



Graduate Theses, Dissertations, and Problem Reports

2001

Space-time block coding with imperfect channel estimates

Dirk Alan Baker
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Baker, Dirk Alan, "Space-time block coding with imperfect channel estimates" (2001). *Graduate Theses, Dissertations, and Problem Reports*. 1103.
<https://researchrepository.wvu.edu/etd/1103>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

SPACE-TIME BLOCK CODING WITH IMPERFECT CHANNEL ESTIMATES

By

Dirk A. Baker

**Thesis submitted to the College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements
for the degree of**

**Master of Science
in
Electrical Engineering**

Approved by

**Dr. Matthew C. Valenti, Committee Chairperson
Dr. Biswajit A. Das
Dr. Larry A. Hornak**

Lane Department of Computer Science and Electrical Engineering

**Morgantown, West Virginia
2001**

**Keywords: Space-time block codes, channel estimation, wireless
Copyright 2001 Dirk A. Baker**

Abstract

SPACE-TIME BLOCK CODING WITH IMPERFECT CHANNEL ESTIMATES

by Dirk A. Baker

Space-time block coding (STBC) is a method that combines diversity and coding without a corresponding increase in bandwidth and with little complexity in the receiver structures, thus making it an ideal candidate for improving wireless communication performance and increasing data rates in systems with bandwidth constraints. The performance of STBC using a Quadrature Phase Shift Keying (QPSK) signal constellation has been shown to provide approximately 10 dB of improvement over the case of uncoded (QPSK) transmission in Rayleigh fading at a bit error rate of 10^{-3} . However, this performance was shown under the assumption that perfect channel state information (CSI) was available at the receiver. In this thesis, the performance of space-time block codes is analyzed when the receiver must rely on noisy, or imperfect, estimates of the channel. The results generated are in the form of bit error rates with varying degrees of errors introduced into the magnitude and the phase of the channel estimates. It is shown that for a QPSK signal constellation the system is robust to errors introduced into the amplitude of the channel estimate, but exhibits extreme performance degradation with relatively minor errors in the phase of the estimate. In fact, as the error in the phase approaches 0.5 radians the performance breaks down completely for even large values of received signal-to-noise ratio. A pilot sequence estimation scheme will be shown through simulation to provide performance within 2 dB of the case of perfect CSI at half the data rate.

ACKNOWLEDGMENTS

The author wishes to thank all the students in the WCRL, who have helped by recommending research papers and through discussions. I would also like to thank my various professors who have provided knowledge and insight into the various topics that have fascinated me in my engineering education. I would also like to thank the members of my committee, Dr. Das and Dr. Hornak, for taking the time to go through my work and for providing feedback.

I would like to extend a special thank you to Dr. Matthew Valenti, who has worked nearly as hard as I have to see that this thesis was completed. He has continuously provided useful feedback, and tactfully pointed out errors. He has given a great deal of time and effort to help me accomplish this task and for that I am grateful.

Finally, I would like to express my heartfelt gratitude to my wife, Siham, and my son Gabriel. Without their support and understanding I would never have been able to accomplish all that I have. This thesis is dedicated to my wife and son, the true joy and happiness in my life.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iii
Table of contents	iv
Introduction	1
Chapter 1 Modulation	3
1.1 Bandpass Signal Notation	4
1.2 Signal Space Concepts	7
1.3 Receiver Structures for Additive White Gaussian Noise Channels	11
1.4 Performance of Modulation Types in AWGN	20
1.5 Chapter Summary	24
Chapter 2 Wireless Channel and Diversity	25
2.1 Large-Scale Path Loss	26
2.1.1 Mean or Exponential Path Loss Model	27
2.1.2 Log-Normal Shadowing	30
2.2 Small-Scale Path Loss	30
2.2.1 Impulse Response Model	32
2.2.2 Time Dispersive nature of the Wireless Channel	33
2.2.3 Time Varying nature of the Wireless Channel	37
2.3 Multiple input Multiple output (MIMO) channel models	40
2.4 Diversity	42
2.5 Chapter Summary	47
Chapter 3 Space-Time Block Codes	48
3.1 General Theory	49
3.2 Performance with Perfect CSI	53
3.3 Chapter Summary	56
Chapter 4 Performance of STBC with Channel Estimation Errors	57
4.1 System Model	58
4.2 Simulation Results	60
4.3 Chapter Summary	69
Conclusion	70
Bibliography	73

Introduction

The field of wireless communications and networks has experienced explosive growth. The demand and purchase of cellular telephones is predicted to soon exceed the purchase and use of traditional wired telephones. The market for handheld devices that boast some form of wireless connectivity has skyrocketed and continued growth is predicted. Along with this rapid growth comes the consumer demand for more and better applications, improved performance, and increased data rates.

All these improvements must be accomplished under a considerable number of constraints. The wireless channel is by its nature random and unpredictable, and in general the performance of a device is poorer over a wireless channel than over a “wired” channel. The spectrum or bandwidth available to the service provider is often limited and the allotment of new spectrum by the federal government is often slow in coming. Also, the power requirements are that devices should use as little power as possible to conserve battery life and keep the products small. The designers for wireless systems face a two-part challenge, increase data rates and improve performance while incurring little or no increase in bandwidth or power.

Several different techniques have been discovered that can improve the performance of a wireless channel by providing multiple copies of a transmitted signal to the receiver. These multiple copies can be sent over different time slots, frequencies, or antennas. These techniques are known collectively as diversity.

Space-time block codes were devised as a means to provide antenna diversity by using multiple transmit and receive antennas. The space-time block codes were found to improve bit error rate performance dramatically without requiring an extensive increase in bandwidth or the design of overly complex receivers. In order to facilitate an understanding of the performance of space-time block codes it is useful to simulate the system to provide results to confirm the performance and validate the theoretical predictions.

It is the goal of this thesis to demonstrate the performance of space-time block codes. Initially, it will be assumed that the channel conditions are known exactly to the receiver

and the system simulated from this point of view. Finally the system will be tested under the assumption that some form of estimation technique was used to generate an imperfect estimate of the channel conditions.

The first two chapters of this thesis will introduce the background material necessary to be able to adequately describe space-time block coding. The first topic covered will be modulation theory. It is here that the concept of signal spaces is developed and the methods to calculate bit error rate performance of various modulation types in an Additive White Gaussian Noise (AWGN) channel. This is the channel model generally associated with a wired channel. The next topic will explain the nature of the wireless channel and the various methods that can be used to develop an accurate statistical model of the channel. The concept of diversity will be discussed in detail and the performance of certain modulation types in a Rayleigh fading channel will be presented.

The final two chapters will introduce space-time block codes and simulation results will be presented to show the bit error rate performance of a Quadrature Phase Shift Keying signal constellation used with space-time block coding. These results will be generated for the case of perfect channel state information at the receiver and the case of channel estimates used at the receiver.

CHAPTER 1: Modulation

Most communication systems in use today are digital communication systems. Whether they work directly with digital data coming out of some type of computing device or the inputs are analog waveforms that are sampled and quantized, the end product to be transmitted over some arbitrary channel is digital data. Most types of channels through which this data must travel, such as telephone wires, fiber optic cables, the atmosphere, etc., have certain characteristics and constraints that force the transmitted data to occupy a particular band of frequencies, or bandwidth. In order to convert the digital data available at the source into a signal that can be efficiently transmitted through the channel the data goes through a process called modulation.

Modulation shifts the spectrum of the digital data, or baseband signal, in such a way as to create a bandpass signal. A bandpass signal is one in which the signal spectrum is non-negligible only about some frequency f_c , called the carrier frequency. For most systems, the carrier frequency is determined by a sinusoidal carrier waveform that is modulated by the baseband waveform to produce the transmitted signal. A common method for accomplishing this is to take a stream of digital information at baseband, filter the waveform, then mix it with a sinusoidal carrier. The sinusoidal carrier, generated by an oscillator, serves to shift the frequency of the baseband waveform into a spectrum suitable for transmission over the channel of interest. Figure 1.1 shows a simplified communication system with no source or channel coding.

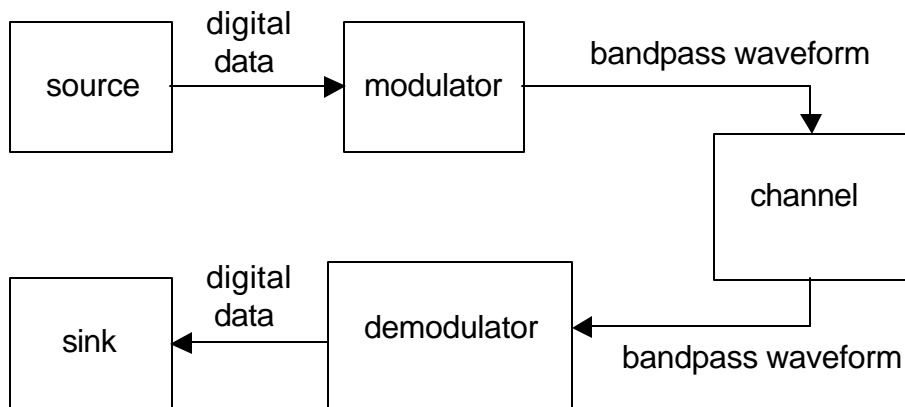


Figure 1.1 General structure of an uncoded communication system

1.1 Bandpass signal notation

It is extremely important to be able to compare and contrast the advantages of different communication systems and techniques in order to be able to choose the best system for a given application. It would be convenient if we could compare two different systems that operate at different frequencies without having to account for the effects introduced by different carrier frequencies. This goal can be accomplished if we can represent bandpass signals as equivalent lowpass signals that retain all the necessary properties for evaluating the performance of the system. The following is a method for generating equivalent lowpass signals from bandpass signals based on the work in [Hay94].

The spectrum of a bandpass signal will be contained in a relatively narrow band of frequencies that are centered about the carrier frequency $\pm f_c$. Figure 1.2 shows an example of the spectrum of an arbitrary signal $s(t)$.

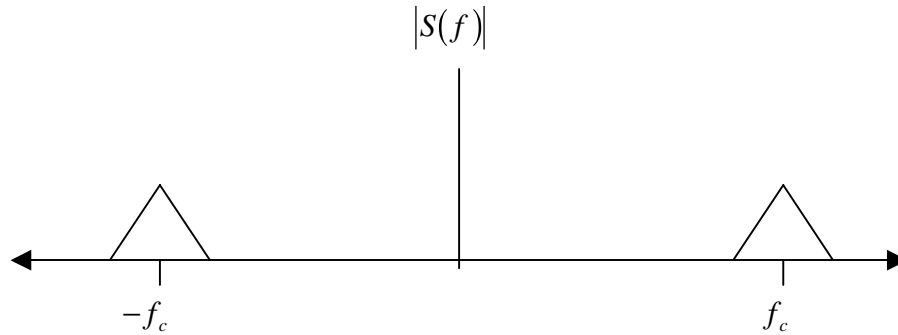


Figure 1.2 Magnitude of the Frequency Response of Bandpass Signal

First we take the positive frequency components in the signal $S(f)$ and build the “analytic” signal

$$S_+(f) = 2u(f)S(f) \quad (1.1)$$

where $u(f)$ is the unit step function and $S(f)$ is the Fourier transform of $s(t)$. To get an expression in the time domain we can take the inverse Fourier transform and get

$$\begin{aligned} s_+(t) &= F^{-1}\{S_+(f)\} \\ s_+(t) &= F^{-1}\{2u(f)\} * F^{-1}\{S(f)\} \\ &\text{where } * \text{ represents convolution} \\ s_+(t) &= \left(\delta(t) + \frac{j}{\pi t} \right) * (s(t)) \\ s_+(t) &= s(t) + \left(s(t) * \frac{j}{\pi t} \right) \end{aligned} \quad (1.2)$$

if we set $\hat{s}(t) = \frac{1}{\pi t} * s(t)$,

where $\hat{s}(t)$ denotes the Hilbert Transform,

then

$$s_+(t) = s(t) + j\hat{s}(t) \quad (1.3)$$

Since $s_+(t)$ is still a bandpass signal we can translate it in frequency to get an equivalent lowpass signal $s_\ell(t)$ which can be described by the following equation

$$s_\ell(t) = s_+(t)e^{-j2\pi f_c t} \quad (1.4)$$

We can now describe $s(t)$ three different ways. These are known as Complex Envelope, Quadrature, and Magnitude and Phase notation. In Complex Envelope notation

$$s(t) = \text{Re}\left[s_\ell(t)e^{-j2\pi f_c t}\right] \quad (1.5)$$

where $s_\ell(t)$ is a complex valued lowpass signal and $\text{Re}[\cdot]$ is the real part of the complex signal.

In Quadrature notation

$$s(t) = x(t)\cos(2\pi f_c t) - y(t)\sin(2\pi f_c t) \quad (1.6)$$

where $x(t)$ and $y(t)$ are real valued lowpass signals called the in-phase and quadrature components of $s(t)$, respectively.

Finally, there is the Magnitude and Phase notation

$$s(t) = a(t)\cos(2\pi f_c t + \theta(t)) \quad (1.7)$$

where $a(t)$ and $\theta(t)$ are real valued lowpass signals known as the magnitude and phase of $s(t)$. These three notations are related to each other and we can move from one representation to the other with the following equations

$$x(t) = a(t) \cos(\theta(t)) \quad (1.8a)$$

$$y(t) = a(t) \sin(\theta(t)) \quad (1.8b)$$

$$a(t) = \sqrt{x^2(t) + y^2(t)} \quad (1.8c)$$

$$\theta(t) = \tan^{-1} \left[\frac{y(t)}{x(t)} \right] \quad (1.8d)$$

$$s_1(t) = x(t) + jy(t) \quad (1.8e)$$

1.2 Signal Space Concepts

In order to develop a general method for the generation, analysis, and reception of signals it is necessary to build a framework wherein signals of varying types can be represented in a similar manner. The idea behind signal spaces is to perform this very task. In order to understand signal spaces and their inherent usefulness we must first define several terms and conditions. A signal, which we will call $x(t)$, is a function of time defined over the interval $[a, b]$. The inner product of two signals is expressed as $\langle x_1(t), x_2(t) \rangle$, where

$$\langle x_1(t), x_2(t) \rangle = \int_a^b x_1(t) x_2^*(t) dt \quad (1.9)$$

Two signals are orthogonal if

$$\langle x_1(t), x_2(t) \rangle = 0 \quad (1.10)$$

The norm of a signal is expressed as $\|x(t)\|$, where

$$\begin{aligned} \|x(t)\| &= [\langle x(t), x(t) \rangle]^{1/2} \\ \|x(t)\| &= \sqrt{\int_a^b |x(t)|^2 dt} \end{aligned} \quad (1.11)$$

The energy of a signal is described by the equation

$$E_x = \|x(t)\|^2$$

$$E_x = \int_a^b |x(t)|^2 dt \quad (1.12)$$

A signal is considered to be normalized if it satisfies the following equations:

$$\|x(t)\| = 1$$

$$E_x = \|x(t)\|^2 = (1)^2 = 1 \quad (1.13)$$

A signal may be expressed as

$$x(t) = \sum_{i=1}^K w_i f_i(t) \quad (1.14)$$

where w_i are constant weights or coefficients and $f_i(t)$ are basis functions from a complete and orthonormal set of basis functions.

The dimensionality of a signal set,

$$S = \{s_1(t), s_2(t), \dots, s_{M-1}(t), s_M(t)\},$$

composed of basis functions from a complete orthonormal set of basis functions,

$$F = \{f_1(t), f_2(t), \dots, f_{k-1}(t), f_k(t)\},$$

is equal to K , the number of basis functions.

A set of basis functions F can be considered to be a complete orthonormal basis for a signal set S if it satisfies three conditions:

- 1) Complete condition: All signals in the set S can be expressed as a linear combination of basis functions from the set F .
- 2) Orthogonal condition: All pairs of basis functions are orthogonal so that

$$\langle f_i(t), f_j(t) \rangle = 0 \quad \forall_{i,j} \text{ where } i \neq j$$

- 3) Normal condition: All basis functions are normalized so that

$$\|f_i(t)\| = 1 \quad \forall i$$

The concept of using basis functions to represent any given signal in a set of signals leads to some simplifications that will make it convenient for analyzing systems.

Now, any given signal can be represented as $s_i(t) = \sum_{k=1}^K w_k f_k(t)$. For different signals in the set the values for w_k will vary but the basis functions remain constant. This implies that the relevant information to distinguish one signal from another will be contained in the coefficients w_k . Therefore, we can represent the different signals in the set by a vector that contains the coefficients of the basis functions, so that $s_i(t)$ is now equivalent to

$$\mathbf{s} = [w_1 \quad w_2 \quad \dots \quad w_{K-1} \quad w_K]$$

Notice that the vector \mathbf{s} above is shown in boldface type, this convention will be used throughout the rest of this thesis whenever we are dealing with vectors.

