CIRCULANT MIMO STRUCTURES

by

Sayed Amir Mirtaheri

B.Sc., Sharif University of Technology, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE
in the School
of
Engineering Science

SIMON FRASER UNIVERSITY
Summer 2005

APPROVAL

Name:	Sayed Amir Mirtaheri
Degree:	Master of Applied Science
Title of thesis:	Circulant MIMO Structures
Examining Committee:	Dr. Bonnie Gray Chair
	Dr. Rodney Vaughan, Senior Supervisor
	Dr. James Cavers, Supervisor
	Dr. Paul Ho, SFU Examiner
Date Approved:	August 05, 2005

SIMON FRASER UNIVERSITY

PARTIAL COPYRIGHT LICENCE

Abstract

Multiple-input, multiple-output (MIMO) systems allow increased capacity over single port antenna systems in the presence of multipath fading environments. The challenging areas in a MIMO system overlap between the propagation channel, the antennas and the signal processing. In this dissertation two aspects of MIMO theory are investigated. Firstly, the effect of systematic correlation on the capacity efficiency is analyzed in detail. This analysis has been undertaken through the introduction of a specific correlated structure, namely the circulant. Compared to the completely random (i.i.d.) structure, the circulant shows a clear capacity increase both in the theoretical Shannon limit and also in the practicable, QAM-included case. This fascinating behavior can be fully explained through investigation of the pdfs of the eigenvalues of the channel matrices. The investigation shows that the capacity increase arises from the more similar eigenvalues of the circulant structure. The empirical pdfs of the eigenvalues are presented and parameters are introduced to compare the similarity between the eigenvalues. Furthermore, the basic hypothesis of parallel channel capacity is clarified with respect to the water-filling of MIMO eigenchannels. Other advantages of the circulant structure, based on its fixed eigenvectors, have been developed. In particular, it is possible to reduce the degradation caused by errors in the channel estimation for circulant channels. While this first aspect of MIMO theory concerns the channel modeling and signal processing, the second aspect focuses on the important practical issue of power allocation between eigenchannels. The optimum power allocation is the non-linear strategy of water filling, but this is expensive in processing power to implement. Therefore it is of interest to decrease the complexity of water filling using a sub-optimum method. It is shown here that it is reasonable to circumvent the complexity of water filling by simply using equal powers. Again, this simplification is feasible because of the arrangement of the eigenvalues in the channel matrix. For differently dimensioned MIMO systems, there is a different power

threshold above which we can substitute equal powers instead of water filling powers. An experimental rule for this power threshold has been derived.

To My Mother

"Science may be described as the art of systematic over-simplification."

— KARL POPPER, *1902-1994*

Acknowledgments

I would like to thank my parents, who taught me how to learn instead of what to learn. Also, I would like to thank my supervisor, Professor Vaughan, for what he taught me, for his great support, and for his patience. I should also thank Professor Cavers and my great friends, and some special thanks to Roozbeh Ghaffari, Maryam Soltanzadeh and Pouya Dehghani.

Contents

List of Tables

List of Figures

Preface

When I began to add to my knowledge about Multi-Input Multi-Output theory, I found that MIMO is exactly showing the spirit of an engineering effort. Instead of trying to remove the interference between different signals which is not feasible, MIMO started to use this classic destructive factor as a possible source of information. After a while, I understood that there is a commonly accepted idea about MIMO systems which says the best performance of MIMO systems can be achieved in uncorrelated situations. This was an ideal assumption far away from the real world.

Since MIMO theory views the interference as a positive factor, instead of negative factor, then could the same be happen to the correlation? The main part of this thesis includes my effort that this idea is true. Systematic correlation can be a constructive factor in the performance of the MIMO system. I have shown these results for a correlated structure v.z. circulant structure which also introduces some other advantages.

The fact that correlation is an inevitable phenomenon in the real world means that instead of trying to remove it, we can put our effort into exploiting it. This thesis work is basically a start for that journey.

Chapter 1

Introduction

Wireless communication systems are in the center of much technical and academic research and development in electrical engineering. One of the primary reasons for this is the convenience of being a member of a communication network when people are mobile. Furthermore, this convenience has become a necessity in recent years. This necessity founds the new era of information networks in which everybody increasingly wants to access information. The result, that of increasing demand for connectivity raises practical difficulties which engineers try to resolve.

Starting from a stationary transmitter and receiver, wireless communications has come a long way to considering Multi-Input Multi-Output (MIMO) systems as an answer to one of its major challenges i.e. spectrum efficiency. The nature of wireless communication in which there is no constrained media to send the data through (in contrast with wired communication) as well as an increasing number of users, are the motivations for seeking new techniques to use the spectrum more efficiently. For a single signal sent through a wireless channel, there are natural phenomena such as scattering and reflection which lead to what is usually called *fading*. Also, with currently used frequencies, the sizes of buildings in the urban areas and their materials cause a significant attenuation causing *shadowing*. Moreover, when the multiple signals share the same channel, the issue of *interchannel interference* is of a great importance such that it completely marginalizes the classic issue of *noise* in communications.

There is always the possibility of using more spectrum to overcome each of these challenges but according to the increasing number of spectrum users, this is no longer a feasible

solution. Multi-Input Multi-Output (MIMO) systems are one of the most promising answers which at least theoretically could overcome many of these issues without any need to use more spectrum. MIMO, which can also be directly deployed for high data rate wireless communications, is based on the simple idea of using multiple antennas in the receiver and transmitter and it has been shown that theoretically (i.e. ideal channel estimation), MIMO capacity increases linearly with the number of antennas.

There are three main areas in the MIMO systems: the signal processing backbone, the antennas and the channel. Among them, the channel is the part where there is less control but more importance in terms of unwanted effects on the signals. The main research tendencies in MIMO are working on signal processing and antenna design parts which are the accessible components. Our knowledge about the channel directly affects our efforts in the two other parts. Because of this, many of the contributions and research in the area of signal processing and antennas are based on the assumption of knowing the channel. In the traditional context of communications, by having the channel information, we usually refer to the receiver, which may have the channel information (for coherent detection) or may not (for non-coherent detection methods like differential detection). But in the wireless communications and specially in the MIMO theory, not only the receiver but also the transmitter ideally needs to know the channel information. Based on this, it is very important to investigate the channel itself. This dissertation looks into the channel and tries to extract some facts about the connections between the channel structure and the MIMO overall behavior.

We will study two different channel structures and compare not only their theoretical capacity efficiency, but also their practical ones, i.e. including digital modulation techniques. The models are the classic i.i.d. channel structure (or completely Gaussian structure) and the circulant Gaussian channel structure. The potential benefits of the general correlated structure will be shown. The distribution of the eigenvalues is the basic mechanism by which these benefits which will be reviewed in detail. We will also introduce the unique and important benefit of the circulant structure in terms of sensitivity to the detection noise.

Finally, the last chapter, the investigation will also extend to the issue of non-linear power allocation in MIMO which is usually called *water filling*. The combination of mathematical characteristics of water filling and the distribution of eigenvalues of the channel matrix is the underlying mechanism for the capacity behavior. The sub-optimum method of equal power allocation is compared with water filling.

Two conference papers have been presented and published from this thesis:

- "Circulant Wired MIMO Structures", IEEE Canadian Conference on Electrical and Computer Engineering, May 2005.
- "Advantages of Circulant MIMO Structures", The IASTED International Conference on Antennas, Radar and Wave Propagation, July 2005.

Chapter 2

Background

Multiple-Input Multiple-Output (MIMO) systems are essentially a new visualization of the classic problem of one channel shared by multiple users. In that classic problem, each user sees the others as the sources of interference and the main goal is to decrease the amount of interference in order to have a better signal to interference plus noise ratio ($SINR = \frac{S}{I+N}$) and finally, a better detection at the receiver. Many smart schemes have arisen from this endeavor. These schemes are in different areas of telecommunication such as coding, digital communication techniques (signal processing), antenna designs and beam forming. However, the underlying problem of interference still exists due to the increasing number of users of mobile communications. Based on these facts, engineers began to consider the possibility of using the interference not as a destructive factor but as a potential beneficial source which carries some information. The mathematical models, followed for a new generation of links called MIMO. The idea is to send different data streams via multiple antennas at the transmitter into the same channel, sounding the channel by multiple antennas at the receiver and finally trying to jointly detect the different streams through signal processing. Though the theoretical basis of full MIMO has been built up, there are practical issues making it a commercially difficult scheme currently.

Unlike wire line channels, in wireless communications the signals are exposed to scattering, diffraction and reflection due to the unconstrained media. It is possible to classify these disturbances as *fading*, *shadowing* and *delay spread*. In addition, fading can be subcategorized to *flat* fading and *frequency selective* fading. Their difference basically relates to the channel environment.

Before going through the above mentioned characteristics it is helpful to consider Figure

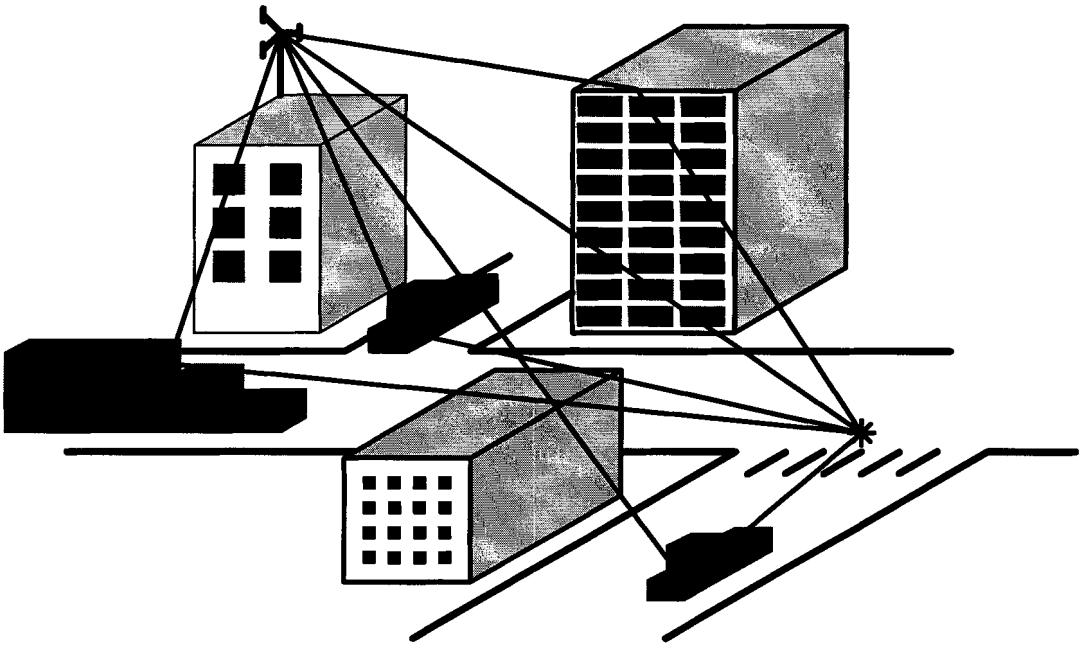

Figure 2.1: A Wireless Multipath Environment.

2.1 which depicts a wireless channel in an urban area. Usually there is a base station responsible for providing service to a mobile subscriber in its area of coverage. But there are also different obstacles, reflectors and scatterers, which could be stationary, such as buildings, or in motion, such as cars and people. So, there is a *multipath* channel between the base station and the user. In other words, the signal has more than one way to get to the receiver and each of these ways introduces different attenuation and propagation delay. Finally the different versions of the desired signal add up at the receiver. The basic challenge is that the channel shows time varying characteristics due to the motion which is in the nature of mobile communications. In this chapter, after providing some introductions about these basic characteristics of wireless channels, we will also overview the MIMO theory and its mathematical basis.

2.1 Fading

In principle, fading in the received signal is caused by the variation with time of either the amplitude or the relative phase, or both, of one or more of the frequency components of the signal. Fading arises from the superposition of different replicas of the signal[1]. In flat fading, these variations are taking place simultaneously in all frequency components of the received signal. In frequency selective fading this is not the case. From [29] consider the transmitted signal generally represented as

$$s(t) = Re\left[s_l(t)e^{j2\pi f_c t}\right] \tag{2.1}$$

where s_l is the complex envelope and f_c is the carrier frequency. Assume that there are n propagation paths each of them having a time-varying propagation delay and an attenuation factor. Thus, the received band-pass signal is

$$x(t) = \sum_n \alpha_n(t)s[t - \tau_n(t)] \tag{2.2}$$

where $\alpha_n(t)$ is the attenuation and $\tau_n(t)$ is the propagation delay of the nth path. From (2.1) and (2.2) we will have

$$x(t) = Re\left(\left\{\sum_n \alpha_n(t)e^{-2\pi f_c \tau_n(t)}s_l[t - \tau_n(t)]\right\}e^{j2\pi f_c t}\right). \tag{2.3}$$

So, the low-pass received signal is

$$r_l(t) = \sum_n \alpha_n(t)e^{-2\pi f_c \tau_n(t)}s_l[t - \tau_n(t)]. \tag{2.4}$$

Here, $r_l(t)$ is the response of the equivalent low-pass channel to the equivalent low-pass signal $s_l(t)$ and therefore, the equivalent low-pass channel could be described by the time-varying impulse response of

$$c(\tau; t) = \sum_n \alpha_n(t)e^{-2\pi f_c \tau_n(t)}\delta[t - \tau_n(t)] \tag{2.5}$$

where $c(\tau; t)$ is the response of the channel at time t to an impulse applied at the time $t - \tau$.

To investigate the concept of fading, let us assume $s_l(t) = 1$ i.e. only the unmodulated

[1]This section has been heavily derived from [29].

carrier has been sent. From (2.4) the received signal reduces to

$$\begin{aligned} r_l(t) &= \sum_n \alpha_n(t) e^{-2\pi f_c \tau_n(t)} \\ &= \sum_n \alpha_n(t) e^{-j\theta_n(t)}. \end{aligned} \tag{2.6}$$

This means that the received signal consists of time-varying phasors with the amplitudes of $\alpha_n(t)$ and the phases of $\theta_n(t)$. Note that $\alpha_n(t)$ is more stable than $\theta_n(t)$ which is sensitive to even small motions in the medium. The randomly variant phases $\theta_n(t)$ may lead to destructive addition of the phasors and result in a very small (practically zero) $r_l(t)$. On the other hand, these phases may cause a constructive superposition of the phasors which leads to a large received signal. This amplitude variations of received signal due to variant phases is exactly what we mean by *fading*.

Since each term inside the summation in (2.6) is a random process and there are large number of these, $r_l(t)$ could be modeled as complex Gaussian random process from the central limit theorem. In that case, $c(\tau;t)$ will be a complex Gaussian random process in t and consequently $|c(\tau;t)|$ is a Rayleigh-distributed random variable at any instant t. This is usually called *Rayleigh fading channel*. When there are few paths or there is a strong line-of-sight path, $|c(\tau;t)|$ is no longer a Rayleigh-distributed random process and can be modeled with the Rice (*Ricean fading*) or Nakagami (*Nakagami fading*) distributions [36]. Figure 2.2 shows a sample Rayleigh faded signal. Since the waves travel with a fixed speed, the horizontal axis could be also considered as the distance normalized by the speed of the receiver. This means that fading could change the SNR rapidly over distances in the order of wavelength. In the following section, we will briefly go through the flat and frequency fading channels. The more detailed description and mathematical results for fading could be found in [41], [4] and [36].

2.1.1 Flat Fading and Frequency Selective Fading Channels

Assume $\Omega(\omega)$ to be the frequency response of the wireless channel [41]. The *coherence bandwidth*, C_C, for such a channel, is the bandwidth over which the frequency response is correlated more than a pre-defined correlation threshold. This correlation threshold does not have a fixed definition and is taken from $e^{-1} = 0.37$ to 0.9 by different authors. Coherence bandwidth is defined usually for the channels with significant dispersion [4].

According to the definition of C_C, a channel is called *flat fading* when the bandwidth of

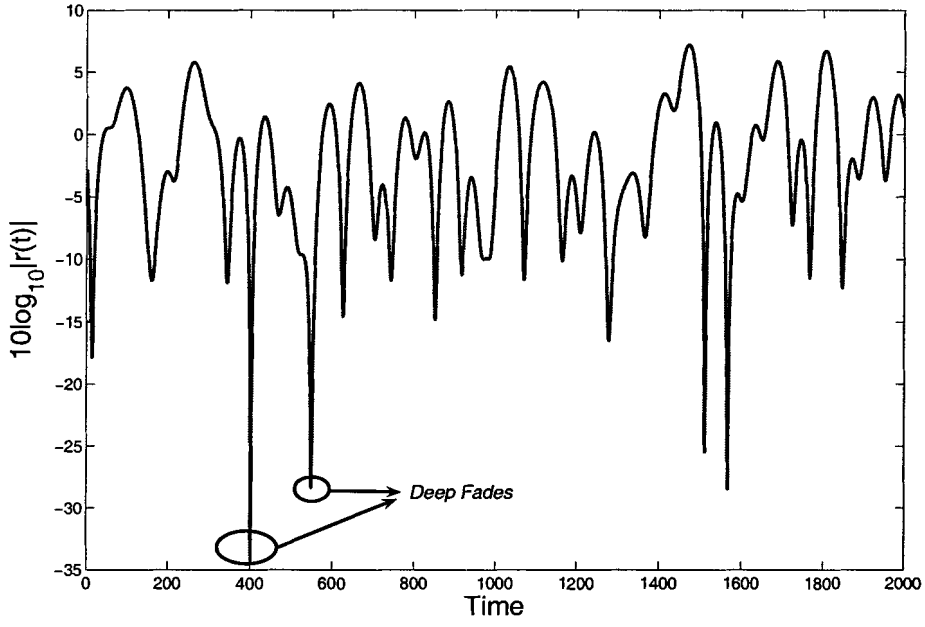

Figure 2.2: The Envelope of a Sample Faded Signal.

the signal is much smaller than the coherence bandwidth. Otherwise, it is a *frequency selective* channel. In some texts, instead of flat fading and frequency selective fading channels, the alternative expressions of *narrowband* and *wideband* channels are being used.

As can be seen in Figure 2.3, in a narrowband channel, the frequency response of the channel is roughly constant. It implies that all the frequency components of the signal experience the same fading pattern, while in the wideband channel (See Figure 2.4), different frequency components undergo different fadings.

A commonly used model for the Rayleigh flat fading channels is the Jakes' scenario [19]. On the other hand, since the inter-symbol interference is a major issue in the frequency selective channels, a different approach is needed in those channels. A proper model for simulation of this kind of channels is a tapped delay line where there is more control to reduce the effect of ISI [9].

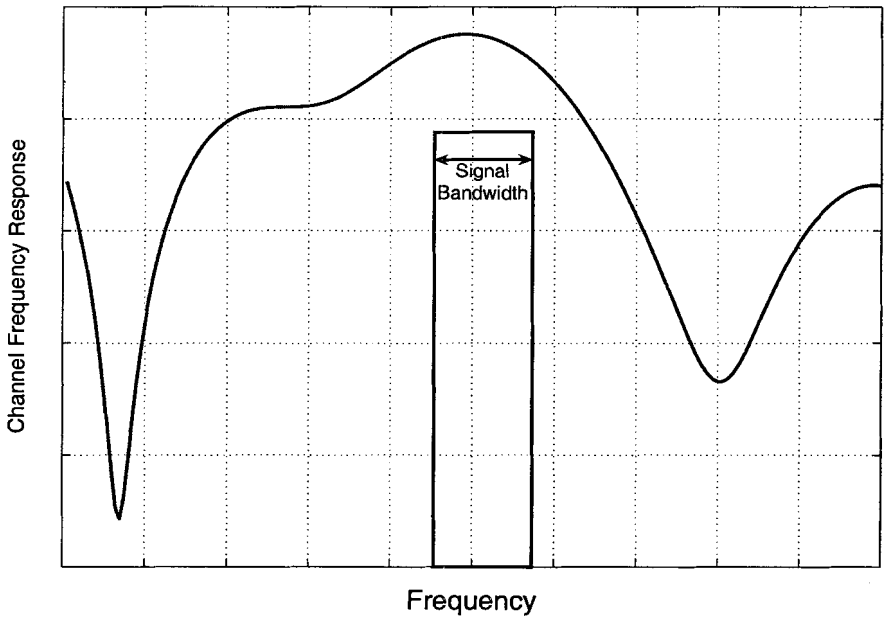

Figure 2.3: Rayleigh Flat Fading or Narrowband Channel.

