UNIVERSITY OF CINCINNATI

Date: July 31, 2006

I, Christopher D. Hicks,

hereby submit this work as part of the requirements for the degree of:

Master of Science

in:

Electrical Engineering

It is entitled:

Performance Analysis of MIMO Linear Precoders/Decoders

This work and its defense approved by:

Chair: Dr. James J. Caffery, Jr.

Dr. H. Howard Fan

Dr. Qing-An Zeng

PERFORMANCE ANALYSIS OF MIMO LINEAR PRECODERS/DECODERS

A thesis submitted to the

Division of Research and Advanced Studies
of The University of Cincinnati

in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE (M.S.)

in the Department of Electrical and Computer Engineering
and Computer Science

July 20, 2006

by

Christopher D. Hicks

B.S., Electrical Engineering, The University of Toledo, 2002

Thesis Advisor and Committee Chair:
Dr. James J. Caffery

Abstract

We address design of optimal MIMO precoders/decoders in the minimum mean square error (MMSE) sense. Prior work assumes full and perfect channel state information (CSI) at the transmitter, whereas we assume that the CSI has been corrupted in some manner. We re-derive new optimal precoders/decoders based on corrupt CSI. It is shown that the received signal cannot be simplified into parallel subchannels as is performed in previous work to greatly simplify that analysis. The error that corrupts the received signal is described and then bounded. Performance analysis is completed by describing the distribution of the error, SNR/SINR and mutual information equations, along with BER and capacity plots.

Acknowledgments

On a professional level I would like to thank Dr. Caffery for his endless patience and guidance in my pursuit of higher learning. I would also like to thank my fellow lab mates for their cultural enrichment, camaraderie and friendship which will never be forgotten.

On a personal level, I am indebted to my lovely wife Amanda, Jesus Christ, Sergey Shtelen, Mom, Grandma Beck and Grandpa Beck. Each of you have taught me many important lessons in your own unique way which I carry with me every day and I am grateful. Thank you.

Table of Contents

List of Figures

Glossary of Acronyms

1G	First Generation
3G	Third Generation
4G	Fourth Generation
AMPS	Advanced Mobile Phone Service
AOA	Angle of Arrival
ATDMA	Advanced Time Division Multiple Access
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
bps	bits per second
BS	Base Station
CDMA	Code Division Multiple Access
CSI	Channel State Information
DOA	Direction of Arrival
EVD	Eigenvalue Decomposition
FIR	Finite Impulse Response
FM	Frequency Modulation
Hz	Hertz
IID	independent identically distributed
IMTS	Improved Mobile Telephone Service
kHz	kilohertz
LOS	Line of Sight
MC-CDMA	Multi Carrier-CDMA
MHz	Megahertz
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
ML	Maximum Likelihood
MS	Mobile Station
OFDMA	Orthogonal Frequency Division Multiple Access

RF	Radio Frequency
SIMO	Single-Input Multiple-Output
SINR	Signal to Interference and Noise Ratio
SISO	Single-Input Single-Output
SLC	Side Lobe Canceler
SNR	Signal to Noise Ratio
STBC	Space-Time Block Code
SVD	Singular Value Decomposition
TDMA	Time Division Multiple Access
UWB	Ultra-Wide Band
WCDMA	Wideband CDMA
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian

Glossary of Mathematical Symbols

u	Scalar
$\mathbf{u}$	Vector
$\mathbf{U}$	Matrix
$(\cdot)^H$	Hermitian of a matrix
$(\cdot)^T$	Transpose of a matrix
$(\cdot)^{\frac{1}{2}}$	Square root of a matrix
$\mathbf{I}_{M_t}$	Identity matrix $M_t \times M_t$
$\mathrm{tr}\,(\cdot)$	Trace of a matrix
$\det(\cdot)$	Determinant of a matrix
$\|\|\cdot\|\|$	Norm of a vector or a matrix
$\mathrm{diag}\,(\cdot)$	Elements placed along the diagonal of a matrix
$\mathrm{vec}\,(\cdot)$	Stack the columns of a matrix into a vector
$\otimes$	Kronecker product
$\log_2(\cdot)$	Base 2 logarithm
$(\cdot)^*$	Moore-Penrose pseudo-inverse

Chapter 1

Introduction

The research trends in wireless communications are continually striving to reliably send more data at a faster rate over a wireless channel. More and more end users are needing to send higher quality data at a faster rate, and those needs are being met through today's research. In order to introduce the thesis topic, this chapter is intended to give a high level overview of the communications industry from its infancy up to the recent research trends that are being followed today so that the reader may have an understanding of the niche in which this thesis exists.

1.1 The Beginning of Radio

Inspired by the lectures of Righi at The University of Bologna, Guglielmo Marconi built the first radio telegraph in 1895. Marconi continued to improve his invention over the years and in 1898 his radio signal bridged the English Channel. In the subsequent years, Marconi integrated cutting edge technology into his radio equipment, such as the magnetic detector which was an improvement over the coherer to detect the radio waves, the use of directional antennas and the rotary spark to increase signal level and reduce interference in duplex receiver circuits, in addition to many others.

The Detroit Police Department was the first organization to make use of wireless technology by installing the first 2 MHz land mobile radiotelephone system in 1921. The widespread use of the new radiotelephone was limited by the number of available channels. Not until the invention

of frequency modulation (FM) in 1933 was there the availability of high quality radio communications.

The first mobile system to be connected to the public telephone network using a fixed number of radio channels over a fixed geographic area was introduced by Bell Systems in 1946. Bell Systems introduced the Personal Correspondence System which operated at 150 MHz with speech channels separated by 120 kHz. Researchers at AT&T introduced the Improved Mobile Telephone Service (IMTS) which also used FM technology. It was soon noted that in order to provide a large number of users with full duplex channels, excessive amounts of bandwidth would be required.

In 1947 Bell Laboratories came up with the cellular concept. The cellular concept helped to overcome the excessive bandwidth needs by dividing the coverage area into smaller cells and by using a subset of the total available channels in each cell. AT&T then proposed the Advanced Mobile Phone Service (AMPS) in 1970, which was the first high capacity analog cellular telephone system.

Since then, the cellular networks have evolved rapidly and today are using the IS-36 (using Time Division Multiple Access - TDMA) and IS-95 (using Code Division Multiple Access - CDMA) standards. The increase in the need for end users to send large quantities of data at a faster rate over wireless links has led to the development of the current Third Generation (3G) wireless technologies. Novel techniques must be developed in order to make use of the limited radio frequency (RF) spectrum and the use of multiple-input multiple-output (MIMO) antenna arrays is one of those techniques [1]. Through the use of multiple antenna arrays at the transmitter and the receiver, multiple copies of the signal will be sent and received which decreases the likelihood of errors.

The Fourth Generation (4G) of wireless communications is on the horizon and will have MIMO technology at its foundation. Some topics that are at the forefront in the research community are Advanced Time Division Multiple Access (ATDMA), Wideband Code Division Multiple Access (WCDMA), Orthogonal Frequency Division Multiple Access (OFDMA), Multi-Carrier CDMA (MC-CDMA), and Ultra-Wide Band (UWB) transmission [2].

The following sections highlight recent trends in the development of wireless communication technologies through the use of Smart Antennas and the inception of MIMO and the benefits.

1.2 Smart Antenna Systems

Due to the limitations of the available RF spectrum, the spatial dimension may be exploited through the use of multiple antennas. When multiple antennas are used at the receiver and when the received signals are decorrelated, the use of smart antennas mitigates signal fluctuations and improves system performance because the receiver may designed so that interference is minimized and signal strength is maximized.

Smart antennas began the 1960s as a proposed measure for electronic warfare in order to counter jamming. Today, they may be categorized into three main groups: switched beam antennas, dynamic phased arrays, and adaptive antenna arrays.

Switched beam antennas are comprised of multiple highly directive, fixed, pre-defined beams, as shown in Fig. 1.1. The beams are usually formed with an array of antennas, through the use of a beamforming network such as a Butler Matrix [3]. By generating multiple orthogonal beams to blanket the cell with coverage [4] the base station (BS) is able to track the mobile station (MS) through the cell according to the received signal strength. Switching from beam to beam may be done through the use of semiconductor switches. Switched beam antennas are advantageous in situations where there is low to moderate co-channel interference. This leads to one major disadvantage which is if the desired user is not centered at the beam's maximum and an interfering signal is present at the beam's center, then the interfering signal will be enhanced more than the desired signal, leading to poor system performance.

Dynamic phased arrays are an improvement upon switched beam antennas because they make use of the direction of arrival (DOA) of the incoming wavefront and steer a beam maximum toward the desired user. In order to continuously steer the beam toward the desired user, a method of tracking is needed.

Adaptive antenna arrays weight the gain of the array to maximize performance measures in order to maximize the received signal. The adaptive system is able to continually determine the angle of arrival (AOA) which gives it the ability to maximize the incoming signal and minimize interference and noise [3], as shown in Fig. 1.2. An early version of an adaptive antenna array is known as a side lobe canceler (SLC). A SLC works to cancel the interference entering the

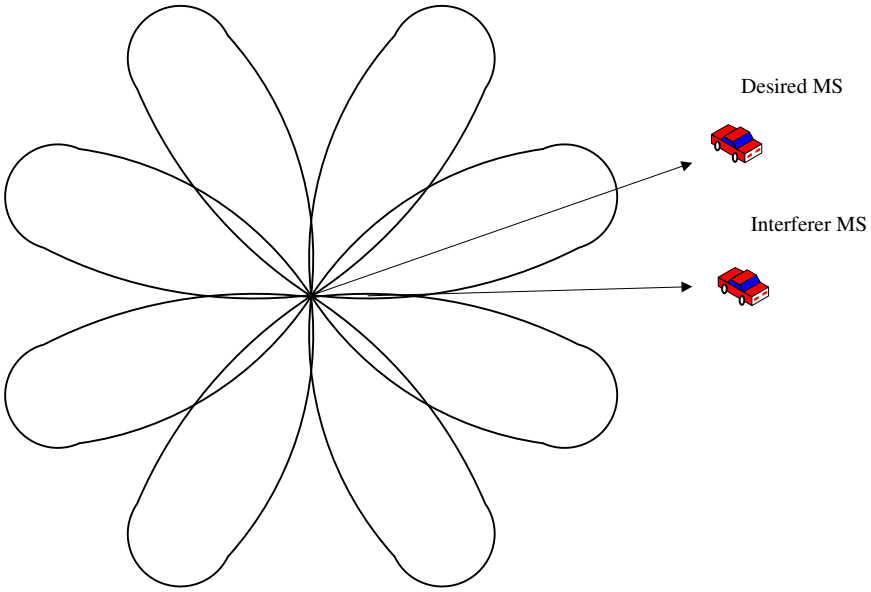


Figure 1.1: A switched beam antenna system.

sidelobes of the antenna. A SLC requires two channels for implementation. The first channel is used to discern the signal of the MS, whereas the second channel is used to pick up the interference and discriminate against the desired signal. Adaptive antenna arrays are more computationally intensive and more expensive to implement than switched beam systems [4].

1.3 MIMO Antenna Systems

The use of multiple antennas allows the exploitation of the multipath fading environment. A MIMO antenna scheme is shown in Fig. 1.3. Due to the multipath propagation of a transmitted radio signal, the received signal is the superposition of multiple copies of the received signals. If there is no line of sight (LOS) component, then the signal is considered to exhibit Rayleigh fading [5].

Quite often the propagation environment varies with time. Due to the time varying nature of the channel and the Rayleigh distribution of the received amplitude, the received channel gain may be small or large. This results in the attenuation or amplification of the signal and steps are taken to alleviate this problem through different methods of diversity. Time diversity is the transmission of the same signal at different time instances, and frequency diversity is the transmission of the same

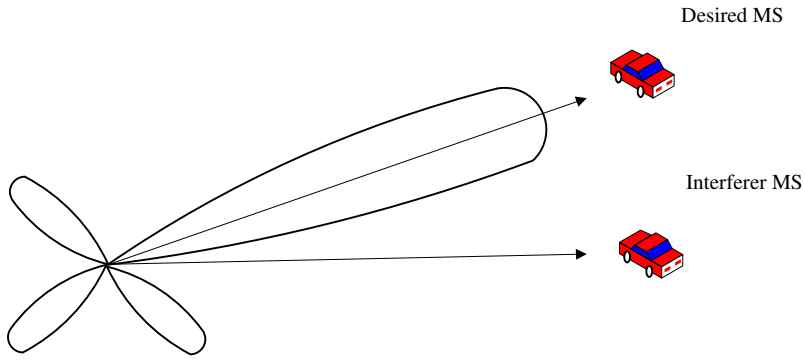


Figure 1.2: An adaptive antenna array.

signal in different frequency bands. In addition to time and frequency diversity, antenna diversity is the exploitation of the independent fading between properly spaced antennas [5].

Antenna diversity may also be known as the exploitation of the spatial dimension. By implementing multiple antennas at the MS one may achieve a performance gain in the downlink with an improved link budget and a higher tolerance to co-channel interference. By adding receive antennas at the BS, the receive diversity improves the quality of the uplink without any cost to the end users who operate the MS [5]. Receiver diversity is also known as single-in multiple-out (SIMO) and is shown in Fig. 1.4.

The exploitation of the spatial dimension may take place at the transmitter as well, known as transmit diversity or MISO (Multiple-input single-output), and is shown in Fig. 1.5. The benefits of transmitter diversity are in accord with those of receiver diversity.

1.3.1 Recent Research Thrust: Precoding and the Utilization of Channel Knowledge

In the event that the transmitter has knowledge of the channel prior to transmission, then it would benefit the system performance measures to make use of this information. Beamforming is one method which uses an adaptive antenna array in which weighted symbols are used to steer the energy in the direction of the receiver. Another method is space-time diversity which is inferior

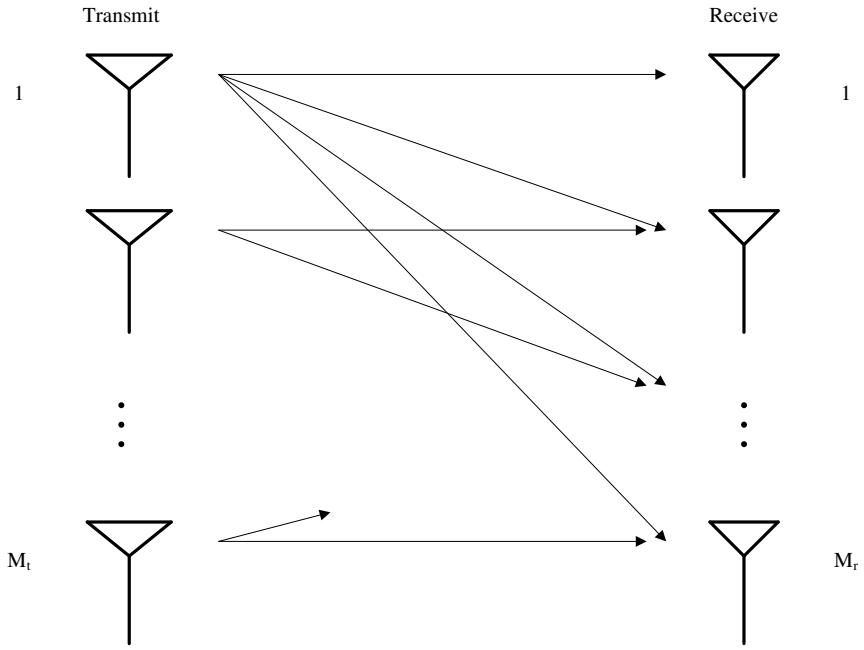


Figure 1.3: Multiple-input multiple-output (MIMO) diagram.

to beamforming because it distributes energy equally across the channel. Through the use of prior knowledge of the channel, the performance of space-time diversity approaches that of beamforming [5].

There are many approaches in how to most efficiently use channel knowledge, also known as channel state information (CSI). CSI of the complete channel realization is known as full CSI and is used in [6], [7], [8] and [9]. Partial CSI refers to the use of statistics of the channel which may be the correlation at the transmitter or receiver, the covariance of the channel or the mean of the channel. Partial CSI using the correlation of the received signals at the transmitter and receiver is used in [10], whereas only transmit fading correlations are used in [11]. Partial CSI comparing the use of amplitude only and the use of phase only is compared in [8]. The covariance of the CSI is used in [12], [10], [13]. In [12], the use of the mean of the CSI is investigated as well.

To capitalize on the benefits of CSI at the transmitter, CSI is used to premulitply the symbols prior to transmission, which may be referred to as weighting, prefiltering, or precoding. The receiver will typically use some form of a maximum likelihood detector which is referred to as a postfilter or a decoder. Here it is referred to as precoding and decoding at the transmitter and receiver, respectively. This thesis will take a closer look at systems using both precoder and a

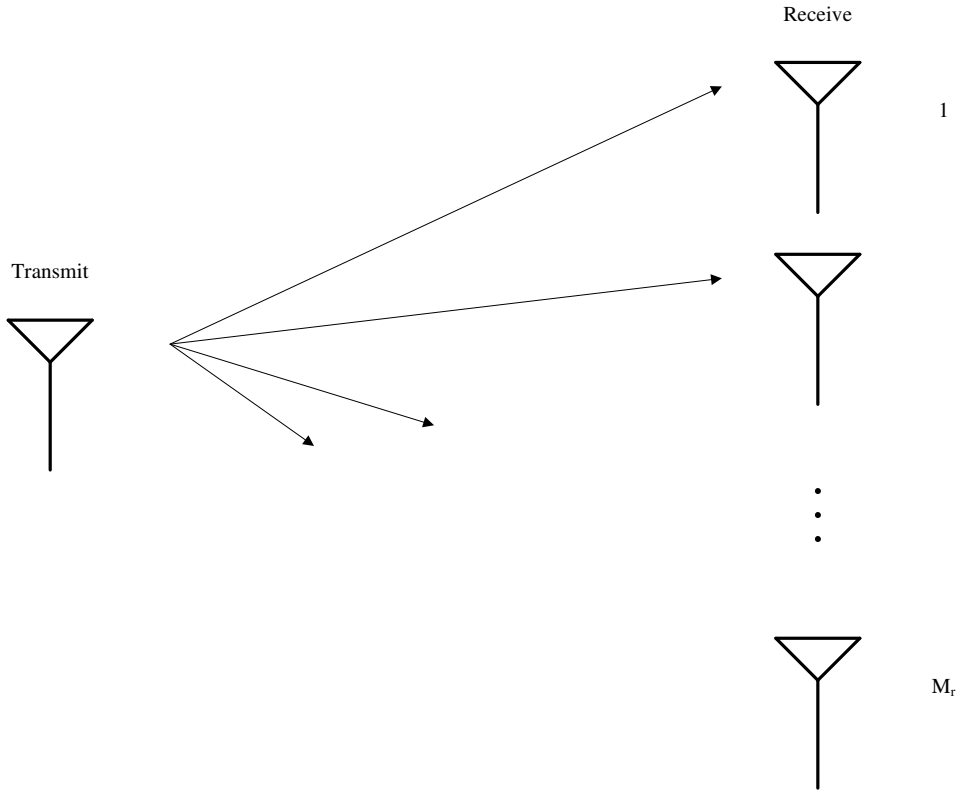


Figure 1.4: Single-in multiple-out (SIMO) diagram.

decoder where full CSI is assumed and analyze the performance loss in the event of imperfect CSI. In mathematical literature the analysis of an error term added to a known value is referred to as perturbation analysis. In effect, the thesis will analyze the performance loss incurred in the computation of the precoder/decoder due to perturbation, or errors in the CSI.

1.4 Thesis Overview and Outline

The outline of the thesis is as follows, Chapter 2 will present background information on the capacity benefits of a MIMO system as well as MIMO precoders/decoders. Chapter 3 explains further the system models whose error analysis is presented in chapter 4. The performance results of the error analysis of chapter 4 is presented in chapter 5. The concluding statements and future work is shown in chapter 6.

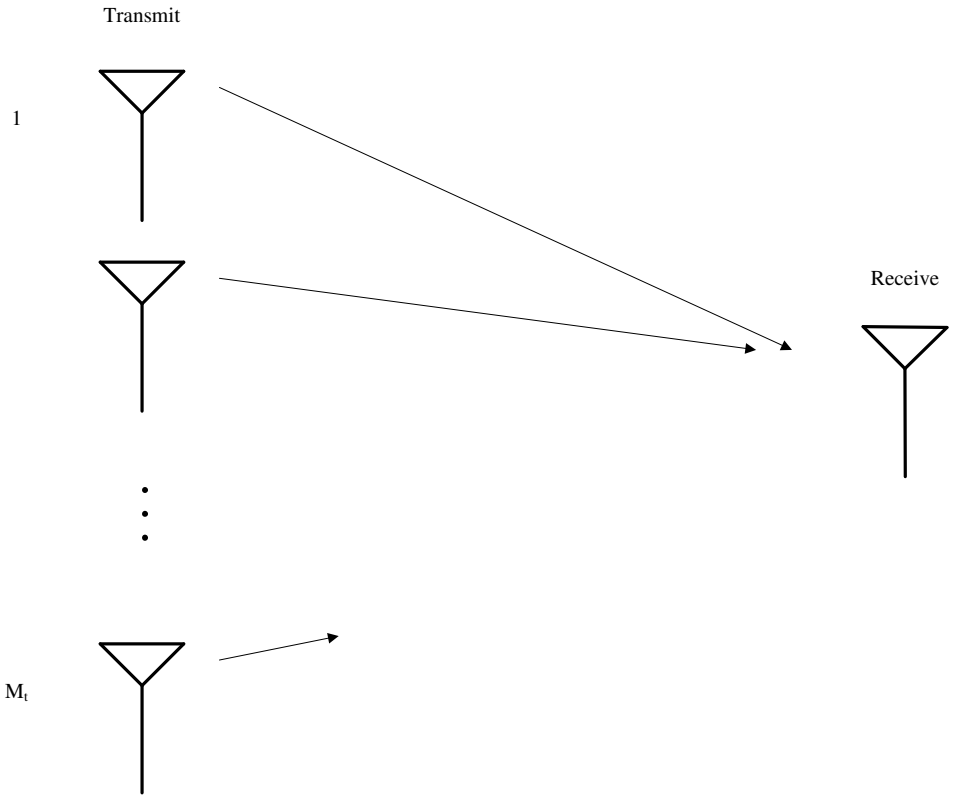


Figure 1.5: Multiple-in single-out (MISO) diagram.

Chapter 2

Background

This chapter presents the benefits of MIMO through an explanation of the performance benefits as well as the theoretical insight behind its usage. Herein lies a description of the wireless channel which introduces the performance gains of a MIMO system, and closes by delving further into the concept of precoding and how that puzzle piece completes the background picture necessary to understand the contribution of this thesis.

2.1 Channel Model

The wireless communications channel is assumed to be the propagation environment between a BS that is elevated from the terrain and devoid of local scatterers, and a MS at ground level which receives its radio waves via reflections, diffraction and scattering as shown in Fig. 2.1. The constructive and destructive vectorial combination of the arriving waves is known as multipath propagation.

A typical MS is modeled as being surrounded by scatterers, resulting in a faded envelope that is devoid of a specular, or line of sight (LOS) component. This type of fading is known as Rayleigh fading. If a LOS component is present then the received signal is dominated by the stronger signal, resulting in a channel modeled as Ricean fading. The channel model may also be described as either flat or frequency selective fading. Flat fading implies that the channel is simply a single tap

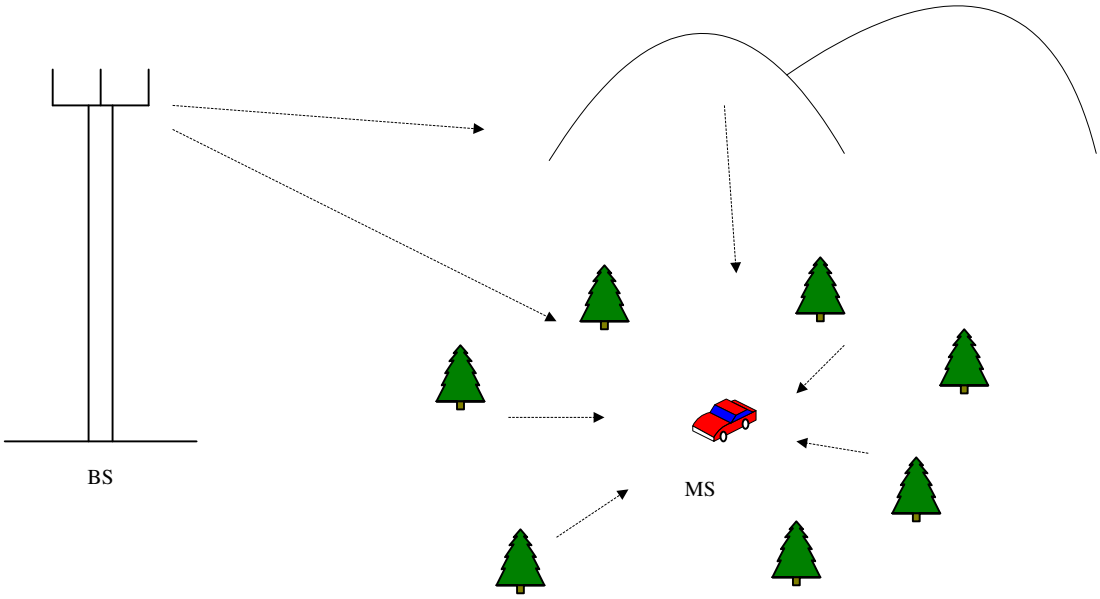


Figure 2.1: Typical MS/BS scenario.

filter with a time-varying gain, whereas frequency selective fading is modeled as a multi-tap filter with a time-varying gain. [14] [15].

2.2 Capacity Benefits of MIMO

The capacity benefits of a MIMO channel are presented in this section. Beginning with the development of the original single-input single-output (SISO) channel capacity developed by Shannon in 1948, and followed with the capacity of a MIMO channel with and without CSI at the transmitter. Throughout this section the channel **H** is assumed to be constant and known to the receiver which is maintained through training and tracking.

2.2.1 SISO Capacity

Capacity is a measurement which quantifies the maximum possible amount of information that may be transmitted across the wireless channel. The transmitted and received signals are random variables, this implies that a measure of the entropy is necessary because it is a measure of the

uncertainty in a random experiment [16]. Given a discrete time, additive white Gaussian noise (AWGN) SISO channel, the received signal is given by

$$z = hx + n, \tag{2.1}$$

where x is the transmitted signal, z is the received signal, h represents the channel, and n is the AWGN. The mutual information is given by the difference between the conditional entropy ($h\left(z|x\right)$) and the differential entropy ($h\left(z\right)$) and is written as [15]

$$I\left(x;z\right) \;\; = \;\; h\left(z\right) - h\left(z|x\right).$$

Mutual information is a measure of information shared between the transmitter and receiver and is given by the base two logarithm of the ratio of a posteriori probability to a priori probability. Given the transmission of a random input x and reception of a random input z, the mutual information is given by [17, 15]

$$I\left(x;z\right) = \sum_{x\in\mathcal{X},z\in\mathcal{Z}} p\left(x,z\right)\log_2\left(\frac{p\left(x,z\right)}{p\left(x\right)p\left(z\right)}\right), \tag{2.2}$$

where $p(x)$ indicates the probability of the random input x. The capacity C, is then defined as the maximization of the mutual information across all possible input distributions as demonstrated by Shannon [18] in his well-known formula

$$C = B\log_2\left(1+SNR\right), \tag{2.3}$$

where B is the bandwidth of the channel, SNR is the signal to noise ratio and the capacity is in bits per second per Hertz (bps/Hz). If the channel is modeled as a random, independent identically distributed (IID - spatially white) channel, the above equation for channel capacity remains valid.

2.2.2 MIMO Capacity

The derivation of the capacity of a SISO channel is easily extended to the MIMO case. The new received signal is given by expanding (2.1) to become

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n},$$

where **z** is the $M_r \times 1$ received signal vector, **x** is the $M_t \times 1$ transmit vector, **H** is a $M_r \times M_t$ matrix representing the channel and **n** is a $M_r \times 1$ vector of AWGN. The mutual information of a MIMO channel is the difference between the conditional entropy and the differential entropy of **z**, and is defined as [15], [1]

$$I(\mathbf{x};\mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z}|\mathbf{x}). \tag{2.4}$$

Since the vectors **x** and **n** are independent, (2.4) simplifies to

$$I(\mathbf{x};\mathbf{z}) = H(\mathbf{z}) - H(\mathbf{n}). \tag{2.5}$$

The maximization of (2.5) simplifies to become the maximization of $H(\mathbf{z})$. The covariance matrix of **z** satisfies the maximization of $H(\mathbf{z})$ when **z** is zero mean circularly symmetric complex Gaussian (ZMCSCG). This implies that **x** must be ZMCSCG as well. The covariance matrix of **z** is given by

$$\begin{aligned} E[\mathbf{z}\mathbf{z}^H] &= \mathbf{R}_{zz} \\ &= \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{R}_{nn}, \end{aligned} \tag{2.6}$$

where $\mathbf{R}_{xx} = E[\mathbf{x}\mathbf{x}^H]$ and $\mathbf{R}_{nn} = E[\mathbf{n}\mathbf{n}^H] = \sigma_{nn}^2\mathbf{I}$. This allows (2.5) to become

$$I(\mathbf{x};\mathbf{z}) = \log_2\left(\det\left(\sigma_{nn}^{-2}\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{I}\right)\right). \tag{2.7}$$

Similar to the SISO derivation of (2.3), the MIMO channel capacity is the maximization of the mutual information (2.7) over all input covariance matrices [15] [1], and is given by

$$C = \max_{\mathbf{R}_{xx}} B\log_2\left(\det\left(\sigma_{nn}^{-2}\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{I}\right)\right). \tag{2.8}$$

In the event that the channel is modeled as a random, IID channel, the above equation for capacity differs. If $M_r = M_t = M$, according the the strong law of numbers the channel covariance matrix becomes

$$\frac{1}{M}\mathbf{H}\mathbf{H}^H \rightarrow \mathbf{I}_M,$$

as M approaches infinity. This results in a revised capacity equation that may be written as

$$C = M\log_2(1 + SNR).$$

This may be interpreted as while the number of transmit and receive antennas increase, the capacity of the random MIMO channel approaches that of the deterministic MIMO channel as well as increases linearly in M [1].

Throughout the rest of the document, only deterministic MIMO channel models will be studied. For a presentation of MISO or SIMO capacity please refer to [1].

CSI Unavailable at the Transmitter

In the event that the channel is unknown to the transmitter, then there is no opportunity to optimize the data ($\mathbf{R}_{xx}$) before transmission. For this reason it is intuitive to allocate equal power to all of the transmit antennas. The capacity of the channel when CSI is not available at the transmitter is

$$C = B\log_2\left(\det\left(\sigma_{nn}^{-2}\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{I}\right)\right) \tag{2.9}$$

By taking the eigenvalue decomposition (EVD) of $\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$, (2.9) may be rewritten as

$$C = B\log_2\det\left(\sigma_{nn}^{-2}\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H + \mathbf{I}\right), \tag{2.10}$$

where $\mathbf{Q}$ is a matrix of eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues. Using the identity $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$ and the property of eigenvectors that $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$ allows (2.10) to reduce to

$$\begin{aligned} C &= B\log_2\det\left(\sigma_{nn}^{-2}\mathbf{\Lambda} + \mathbf{I}\right) \\ &= B\sum_{i=1}^{r}\log_2\left(\sigma_{nn}^{-2}\lambda_i + 1\right), \end{aligned} \tag{2.11}$$

where r indicates the rank of the channel, which implies that $r \leq min(M_r, M_t)$. From (2.11), it is shown that there are r subchannels between the transmitter and the receiver [1]. This shows that in the absence of CSI the capacity approaches

$$C = Br\log_2\left(\sigma_{nn}^{-2}\lambda + 1\right) \tag{2.12}$$

and grows linearly in r. This reveals the appeal of a MIMO system, which shows that without CSI at the transmitter one is able to increase linearly the capacity simply by adding antennas at either end of the link [15].

CSI Available at the Transmitter

If the transmitter is provided with CSI, then it is intuitive to assume that the capacity would increase because the transmitter is able to maximize the $\mathbf{R}_{xx}$ matrix as described in (2.8). This allows access to the subchannels of the channel in order to maximize their effectiveness through the link using linear processing.

To access the subchannel the singular value decomposition (SVD) is used. The SVD is defined as

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{2.13}$$

where $\mathbf{U}$, $\mathbf{\Sigma}$ and $\mathbf{V}$ are matrices of dimension $M_r \times M_r$, $M_r \times M_t$ and $M_t \times M_t$, respectively. $\mathbf{U}$ and $\mathbf{V}$ are composed of the left and right singular vectors, respectively, and $\mathbf{\Sigma}$ is a diagonal matrix of singular values. The singular values are positive numbers ordered such that

$$\sigma_1 \geq \sigma_2 \geq \ldots \sigma_r.$$

The singular values also have the property that $\mathbf{\Sigma}^2 = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues. When performing the SVD on a wide ($M_r < M_t$) matrix, the last $M_t - M_r$ columns of $\mathbf{\Sigma}$ are full of zeros. In the $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ product, these columns of zeros cause the last $M_t - M_r$ columns of $\mathbf{V}$ to be of no importance because they do not alter the $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ product. Similarly if the SVD is performed on a tall matrix, the last $M_r - M_t$ rows of $\mathbf{\Sigma}$ will allow the $M_r - M_t$ columns of $\mathbf{U}$ to be eliminated because they are useless in forming $\mathbf{H}$ from the product $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. The $M_t - M_r$ columns of $\mathbf{V}$ and $\mathbf{\Sigma}$ (wide) and the $M_r - M_t$ columns of $\mathbf{U}$ and rows of $\mathbf{\Sigma}$ (tall) are not needed in the product $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. This is irrelevant information and the integrity of the product $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ is in no way sacrificed. The elimination of the columns and rows of $\mathbf{U}$, $\mathbf{\Sigma}$ or $\mathbf{V}$ will not be discussed further until chapter 3.

By multiplying the transmitted symbols $\mathbf{s}$, prior to transmission with the right singular vectors, the transmitted signal becomes

$$\mathbf{x} = \mathbf{V}\mathbf{s}. \tag{2.14}$$

The transmitted signal corrupted by AWGN prior to reception is

$$\mathbf{y} = \mathbf{H}\mathbf{V}\mathbf{s} + \mathbf{n}. \tag{2.15}$$

Linear processing also takes place at the receiver as the Hermitian of the left singular vectors is left multiplied with (2.15) resulting in the following equation and simplification

$$\begin{aligned} \mathbf{U}^H\mathbf{z} &= \mathbf{U}^H\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{V}\mathbf{s} + \mathbf{U}^H\mathbf{n} \\ \mathbf{z}' &= \mathbf{\Sigma}\mathbf{s} + \mathbf{n}', \end{aligned}$$

which is shown block diagram form in Fig. 2.2. This allows the explicit decomposition of the channel into r parallel SISO channels which are written as follows

$$z_i' = \sigma_i s_i + n_i',$$

where $i = 1 \ldots r$.

The capacity of the channel with the CSI known at the transmitter is the sum of capacities of the r parallel SISO channels and is defined as

$$C = \sum_{i=1}^{r} \log_2 \left(\sigma_{nn}^{-2} \lambda_i \rho_i + 1 \right), \tag{2.16}$$

where $\rho_i = E[\mathbf{s}_i^H \mathbf{s}_i]$ and $\sigma_i^2 = \lambda_i$. This is an improvement upon (2.11) because the user is able to allocate more or less energy in the subchannels, instead of equally allocating energy in all subchannels. This results in a decrease of transmit power consumption.

The objective now becomes the maximization of the energy across the subchannels with respect to ρ, in order to distribute energy to the weaker subchannels [1]. The optimal method of allocating energy is known as the Waterpouring algorithm and is presented in [19] and [20].

2.3 Precoding for MIMO

The ability to provide the transmitter with full or partial CSI allows the exploitation of that knowledge to improve the performance of the system. The transmit schemes implemented at the transmitter depend on factors such as the nature of the channel knowledge, the type of receiver, the performance criterion to be optimized, and the power constraints at the transmitter. Herein the precoding schemes will be grouped according to how much information is provided at the transmitter, i.e. full, partial, or limited.

