{R}$  - is said to be *achievable* if there exists  $\mathcal{C}_k(n), k \in \Sigma$  such that  $P_e(n) \rightarrow 0$ , as  $n \rightarrow \infty$ . For a given coding scheme  $\mathcal{C}$ , the set  $(\mathcal{R}_{\mathcal{C}})$  of all such achievable rate tuple is called an achievable rate region. Capacity region is the closure of the union of all such possible achievable rate regions, i.e.,

$$C_{\text{IMAC}} = \overline{\cup_{\mathcal{C}} \mathcal{R}_{\mathcal{C}}}. \quad (3.3)$$

Here, we characterize the capacity region of the IMAC for a special class of channel matrices; the corresponding IMAC is called *Strongly ordered IMAC*.

**Definition 21** (*Strongly ordered IMAC*). A 4-user IMAC, as shown in Fig.13, is called a *strongly ordered IMAC* if the channel matrices satisfy the following constraints:

$$H_{\tau_k k}^\dagger H_{\tau_k k} \preceq H_{\tau_k j}^\dagger H_{\tau_k j}, \quad 1 \leq k \neq j \leq 2, \quad (3.4)$$

where  $\tau_k = \{(2k-1), 2k\}$  and as previously defined  $H_{\tau_k j} = [H_{(2k-1)j} H_{(2k)j}]$  for  $1 \leq k, j \leq 2$ .

**Remark 1.** It can be easily proved using the definition of partial order between two matrices that equation (3.4) implies the following

$$H_{(2k-1)k}^\dagger H_{(2k-1)k} \preceq H_{(2k-1)j}^\dagger H_{(2k-1)j}; \quad (3.5)$$

$$H_{(2k)k}^\dagger H_{(2k)k} \preceq H_{(2k)j}^\dagger H_{(2k)j}, \quad (3.6)$$

for all  $1 \leq k \neq j \leq 2$ .

*Proof.* We know that,  $1 \leq k \neq j \leq 2$ . Hence there are only two cases possible:

**Case-1:**  $k = 1$  and  $j = 2$ .

The physical interpretation of this case would be demonstrating the direct communication in Cell-1 as discussed in section 2.2 i.e, Users 1 and 2 are communicating with receiver 1.

In this case,

$$H_{\tau_1 1}^\dagger H_{\tau_1 1} \preceq H_{\tau_1 2}^\dagger H_{\tau_1 2}, \quad (3.7)$$

where,  $\tau_1 = \{1, 2\}$ .

Hence,

$$H_{\tau_1 1}^\dagger H_{\tau_1 1} \preceq H_{\tau_1 2}^\dagger H_{\tau_1 2} \quad (3.8)$$

$$\begin{bmatrix} H_{11}^\dagger \\ H_{21}^\dagger \end{bmatrix} \begin{bmatrix} H_{11} & H_{21} \end{bmatrix} \preceq \begin{bmatrix} H_{12}^\dagger \\ H_{22}^\dagger \end{bmatrix} \begin{bmatrix} H_{12} & H_{22} \end{bmatrix} \quad (3.9)$$

$$\begin{bmatrix} H_{11}^\dagger H_{11} & H_{11}^\dagger H_{21} \\ H_{21}^\dagger H_{11} & H_{21}^\dagger H_{21} \end{bmatrix} \preceq \begin{bmatrix} H_{12}^\dagger H_{12} & H_{12}^\dagger H_{22} \\ H_{22}^\dagger H_{12} & H_{22}^\dagger H_{22} \end{bmatrix} \quad (3.10)$$

Now,

$$\begin{bmatrix} H_{12}^\dagger H_{12} - H_{11}^\dagger H_{11} & H_{12}^\dagger H_{22} - H_{11}^\dagger H_{21} \\ H_{22}^\dagger H_{12} - H_{21}^\dagger H_{11} & H_{22}^\dagger H_{22} - H_{21}^\dagger H_{21} \end{bmatrix} \succeq 0 \quad (3.11)$$

Let  $x \in \mathbb{C}^{(M_1+M_2) \times 1}$  be an arbitrary complex vector and  $B \in \mathbb{C}^{(M_1+M_2) \times (M_1+M_2)}$  be an arbitrary P.S.D matrix; then according to the definition of Positive Semi-Definite Matrices [23], i.e  $\forall x \in \mathbb{C}^{(M_1+M_2) \times 1}$ ,  $B$  is said to be positive semidefinite matrix i.e,  $B \succeq 0$ , if  $x^\dagger B x \succeq 0$ .

Now if we set  $x = \begin{bmatrix} x_1 & 0 \end{bmatrix}^\dagger$  where  $x_1$  is arbitrary and,

$$B = \begin{bmatrix} H_{12}^\dagger H_{12} - H_{11}^\dagger H_{11} & H_{12}^\dagger H_{22} - H_{11}^\dagger H_{21} \\ H_{22}^\dagger H_{12} - H_{21}^\dagger H_{11} & H_{22}^\dagger H_{22} - H_{21}^\dagger H_{21} \end{bmatrix}, \quad (3.12)$$

then from the definition of P.S.D, equation (3.11) implies,

$$\begin{bmatrix} x_1^\dagger & 0 \end{bmatrix} \begin{bmatrix} H_{12}^\dagger H_{12} - H_{11}^\dagger H_{11} & H_{12}^\dagger H_{22} - H_{11}^\dagger H_{21} \\ H_{22}^\dagger H_{12} - H_{21}^\dagger H_{11} & H_{22}^\dagger H_{22} - H_{21}^\dagger H_{21} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \succeq 0 \quad (3.13)$$

$$x_1^\dagger \left[ H_{12}^\dagger H_{12} - H_{11}^\dagger H_{11} \right] x_1 \succeq 0 \quad (3.14)$$

Since  $x_1$  is also arbitrary, again from definition of P.S.D matrices we have,

$$H_{11}^\dagger H_{11} \preceq H_{12}^\dagger H_{12} \quad (3.15)$$

Similarly if we set  $x = \begin{bmatrix} 0 & x_2 \end{bmatrix}^\dagger$  where  $x_2$  is arbitrary, It can be proved that

$$H_{21}^\dagger H_{21} \preceq H_{22}^\dagger H_{22} \quad (3.16)$$

**Case-2:**  $k = 2$  and  $j = 1$

In this case there is the direct communication in Cell-2 i.e, Users 2 and 3 are communicating with receiver 2.

Hence,

$$H_{\tau_2 2}^\dagger H_{\tau_2 2} \preceq H_{\tau_2 1}^\dagger H_{\tau_2 1}, \quad (3.17)$$

where,  $\tau_2 = \{3, 4\}$ .

Hence,

$$H_{\tau_3 2}^\dagger H_{\tau_3 2} \preceq H_{\tau_3 1}^\dagger H_{\tau_3 1} \quad (3.18)$$

$$\begin{bmatrix} H_{32}^\dagger \\ H_{42}^\dagger \end{bmatrix} \begin{bmatrix} H_{32} & H_{42} \end{bmatrix} \preceq \begin{bmatrix} H_{31}^\dagger \\ H_{41}^\dagger \end{bmatrix} \begin{bmatrix} H_{31} & H_{41} \end{bmatrix} \quad (3.19)$$

$$\begin{bmatrix} H_{32}^\dagger H_{32} & H_{32}^\dagger H_{42} \\ H_{42}^\dagger H_{32} & H_{42}^\dagger H_{42} \end{bmatrix} \preceq \begin{bmatrix} H_{31}^\dagger H_{31} & H_{31}^\dagger H_{41} \\ H_{41}^\dagger H_{31} & H_{41}^\dagger H_{41} \end{bmatrix} \quad (3.20)$$

Now,

$$\begin{bmatrix} H_{31}^\dagger H_{31} - H_{32}^\dagger H_{32} & H_{31}^\dagger H_{41} - H_{32}^\dagger H_{42} \\ H_{41}^\dagger H_{31} - H_{42}^\dagger H_{32} & H_{41}^\dagger H_{41} - H_{42}^\dagger H_{42} \end{bmatrix} \succeq 0 \quad (3.21)$$

We again use the definition of P.S.D as defined above, for an arbitrary complex vector  $x \in \mathbb{C}^{(M_3+M_4) \times 1}$  and an arbitrary P.S.D matrix  $B \in \mathbb{C}^{(M_3+M_4) \times (M_3+M_4)}$ .

Now if we set  $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}^\dagger$  where  $x_1$  is arbitrary and,

$$B = \begin{bmatrix} H_{31}^\dagger H_{31} - H_{32}^\dagger H_{32} & H_{31}^\dagger H_{41} - H_{32}^\dagger H_{42} \\ H_{41}^\dagger H_{31} - H_{42}^\dagger H_{32} & H_{41}^\dagger H_{41} - H_{42}^\dagger H_{42} \end{bmatrix} \quad (3.22)$$

Then from the definition of P.S.D, equation (3.21) implies,

$$\begin{bmatrix} x_1^\dagger & 0 \end{bmatrix} \begin{bmatrix} H_{31}^\dagger H_{31} - H_{32}^\dagger H_{32} & H_{31}^\dagger H_{41} - H_{32}^\dagger H_{42} \\ H_{41}^\dagger H_{31} - H_{42}^\dagger H_{32} & H_{41}^\dagger H_{41} - H_{42}^\dagger H_{42} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \succeq 0 \quad (3.23)$$

$$x_1^\dagger \left[ H_{32}^\dagger H_{32} - H_{31}^\dagger H_{31} \right] x_1 \succeq 0 \quad (3.24)$$

Since  $x_1$  is also arbitrary, again from definition of P.S.D matrices we have,

$$H_{32}^\dagger H_{32} \preceq H_{31}^\dagger H_{31} \quad (3.25)$$

Similarly, if we set  $x = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}^\dagger$ , where  $x_2$  is arbitrary, It can be proved that

$$H_{42}^\dagger H_{42} \preceq H_{41}^\dagger H_{41} \quad (3.26)$$

Hence Equation (3.5) and (3.6) are the shorthand notations of Equations (3.25), (3.26), (3.15) and (3.16).  $\square$

Now if we practically interpret this definition then, Equation (3.15) means that the gram matrix of  $H_{11}$  is less than or equal to gram matrix of  $H_{12}$ . Now  $H_{11}$  is the channel matrix of the direct link and the  $H_{12}$  is the channel matrix of the interfering link. Hence intuitively we can say that the interferer channel is stronger than the

direct channel. Similarly for equations (3.16), (3.25) and (3.26) we can learn that the interference channel is stronger.

To derive an explicit expression for an achievable rate region and to prove that it is the capacity region, we need several interesting information theoretic results, which we now derive in the following Lemmas.