Having established the equivalence between a signal and its vector representation, a signal space can now be created from the vectors comprising a signal set. As a simple example we will develop the signal space representation for phase shift keying. In phase shift keying the general form for any signal is

$$s_m(t) = \sqrt{\frac{2E_s}{T}} g(t) \cos\left(\frac{2\pi}{M}(m-1)\right) \cos(2\pi f_c t) - \sqrt{\frac{2E_s}{T}} g(t) \sin\left(\frac{2\pi}{M}(m-1)\right) \sin(2\pi f_c t) \quad (1.15)$$

where $g(t)$ is the signal pulse shape and $\frac{2\pi}{M}(m-1)$, $m = 1, 2, \dots, M$, is the phase.

Let's assume that the basis of this signal set consists of the following two functions

$$\begin{aligned} f_1(t) &= \sqrt{\frac{2}{T}} g(t) \cos(2\pi f_c t) \\ f_2(t) &= \sqrt{\frac{2}{T}} g(t) \sin(2\pi f_c t) \end{aligned} \quad (1.16)$$

This means that each signal $s_m(t)$ will be represented as a two dimensional vector of the form

$$\mathbf{s}_m = [s_{m1} \ s_{m2}]$$

$$\mathbf{s}_m = \left[\sqrt{E_s} \cos\left(\frac{2\pi}{M}(m-1)\right) \ \sqrt{E_s} \sin\left(\frac{2\pi}{M}(m-1)\right) \right] \quad (1.17)$$

where $m = 1, 2, \dots, M$.

In the case where $M = 2$ this system is called Binary Phase Shift Keying (BPSK) and when $M = 4$ it is called Quadrature Phase Shift Keying (QPSK), or 4-PSK. The signal space diagram for these two values of M can be seen in the figure 1.3 below.

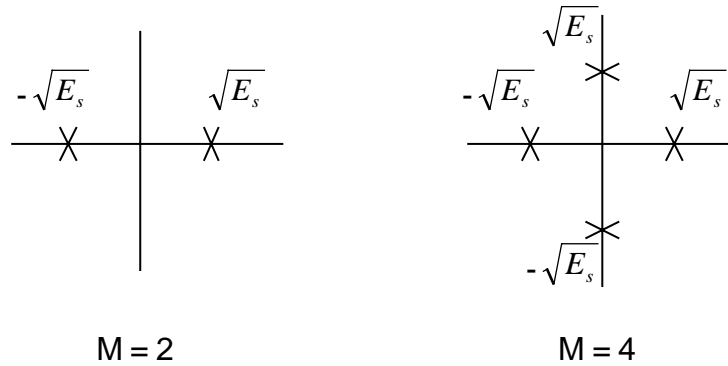


Figure 1.3

The Euclidean distance between two signals, $\mathbf{x}_i(t)$ and $\mathbf{x}_j(t)$, is defined to be

$$d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\| \quad (1.18)$$

where \mathbf{x}_i and \mathbf{x}_j are the vector representations of the respective signals.

Therefore, the Euclidean distance between two signals is nothing more than the distance between the points plotted on a signal space diagram.

In the case of M-PSK this can be shown to be

$$d_{ij} = \sqrt{2E_s \left(1 - \cos\left(\frac{2\pi}{M}(i-j)\right) \right)} \quad \forall_{ij} \quad i \neq j \quad (1.19)$$

The minimum Euclidean distance, denoted by d_{\min} , is the smallest distance between pairs of signal points, or in the case of PSK it is the distance between adjacent signal points. To continue the above example for PSK

$$d_{\min} = \sqrt{2E_s \left(1 - \cos\left(\frac{2\pi}{M}\right) \right)} \quad (1.19)$$

1.3 Receiver Structures for Additive White Gaussian Noise Channels

In order to determine the performance of a given type of modulation it is first necessary to determine the way in which the transmitted signal is to be received. A receiver can be separated into the front end, or demodulator, and the back end, or detector. In this chapter we will examine two types of receiver front ends, one based on the use of correlators and the other on matched filters. We will also discuss two detectors, or decision rules, known as the Maximum a Posteriori and the Maximum Likelihood rule.

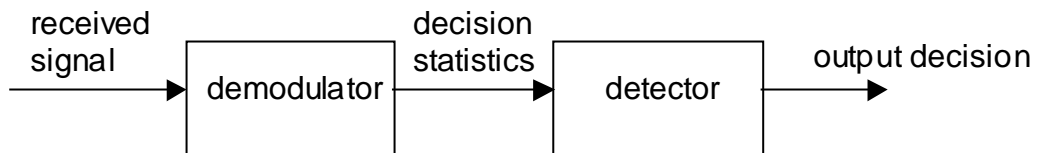


Figure 1.4 Generic diagram of a Receiver

If we want to find an optimum implementation of a receiver we must first establish the representations of signals in the system and then figure out a method to decide which signal was sent given the signal available at the receiver. The receiver will be an implementation of the demodulation and decision rule.

Assume we will transmit a signal $s(t)$ from the set

$$S = \{s_1(t), s_2(t), \dots, s_{M-1}(t), s_M(t)\},$$

where $s(t)$ is non-zero in the interval $[0, T]$. The transmitted signal is composed of weighted basis functions from the set

$$F = \{f_1(t), f_2(t), \dots, f_{K-1}(t), f_K(t)\}.$$

Therefore, each transmitted signal is of the form

$$s_m(t) = \sum_{k=1}^K s_{mk} f_k(t) \quad (1.20)$$

where

$$s_{mk} = \int_0^T s_m(t) f_k(t) dt$$

The signal available to the receiver will be the original signal corrupted by additive white Gaussian noise (AWGN). The received signal will be denoted as

$$r(t) = s(t) + n(t)$$

where $n(t)$ is the AWGN process with two-sided noise spectral density $\frac{N_o}{2}$ and $s(t)$ is the transmitted signal.

The noise can be represented as

$$n(t) = n'(t) + \sum_{k=1}^K n_k f_k(t) \quad (1.21)$$

where

$$n_k = \int_0^T n(t) f_k(t) dt$$

and since $n'(t)$ is orthogonal to all the possible transmitted signals $s_m(t)$, it does not affect performance and can be dropped in further equations. The received signal can now be represented as

$$\begin{aligned}
 r(t) &= \sum_{k=1}^K s_{mk} f_k(t) + \sum_{k=1}^K n_k f_k(t) \\
 r(t) &= \sum_{k=1}^K (s_{mk} + n_k) f_k(t) \\
 r(t) &= \sum_{k=1}^K r_k f_k(t)
 \end{aligned} \tag{1.22}$$

where $r_k = s_{mk} + n_k$

Now we can deal exclusively with the vectors

$$\begin{aligned}
 \mathbf{s} &= [s_1 s_2 \dots s_K] \\
 \mathbf{n} &= [n_1 n_2 \dots n_K] \\
 \mathbf{r} &= [r_1 r_2 \dots r_K] \\
 \mathbf{r} &= \mathbf{s} + \mathbf{n}
 \end{aligned} \tag{1.23}$$

What we need the demodulator to do in the receiver is to provide our detector with the vector \mathbf{r} of sufficient statistics. To accomplish this we can correlate the received signal with each basis function used to generate the transmitted signals. Figure 1.5 is a diagram of this type of correlation demodulator. Figure 1.6 shows an equivalent method, the matched filter demodulator, which accomplishes the same task using a bank of filters that are each matched to a specific basis function. The impulse response of each matched filter is simply the basis function reversed in time and slid over by time T , i.e.

$$h_k(t) = f_k(T - t)$$

For the output of the matched filter to be identical to the output of the correlator, it must be sampled at time $t = T$.

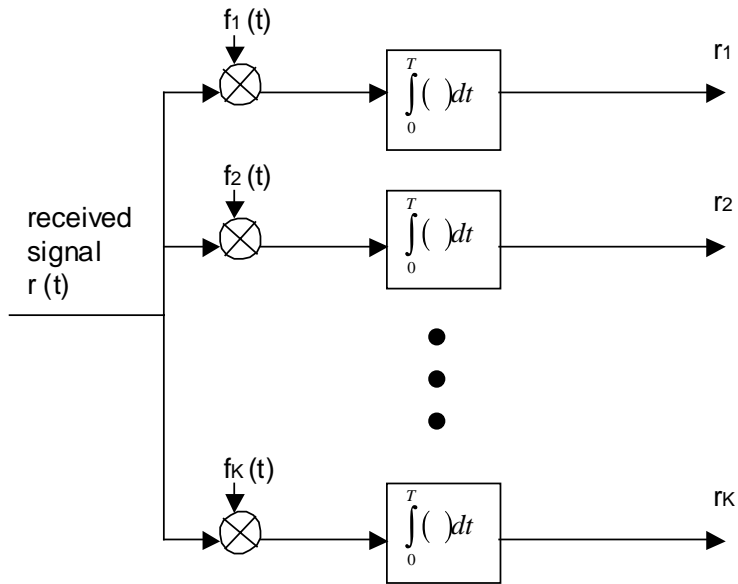


Figure 1.5 Demodulator using a bank of Correlators

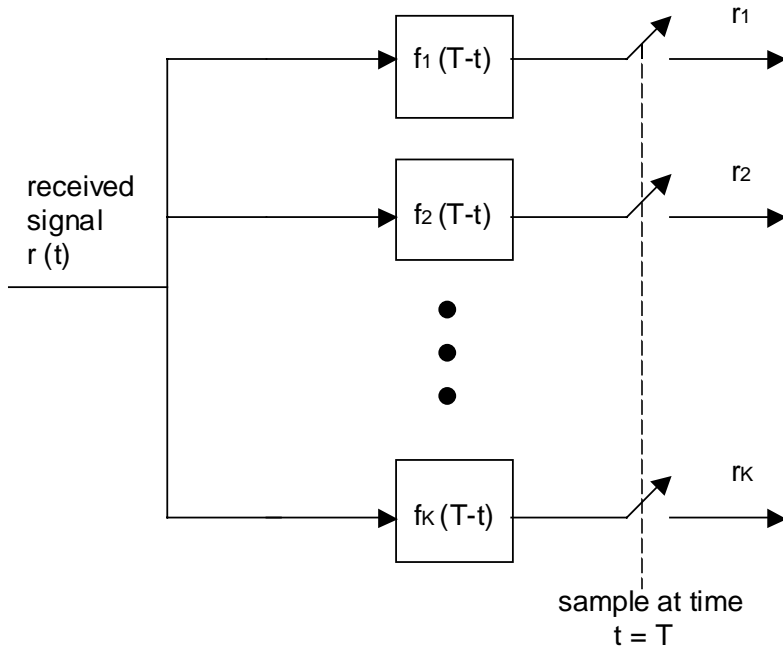


Figure 1.6 Demodulator using a bank of Matched Filters

After the signal has been received and demodulated the remaining vector \mathbf{r} can be used to calculate the estimate of the transmitted signal $\hat{\mathbf{s}}$, and the probability that a given estimate is equal to the original signal \mathbf{s} that was transmitted. If the probability of a symbol error is

$$P_s = \Pr[\hat{\mathbf{s}} \neq \mathbf{s}] \quad (1.24)$$

and the probability that a given signal $s_i(t)$ will be transmitted is p_i , $i = 1, 2, \dots, M$, then the goal of the optimal receiver is to choose an estimate that will minimize P_s . This can be accomplished by choosing as an estimate $\hat{\mathbf{s}} = \mathbf{s}_m$, such that

$$\Pr(\mathbf{s}_m | \mathbf{r}) > \Pr(\mathbf{s}_i | \mathbf{r}) \quad \forall i \neq m$$

or using Bayes rule

$$\Pr(\mathbf{s}_m | \mathbf{r}) = \frac{\Pr(\mathbf{r} | \mathbf{s}_m) \Pr(\mathbf{s}_m)}{\Pr(\mathbf{r})}$$

we can now express

$$\begin{aligned} \Pr(\mathbf{r} | \mathbf{s}_m) \Pr(\mathbf{s}_m) &> \Pr(\mathbf{r} | \mathbf{s}_i) \Pr(\mathbf{s}_i) \\ \Pr(\mathbf{r} | \mathbf{s}_m) p_m &> \Pr(\mathbf{r} | \mathbf{s}_i) p_i \quad \forall i \neq m \end{aligned} \quad (1.25)$$

The last equation above is the maximum a posteriori, or MAP, rule of detection. It requires that the receiver knows the conditional probability that \mathbf{r} was received given that \mathbf{s} was transmitted and the probability that \mathbf{s} would be transmitted. If the individual signal transmission probabilities (p_1, p_2, \dots, p_M) are unknown to the receiver or if they are equal for all p_i then the receiver minimizes P_s by choosing $\hat{\mathbf{s}} = \mathbf{s}_m$ such that

$$\Pr(\mathbf{r} | \mathbf{s}_m) > \Pr(\mathbf{r} | \mathbf{s}_i) \quad \forall i \neq m \quad (1.26)$$

The above equation is known as the maximum likelihood, or ML, decision rule. The MAP and the ML decision rule will make the same decision as to which was the most likely transmitted symbol as long as the a priori probabilities of the transmitted symbols, p_i , are equal.

In order to evaluate the conditional probabilities shown in equations 1.25 and 1.26 we must first find the distributions of \mathbf{s} and \mathbf{n} . The vector \mathbf{s} is composed of scalar

coefficients that are conditionally deterministic and therefore do not contribute to the random nature of the received vector. The noise vector \mathbf{n} is composed of components that are Gaussian with a mean of zero and a variance

$$\sigma_n^2 = \frac{N_o}{2} \quad (1.27)$$

All the components of the noise vector \mathbf{n} are identically distributed and are statistically independent from each other. Now we can find the conditional distribution $p(\mathbf{r}|\mathbf{s})$. The mean of a random variable can be found by taking the expected value of the random variable. If we evaluate each element of the receive vector individually we find that

$$\begin{aligned} E[r_k|s_k] &= E[s_k + n_k|s_k] \\ E[r_k|s_k] &= s_k + E[n_k] \\ E[r_k|s_k] &= s_k \end{aligned} \quad (1.28)$$

where the variance of this conditional distribution is found to be

$$\sigma_{r|s}^2 = \sigma_n^2 = \frac{N_o}{2}$$

Since each noise element is a statistically independent gaussian random variable each component of the receive vector, r_k conditioned on s_k , is also a statistically independent gaussian random variable with a mean of s_k and a variance equal to $\frac{N_o}{2}$. Therefore the probability density function, or pdf, of the vector \mathbf{r} conditioned on \mathbf{s}_m can be broken down into the product of the pdf's of the individual components of the two vectors as shown by

$$p(\mathbf{r}|\mathbf{s}_m) = \prod_{k=1}^K p(r_k|s_{mk}) \quad (1.29)$$

In order to evaluate this it is necessary to know the pdf of a gaussian random variable which is given in [Gar94] as

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \quad (1.30)$$

where σ^2 is the variance of x and m is the mean of x .

Now we can evaluate the conditional pdf by using the pdf of equation 1.30 in equation 1.29 to yield the following

$$\begin{aligned}
 p(\mathbf{r}|\mathbf{s}_m) &= \prod_{k=1}^K \frac{1}{\sqrt{\pi N_o}} \exp\left\{-\frac{(r_k - s_{mk})^2}{N_o}\right\} \\
 p(\mathbf{r}|\mathbf{s}_m) &= \frac{1}{(\pi N_o)^{\frac{K}{2}}} \exp\left\{-\sum_{k=1}^K \frac{(r_k - s_{mk})^2}{N_o}\right\}
 \end{aligned} \tag{1.31}$$

The MAP decision rule can be implemented by designing a system to choose as its estimated symbol, $\hat{\mathbf{s}}$, whichever vector \mathbf{s} that maximizes the conditional pdf of the received vector. This can be expressed in an efficient mathematical form as

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}_m \in S} \{p_m p(\mathbf{r}|\mathbf{s}_m)\} \tag{1.32}$$

where $\hat{\mathbf{s}}$ is the estimate of the transmitted signal vector and S is the set of all possible transmitted signal vectors. The above 'arg max' operator means to set $\hat{\mathbf{s}}$ equal to the possible transmitted signal vector that maximizes the function inside the brackets. If we substitute equation 1.31 into 1.32 and simplify we will get an equation we can use to design an optimal MAP receiver.

