The Capacity of a Non-Coherent channel - Multi-user setting

by

Paul Njoroge

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

January 31, 2002

Author ______________________
Department of Electrical Engineering and Computer Science
January 15, 2002

Certified by ______________________
vanid Tarokh
Thesis Supervisor

Accepted by ______________________
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

Acknowledgments

I would like to thank everyone who helped me write this thesis. In Particular, Professor Vahid Tarokh for his invaluable support and guidance. His counsel and deep insight into issues was very enlightening and inspiring.

I would like to thank my friends, Nii Lartey Dodoo, Danielle Hinton, Timothy Murithi, Carol Fredricks and Kate Baty who offered support and encouragement during my research.

I would like to thank my family which has always been there for me and supported me in many ways during my study. In particular, my mother who has always believed in me and been a source of strength for me at my weakest.

The LIDs department, its staff have been very supportive during my stay as a graduate student. I would like to thank Vincent Chan and DARPA who funded the research. Finally I would like to thank God for seeing me through with my project.

Contents

The Capacity of a Non-Coherent channel - Multi-user setting

by

Paul K. Njoroge

Submitted to the Department of Electrical Engineering and Computer Science

January 31st, 2002

In Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Electrical Engineering and Computer Science

Abstract

The capacity of the non-coherent channel in a two-user-multi-access channel is investigated when a block length of N is used. The work here builds on the already investigated case of a single user. The lower and upper limit for the capacity region in this non-coherent multi access channel is established. In particular, it is shown that the capacity achieving distribution of the signal resulting from the two users is characterized by zero mean and uncorrelated components. It is also shown that a Gaussian distribution for each of the two users does not achieve capacity. As the block length N increases the capacity region approaches that of a coherent channel and that a Gaussian distribution for each of the two users with I.I.D components achieves capacity. Finally, the bounds of the channel cpacity region in a fading environment are discussed. Direction in which follow up work can proceed is briefly mentioned.

Thesis Supervisor: Vahid Tarokh
Title: Professor

1 Introduction

Phase noise in many communication scenarios makes it difficult to discern the phase of a received signal and consequently the whole signal. Research in recent years has focused on novel non-coherent methods which work with unknown phases. The performance of these methods, in many instances, approaches those of coherent communication systems.

In band-pass transmissions coherent phase references are not usually available at the receiver and thus its difficult to obtain the correct phases and information sequence. Non-coherent channels are a good model for band-pass transmission channels. A non-coherent channel in this respect refers to an Additive White Gaussian Noise channel (AWGN) with a random phase rotation.

There are a number of approaches that are used to detect an information sequence transmitted over an AWGN non-coherent channel. One of them employs the use of pseudo-coherent receivers. This is an approximate coherent detection scheme that uses phase synchronization. A phase reference is extracted from the incoming signal sequence and the sequence is passed through a detection scheme that is optimal under coherent detection. Over the years research has shown that the use of the pseudo-coherent detection technique approaches that of an ideal coherent detector when the phase rotation introduced by the channel is constant or slowly varying and the length of the block code is sufficiently large.

A second approach involves the use of non-coherent decoding scheme. These schemes employ two main classes of algorithms. These are the Multiple Symbol Differentiation Detection(MSSD) [1] and the Non-coherent sequence detection (NSD)[2][3]. In these cases the performances are also shown to approach those of ideal coherent detection. Analysis done in case of M-ary phase-shift keying [4] is one of the cases.

In order to investigate the performance of these non-coherent schemes it is important to establish the capacity of a general AWGN non-coherent channel. The capacity of AWGN non-coherent channel has been calculated in the case of input symbols belonging to an M-PSK alphabet[5]. Capacity of the unknown phase channel but employing an orthogonal multiple frequency-shift keying (MFSK) modulation was derived by Butman, Bar-David, Levitt, Lyon and Klass [6]. Capacity of the non-coherent channel was further investigated for spread spectrum systems by Chayat[7] with orthogonal modulation and by Cheun and Stark [8] in relation to multiple access. Zhou, Mei, Xu, and Yao[9] calculated the capacity of a block-wise fading channel whilst Chen and Fuja[10] calculated the capacity a non-coherent differential MPSK (DMPSK) with block length of two.

This thesis discusses the capacity of the non-coherent AWGN in a multi-access setting. In particular, the capacity in a two user setting is investigated. The work done by C. Colavolpe and Ricardo Raheli [11] investigates, in a single user setting, the capacity of a random-phase non-coherent additive white Gaussian noise channel. In their paper, they found that the capacity achieving distribution is non-Gaussian. They also found that for increasing values of the number of symbols transmitted , the capacity asymptotically approaches that of a coherent channel. This two user case investigation is important as it helps develop a bench mark for evaluating multiuser detection schemes in non-coherent AWGN channels.

The investigation in this thesis considers two users transmitting independently. There are no constraints on both input symbols to the users. The situation investigated is meant to mimic the real case situation where signals from two users have random phase noise added to them in addition to the Gaussian noise imposed by the channel. This phase noise can be induced in several ways. In one instance its added during transmission by the RF converters. In other instances, it is added by the digital modem during the generation of reference signals. These reference signals involve the use of primary and secondary reference oscillators which add the phase noise[12]. Yet another cause of the phase noise are signal amplitude fluctuations during amplitude modulation[13].

Results show that in general the distribution which maximizes the average mutual information is not composed of independent and identically distributed zero-mean bivariate Gaussian components as in the coherent case. An asymptotically tight lower bound for the two user capacity region is derived. The bound shows that the capacity bound of the non-coherent AWGN channel approaches that of a coherent channel as $N \to \infty$. This implies that the capacity achieving distribution as $N \to \infty$, for each user, is a zero-mean Gaussian with independent, identically distributed components.

The following sections discuss the system model employed. The capacity achieving input is characterized and a lower bound for the boundary of the two user region is derived. As the block length goes to infinity the asymptotic behavior of the model is analyzed. Finally the lower bounds of the boundary of the capacity region are investigated when fading is employed in the model. A discussion of what the model implies together with future work to extend the model to more than two users follows.

2 The Channel Model

The system model consists of two users transmitting together in a non-coherent manner. The input to the channel is modeled as two vectors $\mathbf{x} = (x_1, x_2, ..., x_N)^T$ and $\mathbf{z} = (z_1, z_2, ..., z_N)^T$. Each vector represents an input of N complex symbols and these two vectors are assumed to be independent. The output is a vector $\mathbf{y} = (y_1, y_2, ..., y_N)^T$ whose components are expressed as $y_k = x_k e^{j\theta_x} + z_k e^{j\theta_z} + w_k$. θ_x and θ_z are the phase shifts introduced by the channel. They represent the phase noise discussed in the introduction. The phase shifts are considered constant over a block of length N and for each user they are modeled as a continuous random variable with uniform distribution. The noise elements w_k are independent and identically distributed (i.i.d.), zero-mean, independent complex Gaussian random variables whose components have a variance of σ^2 for the real dimension. These random variables are collected in a vector $\mathbf{w} = (w_1, w_2, ..., w_N)^T$.

The components of vector $\mathbf{x}$ and $\mathbf{z}$ may also be expressed in polar coordinates as follows, $x_i = r_{xi} e^{j\phi_x}$ and $z_i = r_{zi} e^{j\phi_z}$. Therefore the vectors $\mathbf{r_x} = (r_{x1}, r_{x2}, ..., r_{xN})^T$ and $\phi_\mathbf{x} = (\phi_{x1}, \phi_{x2}, ..., \phi_{xN})^T$ represent the vectors of the radii and the angles of the components of $\mathbf{x}$. In the same manner $\mathbf{r_z} = (r_{z1}, r_{z2}, ..., r_{zN})^T$ and $\phi_\mathbf{z} = (\phi_{z1}, \phi_{z2}, ..., \phi_{zN})^T$ represent the vectors of the radii and the angles of the components of $\mathbf{z}$.

The signal-to-noise ratio(SNR) per information symbol for each user is defined as follows,

$$\gamma_x = \frac{E||\mathbf{x}||^2}{E||\mathbf{w}||^2} = \frac{E||\mathbf{x}||^2}{2N\sigma^2}, \tag{1}$$

$$\gamma_z = \frac{E||\mathbf{z}||^2}{E||\mathbf{w}||^2} = \frac{E||\mathbf{z}||^2}{2N\sigma^2}. \tag{2}$$

Given a variable v, $||\mathbf{v}||$ stands for the Euclidean norm of $\mathbf{v}$.

