Diversity Multiplexing Trade-off in MIMO Systems

H.Srikanth Kamath
Dept of E & C
MIT Manipal

Vinod K.Madan
Dept of E & C
Kalasalingam University

ABSTRACT

This report describes my method of increasing the capacity of any wireless network through the application of MIMO technology. The Shannon capacity of wireless networks has been calculated in different instances by MATLAB. Using Simulink, the various parameters of a wireless Rayleigh Multipath fading channel are calculated. This is done in order to experimentally prove the advantages of a MIMO system in any wireless system. Also to find the right balance before which adding more antennas will not a make a difference

Key Words- Diversity, Tradeoff , MIMO, Antenna

1. INTRODUCTION

Wireless networks, though offering the most important benefit of freeing one from the bond of wires, have always faced some traditional problems. While they have the advantage of longer range communication, they also have some issues with reliability and speed. Both of these problems can be addressed with simple solutions. However these methods which increase reliability and speed come at a price of one another. This chapter discusses the methods and the balance between them.

The most simple and common way to increase the reliability of wireless networks is to increase the diversity of the network. There are three kinds of diversity, namely time diversity, frequency diversity and spatial diversity. Time diversity involves interleaving the input signal before transmitting it. This ensures that in case there is a burst error during transmission then only some bits of the current transmission are destroyed rather than the whole code. This decreases the error probability. Frequency diversity involves frequency hopping which enables security against frequency specific disturbances. Spatial diversity involves placing antennas at a minimum distance of λ/2, which makes sure there multiple paths through which the signal may pass. This helps as each signal may be attenuated differently and at the receiver all the received signals will help in detecting errors if present.

2. DIVERSITY CODING

The simplest form of diversity coding is repetition coding. In this method the same signal is repeated over whichever form of diversity is sought. Diversity in this method can be achieved also over frequency by sending the repeat of the signal over a different frequency. This helps when the channel is selectively coherent.

A. Time Diversity

Time diversity is achieved by averaging the fading of the channel over time. Typically, the channel coherence time is of the order of tens to hundreds of symbols, and therefore the channel is highly correlated across consecutive symbols. To ensure that the coded symbols are transmitted through independent or nearly independent fading gains, interleaving of codewords is required. For simplicity, let us consider a flat fading channel. We transmit a codeword x= [x1... xL] of length L symbols and the received signal is given by

$$y_l = h_l x_l + w_l \qquad l = 1.....L$$

Assuming ideal interleaving so that consecutive symbols x are transmitted sufficiently far apart in time, we can assume that the h are independent. The parameter L is commonly called the number of diversity branches.

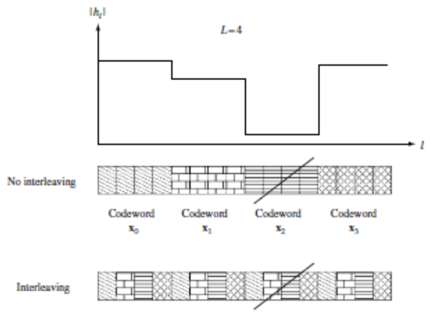


Figure 1 Interleaving

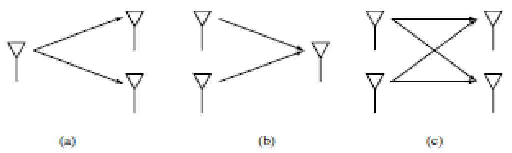
As shown above interleaving is the process of splitting up different code words into block each containing one symbol of each code word. This way if there is a momentary fading in the channel as shown in the fig it affects only one symbol from each of the code words. Thus simple error detection algorithms can reconstruct the original signal.

The rotation code discussed above is specifically designed to exploit time diversity in *fading* channels. In the AWGN channel, however, rotation of the constellation does not affect performance since the i.i.d. Gaussian noise is invariant to rotations. On the other hand, codes that are designed for the AWGN channel, such as linear block codes or convolutional codes, can be used to extract time diversity in fading channels when combined with interleaving.

