MIMO WIRELESS SYSTEMS WITH LIMITED CHANNEL STATE INFORMATION AT THE TRANSMITTER

by

Shyam Prakash Pandula

A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of the
the Requirements for the Degree
of
Doctor of Philosophy
Electrical and Computer Engineering

Committee:

_________________________ Dr.Bernd-Peter Paris, Dissertation Director
_________________________ Dr.Brian L. Mark, Committee Member
_________________________ Dr.Kristine L. Bell, Committee Member
_________________________ Dr.Jill Nelson, Committee Member
_________________________ Dr.Andre Manitius, Chairman, Department of Electrical and Computer Engineering
_________________________ Dr.Lloyd J. Griffiths, Dean, The Volgenau School of Information Technology and Engineering

Date: ______________________ Fall 2008, George Mason University
Fairfax, VA
MIMO WIRELESS SYSTEMS WITH LIMITED CHANNEL STATE INFORMATION AT THE TRANSMITTER

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Shyam Prakash Pandula
Master of Science
George Mason University, 2003
Bachelor of Technology
Nagarjuna University, 2000

Director: Dr. Bernd-Peter Paris, Associate Professor
Department of Electrical and Computer Engineering

Fall 2008
George Mason University
Fairfax, VA
Dedication
Acknowledgments
# Table of Contents

Abstract ........................................... xi

1 INTRODUCTION .................................. 1
   1.1 Motivation ................................... 2
   1.2 Related Work in Literature .................. 4
   1.3 Thesis Contribution and Outline ............. 9

2 MIMO WIRELESS CHANNEL AND PROPERTIES ....... 12
   2.1 SISO Wireless Channel ....................... 13
      2.1.1 SISO signal model ....................... 13
      2.1.2 Channel characteristics ................. 15
      2.1.3 Statistical channel models ............... 18
   2.2 MIMO Wireless Channel ...................... 19
      2.2.1 MIMO channel correlation models ......... 21
      2.2.2 Rayleigh and Ricean fading ............... 23
      2.2.3 MIMO signal models ....................... 24
      2.2.4 Gains of MIMO Channels .................... 26
   2.3 Modeling Channel State Information at the Transmitter ................. 29
      2.3.1 Channel acquisition at the transmitter ... 30
      2.3.2 Dynamic CSIT model ....................... 31

3 MIMO SYSTEM WITH LINEAR PRECODER .......... 35
   3.1 System Model .................................. 36
   3.2 SVD based Linear Precoding .................. 38
   3.3 MIMO Receiver Architectures ................ 42
      3.3.1 Linear decorrelator receiver ............. 43
      3.3.2 Linear MMSE receiver ...................... 44
      3.3.3 Successive interference cancelation ........ 45

4 MIMO RECEIVERS UNDER LIMITED CSIT .......... 48
   4.1 Limited Channel State Information at the Transmitter ... 49
   4.2 Precoding with Limited CSIT .................. 50
4.3 Performance Analysis of Linear Receivers ........................................... 53
  4.3.1 Signal to interference-plus-noise ratio ....................................... 54
  4.3.2 Mean squared error ............................................................... 55
4.4 Performance Analysis of Successive Interference Cancelation .............. 56
  4.4.1 Signal to interference-plus-noise ratio ....................................... 56
  4.4.2 Mean squared error ............................................................... 57
4.5 System Metrics ........................................................................... 57
4.6 Simulation Results ....................................................................... 58
4.7 Chapter Summary ......................................................................... 64
5 PRECODING SCHEMES EXPLOITING COVARIANCE CSIT ....................... 66
  5.1 Precoder Design Criteria ............................................................. 68
  5.2 Asymptotic Random Matrix Theory ................................................ 70
    5.2.1 Uncorrelated Rayleigh MIMO channel ........................................ 72
    5.2.2 Correlated Rayleigh MIMO channel .......................................... 75
  5.3 Maximizing the System Ergodic Capacity ......................................... 77
    5.3.1 Linear precoder design ............................................................ 79
  5.4 Maximizing the Capacity of System with MMSE Receiver .................. 80
    5.4.1 Linear precoder design ............................................................ 82
  5.5 Simulation Results ....................................................................... 85
  5.6 Chapter Summary ......................................................................... 89
6 EXPLOITING MEAN AND COVARIANCE CSIT IN MIMO SYSTEMS WITH DECORRELATOR RECEIVER ......................................................... 91
  6.1 System Model and Design Criterion ................................................ 92
  6.2 Moments of \(\text{SNR}_k\) ................................................................. 95
  6.3 Average Capacity Approximation .................................................. 99
    6.3.1 Via Jensen’s inequality ............................................................. 99
    6.3.2 Via Taylor series expansion ...................................................... 99
  6.4 Linear Precoder Design ............................................................... 100
    6.4.1 Precoder rotation matrix .......................................................... 100
    6.4.2 Power allocation ................................................................. 101
  6.5 Simulations results ..................................................................... 102
  6.6 Chapter Summary ....................................................................... 104
7 EXPLOITING MEAN AND COVARIANCE CSIT VIA APPROXIMATE JOINT DIAGONALIZATION IN STBC MIMO SYSTEMS ......................... 105
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Statistical SISO channel models</td>
</tr>
<tr>
<td>2.2</td>
<td>Statistical MIMO channel models</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Capacity of $M_R \times M_T$ MIMO channels with and without CSIT.</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>Block diagram of a MIMO system.</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>MIMO System with Linear Precoder</td>
<td>36</td>
</tr>
<tr>
<td>3.2</td>
<td>Equivalent decomposed channel</td>
<td>41</td>
</tr>
<tr>
<td>4.1</td>
<td>Capacity of MIMO channel with a non-zero mean and transmit antenna correlation under limited CSIT, SNR=10 dB.</td>
<td>59</td>
</tr>
<tr>
<td>4.2</td>
<td>Capacity of MIMO channel with a non-zero mean and transmit antenna correlation under limited CSIT, SNR=20 dB.</td>
<td>61</td>
</tr>
<tr>
<td>4.3</td>
<td>Capacity of MIMO channel with zero mean and uncorrelated antennas under instantaneous CSIT, SNR=10 dB.</td>
<td>62</td>
</tr>
<tr>
<td>4.4</td>
<td>Capacity of MIMO channel with a non-zero mean and transmit antenna correlation using MMSE receiver.</td>
<td>63</td>
</tr>
<tr>
<td>4.5</td>
<td>MSE of the subchannels with instantaneous CSIT, SNR=10 dB.</td>
<td>63</td>
</tr>
<tr>
<td>4.6</td>
<td>MSE of the subchannels with instantaneous CSIT, SNR=10 dB.</td>
<td>64</td>
</tr>
<tr>
<td>5.1</td>
<td>Ergodic capacity vs SNR. Asymptotic values match the simulations very well.</td>
<td>86</td>
</tr>
<tr>
<td>5.2</td>
<td>SINR of each spatial substream vs SNR. Asymptotic values match the simulations very well.</td>
<td>86</td>
</tr>
<tr>
<td>5.3</td>
<td>Ergodic capacity vs $\phi_i$ at SNR=6 dB. Power constraint: $\sum_i \phi_i = M_T$. Ergodic capacity is concave in $\phi_i$.</td>
<td>87</td>
</tr>
<tr>
<td>5.4</td>
<td>Capacity with an MMSE receiver vs $\phi_i$ at SNR=6 dB. Power constraint: $\sum_i \phi_i = M_T$. System capacity is concave in $\phi_i$.</td>
<td>87</td>
</tr>
<tr>
<td>5.5</td>
<td>Performance of the proposed precoding scheme to maximize system ergodic capacity.</td>
<td>88</td>
</tr>
</tbody>
</table>
5.6 Performance of the proposed precoding scheme to maximize the capacity of system with MMSE receiver. ............................. 89

6.1 Capacity vs SNR without precoding. Taylor series approximation is more accurate than Jensen bound. K-factor=1. ............... 102

6.2 Performance with precoding: Capacity vs SNR. K-factor=0.1. .... 103

6.3 Performance with precoding: Capacity vs K-factor. ...................... 103

7.1 Joint diagonalization of $H^H_m H_m$ and $R_t$ is better than diagonalizing $H^H_m H_m$ or $R_t$ alone. With Golden code using QPSK at SNR= 5 dB. Correlation coefficient $\alpha=0.7$ ................................. 117

7.2 Performance of the precoded system when a minimum distance Golden codeword pair occurs. K-factor=0.1 and Correlation coefficient $\alpha=0.7$ 117

7.3 Performance of the precoded system when a minimum distance ABBA codeword pair occurs. K-factor=0.1 and Correlation coefficient $\alpha=0.7$ 118

7.4 Union bound on the average probability of codeword error using QPSK. ABBA code, K-factor=0.1 and Correlation coefficient $\alpha=0.7$ ........ 119
Abstract

MIMO WIRELESS SYSTEMS WITH LIMITED CHANNEL STATE INFORMATION AT THE TRANSMITTER
Shyam Prakash Pandula, PhD
George Mason University, 2008
Dissertation Director: Dr. Bernd-Peter Paris

Multiple input multiple output (MIMO) systems using multiple transmit and receive antennas are widely recognized as an effective means to significantly improve the spectral efficiency and wireless link reliability. The performance of a MIMO system can be further enhanced when channel knowledge can be made available at the transmitter. Maximum gain can be achieved when perfect channel knowledge is available. In practice, however, perfect channel knowledge at the transmitter is never available because of the time varying nature of the channel and due to unavoidable estimation errors. As a result, the available channel state information at the transmitter (CSIT) is usually outdated instantaneous and/or statistical in nature.

Precoding is a transmitter processing technique that exploits CSIT. A linear precoder functions as a multimode beamformer that spatially directs the signal in orthogonal directions and allocates power based on the available channel knowledge. The main objective of this thesis is to design robust precoding schemes when limited
channel state information is available at the transmitter. Throughout the thesis, perfect channel knowledge is assumed at the receiver.

Our first contribution is a performance analysis for the decorrelator, minimum mean squared error and successive interference cancelation receivers for precoded systems under various degrees of CSIT. We show that significant performance gain can be achieved by precoding even with only a moderate amount of correlation between the available imperfect channel estimate and the current channel.

Next, we address the case when channel covariance information alone is available at the transmitter. The precoder spatially directs the signal based on the covariance information. Using results for random matrix theory, we propose novel iterative power allocation algorithms that maximize the spectral efficiency.

We proceed to analyze the more general case when both the channel mean and covariance information is available at the transmitter. For the MIMO system with decorrelator receiver, we approximate the SNR of each spatial stream by a standard noncentral Chi-squared random variable. Using the moments of the SNR of each subchannel, we obtain a Taylor series approximation for the total average capacity. The obtained approximation is used to design a linear precoder that maximizes the average capacity of the system.

Finally, we present a novel linear precoder design for general space-time block coded (STBC) MIMO systems when both the channel mean and covariance information is available at the transmitter. The linear precoder is designed via approximate joint diagonalization of the channel mean and transmit covariance matrices. The resulting precoder minimizes the Chernoff bound on the pair-wise error probability between a pair of block code words, averaged over channel fading statistics.
Chapter 1: INTRODUCTION

MIMO wireless communication systems has recently emerged as one of the most significant technical breakthroughs in modern communications. The technology figures prominently on the list of recent technical advances with a chance of resolving the bottleneck of traffic capacity in future Internet-intensive wireless networks. Just a few years after its invention the technology seems poised to penetrate large-scale, standards-driven, commercial wireless products and networks such as broadband wireless access systems, wireless local area networks (WLAN), third-generation (3G) networks and beyond [1]. A key feature of MIMO systems is the ability to turn multipath propagation, traditionally a pitfall of wireless transmission, into a benefit for the user. MIMO effectively takes advantage of random fading [2–4] and when available, multipath delay spread [5,6] for multiplying transfer rates. The prospect of many orders of magnitude improvement in wireless communication performance at no cost of extra spectrum (only hardware and complexity are added) is largely responsible for the success of MIMO as a topic for new research. This has prompted progress in areas as diverse as channel modeling, information theory and coding, signal processing and antenna design.
1.1 Motivation

The ergodic capacity of a MIMO Rayleigh fading channel is plotted in Figure 1.1 for different number of antennas. We see that at moderate to high SNR, the capacity of an $n \times n$ MIMO system is about $n$ times the capacity of the single input single output (SISO) system. While benefits of MIMO are realizable when the receiver alone knows the communication channel, these are further enhanced when the transmitter also knows the channel.

In order to exploit the channel knowledge at the transmitter, a transmit processing technique called precoding, which operates on the signal before transmitting from the antennas, can be used. A linear precoder functions as a multimode beamformer, which optimally matches the input signal on one side to the channel on the other side. With perfect CSIT, precoding decouples the channel into parallel spatial substreams. The transmit power can then be distributed optimally across these spatial substreams in order to maximize spectral efficiency. As we show in Figure 1.1, when the transmitter knows the channel, the capacity of the precoded system is noticeably higher than the capacity with no channel knowledge at the transmitter. Hence, exploiting the channel state information at the transmitter (CSIT) is of great practical interest and has been an active research area.

In practical systems, perfect CSIT is rarely available. Typically, transmission of a short training sequence allows estimation of the channel at the receiver with sufficient accuracy. Channel information at the transmitter can be obtained either by means of feedback from the receiver or from previous received measurement by relying on
the reciprocity principle. In either case, there exists a delay from when the channel information is obtained to when it is used by the transmitter, such as a scheduling or a feedback delay [7]. The channel is time-varying and because of the delay, the available instantaneous CSIT is rendered imperfect.

Another form of CSIT involves statistical information, such as the antenna correlations [8] and/or the channel mean [9]. These statistics vary at a much slower rate than the instantaneous channel and, therefore, can be obtained reliably at the transmitter. Statistical CSIT is especially relevant for fast time-varying channels. This thesis focuses on exploiting the available imperfect instantaneous CSIT or statistical CSIT by designing robust precoding and transmit optimization schemes.
1.2 Related Work in Literature

The communication technique employed for transmission over the MIMO channel fundamentally depends on the degree of channel state information available at the transmitter and at the receiver. Clearly, the more the channel information, the better the performance of the system. We assume perfect channel knowledge at the receiver and we briefly review the existing transmission techniques for different degrees of channel state information at the transmitter.

**Transmission techniques with no CSIT**

Transmission techniques that do not require CSIT can be encompassed within two main philosophies: space-time coding and layered architectures. Space-time coding generalizes the classical concept of coding in the temporal domain to coding in both spatial and temporal dimensions [10]. Layered architectures refer to a particular case of a space-time coding when a separate coding scheme is used for each spatial branch [3].

The idea of space-time coding is that the transmitter introduces redundancy in the transmitted signal, both over space and time, that allows the receiver to recover the signal even in difficult propagation situations. The symbols are first encoded and then the encoded data is split into $M_T$ substreams that are simultaneously transmitted using $M_T$ transmit antennas. The received signal is decoded using a maximum likelihood (ML) decoder. STC is very effective as it combines the benefits of forward error correction coding and diversity transmission to provide considerable performance gains.
There are two main types of space-time coding techniques: space-time trellis coding (STTC) and space-time block coding (STBC). STTC may not be practical or cost-effective due to the high complexity of the maximum likelihood (ML) detector. In an attempt to reduce decoding complexity, Alamouti [11] discovered orthogonal STBC for two antennas that can be optimally decoded with a simple linear processing at the receiver. This scheme was latter generalized in [12] to an arbitrary number of antennas.

Spatial multiplexing (SM) can be regarded as a special class of space-time block codes where streams of independent data are transmitted over different antennas, thus maximizing the average data rate over the MIMO system. Each of the $M_T$ transmitted data streams is independently encoded. At the receiver, with the use of interference suppression and interference cancellation, these data streams can be separated and then decoded using conventional decoding algorithms developed for one dimensional codes, leading to a much lower decoding complexity compared to ML decoding. A particular implementation approach for SM is the diagonal layered space time (DLST) architecture, originally proposed by Foschini in [3]. In [13], the less complex horizontal LST (HLST) architecture was considered and the name BLAST was coined standing for Bell-Labs Layered Space-Time architecture.

**Transmission techniques with perfect CSIT**

When perfect channel knowledge is available at the transmitter, the transmission can be adapted to each channel realization so as to improve the link performance. The scheme to exploit the available CSIT depends on various factors like the performance
criterion to be optimized, the receiver architecture, and the power constraints at the transmitter.

A joint linear transmit and receive optimization scheme using the MMSE criterion has been proposed in [5,14] and is further discussed in [15]. The devised optimal transmit and receive filter pair decouples the MIMO channel into \( \min(M_T, M_R) \) eigen subchannels. Power is allocated to these subchannels using the waterfilling algorithm. Several other criteria have been utilized as a means to measure the quality of the link and to design the system, such as the maximization of the signal to interference-plus-noise ratio (SINR) with a zero-forcing (ZF) constraint [16], the minimization of the determinant of the MSE matrix [17], and the minimization of the average bit error rate (BER) [18,19].

An elegant general unifying framework was developed in [19,20] to treat all such criteria by classifying a variety of objective functions into Schur-concave and Schur-convex functions. For Schur-concave objective functions, the channel-diagonalizing structure is always optimal, whereas for Schur-convex functions, an optimal solution diagonalizes the channel only after a very specific rotation of the transmitted symbols. Once the optimal structure of the transmit-receive processing is known, the design problem simplifies and can be formulated within the powerful framework of convex optimization theory, in which a great number of interesting design criteria can be easily accommodated and efficiently solved [21–23].

Transmission can be performed through a sub-set of the available antenna elements
so as to minimize the cost of hardware associated with every additional antenna. Antenna sub-set selection under various optimization criteria has been performed in [24,25]. In [26], the joint transmit receive MMSE design has been optimized using mode selection.

**Transmission techniques with limited CSIT**

Limited CSIT refers to outdated instantaneous or statistical channel information at the transmitter. Outdated instantaneous CSIT arises due to channel variations during the feedback delay or due to channel estimation errors at the receiver. Statistical CSIT refers to channel statistics like mean and covariance that can be made available at the transmitter.

The robust transmission techniques can be classified into stochastic (or Bayesian) and worst case (or maximin) approaches, depending on the way the errors in the CSIT are modeled. The stochastic approach seeks to optimize the average precoder performance assuming that the statistics of the error are known. On the other hand, maximin techniques consider that the error belongs to a predefined uncertainty region, with no inherent statistical assumption, and the final objective is the optimization of the worst system performance for any error in this region. Most of the stochastic precoder designs assume knowledge of either the channel transmit covariance or the channel mean at the transmitter. These include precoders optimal for the channel ergodic capacity, given transmit covariance CSIT [27–29], or mean CSIT [27,30]. Others are based on an error rate criterion with mean CSIT [31–33], or transmit covariance CSIT [34,35]. The precoding solutions for these partial-CSIT cases then
reduce to fixed beam directions at all SNRs, given by the singular- or eigen-vectors of the mean or transmit covariance matrix, with per-beam power allocation, obtained by a numerical water-filling solution dependent on the SNR.

In [36], SM systems in which the matrices containing the left and right singular vectors of the channel matrix form the receiver and the precoder respectively have been considered. The interference among the subchannels when the error in the available CSIT is small has been evaluated using matrix perturbation theory. Performance of SVD based systems with linear receivers under imperfect CSIT has been addressed in [37]. It was noted that imperfect CSIT causes interference among the subchannels which severely degrades performance unless the channel variations and feedback delays are very small. A maximin design of a transmitter in a MIMO channel was proposed in [38,39] for the case of spherical uncertainty regions. In [40], a game-theoretic approach was proposed to maximize the channel capacity for the least favorable channel within a MIMO channel set specified by eigenvalue inequality constraints.

Even though a worst case design guarantees a certain system performance for any possible channel sufficiently close to the estimated one, this approach is extremely pessimistic and translates into a significant increase of the required transmit power. On the other hand, the stochastic design only guarantees a certain system performance averaged over the channels that could have caused the current estimated channel, but avoids the high pessimism inherent to the worst-case design with the consequent saving in transmit power.
1.3 Thesis Contribution and Outline

The following are the main contributions of this thesis.

- We analyze the performance of various MIMO receivers under outdated instantaneous CSIT. The precoder is derived from the singular value decomposition (SVD) of the available channel state information at the transmitter (CSIT) and the receiver is a function of the precoder and the current channel. With perfect CSIT, the $M_R \times M_T$ flat-fading MIMO channel can be decomposed into $\min(M_T, M_R)$ parallel spatial subchannels. However in practice, the available CSIT suffers from delay-induced error due to the channel temporal variations. Using this outdated CSIT for precoding in SM systems causes interference among the subchannels. Performance of the decorrelator, minimum mean squared error (MMSE) and successive interference cancelation (SIC) receivers is analyzed as the reliability of the available CSIT varies. Explicit expressions for the signal to interference-plus-noise ratio (SINR) and the mean squared error (MSE) are derived. We show that significant performance gain achieved by precoding even with a moderate amount of correlation between the available outdated channel estimate and the current channel.

- We address the capacity optimization problem for MIMO wireless channels when the transmit antenna correlation alone is known at the transmitter. Near-optimal linear precoders are designed under the following criteria: Maximizing the system ergodic capacity and maximizing the capacity of a system with MMSE receiver. The proposed precoder designs are based on the results obtained via asymptotic spectral analysis of random matrices. Optimal precoding
beam directions are given by the eigenvectors of the channel covariance matrix. We propose novel iterative power allocation algorithms that maximize the capacity.

- We address the capacity optimization problem for MIMO wireless channels for a more general case when both the channel mean and transmit antenna correlation are known at the transmitter. With a decorrelator receiver, the capacity of the MIMO system is a function of the diagonal elements of an inverted noncentral Wishart distributed matrix. Hence, finding the average capacity is difficult. We simplify the problem by approximating the SNR of each spatial stream by a standard noncentral Chi-squared random variable. The degrees of freedom depend on the number of transmit and receive antennas and the noncentrality parameter depend on the channel mean and covariance matrices. Using the moments of the SNR, we obtain a Taylor series approximation for the average capacity that is significantly better than the commonly used bound via Jensen’s inequality. The obtained Taylor series approximation is used to design a linear precoder that maximizes the total average capacity of the system.