2.2 Shadowing

Different from free space communications where the distance between the transmitter and the receiver is the key parameter determining the signal-to-noise ratio, the existence of huge obstacles such as hills (in rural areas for example) and large buildings (in urban areas for example) causes a phenomenon called *shadowing* in mobile communications.

Here, the locations at the same distance from the transmitter receive different power due to the different attenuations in the different paths. For example, if there was a building between the transmitter and one of these locations, the *shadow* of this building diminishes the average received power for that location. It is known that shadowing loss could be modeled well enough with the log-normal distribution [4] [36].

The reason for this modeling is the multipath nature of the channel, where the sum of different versions of the signal is available at the receiver. We have shown that this sum, according to the central limit theorem, tends to feature the Gaussian distribution. On the other hand, in terms of powers in dB, total path loss is the sum of losses for each of these

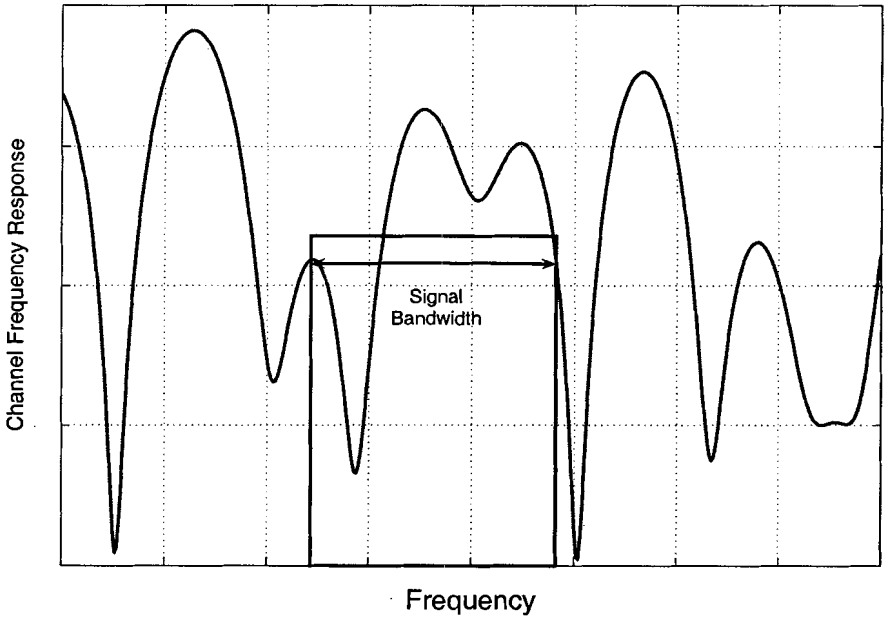

Figure 2.4: Rayleigh Frequency Selective Fading or Wideband Channel.

paths. This sum also converges to a Gaussian distribution as mentioned in the central limit theorem. So, since the average power in dB features a normal distribution, the average power in natural units (e.g. watts) has a log-normal distribution.

Consider $P(r)$ being the average received power at distance r from the transmitter. Then this power in dB, i.e. $P_{dB}(r) = 10\log_{10} P(r)$, is a Gaussian random variable. The parameters of this Gaussian distribution are essentially determined by the environment and also by the communication system, e.g. by antenna configuration. Specifically the mean value ($\mu(r)$) of this Gaussian distribution is a function of distance r. The log-normal pdf of $P(r)$ is

$$f(P(r)) = \frac{1}{\sigma\sqrt{2\pi}P(r)} e^{(ln(P(r))-\mu(r))^2/(2\sigma^2)}. \tag{2.7}$$

Note that by the average power, we mean averaging over a variation of distance. This variation could be small scale due to multipath fading or large scale (in the order of building sizes e.g. 20 meters) due to shadowing.

2.3 Delay Spread

Due to the multipath character of the wireless channels, the different versions of the transmitted signal are received by the receiver with different delays. This phenomenon causes Inter-Symbol Interference (ISI) which degrades the system performance. It is clear that the degradation caused by delay spread in terms of ISI significantly depends on the guard interval duration between the transmitted signals. There are different techniques to overcome ISI. The commonest and traditional scheme is implementation of an adaptive equalizer [29].

However, it is important to formulate the concept of delay as a mathematical quantity. To do this, assume that the impulse response of the channel is given by $h(t)$ [41]. The *power delay profile* is

$$P(t) = \overline{|h(t)|^2} \tag{2.8}$$

where $\overline{|h(t)|^2}$ means averaging the power of impulse response over a region around the receiver (This region is usually small and the reason of averaging is simply to remove the effect of fading).

Based on power delay profile in (2.8), the *mean delay time* is defined as the first moment of $P(t)$, i.e.

$$\overline{t} = \frac{\int tP(t)dt}{\int P(t)dt} = \frac{\int t\overline{|h(t)|^2}dt}{\int \overline{|h(t)|^2}dt}. \tag{2.9}$$

In an static scenario where there is no need to averaging, (2.9) reduces to

$$\overline{t} = \frac{\int tP(t)dt}{\int P(t)dt} = \frac{\int t|h(t)|^2dt}{\int |h(t)|^2dt} \tag{2.10}$$

and the *delay spread* is defined as the standard deviation of the power delay profile, i.e.

$$s = \sqrt{\frac{\int t^2P(t)dt}{\int P(t)dt} - \left(\frac{\int tP(t)dt}{\int P(t)dt}\right)^2} \tag{2.11}$$

or

$$s = \sqrt{\frac{\int t^2\overline{|h(t)|^2}dt}{\int \overline{|h(t)|^2}dt} - \left(\frac{\int t\overline{|h(t)|^2}dt}{\int \overline{|h(t)|^2}dt}\right)^2}. \tag{2.12}$$

Again, for an static scenario we will have

$$s = \sqrt{\frac{\int t^2|h(t)|^2dt}{\int |h(t)|^2dt} - \left(\frac{\int t|h(t)|^2dt}{\int |h(t)|^2dt}\right)^2}. \tag{2.13}$$

Roughly speaking, delay spread in (2.12) is an average time difference between the first and the last versions of the signal arriving at the receiver. More applications and formulations in MIMO theory using delay spread can be found in [41].

2.4 Multi-Input Multi-Output Systems

The basic idea of MIMO systems is using antenna arrays at least at one end of the link[2]. Figure 2.5 demonstrates this scenario. The mathematical analysis of the resultant system shows that MIMO features remarkable benefits.

As shown in the previous sections, fading, shadowing and delay spread are the most important challenges in wireless communications. A primary solution for these problems is using *diversity*. Diversity simply means to diversify the sources of information to be able to overcome the noise and the interference. Although by using more bandwidth, it is possible to have more diversity, the high demand for bandwidth pushes the engineers to develop systems with more diversity in the same bandwidth. MIMO is at least theoretically such a system.

In this section, after a short literature review on MIMO, we will briefly establish the mathematical principles of MIMO and then its information theory background showing that why MIMO leads to a higher capacity efficiency. Also the non-linear power allocation used for MIMO, i.e. water filling, will be reviewed.

2.4.1 Literature Review

Multi-Input Multi-Output systems, which could be viewed as an extension of smart antennas, allow increased capacity efficiency [42] [43] compared to single port antenna systems in the presence of multipath fading environments. In particular when a rich multipath environment is present, i.e. there are many reflections and scattering, MIMO systems provide high capacity with no increase in the bandwidth. It has been shown that MIMO capacity increases linearly with the number of antennas[3] in such an environment [11] [40].[4]

[2]This section has been heavily derived from [41] and [4]

[3]When the number of antennas in the receiver and transmitter are not equal, the capacity is proportional to the smaller of the number of antennas.

[4]This is true when the channel estimation at both the transmitter and the receiver are perfect.

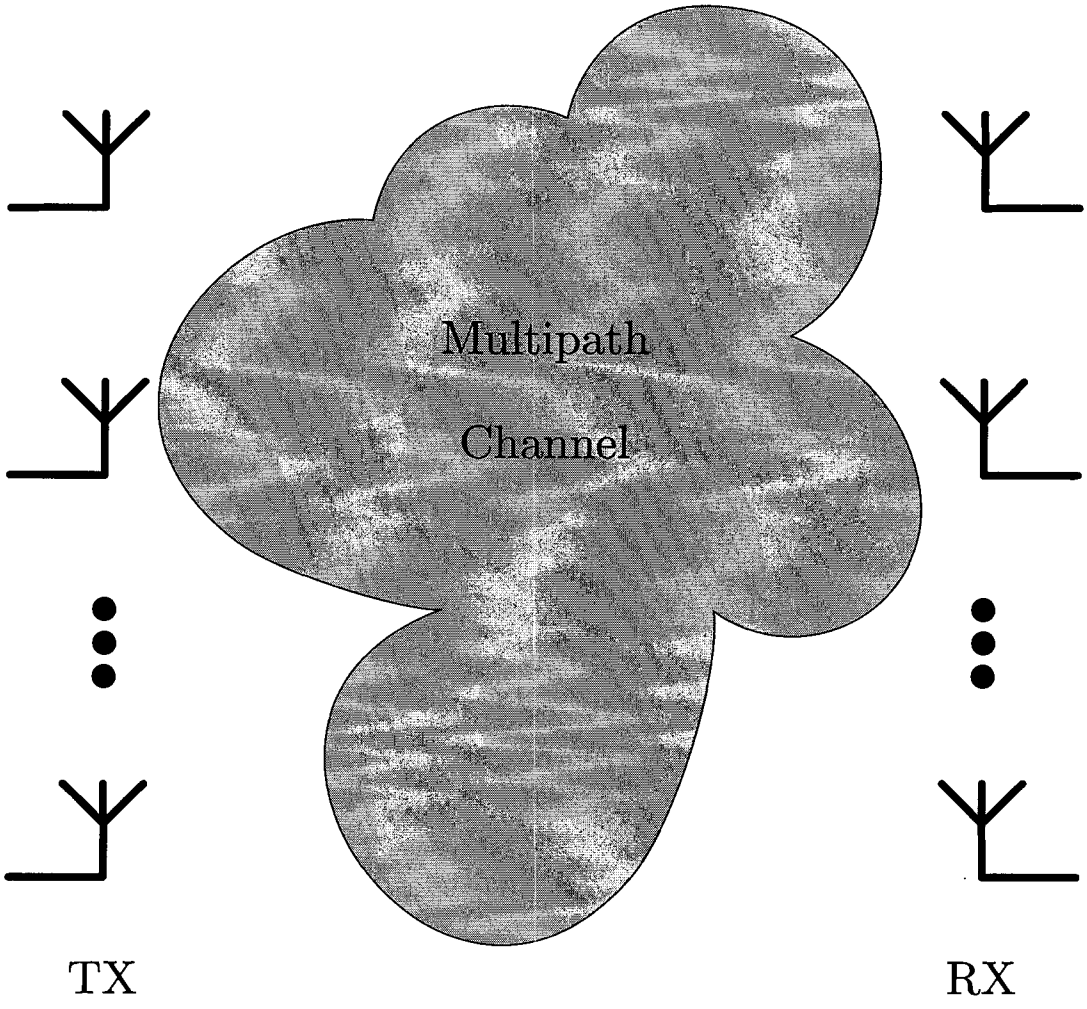

Figure 2.5: Multiple Antennas in Multipath Environment.

The challenging areas in a MIMO system overlap between the propagation channel, the antennas and the signal processing. Outstanding discussions on different aspects of MIMO theory can be found in [4], [13] and also [41].

Usually, MIMO channels are categorized as known and unknown channels. In the case of known channel, there is information about the channel at the transmitter and the transmitter could allocate weights and powers for its different antennas. This is, in fact, the ideal MIMO situation.

For the unknown case, the channel is not known to the transmitter. In this case, the best power allocation strategy is equal power allocation. To increase the capacity efficiency for unknown channels, one suggestion is to use the space-time coding schemes. The analysis in [10] and [11] provides the basis for the BLAST (Bell Laboratories Layered Space-Time) architecture and the space time codes [18] [38]. Some simulations for the performance of BLAST systems (i.e. error probability) could be found in [3] and [23], while the analytical summary is presented in [45] and [46].

The information theoretic approach to find the mean capacity of MIMO is another part of current efforts in expansion of MIMO theory (see e.g. [10] [11] [40] [42] and [34]. The last one introduces the Gaussian approximation for the MIMO capacity distribution). This thesis work in essence lies in this information theoretic part.

The classic view in MIMO theory emphasizes that correlation between the antenna elements in realistic communication channels degrades severely the capacity efficiency of MIMO systems [4] and [32][5]. However, some measurement results show that the capacity degradation might not be as significant as expected [25]. Conversely, Oestges *et. al.* show in their recent works that there are some potential benefits of channel correlation on mean capacity [26] [27]. This fact is the central core of this thesis and will be discussed through introducing a special correlated structure.

2.4.2 Mathematical Principles Of MIMO

Figure 2.6 shows the multi element antenna system and its conversion to a MIMO system through signal processing methods which results in joint data detection at the transmitter[6].

[5] See also [2] [6] [21] [22] [28] [30] and [33]

[6] This section has been heavily derived from [14] and [41]

The nature of MIMO systems is that the data is transmitted over a matrix rather than a vector channel. This provides considerable diversity for the receiver to detect the data. In fact, in an ideal $M \times N$ MIMO system, it is possible to transmit $min(M, N)$ data streams independently and simultaneously over figurative eigenchannels. For an $N \times N$ MIMO system, the high-rate data stream at the transmitter is divided into N data streams with $1/N$-rate and sent to each of the transmitting antennas at the same time. Before broadcasting, these lower-rate data streams are weighted by dynamic weights. After broadcasting, these signals will mix together due to the nature of wireless and mobile channels. At the receiver, similar weights are being applied to the received signal by each of N receiving antennas resulting the detection of $1/N$-rate data streams. Afterwards, putting these lower-rate data streams together will result in the primary high-rate data stream. The mathematical basis for the whole process is similar to the solution of a linear system with N equations and N unknowns.

As will be shown, the weights at the transmitter and the receiver are the eigenvectors of the channel matrix. Theoretically, this will cause each antenna to transmit its data stream over a single scalar channel (i.e. scalar channel for each transmission). However, the extension of MIMO theory for frequency selective fading channels is also available by using coding and signal processing techniques (See [5], [35] and [23] for example).

The maximum number of independent signals which could be transmitted in a MIMO system is the *rank* of MIMO system. This rank is the number of independent equations could be derived from it. Also, this rank is identical to the algebraic rank of the $M \times N$ channel matrix $\boldsymbol{H}$ which is equal to or less than $min(M, N)$. When the rank of a MIMO system is equal to $min(M, N)$, the system is $full - rank$ which means that without coding, the system can give the spectral efficiency expected from the MIMO theory.

To review MIMO information theory, we start with the single-input single-output (SISO) system. For a memoryless 1×1 (SISO) system the capacity efficiency is

$$\frac{C}{B} = E\left\{\log_2\left(1 + \rho|h|^2\right)\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}). \tag{2.14}$$

Here, E means expectation; h is the normalized complex gain of the fixed channel which could be also considered as a realization of a random channel; ρ is the SNR at the transmitter. If there was more than one receiving antenna, i.e. a SIMO system, (2.14) becomes

$$\frac{C}{B} = E\left\{\log_2\left(1 + \rho\sum_{i=1}^{M}|h_i|^2\right)\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}). \tag{2.15}$$

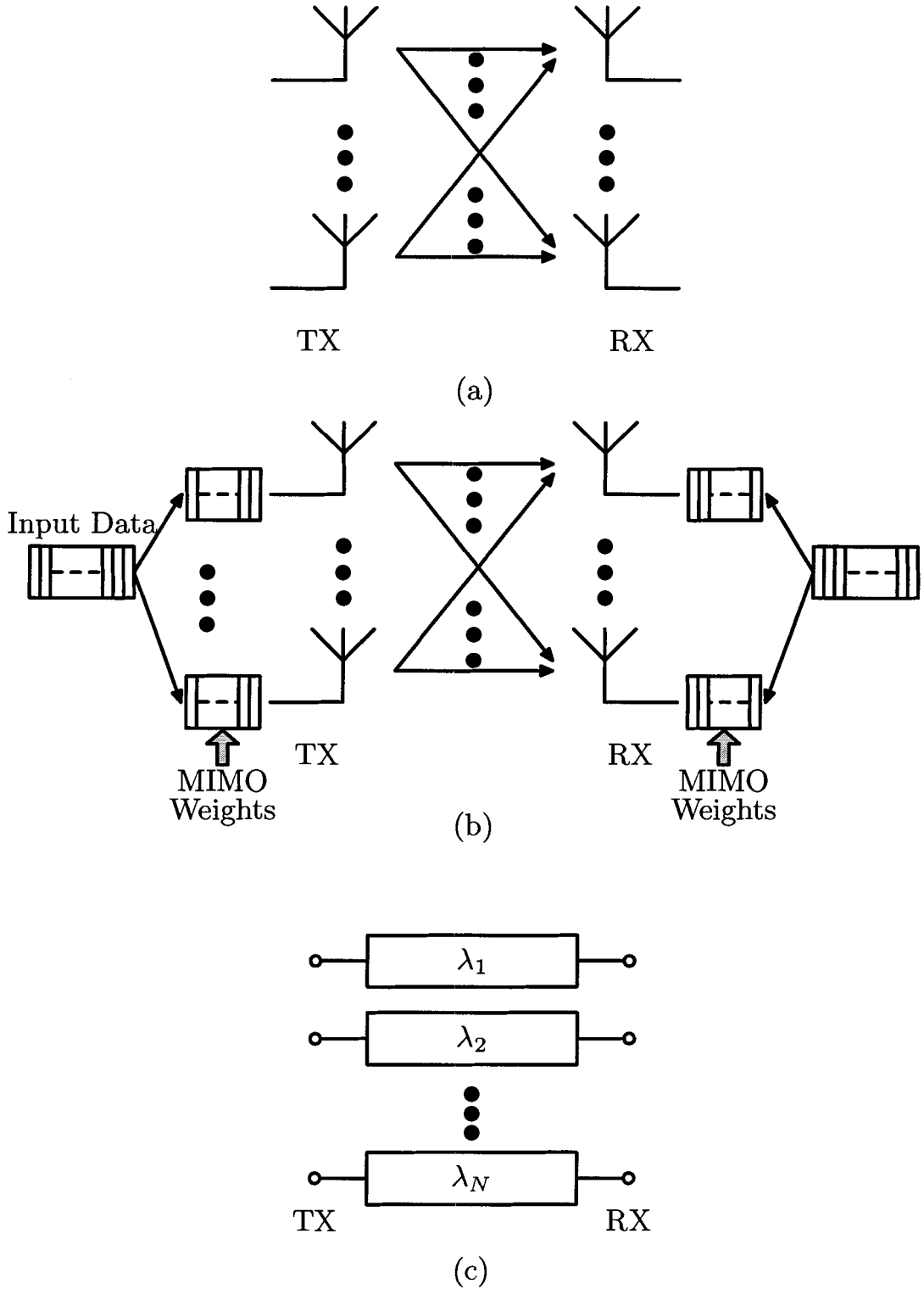

Figure 2.6: (a) Multiple Antennas in Multipath Environment (b) MIMO Signal Processing (c) Equivalent MIMO Eigenchannels.

where M is the number of antennas at the receiver; h_i is the gain for the ith receiving antenna. Note in (2.15) that increasing the number of receiving antennas improves the capacity efficiency in a logarithmic gradient. On the other hand, if we put the diversity at the transmitter where generally there is no information about the channel (MISO system), then the average capacity efficiency will be

$$\frac{C}{B} = E\left\{\log_2\left(1 + \frac{\rho}{N}\sum_{i=1}^{N}|h_i|^2\right)\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}). \tag{2.16}$$

where N is the number of antennas at the transmitter. While in a SIMO system, each receiving antenna sees the channel with the same SNR of ρ which leads to an array gain in the capacity, in MISO this SNR should be divided between the N transmitting antennas and the channel energy can no longer be combined coherently (Recall that there is only one receiving antenna in MISO systems.).