B. Antenna Diversity

To exploit time diversity, interleaving and coding over several coherence time periods is necessary. When there is a strict delay constraint and/or the coherence time is large, this may not be possible. In this case other forms of diversity have to be obtained. Antenna diversity, or spatial diversity, can be obtained by placing multiple antennas at the transmitter and/or the receiver. If the antennas are placed sufficiently far apart, the channel gains between different antenna pairs fade more or less independently, and independent signal paths are created. The required antenna separation depends on the local scattering environment as well as on the carrier frequency. For a mobile which is near the ground with many scatterers around, the channel decorrelates over shorter spatial distances, and typical antenna separation of half to one carrier wavelength is sufficient. For base-stations on high towers, larger antenna separation of several to tens of wavelengths may be required.

We will look at both receive diversity, using multiple receive antennas (single input multiple output or SIMO channels), and transmit diversity, using multiple transmit antennas (multiple input single output or MISO channels). Interesting coding problems arise in the latter and have led to recent excitement in space-time codes. Antenna diversity is of two types, namely receive diversity and transmit diversity.



a)SIMO b) MISO c) MIMO

Figure 2 MIMO Systems

In the first figure there are multiple receive antennas depicting receive diversity. The second figure shows transmit diversity while the third depicts a full MIMO system.

3. FUNDAMENTALS

A. Transmit Diversity: Space Time Codes

Now consider the case when there are L transmit antennas and 1 receive antenna, the MISO channel (Figure 3.11(b)). This is common in the downlink of a cellular system since it is often cheaper to have multiple antennas at the base-station than to have multiple antennas at every handset. It is easy to get a diversity gain of L: simply transmit the same symbol over the L different antennas during L symbol times. At any one time, only one antenna is turned on and the rest are silent. This is simply a repetition code, and, as we have seen in the previous section, repetition codes are quite wasteful of degrees of freedom. More generally, *any* time diversity code of block length L can be used on this transmit diversity system: simply use one antenna at a time and transmit the coded symbols of the time diversity code successively over the different antennas. This provides a coding gain over the repetition code. One can also design codes specifically for the transmit diversity system. There have been a lot of research activities in this area under the rubric of *space-timecoding* and here we discuss the simplest, and yet one of the most elegant, space-time code: the so-called Alamouti scheme. This is the transmit diversity scheme proposed in several third-generation cellular standards. The Alamouti scheme is designed for two transmit antennas; generalization to more than two antennas is possible, to some extent.

B. Alamouti Scheme

With flat fading, the two transmit, single receive channel is written as

y[m] = h1[m]x1[m]+ h2[m]x2[m]+ w[m]

where hi is the channel gain from transmit antenna i. The Alamouti schemetransmits two complex symbols u1 and u2 over two symbol times: at time 1,x1[1] = u1, x2[1] = u2; at time 2, x1[2] =−u*2, x2[2] = u*1. If we assume that the channel remains constant over the two symbol times and set h1= h1[1] =h1[2] h2= h2[1] = h2[2], then we can write in matrix form:

$$[y[1]y[2] = [h1\; h2]\begin{pmatrix} u1 & -u2\,* \\ u2 & u1\,* \end{pmatrix} + \; [w[1]\; w[2]]$$

The Alamouti scheme works for any constellation for the symbols u1_u2, but suppose now they are BPSK symbols, thus conveying a total of two bits over two symbol times. In the repetition scheme, we need to use 4-PAM symbols to achieve the same data rate. To achieve the same minimum distance as the BPSK symbols in the Alamouti scheme, we need five times the energy per symbol. Taking into account the factor of 2 energy saving since we are only transmitting one symbol at a time in the repetition scheme, we see that the repetition scheme requires a factor of 2.5 (4 dB) more power than the Alamouti scheme. Again, the repetition scheme suffers from an inefficient utilization of the available degrees of freedom in the channel: over the two symbol times, bits are packed into only one dimension of the received signal space.

