COMMUNICATION RECEIVERS BASED ON MARKOV MODELS OF THE FADING CHANNEL

by

Michael Louis Bernie Riediger

A thesis submitted to the Faculty of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Electrical and Computer Engineering

University of Manitoba

Manitoba, Canada

July 2002

Copyright © 2002 by Michael Riediger

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-76858-9

Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

COMMUNICATION RECEIVERS BASED ON MARKOV MODELS OF THE FADING CHANNEL

BY

MICHAEL LOUIS BERNIE RIEDIGER

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University

of Manitoba in partial fulfillment of the requirements of the degree

of

Master of Science

MICHAEL LOUIS BERNIE RIEDIGER © 2002

Permission has been granted to the Library of The University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilm Inc. to publish an abstract of this thesis/practicum.

The author reserves other publication rights, and neither this thesis/practicum nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Abstract

This thesis investigates receiver structures for signals transmitted over mobile wireless channels. The specific channel considered is the slow fading, frequency non-selective Rayleigh fading channel. When signals are transmitted over this channel, the received signal can be modeled as the transmitted signal multiplied by a correlated complex time-varying function, plus additive white noise. The complex time-varying function is referred to as fading, and commonly the Jakes-Clarke model is used to describe the statistical nature of the fading.

The correlation in the fading induces a memory dependence on the received signal. For optimum decoding in a digital communication system, the memory in the received signal must be utilized. With the Jakes-Clarke channel, the length of dependence in the received signal is infinite, so true optimum receivers are not practical to implement. Usually the Jakes-Clarke fading is modeled as an ARMA process, and finite memory dependence in the received signal is assumed. Under these conditions, MAP receivers can be implemented, but their complexity grows exponentially with the assumed memory length in the received signal.

An alternative is to model the Jakes-Clarke fading as a *finite state Markov channel* (FSMC). As optimum receiving techniques are already known for decoding on FSMC, good receiver performance can be expected if a FSMC can well represent the statistical properties of the true channel. The state space of the FSMC presented in this thesis are determined by the statistical nature of the complex fading. Although the complexity of FSMC receivers grows exponentially with Markov model order, it only grows linearly with the number of Markov states.

The goal of this research is to compare FSMC receivers to previously established reference receivers in terms of their error performance. Although a 2nd order Markov receiver will generally out perform a 1st order Markov receiver when the number of states is normalized, the increase in performance comes with an exponential increase in receiver complexity. Simulation results suggest that 1st order Markov receivers have a performance comparable to the reference or 2nd order Markov receivers, provided a sufficient number of Markov states is used.

Acknowledgments

Appreciation first goes to my supervisor, Prof. Ed Shwedyk who has consistently provided thorough answers to my questions. He has always given invaluable advice when I needed it, regardless of the context. His disciplined, yet down-to-earth approach to teaching is an inspiration to pursue greater academia.

Financial assistance is greatly acknowledged from the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Faculty of Engineering, the Faculty of Graduate Studies, and the Department of Electrical and Computer Engineering.

Tremendous thanks beyond the price of a coffee goes to David Engel for the many, many hours spent helping me edit this thesis. His ability to understand what I am trying to say and restructuring it so I am actually saying it has been extremely valuable. In addition, I would also like to thank Aimee Betker for further review of this document and for introducing me to the art of falafels.

Finally, I would like to express gratitude to my parents. Without their unconditional support and patience, successful completion of my Masters would not have been possible. This appreciation is given despite my mom never knowing precisely what I am taking in school, my dad wishing I tucked my shirt in more, and both of them abandoning me in the town of Transcona.

Table of Contents

Table of Contents iii

List of Figures vi

List of Tables viii

List of Acronyms ix

Chapter 1 Introduction 1

1.1. Digital Communications 2

1.2. Techniques to Combat Adverse Channel Effects 5

1.3. Objectives 7

1.3.1. Background and Related Work 7

1.3.2. Objectives of this Research 10

1.4. Organization of Thesis 10

Chapter 2 Digital Communications over Wireless Channels 13

2.1. Multipath Phenomena 13

2.1.1. Classification of Multipath Channels 16

2.1.2. Jakes-Clarke Model of the Multipath Channel 21

2.2. Digital Communication over Fading Channels 23

2.2.1. Transmitter and Channel Model 23

2.2.2. Demodulator and Decoder 25

2.3. Channel Simulation 28

2.4. Summary 31
Chapter 3 Memory-Based Detection over Rayleigh Fading Channels 33
3.1. Decoding Criterion 33
3.1.1. MAP Sequence Detection 34
3.1.2. MAP Symbol-By-Symbol Detection 36
3.2. MLSE Receivers 37
3.2.1. Kalman Filtering 38
3.2.2. Prediction-filtering 42
3.3. Results 45
3.3.1. Kalman Filtering 46
3.3.2. Prediction-filtering 49
3.4. Summary 52
Chapter 4 Finite State Markov Channel 54
4.1. Introduction 55
4.2. Finite State Markov Channel (FSMC) 57
4.2.1. FSMC as a Model of the Rayleigh Fading Channel 58
4.2.2. Markov Model Construction 60
4.3. MAP Sequence Detection and Channel State Estimation 65
4.3.1. MAP Sequence Detection 66
4.3.2. Joint MAP Sequence Detection and Channel Estimation 69
4.3.3. MAP Channel State Estimation 71
4.4. Simulation Results 73
4.4.1. Sequence Detection Results 74
4.4.2. Channel Estimation Results 81
4.5. Quadrature Phase Shift Keying over Phase-compensated Channels 86
4.6. Summary 90

Chapter 5 Conclusions and Suggestions for Further Study 91

5.1. Discussion 91

5.2. Suggestions for Further Study 94

Appendix A: Derivation of (2.19) 97

Appendix B: Derivation of (3.19) and (3.20) 99

Appendix C: Derivation of (3.25) and (3.26) 101

Bibliography 103

List of Figures

Figure 1.1: Elements of a traditional communication system 3

Figure 1.2: Simplified communication system used in this thesis.....5

Figure 2.1: A mobile wireless communication environment.....14

Figure 2.2: Mobile and arrival of angle of a single ray 15

Figure 2.3: Example of a power intensity profile and its delay spread 18

Figure 2.4: Example of a Doppler spectrum and its coherence bandwidth 19

Figure 2.5: Normalized Jakes-Clarke autocorrelation 22

Figure 2.6: Baseband Jakes-Clarke Doppler spectrum.....23

Figure 2.7: Generation of sent signal using shaping filter.....24

Figure 2.8: Model for mobile wireless channel 24

Figure 2.9: Model for the frequency non-selective channel 25

Figure 2.10: Matched filter demodulator.....26

Figure 2.11: Discrete-time model of the communication system.....27

Figure 2.12: OOK signal constellation 29

Figure 2.13: Block diagram of the fading simulator.....30

Figure 3.1: Kalman filtering combined with a Viterbi algorithm decoder.....41

Figure 3.2: Prediction-filtering combined with Viterbi algorithm decoder.....44

Figure 3.3: Error performance of Kalman filtering receivers $(f_D T_S = 0.01)$ 47

