ANALYSIS OF SECRECY IN MULTI-USER WIRELESS NETWORK

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

By

Anirban Ghosh

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Electrical and Computer Enggineering

March 2018

Fargo, North Dakota

NORTH DAKOTA STATE UNIVERSITY

Graduate School

Title

ANALYSIS OF SECRECY IN MULTI-USER WIRELESS NETWORK

By

Anirban Ghosh

The supervisory committee certifies that this dissertation complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Sanjay Karmakar
Chair

Dr. Ivan T. Lima Jr.

Dr. Dilpreet Bajwa

Dr. Indrajit Sengupta (GSR)

Approved:

27th March 2018
Date

Dr. Benjamin Bratein
Department Chair

ABSTRACT

We consider an Ergodic Fading Broadcast Channel with one Legitimate receiver and one Eavesdropper (BCoLoE) having arbitrary fading statistics, where the instantaneous Channel State Information (CSI) are known only at the receivers (CSIR). The secrecy capacity of this channel is characterized within 11 bits irrespective of fading statistics and Signal-to-Noise Ratios (SNRs). This is achieved by deriving a new upper bound to the secrecy capacity of the channel and two new lower bounds. The upper bound is derived by approximating Complementary Cumulative Distribution Functions (CCDFs) of the two links by corresponding *staircase functions*. The smaller lower bound, although looser, has a form which can be *analytically* compared with the upper bound and facilitates the approximate secrecy capacity characterization. It is proved that, the so called *Binary Expansion Signaling with Reverse Stripping* (BES-RS) scheme can achieve a secrecy rate larger than both these lower bounds with the help of numerical computation for several BCoLoEs with practical fading statistics.

We further characterize the secrecy capacity of a class of 2-user binary fading interference channel (BFIC) and 2-user layered fading interference channel (LFIC), under the same assumptions as for the wiretap channel. The secrecy capacity region for a *very weak* BFIC turns out to be quadrangular while for LFIC it is polygonal. We explicitly characterize the corner points in both the cases. The converse in either case is proved by dividing the set of upper bounds into two carefully chosen regions depending on the values of ω - the weighting factor of the weighted sum bounds. In case of LFIC each of the regions are also shown to be piece-wise linear. The achievability on the other hand is proved by using capacity optimal code for a layered erasure wiretap channel at both the transmitters and *treating interference as erasure* while decoding the signals at the receivers. In addition, the achievability of the layered case also involves proper assignment of the layers to the two transmitters based on some constraints. We also prove the secrecy capacity of *strong* BFIC and LFIC as zero.

ACKNOWLEDGEMENTS

In the last few years I have been granted the extraordinary privilege to pursue a very interesting and relevant topic of interest, in the field of wireless communications, in a serene and intellectual atmoshpere. In the course of doing so, I have learnt interesting and astonishing facts that I was completely ignorant of. It has been but a very joyful ride forward in my life and a great learning lesson.The work presented in this thesis would not have been possible without the influence, support and encouragement of a number of individuals, mentors and friends with whom I have had the great fortune of interacting and from whom I have had the privilege of learning a lot, not only about my research but also about life.

I would take this oppurtunity to first and foremost thank and express my gratitude to my adviser Dr. Sanjay Karmakar for his constant and generous support, guidance and understanding during my entire tenure. It was only because of his continuous encouragement and motivation that I decided to embark on this long journey. His thoughtful discussions, careful comments and criticism, and mind provoking questions have always helped me remain focussed and work towards my goal whenever I seemed to be drifting from them. I would also like to express my gratitude to Dr. Dilpreet Bajwa, Dr. Ivan T. Lima Jr. and Dr. Indranil Sengupta for agreeing to be a part of my supervisory comittee and constantly helping me whenever I needed some despite their busy schedule.

Due to paucity of space it might not be possible to mention about everybody who have influenced my life or research in some way or the other over the last few years but that in no way should undermine their contribution to say the least. However this thesis would be incomplete without mentioning a few of them. A special thanks to you Mr. Sanjay Patel -you have been like an elder brother and the only one that I could call family here in Fargo. Thank you to all my special friends - Arnab, Ram, Tiku, Amar, Fleming, Shreya, Sujata, Prakshit, Ninad and many more to whom I will always be indebted for their constant support and bearing with my tantrums. Finally, its time to thank my family - my in-laws to be, by beautiful fiancée and my mom and dad who have stood like a pillar of strength and confidence even when I doubted myself. I owe this to you all.

DEDICATION

This is for you *Maa* and *Bapi*.....

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The purpose of any form of communication is to transfer information. But what is information? From a system design perspective we might also be interested to know how much information we are transferring or if it is at all possible to quantify the amount of information that is being transferred and also how fast communication can take place. But before we go into more complicated questions the ones we want to address first are the most primitive ones - what is information and if at all it is possible to quantify it?

In that respect let us now consider a situation - Somebody tells you that tomorrow sun is going to rise in the east. Do you get any information from that? Think about it, you will realise the answer is *no* because it is something that you not only know about but also that it is certainly going to happen. However if somebody tells you that the sun might not come out tomorrow because it's going to be cloudy - is that information? It certainly is because it does not remain cloudy everday in general - so there is a 'cloud' of uncertainity about it being cloudy the next day. Hence intuitively it seems as if, if there is some uncertainity involved in a piece of message then it definitely contains some information. Let us try to get some more intuition about what is information with another example - this time we consider throwing of a fair coin. Do you know before tossing the coin what the output will be, you can just guess but you don't know it for sure. For a fair coin it has equal probability of landing a heads or a tails hence it is difficult to make a guess and hence it has more information (we will find that out shortly), however if it is biased to one side then at least you can make a better guess. Thus it further cements the fact that more the uncertainity about an event more is the information contained in it. From information transmission point of view, revealing some fact to a person who already knows it, is pointless. Hence, information is always accompanied by some amount of uncertainty to the event of interest.

The best way to model uncertainty and thus information is through Random Variables (RV). Information content of a random variable is related to the amount of uncertainty associated with that RV. But how are RV defined and do they have any classification? RV can either be discrete or continuous depending on the values they can take. We define a discrete RV by its Probability Mass Function (PMF), whereas a continuous RV is characterized by its Probability Distribution

Function (PDF). So, its not unreasonable to predict that the information content of a RV should be a function of the PMF or PDF of the RV. A set of such intuitive guidelines helps us to give a mathematical definition for information and to quantify it. In information theory the information contained in a RV is expressed in terms of *entropy*. It is very easy to get confused with entropy used in thermodynamics however although the underlying meaning of entropy in either case remains the same - its a measure of randomness of a system, it is used to define two completely different things in the two contexts. In the next few sections we first show how information is mathematically defined and quantified. Next some other definitions and some basic information theoretic results are presented to provide the setting for information theoretic secure communication. Most of these definitions and results are presented from [1].

1.1. Entropy

We mentioned earlier that the information contained in a random variable is called entropy and the information contained in a RV depends on the uncertainity of the RV. Hence entropy is the measure of uncertainity of a random variable. Let us now denote X as a random variable that can take values from the alphabet $\mathcal{X}$ and has PMF defined as $p(x) = P(X = x) \forall x \in \mathcal{X}$, i.e. the probability that X takes a value $x \in \mathcal{X}$ is $p(x)$. Since we know that probability can take values only in the range $[0, 1]$ hence $p(x) \in [0, 1]$. Thus entropy is defined as follows

Definition 1 *The entropy $H(X)$ of a discrete random variable X is defined by*

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x), \tag{1.1}$$

If the logarithm in 1.1 is to the base 2 then the resulting entropy is measured in *bits* however if it is to the base e then entropy is measured in units of *nats*. If however the logarithm has any other random base b then it is represented as $H_b(X)$. Let us once again get back to the example of coin toss. Let X represent the RV representing toss of a coin and let $P_X(x)$ be its PMF such that $P_X(heads) = p$ and $P_X(tails) = 1 - p$. Hence the entropy of this random variable will be

$$H(X) = -p \log p - (1 - p) \log (1 - p), \tag{1.2}$$

The figure 1.1 shows the variation of entropy with p. If $p = 1$ which essentially corresponds to one always getting heads as an outcome of a toss has zero entropy which intuitively makes sense if you always get heads then its a certain event and there is no uncertainity left in it and hence information content or entropy is zero. Similar explaination goes when $p = 0$ which corresponds to one always getting tails. However if $p = \frac{1}{2}$ then it is the most uncertain situation with one having no idea about the outcome of the toss and hence has maximum entropy of 1 bit. As p varies in between these extreme values entropy follows the graph of figure1.1. We will next show that mathematics supports the explaination that we have just provided and hence validates the nature of the graph in figure 1.1.

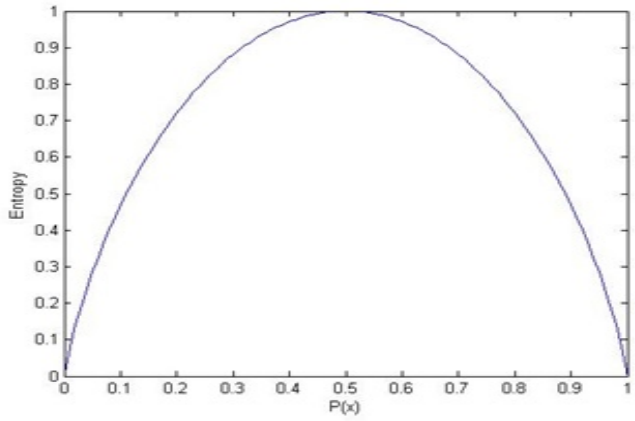

Figure 1.1. Variation of Entropy for a Binary RV.

When $p = 0$ or 1 the entropy can be calculated as follows

$$H(X) = -1 \log 1 - (1-1) \log (1-1) = 0 - 0 = 0, \tag{1.3}$$

where the second log term in equation 1.3 goes to zero because the rate of decrease of x is more than rate of decrease of $\log x$. Next we find the entropy for $p = \frac{1}{2}$,

$$H(X) = -\frac{1}{2} \log \frac{1}{2} - (1 - \frac{1}{2}) \log (1 - \frac{1}{2}) = \frac{1}{2} + \frac{1}{2} = 1, \tag{1.4}$$

Thus we have successfully found a way to measure and quantify information. However we know just to measure discrete RV what about continuous RV. The measurement of continuous RV is not too different. For continuous RV the entropy term is known as *differential entropy*. It is defined as

follows

Definition 2 *The differential entropy $h(X)$ of a continuous random variable X with PDF $f_X(x)$ is defined as*

$$h(X) = -\int_S f_X(x) \log f_X(x) dx, \tag{1.5}$$

where S is the support set of the random variable.

We will next define two more entropy terms *joint entropy* and *conditional entropy* which will be heavily used in our problem solving. When we have to define the entropy of a pair or more than one random variable together then it is known as joint entropy. The formal definition goes as follows

Definition 3 *The joint entropy $H(X,Y)$ of a pair of discrete random variables (X,Y) with a joint distribution is defined as*

$$H(X,Y) = -\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y) \log p(x,y), \tag{1.6}$$

We have a similar definition for joint entropy if the RVs X and Y are continuous where the summation signs are replaced by integration and the PMFs with PDFs. Next we define the conditional entropy of a random variable given another as the expected value of the entropies of the conditional distributions, averaged over the conditioning random variable.

Definition 4 *If $(X,Y) \sim p(x,y)$, the conditional entropy $H(Y|X)$ is defined as*

$$\begin{aligned} H(Y|X) &= \sum_{x\in\mathcal{X}} p(x) H(Y|X=x), \\ &= -\sum_{x\in\mathcal{X}} p(x) \sum_{y\in\mathcal{Y}} p(y|x) \log p(y|x), \end{aligned} \tag{1.7}$$

One of the important properties of conditional entropy which will be of immense use in our derivations later is that conditioning *reduces* entropy. Next we provide the chain rule of entropy

Definition 5 *Let $X_1, X_2, \cdots, X_n$ be drawn according to $p(x_1, x_2, \cdots, x_n)$. Then*

$$H(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} H(X_i | X_{i-1}, \cdots, X_n), \tag{1.8}$$

Next we define the relationship between the entropy, joint entropy and conditional entropy by using the chain rule

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y), \tag{1.9}$$

The naturalness of the definition of joint entropy and conditional entropy is exhibited by the fact that the entropy of a pair of random variables is the entropy of one plus the conditional entropy of the other. The concept of entropy plays the central role in information theory. Most of the other information-theoretic terms are built on the definition of the entropy. Next, we consider another key concept called *mutual information*.

1.2. Mutual Information

Mutual Information is the measure of information that one random variable has about another. In other words it is the measure of the uncertainity that is left about a random variable after knowing the other.

Definition 6 *Consider two random variables X and Y with a joint PMF $p(x,y)$ and the marginal PMFs $p(x)$ and $p(y)$. The mutual information $I(X;Y)$ is given by*

$$I(X;Y) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \tag{1.10}$$

Mutual Information can also be expressed in terms of entropy and conditional entropy as follows

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X), \tag{1.11}$$

Thus mutual information is the reduction in the uncertainity of X due to the knowledge of Y or vice versa. Thus from the relation shown in 1.11 we can easily see that

$$I(X;Y) \leq H(X) \quad \text{and} \quad I(X;Y), \leq H(Y) \tag{1.12}$$

Also since conditioning reduces entropy so from equation 1.11 we can say that $H(X) \geq H(X|Y)$ or $H(Y) \geq H(Y|X)$ and hence $I(X;Y) \geq 0$. Figure 1.2 shows the relation between entropy, conditional entropy, joint entropy and mutual information. Next we try to find out the relationship between mutual information and entropy. So if we try to find out the mutual information between

X with itself then we get

$$I(X;X) = H(X) - H(X|X) = H(X), \tag{1.13}$$

Thus the mutual information of X with itself is same as its entropy. Hence entropy is sometimes also known as *self information*. Since most of the things have been defined already it is now time to focus on the rate of information transfer. Let us consider a simple channel with input X and output Y. The channel is not affected by any external noise or interference hence Y should be same as X, i.e. the receiver should receive whatever the transmitter has sent. The mutual information between X and Y thus becomes

$$I(X;Y) = H(Y) - H(Y|X) = H(X) - H(X|Y) = H(X) - 0 = H(X), \tag{1.14}$$

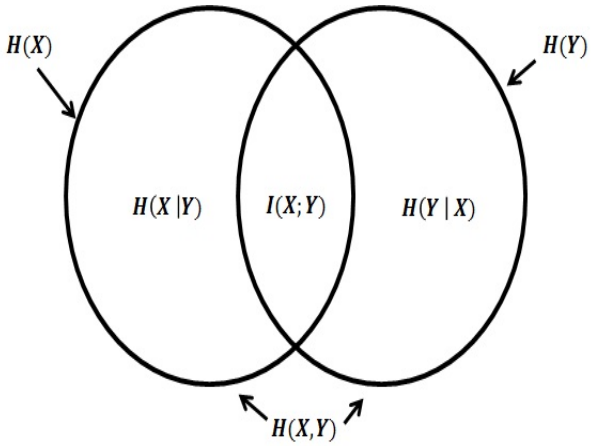

Figure 1.2. Relation Between Entropy, Conditional Entropy, Joint Entropy and Mutual Information

In equation 1.14 $H(X|Y)$ is equal to 0 due to the noiseless nature of the channel, which makes sure that the output is the same as the input. However had X and Y been independent then the mutual information between them would have been 0 as in that case $H(X|Y) = H(X)$. Thus the mutual information between X and Y can vary between 0 and $H(X)$ depending on the relationship between X and Y, where $H(X)$ is the maximum value that $I(X;Y)$ can have since conditional entropy can never be negative and 0 is the minimum that it can have since $I(X;Y) \geq 0$. Thus depending on the channel between X and Y we can also define $I(X;Y)$ as the mutual information of the channel. From the above example, we can predict that the amount of information transferable

is dependent on the mutual information of the channel.We will soon find out that if we try to send information at a rate higher than the mutual information of the channel then the information sent over the channel cannot be received *reliably* by the receiver. In the next section we formally define *channel capacity* - the maximum rate at which information can be reliably transmitted over a channel.

1.3. Channel Capacity

From the previous section it is pretty clear that any communication over a channel must be associated with a rate at which information is being transferred over the channel and possibly that rate is somehow related to the mutual information between the input and the output. Now when that exchange of information, taking place over the channel happens at a rate such that the receiver receives with minimum error then we can say that the rate is *achievable*. But then the question is - is the achievable rate a constant or does it vary? Let us say, for example - Lucy designs a transmission scheme on a channel such that she can achieve a rate of 5 bits per channel use and it turns out that any rate greater than that does not provide reliable means of communication. What does that mean? You can design a transmission scheme where you can achieve a rate upto 5 bits per channel use and nothing more than that. We define *channel capacity* of a channel as the maximum achievable rate on a channel over all possible transmission schemes.

In information theory we characterize the channel capacity in two simple steps: First we find an upper bound to the transmission rate of a channel and then we compare it with the achievable rates derived for various transmission schemes. The different upper bounds derived serves as an upper bound to all achievable rates. Now if any of these upper bounds coincide with the achievable rate derived then that is our *capacity*. In other words capacity is the infimum of all upper bounds and at the same time the supremum of all achievable rates. With this background we are now ready for the formal definition of channel capacity and to find the capacity of some very popular channels which will be of use later in solving the problems.

Definition 7 *We define the information channel capacity of a discrete memoryless channel as*

$$C = \max_{p(x)} \Big[I(X;Y) \Big], \tag{1.15}$$

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

The definition for the continuous channel follows in the same spirit except for the fact that now the PMFs are replaced with PDFs. We next find out the capacity of a Binary Symmetric Channel (BSC) and a Binary Erasure Channel (BEC).

1.3.1. Binary Symmetric Channel (BSC)

The channel shown in figure 1.3 is a binary symmetric channel in which the input symbols are transmitted to the receivers as it is with probability $1-p$ whereas it is complemented with probability p. The channel is binary because it can take inputs only from the field $\{0,1\}$. Thus as you can see that the channel is not reliable and there are chances for error to occur. However we will soon show that even such a channel has a positive capacity rate of reliable transmission.

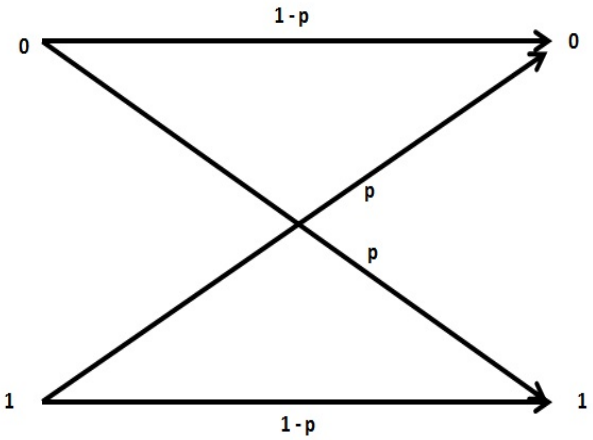

Figure 1.3. A Binary Symmetric Channel.

We first try to find a bound for $I(X;Y)$ and then try to achieve the bound

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X), \\ &= H(Y) - \sum p(x) H(Y|X=x), \\ &= H(Y) - \sum p(x) H(p), \qquad (1.16) \\ &= H(Y) - H(p), \qquad (1.17) \\ &\leq 1 - H(p), \qquad (1.18) \end{aligned}$$

where equation 1.16 follows from the fact that the channel $Y|X$ is a binary channel,i.e. it can take values only in $\{0,1\}$, equation 1.17, is the result of the fact $\sum p(x) = 1$ and $H(p)$ is a constant value for a particular p and is independent of the distribution of X. Finally since $Y \in \{0,1\}$ a binary

field, the maximum entropy it can have is 1. Now that the upper bound is defined we can achieve it if we use Bernoulli($\frac{1}{2}$) distribution for X, i.e. for X both 0 and 1 occur with equal probability. That will ensure that Y will also have a Bernoulli ($\frac{1}{2}$) distribution and hence we can achieve the upper bound and thereby giving the capacity of BSC channel as $1 - H(p)$.

1.3.2. Binary Erasure Channel (BEC)

Next we find the capacity of another channel known as the Binary Erasure Channel as shown in figure 1.4. In this case the input is very similar to BSC it takes values only in $\{0, 1\}$ however the output is a little different in the sense that it can now take a value from the field $\{0, 1, e\}$ where e represents the erased channel state. With this small description of the channel in hand we next try to calculate the capacity of the binary erasure channel as follows

$$\begin{aligned} C &= \max_{p(x)} I(X;Y), \\ &= \max_{p(x)} \Big[H(Y) - H(Y|X) \Big], \\ &= \max_{p(x)} \Big[H(Y) - H(p) \Big], \end{aligned} \tag{1.19}$$

where equation 1.19 is because of the channel configuration as shown in figure 1.4. Well the next step is to find an input distribution so as to find the maximum of equation 1.19.

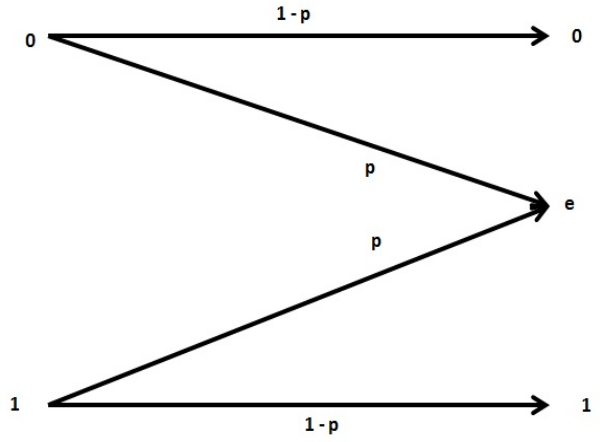

Figure 1.4. A Binary Erasure Channel.

The obvious first guess would be to have the ouput samples $\{0, 1, e\}$ equally distributed so that $H(Y) = \log_2 3$ which is the maximum that one can get for a three sample discrete random variable. But one can find out that there is no such input distribution with which you can get

$H(Y) = \log_2 3$. What we do next is we define an event $E = \{Y = e\}$, then using the chain rule of entropy

$$H(Y) = H(Y, E) = H(E) + H(Y|E), \tag{1.20}$$

Let us define a distribution for X such that $P(X = 1) = \alpha$ and $P(X = 0) = 1 - \alpha$. Then we have

$$H(Y) = H((1-p)(1-\alpha), p, (1-p)\alpha) = H(p) + (1-p)H(\alpha), \tag{1.21}$$

where $P(Y = 1) = (1-p)\alpha$, $P(Y = 0) = (1-p)(1-\alpha)$ and $P(Y = e) = p$ for the above input distribution. Now from equation 1.19 we can further write using equation 1.21

$$\begin{aligned} C &= \max_{\alpha} \Big[\cancel{H(p)} + (1-p)H(\alpha) - \cancel{H(p)}\Big], \\ &= \max_{\alpha}(1-p)H(\alpha), \\ &= (1-p), \end{aligned} \tag{1.22}$$

where the maximum in equation 1.22 is achieved when $\alpha = \frac{1}{2}$. Thus equation 1.22 gives the capacity for a Binary Erasure Channel. The next section discusses about the popular channel models considered for implementing physical layer security, its practicality and the problem considered in this thesis.

1.4. Introduction to Physical Layer Security and Problem Statement

Wireless communication is one of the most ubiquitous discovery of modern technologies. Cellular communication alone is accessible to an estimated 5 billion people, and this is but one, of an array of wireless technologies that have emerged in recent decades. Wireless networks are increasingly used for a very wide range of applications, including banking and other financial transactions, social networking, and environmental monitoring, among many others. For this reason, the security of wireless networks is of critical societal interest. Traditional methods of providing security in such networks are impractical for some emerging types of wireless networks due to the light computational abilities of some wireless devices, such as radio-frequency identification (RFID) tags, certain sensors, etc., or due to the very large scale or loose organizational structure of some networks. Also modern day existing security schemes such as authentication schemes, encryption schemes or identification schemes and so on are based on the assumption of computational hard-

ness, for example the assumption that factoring and computing discrete log is hard. But with the processing power of modern day machines improving exponentially we might soon come across adversaries with immense computing power and then the existing cryptographic algorithms might no longer hold good. For these and other reasons, there has been considerable recent interest in developing methods for secure data transmission that are based on the physical properties of the radio channel (the so-called wireless physical layer). These results are based on information theoretic characterizations of secrecy, which date to some of Claude Shannon's early work on the mathematical theory of communication [2]. Whereas Shannon's work focused on symmetric key encryption systems, perhaps a more relevant development in this area was Aaron Wyner's work on the wiretap channel, which introduced the idea that secrecy can be imparted by the communication channel itself without resorting to the use of shared secret keys [3]. Wyner's work dealt with a *degraded* wiretap channel and in the next section we give an overview of the same as well as the wiretap channel in general.

1.4.1. Introduction to Wiretap Channel

In [3] as shown in figure 1.5 Alice wants to transmit a confidential message to Bob while keeping it secret from Eve. The objective now is two fold: Alice must encode the message M into a codeword X^n of length n such that Bob, having received Y^n can reliably recover the message, i.e. $P\{\hat{M} \neq M\} \xrightarrow[n\to\infty]{} 0$. Note that a codeword of length n uses the channel n times,i.e. $X^n = \{X_1, \cdots, X_n\}$, where X_i is sent in the ith channel use. Similarly, $Y^n = \{Y_1, \cdots, Y_n\}$ and $Z^n = \{Z_1, \cdots, Z_n\}$ describe corresponding channel outputs at the legitimate receiver and eavesdropper respectively. Also the message must be kept secret from Eve. Wyner defined secrecy in terms of *equivocation*, or conditional entropy. He required that $\frac{1}{n}H(M|Z^n) \geq \frac{1}{n}H(M) - \epsilon$,i.e. the knowledge of the channel output Z^n will not decrease the uncertainity about the message M.

Now there are several secrecy criterias in literature. Shannon's work [2] dealt with *perfect secrecy* which is a very stringent criteria which requires the statistical independence of the message M and the channel output Z^n at Eve,i.e. there is no relationship between M and Z^n. Mathematically, $I(M; Z^n) = 0$. Wyner in [3] defined the concept of *weak secrecy* where the mutual information between M and Z^n satisfies the condition $\frac{1}{n}I(M; Z^n) \xrightarrow[n\to\infty]{} 0$. This quantity describes the information leaked about M to Eve in terms of a rate due to the normalization by the block length n. The above definition of secrecy was further strengthened and was termed *strong secrecy*

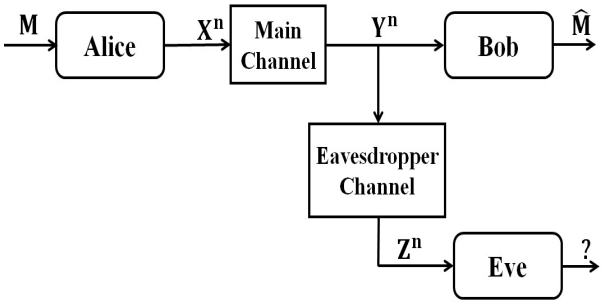

Figure 1.5. Wyner's Wiretap Channel.

in [4] and[5]. Mathematically it was defined as $I(M; Z^n) \xrightarrow[n\to\infty]{} 0$ and the intution was to have the total amount of information leaked to Eve vanish as $n \to \infty$.

Now with all the definition of secrecy for wiretap channel as above it is now time to find out how secrecy capacity can be achieved. Recall that Alice must encode the message into a codeword such that it is useful for Bob to recover the transmitted message (reliability) and at the same time the same codeword is useless for Eve (security). These two requirements seem to be conflicting and it is not obvious that it is possible to achieve both simultaneously. The crucial idea for achieving the secrecy capacity is the following: Instead of using all of the available resources for message transmission, a certain part of them are used for randomization by adding "dummy" messages unknown to Bob and Eve. Specifically, for each confidential message Alice wants to transmit, there are multiple valid codewords and a stochastic encoder chooses one of them uniformly at random. The key idea is now to choose the randomization rate for each confidential message such that it is almost same as Eve's channel capacity. Thus, Eve will be saturated with the useless information carried by the dummy variables, leaving no remaining resources for decoding the confidential message itself.

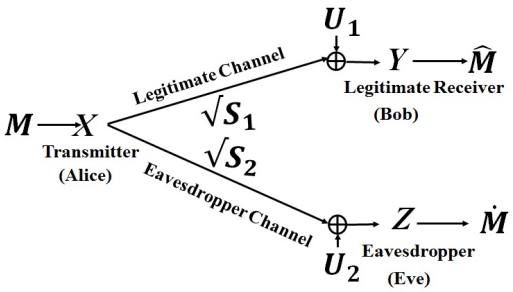

Figure 1.6. A Model for a Gaussian Wiretap Channel.

Figure 1.6 represents a general wire - tap channel. Mathematically it is defined as follows

$$Y = \sqrt{S_1}X + U_1$$
$$Z = \sqrt{S_2}X + U_2 \tag{1.23a}$$

where X is transmitted signal by Alice whereas Y and Z are the signals received at Bob (the intended receiver) and Eve (eavesdropper) respectively. U_1 and U_2 are the additive gaussian noises at the two respective receivers. S_1 and S_2 according to information theoretic terminology are the Channel State Information (CSI). This channel states can be fixed or time varying depending on whether they are constant over time or they vary with time. Moreover there is further classification based on the rate at which the channel states vary. In addition to that the receivers can derive information about the channels from their respective received signals but the transmitter can get information about the channels only if the receivers feedback about the same. So depending on how fast the channel is changing it might or might not be always possible for the receivers to provide a timely update about the channel states to the transmitters. Hence as we can see there can be various classifications of the channels depending on their distributions as well as what information is available where. Realistically speaking in most cases the channel might be varying fast enough to prevent provide a timely update to the transmitter and hence the transmitter might be unaware of the instantaneous channel states. In addition in a wiretap set up it is not justifiable to expect the eavesdropper to provide with a feedback even if it can.

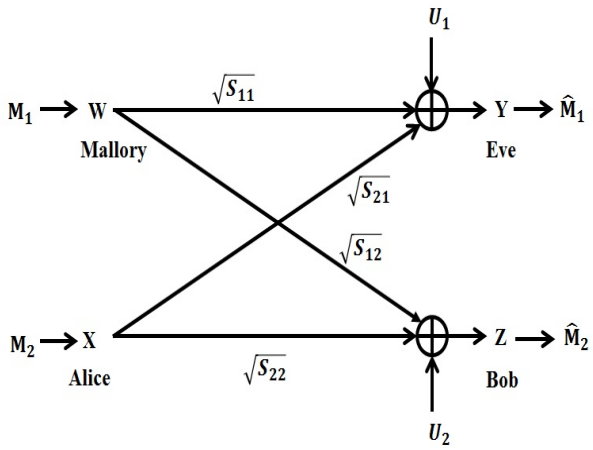

Figure 1.7. A Gaussian Interference Channel Model.

1.4.2. Introduction to Interference Channel

There has been considerable effort in extending and generalizing concepts and results for the wiretap channel to more complex multiuser scenarios which includes the very popular *Interference Channel (IC)* and a little less complicated version of it, the *Z-Interference Channel (ZIC)*. The interference channel (IC) describes the communication scenario in which multiple transmitter-receiver pairs interfere with each other. Each sender is interested only in transmitting information to its designated receiver. However, due to the open nature of the wireless medium, the transmitted signals are received not only by the intended receivers but also by the other users. The interference channel with confidential messages consider two transmitters Mallory and Alice who wish to transmit their confidential messages M_1 and M_2 to their respective receivers Eve and Bob. Because both transmissions interfere with each other, each transmitter must encode and transmit its message in such a way that it is kept secure from the counterpart receiver. This is shown in 1.7. The Z-Interference Channel as shown in figure 1.8, on the other hand is not much different from the

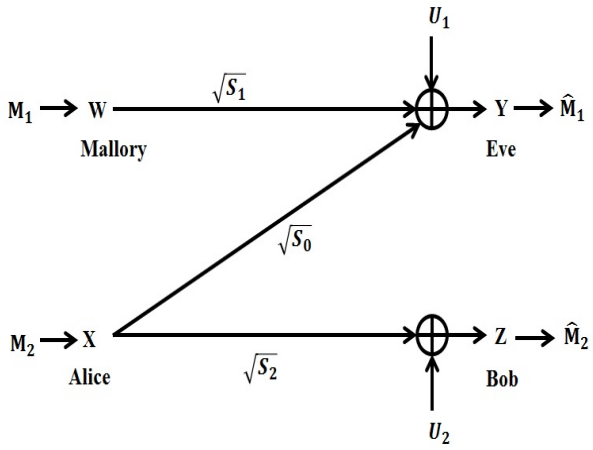

Figure 1.8. A Gaussian Z-Interference Channel Model.

Interference Channel except from the fact that now, due to topological location of the transmitters and receivers the signal from one of the transmitter does not reach the unintended receiver,i.e. one of the interfering links of the interference channel is absent and hence might turn out to be a little simpler than an interference channel in general. A model of such a channel is shown in figure 1.8 where as it can be seen the link between Mallory and Bob is absent and hence any message M_1 transmitted from Mallory is inherently secret. Hence here the concern is to keep M_2 transmitted

from Alice secret. Figure 1.7 shows a general gaussian interference channel. Mathematically, it can be represented as follows

$$\begin{aligned} Y &= \sqrt{S_{11}}W + \sqrt{S_{21}}X + U_1 \\ Z &= \sqrt{S_{12}}W + \sqrt{S_{22}}X + U_2 \end{aligned} \tag{1.24}$$

where W is the signal transmitted by Mallory and intended for Eve while X is the signal transmitted by Alice and intended for Bob. Here Bob is the unintended receiver during communication between Mallory and Eve while Eve plays the role of Bob (unintended receiver) during communication between Alice and Bob. U_1 and U_2 are the additive gaussian noises at the two receivers. $S_{ij} \forall i, j \in \{1, 2\}$ represents the channel state information from Tx_i to Rx_j and as mentioned in the previous subsection can have different properties based on whether it varies over time or is fixed. Besides the CSI might be available at both the receiver and the transmitter, when fedback by the receiver or might just be known at the receiver. With this background we next provide a brief review of practical coding schemes that are used for physical layer security.

1.4.3. Practical Coding

Although most work on physical security are based on non-constructive random-coding arguments to establish theoritical results but are of little practical significance however in recent times designing of practical codes for physical layer security has gained momentum. Low-density parity-check (LDPC) codes with nested graph [6], two-edge type LDPC codes in a scenario where both the main and eavesdropping channel are Binary Erasure [7] has been seen to achieve exactly or close to the theoretical secrecy capacity. Also LDPC codes with puncturing [8] and non-puncturing [9] for Gaussian wiretap channel has yielded significantly close results. A LDPC code based Han Kobayashi scheme was proposed for the Gaussian IC in [10], which has close to capacity performance in the case of strong interference. In [11], a specific coding scheme was designed for the binary-input binary-output Z- Interference Channel (ZIC) using LDPC codes, and an example was shown to outperform time sharing of single user codes. Polar codes have also been shown to provide promising results for Symmetric binary input memoryless degraded wiretap channel with weak secrecy criteria in [12],[13],[14] and with strong secrecy criteria in [15] and for general wiretap channel with strong secrecy criteria in [16]. A polar coding scheme that achieves the Han-Kobayashi inner bound for

the 2-user Interference Channel (IC) was proposed in [17], and [18] used a similar scheme to achieve the Han - Kobayashi region in the 2-user classical-quantum IC. In addition Lattice code has also been used for both Gaussian as well as Rayleigh wiretap channels under the secrecy gain or weak and strong secrecy criteria as in [19],[20],[21],[22].

1.4.4. Problem Statement

With a preface about the existing channel models, their various forms of security and the practical coding for them, we are now ready to describe the progress we have made through the work in this thesis in the general area of secret communication. In chapter 2, we consider a Gaussian fading wiretap channel with no channel state information at the transmitter. This is a very practical situation since in a wireless environment the channel is almost constantly changing and that is captured by the fast fading nature of the channel that we are considering. In addition we consider that the transmitter is not aware of the instantaneous value of the channel, in technical terms we say there is no channel state information at transmitter (CSIT). For this setting we try to find the secrecy capacity. We first find an outer bound for the channel and then achieve secrecy rates within constant gap of the outer bound for an arbitrary distribution of the legitimate and eavesdropper channel. This is the best result known till now for an arbitrary channel distribution in fast fading Gaussian Wiretap channel. After making some significant progress in characterizing the secrecy capacity of fading gaussian wiretap channel within constant bits, we next move on to a more general multi-user scenario. In chapter 3, we next consider the fading IC problem for the same setting as the wiretap channel,i.e. we consider a fast fading channel with arbitrary distribution and no CSIT. Since the problem of determining the secrecy capacity in a fading IC in general is one of the most complicated problems in the interference channel realm, so we first consider a less complicated version of the channel. We consider a binary fading interference channel (BFIC) where fading is characterised by the channel being either present or absent,i.e. erasure or presence of the channel. The channel states being binary can only take values $\{0, 1\}$ and so can the inputs. We define *strong* and *very weak* BFIC in this chapter and characterize the exact secrecy capacity for the same. This idea of solving a fading binary model before embarking on solving the more general real fading channel problem is not at all arbitrary. It has been shown previously to serve as a very potent first step in solving the general gaussian fading problem in [23] and [24] while characterizing the approximate capacity of fading broadcast channel and fading ZIC channel

respectively and in [25] while characterizing the secrecy capacity of fading wiretap channel. With intuitions generated from the Binary model and armed with motivation from [23], [24] and [25] we take the next logical step in the direction of solving the fading gaussian IC problem. In chapter 4, we characterize the secrecy capacity for *strong* and *very weak*, both defined in the chapter, layered fading interference channel (LFIC). In LFIC the fading characteristic of the channel is captured by the random number of channels that are erased or survives at any time instant. In other words the channel state information basically depicts the number of surving layers and is completely random. The thesis is concluded in chapter 5, with strong intuitions for solving the gaussian fading IC problem. It also explores the other directions in which this work can be further extended so as to address the security demands of the time.

