CHANNEL PREDICTION FOR ADAPTIVE MODULATION IN WIRELESS COMMUNICATIONS

Raymond Chan

Thesis submitted to the Faculty of
Virginia Polytechnic and State
University in partial fulfillment of
the requirements for the degree of

Master of Science

In

Electrical Engineering

R. Michael Buehrer, Chair
Brian Woerner
Tim Pratt

July 16, 2003
Blacksburg, Virginia

Keywords: Channel Prediction, Linear Prediction, Adaptive Modulation, Channel Estimation, Rayleigh Fading, Pilot Symbol Assisted Modulation, Forward Error Correction

CHANNEL PREDICTION FOR ADAPTIVE MODULATION IN WIRELESS COMMUNICATIONS

Raymond Chan

(ABSTRACT)

This thesis examines the benefits of using adaptive modulation and coding in terms of spectral efficiency and probability of bit error. Specifically, we examine the performance enhancement made possible by using linear prediction along with channel estimation in conjunction with adaptive modulation. We begin this manuscript with basic fundamentals of our study, followed by a detailed view of simulations, their results, and our conclusions from them. The study includes simulations in slow and moderately fast flat fading Rayleigh channels.

We present our findings regarding the advantages of using predictive measures to foresee the state of the channel and make adjustments to transmissions accordingly.

In addition to finding the general advantages of channel prediction in adaptive modulation, we explore various ways to adjust the prediction algorithm when we are faced with high Doppler rates and fast fading.

By the end of this work, we should have a better understanding of when channel prediction is most valuable to adaptive modulation and when it is weakest, and how we can alleviate the problems that prediction will have in harsh environments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Number *Page*

ACKNOWLEDGMENTS

I would like to express sincere appreciation to Professor Buehrer for his assistance in the preparation of this manuscript and the work it represents. Without his patience and guidance, nothing would have been done. I would also like to thank my committee for taking the time and effort that they did for the review of my work. Also, I would like to thank the students of MPRG for giving me all the help and advice. I should also have to thank my family for giving me the support that they have. And lastly, I thank all of my old teachers who have helped shape me into who I am today. Special thanks to Peter Kinman and Fugen Shinjin.

Chapter 1

INTRODUCTION

1.1 Overview of the Problem

In wireless communications, spectrum is a most precious commodity. It, along with SNR, is what determines the rate at which we can transmit information. For many years, researchers have been looking for ways to either conserve or exploit spectrum for its most efficient use. In the early 90's, TDMA gave us improved spectral efficiency over FDMA. Later that decade, CDMA provided even better spectral efficiencies (from a system point of view).

One of the barriers to wireless systems is Rayleigh fading; a phenomenon that reduces error performance in mobile fading environments. Whether in

In this thesis, we investigate ways to take advantage of Rayleigh fading by means of adaptive modulation and coding [1],[2]. By this method, we mean to change the modulation and/or coding used by the transmitter in response to the changing channel conditions. In essence, it is a way to optimize the transmission scheme according to the state of the channel for a required fidelity. For example, when the channel is in a poor state (i.e., low SNR) we can reduce the signal constellation size in order to improve fidelity. Conversely, when the channel is in a

good state (high SNR) we can increase the signal constellation size in order to increase the data rate achievable.

The same argument can be made for forward error correction coding. In periods of deep fade, we can lower the code rate and make the transmission more resilient to errors. In addition to this, through the use of punctured codes, we may provide flexibility with the lower complexity afforded by a common decoder structure.

1.2 Background

Adaptive modulation for wireless communications has received significant interest in the past five years [1], [3], [4]. It has long been recognized that adaptive modulation provides more efficient use of the channel than fixed modulation schemes. Landline modems have long adapted the modulation scheme to the SNR of the underlying channel. Adaptive modulation has also recently gained momentum in wireless systems. In fact, forms of adaptive modulation are currently implemented in a packet data CDMA standard [3] and in wireless LAN standards such as IEEE 802.11. However, the challenge associated with adaptive modulation is that the mobile channel change with time. Thus, the feedback of channel information becomes the limiting factor in adaptive modulation. To meet this challenge, in this study we examine the use of channel prediction to allow the use of past information to predict the future channel state and adjust the modulation scheme accordingly.

Signal prediction is also not a new topic. In the past, linear prediction was used extensively in speech processing and data

analysis. Algorithms such as ESPIRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) have been explored by Andersen and Roy [5], [6] for wireless fading channels and general signal parameter estimation (direction of arrival, system identification and time series analysis). Hwang and Winters [7] worked on subspace methods to predict the channel in periods of fast fading for high frequency carrier waves. More relevant to our topic is the work done by Duel-Hallen [8] on linear channel prediction. In Duel-Hallen's work, she proposed upgrading adaptive modulation by means of implementing linear prediction to predict the future power levels of the fading channel. Her work has focused on the use of simulation data and the mean square error of the prediction methods. She has worked to study how well prediction works by altering various prediction parameters in her systems. Recently Hu has applied linear prediction to the adaptive modulation problem, although the work is still in its early stages [9]. In this work we will more fully explore one of the uses of linear prediction in adaptive modulation building on the work done in adaptive modulation by Jain in [4].

1.3 Contributions

In this study, we will demonstrate the advantages of using channel prediction in conjunction with adaptive modulation techniques. Specifically, we will show the change in throughput and bit error rate performance when using linear prediction to estimate the future state of the channel. This thesis makes several distinct contributions to the literature:

1. We examine the impact of using FFT-based channel estimation on the performance of linear prediction.

2. We illustrate the impact that linear prediction has on the BER performance of adaptive modulation in different Doppler environments and with reverse link propagation delays.

3. We illustrate the impact that error correction coding has on the performance of adaptive modulation, specifically when using linear prediction.

4. We illustrate the limitations of linear prediction in high Doppler environments and demonstrate the impact that these limitations have on the performance of adaptive modulation. Further, we investigate several methods to improve prediction accuracy and their impact on adaptive modulation.

5. We demonstrate the usefulness of channel prediction and adaptive modulation using actual measured channels.

The organization of this thesis is as follows:

In Chapter 2 we lay the foundation of the system model. We will briefly review modulation and Rayleigh fading. We also introduce our methods of channel compensation and SNR estimation.

In Chapter 3 we introduce adaptive modulation. We observe the effects that imperfect channel estimation has on the performance of adaptive modulation and we present linear prediction as a means of counteracting some of these problems.

Chapter 4 presents the specific issue of high Doppler fading and the ways in which we can compensate for the weaknesses of channel prediction in high speed channels.

Chapter 5 briefly concludes this work.

Chapter 2

THE SYSTEM MODEL

2.1 Introduction

In this chapter, we will outline in detail the system model used in this study. We will begin with a discussion of additive white Gaussian noise, followed by a look at Rayleigh fading. We will then discuss the channel estimation and compensation methods used, as well as the SNR estimation technique employed.

2.2 Noise

In no examination of a communication system is a model produced without the inclusion of noise, or some unwanted signal in addition to the desired information signal. In most, if not all studies (ours included), the term noise will denote additive white Gaussian noise (AWGN). The term ‘additive’ means the interfering signal is combined with the information signal by addition; the term ‘white’ refers to the noise having constant power spectral density; and Gaussian means that the probability density function of samples of the random process is Gaussian.

It is interesting to discuss where AWGN originates. The answer lies in two places, but for the same reason, namely, Brownian motion. One source of noise is from the agitation of charge carriers in a conductor moving through a potential barrier [10]. This type of noise is called shot noise and usually occurs in

electrical devices such as diodes and transistors and generally does not come into play in link level communication models. The other source of noise is thermal noise. This refers to the random motion of electrons due to thermal energy. This random motion at the receiver produces a signal that is independent of the desired signal and is detected in addition to the information signal. Since this noise is produced by a large number of independent sources, it is well modeled by a Gaussian random process. Note that decreasing the temperature of the receiver can reduce the power of the thermal noise but not shot noise.

From this point on, all references to noise will be thermal noise or AWGN.

2.2.1 Noise Power

To see the effect that noise has on the received signal, we are interested in the power of the noise signal relative to the power of the desired signal. The ratio of the two is known as the signal-to-noise-ratio (SNR). The units that we will be using are logarithmic, noted by decibels, or dB. To find the power of the noise, one must know the temperature in Kelvin at the receiver, and the receiver noise bandwidth [10]:

$$N_p = kTB \tag{2.1}$$

where k is Boltzmann's Constant ($1.38*10^{-23}$ W/Hz*K), T is the thermal temperature in Kelvin, and B is the effective noise bandwidth of the receiver in Hz.

The power of the noise in our simulation is such that the received SNR is in the range of 0 to 40 dB. In a later section,

we will see the significance of noise power in conjunction with channel fading.

2.3 Rayleigh Fading

In this section, we will examine the impact that the movement on the mobile receiver (or transmitter) causes on the received signal. The resulting effect (typically called fading) has a significant impact on the received signal strength which can be severely degraded as we will show in the following sections.

2.3.1 Doppler Effect

Fading is caused by a phenomenon known as the Doppler Effect. When dealing with any sort of waves, a receiver's movement in relation to the source of the wave will distort the perceived frequency of that wave.

Consider the following figure from [11]:

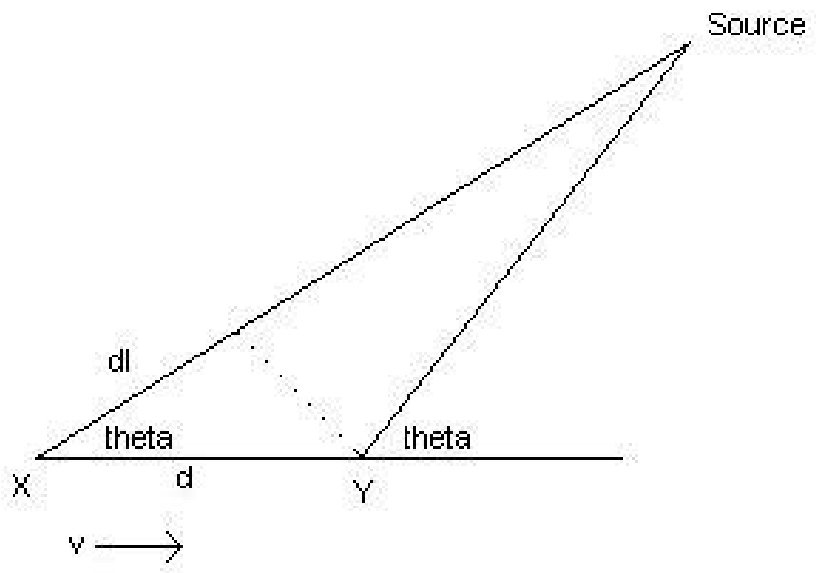

Figure 2.1 - Illustration of Doppler Effect

where a user is at point X, moving toward point Y at velocity v. We make the approximation that the angle of arrival of the received signal is the same at all points during transmission (i.e., the source is very far away). The difference in distance that the transmission must travel between the receiver at points X and Y is: $\Delta l = v\Delta t \cos q$. The resulting phase change between the two points is:

$$\Delta f = \frac{2p\Delta l}{l} = \frac{2p\, v\Delta t}{l} \cdot \cos q \qquad (2.2)$$

where l is the wavelength in meters. The Doppler frequency can then be found as:

$$f_d = \frac{1}{2p} \cdot \frac{\Delta f}{\Delta t} = \frac{v}{l} \cdot \cos q \qquad (2.3)$$

One can see that if the receiver is heading toward the source ($q < 90^o$), the Doppler shift will be positive and the perceived frequency of the signal will be increased. If the receiver is moving away from the source ($q > 90^o$), the Doppler shift will be negative and the perceived frequency will be decreased. If the user is moving perpendicular to the source ($q = 90^o$), there will be no shift in frequency.

The maximum Doppler shift is $\frac{v}{l}$ and this is in units of Hz.

2.3.2 Multipath Channels

In addition to Doppler, our channel model will include multipath distortion. What we mean by this is that there will be several copies of the same signal being picked up by the receiver coming from different angles. This is common in environments such as metropolitan areas, where electromagnetic waves can bounce off buildings or in indoor environments where signals reflect off walls, etc. Each of these components will have their own Doppler shift and phase offset due to different angles-of-arrival and time delays. The combination of these paths will be constructive and destructive due to the different phases causing the signal strength to change with mobile movement. It is this multipath factor that produces the gains and fades in power that we will see in the figures ahead.

2.3.3 Rayleigh Model

With the Doppler shift and multipath, we can now build a model of our Rayleigh channel. From [11], we model the received signal as a sum of non-resolvable multipath components each with independent amplitude, phase and frequency components. The channel is thus modeled in complex baseband as

$$c(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} A_i \cdot e^{j(2p f_{ci} t + f_i)} \qquad (2.4)$$

Where A_i is the amplitude of the i^{th} complex sinusoid which is a constant of 1 in our case, f_i is the random phase uniformly distributed from zero to $2p$, and f_{ci} is the Doppler frequency found as $f_i = f_{\max} \cos q_i$ where f_{max} is the maximum Doppler shift and q is the angle of arrival assumed to be uniformly distributed from 0 to $2p$. The channel is made up of the sum of N complex sinusoids. We set the number of multipaths N to be 32 in all experiments. Note that the resulting channel is a complex Gaussian random process due to the sum of independent sinusoids. The envelope is thus a Rayleigh random variable and thus we call this Rayleigh fading.

In the next few figures, we will look at some of the characteristics of Rayleigh channel.

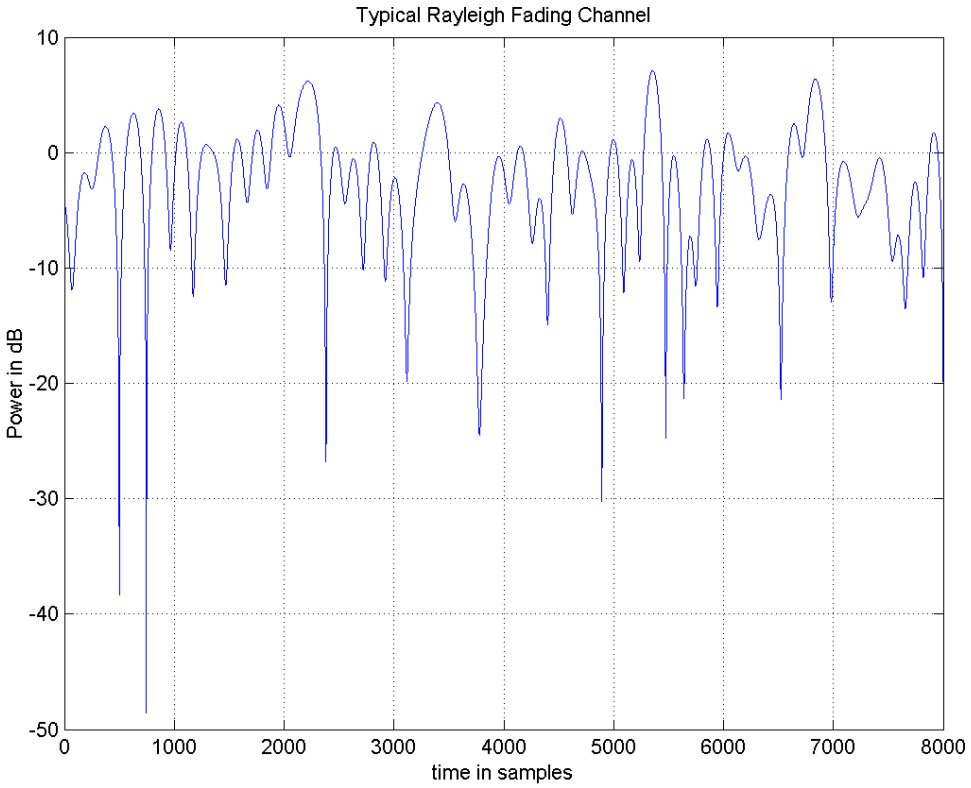

Figure 2.2 - Typical Rayleigh Fading Channel

In Figure 2.2, we plot the envelope (in dB) of a Rayleigh channel taken over a period of 8000 samples at a maximum Doppler frequency of 30 Hz and a sampling rate of 10 kHz. Note the many short periods of fading throughout channel. In this channel, we can lose between 5 and 50 dB of power from our signal.

On the other hand, the Rayleigh channel is somewhat deterministic and samples that are close to one another have similar power levels. This temporal correlation can be exploited as we will later show. But before we discuss why fading is so bad for

transmission, or how we exploit it, we further examine our assumed system model.

2.3.4 Modulation

In this section, we will examine the transmitter and receiver architectures assumed in this work. Specifically, we describe how we take a binary information signal and prepare it for transmission through free space. We will do this by taking a look at a signal transformation known as modulation.

We will begin with a binary signal of zeros and ones. What we would like to do is encode it in a format that will travel through a medium. We do that by encoding electromagnetic sinusoids, since radio waves travel readily through the media of interest.

In simple modulation schemes, it is possible that a binary ‘1’ would be represented by a cosine wave with some phase and frequency while a binary ‘0’ would be represented by that very same wave, but with a 180 degree change in phase. This is called binary phase shift keying, or BPSK. We could use any finite number of phases although we typically restrict ourselves to powers of two (e.g., BPSK, 4-PSK, 8-PSK). Our system will include modulation schemes QPSK and QAM, or Quadrature Amplitude Modulation.

In QAM, we map sets of bits to a complex sinusoid with a particular phase and amplitude. Below is a diagram [Fig 2.3] for 16 QAM, a modulation scheme that uses four information bits to produce one of 16 transmission symbols. Other QAM schemes use a different number of information bits to produce one information

symbol, hence, different modulation schemes can produce different spectral efficiencies as well as different BER curves. We will go into more of this later.

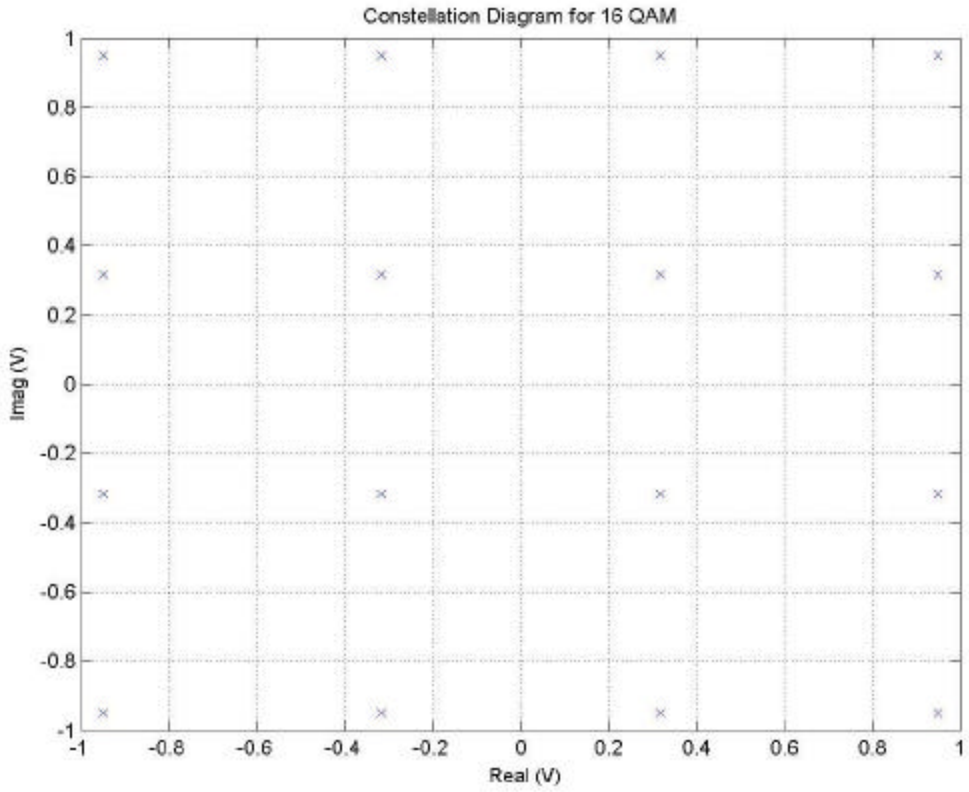

Figure 2.3 - Constellation Diagram for 16 QAM

Figure 2.3 shows all of the possible modulation symbols in what is termed a constellation diagram. The angle of the point in two dimensions represents the phase of the symbol while the magnitude represents the symbol's amplitude. This diagram illustrates the received signal points in the absence of noise. We will call the points in which the bits are mapped to 'constellation points'. The received signal is decoded by mapping the received signal

projected onto this two dimensional plane to the nearest constellation point. In other words, the decoded symbol is the constellation point that is closest to the received symbol.

2.3.5 Effects of Fading and Noise

At the receiver, we can model the signal as:

$$r(t) = c(t) \cdot s(t) + n(t) \qquad (2.5)$$

where $r(t)$ is the received signal, $c(t)$ is the Rayleigh channel signal defined in Equation 2.4 and $n(t)$ is Gaussian noise. Further, as is conventional, we model the signals using complex baseband notation. The information signal is multiplied by the channel, and white Gaussian noise is added to the signal at the receiver.

Let us see what happens to a signal when it is put through a Rayleigh channel.

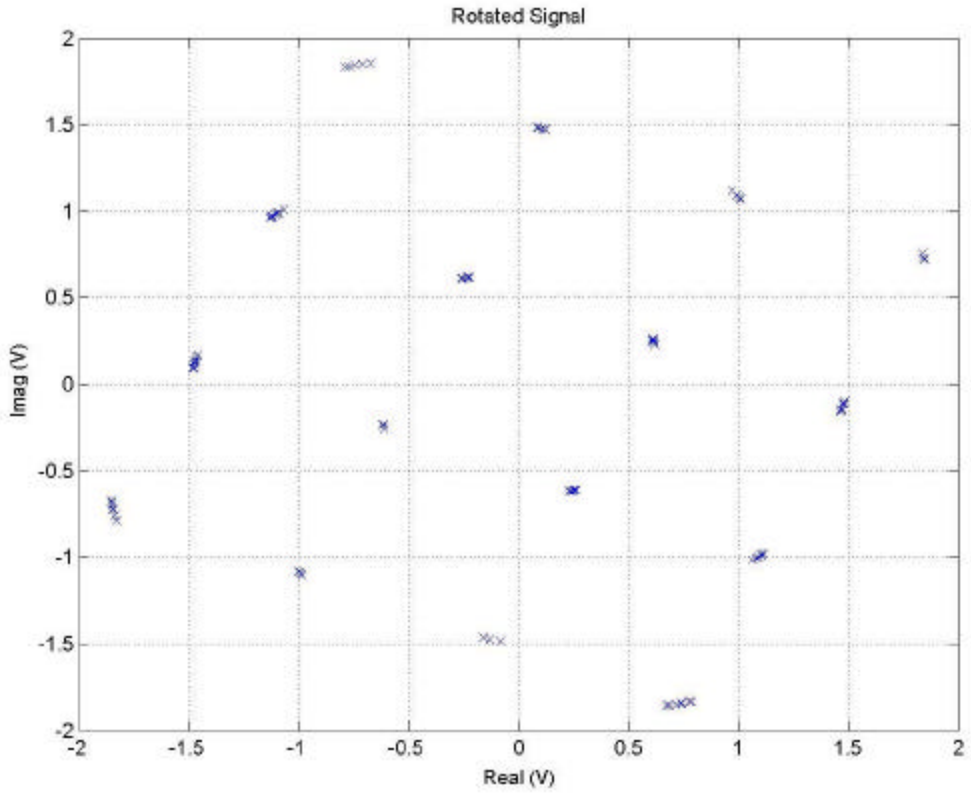

Figure 2.4 - Rotated 16 QAM signal

In Figure 2.4, we have taken the same QAM signal constellation from Figure 2.3 and passed it through a Rayleigh channel with a specific SNR. As can be seen, the received signal constellation is rotated and scaled and slightly perturbed. In order to demodulate this signal properly, one must compensate for the channel by de-rotating the constellation and removing the amplitude modulation (since we are working with QAM which has information in the amplitude). However, this is not the only problem that Rayleigh fading produces.

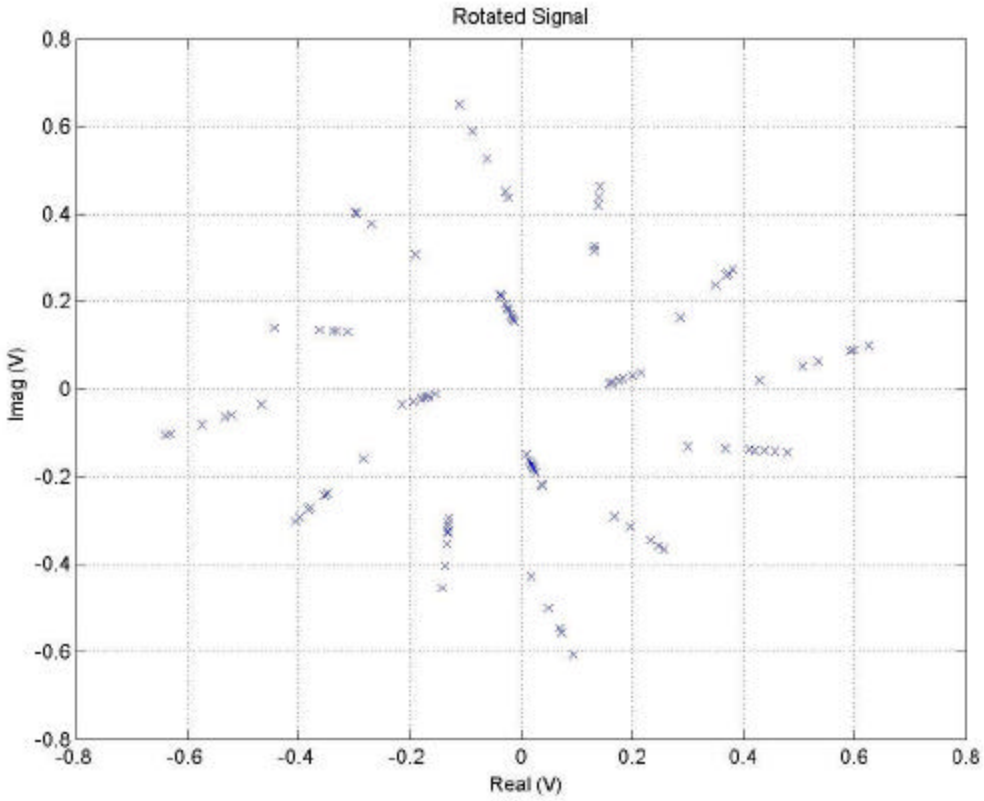

Figure 2.5 - Information Symbols are Rotated and Scaled

In Figure 2.5, we have rotation similar to what we saw in Figure 2.4. However, note that the constellation is scaled down much further than it was in Figure 2.4. This produces an additional problem when noise is introduced to the signal. The smaller the signal constellation, the less noise power is necessary to move the received symbol from the correct decision region to another. This is significant because it will cause symbol errors in a maximum likelihood detector.

Even without noise, amplitude and phase rotations can result in symbols errors if channel distortion is not countered. When a transmission encounters a deep fade and symbols are moved around by noise, symbols errors become more and more likely. However, the channel itself can be estimated and its effects mitigated through a method called Pilot Symbol Assisted Modulation (PSAM). This is a necessary step in the demodulation process when Rayleigh fading is involved.

2.4 Channel Estimation and Compensation

The job at the receiving end of a transmission is to decode the received signal and produce a bit stream that matches the original transmission bit stream at the beginning of the transmission process. Unfortunately, due to the distortion caused by the channel, we cannot do this directly. Instead, we must find a way to adjust the received signal before demodulation occurs. We will do this by means of Pilot Symbol Assisted Modulation (PSAM) [12].

The idea behind PSAM is to insert known symbols in our transmission at set intervals. These symbols are called pilot symbols.

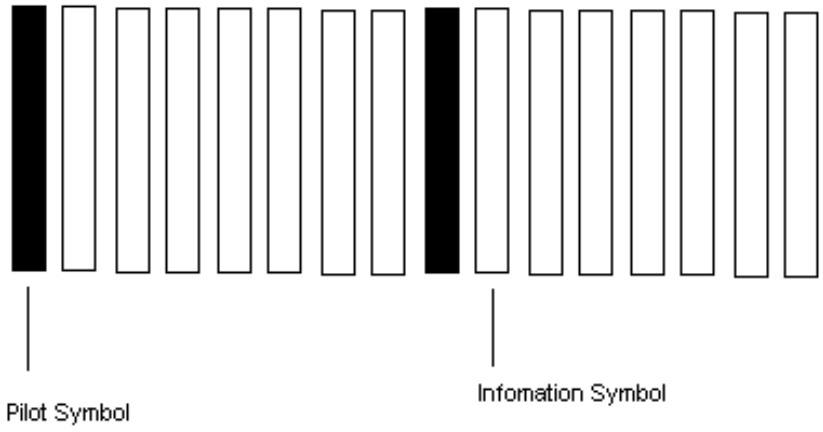

Figure 2.6 - Pilot Symbol Insertions

The purpose of the pilot symbol is to have a way at the receiver to know what the channel value is at the pilot. Using this system, we will have fairly accurate channel samples at a frequency equal to the pilot symbol rate. Using the pilots, we can interpolate the channel values in-between using a proper interpolation method. In this work, we use an FFT-based approach as we shall discuss.

Let us look at this process in more detail. First, there is a Nyquist condition that must be met [12]. Based on this criterion, we can show that the following relationship must be satisfied:

$$f_D T_s \leq \frac{1}{2N} \tag{2.6}$$

where f_D is the maximum Doppler frequency of the channel; T_s is the symbol period, and N is the number of symbols in a frame and

a frame is defined as a set of symbols associated with a single pilot symbol.

Now, let us describe the FFT interpolation procedure. We must first take the pilot symbols out of the received transmission stream. Then we divide each pilot by the expected information symbol, typically the symbol with greatest energy (this is determined beforehand). Next, we take the FFT of that vector [12].

$$G_1(n) = \sum_{l=0}^{2N_p-1} g_1(l) \cdot \exp(-\frac{jp\ nl}{N_p}) \quad (n = 0,1,\ldots,2N_p - 1) \qquad (2.7)$$

Where $2N_p$ is the number of pilot symbols used to create the channel estimate and $g_1(l)$ is the vector of channel samples obtained from dividing the received pilot symbols by the expected pilots. This gives us a vector of channel samples in the frequency domain. Once in the frequency domain, zero padding is used to accomplish interpolation.

$$G_1'(m) = \begin{cases} NG_1(m) & [0 \le m \le N_p - 1] \\ 0 & [N_p \le m \le N_p(2N-1) - 1] \\ NG_1(m - 2N_p(N-1)) & N_p(2N-1) \le m \le 2N_pN - 1 \end{cases} \qquad (2.8)$$

Equation 2.8 [12] shows us the way to implement the interpolation process in the frequency domain where N is the interpolation factor.

The IFFT is then performed with the following equation:

$$g_1'(k) = \frac{1}{2NN_p} \sum_{m=0}^{2NN_p} G_1'(m) \exp\left(\frac{jp\,mk}{NN_p}\right) \quad (k = 0,1,\ldots,2NN_p - 1) \qquad (2.9)$$

The result of Equation 2.9 is then the channel samples that we desire.

To employ these samples in channel compensation, recall Equation 2.5: $r(t) = c(t) \cdot s(t) + n(t)$. What we want to do is to remove the effect of *c(t)* on the received signal using the estimated *c'(t)* obtained from the FFT algorithm. It is the following:

$$z(t) = \frac{c(t) \cdot s(t) + n(t)}{c'(t)} = \frac{c(t) \cdot s(t)}{c'(t)} + \frac{n(t)}{c'(t)} \qquad (2.10)$$

Ideally, $\frac{c(t)}{c'(t)}$ is very close to 1, leaving us with $s(t) + \frac{n(t)}{c'(t)}$ where the noise term is hopefully very small compared to the signal term. Note that this process does not improve SNR since we are multiplying the data and the noise by the same value.

2.4.1 Edge Effects of FFT Estimation

There is one drawback to the FFT estimation method that we just discussed. As shown in Figure 2.7, at the ends of the interpolated channel, there are ripples in the channel estimate due to the finite block size used. This will cause inaccuracies when the estimate is used for compensation. The problem is that there are leakage components in the frequency domain when the symbols are transformed. What is needed in order to counter this is for the time interval of $g_1(l)$ to be an integer multiple of $1/f_D$. Also, applying Equation 2.7 is like truncating in the

frequency domain. This amounts to a loss of information, which is translated into an edge ripple when the IFFT is applied to the signal [13].

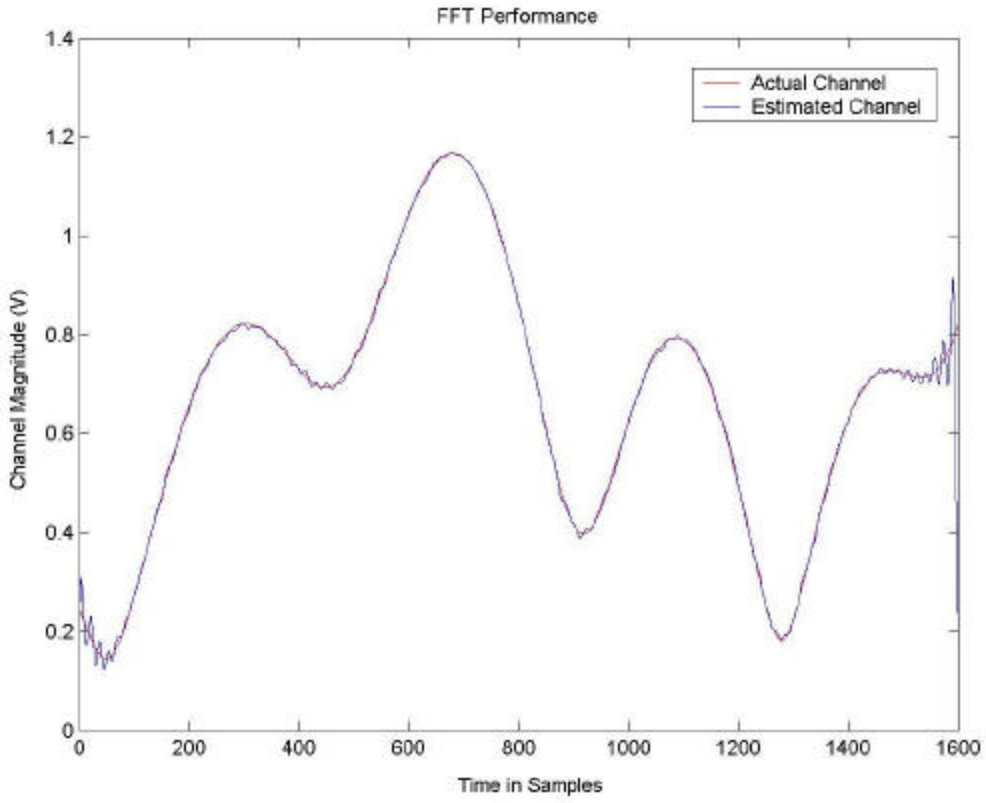

Figure 2.7 – Illustration of FFT Channel Estimation Accuracy

Figure 2.7 illustrates an example of the accuracy of the FFT interpolation algorithm. Note that the channel in the middle of the set is traced fairly accurately, but the channel at the ends is not.

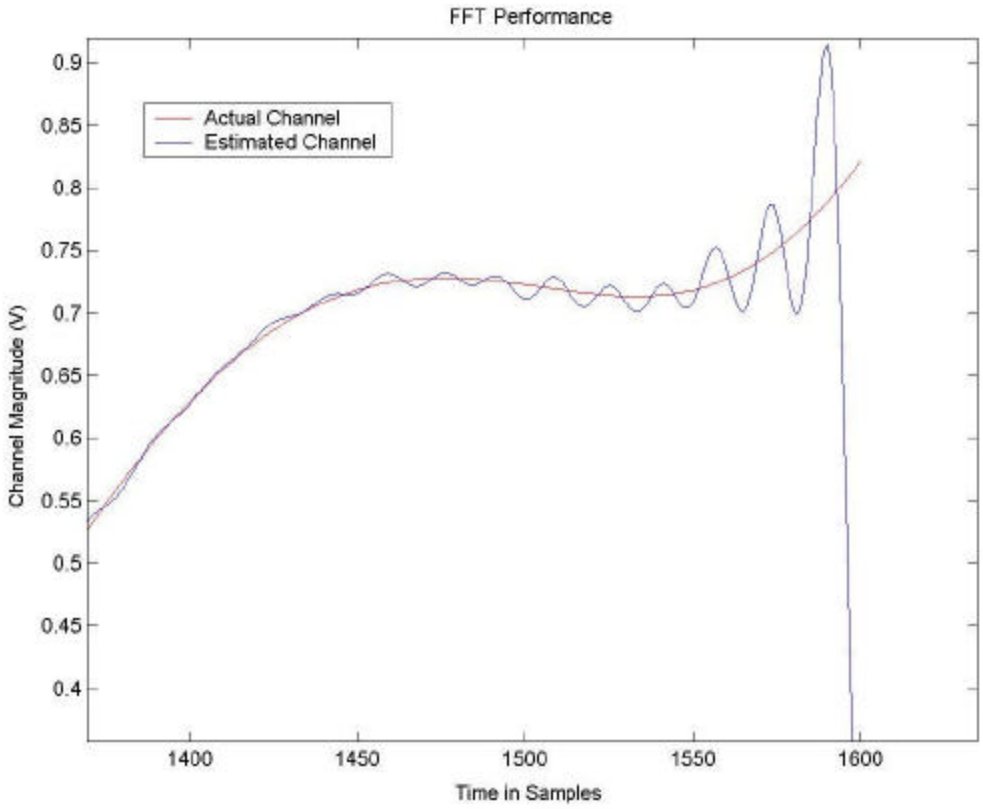

Figure 2.8 - Close-up of the edge effect

In Figure 2.8, we see a close up of the edge effect from the FFT algorithm.