Figure 3.4: Error performance of Kalman filtering receivers $(f_D T_S = 0.05)$ 48

Figure 3.5: Error performance of Kalman filtering receivers ($f_D T_S = 0.1$) 48

Figure 3.6: Error performance of prediction-filtering receivers ($f_D T_S = 0.01$) 50

Figure 3.7: Error performance of prediction-filtering receivers ($f_D T_S = 0.05$) 50

Figure 3.8: Error performance of prediction-filtering receivers ($f_D T_S = 0.1$) 51

Figure 4.1: General example of a Markov process 57

Figure 4.2: Gilbert-Elliott 2-state Markov channel 57

Figure 4.3: Quantization of 1st order density for 4-state 1st order state-space..... 64

Figure 4.4: VQ of 2nd order density for 16-state 2nd order state-space..... 64

Figure 4.5: 2nd order state-space from CP of two 4-state 1st order state-spaces..... 74

Figure 4.6: BERs of 1st order Markov receivers for slow fading channel 75

Figure 4.7: BERs of 2nd order CP Markov receivers for slow fading channel..... 75

Figure 4.8: BERs of 2nd order VQ Markov receivers for slow fading channel..... 76

Figure 4.9: BERs of 1st order Markov receivers for medium fading channel..... 77

Figure 4.10: BERs of 2nd order CP Markov receivers for medium fading channel 78

Figure 4.11: BERs of 2nd order VQ Markov receivers for medium fading channel 78

Figure 4.12: BERs of 1st order Markov receivers for fast fading channel 79

Figure 4.13: BERs of 2nd order CP Markov receivers for fast fading channel..... 80

Figure 4.14: BERs of 2nd order VQ Markov receivers for fast fading channel..... 80

Figure 4.15: CEEs of 1st order Markov receivers for slow fading channel..... 82

Figure 4.16: CEEs of 2nd order VQ Markov receivers for slow fading channel 83

Figure 4.17: CEEs of 1st order Markov receivers for fast fading channel..... 84

Figure 4.18: CEEs of 2nd order VQ Markov receivers for fast fading channel 84

Figure 4.19: QPSK signal constellation 87

Figure 4.20: Quantization of 1st order density for 4-state 1st order state-space..... 88

Figure 4.21: BERs of Markov receivers for phase-compensated channel 88

List of Tables

Table 2.1: Mobile wireless channel classifications 20

Table 3.1: Reference receivers for channels with different fading rates 53

Table 4.1: LBG algorithm used to determine a Markov state-space 63

Table 5.1: Comparing receiver complexity by number of VA states 93

List of Acronyms

AGC	automatic gain control
AM	amplitude modulation
AR	auto-regressive
ARMA	auto-regressive moving average
AWGN	additive white Gaussian noise
BCJR	Bahl-Cocke-Jelinek-Raviv
BER	bit error rate
CP	Cartesian product
CEE	channel estimation error
DPSK	differential phase shift keying
DQPSK	differential quadrature phase shift keying
FIR	finite impulse response
FSMC	finite state Markov channel
GMRS	general mobile radio service
GPS	global positioning system
ISI	inter-symbol interference
LBG	Linde-Buzo-Gray
LCR	level crossing rate

LOS	line of sight
MAP	maximum *a posteriori*
ML	maximum likelihood
MLSE	maximum likelihood sequence estimation
MMSE	minimum mean squared error
OOK	on-off keying
PCU	personal cellular unit
PLL	phase locked loop
PSP	per-survivor processing
PSAM	pilot symbol assisted modulation
PSK	phase shift keying
PTAM	pilot tone assisted modulation
QAM	quadrature amplitude modulation
QOS	quality of service
QPSK	quadrature phase shift keying
SNR	signal-to-noise ratio
SMIA	Steiglitz-McBride iteration algorithm
US	uncorrelated scattering
VA	Viterbi algorithm
VQ	vector quantization
WLAN	wireless land area network
WSS	wide sense stationary

Chapter 1

Introduction

Digital communications, specifically wireless digital communications, have become vital to the transfer of information in the last decade. The physical realization of this wireless technology is getting smaller, more efficient and less expensive for the end user. The scope of possible applications is vast and continues to grow. Particular growth can be noted in the demand for mobile wireless communications. Mobile wireless communication is defined as the ability to communicate from anywhere at anytime. The focus of this thesis is the design of simple communication devices for use in a mobile wireless environment.

The communication path between the transmitter and receiver in a stationary wireless communication environment is often allowed to be partially obstructed. Although this could produce a weak signal, there are ways of improving the quality of the communication. Since both the transmitting and receiving devices are stationary, power is not a limited resource. All that is required to combat a poor channel is an increase in the transmission power. In stationary wireless communications, such as a *wireless land area network* (WLAN), the quality of the channel can be predetermined. This knowledge, coupled with an abundant supply of power, generally results in a satisfactory quality of communication.

For mobile wireless communications, such as a *personal cellular unit* (PCU) or a *general mobile radio service* (GMRS), the adverse effects of a poor communication channel are much more difficult to handle. Since the user is in motion, the negative channel effects are time variant, and power resources are limited. This proves to be a challenging environment in which to maintain a cost efficient, reliable and physically realizable communication system, and hence is the motivation for this research. This chapter reviews digital communication terminology, and presents a few techniques known to increase the quality of communication over mobile wireless channels.

1.1. Digital Communications

There are several elements to consider in a digital communication system, which vary according to the requirements of the application and the resources available. In any communication system there is a correlation between the complexity and the overall quality of the end-to-end communication. For example, a *global positioning system* (GPS) may transfer small quantities of data with a high emphasis on data integrity, whereas a real-time cellular technology demands simple and fast communication with a tolerable level of data corruption. While the building blocks of a communication system are dependent on the design specifications, generally the traditional elements are as shown in Figure 1.1.