To finally sum up the contributions of this thesis, we can state the following :

1. We devise a concrete coding scheme that can achieve physical layer security in any arbitrary fading wiretap channel within a constant number of bits, which is a first of its kind result in wiretap channel.

    - In this, we derive new lower bounds and a looser upper bound such that the gap between them is comparable. Such comparision is unique in wiretap channel.
    - With several numerical examples we show that the upper bound and the achievable rates can be as close to as within 2 bits of each other.

2. We characterize the secrecy capacity for a *strong* and *very weak* fading binary and layered interference channel. To the best of our knowledge, this is the sole result so far which can provide significant intuition to solve the more general real fading interference channel problem with secrecy constraint.

    - In this, we show for both the binary and layered case that an optimal layered wiretap channel code at both the transmitters with proper rates help achieve capacity.
    - It is also shown that the key to achieving capacity is dependent on the proper allocation of layers to the two transmitters based on the channel distribution.

2. SECRECY CAPACITY OF THE FADING BROADCAST CHANNEL WITHIN 11 BITS WITH ONLY CSIR

2.1. Introduction

Secrecy emerges as an additional but natural system design constraint because of the vulnerability of signals to eavesdropping in a wireless network. Multiple legitimate transmit-receive pairs may need to maintain secrecy from one or more potentially malicious receivers in such a network. The simplest communication scenario requiring secrecy is a three node network, where a transmitter (Alice) communicates with a legitimate receiver (Bob) in the presence of an Eavesdropping node (Eve). The maximum rate at which the legitimate pair can communicate information satisfying a suitable secrecy constraint is called the *secrecy capacity* of this network. Earliest formal treatment of the secrecy capacity can be found in [26], [3]; the later considered a special case of the aforementioned three node network, where the eavesdropper obtains a degraded version of the signal received by the legitimate user and called it *wiretap channel* (WC). In this chapter we consider a general version of the wiretap channel called Broadcast Channel with one Legitimate receiver and one Eavesdropper (BCoLoE), where Eve receives its signal via a separate channel that is not necessarily dependent on the legitimate user's channel. Evidently, the individual point-to-point (PTP) capacity of both the receivers in the BCoLoE depends on the statistics of the corresponding channel coefficients. Intuitively, it may seem that secrecy capacity can be achieved only if Bob has a larger PTP capacity than Eve. Interestingly, on a fading BCoLoE that is not necessarily true. If the instantaneous fading coefficients are known at the transmitter, the transmitter can choose to transmit only during those channel realizations when the strength of the legitimate channel is stronger than that of the eavesdropper. However, knowledge of the instantaneous fading coefficients of communication links to both the legitimate user and eavesdropper might be impractical. On one hand, it is unreasonable to assume that the malicious receiver/s will feedback their channel coefficients to the transmitter. On the other hand, the main channel fading coefficients may change sufficiently fast making it impossible for the legitimate receiver to feed it back to the transmitter in a timely manner. This motivates us to address the No-CSIT problem in this chapter. How-

ever, without the instantaneous CSI at the transmitter, the aforementioned selective transmission technique cannot be implemented and devising an optimal transmit-receive strategy becomes more challenging.

We will use the so called *perfect secrecy* criterion, where the *equivocation rate* is assumed to be arbitrarily close to unity or equivalently the rate of information leakage to the malicious user tends to zero asymptotically. This is in contrast to a relatively stronger secrecy constraint [27] which assumes that the total information leakage to the malicious receiver goes to zero asymptotically. The equivocation criterion was used by Wyner [3] to characterize the secrecy capacity of a Discrete Memoryless WC (DM-WC), with a degraded eavesdropper. The restriction of a physically degraded eavesdropper in Wyner's model was later lifted and a two user broadcast channel with both confidential and common messages were considered by Csciszar and Korner in [28]. Characterization of the secrecy capacity for various other channel configurations has since then been an active area of research. For instance, the secrecy capacity of SISO and MIMO BCoLoEs with time-invariant communication links have been characterized in [29] and [30], [31], [32], respectively. The BCoLoE and its different variations was investigated under diffrent type of fading assumptions as well. For instance, [33], [34] and [35] considered slow fading links, [36], [37], [38] assumed block fading links, [39], [40] and [41] treated fast fading links whereas [42], [43] and [44] addressed a *mixed* fading environment, where the legitimate receiver has a fixed channel but the eavesdropper has a fast fading channel.

The authors in [33] studied a slow fading SIMO WC and derived an expression for outage secrecy capacity assuming only main channel CSIT, which was extended in [34] and [35] for the SISO BCoLoE. It was also established later that outage secrecy capacity can be larger than AWGN secrecy capacity and positive outage secrecy capacity is achievable even if the average SNR of the main channel is worse. Yuksel and Erkip derived the achievable DMT of MIMO wiretap channel with Gaussian input and No-CSIT and complete CSIT. In [36], Gopala et al. considered a block fading BCoLoE with asymptotically large coherence period and characterized ergodic secrecy capacity both with complete CSIT and only main channel CSIT. Li et. al. in [39] characterized the secrecy capacity of independent parallel BCoLoE and as a special case of it derived the secrecy capacity of ergodic fading BCoLoE assuming complete CSIT. This result was extended by Liang et. al. [38] to a BC with one confidential message and one common message and by Ekrem et.

al. [41] to the case having two secret messages and one common message. Parallel channels with only main channel CSIT has been studied in [40] in the context of a BC with multiple legitimate receivers, one eavesdropper and where the transmitter either has a common secrect message for all legitimate receivers or independent messages for them. Upper and lower bounds to the secrecy capacity was derived which coincide for reversely degraded channel and channel with asymptotically large number of receivers, respectively.

Motivated by the difficulty of availing instantaneous CSIT in fast fading channels, some recent works assume partial [45], delayed [46] and imperfect [47] CSIT. For instance, in [46] the authors have considered a fast fading BCoLoE with delayed CSIT and characterized the outage throughput of the channel. In [47] the authors has characterized lower and upper bounds to the ergodic secrecy capacity of BCoLoE with imperfect channel state estimates of only the main channel which was extended in [48] to a BC with one common and two secret messages and in [49] to a BC with multiple legitimate receiver and one eavesdropper. In contrast to these papers which assume CSIT in some form or other, research articles which address the No-CSIT secrecy capacity problem is surprisingly scarce. A fast fading BCoLoE with only statistical CSI at the transmitter was considered in [25] and the exact ergodic secrecy capacity was characterized for a class of channels known as the *stochastically degraded* channel. For general fading, only an upper bound to the secrecy capacity was provided. A similar result was also reported in [50], where an achievable secrecy rate expression was also established for Nakagami-m fading channels. However, no comment on the proximity of this rate to the secrecy capacity of the channel was made. Characterization of the No-CSIT secrecy capacity of a fast fading BCoLoE (FBCoLoE) is an open problem till date. Infact, the secerecy capacity of a FBCoLoE even in presence of only main channel CSIT is not known. However, with increasing demand of connectivity while on the move, fast fading scenarios are abundant. Motivated by these factors, in this chapter, we characterize the secrecy capacity of the FBCoLoE approximately within 11 bits with no instantaneous CSIT.

CSI at the transmitter is typically used to implement a transmission scheme which favors the legitimate receiver with respect to the eavesdropper. For instance, the concept of *injection of artificial noise*, was introduced in [51], [52] to achieve secrecy on fading BCoLoE, which was later used in [53] and [54] assuming only main channel CSIT and in [55] assuming complete CSIT. The main idea of noise injection is to transmit some noise signal into the orthogonal subspace

to the receive signal space of the legitimate receiver, where the CSI of at least the main channel is required at the transmitter to compute these subspaces. The transmitted noise thus does not affect the legitimate receiver but only interferes with the eavesdropper. Similarly, in [39] and [36] CSI of both receivers are used to selectively transmit only when main channel is better than the eavesdropper. The transmission scheme in [40] utilizes only the main channel CSI to implement a power allocation strategy that maximizes the average legtimate receiver rate with respect to the average rate of eavesdropper. In absence of CSIT, the challenge is to devise good coding schemes that are not dependent on instantaneous CSIT. In this chapter, we adopt a coding stategy which utilizes statistical properties of channel coefficients. Evidently, a signal transmitted with a particular power on a point-to-point fading channel is received at different SNRs with differnt probability, depending on the distribution of the corresponding channel coefficients. In our scheme, intuitively, informaion symbols are sent at carefully chosen signal levels so that the corresponding received SNRs are good for decoding at the legitimate receiver with more probability than at the eavesdropper. The contribution of this chapter can be summarized as follows:

- We derive an upper bound and two lower bounds to the secrecy capacity of a Fading Broadcast Channel with one Legitimate receiver and one Eavesdropper (FBCoLoE) with arbitrary fading distribution. Only fading statistics of both the channels are assumed at the transmitter and instantaneous CSI are assumed at the receivers.

- We show that the upper bound and the smaller among the two lower bounds are within 11 bits to each other for all channel statistics and Signal-to-Noise Ratios (SNR), which in turn imply that both the lower bounds are within 11 bits to the secrecy capacity of the channel. Thus, the achievable scheme of this chapter provides a guaranted performance.

- We show that the BES-RS [56] scheme can achieve a secrecy rate which is better than both the lower bounds.

- Finally, the aforementioned upper and lower bounds are computed for several BCoLoEs *numerically*.

Since the derivation of the universal gap of 11 bits involves several loose bounding steps we expect the actual gap to be much smaller. Indeed, in all of our numerical examples the actual gap between

the upper and lower bound are found to be 2 bits or less. The rest of the chapter is organized as follows: the Fading Gaussian Wiretap Channel (WC) model is described in Section 2.2 along with the secrecy criterion and the notion of secrecy capacity within a constant number of bits. Section 2.3 summarizes the main contribution of this chapter and also provides an example to illustrate the improved performance of the BES-RS scheme over simple Gaussian coding. The remaining sections contain the proofs and verification of the results presented in section 2.3. It starts with the proof of the new upper bound to the secrecy capacity in section 2.4, which is followed by the proof of the two lower bounds provided in section 2.5. The fact that the upper and lower bounds derived are within a constant number of bits is proved in section 2.6. In section 2.7, we compute these aforementioned bounds numerically for several example BCoLoEs. Finally, we conclude the chapter in Section 2.8.

Notations 1 *We will denote the set of real, complex and natural numbers by* $\mathbb{R}$, $\mathbb{C}$ *and* $\mathbb{N}$, *respectively. Let us denote the distribution of* $B \in \{+1, -1\}$, *a binary antipodal random variable with* $\Pr(B = +1) = p$, *by* $\mathcal{B}(p)$. *For a discrete random variable* A *with realizations coming from* $\{a_1, a_2, \cdots, a_k\}$, *its Probability Mass Function (PMF) will be denoted by* $P_A(.)$, *i.e.,* $P_A(a_i) = \Pr(A = a_i)$, $\forall i$. *For an arbitrary real number* $a \in \mathbb{R}$, $(a)^+$ *represents the maximum of* a *and zero, i.e.,* $(a)^+ = \max\{a, 0\}$. *For a non-negative real number* b, *its logarithm with base* 2 *and* e *will be denoted by* $\log(b)$ *and* $\ln(b)$, *respectively. The distribution of a Circularly Symmetric Complex Gaussian (CSCG) random variable with mean* μ *and variance* σ^2 *will be denoted as* $\mathcal{CN}(\mu, \sigma^2)$. *The distribution of a real Gaussian random variable with mean* μ *and variance* σ^2 *will be denoted by* $\mathcal{RN}(\mu, \sigma^2)$. *We will use "IID" to mean identically and independently distributed.*

2.2. Channel Model and Some Preliminaries

Consider a fast fading Broadcast Channel (BC) with one transmitter (Tx/Alice), one legitimate receiver (Rx_1/Bob) and one unintended/malicious receiver (Rx_2/Eve), as shown in figure 2.1. The relation between the input and outputs of the channel, at time t, can be written as,

$$Y'_t = \sqrt{S_{1t}} e^{j\theta_{1t}} X_t + U'_{1t} \tag{2.1}$$

$$Z'_t = \sqrt{S_{2t}} e^{j\theta_{2t}} X_t + U'_{2t}, \tag{2.2}$$

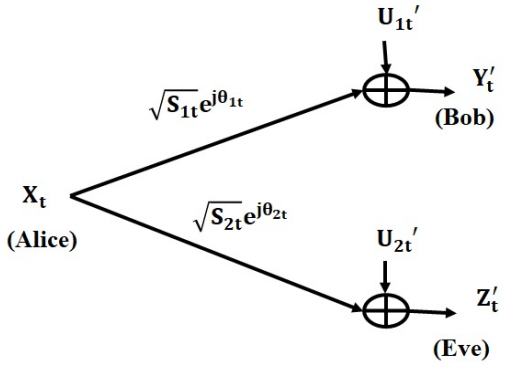

Figure 2.1. Fading Gaussian Wiretap Channel.

where $X_t \in \mathbb{C}$ is the signal transmitted by Alice with unit average power constraint, i.e., $\mathbb{E}(|X|^2) \leq 1$ and $Y_t', Z_t' \in \mathbb{C}$ represent signals received by Bob and Eve, respectively. In Fig. 2.1, $(\sqrt{S_{kt}}, \theta_{kt})$ represents the magnitude and phase pair of the fading coefficient of the k-th receiver's channel. The fading coefficients, i.e., the $(\sqrt{S_{kt}}, \theta_{kt})$ pairs are assumed to change IID as $\left(\sqrt{S}_k, \theta_k\right)$ on every channel use, for each $k \in \{1, 2\}$. Moreover, they are assumed to be independent across users, i.e., $\left(S_{1t}, \theta_{1t}\right)$ is independent of $\left(S_{2t}, \theta_{2t}\right)$, $\forall t \geq 1$. Instantaneous realizations of these random coefficients are assumed to be available only at the respective receivers; the transmitter knows only their statistics. $U_{1t}^{'}$ and $U_{2t}^{'}$ in equations (2.1) and (2.2) represent the additive noise at Bob and Eve, respectively, where $U_{kt}^{'}$'s are IID as $\mathcal{CN}(0, 1)$ both across t and k. Removing the effects of phase from equations (2.1) and (2.2) we get the following alternative input-output relations, [1]

$$Y_t = Y_t' e^{-j\theta_{1t}} = \sqrt{S_{1t}} X_t + U_{1t}; \tag{2.3}$$

$$Z_t = Z_t' e^{-j\theta_{2t}} = \sqrt{S_{2t}} X_t + U_{2t}, \tag{2.4}$$

where $U_{kt} = U_{kt}' e^{-j\theta_{kt}}$ and U_{kt}' are identically distributed for all t and k. All the signals are complex in these equations except the fading magnitudes, $\sqrt{S}_{kt}$'s. As a result, each of these equations can be visualized as a pair of real channels. For instance, Y_t in equation (2.3) can also be written as

$$Y_t = (\sqrt{S_{1t}} X_{r,t}) + U_{r,1t} + j(\sqrt{S_{1t}} X_{q,t} + U_{q,1t}), \tag{2.5}$$

[1] It was shown in [28] that, the secrecy capacity of a Discrete Memoryless BCoLoE can be expressed in terms of mutual information of the legitimate channel and the eavesdropper's channel. Neither of which changes [1] if the outputs are multiplied by scalars. Thus, the secrecy capacity of the BCoLoE does not change if the outputs are multiplied by arbitrary scalars.

where $Y_t = Y_{r,t} + jY_{q,t}$, $X_t = X_{r,t} + jX_{q,t}$ and $U_{1t} = U_{r,1t} + jU_{q,1t}$. It is interesting to note that, only the magnitudes of fading coefficients are sufficient to completely characterize the performance of the channel. Statistical properties of the fading coefficients will be specified in terms of their *Complementary Cumulative Distribution Function* (CCDF), denoted as $\bar{F}_{S_k}(s)$, i.e., $\bar{F}_{S_k}(s) = \Pr(S_k \geq s),\ \forall s \geq 0$. A BCoLoE, as described above and shown in Fig. 2.1, will be referred to as a $(\bar{F}_{S_1}(s), \bar{F}_{S_2}(s))$-BCoLoE.

The secrecy capacity of a $(\bar{F}_{S_1}(s), \bar{F}_{S_2}(s))$-BCoLoE is not known, for general joint distribution of (S_1, S_2). However, if there exists a *stochastic ordering* between the two then the secrecy capacity of the channel can be exactly characterized [25], [50].

Definition 8 (Stochastically stronger channel) *Consider a pair of fading PTP channels - as in equation* (2.3) *- with non-negative fading coefficients* $\sqrt{S_i}$ *and* $\sqrt{S_j}$*, where the CCDFs of* S_i *and* S_j *are denoted by* $\bar{F}_{S_i}(s)$ *and* $\bar{F}_{S_j}(s)$*, respectively. The channel with coefficient* S_i *is called stochastically stronger than that with* S_j*, if*

$$\bar{F}_{S_i}(s) \geq \bar{F}_{S_j}(s),\ \forall s \geq 0. \tag{2.6}$$

On a BCoLoE, if the legitimate user's channel is *stochastically stronger* than that of the Eavesdropper, then it is called a *stochastically degraded* BCoLoE.

The secrecy capacity of a *stochastically degraded* BCoLoE was computed in [25], [50] and is given as:

$$C_s^{st} = \mathbb{E}_{S_1}\Big[\log(1+S_1)\Big] - \mathbb{E}_{S_2}\Big[\log(1+S_2)\Big]. \tag{2.7}$$

Obviously, not all channels are *stochastically degraded*. In fact, there are *stochastically non-degraded* channels for which the above expression has a negative value. For instance, consider a BCoLoE with fading coefficients $\sqrt{S_1}$ and $\sqrt{S_2}$ and their CCDFs as shown in Fig. 2.3. Using the PMF of S_1 and S_2, which is provided in Table 2.2 of section 2.7 we get,

$$\begin{aligned} C_s^{st} &= \sum_{s\in\{10,10^2,10^7\}} \Big[P_{S_1}(s) - P_{S_2}(s)\Big]\log(1+s), \\ &= [-0.4]\log(11) + [0.5]\log(101) + [-0.1]\log(1+10^7) = -0.71. \end{aligned} \tag{2.8}$$

While the secrecy capacity of a general $(\bar{F}_{S_1}(s), \bar{F}_{S_2}(s))$-BCoLoE is an open problem, the authors in [25] and [50][2], derived the following upper bound to the secrecy capacity of the BCoLoE:

$$C_s \leq \log e \int_0^\infty \left[\bar{F}_{S_1}(\gamma) - \bar{F}_{S_2}(\gamma)\right]^+ \frac{d\gamma}{1+\gamma}, \tag{2.9}$$

where C_s will denote the secrecy capacity of the BCoLoE, hereafter. Note that, in contrast to the secrecy capacity expression for a BCoLoE with instantaneous CSI at the transmitter (CSIT) that was derived separately in [39], [41] and [38], the upper bound in equation (2.9) does not have any power allocation scheme at the input. This is because in the derivation of equation (2.9) only CSIR is assumed. Moreover, it is not known how far this bound in (2.9) is from the actual secrecy capacity of the channel with only CSIR. In this chapter, we answer this question by first proposing an achievable scheme for the BCoLoE with CSIR only and then showing that it can achieve the secrecy capacity of the channel approximately within 11 bits. The notion of achieving the secrecy capacity approximately within a constant number of bits is defined in the next subsection, along with the secrecy criterion used in this chapter and a definition of an achievable secrecy rate.

2.2.1. Approximate Secrecy Capacity Within a Constant Number of Bits

Given a message $\mathcal{M}(i)$, $i \in \{1, \cdots, 2^{nr_s}\}$, the transmitter uses a stochastic encoder [28] to convert the message into a codeword $X^n \in \mathcal{C}_s(n)$ and sends it through the channel, where $\mathcal{C}_s(n)$ is the codebook. It is received at Bob as Y^n and at Eve it is received as Z^n. If the estimated message at the legitimate receiver (Bob) is $\mathcal{M}(\hat{i})$, then the probability of detection error can be denoted as $P_e(n) = \Pr(i \neq \hat{i})$. The *secrecy* of this message from Eve is measured in terms of equivocation rate, i.e., $\frac{1}{n}h(\mathcal{M}|Z^n, \mathcal{S}_2)$, where $\mathcal{S}_2 = \{S_{2t}\}_{t=1}^n$.

A *secrecy rate* r_s is said to be *achievable* if there exists a sequence of codebooks $\{\mathcal{C}_s(n)\}$ such that $P_e(n) \to 0$ and

$$\frac{h(\mathcal{M}|Z^n, \mathcal{S}_2)}{h(\mathcal{M})} > 1 - n\delta_n, \tag{2.10}$$

with $\delta_n \to 0$, as $n \to \infty$. In the sequel, unless explicitly mentioned otherwise, *a rate will always mean secrecy rate, i.e., it satisfies* (2.10). The secrecy criterion of (2.10) is well known in the literature as the *weak secrecy* constraint. However, using a *privacy amplification* technique from [57]

[2]The upper bound in [50] is provided in terms of the Probability Density Functions (PDFs) of the channel coefficients and looks different from (2.9). However, it can be easily shown that these two are different forms of the same expression.

Maurer et. al. [27] proved that, any rate that can be achieved under the *weak secrecy* constraint can also be achieved under the more desirable *strong secrecy* [58] constraint. In the above definitions $I(.)$, $h(.)$ and $h(.|.)$ represent mutual information, average differential entropy and conditional average differential entropy, respectively. The supremum - over all possible encoding-decoding strategies - of all achievable secrecy rates is called the secrecy capacity of the channel.

In this chapter, we will characterize the secrecy capacity of BCoLoE approximately within a constant number of bits, where the approximate secrecy capacity is a natural extension of approximate capacity used on channels without any secrecy constraints [59], [60].

Definition 9 (*approximate secrecy capacity) If a coding scheme can achieve a secrecy rate r_s, where $C_s - r_s \leq \mu$ and C_s denotes the secrecy capacity of the BCoLoE, then the secrecy capacity of the channel is said to be achievable approximately within μ bits.*

2.3. Main Results

In this section, we derive an upper bound to the secrecy capacity of the BCoLoE which is different from that in equation (2.9) and a couple of lower bounds as shown in Fig. 2.2. The newly derived upper bound, despite looser than that derived in [25] and [61], is more easily computable for arbitrary channel statistics and has a form comparable to a corresponding lower bound. The lower bounds are actually lower bounds to the achievable secrecy rate of a coding scheme proposed here for the BCoLoE. It is shown that the difference between the new/larger upper bound and the smaller lower bound can not exceed 11 bits.

Theorem 1 *The secrecy capacity, C_s, of the BCoLoE with fading statistics $(\bar{F}_{S_1}(.), \bar{F}_{S_2}(.))$ as shown in Fig. 2.1, is upper bounded as follows:*

$$C_S \leq \log e \int_0^\infty \left[\bar{F}_{S_1}(\gamma) - \bar{F}_{S_2}(\gamma)\right]^+ \frac{d\gamma}{1+\gamma} \triangleq r_{su}, \tag{2.11}$$

$$\leq \sum_{n\in\mathcal{N}} 2[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})] + \log(1+\rho) - \log(e) \int_{\Gamma_0} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma \triangleq \bar{r}_{su}, \tag{2.12}$$

where the set $\mathcal{N}$ is defined as

$$\mathcal{N} = \{n \in \mathbb{N} : \bar{F}_{S_1}(\gamma_n) > \bar{F}_{S_2}(\gamma_{n+1})\}, \tag{2.13}$$

$$\gamma_0 = 0, \quad \gamma_n = \rho 2^{2(n-1)}, \quad n = 1, 2, 3, \cdots \tag{2.14}$$

$$\Gamma_n = [\gamma_n, \gamma_{n+1}), \quad n = 0, 1, 2, \cdots \tag{2.15}$$

with $\rho > 0$ is an arbitrary real number which will be specified in the sequel.

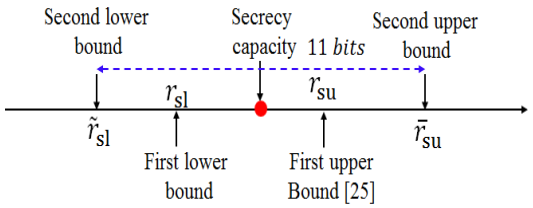

Figure 2.2. Bounds to the Secrecy Capacity.

Evidently, for any given arbitrary CCDF pair, i.e., $(\bar{F}_{S_1}(\gamma), \bar{F}_{S_2}(\gamma))$, the upper bound in equation (2.12) is easier to compute than the integral in (2.11).

Example 1 *Let us consider a BCoLoE with the CCDFs of the fading coefficients as depicted in Fig 2.3. Substituting the values of these CCDFs in equation* (2.11) *we get,*

$$\begin{aligned} C_s &\leq \log(e) \int_{10}^{10^2} \frac{0.4 d\gamma}{1+\gamma}, \\ &= 0.4\Big(\log(1+10^2) - \log(1+10) \Big) = 1.28, \end{aligned} \tag{2.16}$$

Let us also compute the larger upper bound in equation (2.12) *for this example. Substituting $\rho = 20$ (e.g., see Theorem 4) in equation* (2.14)*, we get $\gamma_1 = 20$, $\gamma_2 = 80$ and $\gamma_3 = 320$. Plugging this values in the CCDFs we see that $\bar{F}_{S_1}(\gamma_1) = \bar{F}_{S_1}(\gamma_2) = 0.6, \bar{F}_{S_1}(\gamma_k) = 0$, $\forall k \geq 3$ and $\bar{F}_{S_2}(\gamma_2) = 0.2$, $\bar{F}_{S_2}(\gamma_3) = 0.1$ which when substituted in equation* (2.26)*, we get $\mathcal{N} = \{1, 2\}$ for this channel. Using these values in* (2.12) *we have,*

$$\begin{aligned} C_s &\leq \sum_{n=1}^{2} 2\Big[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1}) \Big] + \log(1+20) \\ &\qquad - \log(e) \int_{0}^{10} \frac{d\gamma}{1+\gamma} - \log(e) \int_{10}^{20} \frac{(0.2) d\gamma}{1+\gamma}, \\ &= 2(0.4) + 2(0.5) + 0.8 \log(21/11) = 2.55. \end{aligned} \tag{2.17}$$

Note that, the fading statistics of the BCoLoE in Example 1 does not satisfy equation (2.6) *and therefore, the channel is not a stochastically degraded BCoLoE. Moreover, it is clear from equation* (2.8) *that the secrect capacity optimal Gaussian input [25] fails to achieve a positive secrecy rate on this BCoLoE. It makes us wonder if there exists an encoding scheme which can achieve a positive secrecy rate on BCoLoEs for all channel statistics including the present one.*

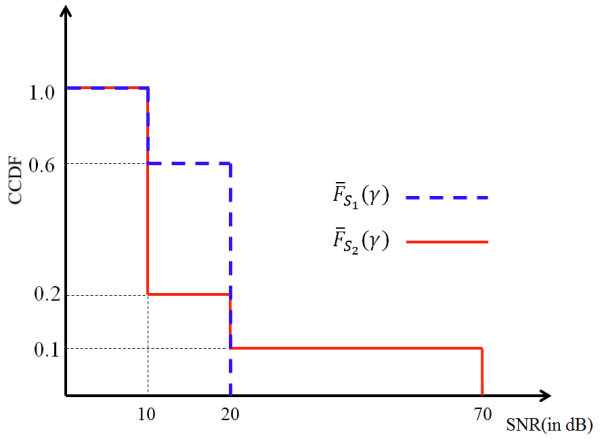

Figure 2.3. Ergodic Capacity of Eavesdropper's Channel is Larger than that of the Legitimate Receiver.

It is insightful to compare the above upper bound to the secrecy capacity of a wiretap channel where both the desired and the eavesdropper's links are binary erasure channels (BECs). The secrecy capacity of such a Binary Erasure Wiretap Channel (BE-WTC) can be easily computed from the result of [62], [28] to be given by

$$C_{BE-WTC} = [\bar{\epsilon}_d - \bar{\epsilon}_e]^+, \tag{2.18}$$

where $\bar{\epsilon}_d = 1 - \epsilon_d$ and $\bar{\epsilon}_e = 1 - \epsilon_e$ and ϵ_d and ϵ_e are the erasure probabilities of the direct and eavesdropper's link, respectively. Comparing with the expression in (2.18), the upper bound in Theorem 1 can be approximately interpreted as the sum of secrecy capacities of several BE-WTCs, where each difference, $[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})]$, can be thought of as the secrecy capacity of the BE-WTC at the n-th layer.

The secrecy capacity upper bound in Theorem 1 can be intuitively explained by visualizing each of the real channels of the legitimate user as a collection of BECs with erasure probabilities

$\{1-\bar{F}_{S_1}(\gamma_n)\}_n$ and each of the real channels [3] of the eavesdropper as a collection of BECs with erasure probabilities $\{1-\bar{F}_{S_2}(\gamma_{n+1})\}_n$. Subsequently, the BECs at layer n of the legitimate channel and the eavesdropper's channel together forms a BE-WTC at layer n. Note that, the secrecy capacity of such a BE-WTC is $[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})]^+$, and therefore, only those BE-WTCs for which the erasure probability of the eavesdropper is larger than that of the legitimate user appears in Theorem 1.

Following this intuition, as an achievable scheme for this channel, we consider the so called *Binary Expansion Signaling* (BES) [56], where each transmitted real symbol is formed by combining several mutually independent antipodal symbols, i.e., if X_r represents the signal transmitted on a real channel then,

$$X_r = \sum_{l=1}^{\infty} x_{rl} 2^{-l}, \forall\ l,\ x_{rl} \in \{+1, -1\}.$$

At the receiver, an estimate $\{\hat{x}_{r,l}\}_{l=1}^{\infty}$ of these antipodal symbols are computed. Thus, the BES scheme effectively converts a real channel into a sequence of binary symmetric channels. The equivalent BSC from $x_{r,n}$ to $\hat{x}_{r,n}$ will be sometime referred to as the n-th layer. The crossover probability and the capacity of a BSC or layer depends on the instantaneous value of the fading state. Averaging over these states we get the average rate of information transmittable via such a BSC or layer. A subset, $\phi \subseteq \mathbb{N}$, of these layers are then carefully chosen to transmit information symbols in such a manner that the information transmitted to the legitimate receiver is maximized and simultaneously, the information leaked to the eavesdropper is minimized. At the receiver, a symbol at layer n can be decoded either by treating all the lower layer symbols as noise, or it can be decoded by first estimating the lower layer symbols and removing their contribution. Clearly, the later is a better scheme and is called BES with *reverse stripping* (BES-RS) scheme. The following theorem provides a lower bound to the secrecy rate achievable by the BES-RS scheme.

Theorem 2 *The secrecy rate, R_{BES}^{RS}, achievable by the BES-RS scheme, where information is transmitted only via layers belonging to $\phi \subseteq \mathbb{N}$, can be lower bounded as:*

$$R_{BES}^{RS} \geq 2\sum_{n\in\phi} \mathbb{E}_{S_1}\Big[\hat{r}_{n,d}(S_1)\Big] + 2\sum_{n\in\phi^c} \mathbb{E}_{S_2}\Big[\hat{r}_{n,d}(S_2)\Big] - \mathbb{E}_{S_2}\big[\log(1+S_2)\big] \triangleq r_{sl}, \tag{2.19}$$

[3] Recall from equation (2.5) that, each of the communication links from transmitter to the legitimate receiver and the eavesdropper is effectively a pair of real channels. This also explains the factor 2 in equation (2.12).

where, denoting the entropy of a $\mathcal{B}(p)$ random variable by $H(p)$ we have

$$\hat{r}_{n,d}(s) = 1 - H(\hat{\epsilon}_d(a_n(s))), \tag{2.20a}$$

$$\hat{\epsilon}_d(a_n(s)) = \min[1/2, \epsilon_d(a_n(s))], \tag{2.20b}$$

$$\epsilon_d(a) = \frac{G(\sqrt{a}(1+2^{-d})) - G(\sqrt{a}(1-2^{-d}))}{\sqrt{a}}, \tag{2.20c}$$

$$G(x) = xQ(x) - \frac{1-\exp(-\frac{x^2}{2})}{\sqrt{2\pi}}, \tag{2.20d}$$

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du, \tag{2.20e}$$

$$a_n(s) = \frac{3}{4} s 2^{-2n} = 3s2^{-2(n+1)}, \tag{2.20f}$$

As mentioned before, the BES scheme effectively converts every real fading channel into a sequence of random Binary Symmetric Channels (rBSCs), where the crossover probability of each such rBSC depends on the instantaneous fading realization. In the computation of the crossover probability $p_{n,d}(s)$ for the n-th layer, all or a selected number of layers below it may appear as interference. In equation (2.20), d represents the distance from the nearest lower layer which appears as interference, while decoding the symbol from the n-th layer.

The first term on the RHS of equation (2.11) approximately represents the rate achievable by the BES-RS scheme on a PTP fading channel. Intuitively, then it is natural to think that the eavesdropper can extract $\sum_{n\in\phi} \mathbb{E}_{S_2}\big(\hat{r}_{n,d}(S_2)\big)$ bits per channel use from the transmitted signal. Therefore, the secrecy capacity will be the difference of the two, i.e., $\sum_{n\in\phi} \mathbb{E}_{S_1}\big(\hat{r}_{n,d}(S_1)\big) - \sum_{n\in\phi} \mathbb{E}_{S_2}\big(\hat{r}_{n,d}(S_2)\big)$. However, to derive the secrecy rate we need an upper bound on the maximum rate extractable by the eavesdropper. Equation (2.39) in the proof of Theorem 2 provides such an upper bound, i.e., the eavesdropper's channel can not extract information at a rate more than $\mathbb{E}_{S_2}\big[\log(1+S_2)\big] - \sum_{n\in\phi^c} \mathbb{E}_{S_2}\big(\hat{r}_{n,d}(S_2)\big)$ from the transmitter for such a signalling scheme. Thus, the transmitter can send secret information at a rate which is larger or equal to the difference of the aforementioned two rates.

In section 2.7, the lower bound from Theorem 2 will be numerically computed for several specific BCoLoEs. In particular, it will be shown in Example 3 that the BES-RS scheme can achieve a secrecy rate which is larger than or equal to 0.14 bits/channel use on the BCoLoE. Comparing

with the upper bound to the secrecy rate computed in (2.16), it is clear that the BES-RS scheme can achieve the secrecy capacity of the BCoLoE of Fig. 2.3 within 1 bit.

In general, it is desirable to compute the upper and lower bounds to the secrecy capacity and compare them. There are two main difficulties at this point: 1) the lower bound of Theorem 2 is not easy to compute when the CCDFs of the channel are continuous functions, i.e., there are infinitely many fading states; and 2) it is not clear if the difference between the upper and lower bounds can be arbitrarily large or not. The first difficulty will be addressed by Theorem 3, where we will derive an alternative lower bound which is easier to compute for any arbitrary fading statistics of the channel. Moreover, the lower bound will be in a similar form to the larger upper bound in Theorem 1. It will be used later in Theorem 4 to prove that the lower bound can not be further than than 11 bits from the upper bound.

Theorem 3 *Denoting the secrecy rate, achievable by the BES-RS scheme on the fast fading wiretap channel with fading statistics* $(\bar{F}_{S_1}(.), \bar{F}_{S_2}(.))$ *as shown in Fig.2.1, by* R_{BES}^{RS}*, it can be lower bounded as:*

$$R_{BES}^{RS} \geq \sum_{n\in\phi} 2[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})] - \log\frac{1+4\rho}{1+\rho} - \int_{\Gamma_0} \frac{\bar{F}_{S_2}(\gamma)d\gamma}{\ln(2)(1+\gamma)} - 2\sum_{n\in\phi}\delta_n(S_1,\gamma_n) - 2\sum_{n\in\phi^c}\delta_n(S_2,\gamma_{n+1}) \triangleq \tilde{r}_{sl}, \tag{2.21}$$

where, ϕ *is the set of layers on which information is transmitted and* $\delta_n(S_k,\alpha)$ *for* $k = 1, 2$ *and* $\alpha > 0$ *are defined as*

$$\delta_n(S_k,\alpha) = \int_{\alpha}^{\infty} f_{S_k}(\gamma) H[\hat{\epsilon}_d(a_n(\gamma))]d\gamma. \tag{2.22}$$

In equation (2.22)*,* $H[p]$ *represents the entropy of a* $\mathcal{B}(p)$ *random variable,* $a_n(\gamma)$ *represents the effective SNR faced by the symbol at the n-th layer and* $\hat{\epsilon}_d(a_n(\gamma))$ *is the quantized crossover probability of the equivalent BSC at the n-th layer (e.g., see equation* (2.20)*). A detailed account of the BES-RS scheme and these parameters are provided in Appendix A.1.*

The secrecy rate achievable by the BES-RS scheme, $R_{\mathrm{BES}}^{\mathrm{RS}}$, is a lower bound to the secrecy capacity of the BCoLoE. As a result, the right hand side of equation (2.21) also serves as a lower bound to the secrecy capacity of the BCoLoE. It is only natural to wonder how far this lower bound

is to the secrecy capacity of the BCoLoE. In absence of an exact expression for the secrecy capacity, this question can be answered only approximately, comparing the lower bound to the upper bound provided in Theorem 1.

Theorem 4 *The secrecy capacity of the BCoLoE with fading statistics* $(\bar{F}_{S_1}(.), \bar{F}_{S_2}(.))$ *as shown in Fig. 2.1, can be achieved by the BES-RS scheme within* 11 *bits.*

Proof 1 (outline) *The Theorem is proved by comparing the upper bound and the lower bound derived in Theorems 1 and 3, respectively. It is shown that the difference between the upper and lower bounds is not more than* 11 *bits if* ρ *is set to* 20*. The theorem then follows from the fact that difference between the lower bound and the secrecy capacity can not be larger than the difference between the lower bound and the upper bound. In section 2.6, we will prove that the difference between the RHS of equation* (2.12) *and the RHS of equation* (2.21) *can be bounded by* 11 *bits if we choose* ρ *appropriately.*

2.4. Proof of Theorem 1

The main idea of the upper bound is to partition the range $[0,\ \infty)$ of the channel magnitudes into several sub-intervals, i.e., $[0,\ \infty) = \cup_{i=0}^{\infty}\Gamma_i$ and then approximate the CCDFs of the two channels carefully so that within each such sub-interval one of them is a stochastically degraded version of the other. Finally, we retain only those sub-intervals where the eavesdropper's channel is stochastically degraded than the main channel.