- We propose a novel linear precoder design based on the minimum pair-wise error probability (PEP) criterion for general space-time block coded (STBC) MIMO systems. The channel state information at the transmitter consists of the channel mean and transmit antenna covariance matrices. Approximate joint diagonalization reveals the “average eigen-structure” shared by the mean and covariance matrices. These average eigenvectors turn out to be the left singular vectors of the precoder matrix. The right singular vectors of the precoder
matrix depend on the space-time code. The singular values correspond to the power allocated to each mode, which depends on the channel statistics and the space-time code. We provide precoder design examples for a non-orthogonal STBCs.

The rest of the thesis is organized as follows. In Chapter 2, we review MIMO wireless channel properties and we introduce a model for channel state information at the transmitter. In Chapter 3, we introduce singular value decomposition based linear precoding and we review various MIMO receiver architectures. In chapters 4, 5, 6 and 7 we present our main contributions. We conclude in Chapter 8 and present suggestions for future research.
Chapter 2: MIMO WIRELESS CHANNEL AND PROPERTIES

A defining characteristic of the mobile wireless channel is the variation of the channel strength over time and frequency. The variations are usually characterized as large-scale fading and small-scale fading [41]. Large-scale fading characterizes the path loss of the signal as a function of distance and shadowing by large objects such as buildings and hills. This occurs as the mobile moves a distance on the order of several wavelengths and is typically frequency independent. Small-scale fading captures the variations due to constructive and destructive interference of the multiple signal paths between the transmitter and receiver. This occurs at the spatial scale on the order of the carrier wavelength and is frequency dependent. Large-scale fading is more relevant to issues such as cell-site planning. Small-scale fading is more relevant to the design of reliable and efficient physical layer communication systems.

In this chapter, we first review the scalar single input single output (SISO) channel models and and their small-scale characteristics. Next, we will discuss the matrix multiple input multiple output (MIMO) channel that arises in wireless links with multiple antennas at the transmitter and receiver. Multiple antennas bring an additional spatial dimension and the MIMO channel model, which builds upon the scalar channel model, has several new parameters.
2.1 SISO Wireless Channel

2.1.1 SISO signal model

Consider a single input single output (SISO) wireless link. The transmitted signal that is launched into the wireless environment arrives at the receiver along a number of distinct paths, referred to as multipaths. The path associated with the $i^{th}$ distinct scatterer is characterized by $a_i(t)$ and $\tau_i(t)$, where $a_i(t)$ represents the amplitude fluctuation introduced to transmitted signal by the $i^{th}$ scatterer at time $t$ and $\tau_i(t)$ is the associated propagation delay. Consider a narrowband signal $x(t)$ transmitted over the wireless channel at a carrier frequency $f_c$, such that

$$x(t) = Re[x_b(t)e^{j2\pi f_c t}],$$

(2.1)

where $x_b(t)$ is the baseband equivalent of the transmitted signal. Assuming a multipath propagation environment with $K$ distinct scatterers, in the absence of noise, the received signal at the channel output is

$$y(t) = \sum_{i=1}^{K} a_i(t)x(t - \tau_i(t))$$

(2.2)

$$= Re\left\{\sum_{i=1}^{K} a_i(t)x_b(t - \tau_i(t))e^{-j2\pi f_c \tau_i}e^{j2\pi f_c t}\right\}$$

Hence, the baseband equivalent received signal is

$$y_b(t) = \sum_{i=1}^{K} a_i^b(t)x_b(t - \tau_i(t)),$$

(2.3)
where \( a_i^k(t) = a_i(t)e^{-j2\pi f_c \tau_i(t)} \).

The input/output relationship in (2.3) is also that of a linear time-varying system, and the baseband equivalent impulse response is

\[
h_b(\tau, t) = \sum_{i=1}^{K} a_i^k(t)\delta(t - \tau_i(t)).
\] (2.4)

The impulse response \( h_b(\tau, t) \) is the channel output at time \( t \) in response to an impulse applied to the channel at \( t - \tau \). The channel is completely characterized by the number of multipath components \( N \) and the path variables: amplitude \( a_i(t) \), delay \( \tau_i(t) \), and phase \( \theta_i(t) = 2\pi f_c \tau_i(t) \). These parameters change randomly with time and are often described statistically. Changes in \( \theta_i(t) \) have a far greater effect on the transmitted signal than changes in \( a_i(t) \) because a small change, such as motion, in scatterer can cause a significant change in the phase \( \theta_i(t) \), but may not cause significant changes in the amplitude \( a_i(t) \).

For digital communication systems, the signal is processed in the sampled domain. Hence, in order to create a useful channel model, it is necessary to convert the continuous-time channel to a discrete-time channel. Sampling at frequency twice the maximum signal bandwidth, Nyquist rate, allows perfect reconstruction of the continuous signal from its samples. In wireless communications, the received signal sometimes needs to be sampled at a slightly higher frequency than the Nyquist rate because of possible bandwidth expansion through the channel. We assume that an appropriate sampling frequency has been chosen. A single time-sample of the channel response at a specific delay is called a channel tap. Depending on the sampling time,
the number of distinguishable channel taps $L$ is often smaller than the number of multipaths ($L \leq K$). From (2.4), the discrete channel response at the discrete time $m$ to a unit impulse applied at time $m - l$ is given as

$$h[m, l] = \sum_{l=1}^{L} h_l[m] \delta[m - l].$$

(2.5)

Where, $h_l[m]$ is the $l^{th}$ complex channel filter tap at time $m$ and is given as

$$h_l[m] = \sum_{i=1}^{N_l} a_i[m] e^{-j2\pi f_c \tau_i[m]}.$$  

(2.6)

Here $N_l$ is the number of paths contributing to the $l^{th}$ channel tap and $l = 1, 2, \ldots, L$.

2.1.2 Channel characteristics

Temporal coherence

An important channel parameter is the time-scale of the variation of the channel filter taps $h_l[m]$. Let $\nu_i$ denote the velocity with which the $i^{th}$ path length is increasing and let $\theta_i$ denote the angle of incidence of the $i^{th}$ path viewed from the receiver. With a signal frequency $f_c$, the Doppler shift on the $i^{th}$ path is given as

$$f_d^i = \frac{f_c}{c} \nu_i \cos(\theta_i),$$

(2.7)

where $c$ is the velocity of light. When the different paths contributing to the $l^{th}$ tap have different Doppler shifts, the magnitude of $h_l[m]$ changes significantly. These variations occur at a time-scale inversely proportional to the largest difference between
the doppler shifts. Doppler spread is given by

$$D_s = \max_{i,j} |f_i^d - f_j^d|,$$  \hspace{1cm} (2.8)

where the maximum is taken over all the paths that contribute significantly to a tap.

The coherence time $T_c$ of a wireless channel is defined as the interval over which the magnitude of $h_i[m]$ changes significantly as a function of $m$ and is related to Doppler spread as

$$T_c \approx \frac{1}{D_s}$$  \hspace{1cm} (2.9)

The channel is called fast fading if the coherence time $T_c$ is much smaller than the delay requirements of the application, and is slow fading if $T_c$ is longer. Thus, whether a channel is fast or slow fading depends not only on the environment but also on the application.

The channel tap gain auto-correlation function characterizes how rapidly the channel decorrelates with time and is defined as

$$\rho_l[\tau] = \mathbb{E}\{h_i^*[m]h_i[m+\tau]\}.$$  \hspace{1cm} (2.10)

A common assumption is that the taps are stationary and statistically independent. Hence, $\rho_l$ depends only on the time difference $\tau$. A popular statistical model for flat fading is Clark’s model [42], which assumes uniformly distributed scatterers on a circle around the antenna. In this model, the tap gain auto-correlation function is given as

$$\rho_l[\tau] = J_0(2\pi f_d \tau),$$  \hspace{1cm} (2.11)
where $J_0(.)$ is the zeroth-order Bessel function of the first kind and $f_d$ is the maximum Doppler frequency. For example, with a signal frequency $f_c = 1$ GHz and velocity $\nu = 30$ m/sec, the value of the tap gain auto-correlation function $\rho_l[\tau]$ decreases from 1 to 0 as the value of the time delay $\tau$ increases from 0 to 3.82ms. Hence, the channel changes within a few milliseconds.

**Spectral coherence**

Another important parameter of a wireless system is the multipath delay spread, $T_d$, defined as the difference in propagation time between the longest and shortest paths with significant energy. Thus,

$$T_d = \max_{i,j} |\tau_i(t) - \tau_j(t)|. \quad (2.12)$$

The delay spread of the channel dictates its frequency coherence. Coherence bandwidth, $W_c$, indicates approximately the frequency separation at which the channel behaves independently and is related to delay spread as

$$W_c \approx \frac{1}{T_d}. \quad (2.13)$$

When the bandwidth of the input is considerably less than $W_c$, the channel is referred to as flat fading or frequency non-selective. In this case, the delay spread $T_d$ is much less than the symbol time, and a single channel filter tap is sufficient to represent the channel. When the bandwidth is much larger than $W_c$, the channel is said to be frequency selective, and it has to be represented by multiple taps. Hence, flat or frequency selective fading is not a property of the channel alone, but of relationship
between the bandwidth of the input and coherence bandwidth $T_d$.

### 2.1.3 Statistical channel models

Because of the often unpredictable time-varying nature of a wireless channel, the channel is modeled as a random process. The probabilistic model for the channel filter taps is based on the assumption that each tap $h_l[m]$ in (2.6) is an aggregate of large number of statistically independent reflected and scattered paths with random amplitudes. The phases of these paths vary rapidly with time and are often modeled as independent uniform random variables in $[0, 2\pi]$.

The contribution of each path in the tap gain can be modeled as a circular symmetric complex random variable. Each tap $h_l[m]$ is sum of large number of such small independent circular symmetric random variables. It follows that by applying the central limit theorem, $h_l[m]$ can reasonably be modeled as a zero-mean circular symmetric complex Gaussian random variable, $h_l[m] \sim \mathcal{CN}(0, \sigma^2_l)$. It is assumed that the variance of $h_l[m]$ is a function of the tap $l$, but independent of time $m$. With the assumed Gaussian probability density, we know that the magnitude $|h_l[m]|$ of the $l^{th}$ tap is a Rayleigh random variable. This model is called Rayleigh fading model and has been verified empirically to be a good fit for many channels, especially when there are many scatterers in the environment and no direct line-of-sight between the transmitter and the receiver.

When there is a direct line-of-sight path or a cluster of strong paths, the channel
will have a non-zero mean often called a specular component. In this case, $h_l[m]$, at least for one value of $l$, can be modeled as

$$h_l[m] = \sqrt{\frac{k}{k+1}} h_{\text{spec}} + \sqrt{\frac{1}{k+1}} \mathcal{CN}(0, \sigma_l^2)$$

with the first term corresponding to the deterministic specular component and the second term corresponding to the aggregation of the large number of reflected and scattered paths. The magnitude of such a random variable is said to have a Ricean distribution. The parameter $k$ is called the $K$-factor and is the ratio of the energy of the specular path to the energy in the scattered paths. As $k$ approaches zero, the Ricean distribution approaches the Rayleigh distribution. On the other hand, as $k$ approaches infinity, only the specular component matters and there is no fading.

Table 2.1: Statistical SISO channel models

<table>
<thead>
<tr>
<th>Statistical SISO channel models:</th>
<th>$h_l[m]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayleigh fading</td>
<td>$h_l[m] \sim \mathcal{CN}(0, \sigma_l^2)$</td>
</tr>
<tr>
<td>Ricean Fading</td>
<td>$h_l[m] = \sqrt{\frac{k}{k+1}} h_{\text{spec}} + \sqrt{\frac{1}{k+1}} \mathcal{CN}(0, \sigma_l^2)$,</td>
</tr>
</tbody>
</table>

### 2.2 MIMO Wireless Channel

We consider a MIMO system with a transmit array of $M_T$ transmit antennas and a receive array of $M_R$ antennas. The MIMO channel comprises of $M_T M_R$ SISO links and the block diagram of such a system is shown in Figure 2.1.

For simplicity, we first examine flat-fading MIMO channels. In this case the
channel has only a single tap. This tap, however, contains $M_T M_R$ elements and is represented by an $M_R \times M_T$ matrix

$$H = \begin{bmatrix} h_{1,1} & \cdots & h_{1,M_T} \\ \vdots & \ddots & \vdots \\ h_{M_R,1} & \cdots & h_{M_R,M_T} \end{bmatrix} \quad (2.15)$$

The $(i, j)$-th element of the matrix $H$, $h_{i,j}$, is the SISO channel between the $j^{th}$ transmit antenna and the $i^{th}$ receive antenna. The time-dependent index $[m]$ has been omitted for clarity. The properties of each $h_{i,j}$ can be described by the same concepts introduced in the previous subsection on SISO channels. The entries of $H$ have all the same statistical properties and hence, the coherence time and the coherence bandwidth of a MIMO channel can be evaluated as in the SISO case.
2.2.1 MIMO channel correlation models

Spatial correlation

One of the distinguishing features of the MIMO channel is the spatial correlation among the elements of the channel matrix $H$. Small antenna spacing is the primary cause of spatial correlation. The $M_T M_R \times M_T M_R$ channel covariance matrix that captures the spatial correlation among all the transmit and receive antennas is defined as

$$ R_0 = \mathbb{E}\{\text{Vec}(H)\text{Vec}(H)^H\} \quad (2.16) $$

$R_0$ is a positive semidefinite Hermitian matrix. Its diagonal elements represent the power gain of the $M_T M_R$ scalar channels, and the off-diagonal elements are the cross-coupling between these scalar channels.

The covariance $R_0$ is often assumed to have a simplified, separable Kronecker structure. The Kronecker model assumes that the covariance of the scalar channels seen from all $M_T$ transmit antennas to a single receive antenna (corresponding to a row of $H$) is the same for any receive antenna (any row) and equals to $R_t$.

Let $h_i^T$ be the $i^{th}$ row of $H$, then for any $i$

$$ R_t = \mathbb{E}\{h_i h_i^T\} \quad (2.17) $$

Similarly, the covariance of the scalar channels seen from a single transmit antenna to all $M_R$ receive antennas (corresponding to a column of $H$) is assumed to be the same for any transmit antenna (any column) and equals to $R_r$. That is, for any $j$

$$ R_r = \mathbb{E}\{h_j h_j^T\}. \quad (2.18) $$
Both covariance matrices $\mathbf{R}_t$ and $\mathbf{R}_r$ are complex Hermitian positive semidefinite. The channel covariance can now be decomposed as

$$\mathbf{R}_0 = \mathbf{R}_t^T \otimes \mathbf{R}_r$$

(2.19)

where $\otimes$ denotes the Kronecker product.

**Temporal correlation**

The channel auto-covariance characterizes how rapidly the channel decorrelates with time. Assuming stationarity, the $M_T M_R \times M_T M_R$ channel auto-covariance matrix depends only on the time difference and is given as

$$\mathbf{R}[n] = \mathbb{E}\{\text{Vec}(\mathbf{H}[m])\text{Vec}(\mathbf{H}[m+n])^H\},$$

(2.20)

where $\mathbf{H}[m]$ and $\mathbf{H}[m+n]$ are the matrix channel samples at time $m$ and $m+n$ respectively. When $n = 0$, this auto-covariance matrix coincides with the channel covariance matrix $\mathbf{R}_0$; when $n$ becomes large, $\mathbf{R}[n]$ eventually decays to zero. For a MIMO channel, the covariance $\mathbf{R}_0$ captures the spatial correlation among all the transmit and receive antennas, while the auto-covariance at a non-zero delay $\mathbf{R}[n]$ captures both channel spatial and temporal correlations.

Based on the premise that the channel temporal statistics can be the same for all antenna pairs, it may be assumed that the temporal correlation is homogeneous and identical for any channel element. Then, the two correlation effects are separable, and the channel auto-covariance becomes their product as

$$\mathbf{R}[n] = \rho[n] \mathbf{R}_0,$$

(2.21)
where $\rho[n]$ is the temporal correlation of a scalar channel (2.11). In other words, all the $M_T M_R$ scalar channels between the $M_T$ transmit and $M_R$ receive antennas have the same temporal correlation function. This temporal correlation is a function of the time difference $n$ and the channel Doppler spread.

### 2.2.2 Rayleigh and Ricean fading

We first define a $M_R \times M_T$ spatially white matrix $H_w$. All the elements of $H_w$ are independent and identically distributed zero-mean circular symmetric complex Gaussian (ZMCSCG) random variables.

When there is no line-of-sight between the transmitter and receiver, the channel is modeled as Rayleigh fading as described in section 2.1.3. Rayleigh fading spatially correlated MIMO channel is given as

$$H = R_t^{1/2} H_w R_r^{1/2}, \quad (2.22)$$

where $R_t$ and $R_r$ are the transmit and receive covariance matrices from (2.17) and (2.18).

In the presence of line-of-sight, the channel is modeled as Ricean fading and the MIMO channel has a non-zero mean corresponding to the specular component. A Ricean fading spatially correlated MIMO channel is given as

$$H = \sqrt{\frac{k}{1+k}} H_{\text{spec}} + \sqrt{\frac{1}{1+k}} R_t^{1/2} H_w R_r^{1/2}, \quad (2.23)$$

where $\sqrt{k/(1+k)} H_{\text{spec}} = \bar{H} = \mathbb{E}\{H\}$ is the specular component of the channel and
$k$ is the $K$-factor as described in section 2.1.3. For a spatially uncorrelated channel
\[ R_0 = R_l^T \otimes R_r = \sigma_h^2 I_{MrMt}. \]

Table 2.2: Statistical MIMO channel models

<table>
<thead>
<tr>
<th>Statistical MIMO channel models: Notation $H_w \sim \mathcal{CN}(0, I_{MrMt})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatially white Rayleigh fading</td>
</tr>
<tr>
<td>$H = \sigma_h H_w$</td>
</tr>
<tr>
<td>Spatially correlated Rayleigh fading</td>
</tr>
<tr>
<td>$H = R_r^{1/2} H_w R_t^{1/2}$</td>
</tr>
<tr>
<td>Spatially white Ricean fading</td>
</tr>
<tr>
<td>$H = \sqrt{\frac{k}{1+k}} H_{spec} + \sqrt{\frac{1}{1+k}} \sigma_h H_w$</td>
</tr>
<tr>
<td>Spatially correlated Ricean fading</td>
</tr>
<tr>
<td>$H = \sqrt{\frac{k}{1+k}} H_{spec} + \sqrt{\frac{1}{1+k}} R_r^{1/2} H_w R_t^{1/2}$</td>
</tr>
</tbody>
</table>

2.2.3 MIMO signal models

In this subsection, we introduce a discrete time signal model for MIMO system in cases of flat-fading and frequency selective fading.

Flat-fading channel

When the bandwidth of the input is considerably less than the channel coherence bandwidth, the channel is referred to as flat fading or frequency non-selective. A single $M_R \times M_T$ matrix channel filter tap, $H$, is sufficient to represent the channel. The received signal vector is given as

\[ y[k] = Hs[k] + n[k] \quad (2.24) \]

where $y[k]$ is the $M_R \times 1$ received signal vector at time $k$, $s[k] = [s_1[k], s_2[k], \ldots, s_{Mt}[k]]^T$ is the $M_T \times 1$ transmitted signal vector at time $k$ and
\( \mathbf{n}[k] \) is the \( M_R \times 1 \) spatio-temporally white ZMCSCG noise vector with variance \( \sigma_n^2 \).

Since the output at any instant of time is independent of inputs at previous times, we can drop the time index \( k \) for clarity and express the input-output relation as

\[
y = \mathbf{Hs} + \mathbf{n}
\]  

(2.25)

**Frequency selective channel**

When the bandwidth of the input is much larger than the channel coherence bandwidth, the channel is said to be frequency selective, and it has to be represented by multiple taps. Each channel tap is represented by a \( M_R \times M_T \) matrix \( \mathbf{H}[l] \) \((l = 0, 1, 1, \ldots, L)\), where \( L \) is the maximum channel length. The received signal vector at time \( k \), \( \mathbf{y}[k] \), is a linear combination of \( \{\mathbf{s}[k], \mathbf{s}[k-1], \ldots, \mathbf{s}[k-L]\} \) and is given as

\[
\mathbf{y}[k] = \sum_{l=0}^{L} \mathbf{H}[l]\mathbf{s}[k-l] + \mathbf{n}[k]
\]  

(2.26)

It is easier to deal with the frequency selective channel by taking a multicarrier approach. By using orthogonal frequency division multiplexing (OFDM) modulation, the frequency selective MIMO channel is transformed into a set of narrow-band, frequency-flat MIMO channels which are easier to equalize.

In a MIMO-OFDM system with \( N \) subcarriers, the individual data streams are first passed through OFDM modulators which perform an IFFT on blocks of length \( N \) followed by a parallel-to-serial conversion. A cyclic prefix (CP) is then prepended and the resulting OFDM symbols are launched simultaneously from the individual
transmit antennas. The CP is essentially a guard interval which serves to eliminate interference between OFDM symbols and turns linear convolution into circular convolution such that the channel is diagonalized by the FFT. At the receiver, the individual signals are passed through OFDM demodulators which first discard the CP and then perform an N-point FFT. The outputs of the OFDM demodulators are finally separated and decoded. A more detailed discussion of the basic principles of MIMO-OFDM can be found in [43]. Choosing the length of the CP to be greater than the length of the channel impulse response guarantees that the frequency-selective MIMO fading channel decouples into a set of parallel frequency-flat MIMO fading channels. Henceforth, we assume a flat-fading channel unless specified otherwise.