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}_m} \left\{ p_m (\pi N_o)^{-\frac{K}{2}} \exp\left\{-\sum_{k=1}^K \frac{(r_k - s_{mk})^2}{N_o}\right\} \right\}$$

now take the natural logarithm of both sides

$$\begin{aligned}\hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \ln \left(p_m (\pi N_o)^{-\frac{K}{2}} \exp \left\{ -\sum_{k=1}^K \frac{(r_k - s_{mk})^2}{N_o} \right\} \right) \right\} \\ \hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \ln(p_m) + \ln \left((\pi N_o)^{-\frac{K}{2}} \right) + \ln \left(\exp \left\{ -\sum_{k=1}^K \frac{(r_k - s_{mk})^2}{N_o} \right\} \right) \right\} \\ \hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \ln(p_m) - \frac{K}{2} \ln(\pi N_o) - \sum_{k=1}^K \frac{(r_k - s_{mk})^2}{N_o} \right\} \\ \hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \ln(p_m) - \frac{K}{2} \ln(\pi N_o) - \frac{1}{N_o} \sum_{k=1}^K (r_k^2 - 2s_{mk}r_k + s_{mk}^2) \right\}\end{aligned}$$

Note that the term $\frac{K}{2} \ln(\pi N_o)$ and the term r_k^2 are common to all possible received signal vectors and can therefore be discarded.

$$\begin{aligned}\hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \ln(p_m) - \frac{1}{N_o} \sum_{k=1}^K (-2s_{mk}r_k + s_{mk}^2) \right\} \\ \hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \ln(p_m) + \frac{2}{N_o} \sum_{k=1}^K s_{mk}r_k - \frac{1}{N_o} \sum_{k=1}^K s_{mk}^2 \right\}\end{aligned}$$

where $\sum_{k=1}^K s_{mk}^2 = E_m$, the energy of \mathbf{s}_m ,

$$\begin{aligned}\hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \frac{N_o}{2} \left(\ln(p_m) + \frac{2}{N_o} \sum_{k=1}^K s_{mk}r_k - \frac{E_m}{N_o} \right) \right\} \\ \hat{\mathbf{s}} &= \arg \max_{\mathbf{s}_m} \left\{ \frac{N_o}{2} \ln(p_m) + \sum_{k=1}^K s_{mk}r_k - \frac{E_m}{2} \right\}\end{aligned} \tag{1.33}$$

This is the equation we will use to build the MAP receiver. In order to get the term z_m ,

$$z_m = \sum_{k=1}^K s_{mk}r_k$$

It is possible for the receiver to use M correlators to correlate the received signal against all possible transmitted signal vectors. However, for a more efficient implementation it is necessary to only correlate the received signal, $r(t)$, with each possible basis function, $f_i(t)$, and then do a vector-matrix multiply to get the inner product between the signal vectors and the received vector. The vector-matrix multiply will follow the equation

$$\mathbf{z} = \mathbf{S}\mathbf{r}$$

where \mathbf{r} is the received vector and \mathbf{S} is a matrix whose rows are the signal vectors of each possible transmitted signal as shown below

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_M \end{bmatrix}$$

The new demodulator outputs will then be the statistics z_m , $m = 1, 2, \dots, M$. At this point, it is necessary for the receiver to take into account the a priori probabilities of the transmitted signals and their respective energies, before choosing the largest of the generated metrics and setting $\hat{\mathbf{s}}$ equal to the transmitted signal corresponding to the chosen metric. Figure 1.7 is a diagram of the complete receiver implementing the MAP decision rule with a bank of K correlators.

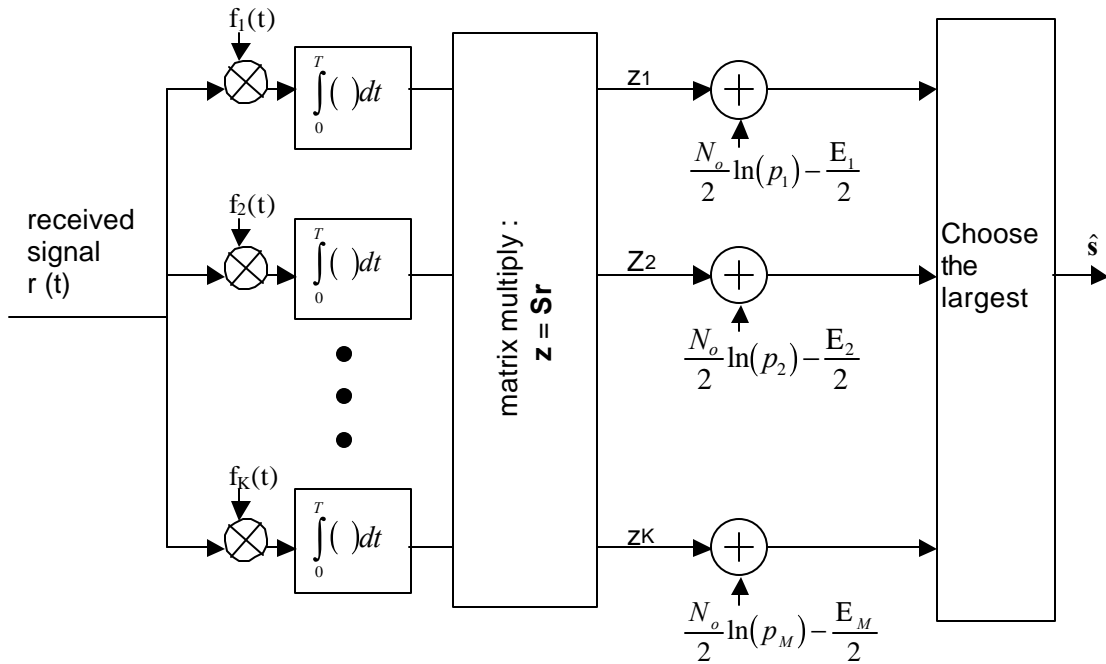


Figure 1.7 Receiver implementing the MAP rule

In the above receiver, if the probabilities that a signal was transmitted are equal or unknown, we can remove the terms that account for the a priori probabilities of the signals, $\frac{N_o}{2} \ln(p_m)$, and create a ML decision rule receiver. In the case that the energies of each transmitted signal are equal, e.g. M-PSK, then we can drop the E_m terms from the receiver. So, for M-PSK with equal p_i the output decision is based solely on the metric z_i .

1.4 Performance of Modulation types in AWGN

The primary function used in this thesis to analyze the performance of various systems is the Bit Error Rate (BER) or Symbol Error Rate (SER) of a given modulation type over a given channel. In this introductory chapter we are analyzing the performance of various

standard modulation types in the presence of AWGN. The SER of a given modulation type is dependent on the Euclidean distance between signal points on its respective signal space diagram. This means it is independent of the choice of basis functions. For many modulation types it is extremely difficult to obtain a closed form solution to the BER or SER of the system because such an expression requires integration over irregular shaped K-dimensional regions. It becomes convenient to use a different technique that can provide an upper bound on the probability of error and is much easier to calculate. The predominant technique to simplify these calculations and provide an upper bound on error probability is known as the Union Bound. A brief explanation of the Union Bound is found in [Pro95], and its relationship to Euclidean distance will now be shown.

Earlier in this chapter it was shown that symbol error probability could be found using the relationship

$$\begin{aligned}
 P_s &= \Pr[\hat{\mathbf{s}} \neq \mathbf{s}] \\
 P_s &= \sum_{\mathbf{s}_i \in \mathcal{S}} \Pr[\hat{\mathbf{s}} \neq \mathbf{s} | \mathbf{s} = \mathbf{s}_i] p_i
 \end{aligned}
 \tag{1.34}$$

The probability inside the summation in equation 1.34 is found by integrating the pdf of the received signal given that signal \mathbf{s}_i was transmitted, which was found in equation 1.31, over all the regions in the signal space except that region corresponding to the i^{th} signal. This relationship can be described by

$$\Pr[\hat{\mathbf{s}} \neq \mathbf{s} | \mathbf{s} = \mathbf{s}_i] = \int_{\bar{R}_i} p(\mathbf{r} | \mathbf{s}_i) d\mathbf{r}
 \tag{1.35}$$

where \bar{R}_i is all the decision regions in the signal space except the one corresponding to \mathbf{s}_i . The union bound is used at this point to approximate the above integral, which is an integral over a large boundary with arbitrary shape, by using a series of smaller integrals over more easily defined regions. In the course of this approximation it is generally necessary to integrate over the same regions more than once, this is the reason this technique is only an upper bound on the error probability. Using this technique will involve

replacing the probability density function with a probability using metrics. The metric used, Z_j , corresponds to the decision metrics used in the MAP or ML decision rules.

So, the union bound yields

$$P_s = \int_{\bar{R}_i} p(\mathbf{r}|\mathbf{s}_i) d\mathbf{r} \leq \sum_{\substack{j=1 \\ i \neq j}}^M \Pr[Z_j \geq Z_i]$$

if we consider the ML decision rule, then

$$Z_j = \sum_{k=1}^K s_{jk} r_k - \frac{E_j}{2}$$

$$\Pr[Z_j \geq Z_i] = \Pr\left[\sum_{k=1}^K s_{jk} r_k - \frac{E_j}{2} \geq \sum_{k=1}^K s_{ik} r_k - \frac{E_i}{2}\right]$$

note that $E_j = \sum_{k=1}^K s_{jk}^2$

$$\Pr[Z_j \geq Z_i] = \Pr\left[\sum_{k=1}^K s_{jk} r_k - \frac{1}{2} \sum_{k=1}^K s_{jk}^2 \geq \sum_{k=1}^K s_{ik} r_k - \frac{1}{2} \sum_{k=1}^K s_{ik}^2\right]$$

$$\Pr[Z_j \geq Z_i] = \Pr\left[\sum_{k=1}^K \left(s_{jk} r_k - \frac{1}{2} s_{jk}^2\right) \geq \sum_{k=1}^K \left(s_{ik} r_k - \frac{1}{2} s_{ik}^2\right)\right]$$

inside the parentheses multiply by -2 and add r_k^2

$$\Pr[Z_j \geq Z_i] = \Pr\left[\sum_{k=1}^K \left(s_{jk}^2 - 2s_{jk} r_k + r_k^2\right) \leq \sum_{k=1}^K \left(s_{ik}^2 - 2s_{ik} r_k + r_k^2\right)\right] \quad (1.36)$$

$$\Pr[Z_j \geq Z_i] = \Pr\left[\sum_{k=1}^K \left(s_{jk} - r_k\right)^2 \leq \sum_{k=1}^K \left(s_{ik} - r_k\right)^2\right]$$

Notice that the terms inside the summation are squared Euclidean distances, therefore

$$\Pr[Z_j \geq Z_i] = \Pr[d^2(\mathbf{r}, \mathbf{s}_j) \leq d^2(\mathbf{r}, \mathbf{s}_i)]$$

we can simplify this to

$$\Pr[Z_j \geq Z_i] = \Pr[d(\mathbf{r}, \mathbf{s}_j) \leq d(\mathbf{r}, \mathbf{s}_i)]$$

now substitute this expression back into the express for P_s

$$\begin{aligned}
P_s &\leq \sum_{i=1}^M p_i \sum_{\substack{j=1 \\ i \neq j}}^M \Pr[Z_j \geq Z_i] \\
P_s &\leq \sum_{i=1}^M p_i \sum_{\substack{j=1 \\ i \neq j}}^M \Pr[d(\mathbf{r}, \mathbf{s}_j) \leq d(\mathbf{r}, \mathbf{s}_i)] \\
\Pr[d(\mathbf{r}, \mathbf{s}_j) \leq d(\mathbf{r}, \mathbf{s}_i)] &= Q\left(\frac{d_{ij}}{\sqrt{2N_o}}\right) \\
P_s &\leq \sum_{i=1}^M p_i \sum_{\substack{j=1 \\ i \neq j}}^M Q\left(\frac{d_{ij}}{\sqrt{2N_o}}\right) \tag{1.37}
\end{aligned}$$

Equation 1.37 is the Union Bound for AWGN channels and provides a fairly tight bound for many modulation types for reasonably large values of Signal to Noise Ratio. In order to calculate bit error rates from the symbol error rate P_s , it is first necessary to know the way in which bits are mapped to the signals in the set. For Phase Shift Keying it is not uncommon to use a gray code to map bits to symbols. This means that adjacent symbols will differ in only one bit position. For example, in QPSK, signal one might map to data (10) and the adjacent signal two would map to data (11). Since it is most likely that a symbol error will occur between adjacent signals then a symbol error will result in only a single bit error. Therefore, the bit error rate for M-ary PSK using the Union Bound can be shown to be

$$P_b \approx \frac{P_s}{\log_2 M} \tag{1.38}$$

Figure 1.8 plots the curves of estimated Bit Error Rate versus received SNR per bit for various types of M-ary Phase Shift Keying. These curves were generated using the Union Bound. As can be seen, the energy efficiency decreases with increasing M, although larger values of M correspond to better bandwidth efficiency.

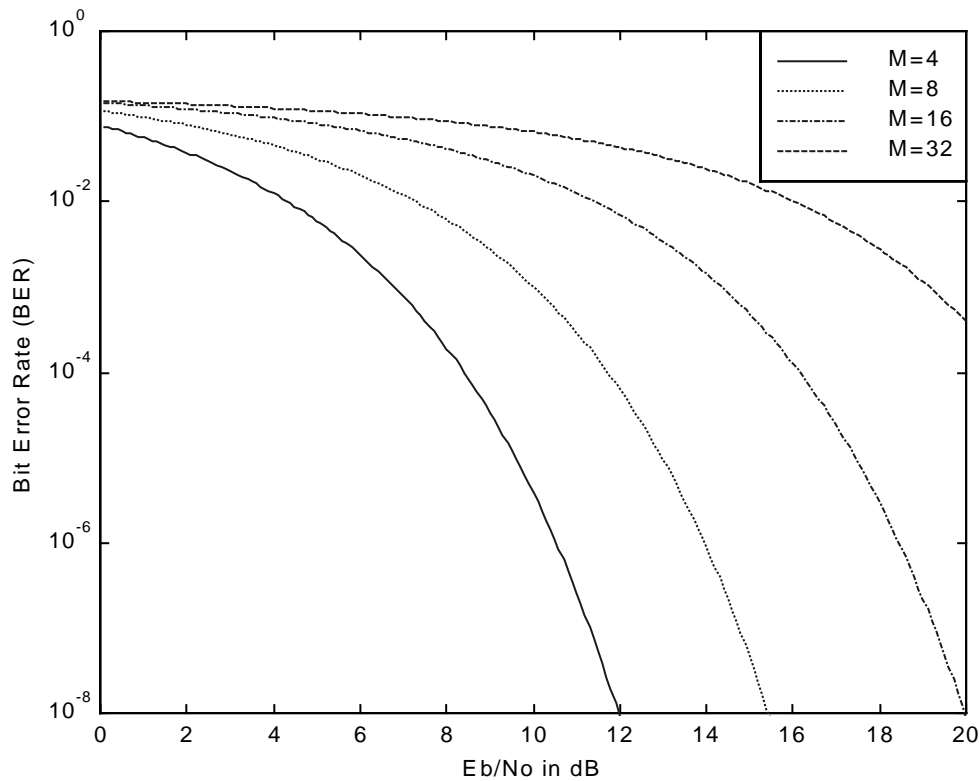


Figure 1.7 Plots BER vs. SNR of M-PSK in AWGN channel

1.5 Chapter Summary

The purpose of this chapter was to introduce the concept of signal spaces and signal vectors and how they are used to design receivers and determine performance of different modulation types. M-ary Phase Shift Keying was defined and the Euclidean distance determined between points in these signal constellations.

Using the concept of signal spaces we derived the optimal receiver for digital modulation over an AWGN channel. We used this receiver structure along with the Euclidean distances to determine the bit error rate performance of different modulation types.

The Union Bound was derived and the expression for the BER of arbitrary modulation types was shown. The Union Bound was then used to calculate the BER of M-ary PSK.

Chapter 2 The Wireless Channel and Diversity

Communications in a wireless environment poses many more problems for the system designer than that posed by a wired system, where the main concern is to compensate for degradation due to additive white Gaussian noise. In the wireless environment we must account for attenuation of the signal over distance, shadowing caused by obstructions between the transmitter and receiver, and fading due to the constructive and destructive interference of multiple reflected paths. Figure 2.1 depicts the effects of the wireless channel on propagating signals in its environment.

There is a large loss in received power that is proportional to the square of the distance between the transmitter and receiver. This loss is normally known as free space loss.

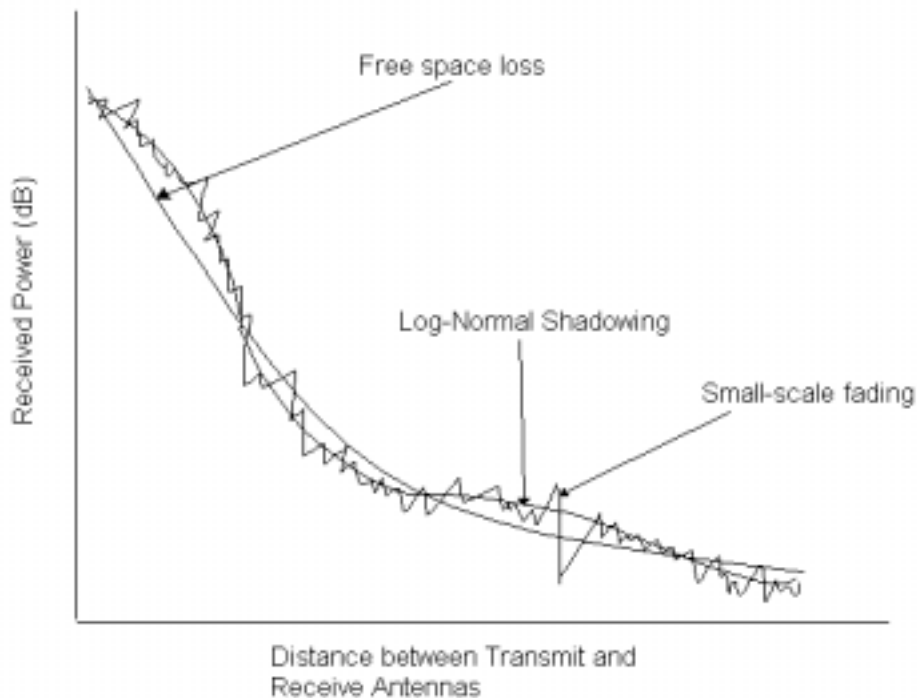


Figure 2.1 Effects on received signal power in a wireless or fading channel

There is also some gradual variation about the free space loss due to variable numbers of objects in the signal path at different locations and times. This large-scale variation will be modeled as Log-Normal shadowing.