C. MIMO

Consider now a MIMO channel with two transmit and two receive antennas. Let hij be the Rayleigh distributed channel gain from transmit antenna j to receive antenna i. Suppose both the transmit antennas and the receive antennas are spaced sufficiently far apart that the fading gains, hij, can be assumed to be independent. There are four independently faded signal paths between the transmitter and the receiver, suggesting that the maximum diversity gain that can be achieved is 4. The same repetition scheme described in the last section can achieve this performance: transmit the same symbol over the two antennas in two consecutive symbol times (at each time, nothing is sent over the other antenna). If the transmitted symbol is x, the received symbols at the two receive antennas are

$$y_l[1] = \; h_{i1}x + \; w_l[1]$$

at time 1 and

$$y_l[2] = \; h_{l2}x + \; w_l[2]$$

at time 2. By performing maximal-ratio combining of the four received symbols, an effective channel with gain $\sum_{i=1}^{2}\sum_{j=1}^{2}|h_{ij}|^2$ is created, yielding a four-fold diversity gain. However, just as in the case of the 2×1 channel, the repetition scheme utilizes the degrees of freedom in the channel poorly; it only transmits one data symbol per two symbol times. In this regard, the Alamouti scheme performs better by transmitting two data symbols over two symbol times. The degree of freedom of a channel is defined as the dimension of the received signal space. In a channel with two transmit and a single receive antenna, this is equal to *one* for every symbol time. The repetition scheme utilizes only half a degree of freedom per symbol time, while the Alamouti scheme utilizes all of it.

As we understand the fundamentals of a basic MIMO system we can see for a system with Nt transmit antennas and Nr receive antennas, diversity and multiplexing can be achieved but not both together. We shall now see the diversity multiplexing trade-off curve and how it varies according to the number of antennas on either side.

The tradeoff curve summarizes succinctly the performance capability of the slow fading MIMO channel. At one extreme where r →0, the maximal diversity gain nt ·nr is achieved, at the expense of very low multiplexing gain. At the other extreme where r→nmin, the full degrees of freedom are attained. However, the system is now operating very close to the fast fading capacity and there is little protection against the randomness of the slow fading channel; the diversity gain is approaching 0. The tradeoff curve bridges between the two extremes and provides a more complete picture of the slow fading performance capability than the two extreme points. For example, adding one transmit and one receive antenna to the system increases the degrees of freedom min(nt, nr) by 1; this corresponds to increasing the maximum possible multiplexing gain by 1. The tradeoff curve gives a more refined picture of the system benefit: for *any* diversity requirement d, the supported multiplexing gain is increased by 1.

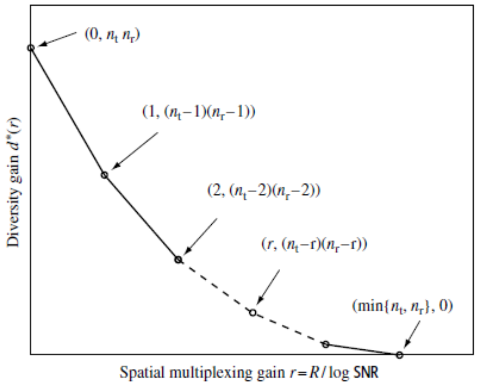


Figure 3: Trade off Curve

The graph shown above is varied according to the number of antennas on either side. This curve is also affected by the type of space time codes used to transmit the data. This is shown in the fig below.