3 Properties of the Capacity Achieving Distribution

This section defines channel capacity (average mutual information) and then discusses the properties of the capacity achieving distribution. It is shown that the average mutual information is maximized when the random vectors $\phi_{\mathbf{x}}$ and $\phi_{\mathbf{z}}$ are independent of $\mathbf{r}_{\mathbf{x}}$ and $\mathbf{r}_{\mathbf{z}}$ respectively and composed of independent uniformly distributed (IUD) components. Since the components of vectors $\mathbf{x}$ and $\mathbf{z}$, are independent the components of the vectors formed from their polar coordinates are also independent.

3.1 Capacity

The information channel capacity , also known as the average mutual information (AMI), is defined as follows

$$I_{\mathbf{nc}} = I(\mathbf{x}, \mathbf{z}; \mathbf{y}) = E\left\{\log_2 \frac{p(\mathbf{y}|\mathbf{x}, \mathbf{z})}{p(\mathbf{y})}\right\}. \tag{3}$$

The capacity-achieving distribution of the vectors $\mathbf{x}$ and $\mathbf{z}$ is characterized by the probability density function $p(\mathbf{x},\mathbf{z})$ which maximizes (3). The random variables $\mathbf{x}$ and $\mathbf{z}$ are independent. We express this distribution as follows,

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x} + \mathbf{z}). \tag{4}$$

The conditional probability density function $p(\mathbf{y}|\mathbf{xz})$ is expressed as follows

$$p(\mathbf{y}|\mathbf{x},\mathbf{z}) = \frac{1}{(2\pi\sigma^2)^N} exp\left\{-\frac{\|\mathbf{x}+\mathbf{z}\|^2}{2\sigma^2} - \frac{\|\mathbf{y}\|^2}{2\sigma^2}\right\} \mathbf{I_o}\left(\frac{|\mathbf{y^T}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right). \tag{5}$$

$\mathbf{I_o}$ is the 0th order modified Bessel function of the first kind. In the case of two users transmitting with the same power the capacity per channel use of the non-coherent channel under consideration is defined as follows,

$$C_{nc} = \frac{max\{I_{nc}\}}{N}. \tag{6}$$

Note that this capacity is the capacity attainable by a single transmitter sending with a power equal to the sum of the powers of user one and two. For comparison purposes we define the capacity per channel use of a coherent channel as $\mathbf{C_c}$. This refers to the case when both vectors $\theta_{\mathbf{x}}$ and θ_z are known.

3.2 Maximum Average Mutual Information

The AMI is maximum when the the random vectors $\phi_{\mathbf{x}}$ and $\phi_{\mathbf{z}}$ are independent of $\mathbf{r_x}$ and $\mathbf{r_z}$ respectively. The phase vectors are also characterized by uniformly distributed components. This is derived in the following section.

Consider a virtual channel obtained from our original channel, having the phase vectors $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{z}}$ as inputs together with $\mathbf{x}$ and $\mathbf{z}$. Let $\mathbf{y}$ be the output. The AMI of this channel can be represented as

$$\begin{aligned}
\mathbf{I_v} &= I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}, \mathbf{x}, \mathbf{z}; \mathbf{y}) && (7) \\
&= I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}, \phi_{\mathbf{x}}, \phi_{\mathbf{z}}, \mathbf{r_x}, \mathbf{r_z}; \mathbf{y}) && (8) \\
&= I(\zeta_{\mathbf{x}}, \zeta_{\mathbf{z}}, \mathbf{r_x}, \mathbf{r_z}; \mathbf{y}). && (9)
\end{aligned}$$

In the above formulae $\zeta_{\mathbf{x}} = (\zeta_{x1}, \zeta_{x2}, ..., \zeta_{xN})^T = (\phi_{x1} + \theta_x, \phi_{x2} + \theta_x, ..., \phi_{xN} + \theta_x)$. And in a similar manner $\zeta_{\mathbf{z}} = (\zeta_{z1}, \zeta_{z2}, ..., \zeta_{zN})^T = (\phi_{z1} + \theta_z, \phi_{z2} + \theta_z, ..., \phi_{zN} + \theta_z)$. The distribution of $\zeta_{\mathbf{x}}$ and $\zeta_{\mathbf{z}}$ are determined by those of θ_x and θ_z respectively as these are among the inputs to the virtual channel and are therefore known.

Using the standard mutual information decomposition

$$I(a; b, c) = I(a; b) + I(a; c|b) \tag{10}$$

We can relate the AMI of the AWGN non-coherent channel to that of the virtual channel. This relation yields the following set of equations,

$$\begin{aligned}
I_{nc} &= I_v - I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z}) && (11) \\
&= I_v - I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{r_x}, \mathbf{r_z}, \phi_{\mathbf{x}}, \phi_{\mathbf{z}}). && (12)
\end{aligned}$$

Equation (12) describes the AMI of the diversity channel in which θ_x and θ_z are input, $\mathbf{y}$ is the output and the receiver has the channels state information represented by $\mathbf{x}$ and $\mathbf{z}$. It follows that the AMI would depend on the joint distribution of $(\mathbf{r_x}, \mathbf{r_z}, \phi_{\mathbf{z}}, \phi_{\mathbf{x}})$. Using Bayes rule this translates to the factors $p(\mathbf{r_x}, \mathbf{r_z})$ and $p(\phi_\mathbf{x}, \phi_\mathbf{z}|\mathbf{r_x}, \mathbf{r_z})$. Using the independence of $\mathbf{x}$ and $\mathbf{z}$ we can rewrite the dependence as collection of the four factors $p(\mathbf{r_x}), p(\mathbf{r_z}), p(\phi_\mathbf{x}|\mathbf{r_x}), p(\phi_\mathbf{z}|\mathbf{r_z})$.

Further analysis shows that the AMI is independent of $p(\phi_\mathbf{x}|\mathbf{r_x})$ and $p(\phi_\mathbf{z}|\mathbf{r_z})$. This follows because the phases ϕ_z, ϕ_x are perfectly known to the receiver. It's thus possible to transform vector $\mathbf{y}$ using the diagonal unitary matrix *diag* $(e^{-j\phi_{x1}}, e^{-j\phi_{x2}}, ..., e^{-j\phi_{xN}})$ *or* diag $(e^{-j\phi_{z1}}, e^{-j\phi_{z2}}, ..., e^{-j\phi_{zN}})$. Both transformations are reversible and they therefore do not modify the AMI[14]. They also do not change the statistics of the noise vector. Therefore the AMI of the diversity channel represented as $I(\theta_\mathbf{x}, \theta_\mathbf{z}; \mathbf{y}|\mathbf{r_x}, \mathbf{r_z}, \phi_\mathbf{x}, \phi_\mathbf{z})$ is independent of $p(\phi_\mathbf{x}|\mathbf{r_x})$ and $p(\phi_\mathbf{z}|\mathbf{r_z})$.

The AMI of the virtual channel is represented as in [15],

$$\begin{aligned} \mathbf{I_v} &= I(\zeta_\mathbf{x}, \zeta_\mathbf{z}, \mathbf{r_x}, \mathbf{r_z}; \mathbf{y}) && (13) \\ &= H(\mathbf{y}) - H(\mathbf{y}|\zeta_\mathbf{x}, \zeta_\mathbf{z}, \mathbf{r_x}, \mathbf{r_z}). && (14) \end{aligned}$$

$H(\mathbf{y})$ and $H(\mathbf{y}|\zeta_\mathbf{x}, \zeta_\mathbf{z}, \mathbf{r_x}, \mathbf{r_z})$ are the entropy and conditional entropy of $\mathbf{y}$ respectively. The conditional entropy is the same as,

$$H(\mathbf{w}) = E\left\{\log_2 \frac{1}{p(\mathbf{w})}\right\} \tag{15}$$

where $p(w)$ is the complex Gaussian distribution of noise given as,

$$p(\mathbf{w}) = \frac{1}{(2\pi\sigma^2)^N} exp\left\{\frac{-||\mathbf{w}||^2}{2\sigma^2}\right\}. \tag{16}$$

Therefore as in [16],

$$H(\mathbf{y}|\zeta_{\mathbf{x}}, \zeta_{\mathbf{z}}, \mathbf{r_x}, \mathbf{r_z}) = N \log_2 \pi e \sigma^2, \tag{17}$$

independently of the joint distribution $\phi_{\mathbf{x}}, \phi_{\mathbf{z}}, \mathbf{r_x}, \mathbf{r_z}$.