The final component of signal degradation is the rapid fluctuation in signal power primarily due to the relative motion between the receiver and transmitter. This is described as small-scale fading and can be modeled in several ways. Each of these three effects will be discussed in this chapter in order to understand and model the wireless channel.

2.1 Large-Scale Path Loss

At a given separation between a transmitter and receiver there is a mean loss in signal strength that is due to the free space path loss. In addition to this there is some additional loss suffered by the signal if it encounters any obstructions before reaching the receiver. The physical phenomena that primarily degrade the signal are reflection, diffraction, and scattering. Reflections occur when the signal, which is an electromagnetic wave, is traveling in a given medium and impinges on another medium with different electrical and/or magnetic properties. When this occurs some portion of the signal will be transmitted and some will be reflected. Diffraction is the principle that allows a signal to reach a receiver that is hidden from the transmitter by some obstruction such as a building. The strength of the signal will decrease rapidly when this situation occurs but some portion of the signal can often still be received at a level that can be processed by the receiver. When the size of an object between a transmitter and receiver is smaller than a wavelength, or if it has a rough surface, the signal that strikes this surface will be scattered.

2.1.1 Mean or Exponential Path Loss Model

To model the mean signal strength at a given distance between a transmitter and receiver there are two options. Measurements can be taken at the prescribed distance in various places to build up an experimental model to describe a particular system, or a free space model can be assumed to be an accurate representation of the mean.

The free space model assumes that there are no obstructions in the signal path and that the signal travels in a single continuous medium. Under these conditions, electromagnetic theory can be used to derive an equation for the received signal power at a given distance from the transmitter. This equation is known as the Friis transmission formula [Che83].

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \quad (2.1)$$

Where $P_r(d)$ is the received power, d is the distance in meters, P_t is the transmitted power, G_t and G_r are the transmit and receive antenna gains, λ is the wavelength in meters, and L is a factor that is related to losses in the receive system, such as transmission line attenuation, but not due to the propagation effects.

The second option is to use experimental data to model the mean path loss in a given system. The path loss of a system is the loss in signal strength from the transmitter to the receiver over a certain distance. This is represented in the equation

$$PL(d) = \frac{P_t}{P_r(d)} \quad (2.2)$$

in decibels this is

$$PL(d)[\text{dB}] = 10 \log_{10} \left(\frac{P_t}{P_r(d)} \right)$$

The Friis free space equation can be used to help calculate the path loss but it is important to note that this equation is only valid in the far field of the antenna. It is convenient to measure the loss at a reference distance near to the transmitting antenna but still in the far field. This distance is referred to as d_0 in the path loss equations. In order to calculate path loss at an arbitrary distance, a ratio between the actual distance and the reference distance is used, as shown in the following

$$P_r(d) = P_r(d_0) \left(\frac{d_0}{d} \right)^n \quad (2.3)$$

The equations for path loss in decibels at a distance d , can now be shown to be

$$\overline{PL}(d) = \overline{PL}(d_0) + 10n \log_{10} \left(\frac{d}{d_0} \right)$$

where

$$\overline{PL}(d_0) = 10 \log_{10} \left(\left(\frac{4\pi d_0}{\lambda} \right)^2 \frac{L}{G_t G_r} \right) \quad (2.4)$$

The path loss exponent, n , in the above equation is dependent on the type of environment in which the signal is propagating. If we assume that the medium is free space then the value of n is two. For practical purposes the value of n must be found for each specific environment, but has been determined experimentally to generally conform to the relation

$$2 \leq n \leq 4$$

Table 2.1 below shows some measured values of n in different environments taken from [Rap96].

Table 2.1 Path Loss Exponents for Different Environments

Environment	Path Loss Exponent, n
Free Space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban area cellular radio	3 to 5
In building line-of-sight	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factories	2 to 3

2.1.2 Log-normal Shadowing

It is important to note that equation 2.4 gives the average path loss of a signal at a distance d between the transmitter to the receiver. The problem with this equation is that it fails to take in to account the differences in terrain that can occur in different systems that operate over the same distance. This leads to differences in measured signal strength loss in various systems even when the separation between the transmitter and receiver are the same. In order to account for this variability, experimental results have shown that accurate path loss calculations can be made if equation 2.4 is modified to include a random variable X_s .

According to [Sk197], X_s can be modeled as a Gaussian random variable (when measured in decibels) with a mean of zero and a standard deviation of σ (also measured in decibels). This yields the log-normal shadowing equation for path loss calculation

$$PL(d)[\text{dB}] = \overline{PL}(d_0) + 10n \log_{10} \left(\frac{d}{d_0} \right) + X_s$$

The values for X_s are found through measurements made at the site of interest and can have standard deviations that are as high as ten decibels.

2.2 Small-Scale Path Loss

While large-scale path loss describes the loss experienced by a signal as it travels over long distances, there are still some very rapid fluctuations that occur over very short distances or during short time intervals. Small-scale fading is the name given to these rapid fluctuations in received signal power. The four main factors that influence small-scale fading are multipath propagation, relative motion between the receiver and transmitter, relative

motion of objects between the receiver and the transmitter, and the relationship between the signal bandwidth and the bandwidth of the channel [Rap96].

Multipath refers to the fact that a signal sent out from a transmitter will encounter many objects that will reflect and/or scatter the signal in various directions. This results in many copies of the original signal reaching the receiver. The different copies will arrive at the receiver at different times, with different signal strengths and phases. When the received signal is composed of many reflected signals and one line-of-sight signal, the envelope of the signal due to fading has a Rician probability density function. When the line-of-sight, or specular, component is not there then the pdf of the received envelope is Rayleigh distributed.

In this thesis small-scale fading will always be assumed to be Rayleigh distributed. This is because the Rayleigh pdf is more mathematically tractable than the Rician pdf and also because Rayleigh fading represents the worst case fading for the purposes of system design. The distribution of the envelope of the Rayleigh faded signal is expressed as

$$p(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left\{-\frac{r^2}{2\sigma^2}\right\} & \text{for } r \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where r is the amplitude of the envelope of the received signal, and $2\sigma^2$ is the pre-detection mean power of the multipath signal. The pdf of several Rayleigh random variables with various values of σ is shown in figure 2.2.

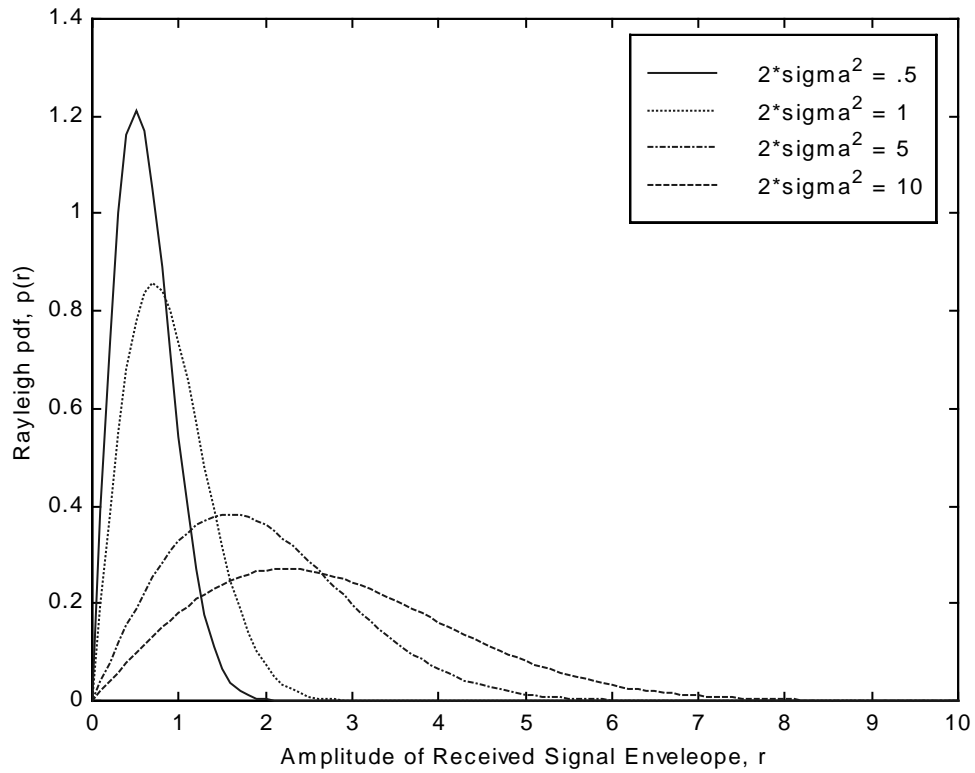


Figure 2.2 Distribution of a Rayleigh random variable

2.2.1 Impulse Response Model

The wireless channel can be modeled as a filter with a time varying impulse response. This impulse response contains all the necessary information needed to account for the small-scale propagation effects. This model can be broken down into two distinct channel

characteristics, the time-varying nature and the time-dispersive nature of the channel. The impulse response of a wireless channel can be shown to be

$$h(t, \tau) = \sum_{i=0}^{N-1} a_i(t) \exp\{j(2\pi f_c \tau_i(t) + \phi_i(t))\} \delta(\tau - \tau_i(t))$$

if we express

$$\theta_i(t) = (2\pi f_c \tau_i(t) + \phi_i(t))$$

and

$$\alpha_i(t) = a_i(t) \exp\{j\theta_i(t)\}$$

then

$$h(t, \tau) = \sum_{i=0}^{N-1} \alpha_i(t) \delta(\tau - \tau_i(t))$$

In the above equation $a_i(t)$ represents the variation in the envelope of the signal and is a Rayleigh random variable, $\theta_i(t)$ is a uniform variable representing phase changes, and $\alpha_i(t)$ is a complex Gaussian random variable.

2.2.2 Time Dispersive Nature of the Wireless Channel

In order to examine the time-dispersive nature of the wireless channel we first assume that the impulse response is time-invariant. The impulse response is no longer a function of time and can be represented as

$$h(t, \tau) = h(\tau)$$

$$h(\tau) = \sum_{i=0}^{N-1} \alpha_i \delta(\tau - \tau_i)$$

Since it is fairly difficult to measure the impulse response $h(\tau)$, a more commonly used measure is the power delay profile, or $|h(\tau)|^2$.

$$|h(\tau)|^2 = \sum_{i=0}^{N-1} a_i^2 \delta(\tau - \tau_i)$$

The channel model is normalized so that the initial delay, τ_0 , is equal to zero. Any delay that occurs after the arrival of the first component is referred to as excess delay. The following parameters are used to describe the behavior of the time-dispersive channel: Maximum Excess Delay, Mean Excess Delay, and RMS Delay Spread.

The Maximum Excess Delay corresponds to the delay associated with the last arriving signal component, or

$$\tau_{N-1} - \tau_0$$

Mean Excess Delay is equivalent to the normalized first moment of the power delay profile. In equation form this is

$$\bar{\tau} = \frac{\int_{-\infty}^{\infty} \tau |h(\tau)|^2 d\tau}{\int_{-\infty}^{\infty} |h(\tau)|^2 d\tau}$$

or

$$\bar{\tau} = \frac{\sum_{i=1}^{N-1} a_i^2 \tau_i}{\sum_{i=0}^{N-1} a_i^2}$$

To calculate the rms delay spread, first take the normalized second moment of the power delay profile

$$\overline{\tau^2} = \frac{\int_{-\infty}^{\infty} \tau^2 |h(\tau)|^2 d\tau}{\int_{-\infty}^{\infty} |h(\tau)|^2 d\tau}$$

or

$$\overline{\tau^2} = \frac{\sum_{i=1}^{N-1} a_i^2 \tau_i^2}{\sum_{i=0}^{N-1} a_i^2}$$

Then use this value, $\overline{\tau^2}$, along with the normalized first moment, $\overline{\tau}$, to find the rms delay spread

$$\sigma_{\tau} = \sqrt{\overline{\tau^2} - (\overline{\tau})^2}$$

All of the parameters dealt with thus far are in the time domain and determine whether or not the wireless channel is time-dispersive. In the frequency domain the channel is said to exhibit frequency-non-selective, also called flat, or frequency-selective fading. The term coherence bandwidth refers to the effective bandwidth of the channel that affects a signal in a similar fashion. In other words the fading is relatively constant over a certain band of frequencies. If the entire frequency content, or bandwidth, of the transmitted signal falls within the coherence bandwidth of the channel then all the frequency components of the signal are effected in the same fashion. This is referred to as flat fading. On the other hand, if the bandwidth of the transmitted signal is larger than the coherence bandwidth of the channel then different frequency components will experience various levels of fading. A channel that acts in this way is known as a frequency-selective channel.

The coherence bandwidth can be calculated from the rms delay spread. Depending on the assumption of how correlated the fading should be, fifty percent or ninety percent

correlation, there are two equations for coherence bandwidth [Rap96]. The two equations are shown below.

$$50\% \text{ correlated: } B_c \approx \frac{1}{5\sigma_\tau}$$

$$90\% \text{ correlated: } B_c \approx \frac{1}{50\sigma_\tau}$$

If B_s is the bandwidth of the transmitted signal and B_c is the coherence bandwidth of the channel, then figure 2.3 shows the different cases of fading.

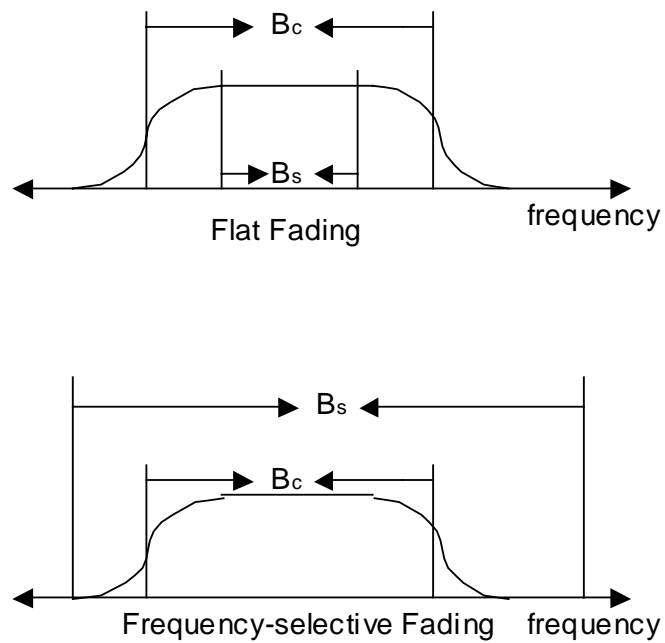


Figure 2.3 Representation of Flat or Frequency-Selective Fading

2.2.3 Time Varying Nature of the Wireless Channel

In the previous section the channel was assumed to be time-invariant in order to examine the time-dispersive nature of the channel. In this section we will assume that all the signal components arrive at the same moment rather than as a series multipath components. The impulse response is no longer a sum of components arriving with variable delays, rather it is a single function of time and can be represented as

$$h(t, \tau) = \alpha(t)\delta(\tau)$$

This is valid if we assume that

$$\sigma_{\tau} = T_s$$

where T_s is the transmitted symbol period and σ_{τ} is the RMS Delay Spread.

In order to understand the time varying parameters of the channel it is necessary to look at the Doppler shift and how it is calculated. Assume that a base station, located at some point z, is broadcasting a signal to a mobile station that is moving from point x to point y with a constant velocity v. This situation is depicted in figure 2.4 below.