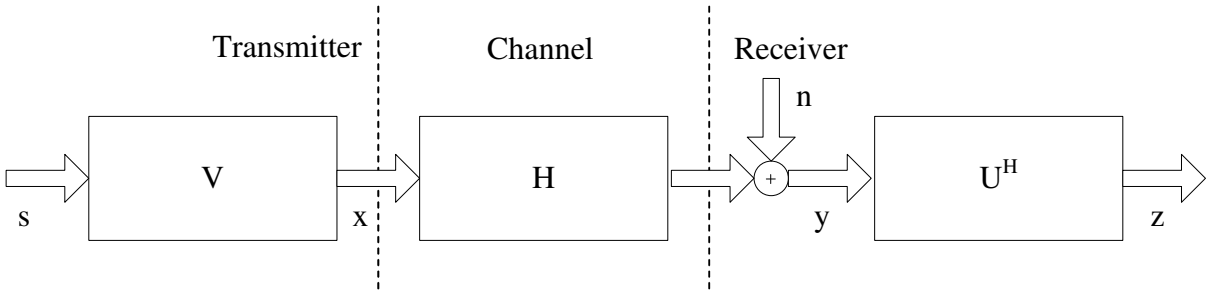


Figure 2.2: Decomposition of the wireless channel.

2.3.1 Full CSI

Using full CSI to precode the data is an ideal assumption but not practical for implementation in the real world. The quantity of information sent to the transmitter on a zero delay feedback channel is unrealistic. Exploitation of the eigenstructure is a common practice in constructing the optimal **F** (precoder) and **G** (decoder). The designs of [21] target the minimization of the MSE of the decoded block of symbols and are further analyzed in Chapter 3. Another method of optimization of a generalized jointly optimum linear precoder/decoder design attempts to minimize the weighted sum of symbol estimation errors with the assumption of constrained transmit power across all antennas [22]. Other research makes use of the assumption of full CSI to obtain optimal linear precoders/decoders, but optimizes the system to a particular performance criterion. Specifically, the work of [23] proposes a minimum bit error rate (MBER) diagonal precoder, whereas [24] maximizes the minimum Euclidean distance of the received constellation under a power constraint.

2.3.2 Partial CSI

Partial CSI makes use of the long term statistics to precode the transmitted signal. Herein partial CSI is referred to as either knowledge of the transmit and receive antenna fading correlations, transmit antenna fading correlations or the channel covariance matrix, all of which allow transmission along the eigenmodes of the channel.

Transmit antenna correlation is caused by the lack of multipath fading at the transmitter. A downlink scenario where the BS is placed high above the ground and sees no local scatterers is an example of transmitter correlation, as previously shown in Fig. 2.1. The lack of multipath fading

which introduces correlation at the transmitter causes the columns of the **H** matrix to be correlated. This MIMO channel may be written as

$$\mathbf{H} = \mathbf{H}_w \mathbf{R}_T^{1/2} \ ,$$

where $\mathbf{R}_T^{1/2}$ is a $M_t \times M_t$ transmit antenna correlation matrix and $\mathbf{H}_w$ is a $M_r \times M_t$ complex matrix with entries that are IID. An optimal linear precoder found through an optimization method may be used, as well as a maximum likelihood (ML) decoder [11].

Similar to [11], [10], [13] and [25] make use of the correlations at the transmitter, but also those at the receiver. The spatial correlations at the transmitter and receiver alter the channel matrix such that it becomes

$$\mathbf{H} = \mathbf{R}_R^{H/2} \mathbf{H}_w \mathbf{R}_T^{1/2},$$

where $\mathbf{R}_R$ and $\mathbf{R}_T$ are the $M_r \times M_r$ and $M_t \times M_t$ receiver and transmitter correlations, respectively.

If the transmit antennas are too correlated, then a beamforming model is derived which makes use of the angle of departure α^T, angle of arrival α^R, and the complex channel amplitude of the Rayleigh fading channel. The channel model which implements beamforming is described by

$$\mathbf{H} = \mathbf{A}\left(\alpha^R\right) \operatorname{diag}\left(\beta\right) \mathbf{A}\left(\alpha^{\mathrm{T}}\right)^{\mathrm{T}},$$

where **A** denotes a steering matrix and β contains the complex amplitudes.

Regardless of the amount of correlation at the antennas the second order fading statistics are used at the transmitter side to create an optimal precoder. The receiver side has the full CSI which enables the use of the minimum MSE criterion to create the optimal decoder. More on this topic may be found in [10], [13] and [25].

Other literature on partial CSI makes use of the input signal covariance in order to calculate the mutual information and optimize the system with respect to the covariance, which is done in [26]. At the same time the authors make use of the transmit antenna correlation in order to determine if beamforming is optimal due to the current state of the transmit antennas. If the transmit antennas are highly correlated, then it is beneficial to use beamforming since it requires using the transmit antenna array to send a single effective stream of data. Otherwise, if the transmit antenna array has

low correlation, this implies that the antennas are independent and it is advantageous to make use of the gains offered by a MIMO system.

2.3.3 Limited CSI

Limited CSI is the process of quantizing the CSI at the receiver and sending a codebook index to the transmitter in order to compute the optimal precoder. Quantization is the process by which an index indicates a specific matrix located in a codebook, where both receiver and transmitter have copies of the same codebook. Upon computation of the ideal precoder at the receiver, the minimum distance between the subspaces spanned by the precoding matrix and the ideal precoders in the codebook, a match is made. By sending only the index of the matched precoder to the receiver, less information is needed to be sent in the feedback loop at the cost of the size of the codebook. The larger the codebook, the more minimum distances must be computed, which increases computation time, which delays the arrival of CSI at the transmitter. For more information on codebook design and vector quantization, refer to [27], [28] and [29].

In the case of [30], [31], and [32], the precoder is restricted to having orthonormal columns and the received signal is decoded with a linear decoder. If a zero forcing or a minimum MSE receiver is used, the decoders are

$$\begin{aligned} \mathbf{G}_{ZF} &= (\mathbf{HF})^{*} \\ \mathbf{G}_{MMSE} &= \left(\mathbf{F}^{H}\mathbf{H}^{H}\mathbf{HF} + \sigma_{nn}^{-2}\mathbf{I}\right)^{-1}\mathbf{F}^{H}\mathbf{H}^{H}, \end{aligned}$$

where * indicates the Moore-Penrose pseudoinverse. The precoder is chosen at the receiver from a codebook that has $N = 2^{D}$ $\mathbf{F}$ matrices to chose from. The size of the codebook is determined by the number of feedback bits used, which in this case is D. The codebook is such that the metric computes a chordal distance and selects a matrix $\mathbf{F}$ from $\mathcal{F} = (\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_N)$.

The chordal distance metric minimizes the distortion metric

$$E[\min_{(\mathrm{i}\in 1,2,\cdots,\mathrm{N})} \frac{1}{2}||\mathbf{V}\mathbf{V}^{\mathrm{H}} - \mathbf{F}_{\mathrm{i}}\mathbf{F}_{\mathrm{i}}^{\mathrm{H}}||_{\mathrm{F}}^{2}], \tag{2.17}$$

where $\mathbf{V}$ is from the EVD of $\mathbf{H}$ and the subscript F indicates the Frobenius norm. The chordal distance is a measure of the distance between two subspaces. The literature which makes use

of the chordal distance measurements within the framework of limited feedback MIMO precoders/decoders ([30], [31], and [32]) also make use of Grassmannian packing in order to compute low distortion codebooks. More information on Grassmannian packing may be found by referencing [33] and [34].

Other limited feedback literature simply sends back to the transmitter a quantized estimate of the channel. The estimate of the channel is then used in order to compute what is referred to as a weighing matrix, $\mathbf{W}$. The weighing matrix is simply a precoder, which serves to precode the space time block codes (STBC) prior to transmission [35]. In some cases the precoding matrix is simply the eigenvectors of $\mathbf{H}^H\mathbf{H}$ [36], or the eigenvectors multiplied with a diagonal waterpouring matrix [37].

2.3.4 Imperfect CSI

Despite the many methods of informing the transmitter of the CSI known at the receiver, full, partial and limited CSI methods are all capable of being corrupted by noise leading to imperfect CSI. The corrupted CSI analysis that is common in the literature is analysis of errors that occur in the process of feeding back the CSI to the transmitter. These errors may be categorized but are not limited to, time delay errors for the information to be transmitted from the receiver to the transmitter, quantization errors in the feedback channel (if limited feedback is used) and channel estimation errors.

Some literature (such as [38], [39], [40], [41] and [42]) deals with errors that may occur during the estimation of the channel. In order to estimate the channel a training sequence may be used which is subject to corruption. This thesis will deal with a description and derivation of equations which encompass all forms of errors that befall CSI, so it is not limited to the analysis of one particular type. This section is included to depict common approaches such as time delay errors, quantization errors and CSI errors.

In order to approach the error analysis associated with imperfect CSI, it is common to include an error term which encompasses all of the imperfections of the CSI. Given a system with block diagonal structure, such as that of [43], which is applicable to a MIMO Orthogonal Frequency

Division Multiplexing (OFDM) system, the received $KM \times 1$ signal is modeled as

$$\mathbf{z} = \mathbf{GHFs} + \mathbf{Gn}, \tag{2.18}$$

where the channel **H** is a frequency selective $KM_r \times KM_t$ block diagonal channel matrix. **F** is a $KM_t \times KM$ precoder matrix which allocates the power across the K subcarriers and M_t antennas. $M \leq \min(M_t, M_r)$ is the number of symbols to be transmitted per subcarrier. **G** is a $KM \times KM_r$ decoder matrix and **n** is a $KM_r \times 1$ noise vector.

Due to the block diagonal structure of the received signal model of (2.18), the MIMO channel may be decomposed into K MIMO frequency flat fading channels by imposing a block diagonal structure on the matrices and vectors of (2.18) such that

$$\begin{aligned}
\mathbf{H} &= \operatorname{diag}\left([\mathbf{H}_1\mathbf{H}_2\cdots\mathbf{H}_\mathrm{K}]\right) \\
\mathbf{F} &= \operatorname{diag}\left([\mathbf{F}_1\mathbf{F}_2\cdots\mathbf{F}_\mathrm{K}]\right) \\
\mathbf{G} &= \operatorname{diag}\left([\mathbf{G}_1\mathbf{G}_2\cdots\mathbf{G}_\mathrm{K}]\right) \\
\mathbf{s}^T &= \left[\mathbf{s}_1^T\cdots\mathbf{s}_K^T\right] \\
\mathbf{z}^T &= \left[\mathbf{z}_1^T\cdots\mathbf{z}_K^T\right].
\end{aligned}$$

This result enables rewriting (2.18) on the subcarrier level as

$$\mathbf{z}_k = \mathbf{G}_k\mathbf{H}_k\mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{n}_k \tag{2.19}$$

where $k = 1 \ldots K$. By storing the MIMO channel response for the k*th* subcarrier in a vector

$$\mathbf{h}_k = \operatorname{vec}\left[\mathbf{H}_\mathrm{k}\right]$$

and making use of the identity [44]

$$\operatorname{vec}(\mathbf{ABC}) = \left(\mathbf{C}^\mathrm{T} \otimes \mathbf{A}\right)\operatorname{vec}(\mathbf{B}),$$

(2.19) may be rewritten as

$$\mathbf{z}_k = \mathbf{A}_k\mathbf{h}_k + \mathbf{G}_k\mathbf{n}_k$$

where $k = 1 \ldots K$ and $\mathbf{A}_k = (\mathbf{F}_k\mathbf{s}_k)^T \otimes \mathbf{G}_k$.

This enables incorporating the channel error $\mathbf{\Delta h}$ into the the channel estimation as

$$\widetilde{\mathbf{H}} = \sqrt{\frac{\sigma_h^2}{\sigma_h^2 + \sigma_{\Delta h}^2}} (\mathbf{h} + \mathbf{\Delta h}) \tag{2.20}$$

where $\sigma_h^2 = E\left[\mathbf{h}^H \mathbf{h}\right]$ and $\sigma_{\Delta h}^2 = E\left[\mathbf{\Delta h}^H \mathbf{\Delta h}\right]$. This error term encompasses the feedback delay and the CSI errors due to additive noise. The work of [43] derives (2.18), (2.19) and (2.20) and also assumes that the channel $\mathbf{H}$ and the channel error $\widetilde{\mathbf{H}}$ are jointly Gaussian in order to design the optimal linear precoder and decoder which are minimized according to the MMSE and the minimization of the bit error rate (BER). In the design process according to the MMSE and BER criterion a cost function is derived, followed by a closed form solution and a description of the asymptotic performance of the linear precoder/decoder developed. Models of the MMSE precoder/decoder are not developed in [43] as they are in Chapter 4 for the MMSE precoder/decoder.

The mismatch between the CSI at the receiver and the CSI at the transmitter is analyzed for broadcast Orthogonal Space Division Multiple Access (OSDMA) as well. One unique aspect of this work is that it focuses on the individual error sources instead of grouping all of the errors into one term, which is done in [45]. The separate development of the impact of imperfect CSI on system performance parameters such as ML channel estimation, the effects of channel estimation errors and the time delay effects of channel estimation are individually considered.

Other literature reveals the development of optimal power adaptation and adaptive modulation for MIMO systems with imperfect CSI [46]. The system model developed used the same precoder/decoder originally designed and developed by [20], based on the SVD of the CSI.

It is common among literature dealing with imperfect CSI at the transmitter and/or receiver to introduce an error term. After the introduction of the error term and a system model, development of performance analysis of SNR and bit error probability [47], or BER analysis and quantization effects [48] or the trade-offs in transmission strategies of space time coding and beamforming and the effects on the information transfer rate [12] may be found.

2.4 Conclusion

This chapter introduces the concept of MIMO and the different forms of precoding. Section 2.3.4 shows that there has been analysis into the more realistic scenarios of imperfect CSI. Despite the research completed in the literature, none develop equations which quantify the amount of error introduced to the system, as will be shown in Chapter 4. Next, Chapter 3 will show the development of the MMSE precoder/decoder designs of [21].

Chapter 3

Design of Optimal Linear Precoders/Decoders

This chapter delves into the design aspect of linear precoders/decoders within a MIMO system. One nice aftereffect of some MIMO precoders is their ability to separate the subchannels into individual SISO channels, which if accomplished, enables streamlined analysis. Through the use of full CSI the optimal designs of the linear precoders/decoders presented in [21] accomplish the isolation of the individual subchannels as well as present closed form solutions to various performance measures. However, partial CSI provides a more realistic approach since providing the transmitter with complete channel knowledge instantaneously is impossible. Performance measures are provided herein for full CSI. This chapter will address the design and performance measures of the above precoders/decoders.

3.1 Optimal Linear Precoders/Decoders using Full CSI

By beginning the design with the minimum MSE, an initial problem occurs on whether to design a decoder tuned to the precoder and channel, or to design a precoder tuned to the channel and decoder. One logical option is to pick the former choice, and design an optimal decoder for a given precoder and channel since the receiver will have complete CSI. Starting from this point a

decoder design is created using the minimum MSE and three different design criterion are created from which six optimization scenarios emerge in order to design multiple precoders. Performance measures are then derived for the precoder/decoder designs.

3.2 System Model

The overall system model has M_r receive antennas and M_t transmit antennas. Every T symbol intervals there are N symbols, denoted as $\mathbf{s}_i$, transmitted by a $N \times 1$ vector. The system begins by precoding the $N \times 1$ vector with $\mathbf{F}$ in order to produce a $M_t \times 1$ transmit vector

$$\mathbf{x}_i = \mathbf{F}\mathbf{s}_i.$$

It is assumed that the discrete time channel has a time varying impulse response, is causal and has finite memory L, which enables writing the input-output equations in block FIR (finite impulse response) form. Using a block by block transmission structure the transmitted and received symbols are stacked into snapshots. A snapshot is one small block of data that is transmitted over one symbol period which is then stacked into large blocks before and after transmission in order to attempt to cancel interblock and intersymbol interference (IBI and ISI, respectively) which are caused by the frequency selectivity of the channel.

The block transmission begins by stacking $P = M + L$ precoded transmit snapshots (M is the number of receive snapshots) into a vector $\mathbf{x}_i$. The structure of the $M_t P \times 1$ transmitted vector $\mathbf{x}_i$ is

$$\mathbf{x}_i = \begin{bmatrix} \mathbf{x}[iP] \\ \vdots \\ \mathbf{x}[(i+1)\,P-1] \end{bmatrix},$$

which is the product of the $M_t P \times N$ precoder and the symbols $\mathbf{s}_i$. The output of the channel prior to corruption by AWGN is

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i.$$

$\mathbf{y}_i$ is the product of the channel and the stacked transmission vector, where $\mathbf{H}$ is the $M_r M \times M_t P$

channel matrix. Similar to $\mathbf{x}_i$, $\mathbf{y}_i$ has the form

$$\mathbf{y}_i = \begin{bmatrix} \mathbf{y}[iP] \\ \vdots \\ \mathbf{y}[(i+1)\,P-1] \end{bmatrix}.$$

The received $M_r M \times 1$ vector $\mathbf{y}$ is corrupted by AWGN with variance σ_{nn}^2. The received signal is $\mathbf{z}_i = \mathbf{y}_i + \mathbf{n}_i$. The received signal is then decoded by a $N \times M_r M$ decoder matrix $\mathbf{G}$. This results in a $N \times 1$ vector of estimated symbols in the receiver which are expressed as

$$\hat{\mathbf{s}}_i = \mathbf{G}\mathbf{z}_i = \mathbf{GHF}\mathbf{s}_i + \mathbf{G}\mathbf{n}_i. \tag{3.1}$$

The transmit and receive snapshots are included for the frequency selective case. If the channel exhibits flat fading, then $P = M = 1$ and $L = 0$ and the above equations collapse to simpler forms, which will be dealt with throughout this thesis since frequency flat fading is assumed.

The precoder/decoder block diagram is shown in Fig. 3.1. The optimal design of the decoder begins with the minimization of the mean square error (MSE). The MSE is a function of both $\mathbf{F}$ and $\mathbf{G}$ and is given by

$$\mathrm{MSE}\left(\mathbf{F},\mathbf{G}\right) \;=\; E\left\{\left(\hat{\mathbf{s}}_i - \mathbf{s}_i\right)\left(\hat{\mathbf{s}}_i - \mathbf{s}_i\right)^H\right\}. \tag{3.2}$$

By substituting in the received signal (3.1) into (3.2), the following is obtained

$$\mathrm{MSE}\left(\mathbf{F},\mathbf{G}\right) \;=\; \left(\mathbf{GHF} - \mathbf{I}\right)\mathbf{R}_{ss}\left(\mathbf{GHF} - \mathbf{I}\right)^H + \mathbf{G}\mathbf{R}_{nn}\mathbf{G}^H, \tag{3.3}$$

where $\mathbf{R}_{ss} = \sigma_{ss}^2\mathbf{I}$ because the transmit symbols are assumed to be white. The noise covariance matrix $\mathbf{R}_{nn} = \sigma_{nn}^2\mathbf{I}$ is positive definite, and the noise $\mathbf{n}_i$ is uncorrelated with the transmitted symbols $\mathbf{s}_i$.

The optimal decoder is found by designing a Weiner receiver [49] which minimizes the $tr\left(\mathrm{MSE}\left(\mathbf{F},\mathbf{G}\right)\right)$, and is given by

$$\mathbf{G}_{opt} = \mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H\left(\mathbf{HF}\mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H + \mathbf{R}_{nn}\right)^{-1}. \tag{3.4}$$

By substituting in the optimal $\mathbf{G}$ of (3.4) into the MSE equation (3.3), the minimum MSE($\mathbf{F},\mathbf{G}$) becomes MSE($\mathbf{F},\mathbf{G}_{opt}$). The resulting MSE($\mathbf{F}$) is minimum in the sense that

$$\mathrm{MSE}\left(\mathbf{F}\right) = \mathrm{MSE}\left(\mathbf{F},\mathbf{G}_{\mathrm{opt}}\right) = \sigma_{\mathrm{ss}}^2\left(\mathbf{I} + \sigma_{\mathrm{ss}}^2\mathbf{F}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{R}_{\mathrm{nn}}^{-1}\mathbf{HF}\right)^{-1}. \tag{3.5}$$

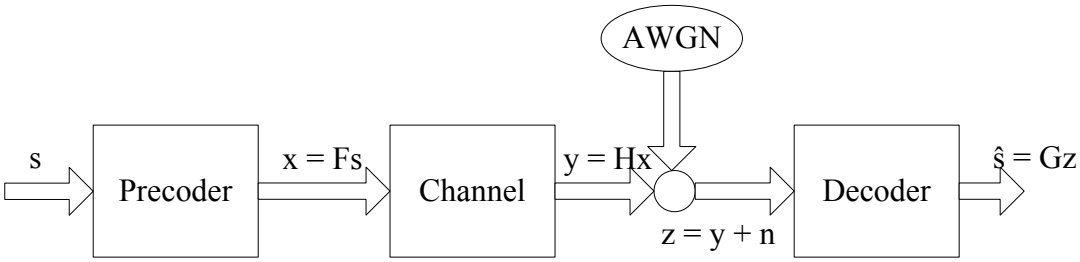


Figure 3.1: A block diagram of the system model.

Since the transmit symbols are white, the previous equation for **G**, (3.4) simplifies to

$$\mathbf{G} = \mathbf{F}^H\mathbf{H}^H\left(\mathbf{HFF}^H\mathbf{H}^H + \mathbf{R}_{nn}\sigma_{ss}^{-2}\right)^{-1}. \tag{3.6}$$

The result of (3.6) requires that the receiver has knowledge of the channel and the noise covariance matrix. If the receiver does not have knowledge of the two matrices then decoder is useless. The most important assumption of the optimal linear precoder/decoder design due to full CSI is that both the receiver and transmitter have complete and perfect channel knowledge.

The optimal designs of the optimal precoder may be found by placing different performance measures which depend on the MSE of **F** (3.5). A trivial solution would be to increase the norm of **F** to infinity, but that is not reasonable. The first constraint limits the transmit power and is found by limiting the expected value of the norm of the transmit vector $E\left[||\mathbf{x}_i||^2\right] = tr\left(\mathbf{FF}^H\right)\sigma_{ss}^2$ which results in

$$tr\left(\mathbf{FF}^H\right)\sigma_{ss}^2 = P_0. \tag{3.7}$$

The second constraint limits the maximum eigenvalue of the transmit covariance matrix. It is another method of limiting the power, and is given by

$$\lambda_{max}\left(\mathbf{FF}^H\right)\sigma_{ss}^2 = L_0. \tag{3.8}$$

The third and final constraint limits the peak power of the transmit symbols and is shown by the following inequality

$$\max_{i,k}\left(||\left[\mathbf{Fs}_i\right]_k||^2\right) \leq \lambda_{max}\left(\mathbf{F}^H\mathbf{F}\right) max_i\left(||\mathbf{s}_i||^2\right). \tag{3.9}$$

The constraint comes from $||\mathbf{s}_i||^2$ since the transmit symbols are bounded in amplitude. This constraint limits signal peak and is constellation independent.

In order to find the optimal **F**, the minimum value of (3.5) must be found. The trace operator and the determinant are two avenues in which the minimum may be found. Other methods of finding minimum values of a matrix are through Lagrange multipliers, the Frobenius, two, or infinity norms, and by minimizing the eigenvalues [50] [51] [52]. In this case, the trace and determinant are used to compute the minimum argument of (3.5), with respect to the three constraints (3.7), (3.8) and (3.9), resulting in six optimal **F** matrices. These six designs have different performance results which may be used to suit the users needs. Obviously, the **F**'s obtained are by no means an all encompassing list of optimal precoder designs.

3.3 Considerations for Non-Square Channels

No reservations have been made yet as to the size of the channel, which is determined by the number of transmit and receive antennas. If the number of receive antennas is less than the number of transmit antennas ($M_r < M_t$), then the channel matrix is wide. If the number of receive antennas is greater than the number of transmit antennas ($M_r > M_t$), then the channel matrix is tall. A by-product of the SVD of the channel matrix **H**, is that for wide or tall matrices some of the matrices contain irrelevant information. In the case of a wide channel matrix, the $M_t - M_r$ columns of $\mathbf{\Sigma}$ and **V** may be eliminated. For a tall channel matrix, the $M_r - M_t$ rows of $\mathbf{\Sigma}$ and the $M_r - M_t$ columns of **U** may be eliminated because those rows and columns are irrelevant to the computation of **H**, as was discussed in Chapter 2

Here is a small example to help in understanding the elimination of irrelevant rows. Recall that the SVD of the channel was defined in (2.13). Recognizing that with a 3×7 **H** matrix, the other matrices of (2.13) are $\mathbf{U} = 3 \times 3$, $\mathbf{\Sigma} = 3 \times 7$, $\mathbf{V}^H = 7 \times 7$, and $rank\,(\mathbf{H}) = rank\,(\mathbf{\Sigma}) = r = 3$.

$$\underbrace{\begin{bmatrix} \mathbf{H} \end{bmatrix}}_{3\times7} = \underbrace{\begin{bmatrix} \mathbf{U} \end{bmatrix}}_{3\times3} \underbrace{\begin{bmatrix} \mathbf{\Sigma} \end{bmatrix}}_{3\times7} \underbrace{\begin{bmatrix} \mathbf{V} \end{bmatrix}^H}_{7\times7}.$$

The previous example shows that in order to fully represent the channel, only 3 dimensions are

necessary. Additionally, the previous result shows only the case of a wide channel matrix, but the same may be shown for a tall channel matrix but is omitted here.

All of the previous results assume a full rank matrix. For the case of a non-full rank channel matrix the necessary information to fully represent **H** reduces to r, where r is the rank of **H** and the following inequality applies $r \leq min\,(M_r, M_t)$.

The above arguments concerning the size and rank of **H** may be restated in terms of the EVD instead of the SVD. However the end result is the same. The EVD of the CSI is presented in [21] as follows

$$\mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H.$$

If further clarification is needed, please refer to [21] for more information.

3.4 Specific Precoder Designs

The following precoders are presented without derivation. For specific details, please refer to the appendices of [21]. Every design results in a precoder **F** to be of the form

$$\mathbf{F} = \mathbf{V}\mathbf{\Phi}, \tag{3.10}$$

where **V** is the same in each, but a new $\mathbf{\Phi}$ is derived. **V** is a matrix containing the right singular vectors of the SVD of the full CSI, and $\mathbf{\Phi}$ is a square, diagonal power allocation matrix.

The first design is obtained by minimizing the trace of (3.5) with respect to the total transmit power constraint given in (3.7). The entries along the diagonal are denoted as ϕ_{ii} and the first design's entries are [21]

$$|\phi_{ii}|^2 = \left(\frac{P_0 + \sum_{k=1}^{M_r} \lambda_{nn}^{-1}}{\sigma_{ss}^2 \sum_{k=1}^{M_r} \lambda_{nn}^{-1/2}} \lambda_{ii}^{-1/2} - \frac{1}{\lambda_{ii} \sigma_{ss}^2} \right)^+ \tag{3.11}$$

where $(.)^+$ is defined as the $max\,(x, 0)$ and the λ_{ii} term denotes the eigenvalues of **H**.

Minimizing the determinant of (3.5) with respect to the total power constraint of (3.7), yields a new $\mathbf{\Phi}$ which is found as [21]

$$|\phi_{ii}|^2 = \left(\frac{P_0 + \sum_{k=1}^{M_r} \lambda_{kk}^{-1}}{M_r \sigma_{ss}^2} - \frac{1}{\lambda_{ii} \sigma_{ss}^2} \right)^+ . \tag{3.12}$$

Moving on to the minimization of the trace of (3.5) with respect to the maximum eigenvalue constraint in (3.8) yields a $\mathbf{\Phi}$ of the form [21]

$$\phi_{ii} = \sqrt{\frac{L_0}{\sigma_{ss}^2}} \ . \tag{3.13}$$

The minimization of the determinant of (3.5) with respect to the maximum eigenvalue constraint in (3.8) yields the same $\mathbf{\Phi}$ as shown in (3.13).

The final two $\mathbf{\Phi}$'s are constrained by the peak power of the transmit symbols (3.9), and are given by [21]

$$|\phi_{ii}|^2 = \frac{P_0}{\sigma_{ss}^2 \sum_k \lambda_{kk}^{-1}} \lambda_{ii}^{-1}, \tag{3.14}$$

and

$$|\phi_{ii}|^2 = \frac{L_0 \lambda_{NN}}{\sigma_{ss}^2} \lambda_{ii}^{-1}. \tag{3.15}$$

These different precoder matrices may be used to suit the optimization needs of the user.

3.5 Equivalent Channel Model

The optimal form of $\mathbf{G}$ shown in (3.4) may be rearranged in order to show that the equivalent channel seen by the data becomes an identity. Beginning with the manipulation of (3.4) is as follows:

$$\begin{aligned} \mathbf{G} &= \mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H \left(\mathbf{H}\mathbf{F}\mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H + \mathbf{R}_{nn}\right)^{-1} \\ &= \sigma_{ss}^2\mathbf{F}^H\mathbf{H}^H \left(\sigma_{ss}\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H\sigma_{ss} + \mathbf{R}_{nn}^{\frac{1}{2}}\mathbf{R}_{nn}^{\frac{1}{2}}\right)^{-1} \\ &= \sigma_{ss}\mathbf{R}_{nn}^{-\frac{1}{2}}\mathbf{F}^H\mathbf{H}^H \left(\sigma_{ss}\mathbf{R}_{nn}^{-\frac{1}{2}}\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H\mathbf{R}_{nn}^{-\frac{1}{2}}\sigma_{ss} + \mathbf{I}\right)^{-1} \mathbf{R}_{nn}^{-\frac{1}{2}}\sigma_{ss}, \end{aligned} \tag{3.16}$$

and by setting $\mathbf{A} = \sigma_{ss}\mathbf{R}_{nn}^{-\frac{1}{2}}\mathbf{H}\mathbf{F} = \mathbf{\Sigma}\mathbf{V}^H\mathbf{V}\mathbf{\Phi} = \mathbf{\Sigma}\mathbf{\Phi}$, the above $\mathbf{G}$ may be written as

$$\mathbf{G} = \mathbf{A}^H \left(\mathbf{A}\mathbf{A}^H + \mathbf{I}\right)^{-1} \mathbf{R}_{nn}^{-\frac{1}{2}}\sigma_{ss}.$$

Now by making use of the matrix inversion lemma $\mathbf{A}^H \left(\mathbf{A}\mathbf{A}^H + \mathbf{I}\right)^{-1} = \left(\mathbf{A}^H\mathbf{A} + \mathbf{I}\right)^{-1} \mathbf{A}^H$, and by setting $\mathbf{\Gamma}' = \left(\mathbf{A}^H\mathbf{A} + \mathbf{I}\right)^{-1} \sigma_{ss}^2$, $\mathbf{G}$ may be rewritten as

$$\begin{aligned}
\mathbf{G} &= \mathbf{\Gamma}'\mathbf{A}^H\mathbf{R}_{nn}^{-\frac{1}{2}}\sigma_{ss}^{-1} \\
&= \mathbf{\Gamma}'\left(\sigma_{ss}\mathbf{R}_{nn}^{-\frac{1}{2}}\mathbf{F}^H\mathbf{H}^H\right)\mathbf{R}_{nn}^{-\frac{1}{2}}\sigma_{ss}^{-1} \\
&= \mathbf{\Gamma}'\mathbf{F}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1} \\
&= \mathbf{\Gamma}'\left(\mathbf{V}\mathbf{\Phi}\right)^H\mathbf{H}^H\mathbf{R}_{nn}^{-1} \\
&= \mathbf{\Gamma}'\mathbf{\Phi}^H\mathbf{V}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}.
\end{aligned}$$

By substituting in $\mathbf{\Gamma} = \mathbf{\Gamma}'\mathbf{\Phi}^H\mathbf{\Lambda}$ the final form of $\mathbf{G}$ is derived

$$\mathbf{G} = \mathbf{\Gamma}\mathbf{\Lambda}^{-1}\mathbf{V}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}. \tag{3.17}$$

It may be inferred from above that $\mathbf{A}$ is a diagonal matrix since it may be written as the product of two diagonal matrices, $\mathbf{A} = \mathbf{\Sigma}\mathbf{\Phi}$. From this $\mathbf{\Gamma}'$ may also be shown to be a diagonal since $\mathbf{\Gamma}' = \left(\mathbf{A}^H\mathbf{A} + \mathbf{I}\right)^{-1} \sigma_{ss}^2$. Furthermore, it may be concluded that $\mathbf{\Gamma}$ is a diagonal matrix since it is the product of three diagonal matrices, $\mathbf{\Gamma} = \mathbf{\Gamma}'\mathbf{\Phi}^H\mathbf{\Lambda}$.

In the computation of the equivalent channel model, if the diagonal matrices $\mathbf{\Phi}$ and $\mathbf{\Gamma}$ are excluded, the overall channel reduces to an $M_r \times M_r$ identity as shown below,

$$\begin{aligned}
\mathbf{\Psi} &= \mathbf{GHF} \\
&= \mathbf{\Lambda}^{-1}\mathbf{V}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{HV} \\
&= \mathbf{\Lambda}^{-1}\mathbf{V}^H[\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H]\mathbf{V} \\
&= \mathbf{I}_{M_r}.
\end{aligned} \tag{3.18}$$

The subscript M_r indicates the dimension of the matrix. If the channel is square, then the size of $\mathbf{\Psi}$ is M_r or M_t. If the channel is wide or tall, then the size of $\mathbf{\Psi}$ reduces to M_r or M_t, respectively. This is significant because it indicates the number of parallel and independent subchannels of the channel that exist in which the data symbols pass through. $\mathbf{\Psi}$ is the effective channel that the symbols see, excluding the diagonal matrices $\mathbf{\Phi}$ and $\mathbf{\Gamma}$. Since the overall channel model reduces to an identity, the gain of the system is $\mathbf{\Gamma}\mathbf{\Phi}$ as shown in Fig. 3.2.

The noise correlation reduces to $\mathbf{\Lambda}^{-1}$ as shown below

$$\begin{aligned}\mathbf{Gnn}^H\mathbf{G}^H &= \mathbf{\Lambda}^{-1}\mathbf{V}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{R}_{nn}\mathbf{R}_{nn}^{-1}\mathbf{H}\mathbf{V}\mathbf{\Lambda}^{-1} \\ &= \mathbf{\Lambda}^{-1}\mathbf{V}^H\left(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H\right)\mathbf{V}\mathbf{\Lambda}^{-1} \\ &= \mathbf{\Lambda}^{-1}.\end{aligned}$$

3.6 Performance Results

The results in the previous section allow a streamlined performance analysis. One nice result of the overall channel reducing to an identity, as shown in equation (3.18). Another is that the system gain is determined by two diagonal matrices $\mathbf{\Gamma}$ and $\mathbf{\Phi}$. Additionally, the MIMO channel reduces to r parallel, independent and ISI free subchannels. Each subchannel has the gain $\phi_{kk}\gamma_{kk}$ and noise variance λ_{kk}^{-1}, where kk indicates the kth diagonal element of the matrix, or the kth subchannel. This allows writing the estimated signal at the kth subchannel as

$$\hat{s}_k = \phi_{kk}\gamma_{kk}\left(\mathbf{s}_i\right)_k + \gamma_{kk}\left(\mathbf{n}\right)_k.$$

The SNR for the kth subchannel is

$$\begin{aligned}\mathrm{SNR}_k &= \frac{\sigma_{ss}^2|\phi_{kk}|^2|\gamma_{kk}^2|^2}{\lambda_{kk}^{-1}|\gamma_{kk}|^2} \\ &= \sigma_{ss}^2|\phi_{kk}|^2\lambda_{kk} \ ,\end{aligned}$$

where γ_{kk}, λ_{kk} and ϕ_{kk} are the diagonal elements of the diagonal matrices $\mathbf{\Gamma}$, $\mathbf{\Lambda}$ and $\mathbf{\Phi}$, respectively.

