

**TWO DIMENSIONAL SPREADING FOR DOUBLY
DISPERSIVE CHANNELS**

BY JOYDEEP ACHARYA

A thesis submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree of
Master of Science
Graduate Program in Electrical and Computer Engineering

Written under the direction of
Professor Roy D. Yates
and approved by

New Brunswick, New Jersey

May, 2005

© 2005

Joydeep Acharya

ALL RIGHTS RESERVED

ABSTRACT OF THE THESIS

Two Dimensional Spreading for Doubly Dispersive Channels

by Joydeep Acharya

Thesis Director: Professor Roy D. Yates

Systems supporting broadband mobile services over wireless channels suffer from dispersion along time and frequency. Hence transmission by replicating information along both these domains leads to diversity gain in each domain. This work proposes a scheme to implement this replication principle, motivated by the Variable Spread Factor-Orthogonal Frequency Code Division Multiplexing (VSF-OFCDM) scheme introduced by NTT-DoCoMo. In this scheme a symbol is transmitted across several subcarriers with a total power constraint and along each subcarrier it is spread with CDMA codewords. The information theoretic bounds on capacity for this scheme are derived for three different scenarios each corresponding to a specific nature of channel state information (CSI) available at the transmitter. The cases are perfect CSI, partial CSI characterized by one bit of channel information per subcarrier and no CSI. The receiver is assumed to have perfect CSI. The optimal codeword and power allocation strategies to achieve these bounds are also derived for the single user and multi-user uplink channel. We show that for perfect CSI, optimal strategy is to transmit along the best subchannel. For imperfect CSI at transmitter, diversity benefits are observed for a large number of subcarriers. It is also observed that, 1 bit feedback per subcarrier is a good scheme, both in terms of achievable rates and implementation complexity.

Acknowledgements

I had often felt that writing an acknowledgement is a somewhat cliched affair, with students predictably thanking their advisor ,colleagues and finally their family members for the guidance ,help and support provided (in that order). Now after having gone through the gruelling process of research, leading up to my masters dissertation, I am beginning to realize the great truism in those sentiments. My work wouldn't have borne fruition without the guidance of my advisor Dr. Roy D. Yates, to whom I owe my academic knowledge, my motivation for research and numerous small but significant lessons about becoming competent and professional. He has been amazingly patient and understanding but also firm and demanding whenever I tended to relapse into a state of laziness and inactivity. My association with him has benefitted me a lot and I am look forward for the coming years of my doctoral research under him.

I would also like to thank Dr. Leonid Razoumov, who introduced the idea of two dimensional spreading, which has ultimately lead to my masters dissertation. He was always willing to share his vast analytical and practical engineering knowledge with me. I can't forget the long hours, Roy, Leo and myself spent discussing the various aspects of the problem and otherwise.

Next I would like to thank the professors and teachers from whom I have received my education over the years. I would specially like to remember Dr. Saswat Chakrabarti of my alma mater Indian Institute of Technology Kharagpur. It is due to him that I first developed an interest in the field of communications.

I have been always lucky to have excellent friends and I would like to thank all of them. Specifically I thank Jasvinder for rescuing me whenever I was stuck up with a problem. Same goes for Hithesh and Chandru. I thank Ruoheng, Ahmed, Amith, Rahul, Kemal and others. A special thanks to Lalitha for her energy and interest in

my work.

Finally I remember my parents with gratitude for the way they have brought me up over the years, for the way they have always urged me to excel in every sphere of life. And the last one goes my kid sister, who apart from being so much fun, has often surprised me with her sagacity and lessons about life that I have learnt from her.

Table of Contents

Abstract	ii
Acknowledgements	iii
List of Tables	viii
List of Figures	ix
1. Introduction	1
2. VSF-OFDCM and Theoretical Two-Dimensional Spreading	5
2.1. Introduction	5
2.1.1. Brief Description of VSF-OFCDM	5
2.1.2. The Proposed Two-Dimensional Spreading Model	7
Notation	7
Single User Transmission	8
Multi User Transmission	9
3. Achievable Rates for Perfect CSI at Transmitter and Receiver	11
3.1. Introduction	11
3.2. Analysis of the Single User Transmission	12
3.2.1. Uniform Fading	15
3.2.2. Rayleigh Fading	17
3.3. Optimal Allocation for Colored Noise	17
3.4. Simulation Results for Single-User	20
3.4.1. Uniform Fading	20
3.4.2. Rayleigh Fading	21

3.5.	Analysis of the Multi-User Transmission	23
3.5.1.	Both Time and Frequency Domain Spreading	24
3.5.2.	Only Time Domain Spreading	25
3.5.3.	Only Frequency Domain Spreading	26
3.5.4.	Some Proposed Multiuser Transmission Strategies	28
	Optimal Single User Transmission	29
	Treating Other User Interference as Noise	30
	Single Dimension Transmission	31
3.5.5.	Simulation Results for Multi-Users	32
4.	Achievable Rates for Imperfect CSI at Transmitter and Perfect CSI	
	at Receiver	34
4.1.	Introduction	34
4.2.	Single user transmission	34
4.2.1.	Perfect CSIR, No CSIT	36
4.2.2.	Perfect CSIR One Bit Quantized CSIT	37
	All n_f subchannels receive $u = 1$	37
	m subcarriers report that $u = 1$ where $1 \leq m < n_f$	38
	All n_f subcarriers report $u = 0$	38
	Uniform Fading Distribution	40
	Rayleigh Fading	41
4.3.	Simulation Results	41
4.3.1.	Uniform Fading	42
4.3.2.	Rayleigh Fading	44
5.	Conclusions and Future Work	45
	Appendix A. Transmission Policy for 1 bit Quantized CSIT	46
	Appendix B. Properties of the Max function of n Exponentials	47

Appendix C. Maximization of $\mathbf{x}^T \mathbf{A} \mathbf{x}$ with $\mathbf{x}^T \mathbf{x} = 1$ and $\mathbf{x} \geq 0$	53
Class 1: \mathbf{x} has no zero element	54
Class 2: \mathbf{x} has one zero element	55
Cases 3 – n : \mathbf{x} has k zero elements, $2 \leq k \leq n$	56
References	58

List of Tables

3.1. Simulation Parameters for Uniform Fading	21
3.2. Variances in MMI Sequence	22
3.3. Theoretical and Simulated Values of Ergodic Capacities for Large n . .	22
3.4. Simulation Parameters for Rayleigh Fading	22

List of Figures

1.1. Two dimensional distribution of the SNR of a double dispersive fading process with $\tau_c = 3\mu sec$ and $f_D = 500Hz$	3
2.1. One Data Symbol is spread over 8 OFCDM symbols placed evenly in 2 subcarriers	6
2.2. Transmitter Block Diagram	8
3.1. Different solutions for $\lambda^n = a\lambda + b$	16
3.2. MMI variations with time for SNR = 1 dB	21
3.3. MMI variations with time for SNR = 10 dB	21
3.4. Water-Filling parameter λ vs Total Power for n subcarriers: $n = 2, 10, 100$	23
3.5. MMI variations with time for SNR = 1 dB	23
3.6. MMI variations with time for SNR = 10 dB	23
3.7. MMI variations with time for different policies	33
3.8. MMI variations with time different iterations of the	33
4.1. MMI for no CSI	42
4.2. Simulation Parameters for Fig 4.1	42
4.3. Plot of λ vs Power allocated for different feedbacks	43
4.4. Simulation Parameters for Fig 4.3	43
4.5. λ vs Power allocated for all one feedback for different number of subcarriers	43
4.6. Simulation Parameters for Fig 4.5	43
4.7. MMI for different types of CSI	44
4.8. Simulation Parameters for Fig 4.7	44
4.9. λ vs Power allocated for all one feedback for different number of subcarriers	44
4.10. Simulation Parameters for Fig 4.9	44
B.1. Plot of $1/x^2$ and $1/k^2$	51

B.2. $f_{h_n^*}(x)$ for various n 52

Chapter 1

Introduction

Wireless channels are characterized by time-varying multipath fading which makes them doubly dispersive [1]. When signals are transmitted over such channels, the multipath effect leads to dispersion in time, while the time varying channel gain associated with each multipath, causes dispersion in frequency. The nature of these dispersive effects increases with the transmitted signal bandwidth and hence poses an acute problem for reliable broadband mobile services. Quantitatively these effects are broadly captured by the *Spread Factor* of the channel, which is $T_m B_d$ where T_m is the rms delay spread and B_d is the Doppler spread of the channel. A higher value of *Spread Factor* indicates a more dispersive channel. Figure 1.1 illustrates the double-dispersive nature of the channel, when an OFDM signal is transmitted. It can be inferred from the received SNR plot, that the channel response depends both upon the exact time duration and the frequency components of the input signal.

A usual approach to reliable communications is orthogonal signalling, in which the transmitted symbols are modulated by a set of orthonormal basis functions, which correspond to the eigen-modes of the channel. For a variety of channels that have low values of Doppler spread B_d , and hence are not frequency dispersive, OFDM achieves orthogonal signalling by decomposing the transmit signal spectrum into narrow-bands. However, as B_d increases, the channel becomes increasingly doubly-dispersive and OFDM tones lose orthogonality at the receiver. Hence in recent years there have been considerable research [2, 3, 4, 5] in the area of finding theoretical basis functions for such channels. The results suggest signaling with short-time Fourier (STF) basis functions. To understand their properties, assume that an STF basis function $s(t)$ extends from time T_0 to time $T_0 + \tau$ and its spectrum $S(f)$ lies between frequencies f_0 and f_1 . Then

the channel gains from T_0 to $T_0 + \tau$ and over frequencies f_0 to f_1 would be highly correlated. Let us call this region specified by (T_0, τ, f_0, f_1) as a *time-frequency coherence subspace*. Thus An STF basis function falls in a *time-frequency coherence subspace* of the channel, within which the channel gains are highly correlated and these gains are independent across different ‘time-frequency’ coherence subspaces of the channel. Interestingly, such signaling leads to the classical *Block Fading* interpretation for even doubly-dispersive channels [2]. This opens the possibility of analyzing transmission over doubly-dispersive channels, in an information theoretic framework. This is the main aim in this thesis.

Apart from theoretical signal design there have also been attempts to design practical systems for transmitting information over doubly dispersive channels [6]. NTT-DoCoMo, a pioneer in this field, has developed a system, based on the signaling scheme called Variable Spread Factor-Orthogonal Frequency Code Division Multiplexing (VSF-OFCDM), [7, 8, 9, 10]. The key idea in VSF-OFCDM is *two-dimensional spreading* of symbols, which is explained as follows: Information symbols are first spread by a CDMA codeword. Each chip of the resultant sequence is allocated to the successive OFCDM symbols in the time domain (called time domain spreading) and to the successive sub-carriers in the frequency domain (called frequency domain spreading). The extent of spreading in time and frequency domains is adaptively adjusted depending upon the double dispersive nature of the channel. Note that fundamental characteristic of VSF-OFCDM is that of diversity as symbols are being replicated in time and across frequency subcarriers. For the multi-user case, VSF-OFCDM also serves as the multiple access scheme for the wireless medium, where different users are code-multiplexed by different CDMA codewords. Extensive simulation studies on VSF-OFCDM [7, 8, 9, 10] have been done and they conclude that this scheme is better than both OFDM and MC-CDMA [11, 12, 7, 9] as a wireless access scheme, in terms of maximizing the available sector throughput in a cellular system. However these studies do not offer any analytical understanding of the VSF-OFCDM scheme. This is another aspect which we focus upon in this thesis.

Motivated by such practical transmission schemes like VSF-OFCDM and theoretical

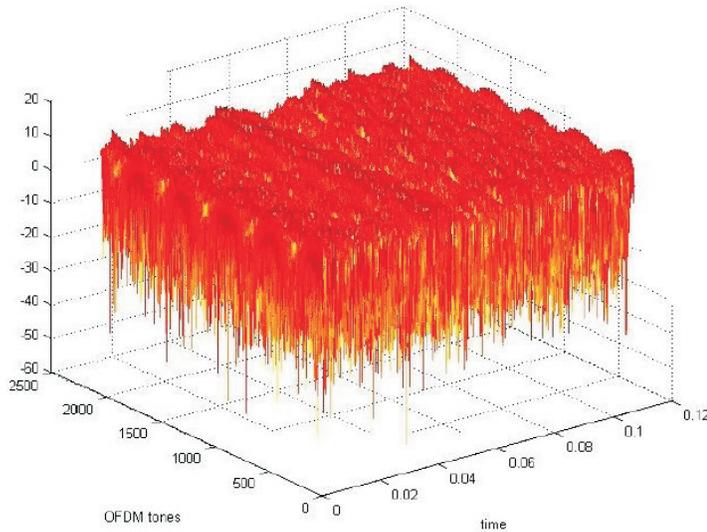


Figure 1.1: Two dimensional distribution of the SNR of a double dispersive fading process with $\tau_c = 3\mu\text{sec}$ and $f_D = 500\text{Hz}$

signal design techniques like STF, we propose a scheme to understand the information theoretic aspects of reliable communication over doubly-dispersive channels. We assume that STF basis functions are available of the transmitter and thus a block-fading model of the doubly-dispersive channel can be considered [2]. In the proposed scheme an information symbol is transmitted across several subcarriers with a total power constraint and along each subcarrier the symbol is spread with CDMA codewords. This feature implements diversity as in VSF-OFCDM. Note that in this context, the *time-frequency coherence subspace*, defined earlier in this section, comprises of the data modulated CDMA chips in time, along one frequency subcarrier. We study the single user and multi-user uplink communications for which our performance metrics are information theoretic achievable rates. This will be discussed in detail in Chapter 2. We derive the capacity bounds and the optimal power and codeword allocation strategies to achieve these bounds. In our work we consider three different scenarios each corresponding to a specific nature of channel state information (CSI) available at the transmitter [13]. The cases are perfect CSI, partial CSI characterized by one bit of channel information per subcarrier and no CSI. The receiver is assumed to have perfect CSI. These assumptions

are consistent with practical systems, where the receiver has an accurate estimate of the channel by processing training symbols. However CSI at transmitter is obtained from receiver feedback, which depending upon the quality of the feedback channel, and receiver resources expended for feedback, can vary from accurate to erroneous.

Chapter 2

VSF-OFDCM and Theoretical Two-Dimensional Spreading

2.1 Introduction

In this chapter, we first describe VSF-OFDCM as a wireless multiple access scheme. The transmitter and the multiplexing schemes are explained. The performance gains of this scheme vis-a-vis that of OFDM and MC-CDMA are discussed. The second part of this chapter, explains the two-dimensional spreading scheme as proposed in Chapter 1.

2.1.1 Brief Description of VSF-OFDCM

The targets for fourth generation wireless access are increased peak throughput for different radio environments and flexible packet access for various data types with various QoS requirements. VSF-OFDCM which is a wireless access scheme was proposed by NTT-DoCoMo [7, 8, 9, 10]. This scheme in conjunction with adaptive coding and modulation, MIMO etc provides data rates up to 100 Mbps in the forward link of a cellular system. In this access scheme data modulated symbols are spread by the spreading sequence, which is the combination of an orthogonal short channelization code and the cell-specific long scrambling code. Each chip of the resultant sequence is allocated to the successive OFDCM symbols in the time domain (called *time domain spreading*) and to the successive sub-carriers in the frequency domain (called *frequency domain spreading*). Therefore, the total spreading factor SF, is expressed as $SF = SF_{\text{Time}} \times SF_{\text{Freq}}$, where SF_{Time} and SF_{Freq} represent the spreading factors in the time and frequency domain spreading, respectively. In VSF-OFDCM, although the data rate is reduced by $1/SF$ due to replication, compared to the non-spreading cases like OFDM, the total data rate is increased by introducing the code multiplexing of different users with

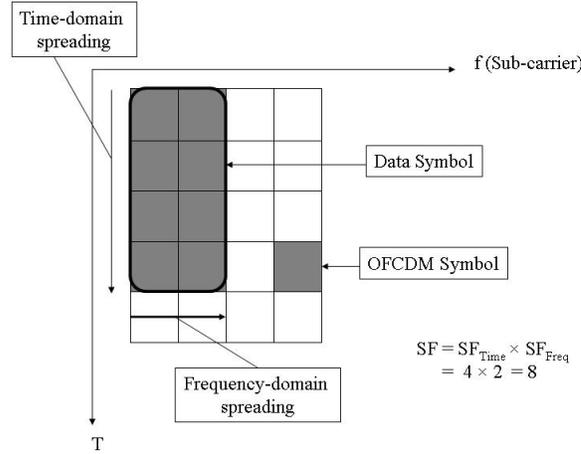


Figure 2.1: One Data Symbol is spread over 8 OFCDM symbols placed evenly in 2 subcarriers

different orthogonal short channelization codes. The concepts of time and frequency domain spreading are explained pictorially in Fig 2.1 [10], where one data symbol is spread in time with spread factor 4 and in frequency with spread factor 2. The parameters SF_{Time} and SF_{Freq} are adaptively controlled as per the cell structure, the channel load and the radio link conditions. For a fixed SF [8, 10] propose to prioritize the time domain spreading, for the following reasons,

- Within a frame duration, which is typically in the order of 0.5 – 1.0msec, channel variation in the time domain is slight.
- Meanwhile the channel variation in the frequency domain increases due to frequency-selective fading.
- Time-domain spreading leads to lower inter-code interference level.