Now, consider having diversity at the both ends of the link, i.e. a MIMO system. The theoretical capacity efficiency limit is [41] [4] [11] [40]

$$\frac{C}{B} = E\left\{\log_2\left[\det\left(\boldsymbol{I}_M + \frac{\rho}{N}\boldsymbol{H}\boldsymbol{H}^H\right)\right]\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}). \tag{2.17}$$

where superscript H means transpose-conjugate; $\boldsymbol{H}$ is the $M \times N$ channel matrix and $\boldsymbol{I}_M$ is the $M \times M$ identity matrix[7]. It has been shown that the capacity efficiency in (2.17) increases linearly with $min(M, N)$. Compared to the logarithmic growth in (2.15) and (2.16), linear growth represents a considerable improvement in the capacity efficiency.

The underlying phenomena for (2.17) is the conversion of one multipath channel to multiple theoretically independent channels. The det operator in (2.17) essentially substitutes the channel matrix $\boldsymbol{H}$ with $min(M, N)$ channels between the transmitting and receiving antennas. On the other hand, properties of the log function will result in the summation of the capacity efficiencies of these $min(M, N)$ channels. This procedure will be completed by imposing unique vectors of weights to each antenna at both ends. These unique weights combine each antenna to create its own eigenchannel. Finally, the multipath channel which mixes the multiple data streams transmitted by N transmitting antennas reduces to $min(M, N)$ independent channels with no interference (theoretically) on each other.

[7]Note that the result in (2.17) is derived under the assumption of equal power uncorrelated sources.

It can be shown that the gain of each of these $min(M, N)$ channels are the eigenvalues of

$$\boldsymbol{G} = \begin{cases} \boldsymbol{H}\boldsymbol{H}^H & \text{for} \quad M \leq N \\ \boldsymbol{H}^H\boldsymbol{H} & \text{for} \quad N > M \end{cases} \tag{2.18}$$

Also, the weights applied at each transmitting and receiving antenna is the eigenvector associated to the eigenvalue of that antenna. According to this, we will call each of the $min(M, N)$ channels an *eigenchannel*.

It is clear that the performance of MIMO systems are sensitive to the distributions of the eigenvalues. Note that $\boldsymbol{H}$ is a random matrix and so, the eigenvalues of $\boldsymbol{G}$ are also random variables. Even the linear growth of the capacity efficiency with the number of antennas depends on these eigenvalues. If there are many small eigenvalues, i.e. many weak eigenchannels, MIMO capacity efficiency degrades severely. However, it is very unlikely to have many small eigenvalues in practice and so, the linear growth in many cases is achievable.

Although (2.17) is the well known formulation for MIMO capacity, it can be rewritten as the function of eigenvalues of $\boldsymbol{G}$ based on the above mentioned argument [41] [40]

$$\frac{C}{B} = E\left\{\sum_{i=1}^{m} \log_2\left(1 + P_i\lambda_i\right)\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}). \tag{2.19}$$

Here, λ_i is the ith eigenvalue of $\boldsymbol{G}$ in (2.18); $m = min(M, N)$ and P_i is the power transmitted to the ith eigenchannel normalized by the noise of channel, so it may be considered as the SNR referred to the transmitter. Since λ_i is the gain of ith eigenchannel, $P_i\lambda_i$ is the SNR at the receiver of the ith eigenchannel. In fact, (2.19) is the sum of single channel capacities which are in the form of Shannon formulation i.e.

$$\frac{C}{B} = \sum \log_2(1 + (\text{SNR at Rx})) \qquad (\text{bits sec}^{-1}\ \text{Hz}^{-1}).$$

The Singular Value Decomposition or SVD builds up another aspect of the MIMO theory. SVD is particularly useful for different interpretations in antenna contexts. Even when $\boldsymbol{H}$ is rectangular, $\boldsymbol{G}$ in (2.18) is a Hermitian matrix in general and will have $m = min(M, N)$ distinct, real positive eigenvalues and the remaining will be zero. The SVD expansion of $\boldsymbol{H}$ itself is

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\,\boldsymbol{V}^H \tag{2.20}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix of real, non-negative singular values. These singular values are equivalent to the square roots of the eigenvalues of $\boldsymbol{G}$. The columns of the orthogonal

matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are the corresponding singular vectors. From (2.20), $\boldsymbol{G}$ can be written as[8]

$$\boldsymbol{G} = \boldsymbol{H}^H \boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Lambda}^H \boldsymbol{U}^H \boldsymbol{U}\boldsymbol{\Lambda}\, \boldsymbol{V}^H = \boldsymbol{V}\boldsymbol{\Lambda}^T \boldsymbol{\Lambda}\, \boldsymbol{V}^H. \tag{2.21}$$

Note that $\boldsymbol{V}$ contains the eigenvectors of $\boldsymbol{G}$. Since $\sqrt{\lambda_i}$ is the singular value of $\boldsymbol{H}$, so

$$\boldsymbol{H}\boldsymbol{V}_i = \sqrt{\lambda_i}\boldsymbol{U}_i \tag{2.22}$$

where $\boldsymbol{V}_i$ is the ith column of $\boldsymbol{V}$ and $\boldsymbol{U}_i$ is the ith column of $\boldsymbol{U}$. If $\boldsymbol{V}_i$ is the transmit weight for the ith transmitting antenna and $\boldsymbol{U}_i^H$ is the receive weight for the ith receiving antenna, the received voltage for this ith eigenchannel will be

$$r_i = \boldsymbol{U}_i^H \boldsymbol{H}\boldsymbol{V}_i = \boldsymbol{U}_i^H \sqrt{\lambda_i}\boldsymbol{U}_i = \boldsymbol{U}_i^H \boldsymbol{U}_i \sqrt{\lambda_i} = \sqrt{\lambda_i}. \tag{2.23}$$

The received power is the square of the received voltage i.e. λ_i. This is the reason why the gain of ith eigenchannel is λ_i. Note again that to have these results, the eigenvectors of $\boldsymbol{G}$ should be applied to the transmitter and the receiver in the forms of $\boldsymbol{V}_i$ and $\boldsymbol{U}_i^H$ respectively.

Finally, it is worth mentioning an important property of $\boldsymbol{G}$ i.e.

$$Trace(\boldsymbol{G}) = \sum_i \lambda_i. \tag{2.24}$$

Since λ_i is the gain of the ith eigenchannel, (2.24) means that the trace of $\boldsymbol{G}$ is the total gain of the MIMO channel.

All of these facts have been used in Figure 2.6. In part (a) of this figure, the real multipath environment with multiple antennas is depicted. In part (b), the signal processing of MIMO is shown. The high-rate input data stream is being fed to MIMO weights of transmitting antennas. These weights are the above mentioned eigenvectors. At the receiver, other eigenvectors are being applied to the received signal leading to the output high-rate data stream. In part (c), the equivalent model of part (b) is illustrated. In fact, all the MIMO signal processing procedures convert (a) to (c) which consists of $min(M, N)$ independent eigenchannels.

[8] Assume $M > N$.

2.4.3 Water Filling Power Allocation

Recall that (2.17) and (2.19) has been derived under the assumption of equal power allocation, i.e. all the eigenchannels receive equal amount of power. It can be shown that knowing the channel, the transmitter can share the power more wisely among the eigenchannels. The optimum power allocation for a MIMO system is a non-linear method which is usually called *water filling* power allocation [[12], Theorem 7.5.1].

Figure 2.7 illustrates the water filling for a 6×6 MIMO system. The first step is determining the threshold D. Note that this threshold does *not* have any physical meaning. D is a mathematical parameter used to determine the power assigned for each of the eigenchannels.

After assigning D, the inverse of the eigenvalues of the matrix $\boldsymbol{G}$ are compared with this threshold. If $1/\lambda_i \geq D$, then the gain of the ith eigenchannel is too small and this eigenchannel will be put away from the communication process (the two last eigenchannels in Figure 2.7 for example). Through this process, the weakest eigenchannels which do not contribute to the communications will be removed. This *weakness* is basically determined by the threshold D. The larger the D is, the more eigenchannels are kept. Note that D itself is dependent on the total available power P and the eigenvalues of $\boldsymbol{G}$.

Assume that the MIMO system is of square dimension i.e. $M = N$ and also

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N. \tag{2.25}$$

Now, consider N' eigenchannels survive the aforementioned cut-off procedure. The power allocated to each of these eigenchannels, P_i, is determined by the water filling rule, i.e.

$$\frac{1}{\lambda_1} + P_1 = \frac{1}{\lambda_2} + P_2 = \ldots = \frac{1}{\lambda_{N'}} + P_{N'} = D. \tag{2.26}$$

The total available power is

$$P = P_1 + P_2 + \ldots + P_{N'}. \tag{2.27}$$

Adding the terms in (2.26) results in

$$N'D = \sum_{i=1}^{N'} \frac{1}{\lambda_i} + \sum_{i=1}^{N'} P_i = \sum_{i=1}^{N'} \frac{1}{\lambda_i} + P \tag{2.28}$$

or

$$D = \frac{1}{N'} \left\{ \sum_{i=1}^{N'} \frac{1}{\lambda_i} + P \right\}. \tag{2.29}$$

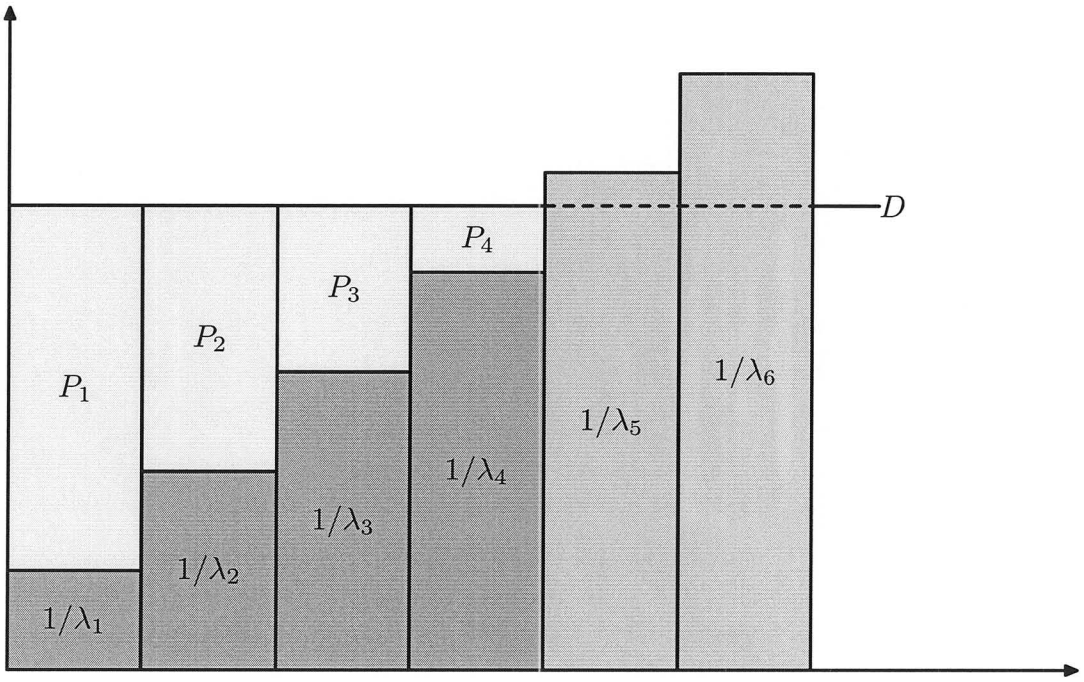

Figure 2.7: Water Filling Strategy.

The threshold D must be determined in an iterative procedure. Finally, as can be seen in Figure 2.7, water filling leads to more power for the stronger eigenchannels and less for the weaker ones.

If we apply the water filling powers to (2.19), it will reduce to

$$\frac{C}{B} = E\left\{\sum_{i=1}^{N'} \log_2\left(1 + P_i\lambda_i\right)\right\} = E\left\{\sum_{i=1}^{N'} \log_2\left(\lambda_i D\right)\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}). \tag{2.30}$$

(2.30) is the theoretical maximum of the capacity efficiency that could be achieved by power allocation in a MIMO system.

Chapter 3

Circulant MIMO Structure

There are different types of *correlated* structures. In fact, although in mathematics and statistics the definition of correlation is unique; in the MIMO context where random matrices exist there could be different definitions for correlation. In the widest sense, by a correlated MIMO structure, we mean a structure in which there is some kind of dependency between the signals coming out.

Two different classes of correlated MIMO structures can be distinguished: systematic and unsystematic. A systematic correlated structure is a structure in which the correlation between the signals is embedded in the channel matrix (i.e. the correlation arises from some random variables being repeated, alone or in linear combinations, in the channel matrix; even if the individual variables values are independent. So, the correlation is distinguishable by just looking into the channel matrix configuration without needing to investigate the elements themselves). In this case, the correlation is in some way fixed or predictable in a sense. On the other hand, an unsystematic correlated structure introduces a random correlation between the signals and does not obey a fixed or pre-defined configuration (i.e. no random variables are repeated; however, propagation conditions may make cause the random variable to become correlated.). So, for this class of correlated structures it is not possible to figure out the correlation just by investigating the channel matrix and we need to

know the elements. Circulant structures are an example of systematic correlated structures:

$$\boldsymbol{H}_{Cir.} = \begin{pmatrix} h_0 & h_1 & \dots & h_{N-1} \\ h_{N-1} & h_0 & \dots & h_{N-2} \\ \vdots & \ddots & \ddots & \vdots \\ h_1 & h_2 & \dots & h_0 \end{pmatrix}.$$

and i.i.d. correlated structure, is an example of unsystematic correlated structure:

$$\boldsymbol{H}_{i.i.d.} = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & \dots & h_{NN} \end{pmatrix}$$

in which h_{ij} s in both matrices are random variables, and elements in i.i.d. have correlation with each other. The claim of this dissertation and specifically this chapter is that some systematic correlated structures lead to a higher capacity efficiency.

Furthermore, among different systematic correlated structures, the circulant structure introduces special advantages. First, the circulant structure is a feasible structure and can have a physical interpretation. Second, the circulant structure is the only structure which has fixed eigenvectors regardless of its elements. These two properties encouraged us to choose it and analyze the different aspects of circulant MIMO structures.

At this point, it should be noted that this section addresses the capacity of a circulant channel and associated signal processing. It is emphasized that the feasibility of a real-world circulant description is separated from the capacity treatment. Nevertheless, some comments are offered regarding cable structures as circulant.

3.1 Circulant Channel Matrix

The MIMO channel matrix is the usual linear model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{3.1}$$

where $\boldsymbol{x}$ is the input to the system and $\boldsymbol{y}$ is the output and $\boldsymbol{n}$ is noise. An $N \times N$ completely random or i.i.d. matrix is

$$\boldsymbol{H}_{i.i.d.} = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & \dots & h_{NN} \end{pmatrix} \tag{3.2}$$

where h_{ij}s are complex (baseband) Gaussian random variables. On the other hand, an $N \times N$ circulant Gaussian matrix is

$$\boldsymbol{H}_{Cir.} = \begin{pmatrix} h_0 & h_1 & \dots & h_{N-1} \\ h_{N-1} & h_0 & \dots & h_{N-2} \\ \vdots & \ddots & \ddots & \vdots \\ h_1 & h_2 & \dots & h_0 \end{pmatrix}. \tag{3.3}$$

Again all the elements are complex Gaussian random variables. From (3.3), the meaning of systematic correlated structure is clear. Here, regardless of the values of h_{ij}, the channel matrix $\boldsymbol{H}_{Cir.}$ clearly shows that there is correlation in the system. It has been shown that a circulant random matrix in (3.3) has the general eigenvalue solution of

$$\lambda_k = h_0 + h_1\rho_k + h_2\rho_k^2 + \ldots + h_{N-1}\rho_k^{N-1} \tag{3.4}$$

in which ρ_k is the kth complex roots of unity, i.e. $\rho_k^{-N} = 1$ (See [16] for example).The corresponding eigenvector is

$$\boldsymbol{V}_k = N^{-1/2}(1, \rho_k, \rho_k^2, \ldots, \rho_k^{N-1})^T \tag{3.5}$$

If $\boldsymbol{H}_{Cir.}$ is a circulant structure, then $\boldsymbol{G}$ in (2.18) is also circulant. So by having a circulant link, the eigenvectors of $\boldsymbol{G}$ are independent of its eigenvalues, i.e. they are fixed.

Note that in all of the results in this thesis, the elements of different random matrices are all complex variables consisting of independent real and imaginary parts which are zero-mean unit-variance real Gaussian random variables.

It turns out to be difficult to arrange a wireless MIMO link (random i.i.d.) to closely approximate a circulant. However, for a wired MIMO link where there is more access to arranging the physical medium parameters, a circulant structure is more feasible. Some analysis and measurement of DSL are given by [15], [39] and [37]. It is evident that an

arrangement of wires (or wire pairs) shown in Figure 3.1 ideally leads to a circulant channel matrix.

For a multi-tier cable, each tier could be arranged and so considered as a 'sub-MIMO'

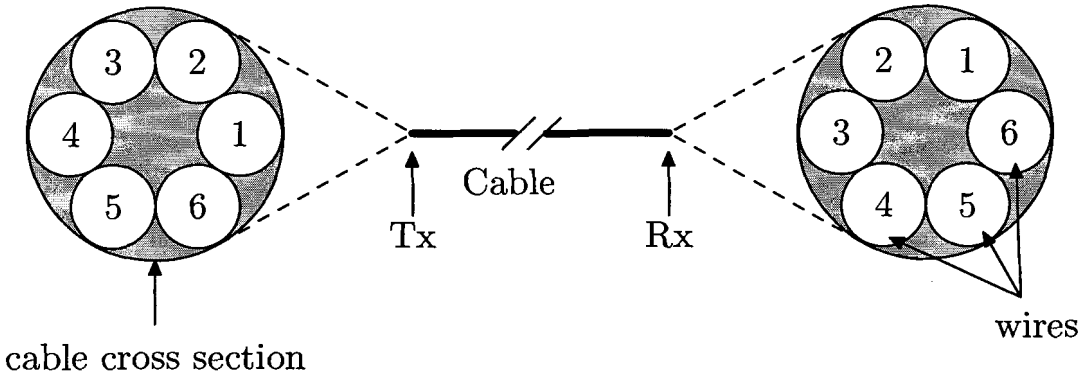

Figure 3.1: Single-tier circulant cable.

circulant system. It can be shown that a combination of sub-MIMO systems comprises an efficient MIMO system if the sub-MIMO systems are well isolated from each other. Figure 3.2 illustrates such an isolated circulant sub-MIMO tier.

Theoretical studies of the capacity of MIMO systems show how correlation degrades MIMO capacity under the assumption of certain channel models [1] and [44][1]. In fact, as we mentioned before, this is a widely accepted characteristic about the correlation in MIMO systems. However some measurement results show that the capacity degradation might not be as significant as expected [25]. Actually, Oestges *et. al.* [26][27] show a potential benefit of finite correlation.