The capacity defined in (2.8) is no longer valid since the precoder/decoder pair are not included. $\mathbf{R}_{zz}$ of (2.6) must be rederived, along with $\mathbf{R}_{nn}$. $\mathbf{R}_{zz}$ and $\mathbf{R}_{nn}$ become

$$\begin{aligned}\mathbf{R}_{zz} &= \mathbf{GHFss}^H\mathbf{F}^H\mathbf{H}^H\mathbf{G}^H + \mathbf{Gnn}^H\mathbf{G}^H, \\ \mathbf{R}_{nn} &= \mathbf{Gnn}^H\mathbf{G}^H.\end{aligned}$$

The mutual information of the overall system becomes

$$I\left(\hat{\mathbf{s}}_i;\mathbf{z}\right) = \log_2\left(\det\left(\mathbf{G}\mathbf{R}_{nn}\mathbf{G}^H\right)^{-1}\left(\mathbf{GHFR}_{ss}\mathbf{F}^H\mathbf{H}^H\mathbf{G}^H + \mathbf{G}\mathbf{R}_{nn}\mathbf{G}^H\right)\right). \tag{3.19}$$

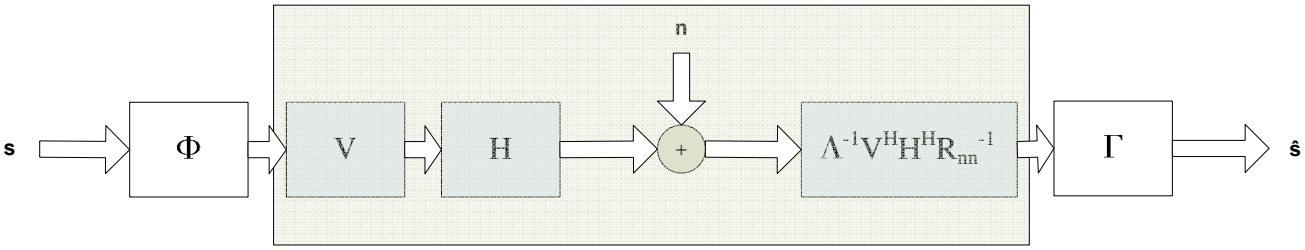


Figure 3.2: Overall channel model.

Capacity is defined as the maximization of the mutual information between the transmitter and receiver across all subchannels. The capacity is given as

$$C = \max_{\mathbf{R}_{ss}} \log_2 \left(\det \left(\left(\mathbf{G}\mathbf{R}_{nn}\mathbf{G}^{H} \right)^{-1} \left(\mathbf{G}\mathbf{H}\mathbf{F}\mathbf{R}_{ss}\mathbf{F}^{H}\mathbf{H}^{H}\mathbf{G}^{H} + \mathbf{G}\mathbf{R}_{nn}\mathbf{G}^{H} \right) \right) \right). \tag{3.20}$$

Since the equivalent channel model reduces to an identity, the information transmitted across all subchannels is parallel and independent. It may also be shown that the mutual information across the parallel and independent r (where r is the rank of **H**) subchannels as

$$I\left(\hat{\mathbf{s}}_i; \mathbf{s}\right) = \frac{1}{r} \sum_{i=1}^{r} \log_2 \left(1 + \mathrm{SNR}_i\right). \tag{3.21}$$

This shows that the mutual information on each of the parallel, independent subchannels form is equivalent, but each subchannel has a different mutual information and hence different capacity. Thus, the total mutual information of the overall system is the summation of the individual subchannels.

3.7 Conclusion

Chapter 2 introduced MIMO and the concept of precoders/decoders. Chapter 3 takes a closer look at a specific precoder/decoder designs, thus this chapter is the completion of the foundational material from which the contribution of the next chapter is built upon. The precoder/decoder designs of this chapter are the same as the imperfect precoders/decoders presented in Chapter 4.

Chapter 4

Performance Analysis of Linear Precoders/Decoders with Imperfect CSI

In the event that the CSI contains errors, how will this affect the overall performance of a MIMO system? Errors may occur in the estimation of the channel, by quantization of the CSI prior to feedback if limited CSI is used, or errors could occur while sending the CSI to the transmitter, or any combination of the above sources of error. It is obvious that there are a plethora of potential locations within a wireless system in which errors could be injected into the system. The concentration herein is limited to the errors in the channel which impact the precoder/decoder performance.

The analysis of imperfect precoders/decoders with full CSI is important because it will enable the design engineer to have an estimate of how robust or how much performance degradation the implemented system will have given varying degrees of error. The purpose of this chapter is to serve as a model on how to estimate the severity of these encounters. The imperfect channel model is introduced, equations are developed which factor in the channel error, a proof on the SVD of the channel as the number of transmit antennas is added is shown and the overall channel model is derived along with performance equations.

4.1 Imperfect Channel

Analysis of the channel matrix $\mathbf{H}$ with errors is an investigatable scenario which may represent any errors in the channel that are introduced on the receiver side, prior to the CSI being fed back to the transmitter side. In order to model the imperfect system a variable $\mathbf{\Delta X}$ is introduced which represents the amount of error added to a system parameter, where $\mathbf{X}$ represents any system parameter.

Beginning with the channel itself is where the error analysis starts. The imperfect channel caused by imperfect CSI is represented as

$$\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{\Delta H}, \tag{4.1}$$

where the imperfect channel is denoted with a tilde, $\mathbf{H}$ is the true channel and $\mathbf{\Delta H}$ is the error. Throughout this thesis a $\widetilde{\mathbf{X}}$ indicates the actual variable plus the error. The error introduced to the system is assumed to be noise that is modeled as uncorrelated, AWGN with variance σ_p^2, i.e., $\mathcal{N}\left(0, \sigma_p^2\right)$. This models the scenario where, by increasing the variance σ_p^2, the error $\mathbf{\Delta H}$ is increased, indicating more errors in the CSI.

The error analysis conducted herein begins with the assumption that errors may first be introduced in the estimation of the channel. The channel estimate comprises the CSI which is used in the decoder and is fed back to the precoder through a low rate duplex channel. If the CSI is imperfect, that implies that the decoder and precoder which are computed from that incorrect data are also imperfect.

The SVD of the imperfect channel is

$$\widetilde{\mathbf{H}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^H. \tag{4.2}$$

The imperfect CSI may also be written in terms of the EVD. Since the noise covariance $\mathbf{R}_{nn}^{-1}$ is assumed to be white and uncorrelated, it becomes $\sigma_{nn}^{-2}\mathbf{I}$ and the EVD of the CSI covariance matrix is

$$\begin{aligned}
\mathbf{R}_{nn}^{-1}\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} &= \sigma_{nn}^{-2}\left(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^H\right)^H\left(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^H\right) \\
&= \sigma_{nn}^{-2}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Sigma}}^2\widetilde{\mathbf{V}}^H \\
&= \sigma_{nn}^{-2}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{V}}^H.
\end{aligned}$$

4.1.1 Imperfect Precoder and Decoder

To consider errors in the knowledge of **H** and their impact on performance, the precoder and decoder designs of Section 3.2 are repeated. Beginning with the minimization of the mean square error (MSE) as a function of both $\widetilde{\mathbf{F}}$ and $\widetilde{\mathbf{G}}$, instead of **F** and **G** is the first step. The MSE is minimized with respect to both $\widetilde{\mathbf{F}}$ and $\widetilde{\mathbf{G}}$ and is given by

$$\mathrm{MSE}\left(\widetilde{\mathbf{F}},\widetilde{\mathbf{G}}\right) \;=\; E\left\{\left(\hat{\mathbf{s}}_i-\mathbf{s}_i\right)\left(\hat{\mathbf{s}}_i-\mathbf{s}_i\right)^H\right\}. \tag{4.3}$$

Substituting the received signal $\left(\hat{\mathbf{s}}_i=\widetilde{\mathbf{G}}\widetilde{\mathbf{H}}\widetilde{\mathbf{F}}\mathbf{s}+\widetilde{\mathbf{G}}\mathbf{n}\right)$ into equation (4.3) yields

$$\mathrm{MSE}\left(\widetilde{\mathbf{F}},\widetilde{\mathbf{G}}\right) \;=\; \left(\widetilde{\mathbf{G}}\widetilde{\mathbf{H}}\widetilde{\mathbf{F}}-\mathbf{I}\right)\mathbf{R}_{ss}\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{H}}\widetilde{\mathbf{F}}-\mathbf{I}\right)^H+\widetilde{\mathbf{G}}\mathbf{R}_{nn}\widetilde{\mathbf{G}}^H, \tag{4.4}$$

where the prior assumptions on the noise and signal covariance matrices hold.

The imperfect decoder is also based on a Weiner receiver [49] which minimizes the $tr\left(\mathrm{MSE}\left(\widetilde{\mathbf{F}},\widetilde{\mathbf{G}}\right)\right)$, and is given by

$$\widetilde{\mathbf{G}}=\mathbf{R}_{ss}\widetilde{\mathbf{F}}^H\widetilde{\mathbf{H}}^H\left(\widetilde{\mathbf{H}}\widetilde{\mathbf{F}}\mathbf{R}_{ss}\widetilde{\mathbf{F}}^H\widetilde{\mathbf{H}}^H+\mathbf{R}_{nn}\right)^{-1}. \tag{4.5}$$

By substituting in the optimal $\widetilde{\mathbf{G}}$ of (4.5) into (4.4), the MSE becomes

$$\mathrm{MSE}\left(\widetilde{\mathbf{F}}\right)=\sigma_{\mathrm{ss}}^2\left(\mathbf{I}+\sigma_{\mathrm{ss}}^2\widetilde{\mathbf{F}}^{\mathrm{H}}\widetilde{\mathbf{H}}^{\mathrm{H}}\mathbf{R}_{\mathrm{nn}}^{-1}\widetilde{\mathbf{H}}\widetilde{\mathbf{F}}\right)^{-1}. \tag{4.6}$$

The imperfect versions of decoders (3.16), (3.17) and the precoder (3.10)may be rewritten as follows in order to express them in a form where the amount of error is quantified in equation form. By rearranging the equations as the decoder plus the terms due to imperfect CSI and as the precoder plus the terms due to imperfect CSI, the system designer may better visualize in equation form the amount of error introduced into the system.

The decoder in (4.5) may be rewritten as

$$\widetilde{\mathbf{G}} \;=\; (\mathbf{G}_1+\mathbf{G}_{\mathbf{\Delta G}_1})(\mathbf{G}_2+\mathbf{G}_{\mathbf{\Delta G}_2})^{-1},$$

where

$$\mathbf{G}_1 \;=\; \mathbf{R}_{ss}\mathbf{F}^H\mathbf{H}^H$$

$$
\begin{aligned}
\mathbf{G}_{\mathbf{\Delta G}_1} &= \mathbf{R}_{ss}\left(\mathbf{F\Delta H}+\mathbf{\Delta FH}+\mathbf{\Delta F\Delta H}\right)^H \\
\mathbf{G}_2 &= \mathbf{HFR}_{ss}\mathbf{F}^H\mathbf{H}^H+\mathbf{R}_{nn} \\
\mathbf{G}_{\mathbf{\Delta G}_2} &= \mathbf{HFR}_{ss}\left(\mathbf{F\Delta H}+\mathbf{\Delta FH}+\mathbf{\Delta F\Delta H}\right)^H \\
&\quad +\mathbf{H\Delta FR}_{ss}\mathbf{\Xi}+\mathbf{\Delta HFR}_{ss}\mathbf{\Xi}+\mathbf{\Delta H\Delta FR}_{ss}\mathbf{\Xi} \\
\mathbf{\Xi} &= \left(\mathbf{FH}+\mathbf{F\Delta H}+\mathbf{\Delta FH}+\mathbf{\Delta F\Delta H}\right)^H.
\end{aligned}
$$

The imperfect decoder from (3.17) may be rewritten as

$$
\widetilde{\mathbf{G}}=\widetilde{\mathbf{\Gamma}}\widetilde{\mathbf{\Lambda}}^{-1}\widetilde{\mathbf{V}}^H\widetilde{\mathbf{H}}\sigma_{nn}^{-2}=\mathbf{G}+\mathbf{\Delta G} \tag{4.7}
$$

where

$$
\begin{aligned}
\mathbf{G} &= \mathbf{\Gamma\Lambda}^{-1}\mathbf{V}^H\mathbf{H}^H\sigma_{nn}^{-2} \\
\mathbf{\Delta G} &= \mathbf{\Gamma\Lambda}^{-1}\left(\left(\mathbf{V\Delta H}+\mathbf{\Delta VH}+\mathbf{\Delta V\Delta H}\right)^H\right)\sigma_{nn}^{-2} \\
&\quad +\mathbf{\Gamma\Delta\Lambda}^{-1}\mathbf{\Pi}+\mathbf{\Delta\Gamma\Lambda}^{-1}\mathbf{\Pi}+\mathbf{\Delta\Gamma\Delta\Lambda}^{-1}\mathbf{\Pi} \\
\mathbf{\Pi} &= \left(\left(\mathbf{VH}+\mathbf{V\Delta H}+\mathbf{\Delta VH}+\mathbf{\Delta V\Delta H}\right)^H\right)\sigma_{nn}^{-2}.
\end{aligned}
$$

The imperfect precoder from (3.10) may be rewritten in a similar fashion

$$
\widetilde{\mathbf{F}}=\widetilde{\mathbf{V}}\widetilde{\mathbf{\Phi}}=\mathbf{F}+\mathbf{\Delta F} \tag{4.8}
$$

where

$$
\begin{aligned}
\mathbf{F} &= \mathbf{V\Phi} \\
\mathbf{\Delta F} &= \mathbf{V\Delta\Phi}+\mathbf{\Delta V\Phi}+\mathbf{\Delta V\Delta\Phi}.
\end{aligned}
$$

In addition, the ϕ_{ii}'s that depend on $\mathbf{\Lambda}$ will be rewritten since $\mathbf{\Lambda}$ becomes $\widetilde{\mathbf{\Lambda}}$. Equations (3.11), (3.12), (3.14) and (3.15) become

$$
\left|\widetilde{\phi}_{ii}\right|^2=\left(\frac{P_0+\sum_{k=1}^{M_r}\widetilde{\lambda}_{nn}^{-1}}{\sigma_{ss}^2\sum_{k=1}^{M_r}\widetilde{\lambda}_{nn}^{-1/2}}\widetilde{\lambda}_{ii}^{-1/2}-\frac{1}{\widetilde{\lambda}_{ii}\sigma_{ss}^2}\right)^+, \tag{4.9}
$$

$$
\left|\widetilde{\phi}_{ii}\right|^2=\left(\frac{P_0+\sum_{k=1}^{M_r}\widetilde{\lambda}_{kk}^{-1}}{M_r\sigma_{ss}^2}-\frac{1}{\widetilde{\lambda}_{ii}\sigma_{ss}^2}\right)^+, \tag{4.10}
$$

$$\left|\widetilde{\phi}_{ii}\right|^2 = \frac{P_0}{\sigma_{ss}^2 \sum_k \widetilde{\lambda}_{kk}^{-1}} \widetilde{\lambda}_{ii}^{-1}, \tag{4.11}$$

and

$$\left|\widetilde{\phi}_{ii}\right|^2 = \frac{L_0 \widetilde{\lambda}_{NN}}{\sigma_{ss}^2} \widetilde{\lambda}_{ii}^{-1}, \tag{4.12}$$

respectively.

4.2 Equivalent Channel Model

The effective channel including the precoder (4.8), the channel (4.1), and the decoder (4.7), while excluding diagonal matrices such as $\widetilde{\mathbf{\Phi}}$ ((4.9), (4.10), (4.11), (4.12)) and $\widetilde{\mathbf{\Gamma}}$ ($\widetilde{\mathbf{\Gamma}} = \widetilde{\mathbf{\Gamma}}' \widetilde{\mathbf{\Phi}}^H \widetilde{\mathbf{\Lambda}}$), becomes

$$\begin{aligned}
\widetilde{\mathbf{\Psi}} &= \widetilde{\mathbf{F}}\mathbf{H}\widetilde{\mathbf{G}} \\
&= \widetilde{\mathbf{\Lambda}}^{-1} \widetilde{\mathbf{V}}^H \widetilde{\mathbf{H}}^H \mathbf{R}_{nn}^{-1} \mathbf{H} \widetilde{\mathbf{V}} && (4.13) \\
&= \sigma_{nn}^{-2} \widetilde{\mathbf{\Lambda}}^{-1} \widetilde{\mathbf{V}}^H \left(\widetilde{\mathbf{U}} \widetilde{\mathbf{\Sigma}} \widetilde{\mathbf{V}}^H \right)^H \left(\mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \right) \widetilde{\mathbf{V}} && (4.14) \\
&= \sigma_{nn}^{-2} \left(\widetilde{\mathbf{\Sigma}}^2 \right)^{-1} \widetilde{\mathbf{V}}^H \widetilde{\mathbf{V}} \widetilde{\mathbf{\Sigma}} \widetilde{\mathbf{U}}^H \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \widetilde{\mathbf{V}} \\
&= \sigma_{nn}^{-2} \widetilde{\mathbf{\Sigma}}^{-1} \widetilde{\mathbf{U}}^H \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \widetilde{\mathbf{V}} && (4.15) \\
&= \sigma_{nn}^{-2} \left(\mathbf{\Sigma} + \mathbf{\Delta\Sigma} \right)^{-1} \left(\mathbf{U} + \mathbf{\Delta U} \right)^H \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \left(\mathbf{V} + \mathbf{\Delta V} \right). && (4.16)
\end{aligned}$$

The imperfect overall channel gain is $\widetilde{\mathbf{\Gamma}} \widetilde{\mathbf{\Psi}} \widetilde{\mathbf{\Phi}}$. The question now becomes: What effect do the $\widetilde{\mathbf{\Sigma}}$, $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{V}}$ have on the overall equivalent channel?

4.3 Effects of Imperfect CSI on the SVD

The goal of this section is to describe in detail a useful SVD theorem which assists in describing the channel model when the CSI is imperfect. In mathematical literature, the analysis of the addition of an error term to a matrix is known as perturbation analysis. A useful tool to describe a perturbed subspace decomposition was found in [53], which considers the case where the error is an uncorrelated AWGN random variable, exactly the scenario here. This theorem assumes that the

variance $E[\mathbf{\Delta H \Delta H}^H]$ is asymptotic as the number of columns approaches infinity and is denoted by $E[\mathbf{\Delta H \Delta H}^H]/r = \sigma_p^2 \mathbf{I}_r$. For a large M_t, the SVD of $\widetilde{\mathbf{H}}$ may be represented as

$$\widetilde{\mathbf{H}}_W \approx \mathbf{U} \left(\mathbf{\Sigma}^2 + M_t \sigma_p^2 \mathbf{I}_{M_r} \right)^{\frac{1}{2}} \widetilde{\mathbf{V}}^H = \hat{\mathbf{H}}_W, \tag{4.17}$$

where σ_p^2 is the variance of the CSI error term or perturbation, and the subscript W indicates a wide matrix (i.e. $M_t \geq M_r$) This shows that for a small σ_p^2 and as the number of columns of $\widetilde{\mathbf{H}}$ are increased, the singular values of $\widetilde{\mathbf{H}}$ increase by an amount approximately equal to $\sigma_p \sqrt{M_t}$. It is important to notice that the matrix of left singular vectors **U**, is not changed. In other words, for a wide **H**, (4.2) becomes

$$\widetilde{\mathbf{H}}_W = \mathbf{U} \widetilde{\mathbf{\Sigma}} \widetilde{\mathbf{V}}^H. \tag{4.18}$$

The matrix of singular values may be written as

$$\widetilde{\mathbf{\Sigma}}_W = \left(\mathbf{\Sigma}^2 + M_t \sigma_p^2 \mathbf{I}_{M_r} \right)^{\frac{1}{2}}. \tag{4.19}$$

The above theorem which appeared in [53] is presented only for wide matrices as the number of columns is increased to infinity. A literature search was undertaken to find a theorem which stated a similar argument for tall matrices, but was not found. Despite the literature search not providing useful results, the symmetry of the SVD regarding the **U** and **V** matrices was enough to postulate a theorem for the contrary asymptotic case where the number of rows approaches infinity. Which may also be found by transposing the above theorem. For a large M_r, the SVD of $\widetilde{\mathbf{H}}$ may be represented as

$$\hat{\mathbf{H}}_T \approx \widetilde{\mathbf{U}} \left(\mathbf{\Sigma}^2 + M_r \sigma_p^2 \mathbf{I}_{M_t} \right)^{\frac{1}{2}} \mathbf{V}^H = \widetilde{\mathbf{H}}_T. \tag{4.20}$$

From the above equation, it may be inferred that for a small σ_p^2 and as the number of rows of $\widetilde{\mathbf{H}}$ are increased, the singular values of $\widetilde{\mathbf{H}}$ increase by an amount approximately equal to $\sigma_p \sqrt{M_r}$. The approximated SVD of $\widetilde{\mathbf{H}}$ may be written in shorthand form as

$$\widetilde{\mathbf{H}}_T = \widetilde{\mathbf{U}} \widetilde{\mathbf{\Sigma}} \mathbf{V}^H, \tag{4.21}$$

followed by the matrix of singular values

$$\widetilde{\mathbf{\Sigma}}_T = \left(\mathbf{\Sigma}^2 + M_r \sigma_p^2 \mathbf{I}_{M_t} \right)^{\frac{1}{2}}. \tag{4.22}$$

Two simulation scenarios were set up to model the error between the actual imperfect channels, $\hat{\mathbf{H}}_W$ and $\hat{\mathbf{H}}_T$ and the approximation to the imperfect channel, $\widetilde{\mathbf{H}}_W$ and $\widetilde{\mathbf{H}}_T$, respectively, for both the cases where M_r and M_t approach infinity. For the wide case, the simulation begins with a random flat fading Rayleigh channel with an antenna configuration of 1×5 and computes the two-norm of the difference between $\hat{\mathbf{H}}_W$ and $\widetilde{\mathbf{H}}_W$ and is averaged over 1000 trials for a given σ_p^2. On the next iteration, a row is added to the channel and the error between $\hat{\mathbf{H}}_W$ and $\widetilde{\mathbf{H}}_W$ is again averaged over 1000 trials and continues until the antenna configuration is square. The simulation then repeats for a different σ_p^2 value. Two different matrix sizes are given (1×5 becomes 5×5 and 1×10 becomes 10×10) in Fig. 4.1 for two different values of σ_p^2.

The same simulation was mimicked for the case of a tall antenna configuration and computed the error between $\hat{\mathbf{H}}_T$ and $\widetilde{\mathbf{H}}_T$. The results for the tall antenna configuration is shown in Fig. 4.2.

Notice in both Fig. 4.1 and 4.2 that the approximation becomes worse the more closer to square the antenna configuration becomes, regardless the size of the σ_p^2. As the ratio of transmit to receive antennas increases to 1, the greater the error between the actual imperfect channel and the approximation to the imperfect channel. In Chapter 5 some plots will be presented which show how well the approximation improves with non-square antenna configurations. This will allow us to quantify how much error is tolerable within a system design.

4.4 Equivalent Channel Model: Wide Case

The results of Section 4.3 may be used as a tool to simplify the equivalent channel model found in (4.16). The matrix of the imperfect left singular vectors may be approximated as $\widetilde{\mathbf{U}} = \mathbf{U}$ as shown in (4.18). Additionally, the amount of error introduced into the matrix of singular values is shown in (4.19). This eliminates two of the three unknowns ($\mathbf{\Delta U}$ and $\mathbf{\Delta \Sigma}$) introduced in (4.2). The only remaining unknown variable of (4.16) is $\mathbf{\Delta V}$ which will be addressed in Section 4.4.1.

Recall from Section 3.3 that in the event of a wide channel model, the irrelevant $M_t - M_r$ columns of $\mathbf{\Sigma}$ and $\mathbf{V}$ are eliminated. This simplification makes the mathematical operations much easier within this section; otherwise, the matrix multiplication of (4.7) and (4.8) would not be possible for a non-square matrix. Earlier in this thesis the subscript, (e.g., $\mathbf{I}_{M_r}$) indicated the

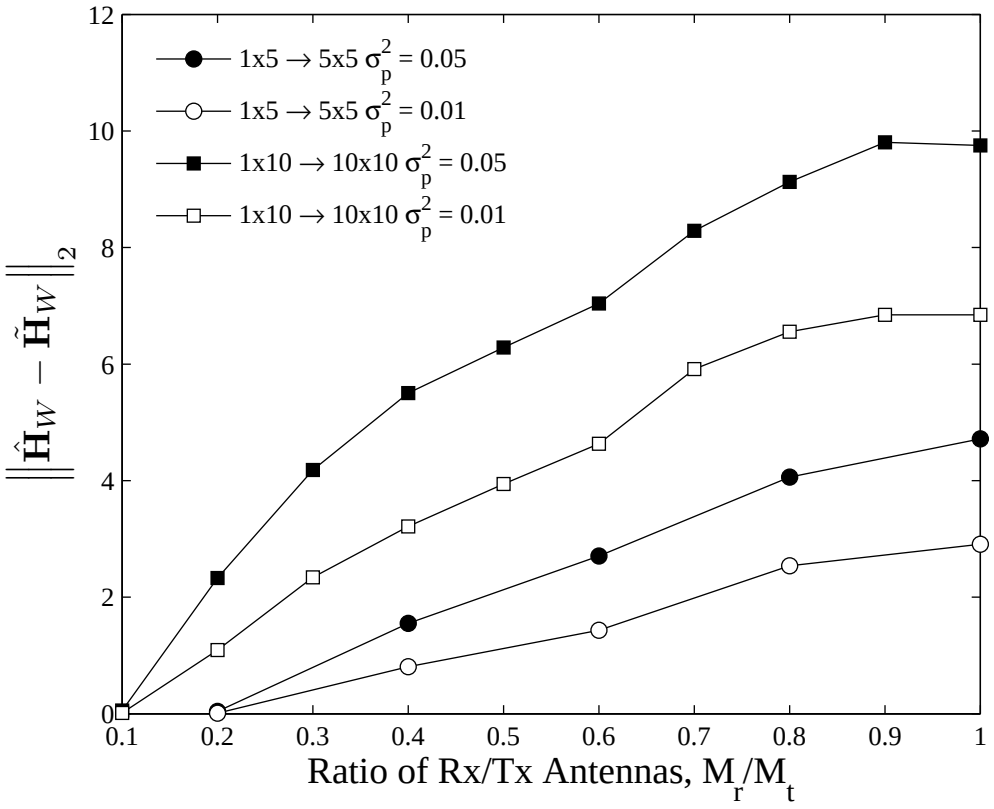


Figure 4.1: SVD approximation error for wide matrices as the antenna configuration becomes square.

dimension of a square matrix, but herein it will also indicate matrices that have had columns or rows eliminated. Specifically for the case of a wide channel model using the approximation $\mathbf{U} = \widetilde{\mathbf{U}}$, the overall channel model becomes

$$\begin{aligned}
\widetilde{\boldsymbol{\Psi}}_{M_r} &= \sigma_{nn}^{-2}\widetilde{\boldsymbol{\Sigma}}_{M_r}^{-1}\widetilde{\mathbf{U}}^H\mathbf{U}\boldsymbol{\Sigma}_{M_r}\mathbf{V}_{M_r}^H\widetilde{\mathbf{V}}_{M_r} \\
&= \sigma_{nn}^{-2}\widetilde{\boldsymbol{\Sigma}}_{M_r}^{-1}\mathbf{U}^H\mathbf{U}\boldsymbol{\Sigma}_{M_r}\mathbf{V}_{M_r}^H\left(\mathbf{V}+\boldsymbol{\Delta}\mathbf{V}\right)_{M_r} \\
&= \sigma_{nn}^{-2}\widetilde{\boldsymbol{\Sigma}}_{M_r}^{-1}\boldsymbol{\Sigma}_{M_r}\mathbf{V}_{M_r}^H\left(\mathbf{V}+\boldsymbol{\Delta}\mathbf{V}\right)_{M_r}. \qquad (4.23)
\end{aligned}$$

This leaves only one unknown, $\boldsymbol{\Delta}\mathbf{V}$. By determining $\boldsymbol{\Delta}\mathbf{V}$, a workable expression for the effect of the error on system performance may be obtained.

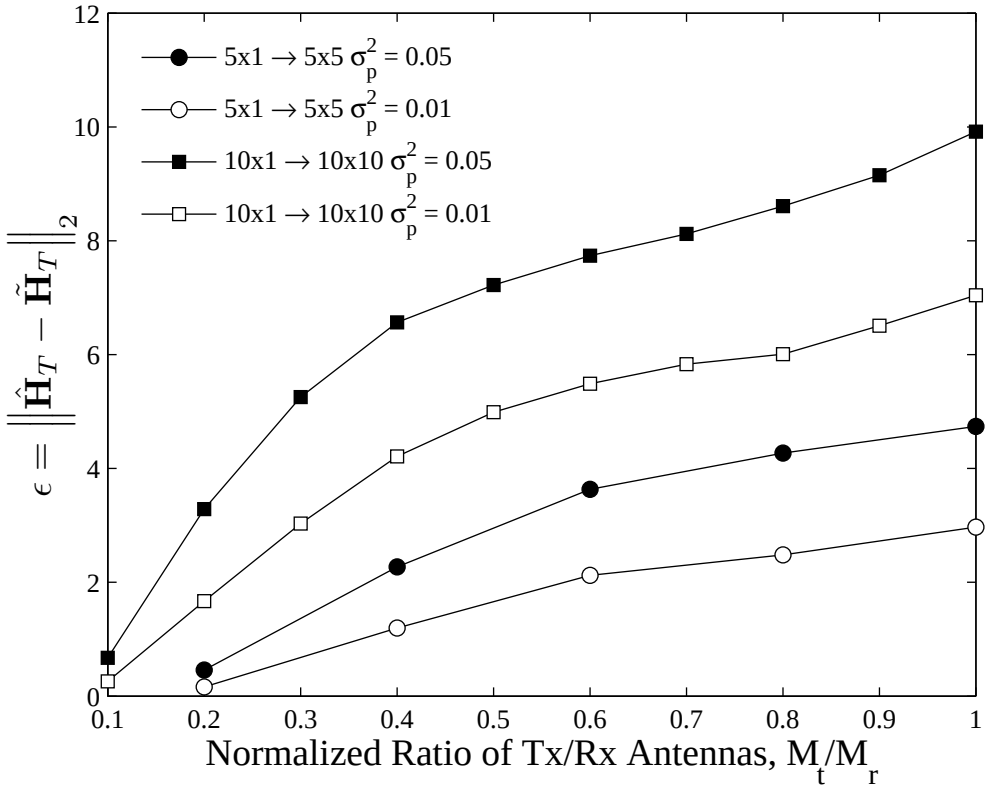


Figure 4.2: SVD approximation error for tall matrices as the antenna configuration becomes square.

4.4.1 Effects of $\Delta\mathbf{V}$

By beginning with the SVD of a wide ($M_t \geq M_r$) $\widetilde{\mathbf{H}}$ and $\mathbf{H}$ of the same size and eliminating the $M_t - M_r$ columns of $\widetilde{\mathbf{\Sigma}}$, $\mathbf{\Sigma}$, $\widetilde{\mathbf{V}}$ and $\mathbf{V}$, the SVD of $\widetilde{\mathbf{H}}$ and $\mathbf{H}$ may be rearranged as follows

$$\begin{aligned} \widetilde{\mathbf{V}}_{M_r} &= \widetilde{\mathbf{H}}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} \\ \mathbf{V}_{M_r} &= \mathbf{H}^H \mathbf{U} \mathbf{\Sigma}_{M_r}^{-1}, \end{aligned}$$

where $\widetilde{\mathbf{V}}_{M_r}, \mathbf{V}_{M_r}, \widetilde{\mathbf{H}}$ and $\mathbf{H}$ are $M_t \times M_r$ and $\mathbf{U}, \widetilde{\mathbf{\Sigma}}_{M_r}^{-1}$ and $\mathbf{\Sigma}_{M_r}^{-1}$ are $M_r \times M_r$. Next, equation (4.1) may be rewritten as

$$\mathbf{\Delta H} = \widetilde{\mathbf{H}} - \mathbf{H}. \tag{4.24}$$

Now the desire is to know $\mathbf{\Delta V}$. We may write $\mathbf{\Delta V}$ in the same fashion as

$$\mathbf{\Delta V} = \widetilde{\mathbf{V}} - \mathbf{V}.$$

Or in this case, it is

$$\begin{aligned} \mathbf{\Delta V}_{M_r} &= \widetilde{\mathbf{V}}_{M_r} - \mathbf{V}_{M_r} \\ &= \widetilde{\mathbf{H}}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} - \mathbf{H}^H \mathbf{U} \mathbf{\Sigma}_{M_r}^{-1}. \end{aligned} \tag{4.25}$$

Now by subtracting and adding the same term ($\mathbf{H}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1}$) from (4.25) yields

$$\mathbf{\Delta V}_{M_r} = \widetilde{\mathbf{H}}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} - \mathbf{H}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} + \mathbf{H}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} - \mathbf{H}^H \mathbf{U} \mathbf{\Sigma}_{M_r}^{-1}. \tag{4.26}$$

By manipulating (4.26) the final expression for $\mathbf{\Delta V}_{M_r}$ becomes

$$\mathbf{\Delta V}_{M_r} = \left(\widetilde{\mathbf{H}} - \mathbf{H}\right)^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} + \mathbf{H}^H \mathbf{U} \left(\widetilde{\mathbf{\Sigma}}_{M_r}^{-1} - \mathbf{\Sigma}_{M_r}^{-1}\right) \tag{4.27}$$

$$= \mathbf{\Delta H}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} + \mathbf{H}^H \mathbf{U} \mathbf{\Delta} \mathbf{\Sigma}_{M_r}^{-1} \tag{4.28}$$

This produces an approximate expression for the matrix $\mathbf{\Delta V}_{M_r}$, which holds for wide matrices only. This does not give an estimation of the size of the elements of the matrix.

It is assumed from the SVD error theorem in Section 4.3 that $\widetilde{\mathbf{U}} = \mathbf{U}$ and the size of $\widetilde{\mathbf{\Sigma}}$ may be inferred from (4.19). The question remains as to the size of $\widetilde{\mathbf{V}}$. The following is an analysis of the elements of $\mathbf{\Delta V}$, which will give insight to the amount of error in $\widetilde{\mathbf{V}}$.

4.4.2 Elements of $\mathbf{\Delta V}$

Taking a closer look at the elements from left to right of (4.23), their properties may be analyzed. The noise standard deviation σ_{nn} is a constant and the matrices $\widetilde{\mathbf{\Sigma}}$ and $\mathbf{\Sigma}$ are both diagonal with zeros on all of the off-diagonal elements. $\mathbf{V}$ is a unitary matrix, which means that the norm of each of the columns is unity and the norm of the matrix itself is unity. The question remains as to the size of the errors in the non-diagonal matrix on the far right of (4.23), $\mathbf{\Delta V}_{M_r}$.

By isolating the matrix product $\mathbf{V}_{M_r}^H \widetilde{\mathbf{V}}_{M_r}$, we may approximate the size of the elements of the error $\mathbf{\Delta V}_{M_r}$. This matrix product may be clarified as

$$\mathbf{Z}_{\mathbf{V}} \equiv \mathbf{V}_{M_r}^H \left(\mathbf{V}_{M_r} + \mathbf{\Delta V}_{M_r}\right) = \mathbf{I}_{M_r} + \mathbf{V}_{M_r}^H \mathbf{\Delta V}_{M_r}. \tag{4.29}$$

Rewriting the right hand side in terms of the individual elements of $\mathbf{V}_{M_r}^H \mathbf{\Delta V}_{M_r}$ and making comparisons to a two dimensional case helps the analysis.

First generalizing the matrices $\mathbf{V}_{M_r}^H \mathbf{\Delta V}_{M_r}$ into vector form

$$\begin{bmatrix} \mathbf{v}_1^H \\ \mathbf{v}_2^H \\ \vdots \\ \mathbf{v}_{M_r}^H \end{bmatrix} \begin{bmatrix} \mathbf{\Delta v}_1 & \mathbf{\Delta v}_2 & \cdots & \mathbf{\Delta v}_{M_r} \end{bmatrix}, \tag{4.30}$$

then taking a closer look at their individual elements yields

$$\begin{bmatrix} \mathbf{v}_1^H \mathbf{\Delta v}_1 & \mathbf{v}_1^H \mathbf{\Delta v}_2 & \cdots & \mathbf{v}_1^H \mathbf{\Delta v}_{M_r} \\ \mathbf{v}_2^H \mathbf{\Delta v}_1 & \mathbf{v}_2^H \mathbf{\Delta v}_2 & \cdots & \mathbf{v}_2^H \mathbf{\Delta v}_{M_r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_{M_r}^H \mathbf{\Delta v}_1 & \mathbf{v}_{M_r}^H \mathbf{\Delta v}_2 & \cdots & \mathbf{v}_{M_r}^H \mathbf{\Delta v}_{M_r} \end{bmatrix}. \tag{4.31}$$

The aim is to approximate or bound the individual elements of the above $M_r \times M_r$ matrix. Two approximations may be made from which the above matrix (4.31) may be simplified. The first approximation may be made by looking at an example that is valid for any dimensions of the vectors $\mathbf{v}_{ij}$, for any columns in $\widetilde{\mathbf{V}}$, regardless of the size of $\widetilde{\mathbf{V}}$. Multiplying two vectors of $\widetilde{\mathbf{V}}$ results in a product equal to zero, but also reveals some useful approximations such as,

$$\begin{aligned} \widetilde{\mathbf{v}}_i^H \widetilde{\mathbf{v}}_j &= \left(\mathbf{v}_i + \mathbf{\Delta v}_i\right)^H \left(\mathbf{v}_j + \mathbf{\Delta v}_j\right) = 0 && (4.32) \\ 0 &= \mathbf{v}_i^H \mathbf{v}_j + \mathbf{v}_i^H \mathbf{\Delta v}_j + \mathbf{\Delta v}_i^H \mathbf{v}_j + \mathbf{\Delta v}_i^H \mathbf{\Delta v}_j && (4.33) \\ 0 &\approx \mathbf{v}_i^H \mathbf{\Delta v}_j + \mathbf{\Delta v}_i^H \mathbf{v}_j. && (4.34) \end{aligned}$$

Recall that $\mathbf{v}_i$ denotes the ith column of the right singular vectors of the true $\mathbf{H}$ and the $\mathbf{\Delta v}_i$ terms are the errors. Thus the terms inside the parenthesis in (4.32) indicate the true value plus the error term due to imperfect CSI. The first term of (4.33) is zero because the vectors are orthogonal due to the unitary structure of $\widetilde{\mathbf{V}}$. The final term of (4.33) is approximated to be zero because it is the inner product of two small terms. The remaining terms of (4.34) show that the $(i,j)^{th}$ element is equal to the negative of the $(j,i)^{th}$ element.