Recall that, an upper bound to the secrecy capacity of an arbitrary fading wiretap channel was derived in [25], [50] and is specified in equation (2.9), where C_s represents the secrecy capacity of the BCoLoE. We expand the intergral from equation (2.9) using the partition from equation (2.15) and get,

$$\begin{aligned}
R_s &\leq \log e \int_0^{\infty} \left(\bar{F}_{S_1}(\gamma) - \bar{F}_{S_2}(\gamma)\right)^+ \frac{d\gamma}{1+\gamma}, \\
&= \log e \sum_{n=0}^{\infty} \int_{\Gamma_n} \left(\bar{F}_{S_1}(\gamma) - \bar{F}_{S_2}(\gamma)\right)^+ \frac{d\gamma}{1+\gamma}, \\
&\leq \log e \int_{\Gamma_0} \left(\bar{F}_{S_1}(\gamma) - \bar{F}_{S_2}(\gamma)\right)^+ \frac{d\gamma}{1+\gamma} + \log e \sum_{n=1}^{\infty} \int_{\Gamma_n} \left(\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})\right)^+ \frac{d\gamma}{1+\gamma}, \quad (2.23)
\end{aligned}$$

where (2.23) follows from the fact that due to the non-increasing nature of the CCDFs of S_1 and S_2, we have

$$\bar{F}_{S_1}(\gamma) - \bar{F}_{S_2}(\gamma) \leq \bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1}),\ \forall \gamma \in \Gamma_n.$$

Writing $\log(e)$ as $\frac{1}{\ln(2)}$ and utilizing the fact that $\bar{F}_{S_k}(\gamma) \leq 1$ for all $\gamma \geq 0$ and $k = 1, 2$, into equation (2.23) we have,

$$\begin{aligned}
R_s \leq & \int_0^{\rho} \frac{\left(1 - \bar{F}_{S_2}(\gamma)\right)}{\ln(2)(1+\gamma)} d\gamma + \sum_{n=1}^{\infty} \int_{\Gamma_n} \frac{\left(\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})\right)^+}{\ln(2)(1+\gamma)} d\gamma, \\
&= \log(1+\rho) - \int_0^{\rho} \frac{\bar{F}_{S_2}(\gamma)}{\ln(2)(1+\gamma)} d\gamma + \sum_{n=1}^{\infty} \int_{\Gamma_n} \frac{\left(\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})\right)^+}{\ln(2)(1+\gamma)} d\gamma, \\
&\leq \log(1+\rho) - \int_0^{\rho} \frac{\bar{F}_{S_2}(\gamma)}{\ln(2)(1+\gamma)} d\gamma + \sum_{n \in \mathcal{N}} [\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})] \int_{\Gamma_n} \frac{d\gamma}{\ln(2)(1+\gamma)}, \quad (2.24) \\
&\leq \log(1+\rho) - \log(e) \int_0^{\rho} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma + \sum_{n \in \mathcal{N}} 2[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})], \quad (2.25)
\end{aligned}$$

where in equation (2.24) we have used the following notation:

$$\mathcal{N} = \{n \in \mathbb{N} : \bar{F}_{S_1}(\gamma_n) \geq \bar{F}_{S_2}(\gamma_{n+1})\}. \quad (2.26)$$

In equation (2.25) we have used the fact that $\int_{\Gamma_n} \frac{d\gamma}{1+\gamma} = \ln(1+\gamma_{n+1}) - \ln(1+\gamma_n) \leq 2\ln(2),\ \forall n \geq 1$.

2.5. Proofs of Theorem 2 and 3

We start by restating in Lemma 1 below, the rate achievable by the BES-RS scheme on a PTP fading channel. Then in subsection 2.5.1, we will prove Theorem 2 and in subsection 2.5.2, we provide the proof for Theorem 3.

Lemma 1 *Consider a PTP fading channel with input $\tilde{X}$ and output $\tilde{T}$, i.e., $\tilde{T} = \sqrt{S}\tilde{X} + U$, where S is a non-negative real random variable with arbitrary distribution and $U \sim \mathcal{CN}(0,1)$. Suppose, a BES scheme is used at the input so that,*

$$\tilde{X} = \tilde{X}_r + j\tilde{X}_q = (\tilde{X}_{r,\phi} + \tilde{X}_{r,\phi^c}) + j(\tilde{X}_{q,\phi} + \tilde{X}_{q,\phi^c}), \quad (2.27)$$

$$\tilde{X}_{\phi} = \tilde{X}_{r,\phi} + j\tilde{X}_{q,\phi}, \quad \tilde{X}_{\phi^c} = \tilde{X}_{r,\phi^c} + j\tilde{X}_{q,\phi^c}, \quad (2.28)$$

$$\tilde{X}_{a,\alpha} = \frac{\sqrt{3}}{2} \sum_{n \in \alpha}^{\infty} x_{a,n} 2^{-n},\ a \in \{r, q\},\ \alpha \in \{\phi, \phi^c\}, \tag{2.29}$$

$$\phi \subseteq \mathbb{N},\ \phi^c = \mathbb{N} \setminus \phi, \tag{2.30}$$

and the components of $\{x_{r,n}, x_{q,n}\}_{n=1}^{\infty}$ *are IID as* $\mathcal{B}(0.5)$ [4]*. Moreover, each symbol is decoded after decoding and removing the lower layers. Then, for any arbitrary* $\phi \subset \mathbb{N}$*,*

$$I(\tilde{X}_\phi; \tilde{T}, S) = I(\tilde{X}_\phi; \tilde{T}|S) \geq 2 \sum_{n \in \phi} \mathbb{E}_S[\hat{r}_{n,d}(S)], \tag{2.31}$$

where $\hat{r}_{n,d}(s)$ *is as defined in equation* (2.20)*.*

Proof 2 *An outline of the proof is provided in Appendix A.1 for completeness. The detailed proof can be found in [56].*

To achieve a secrecy rate the BES-RS scheme of [56] is modified in the following two aspects:

- Information symbols are transmitted only via a carefully selected set, denoted as ϕ, of layers.
- These layers are choosen so the average rate of information transmission is maximized to the legitimate receiver and minimized to the eavesdropper.

2.5.1. Proof of Theorem 2: BES-RS Scheme Adopted to BCoLoE

We know from [28] that, for any choice of input, the difference of mutual information of the legitimate channel and the eavesdropper channel represents an achievable secrecy rate. Thus, if we denote the secrecy rate achievable, using the BES signal $\tilde{X}_\phi$ from equation (2.28) as input, by $R_{\text{BES}}^{\text{RS}}$ then,

$$R_{\text{BES}}^{\text{RS}} = I(\tilde{X}_\phi; \tilde{Y}_\phi|S_1) - I(\tilde{X}_\phi; \tilde{Z}_\phi|S_2), \tag{2.32}$$

where

$$\tilde{Y}_\phi = \sqrt{S_1}\tilde{X}_\phi + U_1, \quad \tilde{Z}_\phi = \sqrt{S_2}\tilde{X}_\phi + U_2. \tag{2.33}$$

[4] In the sequel, $\tilde{(.)}$, $(.)_r$ and $(.)_q$ will be used to indicate a BES signal, real component and imaginary/quadrature component of a signals (or symbols), respectively.

The desired lower bound is subsequently obtained by replacing the first and second mutual information terms by a corresponding lower and an upper bound, respectively. While the lower bound follows directly from Lemma 1, the derivation of the upper bound is based on the following concept.

If we denote the rates supported by $\tilde{X}_\phi$ and $\tilde{X}_{\phi^c}$ on the eavesdropper's channel by r_ϕ and r_{ϕ^c}, respectively then, $r_\phi + r_{\phi^c} \leq C_2$, where C_2 is the ergodic capacity of the eavesdropper channel. As a result, a lower bound to r_{ϕ^c} leads to an upper bound to r_ϕ if we use it in the previous inequality, i.e., $r_\phi \leq C_2 - r_{\phi^c}$.

Note that, the BES-RS scheme used in this chapter transmits only via layers belonging to the set ϕ, i.e., only via $\{x_{r,n}, x_{q,n}\}_{n\in\phi}$, where each component in this sequence is IID as $\mathcal{B}(0.5)$. To complete the proof we now construct another antipodal sequence $\{w_{r,n}, w_{q,n}\}_{n\in\phi^c}$ with components IID as $\mathcal{B}(0.5)$. Moreover, this sequence is independent of $\tilde{X}_\phi$ too. We denote by $\tilde{W}_{\phi^c}$, the following BES signal,

$$\begin{aligned}\tilde{W}_{\phi^c} =& \tilde{W}_{r,\phi^c} + j\tilde{W}_{q,\phi^c},\\ =& \frac{\sqrt{3}}{2}\sum_{n\in\phi^c} w_{r,n}2^{-n} + j\frac{\sqrt{3}}{2}\sum_{n\in\phi^c} w_{q,n}2^{-n}.\end{aligned} \tag{2.34}$$

Now, let us consider a hypothetical scenario where the input to the BCoLoE is $(\tilde{X}_\phi + \tilde{W}_{\phi^c})$, which is identically distributed to $\tilde{X}$ in Lemma 1. So, we invoke Lemma 1 with this input, $S = S_k$ and $U = U_k$ for $k = 1, 2$ to obtain that,

$$2\sum_{n\in\phi^c} \mathbb{E}_{S_2}[\hat{r}_{n,d}(S_2)] \leq I(\tilde{W}_{\phi^c}; \tilde{Z}|S_2), \tag{2.35}$$

$$\begin{aligned}2\sum_{n\in\phi} \mathbb{E}_{S_1}[\hat{r}_{n,d}(S_1)] \leq& I(\tilde{X}_\phi; \tilde{Y}|S_1),\\ =& I(\tilde{X}_\phi; \tilde{Y}_\phi + \sqrt{S_1}\tilde{W}_{\phi^c}|S_1),\end{aligned} \tag{2.36}$$

$$\leq I(\tilde{X}_\phi; \tilde{Y}_\phi|S_1), \tag{2.37}$$

where $\tilde{Z}$ and $\tilde{Y}$ represents the corresponding output at the eavesdropper and the legitimate user, respectively, i.e.,

$$\tilde{Y} = \sqrt{S}_1(\tilde{X}_\phi + \tilde{W}_{\phi^c}) + U_1,\ \tilde{Z} = \sqrt{S}_2(\tilde{X}_\phi + \tilde{W}_{\phi^c}) + U_2.$$

In equation (2.36) we have used the notation from equation (2.33). Equation (2.37) is obtained using the fact that additional independent noise at the receiver can not increase mutual information. Next we will use the fact that, the ergodic capacity of a fading channel with no CSI at the transmitters [63] represents the maximum mutual information between the input and output of the channel, maximized over all possible input. Thus, using the expression for the ergodic capacity of the eavesdropper's link we get,

$$\begin{aligned}
\mathbb{E}_{S_2}[\log(1+S_2)] &\geq I(\tilde{X}_\phi + \tilde{W}_{\phi^c}; \tilde{Z}|S_2), \\
&= I(\tilde{X}_\phi, \tilde{W}_{\phi^c}; \tilde{Z}|S_2), \\
&= I(\tilde{W}_{\phi^c}; \tilde{Z}|S_2) + I(\tilde{X}_\phi; \tilde{Z}|S_2, \tilde{W}_{\phi^c}) \\
&= I(\tilde{W}_{\phi^c}; \tilde{Z}|S_2) + I(\tilde{X}_\phi; \tilde{Z}_\phi|S_2) \\
\text{Or, } I(\tilde{W}_{\phi^c}; \tilde{Z}|S_2) &\leq \mathbb{E}_{S_2}[\log(1+S_2)] - I(\tilde{X}_\phi; \tilde{Z}_\phi|S_2). \qquad (2.38)
\end{aligned}$$

Now, substituting (2.35) into equation (2.38) and rearranging the various terms we get,

$$I(\tilde{X}_\phi; \tilde{Z}_\phi|S_2) \leq \mathbb{E}_{S_2}[\log(1+S_2)] - 2\sum_{n\in\phi^c} \mathbb{E}_{S_2}[\hat{r}_{n,d}(S_2)]. \qquad (2.39)$$

Finally, substituting equations (2.37) and (2.39) into equation (2.32) we get,

$$R_{\text{BES}}^{\text{RS}} \geq 2\sum_{n\in\phi} \mathbb{E}_{S_1}[\hat{r}_{n,d}(S_1)] + 2\sum_{n\in\phi^c} \mathbb{E}_{S_2}[\hat{r}_{n,d}(S_2)] - \mathbb{E}_{S_2}[\log(1+S_2)].$$

2.5.2. Proof of Theorem 3

We prove this theorem by finding further lower bounds to the first two sum-rate terms in the lower bound provided in Theorem 2. This results in a looser lower bound to R_{BES}. However, on one hand, it provides a more easily computable lower bound for arbitrary fading distributioins and on the other hand, it's form makes it easier to compare with the upper bound in Theorem 1, which will be used later to prove an approximate secrecy capacity result. The proof is carried out in two steps, in which we find lower bounds to the first and second terms on the RHS of equation (2.19), respectively.

Step 1: Using the definition from equation (A.4) we get,

$$\begin{aligned}\sum_{n\in\phi}\mathbb{E}_{S_1}[\hat{r}_{n,d}(S_1)] &= \sum_{n\in\phi}\mathbb{E}_{S_1}\Big[1-H[\hat{\epsilon}_d(a_n(\gamma))]\Big],\\ &= \sum_{n\in\phi}\int_0^\infty f_{S_1}(\gamma)\left(1-H\Big[\hat{\epsilon}_d(a_n(\gamma))\Big]\right)d\gamma,\\ &\geq \sum_{n\in\phi}\int_{\gamma_n}^\infty f_{S_1}(\gamma)(1-H[\hat{\epsilon}_d(a_n(\gamma))])d\gamma, \qquad (2.40)\\ &\geq \sum_{n\in\phi}\bar{F}_{S_1}(\gamma_n)-\sum_{n\in\phi}2\delta_n(S_1,\gamma_n), \qquad (2.41)\end{aligned}$$

where equation (2.40) follows from the non-negativity of the integrand and γ_n's as defined in equation (2.14). In equation (2.41) we have used the definition of the CCDF of S_1 and the identity from equation (2.22).

Step 2: Now, for the second term in (2.19), again using the definition from (A.4) we get

$$\begin{aligned}\sum_{n\in\phi^c}\mathbb{E}_{S_2}[\hat{r}_{n,d}(S_2)] &\geq \sum_{n\in\phi^c}\mathbb{E}_{S_2}\Big[1-H[\hat{\epsilon}_d(a_n(\gamma))]\Big], \qquad (2.42)\\ &= \sum_{n\in\phi^c}\int_0^\infty f_{S_2}(\gamma)(1-H[\hat{\epsilon}_d(a_n(\gamma))])d\gamma,\\ &\geq \sum_{n\in\phi^c}\int_{\gamma_{n+1}}^\infty f_{S_2}(\gamma)(1-H[\hat{\epsilon}_d(a_n(\gamma))])d\gamma, \qquad (2.43)\\ &\geq \sum_{n\in\phi^c}\bar{F}_{S_2}(\gamma_{n+1})-\sum_{n\in\phi^c}\delta_n(S_2,\gamma_{n+1}), \qquad (2.44)\\ &\geq \sum_{n=1}^{\infty}\bar{F}_{S_2}(\gamma_{n+1})-\sum_{n\in\phi}\bar{F}_{S_2}(\gamma_{n+1})-\sum_{n\in\phi^c}\delta_n(S_2,\gamma_{n+1}), \qquad (2.45)\end{aligned}$$

where equation (2.43) follows from the non-negativity of the integrand and γ_{n+1}'s as defined in equation (2.14). In equation (2.44) we have used the definition of the CCDF of S_2 and the identity from equation (2.22). Before proceeding further, we simplify the first term on the right hand side of equation (2.45) as follows:

$$2\sum_{n=1}^{\infty}\bar{F}_{S_2}(\gamma_{n+1}) = 2\sum_{k=2}^{\infty}\bar{F}_{S_2}(\gamma_k),$$

$$\geq \frac{1}{\ln(2)} \sum_{k=2}^{\infty} \int_{\Gamma_k} \frac{\bar{F}_{S_2}(\gamma_k)}{(1+\gamma)} d\gamma, \tag{2.46}$$

$$\geq \frac{1}{\ln(2)} \sum_{k=2}^{\infty} \int_{\Gamma_k} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma, \tag{2.47}$$

$$= \frac{1}{\ln(2)} \int_{\gamma_2}^{\infty} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma,$$

$$= \frac{1}{\ln(2)} \int_{0}^{\infty} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma - \frac{1}{\ln(2)} \int_{0}^{\gamma_2} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma,$$

$$= \mathbb{E}_{S_2}[\log(1+S_2)] - \frac{1}{\ln(2)} \int_{0}^{\gamma_2} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma, \tag{2.48}$$

where in equation (2.46) we have used the fact that $\int_{\Gamma_n} \frac{d\gamma}{1+\gamma} = \ln(1+\gamma_{n+1}) - \ln(1+\gamma_n) \leq 2\ln(2),\ \forall n \geq 1$. Equation (2.47) follows from the fact that $\bar{F}_S(\gamma_n) \geq \bar{F}_S(\gamma),\ \forall \Gamma_n, n \in \mathbb{N}$. In equation (2.48) we have used partial integration and assumed that the CCDF decays faster than the increase of $\log(1+\gamma)$ so that

$$\lim_{\gamma \to \infty} \bar{F}_{S_2}(\gamma) \log(1+\gamma) = 0. \tag{2.49}$$

Now, substituting equations (2.48) into equation (2.45) we get

$$\sum_{n \in \phi^c} \mathbb{E}_{S_2}[\hat{r}_{n,0}(S_2)] \geq \mathbb{E}_{S_2}[\log(1+S_2)] - \int_{0}^{\gamma_2} \frac{\bar{F}_{S_2}(\gamma) d\gamma}{\ln(2)(1+\gamma)} - \sum_{n \in \phi} \bar{F}_{S_2}(\gamma_{n+1}) - \sum_{n \in \phi^c} \delta_n(S_2, \gamma_{n+1}). \tag{2.50}$$

The integral-term in equation (2.50) can be simplified more by breaking up the integral into the two seperate integrals and then further upper bounding the CCDF in one of the integrals as is shown below:

$$\int_{0}^{\gamma_2} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma = \int_{0}^{\gamma_1} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma + \int_{\gamma_1}^{\gamma_2} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma, \tag{2.51}$$

$$\leq \int_{0}^{\rho} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma + \int_{\rho}^{4\rho} \frac{1}{(1+\gamma)} d\gamma, \tag{2.52}$$

$$= \int_{\Gamma_0} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)} d\gamma + \ln \frac{1+4\rho}{1+\rho}, \tag{2.53}$$

where in equation (2.51) we have used the fact that $0 < \gamma_1 < \gamma_2$ as mentioned in equation (2.14). In equation (2.52) we have used the fact that maximum value of $\bar{F}_{S_2}(\gamma)$ is 1 and by definition $\gamma_1 = \rho$

and $\gamma_2 = 4\rho$. Now, substituting the inequality from equation (2.53) into (2.50) we get

$$\sum_{n\in\phi^c} \mathbb{E}_{S_2}[\hat{r}_{n,0}(S_2)] \geq \mathbb{E}_{S_2}[\log(1+S_2)] - \int_0^{\rho} \frac{\bar{F}_{S_2}(\gamma)d\gamma}{\ln(2)(1+\gamma)} - \log\frac{1+4\rho}{1+\rho} - \sum_{n\in\phi} \bar{F}_{S_2}(\gamma_{n+1}) - \sum_{n\in\phi^c} \delta_n(S_2, \gamma_{n+1}). \tag{2.54}$$

Finally, substituting equations (2.41) and (2.54) into equation (2.19) we get,

$$R_{\text{BES}}^{\text{RS}} \geq \sum_{n\in\phi} 2\Big[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})\Big] - 2\sum_{n\in\phi} \delta_n(S_1, \gamma_n) - 2\sum_{n\in\phi^c} \delta_n(S_2, \gamma_{n+1}) - \log(e)\int_{\Gamma_0} \frac{\bar{F}_{S_2}(\gamma)}{(1+\gamma)}d\gamma - \log\frac{1+4\rho}{1+\rho}. \tag{2.55}$$

2.6. Proof of Theorem 4: The Constant Gap Result

As explained in the outline of the proof, it is sufficient to prove that the difference, R_D, between the upper bound in equation (2.12) and the lower bound in (2.21) can not be larger than 11.

Subtracting the RHS of equation (2.21) with $\phi = \mathcal{N}$, from the RHS of (2.12) we get,

$$R_D = 2\sum_{n\in\mathcal{N}} \delta_n(S_1, \gamma_n) + 2\sum_{n\in\mathcal{N}^c} \delta_n(S_2, \gamma_{n+1}) + \log(1+4\rho). \tag{2.56}$$

We will import upper bounds to the first two terms from [56], however, will show the main steps here for completeness. From the definition of $\delta_n(S_k, \alpha)$ provided in (2.22) we have,

$$\begin{aligned}
\delta_n(S_1, \gamma_n) &= \int_{\gamma_n}^{\infty} f_{S_1}(\gamma) H[\hat{\epsilon}_0(a_n(\gamma))]d\gamma, \\
&= \sum_{k=n}^{\infty} \int_{\Gamma_k} f_{S_1}(\gamma) H[\hat{\epsilon}_0(a_n(\gamma))]d\gamma, \\
&\leq \sum_{k=n}^{\infty} \int_{\Gamma_k} f_{S_1}(\gamma) H[\hat{\epsilon}_0(a_n(\gamma_k))]d\gamma, &(2.57)\\
&\leq \sum_{k=n}^{\infty} H[\hat{\epsilon}_0(3\gamma_{k-n})] \Pr\left(S_1 \in \Gamma_k\right), &(2.58)\\
&\leq \sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)] \Pr\left(S_1 \in \Gamma_{j+n}\right), &(2.59)
\end{aligned}$$

where in equation (2.57) we have used the fact that $H[\hat{\epsilon}_0(a_n(\gamma))]$ is a decreasing function of γ. As a result, we have $H[\hat{\epsilon}_0(a_n(\gamma))] \leq H[\hat{\epsilon}_0(a_n(\gamma_k))]$ for all $\gamma \in \Gamma_k$. In equation (2.58) we have used the fact that $H[\hat{\epsilon}_0(a_n(\gamma_k))]$ is a constant with respect to γ and we have used the expression for $a_n(.)$ from equation (2.20f) and equation (2.14). Then, using the inequality from (2.59) we get,

$$\begin{aligned}
\sum_{n\in\mathcal{N}} \delta_n(S_1,\gamma_n) \leq & \sum_{n=1}^{\infty} \delta_n(S_1,\gamma_n), \\
\leq & \sum_{n=1}^{\infty}\sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)] \Pr\left(S_1 \in \Gamma_{j+n}\right), \\
= & \sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)] \sum_{n=1}^{\infty} \Pr\left(S_1 \in \Gamma_{j+n}\right), \\
= & \sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)] \bar{F}_{S_1}(\gamma_{j+1}), \\
\leq & \sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)], \qquad (2.60)
\end{aligned}$$

Similarly, it can be shown that

$$\sum_{n\in\mathcal{N}^c} \delta_n(S_2,\gamma_{n+1}) \leq \sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)]. \qquad (2.61)$$

Substituting equations (2.60) and (2.61) into (2.56) we get

$$R_D \leq \log(1+4\rho) + 4\sum_{j=0}^{\infty} H[\hat{\epsilon}_0(3\gamma_j)], \qquad (2.62)$$

where the RHS is independent of channel statistics and is a function of ρ only, since γ_j's are defined in terms of ρ. So, we can numerically minimize the expression on the RHS of equation (2.62) and see that the minimum value of 11 is attained when $\rho = 20$.

2.7. Fading Gaussian Wiretap Channels: Numeric Examples

In this subsection, we will compute the numeric values of the upper and lower bounds provided in (2.11) and (2.19), respectively and the difference between those two, for several example BCoLoEs. For computational convenience, most of the channel statistics will be assumed to have

finite and small number of distinct fading states. Since the secrecy capacity of *stochastically degraded* wiretap channels can be exactly characterized [25], in this section, we will consider channels which are not *stochastically degraded*. Interestingly, for all the channels considered in this section, the difference between these two bounds turns out to be much smaller than 11 bits.

Given the fading statistics, i.e., $(\bar{F}_{S_1}(.), \bar{F}_{S_2}(.))$ of a BCoLoE, the computation of the upper bound in (2.11) is straightforward. For computing the lower bound in equation (2.19), we first need to determine the optimal ϕ - the set of layers, through which information symbols are sent. Recall from section 2.5 that, in the BES-RS scheme of this chapter we send information symbols only via a selected set of layers. The n-th layer is chosen for transmission, i.e., $n \in \phi$, if the corresponding equivalent rBSC (e.g., see the proof of Lemma 1 in Appendix A.1) can on average send more information to the legitimate user than it can to the eavesdropper. We know from equation (2.20a) that the instantaneous rate of information transmission via layer n is $\hat{r}_{n,d}(s)$, for a particular realization s of the fading state on a PTP channel. Thus, we choose ϕ according to the following rule:

$$\phi = \left\{ n : \mathbb{E}_{S_1}\left[\hat{r}_{n,d}(S_1)\right] \geq \max\left(\tau, \mathbb{E}_{S_2}\left[\hat{r}_{n,d}(S_2)\right]\right) \right\}. \tag{2.63}$$

Here, we also impose the condition that, a layer is only used if it can carry more information than τ bits; in our computations, we choose $\tau = 10^{-3}$. Theoretically, an equivalent rBSC of the BES scheme on a PTP fading channel can support a non-zero rate if its corresponding crossover probability is smaller than 0.5. Clearly, the constraint which requires the average rate is larger than τ, is stricter but more practical than the theoretical constraint.

We also use this criterion to determine the maximum number of layers, at each receiver, which can carry an average information more than τ bits, i.e.,

$$n_k^* = \max\left\{ n : \mathbb{E}_{S_k}\left[\hat{r}_{n,d}(S_k)\right] \geq \tau \right\}, \ k = 1, 2. \tag{2.64}$$

Then define the complement of ϕ as follows:

$$\phi^c = [1, \ \max\{n_1^*, n_2^*\}] \setminus \phi. \tag{2.65}$$

For brevity, we will denote the average rate of information that can be transmitted, via layers in ϕ to the legitimate receiver by $\bar{r}_{\phi}^{1}$ and via layers in ϕ^c to the eavesdropper by $\bar{r}_{\phi^c}^{2}$, i.e.,

$$\bar{r}_{\phi}^{1} = 2 \sum_{n \in \phi} \mathbb{E}_{S_1} \Big[\hat{r}_{n,d}(S_1) \Big], \tag{2.66}$$

$$\bar{r}_{\phi^c}^{2} = 2 \sum_{n \in \phi^c} \mathbb{E}_{S_2} \Big[\hat{r}_{n,d}(S_2) \Big]. \tag{2.67}$$

We will also denote the ergodic capacity of the k-th receiver by C_k, for $k = 1, 2$ and the tighter upper bound to the secrecy capacity, i.e., the RHS of (2.11), and the tighter lower bound to $R_{\text{BES}}^{\text{RS}}$, i.e., the RHS of (2.19), by r_{su} and r_{sl}, respectively.

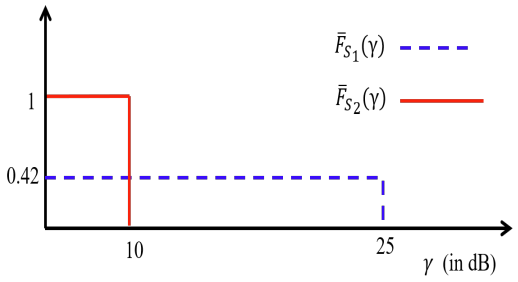

Figure 2.4. Eve's SNR Fixed at 10 dB and Bob's SNR is 25 dB with Probability 0.42.

Example 2 (*Intermittent AWGN Channels) We start with an example where the ergodic capacities of both the PTP links of the BCoLoE are identical and therefore simple Gaussian input can not achieve any positive secrecy rate. We consider a couple of BCoLoEs where the eavesdroppers channel is time-invariant and the legitimate channel has one non-zero random state. Let the Probability Mass Functions (PMFs) of the channel states be as shown in Table 2.1. The corresponding CCDFs for the fading coefficients are depicted in Fig 2.4, substituting these CCDF values in* (2.11) *we get,*

$$\begin{aligned} r_{su} &= 0.42 \big[\log(1 + 10^{2.5}) - \log(1 + 10) \big] = 2.04, \\ \bar{r}_{su} &= \log \frac{(1 + 20)}{(1 + 10)} + 2 \sum_{n=1}^{2} \big[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1}) \big] = 3.01, \end{aligned}$$

On the achievability side, we first compute the values of n_k^* for $k = 1, 2$ and then determine ϕ and ϕ^c for both the cases as shown in Table 2.1. The various components of the lower bound in equation

(2.19) *can be computed as follows:*

$$\bar{r}_{\phi}^{1}= 2(0.42)\sum_{n=3}^{5}\hat{r}_{n,\infty}(10^{2.5}) = 1.51,$$
$$\bar{r}_{\phi^c}^{2}= 2(1)\sum_{n=1}^{2}\hat{r}_{n,2-n}(10) = 2.28,$$
$$C_2= \log(1+10) = 3.46,$$

where $\bar{r}_{\phi}^{1}$ *and* $\bar{r}_{\phi^c}^{2}$ *was defined in* (2.66) *and* (2.67)*, respectively. Substituting these values in equation* (2.19) *we get that a secrecy rate of* $r_{sl} = 0.33$ *bits/channel use or better can be achieved by the BES-RS scheme. It is evident that, the lower and upper bounds are within* 2 *bits from each other. Consequently, we can conclude that the BES-RS scheme can achieve the secrecy capacity of these channels within* 2 *bits.*

Table 2.1. PMFs and Values of Intermediate Parameters for Example 2.

PMF $\backslash s$	0	10	$10^{2.5}$
$P_{S_1}(s)$	0.58	0	0.42
$P_{S_2}(s)$	0	1.0	0

n_1^*	n_2^*	ϕ	ϕ^c	$\bar{r}_{\phi}^{1}$	$\bar{r}_{\phi^c}^{2}$	r_{sl}	r_{su}
5	3	{3,4,5}	{1,2}	1.51	2.28	0.33	2.04

Example 3 (Achievability for Example 1) *Let us consider a BCoLoE where each of the fading link has multiple non-zero states. Assume that the PMF of the two channels are as shown in Table 2.2, the corresponding CCDFs of the channel are shown in Fig. 2.3. Both the upper bounds in Theorem 1 for this channel was also computed and is given in equations* (2.16) *and* (2.17)*, from which we get,* $r_{su} = 1.28$ *bits/channel use. Substituting the CCDFs in equation* (2.64) *we get,* $n_1^* = 5$ *and* $n_2^* = 13$*. However, computation of* $\bar{r}_{\phi}^{1}$ *and* $\bar{r}_{\phi^c}^{2}$ *suggest that we must choose* $\phi = \{1,2,3,4\}$ *and* $\phi^c = \{5,\cdots,13\}$*. Using these parameters we next compute the various components of equation* (2.19) *as follows:*

$$\bar{r}_{\phi}^{1}= 2(.4)\sum_{n=1}^{4}\hat{r}_{n,4-n}(10) + 2(.6)\sum_{n=1}^{4}\hat{r}_{n,4-n}(10^2) = 4.45.$$
$$\bar{r}_{\phi^c}^{2}= 2(.1)\hat{r}_{5,\infty}(10^2) + 2(.1)\sum_{n=5}^{13}\hat{r}_{n,\infty}(10^7) = 1.45.$$
$$C_2= (.8)\log(11) + (.1)[\log(101) + \log(1+10^7)] = 5.76.$$

Finally, substituting these values in (2.19) *we get,* $r_{sl} = 0.31$*. Comparing this lower bound with the upper bound,* $r_{su} = 1.28$*, we see that the upper and lower bounds are within* 1 *bit, which in turn imply that the BES-RS scheme can achieve a secrecy rate within* 1.2 *bits to the secrecy capacity of the channel, as shown in Table 2.2.*

Table 2.2. A BCoLoE with Eavesdropper's Ergodic Capacity Larger than Main Channel Capacity.

PMF$\backslash s$	10	10^2	10^7
$P_{S_1}(s)$	0.4	0.6	0
$P_{S_2}(s)$	0.8	0.1	0.1

n_1^*	n_2^*	ϕ	ϕ^c	$\bar{r}_{\phi}^1$	$\bar{r}_{\phi^c}^2$	r_{sl}	r_{su}
5	13	{1,2,3,4}	{5,$\cdots$,13}	4.45	1.45	0.14	1.28

Example 4 *The CCDFs of the BCoLoE considered in this example is shown in Fig. 2.5 for this example. Substituting the CCDF values into equation* (2.11) *we get*

$$r_{su} = \log(e) \int_{10^2}^{10^5} \frac{(0.7)d\gamma}{1+\gamma} = 6.97.$$

$$\bar{r}_{su} = \log(1+20) - \int_0^{10} \frac{\log(e)d\gamma}{1+\gamma} - .75 \int_{10}^{10^2} \frac{\log(e)d\gamma}{1+\gamma} + 2\sum_{n=2}^{7} \left[\bar{F}_{S_1}(\gamma_n) - \bar{F}_{S_2}(\gamma_{n+1})\right] = 8.63.$$

The PMF for the fading coefficients for this channel is provided in Table 2.3. Substituting the largest fading state from both the links in equation (2.64) *we get,* $n_1^* = 9$ *and* $n_2^* = 4$*. However, computation of* $\bar{r}_{\phi}^1$ *and* $\bar{r}_{\phi^c}^2$ *suggest that we must choose* $\phi = \{3, \cdots, 9\}$ *and* $\phi^c = \{1, 2\}$*. Using these parameters we next compute the various components of equation* (2.19) *as follows:*

$$\bar{r}_{\phi}^1 = 2(.05)\hat{r}_{3,\infty}(10) + 2(.7)\sum_{n=3}^{9} \hat{r}_{n,\infty}(10^5) = 8.31.$$

$$\bar{r}_{\phi^c}^2 = 2(.25)\sum_{n=1}^{2} \hat{r}_{n,2-n}(10) + 2(.75)\sum_{n=1}^{2} \hat{r}_{n,2-n}(10^2) = 3.02.$$

$$C_2 = 0.25\log(1+10) + 0.75\log(1+10^2) = 5.86.$$

Finally, substituting these values in (2.19) *we get,* $r_{sl} = 5.47$*. Interestingly, on the BCoLoE, the difference between the upper and lower bounds are within* 1.5 *bit, which in turn imply that the*

BES-RS scheme can achieve a secrecy rate within 1.5 *bits to the secrecy capacity of the channel, as shown in Table 2.3.*

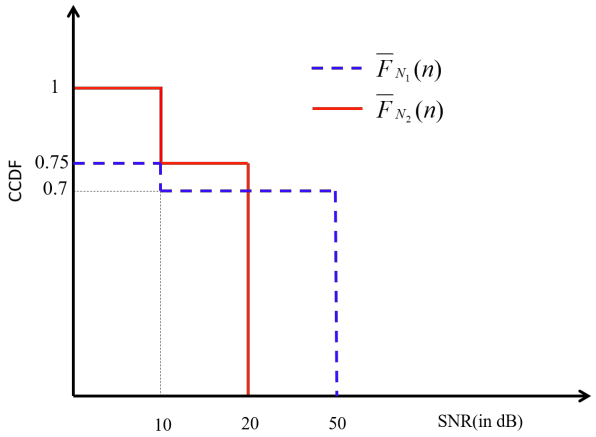

Figure 2.5. CCDFs of CWTC from Example 4

Table 2.3. A BCoLoE with Large Secrecy Capacity.

PMF$\backslash s$	0	10	10^2	10^5
$P_{S_1}(s)$	0.25	0.05	0	0.7
$P_{S_2}(s)$	0	0.25	0.75	0

n_1^*	n_2^*	ϕ	ϕ^c	$\bar{r}_{\phi}^1$	$\bar{r}_{\phi^c}^2$	r_{sl}	r_{su}
4	9	{1,2}	{3,$\cdots$,9}	8.31	3.02	5.47	6.97

Example 5 (*Intermittent AWGN versus Rayleigh Fading Channel) In order to model a practical scenario, next, we consider a BCoLoE where the eavesdropper is assumed to face Rayleigh fading and the legitimate channel has one non-zero SNR with certain probability. Since* $\sqrt{S_2}$ *is Rayleigh distributed, its square has an exponential distribution, which can be expressed as:*

$$\bar{F}_{S_2}(\gamma) = \begin{cases} 1, & \gamma < 0 \\ e^{-\gamma/\Gamma}, & \gamma \geq 0 \end{cases} \tag{2.68}$$

where the average SNR $\Gamma = 16$ *dB for the Rayleigh fading eavesdropper channel. On the other hand, the legitimate channel assume SNRs of* 30 *dB with probability* 0.9 *and* 0 *dB with probability* 0.1. *The CCDFs of the channel are shown in Fig. 2.6. In this case, the upper bound in equation* (2.11) *evaluates to,*

$$r_{su} = \log(e) \int_{4.2}^{10^3} \frac{(0.9 - e^{-(0.03\gamma)})}{1+\gamma} d\gamma = 6.39.$$

On the achievability side, using equation (2.64) *we get,* $n_1^* = 6$. *While theoretically* n_2^* *can be arbitrary large, the expected rate supportable by* n*-th layer for* $n > n_2^* \approx 7$ *is negligible. Computing the average rates supportable by every layer of each user we choose* $\phi = \{1, \cdots, 6\}$ *and* $\phi^c = \{7\}$. *Using these values in* (2.66) *and* (2.67) *we get,*

$$\begin{aligned}
\bar{r}_{\phi}^{1} &= 2(.9)\sum_{n=1}^{6} r_{n,d}(10^3) = 8.20. \\
\bar{r}_{\phi^c}^{2} &= 2\mathbb{E}_{S_2}[r_{7,d}(S_2)] = 0.02. \\
C_2 &= \int_0^d (0.03)e^{-(0.03\gamma)}\log(1+\gamma)d\gamma = 2.7,
\end{aligned} \tag{2.69}$$

where to compute the various components of $\bar{r}_{\phi^c}^{2}$ *in the above equation we have approximated the CCDF,* $\bar{F}_{S_2}(\gamma)$ *by its quantized version as shown using dotted lines in Fig. 2.6. This provides an underestimation of the actual values which can be improved by choosing the step sizes sufficiently small. Finally, substituting these values in* (2.19) *we get the lower bound to* R_{BES}^{RS} *as shown in Table 2.4.*

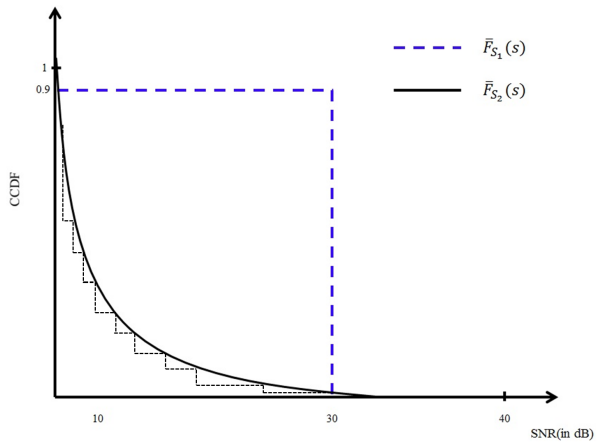

Figure 2.6. CCDFs of Legitimate Receiver and Eavesdropper when Ergodic Capacity of Eavesdropper is Greater than that of the Legitimate Receiver.