Frequency domain spreading is applied mainly for low data rate, low SNR channels, in order to gain frequency diversity gains.

2.1.2 The Proposed Two-Dimensional Spreading Model

In this section we explain the two dimensional spreading scheme to model VSF-OFCDM as mentioned in Section 1. In this scheme each symbol is replicated over n_f subcarriers in frequency and along each subcarrier they are spread by a CDMA code of length n_t chips. The CDMA codes are assumed to be of unit norm. There is a total power constraint on the powers allocated in the n_f subcarriers.

A block fading channel model is assumed for reasons explained in Chapter 1. In the time domain, the channel remains constant for a block of n_t symbols. In the frequency domain, the channel gains on different subcarriers are independent and identically distributed. In this work, two types of channel models have been considered – uniform fading and Rayleigh fading. Although uniform fading is not characteristic of practical wireless channels, we consider this model because it leads to closed form expressions relating various system parameters like total power and water-filling level. These will be explained in Section 3.2. This gives an insight as to how these parameters are related to each other. The same correlation structure is expected to hold for practical Rayleigh fading channels, for which closed form expressions are not available. Details of these fading models appear in Sections 3.2.1 and 3.2.2 respectively.

Notation

Throughout this thesis we use uppercase boldfaced letters to denote matrices, lowercase boldface letters to denote vectors and lowercase letters to denote scalars. In particular we adapt the following notations: the subscript i denotes parameters of user i , the subscript j denotes a index of the j^{th} subcarrier and the subscript k denotes the time slot index along a subcarrier. Each time slot is occupied by an OFCDM symbol.

The symbol b_i denotes the i^{th} user's data symbol. We use $\mathbf{M}_i = \{m_{ijk}\}$, $1 \leq j \leq n_f$, $1 \leq k \leq n_t$ to denote the $n_f \times n_t$ spreading code matrix for the i^{th} user and \mathbf{m}_{ij} to denote the i^{th} user's spreading code along the j^{th} subcarrier. The matrices \mathbf{P}_i and \mathbf{H}_i are the $n_f \times n_f$ diagonal matrices of powers and channel gains respectively for user i , i.e. $\mathbf{P}_i = \text{diag}[p_{i1}, p_{i2}, \dots, p_{in_f}]$ and $\mathbf{H}_i = \text{diag}[h_{i1}, h_{i2}, \dots, h_{in_f}]$ where p_{ij} is the

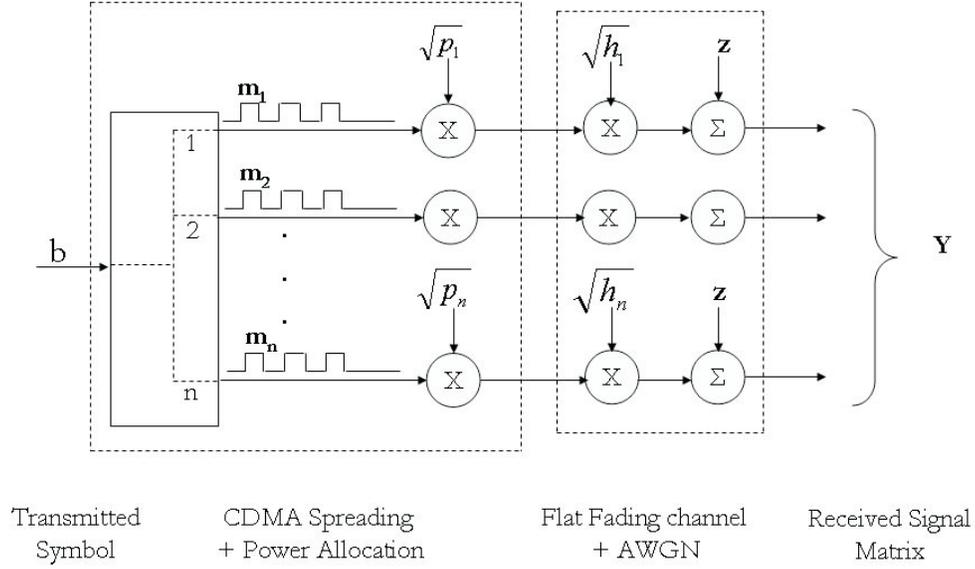


Figure 2.2: Transmitter Block Diagram

power user i puts in the j^{th} subcarrier and $\sqrt{h_{ij}}$ is the channel gain of user i in the j^{th} subcarrier. Note that since \mathbf{P}_i and \mathbf{H}_i are diagonal matrices, the vectors containing their diagonal elements contain the same information. We denote these vectors as \mathbf{p}_i and \mathbf{h}_i respectively. We sometimes use the subscript '0' for a random vector to denote its value, for example \mathbf{h}_0 denotes a particular realization of the channel state random vector. Finally, \mathbf{X}_i , \mathbf{Y}_i are the transmitted and received signal matrices and \mathbf{Z}_i is the AWGN noise matrix for user i . Each of these matrices has dimension $n_f \times n_t$.

Single User Transmission

For the single user transmission equations, we can drop the user index user i for notational simplicity. Let the received signal matrix \mathbf{Y} for the user be denoted as

$$\mathbf{Y} = \mathbf{H}^{\frac{1}{2}} \mathbf{P}^{\frac{1}{2}} \mathbf{M} b + \mathbf{Z}, \quad (2.1)$$

$$= \begin{bmatrix} \sqrt{h_1} & & & \\ & \sqrt{h_2} & & \\ & & \ddots & \\ & & & \sqrt{h_{n_f}} \end{bmatrix} \begin{bmatrix} \sqrt{p_1} & & & \\ & \sqrt{p_2} & & \\ & & \ddots & \\ & & & \sqrt{p_{n_f}} \end{bmatrix} \begin{bmatrix} \mathbf{m}_1^T \\ \mathbf{m}_2^T \\ \vdots \\ \mathbf{m}_{n_f}^T \end{bmatrix} b + \mathbf{Z}, \quad (2.2)$$

where \mathbf{m}_j is the CDMA code along subcarrier j represented as a column vector. Equivalently we can write

$$\mathbf{Y} = \tilde{\mathbf{X}}\mathbf{b} + \mathbf{Z}, \quad (2.3)$$

where

$$\tilde{\mathbf{X}} = \begin{bmatrix} \sqrt{h_1 p_1} \mathbf{m}_1^T \\ \sqrt{h_2 p_2} \mathbf{m}_2^T \\ \vdots \\ \sqrt{h_{n_f} p_{n_f}} \mathbf{m}_{n_f}^T \end{bmatrix}. \quad (2.4)$$

The problem statement is as follows, *Given a particular fading distribution, what is the optimal power allocation and codeword assignment to maximize the achievable rates.* In Chapters 3 and 4, we solve this problem for the cases of perfect channel state information and imperfect channel state information at the receiver respectively.

Multi User Transmission

In this section we introduce multiple users who communicate simultaneously to a common receiver, over a multiple access channel. The received signal matrix at the receiver is given by

$$\mathbf{Y} = \sum_{i=1}^N \mathbf{H}_i^{\frac{1}{2}} \mathbf{P}_i^{\frac{1}{2}} \mathbf{M}_i b_i + \mathbf{Z}. \quad (2.5)$$

As seen from Equation (2.5), the total number of users in the system is N . Let us define the matrix $\tilde{\mathbf{X}}_i := \mathbf{H}_i^{\frac{1}{2}} \mathbf{P}_i^{\frac{1}{2}} \mathbf{M}_i$. This can be written explicitly in terms of \mathbf{m}_{ij} , (user i 's spreading code along the j^{th} subchannel), as

$$\tilde{\mathbf{X}}_i = \begin{bmatrix} \sqrt{h_{i1} p_{i1}} \mathbf{m}_{i1}^T \\ \sqrt{h_{i2} p_{i2}} \mathbf{m}_{i2}^T \\ \vdots \\ \sqrt{h_{in_f} p_{in_f}} \mathbf{m}_{in_f}^T \end{bmatrix}. \quad (2.6)$$

We thus obtain

$$\mathbf{Y} = \sum_{i=1}^N \tilde{\mathbf{X}}_i b_i + \mathbf{z}. \quad (2.7)$$

Note that in both the single-user and multi-user cases, the total number of signalling dimensions are given by $n_{\text{dim}} = n_f n_t$.

The problem statement is as follows, *Given a particular fading distribution, what is the optimal power allocation vector and codeword matrix assignment for all users to maximize the sum of achievable rates.* In Chapters 3 and 4, we solve this problem for the cases of perfect channel state information and imperfect channel state information at the receiver respectively.

Chapter 3

Achievable Rates for Perfect CSI at Transmitter and Receiver

3.1 Introduction

The capacity for fading channels has multiple definitions based on the various assumptions about the nature of fading and availability of the fading knowledge at transmitter and receiver. In this section we assume that the transmitter and receiver are both equipped with the instantaneous values of the channel states, which remain constant during the transmission of a block of $n_f n_t$ symbols. If the channel realization changes from block to block, the ergodic capacity is defined as the average of the achievable rates for each channel realization [14]. For the single subcarrier case with no spreading Goldsmith et al [14] gives the expression for ergodic capacity and a strategy for achievability, which is based upon encoding data by a variable rate, variable power codebook. Caire et al [15] proposes a fixed rate coding scheme, with dynamic power-scaling just prior to transmission. Our problem differs from [14],[15] as we have multiple subcarriers to transmit information. CDMA spreading along each subcarrier is another extra property of our model. In this section we take the approach of [15] to compute bounds on the ergodic capacity of the two dimensional spreading system. The solution of the ergodic capacity maximization problem involves averaging over the distribution of the channel states at the subcarriers to decide the amount of power to allocate to a subcarrier for a given channel state. However any policy that implements this averaging leads to large delays which might not be acceptable for practical application requirements. In order to reduce delays we have to abandon the averaging operation to calculate power and instead consider a fixed power budget to be allocated over the different subchannels,

for a given channel state vector. Thus a key subproblem of the ergodic capacity maximization would be to consider a fixed realization of the channel state vector $\mathbf{h} = \mathbf{h}_0$ and a total power budget P to be allocated over all the subcarriers for that channel state vector. The problem is to allocate this power optimally between the subchannels so as to maximize the mutual information expression between the output and input of the channel for the given value, $\mathbf{h} = \mathbf{h}_0$. For future reference we term this problem as optimization of the maximum mutual information (MMI), for a given channel state.

3.2 Analysis of the Single User Transmission

The single user transmission is governed by Equation (2.3). It is difficult to calculate entropies from the matrices in Equation (2.3). However, we note that a matrix or vector is only a specific arrangement of numbers, it carries no information per se. So the rearrangement of elements of a matrix into a long vector doesn't result in any loss of information. Hence, we form vectors \mathbf{y} , $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{z}}$ by stacking together rows of \mathbf{Y} , $\tilde{\mathbf{X}}$ and \mathbf{Z} respectively. The resulting system becomes

$$\mathbf{y} = \tilde{\mathbf{x}}b + \mathbf{z}. \quad (3.1)$$

Assuming that $E[b^2] = 1$ and that the gains $\tilde{\mathbf{x}}$, are given (fixed), the maximum mutual information (MMI) is given by [16] as,

$$R(\tilde{\mathbf{x}}) = \frac{1}{2} \log \left(1 + \frac{|\tilde{\mathbf{x}}|^2}{\sigma^2} \right), \quad (3.2)$$

where $E[\mathbf{z}\mathbf{z}^T] = \sigma^2 I$. Substituting for $\tilde{\mathbf{x}}$ the expression for the maximum mutual information (MMI) for a given set of signature sequences with unit norm rows ($|\mathbf{M}_j|^2 = 1$ for all j) and a fixed set of channel gains becomes,

$$R(\mathbf{h}, \mathbf{u}) = \frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j=1}^{n_f} h_j p_j(\mathbf{u}) \right), \quad (3.3)$$

$$\text{where } \mathbf{h} = [h_1, h_2, \dots, h_{n_f}], \quad (3.4)$$

where \mathbf{u} is the CSI which is conveyed to the transmitter, called henceforth as CSIT as opposed to CSIR which stands for CSI at receiver. The transmitted powers in

Equation (3.3) depend on \mathbf{u} and is thus denoted by $p_j(\mathbf{u})$. In this section we specifically analyze power allocation policies for $\mathbf{u} = \mathbf{h}$, which the case of perfect CSIT. Note also that CDMA spreading used along subcarriers has no effect on the ergodic capacity, as the terms spreading matrix \mathbf{M}_j doesn't appear in the term $R(\mathbf{h}, \mathbf{u})$. This is consistent with the observation that CDMA doesn't buy us in terms of capacity for single user transmission [17]. The real gains of CDMA spreading would show up in the multi-user case.

Note that our transmission structure spreads signal energy in the dual dimensions of frequency and time which has a similar flavor to MIMO systems which use the dual dimensions of space and time [18].

Let $\mathbf{p}(\mathbf{u}) = [p_1(\mathbf{u}), p_2(\mathbf{u}), \dots, p_{n_f}(\mathbf{u})]$. The ergodic capacity maximization problem can be formulated along the lines of the approach outlined in [15], as

$$C_{\text{CSI}} = \max_{\mathbf{p}(\mathbf{u})} \int \cdots \int R(\mathbf{h}, \mathbf{u}) f(\mathbf{u}, \mathbf{h}) d\mathbf{u} d\mathbf{h}, \quad (3.5)$$

$$\text{s.t.} \sum_{j=1}^{n_f} \int \cdots \int p_j(\mathbf{u}) f(\mathbf{u}) d\mathbf{u} = \bar{P}, \quad p_j(\mathbf{u}) \geq 0, \quad (3.6)$$

where $R(\mathbf{h}, \mathbf{u})$ was defined as the maximum mutual information (MMI) in Equation (3.3). Since $\mathbf{u} = \mathbf{h}$, $f(\mathbf{u}, \mathbf{h}) = f(\mathbf{u}|\mathbf{h}) f(\mathbf{h}) = \delta(\mathbf{u} - \mathbf{h}) f(\mathbf{h})$. Thus the maximization in Equation (3.5) simplifies to:

$$C_{\text{PCSI}} = \max_{\mathbf{p}(\mathbf{h})} \int \cdots \int R(\mathbf{h}, \mathbf{h}) f(\mathbf{h}) d\mathbf{h} \quad (3.7)$$

$$\text{s.t.} \sum_{j=1}^{n_f} \int \cdots \int p_j(\mathbf{h}) f(\mathbf{h}) d\mathbf{h} = \bar{P}, \quad p_j(\mathbf{h}) \geq 0. \quad (3.8)$$

Note that the subscript 'PCSI' denotes perfect CSI in both transmitter and receiver. In every time epoch, when the CSI is revealed to the transmitter, the optimal solution to (3.7) is to transmit only in that subcarrier which has the highest channel gain (henceforth referred to as the *best subchannel*) and not to transmit in the others. In other words transmit in subchannel $i^* = \arg \max\{h_1, h_2, \dots, h_{n_f}\}$. Henceforth the random variable h_{i^*} will be denoted by h_n^* , where n is the number of random variables over which the maximization operation has been performed. In this case is $n = n_f$, the number of subcarriers. Note that i^* may vary from one epoch to another but the

statistics of h_n^* stays the same. The amount of power the transmitter puts in the best subchannel (denoted by $p(h_n^*)$) is obtained by waterfilling over the channel gain and it turns out to be

$$p(h_n^*) = \left(\frac{1}{\lambda} - \frac{\sigma^2}{h_n^*} \right)^+ . \quad (3.9)$$

The value of λ is found by substituting for $p(h_n^*)$ from Equation (3.9) in Equation (3.8), which yields,

$$\int \left(\frac{1}{\lambda} - \frac{\sigma^2}{\gamma} \right)^+ f_{h_n^*}(\gamma) d\gamma = \bar{P}. \quad (3.10)$$

Hence the expression for maximum mutual information, turns out to be

$$R_{\text{PCSI}}(h_n^*) = \frac{1}{2} \log \left(1 + \frac{h_n^* p(h_n^*)}{\sigma^2} \right), \quad (3.11)$$

$$= \frac{1}{2} \log \left(\frac{h_n^*}{\lambda \sigma^2} \right). \quad (3.12)$$

Note that in Equation (3.11) we have slightly abused the notation of MMI and denoted $R(\mathbf{h}, \mathbf{h})$ by $R(h_n^*)$ as in the perfect CSIT problem, h_n^* is the fundamental parameter which determines MMI.

A brief note on the significance of Equation (3.10) is timely at this juncture. The higher the value of h_n^* , the better are the achievable rates. Equation (3.9) shows that transmission takes place only if the channel state h_n^* is above the threshold $\lambda \sigma^2$. So a low value of λ implies that this threshold is low and so transmission could take place in a relatively inferior channel, leading to lower values of achievable rates. However keeping a high value of λ can also reduce the average achievable rate value as the encountered channel state h_n^* may be below the threshold for most of the times and consequently no transmission would take place. Solution of Equation (3.10) gives the optimal value of the threshold λ .

To summarize, the optimal transmission strategy is twofold:

- First choose the best subchannel to put the power in. This corresponds to frequency domain waterfilling.
- Next do waterfilling in time over different realizations of the channel states. This in essence is to do waterfilling over the distribution of h_n^* .