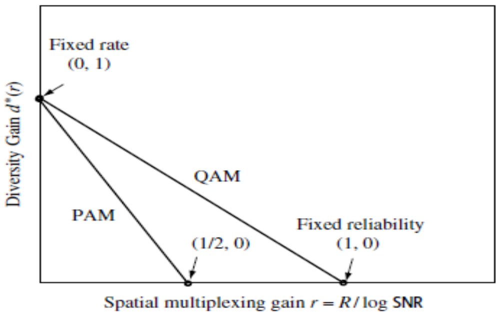


Figure 4: Trade-off curve for different types of modulation

D. Rayleigh Flat Fading Channel

Rayleigh fading is a statistical model for the effect of a propagation environment on a radio signal, such as that used by wireless devices.Rayleigh fading models assume that the magnitude of a signal that has passed through such a transmission medium (also called a communications channel) will vary randomly, or fade, according to a Rayleigh distribution — the radial component of the sum of two uncorrelated Gaussian random variables.Rayleigh fading is viewed as a reasonable model for tropospheric and ionospheric signal propagation as well as the effect of heavily built-up urban environments on radio signals. Rayleigh fading is most applicable when there is no dominant propagation along a line of sight between the transmitter and receiver. It is a reasonable model when there are many objects in the environment that scatter the radio signal before it arrives at the receiver. The central limit theorem holds that, if there is sufficiently much scatter, the channel impulse response will be well-modelled as a Gaussian process irrespective of the distribution of the individual components. If there is no dominant component to the scatter, then such a process will have zero mean and phase evenly distributed between 0 and 2π radians. The envelope of the channel response will therefore be Rayleigh distributed.The requirement that there be many scatterers present means that Rayleigh fading can be a useful model in heavily built-up city centres where there is no line of sight between the transmitter and receiver and many buildings and other objects attenuate, reflect, refract, and diffract the signal. Experimental work in Manhattan has found near-Rayleigh fading there. In tropospheric and ionospheric signal propagation the many particles in the atmospheric layers act as scatterers and this kind of environment may also approximate Rayleigh fading. If the environment is such that, in addition to the scattering, there is a strongly dominant signal seen at the receiver, usually caused by a line of sight, then the mean of the random process will no longer be zero, varying instead around the power-level of the dominant path. Such a situation may be better modelled as Rician fading.

How rapidly the channel fades will be affected by how fast the receiver and/or transmitter are moving. Motion causes Doppler shift in the received signal components. The figures show the power variation over 1 second of a constant signal after passing through a single-path Rayleigh fading channel with a maximum Doppler shift of 10 Hz and 100 Hz. These Doppler shifts correspond to velocities of about 6 km/h (4 mph) and 60 km/h (40 mph) respectively at 1800 MHz, one of the operating frequencies for GSM mobile phones. This is the classic shape of Rayleigh fading. Note in particular the 'deep fades' where signal strength can drop by a factor of several thousand, or 30-40 dB.

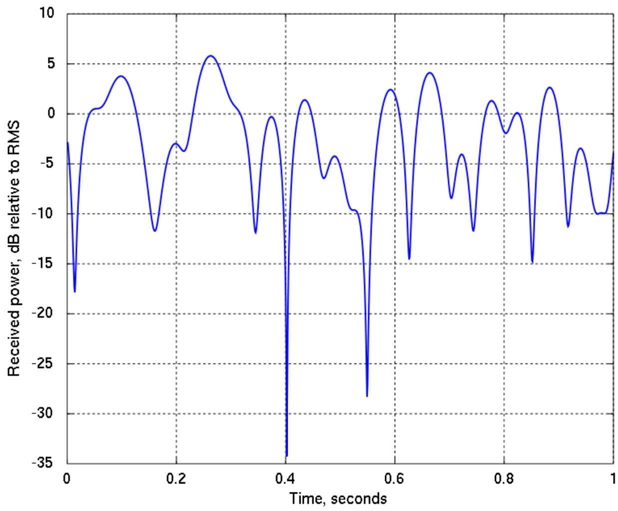


Figure 5: Rayleigh Distribution

Since it is based on a well-studied distribution with special properties, the Rayleigh distribution lends itself to analysis, and the key features that affect the performance of a wireless network have analytic expressions.

4. METHODOLOGY

The first objective to be achieved was to estimate the Shannon capacity of a MIMO system and compare it with the Shannon capacity of a SISO system. The estimate for the SISO system was calculated using the Shannon formula.

This step was iterated for different levels of signal to noise ratio using a loop function, thus obtaining a smooth curve of the capacities at different levels of SNR. The varying levels of SNR are an indication to the power emitted by the transmit antenna. In the next step the Shannon capacity is calculated for a MIMO system by varying the number of antennas. This time around the calculation is repeated and the curve is estimated. The next step involved the calculation of the ergodic capacity and the outage of a MIMO system. For this step a MIMO toolbox was obtained to set the matrices needed for the calculations.