First such result provides an upper bound on the conditional mutual information between the received signal at either of the receivers and a set of transmitters over  $n$ -channel uses. Note that the received signal at  $Rx_j$  over  $n$  channel uses can be written as

$$z_j^n = \sum_{k \in \Sigma} H_{kj}^n x_k^n + u_j^n = H_{\Sigma j}^n x_{\Sigma}^n + u_j^n, \quad j \in \{1, 2\}. \quad (3.27)$$

Here,  $H_{\Sigma j}^n = \begin{bmatrix} H_{1j}^n & H_{2j}^n & H_{3j}^n & H_{4j}^n \end{bmatrix}$ .

**Lemma 1.** *On an IMAC, as shown in Figure 13, with transmitters having signal covariance constraint as in equation (3.2), the conditional entropy of the received signal at  $Rx_j$  have the following upper bound:*

$$I(x_{\alpha}^n; z_j^n | x_{\alpha^c}^n) \leq n \log \left| I_N + \sum_{k \in \alpha} H_{kj} Q_k H_{kj}^{\dagger} \right|. \quad (3.28)$$

*Proof of Lemma 1.* We know that,

$$z_j^n = \sum_{k \in \alpha} H_{kj}^n x_k^n + u_j^n = H_{\alpha j}^n x_{\alpha}^n + u_j^n, \quad j \in \{1, 2\}, \quad \alpha \subseteq \{1, 2, 3, 4\}, \quad \alpha^c = \Sigma \setminus \alpha \quad (3.29)$$

Here for a sequence of column vectors  $x_{k1}, \dots, x_{kn}$ ,  $x_k^n$  represents a tall vector obtained by stacking all the component vectors in a single column, i.e.,  $x_k^n = [x_{k1}^{\dagger} \dots x_{kn}^{\dagger}]^{\dagger}$ . For the sequence of matrices, there is a column wise concatenation of all matrices i.e.  $H_{\alpha j}^n = \begin{bmatrix} H_{1j}^n & H_{2j}^n & H_{3j}^n & H_{4j}^n \end{bmatrix}$  for  $\alpha = \{1, 2, 3, 4\}$ . For a matrix  $A$ , we denote its Tensor product [24] with the  $n \times n$  identity matrix  $I_n$  by  $A^n$ , i.e.,



$A^n = I_n \otimes A$ . Hence  $H_{kj}^n$  is basically shorthand notation for  $(I_n \otimes H_{kj})$ . Now,

$$I(x_\alpha^n; z_j^n | x_{\alpha^c}^n) = h(z_j^n | x_{\alpha^c}^n) - h(z_j^n | x_{\alpha^c}^n x_\alpha^n) \quad (3.30)$$

$$= h\left(\sum_{k \in \alpha} H_{kj}^n x_k^n + u_j^n\right) - h(u_j^n) \quad (3.31)$$

Now if we assume that  $\tilde{x}_k^n$  is the input signal with Gaussian distribution where  $Cov(\tilde{x}_k) = Cov(x_k)$  then from the results of information theory, we know that, entropy maximizes with gaussian distribution and it introduces inequality as follows,

$$I(x_\alpha^n; z_j^n | x_{\alpha^c}^n) \preceq \sum_{t=1}^n h\left(\sum_{k \in \alpha} H_{kj} \tilde{x}_{kt} + u_{jt} \mid \sum_{k \in \alpha} H_{kj}^{(t-1)} \tilde{x}_k^{(t-1)} + u_j^{(t-1)}\right) - h(u_j^n) \quad (3.32)$$

$$\preceq \sum_{t=1}^n h\left(\sum_{k \in \alpha} H_{kj} \tilde{x}_{kt} + u_{jt}\right) - h(u_j^n) \quad (3.33)$$

Above inequality is introduced by removing the conditioning. Now we will use information theoretic result of Entropy of a Gaussian random variable [2]

$$I(x_\alpha^n; z_j^n | x_{\alpha^c}^n) = \sum_{t=1}^n \log |2\pi e \left(\sum_{k \in \alpha} H_{kj} Cov(\tilde{x}_{kt}) H_{kj}^\dagger + I_N\right)| - \log (2\pi e)^{nN} \quad (3.34)$$

$$= \left[n \times \frac{1}{n}\right] \sum_{t=1}^n \log |2\pi e \left(\sum_{k \in \alpha} H_{kj} Q_{kt} H_{kj}^\dagger + I_N\right)| - nN \log(2\pi e) \quad (3.35)$$

Now using the Jensen's Inequality [2] we get,

$$I(x_\alpha^n; z_j^n | x_{\alpha^c}^n) \preceq n \log |2\pi e \left(\sum_{k \in \alpha} H_{kj} \left(\frac{1}{n} \sum_{t=1}^n Q_{kt}\right) H_{kj}^\dagger + I_N\right)| - nN \log(2\pi e) \quad (3.36)$$

Now, as we know that,

$$\frac{1}{n} \sum_{t=1}^n \text{Cov}(x_{\alpha t}) \preceq Q_{\alpha}, \quad \alpha \subseteq \Sigma.$$

where  $Q_{\alpha}$  is a block diagonal matrix with diagonal entries of  $Q_{\alpha t}$ . The inequality occurs because the  $\log \det(\cdot)$  is a monotonically increasing function on the cone of positive definite matrices.

$$I(x_{\alpha}^n; z_j^n | x_{\alpha^c}^n) \preceq nN \log(2\pi e) + n \log \left| \sum_{k \in \alpha} H_{kj} Q_k H_{kj}^{\dagger} + I_N \right| - nN \log(2\pi e) \quad (3.37)$$

In short form,

$$I(x_{\alpha}^n; z_j^n | x_{\alpha^c}^n) \leq n \log |I_N + H_{\alpha j}^n Q_{\alpha} H_{\alpha j}^{n\dagger}| \quad (3.38)$$

In Equation (3.35) the expected value of  $x$  is zero with certain power constraint if we use a non-zero expected value, that leads to inefficient power constraint. It is a typical assumption in information theory.

The significance of Lemma 1 is that it shows the mutual information between the transmitters communicating simultaneously with respective receivers. For example, when  $\alpha = \{1\}$  and so  $\alpha^c = \{2, 3, 4\}$ ; in this case  $I(x_1^n; z_1^n | x_2^n, x_3^n, x_4^n)$  is information between transmitter- 1 and receiver-1 where signal of transmitter 2,3 and 4 is known at receiver 1. The mutual information will be

$$I(x_1^n; z_1^n | x_2^n, x_3^n, x_4^n) \leq n \log |I_N + H_{11} Q_1 H_{11}^{\dagger}|.$$

A similar set of upper bounds were proved in [19] and [17] in the context of 2-user MIMO IC, where there were 2 transmitters in the network.  $\square$

The next lemma basically proves that for two full-rank complex, square channel matrices, satisfying certain condition leads to difference of two mutual information terms less than or equal to zero. The next result is a very important as this result is used in various proofs of section 3.2.1.

**Lemma 2.** *Let  $G, H \in \mathbb{C}^{M \times M}$  are two full-rank, complex matrices and  $x \in \mathbb{C}^{M \times 1}$  is a random vector with arbitrary distribution and  $\tilde{u}_1$  and  $\tilde{u}_2$  are i.i.d. as  $\mathcal{CN}(0, I_M)$  which are also independent of  $x$ . If the matrices  $G$  and  $H$  satisfy the following condition*

$$G^\dagger G \preceq H^\dagger H. \quad (3.39)$$

then,

$$\mathcal{D} \triangleq I(x; Gx + \tilde{u}_1) - I(x; Hx + \tilde{u}_2) \leq 0. \quad (3.40)$$

**Remark 2** (Independence of input power constraint). *The inequality in equation (3.40) with an additional covariance constraint on  $x$  was proved in [22] using the extremal inequality of [25]. The proof given here is independent of any covariance constraint on  $x$  and is therefore more suitable for this analysis because, in subsection 3.2.1 we shall have occasions to use inequality equation (3.40) with a composite vector such as  $x_{\tau_1}^n$  and  $x_{\tau_2}^n$  (e.g., see equation (3.99)).*

*Proof of Lemma 2.* The proof is based on the well known information theoretic *Data Processing Inequality* (Sec. 2.8, [2]). For convenience we shall use the following shorthand in the rest of the proof,  $(G^\dagger G)^{-1} = K_1$  and  $(H^\dagger H)^{-1} = K_2$  and  $K_d = ((G^\dagger G)^{-1} - (H^\dagger H)^{-1})$ , hence  $K_d = K_1 - K_2$ . Using these notations, we have

$$\mathcal{D} = h(Gx + \tilde{u}_1) - h(Hx + \tilde{u}_2) - h(Gx + \tilde{u}_1|x) + h(Hx + \tilde{u}_2|x), \quad (3.41)$$

$$= h(Gx + \tilde{u}_1) - h(Hx + \tilde{u}_2) - h(\tilde{u}_1) + h(\tilde{u}_2), \quad (3.42)$$

$$= h(x + G^{-1}\tilde{u}_1) - h(x + H^{-1}\tilde{u}_2) + \log |G^\dagger G| - \log |H^\dagger H| \quad (3.43)$$

$$- h(\tilde{u}_1) + h(\tilde{u}_2), \quad (3.44)$$

$$= h(x + G^{-1}\tilde{u}_1) - h(x + H^{-1}\tilde{u}_2) \quad (3.45)$$

$$- h(G^{-1}\tilde{u}_1) + h(H^{-1}\tilde{u}_2), \quad (3.46)$$

$$= h(x + \hat{u}_1) - h(x + \hat{u}_2) - h(\hat{u}_1) + h(\hat{u}_2), \quad (3.47)$$

where in last step  $\hat{u}_i \sim \mathcal{CN}(0, K_i)$ , for  $i = 1, 2$ . From the assumption (3.39), it follows that  $K_1 \succeq K_2$ , i.e.,  $Cov(\hat{u}_1) \succeq Cov(\hat{u}_2)$ . Therefore,  $\hat{u}_1$  can be written as  $\hat{u}_1 = \hat{u}_2 + \hat{u}_d$ , where  $\hat{u}_d \sim \mathcal{CN}(0, K_d)$  and is independent of  $\hat{u}_2$ . Substituting that in the above expression we get,

$$\mathcal{D} = h(x + \hat{u}_2 + \hat{u}_d) - h(x + \hat{u}_2) + h(\hat{u}_2) - h(\hat{u}_1), \quad (3.48)$$

$$= [h(x + \hat{u}_2 + \hat{u}_d) - h(\hat{u}_2 + \hat{u}_d)] - [h(x + \hat{u}_2) - h(\hat{u}_2)] \quad (3.49)$$

$$= I(x; x + \hat{u}_2 + \hat{u}_d) - I(x; x + \hat{u}_2) \leq 0, \quad (3.50)$$

where the equality in equation (3.50) follows from the fact that  $\hat{u}_1 = \hat{u}_2 + \hat{u}_d$  and the inequality follows from the well known *Data processing inequality* since  $x \rightarrow x + \hat{u}_2 \rightarrow x + \hat{u}_2 + \hat{u}_d$  forms a Markov chain for any arbitrary  $x$ .  $\square$

Now we need to prove the similar result for full rank complex, non-square matrices. In this case the result of Lemma 2 can be used after converting the matrices into square and invertible.