The second approximation may be made from the two dimensional example shown in Fig. 4.3. The two perpendicular sets of vectors in Fig. 4.3 show that the error of the right singular

vectors, is assumed to be small. As the error terms of $\mathbf{\Delta v}_1$ and $\mathbf{\Delta v}_2$ approach zero, the two pairs of vectors ($\mathbf{v}_1$, $\widetilde{\mathbf{v}}_1$ and $\mathbf{v}_2$, $\widetilde{\mathbf{v}}_2$) approach each other. For small errors, the vectors $\mathbf{\Delta v}_1$ and $\mathbf{\Delta v}_2$ are approximately perpendicular to the vectors they are corrupting, $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively. In other words, the following products are approximately zero

$$\mathbf{v}_i^H \mathbf{\Delta v}_i \approx 0 \tag{4.35}$$

Applying these to approximate the right-hand-side of (4.29) produces

$$\mathbf{Z_V} \approx \mathbf{I} \; + \begin{bmatrix} 0 & \mathbf{v}_1^H\mathbf{\Delta v}_2 & \mathbf{v}_1^H\mathbf{\Delta v}_3 & \cdots & \mathbf{v}_1^H\mathbf{\Delta v}_{M_r} \\ -\mathbf{v}_1^H\mathbf{\Delta v}_2 & 0 & \mathbf{v}_2^H\mathbf{\Delta v}_3 & & \\ -\mathbf{v}_1^H\mathbf{\Delta v}_3 & -\mathbf{v}_2^H\mathbf{\Delta v}_3 & \ddots & & \vdots \\ \vdots & \vdots & & & \mathbf{v}_{M_r-1}^H\mathbf{\Delta v}_{M_r-1} \\ -\mathbf{v}_1^H\mathbf{\Delta v}_{M_r} & \cdots & & -\mathbf{v}_{M_r-1}^H\mathbf{\Delta v}_{M_r-1} & 0 \end{bmatrix}. \tag{4.36}$$

The above matrix may be simplified further, by noting that the off diagonal elements are the dot product of two vectors. Recall that the dot product is the multiplication of two vectors and is denoted as

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^H \cdot \mathbf{b} = ||a|| \cdot ||b|| \cos\theta_{ab}, \tag{4.37}$$

where θ_{ab} is the angle between vectors **a** and **b** and $||\cdot||$ denotes the norm. This allows the off-diagonal elements to be reduced to

$$\mathbf{Z_V} \approx \mathbf{I} \; + \begin{bmatrix} 0 & ||\mathbf{\Delta v}_2|| \cos\theta_{12} & ||\mathbf{\Delta v}_3|| \cos\theta_{13} & \cdots \\ -||\mathbf{\Delta v}_2|| \cos\theta_{12} & 0 & ||\mathbf{\Delta v}_3|| \cos\theta_{23} & \\ -||\mathbf{\Delta v}_3|| \cos\theta_{13} & -||\mathbf{\Delta v}_3|| \cos\theta_{23} & \ddots & \\ \vdots & \vdots & & \end{bmatrix}. \tag{4.38}$$

This shows that each of the off-diagonal elements in column i is bounded by $||\mathbf{v}_i||$ where $i = 1 \ldots M_r$. It is also apparent that the size of the off-diagonal elements are directly related to the amount of CSI error. The above equation (4.38) may be bounded by

$$\mathbf{Z_V} \approx \mathbf{I} \; + \begin{bmatrix} 0 & ||\mathbf{\Delta v}_2|| & ||\mathbf{\Delta v}_3|| & \cdots \\ -||\mathbf{\Delta v}_2|| & 0 & ||\mathbf{\Delta v}_3|| & \\ -||\mathbf{\Delta v}_3|| & -||\mathbf{\Delta v}_3|| & \ddots & \\ \vdots & \vdots & & \end{bmatrix}. \tag{4.39}$$

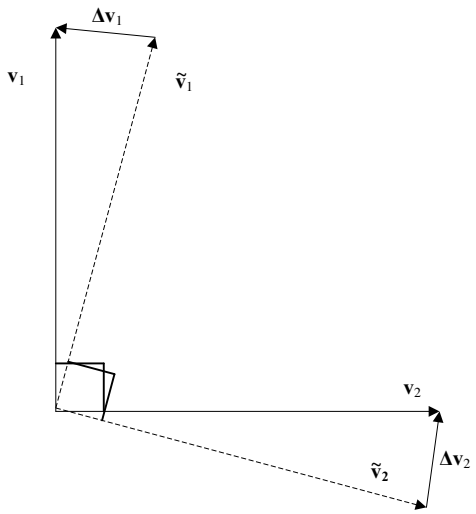


Figure 4.3: A two dimensional example.

4.5 Equivalent Channel Model: Tall Case

The analysis of the previous section may be rederived in order to make use of the SVD theorem of Section 4.3 to apply to tall channel matrices. Recall that the matrix of the imperfect right singular vectors may be approximated as $\widetilde{\mathbf{V}} = \mathbf{V}$, according to the theorem, and the amount of error introduced into the matrix of singular values is shown in (4.22). This eliminates two of the three unknowns ($\mathbf{\Delta\Sigma}$ and $\mathbf{\Delta V}$) introduced in (4.2). The only remaining unknown variable of equation (4.16) is $\mathbf{\Delta U}$ which will be determined next.

In Section 3.3 it was shown that for a wide channel model, the irrelevant $M_t - M_r$ columns of $\mathbf{\Sigma}$, $\widetilde{\mathbf{\Sigma}}$, $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ may be eliminated. For a tall channel model, the $M_r - M_t$ columns of $\mathbf{U}$ and $\widetilde{\mathbf{U}}$ and the $M_r - M_t$ rows of $\mathbf{\Sigma}$ and $\widetilde{\mathbf{\Sigma}}$ may be eliminated. The matrices that have had columns or rows eliminated are denoted with a subscript (e.g., $\mathbf{\Sigma}_{M_t}$),

$$\begin{aligned}
\widetilde{\mathbf{\Psi}}_{M_t} &= \sigma_{nn}^{-2}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1}\widetilde{\mathbf{U}}_{M_t}^{H}\mathbf{U}_{M_t}\mathbf{\Sigma}_{M_t}\mathbf{V}^{H}\widetilde{\mathbf{V}} \\
&= \sigma_{nn}^{-2}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1}\widetilde{\mathbf{U}}_{M_t}^{H}\mathbf{U}_{M_t}\mathbf{\Sigma}_{M_t} \\
&= \sigma_{nn}^{-2}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1}\left(\mathbf{U}+\mathbf{\Delta U}\right)_{M_t}^{H}\mathbf{U}_{M_t}\mathbf{\Sigma}_{M_t}. \qquad (4.40)
\end{aligned}$$

The $\mathbf{\Delta U}$ parameter is the only unknown and will be determined in a manner similar to $\mathbf{\Delta V}$ in the previous section.

4.5.1 Effects of $\Delta\mathbf{U}$

By computing the SVD of a tall ($M_r \geq M_t$) $\widetilde{\mathbf{H}}$ and $\mathbf{H}$ of the same dimension and eliminating the $M_r - M_t$ columns of $\mathbf{U}$, $\widetilde{\mathbf{U}}$, and the $M_r - M_t$ rows of $\mathbf{\Sigma}$ and $\widetilde{\mathbf{\Sigma}}$, the SVD of $\widetilde{\mathbf{H}}$ and $\mathbf{H}$ may be rearranged as follows

$$\begin{aligned}\widetilde{\mathbf{U}}_{M_t} &= \widetilde{\mathbf{H}}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} \\ \mathbf{U}_{M_t} &= \mathbf{H}\mathbf{V}\mathbf{\Sigma}_{M_t}^{-1},\end{aligned}$$

where $\widetilde{\mathbf{U}}_{M_t}, \mathbf{U}_{M_t}, \widetilde{\mathbf{H}}$ and $\mathbf{H}$ are $M_r \times M_t$ and $\mathbf{\Sigma}_{M_t}^{-1}, \widetilde{\mathbf{\Sigma}}_{M_t}^{-1}$ and $\mathbf{V}$ are $M_t \times M_t$. $\mathbf{\Delta U}$ may be rewritten as

$$\mathbf{\Delta U} = \widetilde{\mathbf{U}} - \mathbf{U}.$$

Or in this case, it is

$$\begin{aligned}\mathbf{\Delta U}_{M_t} &= \widetilde{\mathbf{U}}_{M_t} - \mathbf{U}_{M_t} \\ &= \widetilde{\mathbf{H}}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} - \mathbf{H}\mathbf{V}\mathbf{\Sigma}_{M_t}^{-1}. \qquad (4.41)\end{aligned}$$

Again, by subtracting and adding the same term ($\mathbf{H}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1}$) from (4.41) yields

$$\mathbf{\Delta U}_{M_t} = \widetilde{\mathbf{H}}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} - \mathbf{H}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} + \mathbf{H}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} - \mathbf{H}\mathbf{V}\mathbf{\Sigma}_{M_t}^{-1}. \qquad (4.42)$$

After manipulating (4.42) the final expression for $\mathbf{\Delta U}_{M_t}$ becomes

$$\begin{aligned}\mathbf{\Delta U}_{M_t} &= \left(\widetilde{\mathbf{H}} - \mathbf{H}\right)\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} + \mathbf{H}\mathbf{V}\left(\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} - \mathbf{\Sigma}_{M_t}^{-1}\right) \qquad (4.43) \\ &= \mathbf{\Delta H}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} + \mathbf{H}\mathbf{V}\mathbf{\Delta\Sigma}_{M_t}^{-1} \qquad (4.44)\end{aligned}$$

This produces an approximate expression for the matrix $\mathbf{\Delta U}_{M_t}$ which provides the final unknown from (4.40) for a tall channel matrix.

It is assumed from (4.20) that $\widetilde{\mathbf{V}} = \mathbf{V}$ and the size of $\widetilde{\mathbf{\Sigma}}$ may be inferred from (4.22). The question remains as to the size of the errors in $\widetilde{\mathbf{U}}$. The following is an analysis bounding the size of $\mathbf{\Delta U}$, which will give insight to the impact of the error of $\widetilde{\mathbf{U}}$.

4.5.2 Elements of $\Delta\mathbf{U}$

By mimicking the previous analysis which derived the bound for $\mathbf{\Delta V}$, a similar derivation may be completed for the $\mathbf{\Delta U}$ matrix in the case of a tall channel matrix. By isolating the unknown matrices $(\mathbf{U}+\mathbf{\Delta U})_{M_t}^H \mathbf{U}_{M_t}$ of (4.40) an approximation to the size of the error $\mathbf{\Delta U}_{M_t}$ may be made. This matrix product may be clarified as

$$\mathbf{Z_V} \equiv (\mathbf{U}_{M_t}+\mathbf{\Delta U})_{M_t}^H \mathbf{U}_{M_t} = \mathbf{I}_{M_t} + \mathbf{\Delta U}_{M_t}^H \mathbf{U}_{M_t}. \tag{4.45}$$

Again, generalizing the matrices $\mathbf{\Delta U}_{M_t}^H \mathbf{U}_{M_t}$ into vector form

$$\begin{bmatrix} \mathbf{\Delta u}_1^H \\ \mathbf{\Delta u}_2^H \\ \vdots \\ \mathbf{\Delta u}_{M_t}^H \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_{M_t} \end{bmatrix}, \tag{4.46}$$

then taking a closer look at their individual elements produces

$$\begin{bmatrix} \mathbf{\Delta u}_1^H\mathbf{u}_1 & \mathbf{\Delta u}_1^H\mathbf{u}_2 & \mathbf{\Delta u}_1^H\mathbf{u}_3 & \cdots & \mathbf{\Delta u}_1^H\mathbf{u}_{M_t} \\ \mathbf{\Delta u}_2^H\mathbf{u}_1 & \mathbf{\Delta u}_2^H\mathbf{u}_2 & \cdots & & \mathbf{\Delta u}_2^H\mathbf{u}_{M_t} \\ \vdots & \vdots & \ddots & & \vdots \\ \mathbf{\Delta u}_{M_t}^H\mathbf{u}_1 & \mathbf{\Delta u}_{M_t}^H\mathbf{u}_2 & \cdots & & \mathbf{\Delta u}_{M_t}^H\mathbf{u}_{M_t} \end{bmatrix}. \tag{4.47}$$

The goal is to approximate or bound the individual elements of the above $M_t \times M_t$ matrix. Two approximations may be made

$$\mathbf{u}_i^H \mathbf{\Delta u}_j = -\mathbf{\Delta u}_i^H \mathbf{u}_j \quad \text{and} \quad \mathbf{\Delta u}_{\mathrm{i}}^{\mathrm{H}} \mathbf{u}_{\mathrm{i}} \approx 0. \tag{4.48}$$

Again, recall that $\mathbf{u}_i$ is the ith column of the right singular vectors of the true $\mathbf{H}$ matrix and the $\mathbf{\Delta u}_i$ indicates the errors due to imperfect CSI.

Applying the two approximations to (4.45) produces

$$\mathbf{Z_V} \approx \mathbf{I} + \begin{bmatrix} 0 & \mathbf{\Delta u}_1^H\mathbf{u}_2 & \mathbf{\Delta u}_1^H\mathbf{u}_3 & \cdots & \mathbf{\Delta u}_1^H\mathbf{u}_{M_t} \\ -\mathbf{\Delta u}_1^H\mathbf{u}_2 & 0 & \mathbf{\Delta u}_2^H\mathbf{u}_3 & & \\ -\mathbf{\Delta u}_1^H\mathbf{u}_3 & -\mathbf{\Delta u}_2^H\mathbf{u}_3 & \ddots & & \vdots \\ \vdots & \vdots & & & \mathbf{\Delta u}_{M_t-1}^H\mathbf{u}_{M_t-1} \\ -\mathbf{\Delta u}_1^H\mathbf{u}_{M_t} & \cdots & & -\mathbf{\Delta u}_{M_t-1}^H\mathbf{u}_{M_t-1} & 0 \end{bmatrix}. \tag{4.49}$$

Similar to equation (4.36), the above matrix may be simplified through the use of the dot product. The simplification of the above matrix produces

$$\mathbf{Z_V} \approx \mathbf{I} \; + \begin{bmatrix} 0 & \left\|\mathbf{\Delta u}_1^H\right\| & \left\|\mathbf{\Delta u}_1^H\right\| & \cdots \\ -\left\|\mathbf{\Delta u}_1^H\right\| & 0 & \left\|\mathbf{\Delta u}_2^H\right\| & \\ -\left\|\mathbf{\Delta u}_1^H\right\| & -\left\|\mathbf{\Delta u}_2^H\right\| & \ddots & \\ \vdots & \vdots & & \end{bmatrix}. \tag{4.50}$$

This is a significant result because it shows the direct relationship of the size of the CSI errors and the size of the off-diagonal elements for the case of a tall antenna configuration.

4.6 Equivalent Channel Model: Non-Full Rank Case

The final case of the equivalent channel model is for the case of a non-full rank channel **H**. In the event of a non-full rank channel, the effective size of the channel reduces to the rank of **H**, regardless of the antenna configuration, square, wide or tall.

The results of Section 4.3 may be applied to a non-full rank **H** or $\widetilde{\mathbf{H}}$. If the non-full rank channel is wide, the results of Section 4.4 apply. Similarly, if the non-full rank channel is tall, the results of Section 4.5 apply. The effects of a non-full rank channel simply reduces the number of parallel and independent subchannels that exist to be less than the minimum of M_r and M_t.

4.7 Performance Results

The cascade of the **G** and **F** with perfect CSI is given in (3.18) and depicted in Fig. 3.2. When the CSI is perfect it also allows the r suchannels to be written as parallel, independent subchannels. However, when the suboptimal $\widetilde{\mathbf{G}}$ and $\widetilde{\mathbf{F}}$ (equations (4.7) and (4.8), respectively) are cascaded, they do not conveniently reduce to an identity but produce equation (4.13). The corresponding noise correlation for the cascaded precoder/decoder is

$$\begin{aligned} \widetilde{\mathbf{G}}\mathbf{n}\mathbf{n}^H\widetilde{\mathbf{G}}^H &= \widetilde{\mathbf{\Lambda}}^{-1}\widetilde{\mathbf{V}}^H\mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{R}_{nn}\mathbf{R}_{nn}^{-1}\widetilde{\mathbf{H}}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Lambda}}^{-1} \\ &= \sigma_{nn}^{-1}\widetilde{\mathbf{\Lambda}}^{-1}\widetilde{\mathbf{V}}^H\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Lambda}}^{-1} \end{aligned}$$

$$
\begin{aligned}
&= \sigma_{nn}^{-1}\widetilde{\mathbf{\Lambda}}^{-1}\widetilde{\mathbf{V}}^{H}\left(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^{H}\right)^{H}\left(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^{H}\right)\widetilde{\mathbf{V}}\widetilde{\mathbf{\Lambda}}^{-1} \\
&= \sigma_{nn}^{-1}\widetilde{\mathbf{\Lambda}}^{-1}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{\Lambda}}^{-1} \\
&= \sigma_{nn}^{-1}\widetilde{\mathbf{\Lambda}}^{-1}.
\end{aligned}
$$

This results in the estimated signal being written as

$$
\hat{\mathbf{s}} = \begin{bmatrix}
\widetilde{\gamma}_{11}\widetilde{\psi}_{11}\widetilde{\phi}_{11}s_1 + \widetilde{\gamma}_{11}\widetilde{\psi}_{12}\widetilde{\phi}_{22}s_2 + \cdots + \widetilde{\gamma}_{11}\widetilde{\psi}_{1r}\widetilde{\phi}_{rr}s_r + \widetilde{\gamma}_{11}n_1 \\
\widetilde{\gamma}_{22}\widetilde{\psi}_{21}\widetilde{\phi}_{11}s_1 + \widetilde{\gamma}_{22}\widetilde{\psi}_{22}\widetilde{\phi}_{22}s_2 + \cdots + \widetilde{\gamma}_{22}\widetilde{\psi}_{2r}\widetilde{\phi}_{rr}s_r + \widetilde{\gamma}_{22}n_2 \\
\vdots \\
\widetilde{\gamma}_{rr}\widetilde{\psi}_{r1}\widetilde{\phi}_{11}s_1 + \widetilde{\gamma}_{rr}\widetilde{\psi}_{r2}\widetilde{\phi}_{22}s_2 + \cdots + \widetilde{\gamma}_{rr}\widetilde{\psi}_{rr}\widetilde{\phi}_{rr}s_r + \widetilde{\gamma}_{rr}n_r
\end{bmatrix}, \tag{4.51}
$$

where $\widetilde{\gamma}_{rr}$ and $\widetilde{\phi}_{rr}$ are the rrth elements of the diagonal power allocation matrices and $\widetilde{\psi}$ is the approximated channel as shown in (4.13).

The signal received in any of the subchannels may be found by selecting the desired row element of the $r \times 1$ vector of (4.51). For example, the signal recieved at the first subchannel is expressed as

$$
A_1 = \widetilde{\gamma}_{11}\widetilde{\psi}_{11}\widetilde{\phi}_{11}s_1 + \widetilde{\gamma}_{11}\widetilde{\psi}_{12}\widetilde{\phi}_{22}s_2 + \cdots + \widetilde{\gamma}_{11}\widetilde{\psi}_{1r}\widetilde{\phi}_{rr}s_r + \widetilde{\gamma}_{11}n_1. \tag{4.52}
$$

This helps in the computing the SINR at the ith equivalent channel

$$
\alpha_i = \frac{\sigma_{ss}^2|[\widetilde{\gamma}\widetilde{\psi}\widetilde{\phi}]_{i,i}|^2}{\widetilde{\lambda}_i^{-1}\left(|\widetilde{\gamma}_{i,i}\widetilde{\psi}_{i,i+1}\widetilde{\phi}_{i+1,i+1}| + |\widetilde{\gamma}_{i,i}\widetilde{\psi}_{i,i+2}\widetilde{\phi}_{i+2,i+2}| + \cdots + |\widetilde{\gamma}_{i,i}\widetilde{\psi}_{i,r}\widetilde{\phi}_{r,r}|\widetilde{\gamma}_{i,i}n_i\right)^2} \tag{4.53}
$$

where $i, j \leq Min(M_r, M_t)$. Recall that $\widetilde{\lambda}_i$ is the ith eigenvalue of $\widetilde{\mathbf{H}}$.

Finally we may write the channel capacity, as derived in Chapter 3 and shown in (3.20), with the precoder and decoder computed based on imperfect CSI. The channel capacity may be written as

$$
\mathrm{C_{MIMO}} = \max_{\mathbf{R}_{\mathrm{ss}}} \log_2\left(\det\left(\left(\widetilde{\mathbf{G}}\mathbf{R}_{\mathrm{nn}}\widetilde{\mathbf{G}}^{\mathrm{H}}\right)^{-1}\left(\widetilde{\mathbf{G}}\mathbf{H}\widetilde{\mathbf{F}}\mathbf{R}_{\mathrm{ss}}\widetilde{\mathbf{F}}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\widetilde{\mathbf{G}}^{\mathrm{H}} + \widetilde{\mathbf{G}}\mathbf{R}_{\mathrm{nn}}\widetilde{\mathbf{G}}^{\mathrm{H}}\right)\right)\right). \tag{4.54}
$$

4.8 Gaussian Approximated BER and Capacity

Knowledge of the distribution of the CSI errors may be useful in simplifying the performance equations. If the CSI errors are found to be Gaussian in their distribution then they may be treated

in the same manner in which the AWGN is treated with probability of error and the type of detector. This simplifies the performance analysis greatly.

It is desirable to know the distribution of the CSI errors of $\widetilde{\mathbf{H}}$. Histograms of the diagonal and off diagonal elements of the equivalent channel model will be shown. Recall how in the event of perfect CSI, the equivalent channel model, excluding diagonal matrices $\mathbf{\Phi}$ and $\mathbf{\Gamma}$, reduces the to an $M_r \times M_r$ identity as shown in (3.18). The imperfect channel model shown in (4.13) differs from (3.18) only in the inclusion of errors. For small errors, the imperfect channel model of (4.13) should be strongly diagonal with small values on all of the off-diagonal elements. For this reason, the use of a histogram will aid in the development of the results by showing the probability distribution of the diagonal and off-diagonal terms.

A simulation was created which begins by creating 1,000,000 random $\mathbf{H}$ and $\widetilde{\mathbf{H}}$ matrices and computes (4.13). Histograms for the elements of $\widetilde{\mathbf{\Psi}}$ are plotted for various values of σ_p^2. The histograms are plotted ($f\left(p\right)$ and p on the x- and y- axes, respectively) in Fig. 4.4 and Fig. 4.5 for the diagonal and off-diagonal elements, respectively. The mean and variance of the PDF's in Fig. 4.4 are 0.5883, 0.7692, 0.9091 and 0.0154, 0.0112 and 0.0052, respectively. Notice in Fig. 4.4, that as the variance of the CSI error decreases, the distribution of the diagonal elements approaches 1, as expected. The mean and variance of the PDF's in Fig. 4.5 are 0.0001248, 0.0000062, 0.0001339 and 0.0154, 0.0113 and 0.0052, respectively.

Additionally, for each different value of σ_p^2 in Fig. 4.5, the distribution is centered around 0. The PDFs of Fig. 4.4 approaching 1, and the PDFs of Fig. 4.5 remain centered around 0, which is expected. For $\sigma_p^2 = 0$, the PDF would be centered at 1.

If the CSI errors of the overall channel $\widetilde{\mathbf{\Psi}}$ may be shown to be Gaussian then existing performance measures that have been derived in previous works will be applicable to our results. The PDF's shown in Fig. 4.4 display how as the CSI error decreases the mean becomes more closely centered around zero and the variance decreases. Similarly the PDF's of Fig. 4.5 display how an off-diagonal element of the overall channel $\widetilde{\mathbf{\Psi}}$ matrix becomes more closely centered around zero.

Given this information that the CSI error resembles that of Gaussian noise is not enough but will be supported by mathematical expressions. By reviewing the received signal $\hat{\mathbf{s}}$, shown in (4.51), the CSI error may be approximated and a new estimated signal may be shown. Based on

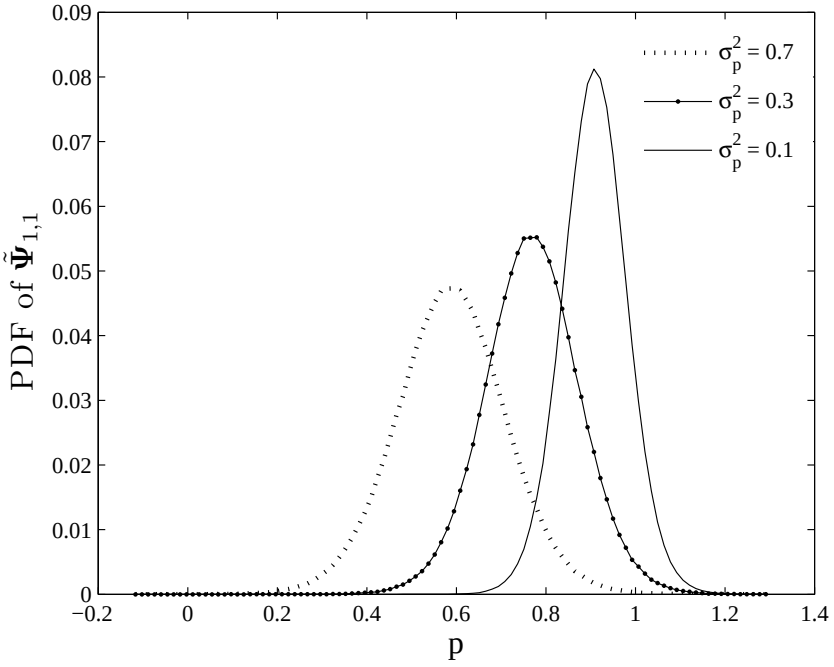


Figure 4.4: PDF of the (1, 1) element of $\widetilde{\mathbf{\Psi}}$ for a 2×6 channel model.

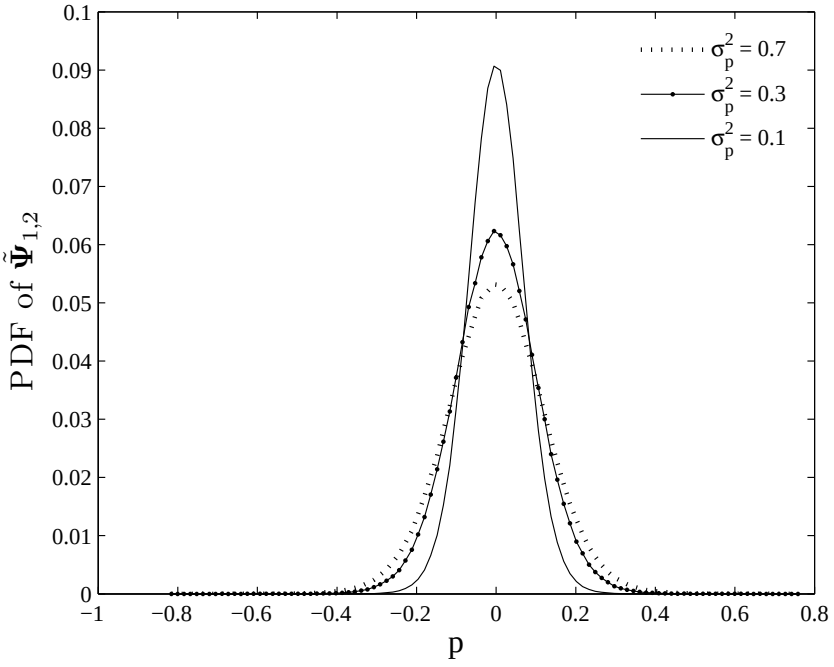


Figure 4.5: PDF of the (1, 2) element of $\widetilde{\mathbf{\Psi}}$ for a 2×6 channel model.

the simulations resulting in the PDF's of Figs. 4.4 and 4.5 the terms due to interfering signals and the AWGN may be lumped into a single AWGN term u, as follows,

$$\hat{\mathbf{s}}' = \begin{bmatrix} \widetilde{\gamma}_{11}\widetilde{\psi}_{11}\widetilde{\phi}_{11}\mathbf{s}_1 + u_1 \\ \widetilde{\gamma}_{22}\widetilde{\psi}_{22}\widetilde{\phi}_{22}\mathbf{s}_1 + u_2 \\ \vdots \\ \widetilde{\gamma}_{rr}\widetilde{\psi}_{rr}\widetilde{\phi}_{rr}\mathbf{s}_1 + u_r \end{bmatrix}. \tag{4.55}$$

Note that u is white Gaussian noise since the interfering s and n (in row 1 s_2 through s_r) are Gaussian. The mean and variance of the interfering terms that are approximated as Gaussian are denoted as (using u_1 as an example)

$$\mathrm{E}\left[\mathbf{u}_1\right] = \widetilde{\gamma}_{22}\widetilde{\psi}_{22}\widetilde{\phi}_{22}\mathrm{E}\left[\mathbf{s}_2\right] + \cdots + \widetilde{\gamma}_{\mathrm{rr}}\widetilde{\psi}_{\mathrm{rr}}\widetilde{\phi}_{\mathrm{rr}}\mathrm{E}[\mathbf{s}_{\mathrm{r}}] + \mathrm{E}\left[\mathrm{n}_1\right]$$

$$\mathrm{Var}\left(\mathbf{u}_1\right) = \left(\widetilde{\gamma}_{11}\widetilde{\psi}_{12}\widetilde{\phi}_{22}\right)^2 \mathrm{Var}\left(\mathbf{s}_2\right) + \cdots + \left(\widetilde{\gamma}_{11}\widetilde{\psi}_{1\mathrm{r}}\widetilde{\phi}_{\mathrm{rr}}\right)^2 \mathrm{Var}\left(\mathbf{s}_{\mathrm{r}}\right) + \mathrm{Var}\left(\mathrm{n}_1\right).$$

Recall that $\mathrm{E}\left(\mathrm{nn}^{\mathrm{H}}\right) = \mathbf{R}_{\mathrm{nn}} = \sigma_{\mathrm{nn}}\mathbf{I}$ and $\mathrm{E}\left(\mathrm{ss}^{\mathrm{H}}\right) = \mathbf{R}_{\mathrm{ss}} = \sigma_{\mathrm{ss}}\mathbf{I}$

Now (4.52) may be rewritten as in a similar fashion as (4.55). Again using the first subchannel,

$$A_1' = \widetilde{\gamma}_{11}\widetilde{\psi}_{11}\widetilde{\phi}_{11}s_1 + u_1. \tag{4.56}$$

Next equation (4.53) may be rewritten as follows,

$$\alpha' = \frac{\sigma_{ss}^2|[\widetilde{\gamma}\widetilde{\psi}\widetilde{\phi}]_{1,1} + [\widetilde{\gamma}\widetilde{\psi}\widetilde{\phi}]_{2,2} + \cdots + [\widetilde{\gamma}\widetilde{\psi}\widetilde{\phi}]_{r,r}|^2}{\widetilde{\lambda}_r^{-1}\left(u_r\right)^2}. \tag{4.57}$$

This new SINR equation will in turn lead to a probability of error equation.

We derive the BER based upon a 16-QAM constellation. The M-QAM constellation was selected based upon its high spectral efficiency in AWGN. Unfortunately, the performance is poor because of channel amplitude and phase variation. Beginning with the probability of error for an M-QAM system we obtain [54],

$$P_{\sqrt{M}} = 2\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3\alpha'}{M-1}}\right). \tag{4.58}$$

From the probablility of error the probability of a symbol error for the system herein may be defined as

$$P_M = 1 - \left(1 - P_{\sqrt{M}}\right)^2 \tag{4.59}$$

To complete the performance equations of the Gaussian Approximation case we now show the mathematical expression for capacity. Similar to the previous capacity equation shown in (4.54) for the case of imperfect CSI we have

$$\mathrm{C}'_{\mathrm{MIMO}} = \max_{\mathbf{R}_{\mathrm{ss}}} \log_2 \left(\det \left(\left(\widetilde{\mathbf{G}} \mathbf{R}_{\mathrm{uu}} \widetilde{\mathbf{G}}^{\mathrm{H}} \right)^{-1} \left(\widetilde{\mathbf{G}} \mathbf{H} \widetilde{\mathbf{F}} \mathbf{R}_{\mathrm{ss}} \widetilde{\mathbf{F}}^{\mathrm{H}} \mathbf{H}^{\mathrm{H}} \widetilde{\mathbf{G}}^{\mathrm{H}} + \widetilde{\mathbf{G}} \mathbf{R}_{\mathrm{uu}} \widetilde{\mathbf{G}}^{\mathrm{H}} \right) \right) \right). \quad (4.60)$$

4.9 Conclusion

This chapter has shown theoretical analysis of the impact of errors in CSI used to compute a decoder/precoder in a MIMO system. This analysis is complete because it applies to full rank channels with both transmitter and receiver diversity. In the next chapter the focus is shifted to the performance of the equations derived within this chapter and their robustness to channel errors, or lack thereof.

Chapter 5

Simulation Results of Linear Precoders/Decoders with Imperfect CSI

The simulation results of linear precoders/decoders determined in the presence of imperfect CSI is analyzed herein. Results are shown that describe the distribution of the error caused by the imperfect CSI by way of BER and capacity plots for various antenna configurations.

5.1 CSI Error

To begin the performance analysis, the first step is to quantify what happens to the vectors of the channel matrix, $\mathbf{H}$, when an error matrix such as $\mathbf{\Delta H}$, is added to it. This first set of plots is shown to verify how in a three dimensional case the angles of the column vectors comprising the $\mathbf{H}$ and $\widetilde{\mathbf{H}}$ matrices differ. By taking a pair of $\mathbf{H}$ and $\widetilde{\mathbf{H}}$ matrices and performing the SVD a comparison between the disturbed and undisturbed matrices may be made. From this comparison the angles between the individual vectors of $\mathbf{U}$, $\widetilde{\mathbf{U}}$, $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ may be shown.

The simulation begins with a three transmit and receive antenna configuration. The random $\mathbf{H}$ and $\mathbf{\Delta H}$ matrices are selected from a Gaussian distribution with zero mean and unit variance. After performing the SVD on the $\mathbf{H}$ and $\widetilde{\mathbf{H}}$, the individual vectors of the $\mathbf{U}$, $\widetilde{\mathbf{U}}$, $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ pairs are plotted against one another with a measure of the angles between the individual vectors listed.

The angles between each of the vectors and their imperfect counterpart are computed using the equation

$$\beta_i = \cos^{-1}\left(\frac{\mathbf{u}_i^H \widetilde{\mathbf{u}}_i}{\|\mathbf{u}_i\| \, \|\widetilde{\mathbf{u}}_i\|}\right) \times \frac{180}{\pi}$$

and is shown in Fig. 5.1. For a CSI error variance of $\sigma_p^2 = 0.01$ the measure of angular error between the vectors of $\mathbf{U}$ and $\widetilde{\mathbf{U}}$ are $\beta_1 = 3.8$, $\beta_2 = 2.1$ and $\beta_3 = 3.1$. For the same CSI error variance the measure of angular error between the vectors of $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ are $\beta_1 = 1.9$, $\beta_2 = 1.8$ and $\beta_3 = 2.4$. This plot showing the angles between the $\mathbf{U}$, $\widetilde{\mathbf{U}}$, $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ pairs is included to reinforce the discussions of Fig. 4.3 in Section 4.4.2 and in Section 4.5.2.