A complete MIMO system was also devised using simulink to simulate a real environment and measure the various parameters involved such bit rate, error probability etc. In a real environment, various factors influence the system. The model was so designed to allow different changes to be made according to the requirements. The model starts with a Bernoulli's bit generator for the data. The raw data is then modulated using any of the modulation techniques possible. The various modulation techniques give varying gains and also utilize the degrees of freedom in varying levels. Thus the block for modulation is kept separately. This block can be changed at any time to suit the user's needs. The modulated data is then passed on to the L-antenna channel. This makes up the transmitter side, whereas on the receiver side the received signal is a sum total of all the signals from all the various antennas. There has to be a proper method to estimate the original signal from the multiple independently faded signals available. I have used the simplest possible method for estimation known as Maximal Ratio Combining. The block was designed for MRC and applied in the model. This is generally applied for AWGN channels, however in our simplistic model it does find relevance.

In telecommunications, maximal-ratio combining is a method of diversity combining in which:
(a) the signals from each channel are added together,
(b) the gain of each channel is made proportional to the rms signal level and inversely proportional to the mean square noise level in that channel.
(c) different proportionality constants are used for each channel. It is also known as ratio-squared combining and predetection combining.

The whole setup is considered on a Rayleigh multipath fading path channel. The channel can also be changed by altering the block. The simulation can also be done for a Rician channel

The Simulink model conceptualised is shown below.

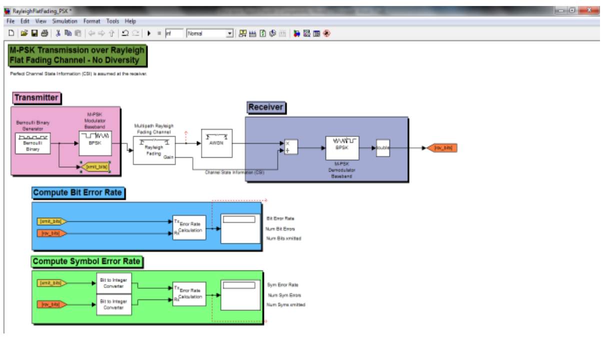


Figure 6:The Simulink model

The model shown above is a simple model involving a transmitter and receiver side. The data is simulated using a Bernoulli bit generator. The data is then modulated using Binary Phase Shift Keying. The modulated signal is then sent over a Rayleigh flat fading channel. It is here we assume that we have perfect channel state information for the sake of simplicity. It is essential here to understand what exactly channel state information is. In wireless communications, channel state information (CSI) refers to known channel properties of a communication link. This information describes how a signal propagates from the transmitter to the receiver and represents the combined effect of, for example, scattering, fading, and power decay with distance. The CSI makes it possible to adapt transmissions to current channel conditions, which is crucial for achieving reliable communication with high data rates in multi-antenna systems. This model is used only to compare with the results obtained in a MIMO model. It should be noticed that the bit error rate and the symbol error rate are being calculated. However for further tabulations we shall only plot the bit error rate with respect to the signal to noise ratio.

The next model designed was with dual antennas thereby an attempt to half the bit error rate. It is shown below.

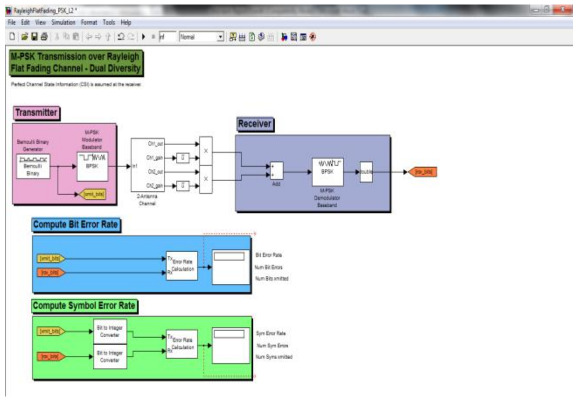


Figure 7

In this model a precoded block is used to simulate the dual antenna channel. Even in this case the knowledge of channel state information is assumed. On the recombining part simple addition has been used.

The next model is the final model which emulates a MIMO system. It consists of a L-antenna channel which can vary the number of antennas. It is shown below.