Generally, in wireless MIMO literature , e.g. [24], by a correlated structure we mean an unsystematic correlated structure. In this class, correlation exists among the antennas (either at the transmitter or at the receiver). So the columns (or rows) of $\boldsymbol{H}$ are independent random vectors, but the elements of each column are correlated with each other and have the same mean and the same covariance matrix. So, it is not possible to extract the correlation pattern unless you have the values of the channel matrix elements, h_{ij}s. For the case of

[1]See also [21], [30], [6] and [33]

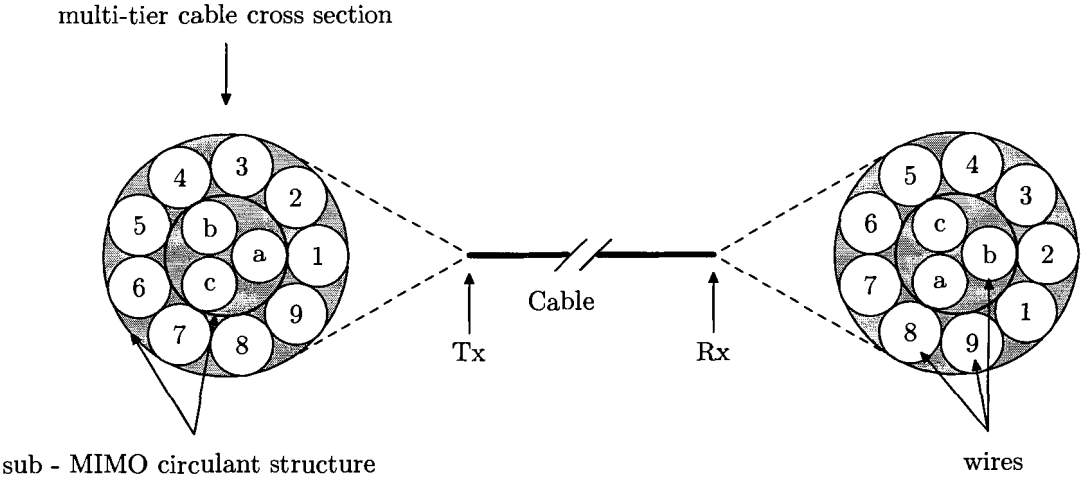

Figure 3.2: Multi-tier cable with sub-MIMO circulant structures.

Rayleigh fading, this implies

$$E\{\boldsymbol{h}_j\} = \boldsymbol{0} \tag{3.6}$$

and the correlation matrix is

$$\boldsymbol{\Sigma} = E\left\{\boldsymbol{h}_j \boldsymbol{h}_j^{\dagger}\right\} \tag{3.7}$$

for $j = 1, \ldots, N$ where $\boldsymbol{h}_j$ is the jth column (or row in case the correlation happens at the transmitter) of $\boldsymbol{H}$. Referred to the above definition, the circulant matrix, $\boldsymbol{H}_{Cir.}$ in (3.3) is not a correlated MIMO structure because its columns (and also its rows) are not independent random vectors. So, with the circulant $\boldsymbol{H}_{Cir.}$, we cannot impose a general correlation structure in $\boldsymbol{\Sigma} = E\left\{\boldsymbol{h}_j \boldsymbol{h}_j^{\dagger}\right\}$ for $j = 1, \ldots, N$ as with the i.i.d. channels. However, the circulant structures show the same potential beneficial impacts on capacity as the standard correlated channels.

3.2 Capacity Advantages of Circulant Channel

Before investigating the capacity advantages of the circulant structure, it would be insightful to see the general capacity behavior of an ideal i.i.d. model. Figure 3.3 demonstrates such

an $N \times N$ MIMO model with different number of transmitting and receiving antennas, N. Note that in this result water filling has been used as the optimum power allocation.

Recall that in an ideal MIMO system, the eigenchannels are perfectly separated from

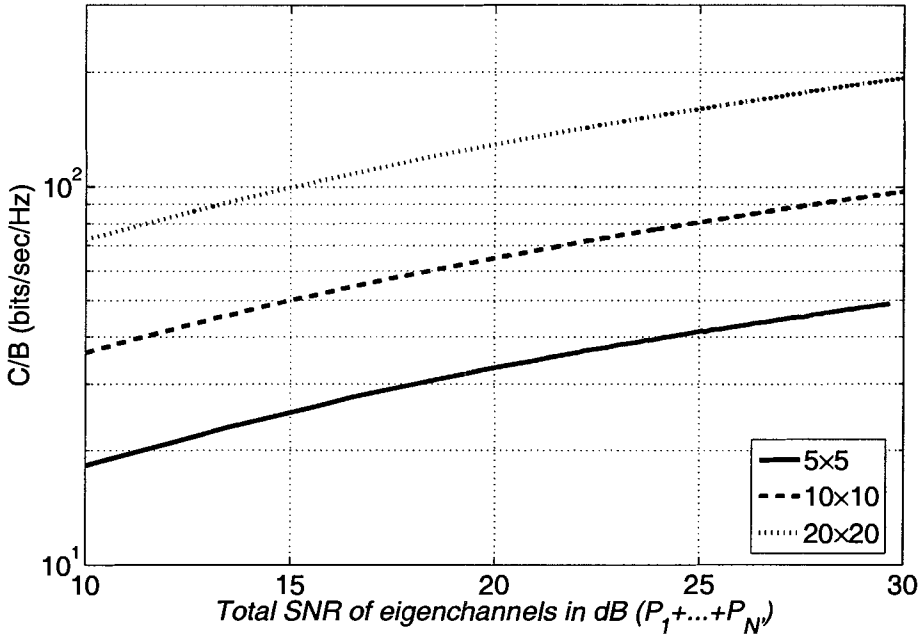

Figure 3.3: MIMO Capacity For Different $N \times N$ i.i.d. MIMO Systems.

each other. This means that theoretically, by sending more power at the transmitter, we will have higher capacity and this is an unbounded increase. In practice, however, because the ideal separation is not achievable, there is some power leakage between the eigenchannels. This power leakage will saturate the capacity efficiency bounding the capacity increased with the transmit power. This leakage will be specially noticeable when there is a large amount of power being sent through the channel. In this case, the perfect separation of different eigenchannels is very difficult.

Also, note that to plot Figure 3.3 we applied (2.30) i.e.

$$\frac{C}{B} = E\left\{\sum_{i=1}^{N'} \log_2\left(1 + P_i\lambda_i\right)\right\} \quad (\text{bits sec}^{-1}\ \text{Hz}^{-1}).$$

where P_is are the normalized powers (or SNRs) found from water filling rule and λ_is are the eigenvalues or the gain of the eigenchannels.

The Shannon capacities for the circulant random channels can be also found from (2.30). Figure 3.4 demonstrates the simulation results for the two models. Like Figure 3.3, the ordinate is $\sum_{N'} P_i$ or total input SNR. The receiver noise is assumed here to be equal for all the eigenchannels. So, the ordinate can be considered as the total SNR for the parallel eigenchannels. The MIMO dimension for this simulation is 20×20.

For small SNRs, the capacity of the completely random channel is higher than the

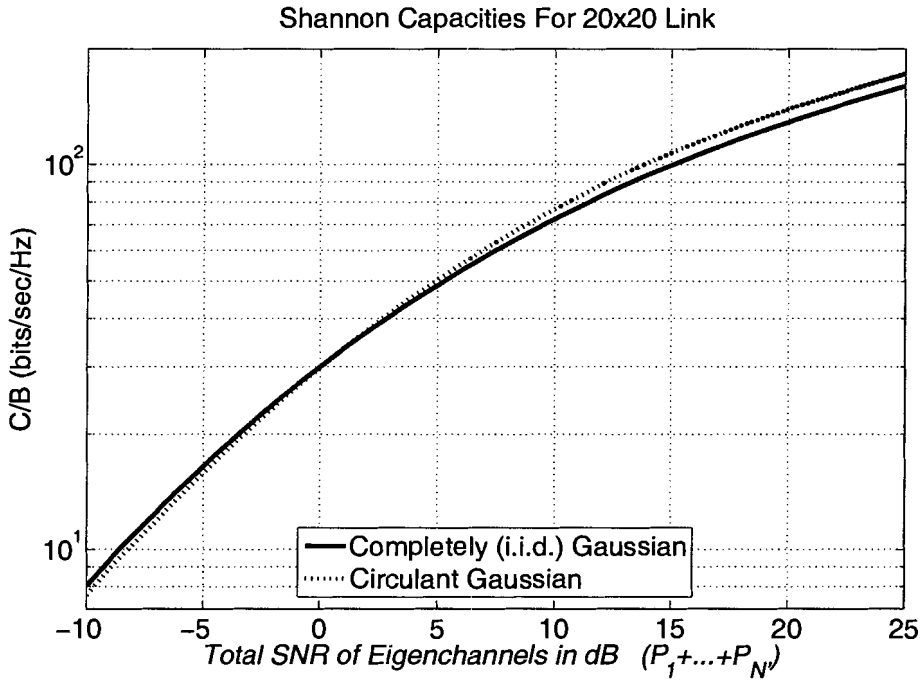

Figure 3.4: The Shannon capacity of the two 20×20 channel types versus the summed SNRs of the eigenchannels referred to the transmitter.

circulant one. For large SNRs, it is vice versa. This shows that despite the notion about destructive effects of correlation, it might have some advantages. Note that in Figure 3.4 the abscissa is in logarithmic scale, so the improvement in the capacity using the circulant structure is considerable, for example, at SNR=20 dB, the extra capacity efficiency is about 10 bits/sec/Hz. From Figure 3.4, the circulant begins to perform better for an SNR above 2 or 3 dB for 20×20 MIMO system.

This interesting behavior basically arises from the distribution of the eigenvalues of the random matrices. Although MIMO systems are usually considered from the space-time coding point of view, this chapter investigates another aspect of MIMO capacity through exploration of the eigenvalues and their pdfs. These investigations which establish the theoretical basis for the abovementioned behavior will be presented later.

Applying digital techniques will degrade the capacities from the theoretical Shannon limit to the practicable possibilities of a digital link. Figure 3.5 shows the effect of using QAM techniques on the single link capacity, along with the Shannon limit. As can be seen from this Figure, in the presence of digital techniques, the capacity does have a limit. While the Shannon limit increases by increasing the power, the QAM-included capacity has a saturation level. From this level, increasing the power will not result in a higher capacity.

There are well-known probabilities of bit error for the uncoded QAM family ([29] and [31]). Since M-QAM is equivalent to $\sqrt{M}$-PAM in the context, the PAM symbol error probability can be used to find the probability of symbol error. Following [29]

$$P_{\sqrt{M}} = 2(1 - \frac{1}{\sqrt{M}})Q\left(\sqrt{\frac{3\log_2 M\gamma_b}{M-1}}\right) \tag{3.8}$$

in which $P_{\sqrt{M}}$ is the probability of symbol error for the $\sqrt{M}$-PAM and γ_b is the SNR per bit. (3.8) could be rewritten as

$$P_{\sqrt{M}} = (1 - \frac{1}{\sqrt{M}})erfc\left(\sqrt{\frac{3\log_2 M\gamma_b}{2(M-1)}}\right). \tag{3.9}$$

So, for a M-QAM, the probability of symbol error could be written as

$$P_M = 1 - (1 - P_{\sqrt{M}})^2. \tag{3.10}$$

To obtain average bit error rate (BER), the usual approximation will be used

$$P_b \approx \frac{P_M}{\log_2 M} \tag{3.11}$$

in which P_b is the probability of a bit error. Figure 3.5 illustrates the results for the commonly used QAM techniques. Note that to find the capacity efficiencies from the BERs, the fixed block length of $L = 500$ is assumed. The technique here otherwise offers an approach to study the impact of the block length, which is not available from the information theoretic equation (3.14). The results are not sensitive to modest variations of this block

length. So, the throughput of correct symbols becomes the practicable capacity and the capacity efficiency is written [31]

$$\frac{C}{B} = \log_2 M(1 - \text{BER(SNR)})^L \tag{3.12}$$

where BER is the bit error rate. So having BER from (3.11) and using it in (3.12), we are able to find the degradation caused by digital modulation techniques. The capacity of Figure 3.5 is actually the throughput of correct bits, and the BER is not specifically defined.

The 7-10dB minimum capacity penalty resulting from using uncoded QAM is clear (See Figure 3.6). This can be reduced by standard Forward Error Correction (FEC).

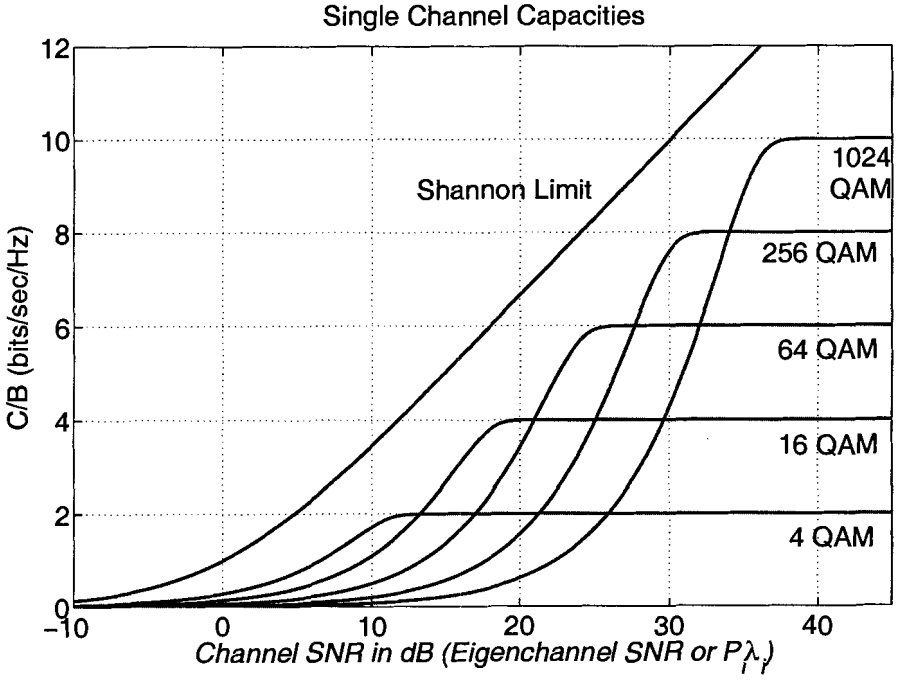

Figure 3.5: Single channel Shannon capacity and limits using QAM versus the SNR at the receiver.

The above argument mainly belongs to the communications theory. However, the degradation caused by using QAM techniques can also be considered from information theory

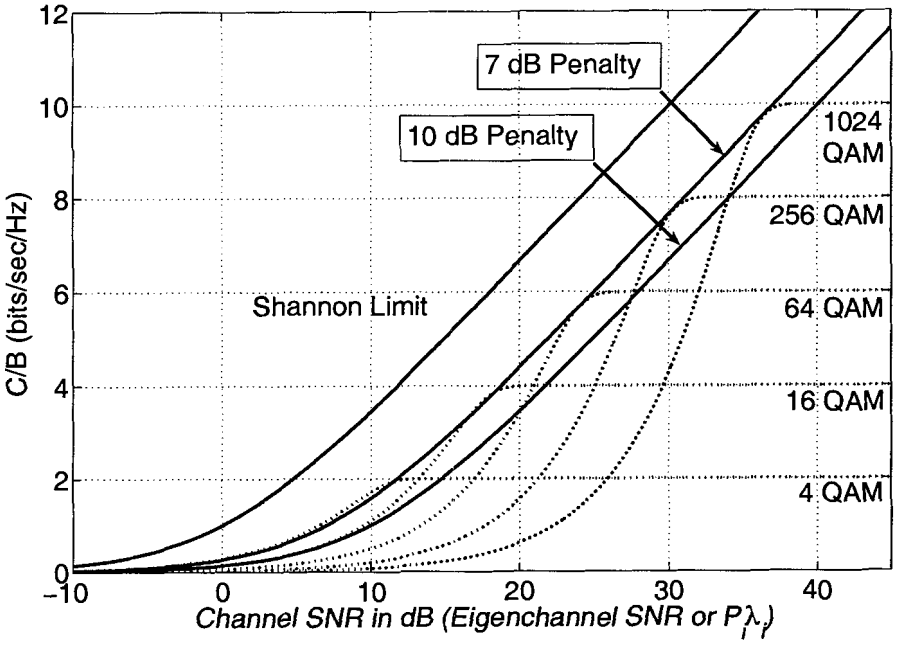

Figure 3.6: Single channel capacity penalties for QAM versus the SNR at the receiver.

point of view. To do this we can use the mutual information, $I(\boldsymbol{y};\boldsymbol{x})$, for the linear model[2]:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}. \tag{3.13}$$

In which $\boldsymbol{y}$ is the output, $\boldsymbol{x}$ is the input and $\boldsymbol{n}$ is the additive white Guassian noise. We will calculate the mutual information for the Gaussian input for different QAM methods to compare with each other and also to the Shannon capacity. Assuming $\sigma_n^2 = 1$, the SNR is $\gamma = \sigma_x^2/\sigma_n^2 = \sigma_x^2$. The Gaussian capacity is

$$C_G(\gamma) = \frac{1}{2}\log_2(1+\gamma). \tag{3.14}$$

The noise pdf is

$$P_n(n) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}n^2\right). \tag{3.15}$$

So, the output pdf, given the input is

$$P_{y|x}(y|x) = P_n(n) = P_n(y-x). \tag{3.16}$$

[2]This section has been heavily derived from Professor Cavers' notes.

To find the mutual information for the M-QAM system, we use the fact that in terms of information theory, the M-QAM is equivalent to two dimensional $\sqrt{M}$-PAM. So we should first find the mutual information for the M-PAM. The input for M-PAM system is

$$\boldsymbol{x} = (..., -3, -1, +1, +3, ...)\Delta \tag{3.17}$$

where the half level spacing, Δ, is determined so that the input variance is γ, i.e.

$$\Delta = \sqrt{\frac{3\gamma}{M^2 - 1}}. \tag{3.18}$$

The constellation for M-PAM is defined with the input variables

$$x_k = \Delta(-M + 1 + 2k) \quad for\ k = 0 : M - 1 \tag{3.19}$$

The average mutual information for M-PAM is then equal to

$$I(\boldsymbol{y}, \boldsymbol{x}) = -\sum_{k=0}^{M-1} \int_{-\infty}^{+\infty} P_{xy}(x_k, y) \log_2 \frac{P_x(x_k)}{P_{x|y}(x_k|y)} dy =$$
$$\sum_{k=0}^{M-1} \int_{-\infty}^{+\infty} P_{y|x}(y|x_k) P_x(x_k) \log_2 \frac{P_{y|x}(y|x_k) P_x(x_k)/P_y(y)}{P_x(x_k)} dy \tag{3.20}$$

Assuming $P_x(x_k) = \frac{1}{M}$ for $k = 0, \ldots M - 1$, we will have

$$I(\boldsymbol{y}, \boldsymbol{x}) = \frac{1}{M} \sum_{k=0}^{M-1} \int_{-\infty}^{+\infty} P_{y|x}(y|x_k) \log_2 \frac{P_{y|x}(y|x_k)}{P_y(y)} dy =$$
$$\frac{1}{M} \sum_{k=0}^{M-1} \int_{-\infty}^{+\infty} P_{y|x}(y|x_k) \log_2 \frac{P_{y|x}(y|x_k)}{\sum_{i=0}^{M-1} P_{y|x}(y|x_i) P_x(x_i)} dy =$$
$$\frac{1}{M} \sum_{k=0}^{M-1} \int_{-\infty}^{+\infty} P_{y|x}(y|x_k) \log_2 \frac{M P_{y|x}(y|x_k)}{\sum_{i=0}^{M-1} P_{y|x}(y|x_i)} dy \tag{3.21}$$

Using (3.16), the mutual information for M-PAM is

$$I(\boldsymbol{y}, \boldsymbol{x}) = \frac{1}{M} \sum_{k=0}^{M-1} \int_{-(M-1)\Delta-3}^{(M-1)\Delta+3} P_n(y - x_k) \log_2 \frac{M P_n(y - x_k)}{\sum_{i=0}^{M-1} P_n(y - x_i)} dy. \tag{3.22}$$

The integral limits are selected to approximate $+\infty$ and $-\infty$. Having the mutual information for M-PAM, we can find the mutual information for M^2-QAM by simply multiplying the M-PAM by a factor of 2. Also, the total amount of power used for M^2-QAM is twice the power used for each of M-PAM systems. The result (for the total γ) is shown in Figure 3.7.