**Lemma 3.** Let  $H_1, H_2 \in \mathbb{C}^{N \times M}$  are full-rank matrices,  $x \in \mathbb{C}^{M \times 1}$  is a random vector with arbitrary distribution and  $\tilde{u}_1$  and  $\tilde{u}_2$  are i.i.d. as  $\mathcal{CN}(0, I_N)$  which are also independent of  $x$ . If the matrices  $H_1$  and  $H_2$  satisfy the following condition

$$H_1^\dagger H_1 \preceq H_2^\dagger H_2, \quad (3.51)$$

then

$$\mathcal{D}_s = I(x; H_1 x + \tilde{u}_1) - I(x; H_2 x + \tilde{u}_2) \leq 0. \quad (3.52)$$

*Proof for Lemma 3.* In the proof, the main idea is to first convert the matrices into square and invertible ones by replacing the zero singular values by infinitesimal singular values, then apply the result of Lemma 2 and finally, take limit to remove the contribution of the additional singular values. This procedure is valid since the set of singular matrices is dense in the space of non-singular square matrices. Note that a similar approach was used in [26] to extend their results on square invertible matrices to the case of non-square channel matrices.

Let us denote  $\tilde{W}_1 = H_1 x + \tilde{u}_1$  and  $\tilde{W}_2 = H_2 x + \tilde{u}_2$ . Hence,

$$D_s = I(x; \tilde{W}_1) - I(x; \tilde{W}_2) \quad (3.53)$$

$$= h(\tilde{W}_1) - h(\tilde{W}_1|x) - h(\tilde{W}_2) + h(\tilde{W}_2|x) \quad (3.54)$$

Now, as  $H_1$  and  $H_2$  are  $N \times M$  matrices and we need to convert them to square and invertible matrices. Using Singular Value Decomposition [24] we get,  $H_1 = V_1 \Sigma_1 U_1^\dagger$  and  $H_2 = V_2 \Sigma_2 U_2^\dagger$ .  $\Sigma_1$  and  $\Sigma_2$  have singular values in its diagonal entries in descending order.  $V_i \in \mathbb{U}^{N \times N}$  and  $U_i \in \mathbb{U}^{M \times M}$  for  $i = 1, 2$ . Since a unitary transformation does not

change the differential entropy of a random vector we have,

$$h(\tilde{W}_i) = h(V_i \Sigma_i U_i^\dagger + \tilde{u}_i) \quad (3.55)$$

$$= h(\Sigma_i U_i^\dagger x + V_i^{-1} \tilde{u}_i) \quad (3.56)$$

$$= h(W_i) \quad (3.57)$$

Let  $u_i = V_i^{-1} \tilde{u}_i$ . Hence,  $W_i = \Sigma_i U_i^\dagger x + u_i$ . Now, depending on the relative values  $N$  and  $M$  we modify the vector  $W_i$  can be written as;

**Case-1:**  $N > M$

In this case, the last  $(N - M)$  rows of  $\Sigma_i$  will have only zeros.

$$W_i = \begin{bmatrix} \hat{\Sigma}_i \\ 0 \end{bmatrix} U_i^\dagger x + u_i \quad (3.58)$$

$$= \begin{bmatrix} \hat{W}_i \\ u_{ie} \end{bmatrix} \quad (3.59)$$

Where,  $\hat{W}_i = \hat{\Sigma}_i U_i^\dagger x + \hat{u}_i$ ;  $u_{ie} \sim \mathcal{CN}(0, I_{N-M})$  and  $\hat{u}_i \sim \mathcal{CN}(0, I_M)$ .  $u_{ie}$  and  $\hat{u}_i$  are mutually independent.  $\hat{\Sigma}_i \in \mathbb{C}^{(M \times M)}$  contains the non-zero diagonal singular values of  $H_i$ .

Since,  $u_{ie}$  is independent of  $\hat{u}_i$ ;

$$h(W_i) = h(\hat{W}_i) + h(u_{ie} | \hat{W}_i) \quad (3.60)$$

$$= h(\hat{W}_i) + h(u_{ie}) \quad (3.61)$$

$$= h(\hat{W}_i) + \log |2\pi e I_{N-M}| \quad (3.62)$$

$$= h(\hat{W}_i) + \log(2\pi e)^{(N-M)} \quad (3.63)$$

$$= h(\hat{W}_i) + (N - M) \log(2\pi e) \quad (3.64)$$

**Case-2:**  $N < M$

In this case, the last  $(M - N)$  columns of  $\Sigma_i$  will have only zeros.

$$\Sigma_i = \begin{bmatrix} \tilde{\Sigma}_i & 0_{N \times p} \end{bmatrix} \quad (3.65)$$

Let  $p = M - N$  and  $\hat{\Sigma}_i = \begin{bmatrix} \tilde{\Sigma}_i + \delta I_N & 0_{N \times p} \\ 0_{p \times N} & \delta I_{p \times p} \end{bmatrix}$ . Now as  $\Sigma_i = \begin{bmatrix} \tilde{\Sigma}_i & 0_{N \times p} \end{bmatrix}$ , we can say that,

$$\hat{\Sigma}_i = \begin{bmatrix} \Sigma_i \\ L \end{bmatrix} \text{ where } L = \begin{bmatrix} 0_{p \times N} & \delta I_{p \times p} \end{bmatrix}. \quad (3.66)$$

Here  $\hat{u}_i = \begin{bmatrix} u_i^\dagger & u_{ie}^\dagger \end{bmatrix}^\dagger \sim \mathcal{CN}(0, I_M)$ .

The differential entropy;

$$h(W_i) = h\left(\Sigma_i U_i^\dagger x + u_i\right) \quad (3.67)$$

$$= h(\hat{\Sigma}_i U_i^\dagger x + \hat{u}_i) - h(LU_i^\dagger x + u_{ie} | \Sigma_i U_i^\dagger x + u_i) \quad (3.68)$$

Now we introduce the limits,

$$h(W_i) = \lim_{\delta \rightarrow 0} \left[ h(\hat{\Sigma}_i U_i^\dagger x + \hat{u}_i) - h(LU_i^\dagger x + u_{ie} | W_i) \right] \quad (3.69)$$

We know that  $\hat{W}_i = \hat{\Sigma}_i U_i^\dagger x + \hat{u}_i$ , also when we apply limits to L it will essentially be zero matrix and hence only  $u_{ie}$  will be there,

$$h(W_i) = \lim_{\delta \rightarrow 0} \left[ h(\hat{W}_i) - h(u_{ie} | W_i) \right] \quad (3.70)$$

Now,  $u_{ie}$  is independent of  $W_i$  as it is independent of  $u_i$  and  $x$ .

Also  $u_{ie} \sim \mathcal{CN}(0, I_{(M-N)})$ ,

$$h(W_i) = \lim_{\delta \rightarrow 0} \left[ h(\hat{W}_i) - (M - N) \log(2\pi e) \right] \quad (3.71)$$

Denoting  $\hat{\Sigma}_i U_i^\dagger$  by  $\hat{H}_i$  it can be proved that,

$$\hat{H}_1^\dagger \hat{H}_1 \preceq \hat{H}_2^\dagger \hat{H}_2 \quad (3.72)$$

For  $N \geq M$ , this fact follows directly from the lemma 1

$$U_1 \Sigma_1^\dagger \Sigma_1 U_1^\dagger \preceq U_2 \Sigma_2^\dagger \Sigma_2 U_2^\dagger \quad (3.73)$$

$$\text{or, } U_1 \begin{bmatrix} \hat{\Sigma}_1^2 \\ 0 \end{bmatrix} U_1^\dagger \preceq U_2 \begin{bmatrix} \hat{\Sigma}_2^2 \\ 0 \end{bmatrix} U_2 \quad (3.74)$$

$$\text{or, } U_1 \hat{\Sigma}_1^\dagger \hat{\Sigma}_1 U_1^\dagger \preceq U_2 \hat{\Sigma}_2^\dagger \hat{\Sigma}_2 U_2^\dagger \quad (3.75)$$

$$\text{or, } \hat{H}_1^\dagger \hat{H}_1 \preceq \hat{H}_2^\dagger \hat{H}_2 \quad (3.76)$$

In the above set of equations (3.74) follows from the definition of  $\Sigma_i$  in (3.65), (3.74) follows from the fact that for any two PSD matrices,  $K_1$  and  $K_2$ ,  $K_1 \prec K_2$  if and only if  $K_1^{\frac{1}{2}} \prec K_2^{\frac{1}{2}}$ , and finally, (3.75) follows from the definition of  $\hat{\Sigma}_i$ .



For  $N < M$  the proof is as follows:

$$U_1 \Sigma_1^\dagger \Sigma_1 U_1^\dagger \preceq U_2 \Sigma_2^\dagger \Sigma_2 U_2^\dagger \quad (3.77)$$

$$\text{or, } U_1 \begin{bmatrix} \tilde{\Sigma}_1^2 & 0_{N \times p} \\ 0_{p \times N} & 0_{p \times p} \end{bmatrix} U_1^\dagger \preceq U_2 \begin{bmatrix} \tilde{\Sigma}_2^2 & 0_{N \times p} \\ 0_{p \times N} & 0_{p \times p} \end{bmatrix} U_2^\dagger \quad (3.78)$$

$$\text{or, } U_1 \begin{bmatrix} \tilde{\Sigma}_1^2 & 0_{N \times p} \\ 0_{p \times N} & 0_{p \times p} \end{bmatrix} U_1^\dagger + \delta U_1 U_1^\dagger \preceq U_2 \begin{bmatrix} \tilde{\Sigma}_2^2 & 0_{N \times p} \\ 0_{p \times N} & 0_{p \times p} \end{bmatrix} U_2^\dagger + \delta U_2 U_2^\dagger; \quad \forall \delta > 0 \quad (3.79)$$

$$\text{or, } U_1 \begin{bmatrix} \tilde{\Sigma}_1^2 & 0_{N \times p} \\ 0_{p \times N} & 0_{p \times p} \end{bmatrix} U_1^\dagger + \delta I_M \preceq U_2 \begin{bmatrix} \tilde{\Sigma}_2^2 & 0_{N \times p} \\ 0_{p \times N} & 0_{p \times p} \end{bmatrix} U_2^\dagger + \delta I_M; \quad \forall \delta > 0 \quad (3.80)$$

$$\text{or, } U_1 \begin{bmatrix} \tilde{\Sigma}_1 + \delta I_N & 0_{N \times p} \\ 0_{p \times N} & \delta I_{p \times p} \end{bmatrix} U_1^\dagger \preceq U_2 \begin{bmatrix} \tilde{\Sigma}_2 + \delta I_N & 0_{N \times p} \\ 0_{p \times N} & \delta I_{p \times p} \end{bmatrix} U_2^\dagger; \quad \forall \delta > 0 \quad (3.81)$$

$$\text{or, } U_1 \begin{bmatrix} (\tilde{\Sigma}_1 + \delta I_N)^2 & 0_{N \times p} \\ 0_{p \times N} & (\delta)^2 I_{p \times p} \end{bmatrix} U_1^\dagger \preceq U_2 \begin{bmatrix} (\tilde{\Sigma}_2 + \delta I_N)^2 & 0_{N \times p} \\ 0_{p \times N} & (\delta)^2 I_{p \times p} \end{bmatrix} U_2^\dagger; \quad \forall \delta > 0 \quad (3.82)$$

$$\text{or, } U_1 \hat{\Sigma}_1^\dagger \hat{\Sigma}_1 U_1^\dagger \preceq U_2 \hat{\Sigma}_2^\dagger \hat{\Sigma}_2 U_2^\dagger \quad (3.83)$$

$$\text{or, } \hat{H}_1^\dagger \hat{H}_1 \preceq \hat{H}_2^\dagger \hat{H}_2 \quad (3.84)$$