2.2.4 Gains of MIMO Channels

MIMO channels have a number of advantages over traditional SISO channels such as the beamforming (or array) gain, the diversity gain, and the multiplexing gain [20]. The beamforming and diversity gains are no exclusive of MIMO channels and also exist in SIMO and MISO channels. The multiplexing gain, however, is a unique characteristic of MIMO channels. Some gains can be simultaneously achieved while others compete and establish a tradeoff as is explained later.

Array gain

Array or beamforming gain is the improvement in SINR that results from a coherent combining effect of wireless signals at the receiver. The coherent combining effect may be realized through spatial processing at the receive antenna array and/or spatial pre-processing at the transmit antenna array. Array gain improves resistance to noise,
thereby improving the coverage and the range of a wireless network. If the BER of a communication system is plotted with respect to the transmitted power or the received power per antenna, the beamforming gain is characterized as a shift of the curve due to the gain in SINR.

**Diversity gain**

Diversity gain is the improvement in link reliability obtained by receiving replicas of the information signal through independently fading links, branches, or dimensions. This type of diversity is clearly related to the random nature of the channel and is closely connected to the specific channel statistics. The basic idea is that with high probability, at least one or more of these links will not be in a fade at any given instant. In other words, the use of multiple dimensions reduces the fluctuations of the received signal and eliminates the deep fades. A MIMO channel with $M_T$ transmit antennas and $M_R$ receive antennas potentially offers $M_T M_R$ independently fading links, and hence a spatial diversity order of $M_T M_R$. If the BER of a communication system is plotted with respect to the transmitted power or the received power per antenna, the diversity gain is characterized as the increase of the slope of the curve in the low BER region.

**Spatial multiplexing gain**

Multiplexing gain is the increase of rate, at no additional power consumption, obtained through the use of multiple dimensions at both sides of the communication link. While the beamforming and the diversity gains can be obtained when multiple dimensions are present at either the transmit or the receive side, multiplexing gain
requires multiple dimensions at both ends of the link. The basic idea is to exploit the multiple dimensions to open up several parallel subchannels within the MIMO channel, also termed channel eigen-modes, which lead to a linear increase in capacity. The multiplexing property allows the transmission of several symbols simultaneously or, in other words, the establishment of several substreams for communication. The number of data streams that can be reliably supported by a MIMO channel equals the minimum of the number of transmit antennas and the number of receive antennas, i.e., \( \min(M_T, M_R) \).

**Tradeoffs between gains**

**Beamforming and Diversity gains**

Beamforming gain is a concept that refers to the combination of multiple copies of the same signal for a specific channel realization regardless of the channel statistics. Diversity gain, however, is directly connected to the statistical behavior of the channel. With multiple receive dimensions, both gains can be simultaneously achieved by a coherent combination of the received signals and there is no tradeoff between them. With multiple transmit dimensions, beamforming gain requires channel knowledge at the transmitter whereas diversity gain can be achieved even when the channel is unknown.

**Beamforming and Multiplexing gains**

Maximum beamforming gain on a MIMO channel implies that only the maximum singular value of the channel should be used. In terms of multiplexing gain, however, the optimum strategy is to use a subset of the channel singular values according to
a water-filling strategy. In other words, maximum beamforming gain requires establishing a single substream for communication, whereas maximum multiplexing gain requires, in general, establishing several simultaneous substreams.

**Diversity and Multiplexing gains**

Traditionally, the design of systems has been focused on either extracting maximum diversity gain or maximum multiplexing gain. Nevertheless, both gains can in fact be simultaneously achieved, but there is a fundamental tradeoff between how much of each type of gain any communication scheme can extract as was shown in [44]. Since the diversity gain is the slope of the BER curve in the high SINR region and the multiplexing gain is related to the achieved rate, the diversity-multiplexing tradeoff is essentially the tradeoff between the error probability and the data rate of a system.

### 2.3 Modeling Channel State Information at the Transmitter

In SISO wireless links, knowledge of the channel at the transmitter is typically used for transmit power control or for adapting the modulation rate. Knowledge of the channel at the transmitter in MIMO wireless systems can be leveraged in additional ways, such as beamforming or precoding, to provide significant performance gains. In this section, we first describe the commonly used techniques for channel acquisition at the transmitter. Next, we review a dynamic model in which channel state information at the transmitter is modeled as a channel estimate and its error covariance.
2.3.1 Channel acquisition at the transmitter

Channel knowledge at the receiver is usually acquired via the transmission of a training sequence that allows the estimation of the channel. Channel acquisition at the transmitter is not directly possible as the signal enters the channel only after leaving the transmitter. The two general techniques for acquiring CSIT are described below [43,45].

Using feedback

In this approach, the channel is measured at the receiver and is sent back to the transmitter on the reverse link. However, an inevitable feedback delay $\delta_t$ exists between channel measurement at the receiver and its use by the transmitter. Reliability of the feedback decreases, unless the feedback delay is much smaller than the channel coherence time: $\delta_t \ll T_c$. Feedback can also be used to send channel statistics that change much slower in time compared to the channel itself. In such cases, the delay requirement for valid feedback can be relaxed significantly.

Using channel reciprocity

The reciprocity principle states that if the time, frequency and antennas for the channel use are the same, then the channels in the forward and reverse links are identical. In real full-duplex communications, however, the forward and reverse links cannot use all identical frequency, time, and space instances. In a time division duplex (TDD) system, the forward and reverse links share the same frequency using different time slots and in a frequency division duplex (FDD) system, the forward and reverse links share the same time slot using different frequencies. The reciprocity principle
may still hold approximately if the difference in any of these dimensions is relatively small, compared to the channel variation across the referenced dimension. In the temporal dimension, this condition implies that any time lag $\delta_t$ between the forward and reverse transmissions must be much smaller than the channel coherence time $T_c$: $\delta_t \ll T_c$. Similarly, any frequency offset $\delta_f$ must be much smaller than the channel coherence bandwidth $B_c$: $\delta_f \ll B_c$, and the antenna location differences on the two links must be much smaller than the channel coherence distance $D_c$.

### 2.3.2 Dynamic CSIT model

Channel knowledge at the transmitter, can be acquired via feedback or by exploiting channel reciprocity. Perfect instantaneous channel knowledge is ideal. However, there exists a delay from when the channel information is measured to when it is used by the transmitter. Because of the temporal channel variation, this time lag affects the reliability of the obtained instantaneous channel measurements.

Channel statistics like the channel mean and covariance (2.19), can be obtained by averaging instantaneous measurements over tens of coherence times. Physical models of wireless channels indicate that these short-term channel statistics remain valid for tens to hundreds coherence intervals, relatively long compared to a transmission interval. Therefore, these statistics are not affected by channel acquisition delay and can be considered reliable.

Exploiting both the instantaneous measurements and the channel statistics, when
available, can provide more gain. A dynamic CSIT model formulated in [7, 46] incorporates both the instantaneous measurements and the channel statistics and is described below.

Suppose that the transmitter has an initial channel measurement $\mathbf{H}_0$ at time 0 along with the channel statistics $\bar{\mathbf{H}}$ and $\mathbf{R}_0$. The main source of irreducible error in channel estimation is the random channel time-variation. Assuming that the channel measurement at time 0 is error-free, the error in the channel estimates depends only on the delay between the initial measurement and its use by the transmitter. The channel at discrete time $n$ can be modeled as

$$\mathbf{H}[n] = \mathbf{H}[n] + \mathbf{E}[n]$$

$$\mathbf{R}_e[n] = \mathbb{E}\{\mathbf{e}[n]\mathbf{e}[n]^T\}, \quad (2.27)$$

where $\mathbf{H}[n]$ is the estimate of the channel at time $n$, $\mathbf{E}[n]$ is the estimation error. $\mathbf{R}_e[n]$ is the error covariance, where $\mathbf{e}[n] = \text{Vec}(\mathbf{E}[n])$.

Assuming unbiased estimates, the estimation error $\mathbf{E}[n]$ can be modeled as a stationary zero-mean complex Gaussian random process. The minimum mean-squared estimate of the channel at time $n$, and its $M_T M_R \times M_T M_R$ error covariance matrix are given as [47]

$$\hat{\mathbf{h}}[n] = \mathbb{E}\{\mathbf{h}[n]/\mathbf{h}[0]\} = \bar{\mathbf{h}} + \mathbf{R}[n]^{\mathbf{H}} \mathbf{R}_0^{-1} (\mathbf{h}[0] - \bar{\mathbf{h}})$$

$$\mathbf{R}_e[n] = \text{Cov}\{\mathbf{h}[n]/\mathbf{h}[0]\} = \mathbf{R}_0 - \mathbf{R}[n]^{\mathbf{H}} \mathbf{R}_0^{-1} \mathbf{R}[n], \quad (2.28)$$
where $\hat{h}[n] = \text{Vec}(\hat{H}[n])$. If we assume homogeneous temporal correlation of the channel (2.21), $R[n] = \rho[n]R_0$, the channel estimate and error covariance matrix can be simplified to

$$
\hat{H}[n] = \rho[n]H_0 + (1 - \rho[n])\bar{H},
$$

$$
R_e[n] = (1 - \rho[n]^2)R_0 = (1 - \rho[n]^2)R_t^T \otimes R_r.
$$

Hence, the CSIT at time $n$ can be modeled as

$$
H[n] = \hat{H}[n] + E[n]
= \rho[n]H_0 + (1 - \rho[n])\bar{H} + \sqrt{(1 - \rho[n]^2)}R_1^{1/2}H_wR_1^{1/2},
$$

(2.30)

where $\rho[n]$ is the temporal correlation of scalar channel. Using Clarks model [42], $\rho[n] = J_0(2\pi f_d n)$, where $J_0(.)$ is the zeroth-order Bessel function of the first kind and $f_d$ is the channel Doppler spread.

In this CSIT model, $\rho[n]$ acts as a channel estimate quality and depends on the time delay $n$. For zero delay, $\rho[n] = 1$, corresponding to perfect CSIT. For a very short delay, $\rho$ is close to 1 and the estimate depends heavily on the initial channel measurement. As the delay increases, $\rho$ decreases in magnitude to zero, reducing the impact of the initial measurement. For example, with a signal frequency $f_c = 1$ GHz and velocity $\nu = 30$ m/sec (approximately 100 Km/hr), the value of the correlation coefficient $\rho$ decreases from 1 to 0 as the value of the time delay $\tau$ increases from 0 to $3.82$ ms. Hence, the instantaneous channel state information becomes obsolete within a few milliseconds for fast moving mobiles. The estimate then moves toward the
channel mean $\bar{H}$, and the error covariance grows toward the channel covariance $R_0$. Therefore, the CSIT ranges between perfect channel knowledge (when $\rho = 1$) and the channel statistics (when $\rho = 0$). In the absence of statistical channel information, we set $\bar{H} = 0$ and $R_t = R_r = \sigma_h I$. 

Chapter 3: MIMO SYSTEM WITH LINEAR PRECODER

Precoding is a signal processing technique at the transmitter that exploits channel state information. Precoding yields a larger capacity, extracts the full diversity and achieves large array gain. Moreover, the overall system complexity may be reduced by using linear precoding/decoding. For a flat-fading channel the optimal precoder is linear. A linear precoder functions as a multimode beamformer that spatially directs data streams along orthogonal directions and allocates power across them based on the available channel knowledge.

If the spatial directions of the data streams matches the channel eigen-directions, there will be no interference among them. Hence, when perfect CSIT is available, precoding decouples the MIMO channel and this allows transmission of non-interfering parallel data streams. With imperfect CSIT, the precoder performs its best to approximately match its spatial directions to the channel eigen-directions, reducing the interference among the date streams.

In this chapter, we first introduce the system model for MIMO systems with precoding. Next, we review the singular value decomposition (SVD) based precoding. With perfect CSIT, SVD based precoder perfectly decouples the channel. Imperfect CSIT results in interference among the transmitted data streams. Finally, we describe
various linear and non-linear receiver architectures for precoded MIMO systems that mitigate the inter-stream interference.

3.1 System Model

Consider the MIMO system with linear precoder, depicted in Figure 3.1. The input bit-stream is coded and modulated to generate a $N \times 1$ symbol vector $x$ whose elements are assumed to be i.i.d Gaussian. The channel dependent precoder linearly processes this vector and outputs a $M_T \times 1$ signal vector $s$ that is launched into the channel using $M_T$ transmit antennas. The received signal is processed by the receive decoder matrix and is finally demodulated. We will not consider coding and modulation design, but rather focus on the design of the precoder and decoder. We will assume that the input symbol stream have been coded and modulated according to well-known schemes.

For a flat-fading channel the input-output relation for the MIMO system shown in Figure 3.1 is given as

$$\hat{x} = Gy = GHFx + Gn,$$

where $F$ is the $M_T \times N$ precoder matrix, $G$ is the $N \times M_R$ receive matrix, $H$ is the
$M_R \times M_T$ channel matrix and

$$\mathbb{E}\{xx^H\} = \sigma_x^2 I; \quad \mathbb{E}\{nn^H\} = \sigma_n^2 I; \quad \mathbb{E}\{xn^H\} = 0.$$  \hspace{1cm} (3.2)

We assume that the MIMO system is limited in the average total transmitted power. Hence the transmit signal vector $s$ must satisfy the following power constraint

$$\text{Tr}(R_s) = \text{Tr}(\mathbb{E}\{ss^H\}) = P.$$  \hspace{1cm} (3.3)

Since, $s = Fx$, the precoder matrix must satisfy

$$\text{Tr}(FF^H) = \frac{P}{\sigma_x^2}.$$  \hspace{1cm} (3.4)

This constraint ensures that the average transmit power from all the transmit antennas per symbol period is constant and is equal to $P$. The specific structure of the precoder $F$ depends on the performance metric to be optimized and on the available CSIT. Some of the common measures of performance of a precoded MIMO communication system are:

**Capacity:**

The capacity is defined as the maximum data rate per unit bandwidth that can be transmitted reliably over the MIMO channel and is given as

$$C = \log_2 \det(I_{M_R} + \gamma H F F^H H^H)$$

$$= \log_2 \det(I_{N} + \gamma F^H H^H HF)$$  \hspace{1cm} (3.5)

where $\gamma = \sigma_x^2/\sigma_n^2$ is the signal-to-noise ratio (SNR).
MSE:
The total average mean square error (MSE) of the system is defined as

\[ \text{MSE} = \mathbb{E}\{\|\hat{x} - x\|^2\} \]  

(3.6)

The smaller the MSE, the better the system.

SINR:
The signal to interference ratio (SINR) experienced by the \( i^{th} \) datastream is given as

\[ \text{SINR}_i = \frac{|(GHF)_{ii}|^2\sigma_x^2}{\sum_{j \neq i} |(GHF)_{ij}|^2\sigma_x^2 + \sum_k |(G)_{ik}|^2\sigma_n^2} \]  

(3.7)

The precoder and decoder matrices are to be designed to reduce the interference and improve the SINR.

3.2 SVD based Linear Precoding

The performance of a MIMO system can be improved if the channel is known at the transmitter. In this section we describe a simple SVD based precoding and decoding scheme which decomposes the MIMO channel into parallel non-interfering SISO subchannels. The transmit power can then be distributed optimally across these spatial subchannels to maximize spectral efficiency. First we briefly review the singular and eigen-value decomposition of matrices [43, 48]. The \( M_R \times M_T \) channel matrix \( \mathbf{H} \) has a singular value decomposition (SVD)

\[ \mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H, \]  

(3.8)
where $\mathbf{U}$ and $\mathbf{V}$ are the $M_R \times r$ and $M_T \times r$ matrices respectively, and satisfy 
$\mathbf{U}^H \mathbf{U} = \mathbf{I}_r$, $\mathbf{V}^H \mathbf{V} = \mathbf{I}_r$ and the $r \times r$ matrix $\mathbf{\Sigma} = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_r\}$ with $r = \min(M_T, M_R)$, $\sigma_i \geq 0$ and $\sigma_i \geq \sigma_{i+1}$, where $\sigma_i$ is the $i^{th}$ singular value. The columns of $\mathbf{U}$ and $\mathbf{V}$ are also known as the left and right singular vectors, respectively.

The eigenvalue decomposition (EVD) of the $M_R \times M_R$ positive semi-definite Hermitian matrix $\mathbf{H}^H \mathbf{H}$ is given as 
$$\mathbf{H}^H \mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H,$$  
where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_r\}$ with $\lambda_i = \sigma_i^2$ for $i = 1, 2, \ldots, r$. We assume that the eigenvalues $\lambda_i$ are ordered so that $\lambda_i \geq \lambda_{i+1}$.

A linear precoder functions as an input shaper and a multiple beamformer with power allocation across the orthogonal beams [46]. Consider the SVD of the precoder matrix 
$$\mathbf{F} = \mathbf{U}_F \mathbf{\Phi} \mathbf{V}_F^H.$$  
$\mathbf{\Phi} = \text{diag}\{\phi_1, \phi_2, \ldots, \phi_r\}$ is the $r \times r$ power allocation matrix. The squared singular values, $\phi_i^2$, correspond to the power allocated to the $i^{th}$ beam. The left singular vectors $\mathbf{U}_F$ are the orthogonal beam directions. The right singular vectors $\mathbf{V}_F$, act as an input-shaping matrix. The optimal input shaping matrix is given by the eigenvectors of the input signal covariance matrix. The precoder beam directions and the power allocation across them depend on the available CSIT and the design criterion.

Consider a ZMCSCG signal vector $\mathbf{x}$ of dimension $r \times 1$, where $r$ is the rank of the
channel $\mathbf{H}$, to be transmitted. Prior to transmission, the symbol vector $\mathbf{x}$ is multiplied by the precoder matrix $\mathbf{F}$ (3.10). Since we assume an uncorrelated input signal vector, $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \sigma_x^2 \mathbf{I}$, we set $\mathbf{V}_F = \mathbf{I}$. When perfect channel knowledge is available at the transmitter, we set the precoder direction matrix $\mathbf{U}_F = \mathbf{V}$. At the receiver, the $M_R \times 1$ received signal vector $\mathbf{y}$ is multiplied by the $r \times M_R$ matrix $\mathbf{U}^H$.

Hence, with perfect CSIT, the precoder matrix $\mathbf{F} = \mathbf{V}\Phi$ and the receive matrix $\mathbf{G} = \mathbf{U}^H$. For this SVD-based linear precoding system, the effective input-output relation in (6.2) becomes

$$
\hat{\mathbf{x}} = \mathbf{U}^H\mathbf{H}\Phi\mathbf{x} + \mathbf{U}^H\mathbf{n}
$$

$$
= \Sigma\Phi\mathbf{x} + \tilde{\mathbf{n}}
$$

(3.11)

where $\tilde{\mathbf{n}}$ is the ZMCSCG $M_R \times 1$ transformed noise vector with covariance matrix $\mathbb{E}\{\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H\} = \sigma_n^2 \mathbf{I}_{M_R}$. Equation (3.11) shows that with perfect channel knowledge at the transmitter, the $M_R \times M_T$ channel matrix $\mathbf{H}$ can be explicitly decomposed into $r = \min(M_T, M_R)$ parallel non-interfering SISO channels as shown in figure 3.2, each satisfying

$$
\hat{x}_i = \phi_i \sqrt{\lambda_i} x_i + \tilde{n}_i, \quad i = 1, 2, \ldots, r.
$$

(3.12)

The signal to noise ratio (SNR) of the $i^{th}$ subchannel is given as $\text{SNR}_i = \gamma \phi_i^2 \lambda_i$, where $\phi_i^2$ is the transmit power allocated to the $i^{th}$ subchannel and $\gamma = \sigma_x^2 / \sigma_n^2$. The capacity of the MIMO channel is the sum of the individual parallel SISO channel...
Figure 3.2: Equivalent decomposed channel

capacities and is given by

\[ C = \sum_{i=1}^{r} \log_2 (1 + \gamma \phi_i^2 \lambda_i). \quad (3.13) \]

The constraint on the total transmitted power in (3.4) reduces to \( \sum_{i=1}^{r} \phi_i^2 = P/\sigma_x^2 \).

Since the transmitter can access the spatial subchannels, it can allocate variable power across the subchannels to maximize the mutual information [43, 49]. The mutual information maximization problem now becomes

\[ \max_{\phi_i} C = \sum_{i=1}^{r} \log_2 (1 + \gamma \phi_i^2 \lambda_i), \]

\[ \text{s.t.} \quad \sum_{i=1}^{r} \phi_i^2 = P/\sigma_x^2. \quad (3.14) \]

The objective function for the maximization is concave in the variables \( \phi_i^2 \) and can
be solved using Lagrangian methods. The optimal power allocation is given as

$$\phi_i^2 = \left( \mu - \frac{1}{\gamma \lambda_i} \right)^+, \quad i = 1, 2, \ldots, r,$$

where

$$\mu = \frac{1}{r\sigma_x^2} \left[ P + \sigma_n^2 \sum_{i=1}^{r} \lambda_i \right]$$

(3.16)

and \((x)^+\) implies \((x)^+ = x\) if \(x \geq 0\) and \((x)^+ = 0\) if \(x < 0\).

The capacity of the MIMO channel when the channel is known to the transmitter is greater that the capacity when the channel is unknown to the transmitter.

### 3.3 MIMO Receiver Architectures

With perfect channel state information, the \(M_R \times M_T\) MIMO channel can be decomposed into \(r = \min(M_T, M_R)\) parallel spatial subchannels using SVD as described in section 3.2. This enables the transmitter to sent independent data streams through the channel (spatial multiplexing) so that they arrive orthogonally at the receiver without interference between the streams. However perfect CSIT is never available in practice because of the time varying nature of the channel and due to unavoidable estimation errors.