Changing to polar coordinates we can express y_i as $v_i e^{\psi}$. Using this new coordinates we get the vectors $\mathbf{v} = (v_1, v_2, ..., v_N), \psi = (\psi_1, \psi_2, ..., \psi_N)$.

$$\begin{aligned} H(\mathbf{y}) &= H(\mathbf{v}, \psi) + \int_0^{+\infty} ... \int_0^{+\infty} p(v_1, v_2, ..., v_N) \log_2 \prod_{i=1}^{n} v_i dv_1 ... dv_N & (18) \\ &= H(\mathbf{v}, \psi) + \sum_{i=1}^{N} \int_0^{+\infty} p(v_i) \log_2 v_i dv_i & (19) \end{aligned}$$

The above equation is the multidimensional form of (33) in [16]. $H(\mathbf{v}, \psi) \leq H(\mathbf{v}) + H(\psi)$, this is shown in [14]. The two sides are equal if and only if $\mathbf{v}$ is independent of ψ.

Taking the equations (11),(14),(17) and the independence of $\mathbf{v}$and ψ we get,

$$I_{nc} \leq H(\mathbf{v}) + H(\psi) + \sum_{i=1}^{N} \int_0^{+\infty} p(v_i) \log_2 v_i dv_i - N \log_2 \pi e \sigma^2 - I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{r_x}, \mathbf{r_z}, \phi_{\mathbf{x}}, \phi_{\mathbf{z}}). \tag{20}$$

The first and third terms of (20) are independent of the conditional distribution$p(\rho|\mathbf{r_x}, \mathbf{r_z}, \zeta_{\mathbf{x}} - \zeta_{\mathbf{z}})$. This follows from the independence of $p(\mathbf{v})$ from $p(\rho|\mathbf{r_x}, \mathbf{r_z}, \zeta_{\mathbf{x}} - \zeta_{\mathbf{z}})$. ρ is the exponent in the polar coordinate version of the following new variable $\mathbf{r}$.

$$\begin{aligned} \mathbf{r} &= \mathbf{x} \mathbf{e}^{\mathbf{j}\theta_{\mathbf{x}}} + \mathbf{z} \mathbf{e}^{\mathbf{j}\theta_{\mathbf{z}}} \\ \mathbf{r} \mathbf{e}^{\mathbf{j}\rho} &= \mathbf{r_x} \mathbf{e}^{\mathbf{j}\zeta_{\mathbf{x}}} + \mathbf{r_z} \mathbf{e}^{\mathbf{j}\zeta_{\mathbf{z}}} & (21) \end{aligned}$$

Note that we can represent r_i as a square root of the following function,

$$f(r_{xi}, r_{zi}, \zeta_{xi} - \zeta_{zi}) = r_{xi}^2 + r_{zi}^2 - 2r_{xi}r_{zi}cos(2\pi - \zeta_{xi} + \zeta_{zi}). \tag{22}$$

In other words $r_i^2 = f(r_{xi}, r_{zi}, \zeta_{xi} - \zeta_{zi})$.

Equation (20) achieves equality when $\mathbf{v}$ and ψ are independent and if their components are independent and uniformly distributed. $H(\psi)$ and $H(\mathbf{v})$ are maximized when they have independent, uniformly distributed components as shown in [15]. In the next two sections it's shown that ψ has independent uniformly distributed components if ρ and hence $\zeta_{\mathbf{x}}$ and $\zeta_{\mathbf{z}}$ have independent uniformly distributed components independent of the distribution of $\mathbf{r}$ i.e $(r_x, r_z, \zeta_x - \zeta_z)$. Note that this implies that $\phi_{\mathbf{x}}$ and $\phi_{\mathbf{z}}$ have independent uniformly distributed components independent of distribution of $\mathbf{r}$.

In the next three sections we prove that I_{nc} is maximum when vectors $\mathbf{r}$ and ρ are independent and ρ has I.U.D components. ρ having I.U.D. components implies that ζ_x and thus ϕ_x has I.U.D. components too. The same applies for ζ_z and ϕ_z - both have I.U.D. components too.

3.2.1 Probability Distributions in Polar coordinates

Before we proceed on to prove the independence of $p(\mathbf{v})$ from $p(\rho|\mathbf{r_x}, \mathbf{r_z}, \zeta_{\mathbf{x}} - \zeta_{\mathbf{z}})$ and that vector ρ has independent uniformly distributed components, independently of the distribution of $\mathbf{r}$, we express the Gaussian probability function $p(\mathbf{y}|\mathbf{x}, \mathbf{z}, \theta_{\mathbf{x}}, \theta_{\mathbf{z}})$ in polar coordinates.

$$p(\mathbf{v}, \psi|\mathbf{r_x}, \mathbf{r_z}, \zeta_{\mathbf{x}} - \zeta_{\mathbf{z}}, \rho) = \prod_{i=1}^{N} \frac{v_i}{2\pi\sigma^2} exp\left\{-\frac{v_i^2 + r_i^2 - 2v_i r_i cos(\psi_i - \rho_i)}{2\sigma^2}\right\} \tag{23}$$

$\mathbf{r}$ is defined in (21) and it's components values are defined as the square root of (22). ρ on the other hand can be defined in reference to ζ_z or ζ_x as follows,

$$
\begin{aligned}
\varpi_i &= cos^{-1}\left(-\frac{r_{zi}^2 - r_{xi}^2 - r_i^2}{2r_{xi}}\right) \\
\rho_i &= \zeta_{xi} - \varpi_i \qquad (24) \\
\varphi_i &= cos^{-1}\left(-\frac{r_{xi}^2 - r_{zi}^2 - r_i^2}{2r_{zi}}\right) \\
\rho_i &= \zeta_{zi} + \varphi_i \qquad (25)
\end{aligned}
$$

With the above definitions the marginal probability density functions are described as follows,

$$
\begin{aligned}
p(\mathbf{v}|\mathbf{r_x}, \mathbf{r_z}, \zeta_\mathbf{x} - \zeta_\mathbf{z}, \rho) &= p(\mathbf{v}|\mathbf{r_x}, \mathbf{r_z}, \zeta_\mathbf{x} - \zeta_\mathbf{z}) \\
&= p(\mathbf{v}|\mathbf{r}) \\
&= \prod_{i=1}^{N} \frac{v_i}{2\pi\sigma^2} exp\left\{-\frac{v_i^2 + r_i^2}{2\sigma^2}\right\} I_o\left(\frac{v_i r_i}{\sigma^2}\right) \qquad (26) \\
p(\psi|\mathbf{r_x}, \mathbf{r_z}, \zeta_\mathbf{x} - \zeta_\mathbf{z}, \rho) &= p(\psi|\mathbf{r}, \rho) \\
&= \prod_{i=1}^{N} \left\{\frac{1}{2\pi} exp\left[-\frac{r_i^2}{2\sigma^2}\right] + \frac{r_i cos(\psi_i - \rho_i)}{\sqrt{2\pi}\sigma} exp\left[-\frac{r_i^2 sin^2(\psi_i - \rho_i)}{2\sigma^2}\right]\right\} \\
&\quad .Q\left\{\left(-\frac{r_i cos(\psi_i - \rho_i)}{\sigma}\right)\right\} \qquad (27)
\end{aligned}
$$

Here $Q(x)$ is the Gaussian Q function defined as follows,

$$
Q(x) = \frac{1}{\sqrt{2\pi}} \int_r^{+\infty} exp\left\{-\frac{t^2}{2}\right\} dt \qquad (28)
$$

3.2.2 Independence of $p(\mathbf{v})$ from $p(\rho|\mathbf{r_x}, \mathbf{r_z}, \zeta_{\mathbf{x}} - \zeta_{\mathbf{z}}) = p(\rho|\mathbf{r})$

In this section, it is shown that the probability density function p(v) is independent of the conditional distribution $p(\rho|\mathbf{r})$. We start by noting that ρ varies in $[0, 2\pi]$ and that

$$p(\mathbf{v}|\mathbf{r}) = \int_{\rho^N} p(\mathbf{v}|\mathbf{r}, \rho)p(\rho|\mathbf{r})d\rho \tag{29}$$

In general $p(\mathbf{v}|\mathbf{r})$ depends on $p(\rho|\mathbf{r})$. However,(26) shows that $p(\mathbf{v}|\mathbf{r})$ is independent of $p(\rho|\mathbf{r})$. As a direct result, it is independent of $p(\zeta_x|\mathbf{r})$ and $p(\zeta_z|\mathbf{r})$. This can be observed by noting the alternative definitions of ρ_i in (24) and (25). As a consequence $p(\mathbf{v}|\mathbf{r})$ is also independent of $p(\phi_z|\mathbf{r})$ and $p(\phi_x|\mathbf{r})$. This property is true also for $p(\mathbf{v})$, as it's obtained by averaging with respect to $\mathbf{r}$.