In the above set of equations (3.78) follows from the definition of  $\Sigma_i$  in (3.65),(3.79) and (3.82) follow from the fact that for any two PSD matrices,  $K_1$  and  $K_2$ ,  $K_1 \prec K_2$  if and only if  $K_1^{\frac{1}{2}} \prec K_2^{\frac{1}{2}}$ , and finally, (3.83) follows from the definition of  $\hat{\Sigma}_i$ .

Using Equations (3.64) and (3.71) in equation (3.54);

$$D = h(\tilde{W}_1) - h(\tilde{W}_2) - h(\tilde{u}_1) + h(\tilde{u}_2) \quad (3.85)$$

$$= h(H_1 x + \tilde{u}_1) - h(H_2 x + \tilde{u}_2) - h(\tilde{u}_1) + h(\tilde{u}_2) \quad (3.86)$$

$$= \lim_{\delta \rightarrow 0} h(\hat{H}_1 x + \tilde{u}_1) - h(\hat{H}_2 x + \tilde{u}_2) \quad (3.87)$$

Using results of Lemma 2 we can prove;  $D \preceq 0$  □

In deriving the converse to the capacity region however, we shall need an upper bound to the difference of entropies of  $n$ -symbol extensions of the channel which is provided by the following Lemma.

**Lemma 4.** *Let  $H_1 \in \mathbb{C}^{N \times M}$  and  $H_2 \in \mathbb{C}^{N \times M}$  are full-rank matrices,  $\{x_t \in \mathbb{C}^{M \times 1}, 1 \leq t \leq n\}$  is a sequence of arbitrary random vectors and  $\tilde{u}_{jt}$  are i.i.d. as  $\mathcal{CN}(0, I_N)$  for all  $1 \leq t \leq n$  and mutually independent across  $1 \leq j \leq 2$ , which are also independent of  $x^n$ . In addition, if the matrices  $H_1$  and  $H_2$  satisfy equation (3.51), then*

$$\mathcal{D}_n = I(x^n; H_1^n x^n + \tilde{u}_1^n) - I(x^n; H_2^n x^n + \tilde{u}_2^n) \leq 0, \quad (3.88)$$

where  $H_i^n = I_n \otimes H_i$  for  $i = 1, 2$ .

*Proof of Lemma 4.* In equation (3.88),  $x^n \in \mathbb{C}^{Mn \times 1}$  is an arbitrary random vector,  $\tilde{u}_j^n \sim \mathcal{CN}(0, I_{nN})$  for  $j = 1, 2$  and are independent of  $x^n$ . Moreover, from the definition of partial order and equation (3.51) it follows that

$$(I_n \otimes H_1)^\dagger (I_n \otimes H_1) \preceq (I_n \otimes H_2)^\dagger (I_n \otimes H_2). \quad (3.89)$$

The statement of the Lemma then follows from the above conditions and Lemma 3. □

In the next section we will derive the capacity region. Initially we define the capacity region and then by using the traditional approach in information theory we derive the converse and the achieveability in the subsections.

### 3.2. Capacity Region

The capacity region here is basically the capacity region of the Interfering MAC and hence it will be the capacity region where all the transmitters can effectively communicate simultaneously to the desired receivers. The next theorem will essentially define the capacity region of the Interfering MAC ( $C_{\text{IMAC}}$ ) of a *strongly ordered* IMAC as intersection of all the outer rate bounds of the achievable rate tuples.

**Theorem 5.** *The capacity region ( $C_{\text{IMAC}}$ ) of a strongly ordered IMAC, as shown in Fig.13, with  $Tx_k$  having  $M_k$  antennas for  $k \in \Sigma$  and  $N$  antennas at both the receivers is given as*

$$C_{\text{IMAC}} = \left\{ \bar{R} : \bigcap_{\beta \subseteq \Sigma} R_\beta \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \beta} H_{kj} Q_k H_{kj}^\dagger \right| \right\}, \quad (3.90)$$

where  $R_\beta = \sum_{k \in \beta} R_k$ .

*Proof of Theorem 5.* In subsections 3.2.1 and 3.2.2 we prove that  $C_{\text{IMAC}} \subseteq C_u$  and  $\mathcal{R}_{\text{CMAC}} \subseteq C_{\text{IMAC}}$ , respectively. It is also clear from the expressions of  $\mathcal{R}_{\text{CMAC}}$  in (3.176) and  $C_u$  in (3.169) that  $\mathcal{R}_{\text{CMAC}} = C_u$ . Combining all of these we have,

$$\mathcal{R}_{\text{CMAC}} \subseteq C_{\text{IMAC}} \subseteq C_u = \mathcal{R}_{\text{CMAC}}. \quad (3.91)$$

Further, it is proved in subsection 3.2.2 that every rate tuple  $\bar{R} \in \mathcal{R}_{\text{CMAC}}$  can be achieved by independent Gaussian coding at all the transmitters and if each receiver does joint decoding of all the messages.  $\square$

**Remark 3** (special case: IC). *Note that if we reduce the number of users in each MAC to one, then the strongly ordered condition of equation (3.4) coincides with the strong in partial order definition in [22] and the capacity region also becomes the same as in [22]. It was shown in [22] that, under a particular covariance constraint, the*

results of [17] forms a subset of the result reported in [22]. In other words, the result of this paper while more general than IC results can also incorporate them as special cases.

### 3.2.1. Converse: Outer bounds to the achievable rate tuples

In this section, we derive all the outer rate bounds of the achievable rate tuples. Now  $R_1 + R_2 + R_3 + R_4$  is the sum rate bound which basically involves all the transmitter communicating. Similarly if we consider any three transmitter communicating simultaneously i.e. one of the transmitter is mute, for example  $R_1 + R_2 + R_3$  i.e.  $R_{4^c}$  then we get outer bound for three such cases i.e.  $R_{1^c}$ ,  $R_{2^c}$  and  $R_{3^c}$ . Similarly we can consider only two transmitters communicating simultaneously, in this case  $R_1 + R_2$  and  $R_3 + R_4$  will be transmitters from same cell and hence will have direct communication link, like a MAC channel. But for  $R_2 + R_3$ ,  $R_2 + R_4$ ,  $R_1 + R_3$  and  $R_1 + R_4$  there will be interfering links included, hence it will be interference channel. Finally we will consider only one transmitter communicating at once, i.e.  $R_1$ ,  $R_2$ ,  $R_3$  and  $R_4$ . Now we will use the previously defined lemmas in order to derive these rate bounds. Let us consider the sum bound first;

**Case-1:** Providing  $x_{\tau_1}^n$  at  $R_2$  i.e.(Providing receiver 2 with information of transmitter 1 and 2

Using Fano's inequality [2] we have,

$$nR_{\Sigma} \leq I(x_{\tau_1}^n; z_1^n) + I(x_{\tau_2}^n; z_2^n) + n\epsilon, \quad (3.92)$$

$$\leq I(x_{\tau_1}^n; z_1^n) + I(x_{\tau_2}^n; z_2^n, x_{\tau_1}^n), \quad (3.93)$$

$$= I(x_{\tau_1}^n; z_1^n) + I(x_{\tau_2}^n; z_2^n | x_{\tau_1}^n) + I(x_{\tau_2}^n; x_{\tau_1}^n), \quad (3.94)$$

$$= I(x_{\tau_1}^n; z_1^n) + I(x_{\tau_2}^n; z_2^n | x_{\tau_1}^n) \quad (3.95)$$

$$= h(z_1) - h(z_1^n | x_{\tau_1}^n) + h(z_2^n | x_{\tau_1}^n) - h(z_2^n | x_{\tau_1}^n x_{\tau_2}^n) \quad (3.96)$$

$$= h(z_1) - [h(z_1^n | x_{\tau_1}^n) - h(z_1^n | x_{\tau_1}^n x_{\tau_2}^n)] + [h(z_2^n | x_{\tau_1}^n) - h(z_2^n | x_{\tau_1}^n x_{\tau_2}^n)] - h(z_1^n | x_{\tau_1}^n x_{\tau_2}^n) \quad (3.97)$$

$$= I(x_{\Sigma}^n; z_1^n) - I(x_{\tau_2}^n; z_1^n | x_{\tau_1}^n) + I(x_{\tau_2}^n; z_2^n | x_{\tau_1}^n), \quad (3.98)$$

$$= I(x_{\Sigma}^n; z_1^n) + I(x_{\tau_2}^n; H_{\tau_2 2}^n x_{\tau_2}^n + u_2^n), \quad (3.99)$$

$$- I(x_{\tau_2}^n; H_{\tau_2 1}^n x_{\tau_2}^n + u_1^n), \quad (3.100)$$

$$\leq I(x_{\Sigma}^n; z_1^n), \quad (3.101)$$

$$\leq n \log \left| I_N + \sum_{k \in \Sigma} H_{k1} Q_k H_{k1}^\dagger \right| + n\epsilon, \quad (3.102)$$

where we have ignored  $n\epsilon$  term from equation (3.93) onwards for convenience, which can any way be discarded if we divide both sides by  $n$  and take  $n \rightarrow \infty$ . Equation (3.95) is because  $I(x_{\tau_2}^n; x_{\tau_1}^n) = 0$ , as they are mutually independent. Equation (3.93), (3.94) and (3.98) follows from the fact that additional information at  $R_2$  does not decrease mutual information and  $x_{\tau_1}^n$  and  $x_{\tau_2}^n$  are mutually independent. Equation (3.99) in an alternative expression for (3.102) using the following equations:

$$H_{\tau_2 j}^n = [H_{3j}^n \ H_{4j}^n] = [I_n \otimes H_{3j} \ I_n \otimes H_{4j}], j = 1, 2,$$

and equation (3.101) follows from Lemma 3. Finally equation (3.102) is obtained from Lemma 4. Now, dividing both sides of equation (3.102) by  $n$  and taking  $n \rightarrow \infty$  we get,

$$R_\Sigma \leq \log \left| I_N + \sum_{k \in \Sigma} H_{k1} Q_k H_{k1}^\dagger \right|, \quad (3.103)$$