To counteract this effect we place additional symbols at each end of the frame, so that the part of the frame that is corrupted by the edge effect is of no consequence, and the center of each frame is free from any negative effects of the edges. This will introduce some delay in our system.

2.5 SNR Estimation

The last function that we will discuss in terms of the system model is the estimation of signal to noise power ratio. The purpose of measuring SNR is to get a more accurate view of the channel state. The signal that we wish to receive and demodulate will be corrupted by both the Rayleigh channel (an effect which changes fairly rapidly) and receiver noise (whose statistics do not change over short intervals). Because of this, we will require the short term signal power (to estimate the effect of Rayleigh fading) and the long term noise power.

There are many papers and articles that discuss SNR estimation. The estimator that will be discussed in this work is based on work done by Gagliardi [14] and Jain [4]. The basic idea is the following: If we have a signal that is being corrupted by noise, we can estimate the signal to noise ratio through 'mean and variance' calculations. First, the average received power after Rayleigh fading is found by taking the square of the mean of the magnitude of the received signal (after removing the modulation). This will serve as the estimated signal power. The noise power will be found using a variance calculation on the received signal (again after removing the effect of modulation). This variance calculation can then be averaged over a long time period.

For PSK modulation, the following equations are used to find mean and variance of the current received block of data:

$$Z = \frac{1}{N}\sum_{i=1}^{N}|r_i| \qquad (2.11a)$$

$$T^2 = \frac{1}{N-1}\left(\sum_{i=1}^{N} \left(|r_i| - |c'_i| \right)^2 \right) \qquad (2.11b)$$

In the above equations, Z represents the mean of the received signal, T^2 represents the variance of the signal (ignoring the effects of modulation and the channel), and N is the number of symbols over which the statistics are estimated. Note that the variance calculation depends on the channel estimate c_i that comes from the FFT Estimator. Further the effect of modulation is eliminated by taking the magnitude of the received signal. In some estimators, the mean of the signal is used instead of the channel sample c_i. However, this leads to biased estimation in high SNR regions with even moderate Doppler rates[4].

For modulation schemes that have multiple power levels like QAM, we have a different set of equations to calculate the mean and variance of the signal:

$$Z = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{r_i}{\hat{d}_i} \right| \qquad (2.12a)$$

$$T^2 = \frac{1}{N-1}\left(\sum_{i=1}^{N} \left| r_i - c'_i \cdot \hat{d}_i \right|^2 \right) \qquad (2.12b)$$

It is worth noting that in the multi-power level schemes, the data symbol itself is needed to remove the effect that multiple power levels would have on the estimator. This is obtained using symbol estimates, $\hat{d}_i$.

In order to make these equations useful, we have to have a way to determine SNR from them. The SNR can be calculated as [4]:

$$\Gamma = \left(\frac{N-3}{N-1} \right) \frac{|Z|^2}{T^2} - \frac{1}{N} \qquad (2.13)$$

Equation 2.13 is simply the signal power divided by the noise power. The units of Γ are linear, so we must take the logarithm in order to get the units into decibels. We should note that Γ is a short term estimate based on one value of T^2 and Z. Since the noise power is constant over long periods of time, we can improve our SNR estimate by creating a long term noise estimate. To take long term average, Jain used the following equation to determine long term variance:

$$T_{long}^2(n) = 0.99 T_{long}^2(n-1) + 0.01 T^2(n) \qquad (2.14)$$

However, this method will take time to converge to a reliable estimate. Furthermore, it weighs the first set of variances a great deal. So in light of this, we decided to modify the equation to suit our needs. We thus have the following to estimate the noise power:

$$T_{long}^2(n) = \frac{n}{n+1} T_{long}^2(n-1) + \frac{1}{n+1} T^2(n) \qquad (2.15)$$

In Equation 2.15, the first few samples are still weighted more, but not as much as 2.14. In this case, the convergence of the SNR estimate will be faster.

2.5.1 SNR Estimator Results

In this section, we will present some results for the SNR estimator. In these simulation runs, we simply wanted to examine the accuracy of the estimator, so we did not include any other receiver functions. For the following charts, our communication system consisted only of a modulator, demodulator, and the SNR estimator in a Rayleigh and AWGN channel.

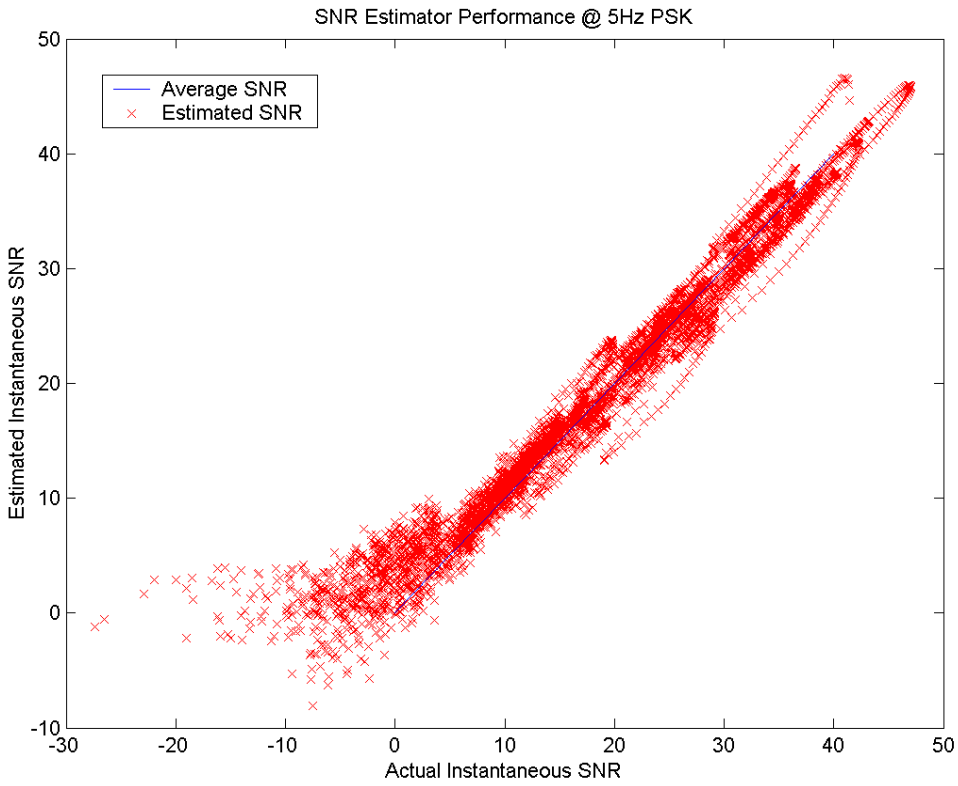

Figure 2.9 – SNR Estimation for PSK at 5Hz Doppler

Figure 2.9 plots the performance (i.e., estimated SNR versus true SNR) of the SNR estimator with perfect channel knowledge at a Rayleigh fading rate of 5Hz. It can be seen that at very low SNR, the estimate is over-estimated, but it gives accurate estimates for SNR at 5dB and higher. The x's on the graph represent the individual estimates of each packet, with its x-coordinate representing the actual SNR of the frame, and the y-coordinate representing the estimated SNR. The closer any x is to the line y=x, the more correct the estimate is to the actual SNR.

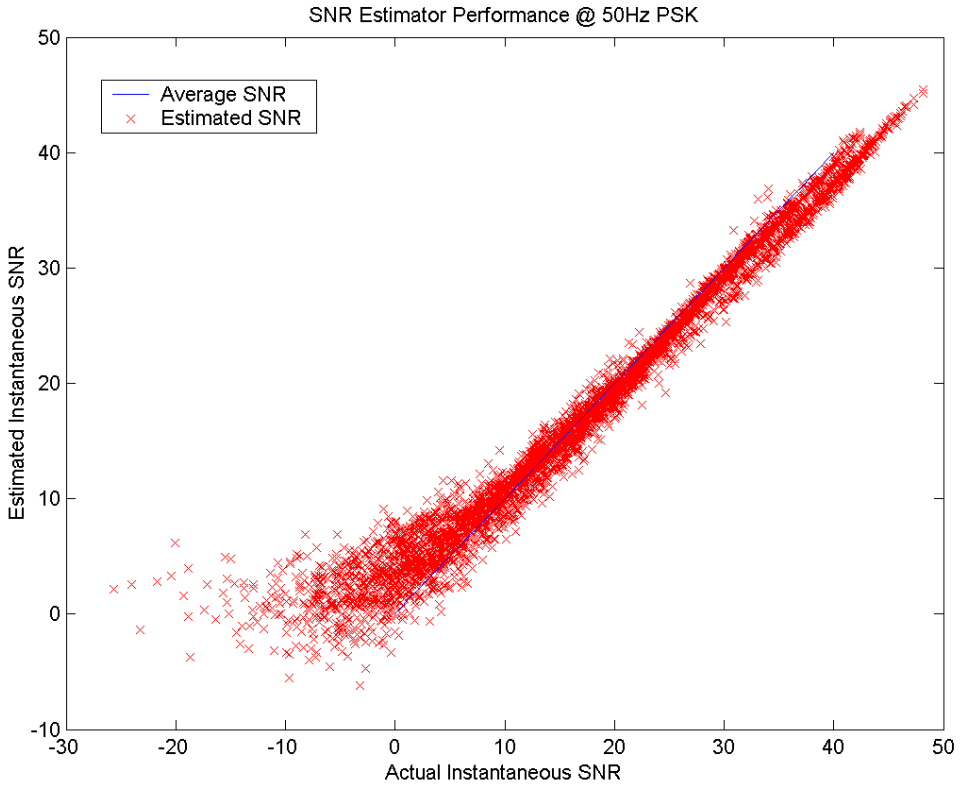

Figure 2.10 – SNR Estimation for PSK at 50Hz Doppler

In Figure 2.10, it can be seen that at higher Doppler, the estimator is still reliable, albeit slightly biased low, due to the faster fading rate. We can see that even though the individual estimates may have a large range at times, the average comes out to be very close to the expected SNR. We see that the SNR estimator works well with PSK modulation. Figures 2.11 and 2.12 show the estimator's performance in QAM schemes.

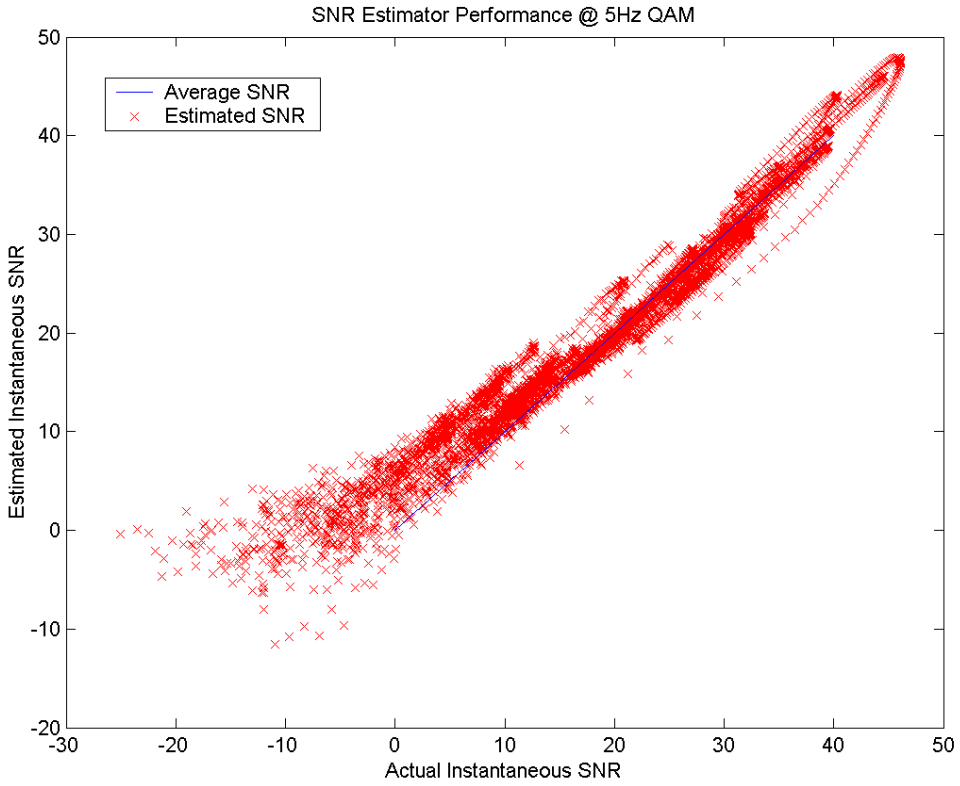

Figure 2.11 – SNR Estimation for QAM at 5Hz Doppler

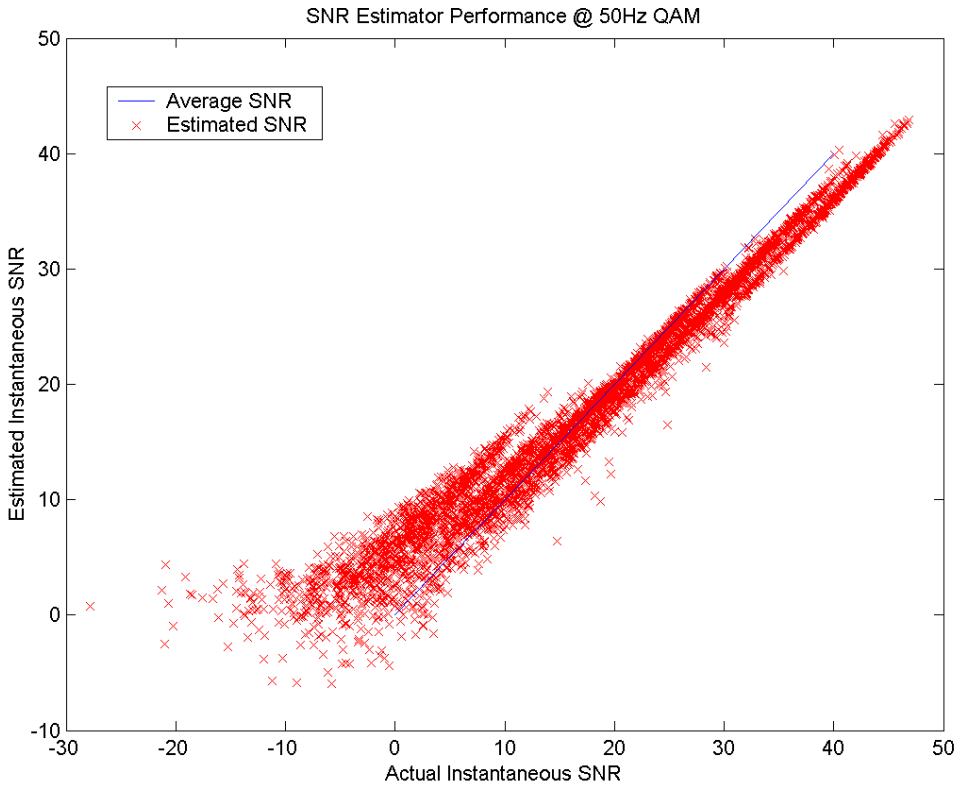

Figure 2.12 – SNR Estimation for QAM at 50Hz Doppler

As shown in Figures 2.11 and 2.12, the SNR estimator is not quite as good for QAM as it is for PSK. For high SNR, the estimates for SNR in QAM are underestimated in the high Doppler case as it is with PSK. For low SNR, the estimate is high. This is because at low SNR, there are a greater number of symbol errors that occur. Those symbols are the input to the estimator. The more reliable the symbol information is, the better the SNR estimate is for QAM schemes. This is why the estimates are better at high SNR. However, this will not prove to be a problem for adaptive systems as we will discuss in the next chapter.

2.6 System Block Diagram

In this section, we will present the block diagram of our adaptive modulation system.

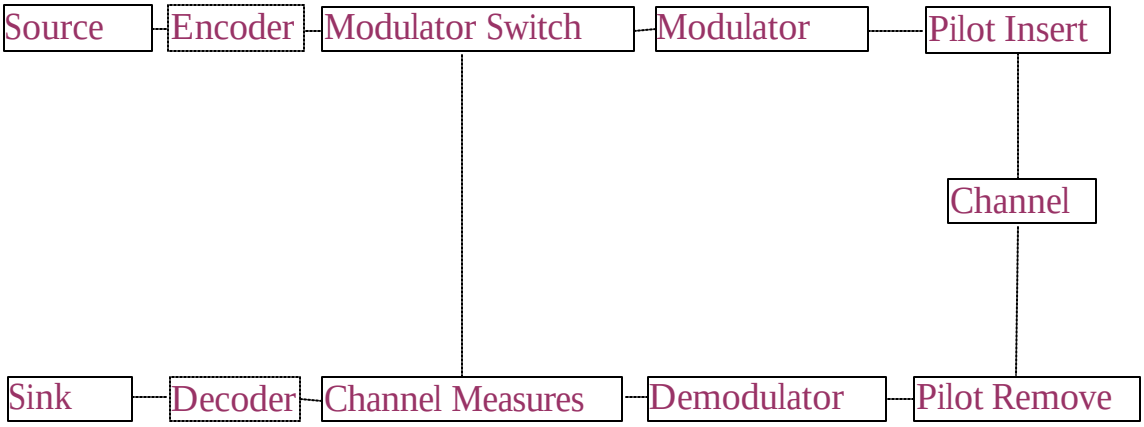

Figure 2.13 – System Block Diagram

Source – Here is where the information bits are produced.

Encoder – The information bits from the **Source** are taken and encoded through an FEC. (Chapter 4)

Modulator Switch – This is the logic that decides which modulation scheme to apply the bits from the **Encoder or Source**. It requires information from the **Channel Measures** block in order to function.

Modulator – Converts the binary data from the **Source** or coded symbols from the **Encoder** into complex sinusoids for transmission.

Pilot Insert – This is the stage where the pilot symbols are inserted into the information stream in order to use PSAM (In a

real system, these symbols would be added as bits before the information reached the **Modulator**. This was done in our simulation for convenience.)

Channel – Rayleigh fading and AWGN are applied to the transmitted signal.

Pilot Remove – At this point, the pilot symbols are removed and stored for later use in the **Channel Measures** block. Also, the channel is also removed from the signal.

Demodulator – Converts the newly compensated signal back into information bits/coded symbols.

Channel Measures – Takes information from the stored pilot symbols and determines the quality of the channel. In predictive adaptation, this is where channel prediction takes place. Information is then sent back to the **Channel Switch** for adaptation decisions.

Decoder – FEC is removed from the symbols. Information bits are the output and are sent to the **Sink**. (Chapter 4)

Sink – BER and spectral efficiency are calculated here when the frame or block is finished being received and demodulated/decoded.

Chapter 3

ADAPTIVE MODULATION

3.1 Introduction

In this chapter, we will introduce the main focus of this work, adaptive modulation. As stated earlier, adaptive modulation is a way to improve the tradeoff between spectral efficiency and bit error rate. We are able to make such optimizations in a Rayleigh channel by exploiting its fading dynamics. Periods of low fade, or high gain, will improve our instantaneous SNR, allowing higher rate modulation schemes to be employed with low probability of error. Periods of high fade will lower the effective SNR and force us to use low rate modulation in order to make transmission more robust.

Additionally, we will discuss ways of making adaptive modulation more effective by incorporating channel prediction into our system. We will investigate various receiver functions such as FFT channel estimation and SNR estimation, as well as transmission delays. We will begin this chapter with analysis in ideal conditions for our three modulation schemes: QPSK, 16QAM and 64QAM.

We will provide results comparing different system setups and investigate the advantages that channel prediction and SNR estimation can give us.

3.2 Adaptation Boundaries

The first topic that we must discuss is how to change the modulation scheme. In other words, we need a way for the system to decide which modulation scheme is best suited for the present (or in the case of delayed feedback - future) conditions. Pons and Dunlop [2] claimed that BER at the receiver would be a good channel metric to decide switching. However, we have decided to use the metric that Pons and Dunlop rejected, which is to estimate the SNR of the link. Reliable BER estimation is difficult over short periods and thus would restrict adaptation rate.

The question now becomes: How will we decide what ranges of SNR will be used for which modulation scheme? The answer lies in the AWGN performance of each modulation scheme.

Recall that we model our received signal as $r(t) = c(t) \cdot s(t) + n(t)$ (Eq. 2.5), where c(t) is the Rayleigh channel, s(t) is the transmitted signal, and n(t) is the noise signal. Recall also that SNR is signal power divided by noise power. In our system, we consider the signal power to be the power of the transmitted signal multiplied by the Rayleigh channel. This resultant signal power is the instantaneous received signal power and can be compared directly to the noise power, thus allowing us to consider the BER in an AWGN channel.

Let us now consider the BER performance of our 3 modulation schemes. From [1] we have equations for their probability of bit error:

$$P_{QPSK}(g) = Q(\sqrt{g}) \tag{3.1}$$

$$P_{16QAM}(g) = \frac{1}{4}\left[Q\left(\sqrt{\frac{g}{5}}\right) + Q\left(3\sqrt{\frac{g}{5}}\right)\right] + \frac{1}{2}Q\left(\sqrt{\frac{g}{5}}\right) \tag{3.2}$$

$$\begin{aligned} P_{64QAM}(g) = {} & \frac{1}{12}\left[Q\left(\sqrt{\frac{g}{21}}\right) + Q\left(3\sqrt{\frac{g}{21}}\right) + Q\left(5\sqrt{\frac{g}{21}}\right) + Q\left(7\sqrt{\frac{g}{21}}\right)\right] \\ & + \frac{1}{6}Q\left(\sqrt{\frac{g}{21}}\right) + \frac{1}{6}Q\left(3\sqrt{\frac{g}{21}}\right) + \frac{1}{12}Q\left(5\sqrt{\frac{g}{21}}\right) + \frac{1}{12}Q\left(7\sqrt{\frac{g}{21}}\right) \\ & + \frac{1}{3}Q\left(\sqrt{\frac{g}{21}}\right) + \frac{1}{4}Q\left(3\sqrt{\frac{g}{21}}\right) - \frac{1}{4}Q\left(5\sqrt{\frac{g}{21}}\right) - \frac{1}{6}Q\left(7\sqrt{\frac{g}{21}}\right) \\ & + \frac{1}{6}Q\left(9\sqrt{\frac{g}{21}}\right) + \frac{1}{12} + Q\left(11\sqrt{\frac{g}{21}}\right) - \frac{1}{12}Q\left(13\sqrt{\frac{g}{21}}\right) \end{aligned} \tag{3.3}$$

In Equations 3.1 – 3.3, g is the SNR, and Q(.) is the Q function,

$$Q(x) = \frac{1}{\sqrt{2p}}\int_{x}^{\infty} e^{-\frac{x^2}{2}}dx$$

Using the above theoretical calculation of BER for each scheme we have the following plot:

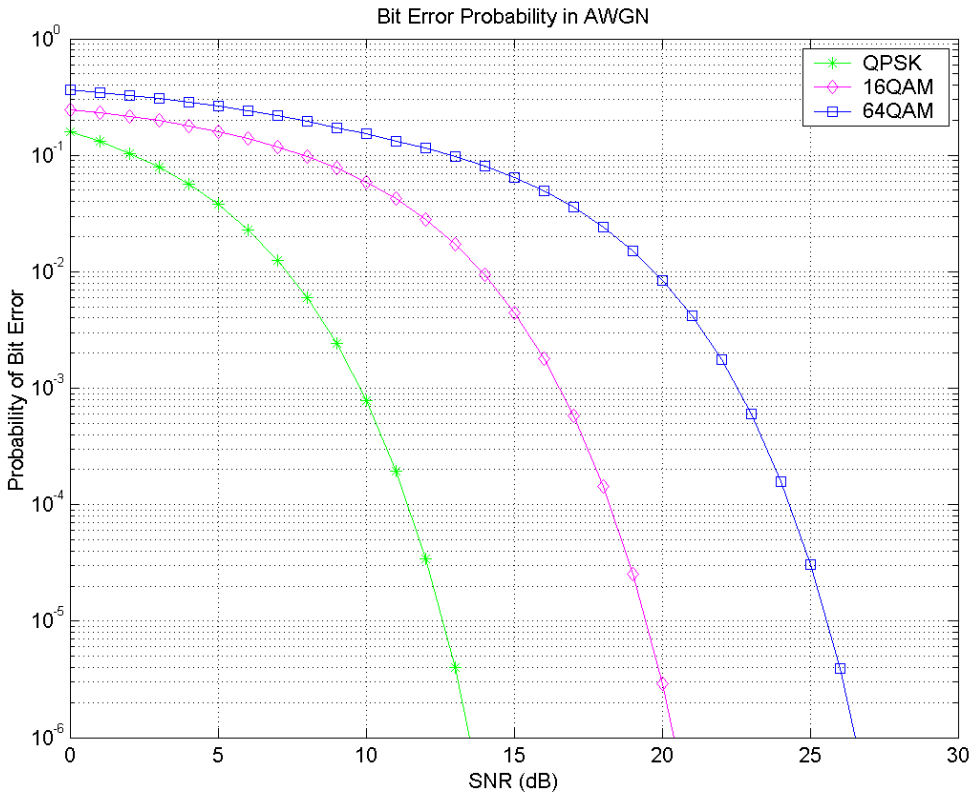

Figure 3. 1 - BER Performance in AWGN

In Figure 3.1, the curves from left to right represent the BER of QPSK, 16QAM and 64QAM in a strictly AWGN channel, respectively. In order to decide the proper switching levels from this plot, we have to decide what our operating point, or desired BER is. In our study, we chose to use 10^{-3} as our operating point. This means that our system will try and keep a BER lower than 10^{-3} with the most spectrally efficient modulation scheme whenever possible. At this point we should define spectral efficiency as the number of information bits encoded on a modulated transmission symbol. For example, QPSK has a spectral efficiency

of 2 bits per symbol, 16QAM has 4 bits per symbol, and 64QAM has 6 bits per symbol.

Therefore, with our operating point, and the given BER plots, we have the following SNR ranges for each modulation scheme:

QPSK	SNR < 17dB
16QAM	17dB <= SNR <= 23dB
64QAM	SNR > 23dB[1]

Table 3.1 – Modulation Schemes to SNR Range

We came to these levels in the following way: At an operating BER of 10^{-3}, there is no modulation scheme that gives us our desired performance at an SNR below 10dB. Therefore, we choose QPSK as it is the most robust. Between 10 and 17dB, there is only one scheme that gives us performance below 10^{-3}, and that is QPSK. Between 17 and 23dB, 16QAM gives us our desired BER at a better spectral efficiency than QPSK. And at SNR higher than 23dB, 64 QAM gives us the best spectral efficiency while providing the desired BER performance.

3.3 Adaptive Modulation

Now that we have established the means by which we switch modulation, we can discuss the theoretical performance of adaptive modulation, both in terms of BER and spectral

[1] In our simulations, 25dB was actually used. We do not expect this to change our results significantly as the ideal simulations fit the theoretical expectations.

efficiency. We shall refer to Torrance and Hanzo [1] for an analysis of adaptive modulation.

First, we have to define the PDF of the fluctuations of the received, instantaneous, Rayleigh amplitude, s. The envelope of a Rayleigh fading channel has a distribution of:

$$F(s,S) = \frac{2\sqrt{s}}{S} e^{\frac{-s}{S}} \tag{3.4}$$

where, S is the average signal power. Next, we need to determine the BER of each modulation scheme. We can analytically determine them by:

$$P_g(S/N) = \int_0^{\infty} P_G(s/N) \cdot F(s,S) ds \tag{3.5}$$

where P_g is the Rayleigh channel BER, and P_G is the BER performance in an AWGN channel. With the above two equations in addition to Equations 3.1-3.3, we can find the BER performance of adaptive modulation as:

$$P_A(S/N) = B^{-1} \cdot \begin{bmatrix} 2\int_{l_1}^{l_2} P_{QPSK}(s/N) \cdot F(s,S) ds \\ +4\int_{l_2}^{l_3} P_{16QAM}(s/N) \cdot F(s,S) ds \\ +6\int_{l_3}^{l_4} P_{64QAM}(s/N) \cdot F(s,S) ds \end{bmatrix} \tag{3.6}$$

where the l_i are the SNR thresholds between the modulation schemes and B is the average spectral efficiency. The values of

l_i can be inferred from Table 3.1. The value of B is computed as follows:

$$B = 2 \cdot \int_{l_1}^{l_2} F(s,S)ds + 4 \cdot \int_{l_2}^{l_3} F(s,S)ds + 6 \cdot \int_{l_3}^{l_4} F(s,S)ds \qquad (3.7)$$

Having established the mathematical foundation behind adaptive modulation, let us look at the results in graphical form.

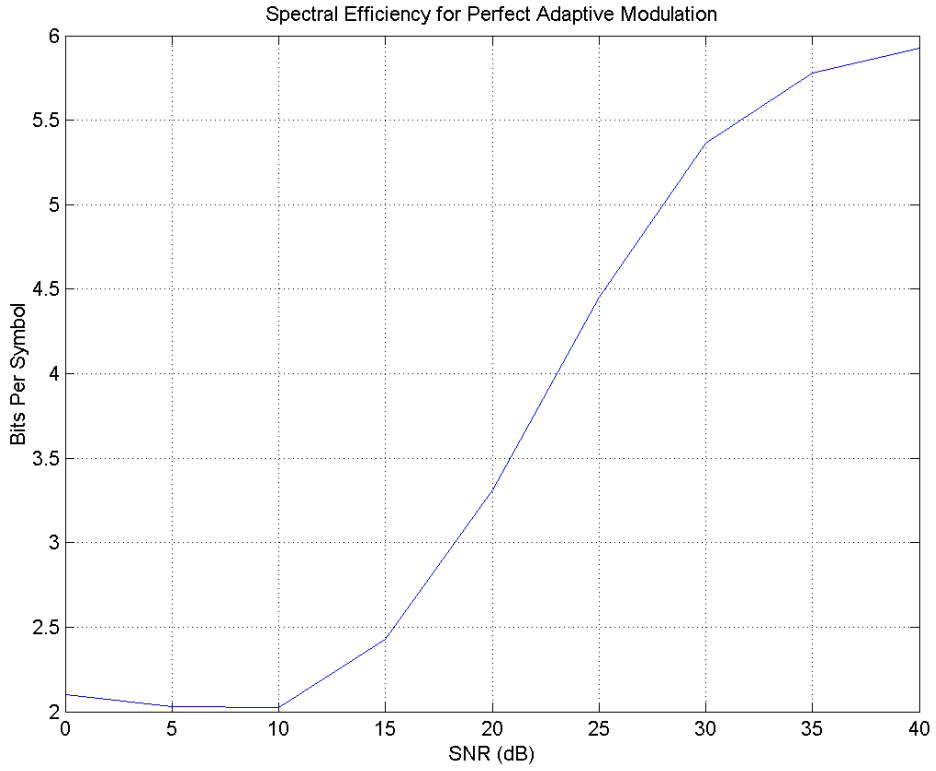

Figure 3. 2 – Spectral Efficiency for Perfect Adaptive Modulation vs. Average SNR for a Rayleigh Channel

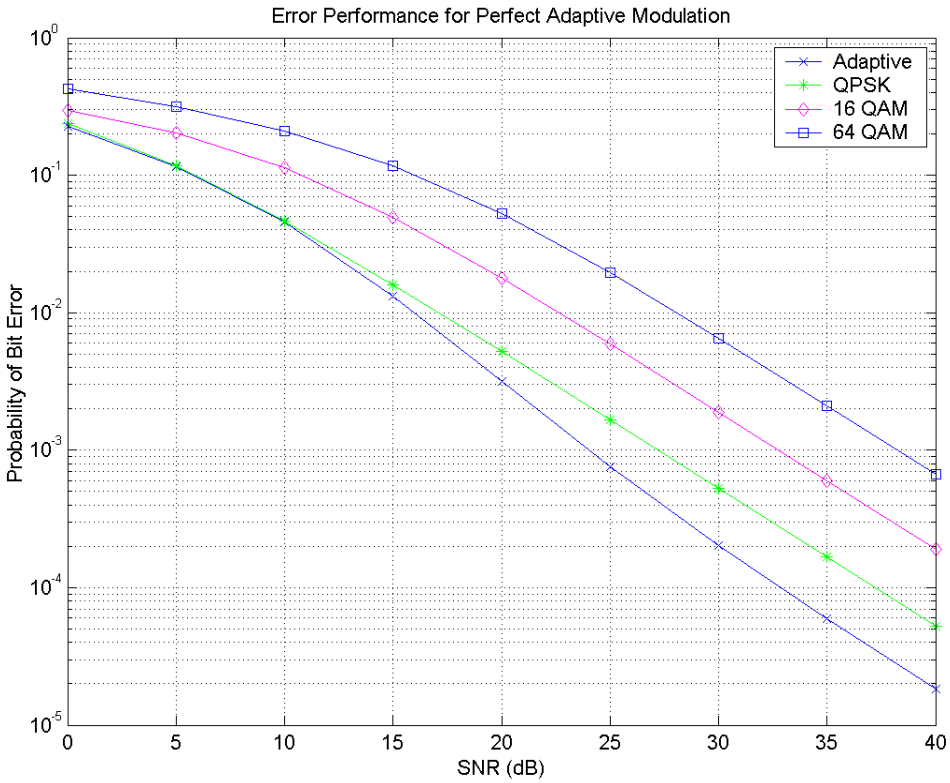

Figure 3. 3 – BER Performance for Perfect Adaptive Modulation vs. Average SNR for a Rayleigh Fading Channel

In Figure 3.2, we have a plot of the spectral efficiency of adaptive modulation versus average SNR in dB. Here, we should define spectral efficiency. In our system, we define spectral efficiency to be the number of bits sent per modulation symbol. We do not take into account whether or not the bits are the correct ones that were sent or not. In other words, we do not concern ourselves with 'goodput'. Because we have set the target BER to a value that we believe the system must operate under, the adaptation system will try to achieve that level of performance.

Note that at low SNR, the system achieves 2 bits per symbol, as QPSK is primarily used. However, as the SNR increases, the throughput also improves steadily, which indicates that we are beginning to use more spectrally efficient modulation schemes. The curve begins to level out at close to 30 dB, as 64QAM becomes the modulation scheme used most often and QPSK is rarely used. As SNR improves, the system is more able to choose more efficient modulation schemes.

The more interesting graphic is the one of Figure 3.3. In Figure 3.3, we see that the performance of adaptive modulation begins by overlapping the QPSK curve. It is analogous to the spectral efficiency curve, as QPSK is the primary modulation scheme used in low SNR. However, as the SNR is increased to 15 dB, we see an interesting result. The performance of adaptive modulation begins to improve beyond what QPSK can provide. Let us examine this result.

Consider a transmission that is encountering a deep fade. Our options here are to use one of three modulation schemes, which differ in spectral efficiency and robustness. If we consider the fading to be extremely deep, perhaps half of all bits will be in error. Here, it is advantageous to send fewer bits because the total number of errors will be decreased, which influences bit error rates much more than total number of bits sent. When the channel is not in a fade, then we want to send as many bits as we can. In this situation, we lower the BER by the increasing the number of bits sent because errors become less frequent.

It is the combination of these two principles that allows the BER performance of adaptive systems to be more robust than static

systems while simultaneously providing better spectral efficiency at most ranges of SNR.

In the next few sections, we will discuss some of the ways adaptive performance is degraded with imperfect system knowledge such as FFT channel estimation and frame delays.

3.4 Performance of Adaptive Modulation in Simulation

Before we discuss the performance of our system in imperfect environments, we should verify that our system simulator functions properly in ideal conditions.