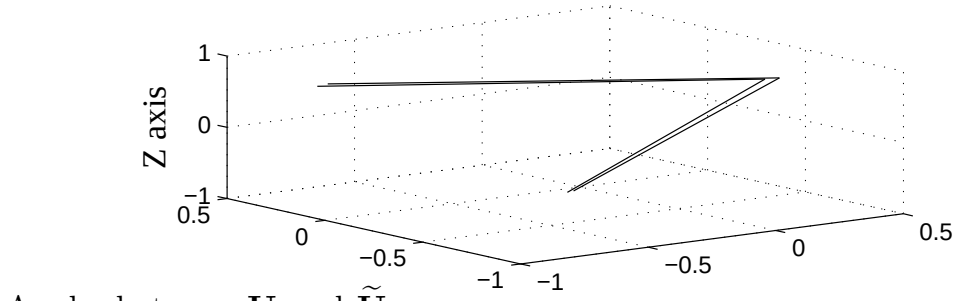
Additional examples of the angular error between the vectors of $\mathbf{U}$, $\widetilde{\mathbf{U}}$ and $\mathbf{V}$, $\widetilde{\mathbf{V}}$ for a larger σ_p^2 are shown in Fig. 5.2 and 5.3. Fig. 5.2 displays the measure of angular error for a CSI error variance of $\sigma_p^2 = 0.05$. The angular error between the vectors of $\mathbf{U}$ and $\widetilde{\mathbf{U}}$ are $\beta_1 = 7.2$, $\beta_2 = 7.2$ and $\beta_3 = 1.3$. The angular error between the vectors of $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ are $\beta_1 = 10.3$, $\beta_2 = 9.0$ and $\beta_3 = 5.2$. Fig. 5.3 displays the measure of angular error for a CSI error variance of $\sigma_p^2 = 0.1$. The angular error between the vectors of $\mathbf{U}$ and $\widetilde{\mathbf{U}}$ are $\beta_1 = 11.6$, $\beta_2 = 18.8$ and $\beta_3 = 20.2$. The angular error between the vectors of $\mathbf{V}$ and $\widetilde{\mathbf{V}}$ are $\beta_1 = 4.3$, $\beta_2 = 28.2$ and $\beta_3 = 28.0$.

From Figs. 5.1, 5.2 and 5.3, it is evident that as σ_p^2 (the CSI error variance) increases, the distance between the column vectors of $\mathbf{U}$, $\widetilde{\mathbf{U}}$ and $\mathbf{V}$, $\widetilde{\mathbf{V}}$ increases. Since $\mathbf{U}$, $\mathbf{V}$ and $\widetilde{\mathbf{U}}$, $\widetilde{\mathbf{V}}$ are from the SVD of the channel matrices $\mathbf{H}$ and $\widetilde{\mathbf{H}}$, respectively, this implies that the column vectors between $\mathbf{H}$ and $\widetilde{\mathbf{H}}$ are similarly corrupted. This discussion shows how a small error will alter the vectors of the $\mathbf{H}$ matrix, but it does not describe the type of distribution of the error.

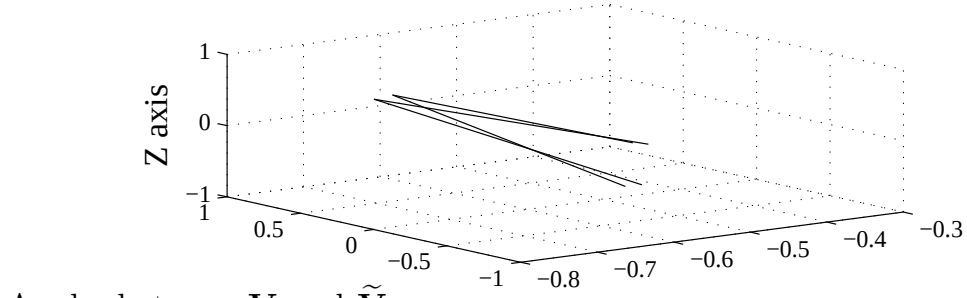
5.1.1 SVD Theorem

The next simulation is presented to show extreme cases of the SVD theorem of Section 4.3 as M_t approaches infinity. Two channel models are compared given different transmit and receive antenna configurations for different values of σ_p^2. The equivalent channel model of (4.13) is compared against the approximation to the equivalent channel model. The approximation to the equivalent channel model is found by substituting (4.28) into (4.23) which yields

$$\widetilde{\mathbf{\Psi}}_{AW} = \sigma_{nn}^{-2} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} \mathbf{\Sigma}_{M_r} \mathbf{V}_{M_r}^H \left(\mathbf{V} + \mathbf{\Delta}\mathbf{H}^H \mathbf{U} \widetilde{\mathbf{\Sigma}}_{M_r}^{-1} + \mathbf{H}^H \mathbf{U} \mathbf{\Delta} \mathbf{\Sigma}_{M_r}^{-1}\right). \tag{5.1}$$

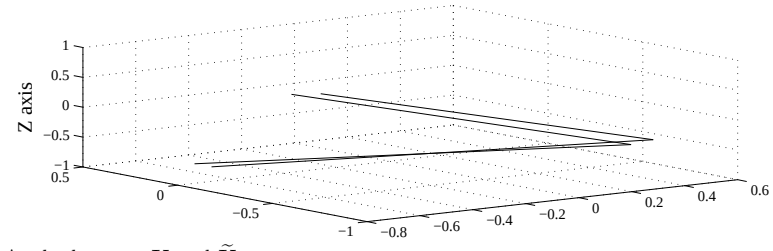


Angles between $\mathbf{U}$ and $\widetilde{\mathbf{U}}$

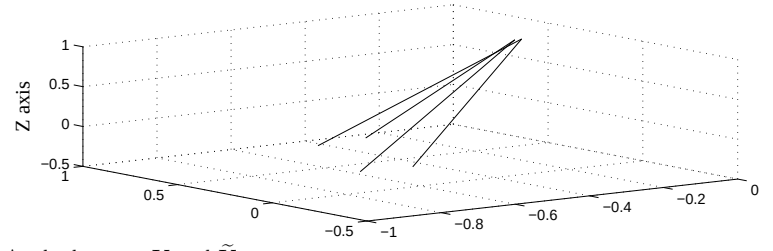


Angles between $\mathbf{V}$ and $\widetilde{\mathbf{V}}$

Figure 5.1: Angles between the vectors of the actual **U** and **V** and error matrices $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{V}}$, respectively.

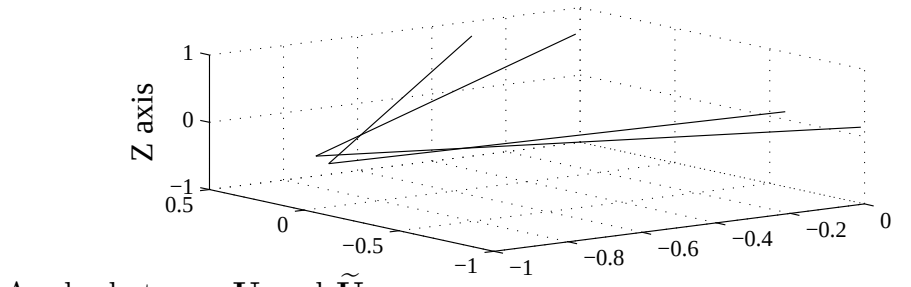


Angles between $\mathbf{U}$ and $\widetilde{\mathbf{U}}$

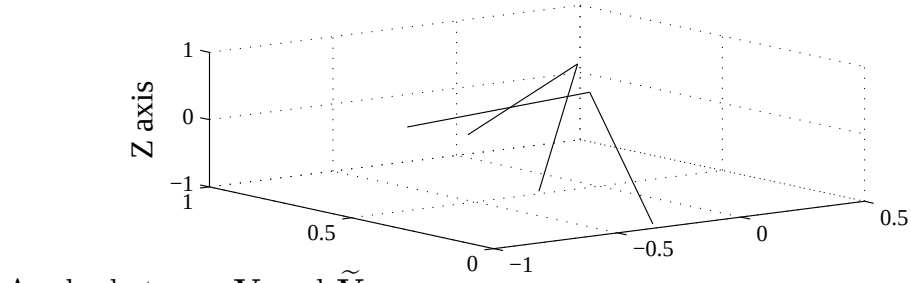


Angles between $\mathbf{V}$ and $\widetilde{\mathbf{V}}$

Figure 5.2: Angles between the vectors of the actual **U** and **V** and error matrices $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{V}}$, respectively.



Angles between $\mathbf{U}$ and $\widetilde{\mathbf{U}}$



Angles between $\mathbf{V}$ and $\widetilde{\mathbf{V}}$

Figure 5.3: Angles between the vectors of the actual **U** and **V** and error matrices $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{V}}$, respectively.

The two norm of the difference between the two channel models is then averaged over 5000 trials and plotted, as shown in Fig. 5.4, where $\widetilde{\mathbf{\Psi}}$, $\widetilde{\mathbf{\Psi}}_{AW}$ indicate the equivalent channel and the approximation to the wide equivalent channel model, respectively. Each point on the x axis indicates a different value of σ_p^2 and the y axis indicates the average of the two norm of the difference between the two channel models, which indicates the amount of error introduced by the $\mathbf{\Delta H}$ term. The different antenna configurations are listed in the upper left corner of the plot. The plot shows that as the channel model becomes wider, the closer the approximation becomes for various values of σ_p^2. In particular notice that the approximation is not good for square matrices, but a 5×7 has a significantly better approximation than the 5×5 and is nearly as good as very wide matrices such as 5×15, especially for small σ_p^2.

The SVD theorem is shown to hold true for a tall channel model as well, when M_r approaches infinity. The setup for the simulation of the tall channel is the same for the case of the wide channel. The theoretical channel model for this simulation was found by substituting (4.44) into

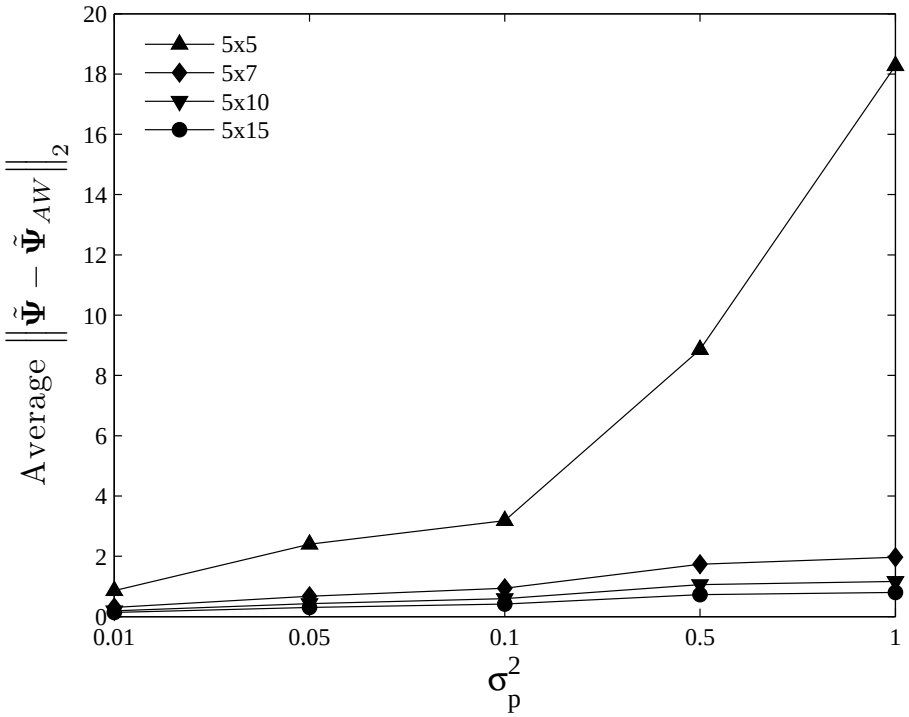


Figure 5.4: SVD theorem example.

(4.40) resulting in

$$\widetilde{\mathbf{\Psi}}_{AT} = \sigma_{nn}^{-2}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1}\left(\mathbf{U} + \mathbf{\Delta}\mathbf{H}\mathbf{V}\widetilde{\mathbf{\Sigma}}_{M_t}^{-1} + \mathbf{H}\mathbf{V}\mathbf{\Delta}\mathbf{\Sigma}_{M_t}^{-1}\right)_{M_t}^{H}\mathbf{U}_{M_t}\mathbf{\Sigma}_{M_t}. \tag{5.2}$$

The result is averaged over 5000 trials and is shown in Fig. 5.5, where $\widetilde{\mathbf{\Psi}}_{AT}$ indicates the approximation to the tall equivalent channel model. Again, each point on the x axis indicates a different value of σ_p^2 and the y axis indicates the average of the two norm of the difference between the two channel models. The different antenna configurations are listed in the upper left corner of the plot. This plot reinforces the idea that as the channel matrix becomes taller, the better the SVD theorem holds. Since the resolution of the non-square antenna configurations is poor, the simulation was repeated but with the square antenna configuration left out. The result of that simulation is shown in Fig. 5.6.

Although the theorem holds for the case of a tall channel, the approximation is not as good as the wide case, notice how much greater the error is in Fig. 5.5 than that of Fig. 5.4. This indicates that setting $\mathbf{U} = \widetilde{\mathbf{U}}$ results in a better approximation to the equivalent channel model for wide

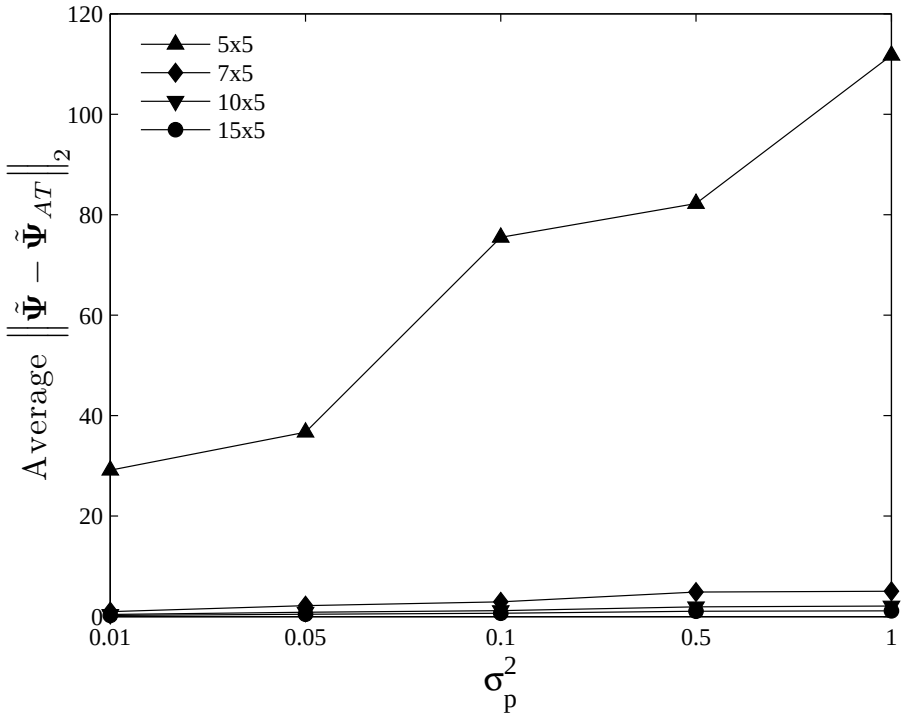


Figure 5.5: SVD theorem example.

channel matrices than setting $\mathbf{V} = \widetilde{\mathbf{V}}$ for the equivalent channel model for tall channel matrices. This is shown to hold true in a similar simulation without the approximations for $\mathbf{\Delta V}$ and $\mathbf{\Delta U}$ as shown in (4.28) and (4.44), which will be shown next.

By looking closer at the equivalent channel model (4.13) assists in the investigation to why the assumption $\mathbf{V} = \widetilde{\mathbf{V}}$ performs worse than the assumption $\mathbf{U} = \widetilde{\mathbf{U}}$. The equivalent channel model is presented again

$$\begin{aligned}
\widetilde{\Psi} &= \widetilde{\Lambda}^{-1}\widetilde{\mathbf{V}}^H\widetilde{\mathbf{H}}^H\mathbf{R}_{nn}^{-1}\mathbf{H}\widetilde{\mathbf{V}} \\
&= \sigma_{nn}^{-1}\widetilde{\mathbf{\Sigma}}^{-2}\widetilde{\mathbf{V}}^H\left(\widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^H\right)^H\left(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\right)\widetilde{\mathbf{V}} \\
&= \sigma_{nn}^{-1}\widetilde{\mathbf{\Sigma}}^{-2}\widetilde{\mathbf{\Sigma}}\underbrace{\widetilde{\mathbf{U}}^H\mathbf{U}}\mathbf{\Sigma}\underbrace{\mathbf{V}^H\widetilde{\mathbf{V}}}. \qquad (5.3)
\end{aligned}$$

From this point one of the two assumptions may be made which leads in two different directions, as indicated by the underbraces, depending on the desired antenna configuration. If a wide antenna configuration is desired, the assumption $\mathbf{U} = \widetilde{\mathbf{U}}$ may be made and (5.3) may be simplified further

as

$$\begin{aligned}\widetilde{\mathbf{\Psi}}_W &= \tilde{\mathbf{\Sigma}}^{-2}\tilde{\mathbf{\Sigma}}\mathbf{U}^H\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\tilde{\mathbf{V}} \\ &= \tilde{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}\mathbf{V}^H\tilde{\mathbf{V}}.\end{aligned}$$

Now if a tall antenna configuration is desired, the contrary assumption $\mathbf{V} = \tilde{\mathbf{V}}$, may be used to simplify (5.3) as follows,

$$\begin{aligned}\widetilde{\mathbf{\Psi}}_T &= \tilde{\mathbf{\Sigma}}^{-2}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{U}}^H\mathbf{\Sigma}\mathbf{V}^H\mathbf{V} \\ &= \tilde{\mathbf{\Sigma}}^{-1}\tilde{\mathbf{U}}^H\mathbf{U}\mathbf{\Sigma}.\end{aligned}$$

Now the desire is to investigate why the error $\left\|\widetilde{\mathbf{\Psi}} - \widetilde{\mathbf{\Psi}}_{AT}\right\|_2$ is greater than the error $\left\|\widetilde{\mathbf{\Psi}} - \widetilde{\mathbf{\Psi}}_{AW}\right\|_2$. Due to the symmetry of the SVD one would guess that they would be relatively close. A separate simulation was setup that mimicked what is shown in Fig. 5.4 and Fig. 5.5, but instead computed

$$\epsilon_W = \left\|\widetilde{\mathbf{\Psi}} - \widetilde{\mathbf{\Psi}}_W\right\|_2 \text{ and } \epsilon_{\mathrm{T}} = \left\|\widetilde{\mathbf{\Psi}} - \widetilde{\mathbf{\Psi}}_{\mathrm{T}}\right\|_2. \tag{5.4}$$

The result is shown in Fig. 5.7, where all of the antenna configurations used are square. Square antenna configurations were used in order to present an unbiased analysis. Notice that the performance of the equivalent channel $\widetilde{\mathbf{\Psi}}_T$, still performs worse than $\widetilde{\mathbf{\Psi}}_W$, for large values of σ_p^2. This shows that the assumption $\mathbf{U} = \tilde{\mathbf{U}}$ is initially poor, but as the antenna configuration becomes taller, the better the approximation holds, which leads to improved performance.

5.1.2 Distribution of CSI Errors

Despite the insight provided in the PDF's of Section 4.8, the analysis is lacking because it does not take into account the statistics of the other elements in the equivalent channel model, $\widetilde{\mathbf{\Psi}}$. The statistics of the other elements differ slightly. The next analysis quantifies a relationship between the mean and variance of each row versus σ_p^2 for different antenna configurations.

In order to describe the mean of the overall channel model, $\widetilde{\mathbf{\Psi}}$ and compare it with σ_p^2 a simulation was setup to show how individual antenna configurations vary with different values of σ_p^2. For a given antenna configuration, the simulation begins by creating random $\mathbf{H}$ and $\mathbf{\Delta H}$

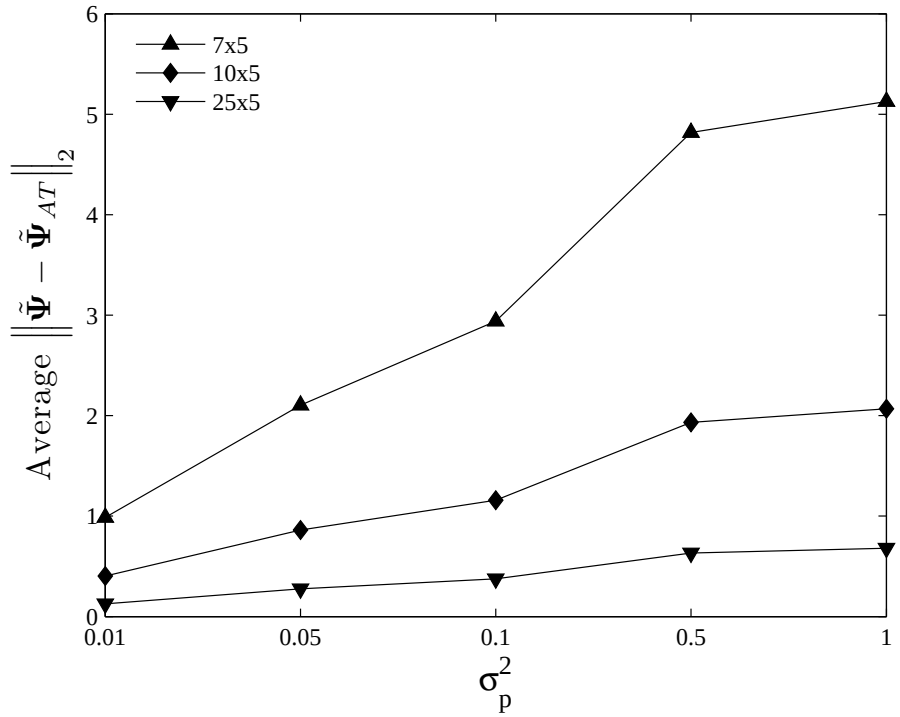


Figure 5.6: SVD theorem example.

matrices for a given σ_p^2 and $\widetilde{\mathbf{\Psi}}$ is computed. The simulation then repeats for 10,000 iterations, storing the individual entries of the overall channel matrix, $\widetilde{\mathbf{\Psi}}$. From the stored values of the overall channel matrix the mean and variance of each entry is computed. Excluding the diagonal element, each entry is then averaged row-wise to produce one mean and one variance for each row of $\widetilde{\mathbf{\Psi}}$. The simulation then repeats for increasing values of σ_p^2.

Specific cases of wide antenna configurations, such as 2×6, 4×10 and 6×12 with σ_p^2 values of $(0, 0.1, 0.2, ...1.0)$ were evaluated. The antenna configurations were selected to also take into consideration the ratio of receive antennas to transmit antennas, which is 0.3333, 0.4 and 0.5 for the aforementioned cases, respectively. The results are shown in Fig. 5.8, Fig. 5.9 and Fig. 5.10 for the 2×6, 4×10 and 6×12 configurations. From the three figures, it may be observed that as the CSI error, σ_p^2 increases, the mean of the diagonal elements decreases, whereas the mean of the off-diagonal elements stays centered at zero, regardless of the size of the antenna configuration.

Within the same simulation to capture the relationship between the means of the overall channel model versus CSI error, the variances of the overall channel model versus CSI error were

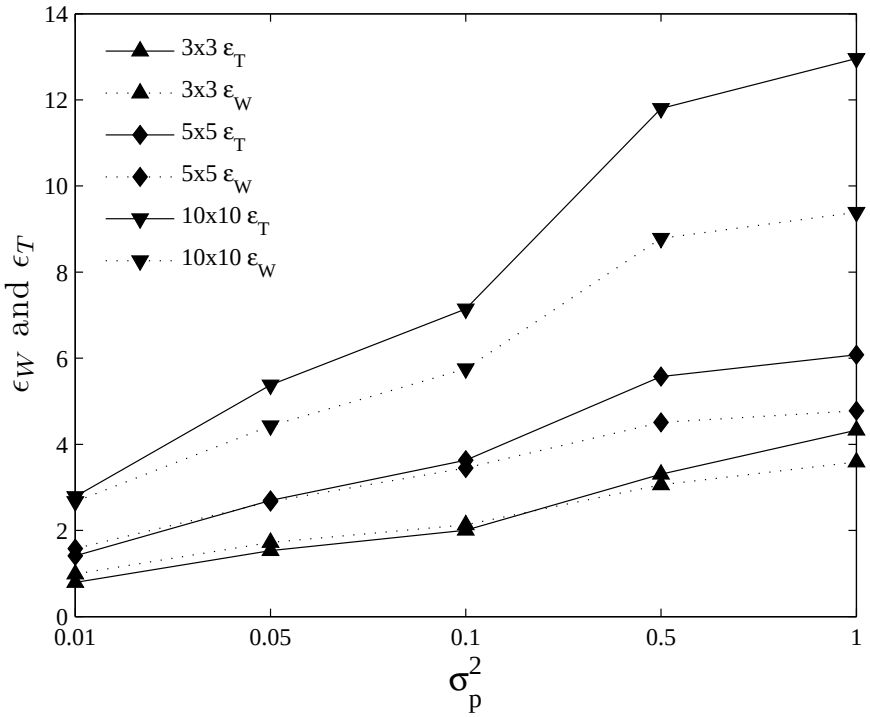


Figure 5.7: SVD theorem example.

captured as well. Figures 5.11, 5.12 and 5.13 show the variances of the diagonal and off diagonal elements of various antenna configurations versus CSI error.

In Fig. 5.11 it is shown that the variance of the diagonal and off diagonal elements increase at the same rate as the CSI error increases. This is an expected result since the Gaussian plots of Figs. 4.4 and 4.5 become wider as the CSI error increases. The same conclusion may be made for the 4×10 and 6×12 cases shown in Figs. 5.12 and 5.13, respectively.

Additional information is revealed in the plots of Figs. 5.12 and 5.13 which shows that different rows of the overall channel model have different variances. This is due to the $\widetilde{\mathbf{\Phi}}$ matrix which is used in the precoder $\widetilde{\mathbf{F}}$ (recall $\widetilde{\mathbf{F}} = \widetilde{\mathbf{V}}\widetilde{\mathbf{\Phi}}$. $\widetilde{\mathbf{\Phi}}$ is computed using the eigenvalues of $\widetilde{\mathbf{H}}$, the eigenvalues are ordered from increasing to decreasing along the diagonal of $\widetilde{\mathbf{\Lambda}}$, the matrix of eigenvalues). This causes more of the power to be allocated to the higher rows found in the equivalent channel model, $\widetilde{\mathbf{\Psi}}$, hence their variance is smaller and the variance of the lower rows is larger. This is why in Fig. 5.12 the variance of the off diagonal elements of row 4 is higher than the other off diagonal rows. Similarly in Fig. 5.13 the off diagonal elements of row 6 are higher than the other off diagonal

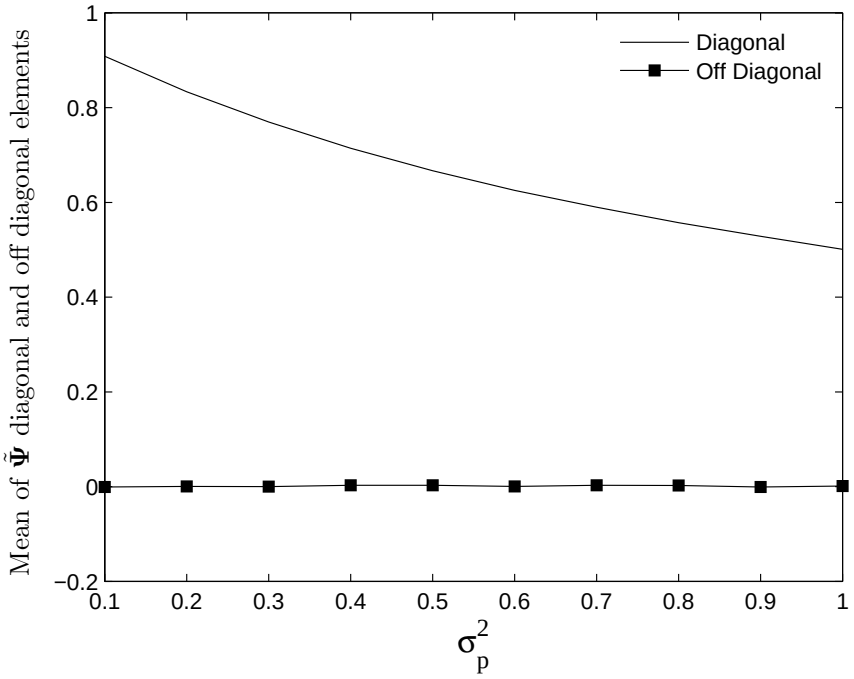


Figure 5.8: Mean versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 2×6 channel model.

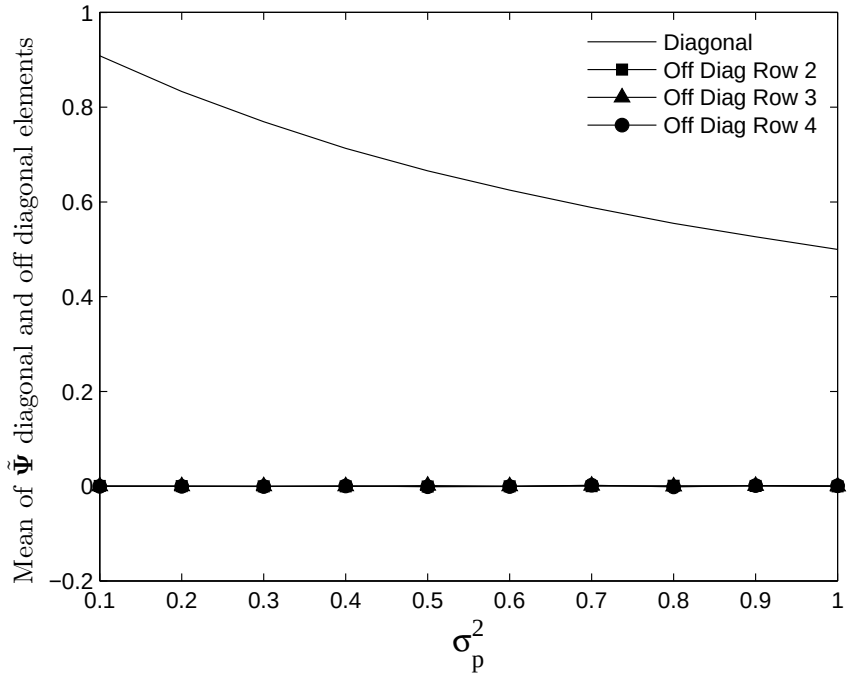


Figure 5.9: Mean versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 4×10 channel model.

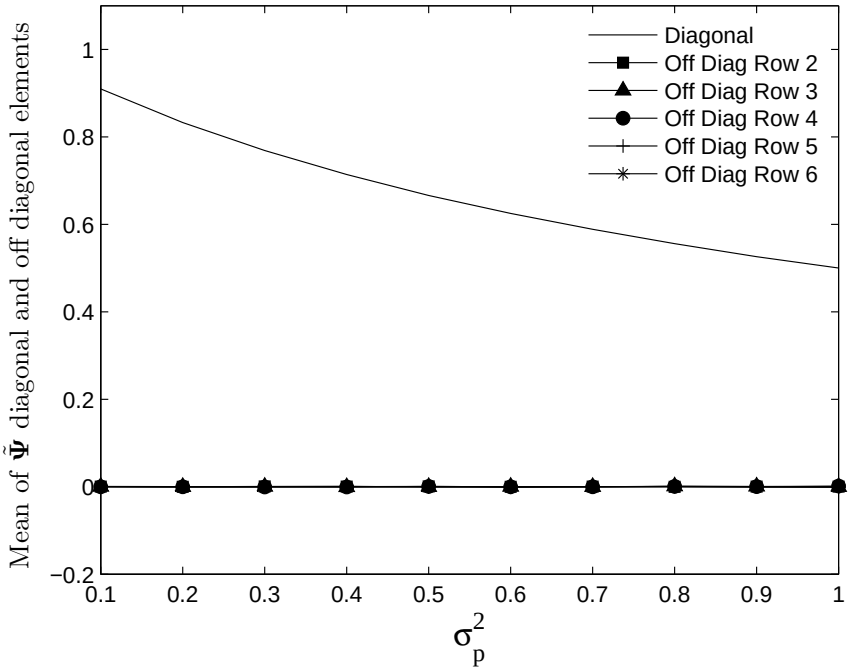


Figure 5.10: Mean versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 6×12 channel model.

rows.

A similar simulation was executed with tall antenna configurations analyzing the means and variances of the overall channel model versus the increasing CSI error. As expected the results were very similar for the cases of 6×2, 10×4 and 12×6. The results for mean versus σ_p^2 are shown below in Figs. 5.14, 5.15, 5.16. The results for variance versus σ_p^2 are shown in 5.17, 5.18 and 5.19.

5.2 BER

One important aspect of a wireless communications system is its' BER performance. Simulations were setup for wide, square, and tall channel models for the designs of Chapter 4, to emphasize the SVD theorem (as both M_r and M_t increase). Simulations were also setup to emphasize the loss in performance as the CSI error (σ_p^2) increased. A flat fading Rayleigh channel was simulated using the MATLAB command RAYLEIGHCHAN, with a sampling frequency of 10kHz and Doppler

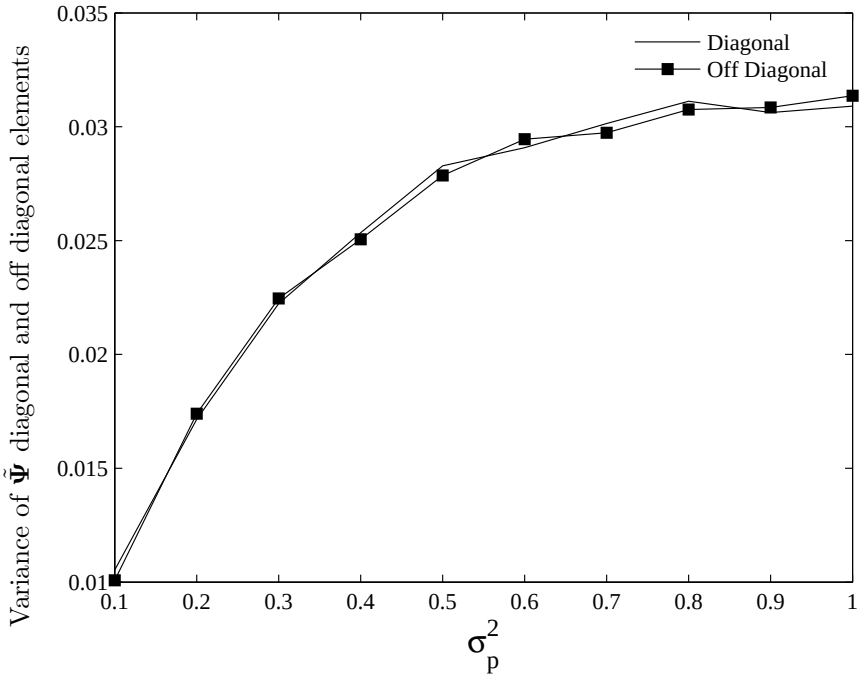


Figure 5.11: Variance versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 2×6 channel model.