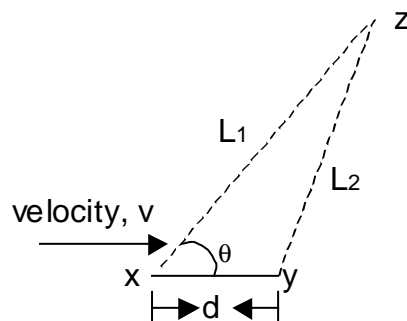


Figure 2.4 Physical description of relative motion between a transmitter and receiver

The first step is to calculate the path length difference between L_1 and L_2 . This difference is labeled Δl . The Doppler frequency, f_d , will now be found using the information in figure 2.4 above.

$$\Delta l = |L_1 - L_2|$$

$$\Delta l = d \cos(\theta)$$

where distance $d = v\Delta t$

$$\Delta l = v\Delta t \cos(\theta)$$

$$\Delta\phi = \frac{2\pi\Delta l}{\lambda}$$

$$\Delta\phi = \frac{2\pi v\Delta t \cos(\theta)}{\lambda}$$

Now calculate the angular frequency measured in radians per second using

$$\omega_d = \frac{\Delta\phi}{\Delta t}$$

$$\omega_d = \frac{2\pi v \cos(\theta)}{\lambda}$$

Then, the Doppler frequency in hertz is

$$f_d = \frac{\omega_d}{2\pi}$$

$$f_d = \frac{v \cos(\theta)}{\lambda}$$

When the mobile channel exhibits a time varying nature the result is a widening of the spectral content of a transmitted signal. The amount of widening, or spreading, is dependent on the Doppler frequency. The Doppler spread, B_D , is the parameter used to describe this broadening of frequency content and is described by

$$B_D = 2f_d$$

The maximum Doppler frequency occurs when $|\cos(\theta)|$ is equal to one and is shown to be

$$f_{d \max} = \pm \frac{v}{\lambda}$$

Coherence time, T_c , is another parameters used to characterize the time varying nature of the channel. It is essentially a statistical measure of the time over which the impulse response of the channel does not vary. In essence, this means that if two signals are received within a time that is less than the coherence time then those signal amplitudes will be highly correlated. The coherence time is inversely proportional to the Doppler spread leading to the relationship

$$T_c \approx \frac{1}{f_{d \max}} \quad (2.5)$$

If the coherence time is defined to be the interval over which the correlation of two signals in time is greater than 0.5 then

$$T_c \approx \frac{9}{16\pi f_{d \max}} \quad (2.6)$$

According to [Rap96], a commonly used method for determining coherence time is to take the geometric mean of equations 2.5 and 2.6 to arrive at

$$T_c = \sqrt{\frac{9}{16\pi f_{d\max}^2}}$$

$$T_c = \frac{0.423}{f_{d\max}}$$

2.3 Multiple input Multiple output (MIMO) Channels

In the previous chapter and thus far in this chapter we have only considered channels with a single input and a single output. In order to analyze or discuss diversity techniques that involve the use of multiple antennas it is necessary to first understand the multiple input-multiple output (MIMO) channel model. Figure 2.5 represents the basic layout of a MIMO channel model.

The major difference between the single input-single output, or SISO channel, and the MIMO channel is the use of vector or matrix notation. In the MIMO channel we no longer have a single input, rather, we have a vector of N input signals. At the output there is a vector of M output signals. To take into account the fading coefficients between transmit and receive antenna pairs it is necessary to use a $N \times M$ matrix, where $\alpha_{n,m}$ is the complex fading gain between transmit antenna n and receive antenna m . The last component to consider is the noise process, which in this case is also a vector. The noise is represented as a vector of M components, each of which is a sample of AWGN.

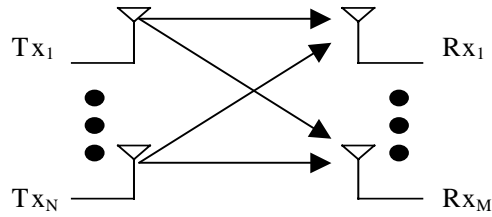


Figure 2.5 Generic model for a MIMO channel

The following shows the design of the input vector \mathbf{s} , output vector \mathbf{r} , noise vector \mathbf{n} , and the fading coefficient matrix \mathbf{H} .

$$\mathbf{s} = [s_1 \quad s_2 \quad \dots \quad s_N]$$

$$\mathbf{H} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1M} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2M} \\ \vdots & \vdots & & \vdots \\ \alpha_{N1} & \alpha_{N2} & \dots & \alpha_{NM} \end{bmatrix}$$

$$\mathbf{n} = [n_1 \quad n_2 \quad \dots \quad n_M]$$

$$\mathbf{r} = [r_1 \quad r_2 \quad \dots \quad r_M]$$

The output is related to the input and the channel parameters by the following equation

$$\mathbf{r} = \mathbf{sH} + \mathbf{n}$$

2.4 Diversity

Diversity is a class of techniques that seek to improve performance in fading channels by providing the receiver with multiple copies of the transmitted signal, including spatial, temporal, and frequency diversity. There are several different methods for generating multiple signal copies. The same signal can be sent to multiple receive antennas, in effect creating multiple channels. This first technique is a form of Spatial diversity. Time, or temporal, diversity is a process by which a signal is repeated any number of times over multiple time slots. Forward error correction (FEC) can be considered a type of temporal diversity. Frequency diversity is achieved by sending a signal simultaneously over multiple frequencies.

In this thesis we are primarily concerned with the idea of spatial diversity, specifically with antenna diversity, and the combination of temporal and spatial diversity. In the wireless channel small-scale fading has been shown to produce highly varying levels of signal degradation over relatively small distances or over small time periods. This can be exploited in a relatively simple and straightforward manner. If two antennas are set a small distance apart it is highly probable that if one of the antennas is in a deep fade the other antenna will have a much stronger signal. If the outputs of these two antennas can be somehow combined in one receiver performance can be greatly improved.

The way in which the multiple received signals are combined determines the performance of the different types of antenna diversity implementations. If the strongest signal is chosen and the weaker dropped completely, this is known as selection combining. If all the received signals are added together to produce a composite signal the technique is called Equal Gain combining. Equal Gain combining performs better than selection combining but the performance could be increased even more if the different receive signals were weighted in such a way that the total received signal to noise ratio was maximized. Maximal Ratio combining seeks to weight the individual received signals in order to generate a signal with the optimal SNR.

The performance of any L-branch diversity technique in a fading channel can be derived using a rather simple process. The first step in this process is to establish the instantaneous SNR on an arbitrary branch ℓ . For Rayleigh fading channels, this instantaneous SNR, known as γ_ℓ , is a random variable with an exponential probability density function (pdf).

$$f_{\gamma_\ell}(\gamma) = \frac{1}{\Gamma_\ell} \exp\left\{\frac{-\gamma}{\Gamma_\ell}\right\} \quad \text{for } \gamma \geq 0$$

where

$$\Gamma_\ell = E\{\gamma_\ell\}$$

From the above pdf we can establish the cumulative distribution function, or CDF, by integrating the pdf.

$$F_{\gamma_\ell}(z) = \int_{-\infty}^z f_{\gamma_\ell}(\gamma) d\gamma$$

$$F_{\gamma_\ell}(z) = 1 - \exp\left\{\frac{-z}{\Gamma_\ell}\right\} \quad \text{for } z \geq 0$$

The above CDF represents the probability that the received SNR on branch ℓ is below a certain threshold defined as z . In equation form this is

$$F_{\gamma_\ell}(z) = \Pr[\gamma_\ell \leq z]$$

The next step is to determine the probability that the received SNR of the signals combined from all L branches falls below the previously established threshold z . The received SNR of all combined branches is referred to as γ_c . In this thesis we are primarily concerned with the performance of Maximal Ratio Combining, or MRC. Therefore, for the calculation of all subsequent parameters in this example we will assume that MRC is the

diversity technique under consideration. The goal of MRC is to maximize the combined received SNR, therefore in this case

$$\gamma_c = \sum_{\ell=1}^L \gamma_\ell$$

The next step is to find the pdf and CDF of γ_c . It is important to note that the pdf's of each of the L branches are assumed to be independent. This fact allows us to calculate the pdf of γ_c as merely the convolution of the L individual pdf's.

$$f_{\gamma_c}(\gamma) = f_{\gamma_1}(\gamma) * f_{\gamma_2}(\gamma) * \dots * f_{\gamma_L}(\gamma)$$

From the theory of probability it has been shown that the sum of L independent and identically distributed exponential random variables is an L-erlang random variable [Gar94]. Therefore if the L branches have an equal average SNR, $\Gamma_\ell = \Gamma$ for all ℓ , then the pdf of γ_c is

$$f_{\gamma_c}(\gamma) = \frac{\gamma^{L-1} \exp\left\{\frac{-\gamma}{\Gamma}\right\}}{\Gamma^L (L-1)!} \quad \text{for } \gamma \geq 0$$

$$F_{\gamma_c}(z) = \int_{-\infty}^z f_{\gamma_c}(\gamma) d\gamma$$

$$F_{\gamma_c}(z) = 1 - \exp\left\{\frac{-z}{\Gamma}\right\} \sum_{k=1}^L \frac{\left(\frac{z}{\Gamma}\right)^{k-1}}{(k-1)!} \quad \text{for } z \geq 0$$

In order to get to a BER for the diversity system it is first necessary to calculate the conditional BER dependent on γ and the type of modulation used.

For this example we will use the BER of BPSK or QPSK, which is

$$P_b(\gamma) = Q(\sqrt{2\gamma})$$

This conditional BER can be averaged over all possible values of γ to arrive at the average BER for the combined signals in the diversity system, thus providing the BER of the diversity system in a fading channel. In our case we are using BPSK or QPSK and MRC, therefore the average BER is

$$\bar{P}_b = E\{P_b(\gamma)\}$$

$$\bar{P}_b = \int_{-\infty}^{\infty} Q(\sqrt{2\gamma}) \frac{\gamma^{L-1} \exp\left\{-\frac{\gamma}{\Gamma}\right\}}{\Gamma^L (L-1)!} d\gamma$$

$$\bar{P}_b = \left(\frac{1-\mu}{2}\right)^L \sum_{k=0}^{L-1} \binom{L-1+k}{k} \left(\frac{1+\mu}{2}\right)^k$$

where

$$\mu = \sqrt{\frac{\Gamma}{1+\Gamma}}$$

In the above example we made the assumption that the average SNR in each branch was equal for all branches. If this is not the case then the BER is found to be the following in [Sim00],

$$\bar{P}_b = \frac{1}{\pi} \int_0^{\pi} \prod_{\ell=1}^L \left(1 + \frac{\Gamma_{\ell}}{\sin^2 \theta}\right)^{-1} d\theta$$

The performance of BPSK modulation in a Rayleigh fading channel can be seen in figure 2.6 below. Notice that with L=2 branch Maximal Ratio Combining the performance

improves by almost 11dB. Using more than two antennas improves performance even more, although the improvement obeys a “law of diminishing returns”, meaning the improvement is less with each additional antenna.

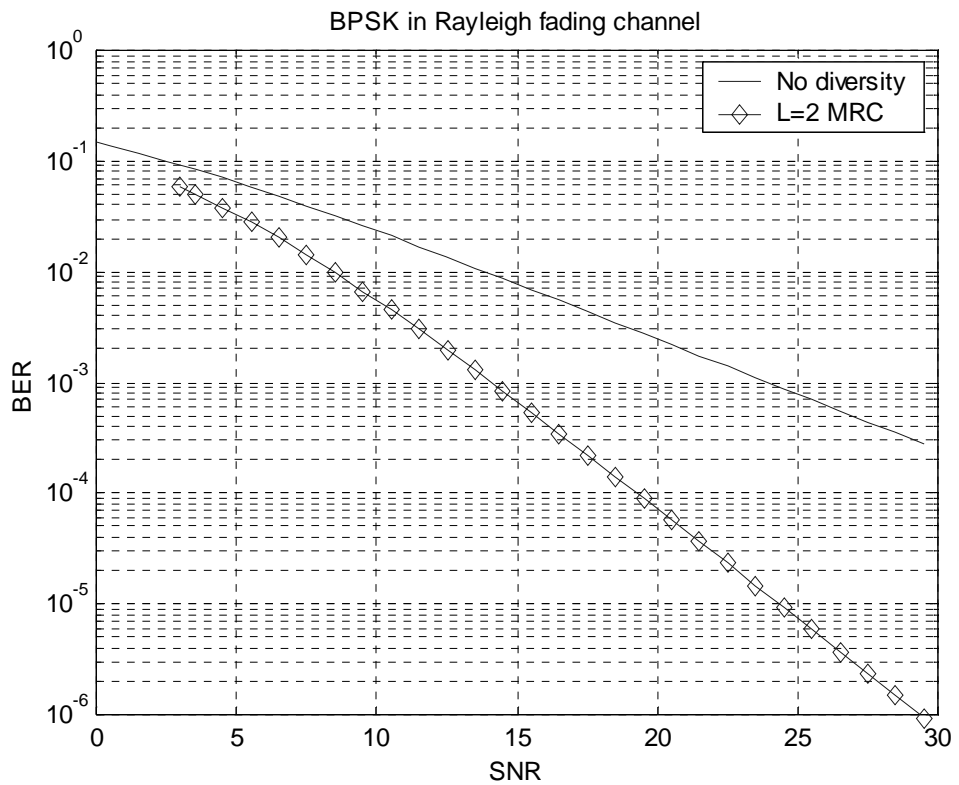


Figure 2.6 Performance of BPSK in Rayleigh fading

2.5 Chapter Summary

In this chapter we described the factors that affect signal propagation in a wireless environment. The concepts of large-scale path loss, or shadowing, and small-scale fading were introduced and used to build a statistical model of the wireless channel that can be used to calculate the performance of different modulation types in fading channels. The type of fading that was of primary concern in this chapter was Rayleigh fading, which could be described as the envelope of a complex Gaussian random variable. The performance of certain modulation types in Rayleigh fading was shown to be much worse than in AWGN channels.

The topic of diversity was also covered and the performance of Maximal Ratio Combining (MRC) was derived. Using two antenna diversity and MRC the performance of BPSK in Rayleigh fading was shown. The improvement using two branch MRC over the case of BPSK with no diversity was shown to be almost 12dB at a bit error rate of 10^{-4} .

CHAPTER 3: Space-Time Block Codes

One of the goals of modern wireless communication systems is to increase the data rates of users without excessive bandwidth expansion. Spectrum is sometimes limited while users continue to demand more data intensive applications, like wireless Internet access, videoconferencing, and streaming multimedia. In order to compensate for the extreme signal degradation that can occur in wireless channels some type of diversity is essential to the functioning of a wireless communication system. It is not uncommon for a wireless system to employ both diversity techniques and channel coding for error detection and/or correction. An effective diversity technique is the use of multiple antennas to provide an improvement in reception without drastically reducing the data rate or increasing the bandwidth. The initial implementations of antenna diversity consisted of using two or more receive antennas. This is a practical technique to use at the base station in a cellular system but not nearly as practical for use in the mobile handsets. If each mobile handset were equipped with multiple receive antennas then the size and cost of each unit would increase beyond what is acceptable to the consumer. If multiple receive antennas cannot be used in the mobiles then diversity gain can only be achieved at the base station.

Research was then being conducted into the feasibility of achieving a suitable diversity gain by using multiple transmitting antennas. A simple scheme for achieving diversity gain without bandwidth expansion using two transmit antennas and any number of receive antennas was devised by S. Alamouti in [Ala98]. Another promising method to provide high data rates, good error performance, and minimal bandwidth expansion is Space-Time Coding. Space-Time coding was introduced in [Tar98] by Tarokh *et al.* The original space-time codes were trellis codes. The codes developed provided maximum diversity gain, dependent on the number of antennas used, and good coding gain, depending on the number of states in the trellis. The complexity of the trellis-based codes is fairly high and increases exponentially with the number of states in the trellis. The performance of trellis-based space-time codes in the presence of channel estimation errors and fading was examined in [Tar99a].

It would be desirable to come up with a method to achieve maximum diversity gain but with minimal decoding complexity. Space-time block codes, as introduced in [Tar99b], are one such method. Space-time block codes are an extension of the simple scheme developed in [Ala98] to use an arbitrary number of transmit and receive antennas. Space-time block codes utilize a block coding method, as opposed to a trellis-based method, to eliminate much of the processing needed at the receiver. Since the block coding requires only linear processing at the receiver, the decoding can be done efficiently and quickly. Space-time block codes can be constructed for any type of signal constellation and provide full diversity gain at half the maximum possible transmission rate allowed by the theory of space-time coding. For real signal constellations, such as Pulse Amplitude Modulation(PAM), space-time block codes provide the maximum possible transmission rate allowed by the theory of space-time coding.

3.1 General Theory

The transmission model for the space-time block code system is taken from [Tar99b] and the rest of this section will define that model. In a space-time block coding system there are n transmit antennas and m receive antennas. At a given time slot t , n signals s_t^i $i = 1, 2, \dots, n$, are sent simultaneously from the n transmit antennas. A block diagram of the transmission side of the system can be seen in figure 3.1.

The signal received at antenna j during time t is

$$r_t^j = \sum_{i=1}^n \alpha_{i,j} s_t^i + n_t^j. \quad (3.1)$$

Where $\alpha_{i,j}$ is the path gain between transmit antenna i and receive antenna j , and n_t^j is the noise at receive antenna j . The channel is assumed to undergo flat-fading and the fading is independent between different transmit antennas. The path gains are considered to be independent samples of a complex Gaussian distribution with a variance of 0.5 per real dimension. The noise at the receiver is independent from the path gains and in the form of

additive Gaussian noise with a mean of zero and a variance equal to $n/(2 \cdot \text{SNR})$, where n is the number of transmit antennas and SNR is a ratio, not in dB.