Every communication system begins with a data source that may be analog or digital in nature. To reduce the amount of data to be transmitted, redundancy is often removed from the source by the source encoder. The entropy of the data source, however, limits the amount of redundancy that can be removed. The fundamentals of entropy and channel capacity were first introduced and discussed by Shannon [44] in 1948. The source decoder at the receiver performs the inverse operation and adds the

redundancy back. Regardless of whether the data source is analog or digital, the output of the source encoder is digital, and typically binary.

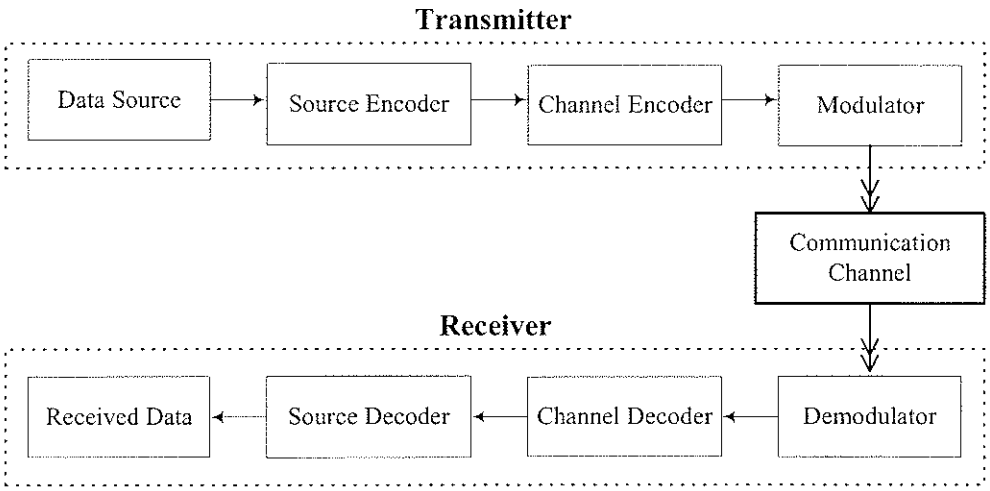

Figure 1.1: Elements of a traditional communication system

The data to be transmitted over the communication channel must be converted to analog form by the modulator, which maps consecutive binary digits (or bits) to waveforms. Since the transmitted signal is analog in nature, it is subject to channel effects and arrives at the receiver in a corrupted form. Different communication environments exhibit different channel effects, with the most common effects being additive thermal noise, attenuation and interference from other communication schemes. In the presence of these degrading effects, the receiver must attempt to decide what waveform was transmitted given the received signal. An error is defined as the incorrect identification of the received waveform. Although the modulator's signal constellation can be chosen to minimize the effects of the channel, due to the random nature of the channel, errors are inevitable. To reduce the number of errors that are propagated through to the end user, redundancy is added to the modulator's input bit stream by the channel encoder.

Ideally, the redundancy added to the transmitted sequence should enable the receiver to readily reconstruct the original data sequence with a limited amount of errors. This role of the channel encoder is known as error control coding. The corresponding channel decoder attempts to detect and repair data that is demodulated in error. Convolutional and linear-block codes are two traditional error control coding schemes. Two such examples of these are turbo codes and low-density parity check codes [6] [18]. The recent popularity of these schemes is primarily due to the simple decoding techniques they are associated with, as well as their near-optimum performance. While it may at first appear that the source encoder and channel encoder perform opposing operations, the redundancy added by the channel encoder differs from that removed by the source encoder in that the added redundancy is actually specific to the corresponding channel.

The traditional building blocks of a digital communication system are shown to be independent in Figure 1.1 for simplification and clarity. In reality there is often cooperation between these blocks in order to optimize the end-to-end system performance. This inter-dependence is evident in the research areas of joint source channel coding [21] [56] and modulation schemes with memory [52].

Source compression and error control coding are not required elements but are often included to enhance system performance. Keeping with the primary focus of this thesis, which is the design of the receiver, source and channel encoding will not be further investigated. Therefore, the vital components of a wireless communication system considered are the modulator, the demodulator, and the final decoder used to interpret the output from the demodulator. The simplified block diagram of this communication system is shown in Figure 1.2.

The fundamental elements of a communication system and the motivation for using each have been stated in this section. Each block in the transmitter and receiver

provide an increase in the quality of communication experienced by the end user. The one block of the system that cannot be changed or modified is the actual channel. The transmitter and receiver must be designed to work with the known channel. Common effects of the communication channel on a transmitted signal are discussed in the next section, as well as methods used to increase communication performance.

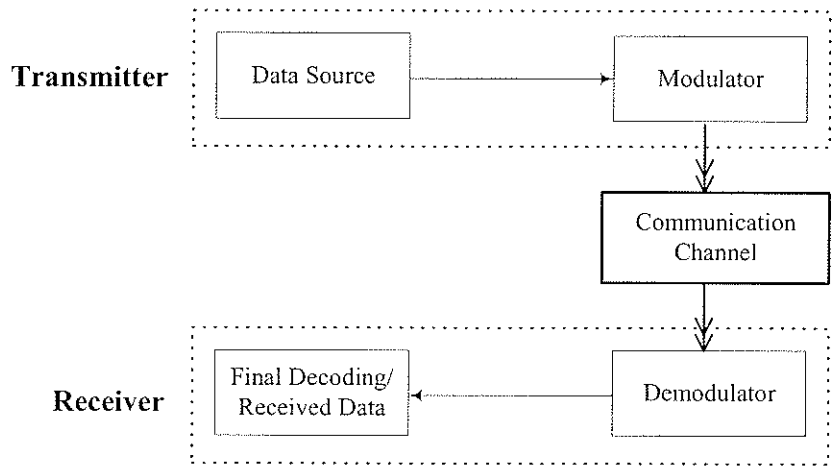

Figure 1.2: Simplified communication system used in this thesis

1.2. Techniques to Combat Adverse Channel Effects

The channel is the inhibiting block in a communication system that prevents error-free performance. In practical communication systems, the properties of the effects that are unique to the channel are often known prior to the system design. Knowledge of the channel characteristics is key to designing a suitable modulator/demodulator pair. Some of the most common adverse channel effects, as well as some practical solutions are discussed in this section.