When imperfect or partial CSIT is available, it is not possible to perfectly decouple the channel and the independent data streams sent on the transmit antennas all arrive cross-coupled at the receiver. We call this interference among the subchannels, inter-stream interference. This section describes the linear and non-linear receiver
architectures that perform spatial equalization so as to reduce or eliminate the inter-stream interference. As described in section 3.1, for a flat-fading channel the system equation is given as

\[ \hat{x} = Gy = GHFx + Gn, \]  

(3.17)

where \( F \) is the \( M_T \times N \) precoder matrix, \( G \) is the \( N \times M_R \) receive matrix and \( H \) is the \( M_R \times M_T \) channel matrix. For the receiver, the effective channel matrix will be \( HF \) and the receiver is assumed to estimate \( HF \) perfectly. Given a precoder matrix \( F \), we next describe various receiver matrix architectures.

### 3.3.1 Linear decorrelator receiver

Imposing a zero forcing (ZF) constraint \( GHF = I \), the decorrelator or zero forcing receiver which simply inverts the effective channel is given by

\[ G_{\text{decorr}} = (HF)^\dagger = (F^H H^H HF)^{-1} F^H H^H, \]  

(3.18)

The output of the decorrelator receiver is given by

\[ \hat{x} = x + (F^H H^H HF)^{-1} F^H H^H n, \]  

(3.19)

where we assume that \( M_R \geq M_T \) and \( H \) has full column rank. The decorrelator receiver decouples the channel matrix into \( M_T \) parallel subchannels with additive noise. Clearly, the noise is enhanced by the decorrelator receiver. Furthermore, the noise is correlated across the subchannels.
3.3.2 Linear MMSE receiver

The decorrelator receiver completely eliminates the inter-stream interference at the expense of noise enhancement. The minimum mean squared error (MMSE) receiver balances the inter-stream interference mitigation with noise enhancement and minimizes the mean squared error which is defined as

\[
\text{MSE} = E\{\|\hat{x} - x\|^2\} = \text{Tr}[\text{MSE}(F, G)],
\]

where the MSE matrix \(\text{MSE}(F, G) = E\{\langle \hat{x} - x \rangle \langle \hat{x} - x \rangle^H\}\). The MMSE receiver matrix \(G_{\text{mmse}}\) is given as

\[
G_{\text{mmse}} = \arg \min \mathbb{E}\{\|\hat{x} - x\|^2\}.
\]

By solving the above minimization problem for a fixed \(F\), we obtain the well known MMSE (Wiener) receiver

\[
G_{\text{mmse}} = R_x F^H H^H \left[ R_n + H F R_x F^H H^H \right]^{-1}
\]

\[
= \left[ \frac{1}{\gamma} I_{M_T} + F^H H^H H F \right]^{-1} F^H H^H
\]

where, \(\gamma = \frac{\sigma^2}{\sigma_n^2}\) since, \(R_x = \sigma_x^2 I_{M_T}\) and \(R_n = \sigma_n^2 I_{M_T}\). In the second equality, we used the following matrix identity

\[
A \left[ B A + R \right]^{-1} = \left[ A R^{-1} B + I \right]^{-1} A R^{-1}
\]

At high signal to noise ratio, \(\frac{1}{\gamma} \to 0\) and the MMSE receiver acts like a decorrelator receiver.
For compactness we denote the linear decorrelator and linear MMSE receivers by the following matrix

\[ G_{\text{linear}} = \left[ \alpha \mathbf{I}_{M_t} + \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right]^{-1} \mathbf{F}^H \mathbf{H}^H \]  

(3.23)

where \( \alpha = 0 \) in case of a decorrelator receiver and \( \alpha = \gamma^{-1} \) in case of an MMSE receiver.

### 3.3.3 Successive interference cancelation

For successive interference cancelation (SIC), the individual signal dimensions are detected sequentially and then canceled from the received signal to form a residual signal. Linear detection is then applied to the residual signal using a reduced signal model that contains only the undetected signals. This process is repeated until all signal components are detected. For a single vector symbol, the algorithm is summarized below [43,50].

**Notation:** Say \( \mathbf{A} \) is an \( n \times n \) matrix. The notation \( \mathbf{A}_{[\cdot:i]} \) denotes an \( n \times (n - i) \) matrix obtained by removing the first \( i \) columns of the matrix \( \mathbf{A} \). The notation \( \mathbf{A}[:,i] \) denotes the \( i^{th} \) column of \( \mathbf{A} \). Say \( \mathbf{b} \) is an \( n \times 1 \) vector. The notation \( \mathbf{b}_{[\cdot:i]} \) denotes an \( (n - i) \times 1 \) vector obtained by removing the first \( i \) elements of the vector \( \mathbf{b} \).

**Successive interference cancellation algorithm**

**Initialization:**

Initial signal model:

\[ \mathbf{y}_1 = \mathbf{H}\mathbf{F}\mathbf{x} + \mathbf{n} \]  

(3.24)
\[ G_1 = \left[ \alpha I + F^H H^H H F \right]^{-1} F^H H^H \]  \hspace{1cm} (3.25)

\[ i = 1 \]

**Recursion:**

**Step 1:** Extract the symbol from the \( i \)th stream,

\[ \hat{x}_i = g_{H}^i \hat{y}_i, \]  \hspace{1cm} (3.26)

where \( g_{i1} \) is the first column of \( G_i^H \). Slice \( \hat{x}_i \) to decode \( x_i \).

**Step 2:** Assume that the decision on \( x_i \) is correct, remodulate to get \( x_i \) and subtract its contribution from the received signal vector \( \hat{y}_i \). The reduced signal model is

\[ \hat{y}_{i+1} = \hat{y}_i - (HF)_{[\cdot,i]} \hat{x}_i \]

\[ = (HF)_{[-i]} x_{[-i]} + n. \]  \hspace{1cm} (3.27)

Obtain the \((M_T - i) \times M_R\) receiver matrix for the reduced signal model as,

\[ G_{i+1} = \left[ \alpha I_{(M_T - i)} + (HF)_{[-i]} (HF)_{[-i]} \right]^{-1} (HF)_{[-i]}^H. \]  \hspace{1cm} (3.28)

**Step 3:** Set \( i = i + 1 \). Return to step 1, extract the next stream and repeat until the vector symbol is decoded.

When \( \alpha = 0 \), the above algorithm is referred to as SVD+SIC-decorrelator and it
is referred to as SVD+SIC-MMSE receiver when $\alpha = \gamma^{-1}$.

In the next chapter, we derive explicit expressions for the signal to interference-plus-noise ratio and the mean squared error of the precoded MIMO system with various receiver architectures. We analyze their performance under various degrees of CSIT.
Chapter 4: MIMO Receivers under Limited CSIT

With perfect channel state information, the MIMO channel can be decomposed into parallel spatial subchannels. However, with limited channel knowledge at the transmitter, it is not possible to decouple the channel perfectly and the independent data streams sent on the transmit antennas all arrive cross-coupled at the receiver. In the previous chapter, we described the linear and non-linear receiver architectures that perform spatial equalization so as to reduce or eliminate the inter-stream interference.

In this chapter, we first revisit the dynamic CSIT model in Chapter 2, and discuss various CSIT scenarios. Next, we describe simple sub-optimal precoding solutions under various degrees of channel knowledge at the transmitter. Finally, we present our performance analysis of the decorrelator, minimum mean squared error (MMSE) and successive interference cancelation receivers as the quantity and the quality of the available CSIT varies. Explicit expressions for the signal to interference-plus-noise ratio (SINR) and the mean squared error are derived. Simulation results are provided to illustrate the significant performance gain achieved by precoding even with a moderate amount of correlation between the available outdated channel estimate and the current channel. However, the performance gain achieved by precoding decreases as the reliability of the available CSIT decreases.
4.1 Limited Channel State Information at the Transmitter

As explained in section 2.3, channel knowledge at the transmitter, can be acquired via feedback or by exploiting channel reciprocity. However, in practice, perfect channel knowledge at the transmitter is rarely available because of the time varying nature of the channel and due to unavoidable estimation errors. As a result, the available channel state information at the transmitter (CSIT) is usually outdated and/or statistical in nature. At time $t$, the outdated instantaneous CSIT with a time lag $\tau$ is denoted as $H_{t-\tau}$. The channel mean $\bar{H}$, the transmit covariance matrix $R_t$ and the receive covariance matrix $R_r$ constitute the statistical CSIT. The channel temporal correlation coefficient, $0 \leq \rho \leq 1$, acts as a channel estimate quality indicator. It is a function of the delay $\tau$ and channel’s time variability. According to Jake’s model [42], for example, $\rho = J_0(2\pi f_d\tau)$, where $J_0$ is the zeroth order Bessel function of the first kind and $f_d$ is the Doppler shift. Next, we discuss various CSIT scenarios that arise depending on the degree of available limited channel knowledge.

**Instantaneous CSIT**

This case arises when only instantaneous channel measurements are fed back from the receiver, and the transmitter has no knowledge of the channel statistics (i.e.,) $\bar{H} = 0$, $R_r = R_t = I$ and $0 < \rho \leq 1$. The channel at time $t$ can be modeled as

$$H = \rho H_{t-\tau} + \sqrt{(1 - \rho^2)}H_w.$$  \hspace{1cm} (4.1)

The elements of $H_w$ are independent and identically distributed zero-mean circular symmetric complex Gaussian (ZMCSCG) random variables.
Covariance CSIT

This case arises when the instantaneous CSIT is unreliable ($\rho = 0$) and when channel covariances alone are known at the transmitter. The channel mean $\bar{H} = 0$ and the channel at time $t$ can be modeled as

$$H = R_{t}^{1/2}H_{w}R_{t}^{1/2}. \quad (4.2)$$

Mean CSIT

This case arises when $\rho = 0$ and when channel mean alone is known at the transmitter.

In this case, $R_{r} = R_{t} = I$ and the channel at time $t$ can be modeled as

$$H = \bar{H} + H_{w}. \quad (4.3)$$

General CSIT

This case arises when both the instantaneous and statistical CSIT is available. In this case, the channel at time $t$ can be modeled as

$$H = H_{m} + \sqrt{(1 - \rho^{2})}R_{t}^{1/2}H_{w}R_{t}^{1/2}, \quad (4.4)$$

where, the effective channel mean $H_{m} = \rho H_{t-r} + (1 - \rho)\bar{H}$.

Next, we describe SVD based sub-optimal precoding solutions for the above described limited CSIT scenarios.

4.2 Precoding with Limited CSIT

When the channel $H$ is perfectly known at the transmitter, the precoder rotates the data streams such that they are sent along the eigen-modes of the MIMO channel
matrix thereby decomposing it into non-interfering parallel subchannels. Therefore, given $\mathbf{H}^H \mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$, the precoder $\mathbf{F} = \mathbf{V} \mathbf{\Phi}$ is optimal. Under a transmit power constraint, the elements of $\mathbf{\Phi}$ are obtained by water-filling across the subchannels.

Under limited CSIT, the channel eigen-directions are not perfectly known. Hence, the precoder must perform its best to decouple the channel by approximately matching its spatial directions to the channel eigen-directions. This can be achieved by precoding on the average channel gain $\mathbb{E}[\mathbf{H}^H \mathbf{H}] = \hat{\mathbf{V}} \hat{\mathbf{\Lambda}} \hat{\mathbf{V}}^H$.

For the instantaneous CSIT case (4.1),

$$\mathbb{E}[\mathbf{H}^H \mathbf{H}] = M_R (1 - \rho^2) \mathbf{I} + \rho^2 \mathbf{H}_{t-\tau}^H \mathbf{H}_{t-\tau}. \quad (4.5)$$

Hence, the optimal precoding beam directions are the eigenvectors of $\mathbf{H}_{t-\tau}^H \mathbf{H}_{t-\tau}$. When the time lag $\tau = 0$, this solution is equivalent to the SVD precoding described in 3.2.

For the covariance CSIT case (4.2),

$$\mathbb{E}[\mathbf{H}^H \mathbf{H}] = \text{tr}(\mathbf{R}_r) \mathbf{R}_t. \quad (4.6)$$

Hence, the optimal precoding beam directions are the eigenvectors of $\mathbf{R}_t$.

For the mean CSIT case (4.3),

$$\mathbb{E}[\mathbf{H}^H \mathbf{H}] = M_R \mathbf{I} + \bar{\mathbf{H}}^H \bar{\mathbf{H}}. \quad (4.7)$$
Hence, the optimal precoding beam directions are the eigenvectors of $\bar{H}^H\bar{H}$.

For the general CSIT case (4.4),

$$E[H^H H] = (1 - \rho^2)\text{tr}(R_r)R_t + H_m^H H_m,$$

(4.8)

where, the effective channel mean $H_m = \rho H_{t-x} + (1 - \rho)\bar{H}$. For this case, the precoding beam directions are the eigenvectors of $(1 - \rho^2)\text{tr}(R_r)R_t + H_m^H H_m$.

Primary objective of this chapter is to illustrate the effect of precoding on limited CSIT and to analyze various receiver architectures that mitigate the resulting effect. In order to simplify further exposition, we consider precoding with equal power allocation ($\Phi = I$). The precoder matrix based on the average channel gain $E[H^H H] = \hat{V}\hat{\Lambda}\hat{V}^H$ is given as

$$F = \hat{V}.$$  

(4.9)

The eigen-value decomposition of the true channel is $H^H H = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$. Due to the mismatch of $\mathbf{V}$ and $\hat{\mathbf{V}}$, the channel is not perfectly decoupled, and the independent data streams sent on the transmit antennas all arrive cross-coupled at the receiver resulting in interference. In the previous chapter, we described linear and non-linear receiver architectures that perform spatial equalization so as to reduce or eliminate the inter-stream interference. Next, we derive MSE, SINR and capacity expressions for those receivers, and analyze their performance under various CSIT scenarios.
4.3 Performance Analysis of Linear Receivers

The linear decorrelator and linear MMSE receivers are given by

\[ G_{\text{linear}} = \left[ \alpha I_{MT} + F^H H^H F \right]^{-1} F^H H^H. \]  \tag{4.10}

When, \( \alpha = 0 \), the above expression represents a decorrelator receiver. When \( \alpha = \gamma^{-1} \), it represents an MMSE receiver. With a linear receiver, the input-output relation of the MIMO system in (3.17) becomes

\[ \hat{x} = G H F x + G n \]

\[ = \left[ \alpha I_{MT} + F^H H^H H F \right]^{-1} F^H H^H H F x + \tilde{n}, \tag{4.11} \]

where \( \tilde{n} = \left[ \alpha I_{MT} + F^H H^H H F \right]^{-1} F^H H^H n \) is the effective noise vector. By replacing \( H^H H \) with its eigenvalue decomposition \( V \Lambda V^H \), and on setting the precoder matrix \( F = \hat{V} \), the system equation (4.11) can be written as

\[ \hat{x} = \left[ \alpha I_{MT} + Q^H \Lambda Q \right]^{-1} Q^H \Lambda Q x + \tilde{n} \]

\[ = Q^H \left[ \alpha I_{MT} + \Lambda \right]^{-1} \Lambda Q x + \tilde{n}, \tag{4.12} \]

where, the \( M_T \times M_T \) matrix \( Q = (q_{ij}) = V^H \hat{V} \). We note that \( Q \) is a random unitary matrix whose statistics depend on the error between the true channel and the available CSIT. As the realibility of the available CSIT increases, the off-diagonal elements of the matrix \( Q \) decrease in magnitude. With perfect CSIT (\( \rho = 1 \)), all the off-diagonal elements vanish and \( Q \) becomes an identity matrix. The second equality in (4.12) is valid only under an equal power allocation assumption. From (4.12), the
input-output relation for the $i^{th}$ subchannel, $i = 1, \ldots, M_T$, is given as

$$
\hat{x}_i = \sum_k q_{ki}^2 \frac{\lambda_k}{\alpha + \lambda_k} x_i + \sum_{j \neq i} \sum_k q_{kj} q_{ki} \frac{\lambda_k}{\alpha + \lambda_k} x_j + \tilde{n}_i.
$$

(4.13)

Where $q_{ki}$ are the elements of the matrix $Q$ and $\lambda_k$ are the eigenvalues of the channel matrix. $x_i$ is the $i^{th}$ element of the input vector $x$, $\tilde{n}_i$ is the $i^{th}$ element of the effective noise vector $\tilde{n}$, $j = 1, \ldots, M_T$ and $k = 1, \ldots, M_T$. In (4.13), the first term is the signal component, the second term is the interference component and the third is the noise component.

### 4.3.1 Signal to interference-plus-noise ratio

We obtain the SINR of the $i^{th}$ subchannel as

$$
\text{SINR}_i = \frac{\left( \sum_k \mathbb{E}[q_{ki}^2] \frac{\lambda_k}{\alpha + \lambda_k} \right)^2}{\sum_{j \neq i} \left( \sum_k \mathbb{E}[q_{kj} q_{ki}] \frac{\lambda_k}{\alpha + \lambda_k} \right)^2 + \gamma^{-1} \left( \sum_k \mathbb{E}[q_{ki}^2] \frac{\lambda_k}{(\alpha + \lambda_k)^2} \right)}.
$$

(4.14)

In the case of perfect CSIT, $V = \hat{V}$ and $Q = I$. Hence $q_{ij} = 1$ for $i = j$, $q_{ij} = 0$ for $i \neq j$ and (4.14) reduces to

$$
\text{SINR}_i = \frac{\left( \frac{\lambda_i}{\alpha + \lambda_i} \right)^2}{\gamma^{-1} \frac{\lambda_i}{(\alpha + \lambda_i)^2}} = \gamma \lambda_i,
$$

(4.15)

where $\gamma = \sigma_x^2 / \sigma_n^2$. 
4.3.2 Mean squared error

With a linear receiver, the mean squared error matrix of the MIMO system is given as [14,15]

\[
\text{MSE} = \mathbb{E}[(\hat{x} - x)(\hat{x} - x)^H] = \sigma_n^2\mathbb{E}[(\alpha\mathbf{I} + \mathbf{F}^H\mathbf{H}^H\mathbf{F})^{-1}].
\] (4.16)

Assuming uniform power allocation among the subchannels and using the precoder matrix based on the available imperfect CSIT i.e \(\mathbf{F} = \hat{\mathbf{V}}\), the MSE matrix becomes

\[
\text{MSE} = \sigma_n^2\mathbb{E}[(\alpha\mathbf{I} + \mathbf{Q}^H\Lambda\mathbf{Q})^{-1}]
= \sigma_n^2\mathbb{E}\{\mathbf{Q}^H[(\alpha\mathbf{I} + \Lambda)^{-1}\mathbf{Q}]\}.
\] (4.17)

The MSE experienced by the \(i^{th}\) subchannel is the \(i^{th}\) diagonal element of the MSE matrix \(\text{MSE}\) and is given as

\[
\text{MSE}_i = \sum_k \mathbb{E}[q_{ki}^2]\frac{\sigma_n^2}{\alpha + \lambda_k}.
\] (4.18)

where the eigenvalues \(\lambda_i\) are the elements of the diagonal matrix \(\Lambda\) and \(q_{ij}\) are the elements of the unitary matrix \(\mathbf{Q}\).

In case of perfect CSIT, \(\mathbf{Q} = \mathbf{I}\) and the MSE experienced by the \(i^{th}\) subchannel reduces to

\[
\text{MSE}_i = \frac{\sigma_n^2}{\alpha + \lambda_i}.
\] (4.19)
4.4 Performance Analysis of Successive Interference Cancelation

The SIC algorithm has been described in section 3.3.3. From (3.28), at the $i^{th}$ iteration, the receiver matrix for the corresponding signal model is given as

$$G_i = Q_{i-(i-1)}^H [\alpha I_{M_F-(i-1)} + \Lambda]^{-1} \Lambda^{1/2} U^H$$  \hspace{1cm} (4.20)

The first column of $G_i^H$ is $g_{i1} = U \Lambda^{1/2} [\alpha I + \Lambda]^{-1} q_i$, where $q_i$ is the $i^{th}$ column of the matrix $Q$.

4.4.1 Signal to interference-plus-noise ratio

From (3.26), the input-output relation for the $i^{th}$ subchannel is

$$\hat{x}_i = g_{i1}^H y_i$$

$$= \sum_k q_{ki}^2 \frac{\lambda_k}{\alpha + \lambda_k} x_i + \sum_{j>i} \sum_k q_{ki} q_{kj} \frac{\lambda_k}{\alpha + \lambda_k} x_j + \tilde{n}_i.$$  \hspace{1cm} (4.21)

Where, $\tilde{n}_i = q_i^H [\alpha I + \Lambda]^{-1} \Lambda^{1/2} U n$ is the effective noise experienced by the $i^{th}$ subchannel. On comparing (4.21) with (4.13), we see that SIC removes the interference from the subchannels that were already detected.

From (4.21) we obtain the SINR of the $i^{th}$ subchannel as

$$\text{SINR}_i = \frac{\left( \sum_k \mathbb{E}[q_{ki}^2] \frac{\lambda_k}{\alpha + \lambda_k} \right)^2}{\sum_{j>i} \left( \sum_k \mathbb{E}[q_{ki} q_{kj}] \frac{\lambda_k}{\alpha + \lambda_k} \right)^2 + \gamma^{-1} \left( \sum_k \mathbb{E}[q_{ki}^2] \frac{\lambda_k}{(\alpha + \lambda_k)^2} \right)^2}.$$  \hspace{1cm} (4.22)

With perfect CSIT, there is no interference among the subchannels and (4.22) reduces
to (4.15).

### 4.4.2 Mean squared error

Using the input-output relation in (4.21), the mean squared error of the $i^{th}$ subchannel is given as

\[
\text{MSE}_i = \mathbb{E}[(\hat{x}_i - x_i)^2] = \sigma_x^2 \left( \sum_k \mathbb{E}[q_{ki}^2] \frac{\lambda_k}{\alpha + \lambda_k} - 1 \right)^2 + \sum_{j>i} \left( \sum_k \mathbb{E}[q_{ki}^2 q_{kj}^2] \frac{\lambda_k}{\alpha + \lambda_k} \right)^2 + \sum_k \mathbb{E}[q_{ki}^2] \alpha \left( \alpha + \lambda_k \right)^2.
\]

On simplification, (4.23), reduces to

\[
\text{MSE}_i = \sum_k \mathbb{E}[q_{ki}^2] \frac{\sigma_n^2}{\alpha + \lambda_k} - \sigma_x^2 \sum_{j<i} \left( \sum_k \mathbb{E}[q_{ki} q_{kj}] \frac{\lambda_k}{\alpha + \lambda_k} \right)^2.
\]

Thus, with SIC the MSE on each subchannel equals the MSE with a linear receiver minus the interference power of the prior detected subchannels. With perfect CSIT, (4.24) reduces to (4.19).