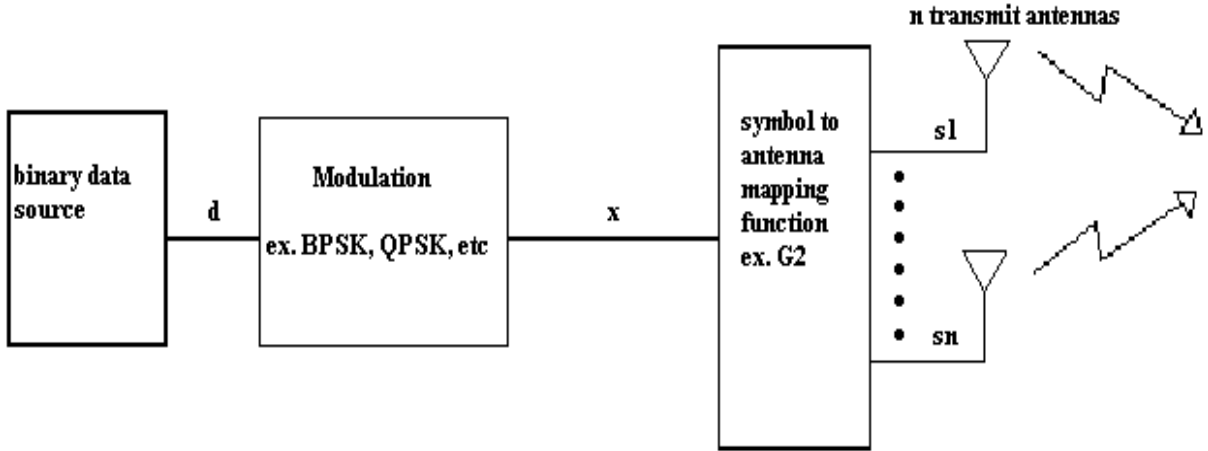


Figure 3.1 Transmission side of Space-Time Block Code system.

The average energy is normalized to be unity for each symbol leaving each of the n transmitting antennas. This gives the energy of the received signal as n and SNR is measured at the receiver. The decoding for this system is rather simple and consists of minimizing the following metric,

$$\sum_{t=1}^l \sum_{j=1}^m \left| r_t^j - \sum_{i=1}^n \alpha_{i,j} s_t^i \right|^2 \quad (3.2)$$

over all possible combinations of transmitted symbols. A block diagram of the receiving side of this system can be seen in figure 3.2.

The encoding process is done based on the data rate the system requires. There is some signal constellation, used for modulation, which maps binary data to real or complex symbols. If there are 2^b symbols in the signal constellation, then $k \times b$ bits will be brought

in to the modulator at one time slot. These $k \times b$ bits will be used to select k symbols that will be sent out over n transmit antennas simultaneously.

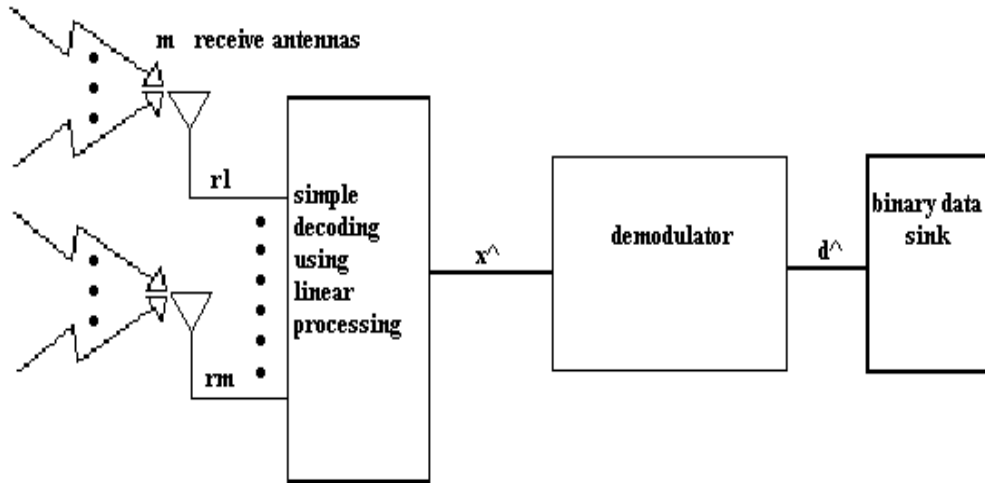


Figure 3.2 Receiving side of Space-Time Block Code system.

The rate of transmission is

$$R = \frac{k}{p}, \quad (3.3)$$

where k is the number of symbols that will have to be decoded and p is the number of time slots it takes to transmit all the symbols. The notation denoting the process by which modulation symbols are mapped to different antennas is a simple $p \times n$ matrix. An example is the encoder matrix

$$G_2 = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix}.$$

The i^{th} row determines the symbols transmitted in time slot i , and the j^{th} column determines the symbols transmitted from antenna j over all time slots. Several other encoder matrices were developed in [Tar99b] for various numbers of transmit antennas. The rate one half matrix used for three antennas is

$$G_3 = \begin{pmatrix} s_1 & s_2 & s_3 \\ -s_2 & s_1 & -s_4 \\ -s_3 & s_4 & s_1 \\ -s_4 & -s_3 & s_2 \\ s_1^* & s_2^* & s_3^* \\ -s_2^* & s_1^* & -s_4^* \\ -s_3^* & s_4^* & s_1^* \\ -s_4^* & -s_3^* & s_2^* \end{pmatrix}.$$

There are several others in [Tar99b] for three and four transmit antennas. The decoding of the Space-time block code is performed by minimizing the metric shown in equation (3.1) above. However, this can be broken down into a simpler form where the metrics can be separated into several equations, each dependent only on a single transmitted symbol. For the specific case of the code defined by G_2 , the metric can be decomposed into two simpler equations. Each one needs only to be evaluated over the possible values that a single symbol can take on, rather than over combinations of symbols. The two equations for this case can be derived as follows:

There are two time slots over which signals will be received at each receive antenna generating two received signals r_1^j and r_2^j . These signals can be shown to be

$$\begin{aligned} r_1^j &= \alpha_{1,j}s_1 + \alpha_{2,j}s_2 + n_1^j \\ r_2^j &= -\alpha_{1,j}s_2^* + \alpha_{2,j}s_1^* + n_2^j \end{aligned} \quad (3.4)$$

This can be shown in matrix form to be

$$\mathbf{r}^j = \begin{pmatrix} r_1^j \\ (r_2^j)^* \end{pmatrix} = \begin{pmatrix} \alpha_{1,j} & \alpha_{2,j} \\ \alpha_{2,j}^* & -\alpha_{1,j}^* \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} n_1^j \\ (n_2^j)^* \end{pmatrix} \quad (3.5)$$

where, for convenience, we have conjugated the equation for r_2^j so that the signals s_1 and s_2 don't need to be conjugated.

If we take the two received signals as shown in equation (3.4) and substitute into equation (3.2) then we have the new decision metric

$$\sum_{j=1}^M \left(|r_1 - \alpha_{11}s_1 - \alpha_{21}s_2|^2 + |r_2 + \alpha_{11}s_2^* - \alpha_{21}s_1^*|^2 \right) \quad (3.6)$$

Next we can use the identity

$$|\xi|^2 = \xi \times \xi^* \quad (3.7)$$

to further expand the previous metric into the following two metrics which can be evaluated separately in order to simplify the decoding structure.

$$\left[\left| \sum_{j=1}^m \left(r_1^j \alpha_{1,j}^* + (r_2^j)^* \alpha_{2,j} \right) \right| - s_1 \right]^2 + \left(-1 + \sum_{j=1}^m \sum_{i=1}^2 |\alpha_{i,j}|^2 \right) |s_1|^2 \quad (3.8)$$

$$\left[\left| \sum_{j=1}^m \left(r_1^j \alpha_{2,j}^* - (r_2^j)^* \alpha_{1,j} \right) \right| - s_2 \right]^2 + \left(-1 + \sum_{j=1}^m \sum_{i=1}^2 |\alpha_{i,j}|^2 \right) |s_2|^2 \quad (3.9)$$

The above two equations are not complex and can be readily implemented. This is the benefit of space-time block codes over the trellis-based space-time codes, they provide maximum diversity gain with little complexity at the receiver.

3.2 Performance with Perfect CSI

In this section the performance of Space-Time Block codes with perfect channel state information(CSI) will be examined. Space-Time Block codes were tested by building a simulation in MATLAB. The simulations were performed using G_2 for the encoding matrix. The modulation types used were Binary Phase Shift Keying (BPSK) and Quadrature Phase Shift Keying (QPSK). The frame size was one hundred bits. The path gains between a particular transmit and receive antenna were assumed to be constant over a frame, and independent between different antennas. What this means is that the path gains are Rayleigh random variables that are uncorrelated between different transmit-receive antenna pairs, and the value for a particular path gain does not change over the transmissions in a single frame. The assumption of constant fading over a frame, called quasi-static fading, is justified if the data rate is high and/or the channel fades relatively slowly. In order to justify the assumption that the fading gains between antennas are

uncorrelated requires that the different antennas be physically separated by approximately ten wavelengths at a base station and three wavelengths at a mobile unit [Ala98]. The path gains were generated by taking samples from two Gaussian random number generators, each with a variance of 0.5, and adding the two together as an in-phase and a quadrature component to yield a complex number. These were multiplied by the transmitted symbols and combined at each receive antenna. The noise term was generated from a Gaussian random number generator with a mean of zero and a variance of $n/(2*SNR)$.

The simulations were tested with two transmit antennas and one receive antenna, and also with two transmit antennas and two receive antennas. The results generated by the simulations are in the form of graphs of Bit Error Rate (BER) versus received signal-to-noise ratio (SNR). In order to produce results with a fairly high degree of accuracy each value of BER for a specific level of SNR was calculated only after at least forty frame errors had occurred.

The results for the first case, two TX and one RX antenna, can be seen in figures 3. Figure 3 shows the bit error rate versus SNR for uncoded BPSK, BPSK using G_2 , and QPSK using G_2 . The improvement of QPSK at 2 bit/s/Hz over the uncoded case is about 7 dB at a bit error rate of 10^{-3} . The improvement of BPSK at 1 bit/s/Hz over the uncoded case is approximately 10 dB at a bit error rate of 10^{-3} . As the SNR increases the improvement of BPSK and QPSK over the uncoded case also increases.

The results of the second case can be seen in figure 3.4. This plot shows the bit error rate of the space-time block code using BPSK, two transmit and two receive antennas. The improvement of BPSK at 1 bit/s/Hz using two receiver antennas over the case of BPSK with just one receive antenna is approximately 11 dB at a BER of 10^{-5} .

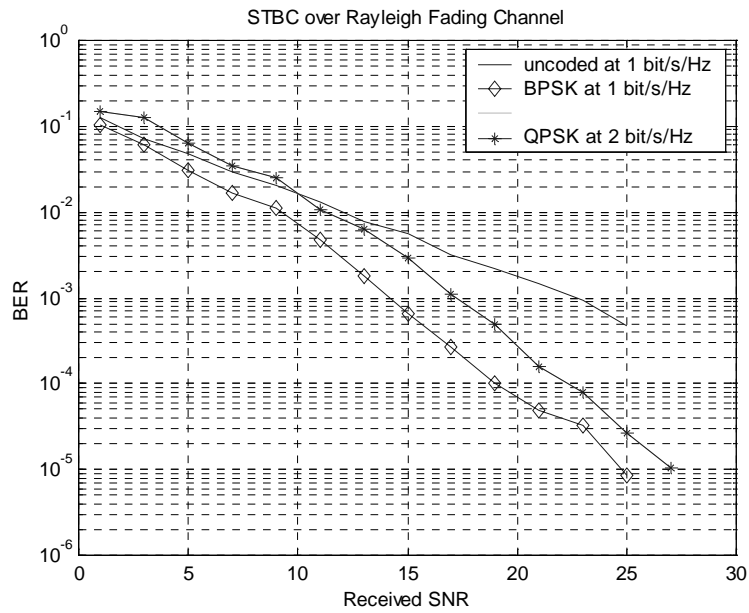


Figure 3.3 STBC over Rayleigh fading using two transmit and one receive antenna

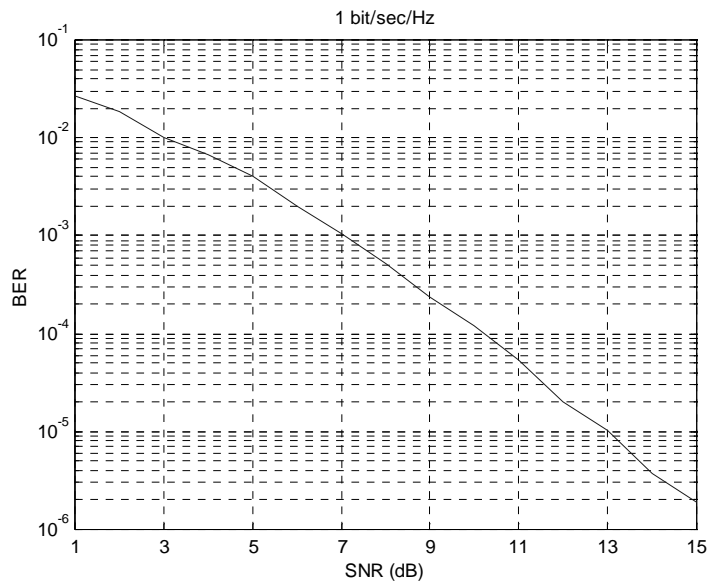


Figure 3.4 STBC over Rayleigh fading using two transmit and two receive antennas

3.3 Chapter Summary

The purpose of this chapter was to introduce the concept of space-time block codes and demonstrate the performance gains obtainable under the assumption that the receiver has access to perfect channel state information. Initially, the basic transmitter and receiver models were explained and the decoding process described. The performance of this system was then shown by simulation. The results of these simulations show that space-time block codes using BPSK and one receive antenna provide an improvement of more than 10dB over the uncoded case at a BER of 10^{-3} . The improvement of a STBC using BPSK and two receive antennas over a STBC using BPSK and one receive antenna, is nearly 11dB at a bit error rate of 10^{-5} .

Chapter 4: STBC with Channel Estimation Errors

In order to adequately demonstrate the performance of a system via simulation, it is necessary to develop an accurate model of the system as it would be physically implemented. The assumption that perfect channel state information would be available to the receiver is inappropriate when simulating a physical system, because in a real system the effects of the channel can never be known exactly. Rather, some form of estimation is performed to find an approximation to the channel. Under the assumption that perfect CSI is available to the decoder, the performance of space-time block codes has been shown in the previous chapter.

In this chapter, the performance of space-time block codes is analyzed under the constraint that the receiver must rely on imperfect estimates of the channel conditions. This is done in order to verify the performance that can be expected in an actual space-time block coded system. The sensitivity and robustness of space-time block codes to varying levels of error in the amplitude and phase of the estimates will be illustrated using a series of simulations. This new information could be used to help design and verify the performance of a channel estimation scheme based on the insertion of pilot sequences into the data stream of each antenna. The imperfect channel estimates will be created by taking the actual fading coefficients introduced into the channel and applying some degree of error into either the magnitude or the phase of those coefficients before decoding. This simulates the inability of an estimation scheme to predict the channel characteristics with perfect accuracy.

Initially the effects of errors in the estimate of the gain and the phase will be analyzed separately for two reasons. First, it would be difficult to represent the resulting changes in BER when compared against the degree of error introduced into both gain and phase simultaneously, especially when we have no clear understanding of what specific effects either type of error will have. Second, it is important to understand what effects each of the two components has on performance when its estimate is not exact. This way it is possible to determine if either gain or phase is more important to the decoding and estimating

process in terms of Bit error rate performance. Also, certain modulation formats may be more or less sensitive to gain or phase estimation errors. After the effects on the system due to the two different types of errors are known, simulations are run with the errors occurring in both the gain and the phase of the estimates simultaneously, to determine the effects the combined errors have on performance.

4.1 System Model

The system model that we will use to analyze the performance of space-time block codes with channel estimation errors can be seen in figure 1. It will consist of two transmit antennas and one receive antenna operating in a Rayleigh fading environment. Symbol mapping will be done using a QPSK or BPSK signal constellation and the generator matrix G_2 , developed in [Tar99b]. The fading coefficient between the i^{th} transmit antenna and the receive antenna is given as

$$c_i = a_i \exp\{j\theta_i\}$$

We ran simulations of the system with errors in both the amplitude and phase of the channel estimates. A channel estimate with phase error will be of the form

$$\hat{c}_i = a_i \exp\{j\theta_i + \phi_i\}$$

where ϕ_i is the error introduced into the phase. An estimate with errors in the amplitude will be of the form

$$\hat{c}_i = K_i a_i \exp\{j\theta_i\}$$

where K_i is the error introduced into the amplitude. The channel is assumed to undergo flat fading and the fading is independent between different transmit antennas. It is also assumed that the fading over a channel is constant over a frame (i.e. quasi-static). The assumption of constant fading over a frame is justified if the data rate is high and/or the channel fades relatively slowly. In order to justify the assumption that the fading gains between antennas are uncorrelated requires that the different antennas be physically separated by approximately ten wavelengths [Ala98]. The path gains are considered to be independent samples of a complex Gaussian distribution.

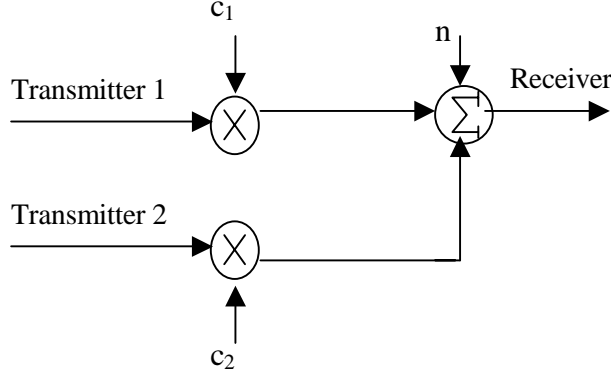


Figure 1: Transmission system consisting of two transmit antennas and one receive antenna.