Amplitude and phase uncertainty in the received signal are common channel effects. This phenomenon is discussed in Chapter 2. There are several methods that can be employed to estimate or track phase and amplitude if the change in phase and/or

amplitude is sufficiently slow relative to the data rate. These methods used to resolve channel amplitude and phase information can be organized into three distinct classes:

i) Receiver recognizes a pre-determined signal that is transmitted periodically,

ii) Receiver estimates phase and amplitude uncertainty regardless of data modulation, and

iii) Receiver incorporates uncertainties using a mathematical channel model, when converting the received signal into data symbols.

Two common examples of i) are *pilot symbol assisted modulation* (PSAM) and *pilot tone assisted modulation* (PTAM). PSAM is a technique where known symbols are interleaved with the data symbols to assist the receiver in identifying the channel phase and amplitude, often using various interpolation methods. Similarly, PTAM is a technique where a known frequency tone is added periodically to data symbols. The main benefit of PSAM over PTAM is the minimal overhead required at the receiver to interpret pilot symbols. While the performance of PSAM is generally better than PTAM, neither of these methods is efficient due to the additional bandwidth requirement for carrying the reference signal [8]. Methods ii) and iii) attempt to improve system performance without the extra bandwidth.

In method ii), estimation and compensation of the phase and amplitude distortion is independent of the modulation scheme. This exclusion of the modulation, however, results in a sub-optimal technique. Two devices that best exemplify this method are the *phase locked loop* (PLL) and *automatic gain controller* (AGC). Another disadvantage of this method is that it is suitable only for channels where the time varying nature is slow enough to be considered continuous over a long time interval [39].

If the nature of the phase and amplitude distortion can be characterized mathematically, the receiver can use this information while processing received data to

perform estimation and detection. This is referred to as memory-based detection, and is presently the most widely researched receiving technique in digital communications. The full incorporation of channel effects and modulation in method iii) makes for an optimal receiving technique, and is hence the method chosen for this research. The specific communication channel considered by this thesis has been well documented previously and has been described adequately as a stochastic process. The description of the random nature of the channel is used in this thesis to improve the communication system's performance.

1.3. Objectives

There are two primary objectives of this thesis. Firstly, it is desired to design and implement memory-based receivers for a typical mobile wireless communication channel. Although several receivers already exist for this type of channel, they are prohibitively complex in design and operation. Secondly, this research aims to model the time varying attenuation and phase distortion as a discrete process. These two objectives are related. If a simple model can describe the essential characteristics of the mobile wireless channel, than this knowledge should aid in the design of a simple receiver.

1.3.1. Background and Related Work

The design of receivers for mobile wireless channels has received considerable attention of late due to the recent growth of the personal communication device industry. To achieve reliable communication, receivers must be able to overcome the multipath propagation effects. Typically receivers treat the multipath propagation effect as a process with memory if the channel coherence time is longer that the symbol interval time. Depending on the fading rate of the channel, different assumptions can be made of

L, the memory length of the process. In general, slow fading channels have a large L. A receiver takes the form of either a prediction filter or a Kalman filter, and the complexity of the receiver grows exponentially with L [25].

The best approach in determining a proper L is not clear, although several methods have been proposed. If the multipath propagation is modeled as being a Gaussian *auto-regressive moving average* (ARMA) process, then the memory length of the received signal is equal to the order of the ARMA model. There are well known techniques for constructing an ARMA model of a process based on either characteristic knowledge or training sequences. These will be discussed in Chapter 3. In [31], Kong and Shwedyk proposed a measure for the memory length of a general random process based on the Kullback-Liebler distance function. For random Gaussian processes this measure is simple to evaluate, and an Lth order approximation of the fading process can be either rejected or validated.

ARMA processes are actually Markov processes with an infinite number of states. It follows then that an alternative to the ARMA representation is to consider the multipath propagation effects to be a Markov process with a finite number of states, and hence treat the channel as a *finite state Markov channel* (FSMC). The five primary aspects for determining the validity of a FSMC model are:

i) The true channel being modeled,

ii) The specific channel characteristic that is being modeled as Markov,

iii) The criterion used to determine the validity of the Markov model,

iv) The order of the Markov model, and

v) The cardinality of and the method used to define the Markov state space.

It should be clear that these five aspects are not independent of one another.

Wang and Chang [52] based their Markov model on the time varying *signal-to-noise ratio* (SNR) of the channel. The criterion they used was the ratio of average

conditional mutual information to the average mutual information between pairs of sequential SNR samples. They concluded that when the number of Markov states is between 50 and 100, a 1st order Markov model of the Rayleigh fading environment is valid.

Chu *et al.* [10] constructed their Markov model on the time varying envelope of the received signal. The motivation to increase the Markov model order was based on the improvement seen in performance of higher order predictors of the complex Gaussian process. Both 1st and 2nd order Markov models were implemented with 16 and 64 states, with the state spaces being chosen to minimize quantization error. The ability for the models to mimic the error recursions experienced on the true channel was the measure used to determine Markov model validity. Using this generative criterion, it was determined that at least a 2nd order model was necessary to represent the true channel.

Similar to the work of Wang and Chang, Zhang and Kassam's [29] methodology was to map the SNR into a state space. The state space was created so that each state would exhibit equal average time duration. The purpose of the design was to represent the Rayleigh fading channel for simulation and performance evaluation. It was determined Markov models built using the method of equal average time duration were superior to models built using equal probability methods and *minimum mean squared error* (MMSE) quantization. The criterion used by Zhang and Kassam was generative in nature, as was the criterion used by Chu *et al.* Numerous other references exist that compare Markov models of the wireless communication channel solely on their ability to mimic the true channel characteristics [38], [50].

It is difficult to relate the FSMC models from the above-mentioned papers as they all differed in true channel, Markov model structure and in the evaluation criterion used. It is intended that this thesis should resolve some of the ambiguity involved in comparing Markov models of the mobile wireless channel.

1.3.2. Objectives of this Research

As mentioned previously, the primary objectives of this thesis are to develop a simple model of the wireless channel and to design efficient communication receivers for use on this channel. It is in this author's opinion that the best way to simplify receiver design is to simplify channel representation. The FSMC model investigated in this paper is less complex than the ARMA models used in previous receiver design. Rather than using Markov models based on functions of the received signal (SNR or envelope), this paper strives to design Markov models based directly on the nature of the received signal. The method for comparing the Markov models in this research will be to evaluate the performance of their FSMC-designed receivers on a simulated true channel. The simulated channel will consist of the transmitter, the channel and the receiver, and the criterion of comparison will be the *bit error rate* (BER) of the FSMC receiver.