### 4.5 System Metrics

Previous subsections focused on the MSE and SINR of the individual substreams. To assess the overall performance of our MIMO systems, we rely on capacity and total MSE. The capacity of the $M_R \times M_T$ MIMO system, $M_R \geq M_T$, is given as [41]

\[
C = \sum_{i=1}^{M_T} \log_2(1 + \text{SINR}_i).
\]
With a linear receiver, SINR\(_i\) is given by (4.14) and with a SIC receiver it is given by (4.22). The total mean square error of the MIMO system is given as

\[
\text{MSE} = \sum_{i=1}^{M_T} \text{MSE}_i.
\] (4.26)

With a linear receiver, MSE\(_i\) is given by (4.18) and with a SIC receiver it is given by (4.24).

### 4.6 Simulation Results

The closed form expressions for SINR and MSE of each subchannel have been obtained in terms of the elements of a random unitary matrix \(Q = (q_{ij}) = V^H \hat{V}\). Since we cannot evaluate the statistics of \(q_{ij}\), we resort to simulation. For simulation purposes, we consider a 4 \(\times\) 4 MIMO system and we analyze the system performance with decorrelator, MMSE and MMSE-SIC receivers. All the plots are averaged over 15,000 channel realizations.

For the precoding and non-precoding systems, the total average capacity as a function of \(\rho\) is plotted in Figures 4.1, 4.2 and 4.3. The mean and transmit covariance matrix used in the simulations are listed in Appendix A.1. Statistical CSIT involves channel mean and transmit covariance. General CSIT involves both the statistical CSIT and instantaneous CSIT. When \(\rho = 0\), CSIT is purely statistical in nature. Precoding based on statistical CSIT alone achieves substantial gains compared to no precoding. For values of \(\rho > 0.4\), exploiting both the instantaneous and statistical channel knowledge provides considerable further improvement in performance. For
Figure 4.1: Capacity of MIMO channel with a non-zero mean and transmit antenna correlation under limited CSIT, SNR=10 dB.
\( \rho < 0.4 \), the instantaneous knowledge becomes unreliable and precoding must be based on channel statistics.

With perfect CSIT, the channel is completely decoupled and there is no interference among the established subchannels. Hence, when \( \rho = 1 \), the performance with all the receivers is identical (4.15). The interference among the subchannels increases as the value of \( \rho \) decreases. The MMSE receiver performs better than the decorrelator receiver because it balances interference mitigation with noise enhancement. SIC further improves the performance of the MMSE receiver.

Figure 4.4 shows the system performance using linear MMSE receiver under various degrees of CSIT. When \( \rho = 0 \), the gain achieved is due to statistical CSIT. The precoding gain increases as the value of \( \rho \) increases. However, instantaneous channel knowledge helps to increase the precoding gain over the statistical CSIT gain only when its sufficiently reliable (\( \rho > 0.4 \)). When \( \rho = 1 \), corresponding to perfect CSIT, maximum precoding gain is achieved.

With a linear receiver, the total mean square error (4.26) is equal to the trace of the MSE matrix given by (4.17). In (4.17), \( Q = V^H \hat{V} \) under limited CSIT, \( Q = V^H \) for non-precoding systems and \( Q = I \) when perfect CSIT is available. Note that in all three cases, \( Q \) is an unitary matrix. Pre and post multiplication by unitary matrices does not affect the trace. Hence, under an uniform power allocation assumption, precoding does not affect the total mean squared error of the system. Even though the sum of the MSEs of all the subchannels (total MSE) remains constant for all values of
Figure 4.2: Capacity of MIMO channel with a non-zero mean and transmit antenna correlation under limited CSIT, SNR=20 dB.
Figure 4.3: Capacity of MIMO channel with zero mean and uncorrelated antennas under instantaneous CSIT, SNR=10 dB.

\[ \rho \], the individual MSE of each subchannel, given by (4.18) and (4.24), varies depending on the interference it experiences. The variation of the MSE of each subchannel as a function of \( \rho \) is shown in Figures 4.5 and 4.6. When the channel is decoupled into independent subchannels, power allocation among the subchannels using the water filling algorithm [43] can be used to decrease the total MSE of the system. When \( \rho = 1 \), there is no interference among the subchannels and the MSE and MSE-SIC receivers perform identically. Under instantaneous CSIT, as the value of \( \rho \) decreases, performance of the precoding systems approaches that of non-precoding systems.
Figure 4.4: Capacity of MIMO channel with a non-zero mean and transmit antenna correlation using MMSE receiver.

Figure 4.5: MSE of the subchannels with instantaneous CSIT, SNR=10 dB.
Figure 4.6: MSE of the subchannels with instantaneous CSIT, SNR=10 dB.

4.7 Chapter Summary

In this chapter, we derived SINR and MSE expressions for decorrelator, MMSE and SIC receivers. The system performance when the data streams are steered along the eigen-modes of the channel through precoding has been analyzed under various degrees of CSIT. Precoding based on imperfect CSIT results in interference among the established subchannels. Due to its interference mitigation capabilities without noise amplification, the SIC receiver exhibits superior performance compared to linear receivers. Exploiting the available instantaneous/statistical CSIT provides substantial capacity gains. Under uniform power allocation, precoding does not effect the total MSE of the system. When the channel is perfectly decoupled, power allocation by water filling can be used to improve capacity or to decrease the total MSE. In latter chapters, we address robust power allocation when the channel is not perfectly...
decoupled.
A linear precoder functions as a multimode beamformer that spatially directs the signal in orthogonal directions and allocates power based on the available channel knowledge and the design criterion. In the previous chapter, we showed that when the MIMO channel $H$ is perfectly known at the transmitter, the precoder rotates the data streams such they are sent along the eigen-modes of the channel matrix thereby decomposing it into non-interfering parallel subchannels. Under limited CSIT, the channel eigen-directions are not perfectly known, and the precoder must perform its best to decouple the channel by exploiting the available channel knowledge. The power allocated across the beams and the precoder beam directions are dictated by the available CSIT and by the design criteria. In this chapter, we consider channel state information in the form of transmit antenna correlation $R_t$ and we address precoder design with the following criteria: (1) maximizing the system ergodic capacity, (2) maximizing the capacity of a system with a MMSE receiver.

Evaluating the ergodic capacity involves evaluating an expectation over a Wishart distributed matrix and the expectation has no closed form solution. Hence, the precoder design under this criterion is carried out by maximizing approximations of the ergodic capacity [28, 29, 45] and the achieved precoding gain depends on the accuracy of the approximation. To the best of our knowledge, precoder design to maximize the
capacity of a system with a MMSE receiver under limited CSIT has never been addressed in the literature. Evaluating the capacity of a system with a MMSE receiver requires knowledge of the signal to interference plus noise ratio (SINR) of each spatial substream. To this extent, we use the moments of SINR for the MMSE receiver derived in [51].

For a statistically characterized system, asymptotic analysis in the large system regime (as the dimensions of the MIMO channel increases without bound) yields a closed form expression for the ergodic capacity [52, 53]. The results obtained in [51] also rely on the asymptotic spectral analysis of random matrices. What makes the asymptotic analysis important is that, due to fast convergence of the empirical eigenvalue distribution to its asymptotic limit, the asymptotic results hold with sufficient accuracy even for MIMO systems with very few antenna elements at both ends. The effectiveness of the asymptotic results enables us to design near-optimal linear precoders.

The rest of this chapter is organized as follows: In section 5.1, we further elaborate our precoder design criteria. In Section 5.2, we introduce the asymptotic random matrix analysis and we review some relevant results. In Sections 5.3 and 5.4, we present our precoder designs. Finally, we present the simulation results.
5.1 Precoder Design Criteria

Consider a MIMO wireless system with $M_T$ transmit antennas and $M_R$ receive antennas. The $M_R \times M_T$ channel matrix with transmit correlation can be written as

$$H = H_w R_t^{1/2},$$

(5.1)

where $R_t$ is the transmit correlation matrix and $H_w$ is a spatially white matrix whose elements are i.i.d ZMCS CG random variables. As described in Section 3.2, assuming uncorrelated inputs, the optimal linear precoder has the following structure: $F = V_F \Phi$. Where, the columns of $V_F$ are the precoder beam directions and $\Phi$ is the diagonal power allocation matrix. Next, we introduce the criteria that are used in this chapter to design precoders that exploit correlation CSIT.

Maximizing the system ergodic capacity:

The mutual information between the input and the output of the correlated MIMO channel with precoding is given as

$$I(H_w) = \log_2 \det \left( I_{M_R} + \gamma H_w R_t^{1/2} F F^H R_t^{H/2} H_w^H \right),$$

(5.2)

where $\gamma$ is the signal-to-noise ratio (SNR). The mutual information $I(H_w)$ is a random variable. The mean of this random variable with optimal precoder $F$ is referred to as ergodic capacity $C_{\text{erg}}$. The capacity-optimal precoder $F$ then is the solution of the
optimization problem

$$\max_{\mathbf{F}} \quad C_{\text{erg}} = \mathbb{E}_{\mathbf{H}_w} \left\{ \log_2 \det \left( \mathbf{I}_{M_T} + \gamma \mathbf{H}_w \mathbf{R}_t^{1/2} \mathbf{F} \mathbf{F}^H \mathbf{R}_t^{H/2} \mathbf{H}_w^H \right) \right\}$$

subject to \( \text{trace} ( \mathbf{F} \mathbf{F}^H ) = P. \) \hfill (5.3)

The above objective function involves evaluating the expectation with respect to \( \mathbf{H}_w \) and the expectation has no closed form solution. However, an upper bound on the ergodic capacity can be obtained via Jensen’s inequality.

**Maximizing the capacity of a system with MMSE receiver:**

For the MMSE receiver, the SINR on the \( k^{th} \), \( k = 1, \ldots, M_T \), spatial stream can be expressed as \[43\]

$$\text{SINR}_k = \frac{1}{\left[ \left( \mathbf{I}_{M_T} + \gamma \mathbf{H}^H \mathbf{H} \right)^{-1} \right]_{kk}} - 1, \quad (5.4)$$

where \( \gamma \) is the SNR. The total average capacity of a MIMO system with MMSE receiver can be approximated as

$$C_{\text{mmse}} = \sum_{k=1}^{M_T} \log_2 \left( 1 + \mathbb{E}\{\text{SINR}_k\} \right). \quad (5.5)$$

Since the channel matrix is random, the mutual information and the mean squared error with a linear receiver are random variables. In the next section, we show that these performance metrics are determined by the distribution of the eigenvalues of
the random channel matrix. In the asymptotic regime, as the dimensions of the random channel matrix increases without bound, these performance metrics tend to be deterministic. Surprisingly, these asymptotic results provide an excellent approximation even when the dimensions of the channel matrix are small. Motivated by this property, we use the asymptotic results for finite dimension precoder design.

5.2 Asymptotic Random Matrix Theory

Towards the goal of evaluating $C_{erg}$ and $\mathbb{E}\{\text{SINR}_k\}$, we review some relevant results from asymptotic random matrix theory. First, we introduce a few technical definitions.

The empirical cumulative distribution function (c.d.f) of the eigenvalues, also referred to as the empirical spectral distribution (ESD), of an $N \times N$ Hermitian matrix $A$ is defined as

$$F_A^N(x) = \frac{1}{N} \sum_{i=1}^{N} 1\{\lambda_i(A) \leq x\}, \quad (5.6)$$

where $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of $A$ and $1\{\text{.}\}$ is the indicator function. If $F_A^N(.)$ converges almost surely as $N \to \infty$, then the corresponding limit (asymptotic ESD) is denoted by $F_A(.)$.

Assume that the empirical eigenvalue distribution of the $N \times N$ Hermitian matrix $A$ converges almost surely.
Its Stieltjes transform is defined as \[52\]

\[
S_A(z) = \mathbb{E}\left[ \frac{1}{X - z} \right] = \int \frac{1}{\lambda - z} dF_A(\lambda).
\]

(5.7)

Its \(\eta\)-transform is defined as

\[
\eta_A(\gamma) = \mathbb{E}\left[ \frac{1}{1 + \gamma X} \right].
\]

(5.8)

Its Shannon transform is defined as

\[
\nu_A(\gamma) = \mathbb{E}[\log(1 + \gamma X)].
\]

(5.9)

Where \(X\) is a random variable whose distribution is the asymptotic ESD of \(A\) while \(\gamma\) is a nonnegative real number.

The rest of this section is organized as follows. We first consider an uncorrelated Rayleigh channel. The empirical eigenvalue distribution of the channel matrix converges to a non-random distribution as the dimensions of the channel matrix increase without bound. We show that, the mutual information is given by the Shannon transform of \(H_wH_w^H\) and the mean squared error is given by the \(\eta\)-transform of \(H_w^H H_w\). Next, we consider correlated Rayleigh channel which is the focus of this chapter. We present asymptotic expressions for the mutual information and the mean squared error. Finally, we present simulation results that demonstrate that the asymptotic results provide an excellent approximation even when the dimensions of the channel matrix are very small. In a latter section, we present a result that relates \(\mathbb{E}\{\text{SINR}_k\}\)
with the asymptotic expression for the mean squared error.

5.2.1 Uncorrelated Rayleigh MIMO channel

Consider an $M_R \times M_T$ Rayleigh MIMO channel matrix $\mathbf{H}_w$ whose entries are i.i.d complex random variables with zero-mean and variance $\frac{1}{M_R}$. The Marcenko-Pasture law [54] states that, as $M_R, M_T \to \infty$ with $\frac{M_T}{M_R} \to \beta$, the empirical eigenvalue distribution of $\mathbf{H}_w\mathbf{H}_w^H$ converges almost surely to a nonrandom limiting distribution with density

$$dF_{\mathbf{H}_w\mathbf{H}_w^H}(x) = \left(1 - \beta\right)^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi x}$$

(5.10)

where

$$a = (1 - \sqrt{\beta})^2 \quad \text{and} \quad b = (1 + \sqrt{\beta})^2.$$

Next, we show that the mutual information and the mean squared error are determined by the eigenvalue distribution of $\mathbf{H}_w\mathbf{H}_w^H$. Using (5.10), in the asymptotic regime, these performance measures can be expressed in terms of Shannon transform and $\eta$-transform.
Mutual information:

The normalized mutual information of the spatially white Rayleigh fading MIMO channel is given by

$$\frac{1}{M_R} C = \frac{1}{M_R} \log_2 \det(\mathbf{I}_{M_R} + \gamma \mathbf{H}_w \mathbf{H}_w^H)$$

$$= \frac{1}{M_R} \sum_{i=1}^{M_R} \log_2(1 + \gamma \lambda_i) \quad (5.11)$$

$$= \int_0^\infty \log_2(1 + \gamma x) dF_{M_R \mathbf{H}_w \mathbf{H}_w^H}(x).$$

Where, $dF_{M_R \mathbf{H}_w \mathbf{H}_w^H}(x)$ is the empirical probability density function of the eigenvalues of $\mathbf{H}_w \mathbf{H}_w^H$

$$dF_{M_R \mathbf{H}_w \mathbf{H}_w^H}(x) = \frac{1}{r} \sum_{i=1}^{M_R} \delta(x - \lambda_i) \quad (5.12)$$

and $\lambda_i$ are the eigenvalues of $\mathbf{H}_w \mathbf{H}_w^H$.

Based on the definition (5.9), as $M_R, M_T \rightarrow \infty$ with $\frac{M_T}{M_R} \rightarrow \beta$, the normalized mutual information of the MIMO channel is related to the Shannon transform of $\mathbf{H}_w \mathbf{H}_w^H$ by

$$\frac{1}{M_R} C = \int_0^\infty \log_2(1 + \gamma x) dF_{\mathbf{H}_w \mathbf{H}_w^H}(x)$$

$$\rightarrow \nu_{\mathbf{H}_w \mathbf{H}_w^H}(\gamma). \quad (5.13)$$

Using the Marcenko-Pastur law in (5.10), the Shannon transform of $\mathbf{H}_w \mathbf{H}_w^H$ is
given as [52]

\[ \nu_{\mathbf{H}_w\mathbf{H}_w^H}(\gamma) = \beta \log \left( 1 + \gamma - \frac{\mathcal{F}(\gamma, \beta)}{4} \right) + M_T \log \left( 1 + \gamma \beta - \frac{\mathcal{F}(\gamma, \beta)}{4} \right) - \frac{M_T \log e \mathcal{F}(\gamma, \beta)}{4\gamma}, \]

(5.14)

with \( \mathcal{F}(\gamma, \beta) = \left( \sqrt{\gamma^2 + 1} - \sqrt{\gamma^2 - 1} + 1 \right)^2. \)

In the asymptotic regime, the dependency of the mutual information on the realization of the random channel \( \mathbf{H}_w \) disappear. Mutual information of the system is rendered non-random and depends on the SNR and the channel dimensions.

**Mean squared error:**

The normalized total mean squared error of the MIMO system with an MMSE receiver is given as

\[ \frac{1}{M_T} \text{MMSE} = \frac{1}{M_T} \text{tr}\{(\mathbf{I}_{M_T} + \gamma \mathbf{H}_w^H \mathbf{H}_w)^{-1}\} \]

\[ = \frac{1}{M_T} \sum_{i=1}^{M_T} \frac{1}{1 + \gamma \lambda_i} \]

(5.15)

\[ = \int_0^\infty \frac{1}{1 + \gamma x} dF_{\mathbf{H}_w^H \mathbf{H}_w}(x). \]

Where, \( dF_{\mathbf{H}_w^H \mathbf{H}_w}(x) \) is the empirical probability density function of the eigenvalues of \( \mathbf{H}_w^H \mathbf{H}_w \) and \( \lambda_i \) are the eigenvalues of \( \mathbf{H}_w^H \mathbf{H}_w \).

Based on the definition (5.8), as \( M_R, M_T \to \infty \) with \( \frac{M_T}{M_R} \to \beta \), the normalized
MMSE of the MIMO system is related to the $\eta$-transform of $H_w^H H_w$ by [52]

\[
\frac{1}{M_T} \text{MMSE} = \int_0^\infty \frac{1}{1 + \gamma x} dF_{M_T H_w^H H_w}(x) \rightarrow \eta_{H_w^H H_w}(\gamma).
\]

(5.16)

Using the Marcenko-Pastur law in (5.10), the $\eta$-transform of $H_w^H H_w$ is given as [52]

\[
\eta_{H_w^H H_w}(\gamma) = 1 - \frac{\mathcal{F}(\gamma, \beta)}{4\gamma}.
\]

(5.17)

Similar to the mutual information, in the asymptotic regime, the MSE is non-random and depends on the SNR and the channel dimensions.

### 5.2.2 Correlated Rayleigh MIMO channel

In this subsection, we consider a MIMO channel with $M_T \times M_T$ transmit correlation matrix $T$. As usual, the entries of $H_w$ are assumed to be i.i.d complex random variables with zero-mean and variance $\frac{1}{M_T}$. The correlated MIMO channel is given by $H = H_w T^{1/2}$. Similar to the uncorrelated channel, for large dimensions, the normalized mutual information of the correlated MIMO channel corresponds to the Shannon transform of $H_w^H T H_w^H$ and is denoted by $\nu_{H_w^H T H_w^H}(\gamma)$. The normalized MSE of the MIMO system corresponds to the $\eta$-transform of $T^{1/2} H_w^H H_w T^{1/2}$ and is denoted by $\eta_{T^{1/2} H_w^H H_w T^{1/2}}(\gamma)$.
Mutual information:

The empirical eigenvalue distribution of $\mathbf{H}_w \mathbf{T} \mathbf{H}_w^H$ converges almost surely, as $M_R, M_T \to \infty$ with $\frac{M_T}{M_R} \to \beta$, to a distribution whose normalized mutual information (Shannon transform) satisfies [52]

$$\nu_{\mathbf{H}_w \mathbf{T} \mathbf{H}_w^H}(\gamma) = \beta \nu_{\mathbf{T}}(\eta \gamma) + \log_2 \frac{1}{\eta} + (\eta - 1) \log_2 e,$$

where $\eta = \eta_{\mathbf{H}_w \mathbf{T} \mathbf{H}_w^H}(\gamma)$ is the corresponding $\eta$-transform and it satisfies

$$\beta(1 - \eta_T(\eta \gamma)) = 1 - \eta.$$

(5.18)

In the above expressions, $\nu_{\mathbf{T}}(\eta \gamma)$ is the Shannon transform and $\eta_T(\eta \gamma)$ is the $\eta$-transform of the matrix $\mathbf{T}$.

Mean squared error:

The empirical eigenvalue distribution of $\mathbf{T}^{1/2} \mathbf{H}_w^H \mathbf{H}_w \mathbf{T}^{1/2}$ converges almost surely, as $M_R, M_T \to \infty$ with $\frac{M_T}{M_R} \to \beta$, to a distribution whose $\eta$-transform satisfies [52]

$$\eta_T(\gamma (1 - \beta + \beta \eta)) = \eta,$$

where $\eta = \eta_{\mathbf{T}^{1/2} \mathbf{H}_w^H \mathbf{H}_w \mathbf{T}^{1/2}}(\gamma)$ corresponds to the normalized total MSE with a linear receiver.