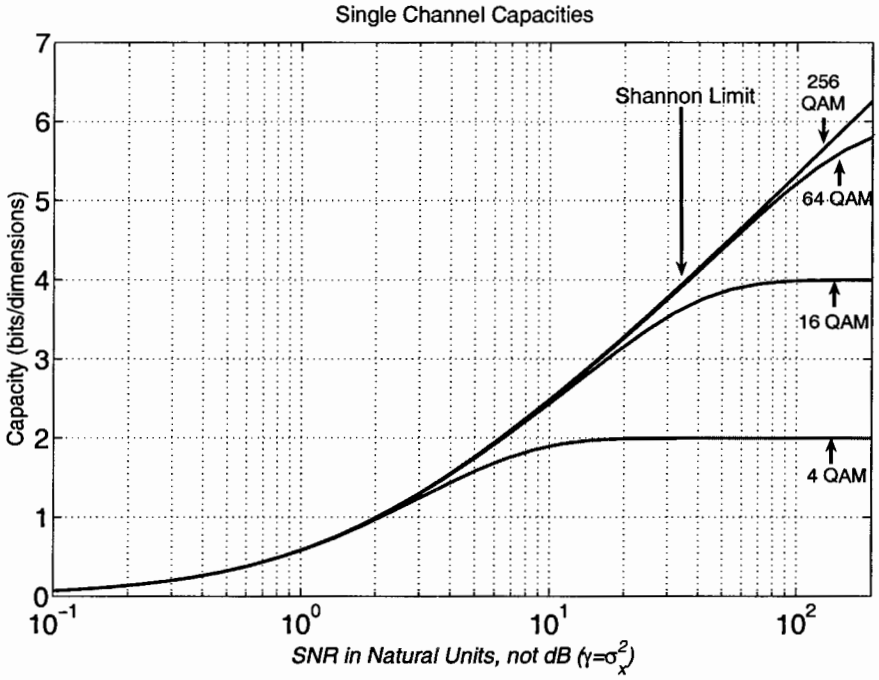

Figure 3.7: Single channel capacity penalties in mutual information for QAM versus the γ.

The channel SNR ($P_i\lambda_i$)	*The technique to be used*
Larger than 27.6 (dB)	256-QAM
Between 20.9 and 27.6 (dB)	64-QAM
Between 13.2 and 20.9 (dB)	16-QAM
Less than 13.2 (dB)	4-QAM

Table 3.1: The digital technique allocation plan

Note that the Shannon limit is similar to the mutual information of 256-QAM for the range of SNRs shown in this figure.

Suppose we choose the QAM for each eigenchannel in a standard optimum way so that we remain as close as possible to the Shannon limit, i.e. the most efficient digital technique map for our MIMO system. Table 3.1 illustrates this QAM selection. Since 1024-QAM is very difficult to implement owing to phase noise limitations we remove it from our consideration.

The capacities of the 20×20 link for the circulant and the completely random cases are shown in Figure 3.8. Again, for small SNRs, the completely random Gaussian link features an advantage whereas for large SNRs, the circulant structure is better.

To investigate the underlying basis for this behavior, in the following section a broad

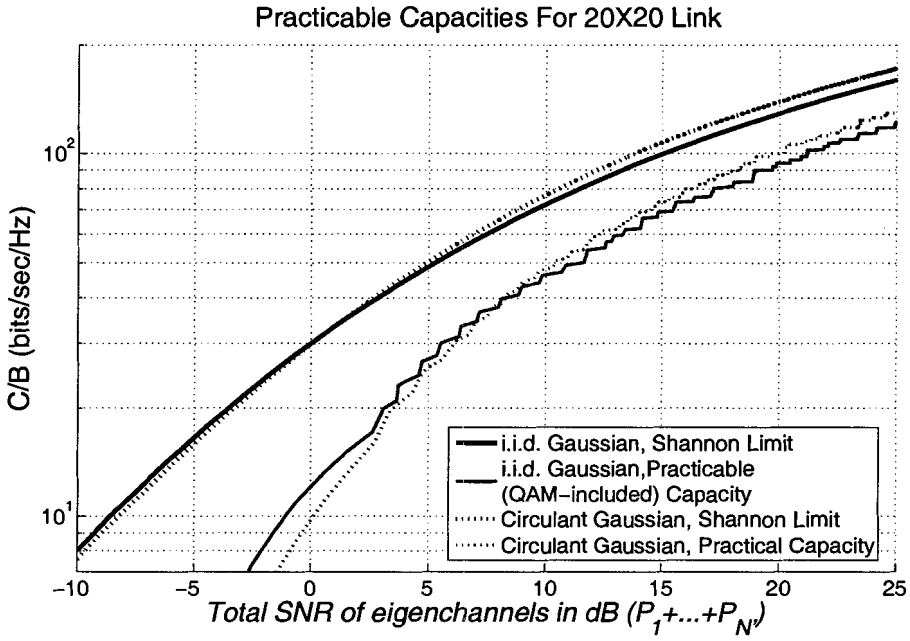

Figure 3.8: Practicable capacity of the two channel types for 20×20 systems using QAM and optimum constellation allocation.

inspection of the distributions of the eigenvalues of the channel matrices is presented. Going through these eigenvalues shows their importance in affecting the MIMO capacity and gives an insight about how this effect takes place.

3.3 Investigation of The Empirical pdfs of The Two Models

The results for the eigenvalues of random matrices (e.g., see [20], [17] and [8]), are for asymptotically large dimensions. Here, we have small matrices compared to asymptotically large numbers. Also, the pdfs of the *sorted* eigenvalues are required here. So instead of an analytical approach, which is quite difficult, a numerical histogram method is used to estimate the pdfs of the ordered eigenvalues of the circulant and completely random (i.i.d.) channel, for a 5×5 system. Note that by eigenvalues we mean the eigenvalues of the Gram matrix $\boldsymbol{G}$ in (2.18).

In Figure 3.9, the pdfs of the ordered eigenvalues of the two channel types are illustrated. The pdfs of the smallest eigenvalues are very similar to the exponential pdf, as expected, while the rest of pdfs are more comparable to, and so could be modeled by, the log-normal pdf.

From the pdfs of the smallest eigenvalues, it is clear that the smallest eigenchannel in the completely random (i.i.d.) structure is weaker than that of the circulant. This is the reason why, for small SNRs, the completely random structure features better capacity. The key factor here is the water filling power allocation cut-off feature. Note that by using water filling, for small SNRs, the weakest eigenchannels are omitted. So in the completely random (i.i.d.) case where the smallest eigenchannels are usually, i.e. in probability, smaller than the circulant ones, and also smaller than the water filling cut-off threshold (see [41]), the capacity is better. In this case, the power is allocated to the stronger eigenchannels only. In the circulant case, conversely, the same power is distributed between not only the stronger eigenchannels, but also some of the weaker ones which contribute little to the total capacity.

While the smallest eigenvalues play the more important role for small SNRs, in the more practical range of SNRs, i.e. for larger SNRs, the similarity of the eigenvalues is the major cause in determining the MIMO capacity. To illustrate this, Figure 3.10 shows the mean of the ordered eigenvalues for the two channel types. It is obvious that the smallest eigenvalues of the completely random channel are smaller than the circulant ones and the largest eigenvalue of the completely random channel is larger than those of the circulant. The eigenvalues of the circulant channel are *more similar* to each other compared to the eigenvalues of the completely random channel.

To clarify the effect of similar eigenvalues on the capacity, we should investigate the information theory aspect of MIMO theory. In fact, the familiar idea about the capacity of

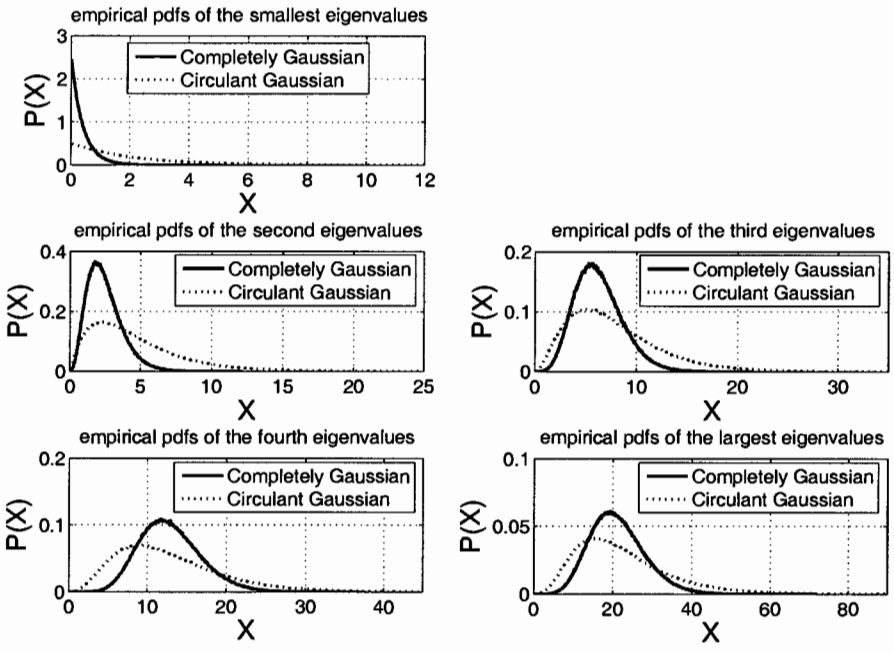

Figure 3.9: The empirical pdfs of the ordered eigenvalues of two 5×5 links.

a MIMO system is that equal eigenchannels leads to maximum capacity. However, this is not always true. As we will show in the following part, only under certain constraints and assumptions, do equal eigenchannels lead to maximum capacity.

Hypothesis: *To maximize the capacity of a MIMO channel when the total power is constant but large enough to use all the eigenchannels through the water filling strategy, all of the eigenvalues should be equal.*

Essentially, this hypothesis is very similar to a proved theorem in information theory about the maximum capacity of parallel channels (See [7] for example). That theorem says the maximum capacity of the parallel channels when there are some constraints on the total available power as well as on the total gain of the channels, is when there are equal channel gains and also equal input powers for the different channels.

However, the MIMO version of this theorem is not unique. It depends on how we define the constraints of the optimization problem. We will clarify the set of constraints for our problem showing that there are slight differences between the MIMO problem and the

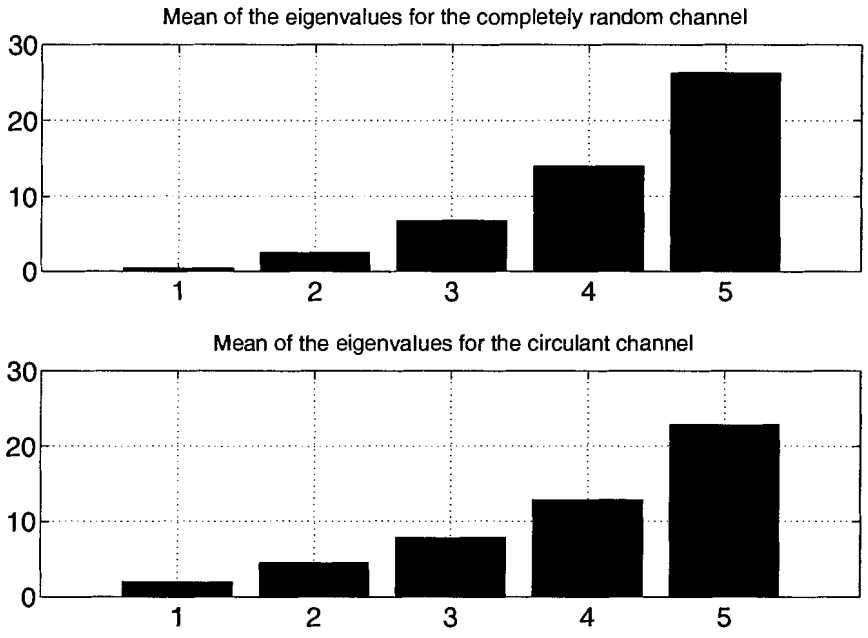

Figure 3.10: Mean of the eigenvalues for the two types of channels.

information theory one in terms of the constraints. Then we will investigate the problem of maximum capacity efficiency for the MIMO systems with water filling showing that this is an interesting open problem. Note that in the investigation we assume that we can determine the eigenvalues of the channel in order to find the combination of the eigenvalues which leads to the maximum capacity. Although this assumption is not necessarily a valid physical assumption (because the eigenvalues and also the random elements of the channel matrix are not controllable), it is necessary for the analysis.

The basic optimization problem is

$$\max\left\{\sum_i \log_2(1+P_i\lambda_i)\right\} \tag{3.23}$$

or equivalently,

$$\max\left\{\log_2(\prod_i(1+P_i\lambda_i))\right\}. \tag{3.24}$$

Since $\log_2(x)$ is an increasing function when $x \geq 1$ and since $1 + \lambda_i P_i \geq 1$ and so $\prod_i (1 + \lambda_i P_i) \geq 1$, it would be enough to maximize the argument of the $\log_2$ function in (3.24). So the problem now reduces to

$$\max_{\lambda_i, P_i} \left\{ \prod_i (1 + P_i \lambda_i) \right\}. \tag{3.25}$$

The first obvious constraint is on the total available power, i.e.

$$\sum_i P_i = P \tag{3.26}$$

in which P_i is the power (or SNR) for the ith eigenchannel and P is the total available power. Also, there is a series of equality constraints arising from the water filling strategy in the MIMO links:

$$\frac{1}{\lambda_1} + P_1 = \frac{1}{\lambda_2} + P_2 = \ldots = \frac{1}{\lambda_N} + P_N = D \tag{3.27}$$

in which D is the cut-off threshold of water filling (See section 2.4.3 and also [41]) and λ_i is the ith eigenvalue. It is important to note that D is *not* a fixed parameter for the different MIMO channels (Although it is a fixed parameter for different eigenchannels in a MIMO channel.). This means that having the same amount of power at the transmitters of different MIMO channels, the value of D may differ from one MIMO system to another depending on their eigenvalues.

Also, we can rewrite the optimization problem in (3.25) according to D:

$$\max_{\lambda_i, P_i} \left\{ \prod_i (1 + P_i \lambda_i) \right\} = \max_{\lambda_i} \left\{ D^N \prod_i \lambda_i \right\} \tag{3.28}$$

To remove D from (3.28) we add up the terms in (3.27)

$$\sum_i \frac{1}{\lambda_i} + \sum_i P_i = ND \tag{3.29}$$

so

$$D = \frac{1}{N} \left\{ \sum_i \frac{1}{\lambda_i} + \sum_i P_i \right\} = \frac{1}{N} \left\{ \sum_i \frac{1}{\lambda_i} + P \right\}. \tag{3.30}$$

Note that (3.30) clearly shows how D is being determined based on the total available power, P, and the eigenvalues of the channel matrix in a recursive process. Using (3.30) the

equality constraints are

$$\frac{1}{\lambda_j} + P_j = \frac{1}{N}\left\{\sum_i \frac{1}{\lambda_i} + \sum_i P_i\right\} \quad for\ j = 1, \ldots, N. \tag{3.31}$$

Note that (3.26) and (3.31) comprise insufficient constraints for (3.25) to have an answer (we can have as large λ_is as we want without violating these constraints). So, we need a constraint on the sum of the eigenvalues, i.e.

$$\sum_i \lambda_i = K \tag{3.32}$$

for some positive constant K. Recall that the eigenvalues are the gains of the eigenchannels, so (3.32) means the total gain of the MIMO link is some finite constant. This is a reasonable assumption for the optimization problem on the capacity of MIMO systems.

It can be shown that equal eigenvalues are always the stationary points for the optimization problem in (3.25) with constraints in (3.26), (3.31) and (3.32). To show these, we can use the Lagrange multiplier method. From (3.30), the optimization problem in (3.28) can be written as

$$\max_{\lambda_i}\left\{D^N \prod_i \lambda_i\right\} = \max_{\lambda_i}\left\{\left(\frac{1}{N}\right)^N \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^N \prod_i \lambda_i\right\}, \tag{3.33}$$

or simply

$$\max_{\lambda_i}\left\{\left\{\sum_i \frac{1}{\lambda_i} + P\right\}^N \prod_i \lambda_i\right\}. \tag{3.34}$$

With the constraint in (3.32), the Lagrange function is

$$h(\lambda_i) = \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^N \prod_i \lambda_i - A\left\{\sum_i \lambda_i - K\right\}. \tag{3.35}$$

where A is the Lagrange multiplier. To find the extremums of h, we should solve

$$\frac{\partial h}{\partial \lambda_i} = 0 \qquad for\ i = 1, \ldots, N, \tag{3.36}$$

or

$$A = \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^N \frac{\prod_i \lambda_i}{\lambda_i} - \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^{N-1} \frac{2\prod_i \lambda_i}{\lambda_i^2} \qquad for\ i = 1, \ldots, N. \tag{3.37}$$

This implies that, for $i \neq j$

$$\left\{\sum_i \frac{1}{\lambda_i} + P\right\}^N \frac{\prod_i \lambda_i}{\lambda_i} - \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^{N-1} \frac{2\prod_i \lambda_i}{\lambda_i^2} = \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^N \frac{\prod_i \lambda_i}{\lambda_j} - \left\{\sum_i \frac{1}{\lambda_i} + P\right\}^{N-1} \frac{2\prod_i \lambda_i}{\lambda_j^2} \tag{3.38}$$

or

$$\frac{R}{\lambda_i} - \frac{2}{\lambda_i^2} = \frac{R}{\lambda_j} - \frac{2}{\lambda_j^2} \tag{3.39}$$

where $R = \left\{\sum_i \frac{1}{\lambda_i} + P\right\}$. It is evident that $\lambda_i = \lambda_j$ is an answer to (3.39). So, equal eigenvalues are always the stationary point for the capacity efficiency in the MIMO system. However, it is very difficult to determine if they are associated with the global maximum, the local maximum or even the minimum. In fact, the optimization problem in (3.25) with constraints in (3.26), (3.31) and (3.32) does not have a unique answer because its behavior is not always convex. Depending on P in (3.26) and K in (3.32), the answer to (3.25) could be equal or non-equal eigenvalues. So, the classic optimization methods such as Lagrange multipliers method can not be used to find its global maximums.