Table 2.4. Eve's Channel Rayleigh Distributed.

n_1^*	n_2^*	ϕ	ϕ^c	$\bar{r}_{\phi}^{1}$	$\bar{r}_{\phi^c}^{2}$	r_{sl}	r_{su}
6	7	{1,⋯,6}	{7}	8.20	0.02	5.52	6.39

Example 6 (*Nakagami-m vs Rayleigh Fading Channel) In this final example we consider both the channel to poses practical CCDFs; We assume that the eavesdropper channel faces Rayleigh Fading*

and the Legitimate channel faces Nakagami-m Fading. The CCDF of the legitimate channel is defined as follows

$$\bar{F}_{S_1}(\gamma) = \begin{cases} 1, & \gamma < 0 \\ 1 - F(m, \frac{m}{\Omega}\gamma^2), & \gamma \geq 0 \end{cases} \tag{2.70}$$

where $F(m, \frac{m}{\Omega}\gamma^2)$ *is the incomplete gamma function with shape factor,* $m = 10$ *and spreading factor,* $\Omega = 500$*. These parameters results in an average SNR of* 13.44 *dB for the main channel. The CCDF of Eve's channel is as specified in equation* (2.68) *with average SNR* $\Gamma = 2dB$*. We first compute the secrecy capacity upper bound using equation* (2.11) *as,*

$$r_{su} = \log(e) \int_0^{10^{2.92}} \frac{\left(1 - F(10, 0.02\gamma^2) - e^{-\gamma/\Gamma}\right)}{1+\gamma} d\gamma = 3.12.$$

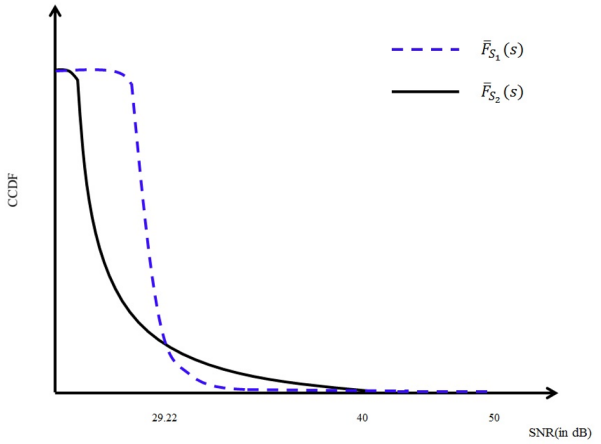

Figure 2.7. CCDFs of Legitimate Receiver and Eavesdropper when Ergodic Capacity of Eavesdropper is greater than that of the Legitimate Receiver.

Since both the CCDFs have infinitely long tails, theoretically, n_k^* *is very large for both* $k = 1, 2$*. However, it turns out that the legitimate receiver can extract non-negligible rates only from layers below* $n_1^* \approx 6$ *and Eve can extract non-negligible rates only from layers below* $n_2^* \approx 7$*. Equations* (2.63) *and* (2.65) *suggests a choice of* $\phi = \{1, \cdots, 6\}$ *and* $\phi^c = \{7\}$*. This choice leads to the following rates:*

$$\bar{r}_{\phi}^1 = 2\sum_{n=1}^{6} \mathbb{E}_{S_1}[r_{n,d}(S_1)] = 3.69.$$

$$\bar{r}_{\phi^c}^2 = 2\mathbb{E}_{S_2}[r_{n,d}(S_2)] = 1.28 * 10^{-4}.$$

$$C_2 = \int_0^{\infty} (0.63)e^{-(0.63\gamma)} \log(1+\gamma)d\gamma = 1.39, \tag{2.71}$$

The average rates - with respect to continuous fading states - in the above equations, are computed using the same method as was stated in Example 5. Comparing the upper and lower bounds as shown in Table 2.5 we see that, for this channel the difference is less than 1 *bit.*

Table 2.5. Nakagami-m vs Rayleigh Fading.

n_1^*	n_2^*	ϕ	ϕ^c	$\bar{r}_{\phi}^1$	$\bar{r}_{\phi^c}^2$	r_{sl}	r_{su}
6	7	$\{1,\cdots,6\}$	$\{7\}$	3.69	$1.28 * 10^{-4}$	2.30	3.12

2.8. Conclusion

In this chapter, we have proved that the BES-RS scheme can achieve the secrecy capacity of an arbitrary BCoLoE within a constant number of bits that is independent of the fading statistics or SNR. Although the new upper bound is looser than a previously derived upper bound, it is easily computable for arbitrary fading distribution. The same holds for the smaller lower bound among the two derived explicit lower bounds to the secrecy rate achievable by the BES-RS scheme on the BCoLoE. The tighter lower and upper bounds however, when computable results in much smaller gap between the two. For all the example BCoLoEs considered in this chapter, which include Rayleigh and Nakagami-m distributions, the difference between the tighter bounds does not exceed 2 bits. This result suggests that there might be a tighter comparable bounds and search of those bounds forms an interesting path of future research. It was shown in [64] in the context of a broadcast channels that, soft-decoding leads to a larger achievable rate for the BES scheme. This result can be used on the BCoLoE as well. However, as mentioned by the authors in [64], the improvement due to soft decision comes at the cost of increased decoding complexity. Moreover, from the numerical computation of the two upper bounds for the example channels, it seems a large part of the gap of 11 bits comes from the looseness of the larger upper bound.

We observed from the simulations that the secrecy capacity that can be achieved is related to the dissimilarity between the channel statistics between the legitimate user and the Eavesdropper. More dissimilarity between the CCDFs results in a larger secrecy capacity.

3. SECRECY CAPACITY OF A CLASS OF BINARY INTERFERENCE CHANNEL

3.1. Introduction

Wireless channels although very attractive due to its mobile nature and ease of implementation suffers from an accute problem related to the security of information being transmitted. Due to the broadcast nature of wireless medium anything that is being transmitted not only reaches the intended receiver but can also reach an unintended receiver, the problem becomes severe when the nature of the information being transmitted is highly sensitive. Information theoretic approach to secure information transfer first appeared in Shannon's work in [26]. It was followed by Wyner's pioneering work on wiretap channel (WTC) [3] where he characterized the secrecy capacity of a *degraded* WTC. The result of Wyner was later generalized by Csiszár and Körner in [28]. After that followed a plethora of work for various WTC models such as Gaussian model [29], WTC with multiple antennas [32],[31],[30], slow fading WTC [65],[33]. In literature we can also find work on fast fading WTC with full channel state information at transmitters (CSIT)[66],[50],[67] or WTC with fixed legitimate channel and fast fading eavesdropper channel with no CSIT [68] or fast fading channel with no CSIT for any of the channels [69],[25]. The survey paper on WTC [70] summarises most of the significant contributions in this regard.

In more recent times efforts has been made to study the security in more complex multi-user network models such as Multiple Access Channels (MAC) [71],[72],[73],[74], Z-Interference Channel (ZIC) [75],[76] and Interference Channel (IC) [77],[78],[79]. Both the ZIC and IC are typically more practical network models. For example, when there are two users communicating in adjacent cells via there respective base stations and are close enough, then they are bound to see interference from each other giving a practical 2-user IC. On the other hand if one of the users is in between the two base stations such that both the base stations can hear it whereas the other user is close to one of the base station but far away from the other such that only the closer base station can hear it then we get a practical 2-user ZIC. A ZIC channel is essentially an IC where one of the interfering links is absent. Due to the inherent complexity in analysing an IC, a coarser metric known as *secure*

degrees of freedom has been used to find the approximate capacity in[80],[81]. However, the results obtained so far are mostly for channels having fixed gain. In a wireless environment where the surroundings is constantly varying it is not practical to assume that the channels remain constant throughout the transmission process. In addition the change in the wireless environment might be too fast to provide a timely feedback to the transmitters about the channel states, hence the transmitters cannot be expected to be aware of the instantaneous channel states and adjust their transmission strategy accordingly. Furthermore, an unintended receiver - if malicious, cannot be expected to provide feedback about its channel state even if it can. So finally taking into account all these factors, we can safely comment that a fast fading wireless channel with no channel state information at the transmitter (CSIT) captures all the practical assumptions needed to describe a wireless channel model. Surprisingly, there has been very few results so far in finding the capacity of such an IC without secrecy constraint and no results at all with secrecy constraint. Our result on binary fading 2-user IC is a step towards filling that void in finding the secrecy capacity for a more general Gaussian 2-user fading IC. The advantages of studying a binary fading model is, it provides a simple physical layer representation for packet wireless network. It is also the most simplistic initial approach in studying the more general fading Gaussian model. This kind of technique is inspired from the determinstic approach introduced in [82] and was used in the study of approximate capacity of fading Gaussian broadcast channel [83], fading Gaussian ZIC channel [24] without secrecy constraint and to find the capacity of certain binary ICs [84] as a first step to the study of fading Gaussian IC.

The rest of the chapter is organised as follows. Section 3.2 formally describes the channel model followed by the introduction of notations used in this chapter. Section 3.3 presents the main findings of this chapter.We further provide some remarks to aid form the intuitive picture about the achievability of the secrecy capacity region for a *very weak* interference channel the definition of which along with a few more will be given in the begining of the section. We state some lemmas and prove them in Section 3.4. These lemmas are later used for the proof of our main results. The proof of the secrecy capacity result for *strong* interference channel is provided in Section 3.5. The proof of the secrecy capacity region result for *very weak* interference channel is divided into two sections where Section 3.6 provides the strategy to achieve the various rate pairs which defines the secrecy capacity region whereas Section 3.7 derives the upper bound of all the rates. This

section also defines two regions and shows that the secrecy capacity region enclosed by the upper bounds from those two regions matches with the achievable secrecy capacity region in Section 3.6. In Section 3.8 we conclude the chapter.

3.2. Channel Model and Some Preliminaries

We consider a 2-user Binary Fading Interference Channel (BFIC) as illustrated in figure 3.1. The channel co-efficient from transmitter Tx_i to receiver Rx_j at time instant t is denoted by $(N_{ij})_t, i, j \in \{1, 2\}$. We assume the channel co-efficients are either 0 or 1, i.e. $N_{ij} \in \{0, 1\}$, and they are distributed as Bernoulli random variables independent from each other and over time. Furthermore we consider the channel co-efficients to be distributed as follows

$$N_{ii} \sim \mathcal{B}(\bar{\epsilon}_{ii}) \text{ and } N_{ij} \sim \mathcal{B}(\bar{\epsilon}_{ij}), \tag{3.1}$$

for $0 \leq \bar{\epsilon}_{ii}, \bar{\epsilon}_{ij} \leq 1, i, j \in \{0, 1\}, i \neq j$. We define $\bar{\epsilon}_{ii} = 1 - \epsilon_{ii}$ and $\bar{\epsilon}_{ij} = 1 - \epsilon_{ij}$.

At each time instant t, the transmit signal at Tx_1 is denoted as W_t whereas that at Tx_2 is denoted as X_t where $W_t, X_t \in \{0, 1\}$ and the received signals at Rx_1 and Rx_2 are respectively denoted as follows

$$\begin{aligned} Y_t &= (N_{11})_t W_t \oplus (N_{21})_t X_t, \\ Z_t &= (N_{12})_t W_t \oplus (N_{22})_t X_t, \end{aligned} \tag{3.2}$$

where all algebraic operations are in a binary field, denoted as, $\mathbb{F}_2$. The channel state information (CSI) at time instant t is denoted by the quadruple

$$N_t = \Big\{(N_{11})_t, (N_{12})_t, (N_{21})_t, (N_{22})_t\Big\}, \tag{3.3}$$

In this chapter we use capital letters to denote random variables (RVs), e.g. $(N_{ij})_t$ is a random variable at time instant t, and small letters denote the realizations, e.g. $(n_{ij)_t}$ is a realization of $(N_{ij})_t$. For a natural number m, the vector N^m represents the following

$$N^m = \Big[N_1, N_2, \cdots, N_m\Big]^T, \tag{3.4}$$

Finally, we write the t - length output vectors as follows

$$(N_{11}W \oplus N_{21}X)^t = \Big[(N_{11}W \oplus N_{21}X)_1, \cdots, (N_{11}W \oplus N_{21}X)_t\Big]^T,$$
$$(N_{12}W \oplus N_{22}X)^t = \Big[(N_{12}W \oplus N_{22}X)_1, \cdots, (N_{12}W \oplus N_{22}X)_t\Big]^T, \tag{3.5}$$

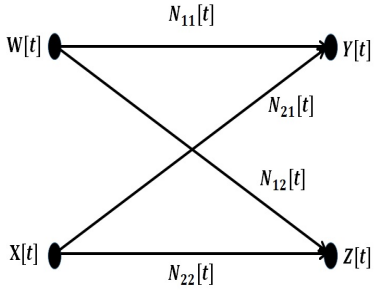

Figure 3.1. 2-User Binary Fading Interference Channel.

We next state the secrecy criterion to be followed in this chapter after defining the decoding error probability. We assume that both the transmitters has secret messages to transmit. Let $\mathcal{M}_i$ be the secret message to be transmitted by transmitter Tx_i where $i = 1, 2$. The transmitters want to transmit at a rate R_i. Now let us suppose that Tx_1 wants to transmit a message $\mathcal{M}_1(k), k \in \{1, 2, \cdots, 2^{nR_1}\}$, then the transmitter chooses a codeword W^n from its codebook $\mathcal{C}_1(n)$. Similarly if Tx_2 wants to transmit a message $\mathcal{M}_2(l), l \in \{1, 2, \cdots, 2^{nR_2}\}$ it does so by choosing a codeword X^n from its codebook $\mathcal{C}_2(n)$. Now both the receivers are assumed to be aware of the coding scheme and the codebooks used by the transmitters. Next let us assume that $\hat{\mathcal{M}}_i$ is the estimate Rx_i makes about the message transmitted from Tx_i on observing the received signal, where $i \in \{1, 2\}$.An error occurs when $\hat{\mathcal{M}}_i \neq \mathcal{M}_i$ and the probability of decoding error is given by

$$\lambda_i = P[\hat{\mathcal{M}}_i \neq \mathcal{M}_i], \quad i \in \{1, 2\}, \tag{3.6}$$

A rate pair (R_1, R_2) is said to be *achievable* if there exists $\mathcal{C}_i(n), i \in \{1, 2\}$ such that $\max(\lambda_1, \lambda_2) \to 0$ and both

$$I(\mathcal{M}_1; Z^n | X^n, N^n) < \delta_1, \text{ and}$$
$$I(\mathcal{M}_2; Y^n | W^n, N^n) < \delta_2, \tag{3.7}$$

is satisfied with arbitrarily small $\delta_1 > 0$ and $\delta_2 > 0$, as $n \to \infty$. Note that the aforementioned conditions are equivalent to an *equivocation* of 1 according to the definitions of [28], since

$$I(\mathcal{M}_1, Z^n | X^n, N^n) < \delta_1,$$

$$\implies h(\mathcal{M}_1) < h(\mathcal{M}_1 | Z^n, X^n, N^n) + \delta_1, \tag{3.8}$$

$$\frac{h(\mathcal{M}_1 | Z^n, X^n, N^n)}{h(\mathcal{M}_1)} > 1 - \delta_1^{'}, \tag{3.9}$$

where the left hand side of the above equation represents the *equivocation* and $\delta_1^{'} = \frac{\delta_1}{h(\mathcal{M}_1)}$ can be made arbitrarily small. Similarly the other condition in (3.7) can be shown to have an *equivocation* of 1 as well. The main contributions of the chapter are stated in the next section.

3.3. Main Results

This section provides the exact characterization of secrecy capacity region for a two user *strong* and *very weak* BFIC. We first provide with some definitions.

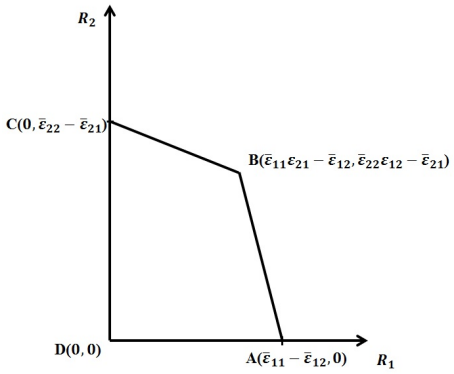

Figure 3.2. The Secrecy Capacity Region of *Very Weak* Binary Fading Interference Channel.

Definition 10 *A fading interference channel is called strong if both the interfering links are probabilistically stronger than their corresponding direct links, mathematically, for BFIC it can be represented as follows :* $\bar{\epsilon}_{ij} > \bar{\epsilon}_{ii}, i \in \{0,1\}, i \neq j$, *where* $\bar{\epsilon}_{ii}$ *is the probability that the direct link is present whereas* ϵ_{ii} *is the probability that the direct link is erased.Similar interpretations can be done for* $\bar{\epsilon}_{ij}$ *and* ϵ_{ij}. *It is called a moderately weak binary fading interference channel if both the direct links are probablistically stronger than their corresponding cross link, mathematically* $\bar{\epsilon}_{ii} \geq \bar{\epsilon}_{ij}, i \in \{0,1\}, i \neq j$. *However if the fraction of the direct link when it does not face any interference is still greater than its corresponding cross link then we say that the channel is a weak*

interference channel, mathematically $\bar{\epsilon}_{ii}\epsilon_{ji} \geq \bar{\epsilon}_{ij}, \forall i,j \in \{1,2\}$ *and* $i \neq j$*. Finally a weak interference channel for which* $\beta_2 > \beta_1$ *is called a very weak interference channel, where* β_2 *and* β_1 *are defined as follows*

$$\beta_1 = \frac{\bar{\epsilon}_{11}\bar{\epsilon}_{21}}{\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21}} \quad \textit{and} \quad \beta_2 = \frac{\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}}{\bar{\epsilon}_{22}\bar{\epsilon}_{12}}, \tag{3.10}$$

With the definitions of β_1 and β_2 as shown in (3.10) we next characterize the secrecy capacity for a *very weak* BFIC in theorem 5 while Lemma 2 gives the secrecy capacity for a *strong* fading interference channel.

Theorem 5 *The secrecy capacity region of a very weak two-user BFIC with no channel state information at transmitter (CSIT),* $\mathcal{C}(\bar{\epsilon}_{11}, \bar{\epsilon}_{12}, \bar{\epsilon}_{21}, \bar{\epsilon}_{22})$ *is the quadrangular region ABCD as shown in figure 3.2 where the co-ordinates of the vertices are as follows*

$$D(0,0);\ A(\bar{\epsilon}_{11} - \bar{\epsilon}_{12}, 0);\ B(\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}, \bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21});\ C(0, \bar{\epsilon}_{22} - \bar{\epsilon}_{21}), \tag{3.11}$$

Remark 1 *The achievability of the rates in points A and C are very intuitive. It just shows that if one of the transmitter is silent then for the other transmitter the interference channel simply reduces to a wiretap channel and it can then use capacity optimal layered erasure wiretap channel code with a single layer to achieve the rates. However it might get a little tricky when both the transmitters are transmitting. Although, it might not be difficult to visualise how the rates in point B can be achieved. Let us take a closer look at figure 3.2, we see that as we move from point A towards point B the rate of transmitter* Tx_1 *slowly decreases while that for transmitter* Tx_2 *increases, which means as* Tx_2 *begins its transmission* Tx_1 *can no longer transmit on its direct link at a rate* $\bar{\epsilon}_{11}$ *and still have its message reliably decoded at its corresponding receiver but must suffer some loss. This makes sense because as* Tx_2 *starts its transmission it will interfere with the signal coming from* Tx_1 *at receiver* Rx_1 *and hence* Rx_1 *cannot decode the whole signal coming from* Tx_1*, if it still transmits at a rate* $\bar{\epsilon}_{11}$*. So we clearly see that unless* Tx_1 *reduces its rate of transmission the signal that it sends cannot be completely decoded by* Rx_1*. But the question is how much should the reduction be. Intuitively it makes sense to think that if* Tx_1 *can transmit during the portion when the interfering link is erased then it reaches* Rx_1 *uninterfered and* Rx_1 *can decode all the information. So the loss in rate that* Tx_1 *suffers is when both direct link and the interfering link to* Rx_1 *is present, i.e.* $\bar{\epsilon}_{11}\bar{\epsilon}_{21}$*. So if its rate of transmission, is* $\bar{\epsilon}_{11} - \bar{\epsilon}_{11}\bar{\epsilon}_{21} = \bar{\epsilon}_{11}\epsilon_{21}$ *then the message can be completely*

decoded by Rx_1. Similar loss in rate will be incurred by Tx_2 as well due to interference of signal from Tx_1 at Rx_2. The secrecy rate achievability is once again based on the usage of optimal code for binary erasure wiretap channel because once a receiver decodes the signal coming from its direct transmitter and subtracts it from the received signal there exists a single layer wiretap channel from the point of view of the other transmitter.

We next state the result for *strong* BFIC in lemma 2.

Lemma 2 *The secrecy capacity of a strong two-user BFIC as shown in figure 3.1 with no channel state information at the transmitter (CSIT) is zero.*

In the following section we explain first the concept of "alignment" and how it does not affect the secrecy capacity region of an interference channel. The section further includes some lemmas along with their proofs which are used later in section 3.5 and 3.7.

3.4. Key Lemmas

In this section we state and prove the lemmas which aids in the proof of the result for *strong* interference channel (lemma 2) and the proof of the converse for *very weak* interference channel (theorem 5). But prior to that we introduce the concept of "alignment".Since the two decoders at the two receivers operate independently so the secrecy capacity region of any interference channel depends only on the marginal distribution of the outputs conditioned on the inputs but not on the joint conditional distribution [1]. With this knowledge we assume for the rest of the chapter that the fading states $(N_{11})_t$ and $(N_{12})_t$ are "aligned" with each other and so is $(N_{22})_t$ with $(N_{21})_t$ such that $P\Big[(N_{ii})_t \cdot (N_{ij})_t = 1\Big] = \min(\bar{\epsilon}_{ii}, \bar{\epsilon}_{ij})\ \forall\ t \geq 0$ and $i, j \in \{1, 2\}, i \neq j$. However the channel states $\{(N_{11})_t, (N_{12})_t\}$ and $\{(N_{22})_t, (N_{21})_t\}$ are independent of each other and the inputs for all t. The above assumption means if the realization of the weaker channel state among $(N_{ii})_t$ and $(N_{ij})_t$ is 1 then the stronger one has to be 1 for all $i, j \in \{1, 2\}, i \neq j$. This however does not change the capacity region as the marginal distribution is not affected by this assumption[24]. We next state and prove the lemmas.

Lemma 3 *Consider n uses of a memoryless channel described by an arbitrary random transformation $P_{Y,Z,T|X,S}$. Let X^n and S^n be the independent input and state sequences respectively. Then the difference of the n-letter entropies can be written as a summation of single letter entropies as*

follows

$$h(Z^n|T^n,S^n)-h(Y^n|T^n,S^n)=\sum_{i=1}^{n}\Big[h(Z_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)-h(Y_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\Big]. \quad (3.12)$$

Proof 3 *The proof of the lemma involves writing the difference of n-letter entropies as a summation of difference of entropies and simple application of chain rule of entropy . The proof is as follows*

$$\begin{aligned}
h(Z^n|T^n,S^n)-h(Y^n|T^n,S^n)&=\sum_{i=1}^{n}\Big[h(Z^i,Y_{i+1}^n|T^n,S^n)-h(Z^{i-1},Y_i^n|T^n,S^n)\Big], &(3.13)\\
&=\sum_{i=1}^{n}\Big[h(Z_i,Z^{i-1},Y_{i+1}^n|T^n,S^n)-h(Z^{i-1},Y_i,Y_{i+1}^n|T^n,S^n)\Big],\\
&=\sum_{i=1}^{n}\Big[h(Z^{i-1},Y_{i+1}^n|T^n,S^n)+h(Z_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\\
&\qquad -h(Z^{i-1},Y_{i+1}^n|T^n,S^n)-h(Y_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\Big], &(3.14)\\
&=\sum_{i=1}^{n}\Big[h(Z_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)-h(Y_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\Big], &(3.15)
\end{aligned}$$

where in equation (3.13) *if the summation is completed by finding the terms for each value of i then we get the difference of n-letter entropy term to the left, equation* (3.14) *follows from simple application of chain rule of entropy. This concludes the proof of the lemma.*

Lemma 4 *Consider the Interference Channel as shown in figure 3.1. If the interference channel satisfies the following constraints* $\bar{\epsilon}_{11}\geq\bar{\epsilon}_{12}$ *and* $\bar{\epsilon}_{22}\geq\bar{\epsilon}_{21}$ *and are "aligned" ,i.e. a realization* 1 *for the weaker channel state would ensure that stronger one is also* 1 *then the input forms the following Markov Chain*

$$W_i\rightarrow\Big\{(lW)^{i-1},\mu(W^n,N^n)\Big\}\rightarrow\Big\{(sW)^{i-1},\mu(W^n,N^n)\Big\},$$

where $l=(N_{11}\bar{N}_{21})$ *and* $s=(N_{12}\bar{N}_{22})$.

Proof 4 *In general* W_i *is correlated to* $\Big\{(lW)^{i-1},\mu(W^n,N^n)\Big\}$ *since the input is not necessarily independently distributed over the several channel uses. So that explains the first part of the chain. Now we know that* $\bar{\epsilon}_{22}\geq\bar{\epsilon}_{21}$ *which implies that* $\epsilon_{22}<\epsilon_{21}$. *Besides since* $\bar{\epsilon}_{11}\geq\bar{\epsilon}_{12}$ *so we can safely*

say that $\bar{\epsilon}_{11}\epsilon_{21} \geq \bar{\epsilon}_{12}\epsilon_{22}$. *Now due to the alignment of the channels we can write*

$$\Big\{(sW)^{i-1}, \mu(W^n, N^n)\Big\} = f\left(\Big\{(lW)^{i-1}, \mu(W^n, N^n)\Big\}\right), \tag{3.16}$$

since we can obtain $\Big\{(sW)^{i-1}, \mu(W^n, N^n)\Big\}$ *from* $\Big\{(lW)^{i-1}, \mu(W^n, N^n)\Big\}$ *by replacing with* 0 *that portion of the signal for which* l *is* 1 *but* s *is not. Next using the definition of data processing inequality we know that for any two sets of correlated random variables* $\{A, B\}$ *and an arbitary function* $f(\cdot)$, *forms a Markov Chain,* $A \rightarrow B \rightarrow f(B)$ *thereby completing the proof of the Markov Chain*

$$W_i \rightarrow \Big\{(lW)^{i-1}, \mu(W^n, N^n)\Big\} \rightarrow \Big\{(sW)^{i-1}, \mu(W^n, N^n)\Big\}, \tag{3.17}$$

We thus conclude the proof of the lemma.

We provide the proof of lemma 2 in the next section.

3.5. Proof of Lemma 2

The proof of the lemma invovles finding an upper bound on the rates, at which if transmitted from the respective transmitters, information can be kept secret from the interfering receiver while the intended receiver can reliably decode it. Let R_i be the rate achievable by transmitter Tx_i over n channel uses, then

$$nR_i = h(\mathcal{M}_i), i = 1, 2, \tag{3.18}$$

where M_i denotes the message at transmitter Tx_i. The derivation of the upper bound for R_1 goes as follows

$$\begin{aligned}
nR_1 - \delta_1 &= h(\mathcal{M}_1) - \delta_1, && (3.19)\\
&\leq h(\mathcal{M}_1|Z^n, X^n, N^n), && (3.20)\\
&\leq h(W^n, \mathcal{M}_1|Z^n, X^n, N^n), && (3.21)\\
&= h(W^n|Z^n, X^n, N^n) + h(\mathcal{M}_1|W^n, Z^n, X^n, N^n), && (3.22)\\
&= h(W^n|Z^n, X^n, N^n), && (3.23)\\
&= h(W^n|X^n, N^n) - I(W^n; Z^n|X^n, N^n),\\
&= h(W^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (3.24)\\
&= h(W^n|N^n) - h(W^n|Y^n, N^n) + h(W^n|Y^n, N^n) - I(W^n; \tilde{Y}^n|N^n),
\end{aligned}$$

$$\leq I(W^n;Y^n|N^n) - I(W^n;\tilde{Y}^n|N^n) + \delta^{'}, \tag{3.25}$$

$$\implies nR_1 - \delta_1^{'} = I(W^n;Y^n|N^n) - I(W^n;\tilde{Y}^n|N^n), \tag{3.26}$$

$$\leq I(W^n;\hat{Y}^n|N^n) - I(W^n;\tilde{Y}^n|N^n), \tag{3.27}$$

$$= h(\hat{Y}^n|N^n) - h(\tilde{Y}^n|N^n), \tag{3.28}$$

$$= \sum_{i=1}^{n}\left[h\Big\{\hat{Y}_i|\hat{Y}^{i-1},\tilde{Y}_{i+1}^n,N^n\Big\} - h\Big\{\tilde{Y}_i|\hat{Y}^{i-1},\tilde{Y}_{i+1}^n,N^n\Big\}\right], \tag{3.29}$$

$$= \sum_{i=1}^{n}\left[h\Big\{\hat{Y}_i|\mathcal{D}_i,N_i\Big\} - h\Big\{\tilde{Y}_i|\mathcal{D}_i,N_i\Big\}\right], \tag{3.30}$$

$$= \sum_{i=1}^{n}\Big[\bar{\epsilon}_{11} - \bar{\epsilon}_{12}\Big]h\Big\{W_i|\mathcal{D}_i,N_i\Big\}, \tag{3.31}$$

$$\leq 0, \tag{3.32}$$

where equation (3.19) follows from (3.18),the secrecy criterion for $Tx_1 - Rx_1$ pair as mentioned in equation (3.8) gives equation (3.20), equation (3.21) is the result of the fact that additional random variable does not reduce entropy. Equation (3.22) follows from the chain rule of entropy. Equation (3.23) occurs because $h(\mathcal{M}_1|W^n,Z^n,X^n,N^n) = 0$ since a receiver is assumed to reliably decode the message that was sent by its corresponding transmitter from the received signal. Equation (3.24) follows from the independence of random variables and $\tilde{Y}^n = (N_{12}W)^n$, $h(W^n|Y^n,N^n) \leq \delta^{'}$ in equation (3.25) and $\delta_1^{'} = \delta_1 + \delta^{'}$ in equation (3.26), Equation (3.27) follows from the fact that independent additive noise cannot increase mutual information, $\hat{Y}^n = (N_{11}W)^n$ in the same equation. $h(\hat{Y}^n|W^n,N^n) = 0$ and $h(\tilde{Y}^n|W^n,N^n) = 0$ results in equation (3.28). Equation (3.29) follows from lemma 3, $\mathcal{D}_i = \{\hat{Y}^{i-1},\tilde{Y}_{i+1}^n,N^{i-1},N_{i+1}^n\}$ in equation (3.30). Since for strong interference $\bar{\epsilon}_{11} \leq \bar{\epsilon}_{12}$ and the entropy of a binary random variable is always non-negative hence we can upper bound (3.31) by zero giving equation (3.32).

Similarly it can be proved that for strong interference the secret rate of transmission from Tx_2 can be upper bounded by 0 as is done for Tx_1 in equation (3.32). So we find that there is no non-negative upper bound to the secret rate of transmission for a *strong* BFIC from either of the transmitters.Thus the overall secrecy rate for a single layer strong fading interference channel is zero. This concludes the proof of lemma 2. We next provide the proof of theorem 5, however the proof is subdivided into a converse and an achievability seperately. The next section provides the achievability part of the proof while section 3.7 provides the converse.

3.6. Achievability

In this section we prove the achievability of the secrecy capacity region given by theorem 5. We see that proving the achievability of the corner points in figure 3.2 would ensure the achievability of the entire region as the rest of the rate pairs on the boundary can be achieved by time sharing. The achievability of point D is trivial and is hence ignored here.

3.6.1. Achievability of Points A and C

When $Tx_2(Tx_1)$ remains silent then BFIC essentially reduces to a single layer erasure wiretap channel where $Rx_2(Rx_1)$ behaves as the eavedropper and $Rx_1(Rx_2)$ as the legitimate receiver. In this scenario using the scheme in [69] $Tx_1(Tx_2)$ can achieve a rate $\{\bar{\epsilon}_{11}-\bar{\epsilon}_{12}\}(\{\bar{\epsilon}_{22}-\bar{\epsilon}_{21}\})$ while $Tx_2(Tx_1)$ remains silent, hence has a zero rate.This ensures the achievability of point $A(C)$.

3.6.2. Achievability of Point B

Let the transmitted signal from both transmitters 1 and 2 be Bernoulli(0.5) distributed as $W = \tilde{W} \sim \mathscr{B}(\frac{1}{2})$ and $X = \tilde{X} \sim \mathscr{B}(\frac{1}{2})$. We will first try to find the rates supported by the direct links when the receivers use the concept of *treating interference as erasure* to reliably decode the signals coming from their corresponding transmitter. Let $r_1^{'}$ represent the rate between $Tx_1 - Rx_1$ and $r_2^{'}$ between $Tx_2 - Rx_2$. So $r_1^{'}$ can be calculated as follows

$$
\begin{aligned}
r_1^{'} &= I(\tilde{W}; \dot{Y}, N), && (3.33)\\
&= I(\tilde{W}; N) + I(\tilde{W}; \dot{Y}|N), && (3.34)\\
&= I(\tilde{W}; \dot{Y}|N), && (3.35)\\
&= \bar{\epsilon}_{21} I(\tilde{W}; N_{11}\tilde{W} \oplus \tilde{X}|N) + \epsilon_{21} I(\tilde{W}; N_{11}\tilde{W}|N),\\
&= \epsilon_{21} I(\tilde{W}; N_{11}\tilde{W} \oplus \tilde{X}|N), && (3.36)\\
&= \bar{\epsilon}_{11}\epsilon_{21} I(\tilde{W}; \tilde{W}),\\
&= \bar{\epsilon}_{11}\epsilon_{21}, && (3.37)
\end{aligned}
$$

where in equation (3.33) $\dot{Y} = N_{11}\tilde{W} \oplus N_{21}\tilde{X}$, equation (3.34) is the result of chain rule of mutual information, equation (3.35) follows from the fact that the inputs are independent of the channel states, equation (3.36) follows from the fact that since interference is treated as erasure so whenever $N_{21} = 1$, the signal is just ignored. Finally equation (3.37) results from the fact that $\tilde{W} \sim \mathscr{B}(\frac{1}{2})$.

Similarly it can be shown that $r_2^{'} = \bar{\epsilon}_{22}\epsilon_{12}$. From [28] we know that on a wiretap channel with input A, legitimate receiver output B, eavesdropper signal C and a channel $P_{BC|A}(\cdot)$, the secrecy capacity achievable is given as

$$\max_{J \to A \to BC} \{I(J;B) - I(J;C)\}, \tag{3.38}$$

It implies for each choice of (J, A) that satisfies the above Markov Chain , a secrecy rate of $\{I(J;B) - I(J;C)\}$ can be achieved on a wiretap channel. Using this result and the fact that interference channel can be thought of as a combination of two wiretap channels, from the point of view of each transmitter, we can find the achievable secrecy rate at each transmitter. Using $J = A \equiv \tilde{W}$, $B \equiv (\dot{Y}, N)$ and $C \equiv (\dot{Z}, \tilde{X}, N)$ where $\dot{Z} = N_{22}\tilde{X} \oplus N_{12}\tilde{W}$ and denoting the corresponding achievable secrecy rate by r_1 we find the achievable secrecy rate at Tx_1 as follows,

$$\begin{aligned}
r_1(\omega) &= I\Big\{\tilde{W}; \dot{Y}, N\Big\} - I\Big\{\tilde{W}; \dot{Z}, \tilde{X}, N\Big\}, \\
&= I\Big\{\tilde{W}; N\Big\} + I\Big\{\tilde{W}; \dot{Y}|N\Big\} - I\Big\{\tilde{W}; N\Big\} - I\Big\{\tilde{W}; \tilde{X}|N\Big\} - I\Big\{\tilde{W}; \dot{Z}|\tilde{X}, N\Big\}, && (3.39) \\
&= I\Big\{\tilde{W}; \dot{Y}|N\Big\} - I\Big\{\tilde{W}; \dot{Z}|\tilde{X}, N\Big\}, && (3.40) \\
&= \bar{\epsilon}_{11}\epsilon_{21} - I(\tilde{W}; N_{12}\tilde{W}|N), && (3.41) \\
&= \bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}, && (3.42)
\end{aligned}$$

where equation (3.39) follows from the chain rule of mutual information, the independence of the channel states with inputs and the inputs among themselves results in equation (3.40), equation (3.41) follows from (3.37). Finally equation (3.42) follows from the fact that $\tilde{W} \sim \mathscr{B}(\frac{1}{2})$. Similarly it can be shown using similar techniques and the result for wiretap channel that achievable secrecy rate for Tx_2 is $\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21}$. The only change in this case will be the fact that $J = A \equiv \tilde{X}$, $B \equiv (\dot{Z}, N)$ and $C \equiv (\dot{Y}, \tilde{W}, N)$.The next section provides the proof of the converse.