Simulation results in Section 5.5, demonstrate that these asymptotic results hold with sufficient accuracy even when the dimensions of the channel matrix are very
small. Hence, we use the expression (5.18) to design the linear precoder that maximizes the system ergodic capacity. The result in (5.20) provides the basis for designing the precoder that maximizes the capacity of the system with MMSE receiver.

5.3 Maximizing the System Ergodic Capacity

The ergodic capacity of finite dimensional MIMO channel with precoding is given as

\[
C_{\text{erg}} = \mathbb{E}_{\mathbf{H}_w} \left\{ \log_2 \det \left( \mathbf{I}_M + \gamma \mathbf{H}_w R_t^{1/2} \mathbf{F} \mathbf{F}^H \mathbf{R}_t^{H/2} \mathbf{H}_w^H \right) \right\}
\]

\[
\cong \nu_{\mathbf{H}_w \mathbf{T} \mathbf{H}_w^H}(\gamma) M_R.
\]

(5.21)

Where \( \mathbf{T} = R_t^{1/2} \mathbf{F} \mathbf{F}^H R_t^{H/2} \), \( \mathbf{R}_t \) is the transmit correlation matrix and \( \mathbf{F} = \mathbf{V}_F \mathbf{\Phi} \) is the precoding matrix. The shannon transform \( \nu_{\mathbf{H}_w \mathbf{T} \mathbf{H}_w^H}(\gamma) \) is given by (5.18).

Let \( \tau_i, i = 1, \ldots, M_T \), be the eigenvalues of the \( M_T \times M_T \) matrix \( \mathbf{T} \). The Shannon transform of the finite dimension matrix \( \mathbf{T} \) can be expressed as

\[
\nu_{\mathbf{T}}(\eta \gamma) = \int \log_2(1 + \eta \gamma \tau)dF_T(\tau)
\]

\[
= \frac{1}{M_T} \sum_{i=1}^{M_T} \log_2(1 + \eta \gamma \tau_i).
\]

(5.22)
The $\eta$-transform of the matrix $T$ can be expressed as

$$
\eta_T(\eta\gamma) = \int \frac{1}{1 + \eta\gamma\tau} dF_T(\tau)
$$

$$
= \frac{1}{M_T} \sum_{i=1}^{M_T} \frac{1}{1 + \eta\gamma\tau_i}.
$$

(5.23)

Since $\eta = \eta_{H_0} T H_0^T(\gamma)$ satisfies $\beta(1 - \eta_T(\eta\gamma)) = 1 - \eta$, using (5.23), we have

$$
\eta = 1 - \beta \left(1 - \frac{1}{M_T} \sum_{i=1}^{M_T} \frac{1}{1 + \eta\gamma\tau_i}\right).
$$

(5.24)

Using (5.22) and (5.24) in the capacity expression (5.21), we obtain the following closed form expression for the ergodic capacity

$$
C_{\text{erg}} = \sum_{i=1}^{M_T} \log_2 \left(1 + \eta\gamma\tau_i\right) - M_R \log_2 \left(1 - \beta \left(1 - \frac{1}{M_T} \sum_{i=1}^{M_T} \frac{1}{1 + \eta\gamma\tau_i}\right)\right)
$$

$$
- M_R \beta \left(1 - \frac{1}{M_T} \sum_{i=1}^{M_T} \frac{1}{1 + \eta\gamma\tau_i}\right) \log_2 e.
$$

(5.25)

In the above expression, $\eta$ can be obtained by solving (5.24) numerically. From (5.25), we observe that $C_{\text{erg}}$ depends on the SNR, the channel dimensions and on the eigenvalues of the matrix $T$. Next we show that, when the precoding beam directions match the eigenvectors of $R_t$, the matrix $T$ is rendered diagonal enabling us to find the capacity optimizing power profile.
5.3.1 Linear precoder design

A linear precoder matrix has the following structure $F = V_F \Phi^{1/2}$. Consider the eigenvalue decomposition $R_t = R_t^{H/2} R_t^{1/2} = V_R A_R V_R^H$. As described in Section 4.2, under covariance CSIT, the optimal precoding beam directions are given by the eigenvectors of $R_t$. Hence, we set $V_F = V_R$ and we have

$$T = R_t^{1/2} FF^H R_t^{H/2}$$

$$= \Lambda_R^{1/2} V_R^H V_F \Phi V_F^H V_R \Lambda_R$$

$$= \Lambda_R \Phi. \quad (5.26)$$

Hence, with $V_F = V_R$, the matrix $T$ is rendered diagonal. The diagonal elements (eigen values) of the matrix $T$ are

$$\tau_i = \lambda_i \phi_i, \quad (5.27)$$

where, $\lambda_i$ are the eigenvalues of $R_t$ and $\phi_i$ is the power allocated to the $i^{th}$ substream. Using (5.27) in (5.25), the objective function $C_{erg}$ turns out to be concave in $\phi_i$. The optimal power allocation is obtained by solving

$$\max_{\phi_i} C_{erg}$$

$$\text{s.t } \sum_{i=1}^{M_t} \phi_i = P. \quad (5.28)$$
Since (5.25) is a function of $\eta$ (5.24), the above optimization problem needs to be solved iteratively. In order to represent the dependence of $\eta$ on $\phi_i$, we denote it as $\eta(\Phi)$. Next we present the algorithm that finds the optimal power allocation $\phi_k$, $i = 1, \ldots, M_T$.

**Iterative power allocation algorithm:**

Initialization: $\phi_i^0 = P/M_T$, $i = 1, \ldots, M_T$ (Equal power allocation).

$n^{th}$ iteration:
1. Find $\eta^n(\Phi^{n-1})$ using (5.24).
2. Use $\eta^n(\Phi^{n-1})$ and find $\phi_i^n$, $i = 1, \ldots, M_T$, according to (5.28).

Iterate until convergence is achieved.

### 5.4 Maximizing the Capacity of System with MMSE Receiver

For a precoded MIMO system with MMSE receiver, the SINR on the $k^{th}$ spatial stream is given by

$$\text{SINR}_k = \frac{1}{[(I_{M_T} + \gamma \tilde{H}^H \tilde{H})^{-1}]_{kk}} - 1,$$

(5.29)

where $\tilde{H} = H_w R_t^{1/2} F$ and $\gamma$ is the SNR. In this section, we focus on designing a linear precoder that maximizes the system capacity

$$C_{\text{mmse}} = \sum_{k=1}^{M_T} \log_2 \left( 1 + \mathbb{E}\{\text{SINR}_k\} \right),$$

(5.30)

under a transmit power constraint. The expectation of $\text{SINR}_k$ with respect to $H_w$ has no close form solution. Fortunately, excellent approximations to the moments of
SINR were recently derived in [51] using random matrix theory. We use the expression for $\mathbb{E}\{\text{SINR}_k\}$ obtained in [51] for designing the precoder that maximizes the system capacity $C_{\text{mmse}}$ in (5.30).

We introduce the following notation: Let $A$ be an $r \times t$ matrix. The $r \times (t-1)$ matrix $A_{(-k)}$ is $A$ with the $k^{th}$ column removed and the $(r-1) \times (t-1)$ matrix $A_{(-k,-k)}$ is $A$ with the $k^{th}$ column and the $k^{th}$ row removed. The $r \times 1$ vector $a_k$ is the $k^{th}$ column of $A$, the $(r-1) \times 1$ vector $a_{k(-k)}$ is $a_k$ with the $k^{th}$ entry removed and $a_{kk}$ is the $(k,k)^{th}$ element of $A$.

From [51], the average SINR on the $k^{th}$ spatial stream is given by

$$
\mathbb{E}\{\text{SINR}_k\} = \frac{M_R - M_T + 1}{M_R} \gamma \Sigma_k + \frac{M_T - 1}{M_R} \gamma \Sigma_k \eta^k + \gamma e_k, \quad (5.31)
$$

where

$$
\Sigma_k = \frac{1}{\left[T^{-1}\right]_{kk}}, \quad e_k = \| (T_{(-k,-k)})^{-1} t_{k(-k)} \|^2 \quad \text{and} \quad T = F^H R_c F. \quad (5.32)
$$

In (5.31), $\eta^k = \eta \hat{H}_{(-k)}^H \hat{H}_{(-k)} (\gamma)$ is the $\eta$-transform of $\hat{H}_{(-k)}^H \hat{H}_{(-k)}$. Since $\hat{H} = H_w T^{1/2}$, we have $\hat{H}_{(-k)}^H \hat{H}_{(-k)} = T_{(-k,-k)}^{H/2} H_{w(-k)}^H H_{w(-k)} T_{(-k,-k)}^{1/2}$. Let $\tau_i, i = 1, \ldots, M_T - 1$, be the eigenvalues of the $(M_T - 1) \times (M_T - 1)$ matrix $T_{(-k,-k)}$. From (5.20), we have
\[ \eta^k = \eta_{T(-k,-k)}(\gamma(1 - \beta + \beta \eta^k)) \]

\[ = \frac{1}{M_T - 1} \sum_{i=1}^{M_T-1} \frac{1}{\gamma T_i(1 - \beta + \beta \eta^k) + 1} \]  

(5.33)

and \( \beta = \frac{M_T - 1}{M_R} \). From (5.31), we observe that \( \mathbb{E}\{\text{SINR}_k\} \) depends on \( T = \mathbf{F}^H \mathbf{R}_t \mathbf{F} \), the channel dimensions and on SNR. Next, we design a precoder \( \mathbf{F} \) that maximizes the system capacity \( C_{mmse} \).

### 5.4.1 Linear precoder design

Since, the optimal precoding beam directions under covariance CSIT are given by the eigenvectors of \( \mathbf{R}_t \), we set \( \mathbf{V}_F = \mathbf{V}_R \) and we have

\[ \mathbf{T} = \mathbf{F}^H \mathbf{R}_t \mathbf{F} \]

\[ = \Phi^{1/2} \mathbf{V}_F^H \mathbf{V}_R \Lambda \mathbf{V}_R^H \mathbf{V}_F \Phi^{1/2} \]  

(5.34)

\[ = \Lambda \mathbf{R} \Phi \]

Hence, with \( \mathbf{V}_F = \mathbf{V}_R \), the matrix \( \mathbf{T} \) is once again rendered diagonal. This enables us to write the expression (5.31) in terms of the eigenvalues of \( \mathbf{R}_t \) and \( \Phi \).

Since, \( t_{k(-k)} \) is the \( k^{th} \) column of the matrix \( \mathbf{T} \) with the \( k^{th} \) entry removed, when
\( \mathbf{T} \) is diagonal \( t_{k(-k)} = 0 \). With \( \mathbf{T} = \mathbf{\Lambda}_r \mathbf{\Phi} \), from (5.32) we obtain

\[
\Sigma_k = \frac{1}{(\mathbf{T}^{-1})_{kk}} = \lambda_i \phi_i \quad \text{and} \quad e_k = \| (\mathbf{T}_{(-k,-k)})^{-1} t_{k(-k)} \|^2 = 0. \tag{5.35}
\]

Since, \( \mathbf{T} \) is diagonal, the eigenvalues of \( \mathbf{T}_{(-k,-k)} \) are given by its diagonal elements. Hence, \( \tau_i = \lambda_i \phi_i \), \( i \neq k \) and \( i = 1, \ldots, M_T \). As a result, the \( \eta \)-transform in (5.33) can be expressed as

\[
\eta^k = \frac{1}{M_T - 1} \sum_{i=1, i \neq k}^{M_T} \frac{1}{\gamma \lambda_i \phi_i (1 - \beta + \beta \eta^k) + 1}. \tag{5.36}
\]

Note that, \( \eta^k \) in the above expression needs to be evaluated numerically. Finally, using (5.35) and (5.36) in (5.31), \( \mathbb{E}\{\text{SINR}_k\} \) can be expressed as

\[
\mathbb{E}\{\text{SINR}_k\} = \frac{M_R - M_T + 1}{M_R} \gamma \lambda_k \phi_k + \frac{M_T - 1}{M_R} \gamma \lambda_k \phi_k \eta^k \]

\[
= Z_k(\mathbf{\Lambda}_R, \mathbf{\Phi}) \gamma \lambda_k \phi_k, \tag{5.37}
\]

where

\[
Z_k(\mathbf{\Lambda}_R, \mathbf{\Phi}) = \frac{M_R - M_T + 1}{M_R} + \frac{M_T - 1}{M_R} \eta^k. \tag{5.38}
\]

Hence, the average SINR on the \( k \)-th spatial stream depends on the eigenvalues of \( \mathbf{R}_t \), the power allocation \( \mathbf{\Phi} \), the signal-to-noise ratio \( \gamma \) and on the dimensions of the channel.

With the precoder direction matrix been set to \( \mathbf{V}_F = \mathbf{V}_R \), maximizing the system capacity \( C_{\text{mse}} = \sum_{k=1}^{M_T} \log_2 \left( 1 + \mathbb{E}\{\text{SINR}_k\} \right) \) reduces to finding the optimal power
allocation. Using the expression for $\mathbb{E}\{\text{SINR}_k\}$ (5.37), $C_{mmse}$ turns out to be concave in $\phi_i$ and the optimal power allocation is by solving

$$
\max_{\phi_k} C_{mmse} = \sum_{k=1}^{M_T} \log_2 \left( 1 + Z_k(\Lambda_R, \Phi) \gamma \lambda_k \phi_k \right),
$$

subject to

$$
\sum_{k=1}^{M_T} \phi_k = P.
$$

This is a water filling problem and the solution is given as

$$
\phi_k = \left[ \frac{1}{M_T} \left( P + \sum_{i=1}^{M_T} \frac{1}{Z_i(\Lambda_R, \Phi) \gamma \lambda_i} \right) - Z_k^{-1}(\Lambda_R, \Phi) \gamma^{-1} \lambda_k^{-1} \right]^+.
$$

Due to the dependence of the term $Z_k(\Lambda_R, \Phi)$ on $\Phi$, the above expression needs to be evaluated iteratively. Next we present the algorithm that finds the optimal power allocation $\phi_k$, $k = 1, \ldots, M_T$.

**Iterative power allocation algorithm:**

Initialization: $\phi_k^0 = P/M_T$, $k = 1, \ldots, M_T$ (Equal power allocation).

$n^{th}$ iteration:

1. Find $Z_k^n(\Lambda_R, \Phi^{n-1})$, $k = 1, \ldots, M_T$, using (5.38) and (5.33).
2. Use $Z_k^n(\Lambda_R, \Phi^{n-1})$ and find $\phi_k^n$, $k = 1, \ldots, M_T$, according to (5.40).

Iterate until convergence is achieved.
5.5 Simulation Results

In this section, we present the computer simulations that have been performed to analyzed the proposed precoder solutions. We consider a MIMO system with $M_T = 3$ and $M_R = 6$. The proposed precoder designs are based on the premise that the asymptotic random matrix results hold with sufficient accuracy even when the dimensions of the system are small. In Figure 5.1, the simulated ergodic capacity is compared with the asymptotic ergodic capacity given by (5.25). The simulated ergodic capacity is obtained by averaging over 20,000 channel realizations. The plot shows that the asymptotic result approximates the simulated values excellently. Figure 5.2 shows the SINR of each substream with an MMSE receiver for a precoded system with equal power allocation ($\Phi = I$). The simulated and the asymptotic SINR, given by (5.37), are compared. The plots confirm the validity of the asymptotic results for finite dimension systems. Also, from Figures (1 and 2), we see that the asymptotic results are far more accurate compared to the commonly used bound obtained via Jensen’s inequality.

Choosing the number of transmit antennas $M_T = 3$ allows us to plot the system capacity as a function of the elements of the $3 \times 3$ power allocation matrix $\Phi$. The asymptotic ergodic capacity by is plotted as a function of $\phi_1$ and $\phi_2$ in Figure 5.3. Given $\phi_1$ and $\phi_2$, $\phi_3 = P - \phi_1 - \phi_2$. The plot shows that the ergodic capacity given by (5.25) is concave in $\phi_i$. The system capacity with an MMSE receiver, given by the objective function in (5.39), is plotted as a function of $\phi_i$ in Figure 5.4. The plot shows that $C_{mmse}$ is concave in $\phi_i$. 
Figure 5.1: Ergodic capacity vs SNR. Asymptotic values match the simulations very well.

Figure 5.2: SINR of each spatial substream vs SNR. Asymptotic values match the simulations very well.
Figure 5.3: Ergodic capacity vs $\phi_i$ at SNR=6 dB. Power constraint: $\sum_i \phi_i = M_T$. Ergodic capacity is concave in $\phi_i$.

Figure 5.4: Capacity with an MMSE receiver vs $\phi_i$ at SNR=6 dB. Power constraint: $\sum_i \phi_i = M_T$. System capacity is concave in $\phi_i$. 
The performance of the proposed precoding scheme to maximize the system ergodic capacity is shown in Figure 5.5. With equal power allocation, the precoder matrix $F = V_R$ is unitary. From (5.3), we see that a unitary precoder does not effect the ergodic capacity. Hence, with equal power allocation the performance is same as that with no precoding. The gain achieved by allocating the power optimally is more pronounced at low SNR. At the SNR increases, the gain achieved by optimal power allocation diminishes and equals equal power allocation. The achieved ergodic capacity at 6dB SNR in Figure 5.5 matches the maxima in the Figure 5.3. Hence, we conclude that the proposed iterative power allocation algorithm in Section 5.3 converges to the optimal $\Phi$. 

Figure 5.5: Performance of the proposed precoding scheme to maximize system ergodic capacity.
Figure 5.6: Performance of the proposed precoding scheme to maximize the capacity of system with MMSE receiver.

Figure 5.6 shows the performance of the proposed precoding scheme to maximize the capacity of the system with MMSE receiver. We see that, transmitting in the optimal directions alone provides considerable capacity gains. Allocating the power optimally further increases the capacity gain. However, this gain diminishes as the SNR increases. The achieved capacity at 6dB SNR in Figure 5.6 matches the maxima in the Figure 5.4. Hence, we conclude that the proposed iterative power allocation algorithm in Section 5.4 converges to the optimal $\Phi$.

5.6 Chapter Summary

In this chapter, we addressed the precoder design to exploit the knowledge of channel covariance information at the transmitter. Near-optimal linear precoders are designed under the following criteria: Maximizing the system ergodic capacity and maximizing
the capacity of the system with MMSE receiver. The proposed precoder designs are based on the results obtained via asymptotic spectral analysis of random matrices. Simulation results confirm the validity of the asymptotic results for finite dimension systems. Under both the design criteria, the optimal precoding beam directions are given by the eigenvectors of $\mathbf{R}_t$. Via simulation examples, the proposed iterative power allocation algorithms are shown to converge to the optimal.
Multiple-input multiple-output antenna systems offer substantial capacity improvements over single-antenna systems and the system performance can be further enhanced when perfect or partial channel state information is available at the transmitter [27, 28]. This chapter addresses the capacity optimization problem for MIMO systems with decorrelator receiver. We assume channel state information in the form of mean $H_m$ and transmit antenna correlation $R_t$. This corresponds to the general CSIT case. The resulting non-zero mean correlated MIMO channel can be modeled as in (4.4). The non-zero mean can also signify the presence of a direct line-of-sight component resulting in Ricean fading channel. Small antenna spacing is the primary cause of transmit correlation.

For MIMO systems with linear receivers, calculating the average system capacity requires knowledge of the statistics of the post-detection signal-to-noise ratio (SNR) of each spatial subchannel. In [55], for MIMO systems using a zero forcing receiver in zero-mean (Rayleigh) correlated channels, the SNR on each decoded subchannel is shown to be a weighted Chi-squared distributed. In [56], an approximation to the distribution of non-central complex Wishart distribution is used to derive the distribution of SNR. For the non-zero mean (Ricean) correlated channels, we approximate
the SNR of each spatial stream by a standard noncentral Chi-squared random variable. The degrees of freedom depend on the number of transmit and receive antennas and the noncentrality parameter depend on the channel mean and correlation matrices.

For system capacity optimization, a bound on the average capacity based on Jensen’s inequality is commonly used [57, 58]. Using the moments of the SNR, we obtain a Taylor series approximation for the average capacity that is significantly better than the Jensen bound. The obtained approximation is used to design a linear precoder that improves the average system capacity by exploiting the available channel state information.

The rest of the chapter is organized as follows: In section 6.1, we introduce the system model and precoder design criterion. In sections 6.2 and 6.3, we derive the approximate moments of the SNR and Taylor series approximation for the capacity of each subchannel. Section 6.4 describes the precoder design. Simulation results are presented in section 6.5.

6.1 System Model and Design Criterion

Consider a MIMO wireless system with $M_T$ transmit antennas and $M_R$ receive antennas. We assume knowledge of the channel mean $H_m$ and transmit correlation $R_t$ at the transmitter. The frequency flat MIMO channel can be modeled as

$$ H = H_m + H_w R_t^{1/2} $$

(6.1)
where $H_w$ is spatially white matrix and has i.i.d zero-mean, unit variance, complex Gaussian elements. As in (4.4), the effective channel mean $H_m$ corresponds to weighted sum of the outdated instantaneous channel $H_{t-\tau}$ and the statistical channel mean $\bar{H}$. We assume uncorrelated receive antennas ($R_r = I$).

A non-zero channel mean signifies the presence of a direct line-of-sight component. Hence, (6.1) can also model a Ricean correlated channel. In such a model, the channel mean $H_m = \sqrt{\frac{K}{\kappa+1}} H_0$ and the transmit correlation matrix $R_t = \frac{1}{\kappa+1} R_0$.

Where, $H_0$ is the normalized channel mean such that $\text{tr}(H_0^H H_0) = M_T M_R$; $R_0$ is the normalized transmit correlation matrix such that $\text{tr}(R_0) = M_T$ and $K$ is the ratio of power in the mean component to the average power in the random components of the channel.