To understand this behavior, the capacity of a 2×2 MIMO link with these constraints are given in Figures 3.11 (convex) and 3.12 (concave). For Figure 3.11

$$\begin{aligned} \lambda_1 + \lambda_2 &= 2 \\ P_1 + P_2 &= 2 \\ \frac{1}{\lambda_1} + P_1 &= \frac{1}{\lambda_2} + P_2 \end{aligned} \tag{3.40}$$

and for Figure 3.12

$$\begin{aligned} \lambda_1 + \lambda_2 &= 5 \\ P_1 + P_2 &= 5 \\ \frac{1}{\lambda_1} + P_1 &= \frac{1}{\lambda_2} + P_2. \end{aligned} \tag{3.41}$$

In both figures, the optimization parameter $\prod_i (1 + P_i\lambda_i)$ (see (3.25)) is shown versus λ_1 and λ_2. It is clear that for the constraints of (3.40), converting the 2×2 channel to a single eigenchannel leads to maximum capacity (In this case the maximum takes place when one of the eigenvalues is equal to zero.). But for the constraints of (3.41), it is equal

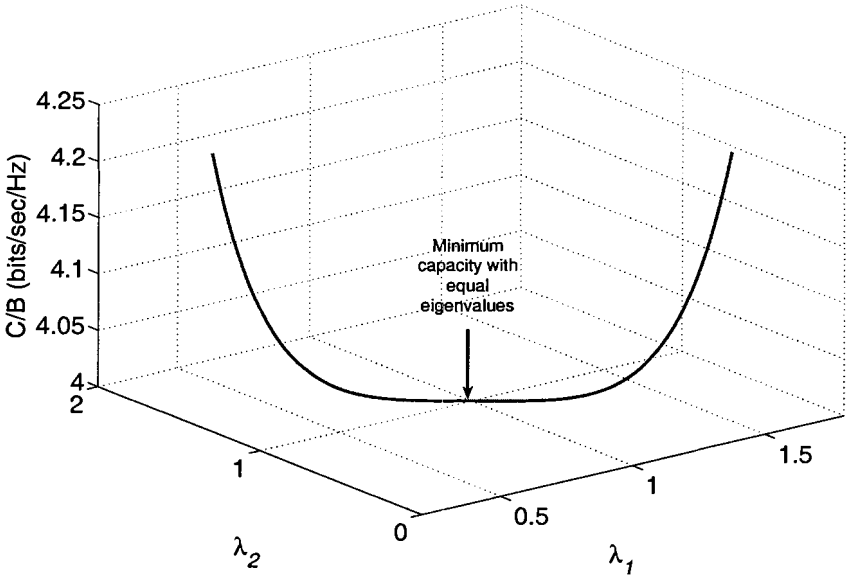

Figure 3.11: Capacity behavior in (3.25) according to the constraints in (3.40).

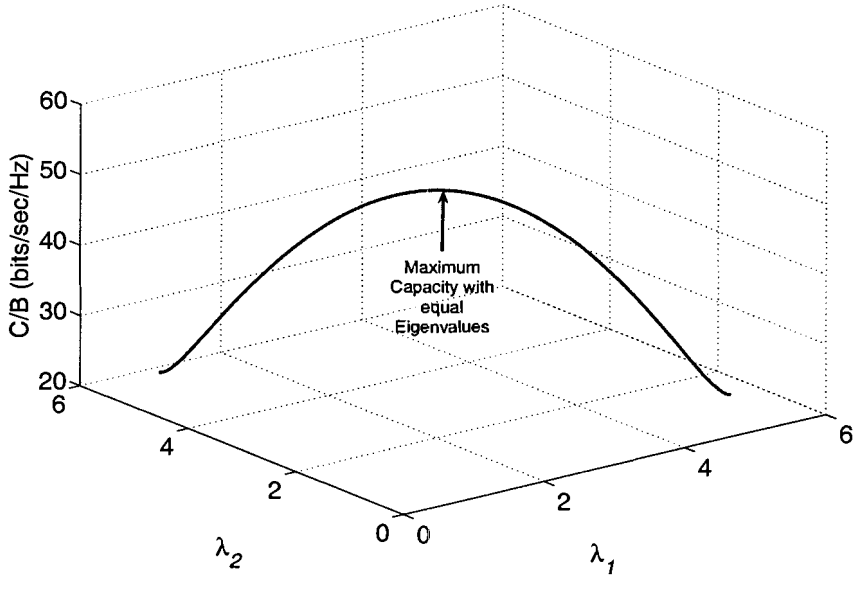

Figure 3.12: Capacity behavior in (3.25) according to the constraints in (3.41).

eigenvalues which results in maximum capacity[3].

The mathematical solution for the open problem above is difficult. However, we can extend our investigation in the 2×2 case for different amount of K (i.e. total gain of the MIMO channel) and P (i.e. the total available power in natural unit). Figure 3.13 shows the result. The area covered by the points is where the equal eigenvalues do not lead to the maximum capacity. In this region, the equal eigenvalue will result in either the minimum capacity or a local maximum (See Figure 3.14 for an example of local maximum). On the other hand, the region covered by the diamonds depicts the area in which the equal eigenvalues leads to the maximum capacity.

From Figure 3.13, we can derive an experimental rule for 2×2 MIMO system. As

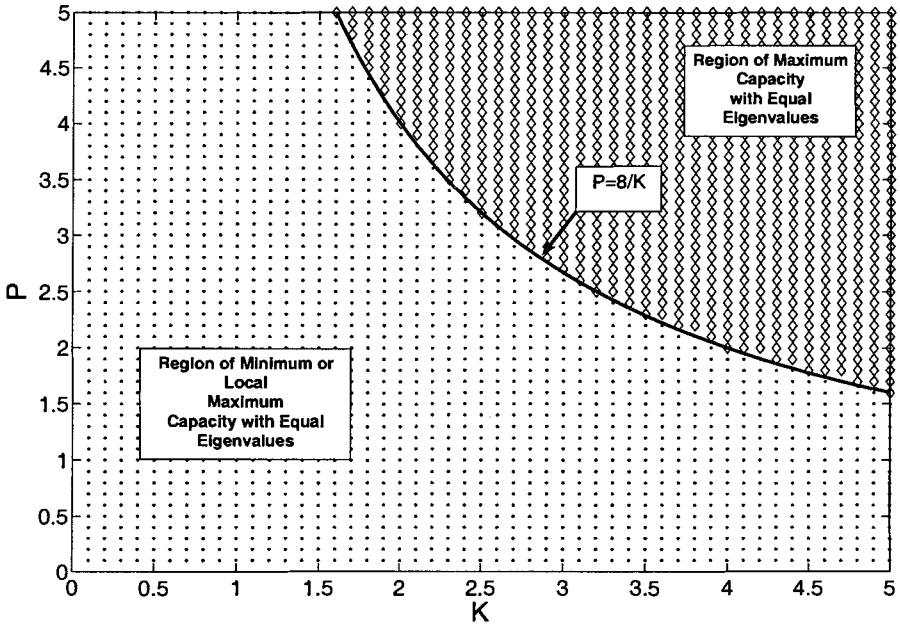

Figure 3.13: Capacity behavior according to K in (3.32) and P in (3.26).

shown in this figure, if we have

$$P \geq \frac{8}{K} \tag{3.42}$$

[3]The experimental results for 2×2 case show that if the combination of P in (3.26) and K in (3.32), i.e. PK is large enough then the equal eigenchannels is the answer to the optimization problem.

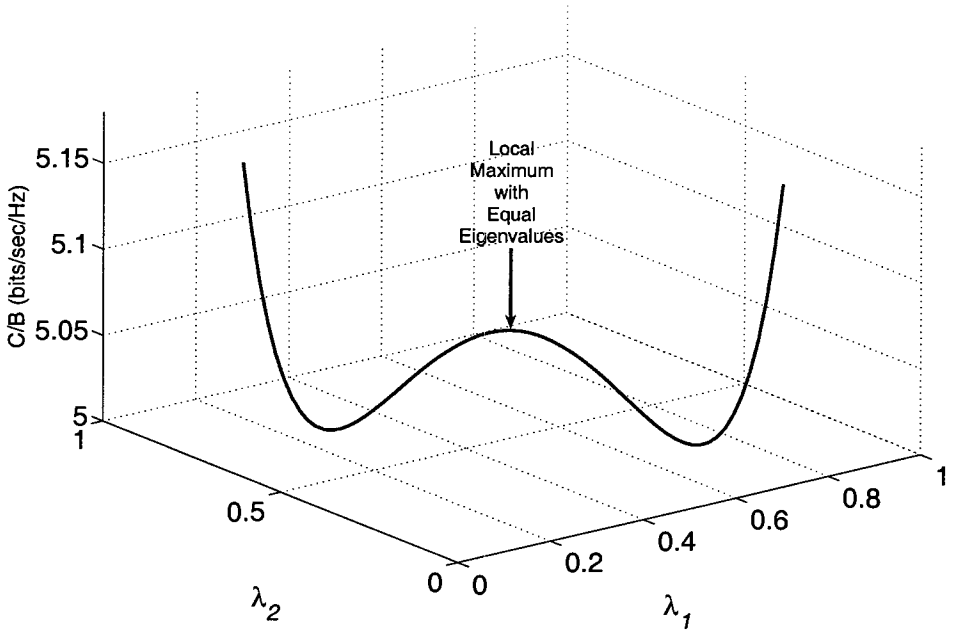

Figure 3.14: An example of local maximum with equal eigenvalues for to $K = 1$ and $P = 5$.

we could be sure that the maximum capacity can be achieved by equal eigenvalues[4]. From this experimental rule (and also clearly from Figure 3.13), it seems that in most of practical constraints where there are strong eigenchannels (large K and also enough total power P), the more similar eigenvalues leads to higher capacity. So, the reason for the circulant having in higher capacity where all the eigenchannels are being used is now clear.

To measure how similar the eigenvalues are, we may introduce different parameters. A parameter which illustrates this similarity is the ratio between the maximum and the minimum eigenvalues, $\lambda_{max}/\lambda_{min}$. The closer to one this parameter is, the larger the capacity. Figure 3.15 shows that the difference between the eigenvalues for the completely random structure is much more than the circulant structure especially when the dimensions of the link become large.

[4]It is noteworthy that the threshold in (3.42), i.e. $PK = 8$ is where the equal eigenvalue capacity for the 2×2 MIMO system is equal to the capacity achieved by shutting down one of the eigenchannels (e.g. $\lambda_1 = 0$) and putting all the power in the other one, i.e. $(1 + PK/4)^2 = 1 + PK$ means $PK = 8$.

This trend also can be seen in the standard deviation of the eigenvalues (which is another parameter to measure the similarity between the eigenvalues). This is always smaller for the circulant than for the i.i.d. case suggesting that the circulant link has more similar eigenvalues (See Figure 3.16). However, $\lambda_{max}/\lambda_{min}$ demonstrates the difference between the eigenvalues more clearly than the standard deviation.

From the hypothesis, the cross over point in Figure 3.4 (and also in Figure 3.8) where the circulant structure begins to feature higher capacity than the i.i.d. structure, is when the total power is large enough to use all the eigenchannels. Before that, the water filling strategy discards the weakest eigenchannels for the i.i.d. structure whereas they remain for the circulant structure, leading to its inferior performance for small SNRs.

As can be seen above, analysis of the pdfs of the eigenvalues gives us an understanding about the behavior of the MIMO system. It helps not only in reasoning the way the mean capacity behaves, but also in providing information about other aspects of the phenomenon. For example, the pdfs in Figure 3.9 show that the variances of the eigenvalues for the circulant case are larger on average than those for the completely random (i.i.d.) case. This means that we should expect a wider range of possible capacities around the mean behavior for the circulant channel. This conclusion can be confirmed by the result in Figure 3.17.

Figure 3.17 shows the 95% percentile of the capacity efficiency. This means that 95% of the time, the capacity efficiency is between the upper and the lower bound. It is clear that these bounds are wider for the circulant case than the i.i.d. case. However, the mean capacity efficiency for the circulant one is higher and it is still right to say that most of the time, the capacity efficiency of the circulant is better than that of the i.i.d.

Before finishing, it is worth mentioning that as far as we have examined, not only the circulant, but also some other systematic correlated structures show the same capacity benefit. For example, *Toeplitz* matrices also have a higher capacity than i.i.d. channels. However, their capacity is smaller than the circulant. It has not been proved yet, but as a guess, it might be correct that in the sets of $N \times N$ random matrices with random elements having same statistics, if we have a pre-defined pattern in the elements (as we have in the circulant or Toeplitz), the capacity efficiency is better than that of the i.i.d. case.

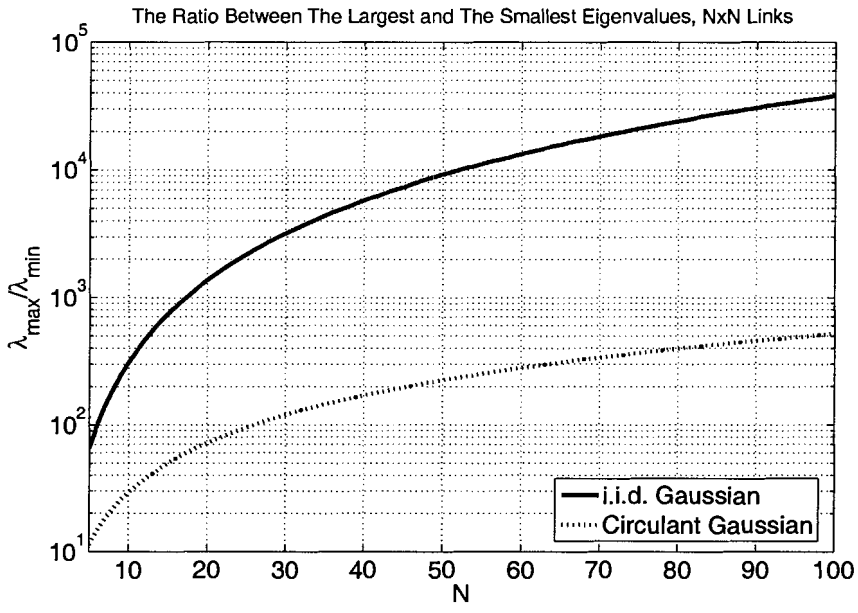

Figure 3.15: $\frac{\lambda_{max}}{\lambda_{min}}$ for the two types of $N \times N$ channels emphasizing the difference between the channel types.

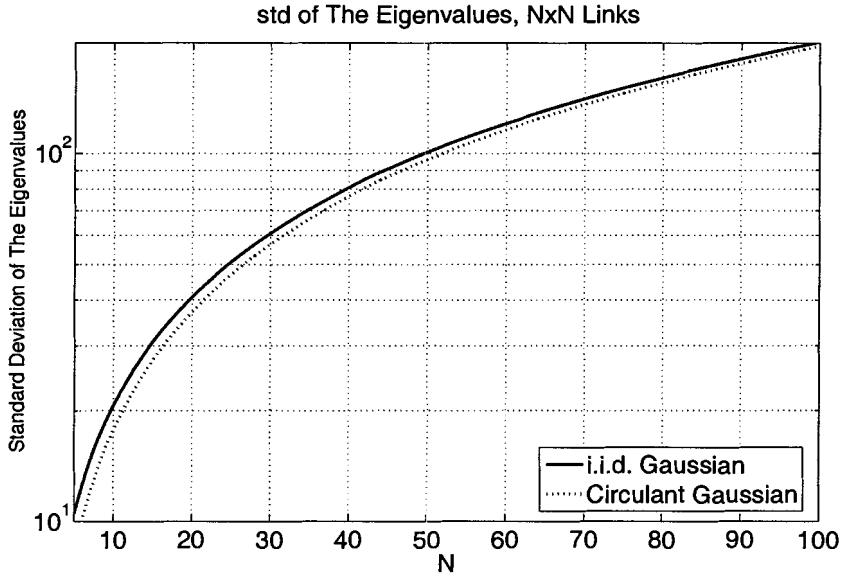

Figure 3.16: The $std(\lambda_i)$ for the two types of $N \times N$ channels.

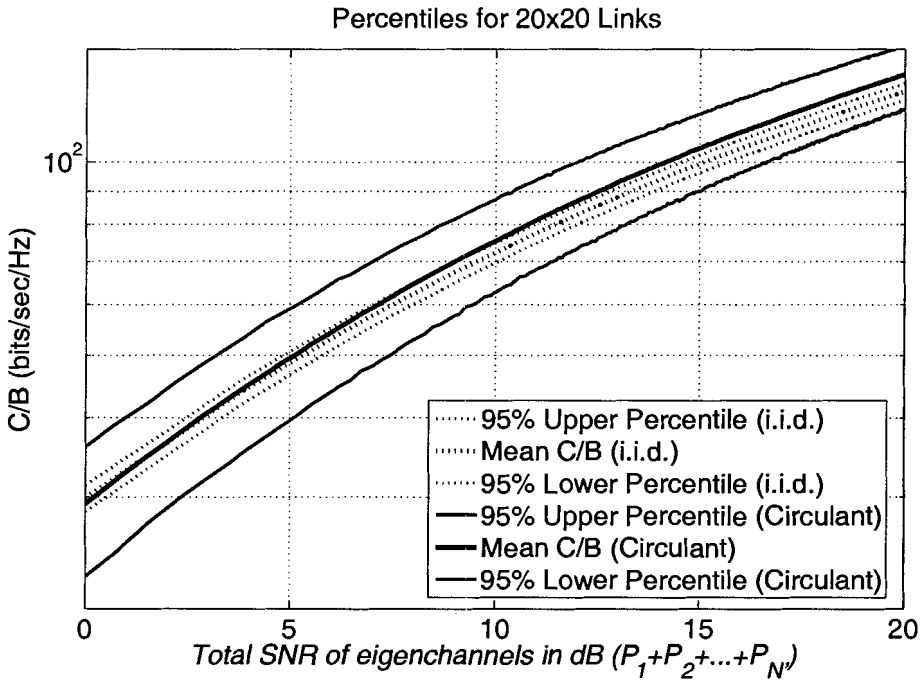

Figure 3.17: The 95% Percentile for the two types of 20×20 channels.

3.4 Signal Processing Advantages of The Circulant Channel

Recall that to have an optimum MIMO system, the transmitter needs to know the eigenvalues in order to assign the power for each of the eigenchannels according to the water filling strategy. However, because of the errors in the estimation of the eigenvalues, this cannot be achieved in practice. For the circulant MIMO system, since the eigenvectors are fixed, we can improve our estimation of the eigenvalues. By choosing the closest fixed eigenvector to the estimated eigenvector at the transmitter, we are able to estimate the eigenvalues and so the power for each of the eigenchannels. But water filling powers are not the only challenging issue in the implementation of a MIMO system, especially when there could be the sub-optimum strategy of equal power allocation (See Chapter 4). The other main drawback is the estimation of the MIMO weights at the transmitter.

In fact, MIMO systems are very sensitive to the errors in the estimation of the eigenvectors. We explained the important role of Singular Value Decomposition (or SVD) in MIMO theory [41] before (See Section 2.4.2). Recall that $\boldsymbol{G}$ in (2.18) is a Hermitian matrix in

general and has N distinct, real positive eigenvalues and the remaining will be zero. The SVD expansion of $\boldsymbol{H}$ itself is

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^H$$

and for $\boldsymbol{G}$

$$\boldsymbol{G} = \boldsymbol{H}^H\boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}\boldsymbol{V}^H.$$

The received voltage for the ith eigenchannel from (2.23) is

$$r_i = \boldsymbol{U}_i^H\boldsymbol{H}\boldsymbol{V}_i = \boldsymbol{U}_i^H\sqrt{\lambda_i}\boldsymbol{U}_i = \boldsymbol{U}_i^H\boldsymbol{U}_i\sqrt{\lambda_i} = \sqrt{\lambda_i}.$$

Now consider we have error in the estimation of the channel matrix at the transmitter. Denote $\widehat{\boldsymbol{H}} = \boldsymbol{H} + \boldsymbol{n}$ where $\boldsymbol{n}$ is the estimation noise. Then instead of

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^H,$$

the estimation of the channel at the transmitter is

$$\widehat{\boldsymbol{H}} = \widehat{\boldsymbol{U}}\widehat{\boldsymbol{\Lambda}}\widehat{\boldsymbol{V}}^H \tag{3.43}$$

where $\widehat{\boldsymbol{V}}$ are the estimated weights (eigenvectors) at the transmitter. So, Assuming that the estimation of the channel at the receiver is perfect (i.e. the receiver knows $\boldsymbol{U}$), the equivalent system will consist of the noisy estimation of the eigenvectors at the transmitter ($\widehat{\boldsymbol{V}}$), the channel matrix ($\boldsymbol{H}$) and the noiseless estimation of the eigenvectors at the receiver ($\boldsymbol{U}$) i.e.

$$\widehat{\boldsymbol{\Lambda}} = \boldsymbol{U}^H\boldsymbol{H}\widehat{\boldsymbol{V}}. \tag{3.44}$$

Here, $\widehat{\boldsymbol{\Lambda}}$ is no longer a diagonal matrix and does have non-zero off-diagonal elements. This means that there is interference (power leakage) between the eigenchannels (which are supposed to be completely isolated in the ideal case). This interference degrades the MIMO performance severely. To calculate the degraded (noisy) i.i.d. and circulant capacities due to the estimation noise, (3.43) can be used. The off-diagonal non-zero elements of $\widehat{\boldsymbol{\Lambda}}$ may be considered as the interference between the eigenchannels. So, we are able to find the noisy capacities for the ith eigenchannel as

$$(\frac{C}{B})_i = \log_2\left(1 + \frac{S_i}{I_i + N_i}\right), \tag{3.45}$$

in which

$$S_i = \left|\widehat{\mathbf{\Lambda}}_{ii}\right|^2 \times P_i$$

where P_i is the power assigned to the ith eigenchannel (either from the water filling strategy or from the equal power allocation). N_i is the noise in this eigenchannel which is assumed to be $N_i = S_i \times 10^{-3}$ i.e. 30 dB below the signal power. Also

$$I_i = \sum_{j=1, j\neq i} \left\{ \left|\widehat{\mathbf{\Lambda}}_{ij}\right|^2 \times P_j \right\}$$

is the interference in the ith eigenchannel.