3.7. Converse

In this section we prove the outer bounds of theorem 5.The derivation of the outer bound is heavily dependent on the proper application of the lemmas introduced in section 3.4 and proper partioning of the ω region. We start with Fano's Inequality and the secrecy criterion to first derive the individual bounds, then we are going to prove the weighted bounds for the two regions described

in subsections 3.7.1 and 3.7.2

$$R_1 + \beta_1 R_2 \leq (\bar{\epsilon}_{11} - \bar{\epsilon}_{12}), \tag{3.43}$$

$$R_1 + \beta_2 R_2 \leq \beta_2(\bar{\epsilon}_{22} - \bar{\epsilon}_{21}). \tag{3.44}$$

The derivation of the upper bound of the achievable rate of transmitter Tx_1 starts with equation (3.18), which followed by the application of the secrecy criterion for the $Tx_1 - Rx_1$ as in equation (3.8) yields (3.45)

$$\begin{aligned}
nR_1 - \delta_1 &\leq h(\mathcal{M}_1|Z^n, X^n, N^n), && (3.45)\\
&\leq h(W^n, \mathcal{M}_1|Z^n, X^n, N^n), && (3.46)\\
&= h(W^n|Z^n, X^n, N^n) + h(\mathcal{M}_1|W^n, Z^n, X^n, N^n), && (3.47)\\
&= h(W^n|Z^n, X^n, N^n), && (3.48)\\
&= h(W^n|X^n, N^n) - I(W^n; Z^n|X^n, N^n),\\
&= h(W^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (3.49)\\
&= h(W^n|N^n) - h(W^n|Y^n, N^n) + h(W^n|Y^n, N^n) - I(W^n; \tilde{Y}^n|N^n),\\
&\leq I(W^n; Y^n|N^n) - I(W^n; \tilde{Y}^n|N^n) + \delta^{'}, && (3.50)\\
\implies nR_1 - \delta_1^{'} &= I(W^n; Y^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (3.51)\\
&= h(Y^n|N^n) - h(Y^n|W^n, N^n) - h(\tilde{Y}^n|N^n), && (3.52)
\end{aligned}$$

where equation (3.46) is the result of the fact that additional random variable does not reduce entropy.Equation (3.47) follows from the chain rule of entropy. Equation (3.48) occurs because $h(\mathcal{M}_1|W^n, Z^n, X^n, N^n) = 0$ as it is assumed that the receiver can reliably decode its own message from the recived signal. Equation (3.49) follows from the independence of the inputs and $\tilde{Y}^n = (N_{12}W)^n$, $h(W^n|Y^n, N^n) \leq \delta^{'}$ in equation (3.50) since the receiver should be able to find out the transmitted codeword from the received signal with less than δ amount of error where $\delta \to 0$ as $n \to \infty$ and $\delta_1^{'} = \delta_1 + \delta^{'}$ in equation (3.51) and equation (3.52) follows from the fact that $h(\tilde{Y}^n|W^n, N^n) = 0$.

Now we simplify the first entropy term of equation (3.52) as follows

$$
\begin{aligned}
h(Y^n|N^n) &= h\Big\{(kW \oplus kX)^n, (lW)^n, (mX)^n|N^n\Big\}, && (3.53)\\
&= h\Big\{(kW \oplus kX)^n|N^n\Big\} + h\Big\{(lW)^n|(kW \oplus kX)^n, N^n\Big\} \\
&\qquad + h\Big\{(mX)^n|(lW)^n, (kW \oplus kX)^n, N^n\Big\}, \\
&\leq h\Big\{(kW \oplus kX)^n|N^n\Big\} + h\Big\{(lW)^n|N^n\Big\} + h\Big\{(mX)^n|N^n\Big\}, && (3.54)\\
&= h\Big\{\bar{Y}^n|N^n\Big\} + h\Big\{(lW)^n|N^n\Big\} + h\Big\{(mX)^n|N^n\Big\}, && (3.55)
\end{aligned}
$$

where in equation (3.53), $k = (N_{11}N_{21})$ and $m = (N_{21}\bar{N}_{11})$ whereas l has been defined before in lemma 4, equation (3.54) follows from the fact that conditioning reduces entropy and $\bar{Y}^n = (kW \oplus kX)^n$ in equation (3.55).The second entropy term of equation (3.52) is expressed in a more compact form as follows

$$h(Y^n|W^n, N^n) = h\Big\{(N_{21}X)^n|N^n\Big\} = h(\tilde{Z}^n|N^n), \tag{3.56}$$

where $\tilde{Z}^n = (N_{21}X)^n$ in equation (3.56). Therefore the bound for R_1 can be written as follows putting the simplifications of equations (3.55) and (3.56) back into (3.52)

$$nR_1 - \delta_1^{'} \leq h\Big\{\bar{Y}^n|N^n\Big\} + h\Big\{(lW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\} + h\Big\{(mX)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}. \tag{3.57}$$

Proceeding as above we can similarly find the bound for R_2 as follows

$$nR_2 - \delta_2^{'} \leq h\Big\{\bar{Z}^n|N^n\Big\} + h\Big\{(tX)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\} + h\Big\{(sW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\}, \tag{3.58}$$

where $\bar{Z}^n = (aW \oplus aX)^n$, with $a = (N_{22}N_{12})$ whereas $t = (N_{22}\bar{N}_{12})$ and s has been defined before in lemma 4. Now finding the weighted sum bound by adding equation (3.57) with ω times of equation (3.58) we get,

$$
\begin{aligned}
&n(R_1 + \omega R_2) - (\delta_1^{'} + \omega\delta_2^{'}) \\
&\leq h\Big\{\bar{Y}^n|N^n\Big\} + \omega h\Big\{\bar{Z}^n|N^n\Big\} + \Bigg[h\Big\{(lW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\}\Bigg] + \omega\Bigg[h\Big\{(sW)^n|N^n\Big\}
\end{aligned}
$$

$$-h\Big\{\tilde{Y}^n|N^n\Big\}\Big] + \Big[h\Big\{(mX)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}\Big] + \omega\Big[h\Big\{(tX)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}\Big], \quad (3.59)$$

Next we will try to simplify the pair of entropies inside each of the square braces seperately using the Marton style expansion as follows,

$$h\Big\{(lW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\} = \sum_{i=1}^{n}\Big[h\Big\{(lW)_i|(lW)^{i-1},\tilde{Y}_{i+1}^n,N^n\Big\} - h\Big\{\tilde{Y}_i|(lW)^{i-1},\tilde{Y}_{i+1}^n,N^n\Big\}\Big], \quad (3.60)$$

$$= \sum_{i=1}^{n}\Big[h\Big\{(lW)_i|\mathcal{D}_i,N_i\Big\} - h\Big\{\tilde{Y}_i|\mathcal{D}_i,N_i\Big\}\Big], \quad (3.61)$$

$$= \sum_{i=1}^{n}\Big[\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}\Big]h\Big\{W_i|\mathcal{D}_i,N_i\Big\}, \quad (3.62)$$

where equation (3.60) follows from lemma 3, $\mathcal{D}_i = \{(lW)^{i-1},\tilde{Y}_{i+1}^n,N^{i-1},N_{i+1}^n\}$ in equation (3.61).

Using the same lemmas as above and the same method of simplification we can simplify the other entropy pair involving W as well to give us the following

$$h\Big\{(sW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\} = \sum_{i=1}^{n}\Big[\bar{\epsilon}_{12}\epsilon_{22} - \bar{\epsilon}_{12}\Big]h\Big\{W_i|\mathcal{E}_i,N_i\Big\}, \quad (3.63)$$

$$\leq -\sum_{i=1}^{n}\bar{\epsilon}_{12}\bar{\epsilon}_{22}h\Big\{W_i|\mathcal{E}_i,N_i\Big\}, \quad (3.64)$$

$$\leq -\sum_{i=1}^{n}\bar{\epsilon}_{12}\bar{\epsilon}_{22}h\Big\{W_i|\mathcal{D}_i,N_i\Big\}, \quad (3.65)$$

where in equation (3.63) $\mathcal{E}_i = \{(sW)^{i-1},\tilde{Y}_{i+1}^n,N^{i-1},N_{i+1}^n\}$, equation (3.65) follows from lemma 4 where we have used $\mu(W^n,N^n) = \{\tilde{Y}_{i+1}^n,N^{i-1},N_{i+1}^n\}$. Thus combining equation (3.62) and (3.65) together we can write the following

$$\Big[h\Big\{(lW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\}\Big] + \omega\Big[h\Big\{(sW)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\}\Big]$$
$$\leq \sum_{i=1}^{n}\Big\{\Big[\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}\Big] - \omega\bar{\epsilon}_{12}\bar{\epsilon}_{22}\Big\}h\Big\{W_i|\mathcal{D}_i,N_i\Big\}, \quad (3.66)$$

Now using the result in equation (3.62) and the fact that $\Big[h\Big\{(mX)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}\Big]$ can be proved to be negative using similar methods as for equation (3.65) we can further simplify equation

(3.57) to find the individual bound on R_1 as follows

$$\begin{aligned} nR_1 - \delta_1^{'} &\leq \sum_{i=1}^{n} \Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21} h\Big\{ (X \oplus W)_i | Y^{i-1}, N^n \Big\} + (\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) h\Big\{ W_i | \mathcal{D}_i, N_i \Big\} \Bigg], \\ &\leq \sum_{i=1}^{n} \Big[\bar{\epsilon}_{11}\bar{\epsilon}_{21} + (\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) \Big], \qquad (3.67) \\ &= \sum_{i=1}^{n} \Big[\bar{\epsilon}_{11} - \bar{\epsilon}_{12} \Big], \qquad (3.68) \end{aligned}$$

where (3.67) follows from the fact that $(\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) > 0$ and that entropy of binary random variables can be upper bounded by 1. On dividing both sides of equation (3.68) by n and letting n approach ∞ we get the bound on R_1 as $\bar{\epsilon}_{11} - \bar{\epsilon}_{12}$ by using the fact that $\delta_1^{'} \to 0$ as $n \to \infty$.

We can similarly simplify equation (3.58) to get the individual bound on R_2 just like as for R_1. Further the entropy terms involving X can be combined in a similar fashion as those for W and simplified to give us the following

$$\begin{aligned} \omega \Bigg[h\Big\{ (tX)^n | N^n \Big\} - h\Big\{ \tilde{Z}^n | N^n \Big\} \Bigg] + \Bigg[h\Big\{ (mX)^n | N^n \Big\} - h\Big\{ \tilde{Z}^n | N^n \Big\} \Bigg] \\ \leq \sum_{i=1}^{n} \Bigg\{ \omega \Big[\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21} \Big] - \bar{\epsilon}_{21}\bar{\epsilon}_{11} \Bigg\} h\Big\{ X_i | \mathcal{C}_i, N_i \Big\}, \qquad (3.69) \end{aligned}$$

where $\mathcal{C}_i = \{(tX)^{i-1}, \tilde{Z}_{i+1}^n, N^{i-1}, N_{i+1}^n\}$ in equation (3.69). So the weighted sum bound of equation (3.59) can be further simplified using equations (3.66) and (3.69) as follows

$$\begin{aligned} & n(R_1 + \omega R_2) - (\delta_1^{'} + \omega\delta_2^{'}) \\ & \leq \sum_{i=1}^{n} \Bigg[h(\bar{Y}_i | \bar{Y}^{i-1}, N^n) + \omega h(\bar{Z}_i | \bar{Z}^{i-1}, N^n) + \Big\{ (\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) - \omega\bar{\epsilon}_{22}\bar{\epsilon}_{12} \Big\} h\Big\{ W_i | \mathcal{D}_i, N_i \Big\} \\ & \qquad + \Big\{ \omega(\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21}) - \bar{\epsilon}_{11}\bar{\epsilon}_{21} \Big\} h\Big\{ X_i | \mathcal{C}_i, N_i \Big\} \Bigg], \qquad (3.70) \\ & \leq \sum_{i=1}^{n} \Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21} h\Big\{ (X \oplus W)_i | Y^{i-1}, N^n \Big\} + \omega\bar{\epsilon}_{22}\bar{\epsilon}_{12} h\Big\{ (X \oplus W)_i | Z^{i-1}, N^n \Big\} \\ & + \Big\{ (\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) - \omega\bar{\epsilon}_{22}\bar{\epsilon}_{12} \Big\} h\Big\{ W_i | \mathcal{D}_i, N_i \Big\} + \Big\{ \omega(\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21}) - \bar{\epsilon}_{11}\bar{\epsilon}_{21} \Big\} h\Big\{ X_i | \mathcal{C}_i, N_i \Big\} \Bigg], \qquad (3.71) \end{aligned}$$

$$= \sum_{i=1}^{n} \Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21} h\Big\{(X \oplus W)_i | Y^{i-1}, N^n\Big\} + \omega \bar{\epsilon}_{22}\bar{\epsilon}_{12} h\Big\{(X \oplus W)_i | Z^{i-1}, N^n\Big\}$$
$$+ \bar{\epsilon}_{22}\bar{\epsilon}_{12}(\beta_2 - \omega) h\Big\{W_i | \mathcal{D}_i, N_i\Big\} + (\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21})\Big\{\omega - \beta_1\Big\} h\Big\{X_i | \mathcal{C}_i, N_i\Big\} \Bigg], \quad (3.72)$$

where equation (3.70) follows by applying the chain rule on the first two entropy terms of (3.59), β_1 & β_2 in equation (3.72) is as has been defined in (3.10).

The weighting factor ω can take any positive real number, i.e. $\omega \in [0, \infty)$. Since we are finding the upper bound for the region when $\beta_2 > \beta_1$ so the the entire range of ω can be subdivided as shown in figure 3.3.

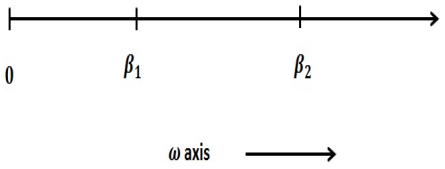

Figure 3.3. The Range of ω Axis with it's Subdivisions.

Now when $\omega \leq \beta_1$ then from equation (3.72) we see that $\omega - \beta_1$ is less than 0 while for $\omega \geq \beta_2$ the term $\beta_2 - \omega$ is less than 0. Thus for these regions one of the possible upper bound can be obtained by eliminating the negative terms. However when $\beta_1 < \omega < \beta_2$ there is no term in equation (3.72) which is negative. With that insight now for further upper bounding the equation in (3.72) we divide the whole range of ω into two overlapping regions. The regions are described as below:

$$\text{Region 1} : \omega \in [0, \beta_2) \&$$
$$\text{Region 2} : \omega \in (\beta_1, \infty) \quad (3.73)$$

We next find the tightest upper bounds for each of the regions and see if they match with the equations (3.43) and (3.44).

3.7.1. Region 1: $\omega \in [0, \beta_2)$

Let us first find the tightest bound when the coefficient of $h\Big\{X_i | Y^{i-1}, N^n\Big\}$ in (3.60) is negative, i.e. when $\omega \in [0, \beta_1]$. For this portion of region 1 the outer bound of equation (3.72) can be further bounded as follows

$$n\Big[R_1+\omega R_2\Big]-\delta^{'}\leq \sum_{i=1}^{n}\Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21}h\Big\{(X\oplus W)_i|Y^{i-1},N^n\Big\}+\omega\bar{\epsilon}_{22}\bar{\epsilon}_{12}h\Big\{(X\oplus W)_i|Z^{i-1},N^n\Big\}$$
$$+\bar{\epsilon}_{22}\bar{\epsilon}_{12}(\beta_2-\omega)h\Big\{W_i|Z^{i-1},N^n\Big\}\Bigg], \quad (3.74)$$
$$\leq \sum_{i=1}^{n}\Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21}+\omega\bar{\epsilon}_{22}\bar{\epsilon}_{12}+\bar{\epsilon}_{22}\bar{\epsilon}_{12}(\beta_2-\omega)\Bigg], \quad (3.75)$$
$$=\sum_{i=1}^{n}(\bar{\epsilon}_{11}-\bar{\epsilon}_{12}), \quad (3.76)$$
$$\implies R_1+\omega R_2\leq(\bar{\epsilon}_{11}-\bar{\epsilon}_{12}), \quad (3.77)$$

where equation (3.74) follows from the fact that $(\omega-\beta_1)<0$ when $\omega\in[0,\beta_1]$ and hence can be ignored for the purpose of upper bound, equation (3.75) is the result of the fact that binary entropy can be upper bounded by 1 and that, all the other terms in (3.74) are non-negative. Equation (3.76) uses the definition of β_2 from (3.10) for simplification. Finally dividing both sides of equation (3.76) by n and letting $n\to\infty$, we get equation (3.77).

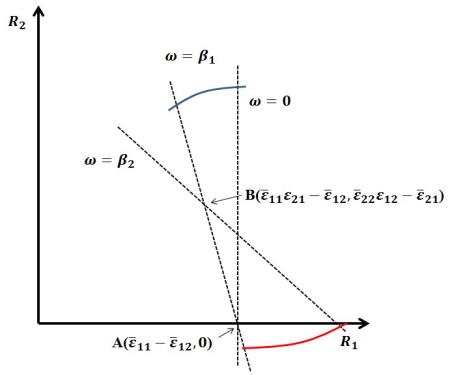

Figure 3.4. The Variation of the Upper Bounds as ω Changes from 0 to β_2 in Region 1.

In figure 3.4 the curved blue line shows how the bounds varies as ω goes from 0 to β_1. Looking at the figure one can say without doubt the slope of the bounds decreases and it slants more towards the R_1 axis as the value of ω moves away from 0 towards β_1 giving the tightest bound when $\omega=\beta_1$. Thus the tightest bound when $\omega\in[0,\beta_1]$ is obtained by replacing ω by β_1 in equation (3.77) and is given as

$$R_1+\beta_1 R_2\leq(\bar{\epsilon}_{11}-\bar{\epsilon}_{12}). \quad (3.78)$$

Now we find the tightest upper bound for the rest of the values of ω in region 1 and compare it with the bound of equation (3.78) to find the tightest bound for the entire portion of region 1. Now when $\omega \in [\beta_1, \beta_2)$ the last two entropy terms in equation (3.72) are both positive, hence the weighted sum bound can be further upper bounded as

$$n\Big[R_1 + \omega R_2\Big] - \delta' \leq \sum_{i=1}^{n} \left[\bar{\epsilon}_{11}\bar{\epsilon}_{21} + \omega\bar{\epsilon}_{22}\bar{\epsilon}_{12} + \bar{\epsilon}_{22}\bar{\epsilon}_{12}\Big\{\beta_2 - \omega\Big\} + (\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21})\Big\{\omega - \beta_1\Big\} \right], \quad (3.79)$$

$$= \sum_{i=1}^{n} \left[(\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) + \omega(\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21}) \right], \quad (3.80)$$

$$\implies R_1 + \omega R_2 \leq \left[(\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}) + \omega(\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21}) \right], \quad (3.81)$$

where equation (3.79) results from the fact that binary entropy can be upper bounded by 1 and the fact that all of the terms are non-negative. In equation (3.80) the definitions of β_1 and β_2 as given by (3.10) is used for the simplification and dividing both sides of equation (3.80) by n and letting $n \to \infty$, gives equation (3.81).

Now as can be seen from equation (3.81) all the bounds pass through the point $(\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}, \bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21})$ but so does the tightest bound of the $\omega \in [0, \beta_1]$ region given by equation (3.78). Now if we refer to figure 3.4 then the red curve shows the trace of the variation of the bounds as ω, the slope, varies from β_1 to β_2. Now looking at the trace it is pretty much clear that the tightest bound for even this subregion of region 1 will be given by equation (3.78) which is obtained by placing $\omega = \beta_1$ in equation (3.81) followed by some simple algebraic simplification. Thus the overall tightest bound for region 1 is given by (3.78) which concludes the proof of the bound in equation (3.43).

3.7.2. Region 2: $\omega \in [\beta_1, \infty)$

Now when ω varies between β_1 to β_2 we have already seen the bound as given by equation (3.81), which does not change even for region 2. The red curve in figure 3.5 traces the variation of the bounds as the slope varies from β_1 to β_2, each of them passing through the point $(\bar{\epsilon}_{11}\epsilon_{21} - \bar{\epsilon}_{12}, \bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21})$. It is then easily seen from the figure that the tightest of the bounds occur when $\omega = \beta_2$ and is same as equation (3.86) after some algebraic simplification.

So with the tightest bound for $\omega \in (\beta_1, \beta_2]$ already derived we will now only concentrate on the bounds for the subregion, $\omega \in [\beta_2, \infty)$. We start from equation (3.72) and find that $(\beta_2 - \omega)$ is

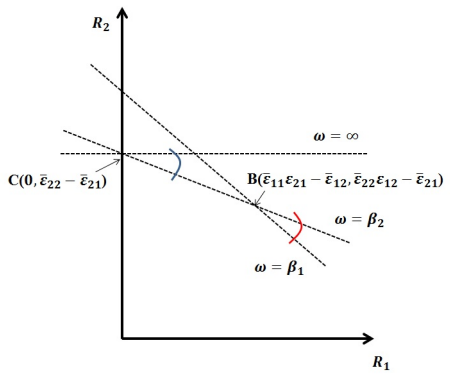

Figure 3.5. The Variation of the Upper Bounds as ω Changes from β_1 to ∞ in Region 2.

less than 0 when ω is in the range $[\beta_2, \infty)$. Thus ignoring the negative term, equation(3.72) can be further bounded as follows

$$n\Big[R_1 + \omega R_2\Big] - \delta' \leq \sum_{i=1}^{n} \Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21} h\Big\{(X \oplus W)_i | Y^{i-1}, N^n\Big\} + \omega \bar{\epsilon}_{22}\bar{\epsilon}_{12} h\Big\{(X \oplus W)_i | Z^{i-1}, N^n\Big\} + (\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21})\Big\{\omega - \beta_1\Big\} h\Big\{X_i | Y^{i-1}, N^n\Big\}\Bigg], \quad (3.82)$$

$$\leq \sum_{i=1}^{n} \Bigg[\bar{\epsilon}_{11}\bar{\epsilon}_{21} + \omega\bar{\epsilon}_{22}\bar{\epsilon}_{12} + (\bar{\epsilon}_{22}\epsilon_{12} - \bar{\epsilon}_{21})\Big\{\omega - \beta_1\Big\}\Bigg], \quad (3.83)$$

$$\leq \sum_{i=1}^{n} \omega(\bar{\epsilon}_{22} - \bar{\epsilon}_{21}), \quad (3.84)$$

$$\implies R_1 + \omega R_2 \leq \omega(\bar{\epsilon}_{22} - \bar{\epsilon}_{21}), \quad (3.85)$$

where in equation (3.83) all the terms being positive the binary entropy for each of the terms has been upper bounded by 1. Equation (3.84) uses the definition of β_1 from (3.10) for simplification. Finally equation (3.84) is obtained by dividing both sides of equation (3.83) by n and then letting $n \to \infty$.

We refer back to figure 3.5 where the blue curve shows the variation of the bounds as ω, the slope changes from β_2 to ∞. It can be said without doubt that the tightest of all the bounds occur when $\omega = \beta_2$ and is given as follows

$$R_1 + \beta_2 R_2 \leq \beta_2(\bar{\epsilon}_{22} - \bar{\epsilon}_{21}), \quad (3.86)$$

Thus equation (3.86) gives the tightest of all the bounds for region 2 and proves the bound of equation (3.44).

Remark 2 *The tightest bound in region* 1*, i.e. when* $\omega \in [0, \beta_2)$ *is obtained when the entropy term* $h\Big\{X_i|Y^{i-1}, N^n\Big\}$ *goes to zero. Similarly the tightest bound for region* 2*, i.e. when* $\omega \in (\beta_1, \infty)$ *is obtained when* $h\Big\{W_i|Z^{i-1}, N^n\Big\}$ *goes to zero. Intuitively speaking it is as if the tightest bound for the two regions are obtained when only one of the users transmit while the other remains silent. We gain nothing extra by transmitting from both the users simultaneously even when we can. Also for the transmission scheme both the entropy terms can never go to zero given the constraints on* β_1 *and* β_2*. This observation might be helpful when we move on to the multi-layer scenario, meaning for a multi-layer scenario it might be logical to give certain layers to user* 1 *and the rest to user* 2 *since giving the same layer to both the users does not give us any additional advantage.*

3.8. Conclusion

It is shown in this chapter that for a fast fading 2-user binary interference channel the secrecy capacity is zero for the class of fading IC defined as *strong*. It is also shown that for the class defined as *very weak* LFIC we get non-zero secrecy rate. The result in this case suggest that carefully transmitting from a transmitter when it does not see interference from the other transmitter and then applying the layered erasure wiretap channel optimal code for binary channel helps in achieving the secrecy rate. The result is important because of the practical channel assumptions wherein a fast fading channel is considered without any channel state information at the transmitter. Also it serves as the first step towards characterizing the secrecy capacity of the real fading IC for which there is limited to almost no result due the complexity of the channel model. However, the solution to this binary problem motivates us to consider the multi-layer scenario, which is solved next in the following chapter and would be a step further towards solving the real fading IC problem.

4. SECRECY CAPACITY OF A CLASS OF LAYERED INTERFERENCE CHANNEL

4.1. Introduction

It has been shown recently in[59] that by allowing multiple transmit-receive pair operate in the same frequency band the overall throughput can be increased significantly. However, the broadcast nature of a wireless channel can cause severe security issues in a multi-user scenario which might supercede the attractive gains that it has to offer if the message being transmitted is of a highly secure nature. The issue of security in such scenarios is becoming an increasingly challenging problem due to the advancement in computational technologies which has evolved manifold in over a decade or so. Information theoretic approach for secret communication started with Shannon's work in [26] which was followed by the pioneering work of Wyner in [3] determining the secrecy capacity of a *degraded* wiretap channel (WTC). This was followed by a plethora of work in discrete memoryless WTC, Gaussian WTC and slow fading WTC in [28], [29], [65], [33]. Some interesting results in fast fading wiretap channel has also been pursued in [66], [50], [68], [25]. However none of these results really treat a multi-user scenario.

Without secrecy constraint, there has been significant work in multi-user scenario too and one such popular multi -user network model is the Interference Channel (IC). After the progress made by Etkin, Tse and Wang recently in characterizing the capacity region of 2-user Gaussiann IC within a single bit in [59] several important results have come to the fore. Some of them are from the perspective of degrees of freedom (DoF) [85], [86], [87] while others for multiple input multiple output (MIMO) Gaussian ICs [24], [88], [89], [90], [91]. The capacity of *fading* ICs has also been studied in [92],[93] where the focus was on scenarios where full channel state information (CSI) is available. Results on fast fading ICs with no CSI at transmitter (CSIT) has also been studied for binary ICs in [84] and layered ICs in [94] and Z-Interference channel in [24]. But surprisingly there has been very few work that addresses the issue of security in ICs. Some of the results as in [80], [81] partially addresses the issue by describing a metric called *secure degrees of freedom* which however is a coarse approximation of secrecy capacity of a channel and also the channel

considered in those works are not fast fading. There is another work as in [78], that finds it for a special IC. Our findings on the other hand paves the way for a more general result. Although there are results with constant channel gain and some partial ones with channel state varying slowly, there seems to be almost no result for fast fading channels and that motivates the current problem. The assumption that is made here is that there is no channel state information at the transmitter. This makes sense once again, since for a fast changing channel the receiver might not be able to provide the transmitter with a timely update about the state of the channel and without feedback from the receiver the transmitter has no other means to know about the state of the channel. In addition, an unintended receiver - if malicious would never provide a feedback about the state of the channel. So taking into account all these factors, a fast fading wireless channel with no channel state information at the transmitter (no CSIT) appears to be the most practical assumption.

Deterministic model of [82] is very useful to gain insight about multi-user network models which paves the way towards characterization of the performance of a Gaussian channel. In the deterministic channel model, a q-bit vector typically models the input and the effect of channel co-efficient is modeled by removing a particular number of least significant bits which remain below the noise floor even after getting multiplied by the channel gain. For example, if the channel gain is 2^n, then only n most significant bits of the input vector will reach the destination while the others will remain below the noise floor and hence get erased. The fading nature of the communication channel is captured by modelling this n to be random and varying which in turn imply that the number of transmitted bits that survive till the destination is also random. If the bits are imagined as layers then on a fading channel an arbitrary number of such layers get erased depending on the random channel co-efficients. Such a layered model was used in [95] as a stepping stone to characterize the approximate capacity of a 2-user fading broadcast channel, then in [96] to characterize the sum capacity of a class of layered erasure one-sided IC and then also in [24] as an intial step before finding the approximate capacity of a fading Z-IC. Our result on layered fading interference channel (LFIC) is a further step forward towards sloving the more general Gaussian fading IC.

The rest of the chapter is organised as follows. Section 4.2 formally describes the channel model followed by the introduction of notations used in this chapter and also states the secrecy criterion to be followed in the solution of this problem. Section 4.3 presents the main findings of this chapter.We further provide some remarks to aid form the intuitive picture about the achievability

of the secrecy capacity region for a *very weak* layered fading interference channel (LFIC) and also provide an outline of the outer bound. The section also includes some definitions about the various types of ICs in the begining. We state some lemmas and prove them in Section 4.4. These lemmas are later used for the proof of our main results. The proof of the secrecy capacity result for *strong* LFIC is provided in Section 4.5. The proof of the secrecy capacity region result for *very weak* LFIC is divided into two sections where Section 4.6 derives the upper bound of all the rates whereas Section 4.7 provides the strategy to achieve the various rate pairs which defines the secrecy capacity region. Section 4.6 is subdivided into three subsections where subsection 4.6.1 derives a weighted sum bound while subsection 4.6.2 describes two regions and defines a superset to the secrecy capacity region of *very weak* LFIC. Subsection 4.6.3 then proves that the region derived in 4.6.2 is a subset of the result given in Section 4.3. Section 4.7 also consists of two subsections 4.7.1 and 4.7.2 each of which derives the achievability scheme to achieve the rate pairs present in region 1 and 2 respectively. Both the regions are described in subsection 4.6.2. We finally conclude the chapter with Section 4.8.

4.2. Channel Model and Some Preliminaries

We consider a 2-user Layered Fading Interference Channel(LFIC) as shown in Fig. 4.1, where each user transmits a certain number of bits, say q, to its desired receiver and the bits transmitted by the transmitters cause interference at the unwanted receivers. To model fading we assume that the number of bits/layers on each link that survive (does not get erased) and reach the destination is random [95]. Let us denote these random numbers corresponding to the different links of the channel by $(N_{ij})_t$, for $i, j = 1, 2$ as shown in Fig. 4.1, where t represents the time index. Clearly, these are discrete random variables and can be completely characterized by their Complementary Cumulative Distribution Functions (CCDF), denoted by $\bar{F}_{(N_{ij})_t}(l)$ for $i, j = 1, 2$ and $0 \leq l \leq q$, where these CCDFs are defined as

$$\bar{F}_{(N_{ij})_t}(l) = P((N_{ij})_t \geq l) = \bar{F}_{N_{ij}}(l), \ \forall l \geq 0, \tag{4.1}$$

We also assume that the various $(N_{ij})_t$'s are mutually independent across i, j and t which together with (4.1) implies that for any fixed $i, j \in \{1, 2\}$, $(N_{ij})_t$'s are IID as N_{ij} for all $t \geq 1$ and thus from now on we will no longer use the time index in the representation of the CCDFs. Except

for the fundamental properties these CCDFs, $\bar{F}_{N_{ij}}(\cdot), i, j = 1, 2$ can be arbitrary. Instantaneous values of N_{ij}'s are assumed to be known only at the respective receivers. Transmitters are assumed to know the statistics, i.e., $\bar{F}_{N_{ij}}(l)$ for $i, j \in \{1, 2\}$ and all $l \geq 0$.

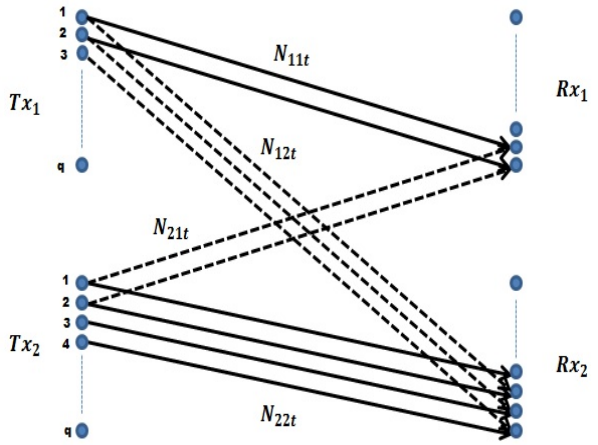

Figure 4.1. 2-User Layered Fading Interference Channel.

We next give a preview of the notations to be used in this chapter for vectors of length q in a binary field represented as $\mathbb{F}_2^q$. The signal transmitted from Tx_1 at time instant t would be represented as W_t while that from Tx_2 as X_t where both W_t and X_t are q-length vectors. Now a q-length vector A^q would mean the sequence$\{A_1, A_2, \cdots, A_q\}$ where each of $A_1, A_2, \cdots, A_q \in \mathbb{F}_2$ whereas a vector $A_{i_1}^{i_2}$ where $i_1 < i_2$ would mean the sequence $\{A_{i_1}, \cdots, A_{i_2}\}$. Next $(A_i)_t$ would represent the i-th component of the q-th length vector, $(A^q)_t$ at time instant t. Similarly $(A_{i_1}^{i_2})_t$ would represent the sequence of binary components at the t-th instant as $\{(A_{i_1})_t, \cdots, (A_{i_2})_t\}$ whereas $(A_{i_1}^{i_2})^t$ would represent the sequence of vectors, $A_{i_1}^{i_2}$ from the first till the t-th instant as $\{(A_{i_1}^{i_2})_1, (A_{i_1}^{i_2})_2, \cdots, (A_{i_1}^{i_2})_t\}$. The sub and superscript within the brackets represent the indices of the vector components while those outside, represent the time indices. For a q-length vector at time instant t, $(A_1)_t$ represents the most significant bit (MSB) while $(A_q)_t$ represent the least significant bit (LSB).

Now superposition of two such binary vectors not necessarily of the same length is modeled by component-wise XOR operation, where the LSBs of the two vectors align with each other; for instance, the XOR of A^m and B^p with $p < m$ will be denoted by $A^m \oplus B^p$, where

$$A^m \oplus B^p = \{A_1, \cdots, A_{m-p}, A_{m-p+1} \oplus B_1, \cdots, A_{m-1} \oplus B_{p-1}, A_m \oplus B_p\}, \tag{4.2}$$

In these notations, the received signals Y_t and Z_t during the t-th time instant at Rx_1 and Rx_2 respectively can be expressed as follows

$$\begin{aligned} Y_t &= (W^{N_{11}} \oplus X^{N_{21}})_t, \\ Z_t &= (W^{N_{12}} \oplus X^{N_{22}})_t, \end{aligned} \tag{4.3}$$

for all $t \geq 1$. Finally, we use capital letters to denote random variables (RVs), e.g. $(A)_t$ is a random variable at time instant t, and small letters denote the realizations, e.g. $(a)_t$ is a realization of $(A)_t$.

We next state the secrecy criterion to be followed in this chapter after defining the decoding error probability. We assume that both the transmitters has secret messages to transmit. Let $\mathcal{M}_i$ be the secret message to be transmitted by transmitter Tx_i where $i = 1, 2$. The transmitters want to transmit at a rate R_i. Now let us suppose that Tx_1 wants to transmit a message $\mathcal{M}_1(k), k \in \{1, 2, \cdots, 2^{nR_1}\}$, then the transmitter chooses a codeword W^n from its codebook $\mathcal{C}_1(n)$. Similarly if Tx_2 wants to transmit a message $\mathcal{M}_2(l), l \in \{1, 2, \cdots, 2^{nR_2}\}$ it does so by choosing a codeword X^n from its codebook $\mathcal{C}_2(n)$. Now both the receivers are assumed to be aware of the coding scheme and the codebooks used by the transmitters. Next let us assume that $\hat{\mathcal{M}}_i$ is the estimate Rx_i makes about the message transmitted from Tx_i on observing the received signal, where $i \in \{1, 2\}$.An error occurs when $\hat{\mathcal{M}}_i \neq \mathcal{M}_i$ and the probability of decoding error is given by

$$\lambda_i = P[\hat{\mathcal{M}}_i \neq \mathcal{M}_i], \quad i = 1, 2, \tag{4.4}$$

A rate pair (R_1, R_2) is said to be *achievable* if there exists $\mathcal{C}_i(n), i = 1, 2$ such that $\max(\lambda_1, \lambda_2) \to 0$ and both

$$\begin{aligned} I(\mathcal{M}_1; Z^n | X^n, N^n) &< \delta_1 \text{ and} \\ I(\mathcal{M}_2; Y^n | W^n, N^n) &< \delta_2, \end{aligned} \tag{4.5}$$

is satisfied with arbitrarily small $\delta_1 > 0$ and $\delta_2 > 0$, as $n \to \infty$. Note that the aforementioned conditions are equivalent to an *equivocation* of 1 according to the definitions of [28], since

$$I(\mathcal{M}_1, Z^n | X^n, N^n) < \delta_1,$$

$$\implies h(\mathcal{M}_1) < h(\mathcal{M}_1|Z^n, X^n, N^n) + \delta_1, \tag{4.6}$$

$$\frac{h(\mathcal{M}_1|Z^n, X^n, N^n)}{h(\mathcal{M}_1)} > 1 - \delta_1^{'}, \tag{4.7}$$

where the left hand side of the above equation represents the *equivocation* and $\delta_1^{'} = \frac{\delta_1}{h(\mathcal{M}_1)}$ can be made arbitrarily small. Similarly the other condition in (4.5) can be shown to have an *equivocation* of 1 as well. The main contributions of the chapter are stated in the next section.