For a flat-fading channel, the input-output relation for the precoded MIMO system is given as

$$\hat{x} = Gy = GHFx + Gn.$$  \hspace{1cm} (6.2)

Where, $x$ is the $M_T \times 1$ normalized transmitted signal vector, $F$ is the $M_T \times M_T$ precoder matrix, $G$ is the $M_R \times M_R$ receive matrix, $\hat{x}$ is the $M_R \times 1$ received signal vector and $n$ is the $M_R \times 1$ complex noise vector. It is assumed that $\mathbb{E}\{n\} = 0$ and $\mathbb{E}\{nn^H\} = \sigma_n^2 I_{M_R}$.

For the receiver, the effective channel matrix will thus be $HF$ and the receiver is assumed to estimate $HF$ perfectly. Imposing a zero forcing constraint $GHF = I$,
the decorrelator or zero forcing receiver which simply inverts the effective channel is \([41,43]\)

\[ G_{\text{decorr}} = (HF)^\dagger = (F^H H^H HF)^{-1} F^H H^H. \] (6.3)

The output of the decorrelator receiver is given by

\[ \hat{x} = x + (F^H H^H HF)^{-1} F^H H^H n, \] (6.4)

where we assume that \( M_R \geq M_T \) and \( H \) has full column rank.

For MIMO systems with a linear receiver, the optimal linear precoder matrix has the following structure \([14,15]\): \( F = B\Phi \), where \( B \) is unitary and \( \Phi \) is the diagonal power allocation matrix. The precoder must satisfy the power constraint \( \text{tr}(FF^H) = 1 \).

For the precoded MIMO system using a decorrelator receiver, the post-detection SNR of the \( k^{th} \) subchannel, \( k = 1, \ldots, M_T \), is easily found to be \([43]\)

\[ \text{SNR}_k = \frac{\gamma}{(B^H H^H HB)^{-1}}_{kk}, \] (6.5)

where, \( \gamma = \phi_k/\sigma^2_n \) and \( \phi_k \) is the \( k^{th} \) diagonal element of \( \Phi \). For a MIMO channel with a non-zero mean and transmit antenna correlation, the matrix \( B^H H^H HB \) is noncentral Wishart distributed. Hence the SNR of each substream is given by the reciprocal of the diagonal elements of an inverted noncentral Wishart distributed matrix.
For the model above, the total average capacity of the MIMO system is given as

\[ C = \sum_{k=1}^{M_T} \mathbb{E}\{ \log_2(1 + \text{SNR}_k) \}. \]  

(6.6)

And a plausible design criterion for the precoder \( \mathbf{F} \) can be stated as

\[ \max_{\mathbf{F}} \quad C \]  

(6.7)

\[ \text{s.t} \quad \text{tr}(\mathbf{F} \mathbf{F}^H) = 1. \]

Solving the above optimization problem involves finding the expectation over the log function and requires the knowledge of the statistics of \( \text{SNR}_k \). This a difficult and probably untractable problem. In order to simplify the problem, next we approximate \( \text{SNR}_k \) by a standard noncentral Chi-squared random variable and obtain its moments.

### 6.2 Moments of \( \text{SNR}_k \)

The \( M_R \times M_T \) matrix \( \hat{\mathbf{H}} = \mathbf{H} \mathbf{B} \) is normally distributed

\[ \hat{\mathbf{H}} \sim \mathcal{C}\mathcal{N}(\mathbf{M}, \mathbf{I}_{M_R} \otimes \mathbf{R}), \]  

(6.8)

where the mean \( \mathbf{M} = \mathbf{H}_m \mathbf{B} \) and the correlation matrix \( \mathbf{R} = \mathbf{B}^H \mathbf{R}_t \mathbf{B} \).

For our further exposition, we introduce the following notation: Let \( \mathbf{A} \) be an \( r \times t \) matrix. The \( r \times (t - 1) \) matrix \( \mathbf{A}_{(-k)} \) is \( \mathbf{A} \) with the \( k_{th} \) column removed and the \( (r - 1) \times (t - 1) \) matrix \( \mathbf{A}_{(-k,-k)} \) is \( \mathbf{A} \) with the \( k_{th} \) column and the \( k_{th} \) row removed. The \( r \times 1 \) vector \( \mathbf{a}_k \) is the \( k_{th} \) column of \( \mathbf{A} \), the \( (r - 1) \times 1 \) vector \( \mathbf{a}_{k(-k)} \) is \( \mathbf{a}_k \) with
the $k^{th}$ entry removed and $a_{kk}$ is the $(k, k)^{th}$ element of $A$. Then, from [51]

$$[A^{-1}]_{kk} = (a_{kk} - a_{k(-k)}^H(A_{(-k,-k)})^{-1}a_{k(-k)})^{-1}$$ (6.9)

Using (6.9), the SNR of the $k^{th}$ subchannel can be expressed as

$$\text{SNR}_k = \frac{\gamma}{[(\tilde{H}^H\tilde{H})^{-1}]_{kk}}$$

$$= \gamma(\tilde{h}_k^H\tilde{h}_k - \tilde{h}_k^H\tilde{H}_{(-k)}(\tilde{H}_{(-k)}^H\tilde{H}_{(-k)})^{-1}\tilde{H}_{(-k)}^H\tilde{h}_k).$$ (6.10)

Denote the SVD of $\tilde{H}_{(-k)} = \tilde{U}\tilde{D}\tilde{V}^H$ as usual. Further, let the $M_R \times (M_R - M_T + 1)$ matrix $\tilde{U}_c$ be the orthogonal compliment of $\tilde{U}$, implying that $\tilde{U}^H\tilde{U}_c = 0_{(M_T-1) \times (M_R-M_T+1)}$ and $U = [\tilde{U} \quad \tilde{U}_c]$ satisfies $U^HU = UU^H = I_{M_R}$. Plugging these into (6.10), we get

$$\text{SNR}_k = \gamma(\tilde{h}_k^HUU^H\tilde{h}_k - \tilde{h}_k^H\tilde{U}\tilde{D}\tilde{V}^H(\tilde{V}\tilde{D}^2\tilde{V}^H)^{-1}\tilde{V}\tilde{D}\tilde{U}^H\tilde{h}_k)$$

$$= \gamma(\tilde{h}_k^HUU^H\tilde{h}_k - \tilde{h}_k^H\tilde{U}\tilde{U}^H\tilde{h}_k)$$

$$= \gamma(\tilde{h}_k^H\tilde{U}_c\tilde{U}_c^H\tilde{h}_k) = \gamma s_k^H s_k,$$ (6.11)

where $s_k = \tilde{U}_c^H\tilde{h}_k$. Next, we obtain the distribution of $s_k$ using the properties of multivariate Normal distribution.

Conditioned on the $M_R \times (M_T - 1)$ matrix $\tilde{H}_{(-k)}$, the $M_R \times 1$ vector $\tilde{h}_k$ is Normally
distributed [59]

\[ \mathbf{h}_k|\mathbf{H}_{(-k)} \sim \mathcal{CN}(\mathbf{m}_k + (\mathbf{H}_{(-k)} - \mathbf{M}_{(-k)})\mathbf{R}_{(-k,-k)}^{-1}\mathbf{r}_k(-k), \mathbf{I}_{M_R} \otimes (r_{kk} - r_k^H \mathbf{R}_{(-k,-k)} r_k(-k))) \]  

(6.12)

As in [51], we denote the correlation between elements of \( \mathbf{h}_k|\mathbf{H}_{(-k)} \) by

\[ \Sigma_k = \frac{1}{[\mathbf{R}^{-1}]_{kk}} = r_{kk} - r_k^H \mathbf{R}_{(-k,-k)}^{-1} r_k(-k). \]  

(6.13)

Using (6.12) and (6.13) we see that, conditioned on \( \mathbf{H}_{(-k)} \), \( s_k \) is Normally distributed,

\[ s_k = \mathbf{U}_c^H \mathbf{h}_k|\mathbf{H}_{(-k)} \sim \mathcal{CN}(\tilde{\mu}_k, \Sigma_k \mathbf{I}_{M_R-M_T+1}), \]  

(6.14)

where

\[ \tilde{\mu}_k = \mathbf{U}_c^H (\mathbf{m}_k + (\mathbf{H}_{(-k)} - \mathbf{M}_{(-k)})\mathbf{R}_{(-k,-k)}^{-1}\mathbf{r}_k(-k)) \]

\[ = \mathbf{U}_c^H (\mathbf{m}_k - \mathbf{M}_{(-k)}\mathbf{R}_{(-k,-k)}^{-1}\mathbf{r}_k(-k)). \]  

(6.15)

The second equality in (6.15) results because \( \mathbf{U}_c^H \mathbf{H}_{(-k)} = \mathbf{U}_c^H \mathbf{U} \mathbf{D} \mathbf{V}^H = \mathbf{0} \).

Hence, conditioned on \( \mathbf{H}_{(-k)} \), \( 2s_k^H s_k / \Sigma_k \) is a standard non-central Chi-squared random variable with \( n = 2(M_R - M_T + 1) \) degrees of freedom and noncentrality parameter \( \lambda_k = 2\|\tilde{\mu}_k\|^2 / \Sigma_k \), where

\[ \|\tilde{\mu}_k\|^2 = (\mathbf{m}_k - \mathbf{M}_{(-k)}\mathbf{R}_{(-k,-k)}^{-1}\mathbf{r}_k(-k))^H \times \mathbf{U}_c \mathbf{U}_c^H (\mathbf{m}_k - \mathbf{M}_{(-k)}\mathbf{R}_{(-k,-k)}^{-1}\mathbf{r}_k(-k)). \]  

(6.16)
Note that $\|\tilde{\mu}_k\|^2$ depends on the random matrix $\tilde{H}_{(\sim k)}$. Assuming $U = [\tilde{U} \ \tilde{U}_c]$ to be a Haar matrix (see  A.2),

$$E\{\tilde{U}_c\tilde{U}_c^H\} = \frac{M_R - M_T + 1}{M_R}I_{M_R}. \quad (6.17)$$

Using (6.17), we have

$$E\{||\tilde{\mu}_k||^2\} = \frac{M_R - M_T + 1}{M_R}(m_k - M_{(\sim k)}R_{(\sim k,\sim k)}^{-1}r_{(\sim k)})(m_k - M_{(\sim k)}R_{(\sim k,\sim k)}^{-1}r_{(\sim k)}). \quad (6.18)$$

By replacing $||\tilde{\mu}_k||^2$ with $||\mu_k||^2 = E\{||\tilde{\mu}_k||^2\}$, we make $\text{SNR}_k$ independent of $\tilde{H}_{(\sim k)}$. This allows us to find closed form expressions for the first four moments of the Chi-squared random variable $\text{SNR}_k$ given as [60,61]

$$E\{\text{SNR}_k\} = \beta(n + \lambda_k), \quad (6.19)$$

$$\text{Var}\{\text{SNR}_k\} = \beta^22(n + 2\lambda_k), \quad (6.20)$$

$$\text{Sk}\{\text{SNR}_k\} = \beta^3\frac{2^{3/2}(n + 3\lambda_k)}{(n + 2\lambda_k)^{3/2}}, \quad (6.21)$$

$$\text{Ku}\{\text{SNR}_k\} = \beta^4\frac{12(n + 4\lambda_k)}{(n + 2\lambda_k)^2}, \quad (6.22)$$

where $\beta = \phi_k\Sigma_k/2\sigma_n^2$, $n = 2(M_R - M_T + 1)$ and $\lambda_k = 2||\mu_k||^2/\Sigma_k$. Next, we obtain an approximation for the average capacity of each subchannel using a Taylor series expansion.
6.3 Average Capacity Approximation

6.3.1 Via Jensen’s inequality

Jensen’s inequality [49] leads to a bound on the average capacity of each subchannel:

$$\mathbb{E}\{\log_2(1 + \text{SNR}_k)\} \leq \log_2(1 + \mathbb{E}\{\text{SNR}_k\}). \quad (6.23)$$

Using (6.19), this bound can be evaluated for given channel statistics $H_m$ and $R_t$.

6.3.2 Via Taylor series expansion

The Taylor series expansion [62] of a function $g(x)$ around $\eta$ is

$$g(x) \simeq g(\eta) + g'(\eta)(x - \eta) + \ldots + g^{(n)}(\eta) \frac{(x - \eta)^n}{n!}, \quad (6.24)$$

and for $\eta = \mathbb{E}\{x\}$

$$\mathbb{E}\{g(x)\} \simeq g(\eta) + g''(\eta) \frac{\sigma^2}{2} + \ldots + g^{(n)}(\eta) \frac{\mu_n}{n!}. \quad (6.25)$$

Thus, using the Taylor series expansion and the first four moments of SINR$_k$, we obtain the following approximation for the capacity of each subchannel

$$\mathbb{E}\{\log_2(1 + \text{SNR}_k)\} \simeq \log_2(1 + \mathbb{E}\{\text{SNR}_k\}) - \frac{\text{Var}\{\text{SNR}_k\} \log_2 e}{2(1 + \mathbb{E}\{\text{SNR}_k\})^2}$$

$$+ \frac{\text{Sk}\{\text{SNR}_k\} \log_2 e}{3(1 + \mathbb{E}\{\text{SNR}_k\})^3} - \frac{\text{Ku}\{\text{SNR}_k\} \log_2 e}{4(1 + \mathbb{E}\{\text{SNR}_k\})^4}. \quad (6.26)$$

The total average capacity of the MIMO system is given by (6.6). Using the above
average capacity approximation, we next derive a linear precoder \( \mathbf{F} = \mathbf{B}\Phi \) that maximizes the average capacity.

### 6.4 Linear Precoder Design

#### 6.4.1 Precoder rotation matrix

The optimal precoder \( \mathbf{F} = \mathbf{B}\Phi \), maximizes the total average capacity \( \sum_{k=1}^{M_T} \mathbb{E} \{ \log_2(1+\text{SNR}_k) \} \). However, finding a precoder rotation matrix \( \mathbf{B} \) that maximizes the capacity using the Taylor series approximation in (6.26) is untractable. Hence, we first obtain the matrix \( \mathbf{B} \) that maximizes \( \sum_k \log_2(1 + \mathbb{E}\{\text{SNR}_k\}) \), since it is the dominating term in the expression for the average capacity. Given \( \mathbf{B} \), the optimal power allocation matrix \( \Phi \) is then obtained by maximizing the Taylor series approximation for the average capacity.

Since, \( \mathbf{M} = \mathbf{H}_m\mathbf{B} \) and \( \mathbf{R} = \mathbf{B}^H\mathbf{R}_t\mathbf{B} \), we have \( \mathbf{m}_k = \mathbf{H}_m\mathbf{b}_k; \mathbf{M}_{(-k)} = \mathbf{H}_m\mathbf{B}_{(-k)};\)
\( r_{kk} = \mathbf{b}_k^H\mathbf{R}_t\mathbf{b}_k; \mathbf{r}_{k(-k)} = \mathbf{B}_{(-k)}^H\mathbf{R}_t\mathbf{b}_k; \mathbf{R}_{(-k,-k)} = \mathbf{B}_{(-k)}^H\mathbf{R}_t\mathbf{B}_{(-k)} \). Using these in (6.19), we obtain

\[
\mathbb{E}\{\text{SNR}_k\} = \beta(n + \lambda_k) = \gamma(\Sigma_k + \|\mu_k\|^2)
\]

\[
= \gamma(c(\mathbf{b}_k - \mathbf{G}_k\mathbf{r}_{k(-k)}))^H\mathbf{H}_m^H\mathbf{H}_m(\mathbf{b}_k - \mathbf{G}_k\mathbf{r}_{k(-k)}) + \mathbf{b}_k^H\mathbf{R}_t(\mathbf{b}_k - \mathbf{G}_k\mathbf{r}_{k(-k)}),
\]

(6.27)

where \( c = (M_R - M_T + 1)/M_R \) and \( \mathbf{G}_k = \mathbf{B}_{(-k)}[\mathbf{B}_{(-k)}^H\mathbf{R}_t\mathbf{B}_{(-k)}]^{-1} \).
The average SNR of each subchannel increases as the magnitude of the elements of the vector $G_k r_{k(-k)}$ decreases. For tractability purposes, we set $G_k r_{k(-k)} = 0$ and the above expression reduces to

$$E\{\text{SNR}_k\} = \gamma b_k^H \left( \frac{M_R - M_T + 1}{M_R} H_m^H H_m + R_t \right) b_k. \quad (6.28)$$

Consider the following decomposition of the average MIMO channel:

$$\left( \frac{M_R - M_T + 1}{M_R} H_m^H H_m + R_t \right) = V \Lambda V^H.$$  

It is well documented that a unitary precoder matrix that decouples the average MIMO channel increases the average capacity [15]. Decoupling of the average MIMO channel can be accomplished by setting $B = V$. Furthermore, setting $B = V$ reduces the magnitude of the off-diagonal elements of the matrix $R_t$ (i.e. $r_{k(-k)}$). This in turn reduces the magnitude of the elements of the vector $G_k r_{k(-k)}$ and results in an increase in the SNR of each subchannel.

### 6.4.2 Power allocation

On setting $B = V$ and using (6.26), the optimal power allocation that maximizes the average capacity is obtained by solving

$$\max_{\phi_k} \sum_k E \{ \log_2 (1 + \text{SNR}_k) \} \quad \text{s.t} \quad \sum_k \phi_k = 1, \quad (6.29)$$

which is a standard waterfilling problem.
6.5 Simulations results

In this section, we present numerical results related to the performance of the proposed capacity approximation and the designed precoder. The simulations are performed for a $4 \times 4$ MIMO system. Figure 6.1 shows the Taylor series approximation for the average capacity and the bound via Jensen’s inequality along with simulation results. At low SNR the Taylor series approximation is very accurate. At 12dB SNR, the Jensen bound over-estimates the capacity by 1.4 bits/sec/Hz while the difference between the Taylor series approximation and the simulated value is less than 0.5 bits/sec/Hz.

Figures 6.2 and 6.3 shows the performance of the proposed precoding scheme. Precoding gain achieved by waterfilling using the Taylor series approximation is higher...
Figure 6.2: Performance with precoding: Capacity vs SNR. K-factor=0.1.

Figure 6.3: Performance with precoding: Capacity vs K-factor.
than the precoding gain achieved by waterfilling using the Jensen bound. The observed improvement results because the Taylor series provides a better approximation of the average capacity. However, as the SNR increases or as the K-factor increases, the gain achieved by waterfilling diminishes and equals equi-power allocation. The CSIT parameters used in the simulations are listed in A.1.

6.6 Chapter Summary

For non-zero mean correlated MIMO channels with decorrelator receiver, we approximate the SNR of each spatial stream by a standard noncentral Chi-squared random variable. The degrees of freedom depend on the number of transmit and receive antennas and the noncentrality parameter depend on the channel mean and correlation matrices. Using the moments of the SNR of each subchannel, we obtain a Taylor series approximation for the total average capacity. The obtained approximation is used to design a linear precoder that maximizes the average capacity of the system. The simulation results illustrate the achieved performance gain.
Chapter 7: EXPLOITING MEAN AND COVARIANCE CSIT VIA APPROXIMATE JOINT DIAGONALIZATION IN STBC MIMO SYSTEMS

Multiple-input multiple-output systems significantly improve the wireless link performance through capacity and diversity gains and the system performance can be further enhanced when perfect or partial channel state information is made available at the transmitter [43]. Conventional space-time codes are designed to obtain maximum diversity gain under the assumption that the transmitter has no knowledge about the channel [10, 12]. Whereas, the precoder is a transmitter processing technique designed to obtain array gain by exploiting the available channel knowledge [14, 15]. Hence, the precoder and STBC combination can improve the system performance by delivering both diversity and array gains.

Linear precoding for STBC systems exploiting either the channel mean CSIT or the transmit correlation CSIT have been studied in [33–35]. For orthogonal-STBC systems, [31] formulates the PEP minimization problem using a general mean and covariance model, but only solves for the uncorrelated case. Precoder design exploiting both channel statistics was addressed in [63] using a convex optimization framework, but several relaxations were employed in order to obtain a solution for the general
In this chapter we propose a novel approach to designing linear precoders that exploit both the channel mean and transmit antenna correlation for general STBC-MIMO wireless systems. Our key contribution is to design a linear precoder via approximate joint diagonalization of the channel mean and transmit correlation matrices. The resulting precoder minimizes the Chernoff bound on the pair-wise error probability between a pair of block code words, averaged over channel fading statistics. The optimal precoder directions and the power distribution depend on the channel statistics and the STBC.

7.1 STBC system Model

Consider a MIMO wireless system with $M_T$ transmit antennas and $M_R$ receive antennas. We assume knowledge of the channel mean $H_m$ and transmit correlation $R_t$ at the transmitter. This corresponds to the general CSIT case (4.4). Assuming a quasi-static, frequency-flat channel, the $M_R \times M_T$ MIMO channel matrix can be written as

$$H = H_m + H_w R_t^{1/2},$$

(7.1)

where $H_w$ is spatially white matrix and has i.i.d zero-mean, unit variance, complex Gaussian elements. As explained in section 6.1, Ricean correlated channels can also be modeled by (7.1).

Assume the information bit stream to be transmitted is encoded into a space-time block codeword $C$ of size $M_T \times T$. Prior to transmission, these codewords are
multiplied by a $M_T \times M_T$ precoder matrix $\mathbf{F}$ in order to adapt the code to the available channel information. The $M_R \times T$ received signal block is given as

$$\mathbf{Y} = \mathbf{HFC} + \mathbf{N},$$

where $\mathbf{N} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$ is an additive complex white Gaussian noise matrix. To maintain a constant total average transmit power, the precoder must satisfy the power constraint $\text{tr}(\mathbf{FF}^H) = 1$.

The receiver is assumed to employ maximum-likelihood (ML) detection. The estimated codeword is

$$\hat{\mathbf{C}} = \arg \min_{\mathbf{C} \in \mathbb{C}} \| \mathbf{Y} - \mathbf{HFC} \|_F^2,$$

(7.3)

where $\mathbb{C}$ is the STBC codebook, and the subscript $F$ denotes the Frobenius norm.