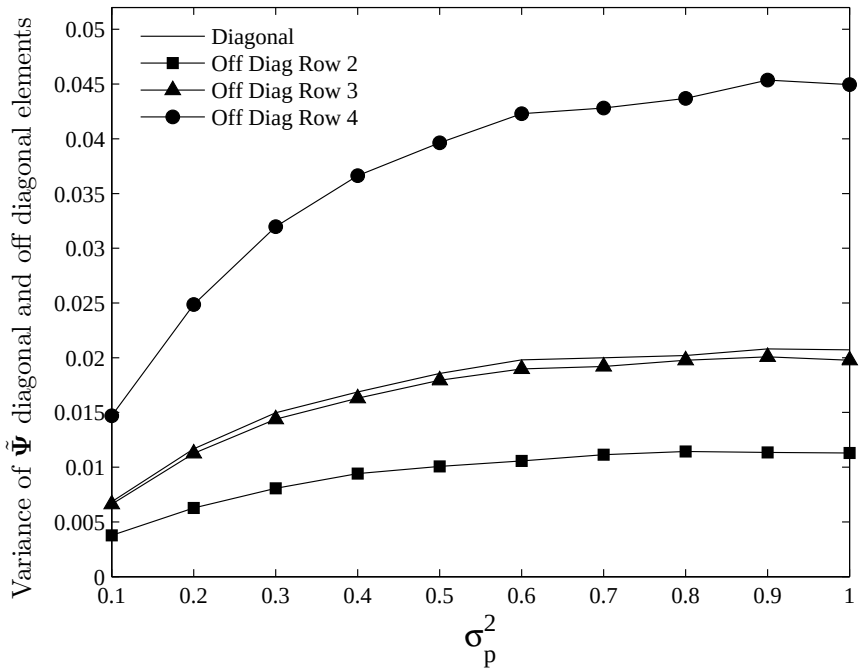


Figure 5.12: Variance versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 4×10 channel model.

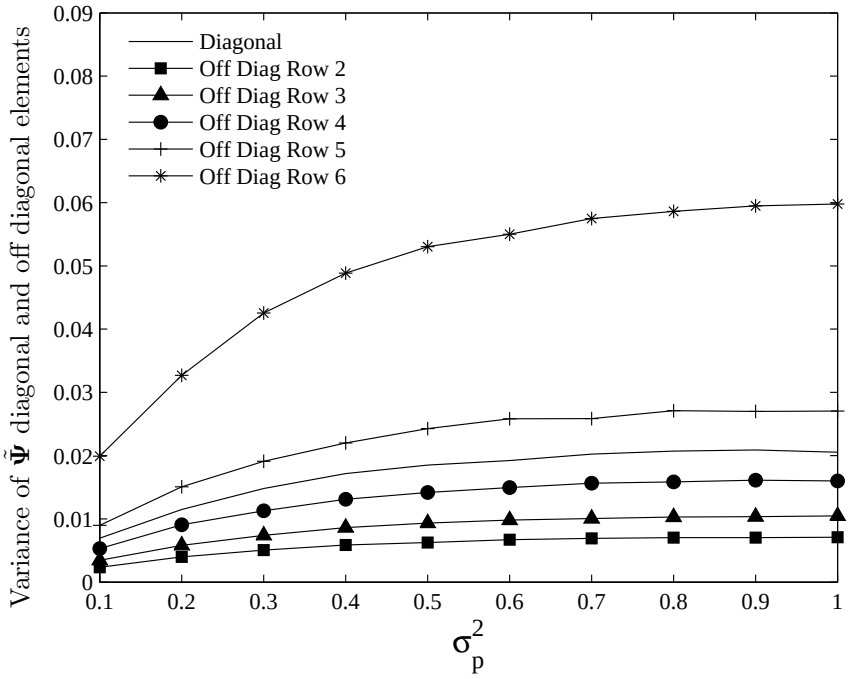


Figure 5.13: Variance versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 6×12 channel model.

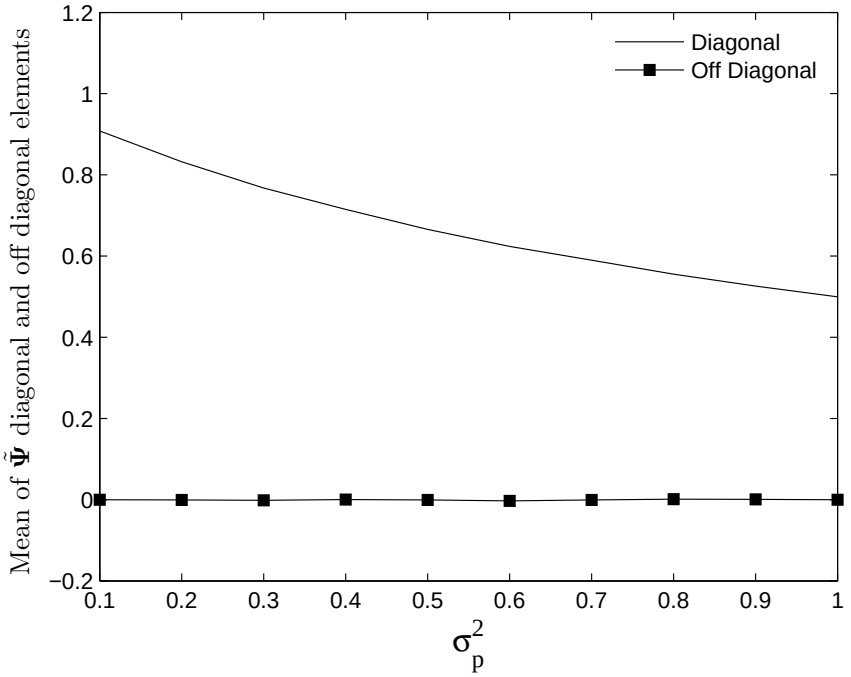


Figure 5.14: Mean versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 6×2 channel model.

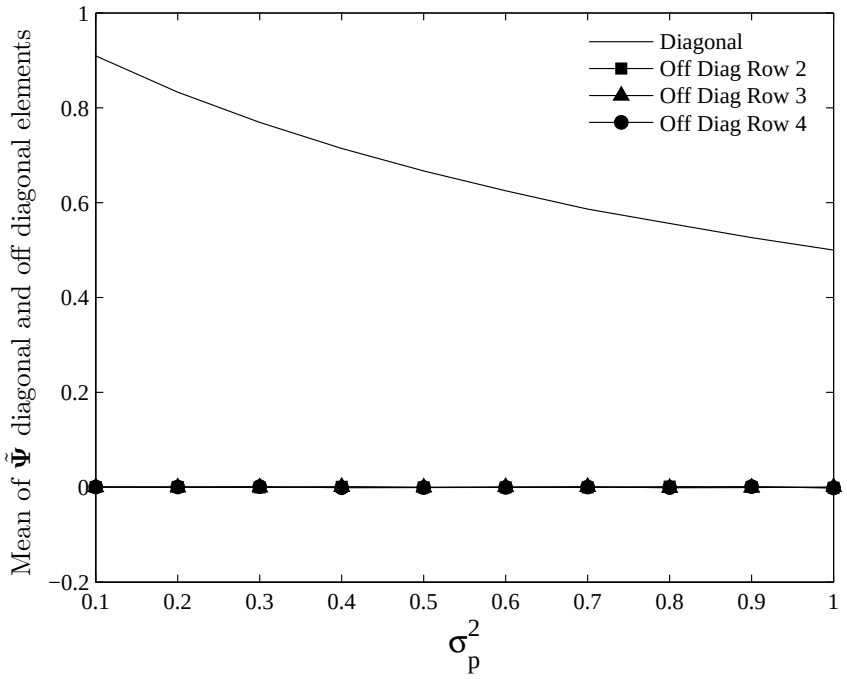


Figure 5.15: Mean versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 10×4 channel model.

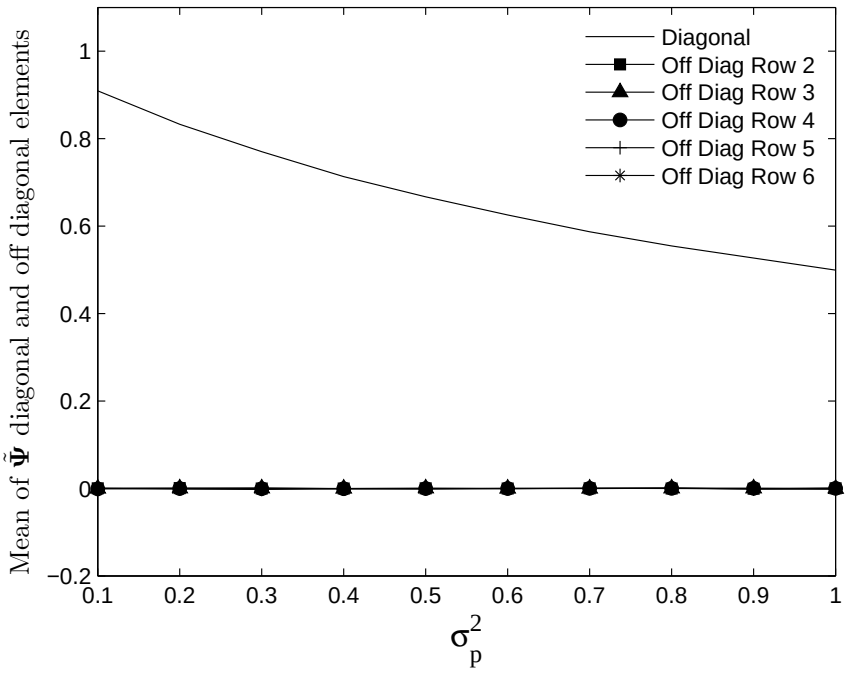


Figure 5.16: Mean versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 12×6 channel model.

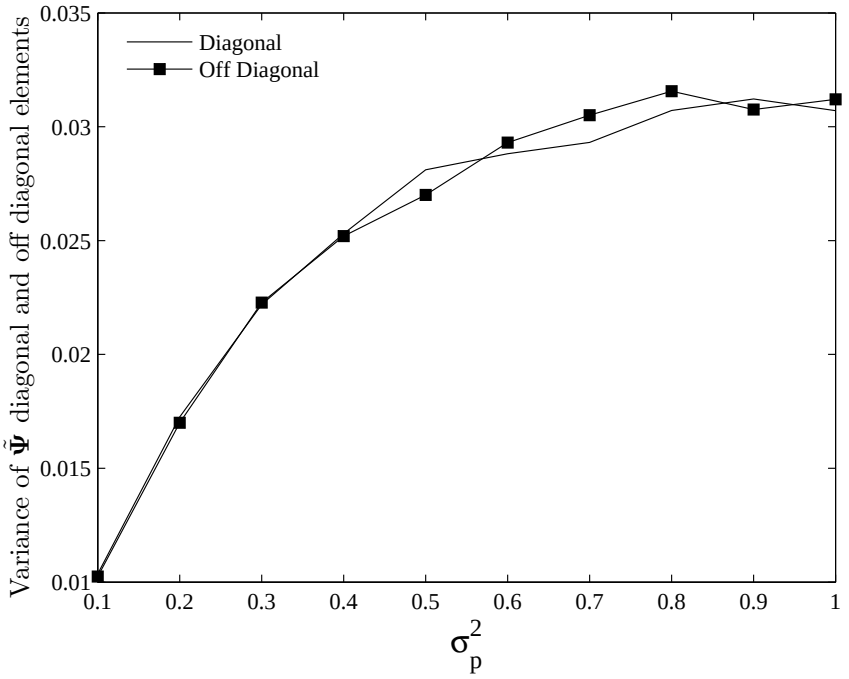


Figure 5.17: Variance versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 6×2 channel model.

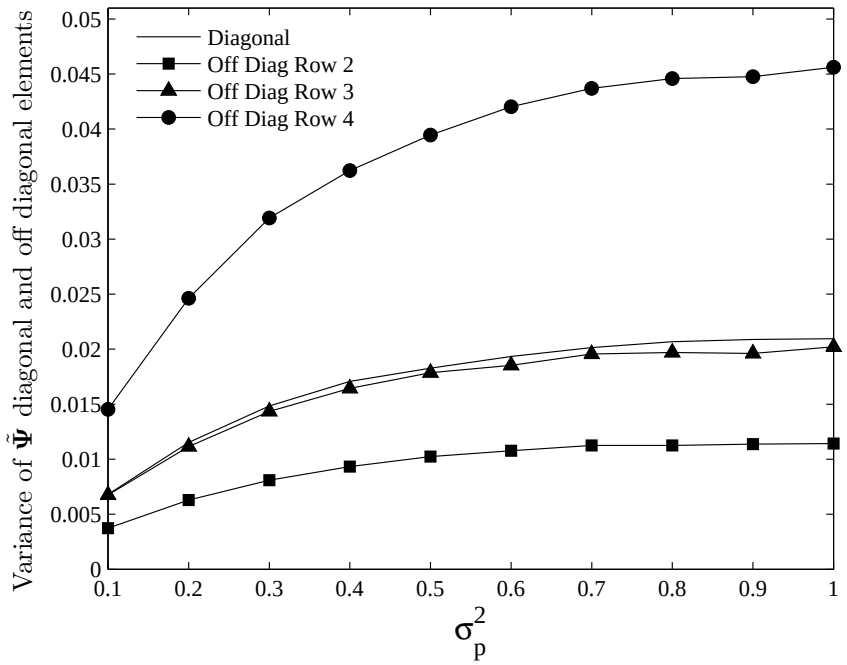


Figure 5.18: Variance versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 10×4 channel model.

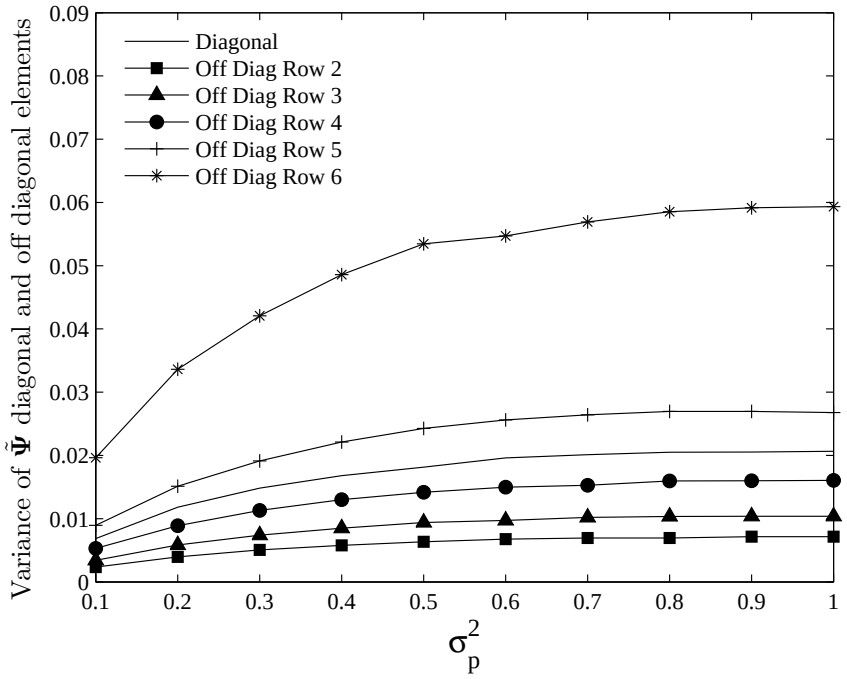


Figure 5.19: Variance versus σ_p^2 of $\widetilde{\mathbf{\Psi}}$ for a 12×6 channel model.

frequency of 80Hz to the symbols. The noise variance is assumed to be unity, since the transmitted symbols are white and the noise is modeled as AWGN. Every transmission period results in 1 symbol being transmitted from each antenna.

In order to be cognizant of the appropriate values to use for the variance of the CSI errors, σ_p^2, a literature search was undertaken on the topic of channel estimation errors. For the simulations of [47], σ_p^2 was set to 0.09, [39] let σ_p^2 be equal to 0.2, 0.6, 0.7 and 0.9, [42] let σ_p^2 = 0.12 and 0.375, [43] let σ_p^2 = 0.02 and 0.625 and [41] varied σ_p^2 between 0 and 1. For this reason various σ_p^2 values were used to simulate CSI error for this section.

The precoder that was used to precode the transmitted symbols (which were already modulated by a 16 QAM constellation) is the same as found in (4.8) using the $\widetilde{\phi}_{ii}$ found in (4.9). This $\widetilde{\phi}_{ii}$ was selected because according to [21] it provides the best compromise between BER and information rate. AWGN corrupts the transmitted signal prior to reception at the receiver where (4.7) is used to decode the received signal. The received signal is then compared with the transmitted signal and the number of bits in error results in the BER curves.

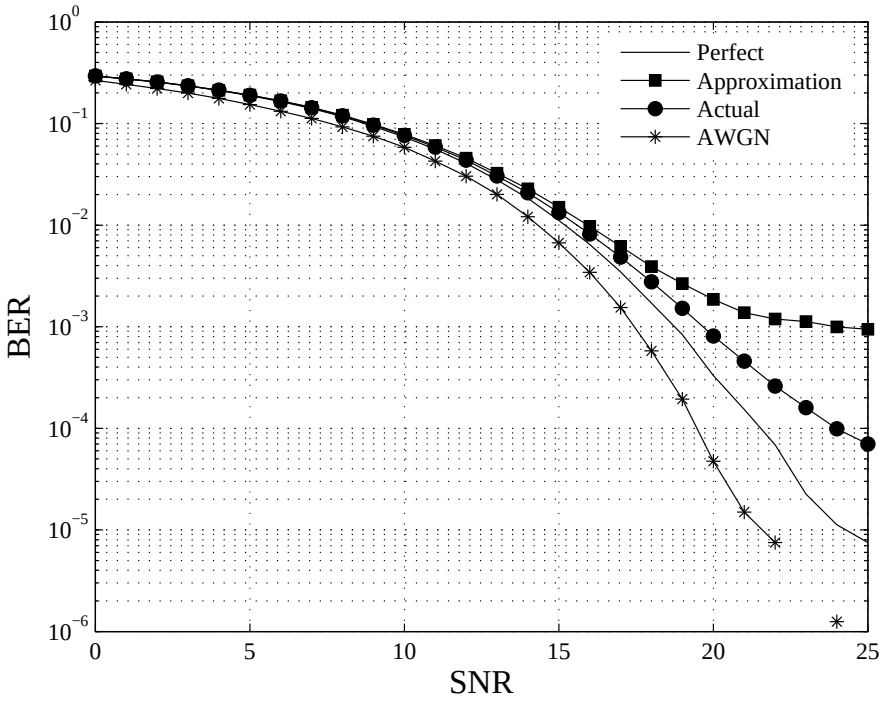


Figure 5.20: BER for a 2×6 channel with $\sigma_p^2 = 0.05$.

The first set of BER curves shown in Fig. 5.20 are for an $M_r \times M_t$ antenna configuration of 2×6. The first of the four curves shown in the figure are the scenario where the CSI is perfect, the second curve uses the approximation of the SVD theorem, the third curve denotes the actual case when the CSI is imperfect. A variance of $\sigma_p^2 = 0.05$ was used to degrade the quality of the CSI. Notice that the two curves with CSI errors are able to perform as well as the curve with no CSI error up to 20dB. This system reaches a BER of 10^{-3} at approximately 18dB, 20dB and 22dB for the perfect CSI, imperfect CSI and the approximation using imperfect CSI, respectively. The fourth curve, which is denoted as AWGN, depicts the BER for an AWGN channel with no fading for a 1×1 channel configuration. It is displayed for comparison purposes only.

The same simulation setup was then used for an increased amount of CSI error. With $\sigma_p^2 = 0.1$, the three curves given (Perfect, Approximate and Actual) in Fig. 5.21 nearly achieve a BER of 10^{-2} at 15dB before the performance degrades. At 10^{-3} the curve with perfect CSI is at 18dB, but the actual curve does not achieve that same level of performance until 25dB and the approximation is not on the plot. This shows a degradation in the performance of the approximation to the wide

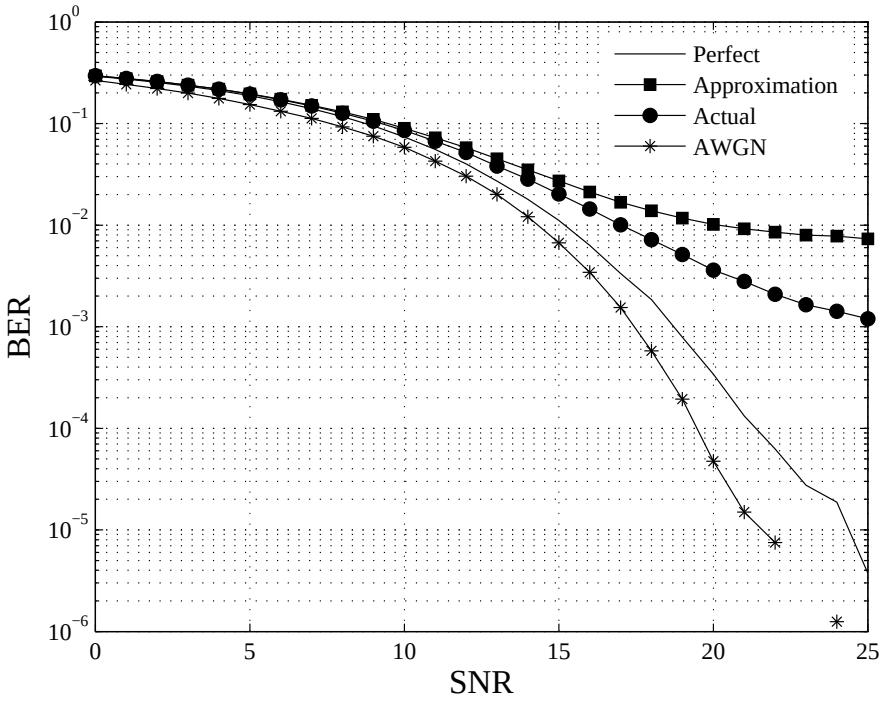


Figure 5.21: BER for a 2×6 channel with $\sigma_p^2 = 0.1$.

channel model when compared to the perfect CSI case.

This issue was explored further with a wider antenna configuration, to investigate how well, if at all, the SVD theorem held as the number of transmit antennas increased. Fig. 5.22 shows a 3×10 with three curves: perfect, approximation and actual with a $\sigma_p^2 = 0.05$. There is a noticeable increase in performance of the approximation since at 10^{-4} there is approximately 1dB of separation between the curve with perfect and the curves with imperfect CSI. The approximation curve of Fig. 5.20 never even achieves a BER of 10^{-4}, which indicates that for the same CSI error variance of $\sigma_p^2 = 0.05$, a wider antenna configuration will perform better, as well as the approximation.

Next, Fig. 5.23 shows a 3×10 antenna configuration with a CSI error variance of $\sigma_p^2 = 0.1$. It is here that the performance gains of the wider antenna configuration break down. At 10^{-2} there is 1dB of separation with the perfect and approximate BER curves. At 10^{-3} a 7dB separation in performance, which shows that with a higher CSI error variance, the gains achieved by wider antenna configuration are lost.

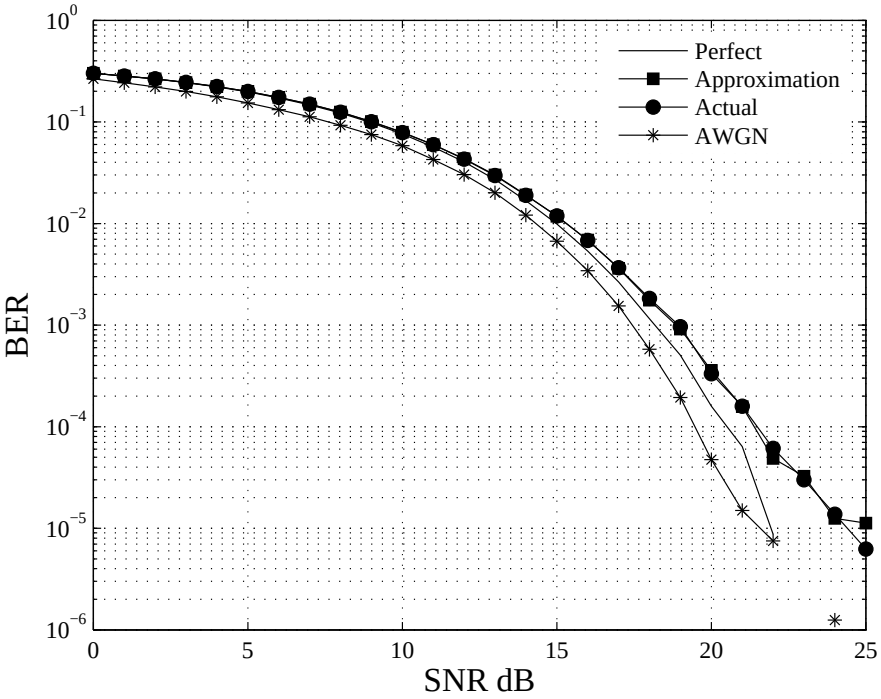


Figure 5.22: BER for a 3×10 channel with $\sigma_p^2 = 0.05$.

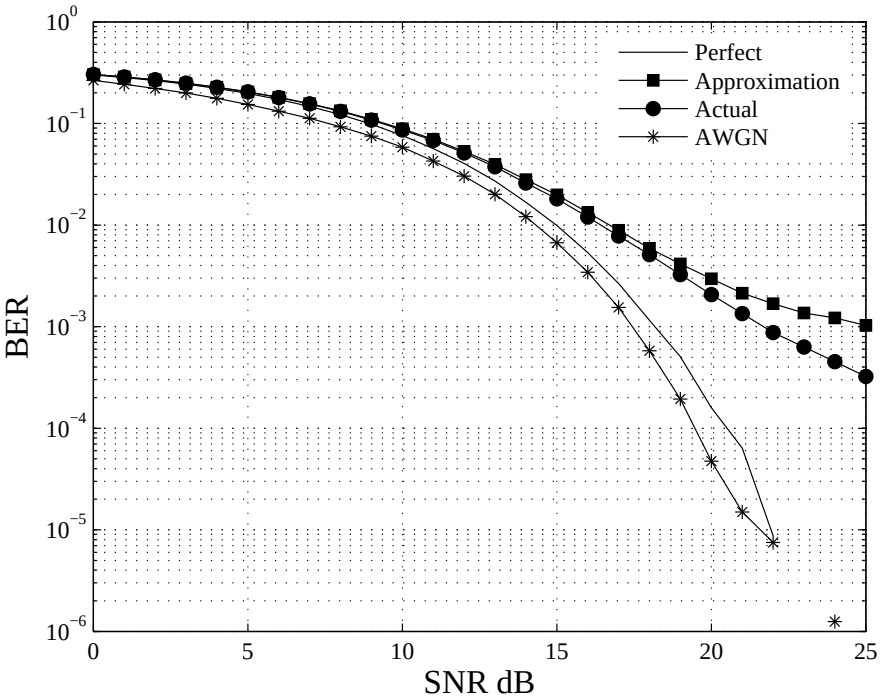


Figure 5.23: BER for a 3×10 channel with $\sigma_p^2 = 0.1$.

By isolating the individual effects of different values of the CSI error variance it allows investigation into the point at which the approximation breaks down due to too much CSI error. Given the 3×10 antenna configuration in Fig. 5.24 the effects of different values of σ_p^2 are shown. At 10^{-2} BER the approximation using $\sigma_p^2 = 0.08$ is less than 1dB away from the perfect curve, the approximation with $\sigma_p^2 = 0.12$ is around 2.5dB away and the approximation with $\sigma_p^2 = 0.16$ is 10dB away. At 10^{-3} only the approximation using σ_p^2 is on the plot and is a little over 2dB away from the curve with perfect CSI.

Next the CSI error was isolated and different antenna configurations were used to show how a wider antenna configuration effected the BER. In Fig. 5.25 the 2×6 case with perfect CSI (where perfect indicates the ideal situation of zero CSI) compared against imperfect CSI with $\sigma_p^2 = 0.1$ with antenna configurations 2×6, 2×10 and 2×14. At 10^{-2} there is 2dB of separation in performance between the approximation using the 2×6 antenna configuration. At 10^{-3} the 2×6 antenna configuration is not in the figure, but there is only 1dB of separation between the approximation using 2×10 and the 2×14 and perfect 2×6 case. At 10^{-4} the approximation with the 2×10 case begins to degrade more and now there is 2dB of separation. The approximation using the imperfect CSI with a 2×14 antenna configuration performs as well as the 2×6 case with perfect CSI throughout the entire range of SNR values. The cases for perfect CSI for antenna configurations of 2×10 and 2×14 are left off of Fig. 5.25 to readily compare approximation and actual curves easily for the other antenna configurations.

Due to the symmetry of the SVD theorem, a converse analysis was performed for tall channel matrices. The performance of this tall approximation was examined under the same simulation conditions as the previous approximation for a wide channel. Various antenna configurations all with $M_r > M_t$ were performed with different values of σ_p^2.

The first simulation result presented for a tall antenna configuration is with a 6×2 channel with a CSI error variance of $\sigma_p^2 = 0.05$, shown in Fig. 5.26. When the BER reaches 10^{-2}, the curve which indicates the case using perfect CSI, the approximation and the actual curves found using the imperfect CSI are all near 12dB. Once the BER reaches 10^{-3} there is significant degradation in the performance since the perfect curve is at 20dB, the actual curve is at 25dB and the approximation is too poor to be on the figure.

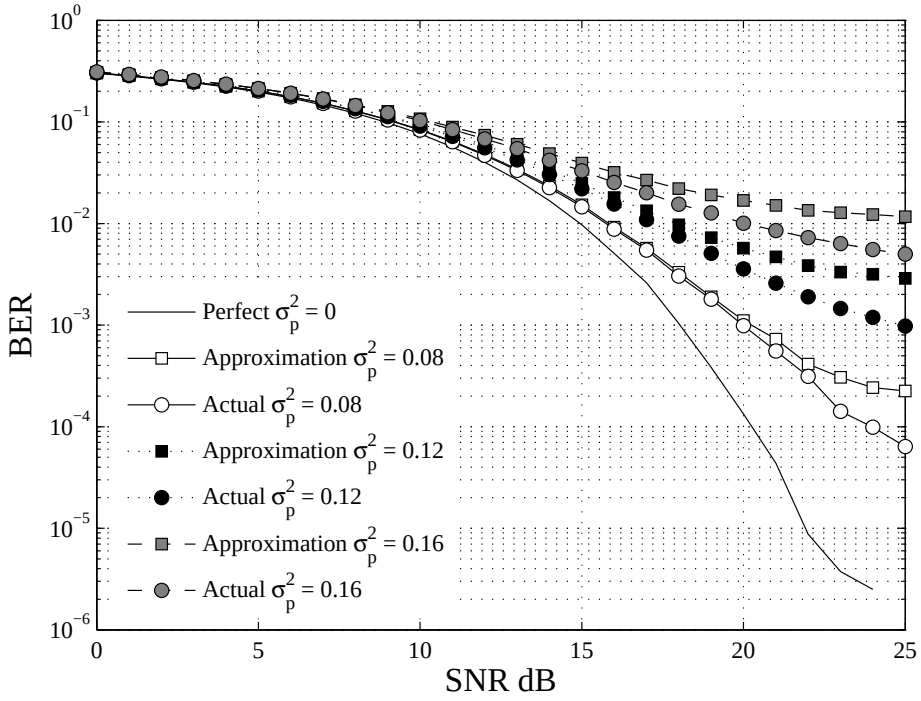


Figure 5.24: BER for a 3×10 channel with multiple σ_p^2.

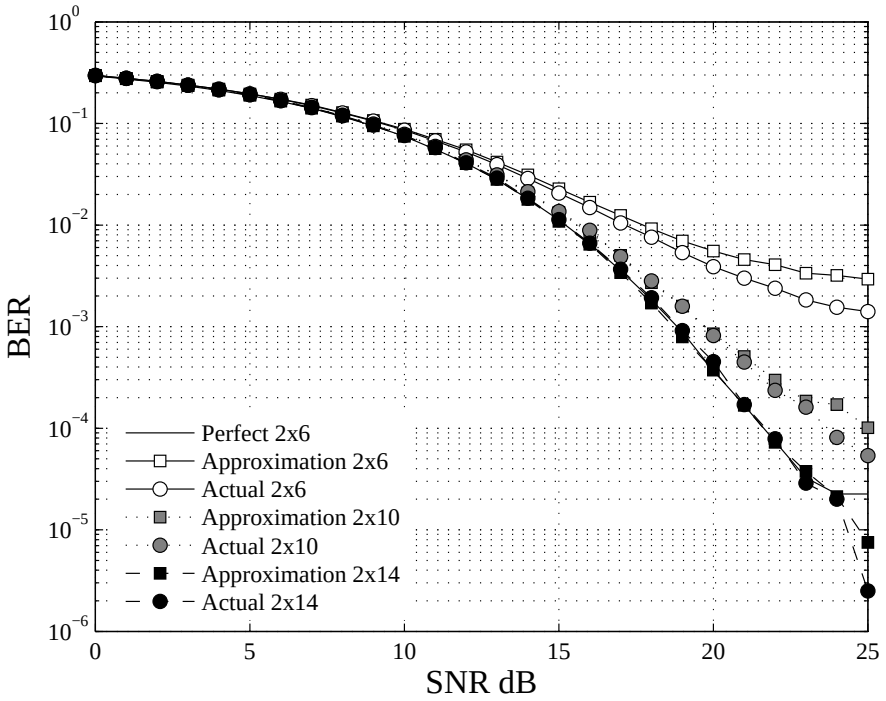


Figure 5.25: BER with a CSI error variance of $\sigma_p^2 = 0.1$ and multiple antenna configurations.

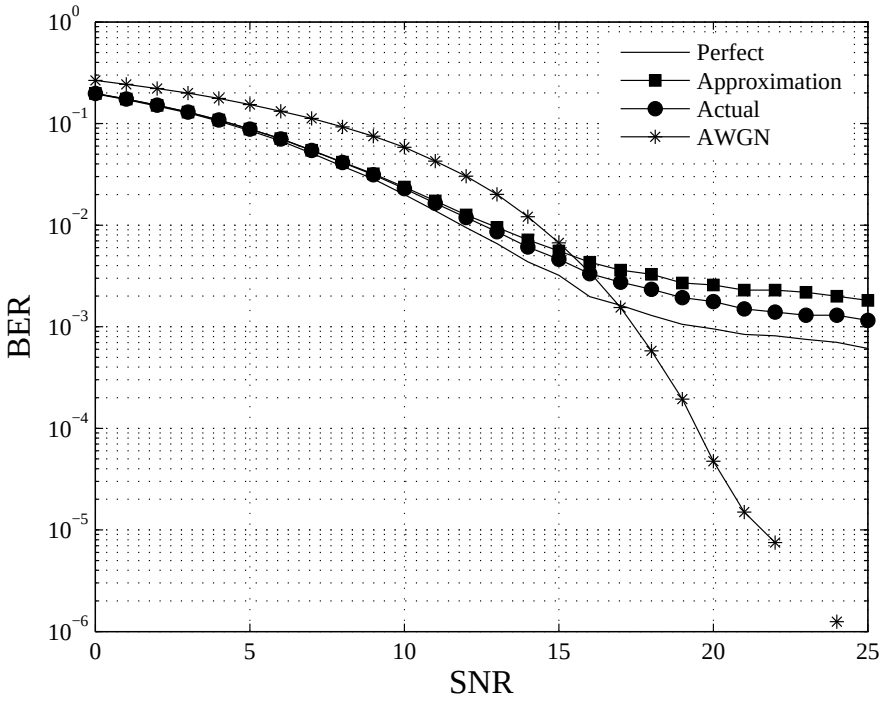


Figure 5.26: BER for a 6×2 channel with $\sigma_p^2 = 0.05$.

By increasing the value of the CSI error variance, σ_p^2, it may be shown how the performance of the approximation for the tall channel degrades. Fig. 5.27 shows a 6×2 antenna configuration with a CSI error variance of $\sigma_p^2 = 0.1$. When the SNR of the system reaches 13dB the BER of the perfect curve is only at 10^{-2} and the approximation does not reach the same level of performance until 17dB. At a BER of 10^{-3} the approximation is not on the figure, which indicates the decline in performance due to the increase of CSI error in addition to the poor quality of the tall approximation compared with that of the wide approximation.