The variance of the path gains is 0.5 per complex dimension. The noise at the receiver is an additive Gaussian noise produced from samples of another Gaussian random variable with a mean of zero and a variance equal to $n/(2 \cdot \text{SNR})$. Here, n is the number of transmit antennas and SNR is the signal to noise ratio at the receiver. The average energy is normalized to be one for each symbol leaving each of the n transmitting antennas. This gives the energy of the received signal as n and SNR is measured at the receiver.

The received signal at time t is

$$r_t = \sum_{i=1}^n c_i s_i^{(t)} + n_t$$

Since we are using generator matrix G_2 , there will be two sets of transmissions for each set of two input symbols. Therefore in matrix notation we can show the received signal as

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$$

$$\mathbf{r} = \mathbf{S}\mathbf{c} + \mathbf{n}$$

The decoding for this system is rather simple and consists of minimizing the following metric,

$$\sum_{t=1}^l \left| r_t - \sum_{i=1}^n c_i s_i^{(t)} \right|^2,$$

over all possible combinations of transmitted symbols.

For the simulations using channel estimates the metric is

$$\sum_{t=1}^l \left| r_t - \sum_{i=1}^n \hat{c}_i s_i^{(t)} \right|^2$$

4.2 Simulation Results

In this section we provide simulation results for the performance of space-time block codes with channel estimation errors as described in the previous sections. Figure 4.1 shows the performance of uncoded QPSK and QPSK using G_2 , under the assumption that perfect CSI is available at the receiver. All further figures are using a QPSK signal constellation, the generator matrix G_2 , and two transmit and one receive antenna. Figures 4.2, 4.3, and 4.4 show the bit error rates of our transmission scheme against the phase error in each channel for fixed levels of received SNR of 10, 20, and 25 dB, respectively. The phase errors are measured in radians, with a maximum phase error of $\pi/4$ radians.

These figures show that as the SNR is increased the system can tolerate a larger degree of error and still retain reasonable performance. However, as the degree of error in the phase approaches $\pi/4$, the system performance breaks down regardless of SNR. This is to be expected as the decision regions for QPSK are defined by boundaries that are $\pi/4$ radians between different signals. Figure 4.5 shows the bit error rate of our system when the channel estimates contain amplitude, or gain, errors as high as 1.5.

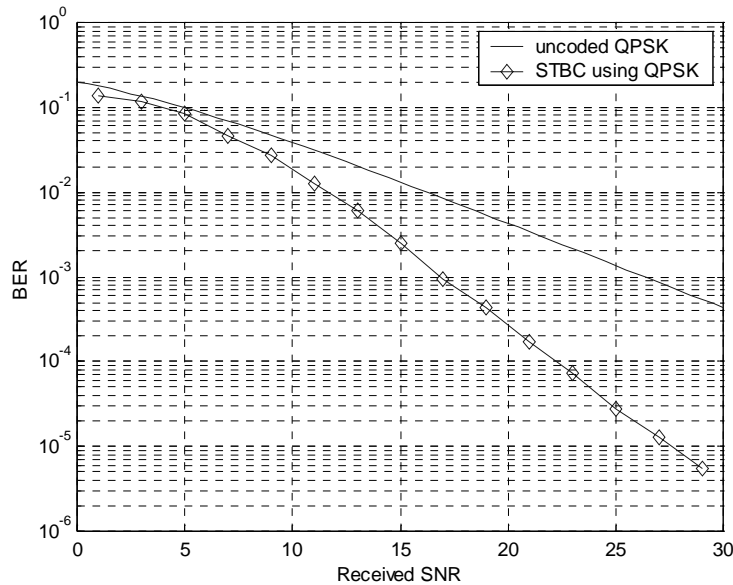


Figure 4.1. Performance of STBC in Rayleigh flat fading with QPSK modulation, perfect CSI, two transmit antennas, and one receive antenna.

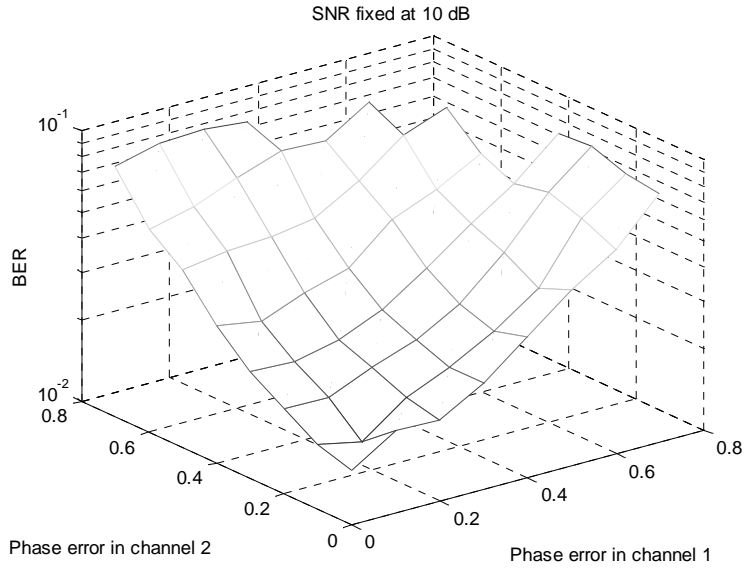


Figure 4.2. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 10 dB, and a maximum phase error of $\pi/4$ radians.

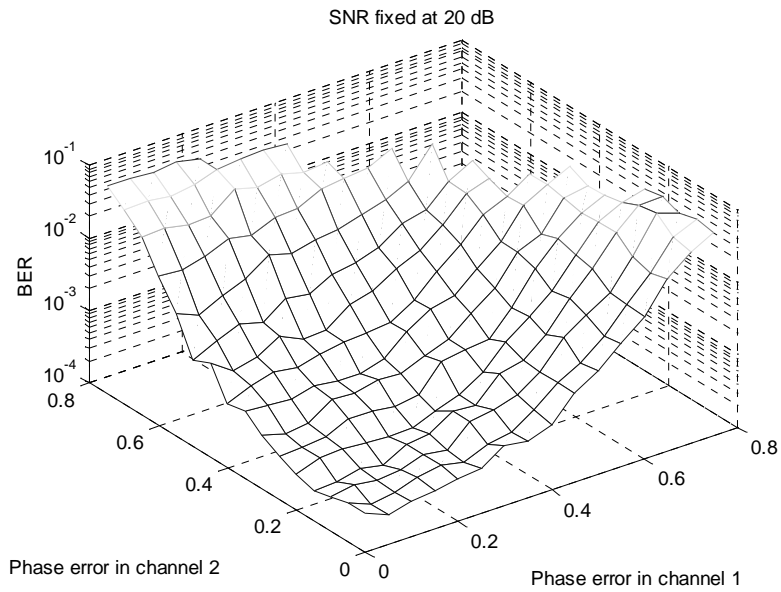


Figure 4.3. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 20 dB, and a maximum phase error of $\pi/4$ radians.

If the gain errors are equal in each channel than there is little performance degradation. It is for this reason that further results concerning gain errors will only be concerned with what we call the normalized gain error. This is simply the ratio of the gain error in channel one to the gain error in channel two, or $\frac{K_1}{K_2}$.

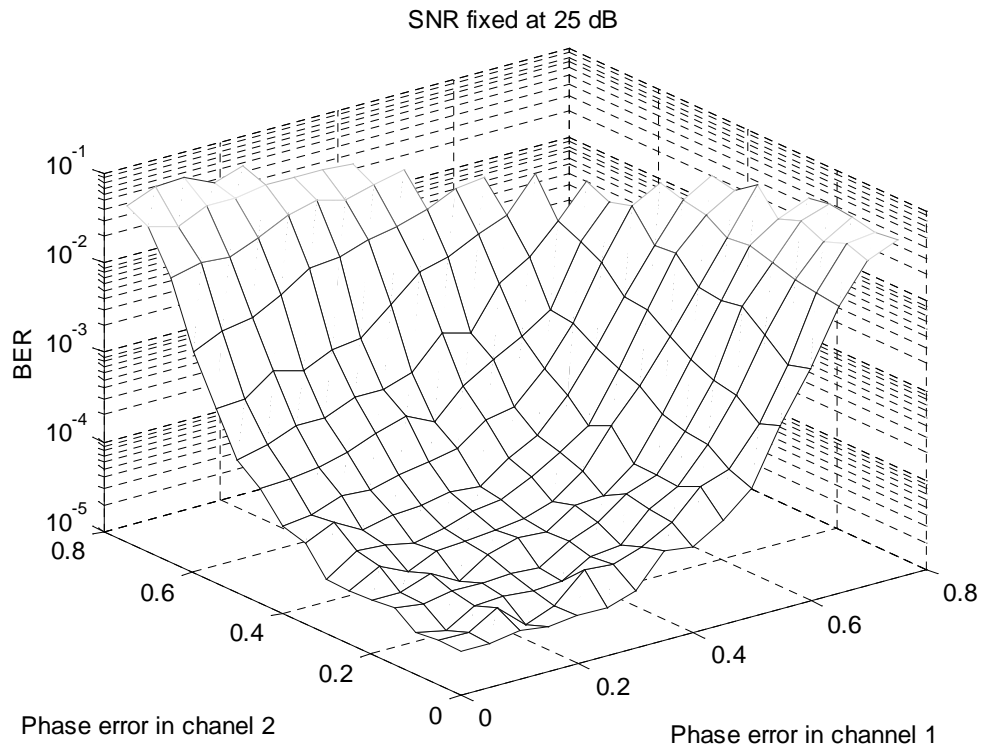


Figure 4.4. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 25 dB, and a maximum phase error of $\pi/4$ radians.

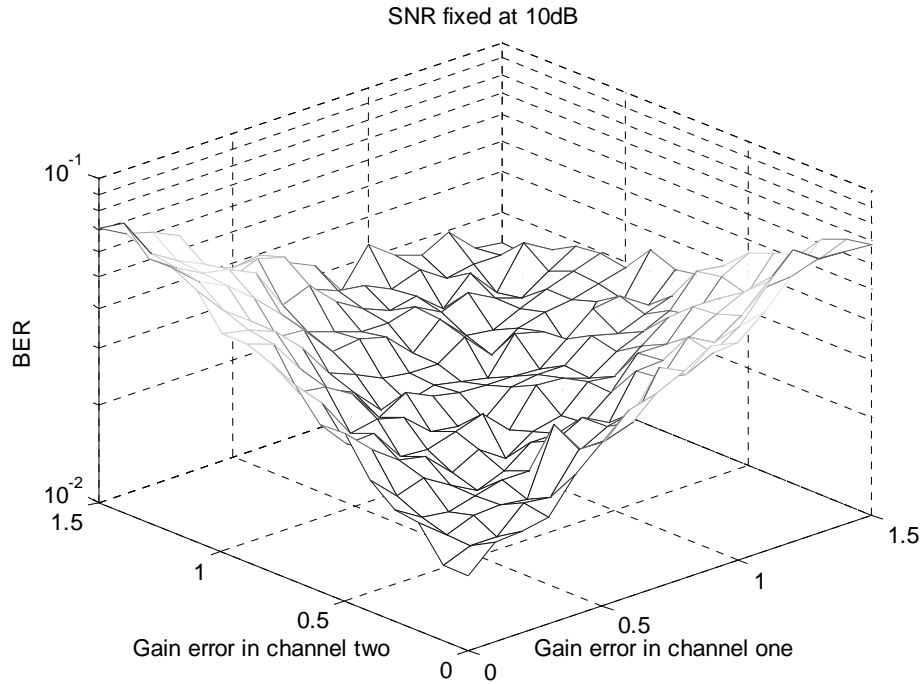


Figure 4.5. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 10 dB, and a maximum gain error of 1.5.

Figures 4.6 and 4.7 show the bit error rates for the normalized gain error at fixed SNR of 10 and 20 dB. The performance becomes extremely degraded only when the difference between the gain errors of each channel differ by an order of magnitude. This shows that a large degree of error can be tolerated in the estimate, especially if the degree of error is relatively equal in each channel.

The results of bit error rate performance with various levels of average phase error per channel are shown in figure 4.8. When the average phase error in both channels exceeds 0.6 radians the performance is not acceptable even at large values of SNR. Figure 4.9 is a plot of BER versus received SNR for several values of average normalized gain error. Regardless of the degree of error in any individual channel, if the normalized error is close to one, the performance is very close to that of having perfect channel estimates. Even when the error in one channel is nearly double that of the error in the other channel, acceptable performance can still be achieved.

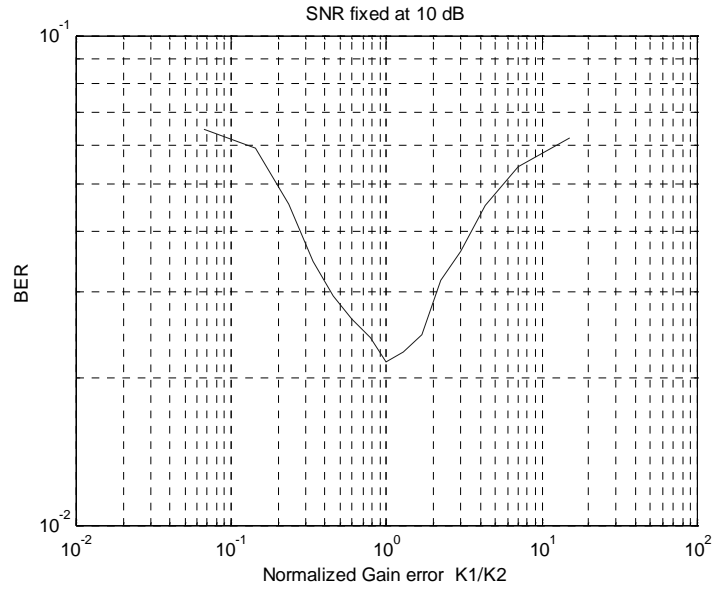


Figure 4.6. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 10 dB, and a range of gain error.

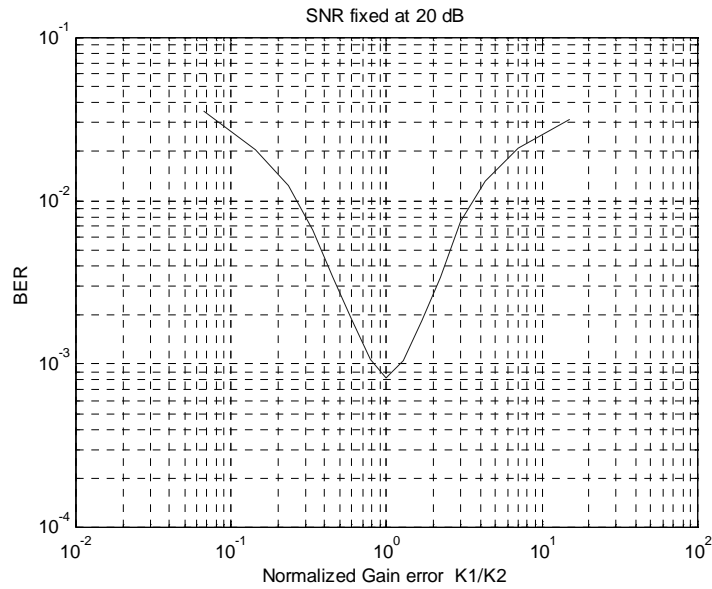


Figure 4.7. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 20 dB, and a range of gain error.

The plot in figure 4.10 shows the bit error rate performance of the system when there are errors present in both the gain and the phase of the channel estimate. This is a plot of BER versus normalized gain error at several values of average phase error per channel. Once the average phase error exceeds approximately 0.5 radians the degree of gain error is irrelevant, because the performance is already too degraded. When there are relatively small phase errors the degree of gain error can be relatively high without a large performance penalty.