4.3. Main Result

In this section the explicit characterization of the secrecy capacity region of a *strong* and a *very weak* LFIC is stated. However we first provide with some definitions so as to ease the understanding of the results.

Definition 11 *A layered intereference channel is called strong if both the interfering links are stronger than their corresponding direct links, mathematically* $P(N_{ij} \geq l) \geq P(N_{ii} \geq l)$, *for all* $l \in \{1, \cdots, q\}$ *and* $i, j = 1, 2; i \neq j$ *where* N_{ii} *represents the unerased fading channel states for the direct links between* Tx_i *and* Rx_i *and* $N_{ij}, i \neq j$ *represents the unerased fading channel states for the cross links between* Tx_i *and* Rx_j. *But if the direct links are stronger than their corresponding interfering link then we call it as a moderately weak LFIC, mathematically,* $P(N_{ii} \geq l) \geq P(N_{ij} \geq l)$, *for all* $l \in \{1, \cdots, q\}$ *and* $i, j = 1, 2; i \neq j$. $N_{ii} - N_{ji}$ *represents the number of uninterfered and unerased direct link layers from* Tx_i *to* Rx_i. *So if these uninterfered and unerased direct links for a transmit receive pair are stronger than their corresponding cross links then we say that the channel is a weak interference channel, mathematically for a LFIC,* $P(N_{ii} - N_{ji} \geq l) \geq P(N_{ij} \geq l)$ *for all* $l \in \{1, \cdots, q\}$ *and* $i, j \in \{1, 2\}; i \neq j$. *Finally we now define two channel parameters* $\beta_1(l)$ *and* $\beta_2(l)$ *as follows*

$$\beta_1(l) = \frac{\alpha_2(l)}{\left[P(T \geq l) - P(N_{21} \geq l)\right]} \text{ and } \beta_2(l) = \frac{\left[P(L \geq l) - P(N_{12} \geq l)\right]}{\alpha_1(l)}, \tag{4.8}$$

where $\alpha_2(l) = [P(N_{21} \geq l) - P(N_{21} - N_{11} \geq l)]$, $T = (N_{22} - N_{12})^+$, $\alpha_1(l) = [P(N_{12} \geq l) - P(N_{12} - N_{22} \geq l)]$ *and* $L = (N_{11} - N_{21})^+$. *So if for a weak interference channel* $\min(\beta_2(l)) > \max(\beta_1(l))$, $\forall l \in \{1, \cdots, q\}$ *then we call it as a very weak layered interference channel.*

With the above definitions we now state the result for a *very weak* layered fading interference channel in theorem 6 and a *strong* layered fading interference channel in lemma 5.

Theorem 6 *The secrecy capacity region of a very weak two-user LFIC with no channel state information at transmitter (CSIT), $\mathcal{C}(N)$, is given by the polygonal region $A_0A_1\cdots A_qB_q\cdots B_1B_0O$ as shown in figure 4.2 where the vertices A_is are given by the co-ordinates*

$$\Big\{\mathbb{E}[N_{11}]-\mathbb{E}[N_{12}]-\sum_{l\in\mathcal{A}(\omega)}\alpha_2(l),\ \sum_{l\in\mathcal{A}(\omega)}[P(T\geq l)-P(N_{21}\geq l)]\Big\}, \tag{4.9}$$

while B_is are given by

$$\Big\{\sum_{l\in\mathcal{B}(\omega)}[P(L\geq l)-P(N_{12}\geq l)],\mathbb{E}[N_{22}]-\mathbb{E}[N_{21}]-\sum_{l\in\mathcal{B}(\omega)}\alpha_1(l)\Big\}, \tag{4.10}$$

where $\mathcal{A}(\omega)$ is defined as

$$\mathcal{A}(\omega)=\Big\{l\in\{1,\cdots q\}|\omega>\beta_1(l)\Big\}, \tag{4.11}$$

and $\mathcal{B}(\omega)$ as

$$\mathcal{B}(\omega)=\Big\{l\in\{1,\cdots q\}|\omega<\beta_2(l)\Big\}, \tag{4.12}$$

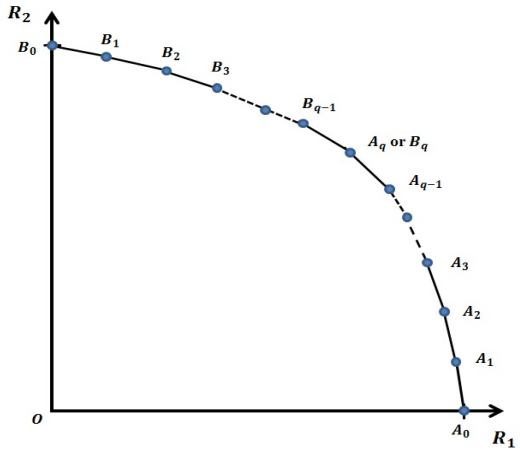

Figure 4.2. Secrecy Capacity Region of *Very Weak* Layered Fading Interference Channel.

Remark 3 *The approach to finding the outer bound for the capacity region involves determination of two carefully defined set of regions $H(\omega)$ and $G(\omega)$, where $H(\omega)$ is given as*

$$H(\omega)=\Bigg\{(R_1,R_2)|R_1+\omega R_2\leq\Big\{\mathbb{E}[N_{11}]-\mathbb{E}[N_{12}]-\sum_{l\in\mathcal{A}(\omega)}\alpha_2(l)\Big\} + \omega\Big\{\sum_{l\in\mathcal{A}(\omega)}[P(T\geq l)-P(N_{21}\geq l)]\Big\}\Bigg\}, \tag{4.13}$$

and $G(\omega)$ *is given as*

$$G(\omega) = \Bigg\{(R_1, R_2) | R_1 + \omega R_2 \leq \Big\{ \sum_{l \in \mathcal{B}(\omega)} [P(L \geq l) - P(N_{12} \geq l)] \Big\} + \omega \Big\{ \mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l \in \mathcal{B}(\omega)} \alpha_1(l) \Big\} \Bigg\}, \tag{4.14}$$

with $\mathcal{A}(\omega)$ *and* $\mathcal{B}(\omega)$ *as defined in equations* (4.11) *and* (4.12)*, respectively. Once these set of regions are determined it will be further shown that the polygonal region as shown in figure 4.2 can also be represented by the following mathematical expression, in other words the following mathematical expression does actually represent the secrecy capacity region of a very weak layered fading interference channel,*

$$\mathcal{R} = \underset{\omega \in [0, \omega_1)}{\cap} H(\omega) \underset{\omega \in (\gamma_1, \infty)}{\cap} G(\omega) \cap [0, \infty)^2, \tag{4.15}$$

where $\omega_1 = \min \{\beta_2(l)\}$ *and* $\gamma_1 = \max \{\beta_1(l)\}$.

Remark 4 *The* $\beta_1(l)$*'s in* $\mathcal{A}(\omega)$ *in equation* (4.11) *are ordered as* $\gamma_q < \gamma_{q-1} < \cdots < \gamma_1$ *with their corresponding permutation as* ν *such that* $\gamma_i = \beta_1(\nu(i)), i = \{1, \cdots, q\}$*.The subscript* i *in* A_i *is same as the subscript used for* $\gamma_i s$*. Also* $\gamma_0 \to \infty$ *and* $\gamma_q = \min \beta_1(l)\ \forall l \in \{1, \cdots, q\}$*. Similarly, the* $\beta_2(l)$*'s in* $\mathcal{B}(\omega)$ *are ordered as* $\omega_1 < \omega_2 < \cdots < \omega_q$ *with corresponding permutation as* τ *such that* $\omega_i = \beta_2(\tau(i)), i = \{1, \cdots, q\}$*. Once again the subscript* i *in the vertices* B_i *are same as the subscripts for* $\omega_i s$*. Finally,* $\omega_0 = 0$ *and* $\omega_q = \max \beta_2(l)\ \forall l \in \{1, \cdots, q\}$*. Thus the entire range of* ω *looks as shown in figure 4.3*

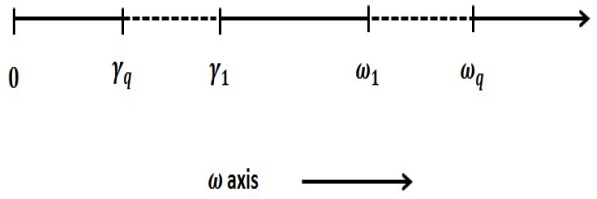

Figure 4.3. The Range of ω With its Subdivisions.

Remark 5 *The parameter* $\alpha_1(l)$ *in* $G(\omega)$ *basically represents the probability that the* l*-th layer of* Tx_1 *is not erased by the cross channel and interferes with a certain layer of* Tx_2 *while* $\alpha_2(l)$ *in* $H(\omega)$ *represents the probability that* l*-th layer of* Tx_2 *interferes with a certain layer of* Tx_1.

Remark 6 *We take a closer look at the co-ordinates of the vertices as given in equations* (4.9) *and* (4.10) *and see if we can make a comment about their achieveability. We see that to achieve the points in* A_i, Tx_2 *should transmit only in those layers which belong to the set* $\mathcal{A}(\omega)$. *As a result* Rx_1 *faces interference only from those layers and hence* Tx_1 *has to reduce its rate on those layers which is confirmed by the presence of the term* $\sum_{l \in \mathcal{A}(\omega)} \alpha_2(l)$ *in the abscissa of* A_i. *Now as we move from* A_0 *towards* A_q *in figure 4.2 along the boundary of the polygon we see that* A_0 *lies on the* R_1*-axis which means* $R_2 = 0$, *hence* $\mathcal{A}(\omega) = \phi$, *in other words* Tx_2 *does not transmit, so* Tx_1 *can fulfill its full capability without facing any interference from* Tx_2. *Next* A_1 *lies to the left of* A_0 *in figure 4.2 meaning* R_1 *co-ordinate is less in* A_1 *compared to in* A_0 *which would make sense as in* A_1 *point,* R_2 *co-ordinate is non-zero meaning* Tx_2 *has started transmitting and hence* Tx_1 *has to reduce its rate due to interference from* Tx_2*'s transmission so that* Rx_1 *can still reliably decode its own message. Similar interpretations can be provided for the points* $B_i s$ *as to how* Tx_2 *utilises its full capacity when* Tx_1 *is silent as in case of point* B_0 *and then slowly reduces its rate as* Tx_1 *starts transmitting on layers belonging to* $\mathcal{B}(\omega)$ *starting from point* B_1 *and moving upto* B_q.

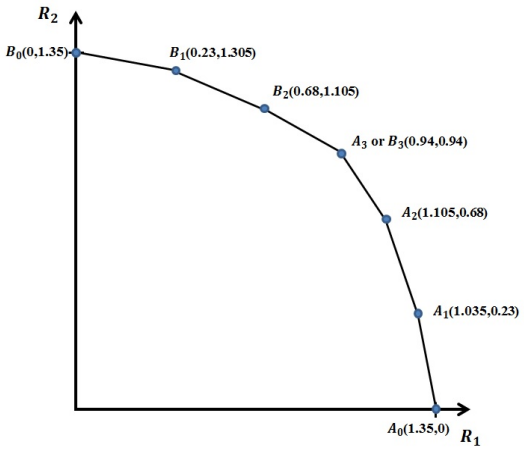

Figure 4.4. The Capacity Region of Example 7.

To further consolidate the understanding of the capacity region given by theorem 6 we next consider a simple example. For the sake of simplicity, we consider a symmetric LFIC such that the direct links N_{11} and N_{22} are distributed similarly as N_1 and the cross links N_{21} and N_{12} are distributed same as N_0. In this symmetric purview $\alpha_1(l)$ and $\alpha_2(l)$ will be same and can be represented by $\alpha(l)$ while $\beta_1(l)$ and $\beta_2(l)$ will be reciprocal to one another and can be represented by $\frac{1}{\beta(l)}$ and

$\beta(l)$ respectively, for all $l \in \{1, 2, \cdots, q\}$. We next state the example and find the secrecy capacity region for it using theorem 6.

Example 7 *Consider a symmetric LFIC with CCDFs of its links as shown in table 4.1. It is clear from the values that* $q = 3$. *Thus the corner points* A_0 *to* A_3 *can be determined from equation* (4.9) *in theorem 6 as* $A_0(1.35, 0)$, $A_1(1.305, 0.23)$, $A_2(1.105, 0.68)$ *and* A_3 *or* $B_3(0.94, 0.94)$ *while the corner points* B_0 *to* B_3 *can be obtained from equation* (4.10) *as* $B_0(0, 1.35)$, $B_1(0.23, 1.305)$, $B_2(0.68, 1.105)$ *and* A_3 *or* B_3, *as has been calculated earlier. Hence the polygonal region as shown in figure 4.4 represents the capacity region for the example as given by theorem 6.*

We next take a step further and define the set of regions $H(\omega)$ *and* $G(\omega)$ *as given by equations* (4.13) *and* (4.14) *respectively, as follows*

$$H(\omega) = \Bigg\{(R_1, R_2) | R_1 + \omega R_2 \leq \Big\{1.35 - \sum_{l \in \mathcal{A}(\omega)} \alpha(l)\Big\} + \omega\Big\{\sum_{l \in \mathcal{A}(\omega)} [P(N_1 - N_0 \geq l) - P(N_0 \geq l)]\Big\}, \tag{4.16}$$

$$G(\omega) = \Bigg\{(R_1, R_2) | R_1 + \omega R_2 \leq \Big\{\sum_{l \in \mathcal{B}(\omega)} [P(N_1 - N_0 \geq l) - P(N_0 \geq l)]\Big\} + \omega\Big\{1.35 - \sum_{l \in \mathcal{B}(\omega)} \alpha(l)\Big\}, \tag{4.17}$$

The $\beta(l)$*'s and* $\frac{1}{\beta(l)}$*'s can then be arranged as shown in figure 4.5, where* ω_1 *and* γ_1 *will be same as* $\beta(2)$ *and* $\frac{1}{\beta(2)}$ *for the example and the secrecy capacity region will be same as the following region as claimed in remark 3*

$$\bigcap_{\omega \in [0, \beta(2))} H(\omega) \bigcap_{\omega \in (\frac{1}{\beta(2)}, \infty)} G(\omega) \cap [0, \infty)^2, \tag{4.18}$$

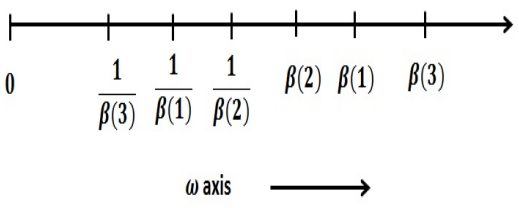

Figure 4.5. The Order of $\beta(l)$s and $\frac{1}{\beta(l)}$s in Example 7.

So from equation (4.18) *we observe that for defining the rate region for the example the bound in* (4.16) *is used when* $\omega \in [0, \beta(2))$ *and the one in* (4.17) *is used when* $\omega \in (\frac{1}{\beta(2)}, \infty)$*.Now if we replace the inequalities in the equations* (4.16) *and* (4.17) *with equality then we get equation of lines which represents the boundary of the rate region. Let us denote these lines by* $\mathcal{M}(\omega)$*.*

Table 4.1. Table Showing Various Parameter Values for Example 7

l	1	2	3
$\bar{F}_{N_1}(l)$	0.9	0.6	0.4
$\bar{F}_{N_0}(l)$	0.3	0.2	0.05
$P(N_1 - N_0 \geq l)$	0.75	0.46	0.28
$P(N_1 - N_0 \geq l) - \bar{F}_{N_0}(l)$	0.45	0.26	0.23
$P(N_0 - N_1 \geq l)$	0.1	0.035	0.005
$\alpha(l)$	0.2	0.165	0.045
$\beta(l)$	2.25	1.58	5.1
$\frac{1}{\beta(l)}$	0.4	0.63	0.2

Now if we go back to the defnition of $\mathcal{A}(\omega)$ *then we see that for all* $\omega \in [0, 0.2]$*,* $\mathcal{A}(\omega) = \phi$ *and then from the bound in* (4.16) *we see that the lines* $\mathcal{M}(\omega)$ *pass through the point* $A_0(1.35, 0)$*. Similarly, the lines obtained when* $\omega \in [0.2, 0.4]$ *passes through the point* $A_1(1.305, 0.23)$ *and* $\mathcal{A}(\omega) = \{3\}$*. Further when* $\omega \in [0.4, 0.63]$*, the lines* $\mathcal{M}(\omega)$ *pass through* $A_2(1.105, 0.68)$ *and* $\mathcal{A}(\omega) = \{1, 3\}$ *while for* $\omega \in [0.63, 1.58)$ *the lines pass through* A_3 *or* $B_3(0.94, 0.94)$ *with* $\mathcal{A}(\omega) = \{1, 2, 3\}$*. Among these lines* $\mathcal{M}(0.2)$ *is special as it passes through both* A_0 *and* A_1*. The same is true for* $\mathcal{M}(0.4)$ *and* $\mathcal{M}(0.63)$*; the former passes through* A_1 *and* A_2 *while the later passes through* A_2 *and* A_3*. Similarly it can also be verified from the bound in* (4.17) *and the definition of* $\mathcal{B}(\omega)$ *that for all* $\omega \in [5.1, \infty)$*,* $\mathcal{B}(\omega) = \phi$ *and the lines* $\mathcal{M}(\omega)$ *pass through the point* $B_0(0, 1.35)$*. These lines pass through* $B_1(0.23, 1.305)$ *for all* $\omega \in [2.25, 5.1]$ *and* $\mathcal{B}(\omega) = \{3\}$*, through* $B_2(0.68, 1.105)$ *for all* $\omega \in [1.58, 2.25]$ *and* $\mathcal{B}(\omega) = \{1, 3\}$ *and through* A_3 *or* $B_3(0.94, 0.94)$ *for all* $\omega \in [0.63, 1.58)$ *where* $\mathcal{B}(\omega) = \{1, 2, 3\}$*. Among these lines* $\mathcal{M}(5.1)$ *is special as it passes through both* B_0 *and* B_1*. The same is true for* $\mathcal{M}(2.25)$ *and* $\mathcal{M}(1.58)$*; the former passes through* B_1 *and* B_2 *while the later passes through* B_2 *and* B_3*.*

The rate region defined by the weighted sum rate bounds in equation (4.16) *corresponding to* $\omega = 0.2, 0.4, 0.63$ *and the bounds in equation* (4.17) *corresponding to* $\omega = 1.58, 2.25, 5.1$ *is the same*

as given by equation (4.18) *because all the constraints except for the ones used here are redundant. This happens because for all* $\omega \in [0, 0.2]$, $\mathcal{M}(\omega)$*s pass through* A_0*. However the slope of* $\mathcal{M}(\omega)$ *increases with increasing* ω *which means for any value of* $\omega < 0.2$ *the line* $\mathcal{M}(\omega)$ *has a slope less than the slope of line segment* A_0A_1 *in figure 4.4. Thus the line* $\mathcal{M}(0.2)$ *gives the tightest bound for* $\omega \in [0, 0.2]$*. The same argument holds for all the other bounds for the other ranges of* ω*. Later in the chapter we show that the region defined by the sets* $H(\omega)$ *and* $G(\omega)$ *for the predefined ranges of* ω *are indeed piece-wise linear. Thus we see that the region defined by equation* (4.18) *is indeed same as the one given by theorem 6 and the claim in remark 3 holds true for the example.*

The next lemma states the result for a *strong* LFIC.

Lemma 5 *The secrecy capacity of a strong two-user LFIC as shown in figure 4.1 with no channel state information at the transmitter (CSIT) is zero.*

The next section provides all the lemmas needed for the various proofs in this chapter and also introduces the concept of "alignment" for a layered interference channel.

4.4. Key Lemmas

In this section we state a number of lemmas and prove them which we use in the later sections extensively. But before that we introduce the concept of channel "alignment" for layered interference channels like we did for the binary version of fading interference channel. The decoders in this case too decodes independently as in the binary case and hence just depends on the marginal distribution of the outputs conditioned on the inputs and not on the joint conditional distribution of the outputs [1]. Thus any arbitrary distribution can be assumed for the direct and the interfering links without changing the secrecy capacity region as long as the marginal distribution is unaffected. With this prelude we next define $F_{\mathcal{T}}^{-1}(x) = \{\inf v : F_{\mathcal{T}}(v) \geq x\}$ and the channel states from each transmitter as [24]

$$(N_{ii})_t = F_{N_{ii}}^{-1}(\Lambda_t) \quad \text{and} \quad (N_{ij})_t = F_{N_{ij}}^{-1}(\Lambda_t), \tag{4.19}$$

where $i, j \in \{1, 2\}$ and $i \neq j$, $t \geq 1$ and Λ_t is uniformly distributed on $[0, 1]$. Also $F_{\mathcal{T}}(t) = P(\mathcal{T} \leq t)$ is the cummulative distribution function for a random variable $\mathcal{T}$. In general the layers for the direct link and the cross link from a particular transmitter looks as shown in figure 4.6, where the filled blocks represent the unerased layers while the unfilled blocks represent the erased layers. So as it can be seen the unerased layers in general are random and does not have any ordering.

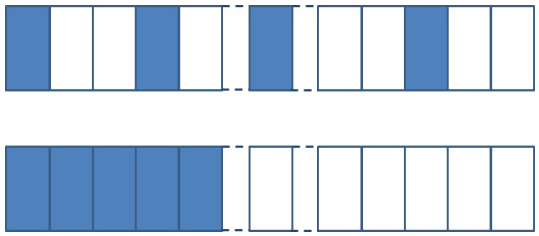

Figure 4.6. The Unaligned Layers between the Direct and the Cross Link from a Transmitter.

The above definition of channel states however ensures that we introduce some ordering to the erased and unerased layers without changing the capacity region. The new definition makes sure that the presence of a particular layer in a weak channel guarantees the presence of the same layer for the stronger channel in other orders the "alignment" ensures that there can be no layer which is present for the weaker channel but not for the stronger one as shown in figure 4.7. So essentially figure 4.7 shows how 4.6 will look after "alignment". It was shown in [24] that this definition of the channel states does not affect the marginal distribution and hence the secrecy capacity region. We are now in a position to state the lemmas used in the rest of the chapter.

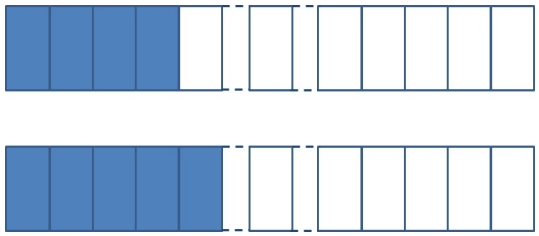

Figure 4.7. The Layers between the Direct and the Cross Link from a Transmitter after being Aligned.

Lemma 6 *Consider n uses of a memoryless channel described by an arbitrary random transformation $P_{Y,Z,T|X,S}$. Let X^n and S^n be the independent input and channel state sequences respectively. Then the difference of the n-letter entropies can be written as a summation of single letter entropies as follows*

$$h(Z^n|T^n,S^n) - h(Y^n|T^n,S^n) = \sum_{i=1}^{n} \Big[h(Z_i|Z^{i-1},Y_{i+1}^n,T^n,S^n) - h(Y_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\Big], \quad (4.20)$$

Proof 1 *The proof of the lemma involves writing the difference of n-letter entropies as a summation of difference of entropies and simple application of chain rule of entropy . The proof is as follows*

$$h(Z^n|T^n,S^n)-h(Y^n|T^n,S^n)=\sum_{i=1}^{n}\Big[h(Z^i,Y_{i+1}^n|T^n,S^n)-h(Z^{i-1},Y_i^n|T^n,S^n)\Big], \tag{4.21}$$

$$=\sum_{i=1}^{n}\Big[h(Z_i,Z^{i-1},Y_{i+1}^n|T^n,S^n)-h(Z^{i-1},Y_i,Y_{i+1}^n|T^n,S^n)\Big],$$

$$=\sum_{i=1}^{n}\Big[h(Z^{i-1},Y_{i+1}^n|T^n,S^n)+h(Z_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)$$

$$-h(Z^{i-1},Y_{i+1}^n|T^n,S^n)-h(Y_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\Big], \tag{4.22}$$

$$=\sum_{i=1}^{n}\Big[h(Z_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)-h(Y_i|Z^{i-1},Y_{i+1}^n,T^n,S^n)\Big], \tag{4.23}$$

where in equation (4.21) *if the summation is completed by finding the terms for each value of* i *then we get the difference of* n*-letter entropy term to the left, equation* (4.22) *follows from simple application of chain rule of entropy. This concludes the proof of the lemma.*

Lemma 7 *Consider the Interference Channel as shown in figure 4.1. If the interference channel satisfies the following constraints* $P(N_{11}\geq l)\geq P(N_{12}\geq l)$ *and* $P(N_{22}\geq l)\geq P(N_{21}\geq l)$ *and are "aligned" ,i.e. presence of a layer for the weaker channel would ensure its presence for the stronger one as well then the input forms the following Markov Chain*

$$W_i\rightarrow\Big\{(W^L)^{i-1},\mu(W^n,N^n)\Big\}\rightarrow\Big\{(W^S)^{i-1},\mu(W^n,N^n)\Big\}, \tag{4.24}$$

where $L=(N_{11}-N_{21})^+$ *and* $S=(N_{12}-N_{22})^+$.

Proof 5 *In general* W_i *is correlated to* $\Big\{(W^L)^{i-1},\mu(W^n,N^n)\Big\}$ *since the input is not necessarily independently distributed over the several channel uses. So that explains the first part of the chain. Now since* $P(N_{11}\geq l)\geq P(N_{12}\geq l)$ *and* $P(N_{22}\geq l)\geq P(N_{21}\geq l)$ *and the channel states are "aligned" so* $\Big\{(W^L)^{i-1},\mu(W^n,N^n)\Big\}$ *can also be written as* $\Big\{(W^S)^{i-1},(W_{S+1}^L)^{i-1},\mu(W^n,N^n)\Big\}$. *Thus we can obtain* $\Big\{(W^S)^{i-1},\mu(W^n,N^n)\Big\}$ *from* $\Big\{(W^L)^{i-1},\mu(W^n,N^n)\Big\}$ *when* $(W_{S+1}^L)^{i-1}$ *fraction of the signal is replaced by* 0 *corresponding to the values of the channel states. Thus*

$$\Big\{(W^S)^{i-1},\mu(W^n,N^n)\Big\}=f\left(\Big\{(W^L)^{i-1},\mu(W^n,N^n)\Big\}\right), \tag{4.25}$$

Now from the definition of data processing inequality we know that for any two sets of correlated random variables $\{A, B\}$ and an arbitary function $f(\cdot)$, there exists a Markov Chain, $A \rightarrow B \rightarrow f(B)$ thereby completing the proof of the Markov Chain

$$W_i \rightarrow \Big\{(W^L)^{i-1}, \mu(W^n, N^n)\Big\} \rightarrow \Big\{(W^S)^{i-1}, \mu(W^n, N^n)\Big\}, \tag{4.26}$$

This concludes the proof of the lemma.

Lemma 8 *For a q- bit layered erasure channel with input X^q and output X^N where N represents the randomly varying channel state having a CCDF $\bar{F}_N(n) = P(N \geq n)\ \forall n \in \{1, \cdots, q\}$, the entropy of the ouput conditioned on the channel state is given by*

$$h(X^N|N) = \sum_{l=1}^{q} P(N \geq l) h(X_l|X^{l-1}), \tag{4.27}$$

Proof 6 *Applying the chain rule of entropy and then reversing the order of summation yields the proof of the lemma. We proceed as follows,*

$$\begin{aligned} h(X^N|N) &= \sum_{n=1}^{q} P_N(n) h(X^n|N=n), \\ &= \sum_{n=1}^{q} P_N(n) h(X^n), \\ &= \sum_{n=1}^{q} \sum_{l=1}^{n} P_N(n) h(X_l|X^{l-1}), \\ &= \sum_{l=1}^{q} \sum_{n=l}^{q} P_N(n) h(X_l|X^{l-1}), \\ &= \sum_{l=1}^{q} P(N \geq l) h(X_l|X^{l-1}), \end{aligned} \tag{4.28}$$

Lemma 9 *Let $X \in \mathbb{F}_2^q$ be random vector and X_Q represent a collection from X whose index belongs to Q,i.e. $X_Q = \{X_l : l \in Q\}$. N is a random variable independent of Q and has a CCDF, $\bar{F}_N(n) = P(N \geq n)\ \forall n \in \{1, \cdots, q\}$ then*

$$h(X^N|X_Q) = \sum_{l \in Q^c} P(N \geq l) h(X_l|X^{l-1}, X_Q), \tag{4.29}$$

Proof 7 *We first apply the result of lemma 8 to expand the term* $h(X^N|X_Q)$ *as follows*

$$\begin{aligned} h(X^N|X_Q) &= \sum_{l=1}^{q} P(N \geq l) h(X_l|X^{l-1}, X_Q), \\ &= \sum_{l \in Q^c} P(N \geq l) h(X_l|X^{l-1}, X_Q), \end{aligned} \tag{4.30}$$

where equation (4.30) *follows from the fact that* $h(X_l|X^{l-1}, X_Q) = 0$ *if* $l \in Q$*. This concludes the proof of the lemma.*

Lemma 10 *Let* A *and* B *be two discrete random variables taking values in* $\{1, 2, \cdots, p\}$ *where* $p \in \mathbb{N}$ *and* $h = \min(A, B)$*, then*

$$\mathbb{E}[h] - \mathbb{E}[A] + \mathbb{E}[A - B]^+ = 0, \tag{4.31}$$

Proof 8 *We will divide it into two cases. In the first case let* $h = A$*, then*

$$\mathbb{E}[h] - \mathbb{E}[A] + \mathbb{E}[A - B]^+ = \mathbb{E}[A] - \mathbb{E}[A], \tag{4.32}$$

$$= 0, \tag{4.33}$$

where equation (4.33) *is the result of the fact that* $min(A, B) = A$ *and* $\mathbb{E}[0] = 0$*. Next we consider the case when* $h = B$*. Then*

$$\mathbb{E}[h] - \mathbb{E}[A] + \mathbb{E}[A - B]^+ = \mathbb{E}[B] - \mathbb{E}[A] + \mathbb{E}[A - B], \tag{4.34}$$

$$= \mathbb{E}[B] - \mathbb{E}[A] + \mathbb{E}[A] - \mathbb{E}[B], \tag{4.35}$$

$$= 0, \tag{4.36}$$

where equation (4.34) *follows from the fact that* $\min(A, B) = B$*, while equation* (4.35) *follows from the linearity property of expectation operator. Finally combining the results of equations* (4.33) *and* (4.36) *we conclude the proof of the lemma.*

We prove the result of *strong* layered fading interference channel given by lemma 5 in the next section.

4.5. Proof of Lemma 5

The proof of the lemma involves finding an upper bound for the individual rates and then showing that these rates cannot be upper bounded by anything other than zero. Let us assume that R_i be the rate of transmission from Tx_i over n channel uses, then

$$nR_i = h(\mathcal{M}_i),\ i = 1, 2, \tag{4.37}$$

where $\mathcal{M}_i$ denotes the message to be transmitted by Tx_i. So now the upper bound on R_1 can be derived as follows

$$\begin{aligned}
nR_1 - \delta_1 &= h(\mathcal{M}_i) - \delta_1, && (4.38)\\
&\leq h(\mathcal{M}_1|Z^n, X^n, N^n), && (4.39)\\
&\leq h(W^n, \mathcal{M}_1|Z^n, X^n, N^n), && (4.40)\\
&= h(W^n|Z^n, X^n, N^n) + h(\mathcal{M}_1|W^n, Z^n, X^n, N^n), && (4.41)\\
&= h(W^n|Z^n, X^n, N^n), && (4.42)\\
&= h(W^n|X^n, N^n) - I(W^n; Z^n|X^n, N^n),\\
&= h(W^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (4.43)\\
&= h(W^n|N^n) - h(W^n|Y^n, N^n) + h(W^n|Y^n, N^n) - I(W^n; \tilde{Y}^n|N^n),\\
&\leq I(W^n; Y^n|N^n) - I(W^n; \tilde{Y}^n|N^n) + \delta^{'}, && (4.44)
\end{aligned}$$

where equation (4.38) follows from (4.37) whereas equation (4.39) follows from the secrecy criterion for $Tx_1 - Rx_1$ pair as mentioned in equation (4.6), equation (4.40) is the result of the fact that additional random variable does not reduce entropy. Equation (4.41) follows from the chain rule of entropy. Equation (4.42) occurs because $h(\mathcal{M}_1|W^n, Z^n, X^n, N^n) = 0$ since a receiver is assumed to reliably decode its message from the received signal. Equation (4.43) follows from the independence of the input random variables and $\tilde{Y}^n = (W^{N_{12}})^n$, $h(W^n|Y^n, N^n) \leq \delta^{'}$ in equation (4.44).Then using the fact $\delta_1^{'} = \delta_1 + \delta^{'}$ in equation (4.45) we further simplify it as follows

$$\begin{aligned}
nR_1 - \delta_1^{'} &\leq I(W^n; Y^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (4.45)\\
&\leq I(W^n; \hat{Y}^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (4.46)
\end{aligned}$$

$$= h(\hat{Y}^n|N^n) - h(\tilde{Y}^n|N^n), \tag{4.47}$$

$$= \sum_{i=1}^{n} \left[h\Big\{\hat{Y}_i|\hat{Y}^{i-1}, \tilde{Y}_{i+1}^n, N^n\Big\} - h\Big\{\tilde{Y}_i|\hat{Y}^{i-1}, \tilde{Y}_{i+1}^n, N^n\Big\} \right], \tag{4.48}$$

$$= \sum_{i=1}^{n} \left[h\Big\{\hat{Y}_i|\mathcal{E}_i, N_i\Big\} - h\Big\{\tilde{Y}_i|\mathcal{E}_i, N_i\Big\} \right], \tag{4.49}$$

$$= \sum_{i=1}^{n} \sum_{l=1}^{q} \Big[P(N_{11} \geq l) - P(N_{12} \geq l) \Big] h\Big\{W_i|\mathcal{E}_i, N_i\Big\}, \tag{4.50}$$

$$\leq 0, \tag{4.51}$$

where (4.46) follows from the fact that independent additive noise cannot increase mutual information, $\hat{Y}^n = (W^{N_{11}})^n$ in the same equation. $h(\hat{Y}^n|W^n, N^n) = 0$ and $h(\tilde{Y}^n|W^n, N^n) = 0$ results in equation (4.47). Equation (4.48) follows from lemma 6, $\mathcal{E}_i = \{\hat{Y}^{i-1}, \tilde{Y}_{i+1}^n, N^{i-1}, N_{i+1}^n\}$ in equation (4.49). Equation (4.50) follows from lemma 8. Since for strong interference $P(N_{11} \geq l) \leq P(N_{12} \geq l)$ and the entropy of a binary random variable is always non-negative hence we can upper bound (4.50) by zero as the value of equation (4.50) would otherwise be negative giving equation (4.51).

Similarly it can be shown using similar technique as above that R_2- the rate of transmission from Tx_2, can be upper bounded by zero as well. This essentially proves that there is no non-negative rate at which when coded we get secret information transmission. Thus the overall secrecy rate for a multi - layered strong fading interference channel is zero. In the next two sections we provide the proof of theorem 6. Its proof consists of two parts - one is the derivation of the upper bound and the other is choice of proper coding technique so as to achieve the upper bound. We prove the converse in the next section.

4.6. Converse

In this section we prove the converse to theorem 6. We divide the proof into three subsections. Subsection 4.6.1 starts with finding the upper bounds for the rate of transmission from Tx_1 and Tx_2 denoted as R_1 and R_2 respectively. In the same subsection we also find the weighted sum bound of the rates for all values of the weighting factor ω. In subsection 4.6.2 we carefully divide the range of ω into two overlapping regions, Region 1 and Region 2 and show that each of those regions are piece wise linear. We also confirm that the region, $\mathcal{R}$, formed by the combination of all the regions defined in this section forms a superset to the secrecy capacity region of *very weak* LFIC. Finally, in subsection 4.6.3, we show that the combination of the two regions along with the the two axes forms a region $\mathcal{R}$, which is subset to the polygonal region defined in theorem 6.