1.4. Organization of Thesis

In Chapter 2 a mathematical model of the mobile wireless channel and the relation of digital communications to this channel is presented. The multipath propagation model is developed by relating the communication system to its physical environment. While there are different mathematical models of the true mobile channel that have been derived to describe the multipath propagation effect, the one that is most often utilized in research is the Jakes-Clarke model [46]. This model has been shown to represent the mobile channel accurately [27], and has therefore been selected to be the true channel in this paper. The digital communication system employed in this research, consisting of the transmitter, receiver and channel, is described in this chapter. Additionally, the concept of sufficient statistics obtained from the proper choice of a

transmitter and receiver pair is introduced. Finally, the method used to generate Jakes-Clarke channel statistics for simulation purposes is presented.

In Chapter 3 some of the currently known demodulation techniques used to handle the effects of the Jakes-Clarke channel are discussed. The criteria used to decode the received signal will be shown and validated. ARMA models of the Jakes-Clarke channel are presented as well as the receivers based on these models. This chapter shows that the complexity of these receivers is considerably large when employed with any practical modulation scheme.

Chapter 4 is the focal point of this thesis. A brief description of Markov processes is given, followed by an introduction of the FSMC and its background. The relationship between the FSMC and the true channel is discussed. This relation is then elaborated on to construct FSMC from the true channel. The actual construction techniques used to develop these FSMC are discussed.

The constructed Markov models of the channel are next implemented in the design of receivers. Specific criteria in the design process yield particular receivers. These receivers are presented and simplified. The results obtained from simulation of these receivers are compared to the performance of known receivers from the previous chapter.

A special case when considering a Jakes-Clarke channel is when there is no phase uncertainty resulting from multipath effects. This situation is discussed, and FSMC are developed based on the time varying envelope of a received signal. It will be shown that, in this case, if a proper modulation scheme is used, there is no advantage to the FSMC-based receiver implementation over the symbol-by-symbol demodulation scheme.

Chapter 5 summarizes this thesis and presents conclusions on the validity of Markov models of the mobile wireless communication channel. The basis of these conclusions is the performance of the Markov-based receivers compared to presently

known receivers. Finally, some suggestions and directions for future research are outlined.

Chapter 2

Digital Communications over Wireless Channels

As a starting point in memory-based receiver design, a mathematical model of the mobile wireless channel must be developed. The channel effects, and more specifically the channel impulse response, will be modeled as a random process. Knowledge of the random channel impulse response can be used to derive the channel's autocorrelation function. The autocorrelation function is the complete description of the mobile wireless channel for the purpose of this thesis. This chapter shows how a mobile wireless channel can be classified according to the signaling rate of a digital communication scheme. This parameter, along with the speed of the mobile, completely describes the communication systems in this thesis.

2.1. Multipath Phenomena

The nature of the physical environment is used to create a mathematical model of the mobile wireless channel. The electromagnetic signal emitted by the transmitter can be viewed as many signals spreading out from a common source. This is due to the wireless nature of the channel. Diffraction, refraction and scattering cause the propagated rays to arrive at the receiver via different paths. Each path will result in a different attenuation and delay in its respective ray. A visual example of a signal

propagating as rays through a wireless communication environment is shown in Figure 2.1.

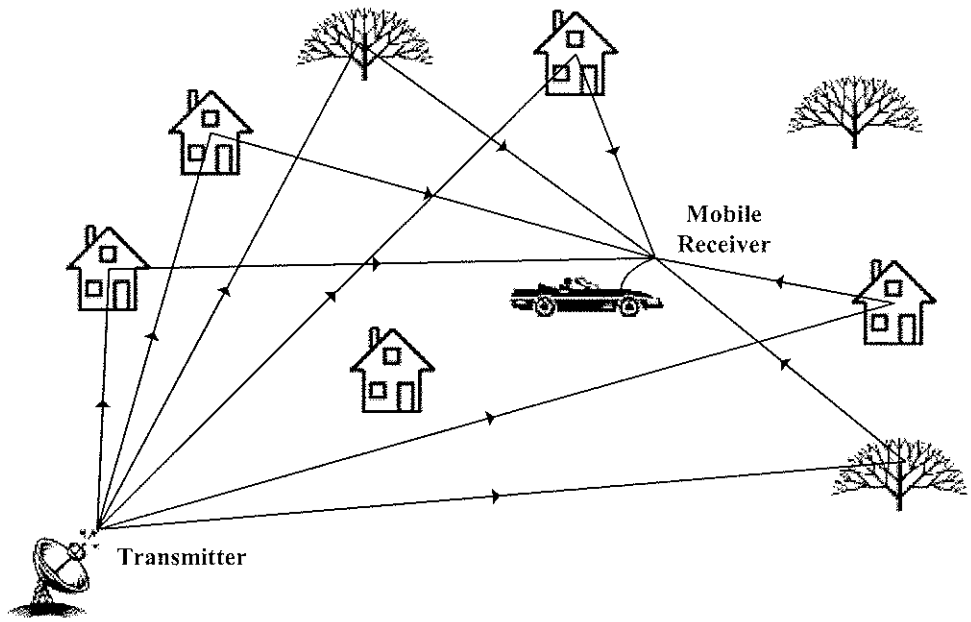

Figure 2.1: A mobile wireless communication environment

To study the behavior of the channel, let $a(t)$ be the baseband signal that is to be transmitted and $s(t)$ be the modulated or passband signal about carrier frequency f_c

$$s(t) = \mathrm{Re}\{a(t)\exp(j2\pi f_c t)\}\,. \tag{2.1}$$

The receiver is in motion and the environment is random in nature, so the attenuations and delays associated with each path will be time varying. There will also be frequency deviations in the incoming rays due to the Doppler effect. If N rays arrive at the receiver, let the attenuation, time delay and Doppler shift associated with the ith ray be $\alpha_i(t)$, $\tau_i(t)$ and $f_i^D(t)$ respectively. Assuming that all rays travel parallel to the ground, the Doppler shift associated with the ith ray will be

$$f_i^D(t) = \frac{v}{c} f_c \theta_i(t)\,. \tag{2.2}$$

$\theta_i(t)$ is the angle of arrival at the mobile receiver with respect to its direction of motion, as shown in Figure 2.2. The velocity v of the mobile is assumed to be constant, and c is the speed of light.