The model has been used to calculate the final results proposed originally. The parameters of the L-antenna block include the SNR and also the number of antennas. Thus the model is run iteratively with different number of antennas at various SNR values. The graphs and results obtained can be seen in the next section of this report.

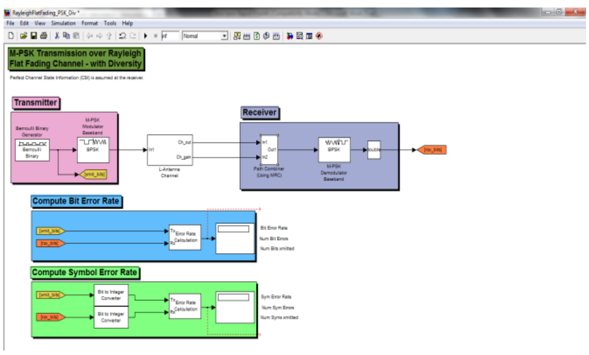


Figure 8

5. RESULT ANALYSIS

The Shannon capacities calculated were depicted in the form of a graph as shown below. The figure clearly shows the effectiveness of a MIMO system over a traditional single antenna system. It can also be inferred from the graph that as the number of antennas increases the capacity of the channel at lower signal to noise ratios.

Thus it can be understood the increase in antennas directly increases the capacity at lower SNRs thus enabling lower input powers for better coverage. Lower signal to noise ratios means the power emitted by the antennas which decreases the power costs. Also the dangers of high exposure to dangerous radiation by the people around the antennas are decreased

Now that the superior channel capacity of a diversity channel has been shown conclusively we move on to the next part of results of showing the lower bit error rates of a diversity channel. Firstly the results of the reference model are presented to be compared with the rest of the later results.

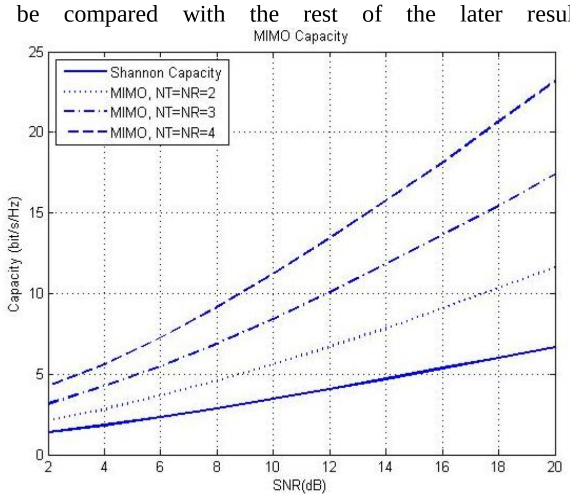


Figure 9

It may be noted that the capacity is higher for lower SNR which helps us decrease the power needed to send the signal.

The complete overview of the different error rates at varying degrees of diversity can been represented in a graphical form for easy comprehension. [fig 4.3]

From the graph it can be noticed that as the diversity increases lower bit bit error rates are achieved at lower SNR. This is main goal stated at the start of this report. This result shows the advantage of MIMO. The power of the original signal can be greatly reduced by incorporating a few more antennas in diversity mode. This increases the reliability of the communication system through redundancy.

Now that we have proved that increasing the number of antennas is beneficial to the system's reliability we have to see whether there is an upper limit on the number of antennas that can be included. Theoretically speaking there is no limit on the number of antennas, however in a real world system the cost of antennas has to be factored in. This is done by the telecommunications company. A point to be noted here is that the decrease in the bit error rate is not constant. Thus after a certain number of antenna the decrease slows down to a very low factor which is where it becomes pointless to go for more diversity. The table given below will give an idea of this. The values in this table have been obtained from the same model.