3.2.3 ψ has I.U.D Components and is Independent of r if ρ has I.U.D. elements

From (28) , we get

$$p(\psi|\mathbf{r}) = \int_{\rho^N} p(\psi|\mathbf{r}, \rho)p(\rho|\mathbf{r})d\rho \tag{30}$$

ρ is independent of $\mathbf{r}$ therefore $p(\rho|\mathbf{r})$ is equal to $p(\rho)$. If we assume that ρ has independent components that are uniformly distributed we get the following,

$$\begin{aligned} p(\rho|\mathbf{r}) &= \prod_{i=1}^{N} p(\rho_i) \\ &= (\frac{1}{2\pi})^N \\ &= p(\rho) \end{aligned} \tag{31}$$

ρ has range $[0 - 2\pi]$. The dependence on $\psi_i - \rho_i$ of the integrand in (30) means that the integrals with respect to ρ and ψ are equal. As shown in [11] we obtain ,

$$\begin{aligned} p(\psi|\mathbf{r}) &= \prod_{i=1}^{N} p(\psi_i) \\ &= (\frac{1}{2\pi})^N \\ &= p(\psi). \end{aligned} \tag{32}$$

The range of ψ is the same as that of ρ. From (23) and (26) it can be deduced that if $p(\rho) = p(\psi) = (\frac{1}{2\pi})^N$ then $p(\mathbf{v}, \psi|\mathbf{r}) = p(\mathbf{v}|\mathbf{r})p(\psi)$. It then follows that ,

$$p(\mathbf{v}, \psi) = \int_{\mathbf{r^N}} p(\mathbf{v}|\mathbf{r})p(\psi)d\mathbf{r}. \tag{33}$$

Thus from (33) we ascertain that $p(\mathbf{v}, \psi) = p(\mathbf{v})p(\psi)$ and therefore the independence of $\mathbf{v}$ and ψ.

3.2.4 Capacity achieving distribution has zero mean uncorrelated components

From the two previous sections, the independence of $\mathbf{r}$ and ρ is shown. It follows from above that both ϕ_x and ϕ_z are independent of $\mathbf{r_x}$ and $\mathbf{r_z}$ respectively. As a consequence the components of both $\mathbf{x}$ and $\mathbf{z}$ exhibit circular symmetry. From these facts, it follows that the capacity-achieving distribution of the joint distribution $\mathbf{x} + \mathbf{z}$ is characterized by zero mean and uncorrelated components as shown below,

$$E\{\mathbf{x} + \mathbf{z}\} = \int\int p(\mathbf{x}, \mathbf{z})(\mathbf{x} + \mathbf{z})d\mathbf{x}d\mathbf{z}$$

$$
\begin{aligned}
&= \int\int p(\mathbf{x},\mathbf{z})p(\mathbf{x}) + p(\mathbf{x},\mathbf{z})p(\mathbf{z})d\mathbf{x}d\mathbf{z} \\
&= \int \mathbf{x}p(\mathbf{x})d\mathbf{x} + \int \mathbf{z}p(\mathbf{z})d\mathbf{z} \qquad (34)
\end{aligned}
$$

We note that the above equations consists of the summation of the mean of $\mathbf{x}$ and $\mathbf{z}$ and both have 0 means. Therefore, (34) integrates to zero. The components are uncorrelated by the nature of transmission. Transmissions of one symbol to the next are independent.

3.2.5 The capacity achieving distribution is not Gaussian

In this section, it is shown that for a transmission of N symbols by each user the capacity achieving distribution is not Gaussian. We first establish the capacity achieving distribution under a constant power constraint. Another constraint that is implicit is the fact that the integral of the probability distribution of $p(\mathbf{x},\mathbf{z})$ with respect to both $\mathbf{z}$ and $\mathbf{x}$ is one. By using Lagrange multipliers and variational techniques we establish the conditions the capacity achieving distribution must satisfy. We then proceed to show that a Gaussian distribution does not satisfy this conditions.

We maximize $p(\mathbf{x},\mathbf{z})$ under the following constraints,

$$\int\int ||\mathbf{x}+\mathbf{z}||^2 p(\mathbf{x},\mathbf{z})d\mathbf{x}d\mathbf{z} = C \qquad (35)$$

$$\int\int p(\mathbf{x},\mathbf{z})d\mathbf{x}d\mathbf{z} = 1. \qquad (36)$$

When maximizing $\mathbf{I_{nc}}p(\mathbf{x},\mathbf{z})$ we are looking for the distribution $p(\mathbf{x},\mathbf{z})$ that satisfies the following condition,

$$\mathbf{I}'_{\mathbf{nc}}(p(\mathbf{x},\mathbf{z})) + \tau u(p(\mathbf{x},\mathbf{z})) = 0 \qquad (37)$$

Lagrange multipliers are used to maximize the above expression and come with the following expression, which is the same as (14) in [11],

$$\int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 p(\mathbf{y}) d(\mathbf{y}) = \lambda - \log_2 e + \mu ||\mathbf{x}+\mathbf{z}||^2 + \int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 p(\mathbf{y}|\mathbf{x},\mathbf{z}) d(\mathbf{y}). \quad (38)$$

Thus any capacity achieving distribution of $(\mathbf{x}+\mathbf{z})$ must satisfy the above equation. Using (5) , we get

$$\begin{aligned} \int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 p(\mathbf{y}|\mathbf{x},\mathbf{z}) d\mathbf{y} &= -N \log_2 2\pi\sigma^2 - \log_2 e \frac{||\mathbf{x}+\mathbf{z}||^2}{2\sigma^2} \\ &\quad - \frac{\log_2 e}{2\sigma^2}(2N\sigma^2 + ||\mathbf{x}+\mathbf{z}||^2) \\ &\quad + \int p(\mathbf{y}|\mathbf{x}+\mathbf{z}) \log_2 I_o(|\mathbf{y}^{\mathbf{T}}(\mathbf{x}+\mathbf{z})^*|) d\mathbf{y}. \quad (39) \end{aligned}$$

The above has been reached by using the following relation,

$$\begin{aligned} \int ||\mathbf{y}||^2 p(\mathbf{y}|\mathbf{x},\mathbf{z}) d\mathbf{y} &= \frac{1}{2\pi} \int_0^{2\pi} \int \int ||\mathbf{y}^2|| p(\mathbf{y}|\mathbf{x},\mathbf{z},\theta_{\mathbf{x}},\theta_{\mathbf{z}}) d\mathbf{y} d\theta_{\mathbf{x}} d\theta_{\mathbf{z}} \\ &= 2N\sigma^2 + ||\mathbf{x}+\mathbf{z}||^2 \quad (40) \end{aligned}$$

(38) can then be represented as follows ,

$$\int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 p(\mathbf{y}) d(\mathbf{y}) = \lambda' + \mu' ||\mathbf{x}+\mathbf{z}||^2 + \int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 \mathbf{I_o}(|\mathbf{y}^{\mathbf{T}}(\mathbf{x}+\mathbf{z})^*|). \quad (41)$$

The following substitutions have been made,

$$\lambda' = \lambda - \log_2 e - N \log_2 2\pi e \sigma^2 \quad (42)$$

$$\mu' = \mu - \frac{\log_2 e}{\sigma^2}. \tag{43}$$

We now proceed to show that $\mathbf{x}+\mathbf{z}$ having a Gaussian distribution cannot achieve capacity because it doesn't satisfy (38). If $\mathbf{x}+\mathbf{z}$ were a Gaussian distribution, it would have independent, uncorrelated components and thus satisfied the condition discussed in 3.2.4.. We can think of the random variable $\mathbf{x}+\mathbf{z}$ being made by the summation of two multi-dimensional Gaussian random variables $\mathbf{x}$ and $\mathbf{z}$. It's probability density function would be

,

$$p(\mathbf{x},\mathbf{z}) = \frac{1}{(2\pi(\sigma_{\mathbf{x}}^{\mathbf{2}}+\sigma_{\mathbf{z}}^{\mathbf{2}})^{\mathbf{N}})} exp\left\{-\frac{||\mathbf{x}+\mathbf{z}||^{\mathbf{2}}}{2(\sigma_{\mathbf{x}}^{\mathbf{2}}+\sigma_{\mathbf{z}}^{\mathbf{2}})}\right\} \tag{44}$$