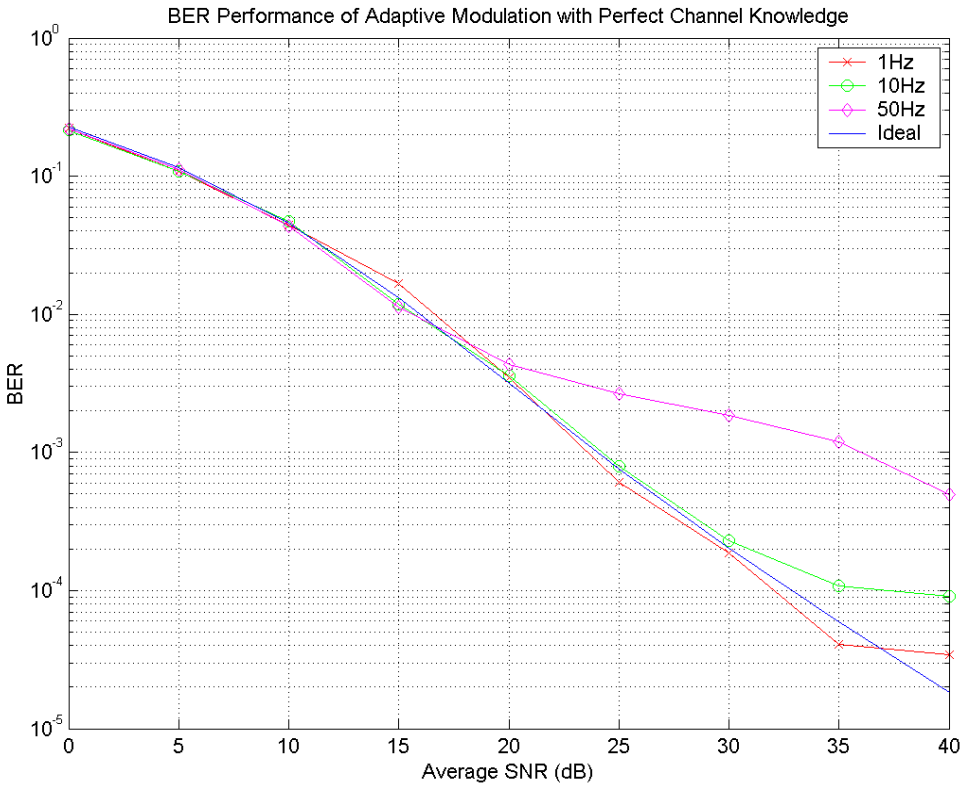

Figure 3. 4 – BER Performance of Adaptive Modulation for Different Doppler Rates

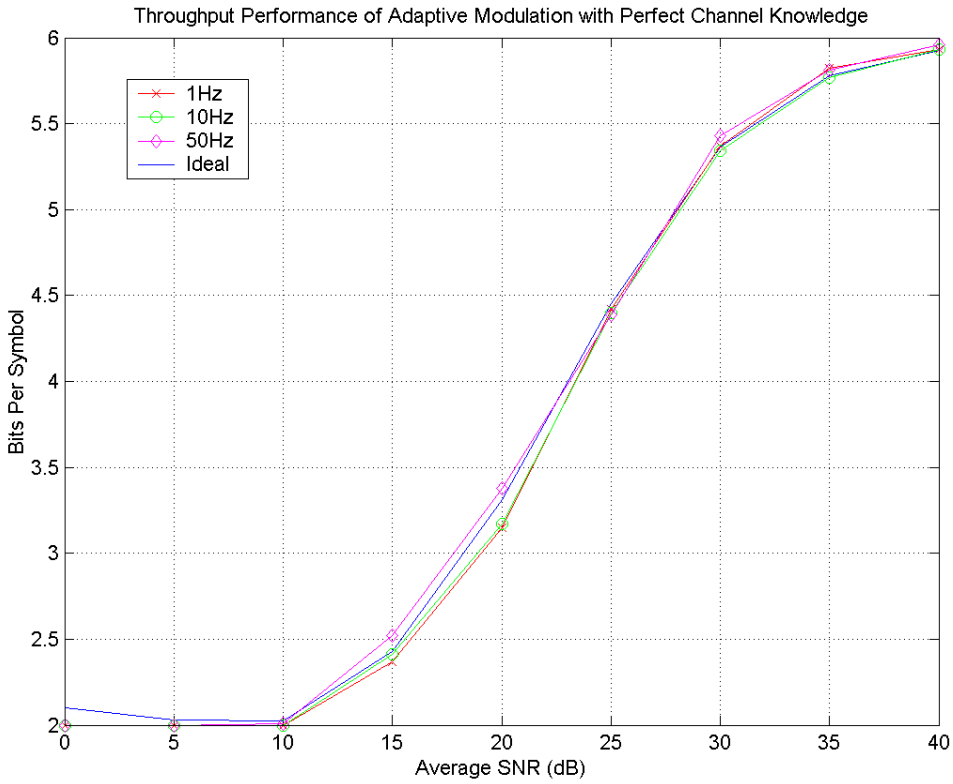

Figure 3. 5 – Spectral Efficiency of Adaptive Modulation

In Figures 3.4 and 3.5 we plot the simulated BER and spectral efficiency of adaptive modulation in situations where we have perfect knowledge of the channel with a frame rate of 640Hz. Note that with increasing Doppler frequencies, the BER degrades significantly. It can be seen that there is a 15dB loss in BER performance at a 50Hz Doppler rate. The reason is due to the more dynamic state of the channel. In the 1Hz situation, we have fairly constant channel conditions, whereas in the 50Hz situation, the power in the present frame may not be the same as the power in the next one. This fluctuation in the channel will

lead to bit errors since the chosen adaptation scheme may no longer be optimal. Additionally, as the Doppler rate increases, the SNR no longer remains static over an individual frame.

Presently, the system measures the immediate SNR and assumes that it will not change between frames. This, of course, is suboptimal. It is possible that we could have a frame with 17.1dB SNR and the following frame could have 16.7dB. If we assume that the next frame still has a 17.1dB SNR, we would be in error and the choice that the system makes would be suboptimal. Situations like this occur in channels of all Doppler, but it happens more frequently in cases of higher Doppler rates. If we could predict the state of the channel, we would be able to alleviate this problem.

When comparing different curves, we should find that Doppler frequency is not the sole factor which determines BER. It is actually a combination of Doppler frequency and the time duration of each frame, i.e. adaptation rate. In the above cases, a symbol rate of 96kHz was assumed, with each frame consisting of 150 symbols. This leads to a frame duration of 1.56ms per frame. With those two figures, we can define a normalized Doppler rate that will give us a more accurate comparison of different situations.

$$D = f_D \cdot T_f \tag{3.8}$$

This will be defined as the normalized Doppler rate. We will have similar channel qualities for systems with similar normalized Doppler. A system with high Doppler and short frames

will have a similar channel with a system that has low Doppler and very long frames times.

In the above cases, all the frame durations are the same, so only the Doppler is variable. The normalized Doppler for the 1, 10, and 50Hz Doppler rates are 1.56E-3, 15.6E-3 and 78.1E-3, respectively.

In Figure 3.5, we see the spectral efficiency of the adaptive scheme. For all Doppler rates, the simulated curves coincide well with the theoretical curves. The reason for this is that statistically, with perfect channel and SNR estimation, we should always have good spectral efficiency curves. The proportion of time that each of the modulation schemes are chosen is independent of Doppler rate. Note that good spectral efficiency does not necessarily indicate good BER performance.

3.4.1 Channel Estimate Impact on Adaptive Modulation

In this section, we will remove the assumption of ideal channel knowledge that we held in the previous section. Here, we should find that our performance curves become even more degraded. In the following plots, we incorporate FFT channel estimation and SNR estimation into our system. The effectiveness of FFT estimation is based on the pilot to information symbol ratio (See Equation 2.6). The more pilots used per frame, the higher Doppler rate our estimator can compensate for. In our simulations, we used 1 pilot per every 15 information symbols.

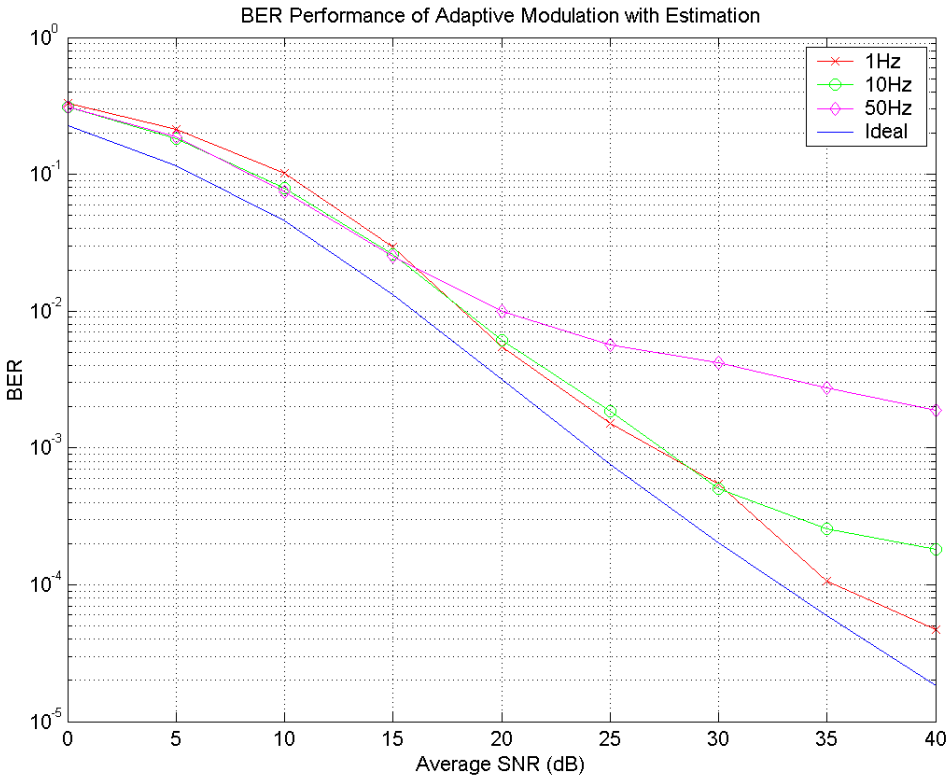

Figure 3. 6 – BER Performance of Adaptive Modulation in Estimated Environment

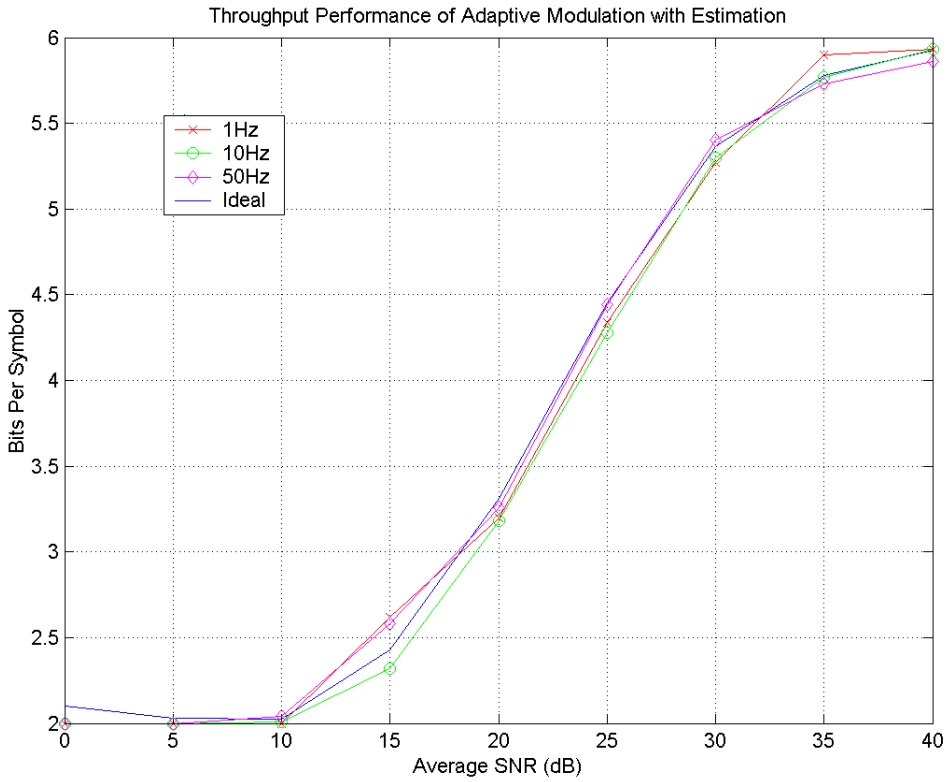

Figure 3. 7 – Spectral Efficiency of System with Estimated Environment

In Figure 3.7, we have a similar result for the spectral efficiency of the system. Since nothing has fundamentally changed the overall statistics of the estimated SNR, we should expect no change in the spectral efficiency. We have more interesting results in Figure 3.6. In the BER curves, we notice that each curve has been shifted to the right by approximately 2-3dB. This is due to the FFT estimation process. When we apply FFT estimation, there is an unavoidable loss of channel information that will degrade the de-rotation process.

Fortunately, this degradation is not severe, as we see only a 3dB loss of performance.

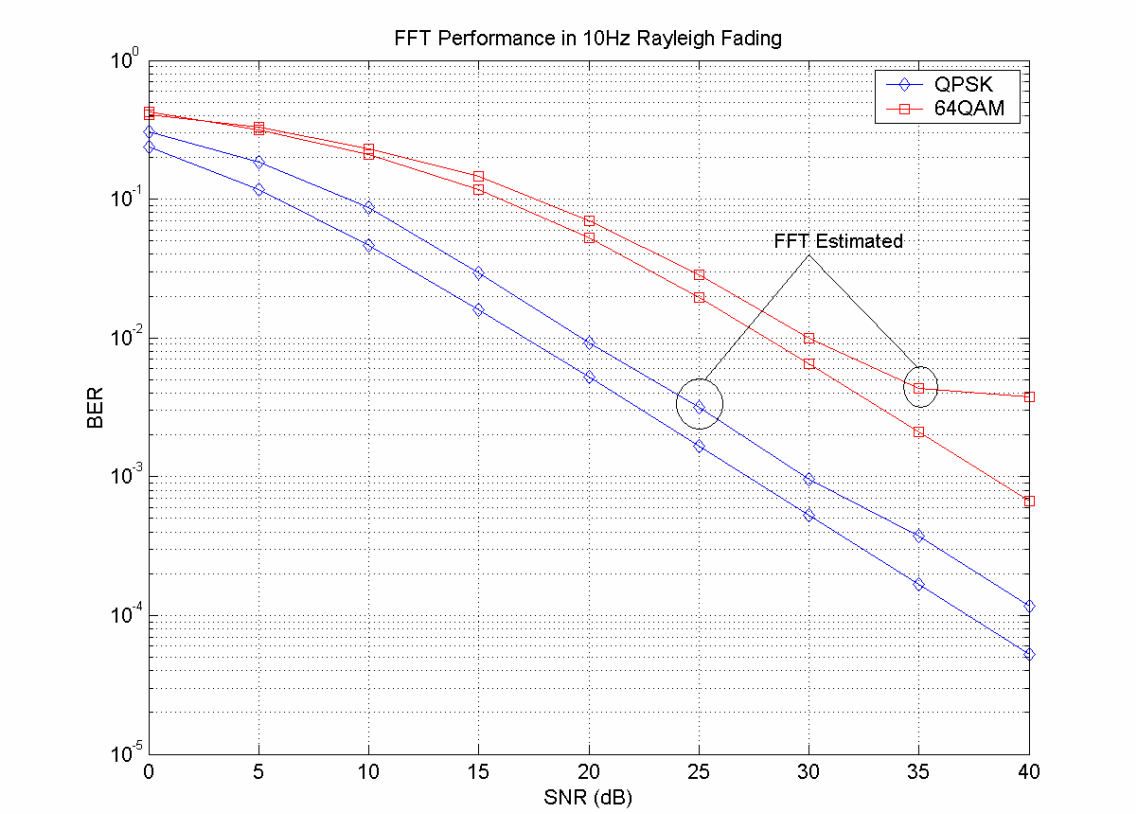

Figure 3. 8 – FFT Degradation in a Rayleigh Channel 10Hz Doppler Rate

In Figure 3.8, we have the theoretical BER performances of QPSK and 64QAM along with their FFT estimated counterparts. Note that there is approximately 3dB difference between the ideal and the actual plots. Also worth noting is the floor that the QAM scheme experiences at high SNR. This performance floor is due to channel estimation.

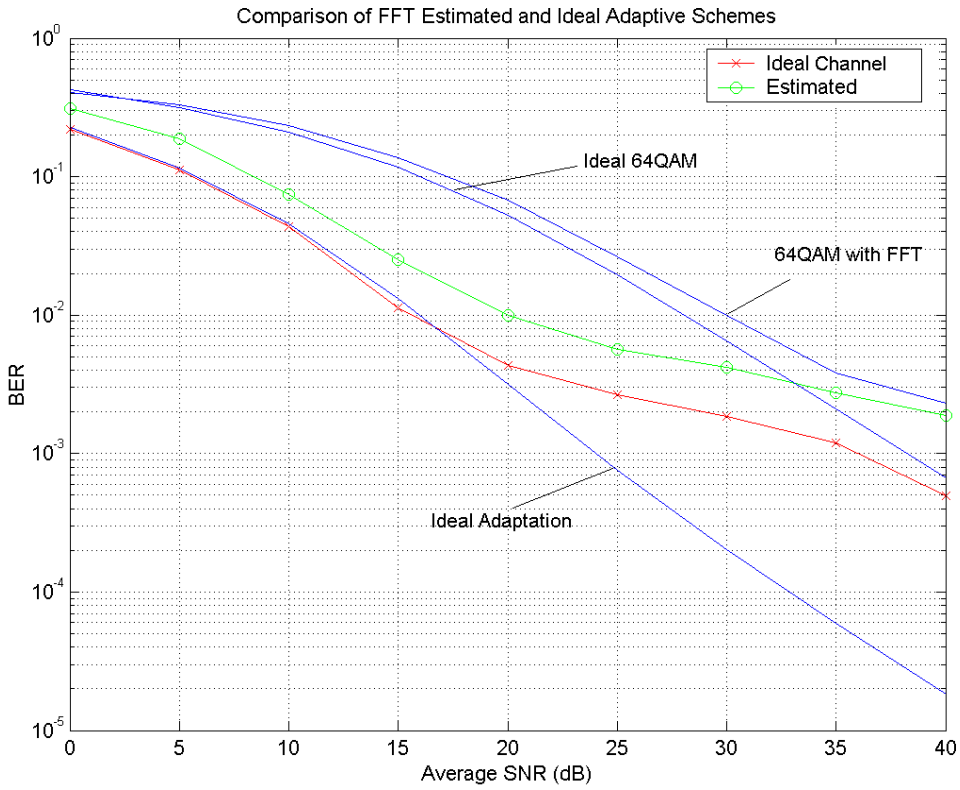

Figure 3. 9 – Comparison between Ideal and Estimated Adaptation at 50Hz Doppler

In Figure 3.9, we have a comparison between the ideal and estimated (SNR and FFT) adaptation systems for a 50Hz Doppler. The shift at low SNR is approximately 3dB, but ends much higher due to the estimated curve hitting a performance floor at high SNR.

Now that we have explored the effects that channel estimation has on our performance, we can move on to how propagation delays can hurt us.

3.4.2 Propagation Delay Impact on Adaptive Modulation

Thus far, we have assumed that there has been no time lag or propagation delay when the receiver relays control information back to the transmitter. In the system so far, there has been instantaneous relay between receiver and transmitter. Now, we will introduce another system imperfection into our simulations. We will introduce a delay in our system that will amount to two frames worth of time. In other words, from the time the receiver transmits information back to the transmitter, two more frames will be on route to the receiver from the transmitter and are not privy to the latest information most recently sent from the receiver.

When we incorporate this factor, the resulting performance with perfect estimation is given in Figure 3.10.

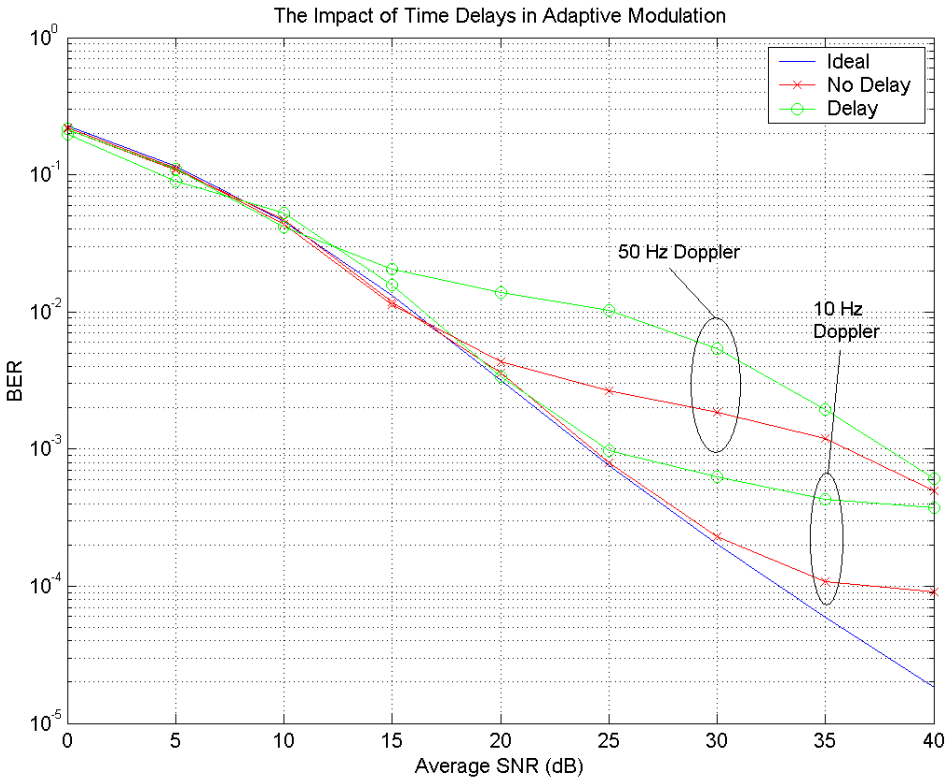

Figure 3. 10 – Comparison of System with No Delay and with Delay

In Figure 3.10, we see a plot comparing the effect that propagation delay has on our performance. The two lower curves represent the system performance at 10 and 50Hz (Normalized Doppler rates 15.6E-3, 78.1E-3 respectively) with no frame delay. The two upper curves represent the system performance at 10 and 50Hz with a 2 frame delay. In this simulation, we see how each ideal curve is degraded by the delay. They seem to be breaking off from their ideal counterparts 5 to 10dB earlier. The explanation is due to the changing nature of the channel. If we have a delay in our system, our SNR estimate will be out-dated.

This is why the performance with 50Hz Doppler shift suffers more from the delay and sooner. With higher Doppler shifts, we have faster variations in the channel. This explains why high Doppler curves show the negative effects of propagation delays sooner than their low Doppler counterparts.

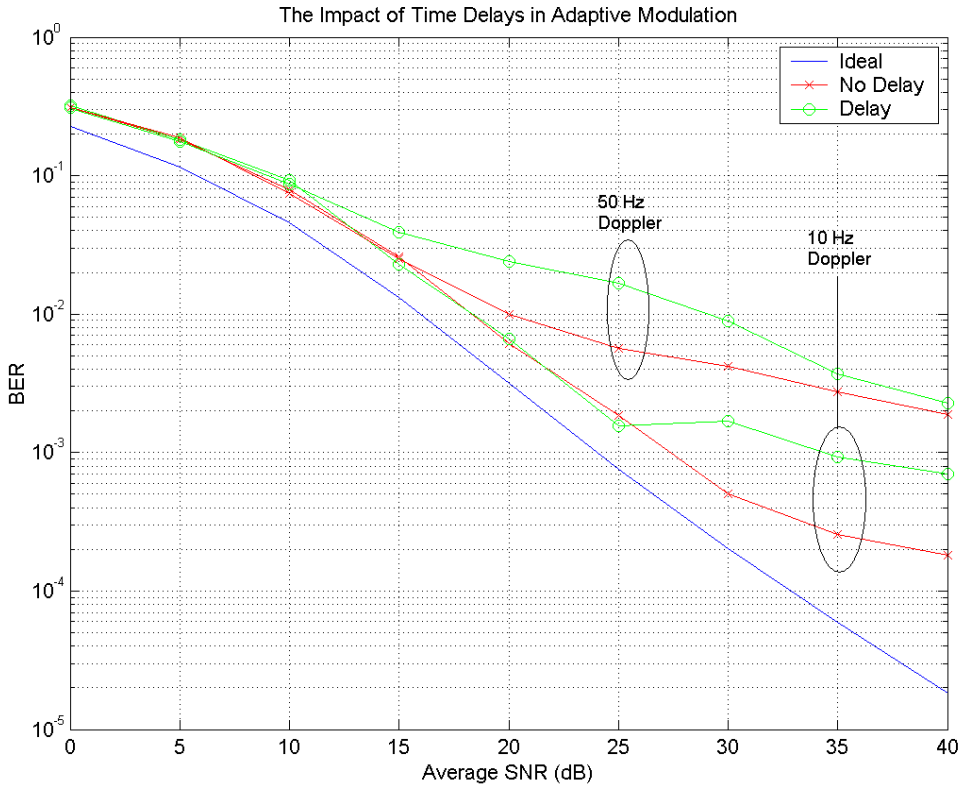

Figure 3. 11 – Delay Impact on System with Channel Estimation

In Figure 3.11, we see the effects that frame delays have on the system with channel estimation. With the channel estimation, the performance curves are simply shifted as they were in the previous section. With delays and the various estimators in

play, we have a performance degradation of 15-20dB for 10Hz and 50Hz Doppler at high SNR values.

However for very low Doppler frequencies, the problem is not so severe.

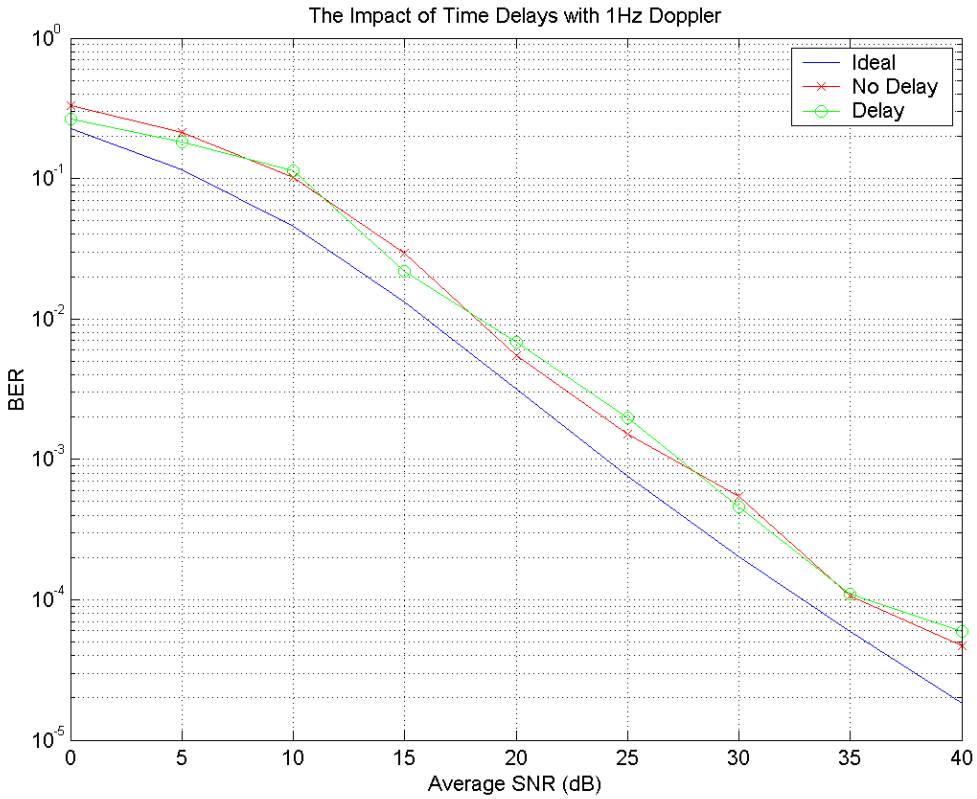

Figure 3. 12 – BER Performance of System with 1Hz Doppler

In Figure 3.12, we see almost no difference between systems with delay and no delay. It would seem that with when the channel is varying sufficiently slowly, propagation delay will not have a significant effect on error rates.

Let us mention briefly the effect on the spectral efficiency that frame delays may cause.

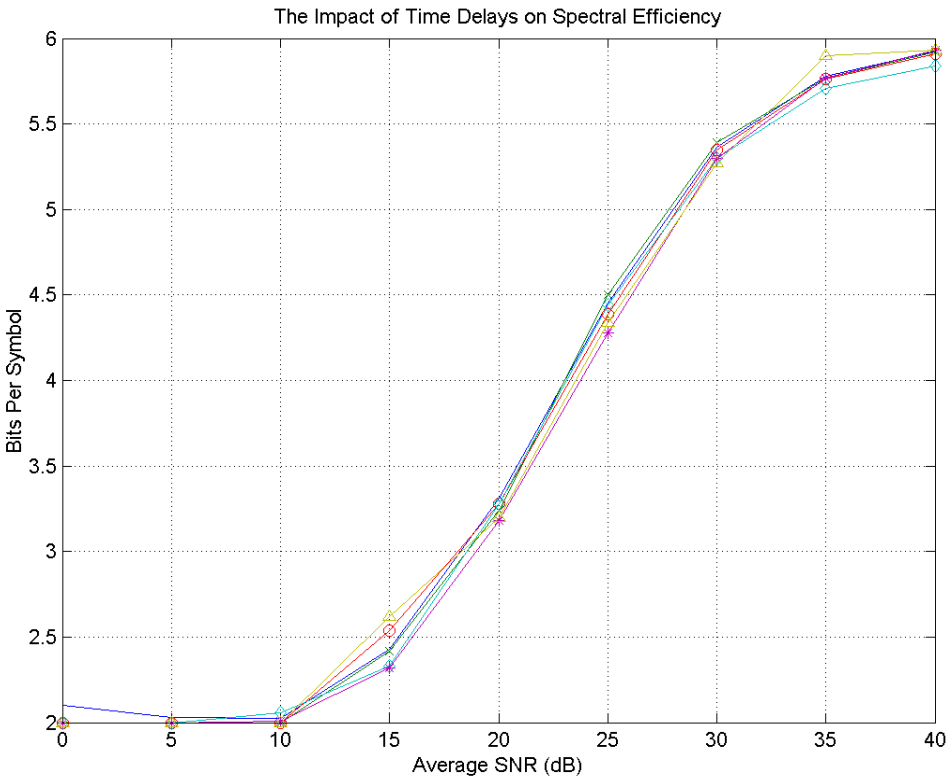

Figure 3. 13 – Spectral Efficiency of Systems with Delay

As Figure 3.13 shows, frame delays do not reduce throughput. The delay will only alter which modulation schemes are used, but will not change the overall proportions used; and the proportions of modulation schemes used are the measure of spectral efficiency. This leads us to conclude that the measure of spectral efficiency is not a good indicator of how well the BER will perform.

In summary of the estimations and delays, they both make negative impacts on the system with respect to ideal performance. For channel estimation, the system has a worse BER no matter what the Doppler shift. However, frame delays hurt only those systems which operate in medium to high Doppler.

The main source of these BER increases is the improper selection of modulation schemes. The system makes these errors because it does not have timely information on the channel. Thus, we propose the use of channel prediction with adaptive modulation.

3.5 Channel Prediction

The idea behind channel prediction is to use past and present channel samples to predict future samples. We will implement a prediction scheme for the specific purpose of anticipating the future power level of the Rayleigh channel. The reasoning behind our desire to predict the future of the Rayleigh channel is that the results of the previous section showed that the old knowledge of the channel is the main cause of degradation with adaptive modulation.

With prediction, we can estimate what the future power level of the channel will be, so that propagation delay will be less of a problem as we will know what the state of the channel will be by the time the transmitter receives the control information from the receiver. Before we present findings for the performance of the system with channel prediction, let us first review the theory behind it.

3.5.1 Linear Prediction

There has been significant work done in the field of linear prediction applied to channel prediction. Eyceoz and Duel-Hallen [15] is one example.

Recall Equation 2.4:

$$c(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} A_i \cdot e^{j(2p f_i t + f_i)}$$

Notice that the channel is dependent on the phases and frequency of each sinusoid only (we have normalized all scatterers to unit power, $A_i = 1$). With this information, we know that the channel is correlated from sample to sample, unlike AWGN. Therefore, we can take advantage of the deterministic properties and predict what the value of the channel will be at a later time.

In order to predict the Rayleigh channel, we will employ spectral estimation according to the Maximum Entropy Method (MEM) followed by linear prediction [15]. The purpose of MEM is to produce a set of coefficients, or poles, to use as input to a linear predictor.

Using MEM, we have the frequency response of the channel modeled by the following:

$$H(z) = \frac{1}{1 - \sum_{j=1}^{p} d_j z^j} \qquad (3.9)$$

In Equation 3.9, the d_j's are the linear prediction coefficients, used by the linear predictor, with there being p coefficients.

The predictor in our case is a simple multiply-and-sum function.

$$c_n' = \sum_{j=1}^{p} d_j c_{n-j} \tag{3.10}$$

In summary, c_n' is a predicted value based on the linear combination of p previous values multiplied by the prediction coefficients. To predict multiple samples in the future, we just treat the latest predicted sample as an actual sample. We can then predict as far as we desire. But of course, the farther we predict, the less accuracy we have since error will accumulate.

Press [16] discusses linear prediction in a different manner. He determines the prediction coefficients through an autocorrelation method. Assuming stationarity of the signal, the autocorrelation can be found as:

$$f_j \equiv \left\langle y_i y_{i+j} \right\rangle \approx \frac{1}{N-j} \sum_{i=1}^{N-j} y_i y_{i+j} \tag{3.11}$$

Using this result, the prediction coefficients can be found through the equation:

$$\sum_{j=1}^{M} f_{|j-k|} d_j = f_k \quad k = (1, \cdots, M) \tag{3.12}$$

where M is the number of poles to be calculated, decided by the user; the more poles, the more accurate the prediction. Of

course, we cannot have more poles than we have channel samples. In our simulations, we used 10 poles with 100 samples of channel information for most of our results. We avoided using many poles, as they increased simulation time significantly.

In actual implementation however, neither of these methods to calculate the prediction coefficients were followed explicitly. Press concedes that the above method is not very good at calculating the autocorrelation figures due to the sensitivity that linear prediction has with the autocorrelation figures. Instead we used a recursive algorithm provided by Press [16] which uses a recursive method to calculate f_k. In any event, this linear prediction algorithm is good for signals that are smooth and oscillatory, which would describe Rayleigh fading at least somewhat accurately.

3.5.2 Channel Prediction Tests

Now that we have established the theory behind our method of linear prediction, we will now demonstrate the effectiveness of it in trial runs.

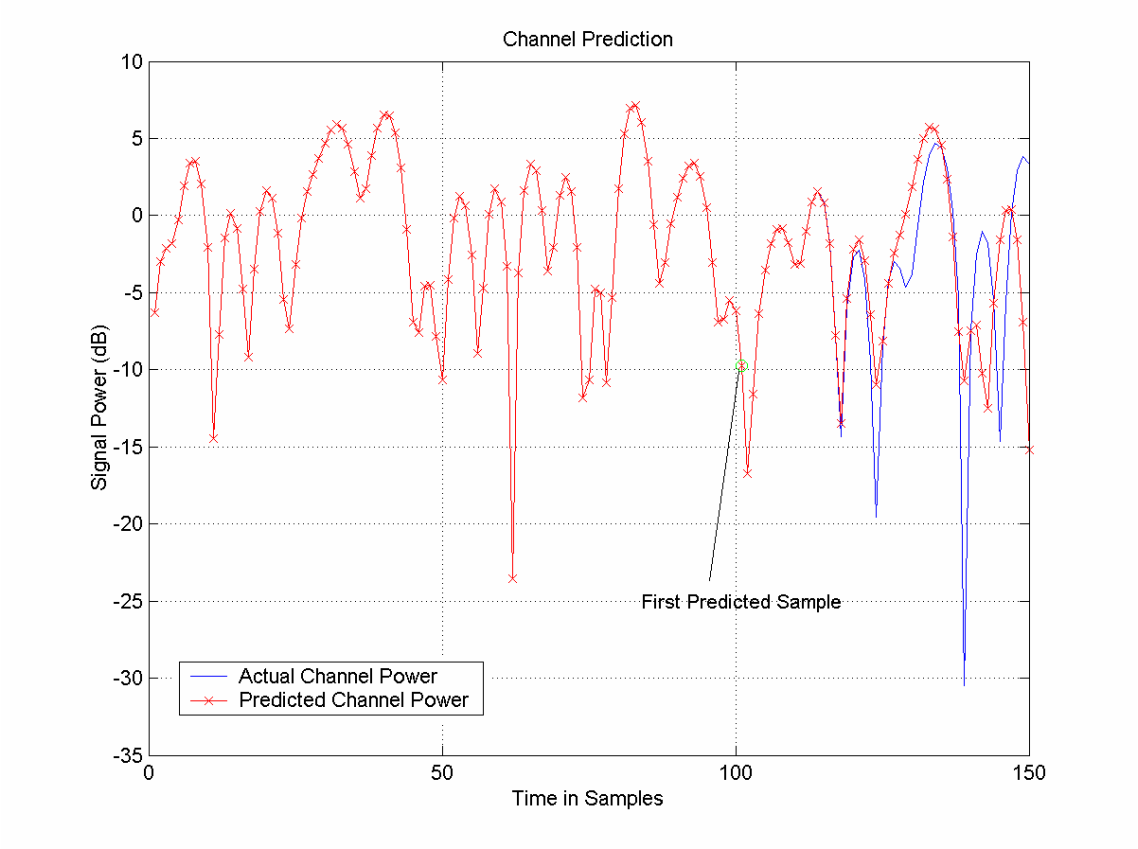

Figure 3. 14 – Channel Prediction Demonstration
(See Table 3.2 for System Parameters)

Figure 3.14 shows the ability of linear prediction in Rayleigh channels. In this example, we used the following specifications to the system:

Doppler Frequency	50Hz
Sampling Frequency	500Hz
Number of Poles	50
Number of Initial Samples	100
Number of Predicted Samples	50

Table 3. 2 – Channel Prediction Specifications

In Table 3.2, we declare various channel specifications for Figure 3.14. With a sampling frequency of 500Hz and 100 initial samples, we have 0.2 seconds of information. We then try to predict the following 50 samples, which amounts to 0.1 seconds of received signal. The figure here shows us that the predictions are good for the first half of the prediction period, or about 0.05 seconds, 25 samples.