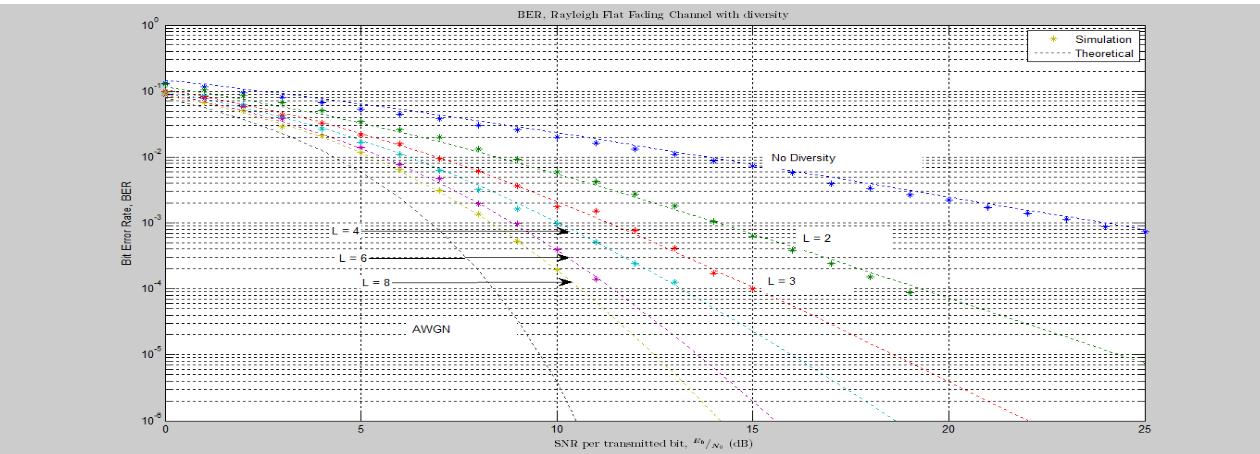


Figure 10

Table 1

Number of Antenna	Bit Error Rate
2	5.8072e-003
10	1.2101e-004
30	1.8230e-005
50	7.5308e-006
70	2.7393e-006
90	6.9941e-007
100	1.0258e-007

The above table gives a fair idea about how the gain in increasing antennas decreases over a certain limit. The fall of the bit error rate decreases and the gain in increasing the antennas goes from a lot to jus negligible. Factoring in the cost of antennas any telecommunication company will have to install jus the optimum number of antennas. The main point to be kept in mind is that the quality of service of the communication system is to be maintained at the lowest possible cost.

6. CONCLUSION

The Shannon capacity is observed to be lowest for a single input single output system.The capacity increases as the signal to noise ratio increases. This means the capacity is directly related to the power emitted by the antenna. As the number of antennas to transmit and receive increases the power needed for the same capacity is lower than for a conventional system. While the Shannon capacity provides an insight into the speeds possible the second part of the experiment has proved the reliability of a MIMO system.The increase in the number of antennas has significantly decreased the bit error rate. A decrease in the bit error rate signifies the reliability of the system to transmit the data safely without any data loss. The Monte Carlo simulations done provide the values required to assess the viability of incorporating a MIMO system in any communication setup. Increasing the number of antennas blindly too does not guarantee a better performance. The cost of the antenna too has to be factored in including any other setup with the antenna such as power supply, modulators and other such subsystems. The Rayleigh flat fading channel has been used as it is the best model to simulate for non-line of sight communication in an urban environment. The same experiment may be repeated for a line of sight urban environment using a Rician model where the variations lie around the main received signal.

7. REFERENCES

Journal / Conference Papers

[1] L. Zheng and D. Tse, "Diversity and Multiplexing: A fundamental tradeoff in multiple antenna channels", *IEEE Trans. on Information Theory*, vol. 49(5), May 2003.

[2] Gerhard Kramer, IvanaMaric and Roy D. Yates, "Cooperative Networks" Journal Foundations and Trends in Networking, Volume 1 Issue 3, August 2006

[3] Giuseppe Caire, PetrosElia and K. Raj Kumar, "Space-Time Coding: an Overview," Invited paper to Journal of Communications Software and Systems, October 2006.

Reference / Hand Books

[1] Tse and Viswanath,P. Fundamentals of Wireless Communication, Cambridge University Press.

[2] Simon Haykin, 1998.Digital Communication, Wiley and Sons.

[3] Rappaport Theodore,S. Wireless Communication, Pearson Publications