It follows that the distribution of $\mathbf{y}$ would be Gaussian with independent components.

$$p(\mathbf{y}) = \frac{1}{(2\pi\sigma_{\mathbf{y}})^{\mathbf{N}}} exp\left\{-\frac{||\mathbf{y}||^{\mathbf{2}}}{2\sigma_{\mathbf{y}}^{\mathbf{2}}}\right\} \tag{45}$$

Above $\sigma_y = \sigma_x + \sigma_z + \sigma$. Using (44) and (45) we can represent the left side of (41) as follows,

$$\begin{aligned} \int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 p(\mathbf{y}) d(\mathbf{y}) &= -N \log_2(2\pi\sigma_y^2) - \frac{\log_2 e}{2\sigma_y^2} \int ||\mathbf{y}||^2 p(\mathbf{y}|\mathbf{x}+\mathbf{z}) d\mathbf{y} \\ &= -N \log_2(2\pi\sigma_y^2) - \frac{\log_2 e}{2\sigma_y^2}(2N\sigma^2 + ||\mathbf{x}+\mathbf{z}||^2) \end{aligned} \tag{46}$$

(46) exhibits a quadratic dependence on $||\mathbf{x}+\mathbf{z}||$, but the right hand-side of (41) is not a quadratic form due to the presence of $\int p(\mathbf{y}|\mathbf{x},\mathbf{z}) \log_2 \mathbf{I_o}(|\mathbf{y^T}(\mathbf{x}+\mathbf{z})^*|)$. Thus $(\mathbf{x}+\mathbf{y})$ that is capacity achieving is not a Gaussian distribution.

3.3 Lower Bound on Channel Capacity Region

In this section, a lower bound on the capacity region for two users transmitting under the channel model, described in section 2, is derived. The argument follows closely that in [11]. First, a review of the channel capacity region for two users in an additive white Gaussian channel is done. Then, it is shown that the capacity region achievable in the non-coherent channel is larger than the capacity region achievable when a particular non-optimal distribution of $\mathbf{x}$ and $\mathbf{z}$ is used. From section 3.2.5, it was noted that a Gaussian distribution for both users does not achieve capacity for a non-coherent channel. Therefore the Gaussian distribution is non-optimal and is used for both users $\mathbf{x}$ and$\mathbf{z}$ when deriving the lower bound.

3.3.1 Capacity of Gaussian Multiple Access Channels

Assume that two transmitters, $\mathbf{X_1}$ and $\mathbf{X_2}$, communicate to a single receiver $\mathbf{Y}$. The received signal at time i is

$$\mathbf{Y_i} = \mathbf{X_{1i}} + \mathbf{X_{2i}} + \mathbf{Z} \tag{47}$$

Where $\mathbf{Z_i}$ is a sequence of independent, identically distributed , zero mean Gaussian random variables with variance N. From [14], it is observed that the capacity region for a Gaussian multiple access channel with two users is the convex hull of the set of rate of pairs satisfying,

$$\mathbf{R_1} \leq I(X_1; Y|X_2), \tag{48}$$

$$\mathbf{R_2} \leq I(X_2; Y|X_1), \tag{49}$$

$$\mathbf{R_1} + \mathbf{R_2} \leq I(X_1, X_2 : Y) \tag{50}$$

for some input distribution $f(x_1), f(x_2)$ satisfying some power constraint $E\{\mathbf{X_1^2}\} \leq P_1$ and $E\{\mathbf{X_2^2}\} \leq P_2$. Calculating the mutual information for this channel yields the following bounds,(the full calculations of are shown in [14])

$$R_1 \leq C\left(\frac{P_1}{N}\right), \tag{51}$$

$$R_2 \leq C\left(\frac{P_2}{N}\right), \tag{52}$$

$$R_1 + R_2 \leq C\left(\frac{P_1 + P_2}{N}\right) \tag{53}$$

3.3.2 Lower Bound

We will assume that $\mathbf{x} + \mathbf{z}$ is Gaussian and that $\mathbf{x}$ and $\mathbf{z}$ are each Gaussian with independent real and imaginary components, each with a variance σ_x^2 and σ_z^2 respectively. Thus we define for each real component $Re(\mathbf{x_i}) \sim \mathbf{N}(\mathbf{0}, \sigma_{\mathbf{x}}^{\mathbf{2}})$ and $Re(\mathbf{z_i}) \sim \mathbf{N}(\mathbf{0}, \sigma_{\mathbf{z}}^{\mathbf{2}})$ such that $Re(\mathbf{x_i} + \mathbf{z_i}) \sim \mathbf{N}(\mathbf{0}, \sigma_{\mathbf{x}}^{\mathbf{2}} + \sigma_{\mathbf{z}}^{\mathbf{2}})$. In a similar manner we define for each imaginary component $Im(\mathbf{x_i}) \sim \mathbf{N}(\mathbf{0}, \sigma_{\mathbf{x}}^{\mathbf{2}})$ and $Im(\mathbf{z_i}) \sim \mathbf{N}(\mathbf{0}, \sigma_{\mathbf{z}}^{\mathbf{2}})$ such that $Im(\mathbf{x_i} + \mathbf{z_i}) \sim \mathbf{N}(\mathbf{0}, \sigma_{\mathbf{x}}^{\mathbf{2}} + \sigma_{\mathbf{z}}^{\mathbf{2}})$. The probability density function of $\mathbf{x} + \mathbf{z}$ is given by (44) and the signal to noise ratio is,

$$\gamma_{xz} = \frac{\sigma_x^2 + \sigma_z^2}{\sigma^2}. \tag{54}$$

The probability density of $\mathbf{y}$ given the input distribution of $\mathbf{x}$ and $\mathbf{z}$ being Gaussian is given by (45). The average mutual information per channel given the Gaussian inputs $\mathbf{x}$ and $\mathbf{z}$ is as calculated as follows,

$$\begin{aligned}
\frac{I_{nc}}{N} &= \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2\frac{p(\mathbf{y}|\mathbf{x},\mathbf{z})}{p(\mathbf{y})}d\mathbf{x}d\mathbf{z}d\mathbf{y} \\
&= -\frac{1}{N}\int p(\mathbf{y})\log_2 p(\mathbf{y})d\mathbf{y} \\
&\quad +\frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 p(\mathbf{y}|\mathbf{x},\mathbf{z})d\mathbf{x}d\mathbf{z}d\mathbf{y} \\
&= \log_2 2\pi e\sigma_y^2 + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 p(\mathbf{y}|\mathbf{x},\mathbf{z})d\mathbf{x}d\mathbf{z}d\mathbf{y} \qquad (55)
\end{aligned}$$

Using (5) and (40) we get the following set of equations,

$$\begin{aligned}
\frac{I_{nc}}{N} &= \log_2 2\pi e\sigma_y^2 + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2(2\pi\sigma^2)^{-N}d\mathbf{x}d\mathbf{y}d\mathbf{z} \\
&\quad +\frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 exp\left\{-\frac{\|\mathbf{x}+\mathbf{z}\|^2}{2\sigma^2}-\frac{\|\mathbf{y}\|^2}{2\sigma^2}\right\}I_o\left(\frac{|\mathbf{y}^{\mathbf{T}}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z} \\
&= \log_2 2\pi e\sigma_y^2 - \log_2(2\pi\sigma^2) + \frac{1}{N}\int\int\int\left\{-\frac{\|\mathbf{x}+\mathbf{z}\|^2}{2\sigma^2}-\frac{\|\mathbf{y}\|^2}{2\sigma^2}\right\}p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 e d\mathbf{x}d\mathbf{y}d\mathbf{z} \\
&\quad +\frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y}^{\mathbf{T}}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z} \\
&= \log_2 2\pi e\sigma_y^2 - \log_2(2\pi\sigma^2) - \frac{\log_2 e}{2\sigma^2 N}\int\int\frac{\|\mathbf{x}+\mathbf{z}\|^2}{2\sigma^2}p(\mathbf{x},\mathbf{z})d\mathbf{x}d\mathbf{z} \\
&\quad -\frac{\log_2 e}{2\sigma^2 N}\int\int\int\frac{\|\mathbf{y}\|^2}{2\sigma^2}p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})d\mathbf{x}d\mathbf{y}d\mathbf{z} \\
&\quad +\frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y}^{\mathbf{T}}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z} \qquad (56)
\end{aligned}$$