By increasing the number of receive antennas and making the antenna configuration 10×3 and reducing the CSI error variance to $\sigma_p^2 = 0.05$ the performance of the system improves, as shown in Fig. 5.28. When the three curves achieve a BER of 10^{-3} the SNR of the curve found using perfect CSI is at 13dB, the actual curve is at 14dB and the approximation is at 15dB. At this level of BER performance there is only a 2dB difference between the perfect curve and the approximation to that perfect curve. However, at 10^{-4} there is a 6dB separation between the perfect (16dB) and approximate (22dB) curves. Comparing these results to that of Fig. 5.26 this shows

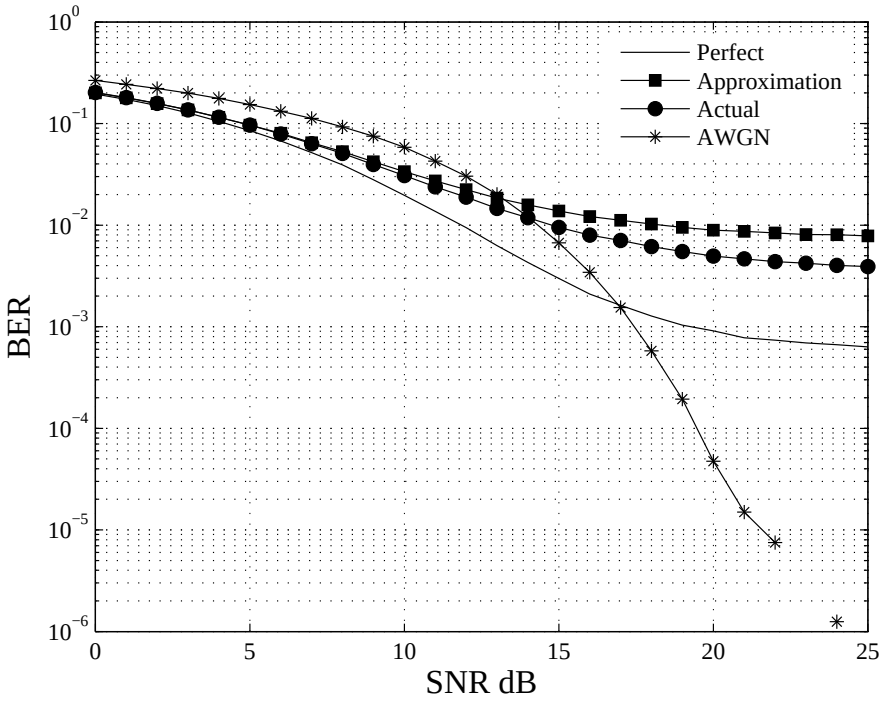


Figure 5.27: BER for a 6×2 channel with $\sigma_p^2 = 0.1$.

that by making the channel taller better BER performance is achieved.

By increasing the CSI error variance of the 10×3 antenna configuration, the gains that were achieved in the previous simulation and shown in Fig. 5.28, are lost. The poor performance which is achieved when the CSI error variance is set to $\sigma_p^2 = 0.1$ is shown in Fig. 5.29. At a BER of 10^{-2} the perfect curve is at 10dB, where as the actual and approximation curves achieve 12dB and 13dB, respectively. The approximation curve performs so poorly that it never achieves a BER of 10^{-3}.

Next, the antenna configuration of 6×2 was used to investigate how well different values of σ_p^2 effected the BER and is shown in Fig. 5.30. The curve computed with perfect CSI achieves a BER of 10^{-2} at 12dB. The approximation using the smallest CSI error variance $\sigma_p^2 = 0.08$ achieves the same level of BER at 14dB. The values of the CSI error variance $\sigma_p^2 = 0.12$ and $\sigma_p^2 = 0.16$ do not achieve the level of 10^{-2} BER. Notice in Fig. 5.30 how as the CSI error variance increases by 0.04 for each pair of curves, each pair is equidistant from each other.

The next group of curves presented uses one value for the CSI error variance and multiple

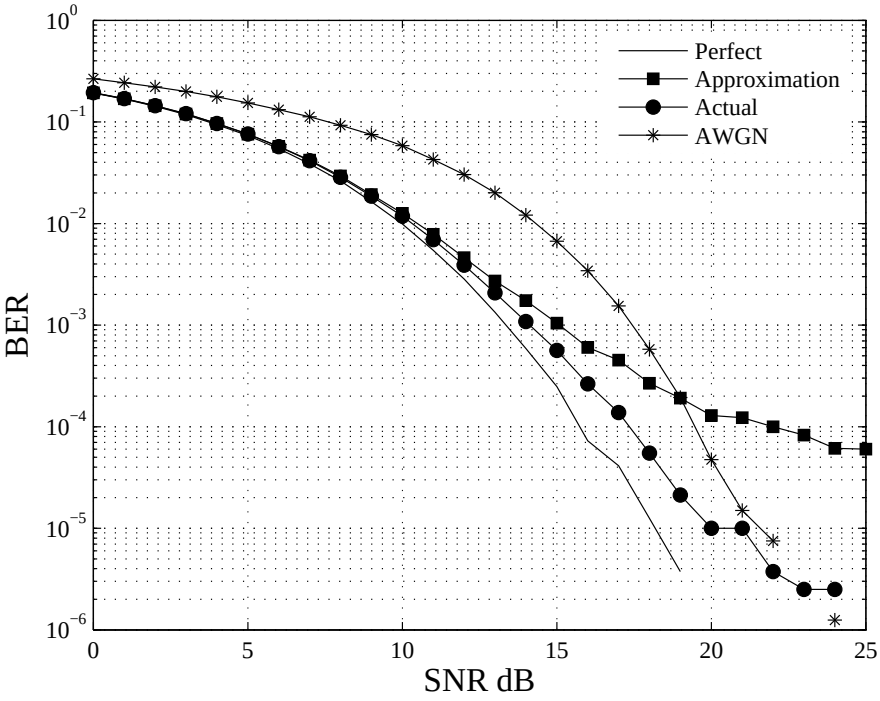


Figure 5.28: BER for a 10×3 channel with $\sigma_p^2 = 0.05$.

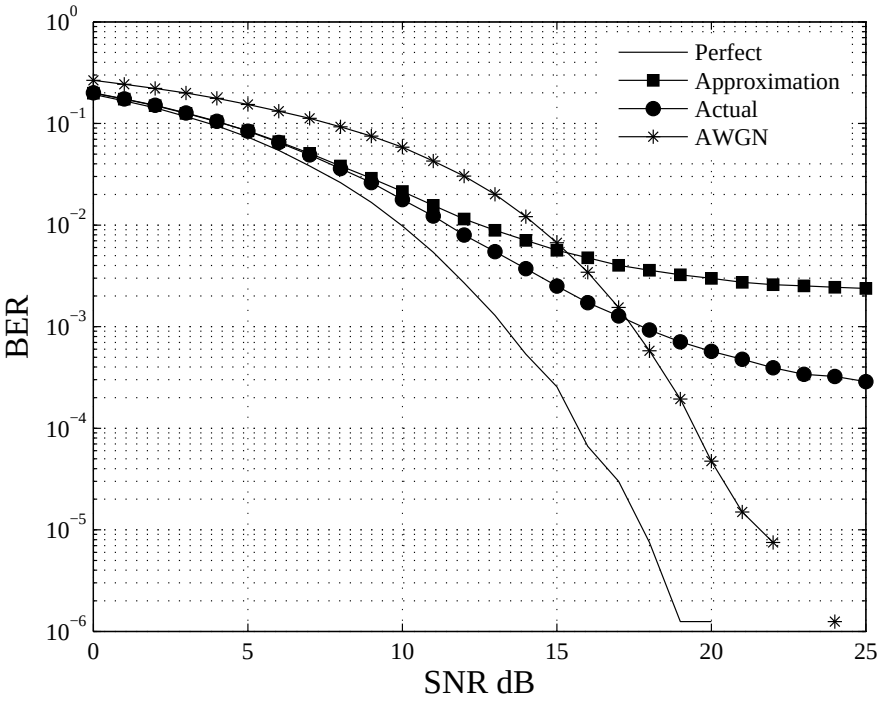


Figure 5.29: BER for a 10×3 channel with $\sigma_p^2 = 0.1$.

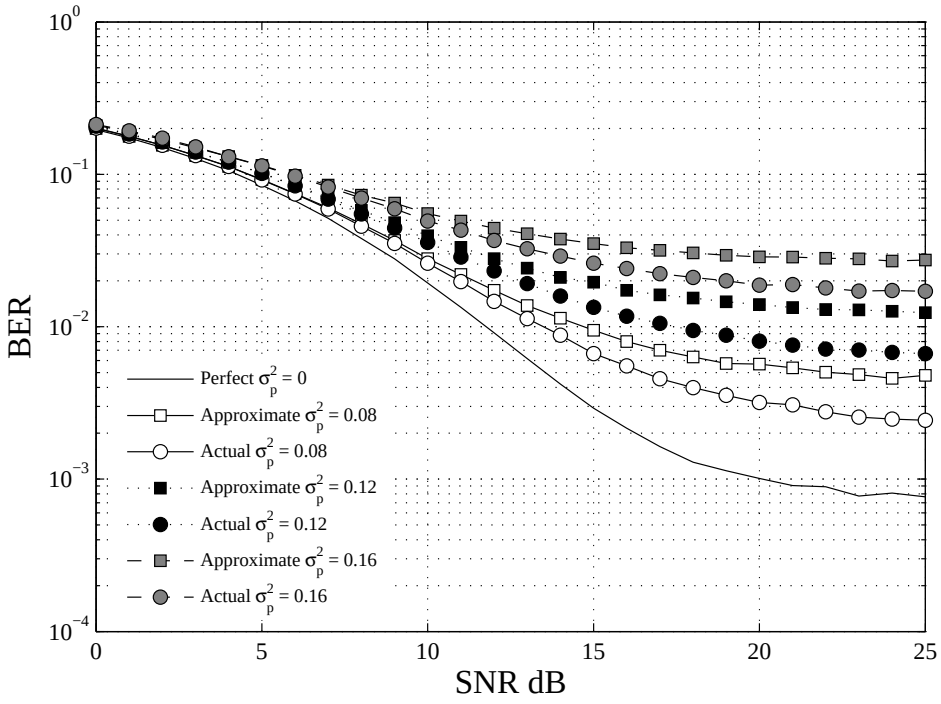


Figure 5.30: BER for a tall channel with many σ_p^2.

antenna configuration pairs. For $\sigma_p^2 = 0.1$ the antenna configurations of 6×2, 10×2 and 14×2 were used to show that as the channel model became taller, the better the approximation becomes, as shown in Fig. 5.31. The perfect curve computed with perfect CSI achieves a BER of 10^{-3} at 21dB. The 6×2 approximation to the perfect curve barely achieves 10^{-2} at the same SNR. The 10×2 approximation achieves a BER of 10^{-3} at 15dB, and the 14×2 approximation achieves a BER of 10^{-3} at 11dB. This shows that as the tall approximation becomes taller, the BER improves, for a given σ_p^2. The cases for perfect CSI for antenna configurations of 10×2 and 14×2 are left off of Fig. 5.31 to readily compare approximation and actual curves easily for the other antenna configurations.

The final scenario is for a square antenna configuration. Fig. 5.32 shows two curves, the perfect and the actual for a $\sigma_p^2 = 0.05$. There is no approximation curve since there was no SVD theorem which allowed an approximation to be made, which leaves only two curves. At 10dB the CSI error does not have an adverse effect on the BER curve, since the two are at the same level 10^{-1}. However, the actual curve does not reach a BER of 10^{-2} within the range of 10dB to 25dB,

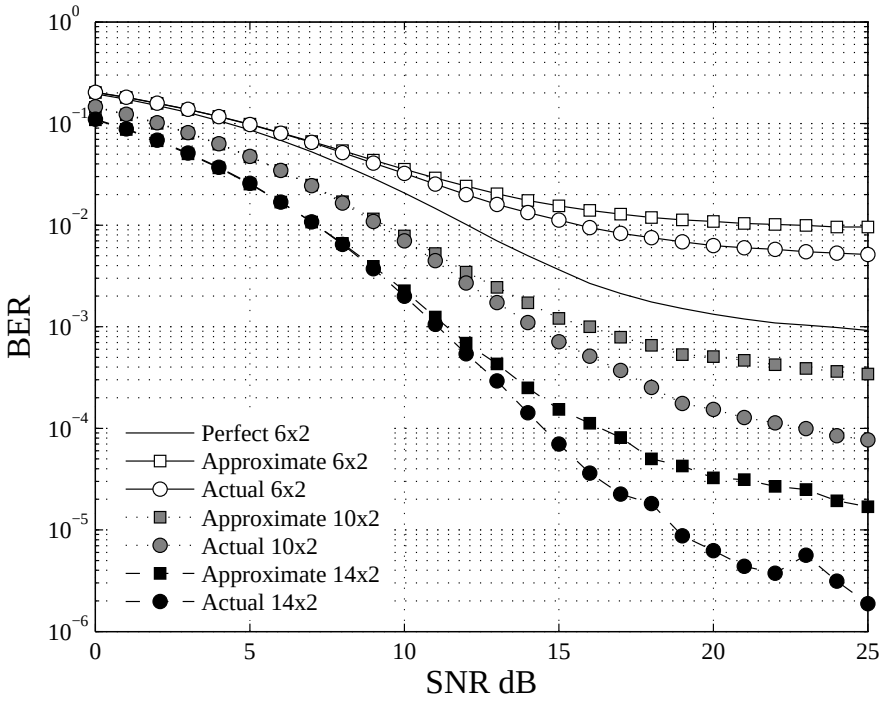


Figure 5.31: BER with a CSI error variance of $\sigma_p^2 = 0.1$ and multiple antenna configurations.

whereas the perfect curve reaches 10^{-2} at 19dB.

To investigate further the effects of imperfect CSI for a square antenna configuration, Fig. 5.33 reveals the BER curves for another 4×4 channel matrix using a CSI error variance of $\sigma_p^2 = 0.1$. The performance of the system for a 4×4 antenna configuration is not able to withstand errors in the CSI, given that there is a 5dB difference in perfect and the actual curves at 10^{-1} BER, which is poor.

To conclude our discussion on BER, we now present the approximation to the equivalent channel model and the Gaussian approximation together on the same figure for both wide and tall antenna configurations. Beginning with Fig. 5.34 it is shown that the performance of the approximations remains relatively equal up to 10dB, but at 15dB the superior performance of the Gaussian approximation is clear. The performance of the equivalent channel model approximation for $\sigma_p^2 = 0.12$ is relatively close to that of the Gaussian approximation for a $\sigma_p^2 = 0.08$.

The tall antenna configuration of 6×2 with both approximations is shown in Fig. 5.35. In this figure there exists separation between the types of approximations at low SNR values. Similar

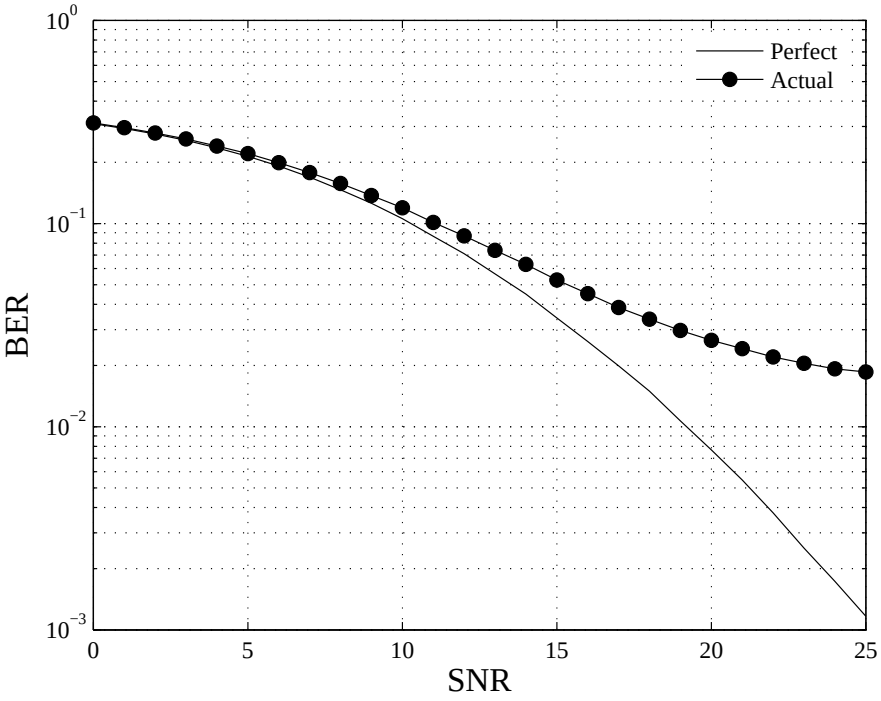


Figure 5.32: BER for a 4×4 channel with $\sigma_p^2 = 0.05$.

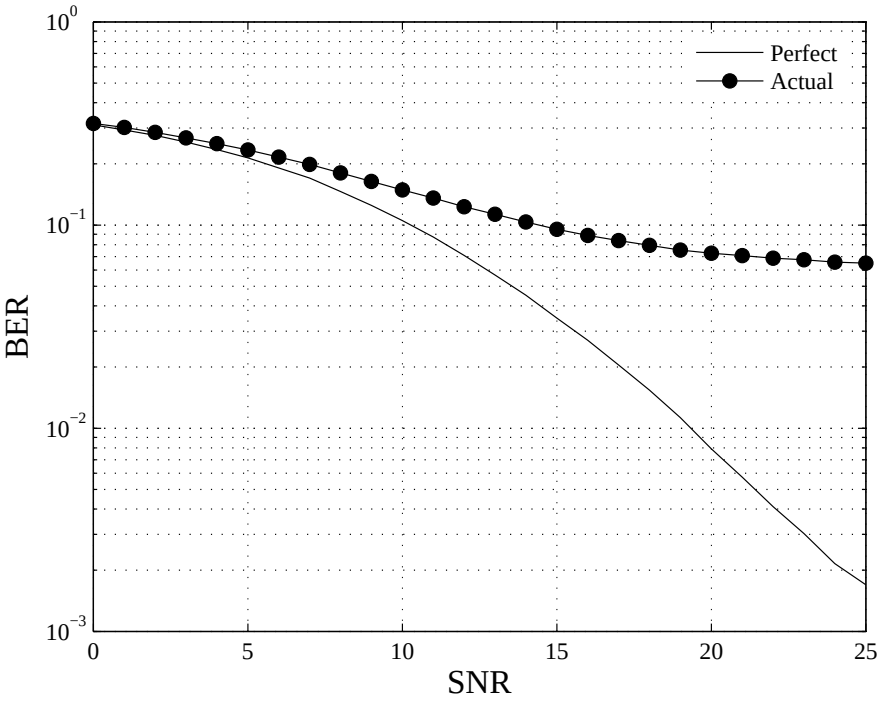


Figure 5.33: BER for a 4×4 channel with $\sigma_p^2 = 0.1$.

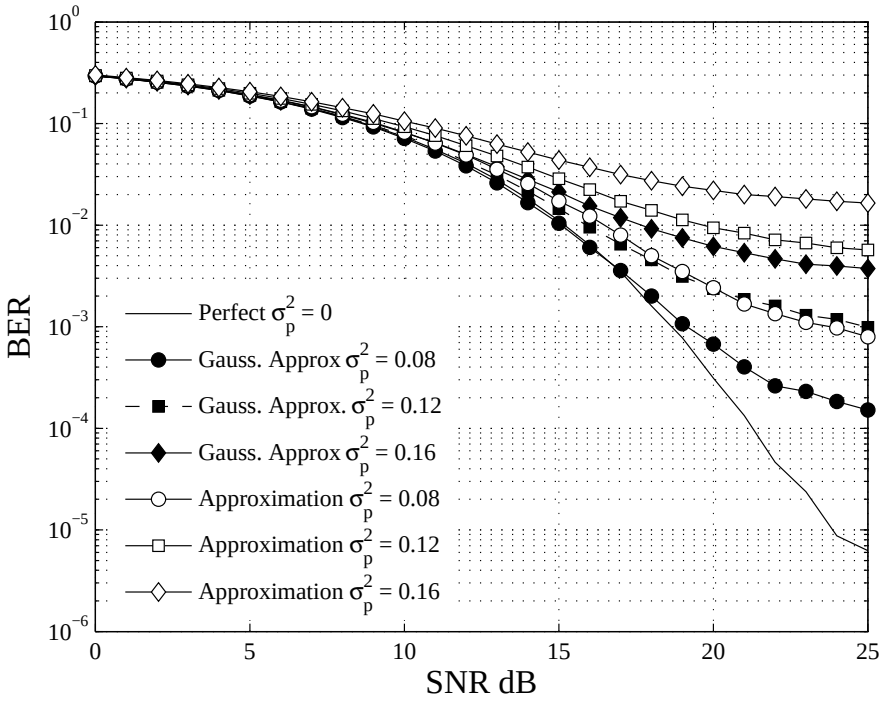


Figure 5.34: BER for a 2×6 channel with both approximations.

to the previous figure, the overall performance of the Gaussian approximation is better and the performance of the Gaussian approximation for $\sigma_p^2 = 0.12$ is equivalent to the approximation to the equivalent channel with a $\sigma_p^2 = 0.08$.

5.3 Capacity

In addition to BER, capacity is also an important system performance measure. Simulations were setup for different antenna configurations to analyze the performance of a precoder/decoder setup under perfect and imperfect conditions. The CSI error variance was set at 0, 0.4 and 0.8 in each case. The simulations begin by first creating a Rayleigh fading channel with the desired fading characteristics, computing the optimal decoder and precoder, then finally using the equation for capacity as given in the perfect (3.20) and imperfect (4.54) cases. The value of capacity computed was then averaged over 10,000 iterations for each SNR/SINR value.

The first capacity plot to be shown is for a wide channel model with an antenna configuration

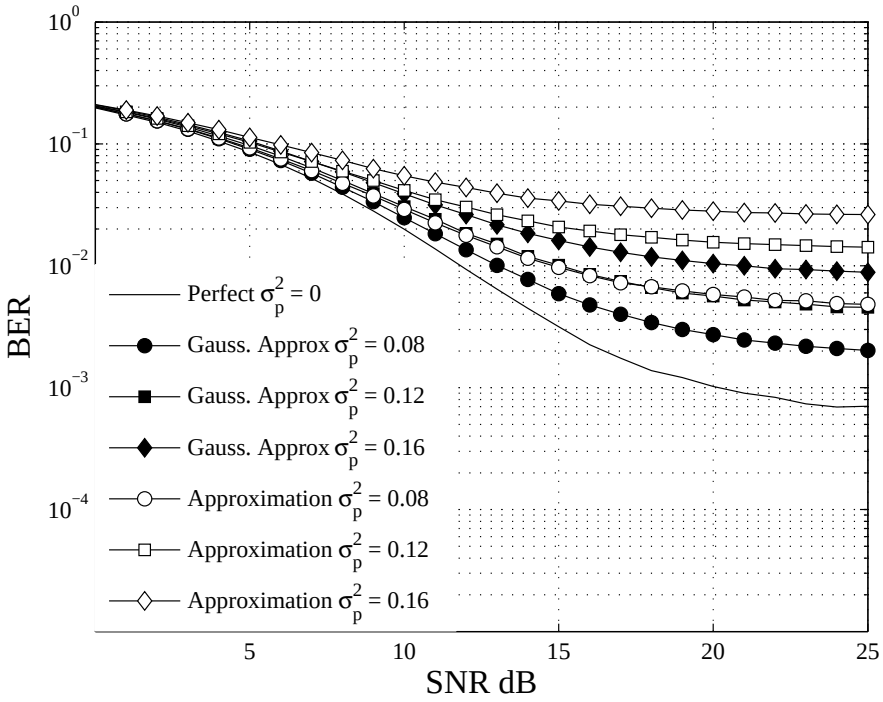


Figure 5.35: BER for a 6×2 channel with both approximations.

of 2×6. The capacity of the system with perfect CSI is given and is shown in Fig. 5.36. In this case the capacity of the system computed using perfect CSI achieves 23bps/Hz at 10db. The capacity of the system computed using CSI with an error variance of $\sigma_p^2 = 0.4$ loses 0.5bps/Hz, and the system computed using CSI with an error variance of $\sigma_p^2 = 0.8$ loses nearly 1bps/Hz.

By increasing the number of receive antennas the overall capacity of the system improves, as shown in Fig. 5.37 which uses a 3×10 antenna configuration. When the capacity of the system is computed using perfect CSI a capacity of 32bps/Hz is achieved at 10dB. Given the same scenario but using imperfect CSI results in a 0.5bps/Hz and 1bps/Hz loss in capacity for the same SNR level for $\sigma_p = 0.4$ and $\sigma_p = 0.8$ respectively.

The final illustration of the capacity of a wide antenna configuration is shown in Fig. 5.38 with a 4×10 antenna configuration. Analyzing the capacity curves at 10dB shows a capacity of 37bps/Hz while using perfect CSI. The subsequent performance degradation losses incurred while using imperfect CSI are 1bps/Hz and 1.5bps/Hz with a CSI error variance of $\sigma_p^2 = 0.4$ and $\sigma_p^2 = 0.8$, respectively.

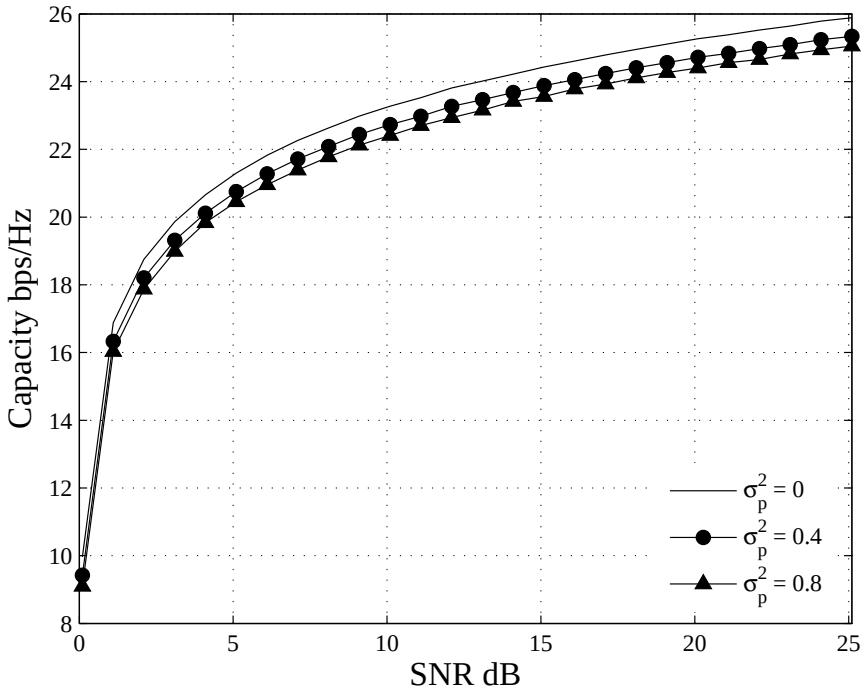


Figure 5.36: Capacity of a 2×6 channel.

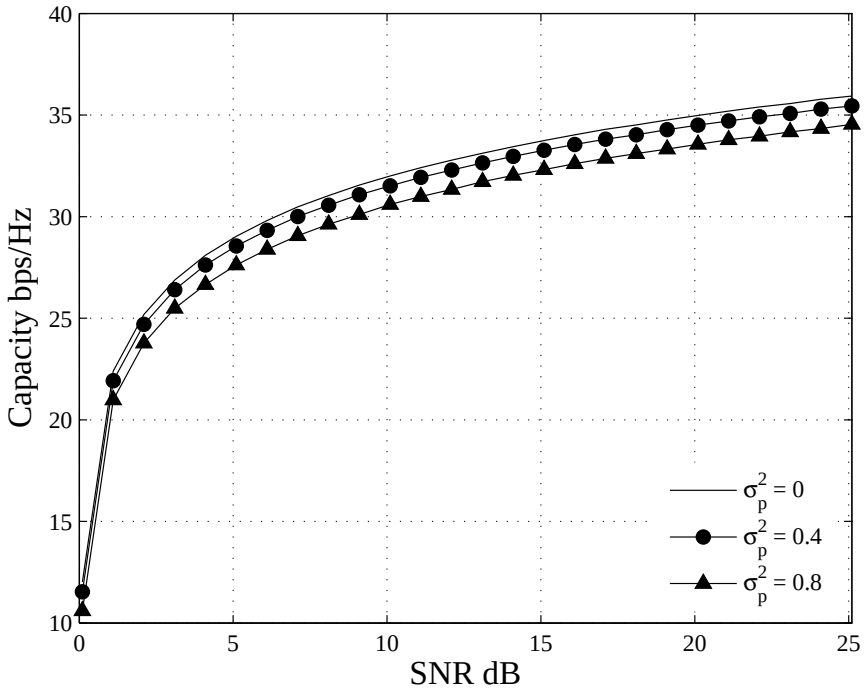


Figure 5.37: Capacity of a 3×10 channel.

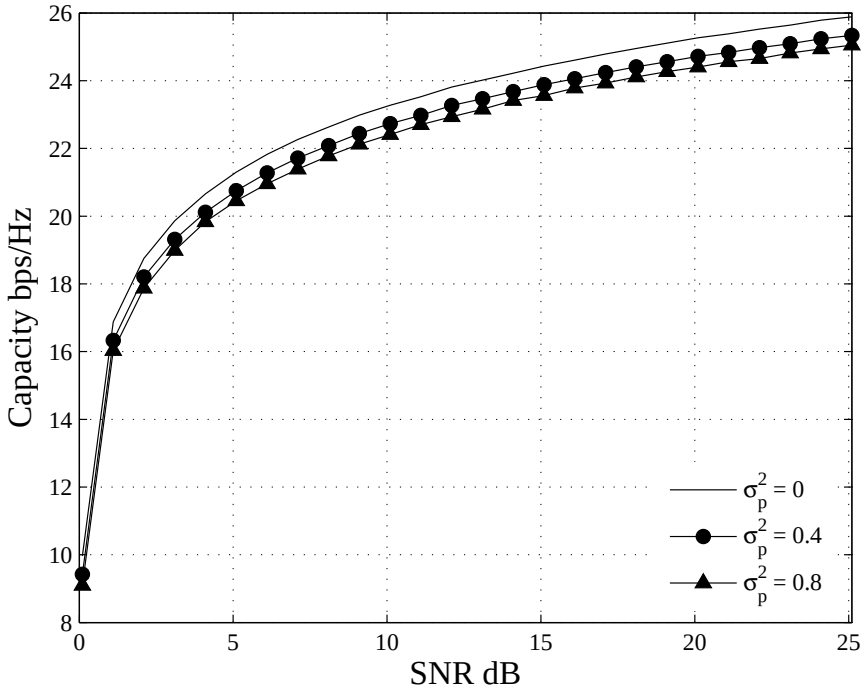


Figure 5.38: Capacity of a 4×10 channel.

By reversing the antenna configurations similar capacity figures were achieved using the same values for σ_p^2. The first capacity plot to be shown is for a tall channel model uses an antenna configuration of 6×2. The capacity of the system with perfect CSI is given and is shown in Fig. 5.39. In this case the capacity of the system computed using perfect CSI achieves 23bps/Hz at 10db. The capacity of the system computed using a CSI error variance of $\sigma_p^2 = 0.4$ loses 0.5bps/Hz, and the system computed using a CSI error variance of $\sigma_p^2 = 0.8$ loses an additional 0.5bps/Hz.

By increasing the number of transmit antennas the overall capacity of the system improves, as shown in Fig. 5.40 which uses a 10×3 antenna configuration. When the capacity of the system is computed using perfect CSI a capacity of 31bps/Hz is achieved at 10dB. Given the same scenario but using imperfect CSI results in a 0.5bps/Hz and 0.5bps/Hz loss in capacity for the same SNR level for $\sigma_p = 0.4$ and $\sigma_p = 0.8$, respectively.

The final illustration of the capacity of a tall antenna configuration is shown in Fig. 5.41 with a 10×4 antenna configuration. Analyzing the capacity curves at 10dB shows a capacity of 37bps/Hz while using perfect CSI. The subsequent performance degradation losses incurred

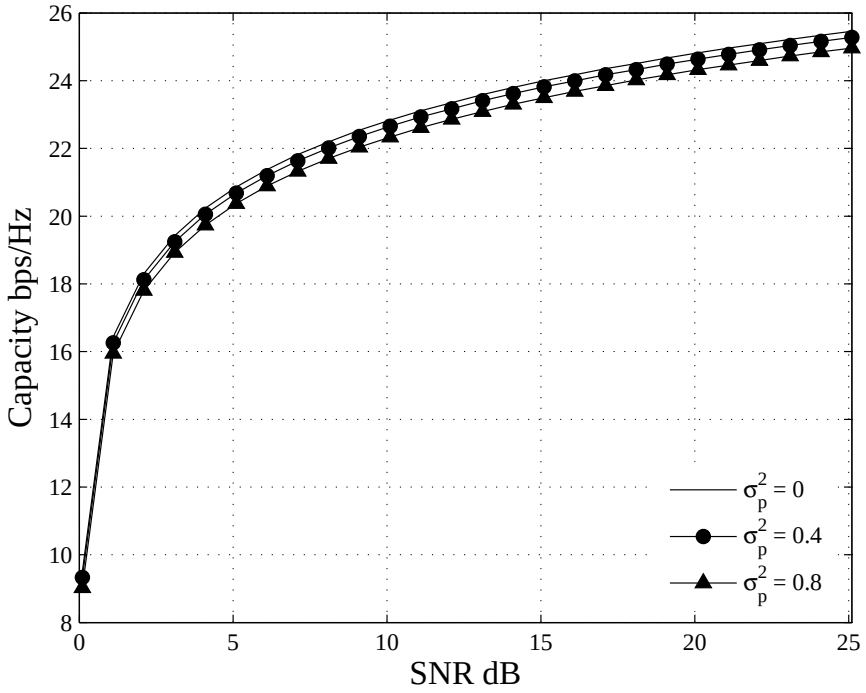


Figure 5.39: Capacity of a 6×2 channel.

while using imperfect CSI are 0.5bps/Hz and 0.5bps/Hz with a CSI error variance of $\sigma_p^2 = 0.4$ and $\sigma_p^2 = 0.8$, respectively.

From the previous six figures it can be concluded that increasing by increasing the number of transmit or receive antennas increases the system capacity. As shown in the figures, as the CSI degrades from perfect to imperfect, it is shown that there is a loss in capacity, but not more than 1.5bps/Hz in any of the cases.

5.4 Gaussian Approximated BER and Capacity Results

The equations of Section 4.8 were then simulated to relate their performance to the performance of the plots in Section 5.2. The first group of results are for the wide cases using different antenna configurations for multiple values of σ_p^2. The first set of plots shown in Fig. 5.42 are for a 2×6 antenna configuration. For the three different σ_p^2 cases, the Gaussian approximations outperform the results of Section 5.2. This is because the off-diagonal elements are approximated as Gaussian,

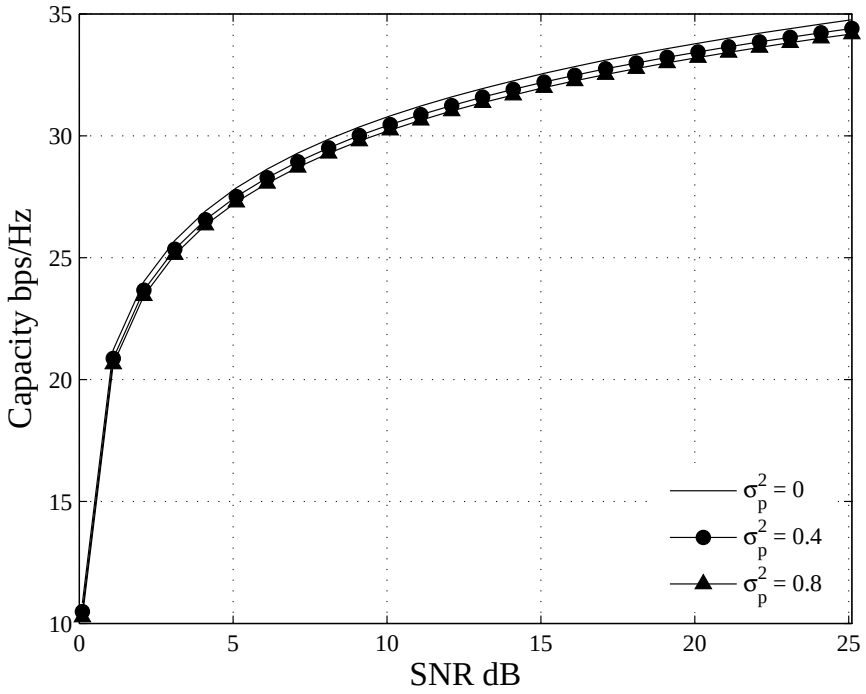


Figure 5.40: Capacity of a 10×3 channel.