This is a challenging scenario for channel prediction. The normalized Doppler (F_d/T_s) rate in this situation is 50Hz/500Hz, which is 0.1. In our simulations, we will assume a normalized Doppler on the order of $8x10^{-3}$. Also, we will not have so many channel coefficients at our disposal, as they take a lot of processing to calculate. We will be using either 10 or 30 poles in our systems. This will be sufficient, as we will show in the following plot:

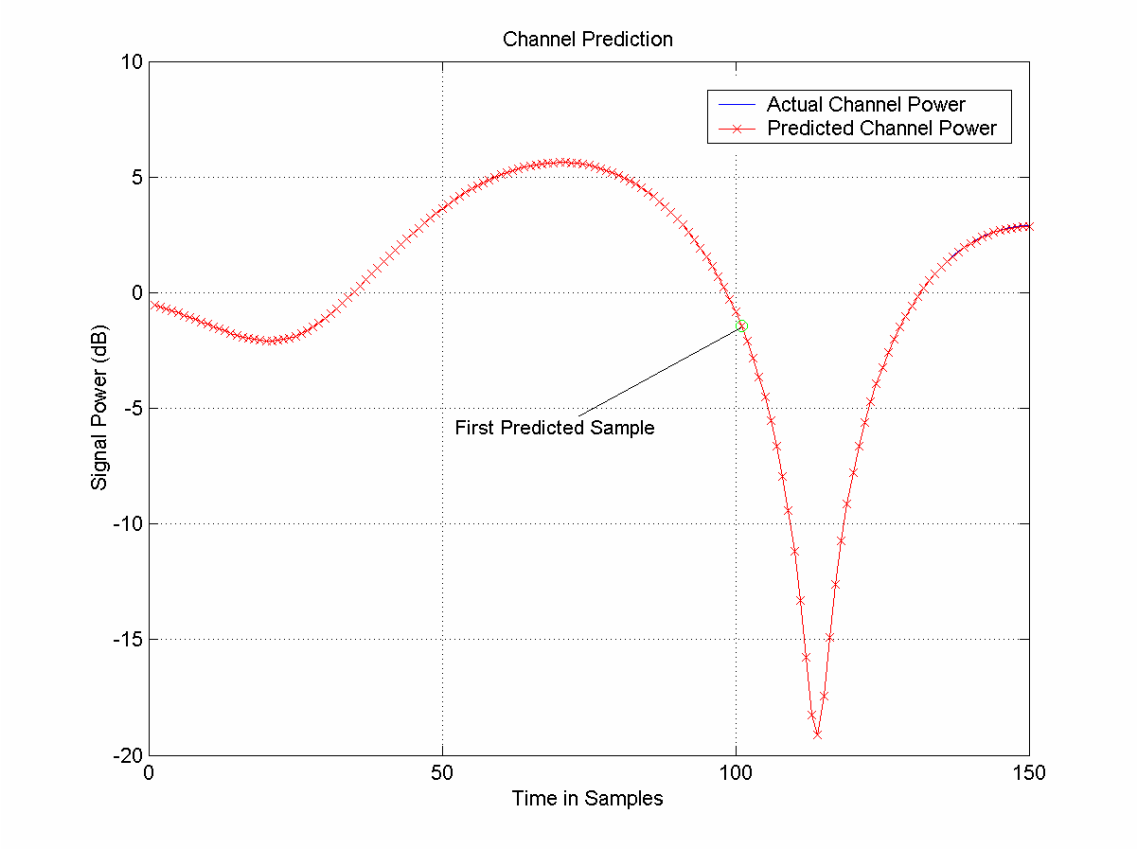

Figure 3. 15 – Channel Prediction with Higher Sampling Rate

In Figure 3.15, we have a plot of channel prediction working in a more realistic case in terms of the situations in which we will run. Clearly, we have much better performance when we have increased the sampling rate to 6000 samples per second (normalized Doppler of 8.3×10^{-3}) while keeping the same number of poles and initial samples. This shows us that the lower the normalized Doppler, the better the performance of the channel predictor. However, in a real simulation environment, we will not have a perfect curve to base our predictions. Instead, we will have corrupted estimates, both from the AWGN noise and the

FFT estimator. To see the viability of our predictor in non-ideal conditions, we will add noise to the initial samples.

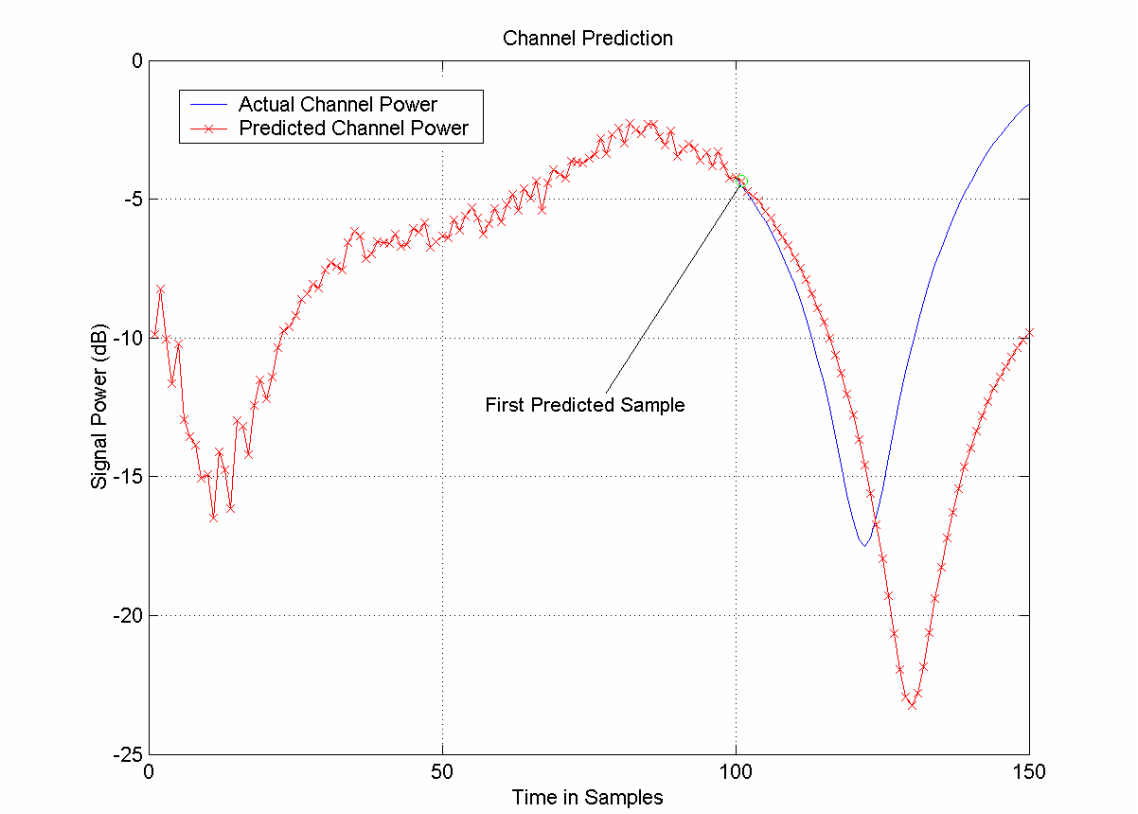

Figure 3. 16 – Prediction of Signal Power with 30dB SNR

Because we do not know precisely the nature of the corruption of the channel estimate we will approximate it by simply adding AWGN. Shown in Figure 3.16 is a demonstration of prediction in 30dB SNR. We can see that with noise, we have very poor prediction capabilities. However, we will later show that even this level of prediction prowess is sufficient to give us an advantage over standard adaptive modulation.

3.5.3 Channel Prediction in Adaptive Modulation

In this section, we will demonstrate the advantages that we will have when we implement channel prediction with adaptive modulation. As we have said earlier, the objective of channel prediction is to offset the delays encountered when we use adaptive modulation in a real system.

When we apply the prediction algorithm in the cases with frame delays, we hope to see the results match the results for the ideal case. But before we start making comparisons between prediction and non-prediction, let us first verify that our model is valid under ideal present conditions.

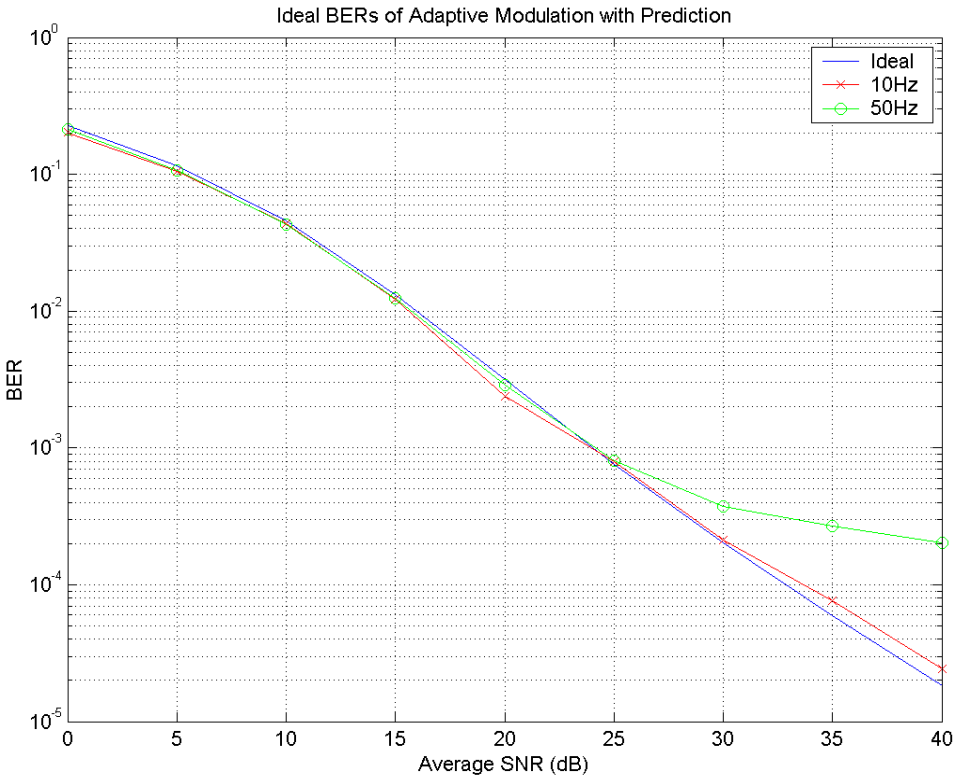

Figure 3. 17 – Ideal Performance of Adaptive Modulation with Prediction

Figure 3.17 shows us the performance of adaptive modulation with channel prediction in an ideal environment with no delay. SNR and channel compensation are also assumed. The prediction specifications include the use of 10 poles and a look-back depth of 100 pilot symbols. Already obvious to us is the fact that the 10Hz Doppler curve coincides with the ideal performance, much like the non-prediction scheme with a Doppler frequency of 1Hz. This is very promising performance. On the other hand, prediction does not seem to have the same effect on the 50Hz

Doppler curve. However, we will see it is much better than not using prediction.

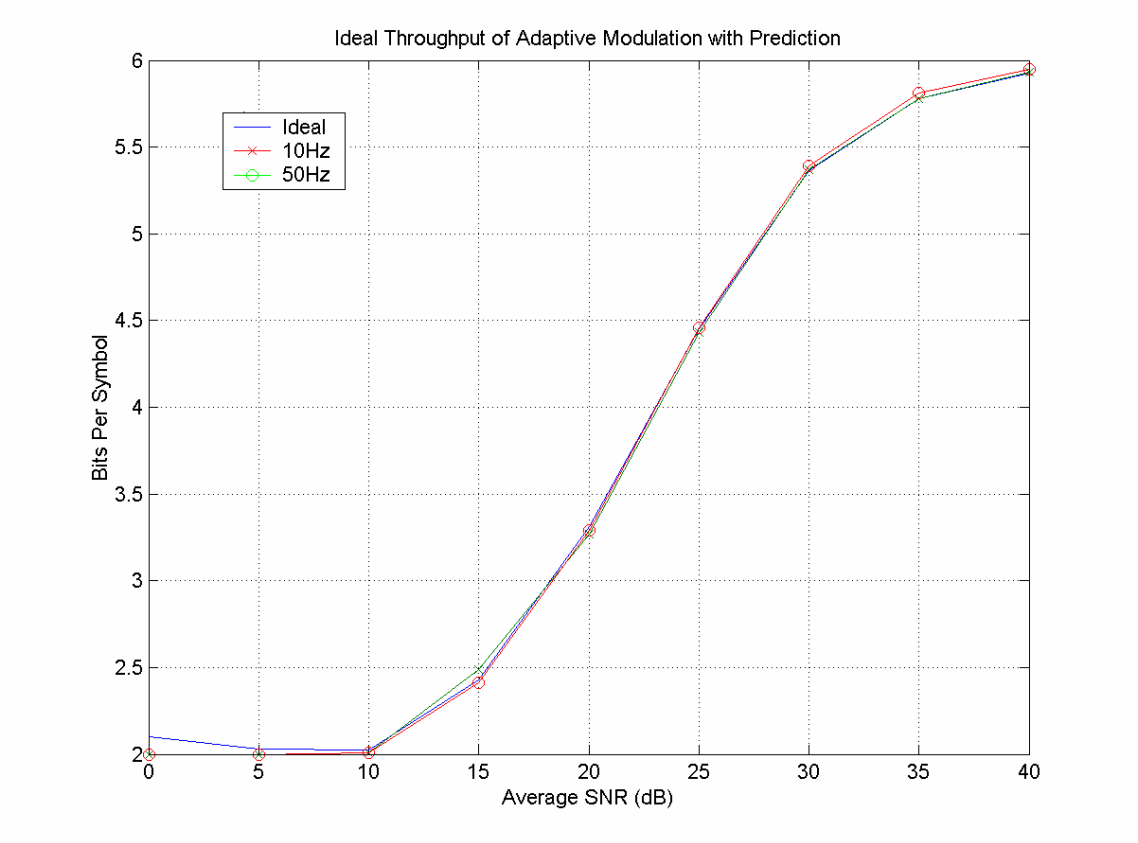

Figure 3. 18 – Ideal Throughput Performance with Prediction with Varying Doppler Rates

The spectral efficiency of adaptive modulation with prediction also coincides with ideal performance as shown in Figure 3.18. This result is what we would expect.

It is instructive to see how the Doppler rate affects the accuracy of the channel predictor. In the following figures, we present charts showing the accuracy of the predictions at

different Doppler rates and average SNR conditions. What we are predicted vs. actual Rayleigh channel power samples.

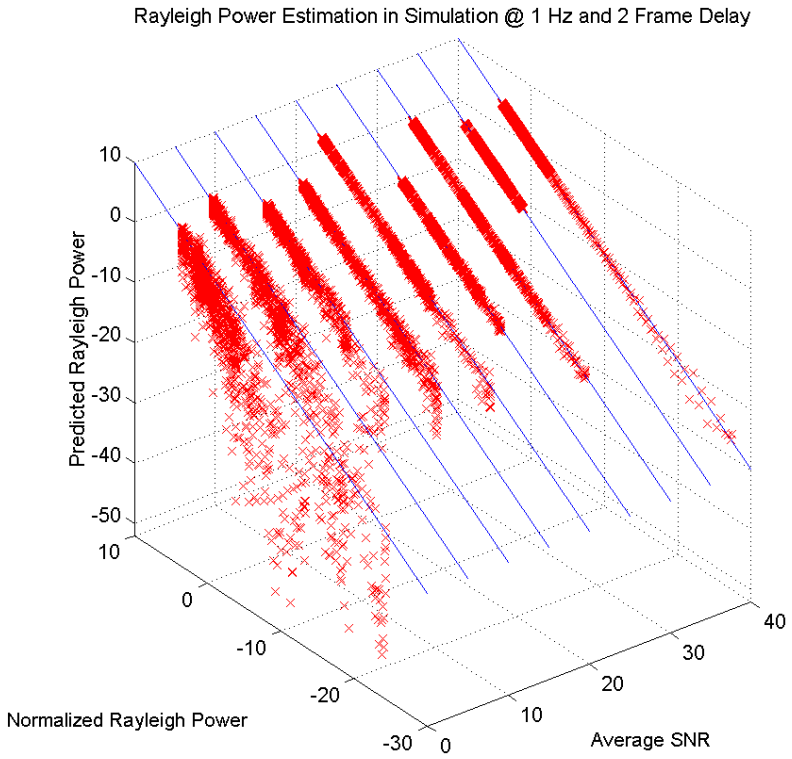

Figure 3. 19 – Prediction Accuracy of Channel Power in 1Hz Doppler

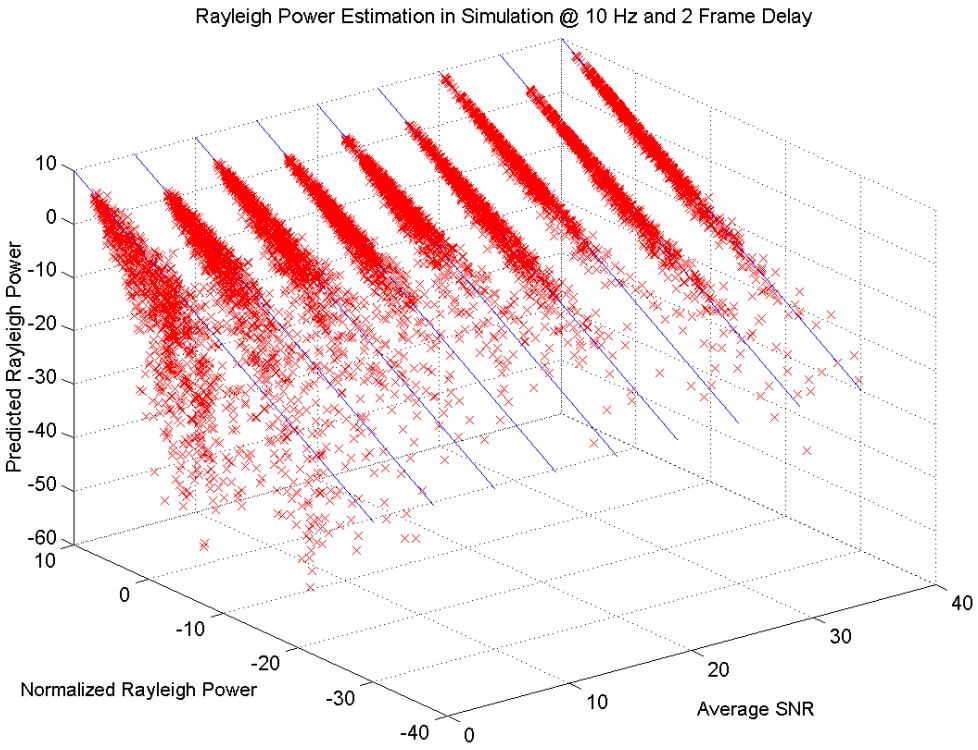

Figure 3. 20 - Prediction Accuracy of Channel Power in 10Hz Doppler

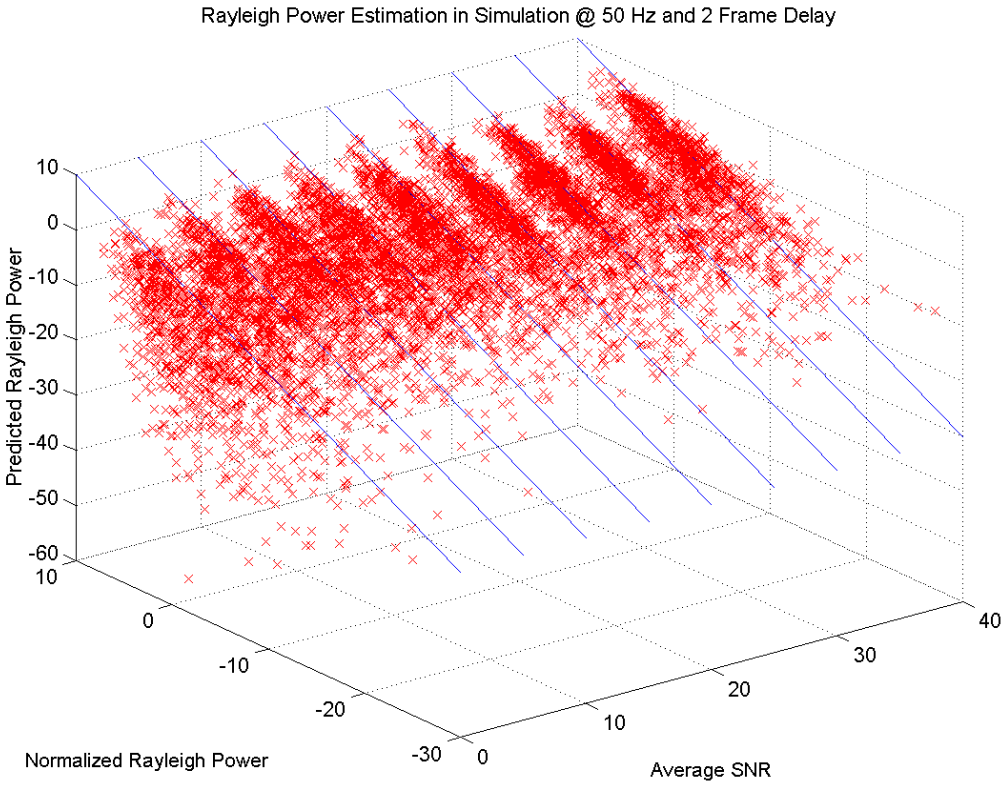

Figure 3. 21 - Prediction Accuracy of Channel Power in 50Hz Doppler

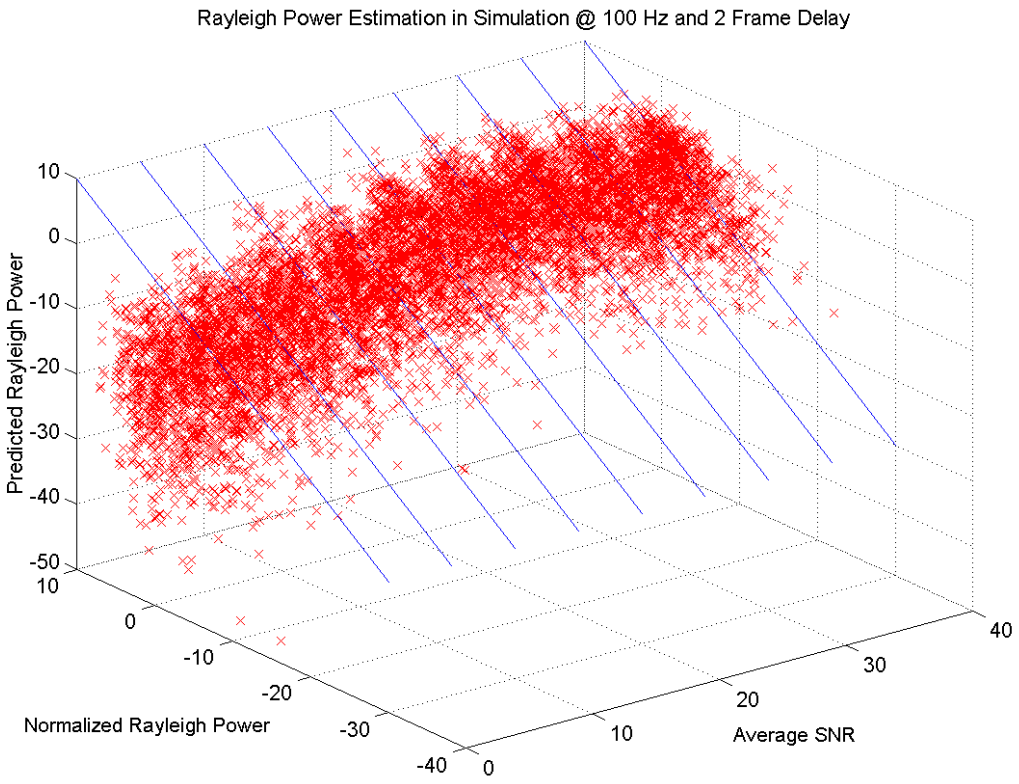

Figure 3. 22 - Prediction Accuracy of Channel Power in 100Hz Doppler

Figures 3.19 – 3.22 show the performance of the prediction of the Rayleigh channel in terms of the strength of the channel with a 2 frame propagation delay. All of these plots incorporate FFT and SNR estimation. Notice that the precision and accuracy increase directly proportional to average SNR and inversely proportional with the Doppler frequency. It is also interesting to note that with the increase in Doppler, the power estimates seem to be underestimated. This is most obvious in the 100Hz Doppler case.

We will see later that this will pose a liability in later simulations.

Let us now examine the performance when we include frame delays with prediction. We will see if channel prediction can do what we intend for it to do when we provide perfect channel samples.

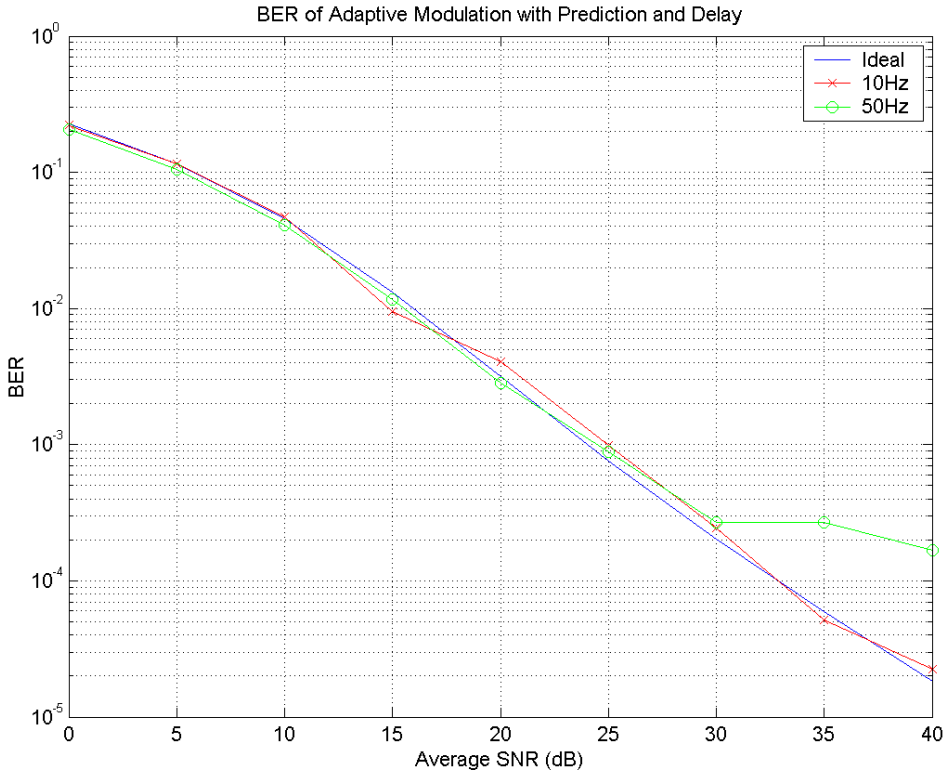

Figure 3. 23 – Prediction Performance with Delay

With no FFT estimation, the resulting performance is given in Figure 3.23. It is approximately the same performance that we obtain with no frame delays. We still have ideal performance with slower Doppler and ideal performance with 50Hz Doppler with

average SNR in the range of 0-30dB. Even the performance at 35 and 40dB are not far removed from the no-delay version.

Now that we are confident with prediction performance in ideal conditions, we should now compare it with non-prediction in non-ideal conditions.

3.5.4 Prediction vs. Non-Prediction in Non-Ideal Conditions

We should at this point compare the results from the non-prediction with those from prediction. We should expect the BER results from the predictive scheme to be superior to the ones from the non-predictive ones in every aspect; the spectral efficiency results from the predictive simulations should be the same as the ones from the non-predictive ones.

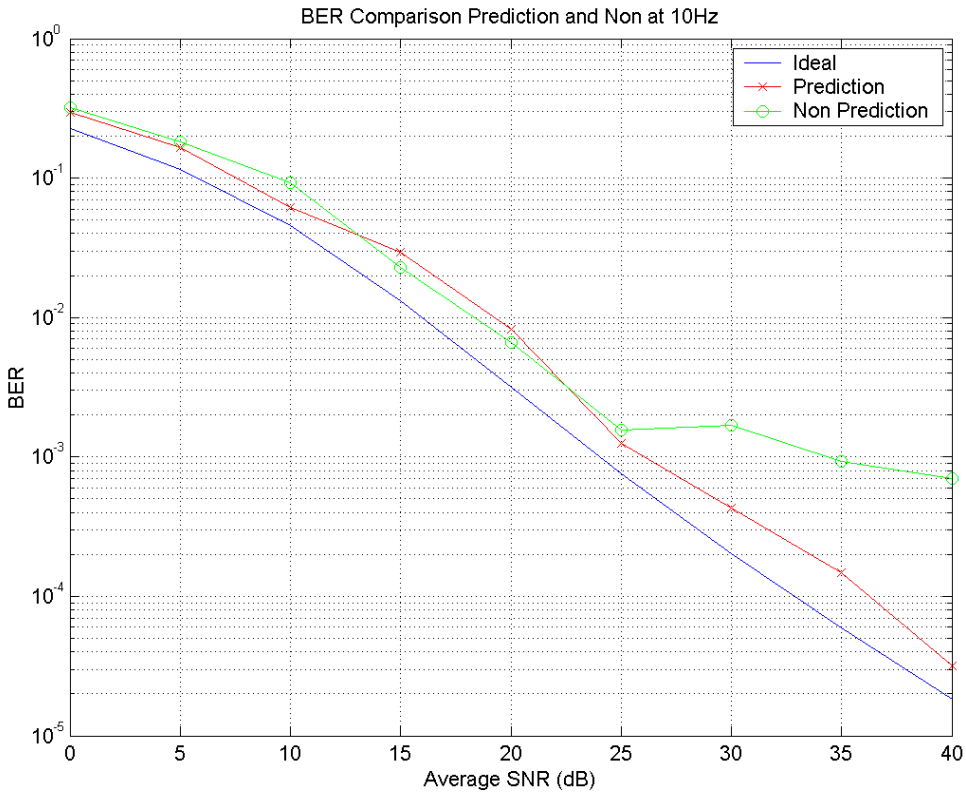

Figure 3. 24 – BER of Prediction and Non-Prediction @ 10Hz Doppler

In Figure 3.24, we should note first that the two approaches have about the same performance in the range of 0-25dB, at which point they deviate from one another. The non-prediction curve seems to bottom out in the BER neighborhood of 10^{-3}, while the predictive curve maintains performance that parallels the ideal case. This is also similar to the non-predictive scheme in 1Hz Doppler, which indicates that the predictive scheme is doing well in this Doppler. At high SNR, there are very large gains with channel prediction.

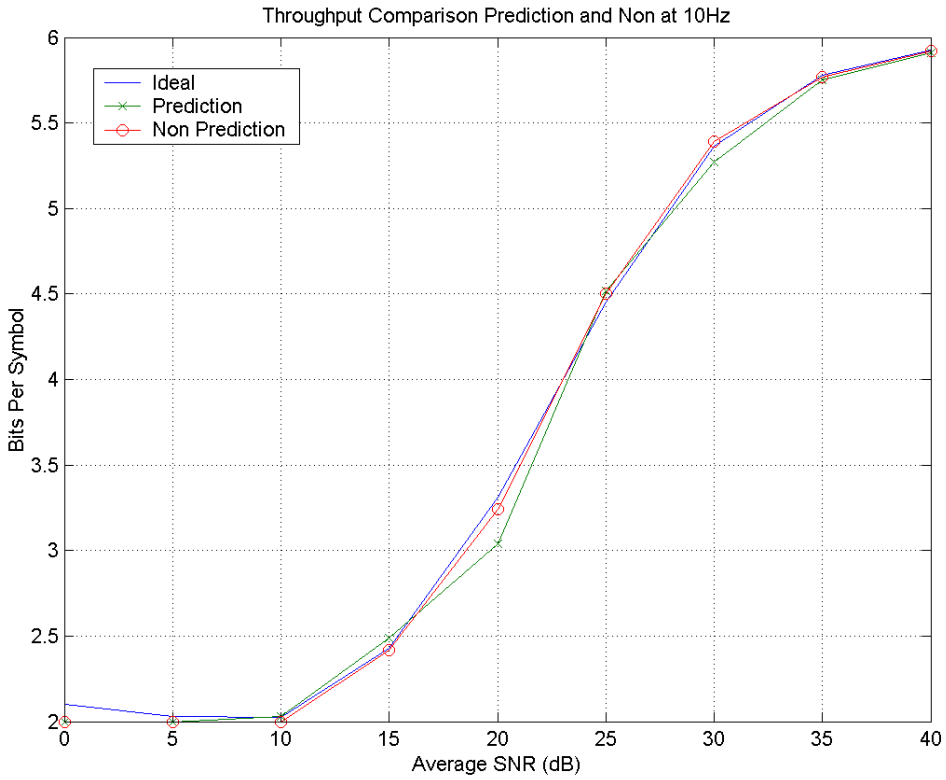

Figure 3. 25 – Throughput Comparison at 10Hz Doppler

In Figure 3.25, we have a comparison of the prediction and non-prediction throughput curves. It seems that they are both running at the correct proportions to match the ideal curve. So at low Doppler, we have a very significant BER gain, but no gain in spectral efficiency. This result is expected.

Let us now increase the Doppler frequency to 50Hz and see if we have similar results.

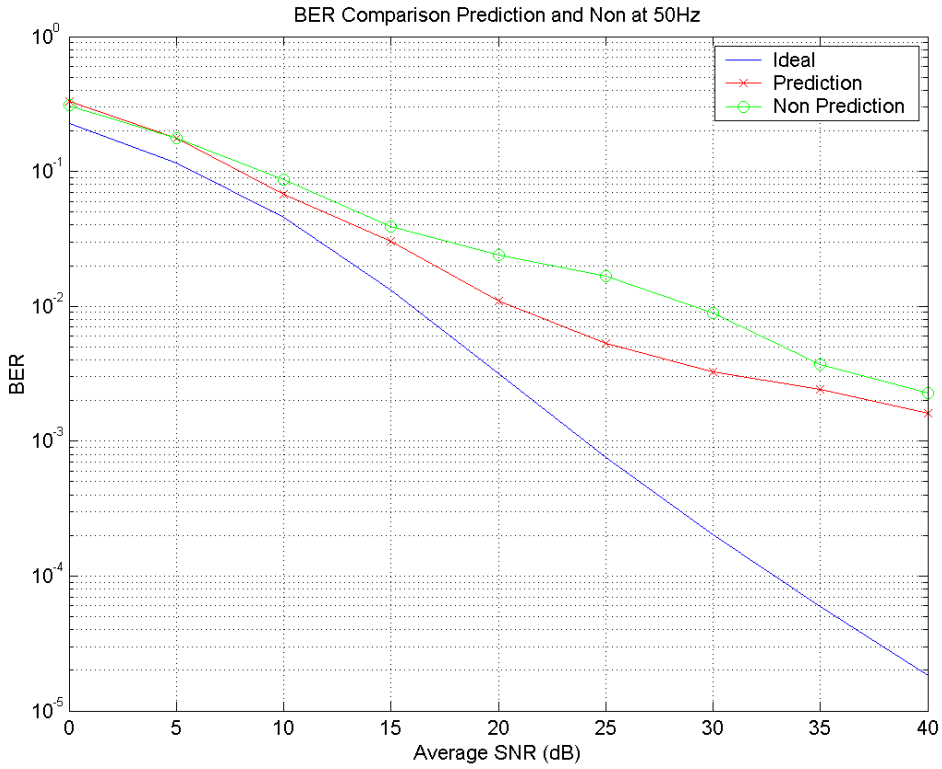

Figure 3. 26 – BER Comparison at 50Hz

At 50Hz Doppler, the improvements are less pronounced, as the prediction system is less effective at high Doppler. Instead of 15dB gains at 10Hz, now we see 5dB gains at 50Hz in high SNR regions. On the other hand, if we look at the medium range SNR, about 20-30dB, we see that the predictive scheme gives us a 10dB advantage. This is a region where in the 10Hz case, the two curves began to break away from each other.

The spectral efficiency performance comparison at 50Hz Doppler is given in Figure 3.27:

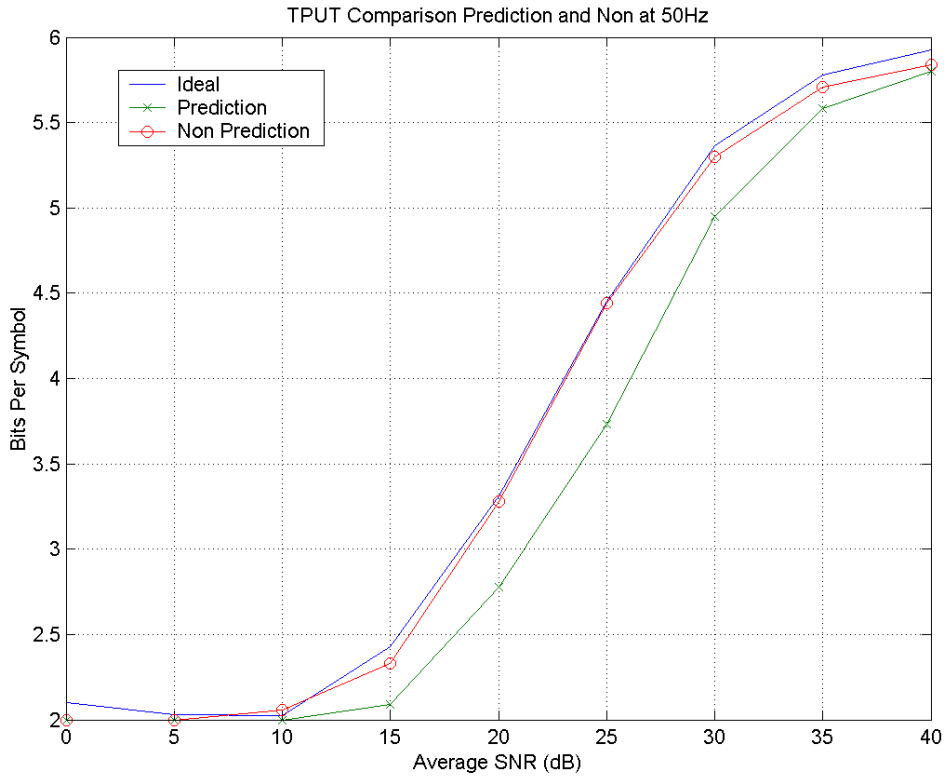

Figure 3. 27 – Throughput Comparison at 50Hz Doppler

This is a surprising comparison at first glance. While the non-predictive scheme matches perfectly with the ideal throughput, our predictive scheme degrades by as much as .75 bits per symbol. Because the only difference between the two simulation runs was the channel predictor, we must conclude that the predictor has come up short with future channel values at high Doppler. If we look back at Figures 3.21 and 3.22, we should recall that at the higher Doppler rates, the channel predictor had a tendency to predict values that were lower than the actual value. If this

occurred consistently throughout the simulation, a decrease in throughput is not surprising.