4.6.1. Derivation of the Weighted Sum Bound

We start this subsection by first finding the individual upper bounds of R_1 and R_2 and then combining them with a weighting factor to find the weighted upper bound. The derivation for R_1 starts from equation (4.37) followed by the application of the secrecy criterion for $Tx_1 - Rx_1$ link in (4.6) giving us the starting equation (4.52)

$$
\begin{aligned}
nR_1 - \delta_1 &\leq h(\mathcal{M}_1|Z^n, X^n, N^n), && (4.52)\\
&\leq h(W^n, \mathcal{M}_1|Z^n, X^n, N^n), && (4.53)\\
&= h(W^n|Z^n, X^n, N^n) + h(\mathcal{M}_1|W^n, Z^n, X^n, N^n), && (4.54)\\
&= h(W^n|Z^n, X^n, N^n), && (4.55)\\
&= h(W^n|X^n, N^n) - I(W^n; Z^n|X^n, N^n),\\
&= h(W^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (4.56)\\
&= h(W^n|N^n) - h(W^n|Y^n, N^n) + h(W^n|Y^n, N^n) - I(W^n; \tilde{Y}^n|N^n),\\
&\leq I(W^n; Y^n|N^n) - I(W^n; \tilde{Y}^n|N^n) + \delta^{'}, && (4.57)\\
\implies nR_1 - \delta_1^{'} &= I(W^n; Y^n|N^n) - I(W^n; \tilde{Y}^n|N^n), && (4.58)\\
&= h(Y^n|N^n) - h(Y^n|W^n, N^n) - h(\tilde{Y}^n|N^n), && (4.59)
\end{aligned}
$$

where equation (4.53) is the result of the fact that additional random variable does not reduce entropy.Equation (4.54) follows from the chain rule of entropy. Equation (4.55) occurs because $h(\mathcal{M}_1|W^n, Z^n, X^n, N^n) = 0$ as a receiver is assumed to reliably decode its message from the received signal. Equation (4.56) follows from the independence of random variables and $\tilde{Y}^n = (W^{N_{12}})^n$, $h(W^n|Y^n, N^n) \leq \delta^{'}$ in equation (4.57) and $\delta_1^{'} = \delta_1 + \delta^{'}$ in equation (4.58). Equation (4.59) is the result of the fact that $h(\tilde{Y}^n|W^n, N^n) = 0$. Now we simplify each of the entropy terms seperately as follows

$$
\begin{aligned}
&h(Y^n|N^n)\\
&= h\Big\{(W^k \oplus X^k)^n, (W^L)^n, (X^K)^n|N^n\Big\}, && (4.60)\\
&= h\Big\{(W^k \oplus X^k)^n|N^n\Big\} + h\Big\{(W^L)^n|(W^k \oplus X^k)^n, N^n\Big\} + h\Big\{(X^K)^n|(W^L)^n, (W^k \oplus X^k)^n, N^n\Big\},
\end{aligned}
$$

$$\leq h\Big\{(W^k \oplus X^k)^n|N^n\Big\} + h\Big\{(W^L)^n|N^n\Big\} + h\Big\{(X^K)^n|N^n\Big\}, \tag{4.61}$$

$$= h\Big\{\bar{Y}^n|N^n\Big\} + h\Big\{(W^L)^n|N^n\Big\} + h\Big\{(X^K)^n|N^n\Big\}, \tag{4.62}$$

where in equation (4.60), $k = \min(N_{11}, N_{21})$, $K = (N_{21} - N_{11})^+$ and L has been defined before in lemma 7, equation (4.61) follows from the fact that conditioning reduces entropy and $\bar{Y}^n = (W^k \oplus X^k)^n$ in equation (4.62). Similarly

$$h(Y^n|W^n, N^n) = h\Big\{(X^{N_{21}})^n|N^n\Big\} = h(\tilde{Z}^n|N^n), \tag{4.63}$$

where $\tilde{Z}^n = (X^{N_{21}})^n$ in equation (4.63). Therefore the bound for R_1 can be written as follows

$$nR_1 - \delta_1^{'} \leq h\Big\{\bar{Y}^n|N^n\Big\} + h\Big\{(W^L)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\} + h\Big\{(X^K)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}, \tag{4.64}$$

Proceeding as above we can similarly find the bound for R_2 as follows

$$nR_2 - \delta_2^{'} \leq h\Big\{\bar{Z}^n|N^n\Big\} + h\Big\{(X^T)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\} + h\Big\{(W^S)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\}, \tag{4.65}$$

where $\bar{Z}^n = (W^m \oplus X^m)^n$, with $m = \min(N_{22}, N_{12})$ whereas $T = (N_{22} - N_{12})^+$ and S has been defined before in lemma 7. Now finding the weighted sum bound by adding equation (4.64) with ω times of equation (4.65) we get,

$$\begin{aligned}
&n(R_1 + \omega R_2) - (\delta_1^{'} + \omega\delta_2^{'}) \\
&\leq h\Big\{\bar{Y}^n|N^n\Big\} + \omega h\Big\{\bar{Z}^n|N^n\Big\} + \Bigg[h\Big\{(W^L)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\}\Bigg] + \omega\Bigg[h\Big\{(W^S)^n|N^n\Big\} \\
&-h\Big\{\tilde{Y}^n|N^n\Big\}\Bigg] + \Bigg[h\Big\{(X^K)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}\Bigg] + \omega\Bigg[h\Big\{(X^T)^n|N^n\Big\} - h\Big\{\tilde{Z}^n|N^n\Big\}\Bigg],
\end{aligned} \tag{4.66}$$

Next we will try to simplify the pair of entropies inside each of the square braces seperately using the Marton style expansion as follows,

$$\begin{aligned}
&h\Big\{(W^L)^n|N^n\Big\} - h\Big\{\tilde{Y}^n|N^n\Big\} \\
&\leq \sum_{i=1}^{n}\Bigg[h\Big\{(W^L)_i|(W^L)^{i-1}, \tilde{Y}_{i+1}^n, N^n\Big\} - h\Big\{\tilde{Y}_i|(W^L)^{i-1}, \tilde{Y}_{i+1}^n, N^n\Big\}\Bigg],
\end{aligned} \tag{4.67}$$

$$= \sum_{i=1}^{n} \left[h\Big\{ (W^L)_i | \mathcal{D}_i, N_i \Big\} - I\Big\{ \tilde{Y}_i | \mathcal{D}_i, N_i \Big\} \right], \tag{4.68}$$

$$= \sum_{i=1}^{n} \sum_{l=1}^{q} \Big[P(L \geq l) - P(N_{12} \geq l) \Big] h\Big\{ (W_l)_i | (W^{l-1})_i, \mathcal{D}_i, N_i \Big\}, \tag{4.69}$$

where equation (4.67) follows from lemma 6, $\mathcal{D}_i = \{(W^L)^{i-1}, \tilde{Y}_{i+1}^n, N^{i-1}, N_{i+1}^n\}$ in equation (4.68) and equation (4.69) is the result of lemma 8. Using the same lemmas as above and the same method of simplification we can simplify the other entropy pair involving W as well to give us the following

$$h\Big\{ (W^S)^n | N^n \Big\} - h\Big\{ \tilde{Y}^n | N^n \Big\}$$

$$\leq \sum_{i=1}^{n} \sum_{l=1}^{q} \Big[P(S \geq l) - P(N_{12} \geq l) \Big] h\Big\{ (W_l)_i | (W^{l-1})_i, \mathcal{E}_i, N_i \Big\}, \tag{4.70}$$

$$\leq \sum_{i=1}^{n} \sum_{l=1}^{q} \Big[P(S \geq l) - P(N_{12} \geq l) \Big] h\Big\{ (W_l)_i | (W^{l-1})_i, \mathcal{E}_i, (W_{S+1}^L)^{i-1}, N_i \Big\}, \tag{4.71}$$

$$\leq \sum_{i=1}^{n} \sum_{l=1}^{q} \Big[P(S \geq l) - P(N_{12} \geq l) \Big] h\Big\{ (W_l)_i | (W^{l-1})_i, \mathcal{D}_i, N_i \Big\}, \tag{4.72}$$

where in equation (4.70) $\mathcal{E}_i = \{(W^S)^{i-1}, \tilde{Y}_{i+1}^n, N^{i-1}, N_{i+1}^n\}$, equation (4.71) is the result of lemma 7. Thus combining equation (4.69) and (4.72) together we can write the following

$$\Big[h\Big\{ (W^L)^n | N^n \Big\} - h\Big\{ \tilde{Y}^n | N^n \Big\} \Big] + \omega \Big[h\Big\{ (W^S)^n | N^n \Big\} - h\Big\{ \tilde{Y}^n | N^n \Big\} \Big],$$

$$\leq \sum_{i=1}^{n} \sum_{l=1}^{q} \Big\{ \Big[P(L \geq l) - P(N_{12} \geq l) \Big] + \omega \Big[P(S \geq l) - P(N_{12} \geq l) \Big] \Big\} h\Big\{ (W_l)_i | (W^{l-1})_i, \mathcal{D}_i, N_i \Big\}, \tag{4.73}$$

Similarly the entropy terms involving X can be combined in a similar fashion and simplified to give us the following

$$\omega \Big[h\Big\{ (X^T)^n | N^n \Big\} - h\Big\{ \tilde{Z}^n | N^n \Big\} \Big] + \Big[h\Big\{ (X^K)^n | N^n \Big\} - h\Big\{ \tilde{Z}^n | N^n \Big\} \Big],$$

$$\leq \sum_{i=1}^{n} \sum_{l=1}^{q} \Big\{ \omega \Big[P(T \geq l) - P(N_{21} \geq l) \Big] + \Big[P(K \geq l) - P(N_{21} \geq l) \Big] \Big\} h\Big\{ (X_l)_i | (X^{l-1})_i, \mathcal{C}_i, N_i \Big\}, \tag{4.74}$$

where $\mathcal{C}_i = \{(X^T)^{i-1}, \tilde{Z}_{i+1}^n, N^{i-1}, N_{i+1}^n\}$ in equation (4.74). So the weighted sum bound of equation (4.66) can be further simplified using equations (4.73) and (4.74) as follows

$$
\begin{aligned}
&n(R_1+\omega R_2)-(\delta_1^{'}+\omega\delta_2^{'})\\
&\leq \sum_{i=1}^{n}\Bigg[h(\bar{Y}_i|\bar{Y}^{i-1},N^n)+\omega h(\bar{Z}_i|\bar{Z}^{i-1},N^n)\\
&+\sum_{l=1}^{q}\Big\{\big[P(L\geq l)-P(N_{12}\geq l)\big]+\omega\big[P(S\geq l)-P(N_{12}\geq l)\big]\Big\}h\Big\{(W_l)_i|(W^{l-1})_i,\mathcal{D}_i,N_i\Big\}\\
&+\sum_{l=1}^{q}\Big\{\omega[P(T\geq l)-P(N_{21}\geq l)\big]+\big[P(K\geq l)-P(N_{21}\geq l)\big]\Big\}h\Big\{(X_l)_i|(X^{l-1})_i,\mathcal{C}_i,N_i\Big\}\Bigg], \qquad (4.75)\\
&\leq n(\mathbb{E}[k]+\omega\mathbb{E}[m])+\sum_{i=1}^{n}\sum_{l=1}^{q}\Bigg[\Big\{\big[P(L\geq l)-P(N_{12}\geq l)\big]-\omega\alpha_1(l)\Big\}h\Big\{(W_l)_i|(W^{l-1})_i,\mathcal{D}_i,N_i\Big\}\\
&\qquad+\Big\{\omega\big[P(T\geq l)-P(N_{21}\geq l)\big]-\alpha_2(l)\Big\}h\Big\{(X_l)_i|(X^{l-1})_i,\mathcal{C}_i,N_i\Big\}\Bigg], \qquad (4.76)\\
&=n(\mathbb{E}[k]+\omega\mathbb{E}[m])+\sum_{i=1}^{n}\sum_{l=1}^{q}\Bigg[\alpha_1(l)\Big\{\beta_2(l)-\omega\Big\}h\Big\{(W_l)_i|(W^{l-1})_i,\mathcal{D}_i,N_i\Big\}\\
&\qquad+\big[P(T\geq l)-P(N_{21}\geq l)\big]\Big\{\omega-\beta_1(l)\Big\}h\Big\{(X_l)_i|(X^{l-1})_i,\mathcal{C}_i,N_i\Big\}\Bigg], \qquad (4.77)\\
&=n(\mathbb{E}[k]+\omega\mathbb{E}[m])+n\sum_{l=1}^{q}\Bigg[\alpha_1(l)\Big\{\beta_2(l)-\omega\Big\}^{+}+\big[P(T\geq l)-P(N_{21}\geq l)\big]\Big\{\omega-\beta_1(l)\Big\}^{+}\Bigg] \qquad (4.78)
\end{aligned}
$$

where equation (4.75) follows by applying the chain rule on the first two entropy terms of (4.66), in equation (4.76) $\alpha_1(l)$ and $\alpha_2(l)$ has already been defined in section 4.3 and the expectation term occurs by application of lemma 8 and the fact that the entropy of a binary random variable can be upper bounded by 1. $\beta_1(l)$ and $\beta_2(l)$ in equation (4.77) is as defined in (4.8). Finally equation (4.78) is the result of using the fact that the entropy of a binary random variable can be upper bounded by 1. Now dividing both sides by n and letting $n\to\infty$, $\delta_1^{'},\delta_2^{'}\to 0$ we get

$$
R_1+\omega R_2\leq \mathbb{E}[k]+\omega\mathbb{E}[m]+\sum_{l=1}^{q}\alpha_1(l)\Big\{\beta_2(l)-\omega\Big\}^{+}+\sum_{l=1}^{q}\big[P(T\geq l)-P(N_{21}\geq l)\big]\Big\{\omega-\beta_1(l)\Big\}^{+}, \qquad (4.79)
$$

The range of ω is divided into two overlapping regions in the next subsection and then the bound of (4.79) is simplified further for each of those regions and then they are shown to be piece-wise linear.

4.6.2. $\mathcal{R}$ is a Superset to the Secrecy Capacity Region of the *Very Weak* LFIC

In this subsection the entire range of ω is divided into two overlapping regions as shown in figure 4.3 and are defined as follows

$$\begin{aligned} &\text{Region 1} : \omega \in [0, \omega_1) \ \& \\ &\text{Region 2} : \omega \in (\gamma_1, \infty) \end{aligned} \tag{4.80}$$

where ω_1 and γ_1 are as defined in remark 4. Now we further simplify (4.79) for each of the above regions and show that the bound for each of the regions is piece-wise linear.

4.6.2.1. Region 1: $\omega \in [0, \omega_1)$

Starting from (4.79) we get

$$\begin{aligned} &R_1 + \omega R_2 \\ &\leq \mathbb{E}[k] + \omega\mathbb{E}[m] + \sum_{l=1}^{q} \alpha_1(l)\Big\{\beta_2(l) - \omega\Big\} + \sum_{l\in\mathcal{A}(\omega)} \big[P(T \geq l) - P(N_{21} \geq l)\big]\Big\{\omega - \beta_1(l)\Big\}, \quad (4.81) \\ &= \mathbb{E}[k] + \omega\mathbb{E}[m] + \sum_{l=1}^{q} \Big\{\big[P(L \geq l) - P(N_{12} \geq l)\big] - \omega\alpha_1(l)\Big\} \\ &\qquad\qquad + \sum_{l\in\mathcal{A}(\omega)} \Big\{\omega\big[P(T \geq l) - P(N_{21} \geq l)\big] - \alpha_2(l)\Big\}, \\ &\leq \left[\sum_{l=1}^{q} \Big[\big\{P(k \geq l) + P(L \geq l)\big\} - P(N_{12} \geq l)\Big] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l)\right] \\ &\qquad\qquad + \omega\left[\mathbb{E}[m] + \sum_{l\in\mathcal{A}(\omega)} \Big[P(T \geq l) - P(N_{21} \geq l)\Big] - \sum_{l=1}^{q} \alpha_1(l)\right], \quad (4.82) \end{aligned}$$

$$= \Big\{\mathbb{E}[N_{11}] - \mathbb{E}[N_{12}] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l)\Big\} + \omega\Big\{\mathbb{E}[m] + \sum_{l\in\mathcal{A}(\omega)} [P(T \geq l) - P(N_{21} \geq l)] - \sum_{l=1}^{q} \alpha_1(l)\Big\}, \tag{4.83}$$

$$= \Big\{\mathbb{E}[N_{11}] - \mathbb{E}[N_{12}] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l)\Big\} + \omega\Big\{\sum_{l\in\mathcal{A}(\omega)} [P(T \geq l) - P(N_{21} \geq l)]\Big\}, \tag{4.84}$$

where $\mathcal{A}(\omega)$ in equation (4.81) is as defined before and it also uses the fact $\omega < \beta_2(l), \forall l \in \{1, \cdots, q\}$ when ω is in region 1. Equation (4.83) results from the fact $\sum_{l=1}^{q} \Big[P(k \geq l) + P(L \geq l)\Big] =$

$\mathbb{E}[N_{11}]$.Equation (4.84) follows from lemma 10 with $A = N_{12}$ and $B = N_{22}$. We can further see that the expression in (4.84) is same as the one in the $H(\omega)$ defined with respect to equation (4.13) and hence region 1 can just be defined by the set $H(\omega)$. In the following lemma we show that region 1 defined by the set $H(\omega)$ as ω varies from $[0, \omega_1)$ is piece-wise linear.

Lemma 11 *For region* 1*, except for* $H(\gamma_0), H(\gamma_1), \cdots, H(\gamma_q)$*, all other* $H(\omega)$*'s are redundant, i.e.*

$$\underset{\omega\in[0,\omega_1)}{\cap} H(\omega) = H(\gamma_0) \cap H(\gamma_1) \cap \cdots \cap H(\gamma_q), \tag{4.85}$$

where $\gamma_i s$ *for all* $i \in \{0, 1, 2, \cdots, q\}$ *and* ω_1 *are all defined in remark 4.*

Proof 9 *For* $i = 0, \cdots, q$*, the boundary of* $H(\gamma_i)$ *and* $H(\gamma_{i+1})$ *intersects at*

$$\Big\{\mathbb{E}[N_{11}] - \mathbb{E}[N_{12}] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l), \sum_{l\in\mathcal{A}(\omega)} [P(T \geq l) - P(N_{21} \geq l)]\Big\}, \tag{4.86}$$

which is denoted as A_i *as in figure 4.2. It can be easily verified from* (4.84)*. If we next define an interval* $\Omega_i = (\gamma_i, \gamma_{i+1})$*, then it is not difficult to see that* $H(\omega) = H(\gamma_i)$*,* $\forall \omega \in \Omega_i$*. Thus we see that the bound* $H(\omega)$ *is redundant to* $H(\gamma_i)$ *and* $H(\gamma_{i+1})$*. Hence the lemma is proved.*

4.6.2.2. Region 2: $\omega \in (\gamma_1, \infty)$

Again starting from equation (4.79) we further simplify the bound for region 2 as follows

$$\begin{aligned}
&R_1 + \omega R_2 \\
&\leq \mathbb{E}[k] + \omega\mathbb{E}[m] + \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l)\Big\{\beta_2(l) - \omega\Big\} + \sum_{l=1}^{q} \big[P(T \geq l) - P(N_{21} \geq l)\big]\Big\{\omega - \beta_1(l)\Big\}, \qquad (4.87)
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}[k] + \omega\mathbb{E}[m] + \sum_{l\in\mathcal{B}(\omega)} \Big\{\big[P(L \geq l) - P(N_{12} \geq l)\big] - \omega\alpha_1(l)\Big\} \\
&\qquad\qquad + \sum_{l=1}^{q} \Big\{\omega\big[P(T \geq l) - P(N_{21} \geq l)\big] - \alpha_2(l)\Big\}, \\
&= \left[\mathbb{E}[k] + \sum_{l\in\mathcal{B}(\omega)} \Big[P(L \geq l) - P(N_{12} \geq l)\Big] - \sum_{l=1}^{q} \alpha_2(l)\right] \\
&\qquad\qquad + \omega\left[\sum_{l=1}^{q} \Big[\big\{P(m \geq l) + P(T \geq l)\big\} - P(N_{21} \geq l)\Big] - \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l)\right], \qquad (4.88)
\end{aligned}$$

$$= \Big\{\mathbb{E}[k] + \sum_{l\in\mathcal{B}(\omega)} [P(L \geq l) - P(N_{12} \geq l)] - \sum_{l=1}^{q} \alpha_2(l)\Big\} + \omega\Big\{\mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l)\Big\}, \tag{4.89}$$

$$= \Big\{ \sum_{l\in\mathcal{B}(\omega)} [P(L \geq l) - P(N_{12} \geq l)]\Big\} + \omega\Big\{\mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l)\Big\}, \tag{4.90}$$

where $\mathcal{B}(\omega)$ in equation (4.87) is as defined before and the fact also used is $\omega > \beta_1(l), \forall l \in \{1, \cdots, q\}$ when ω is in region 2. Equation (4.89) results from the fact $\sum_{l=1}^{q} \Big[P(m \geq l) + P(T \geq l)\Big] = \mathbb{E}[N_{22}]$. Equation (4.90) follows from lemma 10 with $A = N_{21}$ and $B = N_{11}$. The expression in (4.90) is exactly same as the one in $G(\omega)$ defined for equation (4.14) and hence region 2 can essentially be defined by the set $G(\omega)$. In the following lemma we show that region 2 defined by the set $G(\omega)$ as ω varies from (γ_1, ∞) is also piece-wise linear.

Lemma 12 *For region 2, except for $G(\omega_0), G(\omega_1), \cdots, G(\omega_q)$, all other $G(\omega)$'s are redundant, i.e.*

$$\bigcap_{\omega\in(\gamma_1,\infty)} G(\omega) = G(\omega_0) \cap G(\omega_1) \cap \cdots \cap G(\omega_q), \tag{4.91}$$

where $\omega_i s$ for all $i \in \{0, 1, 2, \cdots, q\}$ and γ_1 are as defined in remark 4.

Proof 10 *The proof of lemma 12 is same as the proof of lemma 11 except for the change in notations and the fact that $G(\omega_i)$ and $G(\omega_{i+1})$ for all $i \in \{1, \cdots, q\}$ intersect at*

$$\Big\{ \sum_{l\in\mathcal{B}(\omega)} [P(L \geq l) - P(N_{12} \geq l)], \mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l)\Big\}, \tag{4.92}$$

The details of the proof is omitted to avoid repetition.

Region 1 and region 2 when combined together gives an upper bound for all the values of ω. It thus defines the entire secrecy capacity upper bound region for a *very weak* LFIC, $\mathcal{R}$, and can be mathematically represented as shown in equation (4.15). Now since any achievable secrecy rate pair for the *very weak* LFIC has to either satisfy the bounds in equations (4.84) or (4.90) or both depending on the value of ω, so the secrecy capacity region of this channel will be a subset to the region $\mathcal{R}$.

4.6.3. $\mathcal{R}$ is a Subset to $\mathcal{C}(N)$

In this section we show that the region $\mathcal{R}$ is a subset of the polygonal region $\mathcal{C}(N)$. We complete this proof by method of contradiction where we prove that a if a point lies outside the region defined by $\mathcal{C}(N)$ then it violates one of the bounds among (4.84) or (4.90) depending on the value of ω. This in turn by transposition logic would mean that if a rate pair belongs to $\mathcal{R}$ then it must belong to the region $\mathcal{C}(N)$ which mathematically means

$$\mathcal{R} \subseteq \mathcal{C}(N), \tag{4.93}$$

It is not difficult to see that the general shape of the polygon as shown in figure 4.2 is essentially a combination of several quadrangles and a triangle as shown in figure 4.8. Then comparing the co-ordinates of A_k with A_{k+1} in region 1 (or B_k with B_{k+1} in region 2) we see that A_{k+1} will be above A_k (B_{k+1} will be below B_k) on the polygon and should be shifted a little bit to the left with respect to A_k (to the right with respect to B_k). Further, we represent the boundaries of the piece-wisely linear regions 1 and 2, proved in lemmas 11 and 12, respectively by a straight line, replacing the inequality in their definitions by equality. Let $\mathcal{M}(\gamma_i)$ be the lines representing the boundaries of region 1 while $\mathcal{M}(\omega_i)$ be the lines for region 2 for all $i \in \{1, 2, \cdots, q\}$. So for example $\mathcal{M}(\gamma_q)$ represents the line A_0A_1, $\mathcal{M}(\gamma_{q-1})$ represents the line A_1A_2 and so on in figure 4.8. Similarly $\mathcal{M}(\omega_q)$ represents the line B_0B_1, $\mathcal{M}(\omega_{q-1})$ represents the line B_1B_2 and so on. Then depending on the value of ω, if it is in region 1 then the lines $\mathcal{M}(\gamma_i)$ and $\mathcal{M}(\gamma_{i+1})$ intersects at the point given by equation (4.86) whereas if it is in region 2 then the lines $\mathcal{M}(\omega_i)$ and $\mathcal{M}(\omega_{i+1})$ intersect at the point given by(4.92). The points A_0 and B_0 lie on the R_1 and R_2 axis respectively.

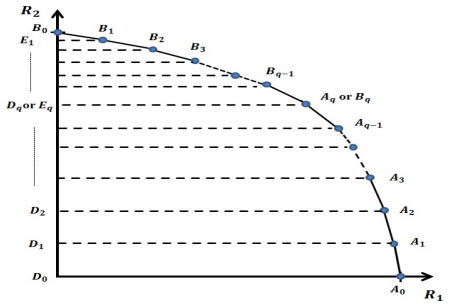

Figure 4.8. Shape of the Polygonal Superset to $\mathcal{R}$.

Now let us consider a rate pair $(\bar{R}_1, \bar{R}_2)$ and let us assume that the rate pair lies outside the polygon shown in figure 4.8. Now if the rate pair lies outside the polygon then it has three options. It can either lie outside all the quadrangles defined by $A_iA_{i+1}D_{i+1}D_i$ where $i \in \{0, 1, \cdots, q-1\}$ but clearly in that case it will violate the the bound from which we obtain the line $\mathcal{M}(\gamma_{q-i})$ by replacing inequality with equality. Otherwise, it can lie outside the quadrangles defined by $B_jB_{j+1}E_{j+1}E_j$ where $j \in \{1, 2, \cdots, q-1\}$ but in that case it will violate the bound from which we obtained the line $\mathcal{M}(\omega_{q-j})$. If the above two conditions are not satisfied then its third option is, it has to lie outside the triangle $B_0E_1B_1$ but in that case it will violate the bound from which we obtained the line $\mathcal{M}(\omega_q)$. This proves that if a rate pair lies outside the polygon, then it violates one of the bounds defining region 1 or 2 and thereby lies outside region $\mathcal{R}$. So now by application of transposition logic we conclude our claim.

Remark 7 *The intersecting point of the bounds of two regions would occur when $\gamma_1 < \omega < \omega_1$ and the bound is given by*

$$
\begin{aligned}
&R_1 + \omega R_2 \\
&\leq \mathbb{E}[k] + \omega\mathbb{E}[m] + \sum_{l=1}^{q} \alpha_1(l)\Big\{\beta_2(l) - \omega\Big\}^+ + \sum_{l=1}^{q} \big[P(T \geq l) - P(N_{21} \geq l)\big]\Big\{\omega - \beta_1(l)\Big\}^+, \quad (4.94)\\
&= \mathbb{E}[k] + \omega\mathbb{E}[m] + \sum_{l=1}^{q} \Big\{\big[P(L \geq l) - P(N_{12} \geq l)\big] - \omega\alpha_1(l)\Big\} \\
&\qquad\qquad + \sum_{l=1}^{q} \Big\{\omega\big[P(T \geq l) - P(N_{21} \geq l)\big] - \alpha_2(l)\Big\}, \quad (4.95)\\
&= \sum_{l=1}^{q} \left[\Big\{P(k \geq l) + P(L \geq l)\Big\} - P(N_{12} \geq l) - \alpha_2(l)\right] \\
&\qquad\qquad + \omega \sum_{l=1}^{q} \left[\big\{P(m \geq l) + P(T \geq l)\big\} - P(N_{21} \geq l) - \alpha_1(l)\right], \\
&= \Big\{\mathbb{E}[N_{11}] - \mathbb{E}[N_{12}] - \sum_{l=1}^{q} \alpha_2(l)\Big\} + \omega\Big\{\mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l=1}^{q} \alpha_1(l)\Big\}, \quad (4.96)
\end{aligned}
$$

where equation (4.94) *is same as* (4.79)*, equation* (4.95) *follows since for the given range of ω both $\Big\{\beta_2(l) - \omega\Big\}$ and $\Big\{\omega - \beta_1(l)\Big\}$ are positive for all $l \in \{1, \cdots, q\}$. Equation* (4.96) *results from the fact $\sum_{l=1}^{q} \Big[P(k \geq l) + P(L \geq l)\Big] = \mathbb{E}[N_{11}]$ and $\sum_{l=1}^{q} \Big[P(m \geq l) + P(T \geq l)\Big] = \mathbb{E}[N_{22}]$. Thus the bounds*

in this range of ω from both the regions intersect at the point

$$\Big\{\mathbb{E}[N_{11}] - \mathbb{E}[N_{12}] - \sum_{l=1}^{q} \alpha_2(l), \mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l=1}^{q} \alpha_1(l)\Big\}, \tag{4.97}$$

In the next section we try to achieve the dominant corner points that forms the vertices of the polygonal region defining the secrecy capacity, thereby concluding the proof of theorem 6.

4.7. Achievability

In the previous section we proved that the secrecy capacity region of *very weak* LFIC is contained in a polygon and it also characterizes the dominant corner points of the polygon. Each side of this polygon joins the two adjacent corner points. From figure 4.8, it is clear that if a rate pair representing the corner points of the polygon can be achieved by some coding scheme then any rate pair on the line joining the corner points can be achieved by time sharing. As a result we try to prove the achievability of the dominant corner points as shown in figure 4.8. We divide this section into two parts, one showing the achievability of the corner points for region 1 and the other for that of region 2.

4.7.1. Achievability of the Dominant Corner Points in Region 1

Tx_1 uses a random codebook such that each entry of the codebook consists of Bernoulli($\frac{1}{2}$), i.e. $W_l = \tilde{W}_l \sim \mathscr{B}(\frac{1}{2}),\ \forall l \in \{1, \cdots, q\}$. On the other hand Tx_2 uses a scheme in which it transmits independent and identically distributed (IID) $\mathscr{B}(\frac{1}{2})$ symbols only on layers that belongs to $\mathcal{A}(\omega)$ and remains silent for the other layers, i.e.$X = \hat{X}$ such that

$$\hat{X}_l : \begin{cases} \hat{X}_l \sim \mathscr{B}(\frac{1}{2}) & \text{if } l \in \mathcal{A}(\omega), \\ \phi & \text{if } l \in \mathcal{A}^c(\omega), \end{cases} \tag{4.98}$$

where $\mathcal{A}^c(\omega) = \{1, \cdots, q\} \backslash \mathcal{A}(\omega)$. We further go ahead and define another random variable $\tilde{X} \in \mathbb{F}_2^q$ from $\hat{X}$ defined above as follows

$$\tilde{X}_l = \begin{cases} \hat{X}_l & \text{if } l \in \mathcal{A}(\omega), \\ \mathscr{B}(\frac{1}{2}) & \text{if } l \in \mathcal{A}^c(\omega), \end{cases} \tag{4.99}$$

This makes all the components of $\tilde{X}$ as IID $\mathscr{B}(\frac{1}{2})$.

Next we find the rates supported by the direct links from each of the transmitters, such that when coded at those rates the intended receiver can decode his own message.

4.7.1.1. Rate of the codebooks at the transmitters

Let $r_1^{'}(\omega)$ represent the rate of the direct link from Tx_1 to Rx_1 and $r_2^{'}(\omega)$ represent the rate of the direct link from Tx_2 to Rx_2. So first we try to calculate $r_1^{'}(\omega)$ as follows

$$
\begin{aligned}
r_1^{'}(\omega) &= I(\tilde{W}; \tilde{W}^{N_{11}} \oplus \hat{X}^{N_{21}} | N), \\
&= h(\tilde{W}^{N_{11}} \oplus \hat{X}^{N_{21}} | N) - h(\tilde{W}^{N_{11}} \oplus \hat{X}^{N_{21}} | \tilde{W}, N), \\
&= h(\tilde{W}^{N_{11}} \oplus \tilde{X}^{N_{21}} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) - h(\tilde{X}^{N_{21}} | \tilde{X}_{\mathcal{A}^c(\omega)}, N), && (4.100) \\
&= h(\tilde{W}^{N_{11}} \oplus \tilde{X}^{N_{21}}_{(N_{21}-N_{11})^+ + 1} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) + h(\tilde{X}^{(N_{21}-N_{11})^+} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) \\
&\qquad - h(\tilde{X}^{N_{21}} | \tilde{X}_{\mathcal{A}^c(\omega)}, N), && (4.101) \\
&= \sum_{l=1}^{q} P(N_{11} \geq l) + \sum_{l \in \mathcal{A}(\omega)} P(N_{21} - N_{11} \geq l) - \sum_{l \in \mathcal{A}(\omega)} P(N_{21} \geq l), && (4.102) \\
&= \mathbb{E}[N_{11}] - \sum_{l \in \mathcal{A}(\omega)} \alpha_2(l), && (4.103)
\end{aligned}
$$

where equation (4.100) follows from the definition of $\tilde{X}$ in (4.99), simple application of chain rule of entropy leads to equation (4.101), equation (4.102) is the result of application of both lemmas 8 and 9 and the fact that both $\tilde{X}$ and $\tilde{W}$ are $\mathscr{B}(\frac{1}{2})$ for all $l \geq 1$.

Similarly we can find $r_2^{'}(\omega)$ as follows

$$
\begin{aligned}
r_2^{'}(\omega) &= I(\hat{X}; \tilde{W}^{N_{12}} \oplus \hat{X}^{N_{22}} | N), \\
&= h(\tilde{W}^{N_{12}} \oplus \hat{X}^{N_{22}} | N) - h(\tilde{W}^{N_{12}} \oplus \hat{X}^{N_{22}} | \hat{X}, N), \\
&= h(\tilde{W}^{N_{12}} \oplus \tilde{X}^{N_{22}} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) - h(\tilde{W}^{N_{12}} | N), && (4.104) \\
&= h(\tilde{W}^{m} \oplus \tilde{X}^{m}, \tilde{W}^{(N_{12}-N_{22})^+}, \tilde{X}^{(N_{22}-N_{12})^+} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) - h(\tilde{W}^{N_{12}} | N), && (4.105) \\
&= h(\tilde{W}^{m} \oplus \tilde{X}^{m} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) + h(\tilde{X}^{(N_{22}-N_{12})^+} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) + h(\tilde{W}^{(N_{12}-N_{22})^+} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) \\
&\qquad - h(\tilde{W}^{N_{12}} | N), && (4.106) \\
&= h(\tilde{W}^{m} | N) + h(\tilde{X}^{(N_{22}-N_{12})^+} | \tilde{X}_{\mathcal{A}^c(\omega)}, N) + h(\tilde{W}^{(N_{12}-N_{22})^+} | N) - h(\tilde{W}^{N_{12}} | N), && (4.107) \\
&= \sum_{l=1}^{q} P(m \geq l) + \sum_{l \in \mathcal{A}(\omega)} P(T \geq l) + \sum_{l=1}^{q} P(N_{12} - N_{22} \geq l) - \sum_{l=1}^{q} P(N_{12} \geq l), && (4.108)
\end{aligned}
$$

$$= \mathbb{E}[m] + \sum_{l \in \mathcal{A}(\omega)} P(T \geq l) - \sum_{l=1}^{q} \alpha_1(l), \tag{4.109}$$

$$= \sum_{l \in \mathcal{A}(\omega)} P(T \geq l), \tag{4.110}$$

where equation (4.104) follows from the definition of $\tilde{X}$ in (4.99), $m = \min\{N_{22}, N_{12}\}$ in equation (4.105) as before, simple application of chain rule of entropy and the fact that the inputs are IID leads to equation (4.106). The first term in (4.107) follows from the fact that $\tilde{W}^m$ is identically distributed as $\tilde{W}^m \oplus \tilde{X}^m$ while the third term follows from the fact that the inputs are independent. $T = (N_{22} - N_{12})^+$ in equation (4.108) and it occurs because of the application of lemmas 8 and lemma 9 and the fact that both $\tilde{X}$ and $\tilde{W}$ are $\mathscr{B}(\frac{1}{2})$ for all $l \geq 1$. Equation (4.110) results from using lemma 10 with $A = N_{12}$ and $B = N_{22}$.

Thus each of the receiver can decode it's own message when their corresponding transmitter uses a random codebook of rate $\mathbb{E}[N_{11}] - \sum_{l \in \mathcal{A}(\omega)} \alpha_2(l)$ for Tx_1 and $\sum_{l \in \mathcal{A}(\omega)} P(T \geq l)$ for Tx_2. We next find the achievable secrecy rate at each of the transmitters and observe if they match with the corner points specified by equation (4.86).

4.7.1.2. Achievable Secrecy Rate

From [28] we know that on a wiretap channel with input A, legitimate receiver output B, eavesdropper signal C and a channel $P_{BC|A}(\cdot)$, the secrecy capacity achievable is given as

$$\max_{J \to A \to BC} \{I(J;B) - I(J;C)\}, \tag{4.111}$$

It implies for each choice of (J, A) that satisfies the above Markov Chain , a secrecy rate of $\{I(J;B) - I(J;C)\}$ can be achieved on a wiretap channel. Using this result and the fact that interference channel can be thought of as a combination of two wiretap channels one at each transmitter we can find the achievable secrecy rate at each transmitter. Using $J = A \equiv \tilde{W}$, $B \equiv (Y, N)$ and $C \equiv (Z, \hat{X}, N)$ and denoting the corresponding achievable secrecy rate by $r_1(\omega)$ we find the achievable secrecy rate at Tx_1 when ω is in region 1 as follows,

$$\begin{aligned} r_1(\omega) &= I\Big\{\tilde{W}; Y, N\Big\} - I\Big\{\tilde{W}; Z, \hat{X}, N\Big\}, \\ &= I\Big\{\tilde{W}; N\Big\} + I\Big\{\tilde{W}; Y|N\Big\} - I\Big\{\tilde{W}; N\Big\} - I\Big\{\tilde{W}; \hat{X}|N\Big\} - I\Big\{\tilde{W}; Z|\hat{X}, N\Big\}, \end{aligned} \tag{4.112}$$

$$= I\Big\{\tilde{W}; Y|N\Big\} - I\Big\{\tilde{W}; Z|\hat{X}, N\Big\}, \tag{4.113}$$

$$= \mathbb{E}[N_{11}] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l) - h\{\tilde{W}^{N_{12}}|\hat{X}, N\Big\}, \tag{4.114}$$

$$= \mathbb{E}[N_{11}] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l) - \sum_{l=1}^{q} P(N_{12} \geq l), \tag{4.115}$$

$$= \mathbb{E}[N_{11}] - \mathbb{E}[N_{12}] - \sum_{l\in\mathcal{A}(\omega)} \alpha_2(l), \tag{4.116}$$

where equation (4.112) follows from the chain rule of mutual information, equation (4.113) is the result of the independence of the inputs, equation (4.114) follows from equation (4.103), lemma 8 is used for equation (4.115) along with the fact that $\tilde{W} \sim \mathscr{B}(\frac{1}{2})$. Finally equation (4.116) matches with the abcissa of (4.86).