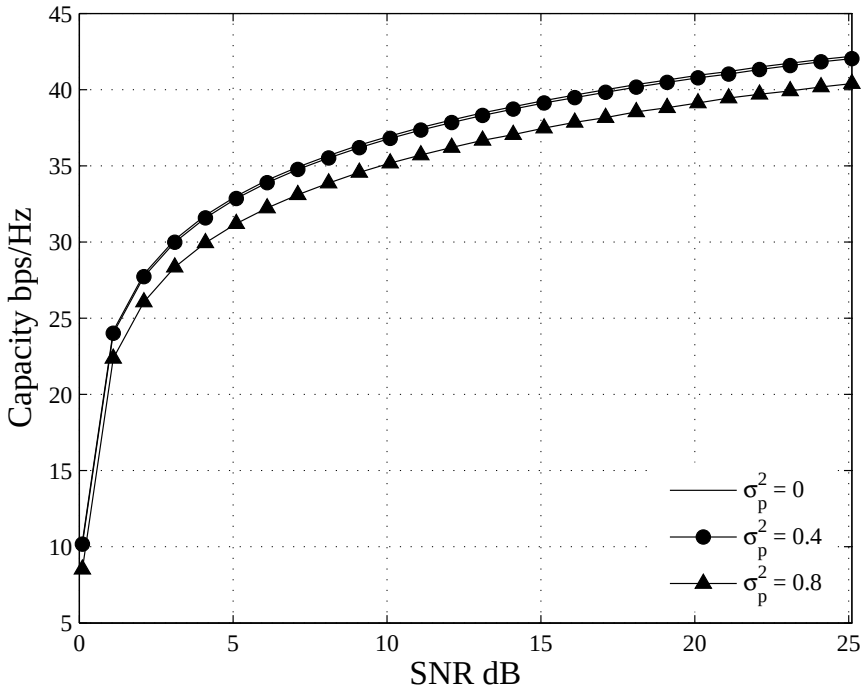


Figure 5.41: Capacity of a 10×4 channel.

which results in the elimination of terms which influence the performance of the system. In the approximation of the BER, the performance of each of the Gaussian approximations outperforms the actual system.

Moving on to a different antenna configuration, such as the 4×10, as shown in Fig. 5.43 shows improved performance over the previous 2×6 case. Each of the Gaussian approximations outperforms its' counterpart using the same σ_p^2 value. Additionally the 4×10 Gaussian approximation using $\sigma_p^2 = 0.08$ achieves a BER of 10^{-4} at 20 dB, whereas the 2×6 Gaussian approximation never achieves a BER of 10^{-4}.

However, by further increasing the size of the antenna configuration from 4×10 to 6×12 does not significantly improve the performance of the BER as did the 2×6 to 4×10 increase. Taking the Gaussian approximation for the case of $\sigma_p^2 = 0.08$ for the 6×12 case, shown in Fig. 5.44 shows a BER of 10^{-2} at a SNR of 14 dB, a BER of 10^{-3} at 17 dB and a BER of 10^{-4} at 21 dB. In Fig. 5.43, the same Gaussian approximation value of $\sigma_p^2 = 0.08$ performs the same. It was previously shown that as the antenna configuration changed from wider to square with this linear MMSE design, that the performance decreased. However, by increasing the number of propagation paths, BER improves. The comparison of the plots of Figs. 5.43 and 5.44 shows that the benefits of more propagation paths is offset by a more square antenna configuration. Where more square indicates a ratio of receive to transmit antennas is closer to 1.

The second group of results compare the approximations for the tall cases using different antenna configurations for multiple values of σ_p^2. Beginning with Fig. 5.45, one may see that the performance of a 6×2 antenna configuration is quite similar to that of Fig. 5.26 which is with a $\sigma_p^2 = 0.05$. Again, similar to Fig. 5.42, the Gaussian approximations outperform the actual plots for each of the various σ_p^2 values.

Increasing the number of antennas and moving to a 10×4 antenna configuration is shown in Fig. 5.46. The increase in the number of propagation paths from transmitter to receiver is the dominating factor in the increased performance from the 6×2 to 10×4 cases. Again, as in Fig. 5.45, the Gaussian approximation plots in Fig. 5.46 outperform their σ_p^2 counterparts.

The last BER approximation plot is for a 12×6 antenna configuration and is shown in Fig. 5.47. There is not a significant performance improvement found by increasing the antenna config-

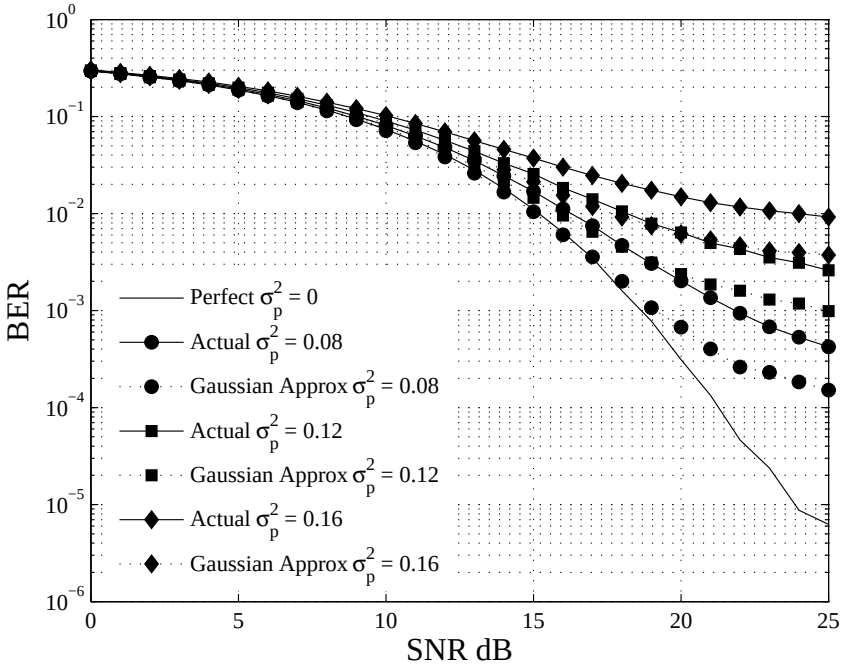


Figure 5.42: BER of the Gaussian approximations and their counterparts of Section 5.2 for a 2×6 antenna configuration.

uration from 10×4 to 12×4. Looking closely at the performance with $\sigma_p^2 = 0.12$ of Fig. 5.46, the BER value at 10^{-2} is 14 dB and at 10^{-3} is 19 dB. The corresponding plot for the 12×6 antenna configuration shown in Fig. 5.47 has a BER of 10^{-2} at 13 dB and 10^{-3} at 20 dB. This shows that for benefit of increasing the number of propagation paths is negated by the poor performance of the approximation to the linear MMSE design. However, similar to the previous Gaussian approximation plots, the 12×6 antenna configuration performs in accord with the previous figures in that the Gaussian approximations outperform their actual counterparts.

It may be concluded that the Gaussian approximations are not a good indicator of the true performance of the linear MMSE designs corrupted with CSI error. It is not a good indicator because in all of the cases as the SNR of the system increased, the BER of the Gaussian approximations decreased at a higher rate than that of their corresponding curves for the particular σ_p^2 values.

Now we will delve into the Gaussian approximations to capacity and show a parallel analysis. Based on the above development of the Gaussian approximation on BER the capacity simulations

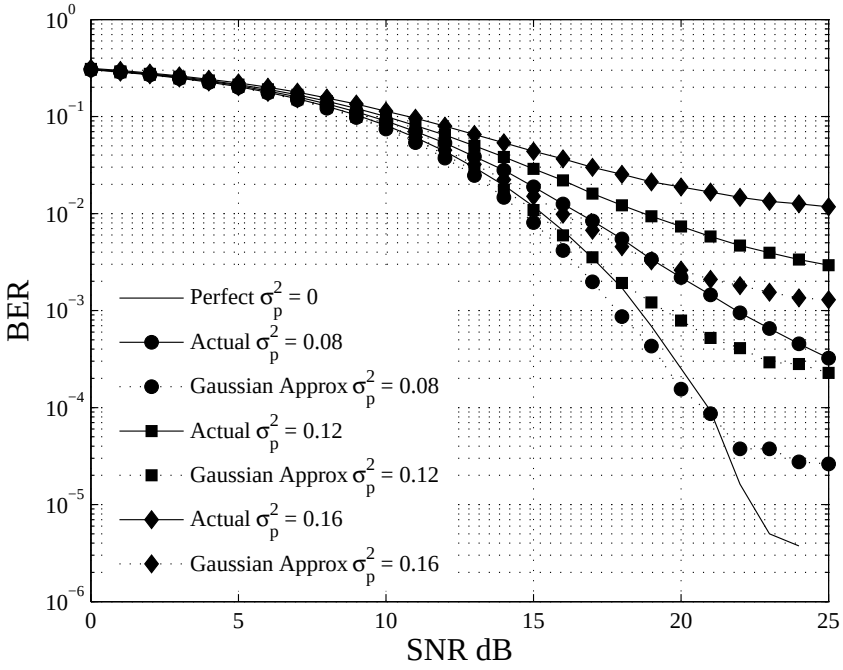


Figure 5.43: BER of the Gaussian approximations and their counterparts of Section 5.2 for a 4×10 antenna configuration.

of the previous section were repeated in order to investigate how well the Gaussian approximation applies to capacity computations. The simulations were carried out identically as in the previous section, with the exception of the addition of the two additional curves indicating the Gaussian approximation to the CSI error for each antenna configuration.

The first plot shown in Fig. 5.48 shows a 2bps/Hz increase using the Gaussian approximations for both CSI error variances between 5 and 25dB with a 2×6 antenna configuration. Increasing to a 4×10 antenna configuration, as shown in Fig. 5.49 shows an increase of nearly 17bps/Hz between 5 and 25dB. This an unreasonably large increase in capacity and shows the limitations of the Gaussian approximation, due to the increase in the number of antennas.

Next, by changing to a tall antenna configuration as shown in Fig. 5.50, similar results are obtained. Using a 6×2 antenna configuration there is a 1bps/Hz increase using the Gaussian approximation than that of the case with perfect CSI from 5 to 25dB. Increasing the size of the antenna configuration to a 10×4 shows an alarming increasing in the capacity of the Gaussian

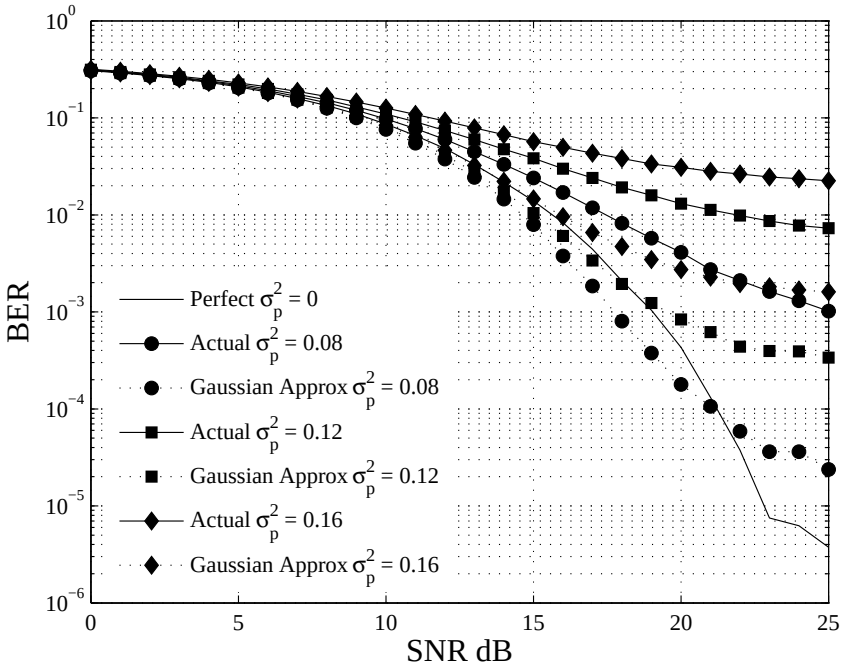


Figure 5.44: BER of the Gaussian approximations and their counterparts of Section 5.2 for a 6×12 antenna configuration.

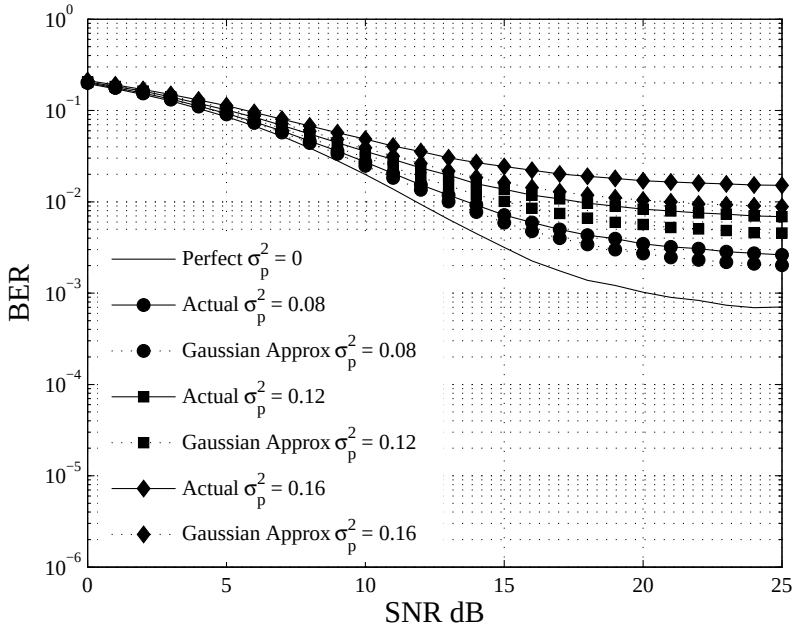


Figure 5.45: BER of the Gaussian approximations and their counterparts of Section 5.2 for a 6×2 antenna configuration.

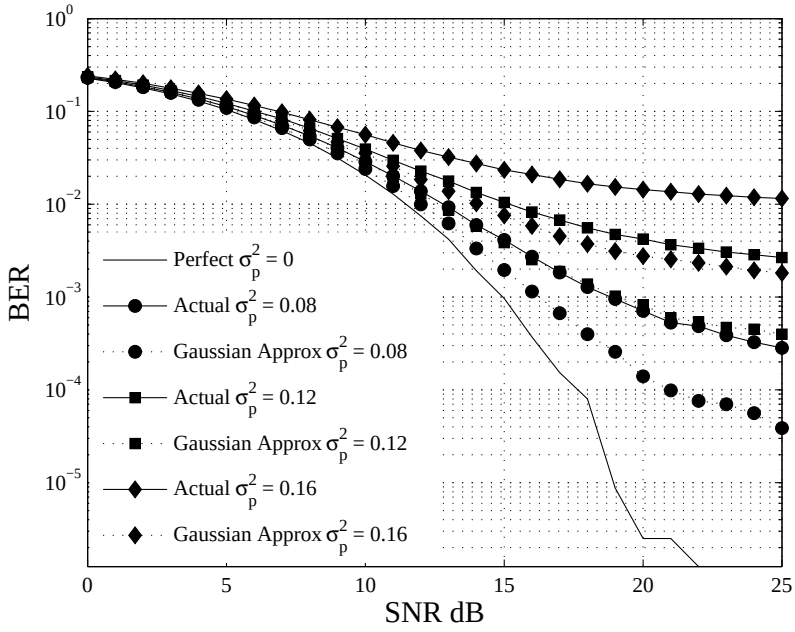


Figure 5.46: BER of the Gaussian approximations and their counterparts of Section 5.2 for a 10×4 antenna configuration.

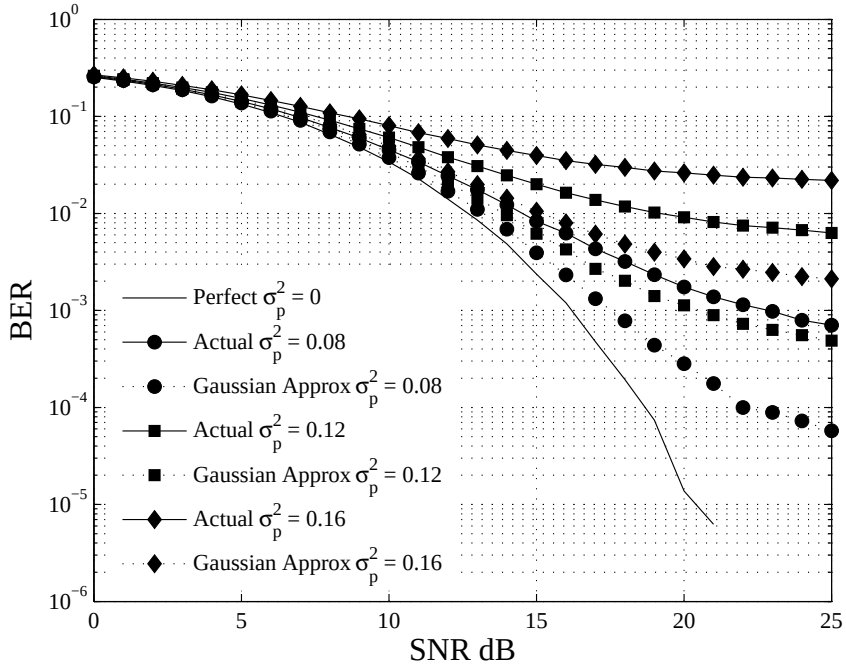


Figure 5.47: BER of the Gaussian approximations and their counterparts of Section 5.2 for a 12×6 antenna configuration.

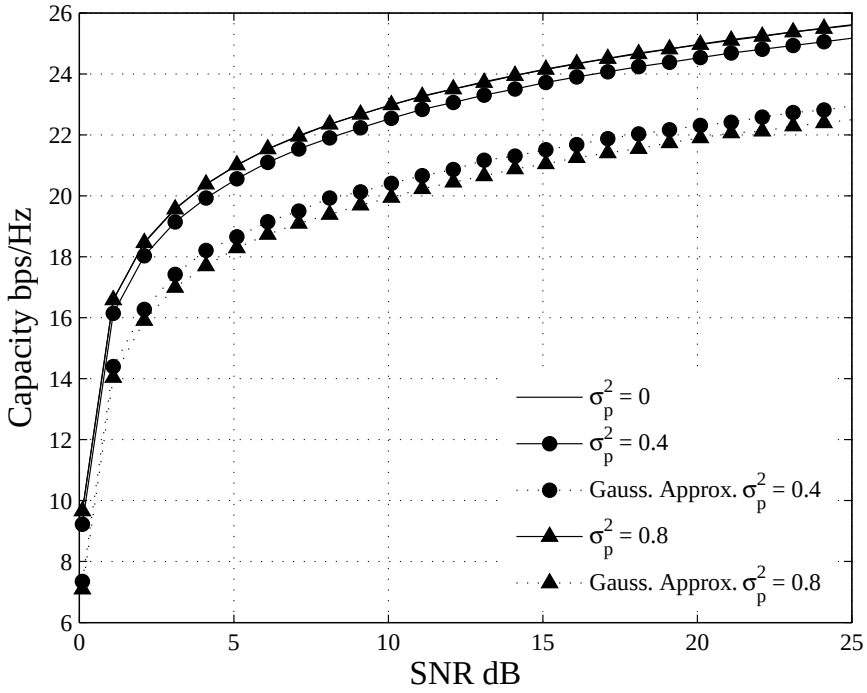


Figure 5.48: Capacity of a 2×6 channel.

approximations. In Fig. 5.51 the Gaussian approximations are 11bps/Hz greater than the case with perfect CSI. Again, similar to Fig. 5.49, Fig. 5.51 shows the limitations of the Gaussian approximation. This is why the Gaussian approximation for the 6×12 and 12×6 antenna configurations were left off, they averaged an increase of 25bps/Hz increase in capacity. This only reinforces the idea that the Gaussian approximation has limitations.

5.5 Conclusion

In this chapter, a performance analysis of linear precoder/decoders was performed. It was shown that the errors disturbing the CSI may be approximated as Gaussian, since their probability distribution is Gaussian-like. The SVD approximation of Section 4.3 is shown to be a better approximation for wide channels than for tall channels. The BER plots indicate the robustness at low SNR values of the SVD approximations. Additionally, the BER plots indicate that as the channel matrix becomes wider or taller and as σ_p^2 decreases, the better the performance of the system. The capacity

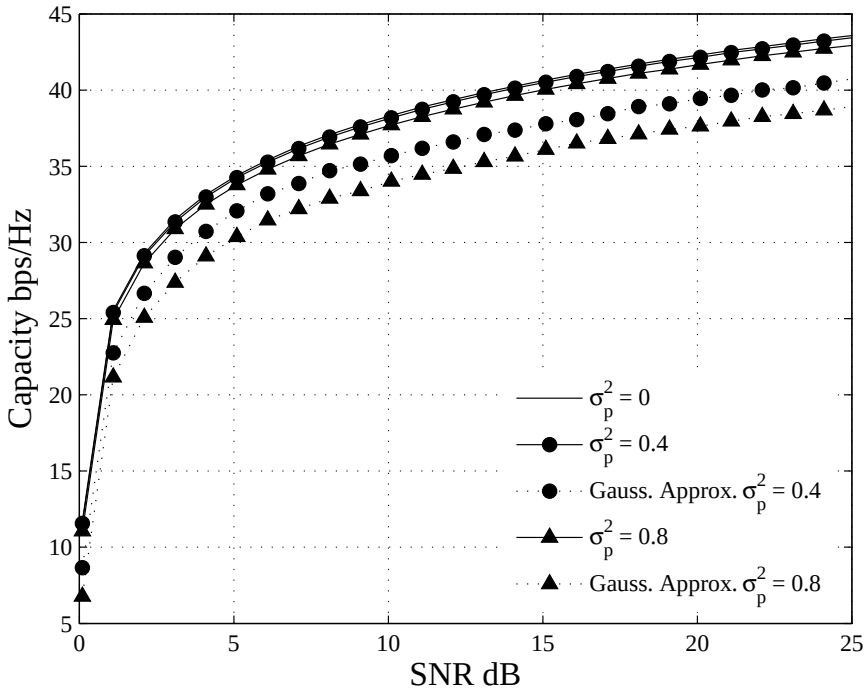


Figure 5.49: Capacity of a 4×10 channel.

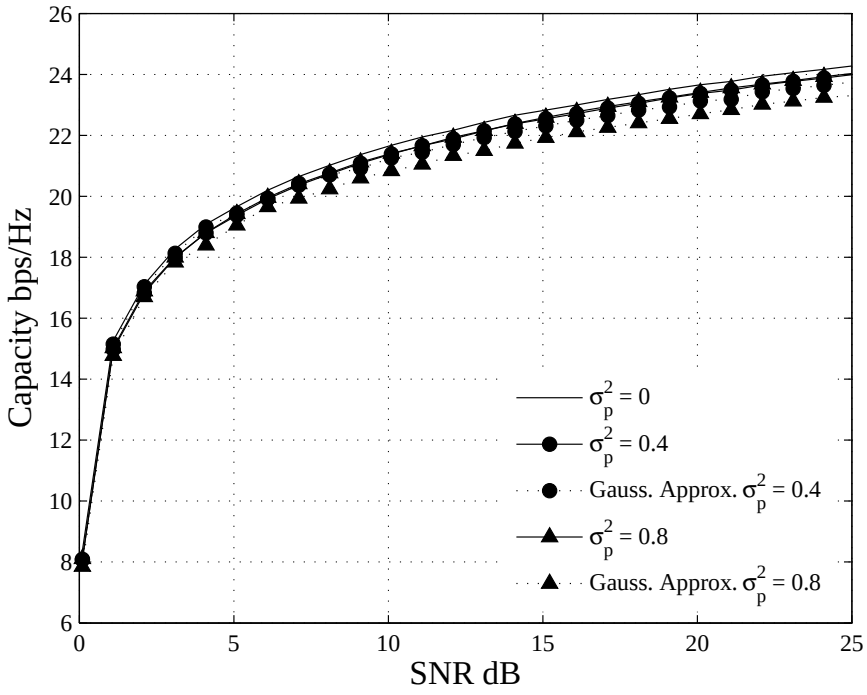


Figure 5.50: Capacity of a 6×2 channel.

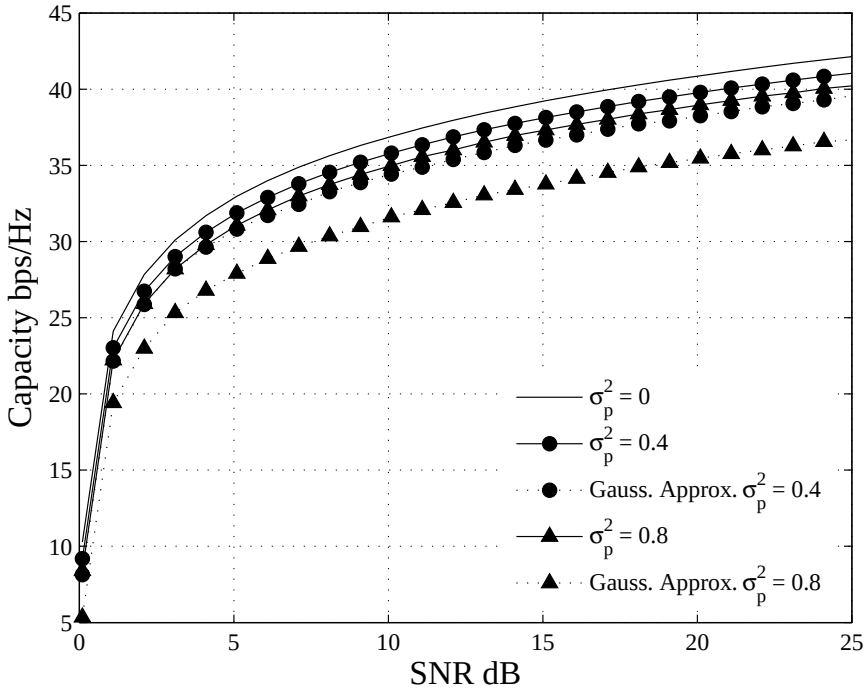


Figure 5.51: Capacity of a 10×4 channel.

figures show how the imperfect CSI matrices decrease the capacity when compared against their undisturbed counterparts because the precoder/decoder matrices are not properly matched to the channel, thereby decreasing capacity. The Gaussian approximations to BER and capacity showed that the Gaussian approximations do hold for small CSI error and relatively smaller antenna configurations.

Chapter 6

Conclusion

This thesis has presented a performance analysis of the a MIMO precoder/decoder design when the CSI is imperfect. This is significant because it enables the designer to answer theoretical questions that cannot be answered through ideal laboratory simulations. This analysis is intended to give a more realistic impact to the sensitivity of linear MMSE precoders/decoders to CSI errors.

The main benefit of MIMO antenna configurations is the increased performance gains. It was shown in chapter 2 that if the CSI is not known at the transmitter, the capacity of the system may be increased linearly by simply by adding antennas at either end of the link. If CSI is available, the simplest precoder/decoder pair ($\mathbf{V}$, $\mathbf{U}^H$) may be used to precode the data, which will in effect cause the capacity to become the sum of r parallel SISO channels.

This lead into the topic of MIMO precoders/decoders and the benefits of full CSI, when available. The particular design introduced in chapter 3 is based on designing the precoder/decoder pair from the MMSE of the received signal. The optimal linear precoder/decoder design using full CSI was then analyzed in the event of imperfect full CSI. A new precoder/decoder design was arrived at, as well as performance equations based upon the imperfect CSI, in addition to considerations for non-square channels. Most importantly, in chapter 4 we were able to analyze exactly how the imperfect CSI effects the off-diagonal elements in the wide and tall antenna configuration scenarios.

Chapter 5 showed detailed analysis of the Gaussian-like probability distributions of the CSI

error in the channel models. Wide and tall channel model approximations were made from the SVD theorem presented in chapter 4. The performance of the approximations was relatively good for wider/taller channel models and low CSI error, but performed poorly with square channel models and high CSI error. The figures of chapter 5 which present capacity curves show that poor CSI does effect capacity, but it is not as large of a factor in the quality of the capacity curves as it is with BER. This shows that the quantity of information sent across a wireless channel is degraded only slightly, whereas the quality of information is significantly degraded.

This body of work has revealed the impact that the imperfect CSI has on the performance of MMSE linear precoders/decoders. The findings indicate that the designs are sensitive to CSI errors, so the practicality is probably limited, unless highly accurate channel estimation is used. Future analysis into the performance of MIMO precoders/decoders may be found in investigations of

- The performance of a multiuser MIMO precoder/decoder design.
- Performance analysis of MIMO precoders/decoders using imperfect partial CSI.
- Performance analysis of MIMO precoders/decoders using imperfect limited CSI.
- The elimination of assumptions such as uncorrelated antennas and the inclusion of their effect in addition to imperfect full/partial/limited CSI designs.

Bibliography

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

[2] S. Glisic, *Advanced Wireless Communications 4G Technologies*. John Wiley and Sons, 2004.

[3] R. Janaswamy, *Radiowave Propagation and Smart Antennas for Wireless Communications*. Kluwer Academic Publishers, 2001.

[4] B. Pattan, *Robust Modulation Methods and Smart Antennas in Wireless Communications*. Prentice Hall PTR, 2000.

[5] E. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2003.

[6] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Transactions on Communications*, vol. 50, pp. 785–799, May 2002.

[7] A. Narula, M. Lopez, M. Trott, and G. Wornell, "Efficient use of side information in mulitple antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1423–1437, October 1998.

[8] K. Mukkavilli, A. Sabharwal, and B. Aazhang, "Design of multiple antenna coding schemes with channel feedback," *Proceedings of Asilomar Conference on Signals, Systems and Computers*, pp. 1009–1013, 2001.

[9] G. Jöngren and M. Skoglund, "Improving orthogonal space-time block codes by utilizing quantized feedback information," *2001 IEEE International Symposioum on Information Theory*, p. 220, June 2001.

[10] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Linear and nonlinear preequalization/equalization for mimo sytems with long-term channel state information at the transmitter," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 373–378, March 2004.

[11] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Communications Letters*, vol. 6, pp. 239–242, June 2002.

[12] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, pp. 2632–2640, September 2001.

[13] O. Simeone, U. Spagnolini, and Y. Bar-Ness, "Linear and non-linear precoding/decoding for mimo systems using the fading correlation at the transmitter," *2003 IEEE Workshop on Signal Processing*, pp. 6–10, June 2003.

[14] G. Stüber, *Principles of Mobile Communication*. Kluwer Academic Publishers, 2001.

[15] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[16] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*. Addison-Wesley Publishing Company, 1994.

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

[18] C. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, July 1948.

[19] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.

[20] I. Telatar, "Capacity of mulit-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, November/December 1999.

[21] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[22] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for mimo channels using the weighted mmse criterion," *IEEE Transactions on Communications*, vol. 49, pp. 2198–2206, December 2001.

[23] O. Berder, P. Rostaing, L. Collin, and G. Burel, "Minimum ber diagonal precoder for mimo systems," *IEEE Communications 2002*, pp. 1–6, 2002.

[24] L. Collin, O. Berder, P. Rostaing, and G. Burel, "Optimal minimum distance based precoder for mimo spatial multiplexing systems," *IEEE Transactions on Signal Processing*, vol. 52, pp. 617–618, March 2004.

[25] O. Simeone and U. Spagnolini, "Combined linear pre-equalization and blast equalization with channel correlation feedback," *IEEE Communications Letters*, vol. 7, pp. 487–489, October 2003.

[26] S. Simon and A. Moustakas, "Optimizing mimo antenna systems with channel covariance feedback," *IEEE Journal on Selected Areas of Communications*, vol. 21, pp. 406–418, April 2003.

[27] K. R. Rao and J. J. Hwang, *Techniques & Standards for Image, Video & Audio Coding*. Prentice Hall, 1996.

[28] M. I. Sezan and R. L. Lagendijk, *Motion Analysis and Image Sequence Processing*. Kluwer Academic Publishers, 1993.

[29] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.

[30] D. Love and R. Heath, "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Transactions on Signal Processing*, vol. 53, pp. 64–73, January 2005.

[31] ——, “Limited feedback precoding for spatial multiplexing systems,” *IEEE Globecom Conference, 2003*, vol. 4, pp. 1857–1861, December 2003.

[32] ——, “Limited feedback precoding for spatial multiplexing systems using linear receivers,” *IEEE Milcom Conference, 2003*, vol. 1, pp. 627–632, October 2003.

[33] J. Conway, R. Hardin, and N. Sloane, “Packing lines, planes, etc.: Packings in grassmannian spaces,” *Experimental Mathematics*, vol. 5, pp. 139–159, 1996.

[34] A. Barg and D. Nogin, “Bounds on packings of spheres in the grassmann manifold,” *IEEE Transactions on Information Theory*, vol. 48, pp. 2450–2454, September 2002.

[35] G. Jöngren and M. Skoglund, “Quantized feedback information in orthogonal space-time block coding,” *IEEE Transactions on Information Theory*, vol. 50, pp. 2473–2482, October 2004.

[36] W. Santipach and M. Honig, “Achievable rates for mimo fading channels with limited feedback and linear receivers,” *IEEE International Symposium on Spread-Spectrum Technology and Applications*, pp. 1–6, August/September 2004.

[37] ——, “Asymptotic performance of mimo wireless channels with limited feedback,” *MILCOM 2003*, vol. 1, pp. 141–146, October 2003.

[38] E. Baccarelli and M. Biagi, “Power-allocation policy and optimized design of multiple-antenna systems with imperfect channel estimation,” *IEEE Transactions on Vehicular Technology*, vol. 53, pp. 136–147, January 2004.

[39] ——, “Performance and optimized design of space-time codes for mimo wireless systems with imperfect channel estimates,” *IEEE Transactions on Signal Processing*, vol. 52, pp. 2911–2925, October 2004.

[40] A. Pascual-Iserte, A. Pérez-Neira, and M. Lagunas, “On power allocation strategies for maximum signal to noise and interference ratio in an ofdm-mimo system,” *IEEE Transactions on Wireless Communications*, vol. 3, pp. 808–822, May 2004.

[41] C. Martin, S. Bergman, and B. Ottersten, “Simple spatial mutiplexing based on imperfect channel estimates,” *IEEE International Conference on ASSP*, vol. 4, pp. 713–716, May 2004.

[42] G. V. F. Rey, M. Lamarca, “A robust transmitter design for mimo multicarrier systems with imperfect channel estimates,” *2003 IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 546–550, 2003.

[43] F. Rey, M. Lamarca, and G. Vazquez, “Robust power allocation algorithms for mimo ofdm systems with imperfect csi,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 1070–1085, 2005.

[44] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons, 1999.

[45] G. Primolevo, O. Simeone, and U. Spagnolini, “Effects of imperfect channel state information on the capacity of broadcast osdma-mimo systems,” *IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 546–550, July 2004.

[46] P. Ivanis and D. Drajic, “Combined optimal power allocation and adaptive modulation for mimo systems with imperfect csi,” *IEEE TELSIKS 2003*, vol. 1, pp. 167–170, October 2003.

[47] T. Weber and M. Meurer, “Imperfect channel state information in mimo transmission,” *IEEE Vehicular Technology Conference*, vol. 2, pp. 693–697, May 2004.

[48] Y. Song and S. Blostein, “Adaptive modulation for mimo systems with imperfect channel knowledge,” *IEEE Biennial Symposium on Communications*, pp. 428–430, May 2004.

[49] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation/Detection Theory*. Prentice-Hall, Inc, 1993.

[50] J. Yang and S. Roy, “On joint transmitter and receiver optimization for multiple-input-multiple-output (mimo) transmission systems,” *IEEE Transactions on Communications*, vol. 42, pp. 3221–3231, December 1994.

[51] A. Scaglione, G. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers part i: Unification and optimal design," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

[52] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Prentice-Hall, 1987.

[53] A. VanDerVeen, E. Deprettere, and A. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," *Proceedings of the IEEE*, vol. 81, pp. 1277–1308, 1993.

[54] J. G. Proakis, *Digital Communications*. McGraw-Hill, 2001.

[55] N. Al-Dhahir and S. Diggavi, "Guard sequence optimization for block transmission over linear frequency-selective channels," *IEEE Transactions on Communications*, vol. 50, pp. 938–946, 2002.

[56] H.-C. Wu and G. Gu, "Analysis of intercarrier and interblock interferences in wireless ofdm systems," *IEEE Globecom Conference, 2003*, vol. 2, pp. 784–788, 2003.

[57] A. Stamoulis, G. Giannakis, and A. Scaglione, "Block fir decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Transactions on Communications*, vol. 48, pp. 69–83, 2001.

[58] N. Al-Dhahir and J. Cioffi, "Block transmission over dispersive channels: Transmit filter optimization and realization, mmse-dfe receiver performance," *IEEE Transactions on Information Theory*, vol. 42, pp. 137–160, 1996.

[59] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge Universiy Press, 1985.

[60] J. Goldberg, *Matrix Theory with Applications*. McGraw-Hill, Inc, 1991.

[61] G. Golub and C. VanLoan, *Matrix Computations*. The Johns Hopkins University Press, 1996.