To further support the notion that the channel predictor was indeed giving lower channel values, we ran a few tests at 25dB average SNR to compare predicted SNR and actual SNR at high and low Doppler for adaptive systems with channel prediction.

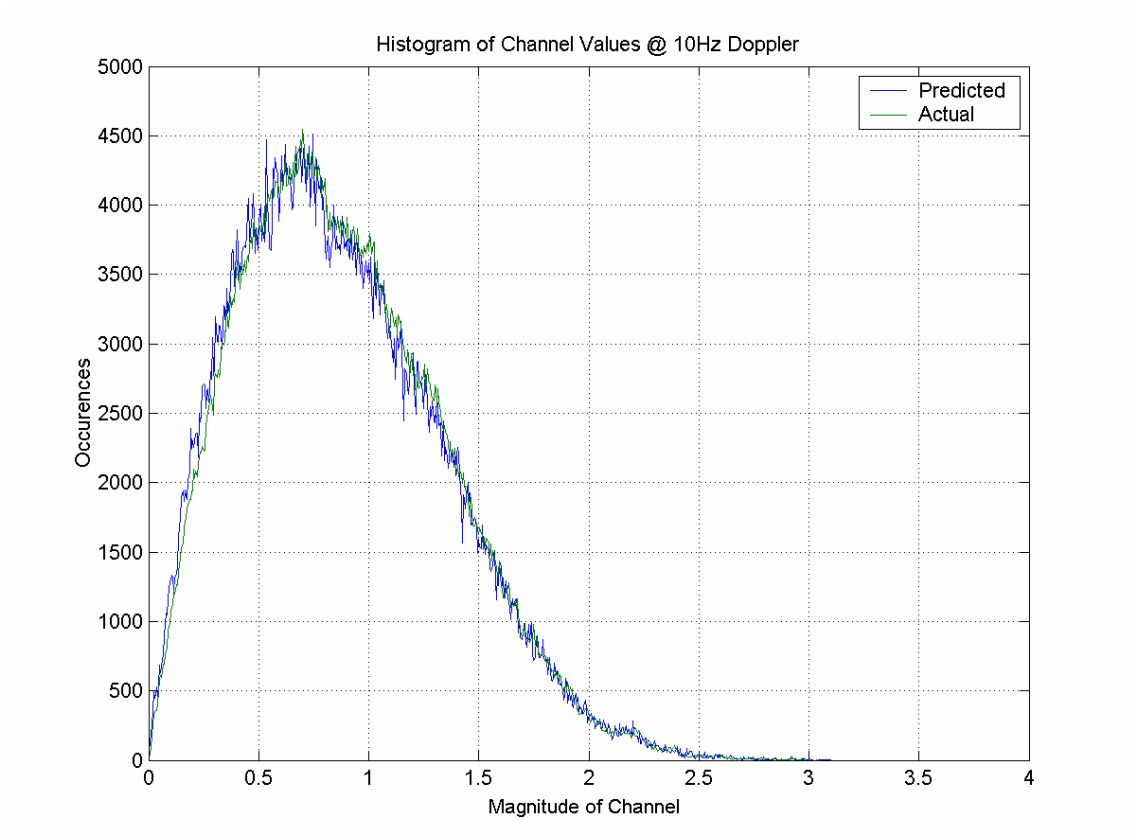

Figure 3. 28 – Predicted and Actual Channel Values @ 10Hz

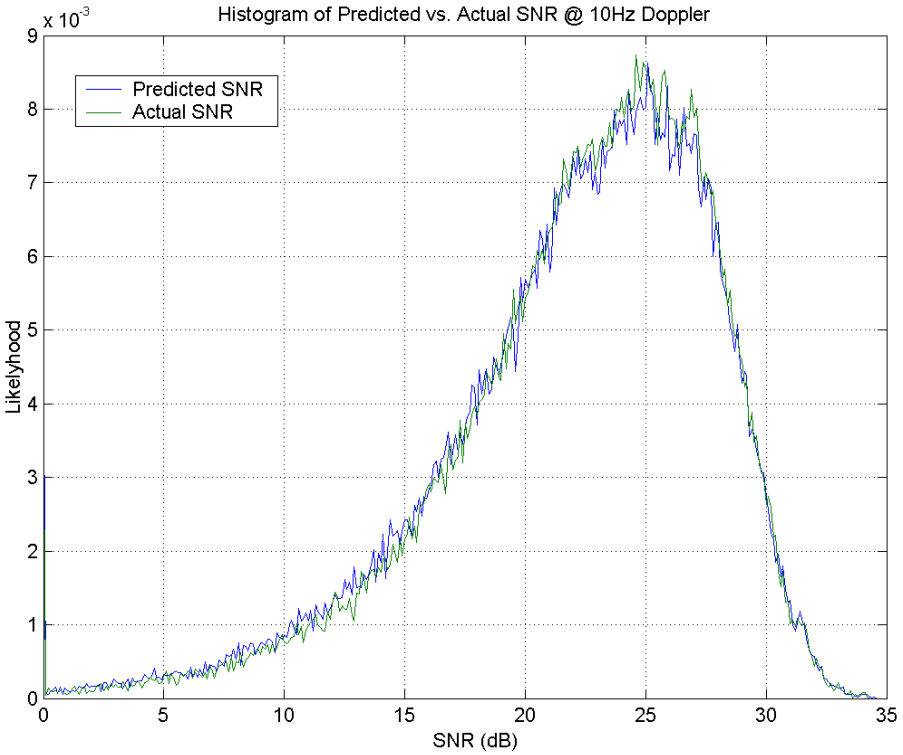

Figure 3. 29 – Predicted and Actual SNR @ 10Hz

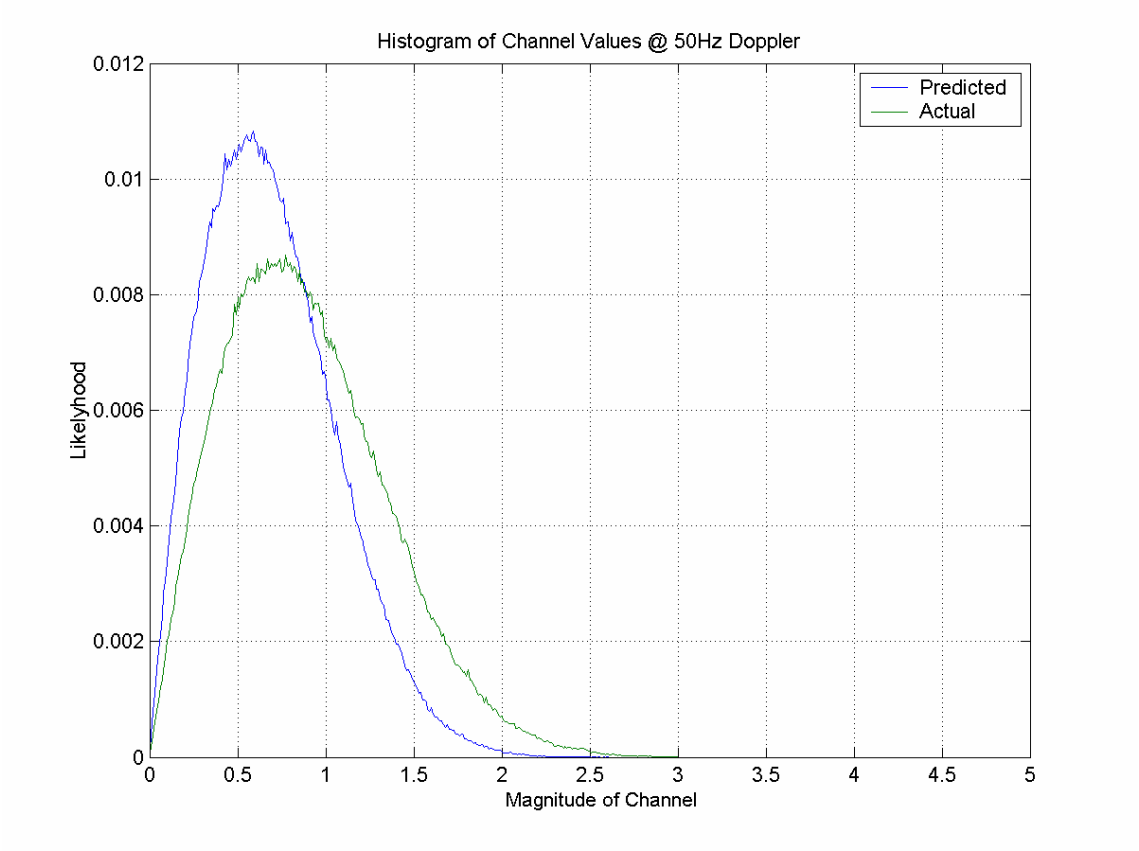

Figure 3. 30 – Predicted and Actual Channel Values @ 50Hz

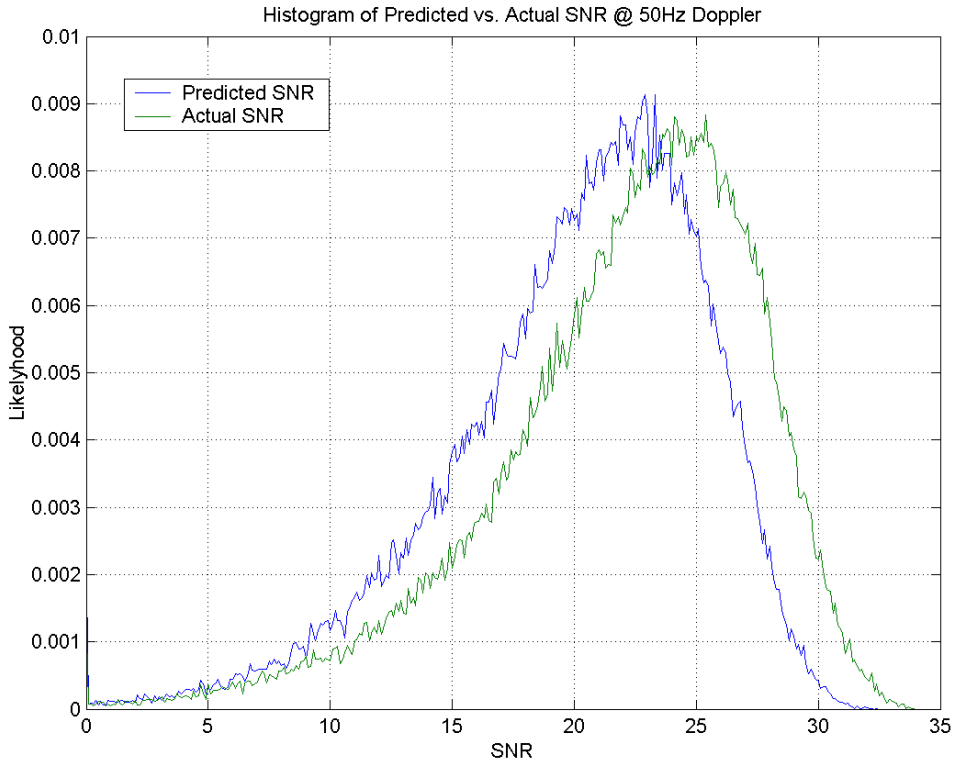

Figure 3. 31 – Predicted and Actual SNR Values at 50Hz

In Figures 3.28 – 3.31, we have plotted histograms of the values at the output of the channel predictor. We have kept track of both the channel samples and the resultant SNR values. This should decisively convince us that our hypothesis of the predictor losing precision and accuracy at high Doppler is true.

In Figures 3.28 and 3.29, we see that the statistics for the predicted and actual channel values are the same. This tells us that the predictor is accurate for low Doppler, as we saw in

Figures 3.19 and 3.20. Accurate and precise prediction leads to high throughput.

However, in Figures 3.21 and 3.22, we see that the predicted values for the channel and the SNR have deviated quite a bit. In the channel samples comparison, we see that the prediction does not give us as large a variance in values as the actual channel has; a smaller variance indicates a smaller power. In the SNR histogram envelope, we see that the SNR curve for the prediction is to the left of the actual curve by 3dB. This is an interesting result because in Figure 3.27, the throughput curve for prediction is to the right of the ideal curve by about 3dB.

3.6 Conclusions

In this chapter, we discussed the advantages of adaptive modulation over the use of static modulation systems. We showed that better BER can be achieved simultaneously with better spectral efficiency for certain system parameters. We then continued to incorporate imperfections into our simulations. FFT and SNR estimation were included and we saw that the performance was degraded by approximately 3dB. When we introduced feedback delays, we saw that high Doppler simulations suffered severe BER performance losses. This led us to use channel prediction in our adaptation system.

With channel prediction, we gained significant improvement in BER across all Doppler rates. However, in terms of spectral efficiency, we observed a lower throughput. This was due to the channel predictor biasing the prediction at high Doppler rates.

We will begin Chapter 4 with a discussion of ways to mitigate this bias in the channel predictor.

Chapter 4

IMPROVING ADAPTIVE MODULATION IN HIGH DOPPLER RATE ENVIRONMENTS

4.1 Introduction

In Chapter 3, we introduced a way of improving adaptive modulation through channel prediction. While it offered improved BER performance over non-prediction in all Doppler rates, it gave us a poorer throughput performance at higher Doppler. Specifically, it was shown that at high Doppler rates, the linear predictor exhibits degraded performance which in turns limits the performance of adaptive modulation. In this chapter, we investigate ways to improve upon adaptive modulation at high Doppler rates. We will begin by looking into the bias problem of linear prediction. We hope to be able to mitigate the problem of the biased predictor. By improving the predictor we hope to improve the performance of adaptive modulation.

4.2 Bias in Linear Prediction

According to Press and Rybicki [16], [17], there is an inherent bias in linear prediction. In our experiments, we concluded that this bias is most prevalent in cases of high signal fluctuations, i.e. high Doppler. The reason for this lies in the output of the linear prediction coefficient generator. At high Doppler rates, the channel poles that are calculated from the generator

generally sum to a figure less than 1. This is a problem when considering how the prediction coefficients are combined with sample data. Recall Equation 3.10:

$$c_n^{'} = \sum_{j=1}^{p} d_j c_{n-j}$$

When multiplying and summing the result, the mean of the predicted signals will be lower than the mean of the actual signal which means the predicted samples will also be lower than the actual samples. In other words, the predictor lacks accuracy in cases of high Doppler. We will present three different solutions in trying to solve this issue and examine the improvement (or lack thereof) they provide to adaptive modulation.

4.2.1 Improved Linear Prediction

We mentioned in Chapter 3 that we could adjust the number of prediction poles and number of input samples to affect the accuracy of prediction in the linear predictor. However, we provided no evidence of any improvement. In this section, we will perform various adaptive modulation simulations and demonstrate the impact of increasing the number of poles and samples. Also, we examine the effect of decreasing the frequency of the prediction samples on the accuracy of the predictions. Duel-Hallen [15] acknowledges that by decreasing the sample frequency of the predictor, we can increase the depth of prediction. In other words, the smaller the predictor sample frequency, the further into the future we can predict. All of these variations in channel prediction will be explored in this section.

In Chapter 3, we simulated with a consistent number of poles, training samples, and predictor sample frequency. We used 10 poles, 100 training samples, and used every 16^{th} symbol as a training sample, the same symbols used by the FFT estimator for channel estimation. We will first change the sampling frequency of the predictor from having 10 samples per frame to 5 and 1 sample per frame. Figure 4.1 illustrates the improvement that we can get at a normalized Doppler rate of $f_d*T_f = 78x10^{-3}$ (50Hz Doppler frequency and 150 symbol frames):

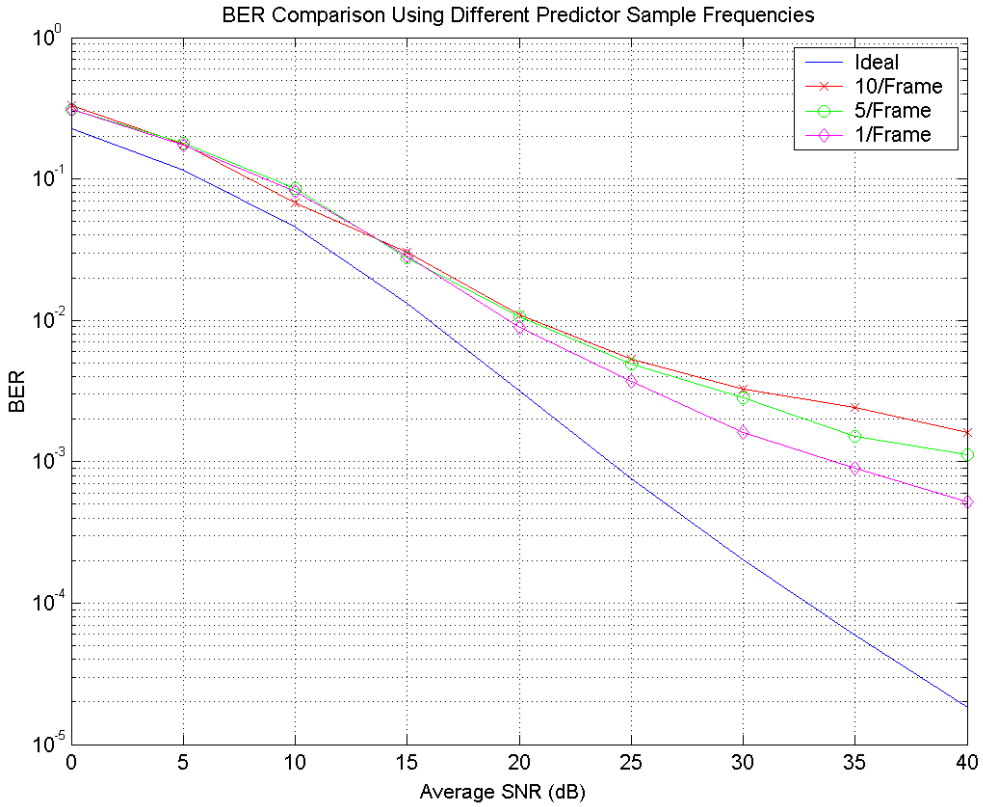

Figure 4. 1 - BER Comparison between Different Prediction Sample Rates

Figure 4.1 shows us the improvement in BER by decreasing the number of pilots per frame that the predictor uses. By keeping the number of training samples the same (100), using 10, 5 and 1 sample per frame for prediction corresponds to storing information from 10, 20, and 100 previous frames, respectively. Let us see if there is any gain to be found in spectral efficiency.

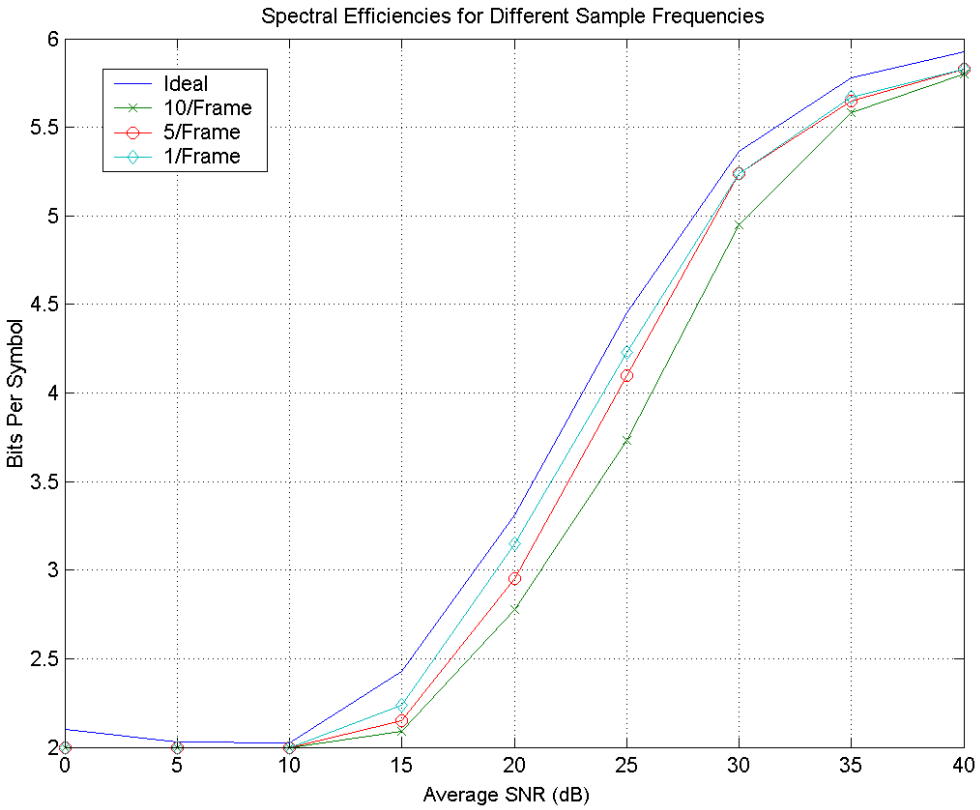

Figure 4. 2 – Spectral Efficiency for Different Sampling Frequencies

In Figure 4.2, we see that along with the improvement in BER, we also have improvement in spectral efficiency. It is clear that we can get a more accurate prediction of future channel values by sampling past values more slowly, spacing out training samples in time. However, we do not reach the ideal spectral efficiency curve in any of the frequencies used.

Let us conjecture on why the lower sampling frequency would improve our performance. A consequence of using fewer pilots per frame means that that we do not need to predict as many samples into the future for the same amount of propagation delay. When we were using 10 samples per frame, we had to predict 30 samples into the future in order to support a propagation delay of two frames; ten samples per frame and three frames of information means that we have to predict 30 samples ahead. Now when we use only 5 samples per frame, we need only predict 15 samples into the future in order to support our delay. If we use one sample per frame, we only need to predict three samples ahead. The significance of this is the following: since we have already shown that linear prediction is not perfect (Section 3.5.3), the more samples into the future we predict, the more error is propagated through our result. The fewer samples it has to look ahead, the smaller the accumulated error will become. However, note that we must still sample fast enough satisfy Nyquist's criterion.

Now that we have confirmed Duel-Hallen's claim that slower prediction sampling frequencies lead to better future predictions, let us see what happens when we adjust the number of poles and training samples. Let us begin using 5 samples per frame as our rate of prediction sampling.

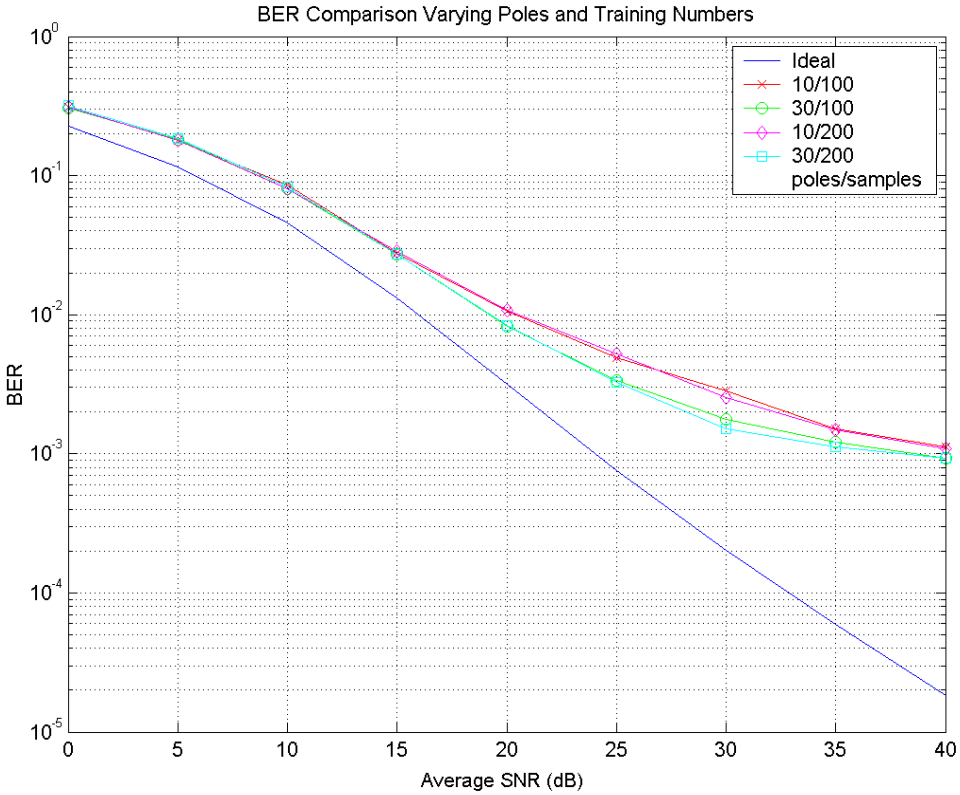

Figure 4. 3 – BER Comparison between Varying Number of Poles and Samples

In Figure 4.3, we varied the number of poles between 10 and 30 poles, and used 100 and 200 training samples. At first glance it does not seem that there is much of a difference, and indeed there is not. However, from the graph, we can see an interesting result. It would seem that increasing the number of poles has a much larger impact on performance than increasing the number of sample data. The difference between the two curves representing the use of 30 poles is negligible. The 30-pole, 200-sample curve is only minutely better than the 30-pole, 100-sample curve. Also, at high SNR (~40dB), the two sets of curves seem to

converge. Let us examine the results of the spectral efficiency curves and see if we find similar gains.

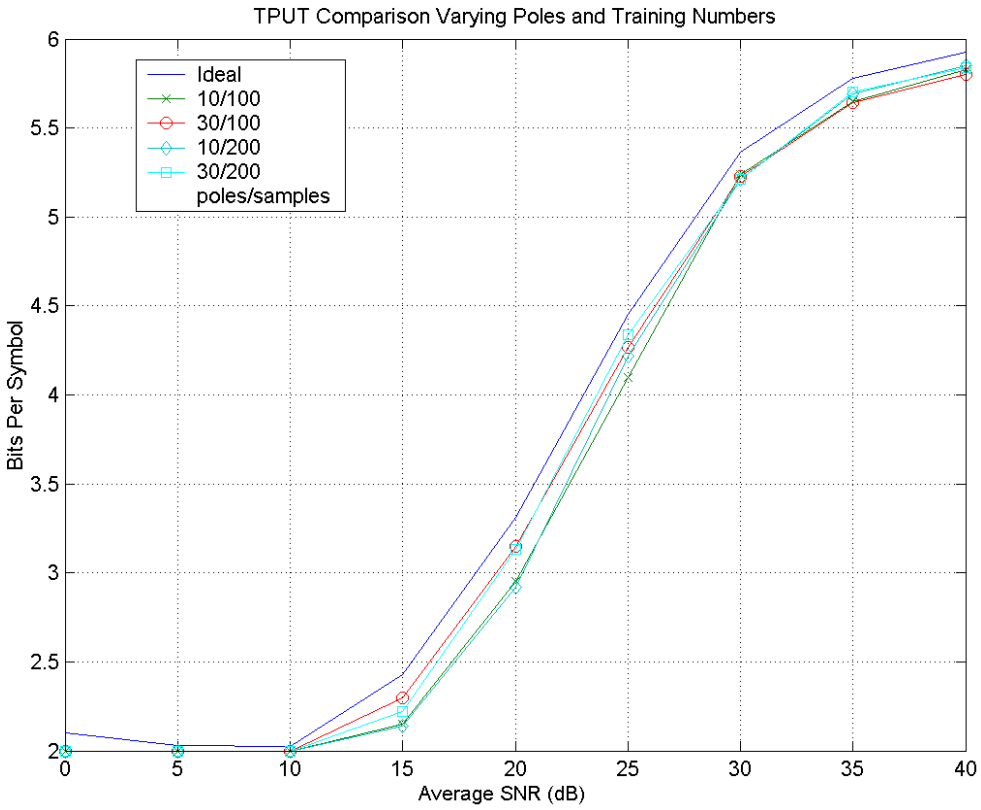

Figure 4. 4 – Spectral Efficiency when Varying Parameters

Indeed, we do find that we have a slightly better spectral efficiency when we use 30 poles over 10 at medium SNR. In cases of high and low average SNR, we find that neither poles nor training samples make a difference in performance, similar to what we found in the BER curves.

It is important to observe that an increase in training samples does not seem to have any significant effect on performance. The explanation revolves around the functions of the poles and training samples. The purpose of training samples is to determine the pole values. The more training samples we use, the more accurate the poles are. However, the poles are directly responsible for the prediction process. The more poles we use, the more past data can be incorporated into the next prediction point. With ten poles and 200 training samples, we are limited to ten samples of past information to predict the next sample. With more poles, we have more past information to use, and this leads to better performance. The simulation results seem to agree with this thinking. The fact that we have little improvement with the increase of training samples tells us that 100 training samples provides very good estimation of the poles to use in prediction.

We should remind the reader that the previous plots were generated by using 5 pilots per frame in the prediction process. Let us see what happens when we use a more efficient frequency when predicting: One sample per frame.

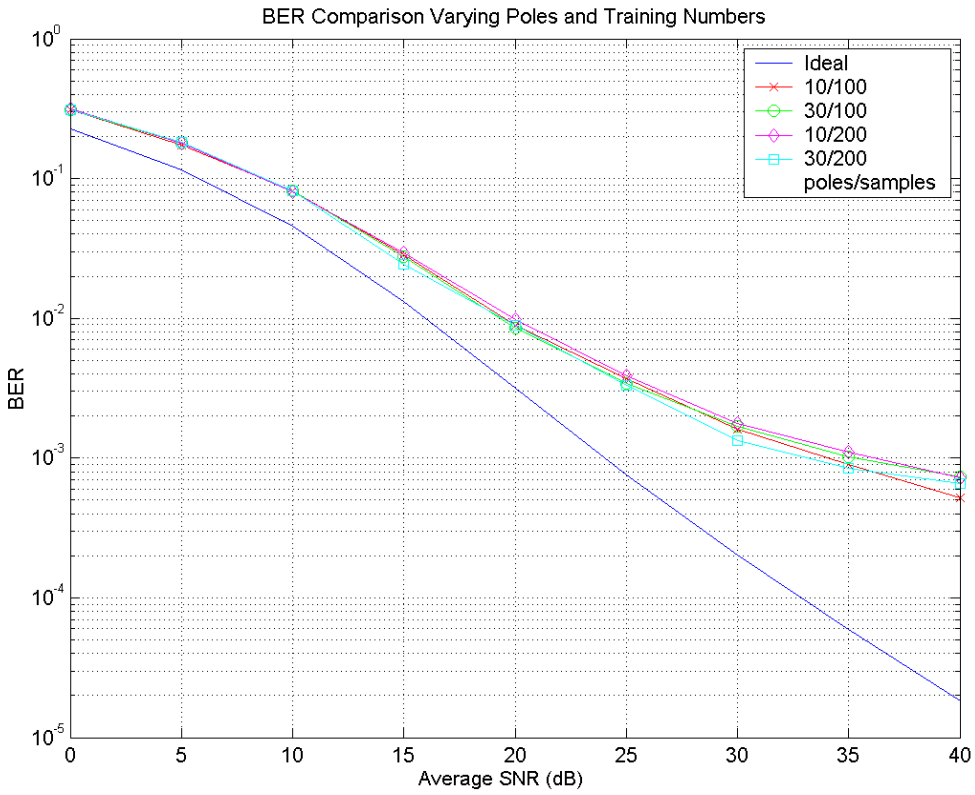

Figure 4. 5 – Performance of Varying Parameters in Low Prediction Frequency

In Figure 4.5, we find that using only one sample per frame renders changing the number of poles and training samples irrelevant. We can see that the performances of all four curves are very close to one another. However, if we look closely we can see that the 10-pole, 100-sample case has the best performance. This can most likely be attributed to simulation noise at high SNR. If we take a careful look at the plots, we can see that up to about 30dB, the results we see are consistent with our results in Figure 4.3.

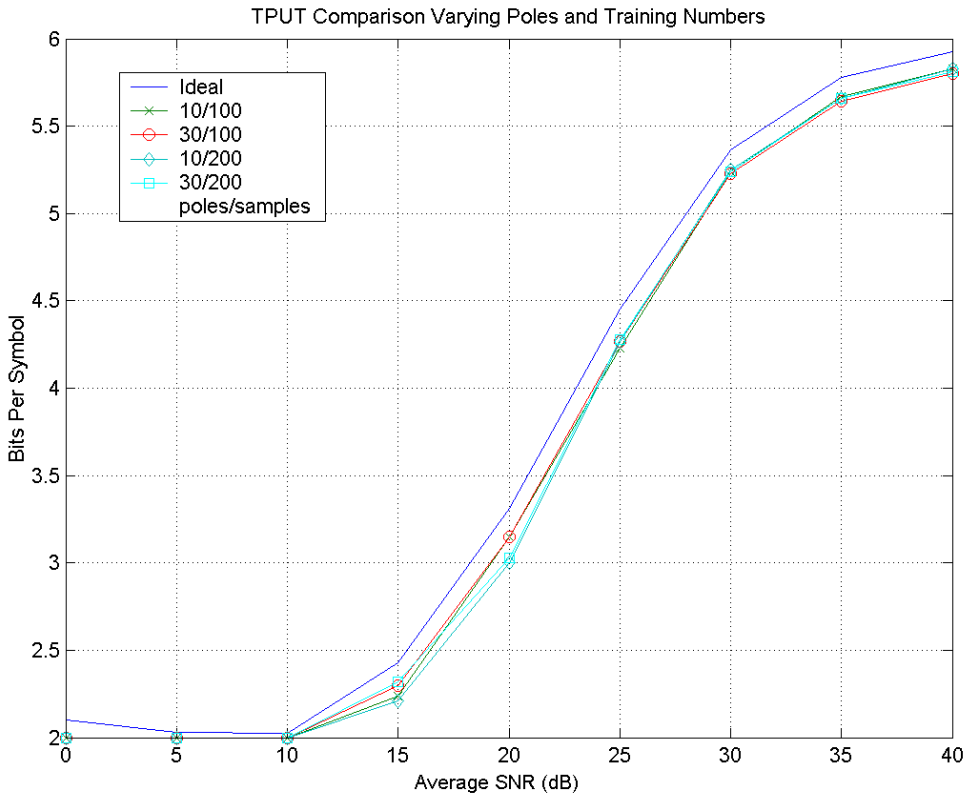

Figure 4. 6 – Spectral Efficiency of Varying Parameters at 1 Pilot/Frame

The spectral efficiency curve that we obtain by varying the prediction parameters shows us that when running one prediction pilot per frame, we see near identical performance over each permutation of poles and training sample numbers. It is important to note that it is about 0.1 bit per symbol short of the ideal performance characteristic.

In summary, increasing the number of poles and training samples does have a positive impact on adaptive modulation performance. However, the impact that it will have depends more upon the sampling frequency that is used by the predictor in acquiring

channel samples to use for training. When we use 5 pilots per frame, we can see a small gain in both BER and throughput, but when we use only 1 pilot per frame, increasing poles and training samples provide negligible gain. This combined with the earlier simulations varying only the sampling frequency tells us that the sampling rate is the most significant variable in linear prediction among the three variables examined.

4.2.2 Manipulating Prediction Coefficients

Now that we have shown that we can get gains by making better choices in the number of poles/training samples and prediction pilot frequency, we will move on to try and mitigate the bias of the predictor. As stated earlier by [16] and [17], the underlying problem of the predictor is the fact that the poles generated by the MEM algorithm do not sum to 1. Instead, they sum to a number slightly less than one. In this section, we intend to examine the option of manipulating the prediction poles (or coefficients) in such a way that the predictive properties are still present, but the sum of the poles will add to 1. We propose to add to the poles the difference it would take to make them sum to 1. Following this line of thinking, we require an intelligent method of adding weights to the poles. We presently have two ways of adding weights. We could divide the difference from unity of the poles and evenly distribute them along all poles.

Another approach is to add the difference to only one pole. The pole that would receive the correction factor would be the one that corresponds to the most recent channel sample. The reason why we would choose that pole is because it is the sample that

would be most correlated with the next sample to be predicted, thus minimizing any variation in immediate prediction.

Experiments were performed to examine the impact of these two methods. The first set of plots corresponds to distributing the residual value to all poles.