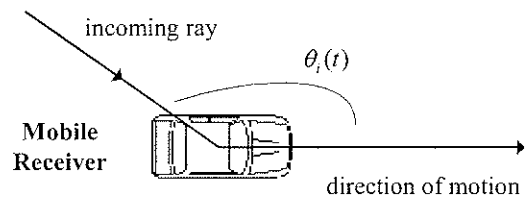

Figure 2.2: Mobile and arrival of angle of a single ray

The passband signal arriving at the receiver is the sum of all the incoming rays and can be written as

$$r_{PB}(t) = \sum_{i=1}^{N} \mathrm{Re}\{a[t-\tau_i(t)]\alpha_i(t)\exp\{j2\pi[f_c + f_i^D(t)][t-\tau_i(t)]\}\}. \tag{2.3}$$

The baseband equivalent of the received signal is

$$r(t) = r_{BB}(t) = \sum_{i=1}^{N} a[t-\tau_i(t)]\alpha_i(t)\exp\{j2\pi(f_i^D(t)[t-\tau_i(t)] - f_c\tau_i(t))\}. \tag{2.4}$$

If the transmitted signal is an unmodulated carrier then the received signal simplifies to

$$r(t) = \sum_{i=1}^{N} \alpha_i(t)[\cos\{\phi_i(t)\} + j\sin\{\phi_i(t)\}] \tag{2.5}$$

where

$$\phi_i(t) = 2\pi\{f_i^D(t)[t-\tau_i(t)] - f_c\tau_i(t)\}. \tag{2.6}$$

Alternatively, if the transmitted baseband signal is a Dirac delta function $\delta(\cdot)$, the received signal is determined to be

$$r(t) = h_c(t;\tau) = \sum_{i=1}^{N} \delta[t-\tau_i(t)]\alpha_i(t)\left[\cos\{\phi_i(t)\} + j\sin\{\phi_i(t)\}\right] \tag{2.7}$$

where $h_c(t;\tau)$ is the channel impulse response at time t when an impulse is applied at time $t-\tau$. The attenuations, delays and angles of arrival of the rays are unpredictable and hence considered random. In addition, they can be assumed to be statistically independent due to uncertainties in the environment. The propagation can then be considered *uncorrelated scattering* (US), and the central limit theorem can be applied. The central limit theorem states that the sum of N independent random variables will tend to a Gaussian distribution, provided N is large. Furthermore, the scattering can be considered a *wide sense stationary* (WSS) random process.

If the angles of arrival have a uniform distribution (as is often the case in practice) then the channel impulse response can be considered a zero mean, complex Gaussian random variable. The magnitude (or envelope) of this random variable is therefore Rayleigh distributed, and the channel is identified as a Rayleigh channel. In the case where there is a ray much stronger than the rest (referred to as a *line of sight* (LOS) path), the channel impulse response will be a non-zero mean, complex Gaussian random variable, and the envelope will be Rician distributed. For this thesis we are concerned only with the Rayleigh fading channel.

Although the density of the received envelope is important to the description of a mobile wireless channel, more information is required to describe how the channel varies with time. Relating the random channel impulse response to the digital communication system will further assist in channel classification.

2.1.1. Classification of Multipath Channels

When characterizing the mobile wireless channel, the channel impulse response is considered as a random process. A mobile wireless channel can be classified as one of

four distinct types by relating the channel impulse response to the signaling rate of the digital communication scheme. Using the procedure from [39], the autocorrelation function of the channel impulse response can be defined and analyzed as follows:

$$R_{h_c}(\tau_1,\tau_2;\Delta t) \equiv \frac{1}{2}\xi\left[h_c^*(t.\tau_1)h_c(t+\Delta t.\tau_2)\right]$$

$$= \frac{1}{2}\xi\left[h_c^*(t.\tau_1)h_c(t+\Delta t.\tau_1)\right]\delta(\tau_1-\tau_2)$$

$$= R_{h_c}(\tau_1;\Delta t)\delta(\tau_1-\tau_2)\,, \qquad (2.8)$$

where $\xi[\cdot]$ is the expectation operator and $(\cdot)^*$ denotes the complex conjugate. $R_{h_c}(\tau;\Delta t)$ represents the average power output as a function of the time delay τ, and the difference in observation time Δt. The *power intensity profile* (or power delay profile) $R_{h_c}(\tau;0)$ results when Δt is equal to zero. This special case describes how the power of an impulse applied to the channel is distributed in time when arriving at the receiver. An example power intensity profile is shown in Figure 2.3. The time interval over which most of the power arrives is called the delay spread T_D. If the delay spread is considerably smaller than the time required to transmit one digital symbol T_S, then the overall delay effect can be disregarded. If the length of the delay spread is comparable to that of the symbol interval, consecutive signals will "leak" into each other's time intervals. This phenomenon is referred to as *inter-symbol interference* (ISI).

A different perspective is obtained by looking in the frequency domain. Consider the Fourier transform partner of the power intensity profile, $S_{h_c}(\lambda) \equiv \Im\{R_{h_c}(\tau;0)\}$. The bandwidth over which $S_{h_c}(\lambda)$ is a non-zero constant is referred to as the *coherence bandwidth* B_C, and is the inverse of the delay spread. For each symbol to arrive without distortion in its power spectrum, the spectrum of the signal must fall within the

coherence bandwidth. If the signal bandwidth is larger than the coherence bandwidth, the channel is referred to as frequency selective, and ISI occurs. Otherwise the channel is deemed frequency non-selective.

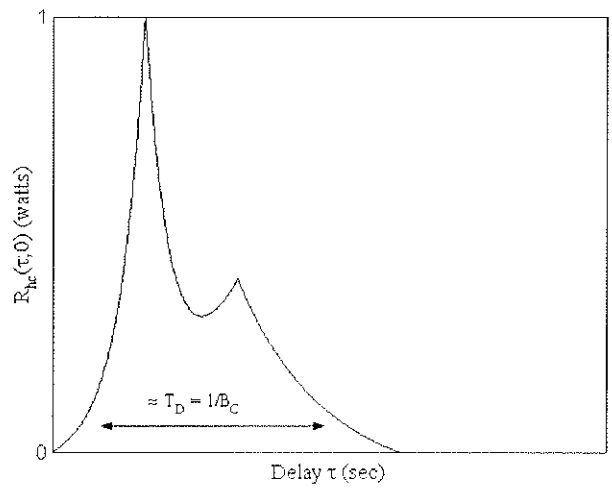

Figure 2.3: Example of a power intensity profile and its delay spread

Whether a channel is frequency selective or non-selective is a consequence of the wireless nature of the communication channel. Simply stated, if the multiple rays have comparable delays then the channel is frequency non-selective. The second means of characterization is derived from the communications channel's mobile aspect. By setting the time delay τ equal to zero in $R_{h_c}(\tau;\Delta t)$, and taking the Fourier transform, we obtain the Doppler power spectrum $S_D(f)$.