**Case-2:** Providing  $x_{\tau_2}^n$  at  $R_1$ . Using the similar steps as previous case we get the following:

$$nR_\Sigma \leq I(x_{\tau_1}^n; z_1^n) + I(x_{\tau_2}^n; z_2^n) + n\epsilon, \quad (3.104)$$

$$\leq I(x_{\tau_2}^n; z_2^n) + I(x_{\tau_1}^n; z_1^n, x_{\tau_2}^n), \quad (3.105)$$

$$= I(x_{\tau_2}^n; z_2^n) + I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) + I(x_{\tau_2}^n; x_{\tau_1}^n), \quad (3.106)$$

$$= I(x_{\tau_2}^n; z_2^n) + I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) \quad (3.107)$$

$$= h(z_2) - h(z_2^n | x_{\tau_2}^n) + h(z_1^n | x_{\tau_2}^n) - h(z_1^n | x_{\tau_1} x_{\tau_2}^n) \quad (3.108)$$

$$= h(z_2) - [h(z_2^n | x_{\tau_2}^n) - h(z_2^n | x_{\tau_2} x_{\tau_1}^n)] + [h(z_1^n | x_{\tau_2}^n) - h(z_1^n | x_{\tau_2} x_{\tau_1}^n)] - h(z_2^n | x_{\tau_2}^n x_{\tau_1}^n) \quad (3.109)$$

$$= I(x_{\Sigma}^n; z_2^n) - I(x_{\tau_1}^n; z_2^n | x_{\tau_2}^n) + I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n), \quad (3.110)$$

$$= I(x_{\Sigma}^n; z_2^n) + I(x_{\tau_1}^n; H_{\tau_1 1}^n x_{\tau_1}^n + u_1^n), \quad (3.111)$$

$$- I(x_{\tau_1}^n; H_{\tau_1 2}^n x_{\tau_1}^n + u_2^n), \quad (3.112)$$

$$\leq I(x_{\Sigma}^n; z_2^n), \quad (3.113)$$

$$\leq n \log \left| I_N + \sum_{k \in \Sigma} H_{k2} Q_k H_{k2}^\dagger \right| + n\epsilon, \quad (3.114)$$

Hence we get,

$$R_\Sigma \leq \log \left| I_N + \sum_{k \in \Sigma} H_{k2} Q_k H_{k2}^\dagger \right|, \quad (3.115)$$

Combining equation (3.103) and (3.115) we have the sum rate upper bound

$$R_\Sigma \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \Sigma} H_{kj} Q_k H_{kj}^\dagger \right|, \quad (3.116)$$

Let us now consider outer bounds on sum of 3 users rates, such as  $R_1 + R_2 + R_3$ . Again, using Fano's inequality we have

$$nR_{\{4\}^c} \leq I(x_{\tau_1}^n; z_1^n) + I(x_3^n; z_2^n) + n\epsilon, \quad (3.117)$$

Next, we provide the side information  $x_4^n$  to both  $R_1$  and  $R_2$  and  $x_{\tau_1}^n$  to only  $R_2$  to obtain the following

$$nR_{\{4\}^c} \leq I(x_{\tau_1}^n; z_1^n | x_4^n) + I(x_3^n; z_2^n | x_{\tau_1}^n, x_4^n), \quad (3.118)$$

$$= I(x_\Sigma^n; z_1^n | x_4^n) - I(x_{\tau_3}^n; z_1^n | x_{\tau_1}^n, x_4^n) \quad (3.119)$$

$$+ I(x_3^n; z_2^n | x_{\tau_1}^n, x_4^n), \quad (3.120)$$

$$= I(x_\Sigma^n; z_1^n | x_4^n) + I(x_3^n; H_{32}^n x_3^n + u_2^n) \quad (3.121)$$

$$- I(x_3^n; H_{31}^n x_3^n + u_1^n) \quad (3.122)$$

$$\leq I(x_\Sigma^n; z_1^n | x_4^n) \quad (3.123)$$

$$\leq n \log \left| I_N + \sum_{k \in \{4\}^c} H_{k1} Q_k H_{k1}^\dagger \right| + n\epsilon, \quad (3.124)$$

where equation (3.118) follows from the independence of  $x_{\tau_1}^n$  to  $x_4^n$  and  $x_3^n$  to  $x_{\tau_1}^n$  and  $x_4^n$ , (3.119) is chain rule of mutual information, (3.121) is just a simpler expression of (3.119) and the first inequality in (3.123) follows from Lemma 4. Finally, the last inequality in equation (3.124) follows from Lemma 1.

Alternatively, providing the side information of  $x_4^n$  to both  $R_1$  and  $R_2$  and  $x_3^n$  to only  $R_1$  we get from equation (3.117)

$$nR_{\{4\}^c} \leq I(x_{\tau_1}^n; z_1^n, x_{\tau_2}^n) + I(x_3^n; z_2^n, x_4^n), \quad (3.125)$$

$$= I(x_{\tau_1}^n; x_{\tau_2}^n) + I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) + I(x_3^n; x_4^n) \quad (3.126)$$

$$+ I(x_3^n; z_2^n | x_4^n), \quad (3.127)$$

$$\leq I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) + I(x_3^n; z_2^n | x_4^n), \quad (3.128)$$

$$= I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) + I(x_{\Sigma}^n; z_2^n | x_4^n) - I(x_{\tau_1}^n; z_2^n | x_{\tau_2}^n), \quad (3.129)$$

$$= I(x_{\Sigma}^n; z_2^n | x_4^n) + I(x_{\tau_1}^n; H_{\tau_1}^n x_{\tau_1}^n + u_1^n) \quad (3.130)$$

$$- I(x_{\tau_1}^n; H_{\tau_1 2}^n x_{\tau_1}^n + u_2^n), \quad (3.131)$$

$$\leq I(x_{\Sigma}^n; z_2^n | x_4^n) \quad (3.132)$$

$$\leq n \log \left| I_N + \sum_{k \in \{4\}^c} H_{k2} Q_k H_{k2}^\dagger \right| + n\epsilon, \quad (3.133)$$

where equations (3.126), (3.128), (3.129) and (3.130) are obtained via basic properties of information theory as before. The first and second inequalities of equation (3.132) and (3.133) follows from Lemma 4 and 1, respectively.

Now, dividing both sides of equation (3.124) and (3.133) by  $n$ , taking  $n \rightarrow \infty$  and subsequently combining them we have

$$R_{\{4\}^c} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{4\}^c} H_{kj} Q_k H_{kj}^\dagger \right|, \quad (3.134)$$

Using a similar approach, the following upper bounds can also be proved:

For the case of  $R_1 + R_3 + R_4$ , Using the Fano's inequality we have;

$$nR_{\{2\}^c} \leq I(x_{\tau_2}^n; z_2^n) + I(x_1^n; z_1^n) + n\epsilon, \quad (3.135)$$



Next, we provide the side information  $x_2^n$  to both  $R_1$  and  $R_2$  and  $x_1^n$  to only  $R_2$  to obtain the following

$$nR_{\{2\}^c} \leq n \log \left| I_N + \sum_{k \in \{2\}^c} H_{k1} Q_k H_{k1}^\dagger \right| + n\epsilon, \quad (3.136)$$

Now, we provide the side information  $x_2^n$  to both  $R_1$  and  $R_2$  and  $x_{\tau_2}^n$  to only  $R_1$  to obtain the following

$$nR_{\{2\}^c} \leq n \log \left| I_N + \sum_{k \in \{2\}^c} H_{k2} Q_k H_{k2}^\dagger \right| + n\epsilon, \quad (3.137)$$

Now, dividing both sides of equation (3.136) and (3.137) by  $n$ , taking  $n \rightarrow \infty$  and subsequently combining them we have

$$R_{\{2\}^c} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{2\}^c} H_{kj} Q_k H_{kj}^\dagger \right|. \quad (3.138)$$

For the case of  $R_1 + R_2 + R_4$ , Using the Fano's inequality we have;

$$nR_{\{3\}^c} \leq I(x_{\tau_1}^n; z_1^n) + I(x_4^n; z_2^n) + n\epsilon, \quad (3.139)$$

Next, we provide the side information  $x_3^n$  to both  $R_1$  and  $R_2$  and  $x_{\tau_1}^n$  to only  $R_2$  to obtain the following

$$nR_{\{3\}^c} \leq n \log \left| I_N + \sum_{k \in \{3\}^c} H_{k1} Q_k H_{k1}^\dagger \right| + n\epsilon, \quad (3.140)$$

Now, we provide the side information  $x_3^n$  to both  $R_1$  and  $R_2$  and  $x_4^n$  to only  $R_1$  to

obtain the following

$$nR_{\{3\}^c} \leq n \log \left| I_N + \sum_{k \in \{3\}^c} H_{k2} Q_k H_{k2}^\dagger \right| + n\epsilon, \quad (3.141)$$

Now, dividing both sides of equation (3.140) and (3.141) by  $n$ , taking  $n \rightarrow \infty$  and subsequently combining them we have

$$R_{\{3\}^c} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{3\}^c} H_{kj} Q_k H_{kj}^\dagger \right|, \quad (3.142)$$

For the case of  $R_2 + R_3 + R_4$ , Using the Fano's inequality we have;

$$nR_{\{1\}^c} \leq I(x_{\tau_2}^n; z_1^n) + I(x_2^n; z_2^n) + n\epsilon, \quad (3.143)$$

Next, we provide the side information  $x_1^n$  to both  $R_1$  and  $R_2$  and  $x_2^n$  to only  $R_2$  to obtain the following

$$nR_{\{1\}^c} \leq n \log \left| I_N + \sum_{k \in \{1\}^c} H_{k1} Q_k H_{k1}^\dagger \right| + n\epsilon, \quad (3.144)$$

Now, we provide the side information  $x_1^n$  to both  $R_1$  and  $R_2$  and  $x_{\tau_2}^n$  to only  $R_1$  to obtain the following

$$nR_{\{1\}^c} \leq n \log \left| I_N + \sum_{k \in \{1\}^c} H_{k2} Q_k H_{k2}^\dagger \right| + n\epsilon, \quad (3.145)$$

Now, dividing both sides of equation (3.144) and (3.145) by  $n$ , taking  $n \rightarrow \infty$  and

subsequently combining them we have

$$R_{\{1\}^c} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{1\}^c} H_{kj} Q_k H_{kj}^\dagger \right|, \quad (3.146)$$

There are six possible bounds on sum of two rates; we divide these into two categories depending on whether both the rates belong to the same MAC or not, i.e., 1) bounds on  $(R_1 + R_2)$  and  $(R_3 + R_4)$  and 2) the remaining four bounds. First let us consider the bound on  $(R_1 + R_2)$ ; From Fano's inequality we have,

$$nR_{\{1,2\}} \leq I(x_{\tau_1}^n; z_1^n) + n\epsilon \quad (3.147)$$

$$\leq I(x_{\tau_1}^n; z_1^n, x_{\tau_2}^n) + n\epsilon \quad (3.148)$$

$$= I(x_{\tau_1}^n; x_{\tau_2}^n) + I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) + n\epsilon \quad (3.149)$$

$$= I(x_{\tau_1}^n; z_1^n | x_{\tau_2}^n) + n\epsilon \quad (3.150)$$

$$\leq n \log \left| I_N + \sum_{k \in \{1,2\}} H_{k1} Q_k H_{k1}^\dagger \right| + n\epsilon, \quad (3.151)$$

where equation (3.151) follows from Lemma 1.