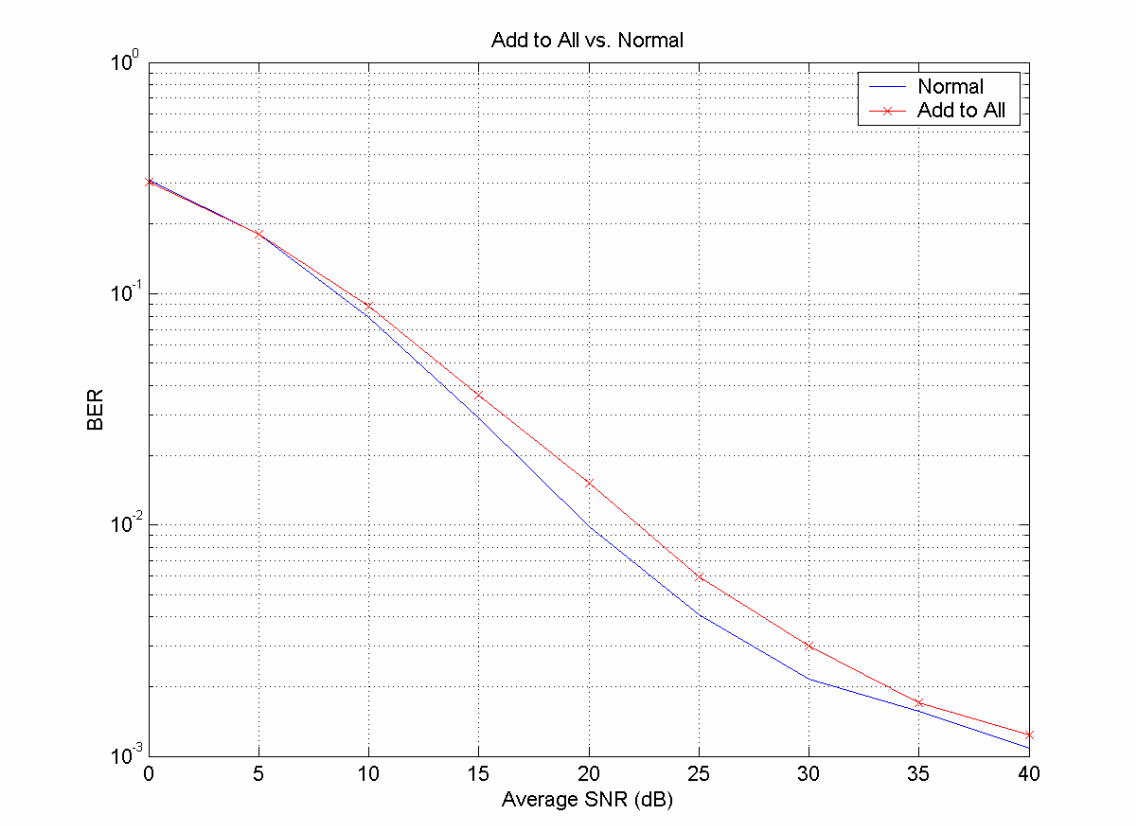

Figure 4. 7 – Add to All Poles Method vs. No Correction

The plots shown in Figure 4.7 show us that by using the Add to All poles (or coefficients) method gives us poorer BER performance than the uncorrected version. We suspect that the

reason for this is because the uncorrected version is choosing less efficient modulation schemes than the corrected version.

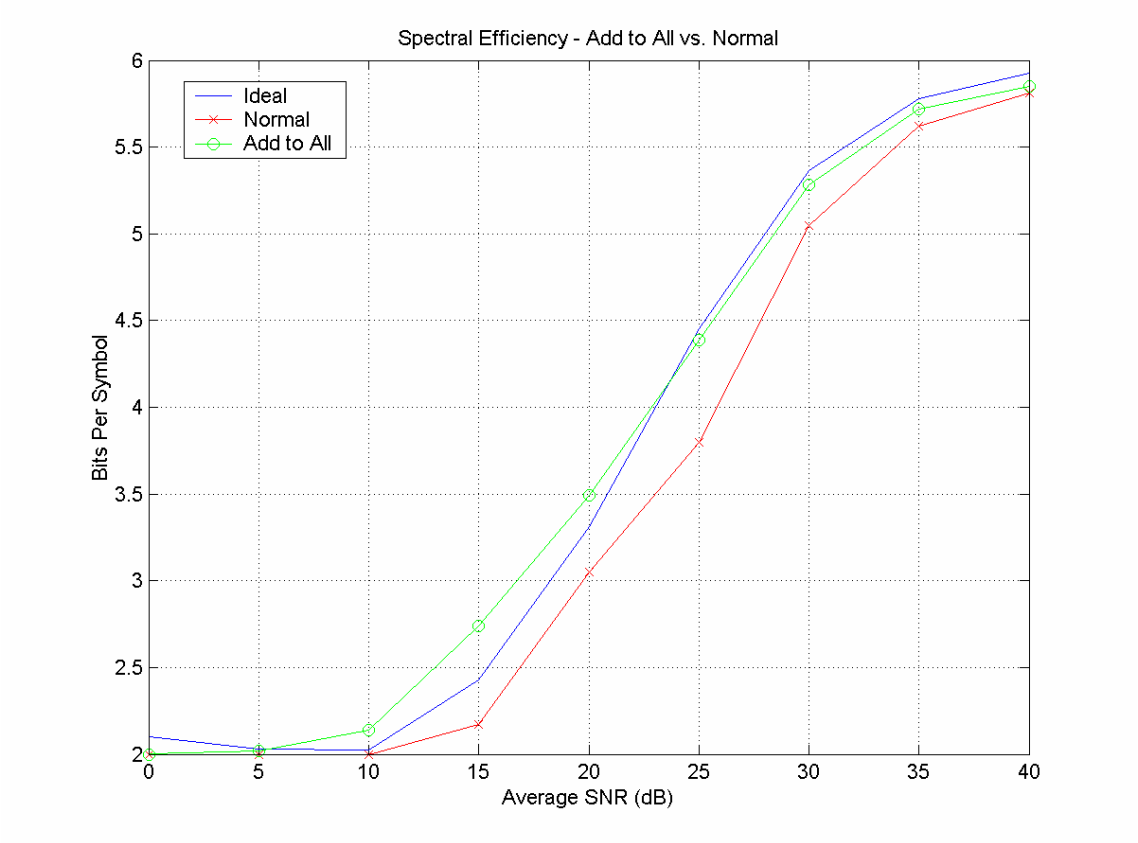

Figure 4. 8 – Spectral Efficiency Comparison for Add-to-All and Normal Prediction

In Figure 4.8, we plot the spectral efficiency curves for the Add-to-All correctional type and the normal version of prediction. The efficiency of the Add-to-All correctional method does outperform the normal method. This would explain why the BER of the normal method is better than that of the corrected method. We are trading throughput for BER performance.

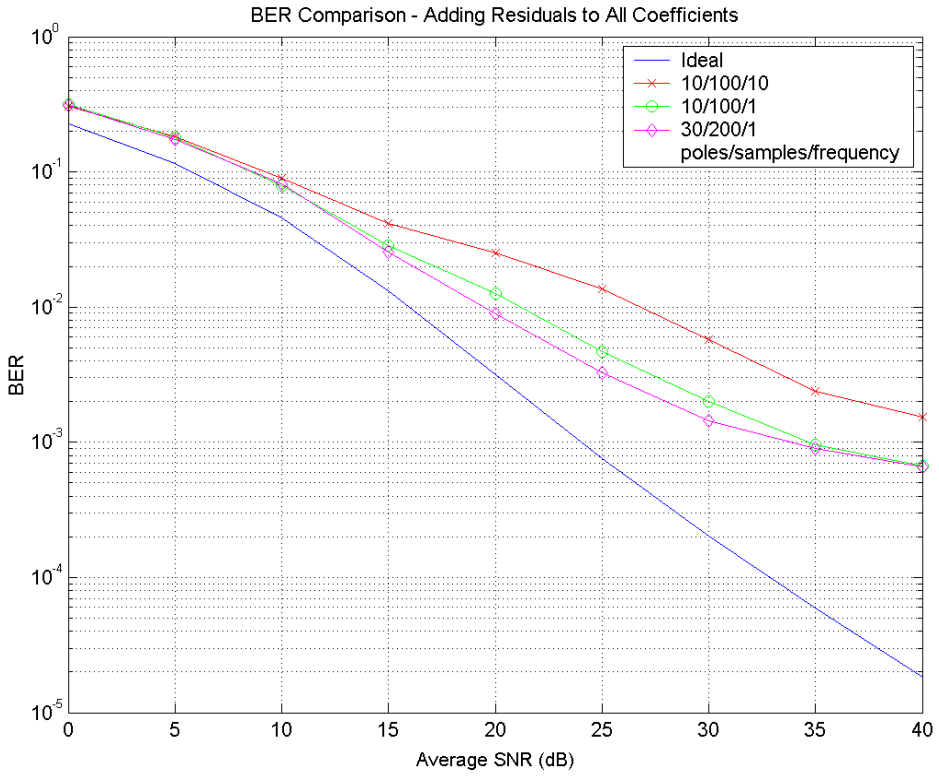

Figure 4. 9 – Comparison of BER When Adding Residuals to All Poles

Figure 4.9 shows the BER performance when forcing the poles to equal 1 while adjusting the prediction coefficients. Not surprisingly, using 30 poles, 200 training samples and 1 pilot per frame yields the best performance.

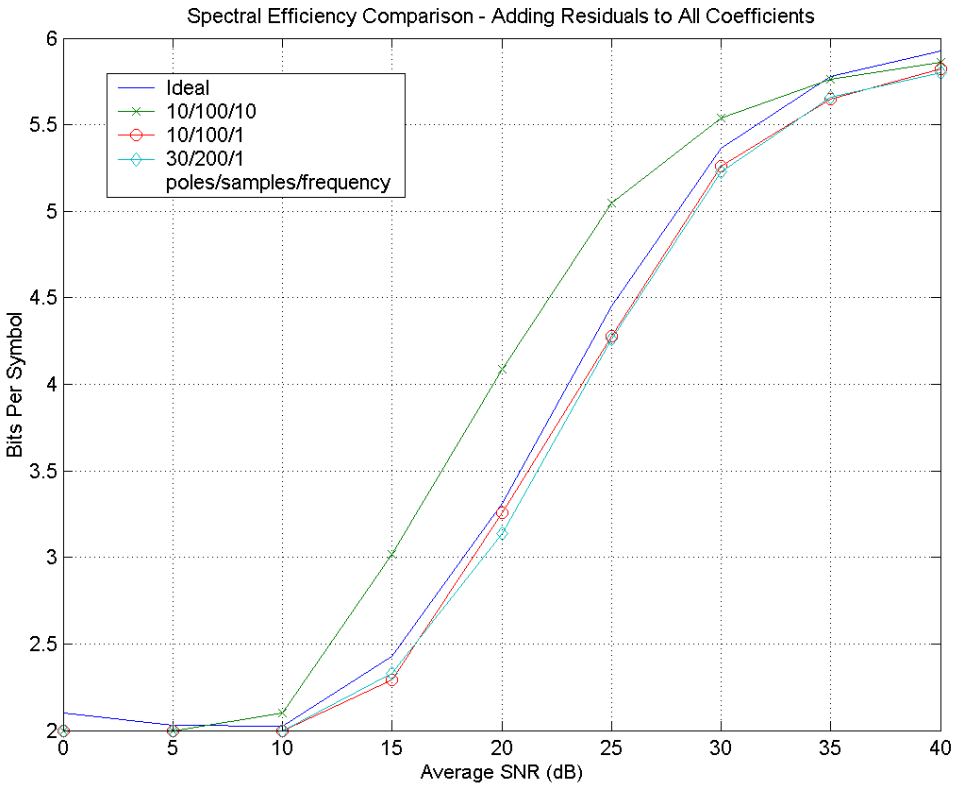

Figure 4. 10 – Throughput Comparison

Figure 4.10 shows us the throughput comparison among the different ways of implementing this method of correcting linear prediction. In terms of spectral efficiency, the 10 pole, 100 training sample, 10 pilots per frame case gives us much better performance than do the other two. In fact, it gives us performance that surpasses what we would expect with ideal performance. This is not such a surprising result considering the BER performance of the 10/100/10 (poles/training samples/pilots per frame) curve is much worse than the other two. On the other hand, it seems that although the 30/200/1 scheme is better for BER than the 10/100/1 scheme, it yields the same

spectral efficiency. What we are seeing here is a tradeoff between BER and spectral efficiency. That is the linear predictor is over-estimating the SNR which leads to the use of 64QAM or 16QAM at lower SNR values. This increases the spectral efficiency but degrades power efficiency.

Let us move on to the scenario where we add the residual to the first coefficient only.

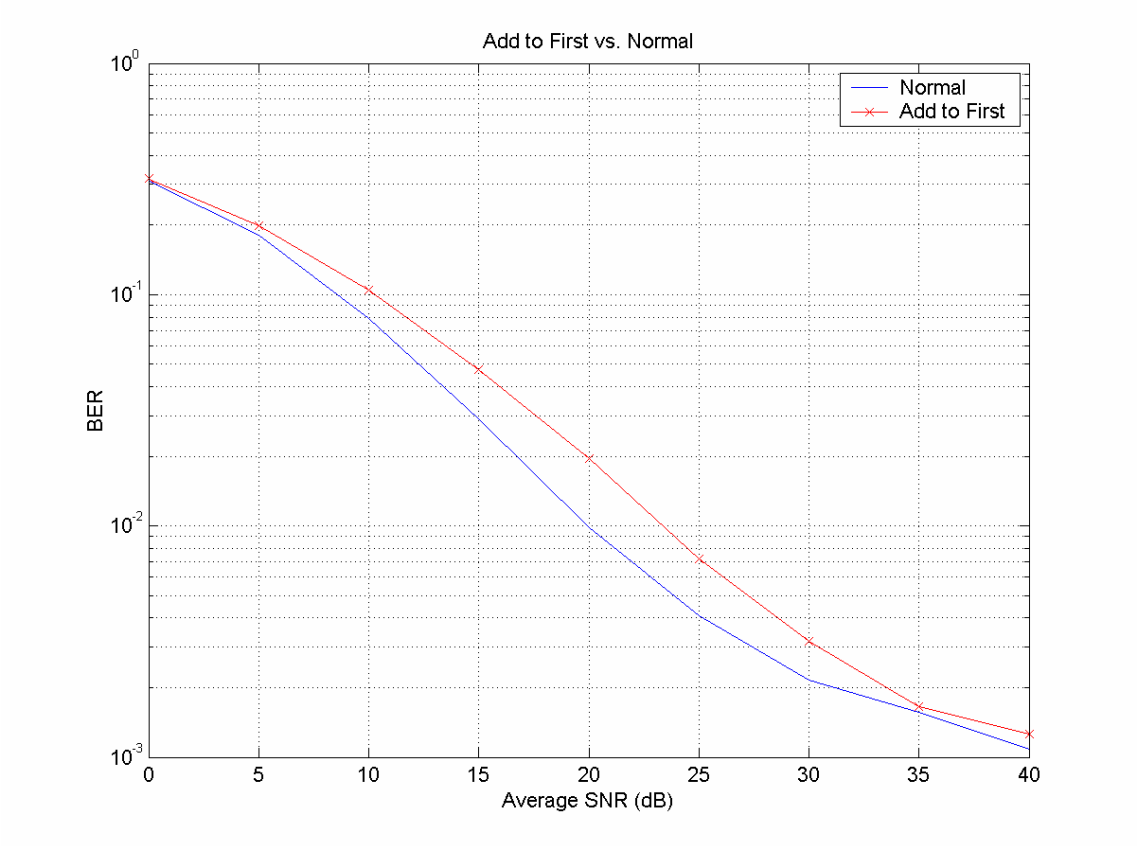

Figure 4. 11 – BER Performance of Add-to-First and no Correction

In Figure 4.11, like the Add-to-All correction method, the Add-to-First method underperforms compared to the no correction

method. Again, let us plot the performance of the spectral efficiency and determine if we have a tradeoff similar to what we had in Figure 4.9 - 4.10.

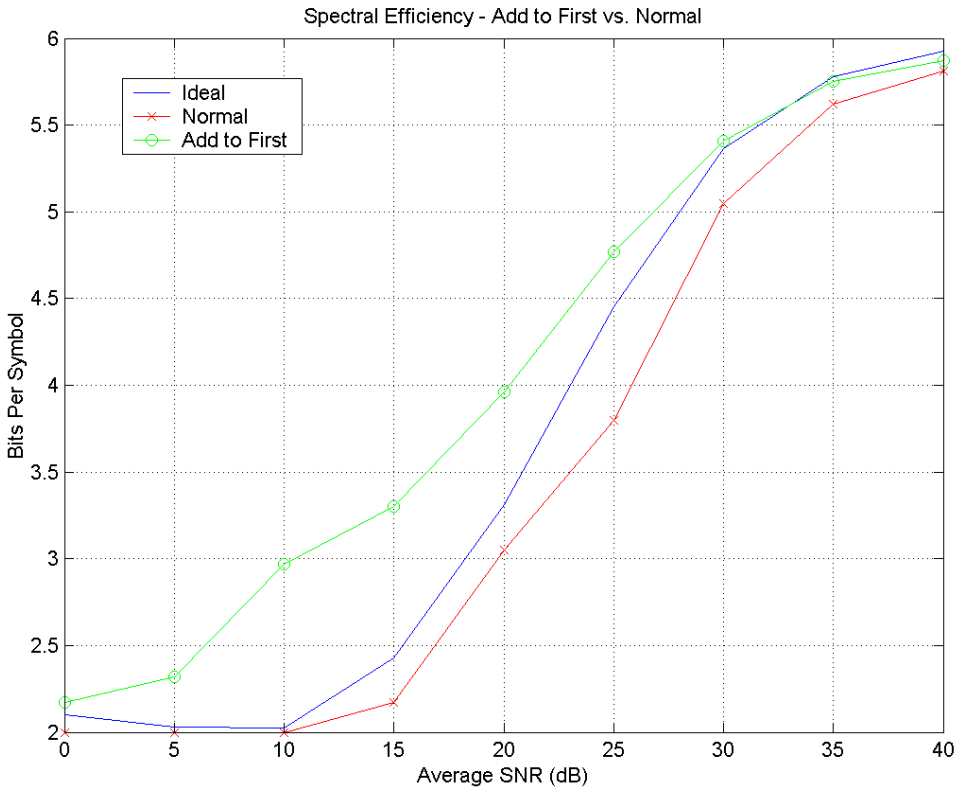

Figure 4. 12 – Spectral Efficiency of Correctional Method and Normal Method

Figure 4.12 shows us that there is a tradeoff between efficiency and error performance. It would seem that the Add-to-First method tends to over-estimate the power of the channel and chooses higher order modulation schemes at low average SNR. This would explain the large gap in performance in the BER plot at low and medium average SNR.

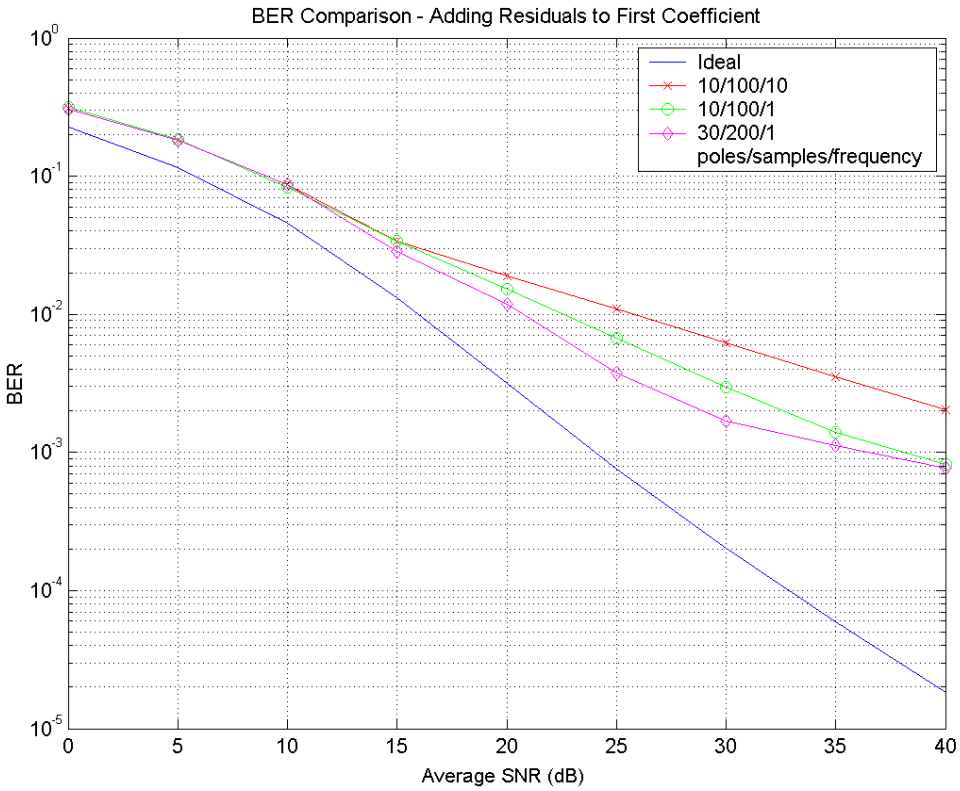

Figure 4. 13 – BER of Adding to the First Pole Only

In Figure 4.13, we plot the performance of adding residuals to the first coefficient in systems where we vary the number of poles and training samples. The most sophisticated scheme yields the best return in terms of BER (30 coefficients and 200 training samples).

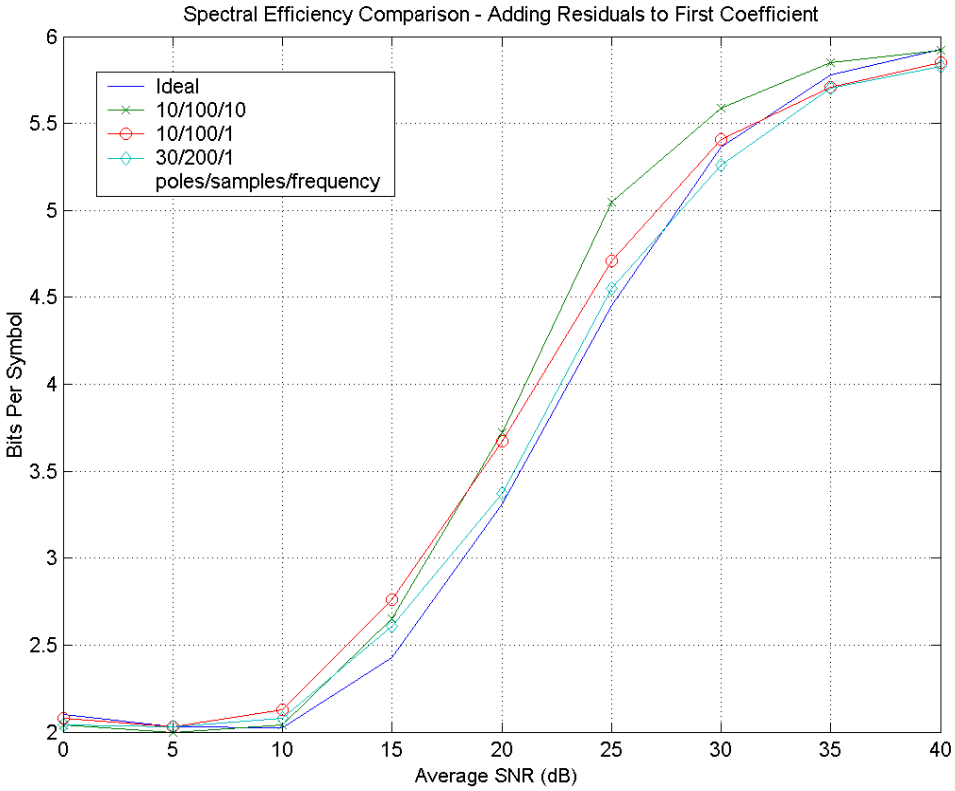

Figure 4. 14 – Throughput of Adding to the First Pole

Looking at Figure 4.14 we find the spectral efficiency, however, does change. In this case, the spectral efficiency in general is greater than the ideal performance in all cases. However, at high SNR, we see that all curves seem to level out beneath the ideal performance. Unlike the all-poles addition case, the most sophisticated prediction scheme in this case yields the same performance as the ideal performance at medium SNR. It would seem that to obtain performance near that of ideal spectral efficiency, one must combine the 30-pole scheme with adding residuals to the first pole.

Let us now compare the strongest correctional methods alongside the non-corrected simulations to see if we gain any increases in performance.

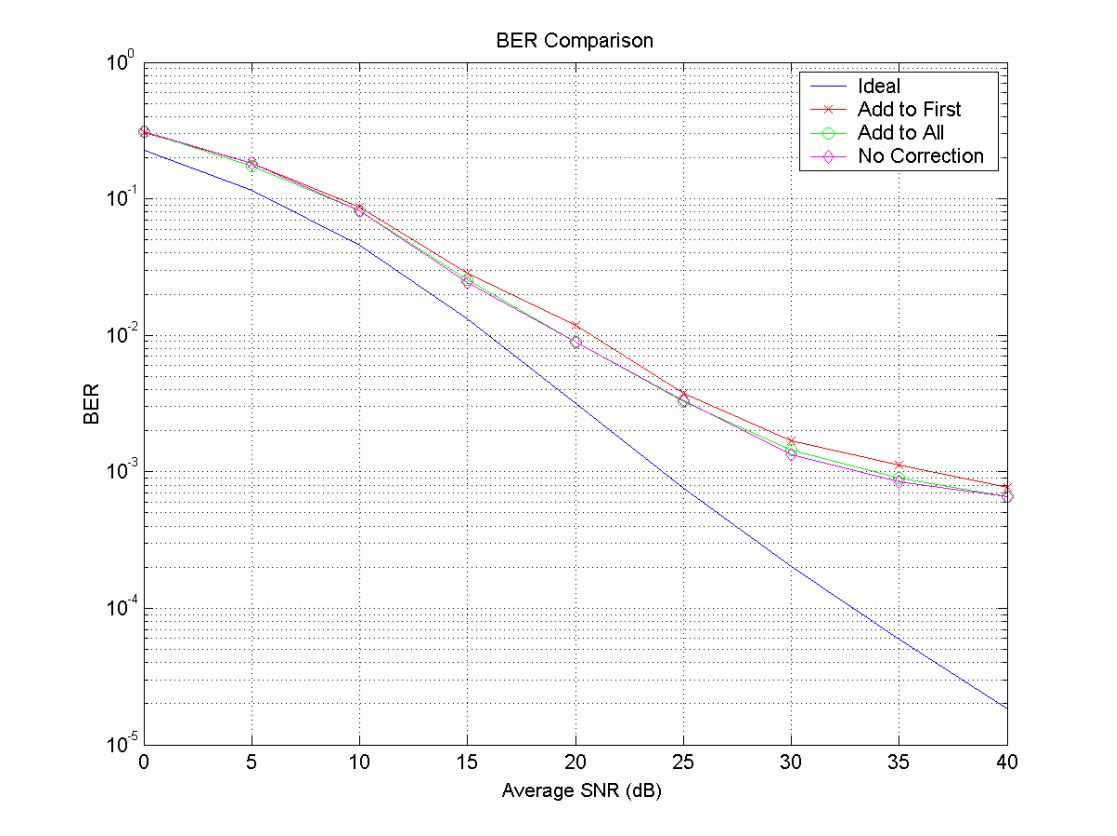

Figure 4. 15 – BER Comparison of Different Correction Methods

In Figure 4.15, we see the performances of the different correctional methods for linear prediction. While they all have relatively the same performance, the best performance comes through not correcting at all. The worst performance is through using the Add-to-First method. The Add-to-All method is only slightly inferior to not correcting at all.

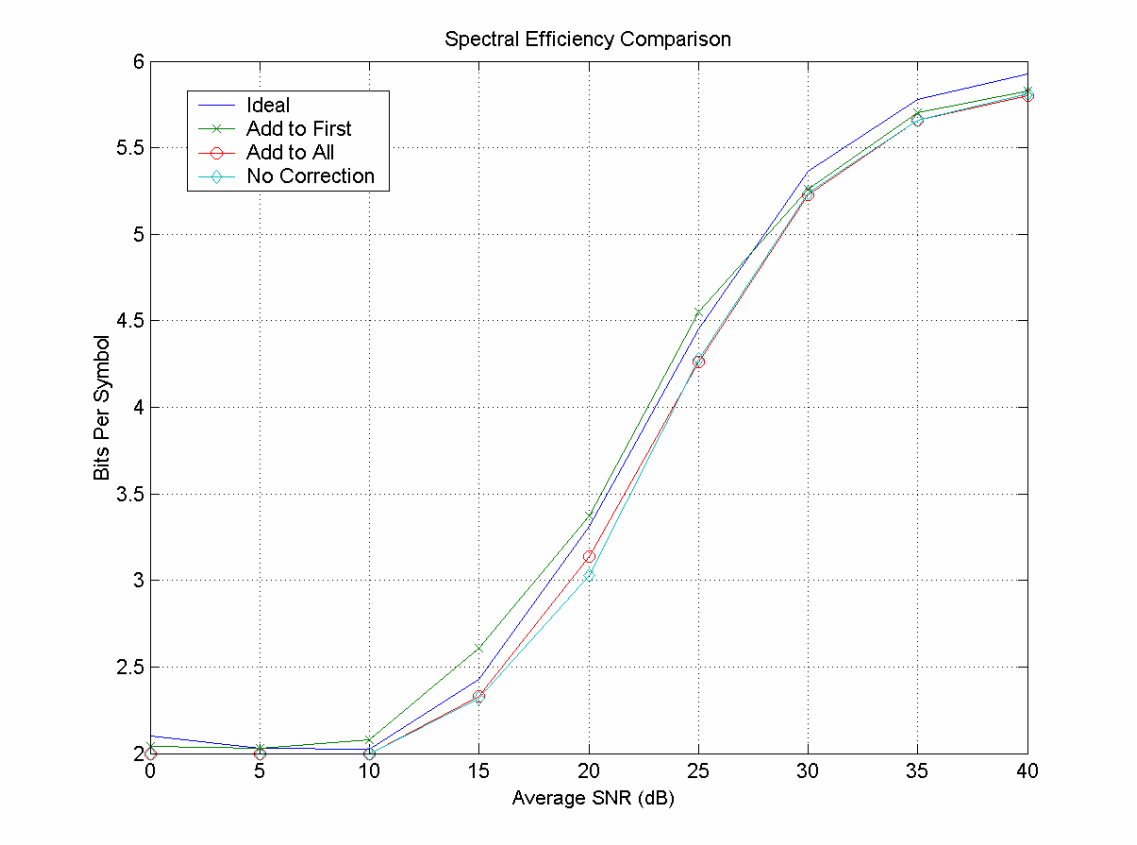

Figure 4. 16 – Throughput Comparison in 3 Correction Schemes

Figure 4.16 shows us the spectral efficiency plots of the three correctional methods that have been discussed. Not surprisingly after seeing the BER curves, the Add-to-All method yields nearly the same efficiency as the non-corrected version. Also expected is the throughput of the Add-to-One plot being more efficient than the other two. However, the efficiency at high SNR in all cases still falls short of ideal.

In summary of this section, we conclude that even though the non-corrected method of linear prediction is slightly under the ideal spectral efficiency performance curve, it is close enough such

that altering the prediction poles in this manner will not improve BER and spectral efficiency simultaneously. Using the Add-to-First method was successful in improving the spectral efficiency, but degraded the BER whereas the Add-to-All method did not improve much of either BER or efficiency.

At this point we can only conjecture why adding the value of difference between the poles and 1 among all poles did not do much to change the performance of adaptive modulation. We believe it is because the residual value is so small that adding it across all poles (especially over 30) amounts to very little change at all. In testing simulations, we found that often times, the residual value would be on the order of 10^{-2} or smaller. Divide that by 30 and we would not expect predicting 3 samples in the future to be changed dramatically.

Perhaps this is why adding to the most recent pole resulted in more significant changes. Adding to the first pole allowed a higher weighting on it, making future predictions pull toward it. So if on average, we had low predictions, this way of correcting would pull them higher.

4.2.3 Mean Subtraction

In the previous section, we developed a way to adjust the bias in linear prediction by adjusting the linear prediction coefficients, or poles directly. In this section we will explore a more conventional approach that involves removing the mean from the data set and returned in the predicted signal. Press and Rybicki [16], [17] propose that in order to remove the bias in linear prediction, a correctional factor should be subtracted

from the data set before the prediction coefficients are calculated. The prediction method should be used with the adjusted data and said coefficients followed by adding the correctional factor into the predicted data. Now we have to decide what the correctional factor should be. Press and Rybicki state that although it would be intuitive for the correctional factor to be the mean of the data set, it is actually an autocorrelation-weighted mean [17].

$$m = \frac{E^T \left[S + N\right]^{-1} y}{E^T \left[S + N\right]^{-1} E} \tag{4.1}$$

where E is a column vector of unit components of the same length as the training data. S and N are autocorrelation matrices of the signal and the noise, both we must assume to have long term knowledge. They are found in the following way [15]:

$$R_{ss}(T_s) = J_0(2p\, f_d T_s) \tag{4.2a}$$

$$S = \begin{pmatrix} 1 & R_{ss}(2T_s) & \cdots & R_{ss}(NT_s) \\ R_{ss}(2T_s) & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & R_{ss}(2T_s) \\ R_{ss}(NT_s) & \cdots & R_{ss}(2T_s) & 1 \end{pmatrix} \tag{4.2b}$$

$$N = \begin{pmatrix} s^2 & & 0 \\ & \ddots & \\ 0 & & s^2 \end{pmatrix} \tag{4.3}$$

where J_0 represents the first order Bessel function, T_s is the symbol period, and s^2 is the variance of AWGN.

Let us observe the performance of the Mean subtraction Method as compared to no correction.

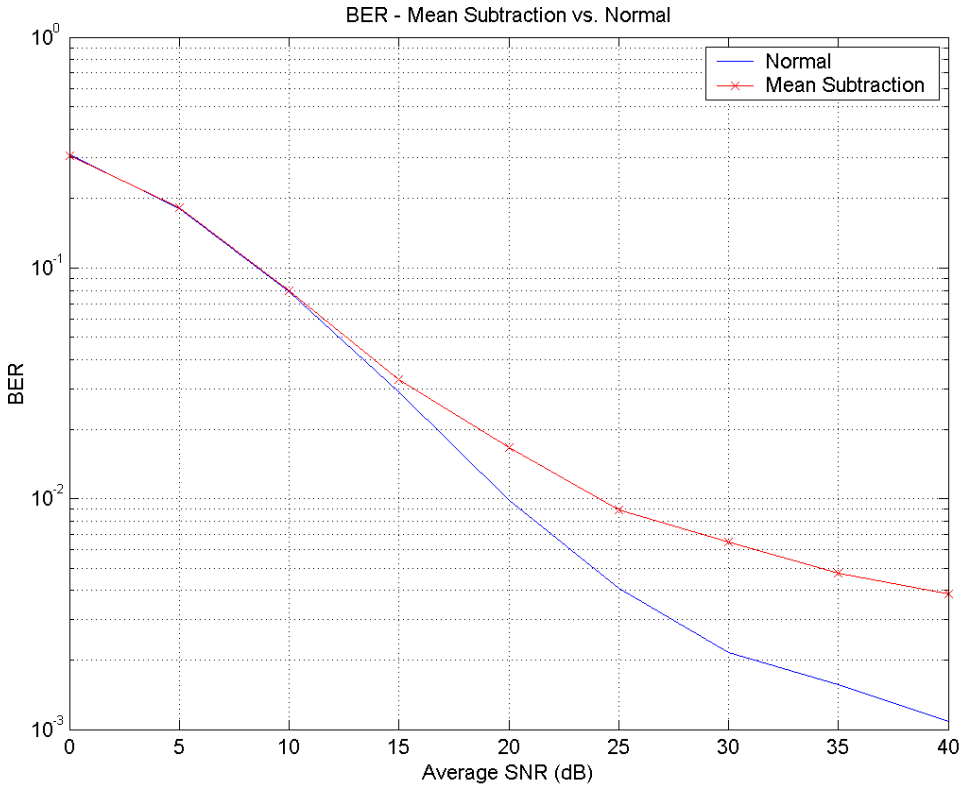

Figure 4. 17 – BER Performance between Mean Subtraction and No Correction

In Figure 4.17, we see that there is a very large deviation in performance between the Mean Subtraction method and the normal technique to perform channel prediction. This method has given us the greatest decrease in performance as compared to the coefficient manipulation techniques. Let us see if there is a greater increase in spectral efficiency than the other two methods.