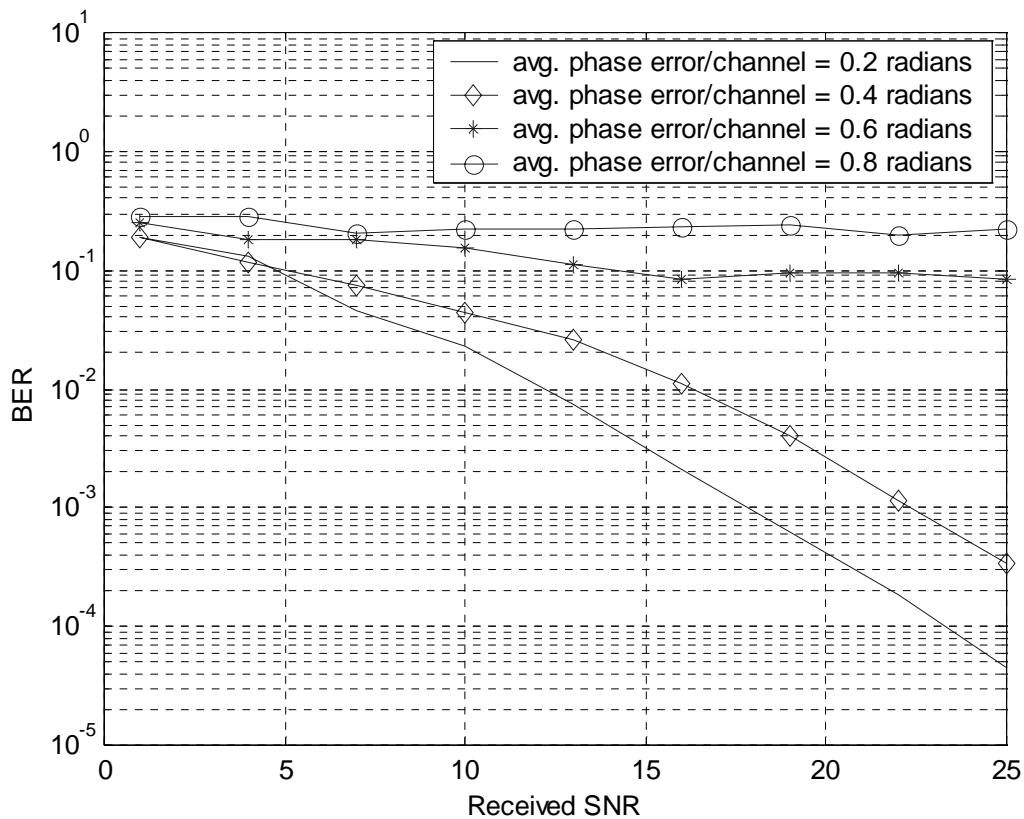


Figure 4.8. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, and various levels of phase errors per channel.

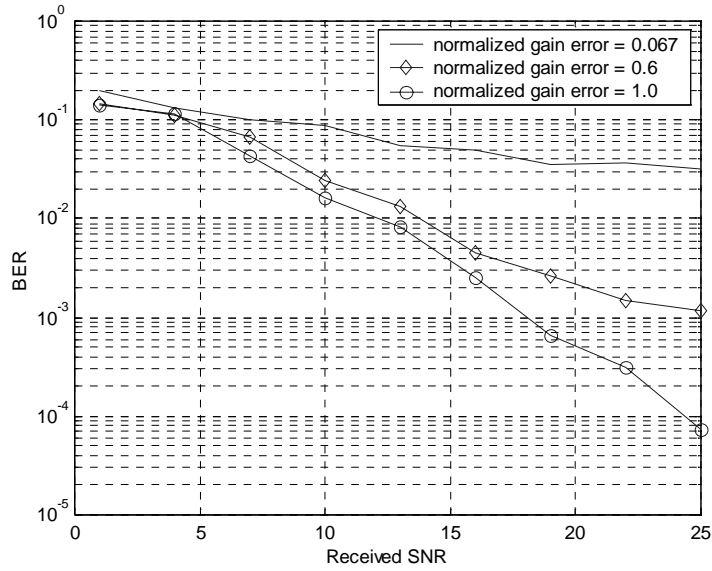


Figure 4.9. Performance of STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, and various levels of gain errors per channel.

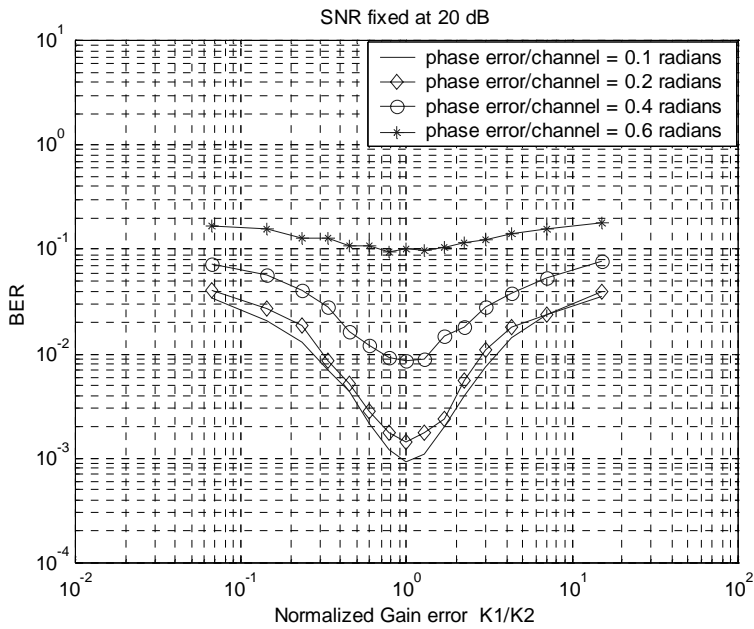


Figure 4.10. BER versus normalized gain error for STBC in Rayleigh flat fading with QPSK modulation, two transmit antennas, one receive antenna, a fixed SNR of 20 dB and various levels of phase errors per channel.

The last figure in this chapter, figure 4.11, shows the performance of STBC with QPSK signal constellation that used a running average pilot sequence estimation scheme. The running average pilot sequence estimation scheme uses two known symbols at the beginning of each block of four symbols. After the two known symbols have been transmitted they are used to calculate an estimate of the channel fading coefficients. Then this estimate is used in the decoding of the following two symbols. Since the fading is assumed to be quasi-static, the fading coefficients do not change over a frame of length 60 bits. After an estimate has been generated for a four symbol block, it is averaged with the previous estimates of the frame. This is why it is called a running average estimation. These results show that this estimation scheme performs with only a 2dB loss when compared to the case of perfect CSI. However, this performance comes at the cost of reducing the data rate by half.

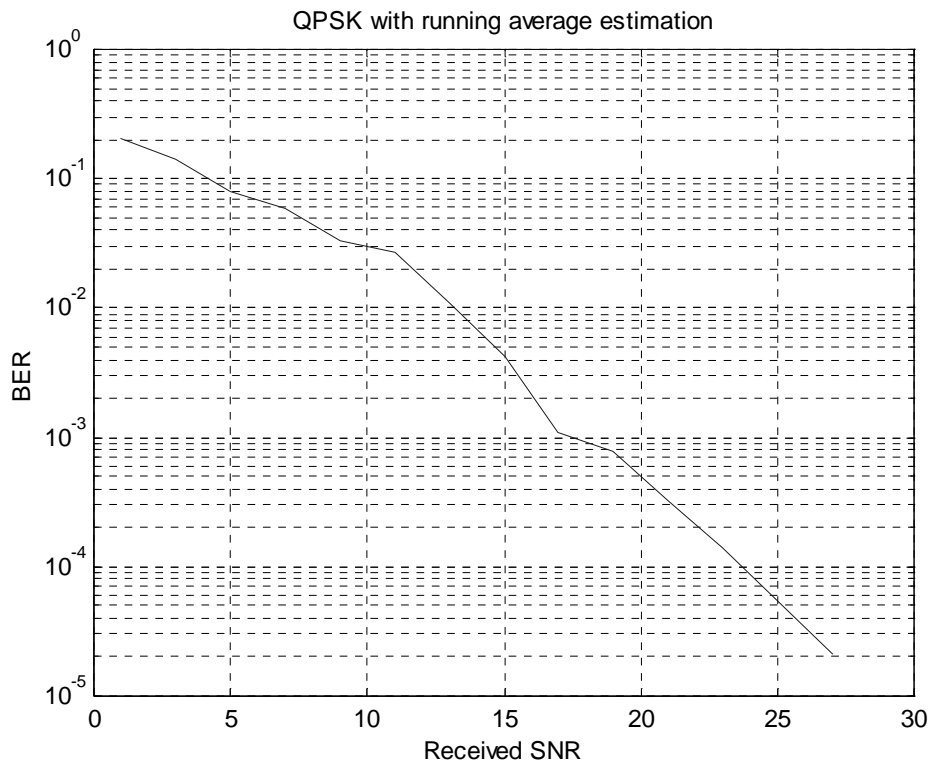


Figure 4.11 Rate $\frac{1}{2}$ QPSK using channel estimates derived from known transmitted symbols and using static fades over frame of 60 bits

4.3 Chapter Summary

In this chapter we have shown the performance of space-time block codes when decoded using imperfect estimates of the channel. For the case of a two transmit antenna system employing a QPSK constellation we have shown that errors in the amplitude of the channel estimate have a relatively minor effect on the bit error rate performance. If the amount of gain error in each channel is approximately the same there is almost no performance degradation. However, errors in the phase have the predominant effect, as would be expected when using Phase Shift Keying. The amount of error that can be tolerated in the phase of the channel estimate before the performance completely breaks down, is approximately 0.5 radians. Even when the level of error in the phase is 0.4 radians the performance has been greatly degraded.

A pilot sequence estimation scheme was used to generate the channel estimates in a simulation of STBC using QPSK. This scheme reduced the data rate by $\frac{1}{2}$ but performed with only a 2dB loss when compared to the case where perfect CSI is available to the receiver.

CONCLUSION

The first two chapters of this thesis were an attempt to provide a sufficient background to understand space-time block coding. The first chapter discussed modulation and introduced the concept of signal spaces and how the design of demodulators and detectors are dependent on the signal space representation of any modulation format. The second chapter was devoted to the introduction and explanation of the wireless channel, the effects of the channel on performance, methods to model the channel statistically, and methods such as diversity to improve performance.

The third chapter was an introduction to the concept of space-time block coding. Its purpose was to introduce the system model and clarify any assumptions that were made in the creation of a descriptive model. It was shown that space-time block codes provide the maximum diversity gain at high data rates and with little decoding complexity. However, all these results were shown based on the assumption that perfect CSI was available to the receiver.

The fourth chapter of the thesis first attempted to explain the necessity of analyzing space-time block codes with channel estimation errors. In order to have a reasonable predictor of a system, as it would be deployed, it is necessary to build a model that as closely resembles the actual conditions that would be present in the deployed environment. This “realistic” model requires that certain assumptions that were made to prove the theory behind STBC be cast aside. The predominant assumption that would be invalid in a physically realizable system is the presence of perfect channel state information (CSI) at the receiver. An actual implementation of a STBC system would rely on some form of channel estimation scheme to provide the receiver with imperfect estimates of the channel fading coefficients. In this chapter we created estimates of the channel that contained errors in the phase of the estimate, the gain of the estimate, or both, and used these incorrect, or noisy estimates to demodulate and detect the transmitted signals.

The results of these simulations showed that the presence of errors in the phase of the channel estimate severely degraded the performance of the STBC using QPSK. Specifically, when the error in the phase of the estimate approached 0.5 radians the performance was degraded to the point where the detector produced the wrong output nearly half the time even at high signal to noise ratio. When the errors in the gain of the estimate were roughly equivalent for each channel, the performance was not really affected. At this point we defined the normalized gain error to be the gain error in the estimate of the first channel divided by the gain error in the estimate of the second channel. When the normalized gain error was unity the performance was similar to that shown for the case of perfect CSI. At a normalized gain error of 0.6, the error in channel two was nearly twice as large as the error in channel one, there was a performance loss of about 7dB at a bit error rate of 10^{-3} . This work was published in a conference paper for the 2001 Virginia Tech Symposium on Wireless Personal Communications [Bak01].

The last results shown in chapter four were the performance of STBC using an estimation scheme based on the use of pilot sequences. The system used two known symbols at the beginning of each four-symbol block to gain an estimate of the channel. Since the channel was assumed to exhibit quasi-static fading, the fades for each channel are constant over a frame. For this reason the channel estimate for each block of symbols was averaged with the previous estimates to improve the estimate. The results show that this estimation scheme performs with only a 2dB loss when compared to the case of perfect CSI. However, this performance comes at the cost of reducing the data rate by half.

The following is a brief discussion of possible avenues for future research. In this thesis, the only signal constellation used to test the robustness of STBC against channel estimation errors was QPSK. M-ary Shift Keying using values of M other than four should be tested but will most likely confirm the results already generated to be fairly consistent for PSK with any value of M. It would be more beneficial to test the codes with several different types of modulation, such as Quadrature Amplitude Modulation (QAM), Frequency Shift Keying (FSK), or Minimum Shift Keying. This will show the different dependencies

different modulation types have on magnitude of phase errors. In the case of QPSK, analyzed in chapter four it was the phase of the estimate that had the largest impact on bit error rate performance. Therefore, it is likely that QAM would suffer increased performance degradation due to error in the amplitude of the estimates because in QAM information is in the magnitude and the phase of the transmitted signal.

In this thesis we ran a simulation using pilot sequences to perform what we called a running average estimation of the channel. In the future it would be beneficial to design an estimation scheme based on the performance of a particular modulation type with channel estimation errors. This would involve analyzing the performance of the modulation type against estimation errors similar to what was done for QPSK in chapter four and taking into account those results in the design of the estimation scheme. Research in this area could perhaps generate new types of estimation schemes that either entail greater performance or minimized complexity.

Since it was shown in this thesis that the performance loss was primarily due to errors in the phase of the channel estimate it would be prudent to build a space-time block code system with some form of Differential Phase Shift Keying (DPSK). In a DPSK system it is not necessary for the decoder to know the exact phase of a received signal only the difference in phase between the current and previous signals. This could serve to alleviate some of the loss in performance due to errors in the phase of the channel estimate. Work of this nature was presented in a recent paper by [Hug00].

In chapter four the performance of space-time block codes with channel estimation errors was analyzed under the assumption that the channel fading conditions changed slowly enough that the fading coefficients were constant over a frame. It would be an interesting topic to vary the rate at which the channel fades and measure the impact on performance when there is a given level of error in the channel estimates.

In chapters three and four of this thesis it was assumed that the fading that occurred over each transmitter-receiver pair was independent. In a physical system it may not be possible to place the antennas at the minimum required distance apart, thus leading to correlated fading between the different antennas. For this reason, the performance of the system with

channel estimation errors should be shown under the condition that correlation exists between the different fading coefficients.

The pilot estimation sequence in this thesis that was used in chapter four reduced the data rate by half. It would be beneficial to conduct further simulations at data rates closer to one, to measure the amount of performance loss as less pilot symbols are used and the data rate increases.

Future research should also focus on iterative decoding and estimation [Val01]. This is a process where pilot symbols are used to estimate the channel, then decoding is done using the channel estimates. The decoded data is then fed back into the estimator and the data used as pilot symbols to refine the estimate of the channel. The refined estimate is then used to decode the data a second time. This process can be repeated for any number of iterations until the performance reaches a desirable level. This idea leads naturally to the concept of turbo space-time codes.

BIBLIOGRAPHY

- [Ala98] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE Journal on Selected Areas in Commun.*, vol 16, pp.1451-1458, Oct. 1998.
- [Bak01] D. Baker and M. C. Valenti, "The Impact of Channel Estimation Errors on Space-Time Block Codes," in *Proc., Virginia Tech Symposium on Wireless Personal Communications*, Blacksburg, Virginia, June 2001, to appear.
- [Che83] D. K. Cheng, *Field and Wave Electromagnetics*, Addison-Wesley, New York, 1983.
- [Gar94] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, Second Edition, Addison Wesley, 1994.
- [Gue99] J. Guey, M. Fitz, M. Bell, and W. Kuo, "Signal design for transmitter diversity over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 527-537, April 1999.
- [Hay94] S. Haykin, *Communication Systems*, J. Wiley & Sons, New York, 1994.
- [Hug00] B. L. Hughes, "Differential Space-Time Modulation," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2567-2578, Mar. 1998.
- [Jak74] W.C. Jakes, Jr.(Ed.) *Microwave Mobile Communications*. J. Wiley & Sons, New York, 1974.
- [Lee94] E. A. Lee, D. G. Messerschmitt, *Digital Communications, Second Edition*, Kluwer Academic Publishers, 1994.
- [Lin00] X. Lin, R. S. Blum, "Improved space-time codes using serial concatenation," *IEEE Communications Letters*, vol. 4, no. 7, pp. 221-223, July 2000.
- [Pro95] J. G. Proakis, *Digital Communications*, McGraw-Hill, Inc., 1995.
- [Rap96] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, 1996.
- [Sei91] S. Y. Seidel, T. S. Rappaport, S. Jain, M. Lord, and R. Singh, "Path loss, scattering and multipath delay statistics in four European cities for digital cellular and microcellular radiotelephone," *IEEE Trans. Veh. Tech.*, vol. 40, pp. 721-730, Nov. 1991.

- [Sim00] M. K. Simon, M. Alouini, *Digital Communications over Fading Channels*, J. Wiley & Sons, New York, 2000.
- [Sk197] B. Sklar, "Rayleigh Fading Channels in Mobile Digital Communication Systems Part I: Characterization," *IEEE Commun. Mag.*, pp 136-146, Sep. 1997.
- [Tar98] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.
- [Tar99a] V. Tarokh, A. Naguib, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Comm.*, vol. 47, no. 2, pp. 199-207, Feb. 1999.
- [Tar99b] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE JSAC*, vol. 17, no. 3, pp. 451-460, March 1999.
- [Wit93] A. Wittenben, "Base station modulation diversity for digital SIMULCAST," in *Proc. IEEE Vehicular Tech. Conf.*, pp. 505-511, May 1993.
- [Val01] M.C. Valenti and B.D. Woerner, "Iterative Channel Estimation and Decoding of Pilot Symbol Assisted Turbo Codes over Flat-fading Channels," *IEEE Journal on Selected Areas in Commun.*, to appear, 2001.