Further, from the definition of ordered strong interference channel we have,

$$H_{\tau_1 1}^\dagger H_{\tau_1 1} \preceq H_{\tau_1 2}^\dagger H_{\tau_1 2}; \quad (3.152)$$

$$\text{or, } (Q_{12}^{\frac{1}{2}})^\dagger H_{\tau_1 1}^\dagger H_{\tau_1 1} Q_{12}^{\frac{1}{2}} \preceq (Q_{12}^{\frac{1}{2}})^\dagger H_{\tau_1 2}^\dagger H_{\tau_1 2} Q_{12}^{\frac{1}{2}}; \quad (3.153)$$

$$\text{or, } \log \left| I + (Q_{12}^{\frac{1}{2}})^\dagger H_{\tau_1 1}^\dagger H_{\tau_1 1} Q_{12}^{\frac{1}{2}} \right| \leq \log \left| I + (Q_{12}^{\frac{1}{2}})^\dagger H_{\tau_1 2}^\dagger H_{\tau_1 2} Q_{12}^{\frac{1}{2}} \right|; \quad (3.154)$$

$$\text{or, } \log \left| I_N + H_{\tau_1 1} Q_{12} H_{\tau_1 1}^\dagger \right| \leq \log \left| I_N + H_{\tau_1 2} Q_{12} H_{\tau_1 2}^\dagger \right|; \quad (3.155)$$

$$\text{or, } \log \left| I_N + \sum_{k \in \{1,2\}} H_{k1} Q_k H_{k1}^\dagger \right| \leq \log \left| I_N + \sum_{k \in \{1,2\}} H_{k2} Q_k H_{k2}^\dagger \right|, \quad (3.156)$$

where (3.153) follows from definition of partial order, (3.154) from the fact that

$\log \det(\cdot)$  is a monotonically increasing function in the cone of positive semi definite matrices and (3.155) follows from the following  $\log \det(\cdot)$  identity:

$$\log |I + AB| = \log |I + BA|, \text{ and } Q_{12} = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \quad (3.157)$$

is a positive semi-definite (psd) matrix and therefore, so is  $Q_{12}^{\frac{1}{2}}$ , i.e.,  $(Q_{12}^{\frac{1}{2}})^{\dagger} = Q_{12}^{\frac{1}{2}}$ . Equation (3.156) follows from the notational definition of  $H_{\tau_1 j}$ , i.e.,  $H_{\tau_1 j} = [H_{1j} \ H_{2j}]$ .

Now, dividing equation (3.151) by  $n$  and taking the limit of  $n \rightarrow \infty$ , and combining the resulting equation with (3.156) we get

$$R_{\{1,2\}} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{1,2\}} H_{kj} Q_k H_{kj}^{\dagger} \right|, \quad (3.158)$$

Similarly, we can prove that

$$R_{\{3,4\}} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{3,4\}} H_{kj} Q_k H_{kj}^{\dagger} \right|, \quad (3.159)$$

Next we turn to the second set of bounds. Say the following bound,

$$nR_{\{1,4\}} \leq I(x_1^n; z_1^n) + I(x_4^n; z_2^n) + n\epsilon \quad (3.160)$$

Providing, the side information  $x_2^n, x_3^n$  to  $R_1$  and  $x_2^n, x_1^n, x_3^n$  to  $R_2$  and following similar

technique as before we get,

$$nR_{\{1,4\}} \leq I(x_{\Sigma}^n; z_1^n | x_2^n, x_3^n) + I(x_4^n; H_{42}^n x_4^n + u_2^n) \quad (3.161)$$

$$- I(x_4^n; H_{41}^n x_4^n + u_1^n) \quad (3.162)$$

$$\leq I(x_{\Sigma}^n; z_1^n | x_2^n, x_3^n), \quad (3.163)$$

$$\leq n \log \left| I_N + \sum_{k \in \{1,4\}} H_{k1} Q_k H_{k1}^{\dagger} \right| + n\epsilon, \quad (3.164)$$

where in equation (3.163) we have used Lemma 4 along with equation (3.6) with  $k = 2$  and in (3.165) Lemma 1.

From equation (3.160) we can also have

$$nR_{\{1,4\}} \leq n \log \left| I_N + \sum_{k \in \{1,4\}} H_{k2} Q_k H_{k2}^{\dagger} \right| + n\epsilon, \quad (3.165)$$

Combining (3.163) and (3.165) we have,

$$R_{\{1,4\}} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \{1,4\}} H_{kj} Q_k H_{kj}^{\dagger} \right| + n\epsilon, \quad (3.166)$$

Similarly we can also get

$$R_{\beta} \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \beta} H_{kj} Q_k H_{kj}^{\dagger} \right|, \quad (3.167)$$

for  $\beta \in \{\{1, 3\}, \{2, 3\}, \{2, 4\}\}$ . Using a similar method as in the derivation of (3.151) and (3.156) we get the following

$$R_k \leq \min_{1 \leq j \leq 2} \log \left| I_N + H_{kj} Q_k H_{kj}^{\dagger} \right|, \quad \forall k \in \Sigma. \quad (3.168)$$

Finally, if we define the set of rate tuples satisfying all of the above constraints by  $C_u$ , i.e.,

$$C_u = \left\{ \bar{R} : \bigcap_{\{\beta \subseteq \Sigma\}} R_\beta \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \beta} H_{kj} Q_k H_{kj}^\dagger \right| \right\}, \quad (3.169)$$

then clearly,

$$C_{\text{IMAC}} \subseteq C_u. \quad (3.170)$$

Hence we have derived all the possible outer rate bounds in this section. Then we define a set  $C_u$  which is a set of rate tuples satisfying all the above constraints, which essentially means that in a four dimensional planar structure, each rate bound will form a plane according to its constraint and intersection of all such planes will form a closed structure which will be  $C_u$ . But it is obvious that  $C_{\text{IMAC}}$  will be subset of  $C_u$ . Now in next section we will work on the achievability. Here our task is to show that independent Gaussian coding at each transmitter and joint decoding of the messages at the receivers can achieve all the rate pairs that satisfy all the aforementioned constraints.

### 3.2.2. Achievability

As we are considering a cellular structure here, Cell-1 has two users, Transmitter 1 and 2 and one Base Station, receiver 1 with which the users intend to communicate. Similarly Cell-2 has two users, Transmitter 3 and 4 and a Base Station, receiver 2 with which they intend to communicate. Hence they form direct links of communication within the cell and hence form a MAC channel in the individual cell. Now let us define  $R_{\text{MAC}-R_1}$  as the rate region of MAC channel formed with receiver 1 in Cell-1. Similarly  $R_{\text{MAC}-R_2}$  is the rate region of MAC channel formed with receiver 2 in Cell-2.

Now we will use the results of Discrete Memoryless Channel (DMC) which we discussed in section 1.5.1, for any given distribution on  $x_\Sigma$  satisfying  $p(x_\Sigma) = \{p(x_1)p(x_2)p(x_3)p(x_4)\}$ ,

$$\mathcal{R}_{\text{MAC}-R_j} = \left\{ \bar{R} : \bigcap_{\alpha \subseteq \Sigma} (R_\alpha \leq I(x_\alpha; z_j | x_{\alpha^c})) \right\}, \quad (3.171)$$

represents an achievable rate region for the 4-user MAC formed by all 4 transmitters and  $\text{Rx}_j$ , for  $1 \leq j \leq 2$ , where  $x_\alpha = [x_{i_1}^\dagger, x_{i_2}^\dagger, \dots]^\dagger$ ,  $\alpha = \{i_1, i_2, \dots\}$ .

Now we define a Compound MAC region which is basically intersection of rate regions of two MACs. This is important because this intersection will be the region where all the four transmitter are communicating. It is defined as follows:

$$\mathcal{R}_{\text{CMAC}} = \mathcal{R}_{\text{MAC}-R_1} \cap \mathcal{R}_{\text{MAC}-R_2}. \quad (3.172)$$

Clearly, if  $\bar{R} \in \mathcal{R}_{\text{CMAC}}$ , then  $\text{Tx}_k$  can communicate at a rate of  $R_k$  to both  $\text{Rx}_1$  and  $\text{Rx}_2$  with probability of error going to zero  $\forall k \in \Sigma$ . In particular,  $\text{Tx}_1$  and  $\text{Tx}_2$  can communicate to  $\text{Rx}_1$  at rates  $(R_1, R_2)$  and  $\text{Tx}_3$  and  $\text{Tx}_4$  can communicate to  $\text{Rx}_2$  at rates  $(R_3, R_4)$ , respectively, which in turn imply that  $\bar{R} \in \mathcal{R}_{\text{CMAC}}$ . Since this is true for all rate tuples  $\bar{R} \in \mathcal{R}_{\text{CMAC}}$  we have

$$\mathcal{R}_{\text{CMAC}} \subseteq C_{\text{IMAC}}. \quad (3.173)$$

In what follows, we evaluate the mutual information terms of (3.171) in order to derive an explicit expression for  $\mathcal{R}_{\text{CMAC}}$ , which serves as an achievable rate region for the IMAC by equation (3.173). Incorporating the input covariance constraint of (3.2), we assume that each transmitter uses i.i.d. Gaussian code books, i.e., the

inputs are distributed as

$$x_k \sim \mathcal{CN}(0, Q_k), \quad 1 \leq k \leq 4. \quad (3.174)$$

Computation of  $I(x_\alpha; z_j | x_{\alpha^c})$ , for any  $\alpha \subseteq \Sigma$  with such distributions is straight forward and can be evaluated as

$$I(x_\alpha; z_j | x_{\alpha^c}) = \log \left| I_N + \sum_{k \in \alpha} H_{kj} Q_k H_{kj}^\dagger \right|, \quad (3.175)$$

for all  $\alpha \subseteq \Sigma$  and  $1 \leq j \leq 2$ . Substituting this expression into equation (3.171) we obtain the achievable rate region for the IMAC as follows:

$$\mathcal{R}_{\text{CMAC}} = \left\{ \bar{R} : \bigcap_{\alpha \subseteq \Sigma} R_\alpha \leq \min_{1 \leq j \leq 2} \log \left| I_N + \sum_{k \in \alpha} H_{kj} Q_k H_{kj}^\dagger \right| \right\}. \quad (3.176)$$

The converse and the achievability results match in this case and hence we can say that this is the desired capacity region of the strongly ordered IMAC and the capacity is the intersection of all the sum rate constraints derived before. Practically, the capacity region will be the intersection of planes formed by the rate bounds, in the four dimensional co-ordinate system.