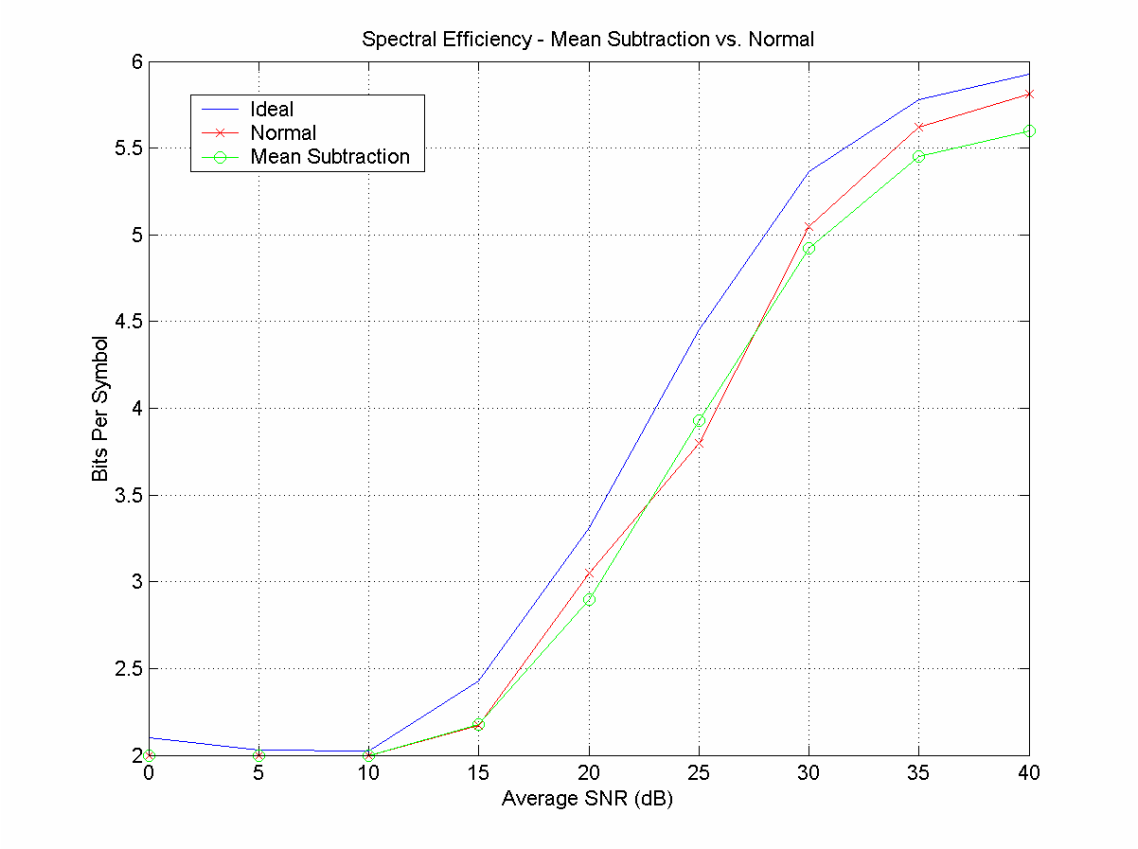

Figure 4. 18 - Spectral Efficiency of Mean Subtraction and Normal Prediction

The spectral efficiency result that we have in Figure 4.18 is very surprising. We expected that with the decrease in BER, we would have a much better spectral efficiency performance with Mean Subtraction. On the contrary, we have roughly the same efficiency performance as the non corrective scheme. This is showing us that the mean subtraction method is not a proper correction scheme for our purposes. We run a plot of the predicted channel powers against a plot of the actual channel powers to see if they agree.

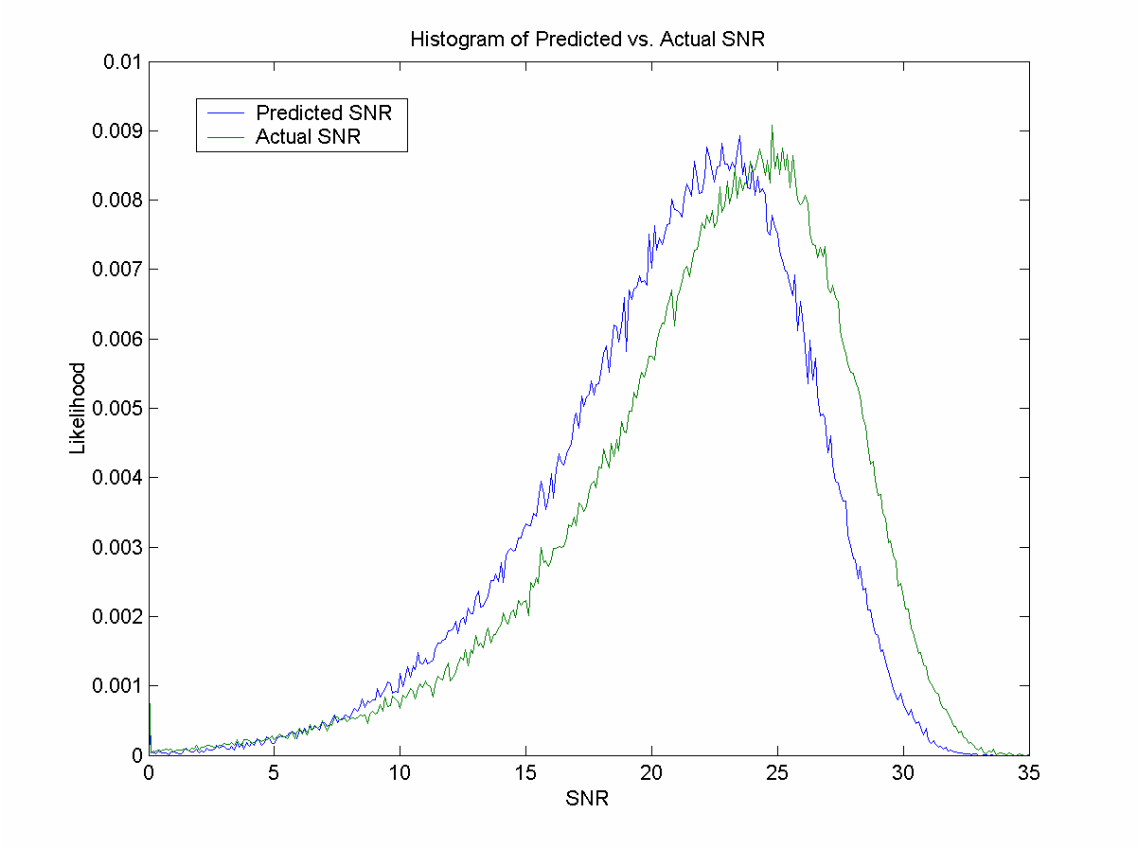

Figure 4. 19 – Histogram of Predicted and Actual Instantaneous SNR for Mean Subtraction at 25dB Average SNR and 50Hz Doppler

In Figure 4.19, we see that the Mean Subtraction method does not eliminate the bias in the predictor. What we interpret this is that Mean Subtraction is not doing anything for us. One reason this may be the case is due to what we are predicting and what this correctional algorithm is meant to do.

The predictor that we employ extrapolates the future channel samples in the I and Q channels of the signal separately. These channels both have zero mean, so using a Mean Subtraction method

here would not be of value. However, let us attempt to use this method while lowering the predictor's sampling frequency and see if we receive better results.

In testing the mean subtraction method in combination with slower prediction sampling, we used a 50Hz Doppler fading channel and two different prediction schemes: one with 10 poles, 100 training samples, and one pilot per frame, and another with 30 poles, 200 training samples, and one pilot per frame. Here are the results that the simulations have produced.

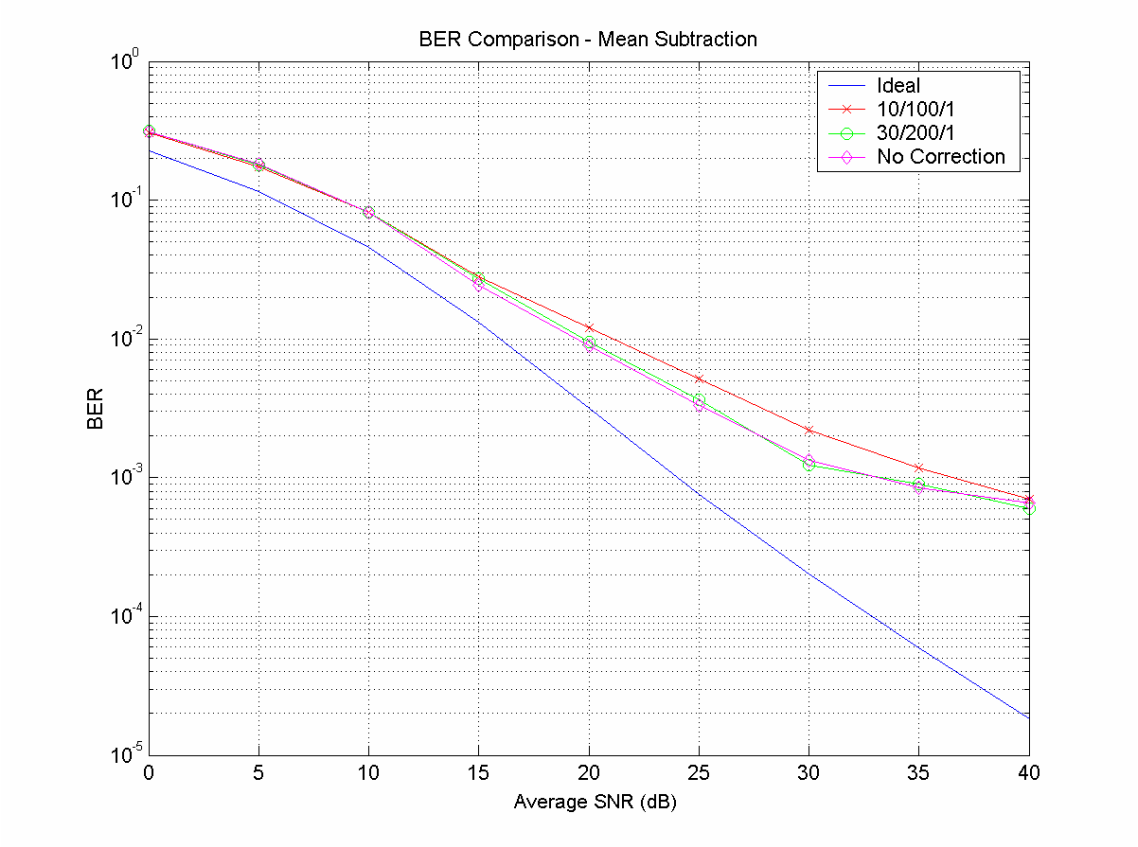

Figure 4. 20 – BER Comparison of Mean Subtraction

In Figure 4.13, we see that the 10 pole corrected case has significantly worse BER at medium and high SNR, however at 40dB SNR, we see convergence of all curves. The 30-pole case lines up exactly where the no correction curve is. This is very similar to the Add-to-First algorithm that we used in the previous section. But it also shows us that with a lower predictor sampling frequency, we have comparable results to not correcting the predictor, rather than having worse performance.

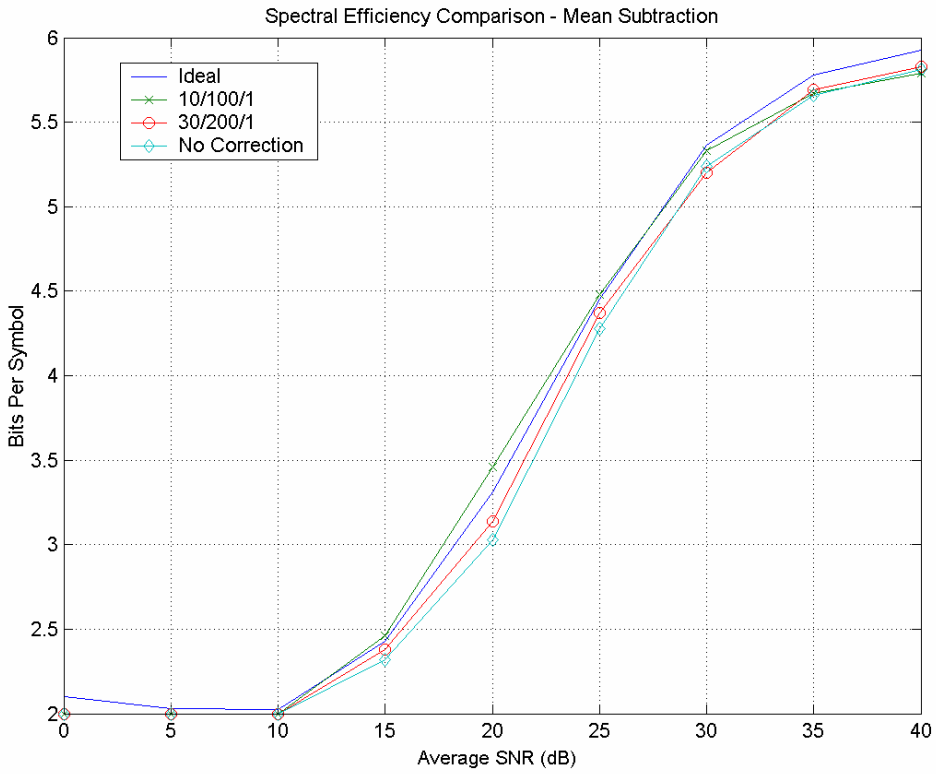

Figure 4. 21 – Throughput Comparison of Mean Subtraction

In Figure 4.14, we have a plot of spectral efficiency curves for mean-subtraction and no correction. As we expect, the weakest

case in terms of BER is the strongest one in terms of throughput. At medium SNR we see that the corrected linear prediction performs better than the one having no correction. As with the BER performance, the spectral efficiency for the Mean Subtraction method has gained enough to match performance of no correction. We believe this is due to the resistance of error propagation in the lowered sampling rate. What Mean Subtraction may be doing is adjusting the coefficients in such a way that it becomes severely suboptimal when predicting many samples ahead, but does not change the predictor in the short term.

Let us now compare the Mean Subtraction method to the Add-to-First method when using one sample per frame of information for the predictor.

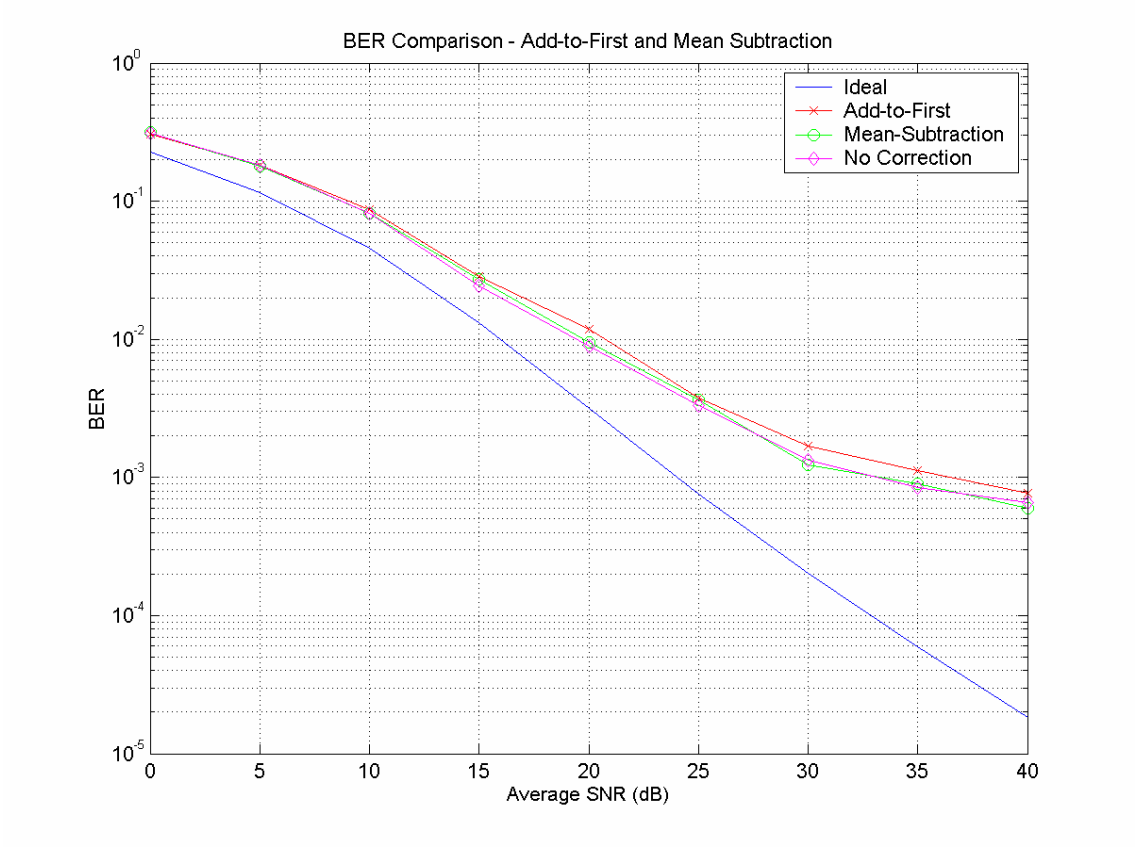

Figure 4. 22 – BER Comparison Add-to-First and Mean Subtraction

In Figure 4.22, we have the BER comparison between manipulating poles and subtracting means. It is clear that subtracting the mean gives us a better BER than manipulating poles in terms of BER, and we will now expect that the Add-to-First method will have better spectral efficiency.

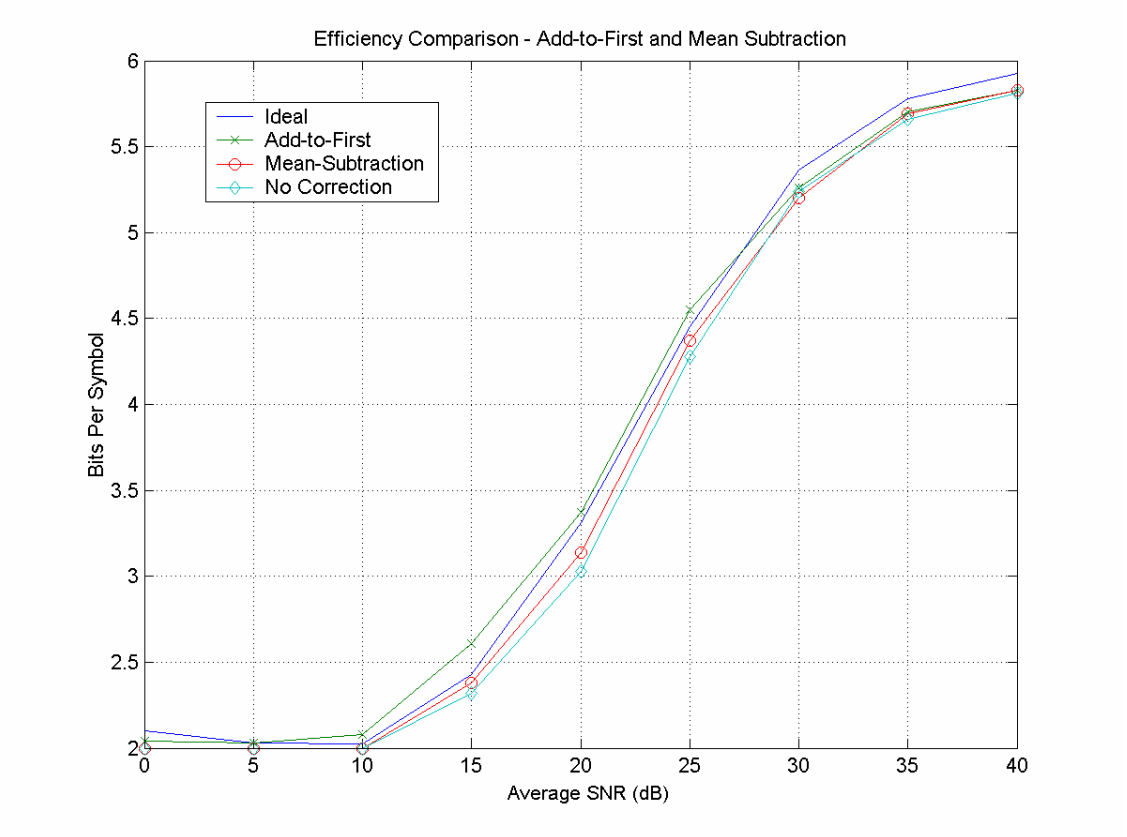

Figure 4. 23 – Throughput Comparison Add-to-First and Mean Subtraction

And in Figure 4.23, we do indeed get better efficiency with pole manipulation than we get with mean subtraction. Mean subtraction gives slightly better efficiency than no correction at all.

To summarize this section, it would seem that there is no clear way to make significant improvements in BER without taking a significant loss in spectral efficiency. However, by using the mean subtraction method of prediction, we can have a very small increase in spectral efficiency while simultaneously maintaining the BER. In the pole manipulation methods, we make a tradeoff

between BER and throughput; we degrade BER in order to increase the throughput.

4.2.4 Parallel Prediction

We have demonstrated that our adaptive modulation system when implementing linear channel prediction produced better results when the predictor samples the channel more slowly. Performance curves indicate that sample rates as low as one channel sample per frame can provide better BER and spectral efficiency. However, in our system, we use the pilot symbols from the FFT estimator to provide this information to the predictor, and the slower sampling rate means that there is information that is not used by the predictor. In this section, we are going to explore the possible benefits of using parallel prediction.

In the superior prediction scheme, as demonstrated in Section 4.2.1, the best performance resulted from using only one prediction pilot per frame. Therefore, there are many pilots that carry information that are not being used in our predictor. In parallel prediction, we try to harness these samples alongside the sample that is being used in the normal case. If we used one prediction sample in the normal predictor, the parallel predictor we will use two predictors and separate samples for each one. Figure 4.17 illustrates the difference between normal prediction and parallel prediction.

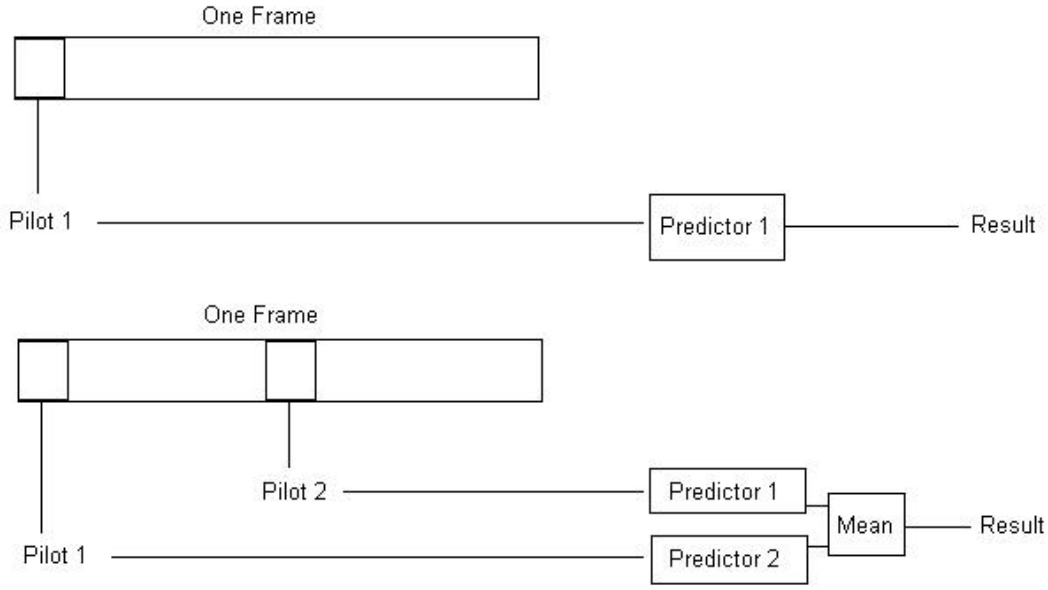

Figure 4. 24 – Diagram of Normal Prediction and Parallel Prediction

As shown in Figure 4.24, in parallel prediction, we take two samples from a frame and run them through two separate but identical predictors. The results from those predictors are averaged out and the result is the mean of the two (or more) predictors.

In the following figures, we will see how the parallel prediction idea affects performance.

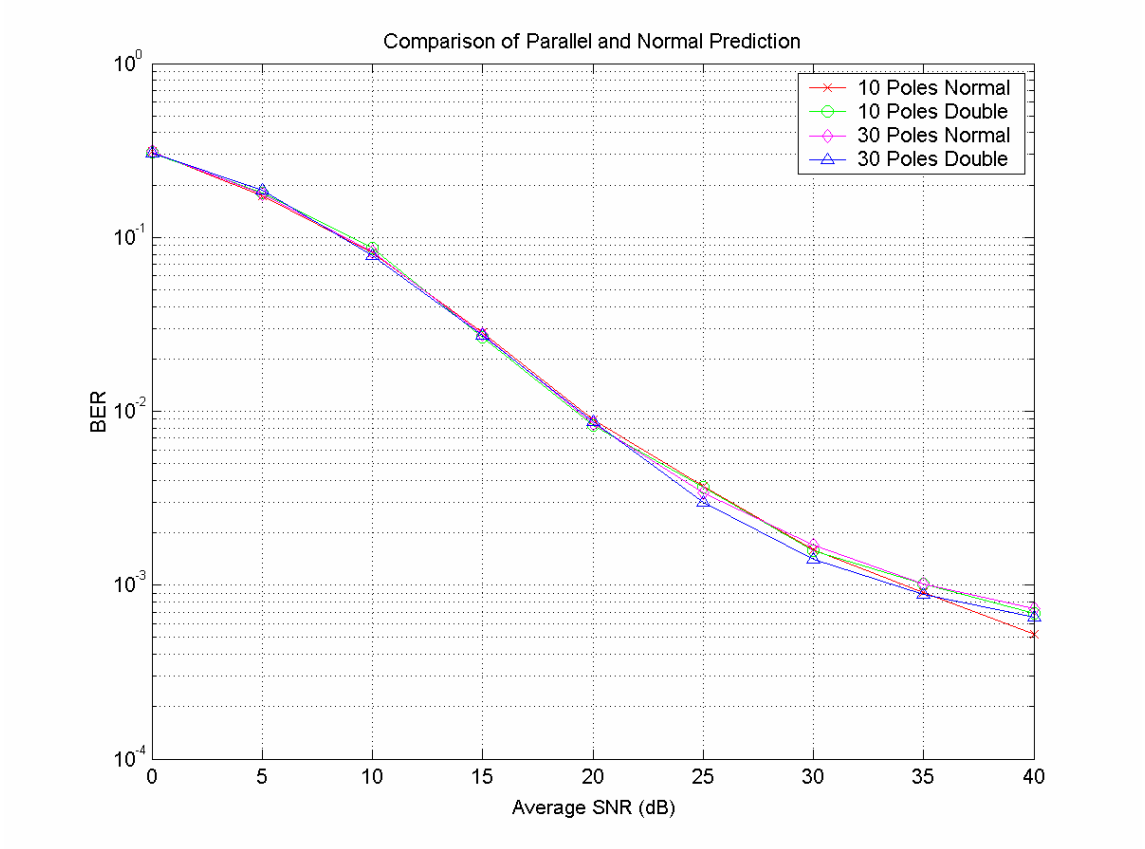

Figure 4. 25 – BER Performance between Normal and Parallel Prediction

Figure 4.25 shows us a set of BER performance curves for regular and parallel prediction. In this plot, the difference between the normal and parallel are the samples chosen to be the inputs of the predictor. In the normal version, the 5th sample (of 10) was collected for the predictor, but in the parallel version, the 3rd and the 7th bits were taken out. Now that we have detailed the difference, we can go on to discuss the result. It is clear that the performances of all curves are virtually identical up to medium-high SNR. The 30-pole double prediction case seems to be the best performing curve up to the last point, where the 10-pole

normal case seems to outperform. However, we believe this to be a result of simulation noise. So it would seem that double prediction has very little impact, if any. We say this because if there should be any significant benefits for the 30 pole case, we should also see it for the 10 pole case, which we do not. This leads us to believe that parallel prediction does not improve BER performance.

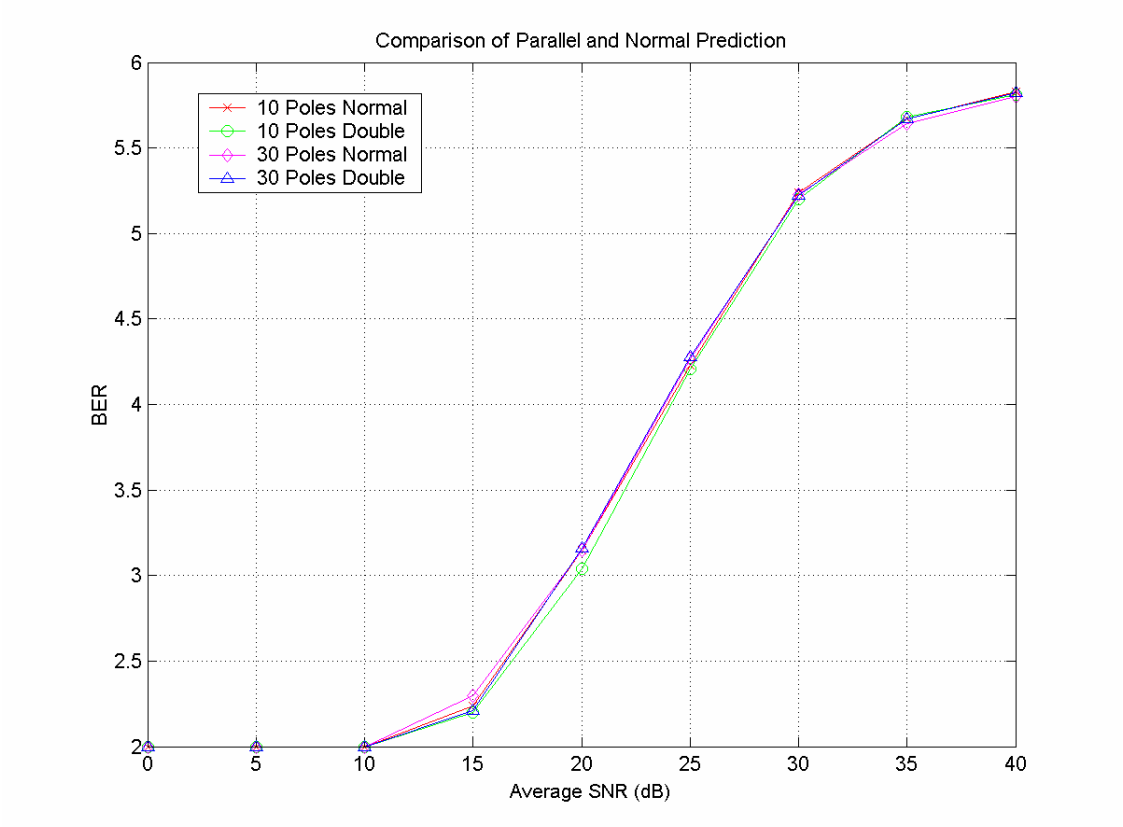

Figure 4. 26 – Spectral Efficiency Performance Normal/Parallel Prediction

In Figure 4.26, we see that the spectral efficiency between all curves is exactly the same. Parallel prediction does not alter the spectral efficiency in any way.

The previous case was a simulation of a system that did not alter the prediction poles in any way. Now let us see what happens when we run the Add-to-First pole alteration method.

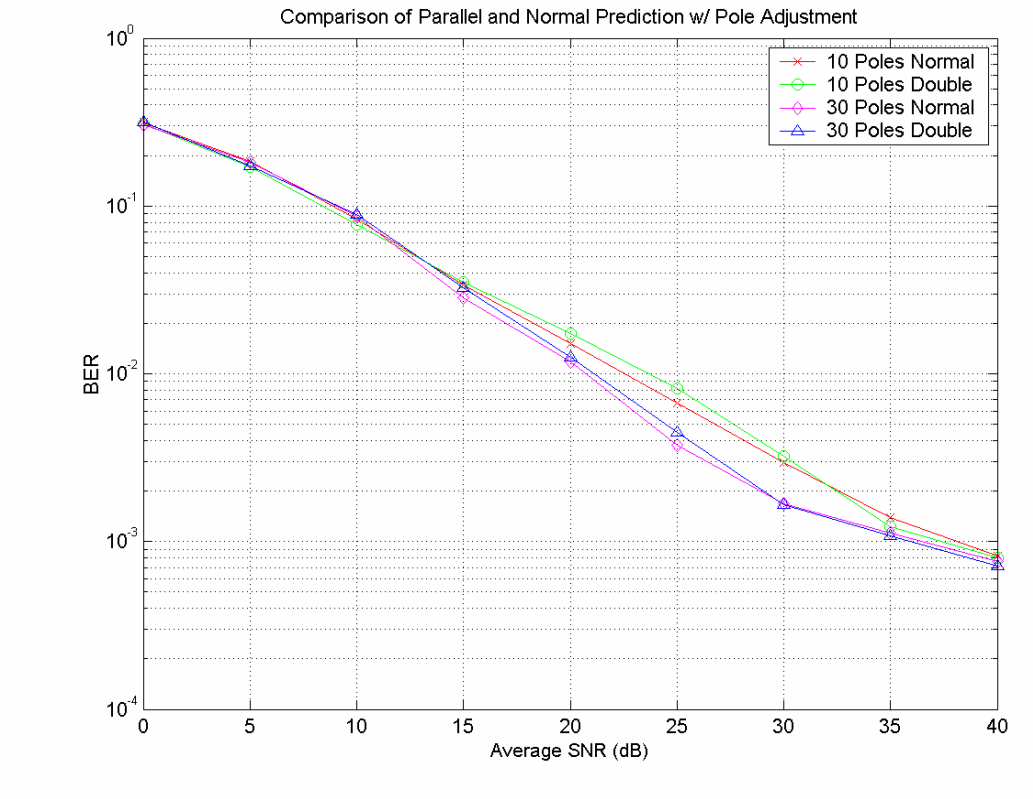

Figure 4. 27 – BER Performance with Pole Adjustment

When we include pole adjustment, we see that there is a large separation between the 10 pole and the 30 pole runs. However, it would seem here that like in Figure 4.27 the parallel prediction does not do much in terms of benefits. Both sets of curves show us that parallel prediction does not buy us anything in terms of BER. Let us move on to throughput analysis.

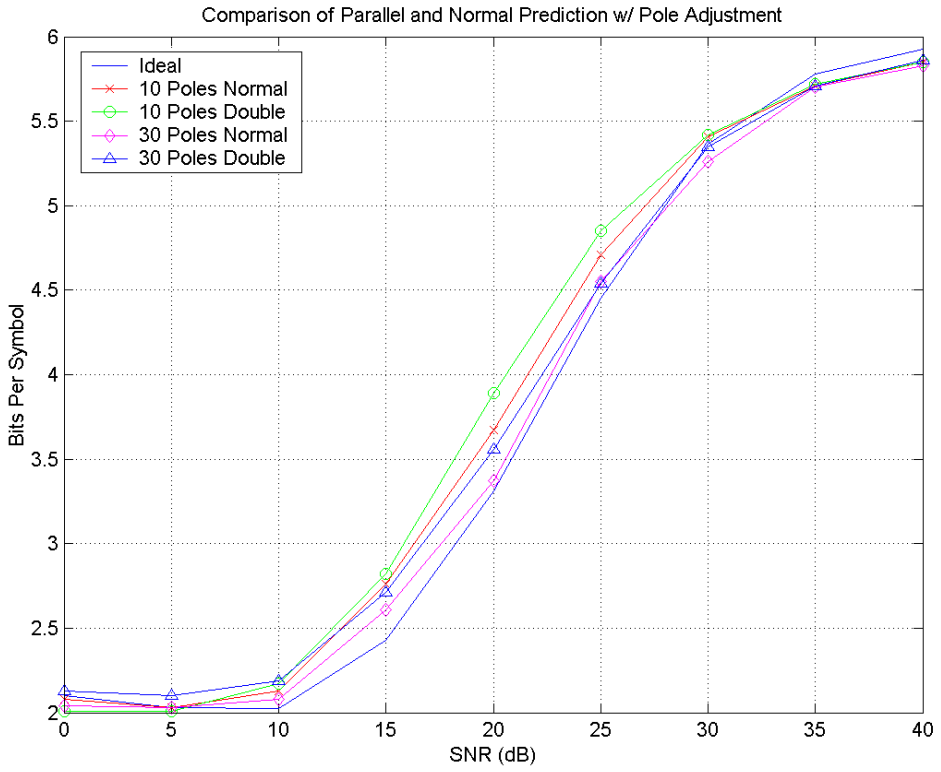

Figure 4. 28 – Spectral Efficiency with Pole Adjustment vs. Average SNR

In Figure 4.28, we have a more interesting result. It would seem that while our error rates have not changed much with parallel prediction, our spectral efficiency has. If we run pole adjustments with parallel prediction, we get a significant increase in throughput at lower and medium SNR. If we look at the BER at those same SNR ranges, they are virtually the same. This leads us to believe that there is a throughput advantage when we use parallel prediction.

4.3 Forward Error Correction Coding

The final area of performance enhancement that we will consider in our study of adaptive modulation is the use of forward error correction (FEC) coding. FEC coding has been used in many data transfer applications for decades now, ranging from wireless communications to optical disk reading [18]. The varieties of FEC codes include Reed-Solomon codes to BCH codes to convolutional codes to turbo codes. We will be using convolutional codes in our study.

In modern digital communications, virtually all data systems use some sort of FEC coding to guard against errors due to noise or interference. In general, convolutionally coded systems perform well in environments where bit errors occur independently, such as in AWGN channels. If a bit is flipped, there is information concerning that bit that is spread among adjacent bits. The decoding process is able to correct these isolated errors. However, should the channel experience a fade, then it is possible that several bits in succession will be in error and the decoder will not be able to correct them. To combat this, systems typically include an interleaver to scramble the bits after encoding and before decoding. This will help to redistribute errors caused by fades provided that the fade duration is shorter than the interleaver length.

The effects of coding have been documented extensively in research. However, we would like to examine the effects of coding on an adaptive modulation scheme in a Rayleigh fading channel. We would like to see how the performance of the system reacts to using prediction and adaptive modulation at the frame

level and coding at the block level (a block consisting of many frames). We have seen from Chapter 3 that higher Doppler rates decrease the effectiveness that prediction gives us. However, higher Doppler rates also provide coded systems an advantage since the interleaver will be able to distribute errors due to fades more effectively [19].

When we incorporated coding into our system, we used coding blocks of 4500 bits and a k = 7, r = ½ convolutional code. Also, rather than using hard decision decoding, we chose to use soft decision decoding. Soft decisions can provide a gain of up to 3dB over hard decisions [18].

Below are plots of our adaptive modulation system with coding.