In order to ease this sensitivity, we can exploit the fact that circulant matrices have fixed eigenvectors as we have seen in (3.5). Assume a 5×5 circulant MIMO structure. From (3.5) the 5 eigenvectors are fixed. Figure 3.18 depicts the 5 directions of these fixed eigenvectors (Note that for true demonstration we need 5 dimensions which is not possible to draw). To compensate degradation caused by errors in the estimation of the eigenstructure, the noisy estimation of each of the eigenvectors, $\widehat{\boldsymbol{V}}_i$, is being compared with these directions and the closest fixed eigenvector is chosen. So, the optimum estimation of the eigenvectors at the transmitter can be found. Using this estimation, the optimum estimation of the eigenvalues for the allocation of the water filling powers is possible. Using this strategy, Figure 3.19 shows the results of this compensation for practical SNRs.

As can be seen from the Figure 3.19, although in the ideal noiseless cases the capacity efficiency of the circulant is better, the estimation noise degrades the performance of the circulant more severely than that of the completely random (i.i.d.) structure. However, using the above technique, the performance of the noisy circulant MIMO systems can improve considerably while there is no way to recover the i.i.d. capacity.

It is also obvious that by increasing the dimensions of the MIMO system the improvement due to the fixed eigenvectors decreases. This is because of the increasing possibility of a wrong decision in picking up the correct fixed eigenvector (See Figure 3.18).

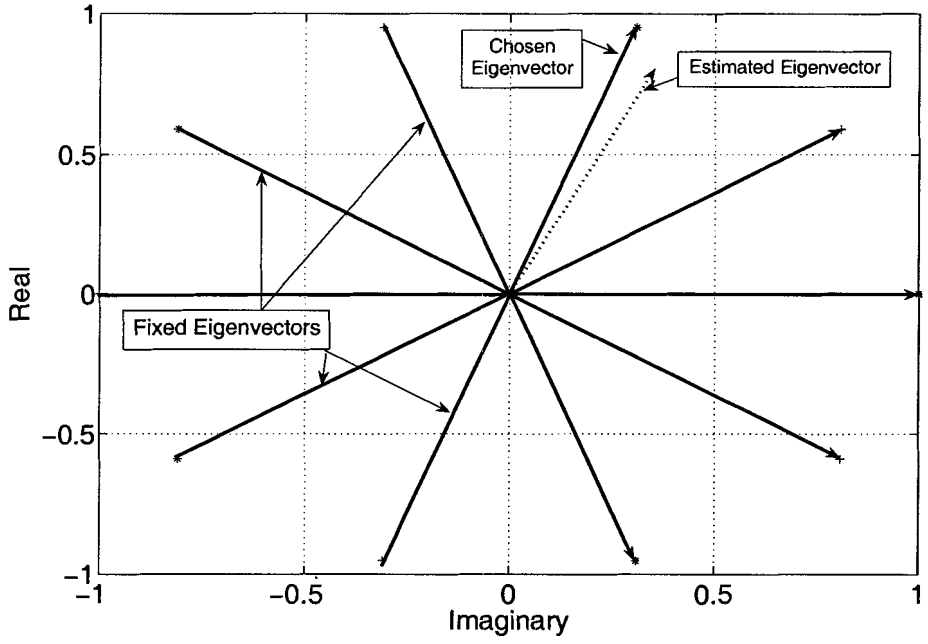

Figure 3.18: Figurative strategy to chose the transmitter weights from the roots of unity based on the estimated eigenvectors.

3.5 Further Aspects

3.5.1 Rearranging an i.i.d.

The fact that the circulant structures give higher capacity, encourages us to try the possibility of rearranging an i.i.d. structure to a more circulant one. We have tried to define a metric showing that *how circulant* a random matrix is and trying to rearrange it based on this metric.

To define this metric, we pick up the first row of the random (i.i.d.) matrix, $\boldsymbol{H}$, as the first row of the new rearranged matrix $\boldsymbol{H}_{new}$ (which is supposed to be more circulant). Then we compare all the permutations of the elements in the second row of $\boldsymbol{H}$ with the circularly shifted version of its first row. Our metric is the distance between these permutations and the circularly shifted version of the first row. We pick up the permutation with the smallest distance as the second row of $\boldsymbol{H}_{new}$. We go on to other rows and through the same process build up the rearranged matrix. Although this method seems to be able to give us a more

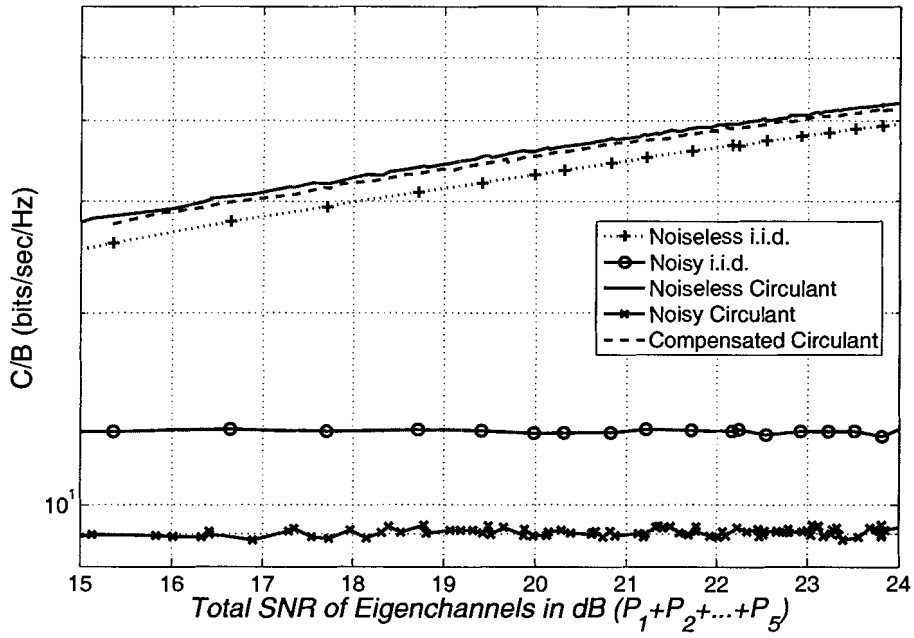

Figure 3.19: Considerable improvement in capacity exploiting the fixed eigenvectors of the 5×5 circulant structures.

circulant structure, the capacity efficiency does not change from $\boldsymbol{H}$ to $\boldsymbol{H}_{new}$. This might be because of the fact that any rearrangement of the i.i.d. structure could be considered as another i.i.d. structure at least in terms of 'mean behavior'.

3.5.2 Figure of Merit

It would be nice to introduce a metric for random matrices showing that how good they are in terms of capacity efficiency. This figure of merit should illustrate two facts. Firstly, it should illustrate how similar the eigenvalues of the random matrix are. Secondly, it should illustrate how much the total gain of the MIMO system is, using this channel matrix. Unfortunately there is no single value showing these two at once, but we can use a two dimensional merit plane. For the first metric, both $\frac{\lambda_{max}}{\lambda_{min}}$ and the standard deviation of the eigenvalues could be used. We chose the standard deviation due to its rather smaller value

(Compare Figure 3.15 and Figure 3.16). For the second metric, the trace of $\boldsymbol{G}$ in (2.18) which is the total gain of the MIMO system is chosen.

For three different cases, i.i.d., circulant and Toeplitz matrices these two figures of merits are shown in Table 3.2 and in Figure 3.20 as well.

Figure 3.20 essentially depicts a plane of merit. Theoretically we prefer to be in the

Link Type	*Trace of G*	*Standard Deviation of Eigenvalues*
i.i.d.	799.9980	40.7841
Toeplitz	796.2739	38.6328
Circulant	801.4081	37.1015

Table 3.2: The figure of merits for 20×20 links.

left upper region of this plane. In this region the total gain of the MIMO channel is large while the standard deviation of the eigenvalues is less, and these two assure us of having good capacity efficiency. However, note that the sensitivity of the MIMO capacity to small variations in standard deviation (horizontal axis) is considerably larger than its sensitivity to small changes in total gain (vertical axis). So, although the total gain of Toeplitz matrix is a bit smaller than i.i.d. one, because of its smaller standard deviation, the performance of Toeplitz is better than i.i.d. It is clear from the Figure and confirmed above that the circulant matrix enjoys better figures of merits compared to i.i.d. and Toeplitz (See Figure 3.21).

3.5.3 Further Work

Currently, I am trying to solve the open problem mentioned in this chapter using computational methods. While it seems to be difficult to find a purely analytical method for this optimization problem, it is easier to simulate different $N \times N$ MIMO systems and find the maximum capacity for different K in (3.32) and P in (3.26). The investigations are time-taking. It seems that the behavior are similar to the behavior of 2×2 case, i.e. if the total gain of the MIMO system is K, we can find the factor, $A(N)$, which is a function of N, the dimension of the $N \times N$ MIMO system. Then, having the total amount of power, P, at the transmitter which satisfies

$$P \geq \frac{A(N)}{K}, \tag{3.46}$$

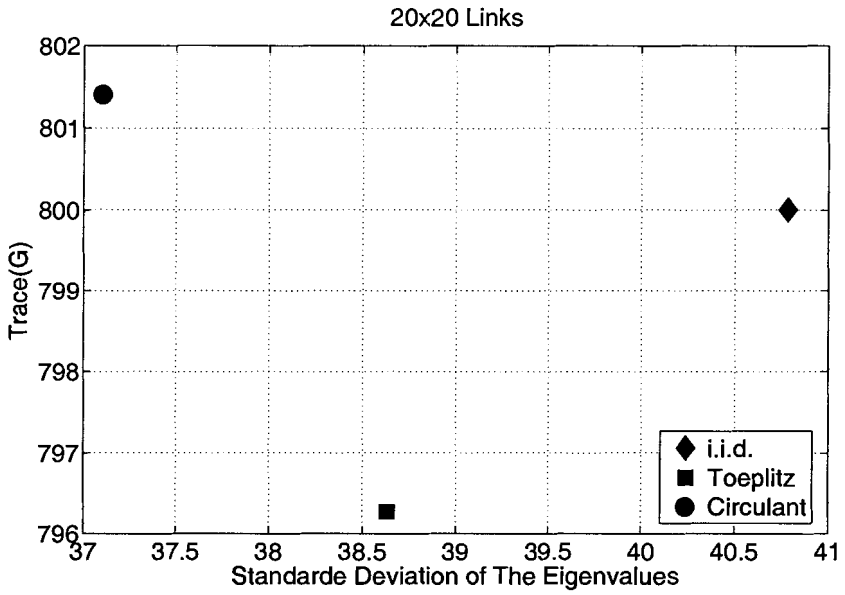

Figure 3.20: Figure of merits for three different types of 20×20 links.

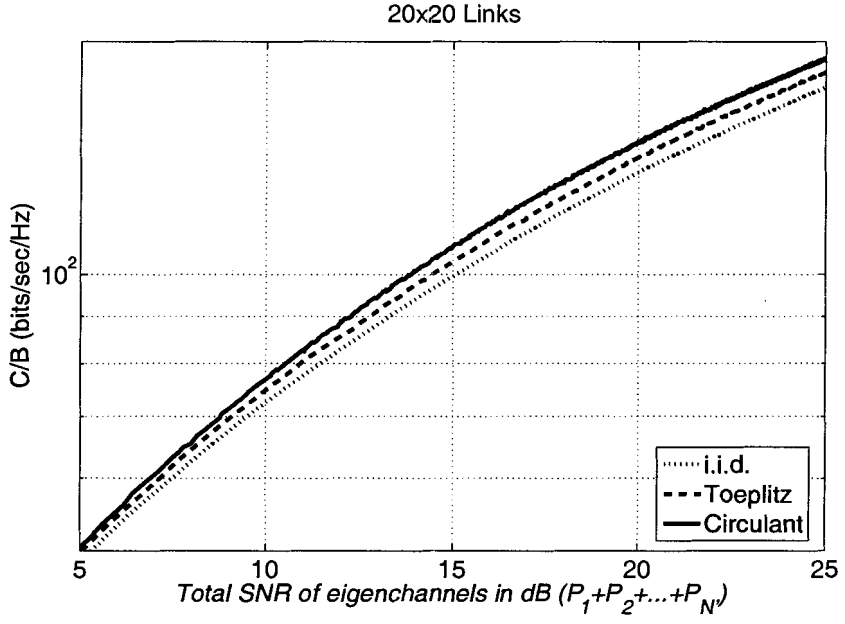

Figure 3.21: Capacity efficiency for three different types of 20×20 links.

means that equal eigenvalues will lead to the maximum capacity.

Another expansion of this thesis work is applying a coding scheme (space-time codes). This is another step toward a full MIMO system. Experimental results also could be done whenever equipment becomes available. Through these experiments we will be able to confirm our results about correlated structures.

Chapter 4

Equal Powers and Water Filling Powers

4.1 Sub-optimum Equal Power Solution

The water filling strategy introduced in Section 2.4.3 is one of the main and challenging elements of MIMO theory. The significance of water filling is even more important and complicated in terms of implementation. In particular, water filling requires a communication protocol to set up the eigenchannels.

This chapter seeks to observe how much degradation in the capacity efficiency occurs by simply omitting water filling. The total power allocated by the water filling is now equally divided between the eigenchannels. Figures 4.1 and 4.2 show the results for two channel types and the two power allocation strategies.

From these figures, it is apparent that for large SNRs the degradation arising from equal power allocation is negligible. This is a significant but reasonable result since for large SNRs the differences between the water filling powers assigned to different eigenchannels become less, and so the capacity converges to that of the equal power allocation scheme. We will investigate this claim with a more quantitative approach below. Also, it is noteworthy that this convergence is faster for the circulant case again. In fact, the more similar eigenchannels of the circulant structure helps the equal power capacity to converge faster to water filling capacity (because the water filling powers are more similar to each other in the circulant case compared to the i.i.d. case.). It is apparent that these two methods of

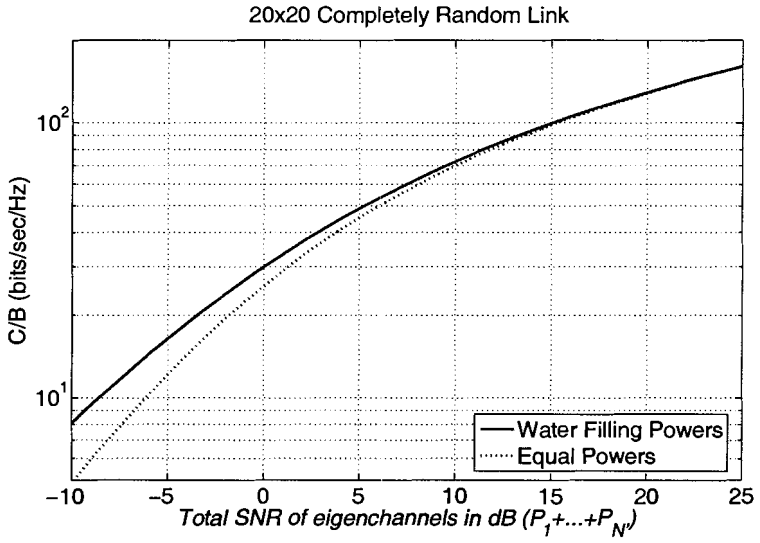

Figure 4.1: The Shannon capacity of equal powers and optimum water filling powers for the completely random link.

power allocation for the $N \times N$ link converge later, i.e. for larger SNRs when N is larger.

The fact that in large dimension MIMO systems (like 20×20 case) we can use equal powers is also true when we apply digital communications techniques. Figures 4.6 and 4.7 illustrate the practicable capacities (capacities using QAM techniques). It is clear that allocation of equal powers is still reasonable for most of the SNRs. Again if we compare the converging behavior of the two methods of power allocation, we can observe that this convergence is faster for the circulant case (for the same reason). From Figure 4.7 the advantage of circulant structures without the complexities of water filling is clear especially for large SNRs.

So for the 20×20 case, in the (total) SNR range over 5 dB, it would be sensible to use equal powers instead of water filling powers. Again, one of the best ways to explain this behavior of the capacity is to look into the eigenvalues. For this, it would be helpful to find the mean eigenvalues for 20×20 Gaussian (i.i.d.) link. These mean values from simulations, are depicted in Figure 4.3 and also mentioned in Table 4.1. From the Figure and the Table, it is clear that the eigenvalues, changing from $\lambda_{1mean} \approx 140$ to $\lambda_{20mean} \approx 0.1$, have a considerable variance. In fact the standard deviation of the eigenvalues for this link (20×20 Gaussian link) is 40.78 which is significant in the light of discussion below. One

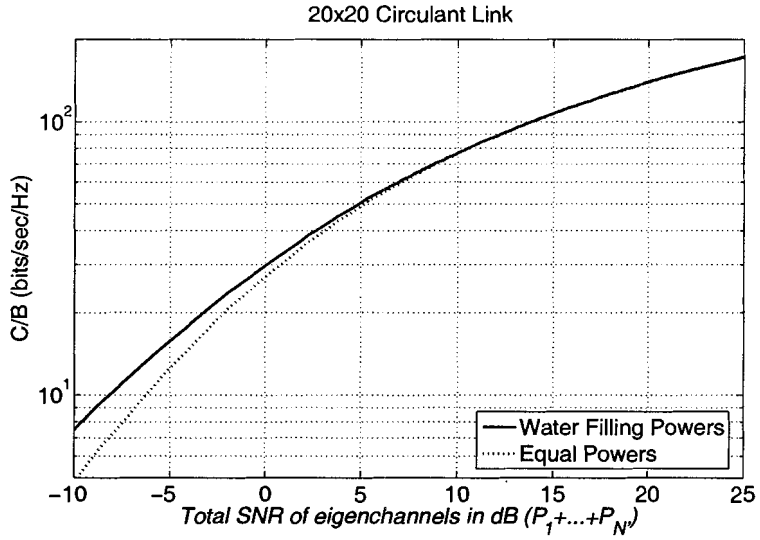

Figure 4.2: The Shannon capacity of equal powers and optimum water filling powers for the circulant link.

may deduce from this large variance that since the eigenvalues are very different from each other (as can be seen clearly in Figure 4.3), then the eigenchannels are very different too, and from water filling they should receive very different amounts of power resulting in an optimum capacity efficiency. So the optimum capacity depends on accurately finding of these very dissimilar eigenchannel powers. Following this argument, the above-mentioned suggestion that we can neglect the water filling and simply use equal powers for different eigenchannels must be false.