## CHAPTER 4. PRACTICAL APPLICATIONS OF OUR RESULTS

The practical application of our results is discussed in this section. As we know that, interference is a fundamental phenomenon in wireless communication networks, uncoordinated interference reduces wireless network throughput. As a result, it is essential to understand and manage interference to achieve the highest network performance. Conventional approaches to deal with interference are

1. Avoiding interference through orthogonalization of the shared time/frequency resource,
2. Treating other transmitters signals as noise or decoding interference.

These strategies have been studied extensively and adapted to contemporary wireless systems such as cellular and wireless local area networks (WLANs). Although these approaches control interference without system overhead, it turns out that they are not the optimal in most network configurations, except in certain special cases. For example consider the  $K$ -user interference channel where  $K$  transmitters send data to their corresponding receivers in a shared wireless medium. When the interference power is of the order of the power of the signal of interest, the traditional interference management approaches have resulted in (at best) achieving the same data rate order as the rate of a single communication link because of their inefficient usage of the spectrum.

It can be shown that the information theoretic results can give the better results compared to conventional approaches. Now if we consider the method of avoiding interference through orthogonalization scheme i.e. Orthogonal multiple access schemes, such as OFDM. This tradeoff can be achieved by varying the number of sub-carriers allocated to each user. Then the capacity region  $\mathcal{C}$  characterizes the optimal tradeoff achievable by any multiple access scheme.

Now it is hard to practically characterize the capacity region derived in section 3.2. To compare the results of two methods:

1. Method-1: Information theoretic rate bounds derived in section 3.2, and
2. Method-2: Rate bounds achieved by Orthogonal multiple access method.

we will derive the capacity of two Cells differently with two different methods and then compare them individually.

Method-1 results are derived in chapter 3. Now we will derive the results for Method-2. Consider an orthogonal scheme that allocates a fraction  $\alpha$  of the degrees of freedom to cell-1 and the rest,  $1 - \alpha$ , to cell-2. It is irrelevant for the capacity analysis whether the partitioning is across frequency or across time, since the power constraint is on the average across the degrees of freedom. If the received power of user-k is  $P_k$ , the amount of received energy for user-1 in cell-1 is  $\frac{P_1}{\alpha}$  joules per degree of freedom, similarly for other users. The maximum rate user-1 can achieve over the total bandwidth  $W$  is

$$R_1 = \alpha W \log \left( 1 + \frac{P_1}{\alpha N_0} \right) \text{ bits/s} \quad (4.1)$$

Similarly, the maximum rate user-2 can achieve is

$$R_2 = \alpha W \log \left( 1 + \frac{P_2}{\alpha N_0} \right) \text{ bits/s} \quad (4.2)$$

and sum rate will be

$$R_1 + R_2 = \alpha W \log \left( 1 + \frac{P_1 + P_2}{\alpha N_0} \right) \text{ bits/s} \quad (4.3)$$

Similarly for Cell-2

$$R_3 = (1 - \alpha)W \log \left( 1 + \frac{P_3}{(1 - \alpha)N_0} \right) \text{ bits/s} \quad (4.4)$$

$$R_4 = (1 - \alpha)W \log \left( 1 + \frac{P_4}{(1 - \alpha)N_0} \right) \text{ bits/s} \quad (4.5)$$

$$R_3 + R_4 = (1 - \alpha)W \log \left( 1 + \frac{P_3 + P_4}{(1 - \alpha)N_0} \right) \text{ bits/s} \quad (4.6)$$

Here,  $W$  is the bandwidth and  $\alpha$  is a constant that varies from 0 to 1. Varying  $\alpha$  from 0 to 1 yields all the rate pairs achieved by orthogonal schemes.

In the simulations, We have considered  $W = 20Hz$ . For Method-1, We have derived all the rate bounds using equations from section 3.2. The matlab code for all rate bounds with Method-1 and its output is in Appendix. A.

For Method-2, first we derive  $R_1, R_2, R_3, R_4, R_1 + R_2$  and  $R_3 + R_4$ , then by adding the rates with other combinations we can derive all the bounds. The matlab code for Method-2 and its output is in Appendix. B.

The easy way to compare the results is by comparing the graphs of 2 user links from Cell-1 and Cell-2. The graph for the Cell-1 formed by equations derived by rate tuple  $(R_1, R_2)$  from Method-1 and Method-2. Figure 14 and Figure 15 shows these graphs for Cell-1 and Cell-2 respectively.

As it is evident from Figures 14 and 15, the graphs from Method-1 have greater capacity region than the graphs from Method-2. This shows that the information theoretic results give better results than orthogonal multiple access method which is practically implemented. We are working on getting the four dimensional graphs for better comparison results.

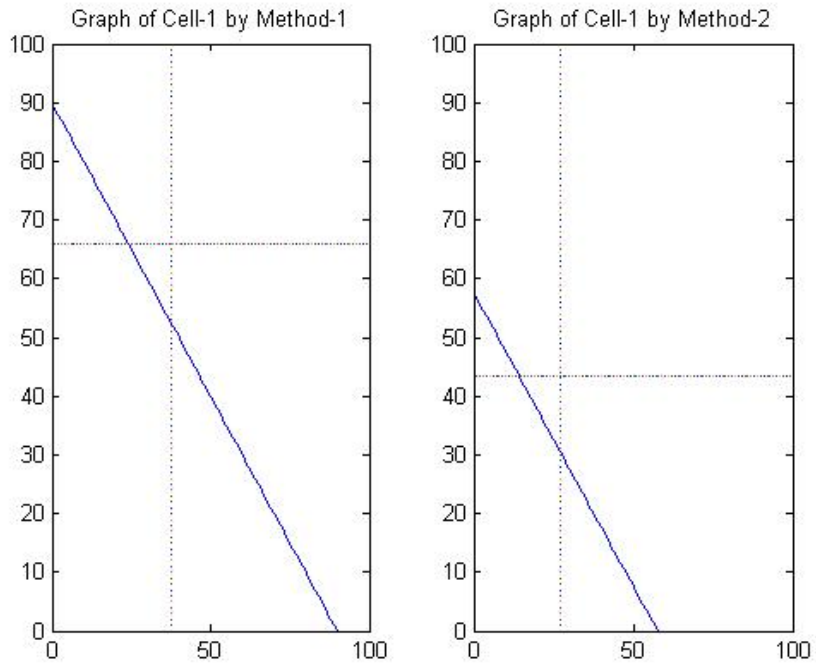


Figure 14. The comparison of graphs by two methods in Cell-1

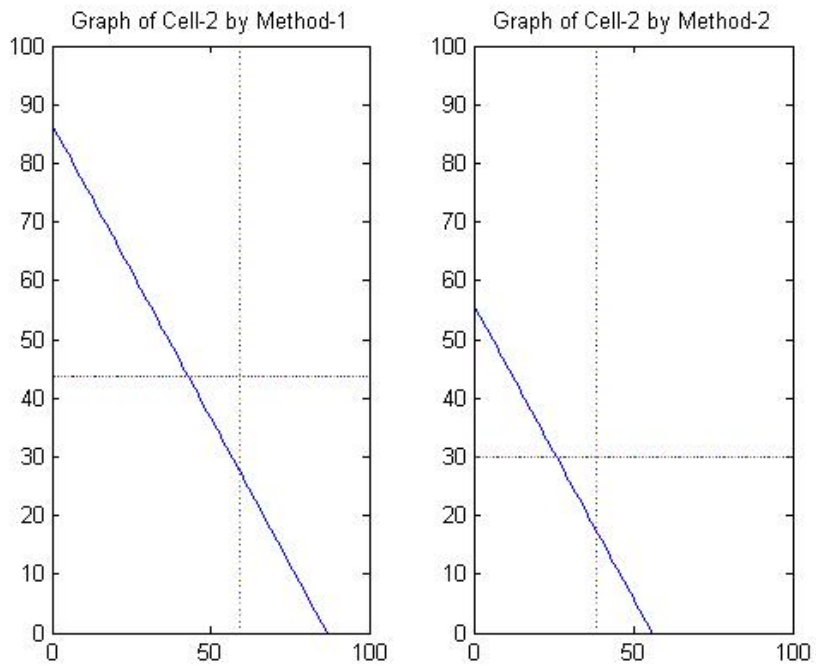


Figure 15. The comparison of graphs by two methods in Cell-2

## CHAPTER 5. CONCLUSION

We characterize the capacity region of the class of *strongly ordered* IMAC, with arbitrary number of antennas at each transmitter and arbitrary but same number of antennas at both the receivers. Our results prove that while characterizing the capacity region the two tradeoffs will turn out to be unique, i.e. Interference channel problem and the MAC channel problem are the same. Hence there is a homogeneity between the two types of tradeoffs. The results of two user MIMO strong interference channel is the special case of four user strongly ordered IMAC.

In this work, we have tried to show that information theoretic result can provide optimal results than the methods used presently in practical applications.

Characterizing the capacity region for non-strong channel matrices can be an interesting direction of future research.

## REFERENCES

- [1] J.G. Proakis and M. Salehi, *Fundamentals of Communication Systems*, Pearson Prentice Hall, 2005.
- [2] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, 2nd edition, John Wiley and Sons, New Jersey, 2006.
- [3] C. E. Shannon, "Two-way communication channel," *Proc. of 4th Berkeley symp. Mathematical statistics and probability, Berkeley, CA*, vol.1, Mar, 1961, pp. 611-644.
- [4] A.B. Carleial, "Interference Channels," *IEEE Trans. on Inform. Th.*, vol. 24, Jan, 1978, pp. 60-70.
- [5] A. B. Carleial, "A case where interference does not reduce the capacity," *IEEE Trans. on Inform. Th.*, vol. 21, Mar, 1979, pp. 228-231.
- [6] H. Sato, "The capacity of Gaussian interference channel under strong interference," *IEEE Trans. on Inform. Th.*, vol. 48, Sep, 2002, pp. 2450-2454.
- [7] T. S. Han and K. Kobayashi, "A new achievable region for the interference channel," *IEEE Trans. on Inform. Th.*, vol. 27, Jan, 1981, pp. 49-60.
- [8] M. H. M. Costa and A. A. El Gamal, "The Capacity Region of the Discrete Memoryless Interference Channel with Strong Interference," *IEEE Trans. on Inform. Th.*, vol. 33, Sep, 1987, pp. 710-711.
- [9] H. Sato, "On degraded Gaussian Two-user channels," *IEEE Trans. on Inform. Th.*, vol. 24, Nov., 1978, pp. 637-640.