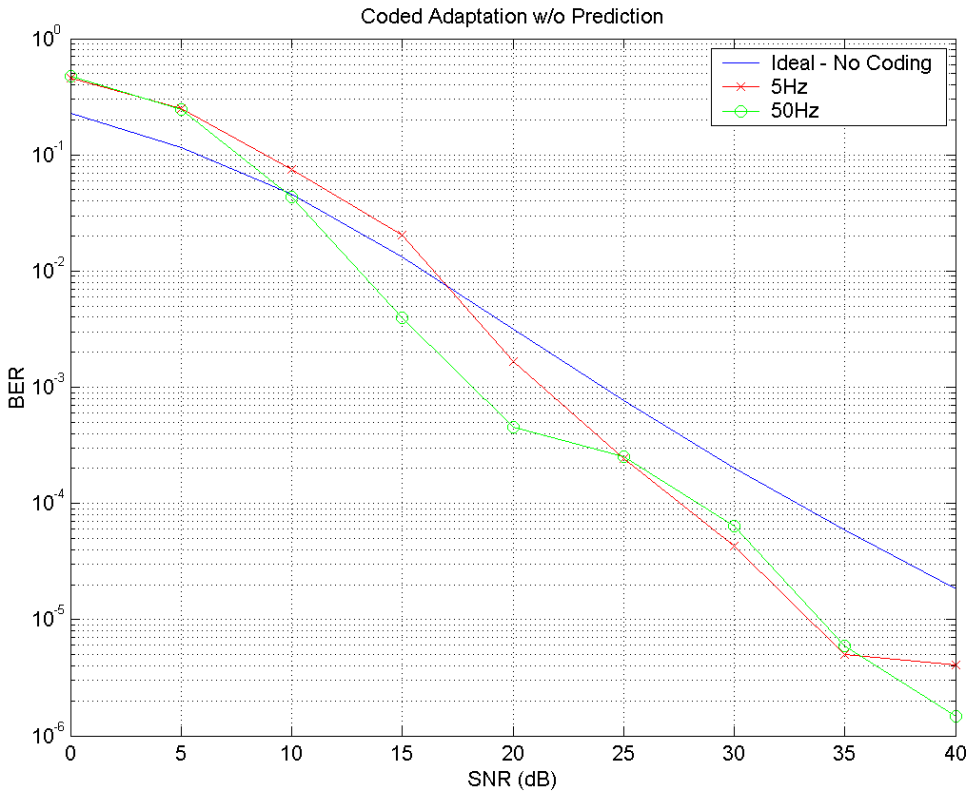

Figure 4. 29 – Coded Adaptation w/o Prediction vs. Average SNR

Figure 4.29 plots the BER performance of our adaptive modulation system with coding along with the ideal adaptive modulation case without coding. We observe at low average SNR the system operating at high Doppler rate outperforms the one in low Doppler. However, as average SNR increases, the performance becomes the same. This is due to the fading channel being the main error source rather than the receiver noise. It would appear that the coding advantage with high Doppler fading has been exhausted in high average SNR.

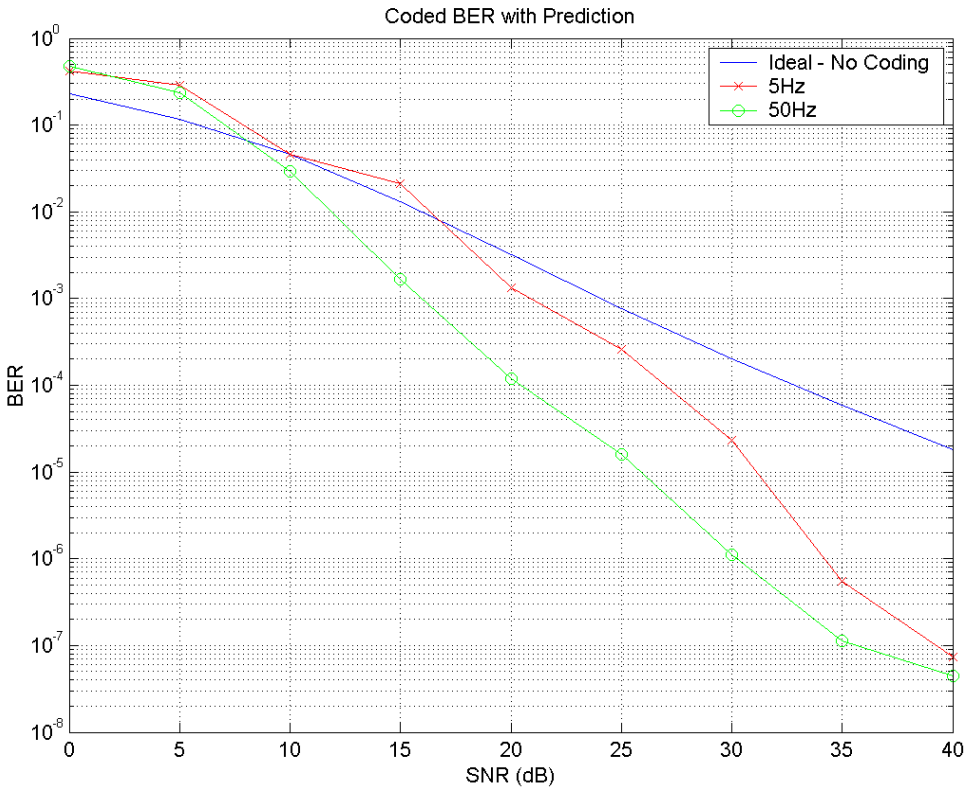

Figure 4. 30 – Coded BER with Prediction vs. Average SNR

Figure 4.30 plots the performance of the adaptive modulation system when prediction is applied. We immediately see that there is no overlap in the curves as there was in Figure 4.29. It would seem that the prediction was able to improve the performance of the adaptive modulation at 50Hz Doppler. Combined with the coding gain at 50Hz, the performance is significantly better than 5Hz. Figure 3.26 illustrates the difference between prediction and non-prediction at 50Hz Doppler. The 5dB gain we have with prediction in Figure 3.26 seems to correspond with the 5dB gain we see here in the coded system. And it appears that

the 5dB gain occurs in the SNR ranges of 25dB and above, exactly where the non-prediction case demonstrated even performance between 5 and 50Hz.

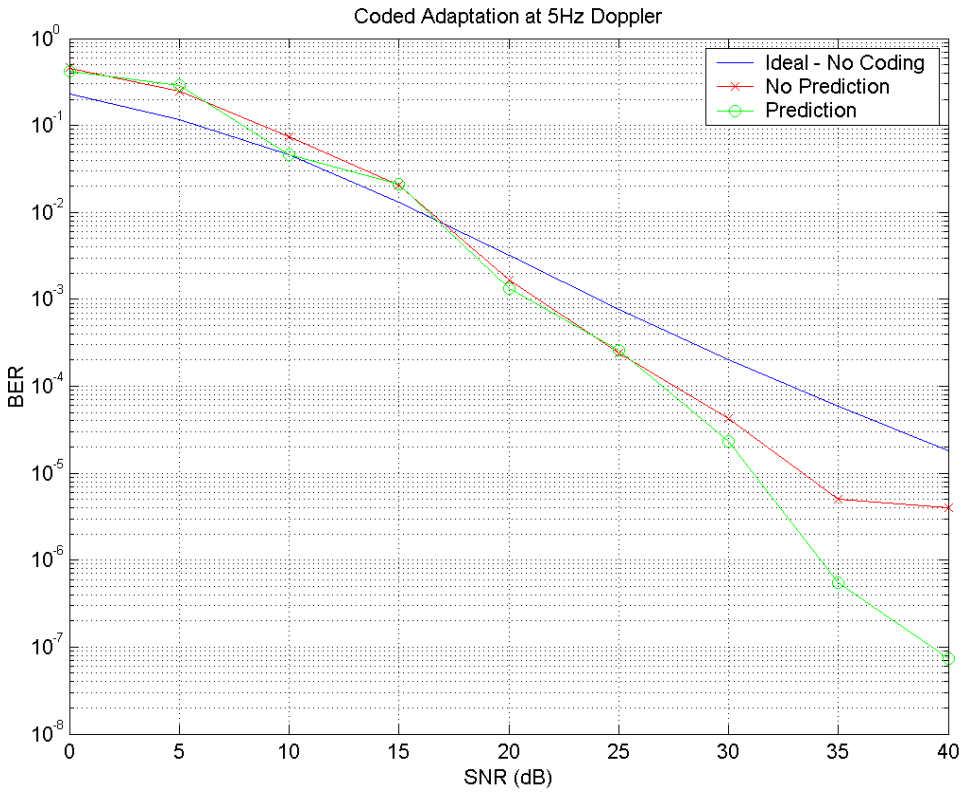

Figure 4. 31 – Coded BER at 5Hz vs. Average SNR

In Figure 4.31, we show a comparison between prediction and non-prediction when applied to the coded system at low Doppler rate. We immediately notice that they provide equal performance up to approximately 30dB, where the non-prediction curve reaches an error floor, but the prediction curve does not. Figure 3.24 shows that adaptive modulation with and without prediction have

the same performance up to 25dB average SNR. Here in the 5Hz coded situation, we see that they deviate around 30dB, which is where a 5Hz comparison without coding would deviate.

Let us now see how prediction and non-prediction measure up against one another at high Doppler.

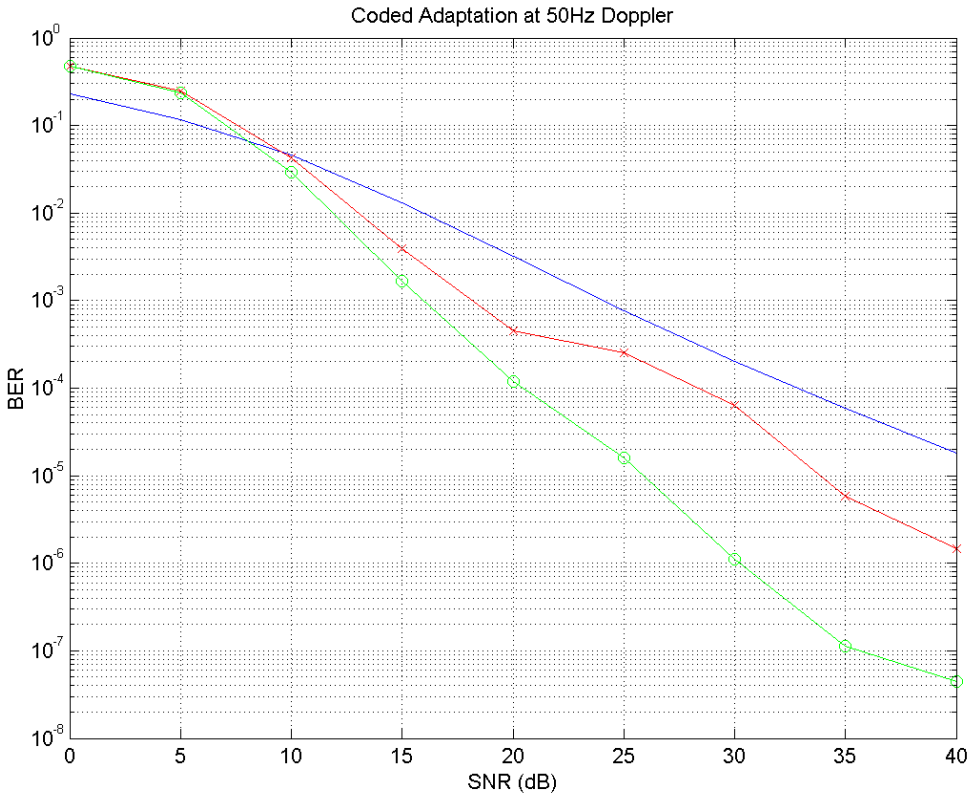

Figure 4. 32 – Coded BER Performance at 50Hz vs. Average SNR

Figure 4.32 compares the performance between prediction and non-prediction at 50Hz Doppler. It would seem that between 10 and 15dB average SNR, we observe a deviation, with the prediction

case being superior to no-prediction. If we refer to Figure 3.26, this is the range in which the uncoded adaptive system deviated as well. Unlike the low Doppler situations, we have gains up to 10dB when running at high Doppler and medium SNR ranges, as opposed to no gains until high SNR. The coding has become more powerful with the in-frame variations from high Doppler channels and has also benefited from prediction allowing better adaptation choices to be made.

To show that the gains from high Doppler and prediction are coming from the proper choices in adaptation and not from being spectrally inefficient, we present the throughput curves of the four cases we presented above.

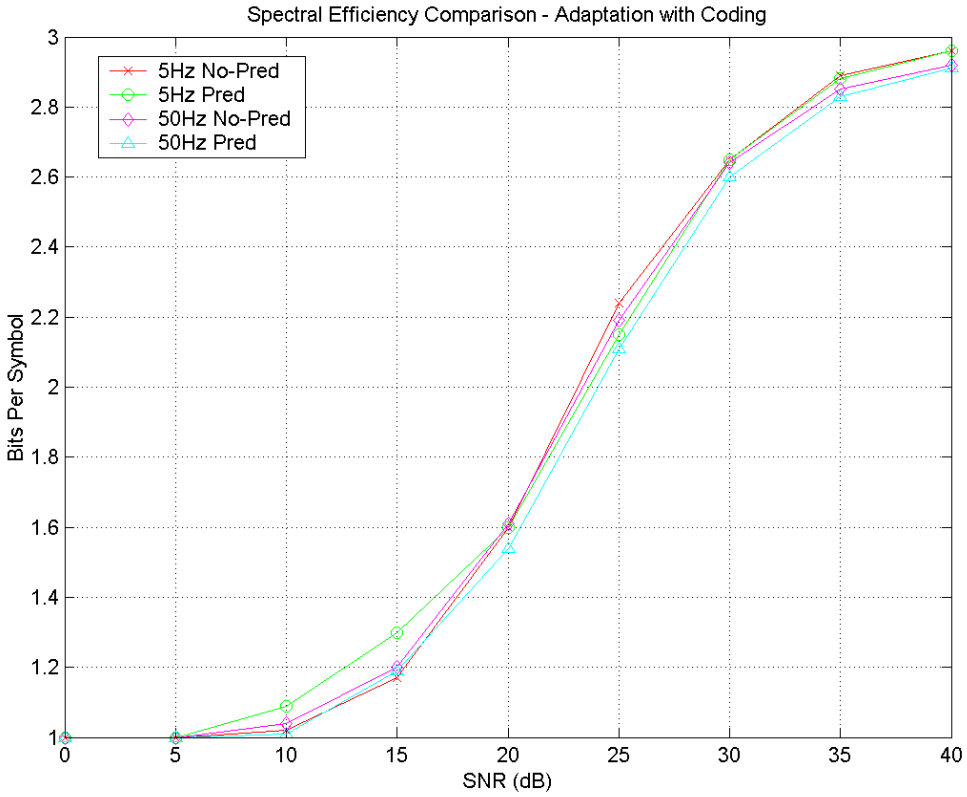

Figure 4. 33 – Spectral Efficiency of Coded Adaptive Modulation vs. Average SNR

As we can see in Figure 4.33, all of the throughput curves essentially overlap one another. This confirms that the performance gains of the predictive simulations have not come from continuously choosing lower modulation schemes. However, we do note that the spectral efficiency here is half of what it was without coding. In summary prediction is particularly powerful when coding is used. Coding tends to work best in moderate to high fading environments with a fixed interleaver depth. Since

prediction benefits the link the most in just these scenarios, it provides substantial improvement in coded cases.

4.4 System Performance with Real Channel Data

Throughout this thesis, we have relied on simulated channel values using the Jakes Model. This section presents the use of our system with data collected from real life situations. As it is a real channel, and not one that is being generated using a Rayleigh model, we may not have a channel that fits a Rayleigh distribution. Figure 4.34 depicts a small time sample of the data:

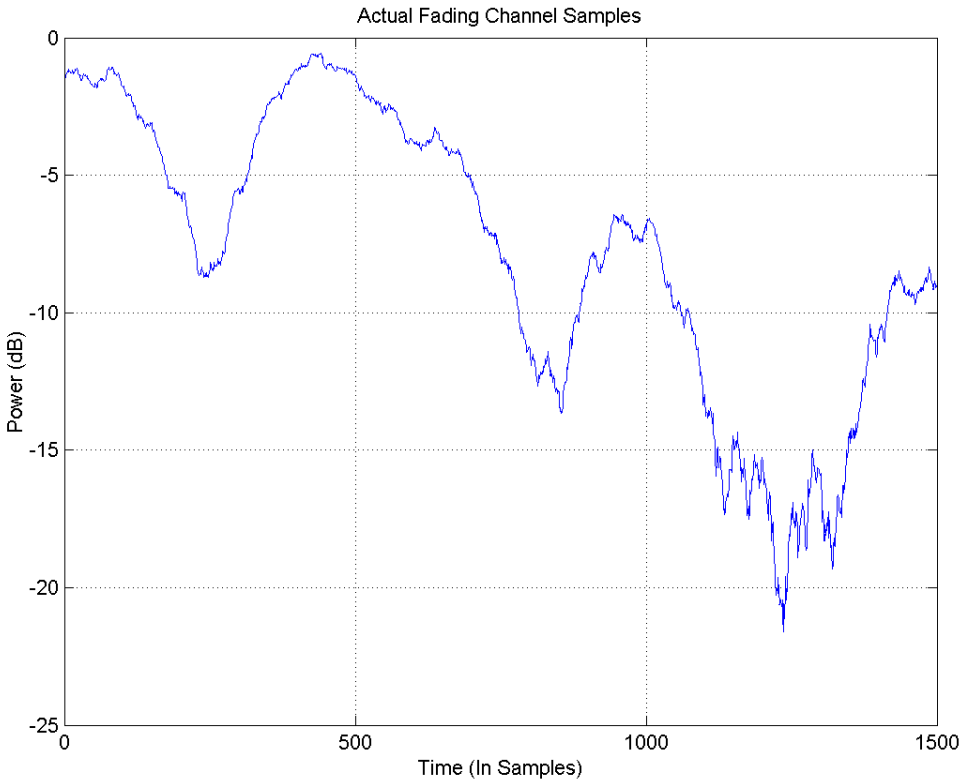

Figure 4. 34 – Real Channel Data

The data that Figure 4.34 shows is noisier than the channel data that was shown in Chapter 2. From our sample rate, we can determine the Doppler rate of this channel by taking the FFT of the channel. Figure 4.35 shows the result.

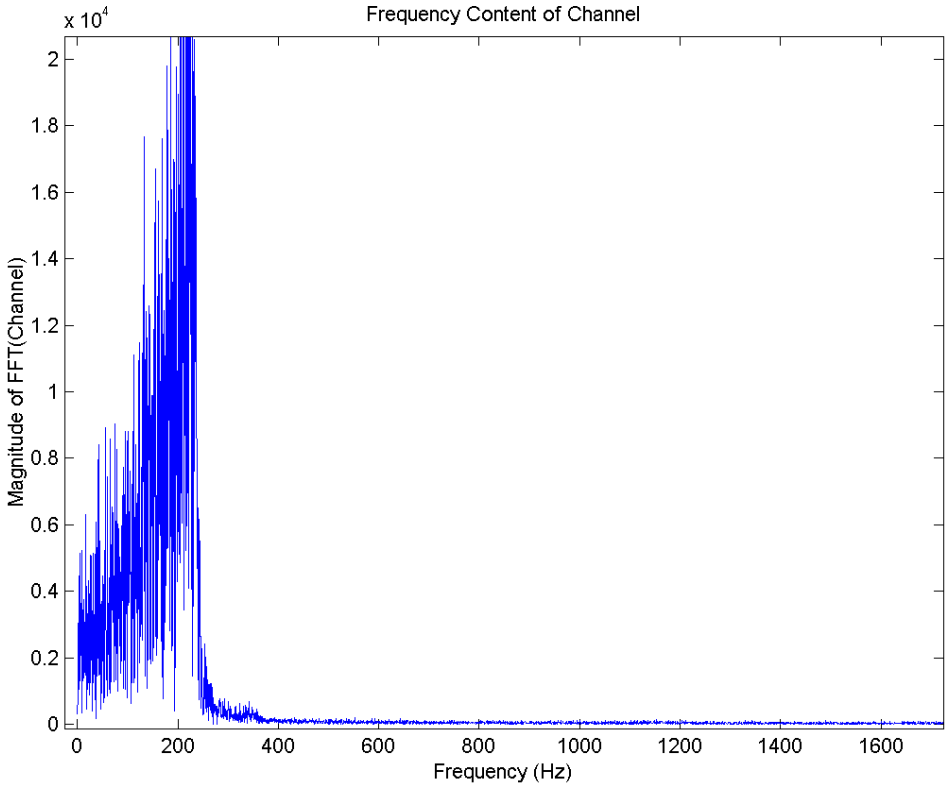

Figure 4. 35 – Frequency Content of the Channel

From Figure 4.35, we see that the Doppler rate of our channel is somewhere between 200Hz and 250Hz, which is faster than anything that we have simulated. To make this data work for models that we have used in simulation, we shall interpolate this data by a factor of 5 to give us a fading rate of 40Hz to 50Hz.

To confirm that this is a Rayleigh channel, we present a histogram of the distribution of the sampled data.

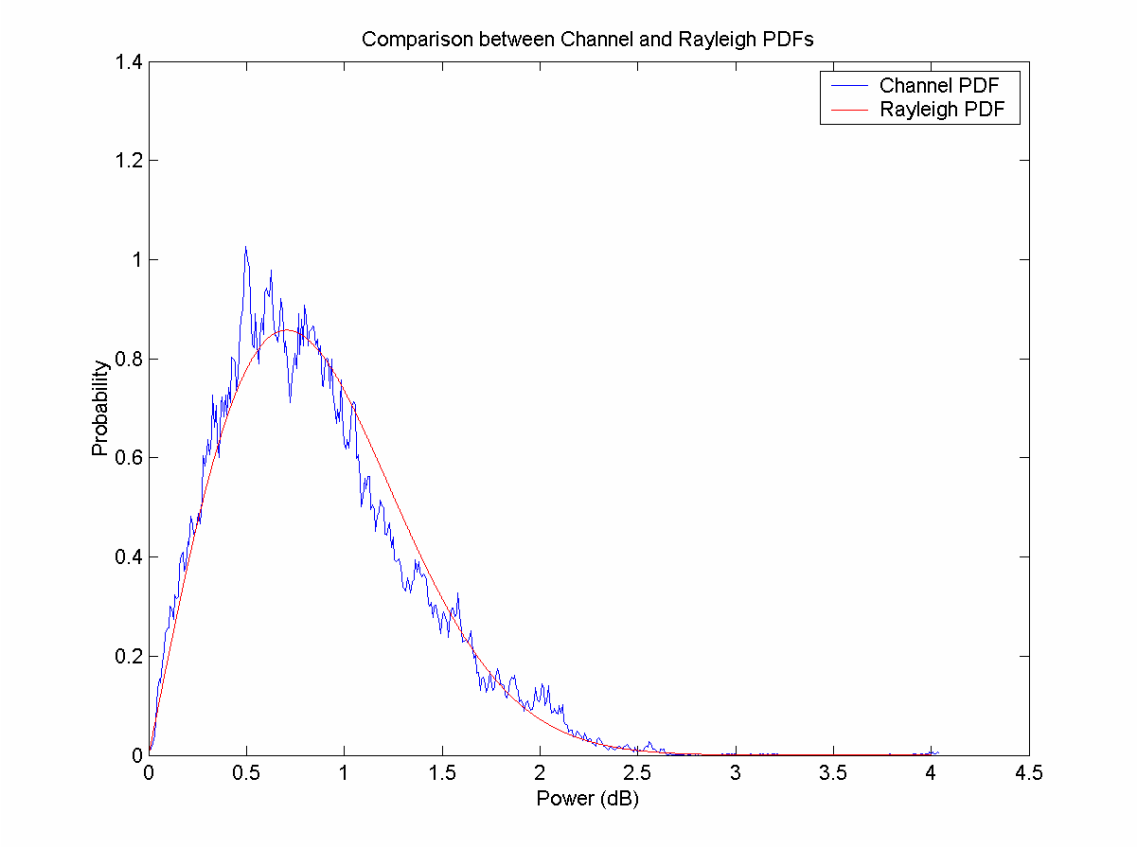

Figure 4. 36 – Histogram Showing the Distribution of the Channel is Rayleigh

Figure 4.36 plots the distribution of our Rayleigh channel. Although the curve is not perfectly Rayleigh, the channel is at least somewhat Rayleigh distributed.

Now that we have discussed the characteristics of the data that we used to create our channel, we can now run our system with the actual interpolated channel data.

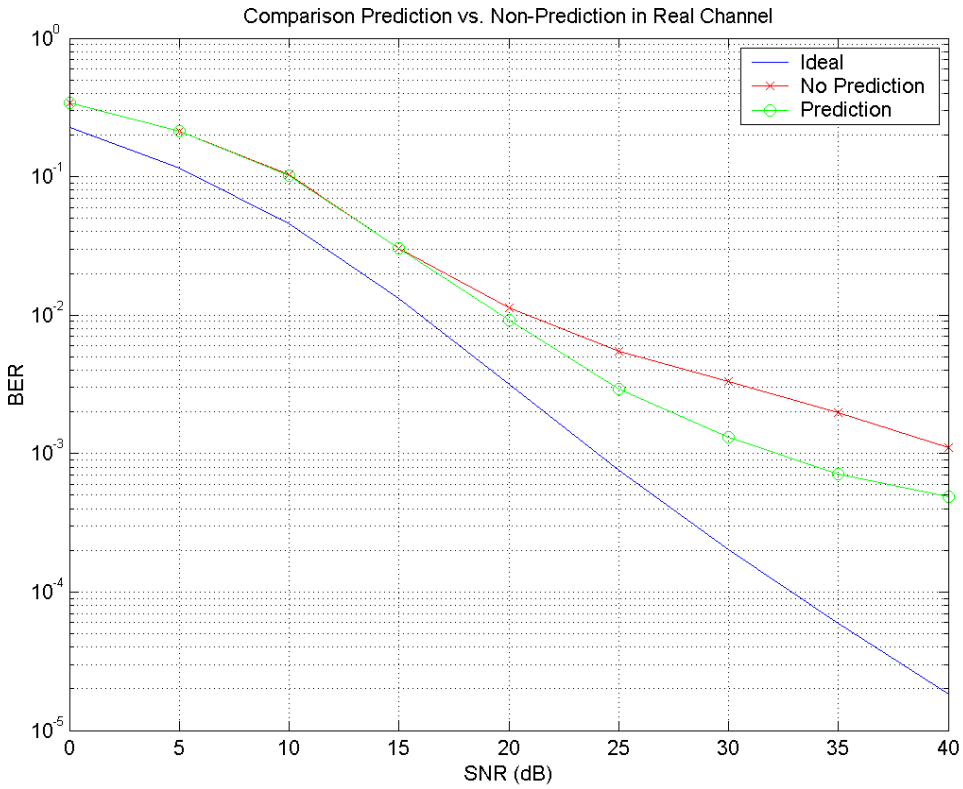

Figure 4. 37 – Predictive and Non-Predictive Adaptive Modulation in Real Channel Data

In Figure 4.37, we have a pair of performance curves that show the processing gain of predictive adaptive modulation. This plot shows that the prediction algorithm works similarly for this channel as it does for the simulated one.

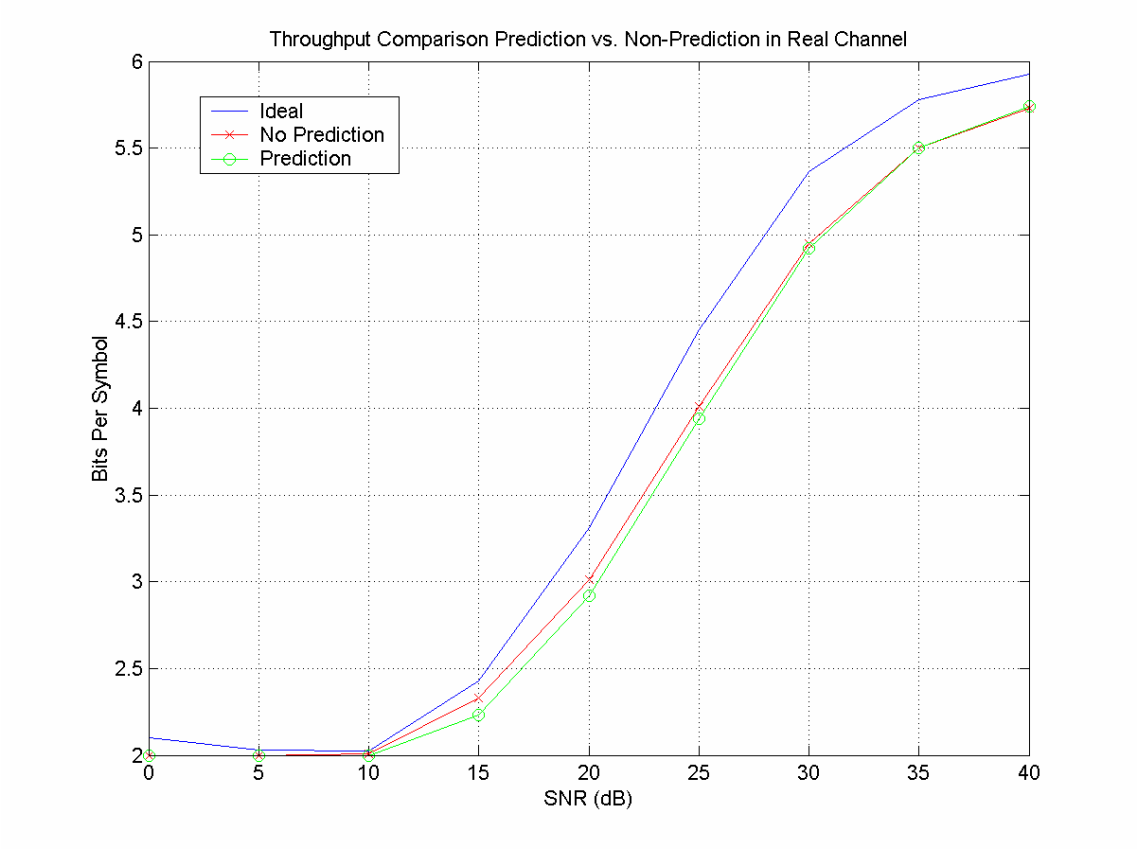

Figure 4. 38 – Throughput Comparison Between Prediction and Non-Prediction in real Channel Data

Figure 4.38 shows us the comparative spectral efficiency performance between adaptive modulation with prediction, without prediction, and the ideal performance. Clearly we see that the adaptive schemes are not attaining the efficiency of the ideal curve. The reason for this is not due to the predictor, as we note that the non-predictive adaptive modulation system suffers the same spectral efficiency shortcoming. The problem is with the SNR estimator. When the SNR estimator in the simulator is

set to perfectly estimate the SNR, we find that the efficiency problem is avoided.

4.5 Conclusions

In Chapter 4, we have presented the problem of bias in linear prediction. This bias was causing the spectral efficiency of the predictive systems to be significantly lower than that of non-predictive systems. We followed this by presenting various solutions to the bias problem. Included was slowing down the sampling rate of the predictor. Also, we looked at manipulating the poles of the predictor by changing their values. We further tried subtracting an auto-correlation weighted mean from the data set. And lastly, we tried to use multiple prediction paths in hopes of utilizing more data. From what we have done, slowing down the predictor sampling rate was the most effective.

The second part of the chapter dealt with adding convolutional coding to the adaptive system. We found that with coding, we could significantly improve the BER performance of our adaptive systems, especially at high Doppler rates. However, we had to cut our spectral efficiency in half in order to have these improvements. Additionally, we found that prediction provided larger gains in a coded system than in an uncoded system.

To summarize, we were able to overcome the bias problem by changing the sampling rate on the predictor and by using FEC, we greatly improved BER performance at the cost of half our spectral efficiency.

Chapter 5

CONCLUSIONS

5.1 Conclusions

In our work we have explored the improvements to adaptive modulation by using linear prediction to predict future channel power levels in periods of fast fading. The reason that it was worthwhile to explore predictive algorithms was to overcome the propagation delay in the feedback channel. We have shown that we can gain significant improvements in BER especially at higher Doppler rates. We explored the effects that SNR and FFT estimation had on our system as well.

At high Doppler rates, we observed that prediction had inherent bias difficulties and we proceeded to provide solutions to those problems. We presented methods that involved altering prediction coefficients, subtraction means of sample data, and changing the sampling frequency of the predictor.

Coding was also a subject of observation in this thesis. We saw that coding operated more effectively at high Doppler rates and could overcome the shortcomings of adaptive modulation in adverse channels.

The highlight of this thesis was the ability of our system to be able to operate on real channel data. Through the use of real channel samples, we were able to show the workability of our

system in a real environment. We showed that with real channels, we would be able to reap benefits from adaptive modulation with prediction over normal adaptive modulation and static modulation.

5.2 Future Research

There is still a lot of work that can be done with adaptive systems and linear prediction.

1) Our channel model consisted of a flat fading channel. It would be an interesting study to try prediction on frequency selective channels.

2) Linear prediction was the prediction method of choice in this paper. There are several other prediction methods that may be used. They include subspace methods, ESPIRIT, MUSIC, or nonlinear prediction.

3) Instead of adapting modulation, future research could include adaptive coding. Rather then changing modulation schemes, we could keep modulation static while changing coding rates.

4) In all simulations in this work, we assumed that the control channels were error free. It would be of interest to see how performance would be affected with an imperfect control channel.

We only dealt with a single link which does not include interference. A study on multiple users would be of interest in simulation.

BIBLIOGRAPHY

[1] J.M. Torrence, L. Hanzo, “Upper bound performance of adaptive modulation in a slow Rayleigh fading channel,” *IEEE Electronics Letters*, Vol. 32, pg 718, April 1996

[2] J. Pons, J. Dunlop, “Bit Error Rate Link Adaptation for GSM,” *The Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 1998, Volume: 3 , 8-11 Sep 1998 Page(s): 1530 -1534 vol.3

[3] P. Bender, et al, “CDMA/HDR: A Bandwidth Efficient High Speed Wireless Data Service for Nomadic Users”, *Communications Magazine*, IEEE, Vol. 38, No. 7, July 2000, pgs. 70-77

[4] P. Jain “On the Impact of Channel and Channel Quality Estimation on Adaptive Modulation” December 2002

[5] J. B. Andersen, et al, “Prediction of Future Fading Based on Past Measurements” *Vehicular Technology Conference*, VTC 99, Vol. 1, pg 151 – 155

[6] R. Roy, T Kailath, “ESPIRIT – Estimation of Signal Parameters via Rotational Invariance Techniques”, *IEEE Transaction on Acoustics, Speech and Signal Processing*, Vol. 37 No. 7, July 1989, pgs 984-995

[7] J.K. Hwang, J.H. Winters, “Sinusoidal Modeling and Prediction of Fast Fading Processes” *IEEE Globecom 98*, Sydney, Nov 1998 pgs 667-672

[8] A. Duel-Hallen, S. Hu, H. Hallen, “Long Range Prediction of Fading Signals” *IEEE Signal Processing Magazine*, May 2000, pg 62 - 75

[9] S. Hu, A. Duel-Hallen, “Combined Adaptive Modulation and Transmitter Diversity using Long Range Prediction for Flat Fading Mobile Radio Channels”, *GlobeCom 2001*, Vol. 2, Nov 2001, pgs. 1256 – 1261

[10] I. A. Glover, P.M. Grant, Digital Communications, Prentice Hall, 1998

[11] Theodore S. Rappaport, Wireless Communications, Prentice Hall, 2002

[12] Okamoto, et al, “Rayleigh Fading Compensation for 16 QAM Using FFT,” *IEEE Transactions On Vehicular Technology*, Vol. 48, No. 5, pg 1626, September 1999

[13] Okamoto, et al, “Rayleigh fading Compensation for QAM by Using FFT,” *Personal, Indoor and Radio Communications, 1996; Seventh international Symposium on PIMRC*, ’96, vol. 3, 15-18 Oct 1996, pg 1079-1082

[14] R.M. Gagliardi and C.M. Thomas, “PCM Data Reliability Monitoring Though Estimation of Signal-to-Noise Ratio”, *IEEE Transactions on Communications*, Vol. COM-16 pgs 479-486 June 1968

[15] T. Eycoez, A. Duel-Hallen, H. Hallen, “Prediction of Fast Fading Parameters by Resolving the Interference Pattern,” *Signals, Systems & Computers, Conference Record of the Thirty-First Asilomar Conference on*, Volume: 1, 2-5 Nov 1997 Page(s): 167 -171

[16] W. H. Press, et al., Numerical Recipes in C, Press Syndicate for the University of Cambridge, 1997

[17] G. B. Rybicki, W. H. Press, “Interpolation, Realization and Reconstruction of Noisy, Irregularly Sampled Data”, *The Astrophysical Journal*, vol 398, Oct 1992 pg 169-176

[18] S. Lin, D.J. Costello, Error Control Coding: Fundamentals and Applications, Prentice Hall, 1983.

[19] J. Hagenauer, “Viterbi Decoding of Convolutional Codes for Fading and Burst Channels”, in *Proc. 1980 Int. Zurich Sem*., March 1980 Pg. G2.1 – G2.7

VITA

Raymond Chan was born and raised in Cleveland, OH. In 2001, he received his Bachelor of Science degree in Electrical Engineering from Case Western Reserve University. Raymond proceeded to attend Virginia Tech to pursue a Masters degree in August of 2001. He joined the Mobile and Portable Radio Research Group at the university under Dr. R. Michael Buehrer in June of 2002. Raymond has presently accepted a position at Northrop Grumman in Maryland.

He has been down with ICHIRO since 2001.