If we define the capacity of a coherent channel $I(\mathbf{x},\mathbf{z};\mathbf{y}) = C_c$ where C_c is defined as follows,

$$C_c = \log_2(1+\gamma_{\mathbf{xz}}) \qquad (57)$$

we can write equation (56) as follows,

$$\begin{aligned}\frac{I_{nc}}{N} &= \log_2(1+\gamma_{\mathbf{xz}}) - 2(\gamma_{\mathbf{xz}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y^T}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z} \\ &= C_c - 2(\gamma_{\mathbf{xz}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y^T}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z} \qquad (58)\end{aligned}$$

Therefore the lower boundary of the channel capacity region of the non-coherent channel for the two user case is given by the convex hull of the set of rate of pairs satisfying,

$$\begin{aligned}R_1 &\leq C_{cx} - 2(\gamma_{\mathbf{x}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x})p(\mathbf{x})\log_2 I_o\left(\frac{|\mathbf{y^T}\mathbf{x}^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y} \qquad (59) \\ R_2 &\leq C_{cz} - 2(\gamma_{\mathbf{z}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{z})p(\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y^T}\mathbf{z}^*|}{\sigma^2}\right)d\mathbf{z}d\mathbf{y} \qquad (60) \\ C_{nc} &\leq C_c - 2(\gamma_{\mathbf{xz}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y^T}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z} \qquad (61)\end{aligned}$$

Equations (59) and (60) have been derived in [14] where C_{cx} and C_{cz} are defined as follows,

$$C_{cx} = \log_2(1+\gamma_{\mathbf{x}}), \qquad (62)$$

$$C_{cz} = \log_2(1+\gamma_{\mathbf{z}}). \qquad (63)$$

$$(64)$$

The integrals in (59) , (60) and (61) are numerically calculated. It turns out that only one calculation that sweeps over a range of SNR is necessary to establish the boundary for a given SNR of the variable given by $\mathbf{x}+\mathbf{z}$. Assuming the noise has constant power evaluating the lower bound capacity for a range of power values for the random variable $\mathbf{x}+\mathbf{z}$ implicitly yields the capacity for both $\mathbf{x}$ and $\mathbf{z}$. This occurs because the sum of their powers is less than or equal to that of $\mathbf{x}+\mathbf{z}$. The integral in (61) is numerically calculated using the results of Appendix B in [14].

Figure 1 shows the proposed lower bound on channel capacity for various values of N. The capacity of the coherent channel is given for comparison purposes. The lower bound curve on the channel capacity region is given in Figure 2 together with the curve for a coherent case. User one and two transmit using the same power and the SNR as seen by $\mathbf{x}+\mathbf{z}$ is given by (54). If we let the $SNR = 5dB$, we obtain the region in Figure 2.

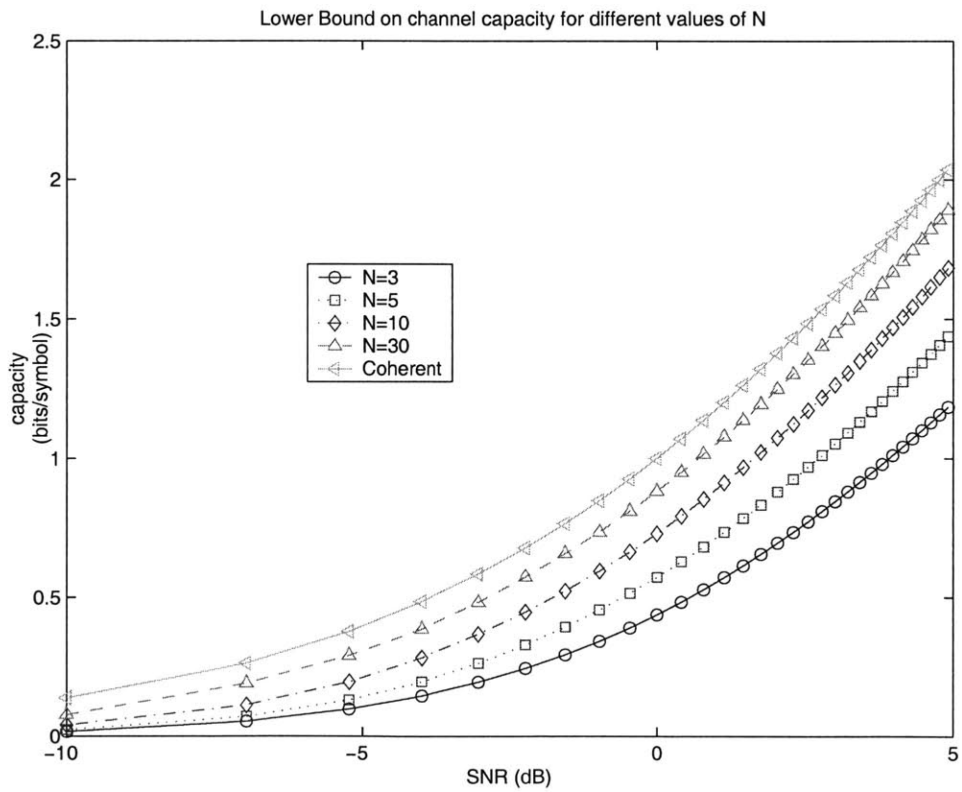

Figure 1: Lower bound on channel capacity for different values of N

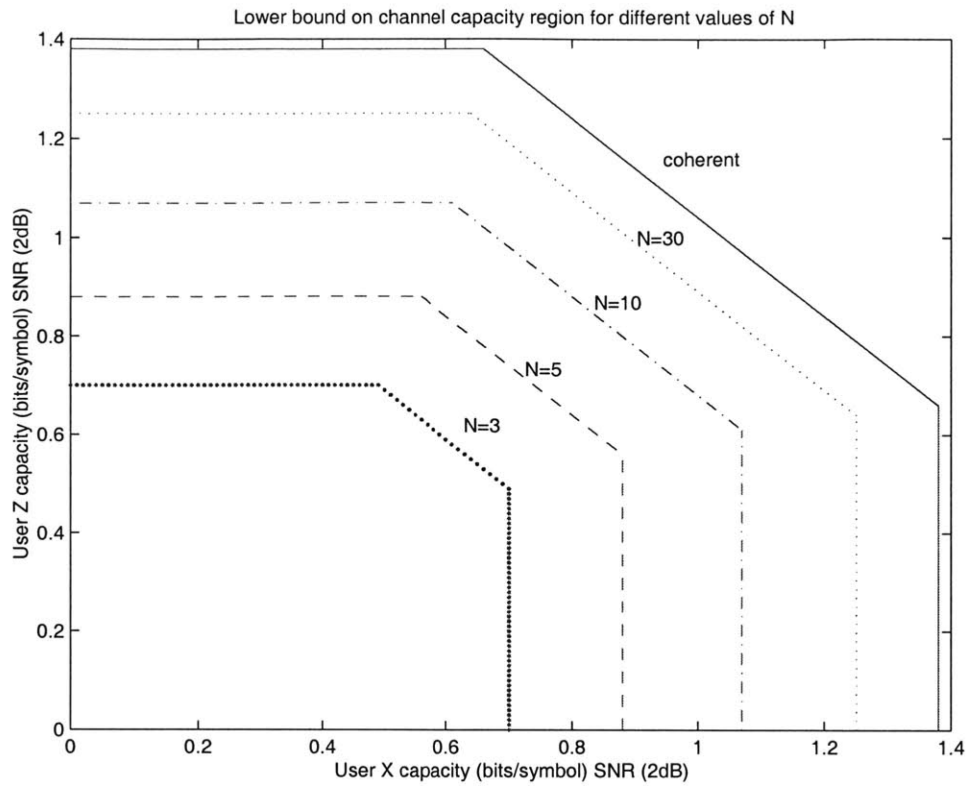

Figure 2: Lower bound on channel capacity region for different values of N

3.4 Upper Bound on Channel Capacity Region

In the previous section, we came up with the lower bound area of the channel capacity region for two users. In this section, we show that as $N \to \infty$ the area of the capacity region approaches that of the coherent channel. We again follow closely the arguments presented in [14] in establishing the asymptotic behavior of this model.