- [10] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels," *IEEE Trans. on Inform. Th.*, vol. 25, Mar, 1979, pp. 228-231.
- [11] A. A. El Gamal and M. H. M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. on Inform. Th.*, vol. 28, Mar, 1982, pp. 343-346.
- [12] S. Karmakar and M.K. Varanasi, "Capacity region of a class of strong MIMO IC," *Information Sciences and Systems (CISS), 2014 48th Annual Conference*, March, 2014, pp. 1-6.
- [13] V. S. Annapureddy and V. V. Veeravalli, "Sum Capacity of the Gaussian Interference Channels in the low interference regime," *IEEE Trans. on Inform. Th.*, vol. 55, Jul, 2010, pp. 3032-3050.
- [14] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. on Inform. Th.*, vol. 55, Feb, 2009, pp. 620-643.
- [15] X. Shang and G. Kramer and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. on Inform. Th.*, vol. 55, Feb, 2007, pp. 689-699.
- [16] S. Vishwanath and S. A. Jafar, "On the capacity of vector Gaussian interference channels," *Proc. of Inform. Th. Workshop*, San Antonio, TX, Oct, 2004, pp 365 - 369.
- [17] X. Shang and B. Chen and G. Kramer and H. V. Poor, "Capacity regions and sum-rate capacities of vector Gaussian interference channels," *IEEE Trans. on Inform. Th.*, vol. 56, no. 10, Oct, 2010, pp. 5030 - 5044.

- [18] R.H. Etkin, D.N.C. Tse and Hua Wang, A.J. Goldsmith and P. P. Varaiya, “Gaussian Interference Channel Capacity to Within One Bit,” *IEEE Trans. on Inform. Th.*, vol. 54, Dec, 2008, pp. 5534-5562.
- [19] S. Karmakar and M. K. Varanasi, “The Capacity region of the MIMO interference channel and its reciprocity to within a constant gap,” *IEEE Trans. on Inform. Th.*, vol. 59, Aug, 2013, pp. 4781-4797.
- [20] A. Chaaban, A. Sezgin, B. Bandemer and A. Paulraj, *IEEE Trans. on Inform. Th.*, Springer Berlin Heidelberg, 2011, pp. 87-96.
- [21] S. Karmakar and N. Modi, “Capacity region of two interfering MIMO Multiple Access channels with strong interference,” *Information Sciences and Systems (CISS), 2015 49th Annual Conference on*, Mar, 2015, pp. 1-6.
- [22] S. Karmakar and M.K. Varanasi, “Capacity region of a class of strong MIMO IC,” *Proc. of 48th Annual Conf. on Inf. Sc. and Syst. (CISS), Princeton, NJ, USA*, Mar, 2014.
- [23] R. A. Horn and C. R. Jhonson, *Matrix analysis*, vol. 1, Cambridge Univ. Press, 1990.
- [24] Meyer and D. Carl, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics, 2000.
- [25] T. Liu and P. Viswanath, “An Extremal Inequality Motivated by Multiterminal Information-Theoretic Problems,” *IEEE Trans. on Inform. Th.*, vol. 53, May, 2007, 1839-1851.
- [26] H. Weingarten, Y. Steinberg and S. Shamai, “The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel,” *IEEE Trans. on Inform. Th.*, vol. 52, Sep, 2006, pp. 3936-3964.



## APPENDIX A

The matlab code for Method-1 and its output is as follows:

```
%R1+R2+R3+R4
R12341=log(det(I+(H11*Q*H11.')+(H21*Q*H21.')+(H31*Q*H31.')+(H41*Q*H41.')));
R12342=log(det(I+(H12*Q*H12.')+(H22*Q*H22.')+(H32*Q*H32.')+(H42*Q*H42.')));
R7=min(R12341,R12342)*W;

%R1+R2+R3
R1231=log(det(I+(H11*Q*H11.')+(H21*Q*H21.')+(H31*Q*H31.')));
R1232=log(det(I+(H12*Q*H12.')+(H22*Q*H22.')+(H32*Q*H32.')));
R14=min(R1231,R1232)*W;

%R1+R3+R4
R1341=log(det(I+(H11*Q*H11.')+(H31*Q*H31.')+(H41*Q*H41.')));
R1342=log(det(I+(H12*Q*H12.')+(H32*Q*H32.')+(H42*Q*H42.')));
R12=min(R1341,R1342)*W;

%R1+R2+R4
R1241=log(det(I+(H11*Q*H11.')+(H21*Q*H21.')+(H41*Q*H41.')));
R1242=log(det(I+(H12*Q*H12.')+(H22*Q*H22.')+(H42*Q*H42.')));
R13=min(R1241,R1242)*W;

%R2+R3+R4
R2341=log(det(I+(H21*Q*H21.')+(H31*Q*H31.')+(H41*Q*H41.')));
R2342=log(det(I+(H22*Q*H22.')+(H32*Q*H32.')+(H42*Q*H42.')));
R15=min(R2341,R2342)*W;

%R1+R2
R121=log(det(I+(H11*Q*H11.')+(H21*Q*H21.')));
R122=log(det(I+(H12*Q*H12.')+(H22*Q*H22.')));
R5=min(R121,R122)*W;

%R3+R4
R341=log(det(I+(H31*Q*H31.')+(H41*Q*H41.')));
R342=log(det(I+(H32*Q*H32.')+(H42*Q*H42.')));
```

```

R6=min(R341,R342)*W;

%R1+R3
R131=log(det(I+(H11*Q*H11.')+(H31*Q*H31.')));
R132=log(det(I+(H12*Q*H12.')+(H32*Q*H32.')));
R8=min(R131,R132)*W;

%R1+R4
R141=log(det(I+(H11*Q*H11.')+(H41*Q*H41.')));
R142=log(det(I+(H12*Q*H12.')+(H42*Q*H42.')));
R9=min(R1241,R1242)*W;

%R2+R3
R231=log(det(I+(H21*Q*H21.')+(H31*Q*H31.')));
R232=log(det(I+(H22*Q*H22.')+(H32*Q*H32.')));
R10=min(R231,R232)*W;

%R2+R4
R241=log(det(I+(H21*Q*H21.')+(H41*Q*H41.')));
R242=log(det(I+(H22*Q*H22.')+(H42*Q*H42.')));
R11=min(R241,R242)*W;

%R1
R011=log(det(I+(H11*Q*H11.')));
R012=log(det(I+(H12*Q*H12.')));
R1=min(R011,R012)*W

%R2
R21=log(det(I+(H21*Q*H21.')));
R22=log(det(I+(H22*Q*H22.')));
R2=min(R21,R22)*W;

%R3
R31=log(det(I+(H31*Q*H31.')));
R32=log(det(I+(H32*Q*H32.')));
R3=min(R31,R32)*W;

%R4
R41=log(det(I+(H41*Q*H41.')));
R42=log(det(I+(H42*Q*H42.')));

```

```
R4=min(R41,R42)*W;
```

```
N(1,:)= [R1 R2 R3 R4 R5 R6 R7 R8 R9 R10 R11 R12 R13 R14 R15]
```

In the above code, ' $N$ ' is the output where the order of the rate bounds is as follows:

$$[R_1 \ R_2 \ R_3 \ R_4 \ R_1 + R_2 \ R_3 + R_4 \ R_1 + R_2 + R_3 + R_4 \ R_1 + R_3 \ R_1 + R_4 \ R_2 + R_3 \ R_2 + R_4 \ R_1 + R_3 + R_4 \ R_1 + R_2 + R_4 \ R_1 + R_2 + R_3 \ R_2 + R_3 + R_4]$$

$Q$  is the power constraint matrix. For this simulation I have considered two antennas at transmitter and the receiver. Hence  $I$  is a  $2 \times 2$  matrix and the channel matrices are randomly generated  $2 \times 2$  matrices.

The output for particular set of random channel matrices is as follows:

```
N = Columns 1 through 10
```

```
37.3412    66.0080    58.9239    43.7388    90.0089    86.7171    ...  
116.7628    82.5857    103.5614    101.5780
```

```
Columns 11 through 15
```

```
96.7876    93.0552    103.5614    108.3696    112.2398
```

## APPENDIX B

The matlab code for Method-2 and its output is as follows:

```
alpha=0.5
%Cell-1
%R1
C011=alpha*W*log(det(I+(H11*Q*H11.').)/alpha));
C012=alpha*W*log(det(I+(H12*Q*H12.').)/alpha));
C1=min(C011,C012);
%R2
C21=alpha*W*log(det(I+(H21*Q*H21.').)/alpha));
C22=alpha*W*log(det(I+(H22*Q*H22.').)/alpha));
C2=min(C21,C22);
%R1+R2
C51=alpha*W*log(det(I+((H11*Q*H11.').+(H21*Q*H21.')).)/alpha));
C52=alpha*W*log(det(I+((H11*Q*H11.').+(H21*Q*H21.')).)/alpha));
C5=min(C51,C52);
%Cell-2
%R3
C31=(1-alpha)*W*log(det(I+(H31*Q*H31.').)/(1-alpha));
C32=(1-alpha)*W*log(det(I+(H32*Q*H32.').)/(1-alpha));
C3=min(C31,C32);
%R4
C41=(1-alpha)*W*log(det(I+(H41*Q*H41.').)/(1-alpha));
C42=(1-alpha)*W*log(det(I+(H42*Q*H42.').)/(1-alpha));
C4=min(C41,C42);
%R3+R4
C61=(1-alpha)*W*log(det(I+((H31*Q*H31.').+(H41*Q*H41.')).)/(1-alpha));
C62=(1-alpha)*W*log(det(I+((H31*Q*H31.').+(H41*Q*H41.')).)/(1-alpha));
C6=min(C61,C62);
```

```

%R1+R2+R3+R4
C7=C5+C6;
%R1+R3
C8=C1+C3;
%R1+R4
C9=C1+C4;
%R2+R3
C10=C2+C3;
%R2+R4
C11=C2+C4;
%R1+R3+R4
C12=C1+C6;
%R1+R2+R4
C13=C5+C4;
%R1+R2+R3
C14=C5+C3;
%R2+R3+R4
C15=C2+C6;
M(1, :)= [C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15]

```

The output for particular set of random channel matrices is as follows, for  $\alpha = 0.5$ :

```

M = Columns 1 through 10

    27.1488    43.3895    38.7082    30.0800    57.6710    55.9576    ...
    113.6286    65.8570    57.2288    82.0977

Columns 11 through 15

    73.4695    83.1064    87.7510    96.3791    99.3472

```