$$S_D(f) \equiv \Im\{R_{h_c}(0;\Delta t)\} \tag{2.9}$$

The Doppler power spectrum illustrates the spread of a single tone transmitted over the channel. In the ideal case the receiver is stationary and the Doppler power spectrum is a single impulse. If the receiver is in motion, the Doppler effect is observed, and the Doppler power spectrum shows an increase in bandwidth. The approximate

amount of bandwidth $S_D(f)$ occupies is the Doppler spread B_D, and is linearly proportional to the speed of the mobile. The inverse of B_D is the *coherence time* T_C of the channel and is a measure of correlation in the received signal. A high correlation will be referred to as slow fading and a low correlation will be referred to as fast fading. An example of a Doppler spectrum is shown in Figure 2.4.

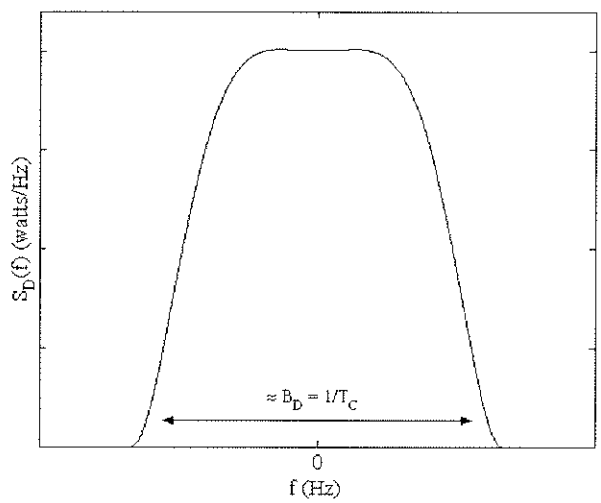

Figure 2.4: Example of a Doppler spectrum and its coherence bandwidth

To review, there are two main parameters for classification of a mobile wireless channel. The relation of the data symbol rate T_S to the delay spread of the channel T_D determines whether the channel is frequency selective or non-selective. In addition, the relation between the coherence time T_C and the data symbol rate T_S determines the relative fading rate of the channel.

Table 2.1 summarizes the possible mobile wireless channel classifications. It is worth noting that a faster mobile will exhibit a larger maximum Doppler shift, corresponding to a larger Doppler spread, and a faster fading channel.

Table 2.1: Mobile wireless channel classifications

	$T_S << T_C$	$T_S >> T_C$
$T_S << T_D$	• Frequency selective • Slow fading	• Frequency selective • Fast fading
$T_S >> T_D$	• Frequency non-selective • Slow fading	• Frequency non-selective • Fast fading

Although there are several classes of mobile wireless channels, it is important to conduct research using a channel that is meaningful with respect to the objective of this thesis. As mentioned earlier, channels that are frequency selective produce an ISI effect. Although instances of these channels occasionally occur in practice, ISI can be dealt with by simply extending the receivers designed for frequency non-selective channels [12]. For this reason, the channel investigated in this thesis will be frequency non-selective. Note that although ISI will not occur due to the channel since it is frequency non-selective, no assumptions have been made yet about ISI due to the receiving technique.

In addition to being frequency non-selective, the channel will be a slow fading channel. Although slow or fast fading is a relative term, when a channel exhibits fast fading, there is generally no correlation in the channel between successive signal intervals. If there is no correlation in the channel then there is no memory in the channel, and the optimum receiver is known and easy to implement [53]. Since correlation in the fading channel corresponds with memory and memory is the impetus for this thesis, it is only natural to consider slow fading channels. To investigate a slow fading, frequency non-selective mobile wireless channel, a specific Doppler spectra must be assumed. This thesis employs the spectra associated with the Jakes-Clarke channel. For simplicity, the term "fading channel" will be used exclusively hereafter in lieu of "slow fading, frequency non-selective mobile wireless channel".

2.1.2. Jakes-Clarke Model of the Multipath Channel

There are several assumptions that can be made when investigating the multipath model of the fading channel that help simplify (2.5) [45]. First, it is assumed the attenuation, delay, angles of arrival and Doppler shifts are individually WSS ($\alpha_i(t) = \alpha_i$, $\tau_i(t) = \tau_i$, $\theta_i(t) = \theta_i$ and $f_i^D(t) = f_i^D$). It is now obvious that the distribution of ϕ_i is uniform over $(0,2\pi]$, as $f_c\tau_i$ in (2.6) will be much greater than 1. Furthermore, it is assumed that the incoming rays are of equal strength ($\alpha_i = \alpha$). The received signal from (2.5) can now be reduced to

$$r(t) = \sum_{i=1}^{N} \alpha[\cos\{\phi_i\} + j\sin\{\phi_i\}]. \tag{2.10}$$

To proceed further, a probability density is required for the angle of arrival θ_i of the rays. If the density of θ_i is uniform over $(0,2\pi]$ then the scattering is referred to as isotropic scattering. Recall that the received signal in (2.10) is when the transmitted signal is an unmodulated carrier. $R_{h_c}(0;\Delta t)$ is now determined by finding the autocorrelation of the in-phase and quadrature components of the received signal. The autocorrelation of each component of the fading ($R_g(\tau) \equiv R_{h_c}(0;\Delta t)$) is identical and becomes

$$R_g(\tau) = (\sigma^2/2)J_o(2\pi f_D\tau) \tag{2.11}$$

where σ^2 is the total power in the fading, and $J_o(\cdot)$ is the zeroth order Bessel function of the first kind. This model is known as the Jakes-Clarke model of the fading channel. The baseband Doppler spectra $S(f) \equiv S_D(f)$ for each of the in-phase and quadrature components that follows is

$$S(f) = \Im\{R_g(\tau)\} = \begin{cases} \dfrac{\sigma^2}{4\pi f_D} \dfrac{1}{\sqrt{1-(f/f_D)^2}} & |f| \le f_D \\ 0 & |f| > f_D \end{cases} . \tag{2.12}$$

The in-phase and quadrature components of the fading are uncorrelated. The normalized Jakes-Clarke autocorrelation is shown in Figure 2.5 and the baseband Doppler spectra is shown in Figure 2.6 respectively.