As shown in (7) we can express the average mutual information of the virtual channel as follows,

$$\begin{aligned}
\mathbf{I_v} &= I(\theta_\mathbf{x}, \theta_\mathbf{z}, \mathbf{x}, \mathbf{z}; \mathbf{y}) \\
&= I(\theta_\mathbf{x}, \theta_\mathbf{z}; \mathbf{y}) + \mathbf{I}(\mathbf{x}, \mathbf{z}; \mathbf{y} | \theta_\mathbf{x}, \theta_\mathbf{z})
\end{aligned}$$

$$= \quad I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}) + \mathbf{I_c}. \tag{65}$$

I_c is the capacity of the additive white Gaussian coherent channel where θ_x and θ_z are known. The first term of (36) is the average mutual information of the diversity channel mentioned in Section 3.2 except in this case the channel state information is not known. Substituting (36) into (11) we get,

$$I_{nc} = I_c + I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}) - \mathbf{I}(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z}) \tag{66}$$

Lets consider the Gaussian distribution $p(\mathbf{x}, \mathbf{z})$ in (44). It has been shown in section (3.3) that,

$$\begin{aligned} C_{nc} &\geq \frac{I_{nc}}{N} \\ &= \frac{I_c}{N} + \frac{I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y})}{N} + \frac{I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z})}{N} \\ &= C_c + \frac{I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y})}{N} + \frac{I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z})}{N} \end{aligned} \tag{67}$$

Being that $x + z$ is Gaussian the resultant $\mathbf{y}$ is also Gaussian and thus independent of θ_x θ_z. This implies that $I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}) = \mathbf{0}$. Therefore (67) reduces to

$$C_{nc} \geq C_c - \frac{I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z})}{N}. \tag{68}$$

It was shown in section (3.2.3) that ψ has I.U.D. components if ρ has independent components. This further implied that θ_x and therefore ϕ_x had I.U.D. components. The same argument applied for ϕ_z. This property of the capacity achieving distribution allows us to conclude that $I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}) = \mathbf{0}$ for any distribution that achieves capacity.

It follows therefore that $C_{nc} \leq C_c$. This can be readily observed from equation (66) where the second term is zero as explained above and the third term is positive [14]. Using (68) we end up with the following bound for $\mathbf{C_{nc}}$.

$$C_c - \frac{I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z})}{N} \leq C_{nc} \leq C_c \tag{69}$$

As in [14] we now consider the following diversity channel,

$$y = zu + w \tag{70}$$

Where $\mathbf{z}$ is input and and $\mathbf{y}$ is the output. If we define $\mathbf{u} = \mathbf{x} + \mathbf{y}$ the capacity of this channel is given as ,

$$C \leq \log_2(1 + N\gamma). \tag{71}$$

If we define $\mathbf{z} = \mathbf{e}^{\mathbf{j}\mathbf{f}(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}, \mathbf{x}, \mathbf{z})}$ we end up with

$$\begin{aligned} I(f(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}, \mathbf{x}, \mathbf{z}); \mathbf{y}|\mathbf{x}, \mathbf{z})) &\leq \log_2(1 + N\gamma) \\ I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}, \mathbf{x}, \mathbf{z}; \mathbf{y}|\mathbf{x}, \mathbf{z}) &\leq \log_2(1 + N\gamma) \\ I(\theta_{\mathbf{x}}, \theta_{\mathbf{z}}; \mathbf{y}|\mathbf{x}, \mathbf{z}) &\leq \log_2(1 + N\gamma) \end{aligned} \tag{72}$$

Now incorporating (69) we get the following new bounds for C_{nc},

$$C_c - \frac{\log_2(1+N\gamma_{xz})}{N} \leq C_{nc} \leq C_c \tag{73}$$

Using the squeeze theorem we note that as $N \to \infty$, the value of C_{nc} tends to C_c which is the capacity of a coherent channel. Because the lower bound tends to C_{nc} which in turn tends to C_c the Gaussian distribution used to calculate the lower bound achieves capacity as $N \to \infty$.

4 Fading non-coherent channel

In this section, the lower bounds of the capacity region considering the case of a non-coherent channel with fading is explored. This is motivated by the fact that propagation through wireless channels is characterized by various effects such as multi-path fading[17]. Precise mathematical description of fading poses an intractable analysis problem. Therefore several relatively simple but accurate models are used to represent fading channels.

In this section I will use the Rayleigh distribution to model multi-path fading. The probability density function of a Rayleigh variable is (3) in [18]

$$p(\gamma) = \frac{1}{\tau} e^{-\frac{\gamma}{\tau}}. \tag{74}$$

Where γ is the signal to noise ratio and τ represents the average power of γ i.e $\tau = E\{\gamma\}$.

The system model presented in section two now becomes,

$$\mathbf{y} = \mathbf{a}\mathbf{x}e^{j\theta_{\mathbf{x}}} + \mathbf{b}\mathbf{z}e^{j\theta_{\mathbf{z}}} + \mathbf{w}, \tag{75}$$

where **a** and **b** represent the fading coefficients of both users. It has been assumed in this model that the two users signal experience different fading. We note that γ_x and γ_z vary in time due to the Rayleigh fading. Therefore, $\gamma_x + \gamma_z$ also varies in time. To calculate the channel capacity in a fading environment we have to do it in an average sense. Therefore to find the lower boundary region of a non-coherent fading channel we do it in an average sense.

As in section 3.2.2. the lower boundary area of the fading non-coherent channel for a two user case will be given by the convex hull of the set of rate of pairs satisfying,

$$\begin{aligned}
R_1 &\leq \int_0^\infty p(\gamma_x)\left(C_{cx} - 2(\gamma_{\mathbf{x}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x})p(\mathbf{x})\log_2 I_o\left(\frac{|\mathbf{y}^{\mathbf{T}}\mathbf{x}^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}\right)d\gamma_{\mathbf{x}} \\
R_2 &\leq \int_0^\infty p(\gamma_z)\left(C_{cz} - 2(\gamma_{\mathbf{z}})\log_2 e + \frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{z})p(\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y}^{\mathbf{T}}\mathbf{z}^*|}{\sigma^2}\right)d\mathbf{z}d\mathbf{y}\right)d\gamma_{\mathbf{z}} \\
C_{nc} &\leq \int_0^\infty p(\gamma_x + \gamma_z)C_c - 2(\gamma_{\mathbf{xz}})\log_2 e + \ldots \\
&\quad \ldots\frac{1}{N}\int\int\int p(\mathbf{y}|\mathbf{x},\mathbf{z})p(\mathbf{x},\mathbf{z})\log_2 I_o\left(\frac{|\mathbf{y}^{\mathbf{T}}(\mathbf{x}+\mathbf{z})^*|}{\sigma^2}\right)d\mathbf{x}d\mathbf{y}d\mathbf{z}d(\gamma_{\mathbf{x}} + \gamma_{\mathbf{z}})
\end{aligned} \tag{76}$$

Figure 3 shows the lower boundary of the capacity region for each value of N in a fading environment when both users are transmitting at 2dB. It can be seen that the lower boundary of the channel capacity region for a given value of N is lower in a fading channel compared to that of a non fading channel. This can be better seen by looking at figure 4 which shows the capacity regions lower boundaries for different values of N for both fading and non-fading channels when both users transmit at $2dB$. This is expected as the channel capacity in a Rayleigh fading environment is always lower than that of a Gaussian noise environment without fading[18].

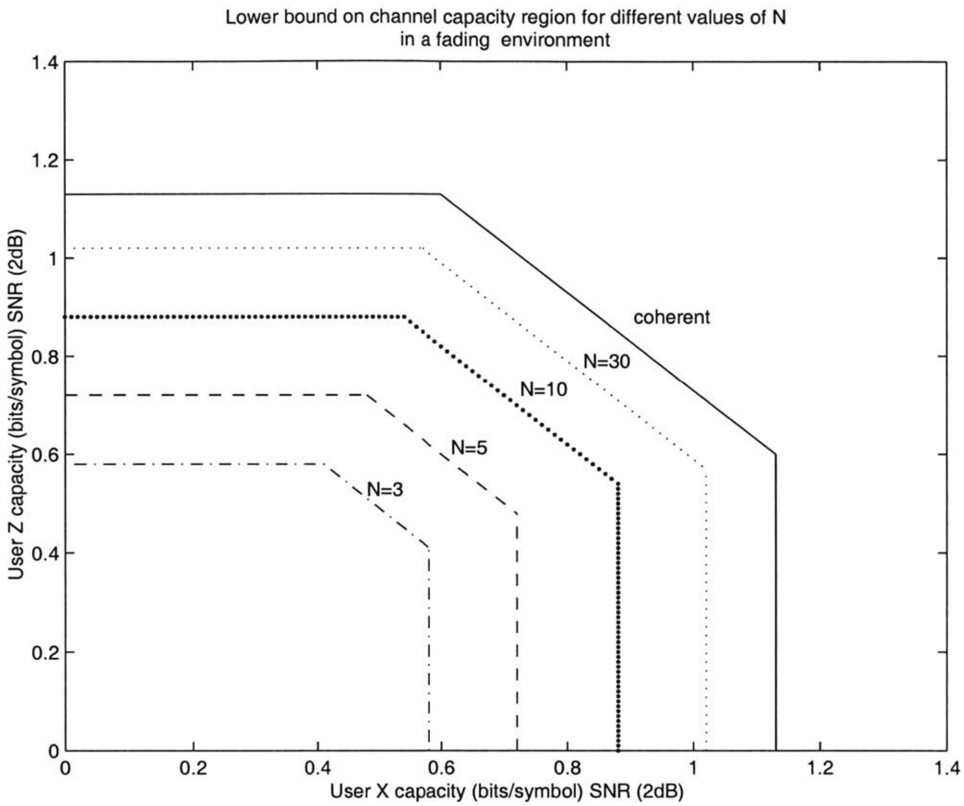

Figure 3: Lower bound on channel capacity region for different values of N for a fading non-coherent channel