A note about the statistics of best subchannel is appropriate at this point. Let $f_h(\gamma)$ and $F_h(\gamma)$ be the PDF and CDF of h respectively, where h is the random variable representing the channel coefficient in a subcarrier. Then assuming i.i.d. channel coefficients across subcarriers, the PDF of h_n^* is given by

$$f_{h_n^*}(\gamma) = n [F_h(\gamma)]^{n-1} f_h(\gamma). \quad (3.13)$$

Let us now consider some specific cases of fading and investigate how capacity can be evaluated.

3.2.1 Uniform Fading

Let h_i , $1 \leq i \leq n_f$, be i.i.d. uniformly distributed on $[0, A]$. This type of fading distribution is not experienced in real world channels but nevertheless is of theoretical importance. Under uniform fading assumptions the Equations (3.10) and (3.11) can be evaluated to yield closed form expressions relating λ and \bar{P} . This yields an analytical understanding of the nature of their mutual dependencies. This in turn helps to predict similar dependencies between these variables for more realistic fading scenarios, for which the direct solutions of Equations (3.10) and (3.11) do not lead to closed form expressions. Hence, uniform fading assumptions have been used in the works of Shamai and Caire [15] and Mecking [19].

For uniform fading, the pdf of the random variable h_n^* is

$$f_{h_n^*}(\gamma) = \begin{cases} n\gamma^{n-1}/A^n & 0 < \gamma < A, \\ 0 & \text{otherwise.} \end{cases} \quad (3.14)$$

Substituting for $f_{h_n^*}(\gamma)$ in Equation (3.10) and evaluating the integral yields

$$\left(\frac{1}{n-1} \frac{\sigma^{2n}}{A^n}\right) \lambda^n - \left(\bar{P} + \frac{n}{n-1} \frac{\sigma^2}{A}\right) \lambda + 1 = 0. \quad (3.15)$$

Note that the limits of the integral for which the integrand is defined and non-zero is $[\lambda\sigma^2, A]$. Thus a necessary condition for the existence of a solution is that $\lambda\sigma^2 \leq A$ or $0 \leq \lambda \leq A/\sigma^2$. Equation (3.15) can be solved numerically or via Matlab for given values of \bar{P} , A and n . However for any general values of these parameters is the existence of a

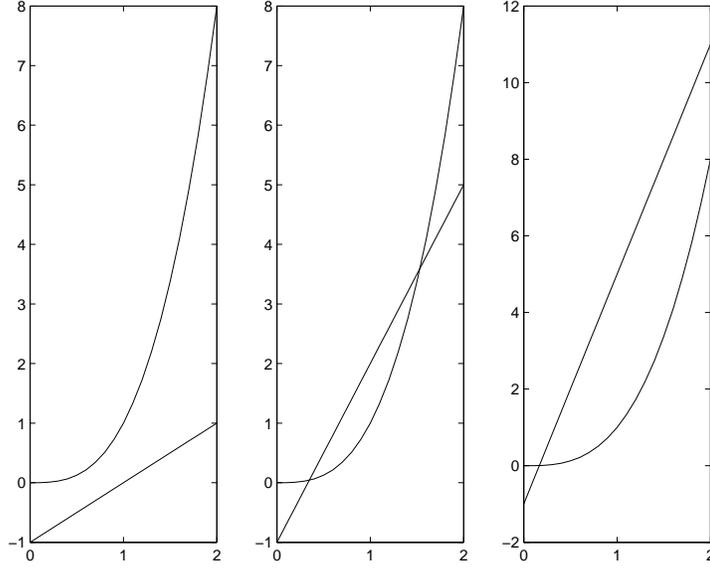


Figure 3.1: Different solutions for $\lambda^n = a\lambda + b$

solution always guaranteed? To answer this question consider the function

$$g(\lambda) = \left(\frac{1}{n-1} \frac{\sigma^{2n}}{A^n} \right) \lambda^n - \left(\bar{P} + \frac{n}{n-1} \frac{\sigma^2}{A} \right) \lambda + 1. \quad (3.16)$$

Now note that

$$g(0) = 1, \quad (3.17)$$

$$g\left(\frac{A}{\sigma^2}\right) = -\frac{A\bar{P}}{\sigma^2}. \quad (3.18)$$

Since $g(0)$ and $g(A/\sigma^2)$ are of opposing signs, the equation $g(\lambda) = 0$ has an odd number of solutions in the range $[0, A/\sigma^2]$. We can further show that there is exactly one solution. Equation (3.15) can be recast in the form $\lambda^n = a\lambda + b$ where $b < 0$. So the real solutions of the equation (3.15) are the points of intersection of the curves $y = \lambda^n$ and $y = a\lambda + b$. Now exactly one of three cases can arise as shown in Figure 3.1. Hence the number of real solutions is either 0, 1, or 2. However it was proved that there are an odd number of solutions in the interval $[0, A/\sigma^2]$. Hence there must be exactly one real solution in this interval. There might be another solution in the interval $[A/\sigma^2, \infty]$ but this is not of interest in the given problem. The MMI is found out from substituting for λ in equations (3.9) and (3.11).

Note that the capacity C_{CSI} is a function of n , the number of subcarriers and it is interesting to see the limiting behavior of the capacity as n tends to infinity. We observe that as $n \rightarrow \infty$, $f_{h_n^*}(\gamma) \rightarrow \delta(\gamma - A)$. Hence the capacity is same as that of a channel with constant gain, A :

$$C_{\text{CSI-Const}} = \frac{1}{2} \log \left(1 + \frac{A\bar{P}}{\sigma^2} \right). \quad (3.19)$$

3.2.2 Rayleigh Fading

Let the channel gains be Rayleigh distributed. Then the h 's are exponentially distributed with PDF and CDF expressions given by,

$$f_h(\gamma) = \frac{1}{\gamma_a} e^{-(\gamma/\gamma_a)}, \quad F_h(\gamma) = 1 - e^{-(\gamma/\gamma_a)}, \quad \gamma \geq 0. \quad (3.20)$$

From Equations (3.13) and (3.20),

$$f_{h_n^*}(\gamma) = \frac{n}{\gamma_a} \left(1 - e^{-(\gamma/\gamma_a)} \right)^{n-1} e^{-(\gamma/\gamma_a)}, \quad \gamma \geq 0. \quad (3.21)$$

Substituting for $f_{h_n^*}(\gamma)$ in Equation (3.10) does not lead to a closed form equation relating λ to \bar{P} . The integral is solved numerically and the results are explained in Section 3.4.2. Recalling that the optimal policy is to waterfill over the distribution of h_n^* , it is instructive to investigate how the the distribution of h_n^* varies with n . Appendix B, presents a detailed analysis of the nature of the random variable h_n^* , and how its statistics change with n , the number of subcarriers. One of the results from Appendix B is that $P[h_m^* > \gamma] > P[h_n^* > \gamma]$ for all $m > n$. This means that probabilistically the actual value of h_n^* is going to be larger as we increase the number of subcarriers. Since the ergodic capacity is directly proportional to the value of h_n^* , as shown in Equation (3.11), the capacity increases as the number of subcarriers is increased. This is shown numerically in Section 3.4.2.

3.3 Optimal Allocation for Colored Noise

In this section, we derive optimal power allocation policies when the additive noise term in Equation (2.1) is assumed to have a colored spectrum. Such a scenario can arise in

multi-user transmission, when the receiver for user i models the interference from other users as Gaussian noise. This scenario is the generalization of the white noise case considered in Section 3.2.

To simplify the analysis, we consider a specific model where no spreading is done along time domain. This allows us to formulate the capacity maximization problem in a way different from the approach followed in Section 3.2 where a long vector $\tilde{\mathbf{x}}$ was formed by stacking the columns of $\tilde{\mathbf{X}}$. The new approach retains the decoupling of the matrices \mathbf{H} and \mathbf{P} , whereas in previous approach of Section 3.2, they were coupled in the matrix $\tilde{\mathbf{X}}$. Since the dimensionality along the time domain (number of spreading symbols) is now 1, the matrix channel of Equation (2.1) can be reduced to a vector channel as follows,

$$\mathbf{Y} = \mathbf{H}^{\frac{1}{2}} \mathbf{P}^{\frac{1}{2}} b + \mathbf{Z}, \quad (3.22)$$

$$\mathbf{y} = \mathbf{H}^{\frac{1}{2}} \mathbf{p}^{\frac{1}{2}} b + \mathbf{z}, \quad (3.23)$$

where $\mathbf{p}^{\frac{1}{2}}$ is the power vector $[\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_{n_f}}]$ formed by the diagonal entries of matrix \mathbf{P} . Vectors \mathbf{y} and \mathbf{z} are similarly defined as the diagonal vectors of the matrices \mathbf{Y} and \mathbf{Z} respectively. Note that \mathbf{p} is dependent upon the channel state vector $\mathbf{h} = [h_1, h_2, \dots, h_{n_f}]$ and is henceforth referred to as $\mathbf{p}(\mathbf{h})$.

Let the noise covariance matrix be given by $\mathbf{S}_{\mathbf{zz}} = E[\mathbf{zz}^T]$. Following the approach of Section 3.2, we can define the maximum mutual information (MMI), for a given set of channel coefficients \mathbf{h} and power vector $\mathbf{p}(\mathbf{h})$, as

$$R_{\text{PCSI}}(\mathbf{h}, \mathbf{h}) = \frac{1}{2} \log \left(\det \left[\mathbf{S}_{\mathbf{zz}} + \mathbf{H}^{\frac{1}{2}} \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T (\mathbf{H}^{\frac{1}{2}})^T \right] \right). \quad (3.24)$$

The ergodic capacity maximization of Equation (3.24) can be written as,

$$C_{\text{PCSI}} = \max_{\mathbf{p}(\mathbf{h})} \int \dots \int R_{\text{PCSI}}(\mathbf{h}, \mathbf{h}) f(\mathbf{h}) d\mathbf{h} \quad (3.25)$$

$$\text{s.t.} \int \dots \int \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) f(\mathbf{h}) d\mathbf{h} = \bar{P}, \quad p_j(\mathbf{h}) \geq 0. \quad (3.26)$$

As seen in Section 3.2, the above maximization is equivalent to finding the direction of the optimal square root power vector $\mathbf{p}^{\frac{1}{2}}(\mathbf{h})$ and its norm given a particular realization of the channel state vector \mathbf{h} . The direction can be found out by maximizing the

argument of the logarithm function in Equation (3.25), which can be simplified as,

$$\max \det \left[\mathbf{S}_{\mathbf{z}\mathbf{z}} + \mathbf{H}^{\frac{1}{2}} \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T (\mathbf{H}^{\frac{1}{2}})^T \right], \quad (3.27)$$

$$= \max \det \left[\mathbf{I} + \mathbf{S}_{\mathbf{z}\mathbf{z}}^{-1} \mathbf{H}^{\frac{1}{2}} \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T (\mathbf{H}^{\frac{1}{2}})^T \right], \quad (3.28)$$

$$= \max \left(1 + \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T (\mathbf{H}^{\frac{1}{2}})^T \mathbf{S}_{\mathbf{z}\mathbf{z}}^{-1} \mathbf{H}^{\frac{1}{2}} \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) \right). \quad (3.29)$$

In the above derivation we have used the result $\det[I + AB] = \det[I + BA]$. Now the ergodic capacity maximization is not mathematically tractable and hence we focus on the key subproblem mentioned in Section 3.1. We seek to optimize the MMI for a given instance of the channel state vector $\mathbf{h} = [h_1, \dots, h_{n_f}] = \mathbf{h}_0$, under a total power constraint \bar{P} . The maximization problem can thus be posed as,

$$\max_{\mathbf{p}^{\frac{1}{2}}(\mathbf{h})} \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T (\mathbf{H}^{\frac{1}{2}})^T \mathbf{S}_{\mathbf{z}\mathbf{z}}^{-1} \mathbf{H}^{\frac{1}{2}} \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) \quad (3.30)$$

$$\text{s.t. } \mathbf{p}^{\frac{1}{2}}(\mathbf{h})^T \mathbf{p}^{\frac{1}{2}}(\mathbf{h}) = \bar{P}, \quad \mathbf{p}(\mathbf{h}) > 0 \quad (3.31)$$

Note that without the constraint $\mathbf{p}(\mathbf{h}) > 0$ the maximization in Equation (3.30) is equivalent to a standard Rayleigh quotient problem [20] of maximizing $R(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} / \mathbf{x}^T \mathbf{x}$. The solution is to let \mathbf{x} be that eigenvector of \mathbf{A} which corresponds to the largest eigenvalue (henceforth referred to as the *maximum eigenvector*). When the constraint $\mathbf{x} > 0$ is added, the problem is, in general, non-trivial to solve analytically. The *maximum eigenvector* solution can be wrong as the components of the *maximum eigenvector* can be negative, thus violating the positivity constraint. The most general solution is outlined in Appendix C. It is shown that the optimal \mathbf{x} lies in the space of the eigenvectors of all the principal sub-matrices of \mathbf{A} , with zeros padded to these eigenvectors to produce the vector \mathbf{x} of length n . However there are specific cases for which the solution again turns out to be the *maximum eigenvector*. One such case is when \mathbf{A} is *positive*, which means that all the entries of \mathbf{A} are positive (this is different from positive definiteness). For this matrix the Perron-Frobenius theorem, [20] states that the *maximum eigenvector* has all positive components and hence the positivity constraint is automatically satisfied. Another case is when $\mathbf{A} \in \mathcal{R}^{2 \times 2}$, in which case due to the orthogonality of the eigenvectors in a two dimensional space, one of the eigenvectors

will always have both entries as positive (or both negative) and that would correspond to the *maximum eigenvector*

In the present case, note that $\mathbf{A} = (\mathbf{H}^{\frac{1}{2}})^T \mathbf{S}_{\mathbf{z}\mathbf{z}}^{-1} \mathbf{H}^{\frac{1}{2}}$, where \mathbf{H} is diagonal. Depending upon the structure of $\mathbf{S}_{\mathbf{z}\mathbf{z}}^{-1}$, the matrix \mathbf{A} might or might not be positive. If \mathbf{A} is positive the optimal $\mathbf{p}^{\frac{1}{2}}(\mathbf{h})$ turns out to be the eigenvector corresponding to the maximum eigenvalue \mathbf{A} , the effective channel gain and background noise matrix.

Note that for $\mathbf{S}_{\mathbf{z}\mathbf{z}} = I$, aligning $\mathbf{p}^{\frac{1}{2}}(\mathbf{h})$ along the *maximum eigenvector* corresponds to transmitting along the *best subchannel* as derived in Section 3.2. Note that even in the most general case, the optimal strategy is transmission along a single dimension, the direction along the *maximum eigenvector* of one of the principal sub-matrices. This direction can be interpreted as the *best effective subchannel*.

3.4 Simulation Results for Single-User

In this section we take some specific fading scenarios and try to evaluate the achievable rates. In this section we plot the quantity maximum mutual information (MMI) as defined in Section 3.2. At each time instant an independent realization of the channel is generated, and the transmitter goes through the steps mentioned in Section 3.2. For the Rayleigh Fading case we also show the variation of λ with total power.

3.4.1 Uniform Fading

The fading is assumed to be uniform in the range $[0, A]$. The various simulation parameters are listed in Table 3.1. Note that the term SNR is used to denote the quantity \bar{P}/σ^2 , where \bar{P} is the average power, as defined in Equation (3.6). It is not the instantaneous signal to noise ratio in a particular subchannel for a given fading state, which would be $p_j(\mathbf{h})/\sigma^2$ or $p_j(\mathbf{u})/\sigma^2$, depending upon the nature of CSI at transmitter.

Figures 3.2 and 3.3 plot the variations in the (MMI) sequence for two different values of SNR. We make the following observations.

- The MMI (and hence the ergodic capacity) increases as the number of subcarriers increase as the distribution of $f_{h_n^*}(\gamma)$ shifts to the right with more subcarriers.

Parameter	Value
A	1
No. of Subcarriers n_f	2,10,100
SNR	1 dB, 10 dB

Table 3.1: Simulation Parameters for Uniform Fading

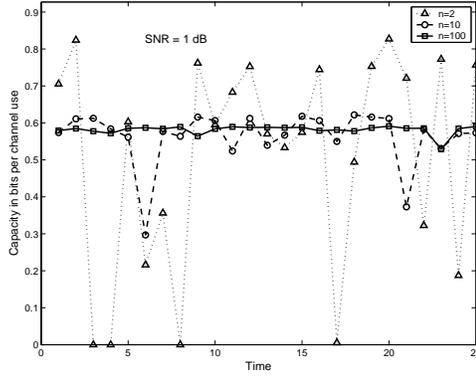


Figure 3.2: MMI variations with time for SNR = 1 dB

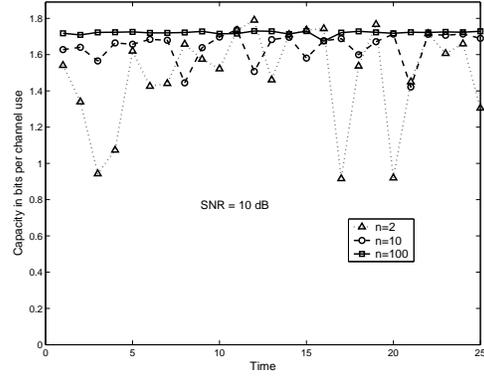


Figure 3.3: MMI variations with time for SNR = 10 dB

- The fluctuations in MMI decrease as the number of subcarriers increase, as the distribution of $f_{h_n^*}(\gamma)$ gets more narrow with more subcarriers. The variances of the MMI sequence along time are recorded in Table 3.2.
- In the regime of large number of subcarriers, the simulation results match the theoretical values of capacity, $C_{\text{CSI-Const}}$, obtained by substituting $A = 1$ and SNR = 1 dB, or SNR = 10 dB in Equation (3.19). The simulation capacity $\bar{C}(100)$ corresponds to the time average for 25 time instants, of the MMI sequence, for 100 subcarriers. The values are recorded in Table 3.3.