### 7.2 Design Criterion

The probability that the transmitted codeword $\mathbf{C}^k$ is erroneously decoded as another codeword $\mathbf{C}^l$ is referred to as pair-wise error probability (PEP). It plays a crucial role in bounding the average error probability via union bound. Let $\mathbf{E}(k, l) = \mathbf{C}^k - \mathbf{C}^l$ be the codeword difference matrix. We define $\Omega$ as the set of non-zero codeword difference matrices $\Omega = \{k, l : \mathbf{E}(k, l) \neq \mathbf{0}\}$. For $\mathbf{E} \in \Omega$, with ML detection and applying the Chernoff bound, the PEP can be tightly upper bounded by [43]

$$\text{PEP}(\mathbf{H}, \mathbf{E}, \mathbf{F}) \leq \exp\left(-\frac{\gamma}{4}\|\mathbf{HFE}\|_F^2\right),$$

(7.4)
where $\gamma = 1/\sigma_n^2$ is the SNR. As we will see, the Chernoff bound allows closed form analysis of the PEP.

We are interested in the average performance over channel statistics and we aim to design a precoder that minimizes the worst-case PEP over all $E \in \Omega$ averaged over channel statistics. The design criterion translates into the following min-max problem

$$
F = \arg \min_F \max_{E \in \Omega} \{ \mathbb{E}_H[\text{PEP}(H, E, F)] \} \tag{7.5}
$$

s.t $\text{tr}(FF^H) = 1$.

Given the pdf of the channel matrix

$$
g(H) = \exp \left( - \frac{\text{tr}[(H - H_m)^H R_t^{-1}(H - H_m)]}{\pi^{M_T M_R} \det(R_t)^{M_R}} \right), \tag{7.6}
$$

after averaging (7.4) over the channel statistics, we obtain the following bound on the average PEP [31, 63]

$$
\text{PEP}(E, F) \leq \frac{\exp[\text{tr}(H_m W^{-1}H_m^H)]}{\det(W)^{M_R}} \det(R_t)^{M_R} \exp[-\text{tr}(H_m R_t^{-1}H_m^H)], \tag{7.7}
$$

where

$$
W = \gamma R_t FE E^H F^H R_t + R_t. \tag{7.8}
$$

Using the Woodbury matrix identity we find

$$
W^{-1} = R_t^{-1} - \frac{\gamma}{4} R_t FE (I + \frac{\gamma}{4} E^H F^H R_t FE)^{-1} E^H F^H. \tag{7.9}
$$
Further, using (7.9) in (7.7) and on simplification, the Chernoff bound on the average PEP can be expressed concisely as

\[ \text{PEP}(E, F) \leq \frac{\exp\left(\frac{\gamma}{4} \text{tr}(X)\right)}{\det(Y)^{M_R}}, \]  

(7.10)

where

\[ X = \left[ (E^H F^H H_m^H H_m F E) (I + \frac{\gamma}{4} E^H F^H R_t F E)^{-1} \right] \]

and

\[ Y = [I + \frac{\gamma}{4} E^H F^H R_t F E]. \]

On taking the logarithm of (7.10), the precoder design criterion can be stated equivalently as

\[ F = \arg \max_F \min_{E \in \Omega} \left[ \frac{\gamma}{4} \text{tr}(X) + M_R \log \det(Y) \right] \]

(7.11)

s.t \[ \text{tr}(FF^H) = 1. \]

This expression provides the basis for further analysis. Expression (7.11) suggests that optimization involves a trade-off between maximizing \( \text{tr}(X) \) and \( \det(Y) \). We tackle this problem by applying approximate joint diagonalization.

### 7.3 Precoder Design

#### 7.3.1 Optimal precoder structure

In this section, we derive a precoder that maximizes the function \( J = \frac{\gamma}{4} \text{tr}(X) + M_R \log \det(Y) \) over the precoder \( F \), for a given codeword difference matrix \( E \in \Omega \). Later we specialize it to the codeword difference matrix \( E \) that results in the worst
Consider the following SVDs: $H_m^H m = V_H \Lambda_H V_H^H$, $R_t = V_R \Lambda_R V_R^H$; $E E^H = V_E \Lambda_E V_E^H$; $F F^H = V_F \Phi V_F^H$. Using the Hadamard inequality, we have

$$\det(Y) = \det[I + \frac{\gamma}{4} E^H F^H R_t F E]$$

$$\leq \det[I + \frac{\gamma}{4} \Lambda_E \Phi \Lambda_R].$$

The equality occurs in (7.12) if and only if $F = V_R \Phi^{1/2} V_E^H$.

Also, it can be shown that

$$\text{tr}(X) = \text{tr}[(E^H F^H H_m^H H_m F E)(I + \frac{\gamma}{4} E^H F^H R_t F E)^{-1}]$$

$$\leq \text{tr}\left[(\Lambda_E \Phi \Lambda_H)(I + \frac{\gamma}{4} \Lambda_E \Phi \Lambda_R)^{-1}\right].$$

Further, $\text{tr}(E^H F^H H_m^H H_m F E) = \text{tr}(\Lambda_E \Phi \Lambda_H)$ when $F = V_R \Phi^{1/2} V_E^H$ and $\text{tr}(E^H F^H R_t F E) = \text{tr}(\Lambda_E \Phi \Lambda_R)$ when $F = V_R \Phi^{1/2} V_E^H$. Hence, for equality in (7.13) to occur, the eigenvectors of $F F^H$ must match the eigenvectors of $H_m^H H_m$ and $R_t$ simultaneously. That is, if we can find an orthogonal matrix $B$ that simultaneously diagonalizes the matrices $B^H H_m^H H_m B$ and $B^H R_t B$, then the precoder matrix $F$ that maximizes the function $J = \frac{\gamma}{4} \text{tr}(X) + M_R \log \det(Y)$ is given as

$$F = B \Phi^{1/2} V_E^H.$$
The diagonal matrix $\Phi$ containing the eigenvalues of $\mathbf{F}\mathbf{F}^H$ is the power allocation matrix. Under the constraint $\sum_i \phi_i = 1$, the available transmit power can be optimally distributed among the modes.

Next we describe approximate joint diagonalization. Approximate joint diagonalization involves finding an orthogonal matrix $\mathbf{B}$ such that $\mathbf{B}^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{B}$ and $\mathbf{B}^H \mathbf{R}_i \mathbf{B}$ are simultaneously as diagonal as possible.

### 7.3.2 Approximate joint diagonalization

Two $p \times p$ matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ are said to be simultaneously diagonalizable if there exists an orthogonal matrix $\mathbf{B}$ such that $\mathbf{B}^T \mathbf{A}_1 \mathbf{B}$ and $\mathbf{B}^T \mathbf{A}_2 \mathbf{B}$ are diagonal. Any two p.d symmetric matrices are simultaneously diagonalizable only if they commute (i.e $\mathbf{A}_1 \mathbf{A}_2 = \mathbf{A}_2 \mathbf{A}_1$). If $\mathbf{A}_1$ and $\mathbf{A}_2$ are not simultaneously diagonalizable (do not commute), it is always possible to find an orthogonal $p \times p$ matrix $\mathbf{B}$ which makes them simultaneously “as diagonal as possible” by optimizing with respect to a joint diagonality criterion.

Define $\text{off}(\mathbf{A}) = \sum_{1 \leq i < j \leq p} |a_{ij}|^2$, where $a_{ij}$ are the off-diagonal elements of the $p \times p$ matrix $\mathbf{A}$. The following functions act as measures of simultaneous deviation of the matrices $\mathbf{B}^T \mathbf{A}_1 \mathbf{B}$ and $\mathbf{B}^T \mathbf{A}_2 \mathbf{B}$ from diagonality.

$$
\Psi_1(\mathbf{B}^T \mathbf{A}_1, \mathbf{B}^T \mathbf{A}_2 \mathbf{B}) = \text{off}(\mathbf{B}^T \mathbf{A}_1 \mathbf{B}) + \text{off}(\mathbf{B}^T \mathbf{A}_2 \mathbf{B})
$$

(7.15)
\[ \Psi_2(B^T A_1 B, B^T A_2 B; n_1, n_2) = \frac{[\det(\text{diag}(B^T A_1 B))]^{n_1}[\det(\text{diag}(B^T A_2 B))]^{n_2}}{[\det(B^T A_1 B)]^{n_1}[\det(B^T A_2 B)]^{n_2}}, \tag{7.16} \]

where \( n_1 \) and \( n_2 \) are positive constants. Approximate joint diagonalization involves finding the orthonormal matrix \( B \) that minimizes the functions \( \Psi_1 \) or \( \Psi_2 \). The approximate joint diagonalization problem of minimizing the function \( \Psi_1 \) is solved by iterative Jacobi techniques in [64,65]. The \( \mathcal{FG} \)-algorithm described in [66] iteratively finds an orthogonal matrix \( B \) such that \( \Psi_2 \) is minimum.

On setting \( A_1 = H_m^H H_m \) and \( A_2 = R_t \), using approximate joint diagonalization, we obtain an \( M_T \times M_T \) orthogonal matrix \( B \) such that \( B^H H_m^H H_m B \) and \( B^H R_t B \) are simultaneously as diagonal as possible. Approximate joint diagonalization reveals the “average eigen-structure” shared by the matrices \( H_m^H H_m \) and \( R_t \). These average eigenvectors can be used as the near-optimal left singular vectors of the precoder matrix in (7.14).

### 7.3.3 Power allocation

Using the precoder \( F = B \Phi^{1/2} V_E^H \) and assuming a perfect joint diagonalization, i.e., \( B^H H_m^H H_m B = \Lambda_H \) and \( B^H R_t B = \Lambda_R \), renders the power allocation problem tractable. Specifically, we have

\[ \text{tr}(X) = \text{tr} \left[ (\Lambda_E \Phi \Lambda_H) \left( I + \frac{\gamma}{4} \Lambda_E \Phi \Lambda_R \right)^{-1} \right] \tag{7.17} \]
det(Y) = det(I + \frac{\gamma}{4} \Lambda_E \Phi \Lambda_R) \quad (7.18)

under the assumption of perfect joint diagonalization. Using (7.17) and (7.18) in (7.11), the optimal power allocation can be obtained by solving the following optimization problem

\[
\max_{\phi_i} \sum_i \left[ \frac{\gamma}{4} \lambda_{Ei} \lambda_{Hi} \phi_i + M_R \log \left( 1 + \frac{\gamma}{4} \lambda_{Ei} \lambda_{Ri} \phi_i \right) \right]
\]

s.t. \quad \sum_i \phi_i = 1. \quad (7.19)

The power allocated to each mode is a function of the eigen values of the channel mean, transmit correlation and \( EE^H \). In order to arrive at (7.19), we assumed perfect joint diagonalization of \( H_m^H H_m \) and \( R_t \). But, since the matrices \( H_m^H H_m \) and \( R_t \) do not commute, its not possible to jointly diagonalize them perfectly and approximate joint diagonalization by minimizing (7.15) or (7.16) is a near-optimum approximation. Clearly, with a perfect joint diagonalization assumption and on setting \( F = B \Phi V_E^H \), we obtain an upperbound on \( J = \frac{\gamma}{4} \text{tr}(X) + M_R \log \det(Y) \). We distribute the power optimally with respect to this bound.

### 7.4 Design Examples and Simulation Results

In this section, we give examples of designing a linear precoder for MIMO systems using the following non-orthogonal STBCs: A 2 \( \times \) 2 Golden Code and 4 \( \times \) 4 ABBA code.
7.4.1 Golden code

A codeword $C$ belonging to the Golden Code [67] has the following form

$$C = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha(c_1 + \theta c_2) & \alpha(c_3 + \theta c_4) \\ i\tilde{\alpha}(c_3 + \theta c_4) & \tilde{\alpha}(c_1 + \theta c_2) \end{bmatrix},$$

where $c_1, c_2, c_3, c_4$ are QAM symbols, $\theta = \frac{1+\sqrt{5}}{2}$, $\tilde{\theta} = \frac{1-\sqrt{5}}{2}$, $\alpha = 1 + i - i\theta$ and $\tilde{\alpha} = 1 + i - i\tilde{\theta}$. These codes are attractive because they provide full diversity and they achieve full transmission rate.

7.4.2 ABBA code

The structure of the quasi-orthogonal ABBA codeword [68] we consider and the corresponding codeword error product matrix are as follows

$$C = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \\ -c_2^* & c_1^* & -c_4 & c_3^* \\ -c_3 & c_4 & c_1 & c_2 \\ -c_4^* & c_3^* & -c_2 & c_1^* \end{bmatrix}; \quad EE^H = \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ b & 0 & a & 0 \\ 0 & b & 0 & a \end{bmatrix},$$

where $E = C - \hat{C}$ (C is the transmitted codeword and $\hat{C}$ is the detected codeword), $a = \sum_{i=1}^4 |c_i - \hat{c}_i|^2$ and $b = 2\text{Re}\{(c_1 - \hat{c}_1)(c_3 - \hat{c}_3) + (c_2 - \hat{c}_2)(c_4 - \hat{c}_4)\}$. These codes provide partial diversity, however, they achieve full transmission rate.

The largest PEP results when the distance between the codeword pair $(C, \hat{C})$ is minimum. If the occurrence probability of the minimum distance codeword pairs is not small, these pairs will dominate the error performance. Thus, the precoder
designed to minimize the largest average PEP will lead to a reasonable overall performance gain.

For STBCs, the distance between any two codewords is minimum when they differ by only one symbol. For such minimum distance Golden codeword pairs, the codeword error product matrix $\mathbf{E}\mathbf{E}^H$ is diagonal. The diagonal elements equal $[\frac{d_{\text{min}}^2}{5}|\alpha \theta|^2; \frac{d_{\text{min}}^2}{5}|\bar{\alpha} \bar{\theta}|^2]$, where $|\cdot|$ denotes the absolute value and $d_{\text{min}}$ is the minimum distance of the constellation. For minimum distance ABBA codeword pairs, $a = d_{\text{min}}^2$, $b = 0$ and $\mathbf{E}\mathbf{E}^H = d_{\text{min}}^2 \mathbf{I}_4$. Hence, for the considered non-orthogonal STBCs, the right singular vectors of the optimal precoder matrix will be an identity matrix. The solution for the min-max problem in (7.5) is obtained by setting $\mathbf{V}_E = \mathbf{I}$ in (7.14) and by optimally allocating power as per (7.19).

In Figure 7.1, under equal power allocation, the performance of the joint diagonalizing precoder is compared to the precoder obtained by diagonalizing $\mathbf{H}_m^H \mathbf{H}_m$ or $\mathbf{R}_t$ alone. At low values of $K$, the the channel approaches Rayleigh fading and the joint diagonalization approaches diagonalizing $\mathbf{R}_t$. At a high K-factor, the channel mean dominates and the joint diagonalization approaches diagonalizing $\mathbf{H}_m^H \mathbf{H}_m$. When a minimum distance codeword pair occurs, the PEP as a function of the SNR with different constellations is depicted in Figures 7.2 and 7.3 for the Golden and ABBA codes respectively. The average precoding gain is around 4 dB for Golden code and is around 2.2 dB with ABBA code at medium SNRs and diminishes at high SNRs. At high SNRs, the power allocation approaches equi-power leading to diminished precoding gain.
Assuming that all the codewords $\mathbf{C} \in \mathbb{C}$ are equally likely to be transmitted, an upper bound (union bound) on the average codeword error probability can be written as

$$
\bar{P}_c \leq \sum_{\mathbf{C} \in \mathbb{C}} \frac{1}{|\mathbf{C}|} \sum_{\mathbf{C} \neq \mathbf{C}} \text{PEP}(\mathbf{E}, \mathbf{F}),
$$

where $|\mathbf{C}|$ is the size of the codebook. $\text{PEP}$ is the PEP averaged over the channel statistics and is given by (7.10). The average codeword error probability $\bar{P}_c$ as a function of SNR is plotted in Figure 7.4. An average precoding gain of 1.8 dB is achieved by the system. The largest PEP results when the minimum distance codeword pairs occur. However, a higher precoding gain is observed for such pairs because, the right singular vectors of the precoder matches the eigenvectors of the codeword error product matrix. The CSIT parameters used in the simulations are listed in A.1.

7.5 Chapter Summary

We have proposed a novel approach to designing linear precoders that exploit channel state information at the transmitter in STBC-MIMO wireless systems. The channel knowledge consists of channel mean and transmit antenna correlation. By performing approximate joint diagonalization, we find an orthonormal change of basis which makes the channel mean and transmit correlation matrices as diagonal as possible. The optimal precoder minimizes the worst PEP averaged over channel fading and is a function of the channel statistics and the STBC. Our precoder design works with
Figure 7.1: Joint diagonalization of $H_m^H H_m$ and $R_t$ is better than diagonalizing $H_m^H H_m$ or $R_t$ alone. With Golden code using QPSK at SNR= 5 dB. Correlation coefficient $\alpha=0.7$

Figure 7.2: Performance of the precoded system when a minimum distance Golden codeword pair occurs. $K$-factor=0.1 and Correlation coefficient $\alpha=0.7$
Figure 7.3: Performance of the precoded system when a minimum distance ABBA codeword pair occurs. $K$-factor=0.1 and Correlation coefficient $\alpha=0.7$

all STBCs. Simulation results confirm the valuable precoding gain achieved by exploiting the available channel state information at the transmitter.
Figure 7.4: Union bound on the average probability of codeword error using QPSK. ABBA code, $K$-factor=0.1 and Correlation coefficient $\alpha=0.7$
Chapter 8: CONCLUSION AND FUTURE DIRECTIONS

The main conclusion of this thesis is that the performance of a MIMO system can be enhanced by exploiting the available channel state information at the transmitter. The thesis has focused on designing robust precoding schemes under various degrees of channel knowledge at the transmitter. This concluding chapter presents a summary of the thesis main results followed by a discussion of future research directions.

8.1 Summary

We analyzed the performance of linear and non-linear MIMO receivers for precoded systems. Precoding based on imperfect CSIT results in interference among the spatial subchannels. Assuming equal power allocation among the subchannels, we derived explicit expressions for the signal to interference-plus-noise ratio and the mean squared error. We demonstrated that exploiting the available instantaneous/statistical CSIT provides substantial capacity gains.

We designed near-optimal precoders that exploit channel covariance information at the transmitter. The proposed precoder designs are based on the results obtained via asymptotic spectral analysis of random matrices. Optimal precoding beam directions are given by the eigenvectors of the channel covariance matrix. We proposed novel
iterative power allocation algorithms that maximize the capacity.

We addressed the capacity optimization problem for MIMO wireless channels with a non-zero mean and transmit antenna correlation. With a decorrelator receiver, the capacity of the MIMO system is a function of the diagonal elements of an inverted noncentral Wishart distributed matrix. Hence, finding the average capacity is difficult. We simplify the problem by approximating the SNR of each spatial stream by a standard noncentral Chi-squared random variable. Using the moments of the SNR, we obtain a Taylor series approximation for the average capacity. The obtained Taylor series approximation is used to design a linear precoder that maximizes the total average capacity of the system.

We presented a novel linear precoder design based on the minimum pair-wise error probability criterion for general space-time block coded MIMO systems. The channel state information at the transmitter consists of the channel mean and transmit antenna correlation matrices. Approximate joint diagonalization reveals the “average eigen-structure” shared by the mean and correlation matrices. These average eigenvectors turn out to be the optimal left singular vectors of the precoder matrix. The right singular vectors of the precoder matrix depend on the space-time code. The singular values correspond to the power allocated to each mode, which depends on the channel statistics and the space-time code. A precoder design example for a non-orthogonal STBC and simulation results are provided to illustrate the achieved performance gain.
8.2 Future Directions

The focus of this thesis has been single-user MIMO systems. Further research is needed to extend the proposed precoder designs to a multi-user scenario where multiple antennas are deployed at the base station to support multiple users with one or more antennas per user terminal.

Bounds on the ergodic capacity of Ricean correlated MIMO channel have been obtained in [69]. Using these results for precoder design would be an avenue for further research. For a Ricean correlated MIMO channel, finding the moments of SINR of each spatial subchannel for the MMSE receiver is an open problem.

Exploiting other types of CSIT like channel condition number, channel K-factor and channel phase distribution is a rich area of research.
Bibliography
Bibliography


Appendix A: Appendix

A.1 Simulation Parameters

The CSIT parameters used in the simulations are listed below. Numbers are rounded to two digits after decimal point.

For a $4 \times 4$ channel, the mean used is

$$H_m = \sqrt{\frac{10K}{K+1}} \begin{bmatrix} \ .3 + .12i & .14 + .45i & -.02 + .16i & -.002 + .06i \\ -.17 - .001i & -.06 - .13i & .21 - .57i & .22 + .29i \\ .001 - .07i & .12 + .02i & .56 + .24i & .01 - .07i \\ -.19 + .3i & .16 - .2i & .13 + .05i & -.27 - .03i \end{bmatrix},$$

and the transmit covariance matrix used is

$$R_t = \frac{1}{K+1} \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix}.$$

For $3 \times 3$ and $3 \times 6$ channels, the transmit covariance matrix used is

$$R_t = \frac{1}{K+1} \begin{bmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{bmatrix}.$$
A.2 Harr Matrix

An $n \times n$ random matrix $U$ is a Haar matrix if it is uniformly distributed on the set, $U(n)$, of $n \times n$ unitary matrices [52]. For such matrices,

$$E\{|U_{ij}|^2\} = \frac{1}{n}$$

and

$$E\{U_{ij}U_{kj}\} = 0 \text{ for } i \neq k.$$ 

Hence, assuming the $M_R \times M_R$ unitary matrix $U = [\tilde{U} \quad \tilde{U}_c]$ to be a Haar matrix, for the $M_R \times (M_R - M_T + 1)$ matrix $\tilde{U}_c$ we obtain

$$E\{\tilde{U}_c \tilde{U}_c^H\} = \frac{M_R - M_T + 1}{M_R} I_{M_R}.$$