Next using the same result of [28] we try to find the secrecy rate achievable by Tx_2 but now we use $J = A \equiv \hat{X}$, $B \equiv (Z, N)$ and $C \equiv (Y, \tilde{W}, N)$, denoting the rate as $r_2(\omega)$ we find the rate as follows,

$$\begin{aligned} r_2(\omega) &= I\Big\{\hat{X}; Z, N\Big\} - I\Big\{\hat{X}; Y, \tilde{W}, N\Big\}, \\ &= I\Big\{\hat{X}; Z|N\Big\} - I\Big\{\hat{X}; Y|\tilde{W}, N\Big\}, && (4.117) \\ &= \sum_{l\in\mathcal{A}(\omega)} P(T \geq l) - h\{\hat{X}^{N_{21}}|N\Big\}, && (4.118) \\ &= \sum_{l\in\mathcal{A}(\omega)} P(T \geq l) - h\{\tilde{X}^{N_{21}}|\tilde{X}_{\mathcal{A}^c(\omega)}, N\Big\}, && (4.119) \\ &= \sum_{l\in\mathcal{A}(\omega)} P(T \geq l) - \sum_{l\in\mathcal{A}(\omega)} P(N_{21} \geq l), && (4.120) \\ &= \sum_{l\in\mathcal{A}(\omega)} \Big[P(T \geq l) - P(N_{21} \geq l)\Big], && (4.121) \end{aligned}$$

where equation (4.117) follows from application of chain rule of mutual information and then cancellation of terms due to independence as in the other case,equation (4.118) follows from equation (4.110),equation (4.119) follows from the definition of $\tilde{X}$ in (4.99), application of lemma 9 along with the fact that $\tilde{X} \sim \mathscr{B}(\frac{1}{2})$ gives equation (4.120). Finally equation (4.121) matches with the ordinate in (4.86). Thus we see that all the corner points given by (4.86) in region 1 can be achieved using the above strategy.

4.7.2. Achievability of the Dominant Corner Points in Region 2

Tx_1 in region 2 uses a strategy similar to Tx_2 in region 1, i.e. it uses the statistics of the channel to find $\mathcal{B}(\omega)$ and then transmits IID $\mathscr{B}(\frac{1}{2})$ symbols only on those layers belonging to $\mathcal{B}(\omega)$, while remaining silent for the rest. Mathematically, $W = \hat{W}$ such that

$$\hat{W}_l : \begin{cases} \hat{W}_l \sim \mathscr{B}(\frac{1}{2}) & \text{if } l \in \mathcal{B}(\omega), \\ \phi & \text{if } l \in \mathcal{B}^c(\omega), \end{cases} \tag{4.122}$$

where $\mathcal{B}^c(\omega) = \{1, \cdots, q\} \backslash \mathcal{B}(\omega)$. We further go ahead and define another random variable $\tilde{W} \in \mathbb{F}_2^q$ from $\hat{W}$ defined above as follows

$$\tilde{W}_l = \begin{cases} \hat{W}_l & \text{if } l \in \mathcal{B}(\omega), \\ \mathscr{B}(\frac{1}{2}) & \text{if } l \in \mathcal{B}^c(\omega), \end{cases} \tag{4.123}$$

This makes all the components of $\tilde{W}$ as IID $\mathscr{B}(\frac{1}{2})$.

Tx_2 on the other hand uses a point-to-point (PTP) layered erasure channel capacity achieving code, $X_l = \tilde{X}_l \sim \mathscr{B}(\frac{1}{2})$, $\forall l \in \{1, \cdots, q\}$. Next Tx_1 uses a random codebook of rate $\sum_{l \in \mathcal{B}(\omega)} P(L \geq l)$ where $L = (N_{11} - N_{21})^+$ and Tx_2 uses a codebook of rate $\mathbb{E}[N_{22}] - \sum_{l \in \mathcal{B}(\omega)} \alpha_1(l)$. We will first verify that each receiver can decode the message coming from its direct transmitter at the above rates.

4.7.2.1. Rate of the codebooks at the transmitters

Now let us say that Rx_1 can decode any message coming at a rate $r_1^{''}(\omega)$ or less from Tx_1 while Rx_2 can decode any message coming at a rate $r_2^{''}(\omega)$ or less from Tx_2. Now,

$$\begin{aligned} r_1^{''}(\omega) &= I(\hat{W}; \tilde{X}^{N_{21}} \oplus \hat{W}^{N_{11}} | N), \\ &= h(\tilde{X}^{N_{21}} \oplus \hat{W}^{N_{11}} | N) - h(\tilde{X}^{N_{21}} \oplus \hat{W}^{N_{11}} | \hat{W}, N), \\ &= h(\tilde{W}^{N_{11}} \oplus \tilde{X}^{N_{21}} | \tilde{W}_{\mathcal{B}^c(\omega)}, N) - h(\tilde{X}^{N_{21}} | N), \qquad (4.124) \\ &= h(\tilde{W}^k \oplus \tilde{X}^k, \tilde{W}^{(N_{11}-N_{21})^+}, \tilde{X}^{(N_{21}-N_{11})^+} | \tilde{W}_{\mathcal{B}^c(\omega)}, N) - h(\tilde{X}^{N_{21}} | N), \qquad (4.125) \\ &= h(\tilde{W}^k \oplus \tilde{X}^k | \tilde{W}_{\mathcal{B}^c(\omega)}, N) + h(\tilde{W}^{(N_{11}-N_{21})^+} | \tilde{W}_{\mathcal{B}^c(\omega)}, N) + h(\tilde{X}^{(N_{21}-N_{11})^+} | \tilde{W}_{\mathcal{B}^c(\omega)}, N) \\ &\qquad - h(\tilde{X}^{N_{21}} | N), \qquad (4.126) \end{aligned}$$

$$= h(\tilde{X}^k|N) + h(\tilde{W}^{(N_{11}-N_{21})^+}|\tilde{W}_{\mathcal{B}^c(\omega)}, N) + h(\tilde{X}^{(N_{21}-N_{11})^+}|N) - h(\tilde{X}^{N_{21}}|N), \tag{4.127}$$

$$= \sum_{l=1}^{q} P(k \geq l) + \sum_{l\in\mathcal{B}(\omega)} P(L \geq l) + \sum_{l=1}^{q} P(N_{21} - N_{11} \geq l) - \sum_{l=1}^{q} P(N_{21} \geq l), \tag{4.128}$$

$$= \mathbb{E}[k] + \sum_{l\in\mathcal{B}(\omega)} P(L \geq l) - \sum_{l=1}^{q} \alpha_2(l), \tag{4.129}$$

$$= \sum_{l\in\mathcal{B}(\omega)} P(L \geq l), \tag{4.130}$$

where (4.124) follows from the fact that $\tilde{W}_l$ is same as $\hat{W}_l$ if $l \in \mathcal{B}(\omega)$, $k = \min\{N_{11}, N_{21}\}$ in equation (4.125) as defined before,simple application of chain rule of entropy and the fact that the inputs are IID and are independent of each other leads to equation (4.126). The first term in (4.127) follows from the fact that $\tilde{X}^k$ is identically distributed as $\tilde{W}^k \oplus \tilde{X}^k$ while the third term follows from the independence of $\tilde{W}$ and $\tilde{X}$. Equation (4.128) is the result of lemmas 8 and 9 and the fact that $\tilde{X}_l, \tilde{W}_l \sim \mathscr{B}(\frac{1}{2})$, $\forall l \in \{1, \cdots, q\}$. Finally equation (4.130) follows from lemma 10 with $A = N_{21}$ and $B = N_{11}$. Next we prove the same for Rx_2 as follows

$$\begin{aligned}
r_2^{''}(\omega) &= I(\tilde{X}; \tilde{X}^{N_{22}} \oplus \hat{W}^{N_{12}}|N), \\
&= h(\tilde{X}^{N_{22}} \oplus \hat{W}^{N_{12}}|N) - h(\tilde{X}^{N_{22}} \oplus \hat{W}^{N_{12}}|\tilde{X}, N), \\
&= h(\tilde{X}^{N_{22}} \oplus \tilde{W}^{N_{12}}|\tilde{W}_{\mathcal{B}^c(\omega)}, N) - h(\tilde{W}^{N_{12}}|\tilde{W}_{\mathcal{B}^c(\omega)}, N), \qquad (4.131) \\
&= h(\tilde{X}^{N_{22}} \oplus \tilde{W}^{N_{12}}_{(N_{12}-N_{22})^+ +1}|\tilde{W}_{\mathcal{B}^c(\omega)}, N) + h(\tilde{W}^{(N_{12}-N_{22})^+}|\tilde{W}_{\mathcal{B}^c(\omega)}, N) \\
&\qquad - h(\tilde{W}^{N_{12}}|\tilde{W}_{\mathcal{B}^c(\omega)}, N), \qquad (4.132) \\
&= \sum_{l=1}^{q} P(N_{22} \geq l) + \sum_{l\in\mathcal{B}(\omega)} P(N_{12} - N_{22} \geq l) - \sum_{l\in\mathcal{B}(\omega)} P(N_{12} \geq l), \qquad (4.133) \\
&= \mathbb{E}[N_{22}] - \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l), \qquad (4.134)
\end{aligned}$$

where (4.131) follows from the fact that $\tilde{W}_l$ is same as $\hat{W}_l$ if $l \in \mathcal{B}(\omega)$, simple application of chain rule of entropy gives equation (4.132). Equation (4.133) follows from the fact that $\tilde{X}_l$ & $\tilde{W}_l \sim \mathscr{B}(\frac{1}{2})$, $\forall l \in \{1, \cdots, q\}$ and the application of lemmas 8 and 9. Finally equation (4.134) confirms that Rx_2 can decode everything if Tx_2 uses a codebook of rate $\mathbb{E}[N_{22}] - \sum_{l\in\mathcal{B}(\omega)} \alpha_1(l)$.

4.7.2.2. Achievable Secrecy Rate

As in case of region 1 if we use the same concept and same A, B and C as was used for Tx_1 in 4.7.1.2 then if $r_1(\omega)$ represents the achievable secrecy at Tx_1, then after the initial simplification involving chain rule of mutual information and independence of the inputs we get,

$$\begin{aligned} r_1(\omega) &= I\Big\{\hat{W}; \hat{W}^{N_{11}} \oplus \tilde{X}^{N_{21}} | N\Big\} - I\Big\{\hat{W}; \hat{W}^{N_{12}} | N\Big\}, \\ &= \sum_{l \in \mathcal{B}(\omega)} P(L \geq l) - h(\tilde{W}^{N_{12}} | \tilde{W}_{\mathcal{B}^c(\omega)}, N), \qquad (4.135) \\ &= \sum_{l \in \mathcal{B}(\omega)} P(L \geq l) - \sum_{l \in \mathcal{B}(\omega)} P(N_{12} \geq l), \qquad (4.136) \end{aligned}$$

where (4.135) follows from the fact that $\tilde{W}_l$ is same as $\hat{W}_l$ if $l \in \mathcal{B}(\omega)$ and also from equation (4.130), simple application of lemma 9 and the fact that $\tilde{W}_l \sim \mathscr{B}(\frac{1}{2}),\ \forall l \in \{1, \cdots, q\}$ gives us (4.136) which is same as the abcissa of the dominant corner points of region 2 given by equation (4.92).

Next let us represent the achievable secrecy rate of Tx_2 as $r_2(\omega)$, then applying the same strategy as in region 1 for Tx_2 and after the initial simplification as before we get,

$$\begin{aligned} r_2(\omega) &= I\Big\{\tilde{X}; \tilde{X}^{N_{22}} \oplus \hat{W}^{N_{12}} | N\Big\} - I\Big\{\tilde{X}; \tilde{X}^{N_{21}} | N\Big\}, \\ &= \mathbb{E}[N_{22}] - \sum_{l \in \mathcal{B}(\omega)} \alpha_1(l) - h(\tilde{X}^{N_{21}} | N), \qquad (4.137) \\ &= \mathbb{E}[N_{22}] - \sum_{l \in \mathcal{B}(\omega)} \alpha_1(l) - \sum_{l=1}^{q} P(N_{21} \geq l), \qquad (4.138) \\ &= \mathbb{E}[N_{22}] - \mathbb{E}[N_{21}] - \sum_{l \in \mathcal{B}(\omega)} \alpha_1(l), \qquad (4.139) \end{aligned}$$

where (4.137) follows from equation (4.134), equation (4.138) follows from lemma 8 and the fact that $\tilde{X}_l \sim \mathscr{B}(\frac{1}{2}),\ \forall l \in \{1, \cdots, q\}$. Equation (4.139) matches with the ordinate of the corner point given by (4.92). Thus we see that all the corner points given by (4.92) in region 2 can be achieved using the above strategy.

Moreover the intersection point of the two regions given by equation (4.97) can be achieved by using any one of the strategies used for the two regions.

4.8. Conclusion

In this chapter we have characterized the exact secrecy capacity of a *strong* and *very weak* layered fading IC. This result is a first of its kind in fading interference channel where the transmitters do not have any knowledge about the channel states. For the *strong* LFIC we determined that it is not possible to get any secrecy rate if the interfering link is stronger (as per our definition) than the direct link. While for the *very weak* case we found that a trade off in the rate of transmission between the two transmitters can help achieve a positive secrecy rate. The achievability involves careful distribution of the layers among the transmitters depending on the channel states. It is followed by usage of a layered wiretap channel optimal code at both the transmitters and *treat interference as erasure* to complete the scheme. This result suggests that although inherently more complicated than the binary case in the previous chapter, the intuition obtained from the less complicated binary helped in extending the result of the binary to multi-layer scenario. The hope is may be the intuitions from the binary and layered fading IC can help in solving the real fading interference channel problem.

5. CONCLUSION AND FUTURE WORKS

Information theoretic secrecy model are mostly build of the very popular wiretap channel (WC) model, first introduced by Wyner in [3]. Ever since then there has been several efforts to solve the problem under various assumptions. The wiretap channel with a single transmitter and two receivers- one legitimate and the other eavesdropper, might appear to be a relatively simple channel. However the simplicity reduces when the channel starts varying with time [1] and the transmitter is no longer aware of the instantaneous channel states. It is no wonder then that although conserted efforts has been made to solve the problem on fading gaussian wiretap channel, except for some special cases [25],[50] it was still largely open. In this thesis we address that problem and characterize secrecy capacity of a fading gaussian wiretap channel within constant number of bits for any arbitrary channel distribution. In the absence of an universal coding scheme that can achieve the secrecy rates for any kind of channel distribution, this is the best effort that has been made so far in characterizing the secrecy capacity. Figure 5.1 shows a practical wiretap channel. For such a channel we show that secrecy can be achieved within 11 bits of the upper bound irrespective of the channel distribution. However our simulation results show that in most of the cases the gap between achievability and the upper bound is within a couple of bits. This simulation result is further strengthened by some of the examples that we considered in chapter 2.

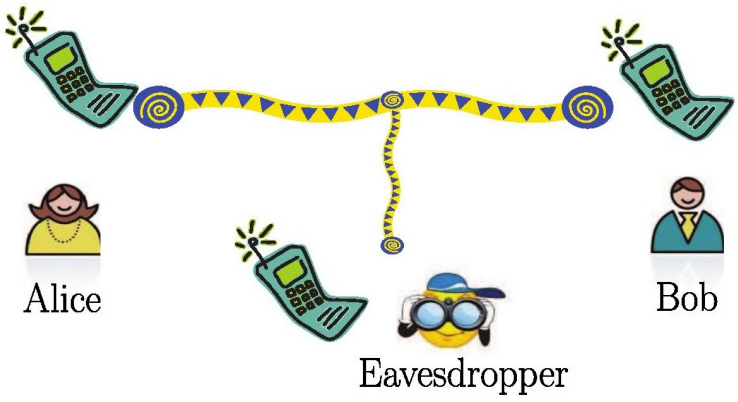

Figure 5.1. A Practical Wiretap Channel.

[1] In techincal terms such a channel is called a *fading* channel.

From the single user model we move on to the multi-user model next as we consider a 2-user interference channel (IC). Interference channel is one of the known complicated models used to study secrecy from information theoretic point.

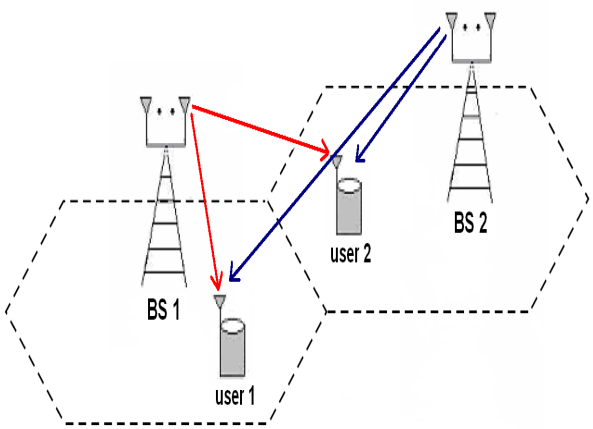

Figure 5.2. A Practical Interference Channel.

It consists of two transmitters and two receivers. The transmit receive pairs are meant to talk to each other, but due to the broadcast nature of wireless medium they get interfered from the other receiver. Figure 5.2 shows a practical interference channel scenario. As can be seen in this scenario BS1 is talking to its nearest user, user1 but user2 can also hear him due its proximity, similarly user1 can also listen from BS2 as BS2 communicates with user2. Just like wiretap channel several work caters to the problem of interference channel under several assumptions. However due to the associated complexity that comes with a fading channel with no CSIT, no work is present under these conditions. In order to reduce the complexity we consider a binary fading version of the real fading IC in chapter 3. In that chapter we characterize the exact secrecy capacity for some custom defined interference channel called the *strong* and *very weak* binary fading IC. We extend the binary result to a multi-layer scenario in chapter 4 and characterize the capacity for a *strong* and *very weak* fading IC. This kind of layered approach to finding the capacity of a real fading channel has been used previously for a fading broadcast channel [56], fading ZIC [24] and fading WC as seen in chapter 2.

The same approach has been taken here with an eye to solving the real fading problem for interference channel as some future research effort. The current result also provides motivation to solve the fading interference channel problem for the other regions defined in this thesis. It might

be interesting to investigate if the current results provide some intuitions to solve the problem for the other regions too. Due to the complicated nature of interfernce channel people often study a variation of it where one of the interfering link is absent as shown in figure 5.3. Such a channel is called the Z-Interference Channel (ZIC) due to its topological architecture.

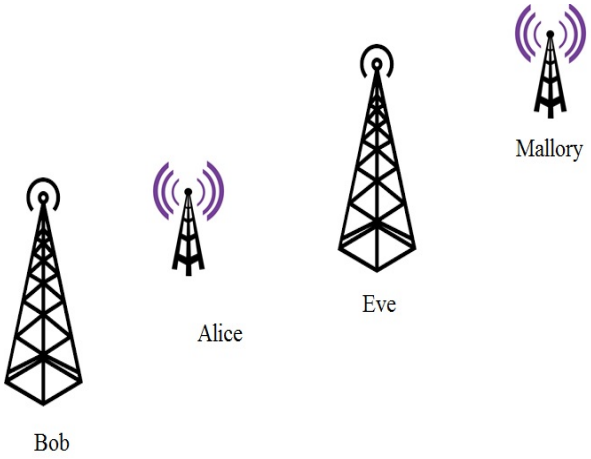

Figure 5.3. A Practical Z - Interference Channel.

Figure 5.3 shows a practical ZIC. In this figure Alice is in between Bob and Eve and hence both of them can listen to Alice creating a security concern whenever Alice wants to communicate with Bob. However Mallory is closer to Eve but far away from Bob, so Bob cannot listen to her. The communication between Mallory and Eve is thus inherently secured. Such a variation of the interference channel under the same channel assumptions as has been considered in this thesis might provide for another interesting future research problem.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley and Sons Inc, 1991.

[2] C. E. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, Oct 1949.

[3] A. Wyner, "The wiretap channel," *The Bell System Technical Journal*, vol. 54, pp. 1355–1387, 1975.

[4] I. Csiszár, "Almost independence and secrecy capacity," *Problemy Peredachi Informatsii*, vol. 32, no. 1, pp. 48–57, 1996.

[5] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *International Conference on the Theory and Applications of Cryptographic Techniques.* Springer, 2000, pp. 351–368.

[6] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J. M. Merolla, "Applications of ldpc codes to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, Aug 2007.

[7] V. Rathi, M. Andersson, R. Thobaben, J. Kliewer, and M. Skoglund, "Performance analysis and design of two edge-type ldpc codes for the bec wiretap channel," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 1048–1064, Feb 2013.

[8] D. Klinc, J. Ha, S. W. McLaughlin, J. Barros, and B. J. Kwak, "Ldpc codes for the gaussian wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 532–540, Sept 2011.

[9] M. Baldi, M. Bianchi, and F. Chiaraluce, "Coding with scrambling, concatenation, and harq for the awgn wire-tap channel: A security gap analysis," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 883–894, June 2012.

[10] S. Sharifi, A. K. Tanc, and T. M. Duman, "Implementing the han kobayashi scheme using low density parity check codes over gaussian interference channels," *IEEE Transactions on Communications*, vol. 63, no. 2, pp. 337–350, Feb 2015.

[11] S. Shari, A. K. Tanc, and T. M. Duman, "Ldpc code design for binary-input binary-output z interference channels," in *IEEE International Symposium on Information Theory*, June 2015, pp. 1084–1088.

[12] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, Oct 2011.

[13] E. Hof and S. Shamai, "Secrecy-achieving polar-coding," in *IEEE Information Theory Workshop*, Aug 2010, pp. 1–5.

[14] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested polar codes for wiretap and relay channels," *IEEE Communications Letters*, vol. 14, no. 8, pp. 752–754, August 2010.

[15] E. Şaşoğlu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *IEEE International Symposium on Information Theory*, July 2013, pp. 1117–1121.

[16] T. C. Gulcu and A. Barg, "Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component," in *IEEE Information Theory Workshop*, April 2015, pp. 1–5.

[17] L. Wang, *Channel Coding Techniques for Network Communication*, 2015. [Online]. Available: https://books.google.com/books?id=2ZjOjgEACAAJ

[18] C. Hirche, C. Morgan, and M. M. Wilde, "Polar codes in network quantum information theory," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 915–924, Feb 2016.

[19] F. Lin and F. Oggier, "A classification of unimodular lattice wiretap codes in small dimensions," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3295–3303, June 2013.

[20] J. C. Belfiore and F. Oggier, "Lattice code design for the rayleigh fading wiretap channel," in *IEEE International Conference on Communications Workshops*, June 2011, pp. 1–5.

[21] L. C. Choo, C. Ling, and K. K. Wong, "Achievable rates for lattice coded gaussian wiretap channels," in *IEEE International Conference on Communications Workshops*, June 2011, pp. 1–5.

[22] C. Ling, L. Luzzi, J. C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the gaussian wiretap channel," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6399–6416, Oct 2014.

[23] D. N. C. Tse and R. D. Yates, "Fading broadcast channels with state information at the receivers," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3453–3471, June 2012.

[24] Y. Zhu and D. Guo, "Ergodic fading z-interference channels without state information at transmitters," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2627–2647, May 2011.

[25] M. K. Hossan, S. Karmakar, and A. Ghosh, "On the secrecy capacity of fading gaussian wiretap channel," in *IEEE 14th Canadian Workshop on Information Theory*, 2015, pp. 36–40.

[26] C. E. Shannon, "Two-way communication channel," in *Proceedings of 4th Berkeley Symposium Mathematical Statistics and Probability*, vol. 1, Mar, 1961, pp. 611–644.

[27] U. M. Maurer, S. Wolf, and E. B. Preneel, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proceedings 19th International Conference on the Theory and Application of Cryptographic Techniques*, 2000, pp. 351–368.

[28] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24(3), pp. 339–348, May, 1978.

[29] S. K. Leung-Yan-Cheong and M. E. Hellman, "The gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, pp. 451–456, Jul, 1978.

[30] T. Liu and S.Shamai, "A note on the secrecy capacity of the multiple antenna wiretap channel," *IEEE Transaction on Information Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[31] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wire-tap channel," *IEEE Transactions on Information Theory*, vol. 57, pp. 4961–4971, Aug, 2011.

[32] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas i: The MISOME wiretap channel," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[33] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channel," in *IEEE International Symposium on Information Theory*, Sep. 2005, pp. 2152–2155.

[34] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *IEEE International Symposium on Information Theory*, 2006, pp. 356–360.

[35] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

[36] P. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, pp. 5059–5067, Nov, 2008.

[37] Y. Liang, L. Lai, H. V. Poor, and S. Shamai, "A broadcast approach for fading wiretap channels," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 842–858, 2014.

[38] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[39] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *44th Annual Allerton Conference*, 2006, pp. 1–5.

[40] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[41] E. Ekrem and S. Ulukus, "Ergodic secrecy capacity region of the fading broadcast channel," in *IEEE International Conference on Communications*, 2009, pp. 1–5.

[42] J. Li and A. Petropulu, "On the ergodic secrecy rate for gaussian miso wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1176–1187, Apr. 2011.

[43] S. Shafiee and S. Ulukus, "Achievable rates in gaussian MISO channels with secrecy constraints," in *IEEE International Symposium on Information Theory*, 2007, pp. 2466–2470.

[44] Z. Li, R. Yates, and W. Trappe, "Achieving secret communication for fast rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2792–2799, 2010.

[45] A. Hyadi, Z. Rezki, and M. S. Alouini, "Secure multiple-antenna block-fading wiretap channels with limited CSI feedback," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6618–6634, 2017.

[46] J. Hu, W. Yang, N. Yang, X. Zhou, and Y. Cai, "On-off-based secure transmission design with outdated channel state information," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6075–6088, 2016.

[47] Z. Rezki, A. Khisti, and M. S. Alouini, "On the secrecy capacity of the wiretap channel with imperfect main channel estimation," *IEEE Transactions on Communications*, vol. 62, no. 10, pp. 3652–3664, 2014.

[48] A. Benfarah, S. Tomasin, and N. Laurenti, "Parallel BCC with one common and two confidential messages and imperfect CSIT," in *IEEE Globecom Workshops*, 2014, pp. 1373–1378.

[49] A. Hyadi, Z. Rezki, A. Khisti, and M. S. Alouini, "Secure broadcasting with imperfect channel state information at the transmitter," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2215–2230, 2016.

[50] P. H. Lin and E. Jorswieck, "On the fast fading gaussian wiretap channel with statistical channel state information at the transmitter," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 46–58, 2016.

[51] R. Negi and S. Goel, "Secret communication using artificial noise," in *IEEE 62nd Vehicular Technology Conference, 2005.*, vol. 3, 2005, pp. 1906–1910.

[52] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, 2008.

[53] A. L. Swindlehurst, "Fixed sinr solutions for the mimo wiretap channel," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 2437–2440.

[54] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 3831–3842, 2010.

[55] Q. Xiong, Y. Gong, and Y.-C. Liang, "Achieving secrecy capacity of MISO fading wiretap channels with artificial noise," in *FJOIEEE Wireless Communications and Networking Conference*, 2013, pp. 2452–2456.

[56] D. N. C. Tse and R. D. Yates, "Fading broadcast channels with state information at the receivers," *IEEE Transactions on Information Theory*, vol. 58, pp. 3453–3471, Jun, 2012.

[57] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM Journal on Computing*, vol. 17, pp. 210–229, 1988.

[58] M. Bloch and J. Barros, *Physical-Layer security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[59] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, pp. 5534–5562, Dec, 2008.

[60] S. Karmakar and M. K. Varanasi, "The capacity region of the MIMO interference channel and its reciprocity to within a constant gap," *IEEE Transactions on Information Theory*, vol. 59, pp. 4781–4797, Aug, 2013.

[61] P. H. Lin, E. A. Jorswieck, and R. F. Schaefer, "On ergodic fading gaussian interference channels with statistical csit," in *IEEE Information Theory Workshop*, 2016, pp. 454–458.

[62] M. K. Hossan and S. Karmakar, "Secrecy capacity of the ergodic layered erasure wiretap channel," in *Proceedings of the 49th Annual Conference on Information Systems and Sciences*, march 18 to 20, 2015.

[63] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[64] R. D. Yates and J. Lei, "Gaussian fading broadcast channels with csi only at the receivers: An improved constant gap," in *IEEE International Symposium on Information Theory Proceedings*, 2011, pp. 2969–2973.

[65] P. K. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4687–4698, Oct 2008.

[66] Z. Rezki, A. Khisti, and M. Alouini, "On the ergodic secrecy capacity of the wiretap channel under imperfect main channel estimation," in *45th Asilomar Conference on Signals, Systems and Computers*, Nov, 2011, pp. 952–957.

[67] J. Li and A. Petropulu, "On the ergodic secrecy rate for gaussian miso wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1176–1187, Apr. 2011.

[68] Z. Li, R. Yates, and W. Trappe, "Achieving secret communication for fast rayleigh fading channels," *IEEE Transactions on Information Theory*, vol. 9, no. 9, pp. 2792–2799, Sep., 2010.

[69] M. K. Hossan and S. Karmakar, "Secrecy capacity of the ergodic layered erasure wiretap channel," in *49th Annual Conference on Information Sciences and Systems*, March 2015, pp. 1–5.

[70] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1550–1573, Third 2014.

[71] N. Liu and W. Kang, "The secrecy capacity region of a special class of multiple access channels," in *Proceedings IEEE International Symposium on Information Theory*, Jul, 2011, pp. 623–627.

[72] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in *46th Annual Allerton Conference on Communication, Control, and Computing*, Sep, 2008, pp. 1014–1021.

[73] E. Tekin and A. Yener, "The general gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2735–2751, June 2008.

[74] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, March 2008.

[75] Z. Li, R. D. Yates, and W. Trappe, "Secrecy capacity region of a class of one-sided interference channel," in *IEEE International Symposium on Information Theory*, July 2008, pp. 379–383.

[76] P. Mohapatra, C. R. Murthy, and J. Lee, "On the secrecy capacity region of the two-user symmetric z interference channel with unidirectional transmitter cooperation," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 572–587, March 2017.

[77] J. Chen, "New results on the secure capacity of symmetric two-user interference channels," in *54th Annual Allerton Conference on Communication, Control, and Computing*, Sept 2016, pp. 524–531.

[78] C. Geng, R. Tandon, and S. A. Jafar, "On the symmetric 2-user deterministic interference channel with confidential messages," in *IEEE Global Communications Conference*, Dec 2015, pp. 1–6.

[79] R. D. Yates, D. Tse, and Z. Li, "Secret communication on interference channels," in *2008 IEEE International Symposium on Information Theory*, July 2008, pp. 374–378.

[80] J. Xie and S. Ulukus, "Secure degrees of freedom of k -user gaussian interference channels: A unified view," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2647–2661, May 2015.

[81] C. Geng and S. A. Jafar, "Secure gdof of k-user gaussian interference channels: When secrecy incurs no penalty," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1287–1290, Aug 2015.

[82] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[83] D. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proceedings IEEE International Symposium on Information Theory*, June-July, 1997.

[84] A. Vahid, M. A. Maddah-Ali, A. S. Avestimehr, and Y. Zhu, "Binary fading interference channel with no csit," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3565–3578, June 2017.

[85] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 3425–3441, Aug, 2008.

[86] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, pp. 4566–4592, Sep, 2010.

[87] S. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric gaussian K user interference channel," *IEEE Transactions on Information Theory*, vol. 56, pp. 3297–3303, Jul, 2010.

[88] S. A. Jafar and S. Shamai, "Degrees of freedom region of the MIMO x channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 151–170, Jan, 2008.

[89] C. S. Vaze, S. Karmakar, and M. K. Varanasi, "On the generalized degrees of freedom region of the MIMO interference channel with No CSIT," in *Proceedings IEEE International Symposium on Information Theory*, July, 2011.

[90] C. S. Vaze and M. K. Varanasi, "The degrees of freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," Sept. 2009, available Online: http://arxiv.org/abs/0909.5424.

[91] P. A. Parker, D. W. Bliss, and V. Tarokh, "On the degrees-of-freedom of the MIMO interference channel," in *Proceedings of Information Sciences and Systems*, Mar, 2008, pp. 62–67.

[92] D. Tuninetti, "Gaussian fading interference channels: Power control," in *42nd Asilomar Conference on Signals, Systems and Computers*, Oct 2008, pp. 701–706.

[93] L. Sankar, X. Shang, E. Erkip, and H. V. Poor, "Ergodic fading interference channels: Sum-capacity and separability," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2605–2626, May 2011.

[94] Y. Zhu and C. Shen, "On layered erasure interference channels without csi at transmitters," in *IEEE International Symposium on Information Theory*, July 2016, pp. 710–714.

[95] D. Tse, R. Yates, and Z. Li, "Fading broadcast channels with state information at the receivers," in *46th Annual Allerton Conference on Communication, Control, and Computing*, Sept 2008, pp. 221–227.

[96] V. Aggarwal, L. Sankar, A. R. Calderbank, and H. V. Poor, "Ergodic layered erasure one-sided interference channels," in *IEEE Information Theory Workshop*, Oct 2009, pp. 574–578.

APPENDIX

A.1. Proof of Lemma 1

To prove this Lemma, we first establish a rate achievable by the BES scheme [56] on a real channel. Since, a real fading complex channel is equivalent to two real fading real channels, the desired result follows just by multiplying the achievable rate of the real channel by two. So, in what follows, we consider a real PTP channel.

The BES scheme expresses a real symbol - to be transmitted - as a weighted sum of several Binary antipodal symbols, i.e.,

$$\tilde{X}_r = \frac{\sqrt{3}}{2} \sum_{n=1}^{\infty} x_{r,n} 2^{-n}, \tag{A.1}$$

where $x_{r,n} \in \{+1, -1\}$ for all $n \geq 1$ and $\frac{\sqrt{3}}{2}$ is a normalizing constant. The corresponding output at the receiver on a PTP real fading channel with the square of the magnitude of the instantaneous channel coefficient denoted by S, can be written as

$$\begin{aligned} \tilde{T}_r &= \frac{\sqrt{3S}}{2} \sum_{n=1}^{\infty} x_{r,n} 2^{-n} + U_r \\ \implies \tilde{T}'_r &= \sum_{n=1}^{\infty} x_{r,n} 2^{-n} + \frac{2}{\sqrt{3S}} U_r, \end{aligned} \tag{A.2}$$

where in the last equation $\tilde{T}'_r$ is obtained by scaling the output by the instantaneous channel magnitude which is known at the receiver, U_r's are IID as $N(0, 0.5)$. In this scheme, $x_{r,n}$ is typically referred to as the symbol transmitted via the n-th layer. Estimates of the antipodal symbols $\{x_{r,n}\}_{n=1}^{\infty}$, denoted by $\{\hat{x}_{r,n}\}_{n=1}^{\infty}$ are then extracted from the received real symbol, $\tilde{T}'_r$, in the following manner:

$$\sum_{n=1}^{\infty} \hat{x}_{r,n} 2^{-n} = \max(-1, \min[1, \tilde{T}'_r]), \tag{A.3}$$

where the above equation yields a unique solution for $\{\hat{x}_{r,n}\}_{n=1}^{\infty}$, since any real number with modulus less than or equal to one has a unique antipodal expansion. In this decoding scheme, an antipodal symbol $x_{r,n}$ is estimated at the receiver as $\hat{x}_{r,n}$ and therefore, at any given channel use, the real fading channel is equivalent to a collection of *random* Binary Symmetric Channels (rBSCs); one

for each value of n, i.e., $x_{r,n} \to \hat{x}_{r,n}$ for all $n \geq 1$. Subsequently, the crossover probabilities of all such rBSCs can also be computed analytically and depends on the instantaneous value of S. It was shown in Lemma 5 of [56] that if the crossover probability of the n-th layer for a instantaneous channel $S = s$ be denoted by $p_{n,d}(s)$, then it can be upper bounded as $p_{n,d}(s) \leq \hat{\epsilon}_d(a_n(s))$, where the expression for $\hat{\epsilon}_d(a_n(s))$ is provided in equation (2.20b).

As stated earlier, in the above formulation, d represents the distance from the nearest lower layer which appears as interference. For instance, for the decoding scheme of (A.3), $d = 0$, because all the layers below the n-th layer apprears as interference to it.

The cross over probability $p_{n,d}(s)$, and therefore the rate supportable via a rBSC, denoted by $r_{n,d}(s)$, is a function of the instantaneous channel state, $S = s$. Consequently, the average rate - averaged over all channel states - achievable through the n-th layer of a real fading channel via coding across time is given as

$$\begin{aligned} \mathbb{E}_S[r_{n,d}(s)] &= \mathbb{E}_S[1 - H(p_{n,d}(s))], \\ &\geq \mathbb{E}_S[1 - H(\hat{\epsilon}_d(a_n(s)))] = \mathbb{E}_S[\hat{r}_{n,d}(s)], \end{aligned} \tag{A.4}$$

where $\hat{r}_{n,d}(s)$ is defined as,

$$\hat{r}_{n,d}(s) \triangleq 1 - H(\hat{\epsilon}_d(a_n(s))), \tag{A.5}$$

and $H(p)$ represents the entropy of a $\mathcal{B}(p)$ random variable. Consequently, for any given set of these mutually independent antipodal symbols such as $\{x_{r,n} : n \in \phi\}$, where ϕ is an arbitrary subset of $\mathbb{N}$, can achieve a overall rate which is greater than or equal to

$$\sum_{n \in \phi} \mathbb{E}_S[\hat{r}_{n,d}(s)]. \tag{A.6}$$

It can be concluded from equation (2.20) that the upper bound to the crossover probability $\epsilon_d(a)$ is a decreasing function of d, which in turn implies that $\hat{r}_{n,d}(s)$ is an increasing function of d. Therefore, better rates can be achieved if the value of d can be made larger, which indeed can be done by a modified version of the BES scheme called the BES with *reverse stripping* (BES-RS) [56]. In the BES-RS scheme, symbols are decoded starting from the deepest layer and before decoding a higher layer symbol, contribution from lower-layer symbols are stripped off. As a result, d may

have non-zero values leading to a smaller crossover probability which in turn results in a better average rate that can be achieved through each layer in the presence of lower layer interference. In particular, if all the symbols are intended for the same receiver, all the lower layer symbols can be stripped off before decoding each layer, resulting in $d = \infty$.

Since a complex channel like those in (2.3), can be visualized as a pair of two real channels the rate achievable via layer n of such a channel will be twice of what is shown in equation (A.6). This is summarized in Lemma 1, which follows from Lemma 5 and Theorem 4 of [56].