3.4.2 Rayleigh Fading

Each subcarrier undergoes independent Rayleigh Fading so the distribution of the h_i s are exponential. The mathematical preliminaries of this case has been presented in Sections 3.2.2 and Appendix B. The various simulation parameters are listed in Table 3.4. Equation (3.10) relates λ to \bar{P} and has to be solved numerically. Figure (3.4) gives the results of this numerical computation. The following observations are readily made

	SNR 1 dB	SNR 10 dB
n=2	0.0612	0.0726
n=10	0.0024	0.0068
n=100	5.709e-5	1.217e-4

Table 3.2: Variances in MMI Sequence

	$C_{\text{CSI-Const}}$	$\overline{C}(100)$
SNR 1 dB	0.5878	0.5835
SNR 10 dB	1.7297	1.7223

Table 3.3: Theoretical and Simulated Values of Ergodic Capacities for Large n

Parameter	Value
γ_a	1
No. of Subcarriers n_f	2,10,100
SNR	1 dB, 10 dB

Table 3.4: Simulation Parameters for Rayleigh Fading

- For the same value of n , a decrease in Total Average Power \overline{P} , leads to an increase in the value of λ , which means an increase in the threshold above which transmission takes place. This implies that when the \overline{P} less then the transmitter has to wait for a really good channel to transmit reliably. This also suggests that the transmission frequency would be less and transmission would be bursty in nature.
- For the same value of Total Power, increasing the number of subcarriers leads to an increase in λ , which raises the threshold. This is because by increasing subcarriers the probability of encountering a higher value of the encountered channel state h_n^* is increased.

Figures 3.5 and 3.6 show how the MMI sequence varies with time. The MMI sequence values (and hence the ergodic capacity) increases with number of subcarriers and increasing SNR. For SNR = 1 dB and $n = 2$ there are some instances when there is no transmission as the actual value of encountered channel state h_n^* is below the threshold λ .

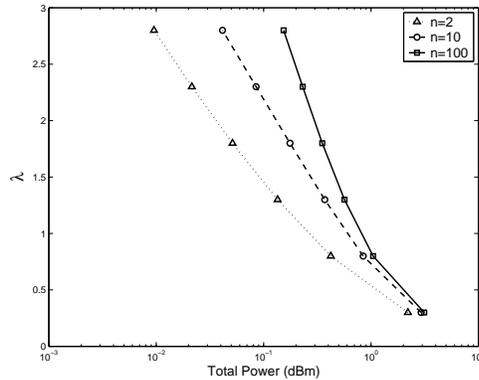


Figure 3.4: Water-Filling parameter λ vs Total Power for n subcarriers: $n = 2, 10, 100$

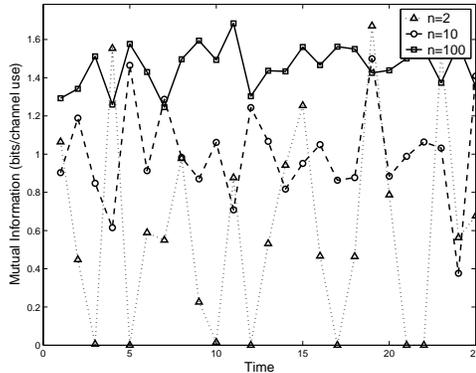


Figure 3.5: MMI variations with time for SNR = 1 dB

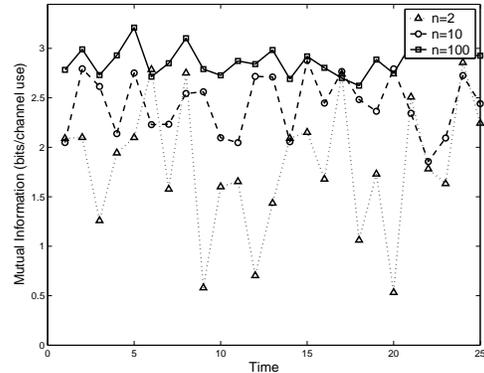


Figure 3.6: MMI variations with time for SNR = 10 dB

3.5 Analysis of the Multi-User Transmission

We assume that there are N users in the system. The multi-user transmission is governed by Equation (2.5). The metric we consider is maximization of the sum capacity. The problem of multi-user codeword assignment and power control has been studied for various contexts. Verdu [21] derived the capacity region for the single carrier CDMA symbol synchronous transmission, in a AWGN channel. Based on the work of [21] Rupp et al [22] gave the structure of the optimal codewords when all users had equal power constraints. Viswanath et al [23] extended the results for users with unequal power constraints and also formulated a procedure for constructing these codewords. Based upon the approach of [23] Kaya and Ulukus [24] solved the problem of optimal codeword selection and power allocation for a single carrier Rayleigh flat-fading channel. For

MIMO channels, the problem becomes transmission over multi-user vector channels. There has been considerable work in trying to find optimal strategies for such types of communications. The results suggest simultaneous water-filling by all the users, where each user treats the interference from other users as background noise [25] [26] [27]. This kind of solution is particularly appealing as it provides intuitive explanations about the nature of the problem and can be implemented in a distributed fashion by all the users. Our problem as stated in Section 2.1.2, involves transmission over multi-user matrix channels and the optimization is over power vectors and spreading matrices of multiple users. The problem is discussed in the next section.

3.5.1 Both Time and Frequency Domain Spreading

We start with the basic transmission equation as given in Equation 2.7, which is reproduced here for convenience:

$$\mathbf{Y} = \sum_{i=1}^N \tilde{\mathbf{X}}_i b_i + \mathbf{z}. \quad (3.32)$$

Recall that $\tilde{\mathbf{X}}_i$, was the matrix with the i^{th} row containing the symbol b_i times the power factor p_i modulated by the CDMA code along the i^{th} subcarrier. Using the argument of Section 3.2, we can stack the columns of $\tilde{\mathbf{X}}_i$, to form the long vector

$$\tilde{\mathbf{x}}_i = \left[\sqrt{h_{i1}p_{i1}}\mathbf{m}_{i1}^T \quad \sqrt{h_{i2}p_{i2}}\mathbf{m}_{i2}^T \quad \cdots \quad \sqrt{h_{in_f}p_{in_f}}\mathbf{m}_{in_f}^T \right]^T. \quad (3.33)$$

The matrix transmission in Equation 3.32 can thus be translated into a vector transmission given by,

$$\mathbf{y} = \sum_{i=1}^N \tilde{\mathbf{x}}_i b_i + \mathbf{z}, \quad (3.34)$$

$$= \mathbf{X}\mathbf{b} + \mathbf{z}. \quad (3.35)$$

In this case, matrix \mathbf{X} is of dimension $N \times n_{\text{dim}}$, where $n_{\text{dim}} = n_f n_t$ is the total number of available dimensions for signalling. In addition, \mathbf{X} has vector $\tilde{\mathbf{x}}_i$ as its i^{th} column and \mathbf{b} is the information symbol vector of size $N \times 1$. Recall that \mathbf{X} depends upon the power vectors $\mathbf{p}_i(\mathbf{h}_i)$ and the spreading code matrices \mathbf{M}_i for all the i users, $1 \leq i \leq N$. The maximum mutual information (MMI), for a given set of channel gains, is known

to be [21]

$$R_{\text{SUM-CSI}}(\mathbf{X}) = \frac{1}{2} \log (\det [\sigma^2 I_{n_{\text{dim}}} + \mathbf{X}\mathbf{X}^T]) \quad (3.36)$$

$$= \frac{1}{2} \log (\det [\sigma^2 I_N + \mathbf{X}^T \mathbf{X}]), \quad (3.37)$$

where the notation I_N denotes the $N \times N$ identity matrix. Note that the matrix $\mathbf{X}^T \mathbf{X}$ is of dimension $N \times N$ where N is the number of users. Let $[\mathbf{X}^T \mathbf{X}]_{ij}$ be the ij^{th} term of this matrix. This term gives the value of the cross correlation between the transmitted signals, of user i and user j , over all subchannels and can be expressed as

$$[\mathbf{X}^T \mathbf{X}]_{ij} = \begin{cases} \sum_{k=1}^{n_f} h_{ik} p_{ik} & i = j, \\ \sum_{k=1}^{n_f} \sqrt{h_{ik} h_{jk}} \sqrt{p_{ik} p_{jk}} \mathbf{m}_{ik}^T \mathbf{m}_{jk} & \text{otherwise.} \end{cases} \quad (3.38)$$

The ergodic sum capacity maximization problem can now be posed as,

$$C_{\text{SUM-CSI}} = \max_{\mathbf{X}} \int \cdots \int R_{\text{SUM-CSI}}(\mathbf{X}) f(\mathbf{X}) d\mathbf{X} \quad (3.39)$$

$$\text{s.t.} \int \cdots \int \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i)^T \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i) f(\mathbf{h}_i) d\mathbf{h}_i = \bar{P}_i, \quad 1 \leq i \leq N \quad (3.40)$$

$$p_{ij}(\mathbf{h}_i) \geq 0, \quad 1 \leq i \leq N, 1 \leq j \leq n_f. \quad (3.41)$$

In Equation (3.39) the optimization is shown over \mathbf{X} for sake of brevity. What it means is that the optimization is over the power vectors $\mathbf{p}_i(\mathbf{h}_i)$ and spreading matrices $\mathbf{M}_i(\mathbf{h}_i)$ for all the i users. However it is very difficult to find the optimal solution in analytical closed form. It is still an unsolved problem for any general n_t, n_f and N . In Section we outline some multi-user transmission schemes that are based on practical considerations like ease of implementation and calculate the achievable rates. In the optimal transmission strategy, transmitting along different subcarriers and spreading along each subcarrier are two separate optimizations. To understand how each individually affect the sum capacity, let us consider two special cases,

3.5.2 Only Time Domain Spreading

In this case, every user modulates their symbols with CDMA codewords and transmits the signal over a single subcarrier. Note that value of $n_f = 1$. Thus the optimization

of sum capacity is over the transmitted power values and codewords of the various users. The matrix channel of Equation (2.5) with output matrix \mathbf{Y} reduces to a vector channel, with vector output \mathbf{y} . The transmission model thus becomes

$$\mathbf{y} = \sum_{i=1}^N \sqrt{h_i p_i} \mathbf{m}_i b_i + \mathbf{z}. \quad (3.42)$$

This is the same model considered in [24], and hence the optimal solution is known. We reproduce Equation (13) of [24], which states the jointly optimal power and signature sequence allocation policy,

$$p_i^*(\mathbf{h}) = \begin{cases} \left(\frac{1}{\lambda_i} - \frac{\sigma^2}{h_i} \right), & \text{iff } i \in \Omega \\ 0, & \text{otherwise} \end{cases} \quad (3.43)$$

$$\mathbf{m}_i^{*T} \mathbf{m}_j^* = 0, i \neq j, \text{ for all } i, j \in \Omega \quad (3.44)$$

$$\Omega = \{i : \gamma_{[i]} > \sigma^2, i \leq \min\{N, n_t\}\}, \quad (3.45)$$

where $\gamma_i = h_i/\lambda_i$. The optimal policy states that only those users whose normalized channel gains γ_i are above the threshold σ^2 should be allowed to transmit in orthogonal channels. The number of such users can't exceed either n_t , the length of the spreading code in time or N , the total number of users in the system. The problem then reduces to independent single user transmissions of [14] for which the optimal solution for each user is to waterfill over the channel fading distribution.

3.5.3 Only Frequency Domain Spreading

We now consider the other extreme, when there is no CDMA spreading along the time domain ($n_t = 1$) and the users transmit their information along the n_f subcarriers. The matrix channel of Equation (2.5) with output matrix \mathbf{Y} again reduces to a vector channel, with vector output \mathbf{y} (as in Section 3.5.2) and the transmission model becomes

$$\mathbf{y} = \sum_{i=1}^N \mathbf{H}_i^{\frac{1}{2}} \mathbf{p}_i^{\frac{1}{2}} b_i + \mathbf{z}, \quad (3.46)$$

where $\mathbf{p}_i^{\frac{1}{2}} = [\sqrt{p_{i1}}, \sqrt{p_{i2}}, \dots, \sqrt{p_{in_f}}]$, the power vector of the i^{th} user. The corresponding maximum mutual information (MMI), for given channel state matrices for all the

users, can be expressed as

$$R_{\text{SUM-CSI}}(\mathbf{h}_1, \dots, \mathbf{h}_N) = \log \left(\det \left[I + \sum_{i=1}^N \mathbf{H}_i^{\frac{1}{2}} \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i) \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i)^T (\mathbf{H}_i^{\frac{1}{2}})^T \right] \right). \quad (3.47)$$

The ergodic sum capacity maximization problem can be expressed as,

$$C_{\text{SUM-CSI}} = \max_{\mathbf{p}_1, \dots, \mathbf{p}_N} \int \dots \int R(\mathbf{h}_1, \dots, \mathbf{h}_N) f(\mathbf{h}_1, \dots, \mathbf{h}_N) d\mathbf{h}_1 \dots d\mathbf{h}_N \quad (3.48)$$

$$\int \dots \int \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i)^T \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i) f(\mathbf{h}_i) \mathbf{h}_i \leq \bar{P}_i \text{ for all } i, \quad (3.49)$$

$$p_{ij}(\mathbf{h}_i) \geq 0, \quad 1 \leq i \leq N, 1 \leq j \leq n_f. \quad (3.50)$$

At this stage, we again consider the key subproblem of optimizing the MMI, for a given realization of the channel state vectors $\mathbf{h}_1, \dots, \mathbf{h}_N$, under a total power constraint P , for each of the N users,

$$\max R_{\text{SUM-CSI}} = \max_{\mathbf{p}_1, \dots, \mathbf{p}_N} R(\mathbf{h}_1, \dots, \mathbf{h}_N) \quad (3.51)$$

$$\mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i)^T \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i) = P, \text{ for all } i, \quad (3.52)$$

$$p_{ij}(\mathbf{h}_i) \geq 0, \quad 1 \leq i \leq N, 1 \leq j \leq n_f. \quad (3.53)$$

However this problem is non-convex in the power vectors $\mathbf{p}_i(\mathbf{h}_i)$, and hence a computationally efficient algorithmic solution like those proposed in [26] doesn't exist. In fact the problem is stated in a slightly different form in [27] where the authors claim that due to the non-convex nature of the problem, the simultaneous single-user iterative water-filling over the interference plus AWGN noise spectrum, need not converge to a global optimum.

Let us investigate the reason behind this phenomenon and how our scheme differs from those in [26] [25] for which an iterative water-filling procedure was optimal. In our proposed scheme of Chapter 2, we can view the transmitted signal vector as $\mathbf{x}_i = \mathbf{p}_i b_i$ (in absence of CDMA spreading). Note that this definition of \mathbf{x}_i restricts it to a set of positive real numbers, whereas the transmitted signals in general belong to the space of \mathcal{C}^n . Hence the transmitted signal in subcarrier i should be strictly expressed as $\sqrt{p_i} e^{-j\phi_i} b_i$. Note however that the phase part is immaterial for the capacity formulation and hence there is no loss of information in representing \mathbf{x}_i as $\mathbf{p}_i b_i$. Thus same information symbol

is transmitted in all the subcarriers and thus the transmitted signal is constrained to lie in an one-dimensional space ($\text{rank}(E[\mathbf{x}_i\mathbf{x}_i^T]) = 1$). This constraint is non-convex in nature. Whereas in [26] the transmitted signal vector \mathbf{x}_i , is such that different symbols are sent over different subcarriers, for which the transmitted covariance matrix had full rank. Hence the main problem with our scheme is that since the transmitted signal lies in an one-dimensional space it can't water-fill over all the other dimensions. This also implies that, the achievable rate performance of our scheme is going to be strictly sub-optimal to that in [26] as we have extra constraints in transmission.

While this constraint of one dimensional signalling may seem artificial in nature, we would like to point out that it is only so because the transmitter i was assumed to possess complete knowledge of channel gain vector \mathbf{h}_i . In fact our scheme is an example of full-diversity mode (as we do symbol replication in all subchannels) and diversity doesn't give higher rates for complete channel state information at the transmitter (CSIT). However complete CSIT might not be possible due to reasons cited in Chapter 4. In case of incomplete CSIT, there is always the possibility of outage when the instantaneous mutual information of the channel falls below the transmitted rate. In such cases diversity reduces the outage probability at the cost of lower transmitted rates. In fact it was shown in [28] that there is always a fundamental tradeoff in using available transmit dimensions for diversity and for multiplexing, i.e. transmitting independent information over different dimensions. We thus expect our scheme to perform better in cases of little or no CSIT, over transmission schemes which use all dimensions for sending independent information.