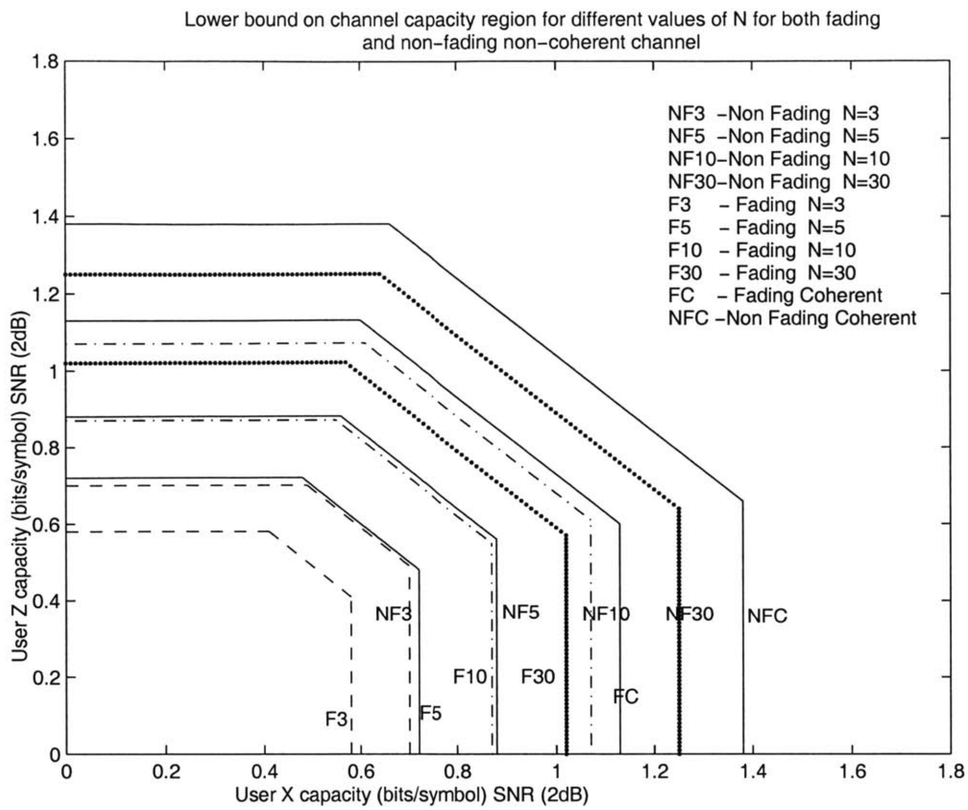

Figure 4: Lower bound on channel capacity region for different values of N for both fading and non-fading non-coherent channels

5 Conclusions

This thesis set out to find a lower and upper limit for the capacity region in a multi-access setting involving two users transmitting in a random-phase additive white Gaussian noise channel. The case for a single user transmitting in a non-coherent channel has been investigated in [14]. That investigation is carried on to a multi-user setting. The investigation is carried out in a similar manner to [14] and the results are similar. In particular, it is shown that the capacity achieving distribution for the two users is characterized by zero mean and uncorrelated components. It is also shown that a Gaussian distribution for each of the two users does not achieve capacity because it doesn't satisfy the properties of a capacity achieving distribution. Using a Gaussian distribution for each of the two users, a capacity region that is a lower bound is derived. It is shown that as the block length increases the capacity region approaches that of a coherent channel and that a Gaussian distribution for each of the two users with I.I.D components achieves capacity. Since the two users have both Gaussian distributions the joint distribution of $\mathbf{x}, \mathbf{z}$ can be thought of as a Bivariate Gaussian distribution at the component level. Finally, the case of fading is introduced into the model. Its found that the lower boundaries for the capacity regions for a given block length is lower in case of fading compared to a non-fading environment.

Further work will involve extending this model to a N user setting. From the results gathered from this thesis and prior work, it seems likely that the capacity region of the non-coherent channel as the block length increases will approach that of a coherent channel. Moreover, the results suggests that the capacity achieving distribution as the $N \to \infty$ will be a multivariate Gaussian. Further work will also involve evaluating this model with other fading models. For instance, fading phases may be considered as well as phenomena such as shadowing.

References

[1] D. Divsalar and M. K. Simon, "Maximum-likelihood differential detection of uncoded and trellis coded amplitude phase modulation over AWGN and fading channels - metrics and performance," *IEEE Transactions on Communications*, vol. 42, pp.76-89, January 1994.

[2] G. Colavolpe and R.Raheli,"Non-coherent sequence detection," accepted for publication in *IEEE Transactions on Communications*, submitted in august 1997

[3] G. Colavolpe and R. Raheli,"Noncoherent sequence detection of continuous phase modulations" accepted for publication in *IEEE Transactions on Communications*, submitted in August 1997.

[4] G. Colavolpe and R. Raheli,"Performance analysis of noncoherent sequence detection of M-ary PSK," in *IEEE International Conference on Communications.,(ICC '98)*,Atlanta , Georgia, U.S.A., June 1998,pp 438-443.

[5] M. Peleg and S. Shamai(Shitz),"On the capacity of the blockwise incoherent MPSK channel", *IEEE Transactions on Communications*, Vol.46,pp. 603-609, May 1998.

[6] S.A. Butman, I. Bar-David, B. K. Levitt, R. F. Lyon, and M.j. Klass,"Design criteria for noncoherent Gaussian channels with MFSK signal coding," *IEEE Transactions on Communications*, vol. COM-24, pp.1078-1088, oct. 1976.

[7] N. Chayat, "Turbo codes for incoherent M-ary orthogonal signaling," in Proc, *Nineteenth Conv. Electrical and Electronics Engineers* in Israel, Jerusalem, Israel,Nov. 5-6, 1996, pp. 471-474.

[8] Cheun and W.E. Stark,"Probability of error in frequency-hop spread spectrum multiple-access communications systems with noncoherent reception," *IEEE Transactions on Communications*, vol. 39,pp. 1400-1409, Sept. 1991.

[9] S. Zho, S.Mei, X. Xu, and Y. Yao, " Channel capacity of fast fading channels," in *IEEE Vehicular Technology Conference.*, Phoenix, AZ, May 1997, pp. 421-425.

[10] S. I. Chen and T. Fuja,"Channel capacity and cutoff rates for DPSK on fading channels," in *Conference on Information Science and Systems*, vol. 1, Princeton, NJ, 1996, pp.375-380.

[11] G. Colavolpe and R. Raheli, "The capacity of noncoherent channels ,"submitted to *IEEE Transactions on Communications*, January 1999.

[12] Z, Rodgers, R.L. Peterson,"Digital Communications and Spread Spectrum Systems" MacMillan publishing company, 1985.

[13] B. Sklar."Digital Communications Fundamentals and applications",Prentice Hall, 1988.

[14] T.M. Cover and J.A. Thomas, *"Elements of Information Theory."* New york: wiley, 1991.

[15] C.E. Shannon and W. Weaver, *"The Mathematical Theory of Communication"*. Urbana, IL: University of Illinois Press, 1963.

[16] A.D. Wyner, "Bounds on communication with polyphase coding",Bell syst. Tech. J., Vol. XLV,pp.523-559, April 1966.

[17] M.K. Simon and MS. Alouini, *"Digital communication over Fading channels"*, Wiley,2001.

[18] C.Y. Lee,"Estimate of Channel Capacity in Rayleigh Fading Environment," in*IEEE Transactions on Vehicular Technology*, vol. 39,No. 3, August 1990.