However, this analysis is incorrect! Recall that in the water filling strategy, the power is being distributed between the eigenchannels according to the inverses of the eigenvalues i.e. $1/\lambda$, and not the eigenvalues themselves. Figure 4.4 shows these inverses. The inverses of the eigenvalues are more similar than the eigenvalues themselves. The inverses are in the approximate range of 0.01 to 10 for the 20×20 link and their standard deviation is approximately 2.23. Moreover, if we neglect the largest inverse (which is being removed from the communications process by water filling cut-off for most of the SNRs) this standard deviation will reduce to 0.38.

Both of these numbers are very small comparing to the standard deviation of the eigenvalues themselves, which was noted above to be 40.78. This clearly shows that although

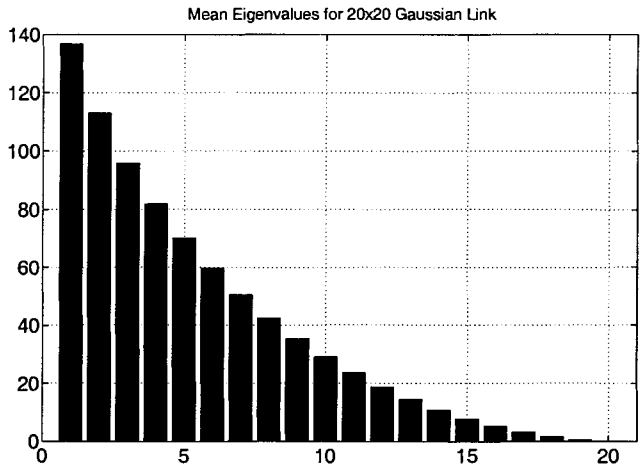

Figure 4.3: The mean eigenvalues for 20×20 Gaussian link.

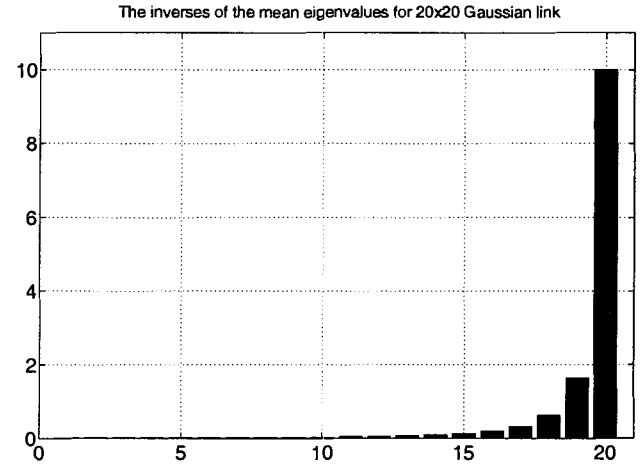

Figure 4.4: The inverses of the mean eigenvalues for 20×20 Gaussian link.

$\lambda_{20mean} = 0.1000$	$\lambda_{19mean} = 0.6095$
$\lambda_{18mean} = 1.6077$	$\lambda_{17mean} = 3.1058$
$\lambda_{16mean} = 5.1172$	$\lambda_{15mean} = 7.6567$
$\lambda_{14mean} = 10.7379$	$\lambda_{13mean} = 14.3904$
$\lambda_{12mean} = 18.6346$	$\lambda_{11mean} = 23.5098$
$\lambda_{10mean} = 29.0657$	$\lambda_{9mean} = 35.3625$
$\lambda_{8mean} = 42.4751$	$\lambda_{7mean} = 50.5059$
$\lambda_{6mean} = 59.5914$	$\lambda_{5mean} = 69.9440$
$\lambda_{4mean} = 81.8690$	$\lambda_{3mean} = 95.8994$
$\lambda_{2mean} = 113.1169$	$\lambda_{1mean} = 136.7358$

Table 4.1: The mean eigenvalues.

MIMO Dimension	std of the inverse eigenvalues	D
10×10	≈ 0.25	≈ 1.25
11×11	≈ 0.27	≈ 1.18
12×12	≈ 0.28	≈ 1.06
13×13	≈ 0.30	≈ 0.98
14×14	≈ 0.31	≈ 0.96
15×15	≈ 0.32	≈ 0.88
16×16	≈ 0.33	≈ 0.83
17×17	≈ 0.35	≈ 0.78
18×18	≈ 0.36	≈ 0.74
19×19	≈ 0.37	≈ 0.66
20×20	≈ 0.38	≈ 0.63

Table 4.2: The standard deviations of the inverses of the eigenvalues except for the smallest eigenvalues (largest inverse), and the water filling cut-off threshold, D, from which equal power capacities are very similar to the water filling ones.

the eigenvalues themselves show a large range of values, the inverses of them change in a smaller range and so the inverses of the eigenvalues are more similar to each other. These more similar values will lead to more similar powers for the different eigenchannels. So, in the water filling process for large dimension MIMO link (such as 10×10 or more), the amounts of power assigned to the channels are not very different from equal powers. This is especially true when the largest inverses (i.e. the smallest eigenvalues) are being omitted through cut-off.

If we expand our investigation to $N \times N$ MIMO systems from $N = 10$ to $N = 20$, we will be able to find an experimental rule. This rule will tell us from which cut-off threshold, D, if we use equal powers instead of water filling powers, the capacity efficiency would be

reasonably similar[1]. To find this rule, we should first extract the standard deviation of the inverses of the eigenvalues (except for the smallest eigenvalue). These standard deviations are illustrated in Table 4.2.

On the other hand we need to find the special water filling threshold, D, for our range of dimensions from where the Shannon limit capacities, with equal and water filling powers, are similar. These values of D are also shown in Table 4.2.

Now, if we depict the Ds in Table 4.2 vs. the *std*s in a plot, we recognize a nearly linear behavior (See Figure 4.5). Least mean square analysis gives us the experimental rule for the desired threshold according to the standard deviation will be

$$D \approx -4.72std + 2.41. \tag{4.1}$$

From (4.1), having the standard deviation of the inverse eigenvalues (except for the smallest one) we can find the approximate amount of D from where there is no need to use water filling. For the abovementioned range of MIMO dimensions, i.e. from 10×10 to 20×20, these amounts of D are associated with the total SNR of 9 to 11 dB at the transmitter.

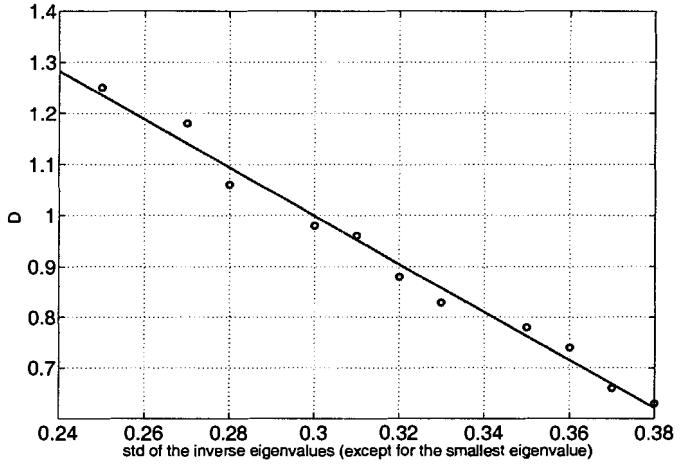

Figure 4.5: The linear interpolation of Ds and *std*s in Table 4.2.

[1]By reasonably similar we mean the difference of 2 bits/sec/Hz or less in capacity efficiencies.

4.2 Comparison of SISO and MIMO

It would be interesting if we compare SISO and MIMO with another approach. As seen before, in MIMO systems, the communication channel is being converted to multiple separated eigenchannels. In general, these eigenchannels have different gains. So, the strongest eigenchannel, i.e. the one associated with the largest eigenvalue, is of interest.

To be able to compare the practicable capacities of our MIMO system with a SISO (Single-Input Single-Output) one, the practical capacity of the strongest eigenchannel is also depicted. To plot these single eigenchannel capacities in Figures 4.6 and 4.7, the total amount of power distributed between the different eigenchannels in the MIMO case is being assigned to the strongest eigenchannel. It is clear and reasonable that for small SNRs, putting all of the power in the strongest channel is better than distributing this small amount of power between many channels, while for large SNRs the MIMO capacities are significantly more than the best that could be achieved by a single channel.

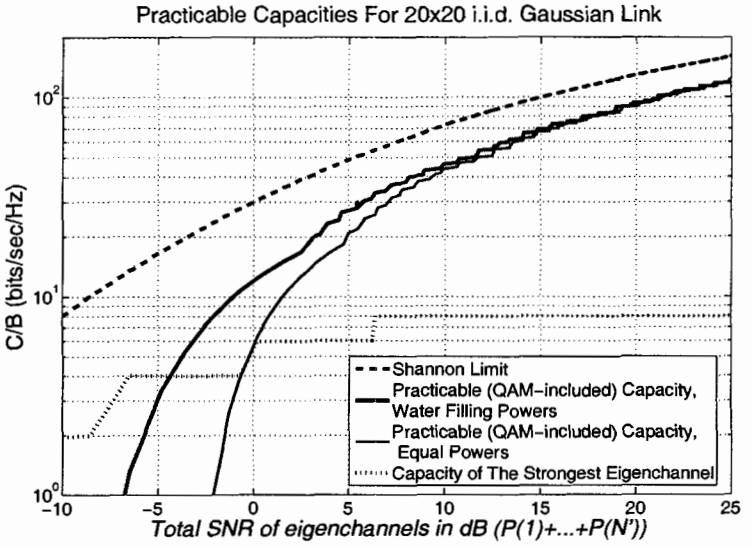

Figure 4.6: The practicable capacity of equal powers and optimum water filling powers for the completely (i.i.d.) random link using QAM techniques.

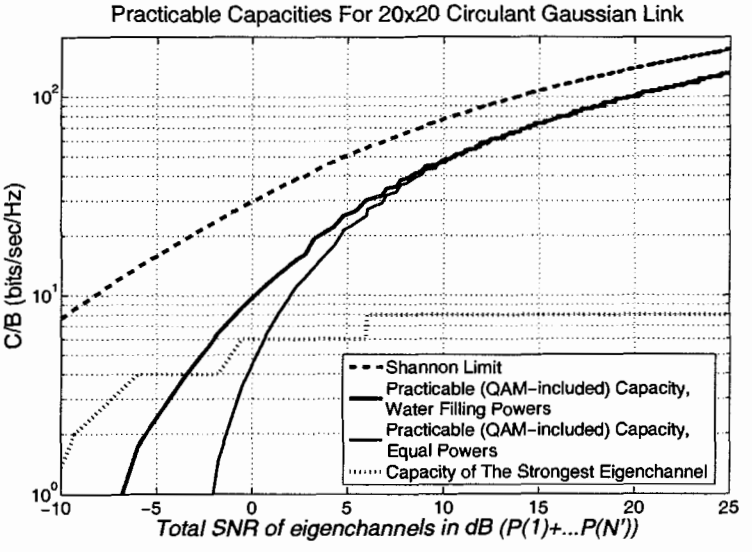

Figure 4.7: The practicable capacity of equal powers and optimum water filling powers for the circulant random link using QAM techniques.

Chapter 5

Summary

The main goal in this thesis was to investigate the different characteristics of the communication channel through analysis of the eigenvalues of the channel matrix. In this analysis we covered different aspects of MIMO systems such as information theory, signal processing and power allocation strategy.

Correlation within the channel matrix is often unavoidable. Of particular fascination, was the systematic correlation of the circulant structure. This structure has the maximum MIMO capacity compared to other structures considered, such as i.i.d. and Toeplitz. The fact that systematic correlation can increase the capacity efficiency means that instead of removing the correlation (which is not possible) we may be able to exploit it. The higher capacity of the circulant was explained by analysis of the pdfs of the eigenvalues. The basic hypothesis on the MIMO capacity was clarified and also different facets of this hypothesis were presented. Furthermore, signal processing advantages of the circulant were developed, again based on eigenstructure characteristics. Figures of merit from the eigenstructure were proposed. Systematic, correlated MIMO channels usually had better figures of merit corresponding to higher capacities compared to that of the i.i.d. channel.

Further to the channel modeling aspect, the power allocation issue in a MIMO system was investigated. Again, analysis of the eigenvalues and their inverses illustrated the possibility of avoiding water filling, which would be an advantage in full MIMO implementation. The research contributions of this thesis provide steps towards the feasibility of the MIMO systems for digital communications.

Bibliography

[1] R. Nabar A. Paulraj and D. Gore. *Introduction to Space-Time Wireless Communications.* Cambridge University Press, 2003.

[2] A. Abdi and M. Kaveh. A space-time correlation model for multielement antenna systems in mobile fading channels. *IEEE Journal on Selected Areas in Communications*, 20(3):550–560, April 2002.

[3] N. Boubaker, K. B. Letaief, and R. D. Murch. Performance of blast over frequency-selective wireless communication channels. *IEEE Transactions on Communications*, 50(2):196–199, February 2002.

[4] James K. Cavers. *Mobile Channel Characteristics.* Kluwer Academic Publishers, 2000.

[5] W. Choi, K. Cheong, and J. M. Cioffi. Iterative soft interference cancellation for multiple antenna systems. In *Proc. of Wireless Communications and Networking Conference*, volume 1, 23-28, pages 304–309, September 2000.

[6] C. N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela. Capacity scaling in mimo wireless systems under correlated fading. *IEEE Transactions on Information Theory*, 48(3):637–650, March 2002.

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

[8] A. Edelman. Eigenvlues and condition numbers of random matrices. *PhD Thesis, MIT,(ps file at http://math.mit.edu/ edelman/thesis)*, 1989.

[9] Stefan A. Fechtel. A novel approach to modeling and efficient simulation of frequency-selective fading radio channels. *IEEE Journal on Selected Areas in Communications*, 11(3):422–431, April 1993.

[10] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multiple antennas. *Bell Laboratories Technical Journal*, 1(2):41–59, Autumn 1996.

[11] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3):311–335, March 1998.

[12] R. Gallager. *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[13] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj. Outdoor mimo wireless channels: Models and performance prediction. *IEEE Transactions on Communications*, 50(12):1926–1934, December 2002.

[14] D. Gesbert, M. Shafi, Da shan Shiu, P. J. Smith, and A. Naguib. From theory to practice: An overview of mimo space-time coded wireless systems. *IEEE Journal on Selected Areas in Communications*, 21(3):281–302, April 2003.

[15] G. Ginnis and J.M. Cioffi. Vectored-dmt: A fext canceling modulation scheme for coordinating users. In *Proc. of IEEE International Conference on Communications*, volume 1, 11-14, pages 305–309, June 2001.

[16] R. M. Gray. *Toeplitz and Circulant Matrices: A Review.* pdf file at http://ee.stanford.edu/'gray/toeplitz.pdf, 2002.

[17] U. Haagerup and S. Thorbjoernsen. Random matrices with complex gaussian entries. *Aalborg University Report, Preprint 1-49*, 1998.

[18] B. Hassibi and B. M. Hochwald. High-rate codes that are linear in space and time. *IEEE Transactions on Information Theory*, 48(7):1804–1824, July 2002.

[19] W. C. Jakes. *Microwave Mobile Communications.* IEEE Press Reprint, 1994.

[20] A. T. James. Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Stat.*, 35:475–501, 1964.

[21] M. Kang and M. S. Alouini. Impact of correlation on the capacity of mimo channels. In *Proc. of IEEE International Conference on Communications*, volume 4, 11-15, pages 2623–2627, May 2003.

[22] S. Loyka and G. Tsoulos. Estimating mimo system performance using the correlation matrix approach. *IEEE Communications Letters*, 6(1):19–21, January 2002.

[23] A. Lozano and C. Papadias. Layered space-time receivers for frequency-selective wireless channels. *IEEE Transactions on Communications*, 50(1):65–73, January 2002.

[24] M. Z. Win M. Chiani and A. Zanella. On the capacity of spatially correlated mimo rayleigh-fading channels. *IEEE Transactions on Information Theory*, 49(10):2363–5371, October 2003.

[25] C. C. Martin, J. H. Winters, and N. R. Sollenberger. Multiple-input multiple-output (mimo) radio channel measurements. In *Proc. of IEEE Vehicular Technology Conference*, volume 2, 24-28, pages 774–779, Fall 2000.

[26] C. Oestges, B. Clerckx, D. Vanhoenacker-Janvier, and A. J. Paulraj. Impact of diagonal correlations on mimo capacity: Application to geometrical scattering models. In *Proc. of IEEE Vehicular Technology Conference*, volume 1, 6-9, pages 394–398, October 2003.

[27] C. Oestges and A. J. Paulraj. Beneficial impact of channel correlations on mimo capacity. *Electronics Letters*, 40(10):606–608, May 2004.

[28] A. Paulraj, R. Nabar, and D. Gore. *Introduction To Space-Time Wireless Communications.* Cambridge University Press, 2003.

[29] John G. Proakis. *Digital Communications.* McGraw-Hill Higher Education, 2001.

[30] V. Raghavan and A. M. Sayeed. Mimo capacity scaling and saturation in correlated environments. In *Proc. of IEEE International Conference on Communications*, volume 5, 11-15, pages 3006–3010, May 2003.

[31] P. F. Driessen S. Catreux and L. J. Greenstein. Data throughputs using multiple-input multiple-output (mimo) techniques in a noise-limited cellular environment. *IEEE Transactions on Wireless Communications*, 1(2):226–235, April 2002.

[32] D. S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Transactions on Communications*, 48(3):502–513, March 2000.

[33] P. J. Smith, L. M. Garth, and S. Loyka. Exact capacity distributions for mimo systems with small number of antennas. *IEEE Communications Letters*, 7(10):481–483, October 2003.

[34] P. J. Smith and M. Shafi. On gaussian approximation to the capacity of wireless mimo systems. In *Proc. of IEEE International Conference on Communications*, volume 1, pages 406–410, May 2002.

[35] D. K. C. So and R. S. Cheng. Detection techniques for v-blast in frequency selective fading channels. In *Proc. of Wireless Communications and Networking Conference*, volume 1, 17-21, pages 487–491, March 2002.

[36] Gordon L. Stuber. *Principles of Mobile Communication.* Kluwer Academic Publishers, 2000.

[37] G. Taubock T. Magesacher, W. Henkel and T. Nordstrom. Cable measurements supporting xdsl technologies. *Journal e&i Elektrotechnik und Informationstechnik*, 199(2):37–43, February 2002.

[38] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, March 1998.

[39] G. Taubock and W. Henkel. Mimo systems in the subscriber-line network. In *Proc. of the 5th International OFDM-Workshop*, pages 18.1–18.3, June Hamburg, 2000.

[40] E. Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications, ETT*, 10(6):585–596, November 1999.

[41] Rodney Vaughan and Joergen B. Andersen. *Channels, Propagation and Antennas for Mobile Communications.* The Institution of Electrical Engineers, 2003.

[42] J. H. Winters. On the capacity of radio communication system with diversity in rayleigh fading environment. *IEEE Journal on Selected Areas in Communications*, 5(5):871–878, June 1987.

[43] J. H. Winters, J. Salz, and R. D. Gitlin. The impact of antenna diversity on the capacity of wireless communication systems. *IEEE Transactions on Communications*, 42(234):1740–1751, February/March/April 1994.

[44] C. S. Withers and R. G. Vaughan. Distribution and percentiles of channel capacity for multiple arrays. In *Proc. 11th Virginia Tech/MPRG Symposium on Wireless Communications*, pages 141–152, June 2001.

[45] A. Zanella, M. Chiani, M. Z. Win, and J. H. Winters. Symbol error probability of high spectral efficiency mimo systems. In *Proc. of Conference on Information Sciences And Systems*, March 2002.

[46] X. Zhu and R. D. Murch. Performance analysis of maximum likelihood detection in a mimo antenna system. *IEEE Transactions on Communications*, 50(2):187–191, February 2002.