3.5.4 Some Proposed Multiuser Transmission Strategies

For the single user case, our performance metric had been ergodic sum capacity maximization. For the multi-user case we were concerned with the subproblem of optimizing the MMI for a given channel state $\mathbf{h}_1, \dots, \mathbf{h}_N$ for all the N users. In Section (3.5.3) we showed that this is still an open problem. Since the optimal scheme is unknown, in this section we propose some practical schemes for multi-user communications. The

motivation is to maximize the achievable rates and also to reduce the complexity involved in implementation of these schemes in practice. There is an inherent trade-off as the optimal schemes often lead to practical difficulties (like long delays for policies that optimize ergodic capacity). The metric for comparing these policies is the value of MMI they yield for given channel state vectors $\mathbf{h}_1, \dots, \mathbf{h}_N$. For a given policy we thus seek to optimize the MMI

Optimal Single User Transmission

In this strategy the transmitters of the N users are aware of their individual channel gains in all the n_f subchannels and each transmitter transmits along its own best subchannel. The motivation of this scheme is the simplicity of its implementation as the various transmitters do not have to consider the channel gains of the other users and formulate a joint strategy. On the other hand this may lead to increasing interference if many users pick the same subchannel to transmit. For example let us consider the two-user two-subchannel case for simplicity. Hence $n_f = N = 2$ and also $n_t = 1$, since no spreading is considered. Let the channel gains of the i^{th} user on the j^{th} subchannel be $\sqrt{h_{ij}}, 1 \leq i, j \leq 2$. The maximum mutual information, for a given set of channel gains depend upon whether both users transmit in the same subcarrier or not. If they transmit in different subcarriers, the MMI is

$$R_d(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \log \left(1 + \frac{h_2^{*(1)} P}{\sigma^2} \right) + \frac{1}{2} \log \left(1 + \frac{h_2^{*(2)} P}{\sigma^2} \right), \quad (3.54)$$

where $h_2^{*(1)} = \max[h_{11}, h_{12}]$ and $h_2^{*(2)} = \max[h_{21}, h_{22}]$ and the subscript 2 shows that the maximization is over two subcarriers. This is done to make the notation consistent with those in Section 3.2. If both users decide to transmit on the same subchannel, then the MMI is

$$R_s(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \log \left(1 + \frac{h_2^{*(1)} P}{\sigma^2} + \frac{h_2^{*(2)} P}{\sigma^2} \right). \quad (3.55)$$

It can be seen that $R_d(\mathbf{h}_1, \mathbf{h}_2) > R_s(\mathbf{h}_1, \mathbf{h}_2)$ and this is due to interference. Note that this scheme can be easily extended to incorporate more users.

Treating Other User Interference as Noise

In this case, each user treats the signal of all other user's as noise and waterfills over the background noise plus interference spectrum, in an iterative way. As discussed in Section 3.5.3 this method is the optimal policy if each user's transmit covariance matrix was full rank, but no comment about its optimality can be made for the present case, when the transmit covariance matrix is rank one. Still this policy takes the other users' interference into account and would intuitively do better than the single user transmission as outlined in Section 3.5.4. This is shown in the simulations of Section 3.5.5. From a practical viewpoint, the extra complexity in the implementation of this scheme is not substantial. The receiver which has all the received signals, can broadcast the received spectrum, using a feedback channel and each user can subtract his own signal from this spectrum to obtain the interference to him, from other users. Such ideas have been proposed in [27] in the context of interference avoidance.

The optimal power vectors can be found out via the strategy outlined in Appendix C. However in the two user case the problem is simplified as discussed in Section 3.3 as the existence of a positive power vector is always guaranteed. The optimization problem for the first user can be stated from Equations (3.51) and Equation (3.30) as

$$\max_{\mathbf{p}(\mathbf{h}_1)} \log \left(\det \left[S_{zz1} + \mathbf{H}_1^{\frac{1}{2}} \mathbf{p}_1^{\frac{1}{2}}(\mathbf{h}_1) \mathbf{p}_1^{\frac{1}{2}}(\mathbf{h}_1)^T (\mathbf{H}_1^{\frac{1}{2}})^T \right] \right), \quad (3.56)$$

$$\text{where } S_{zz1} = I_2 + \mathbf{H}_2^{\frac{1}{2}} \mathbf{p}_2^{\frac{1}{2}}(\mathbf{h}_2) \mathbf{p}_2^{\frac{1}{2}}(\mathbf{h}_2)^T (\mathbf{H}_2^{\frac{1}{2}})^T. \quad (3.57)$$

The optimization for the second user can be similarly expressed. This policy is implemented in an iterative fashion. It can be outlined as

1. Initialize $\mathbf{p}_2(\mathbf{h}_2) = \mathbf{p}_0$,

2. Repeat

for $i = 1$ to 2, $j \neq i$,

$$S_{zzi} = I + \mathbf{H}_j^{\frac{1}{2}} \mathbf{p}_j^{\frac{1}{2}}(\mathbf{h}_j) \mathbf{p}_j^{\frac{1}{2}}(\mathbf{h}_j)^T (\mathbf{H}_j^{\frac{1}{2}})^T,$$

$$\mathbf{p}_i(\mathbf{h}_i) = \max_{\mathbf{p}(\mathbf{h})} \log \left(\det \left[S_{zzi} + \mathbf{H}_i^{\frac{1}{2}} \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i) \mathbf{p}_i^{\frac{1}{2}}(\mathbf{h}_i)^T (\mathbf{H}_i^{\frac{1}{2}})^T \right] \right),$$

end

until the MMI converges.

Section 3.5.5 shows that this procedure done iteratively settles down to a fixed point. Note that this scheme too can be easily extended to incorporate more users.

Single Dimension Transmission

In this section we refine the optimal single user policy as defined in Section 3.5.4 to take care of the issues of interference which reduce the MMI. The key idea is that users should transmit in only one subchannel. This is motivated by single user optimization results. The policy is to allocate a specific subchannel to a user.

To understand this better let us again consider the two-user two-subchannel case. For each given realization of channel state vectors $\mathbf{h}_1 = [h_{11}, h_{12}]$ and $\mathbf{h}_2 = [h_{21}, h_{22}]$, there are four possible ways in which these two users can transmit the information, as each of them has a choice of picking up any of the two subcarriers for transmission. There are many possibilities depending upon the relative values of $[h_{11}, h_{12}, h_{21}, h_{22}]$. If both users pick different subcarriers interference is avoided, however one of the users may experience a bad channel. Also a user might have a high gain in one subchannel and a very low gain in the other and so it makes sense to let that user transmit in the former subchannel, even if the other user had a higher value of channel gain in the same subchannel. Our contention is that it is not possible to predict which method is best and hence a brute force method which calculated the value of MMI for each of these four possible transmission options have to be calculated and the option yielding the highest value is to be selected. The achievable MMI values for these four schemes are given in Equations (3.58) to (3.61). The notation $R_{ij}(\mathbf{h}_1, \mathbf{h}_2)$, denotes the MMI value obtained when user 1 transmits in subchannel i and user 2 transmits in subchannel j .

These are given by,

$$R_{11}(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \log \left(1 + \frac{h_{11}P}{\sigma^2} + \frac{h_{21}P}{\sigma^2} \right), \quad (3.58)$$

$$R_{12}(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \log \left(1 + \frac{h_{11}P}{\sigma^2} \right) + \frac{1}{2} \log \left(1 + \frac{h_{22}P}{\sigma^2} \right), \quad (3.59)$$

$$R_{21}(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \log \left(1 + \frac{h_{12}P}{\sigma^2} \right) + \frac{1}{2} \log \left(1 + \frac{h_{21}P}{\sigma^2} \right), \quad (3.60)$$

$$R_{22}(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \log \left(1 + \frac{h_{12}P}{\sigma^2} + \frac{h_{22}P}{\sigma^2} \right). \quad (3.61)$$

Thus the policy checks for the maximum value $R_{i^*j^*}(\mathbf{h}_1, \mathbf{h}_2)$ and allocates first user to subchannel i^* and second user to subchannel j^* . In principle this is complex to implement and the complexity increases exponentially with the number of users and number of subchannels. Also the computation of $R_{i^*j^*}(\mathbf{h}_1, \mathbf{h}_2)$ needs global knowledge of the channel state vectors of all users and can be best performed at the receiver, who then has to inform the transmitters by a feedback mechanism.

3.5.5 Simulation Results for Multi-Users

In this section we provide a simulation results to compare the performances of the schemes proposed in Sections 3.5.4 to 3.5.4. We assume a system with two users and two subchannels with no time domain spreading. The channel is assumed to be Rayleigh fading with average value of unity. Figure 3.7 plots how the MMI values varies with time for each of the three policies. For the iterative policy, five iterations were performed and the initial choice of power vector was At each time an independent realization of all the four channel states is generated. Figure 3.7 shows that the optimal single user policy performs worse. This isn't surprising because this policy didn't take into the effect of the other users. The figure shows that the optimal single dimension policy performs best but interestingly the iterative policy of achieves identical rates as the optimal single dimension policy in most cases. It shows that the iterative policy for the two user case, eventually leads to each user transmitting in a different subchannel as in the optimal single dimension policy.

We need to specify an initial power vector \mathbf{p}_0 for the iterative policy. But it was seen in simulations that the performance of the iterative policy is invariant to the initial

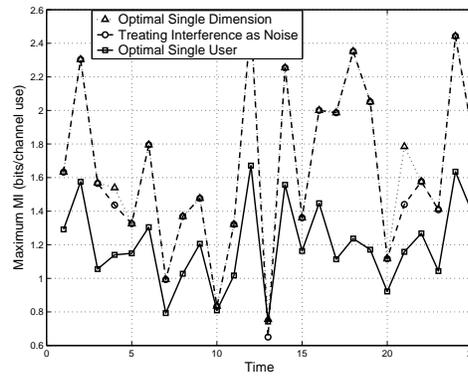


Figure 3.7: MMI variations with time for different policies

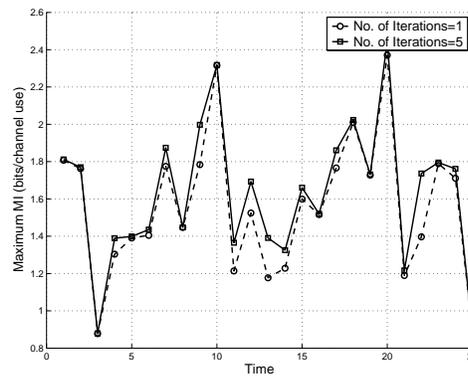


Figure 3.8: MMI variations with time different iterations of the

choice of \mathbf{p}_0 . Similar results were reported in [26] for their iterative algorithm. However in [26] the iterative policy maximized a concave objective over set of positive semi definite matrices \mathbf{Q} with convex constraints for which the invariance of the initial choice of \mathbf{Q} is well established. In this case there are no nice convexity properties, however the same result hold. Also as the number of iterations increase, the performance of the iterative policy becomes better. This is shown by plotting the MMI values for this algorithm for two different number of iterations, namely one and five. The results are shown in Figure 3.8. The MMI for one iteration is marginally less than for MMI five. Note that in case the MMI values do not exhibit a large variation between one and five iterations. This is because we are considering only two users and the iterative policy quickly converges.

Chapter 4

Achievable Rates for Imperfect CSI at Transmitter and Perfect CSI at Receiver

4.1 Introduction

In Chapter 3 it was assumed that the transmitter had perfect knowledge of the instantaneous channel coefficients. While this leads to elegant water-filling solutions for power allocation, perfect CSI might be difficult to achieve at the transmitter. In most practical systems the transmitter sends a pre-determined training sequence and the receiver estimates the channel coefficients. By transmitting symbols during the coherence time of the channel, the channel may be estimated to any degree of accuracy. The receiver then transmits the channel gain estimates to the transmitter, using a feedback channel. Depending upon the quality of the feedback channel, the estimates may be received in error at the transmitter. Also due to delay in the feedback loop and time variability of the channel, the estimates at the transmitter at any instant may not correspond to the actual channel gains, at that instant. In this section, we thus assume that perfect CSI is not available at the transmitter and investigate the effects upon ergodic channel capacity, due to this cause. Quantitatively, we assume that the true channel state is given by the random vector \mathbf{h} and the information about the channel state at the transmitter is given by the random vector \mathbf{u} .

4.2 Single user transmission

The optimal solution depends on the nature of the relationship between \mathbf{u} and \mathbf{h} as characterized by $f(\mathbf{u}, \mathbf{h})$ or more specifically by the conditional density $f(\mathbf{h}|\mathbf{u})$ [15]. This is illustrated as follows. From Equation (3.5) the optimization problem can be

written as

$$\max_{p(\mathbf{u})} \int \cdots \int R(\mathbf{h}, \mathbf{u}) f(\mathbf{h}|\mathbf{u}) d\mathbf{h}, \quad (4.1)$$

$$\text{s.t. } \sum_{j=1}^{n_f} \int \cdots \int p_j(\mathbf{u}) f(\mathbf{u}) d\mathbf{u} = \bar{P}, \quad p_j(\mathbf{u}) \geq 0, \quad (4.2)$$

where $R(\mathbf{h}, \mathbf{u})$ is the MMI, (defined first in Section 3.2) whose expression is given by,

$$R(\mathbf{h}, \mathbf{u}) = \frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j=1}^{n_f} h_j p_j(\mathbf{u}) \right). \quad (4.3)$$

The Lagrangian for the optimization problem can be stated in the most general form as

$$J = \int \cdots \int R(\mathbf{h}, \mathbf{u}) f(\mathbf{h}|\mathbf{u}) f(\mathbf{u}) d\mathbf{h} d\mathbf{u} - \lambda \sum_{j=1}^{n_f} \int \cdots \int p_j(\mathbf{u}) f(\mathbf{u}) d\mathbf{u}. \quad (4.4)$$

Differentiating with respect to $p_j(\mathbf{u})$ we obtain

$$\frac{\partial J}{\partial p_j(\mathbf{u})} = s_j(\mathbf{u}) - \lambda \leq 0, \quad \text{for all } j, \quad (4.5)$$

$$\text{where } s_j(\mathbf{u}) = \frac{1}{2} \int \cdots \int \left(\frac{h_j}{\sigma^2 + \sum_{k=1}^{n_f} h_k p_k(\mathbf{u})} \right) f(\mathbf{h}|\mathbf{u}) d\mathbf{h}, \quad (4.6)$$

for all j , with equality holding in Equation (4.5) when $p_j(\mathbf{u}) \neq 0$. For any revealed CSI \mathbf{u} solving integral equation (4.5) yields an expression for $p_j(\mathbf{u})$ for all j . The values of $p_j(\mathbf{u})$ thus obtained, can be substituted in Equation (4.2) to solve for λ .

The maximum value of $s_j(\mathbf{u})$ occur when $p_j(\mathbf{u}) = 0$ for all j . Let's denote this maximum value by $s_j^*(\mathbf{u})$. Thus

$$s_j^*(\mathbf{u}) = \frac{1}{2\sigma^2} \int \cdots \int h_j f(\mathbf{h}|\mathbf{u}) d\mathbf{h}. \quad (4.7)$$

To understand Equation (4.7) note that λ is related to the water level. In fact, for perfect CSI, $1/\lambda$ is the water level as shown in Equation (3.9). For a given \bar{P} and λ suppose CSI revealed is $\mathbf{u} = \mathbf{u}_0$. Now given subcarrier j with associated channel state value tuple (h_j, u_{0j}) , we want to determine if transmission should take place in that subcarrier or not. From Equation (4.7), we evaluate the value of $s_j^*(\mathbf{u})$, and compare it to λ . Only if $s_j^*(\mathbf{u}) > \lambda$, there is transmission in the j^{th} subchannel. Alternatively for

a given λ and revealed channel state $\mathbf{u} = \mathbf{u}_0$, users stop transmission below a threshold power \bar{P}_{Th} . The value of \bar{P}_{Th} depends upon the conditional fading distribution $f(\mathbf{h}|\mathbf{u})$.

Let us consider different types of CSI at transmitter and try to evaluate the optimal policy.

4.2.1 Perfect CSIR, No CSIT

In this case \mathbf{u} and \mathbf{h} are independent random variables. Hence $f(\mathbf{u}, \mathbf{h}) = f(\mathbf{u})f(\mathbf{h})$. It can be shown that [29] the capacity maximizing policy is constant power allocation across all subchannels i.e.

$$p_j(\mathbf{u}) = \frac{\bar{P}}{n_f}. \quad (4.8)$$

Thus the capacity expression in (3.5) can be upper bounded by:

$$C_{\text{NCSI}} = E_{\mathbf{h}} \left[\frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j=1}^{n_f} h_j \left(\frac{\bar{P}}{n_f} \right) \right) \right] \quad (4.9)$$

$$\leq \frac{1}{2} \log \left(1 + \frac{1}{n_f \sigma^2} \sum_{j=1}^{n_f} E_{\mathbf{h}} [h_j] \bar{P} \right) \quad (4.10)$$

$$= \frac{1}{2} \log \left(1 + \frac{\bar{h} \bar{P}}{\sigma^2} \right). \quad (4.11)$$

This is the AWGN channel capacity if we consider the normalized Rayleigh channel i.e $\bar{h} = 1$. This capacity is achievable by assigning uniform power to each subcarrier. This can be seen by examining the capacity expression for $p_j(\mathbf{u}) = \bar{P}/n_f$, which is,

$$C_{\text{NCSI}} = E_{\mathbf{h}} \left[\frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \frac{1}{n_f} \sum_{j=1}^{n_f} h_j \bar{P} \right) \right] \quad (4.12)$$

$$\longrightarrow \frac{1}{2} \log \left(1 + \frac{\bar{h} \bar{P}}{\sigma^2} \right), \quad (4.13)$$

with equality holding for large n_f . The law of large numbers ensured that the capacity is achievable for a large number of subcarriers. This can be looked upon as the *diversity advantage*. Note that for a single subcarrier the corresponding expression with uniform power allocation is $\log(1 + \bar{h} \bar{P} / \sigma^2) / 2$, and there is no way to actually achieve this bound.

4.2.2 Perfect CSIR One Bit Quantized CSIT

Suppose u_k be an one bit quantized information about the channel state at subcarrier k [15]. The feedback simply indicates if the value of the channel gain h_k is above or below a certain threshold h_{Th} i.e.

$$u_k = \begin{cases} 0 & h_k < h_{Th}, \\ 1 & h_k \geq h_{Th}. \end{cases} \quad (4.14)$$

The optimal transmission strategies can be derived by solving Equation (4.4). However we can list the various possible types of feedback that can arise with such a CSIT, and in all such cases the solution can be deduced by simpler methods, which rely upon the symmetry of the problem. The types of possible feedbacks can be grouped into three categories, which are listed below,

All n_f subchannels receive $u = 1$

From such a feedback the transmitter can't differentiate between the channel qualities in the different subcarriers. The optimal solution is to transmit along all the subcarriers at the same power. We shall denote this power as p_{n_f} (the subscript denotes the n_f subcarriers receive unity feedback) The value of p_{n_f} is determined from equation (4.5) by solving $\partial J / \partial p_j(\mathbf{u}) = 0$, for any $j \in \{1, \dots, n_f\}$. Without loss of generality we choose $j = 1$

$$\frac{1}{2} \int \dots \int \left(\frac{h_1}{\sigma^2 + p_{n_f} \sum_{k=1}^{n_f} h_k} \right) f(\mathbf{h} | \mathbf{u} = 1) d\mathbf{h} = \lambda. \quad (4.15)$$

As discussed in Section 4.2, p_{n_f} is non-zero if

$$\frac{1}{2\sigma^2} \int \dots \int h_1 f(\mathbf{h} | \mathbf{u} = 1) d\mathbf{h} = \frac{1}{2\sigma^2} \int h_1 f(h_1 | u_1 = 1) dh_1 \geq \lambda. \quad (4.16)$$

Following the notation of Section 4.2 we denote the last integral by,

$$s_j^*(\mathbf{u} = 1) = \frac{1}{2\sigma^2} \int h_1 f(h_1 | u_1 = 1) dh_1. \quad (4.17)$$

We recall at this point that if all subcarriers receive $u = 1$, they transmit information only if $s_j^*(\mathbf{u} = 1) > \lambda$. Note that λ depends on the fading distribution and h_{Th} .

m subcarriers report that $u = 1$ where $1 \leq m < n_f$

The transmitter, based on the feedback classifies the subcarriers into two groups, one each for the two types of feedback. From symmetry arguments it can't distinguish between the channel states of subcarriers within the same group and hence has to allocate equal power to all subcarriers within the same group. It can be shown that only the m subcarriers which report $u = 1$ (they correspond to the subcarriers with the best conditional value of channel gains) should transmit and the rest $n_f - m$ subcarriers should stop transmission. For a proof see Appendix A. Thus the power p_m (the subscript denotes the m subcarriers receive unity feedback) is determined from equation (4.5). We assume without loss of generality that user 1, receives feedback, $u_1 = 1$ and solve $\partial J/\partial p_j(\mathbf{u}) = 0$ for $j = 1$.

$$\frac{1}{2} \int \cdots \int \left(\frac{h_1}{\sigma^2 + p_m \sum_{k=1}^m h_k} \right) f(\mathbf{h}|\mathbf{u} = 1) d\mathbf{h} = \lambda. \quad (4.18)$$

In Equation (4.18) only m integrals need be evaluated as the rest integrate to one. As discussed in Section 4.2 p_m is non-zero if

$$\frac{1}{2\sigma^2} \int \cdots \int h_1 f(\mathbf{h}|\mathbf{u} = 1) d\mathbf{h} = \frac{1}{2\sigma^2} \int h_1 f(h_1|u_1 = 1) dh_1 \geq \lambda, \quad (4.19)$$

where we recognize the last integral being $s_j^*(\mathbf{u} = 1)$ as defined in Equation (4.17). This means that if m subcarriers receive $u = 1$, they transmit information only if $s_j^*(\mathbf{u} = 1) > \lambda$. Thus the upper bound on feasible λ which allows transmission is $s_j^*(\mathbf{u} = 1)$ for all $m \neq 0$.

All n_f subcarriers report $u = 0$

This case (henceforth also referred to as the *all zero case*) is similar to case when all subcarriers reported $u = 1$, in the sense that the transmitter can't differentiate between the subcarriers, even though it is clear that all channel states are worse than the $u = 1$ case. Solution is similar to first situation. All users transmit at the same power p_0 which again is determined from equation (4.5) by solving for $\partial J/\partial p_j(\mathbf{u}) = 0$ for $j = 1$

$$\frac{1}{2} \int \cdots \int \left(\frac{h_j}{\sigma^2 + p_0 \sum_{k=1}^{n_f} h_k} \right) f(\mathbf{h}|\mathbf{u} = 0) d\mathbf{h} = \lambda. \quad (4.20)$$

As discussed in Section 4.2 p_0 is non-zero if

$$\frac{1}{2\sigma^2} \int \cdots \int h_j f(\mathbf{h}|\mathbf{u} = 0) d\mathbf{h} = \frac{1}{2\sigma^2} \int h_j f(h_j|u_j = 0) dh_j \geq \lambda. \quad (4.21)$$

Following the notation of Section 4.2 we denote the last integral by,

$$s_j^*(\mathbf{u} = 0) = \frac{1}{2\sigma^2} \int h_1 f(h_1|u_1 = 0) dh_1. \quad (4.22)$$

We recall at this point that if all subcarriers receive $u = 0$, they transmit information only if $s_j^*(\mathbf{u} = 0) > \lambda$.

Let us investigate Equation (4.5) and see how the computation of the m integrals $1 \leq m \leq n_f$ might be simplified. Let us first consider the case when at least one subcarrier reports unity feedback and lets assume without loss of generality that the first subcarrier is one of them. Hence if m subcarriers report unity feedback, (where $m \neq 0$), Equation (4.5) can be written as

$$\lambda = \frac{1}{2} \int \cdots \int \left(\frac{h_1}{\sigma^2 + p_m \sum_{k=1}^m h_k} \right) f(\mathbf{h}|\mathbf{u} = 1) d\mathbf{h}, \quad (4.23)$$

$$= \frac{1}{2} \int \cdots \int \left(\frac{h_1}{\sigma^2 + p_m h_1 + p_m \sum_{k=2}^m h_k} \right) f(\mathbf{h}|\mathbf{u} = 1) d\mathbf{h}, \quad (4.24)$$

$$= \frac{1}{2} \int \cdots \int \left(\frac{h_1}{\sigma^2 + p_m h_1 + p_m h_S} \right) f(h_1|\mathbf{u} = 1) f(h_S|\mathbf{u} = 1) dh_1 dh_S, \quad (4.25)$$

where the terms h_S and $f(h_S|\mathbf{u} = 1)$ denote the random variable $\sum_{k=2}^m h_k$ and its PDF respectively.

Finally we can make a simple observation from Equations (4.16), (4.19) and (4.21) regarding the thresholds values $s_j^*(\mathbf{u} = 1)$ and $s_j^*(\mathbf{u} = 0)$ which determine if the powers $p_m, m \neq 0$ and p_0 are positive. We observe that,

$$s_j^*(\mathbf{u} = 1) = \int_{h_{Th}}^{\infty} h f(h|u = 1) dh > h_{Th} \int_{h_{Th}}^{\infty} f(h|u = 1) dh = h_{Th}, \quad (4.26)$$

$$s_j^*(\mathbf{u} = 0) = \int_0^{h_{Th}} h f(h|u = 0) dh < h_{Th} \int_0^{h_{Th}} f(h|u = 0) dh = h_{Th}. \quad (4.27)$$

This shows that $s_j^*(\mathbf{u} = 1) > s_j^*(\mathbf{u} = 0)$. This can be explained as follows: Recall that transmitted powers $p_j(\mathbf{u})$ are positive only if $s_j^*(\mathbf{u}) > \lambda$. Consider a value of $\lambda = \lambda_0$ higher than $s_j^*(\mathbf{u} = 0)$. A high value of λ_0 implies that the water level is low and hence the total power \mathbf{P} is low. So the transmitter should be more prudent in allocating power

and hence there is no transmission in the *all zero case* as this corresponds to the worst channel state among all types of feedbacks. However if $\lambda_0 < s_j^*(\mathbf{u} = 1)$, then there is transmission over the subcarriers that receive $u = 1$ feedback.

Let us consider specific fading distributions and evaluate the various expressions developed in this Section.

Uniform Fading Distribution

A simple closed form expression for $f(h_S|\mathbf{u})$ does not exist. Hence let us consider the two subcarrier case for simplified analysis. Fading is uniform in the range $[0, A]$. Let $\{u_1, u_2\}$ be the feedbacks in the two subcarriers, where $u_1, u_2 \in \{0, 1\}$. Let the threshold h_{Th} of Equation (4.14) be a . Let the power transmitted in the subcarriers 1 and 2, under feedback $= \{u_1, u_2\}$ be $p_1(u_1, u_2)$ and $p_2(u_1, u_2)$ respectively. The following conditions hold

$$p_1(0, 0) = p_2(0, 0) = p_0, \quad \Pr[p_0] = (a/A)^2, \quad (4.28)$$

$$p_1(1, 0) = p_2(0, 1) = p_1, \quad \Pr[p_1] = a(A - a)/A^2, \quad (4.29)$$

$$p_1(1, 1) = p_2(1, 1) = p_2, \quad \Pr[p_2] = (A - a/A)^2, \quad (4.30)$$

$$p_1(0, 1) = p_2(1, 0) = 0. \quad (4.31)$$

The total power equation, Equation (4.2) can be expanded as,

$$2p_0 \left(\frac{a}{A}\right)^2 + 2p_1 \frac{a(A - 1)}{A^2} + 2p_2 \left(\frac{A - a}{A}\right)^2 = \bar{P}. \quad (4.32)$$

The values of p_0, p_1 and p_2 can be found by evaluating Equations (4.15), (4.18) and (4.20) and the results are

$$\frac{1}{2p_0^2 a^2} \left[\frac{a^2}{2} \log(\sigma^2 + 2ap_0) - \frac{a^2}{2} \log(\sigma^2 + ap_0) \right] = \lambda, \quad (4.33)$$

$$\frac{1}{2p_1(A - 1)} \left[(A - a) - \frac{\sigma^2}{p_1} \{ \log(\sigma^2 + Ap_1) - \log(\sigma^2 + ap_1) \} \right] = \lambda, \quad (4.34)$$

$$\frac{1}{2p_2^2(A - a)^2} \left[\frac{A^2}{2} \log(\sigma^2 + 2Ap_2) - \frac{a^2}{2} \log(\sigma^2 + ap_2 + Ap_2) \right. \quad (4.35)$$

$$\left. - \frac{A^2}{2} \log(\sigma^2 + ap_2 + Ap_2) + \frac{a^2}{2} \log(\sigma^2 + 2ap_2) \right] = \lambda. \quad (4.36)$$

We can also evaluate Equations (4.16), (4.19) and (4.21) to obtain the thresholds for λ for which the transmitted powers are positive. The results are

- p_1 and p_2 are positive for $\lambda \leq s^*(\mathbf{u} = 1) = (A + a)/4 = 0.375$.
- p_0 is positive for $\lambda \leq s^*(\mathbf{u} = 0) = a/4 = 0.125$.

Note that $s^*(\mathbf{u} = 1) > s^*(\mathbf{u} = 0)$ as predicted earlier in this Section. Simulation results for this case will be provided in Section 4.3.

Rayleigh Fading

Unlike the case of uniform fading, there exists a closed expression for $f(h_S|\mathbf{u})$, for most types of feedback. Assume the mean of the Rayleigh fading to be unity, ie $\gamma_a = 1$. Now h_S is dependent on the number of terms in the summation, so let us denote h_S by S_k which means that it is the sum of k i.i.d. exponential random variables, where $1 \leq k \leq n_f$. The following expression for the PDF of $S_k|\mathbf{u} = 1$ holds

$$f_{S_k|\mathbf{u}=1}(s) = \frac{(s - ka)^{n-1} e^{-(s-ka)}}{(n-1)!}. \quad (4.37)$$

However for the all zero case, $f_{S_{n_f}|\mathbf{u}=0}(s)$ does not have a closed form expression.

4.3 Simulation Results

In this section we take the uniform and rayleigh fading scenarios and numerically evaluate the achievable rates. In this section we plot the quantity maximum mutual information (MMI) as defined in Section 3.2. At each time instant an independent realization of the channel is generated, and the transmitter goes through the steps mentioned in Section 3.2. At this juncture it is appropriate to recall two key definitions which would help us to interpret the numerical results correctly.

- \bar{P} as defined in Equation (4.2) is the total average power into the system, i.e. the sum of the average powers in all the subcarriers, where the averaging operation is done over different realizations of channel states information variable \mathbf{u} , available at the transmitter.

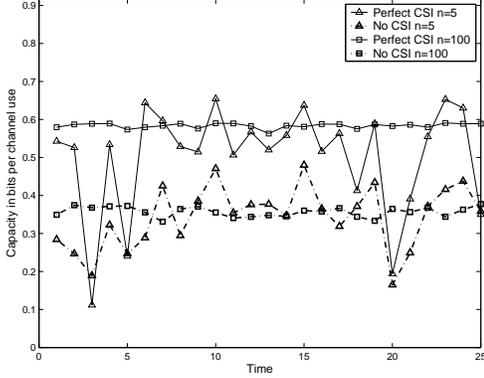


Figure 4.1: MMI for no CSI

Parameter	Value
A	1
a	0.5
SNR	1 dB
# Subcarriers n_f	5,100

Figure 4.2: Simulation Parameters for Fig 4.1

- λ is inversely related to the water-filling level. Hence λ and \bar{P} are also inversely related to each other.

4.3.1 Uniform Fading

Figure (4.1) plots the MMI as a function of time for two different values of n_f , namely $n_f = 5$ and $n_f = 100$. The MMI sequence obtained with perfect CSI is also plotted to give a basis for comparison. Since more number of subcarriers leads to better averaging effect (refer Equation (4.9)), the MMI sequence for $n_f = 100$ has less fluctuations. In fact, its value is close to 0.3522, which is the theoretical value obtained from substituting parameters from Table 4.2 in Equation (4.9).

Figure 4.3 plots the powers p_0, p_1 and p_2 as derived in Equation (4.33) for two subcarriers. The simulation parameters are listed in Table 4.4. It is seen that for the same value of λ , which corresponds to the same total average power \bar{P} , the ordering in the magnitudes of the three power values are: $p_1 \geq p_2 > p_0$. Let us recall the definitions of these power levels to understand the results. p_0 is the power allocated to each subcarrier in the all zero state and hence has the least value. p_1 corresponds to power in subcarrier 1 in case of (1,0) feedback or power in subcarrier 2 in case of (0,1) feedback. In both cases, power is transmitted in only one channel. p_2 corresponds to individual power in both channels for the (1,1) feedback. Intuitively p_1 should be greater than or equal to p_2 , since there are 2 subcarriers transmitting power p_2 and 1 subcarrier transmitting power p_1 . This is verified by the simulations and Figure 4.3

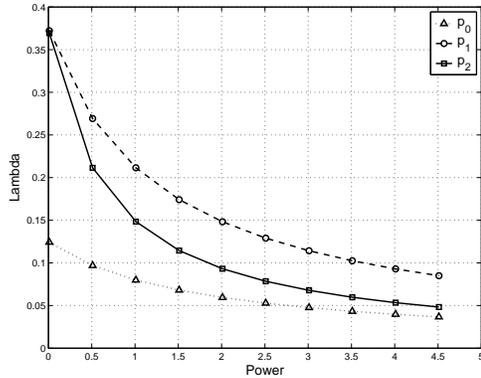


Figure 4.3: Plot of λ vs Power allocated for different feedbacks

Parameter	Value
A	1
a	0.5
Noise	0 dB
# Subcarriers n_f	2

Figure 4.4: Simulation Parameters for Fig 4.3

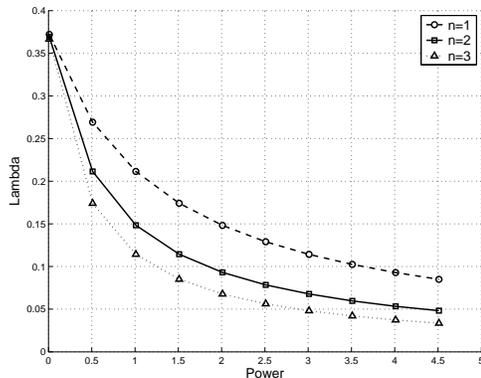


Figure 4.5: λ vs Power allocated for all one feedback for different number of sub-carriers

Parameter	Value
A	1
a	0.5
Noise	0 dB
# Subcarriers n_f	1,2,3

Figure 4.6: Simulation Parameters for Fig 4.5

shows that for a range of λ values $p_1 \sim 2p_2$. These simulations match closely with the predicted cut-off values of λ as discussed under the uniform fading of Section 4.2.2, namely $\lambda = 0.375$ for p_1 and p_2 and $\lambda = 0.125$ for p_0 .

Figure 4.5 refers to the *all one feedback* case when the number of subcarriers are $n_f = 1, 2$ and 3. It is seen that, for a fixed λ , which corresponds to a fixed \bar{P} , the high the value of n_f corresponds to low value of power, p_{n_f} . The intuitive explanation is that the same power \bar{P} is getting divided into more subcarriers.

Finally Figure 4.7 plots the MMI variations with time for all the three types of CSI discussed till now. We observe that for $n_f = 2$, the achievable rates with 1 bit per subcarrier feedback is close to the perfect CSI case. This suggests that the 1 bit per subcarrier scheme is a good practical scheme for transmitter feedback.

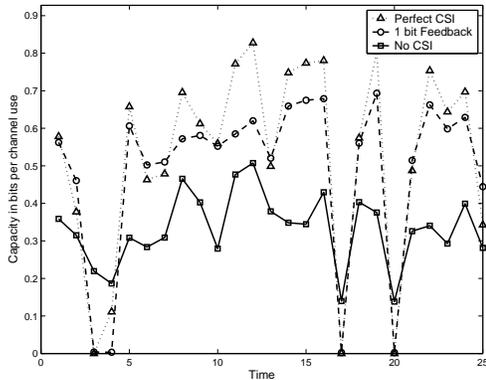
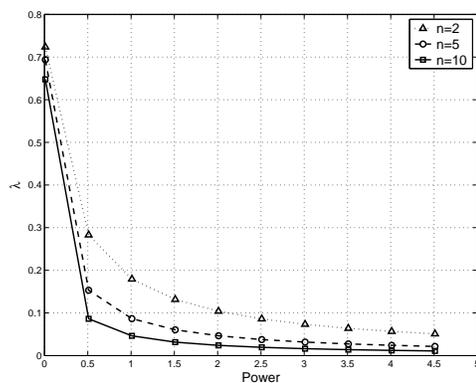


Figure 4.7: MMI for different types of CSI

Parameter	Value
A	1
a	0.5
SNR	1 dB
# Subcarriers n_f	2

Figure 4.8: Simulation Parameters for Fig 4.7

Figure 4.9: λ vs Power allocated for all one feedback for different number of subcarriers

Parameter	Value
γ_a	1
a	0.5
Noise	0 dB
# Subcarriers n_f	2,5,10

Figure 4.10: Simulation Parameters for Fig 4.9

4.3.2 Rayleigh Fading

In this section the simulation results for Rayleigh Fading are presented. Figure (4.9) pertains to the power allocation when all subcarriers receive unity feedback. It plots the variation of λ with power p_n , the power allocated in each subcarrier as the number of subcarriers n is increased. Note that similar results for uniform fading were shown in Figure (4.5). For a fixed value of λ , which corresponds to a fixed \bar{P} , increasing the number of subcarriers, reduced the power allocated in each subcarrier.

Chapter 5

Conclusions and Future Work

This report provides a comprehensive information theoretic analysis a two dimensional spreading scheme which was motivated by the VSF-OFCDM system. It develops the transmission model for single user, and studies optimal power allocation and transmission schemes, under three different cases of channel state knowledge at the transmitter. For perfect CSI at transmitter the optimal single user policy is to transmit in the best subcarrier and the transmit power is obtained by waterfilling over the distribution of the channel state in the best subcarrier. For no CSI the optimal policy is equal power allocation in all subcarriers. For 1 bit per subcarrier feedback the optimal policy turned out to be to transmit in all the subcarriers that are above the threshold. The work also looks at the multi-user scenario and proposes some practical transmission schemes and investigates the achievable sum rates of these schemes.

There are several future directions of research in this area. The multiuser scenario, especially for the imperfect transmitter CSI case, needs to be addressed. Also outage behavior and delay aspects of the proposed model have to be carefully investigated. On a more practical note, since the transmission spans both time and frequency dimensions, scheduling between various transmissions has to be studied, in order to implement the best subchannel or best user policies which were shown to be theoretically optimum.

Appendix A

Transmission Policy for 1 bit Quantized CSIT

Consider the case when $n_f = 2$. The proof can be easily generalized for higher n_f . Now

$f(\mathbf{h}|\mathbf{u}) = \prod_{i=1}^2 f(h_i|u_i)$ from assumptions about the channel and the CSIT. Thus

$$\frac{\partial J}{\partial p_1(\mathbf{u})} = \frac{1}{2} \int \int \frac{h_1}{\sigma^2 + h_1 p_1(\mathbf{u}) + h_2 p_2(\mathbf{u})} f(h_1|u_1) f(h_2|u_2) dh_1 dh_2 - \lambda \quad (\text{A.1})$$

$$\frac{\partial J}{\partial p_2(\mathbf{u})} = \frac{1}{2} \int \int \frac{h_2}{\sigma^2 + h_1 p_1(\mathbf{u}) + h_2 p_2(\mathbf{u})} f(h_1|u_1) f(h_2|u_2) dh_1 dh_2 - \lambda, \quad (\text{A.2})$$

Equivalently,

$$\frac{\partial J}{\partial p_1(\mathbf{u})} = \frac{1}{2} \int \int h_1 g(h_1, h_2, u_1, u_2) dh_1 dh_2 - \lambda \quad (\text{A.3})$$

$$\frac{\partial J}{\partial p_2(\mathbf{u})} = \frac{1}{2} \int \int h_2 g(h_1, h_2, u_1, u_2) dh_1 dh_2 - \lambda. \quad (\text{A.4})$$

Note that the function $g(h_1, h_2, u_1, u_2)$ is same for the integrals. Now consider $u_1 = 1, u_2 = 0$. Then

$$\frac{\partial J}{\partial p_1(\mathbf{u})} = \frac{1}{2} \int_{h_{\text{Th}}}^{\infty} \int_0^{h_{\text{Th}}} h_1 g(h_1, h_2, u_1, u_2) dh_2 dh_1 - \lambda \quad (\text{A.5})$$

$$> \frac{h_{\text{Th}}}{2} \int_{h_{\text{Th}}}^{\infty} \int_0^{h_{\text{Th}}} g(h_1, h_2, u_1, u_2) dh_2 dh_1 - \lambda. \quad (\text{A.6})$$

Similarly, it can be shown that

$$\frac{\partial J}{\partial p_2(\mathbf{u})} < \frac{h_{\text{Th}}}{2} \int_{h_{\text{Th}}}^{\infty} \int_0^{h_{\text{Th}}} g(h_1, h_2, u_1, u_2) dh_2 dh_1 - \lambda. \quad (\text{A.7})$$

Hence,

$$\partial J / \partial p_1(\mathbf{u}) > \partial J / \partial p_2(\mathbf{u}). \quad (\text{A.8})$$

But from the following basic conditions

$$\frac{\partial J}{\partial p_k(\mathbf{u})} = 0 \text{ when } p_k(\mathbf{u}) > 0 \quad (\text{A.9})$$

$$\frac{\partial J}{\partial p_k(\mathbf{u})} \leq 0 \text{ when } p_k(\mathbf{u}) = 0, \quad (\text{A.10})$$

It can be stated that $p_1(\mathbf{u}) > 0$ and $p_2(\mathbf{u}) = 0$.

Appendix B

Properties of the Max function of n Exponentials

Let X_n^* be the random variable denoting the maximum of n iid random variables X_i , $i = 1, 2, \dots, n$, each with probability distribution function $f_X(x)$ and cumulative distribution function $F_X(x)$. Let the pdf and cdf of X_n^* be denoted as $f_{X_n^*}(x)$ and $F_{X_n^*}(x)$ respectively. From probability theory,

$$f_{X_n^*}(x) = n [F_X(x)]^{n-1} f_X(x), \quad (\text{B.1})$$

$$F_{X_n^*}(x) = [F_X(x)]^n. \quad (\text{B.2})$$

Let $X = h$, be an exponentially distributed random variable. The pdf of h is given by

$$f_h(\gamma) = \frac{1}{\gamma_a} e^{-(\gamma/\gamma_a)} \quad F_h(\gamma) = 1 - e^{-(\gamma/\gamma_a)} \quad (\text{B.3})$$

From Equations (B.1) and (B.3), the distribution of h_n^* , the maximum of n iid exponentials can be written down as

$$f_{h_n^*}(\gamma) = \frac{n}{\gamma_a} \left(1 - e^{-(\gamma/\gamma_a)}\right)^{n-1} e^{-(\gamma/\gamma_a)}, \quad (\text{B.4})$$

$$F_{h_n^*}(\gamma) = \left(1 - e^{-(\gamma/\gamma_a)}\right)^n. \quad (\text{B.5})$$

Lemma 1 $F_{h_m^*}(\gamma)$ is stochastically greater than $F_{h_n^*}(\gamma)$ for $m > n$

Proof:

$$F_{h_m^*}(\gamma) = \left(1 - e^{-(\gamma/\gamma_a)}\right)^m \quad (\text{B.6})$$

$$F_{h_n^*}(\gamma) = \left(1 - e^{-(\gamma/\gamma_a)}\right)^n \quad (\text{B.7})$$

$$F_{h_m^*}(\gamma) < F_{h_n^*}(\gamma) \text{ for all } \gamma \quad [\because \text{abscissa} < 1 \text{ and } m > n] \quad (\text{B.8})$$

This implies that $P[h_m^* > \gamma] > P[h_n^* > \gamma]$

Lemma 2 Let $E[h_n^*]$, $E[h_n^{*2}]$ and $Var[h_n^*]$ denote the mean, second moment and variance of h_n^* respectively. Then

$$\begin{aligned} E[h_n^*] &= \gamma_a \sum_{k=1}^n \frac{1}{k} \\ E[h_n^{*2}] &= 2\gamma_a^2 \sum_{k=1}^n \frac{1}{k} \sum_{l=1}^k \frac{1}{l} \\ Var[h_n^*] &= \gamma_a^2 \sum_{k=1}^n \frac{1}{k^2} \end{aligned}$$

Proof:

1. The expression for mean of h_n^* is

$$E[h_n^*] = \int_0^\infty n\gamma \left(1 - e^{-(\gamma/\gamma_a)}\right)^{n-1} \frac{1}{\gamma_a} e^{-(\gamma/\gamma_a)} d\gamma. \quad (\text{B.9})$$

Substituting for $x = 1 - e^{-(\gamma/\gamma_a)}$ the integral becomes

$$E[h_n^*] = n\gamma_a \int_0^1 -\log(1-x)x^{n-1} dx. \quad (\text{B.10})$$

Using Taylor Series Expansion $-\log(1-x) = x + x^2/2 + x^3/3 + \dots \infty$, and evaluating the series of integrals we obtain

$$E[h_n^*] = n\gamma_a \left[\frac{1}{n+1} + \frac{1}{2(n+2)} + \frac{1}{3(n+3)} + \dots \infty \right] \quad (\text{B.11})$$

$$= \gamma_a \left[\frac{n}{n+1} + \frac{n}{2(n+2)} + \frac{n}{3(n+3)} + \dots \infty \right]. \quad (\text{B.12})$$

Note that the k th term is $n/k(n+k) = 1/k - 1/(n+k)$. Regrouping the terms, we obtain,

$$E[h_n^*] = \gamma_a \left[\left(1 + \frac{1}{2} + \dots \infty\right) - \left(\frac{1}{n+1} + \frac{1}{n+2} + \dots \infty\right) \right] \quad (\text{B.13})$$

$$= \gamma_a \sum_{k=1}^n \frac{1}{k}, \quad \text{Hence Proved.} \quad (\text{B.14})$$

2. The expression for the second moment of h_n^* is

$$E[h_n^{*2}] = \int_0^\infty n\gamma^2 \left(1 - e^{-(\gamma/\gamma_a)}\right)^{n-1} \frac{1}{\gamma_a} e^{-(\gamma/\gamma_a)} d\gamma. \quad (\text{B.15})$$

Substitute $x = -\gamma/\gamma_a$ and use the binomial expansion

$$(1 - e^{-x})^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k e^{-kx}. \quad (\text{B.16})$$

This expression becomes

$$E \left[h_n^{*2} \right] = n\gamma_a^2 \int_0^\infty x^2 \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k e^{-kx} e^{-x} dx \quad (\text{B.17})$$

$$= n\gamma_a^2 \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k \int_0^\infty x^2 e^{-(k+1)x} dx. \quad (\text{B.18})$$

Evaluating the inner integral by parts, we obtain

$$E \left[h_n^{*2} \right] = 2n\gamma_a^2 \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k \frac{1}{(k+1)^3}. \quad (\text{B.19})$$

To simplify the RHS of Equation (B.19) consider the following approach:

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k v^k = (1-v)^{n-1} \quad (\text{B.20})$$

Integrating both sides wrt v from the limits 0 to u and then dividing both sides by u , yields,

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k \frac{u^k}{k+1} = \frac{1}{n} \left[\frac{1 - (1-u)^n}{u} \right]. \quad (\text{B.21})$$

Integrating both sides wrt u from the limits 0 to x and then dividing both sides by x , yields

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k \frac{x^k}{(k+1)^2} = \frac{1}{x} \int_0^x \frac{1}{n} \left[\frac{1 - (1-u)^n}{u} \right] du. \quad (\text{B.22})$$

Integrating both sides wrt x from the limits 0 to 1 yields

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k \frac{1}{(k+1)^3} = \frac{1}{n} \int_0^1 \frac{1}{x} \left(\int_0^x \frac{1 - (1-u)^n}{u} du \right) dx. \quad (\text{B.23})$$

The inner integral can be evaluated by substituting $1-u=v$ and using the summation formula for a geometric progression and a similar procedure can be adapted for the outer integral. The result is

$$\frac{1}{n} \int_0^1 \frac{1}{x} \left(\int_0^x \frac{1 - (1-u)^n}{u} du \right) dx = \frac{1}{n} \sum_{k=1}^n \frac{1}{k} \sum_{l=1}^k \frac{1}{l}. \quad (\text{B.24})$$

From Equations (B.19),(B.23) and (B.24), it follows that

$$E \left[h_n^{*2} \right] = 2\gamma_a^2 \sum_{k=1}^n \frac{1}{k} \sum_{l=1}^k \frac{1}{l} \quad \text{Hence Proved.} \quad (\text{B.25})$$

3. The variance of h_n^* can be calculated as

$$\text{Var}[h_n^*] = E[h_n^{*2}] - (E[h_n^*])^2 \quad (\text{B.26})$$

$$= 2\gamma_a^2 \sum_{k=1}^n \frac{1}{k} \sum_{l=1}^k \frac{1}{l} - \gamma_a^2 \sum_{k=1}^n \sum_{l=1}^n \frac{1}{k} \frac{1}{l} \quad \text{From Parts 1 and 2} \quad (\text{B.27})$$

$$= \gamma_a^2 \sum_{k=1}^n \frac{1}{k} \left(\sum_{l=1}^k \frac{1}{l} - \sum_{l=k+1}^n \frac{1}{l} \right). \quad (\text{B.28})$$

Let us denote $\text{Var}[h_n^*]/\gamma_a^2$ by the shortened notation $V(n)$. Now note that

$$V(n+1) = \sum_{k=1}^{n+1} \frac{1}{k} \left(\sum_{l=1}^k \frac{1}{l} - \sum_{l=k+1}^{n+1} \frac{1}{l} \right) \quad (\text{B.29})$$

$$= \sum_{k=1}^n \frac{1}{k} \left(\sum_{l=1}^k \frac{1}{l} - \sum_{l=k+1}^{n+1} \frac{1}{l} \right) + \frac{1}{n+1} \left(\sum_{l=1}^{n+1} \frac{1}{l} \right) \quad (\text{B.30})$$

$$= \sum_{k=1}^n \frac{1}{k} \left(\sum_{l=1}^k \frac{1}{l} - \sum_{l=k+1}^n \frac{1}{l} - \frac{1}{n+1} \right) + \frac{1}{n+1} \left(\sum_{l=1}^{n+1} \frac{1}{l} \right) \quad (\text{B.31})$$

$$= V(n) + \frac{1}{n+1} \left(\sum_{l=1}^{n+1} \frac{1}{l} - \sum_{k=1}^n \frac{1}{k} \right) \quad (\text{B.32})$$

$$= V(n) + \frac{1}{(n+1)^2}. \quad (\text{B.33})$$

This gives a simple recursive formula of $V(n)$. Since $V(1) = 1$, the recursion can be solved to yield

$$V(n) = \sum_{k=1}^n \frac{1}{k^2}. \quad (\text{B.34})$$

From Equation (B.34) and the definition of $V(n)$, it follows that

$$\text{Var}[h_n^*] = \gamma_a^2 \sum_{k=1}^n \frac{1}{k^2} \quad \text{Hence Proved.} \quad (\text{B.35})$$

Let us now discuss the implications of the Lemma 2.

- The mean $E[h_n^*]$ is directly proportional to the n th partial sum of the Harmonic Series, denoted by \mathcal{H}_n . It is known to diverge albeit very slowly. It can be shown that

$$\log(n) + \frac{1}{2} + \frac{1}{2n} \leq \mathcal{H}_n \leq \log(n) + 1. \quad (\text{B.36})$$

Or $\mathcal{H}_n \sim \log(n)$. Hence as $n \rightarrow \infty$, $E[h_n^*] \rightarrow \infty$.

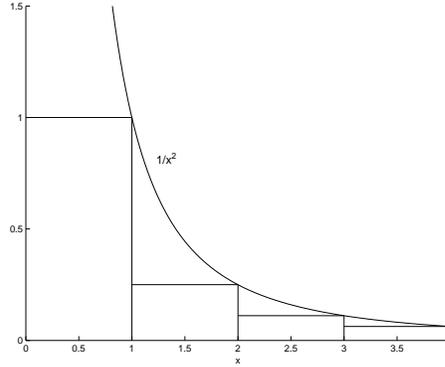


Figure B.1: Plot of $1/x^2$ and $1/k^2$

- The variance $\text{Var}[h_n^*]$ is a convergent sequence of n . This is made clear by Figure B.1. Note that

$$\text{Var}[h_n^*] = \gamma_a^2 \sum_{k=1}^n \frac{1}{k^2} \quad (\text{B.37})$$

$$< \gamma_a^2 \left(1 + \int_1^\infty \frac{1}{x^2} dx \right) \quad (\text{B.38})$$

$$= \gamma_a^2 \left(2 - \frac{1}{n} \right). \quad (\text{B.39})$$

Thus as $n \rightarrow \infty$, $\text{Var}[h_n^*] < 2\gamma_a^2$.

Let us now investigate the variation of the PDF $f_{h_n^*}(x)$ with n . We have already established (Lemma 2), the mean and variance of h_n^* . It can be shown by differentiating $f_{h_n^*}(\gamma)$ that it achieves a global maxima at $\gamma_{n,max} = \gamma_a \log(n)$ with the maximum value $(1 - 1/n)^{n-1}$. The following observations can be made

- As n increases $\gamma_{n,max} \rightarrow E[h_n^*]$
- The maximum value $(1 - 1/n)^{n-1}$ is a decreasing function of n and as $n \rightarrow \infty$ this tends to the finite limit $1/e$.

These findings are consistent with the fact that the mean tends to ∞ but the variance is finite for $n \rightarrow \infty$. Finally Figure (B.2), shows how the PDFs change with n .

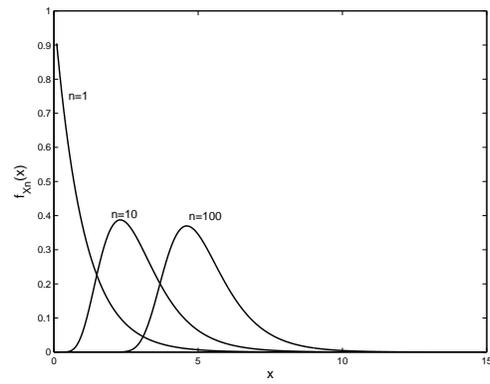


Figure B.2: $f_{h_n^*}(x)$ for various n

Appendix C

Maximization of $\mathbf{x}^T \mathbf{A} \mathbf{x}$ with $\mathbf{x}^T \mathbf{x} = 1$ and $\mathbf{x} \geq 0$

The optimization problem, we seek to solve is

$$\max_{\substack{\mathbf{x}^T \mathbf{x} = 1 \\ \mathbf{x} \geq 0}} \mathbf{x}^T \mathbf{A} \mathbf{x}, \quad (\text{C.1})$$

where $\mathbf{A} \in \mathcal{R}^{n \times n}$ and $\mathbf{x} \in \mathcal{R}^n$. Note that without the positivity constraint $\mathbf{x} \geq 0$, the solution is given by the maximization of the classical Rayleigh Quotient expression [20]. The solution is to choose \mathbf{x} as the eigenvector of \mathbf{A} which has the highest eigenvalue (henceforth referred to as the *maximum eigenvector*). With the addition of the positivity constraint this may no longer be true as the *maximum eigenvector* might have negative components.

There are specific matrices \mathbf{A} for which the optimal \mathbf{x} is again the *maximum eigenvector*. One such case is when \mathbf{A} is *positive*, which means that all the entries of \mathbf{A} are positive (this is different from positive definiteness). For this matrix the Perron-Frobenius theorem, [20] states that the *maximum eigenvector* has all positive components and hence the positivity constraint is automatically satisfied. Another case is when $\mathbf{A} \in \mathcal{R}^{2 \times 2}$, in which case due to the orthogonality of the eigenvectors in a two dimensional space, one of the eigenvectors will always have both entries as positive (or both negative) and that would correspond to the *maximum eigenvector*.

To solve the problem for any arbitrary \mathbf{A} , we write the Lagrangian for the problem in Equation (C.1) as,

$$\mathcal{L} = \mathbf{x}^T \mathbf{A} \mathbf{x} + \lambda \mathbf{x}^T \mathbf{x} + \boldsymbol{\mu}^T \mathbf{x}, \quad (\text{C.2})$$

$$\boldsymbol{\mu} \geq 0 \quad (\text{C.3})$$

$$\lambda \text{ is unconstrained}, \quad (\text{C.4})$$

where λ is a scalar and $\boldsymbol{\mu} \in \mathcal{R}^n$. Taking the partial derivative w.r.t. \mathbf{x} yields,

$$\frac{\partial \mathcal{L}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x} + 2\lambda\mathbf{x} + \boldsymbol{\mu} = 0, \quad (\text{C.5})$$

$$\mu_i x_i = 0, \quad 1 \leq i \leq n, \quad (\text{C.6})$$

$$\mathbf{x}^T \mathbf{x} = 1, \quad (\text{C.7})$$

$$\mathbf{x} \geq 0, \boldsymbol{\mu} \geq 0. \quad (\text{C.8})$$

A broad outline of the solution is provided here. Note that for the purposes of the solution, each individual entry of \mathbf{x} can be thought of being in binary state i.e. either zero or positive. The number of zero elements in the optimal \mathbf{x} can range from 0 to $n - 1$. We now introduce some new definitions. Let all vectors having k zeros be said to belong to the k^{th} class. Each n dimensional vector will have to satisfy certain constraints depending on which class it belongs to. To solve, we first assume that \mathbf{x} belongs to the first class. We then then solve Equation (C.5) and find the vectors which are the stationary points of the Lagrangian. We check if these solutions satisfy the required constraints. If a vector does so, we term it *feasible*. We form the the set of all feasible vectors and call it the *feasible solution set*. We then repeat the procedure for all the k classes. At each class, we potentially keep on adding to the feasible solution set. At the end we calculate the value of the objective $\mathbf{x}^T \mathbf{A}\mathbf{x}$ for all vectors in the feasible solution set and select the vector which maximizes the objective. These are explained in detail below, by considering the different classes.

Class 1: \mathbf{x} has no zero element

In this case the constraints are:

$$x_i > 0, \quad 1 \leq i \leq n \quad (\text{C.9})$$

$$\mu_i = 0, \quad 1 \leq i \leq n, \text{ from Equation (C.6)} \quad (\text{C.10})$$

For these constraints, Equation (C.5) yields,

$$\mathbf{A}\mathbf{x} = -\lambda\mathbf{x}. \quad (\text{C.11})$$

Hence the stationary points of the Lagrangian are given by the eigenvectors of the original matrix \mathbf{A} . The eigenvectors are assumed to be of unit-norm. Amongst these stationary points the feasible vectors are those, for which constraints in Equation (C.9) are satisfied. We add these to the feasible solution set.

Class 2: \mathbf{x} has one zero element

Note that the zero element can occur in any of the n places. We consider all these *sub-classes* one by one. Let us begin with $x_1 = 0$. The constraints become

$$x_i > 0, \quad 2 \leq i \leq n, \quad (\text{C.12})$$

$$\mu_1 > 0, \quad (\text{C.13})$$

$$x_1 = 0 \quad \text{and} \quad \mu_i = 0, \quad 2 \leq i \leq n. \quad (\text{C.14})$$

For this parameters the Equation (C.5) can be expressed as,

$$\mathbf{A}\mathbf{x} + \lambda\mathbf{x} = -\frac{1}{2}\boldsymbol{\mu}, \quad (\text{C.15})$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} + \lambda & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & & a_{nn} + \lambda \end{bmatrix} \begin{bmatrix} 0 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} \mu_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (\text{C.16})$$

This can be rewritten into an eigenvalue problem and a linear equation as follows,

$$\mathbf{a}_1^T \mathbf{x}_1 = -\frac{1}{2}\mu_1, \quad (\text{C.17})$$

$$\mathbf{A}_1 \mathbf{x}_1 = -\lambda \mathbf{x}_1, \quad (\text{C.18})$$

where,

$$\mathbf{a}_1 = [a_{12}, a_{13}, \cdots, a_{1n}], \quad (\text{C.19})$$

$$\mathbf{x}_1 = [x_2, x_3, \cdots, x_n], \quad (\text{C.20})$$

$$\mathbf{A}_1 = \begin{bmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{bmatrix}. \quad (\text{C.21})$$

Note that $\mathbf{A}_1 \in \mathcal{R}^{(n-1) \times (n-1)}$ and $\mathbf{a}_1, \mathbf{x}_1 \in \mathcal{R}^{(n-1)}$. The procedure is now simple. Solve Equation (C.18) and calculate the unit-norm eigenvectors \mathbf{x}_1 . Then,

1. Check if all elements of \mathbf{x}_1 are positive. This is the constraint in Equation (C.12).
2. Calculate μ_1 from Equation (C.17) and check if it is positive. This is the constraint given in Equation (C.13).

If \mathbf{x}_1 satisfies both these constraints then form $\mathbf{x} = [0, \mathbf{x}_1]$ and add \mathbf{x} to the feasible solution set.

Now consider $x_k = 0$, $k \neq 1$. Then $\mu_k > 0$ and $\mu_j = 0$, $j \neq k$. Flip the first and the k^{th} rows of \mathbf{A} and first and the k^{th} elements of both the vectors \mathbf{x} and $\boldsymbol{\mu}$. The resulting system of equations has the same structure as the $x_1 = 0$ case, and can be solved.

Cases 3 – n : \mathbf{x} has k zero elements, $2 \leq k \leq n$

The procedure outlined in Section C can be easily generalized to consider all the remaining cases. For any k , there are $\binom{n}{k}$ ways in the k zeros can be distributed. in \mathbf{x} . Consider all these sub-classes one at a time. Start with the case when $x_1 = x_2 = \dots = x_k = 0$ and hence $\mu_1 > 0$, $\mu_2 > 0, \dots, \mu_k > 0$. Define the vector $[x_{k+1}, \dots, x_n] \in \mathcal{R}^{n-k}$ as \mathbf{x}_k . Then it can be shown that the problem reduces to an eigenvalue problem of the matrix $\mathbf{A}_k \in \mathcal{R}^{(n-k) \times (n-k)}$ and a system of k linear equations as follows,

$$\mathbf{A}_k \mathbf{x}_k = -\lambda \mathbf{x}_k, \tag{C.22}$$

$$\mathbf{a}_{1k+1}^T \mathbf{x}_k = -\frac{1}{2} \mu_1, \tag{C.23}$$

$$\vdots \tag{C.24}$$

$$\mathbf{a}_{kk+1}^T \mathbf{x}_k = -\frac{1}{2} \mu_k, \tag{C.25}$$

where the matrix \mathbf{A}_k is formed by deleting the first k rows and columns of matrix \mathbf{A} and $\mathbf{a}_{jk+1} = [a_{jk+1}, \dots, a_{jn}]$, $1 \leq j \leq k$. The unit-norm eigenvalues of \mathbf{A}_k are computed and the positivity constraints of the elements of \mathbf{x}_k and the values of $\mu_1 \dots \mu_k$ are checked. The vectors which satisfy these constraints are appended initially by k zeros to form feasible \mathbf{x} s. Now for the remaining $\binom{n}{k} - 1$ sub-classes, we can use the flipping

argument as used in Section C and get a similar structure of equations as considered above, which can be solved

After this having formed the feasible solution set, we compute the value of the objective $\mathbf{x}^T \mathbf{A} \mathbf{x}$ for each vector in this set and select the vector which maximizes the objective.

Note that this method solves the optimization problem for any arbitrary matrix \mathbf{A} . However the complexity of the algorithm increases exponentially with n .

References

- [1] J. Proakis. *Digital Communications Fourth Edition*. McGraw Hill.
- [2] K. Liu, T. Kadous, and A. M. Sayeed. Orthogonal time-frequency signaling over doubly dispersive channels. *IEEE Trans. Info Theory*, 50(11):2583–2603, Nov 2004.
- [3] W. Kozek and A. F. Molisch. Nonorthogonal pulse shapes for multicarrier communications in doubly dispersive channels. *IEEE J. Select. Areas Commun.*, 16:1579–1589, Oct 1998.
- [4] R. Haas and J. C. Belfiore. Nonorthogonal pulse shapes for multicarrier communications in doubly dispersive channels. *Wireless Personal Comm.*, 5:1–18, 1997.
- [5] T. Strohmer. Approximation of dual gabor frames, window decay, and wireless communications. *Appl. Comp. Harm. Anal.*, 11(2):243–262, 2001.
- [6] A. Persson, T. Ottosson, and E. Strom. Time-frequency localized CDMA for downlink multi-carrier systems. *Spread Spectrum Techniques and Applications, 2002 IEEE Seventh International Symposium on*, 1, 2002.
- [7] H. Atarashi, N. Maeda, S. Abeta, and M. Sawahashi. Broadband packet wireless access based on VSF-OFCDM and MC/DS-CDMA. *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, 3, September 2002.
- [8] N. Maeda, Y. Kishiyama, H. Atarashi, and M. Sawahashi. Variable spreading factor-OFCDM with two dimensional spreading that prioritizes time domain spreading for forward link broadband wireless access. *IEEE Trans. on Vehic. Tech.*, 1, April 2003.
- [9] N. Maeda, H. Atarashi, S. Abeta, and M. Sawahashi. Throughput comparison between VSF-OFCDM and OFDM considering effect of sectorization in forward link broadband packet wireless access. *IEEE Trans. on Vehic. Tech.*, 1, September 2002.
- [10] Broadband packet wireless access flexibly supporting cellular system and local areas and its experiments. Technical report, Wireless Access Laboratory, Wireless Laboratories, NTT DoCoMo Inc., October 2003.
- [11] N. Yee, J. Linnartz, and G. Fettweis. Multi-carrier CDMA in indoor wireless radio networks. *PIMRC '93, Yokohama, Japan*, 1993.
- [12] S. Abeta, H. Atarashi, and M. Sawahashi. Forward link capacity of coherent DS-SS-CDMA and MC-SS-CDMA broadband packet wireless access in a multi-cell environment. *IEEE Trans. on Vehic. Tech.*, 5, September 2000.

- [13] E. Biglieri, J. Proakis, and S. Shamai (Shitz). Fading channels: Information-theoretic and communications aspects. *IEEE Trans. Info Theory*, 44:2619–2692, Oct. 1998.
- [14] P. P. Varaiya and A. J. Goldsmith. Capacity of fading channels with channel side information. *IEEE Trans. Info Theory*, 43(6):1986 – 1992, Nov 1997.
- [15] G. Caire and S. Shamai (Shitz). On the capacity of some channels with channel state information. *IEEE Trans. Info Theory*, 46(6):2007–2019, Sept 1999.
- [16] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley and Sons, 1991.
- [17] S. Verdu. *Multuser Detection*. Cambridge University Press, Cambridge, UK, 1998.
- [18] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE J. Sel. Areas Commun.*, 21, June 2003.
- [19] M. Mecking. Resource allocation for fading multiple-access channels with partial channel state information. *Communications, 2002. ICC 2002. IEEE International Conference on*, 3:1419 – 1423, April - May 2001.
- [20] G. Strang. *Linear Algebra and its Applications*. International Thompson Publishing, Singapore, 1988.
- [21] S. Verdu. Capacity region of gaussian CDMA channels: The symbol-synchronous case. *Proceedings of the 24th Allerton Conference on Communication, Control and Computing*, 1989.
- [22] M. Rupf and J. L. Massey. Optimum sequence multisets for synchronous code-division multiple-access channels. *IEEE Trans. Info Theory*, 40(4):1226–1266, July 1994.
- [23] P. Viswanath and V. Anantharam. Optimal sequences and sum capacity of synchronous CDMA systems. *IEEE Trans. Info Theory*, 45(6):1984–1991, September 1999.
- [24] O. Kaya and S. Ulukus. Jointly optimal power and signature sequence allocation for fading cdma. In *Proc. IEEE GLOBECOMM 2003*, volume 4, pages 1872–1876, Dec 2003.
- [25] P. Viswanath, D.N.C. Tse, and V. Anantharam. Asymptotically optimal water-filling in vector multiple-access channels. *IEEE Trans. Info Theory*, 47:241–267, January 2001.
- [26] W. Yu, W. Rhee, S. Boyd, and J.M. Cioffi. Iterative water-filling for gaussian vector multiple-access channels. *IEEE Trans. Info Theory*, 50:145–152, January 2004.
- [27] D. Popescu and C. Rose. *Interference Avoidance Methods for Wireless Systems*. Kluwer Academic Publishers.

- [28] L. Zheng and D.N.C. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEEE Trans. Info Theory*, 49:1073–1096, May 2003.
- [29] L. H. Ozarow, S. Shamai, and A. D. Wyner. Information theoretic considerations for cellular mobile radio. *IEEE Trans. on Vehic. Tech.*, 43(2):1527 – 1539, May 1994.