It is worth mentioning that the Jakes-Clarke Doppler spectrum (2.12) evaluated at $f = \pm f_D$ results in infinity due to the Jakes-Clarke model assuming that the incoming rays are in a two-dimensional plane. In reality this will not happen, and if the model is extended to include three-dimensional space, the resulting Doppler spectra is similar but does not go to infinity.

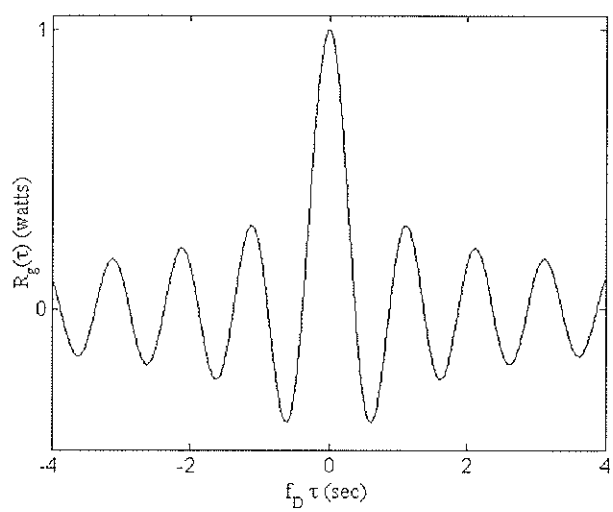

Figure 2.5: Normalized Jakes-Clarke autocorrelation

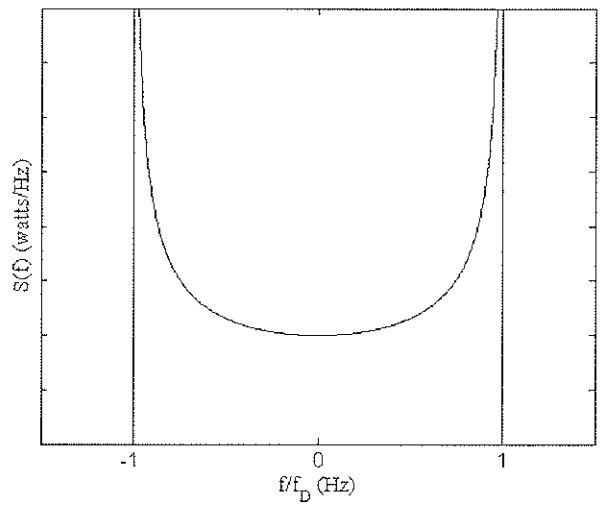

Figure 2.6: Baseband Jakes-Clarke Doppler spectrum

It has been shown in this section that for a large number of rays the channel impulse response is a complex Gaussian process, and that by stating the channel is Jakes-Clarke, the autocorrelation of the channel impulse response is also known. With the channel model established further investigation of digital communications over the mobile wireless channel can now take place.

2.2. Digital Communication over Fading Channels

The discussion of digital communication over the fading channel must first include an introduction of a few important concepts: the generation of the analog signal to be transmitted, the effect of the fading channel on the sent signal, and the model of the signal seen at the receiver. All analysis will be done from a baseband perspective.

2.2.1. Transmitter and Channel Model

Communication using a set of M possible transmission signals is known as M-ary communication. Each signal (representing $\log_2 M$ bits) is represented by a complex

value a_k, where a_k is drawn from a set with cardinality M. The entire information-bearing sequence of length N is represented by $\{a_0, a_1, \ldots, a_{N-1}\}$. When linear modulation is employed, the signal to be transmitted can be represented as the output of a filter excited by a train of impulses weighted by $\{a_0, a_1, \ldots, a_{N-1}\}$. The process of converting digital symbols to transmission signals is shown in Figure 2.7. The filter $f(t)$ is referred to as a shaping filter. With a proper choice of filter and appropriate demodulation, zero ISI can be achieved [39].

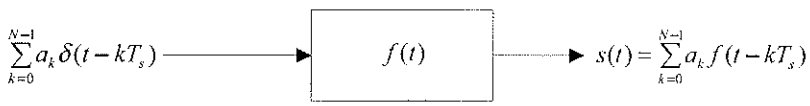

Figure 2.7: Generation of sent signal using shaping filter

Since the channel impulse response has already been described as a random process, the channel can be modeled as a time varying filter. The input to the filter is the transmitted signal $s(t)$. The receiver observes the filtered signal plus additive noise $n(t)$. The noise is white and Gaussian with variance $N_o/2$. The entire communication process from transmitter to receiver is shown in Figure 2.8.

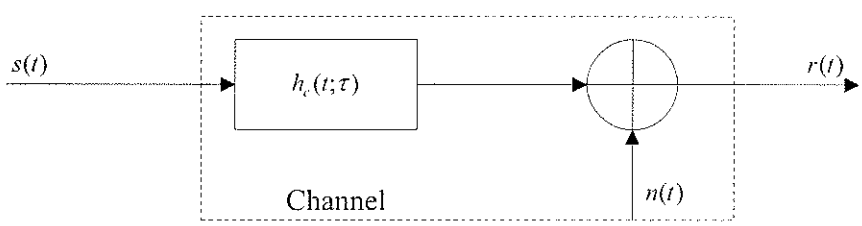

Figure 2.8: Model for mobile wireless channel

For a general mobile wireless channel, the mathematical equivalent of the above process is

$$r(t) = \int_{-\infty}^{\infty} h_c(t;\tau)s(t-\tau)d\tau + n(t)\,. \tag{2.13}$$

Because the channel investigated in this thesis is frequency non-selective, the channel impulse response $h_c(t;\tau)$ reduces to $h_c(t;0)\delta(\tau)$. Therefore, the resultant received signal is

$$\begin{aligned} r(t) &= \int_{-\infty}^{\infty} h_c(t;0)\delta(\tau)s(t-\tau)d\tau + n(t) \\ &= h_c(t;0)s(t) + n(t) \\ &= g(t)s(t) + n(t) \\ &= g(t)\sum_{k=0}^{N-1} a_k f(t-kT_S) + n(t) \end{aligned} \tag{2.14}$$

where $g(t) \equiv h_c(t;0)$. It is now apparent that the received signal produced from a frequency non-selective fading channel is the sent signal multiplied by a random time function, plus the standard additive noise. The multiplicative time function $g(t)$ is complex inducing an unknown gain and phase on the transmitted signal. This model of the channel is shown in Figure 2.9.

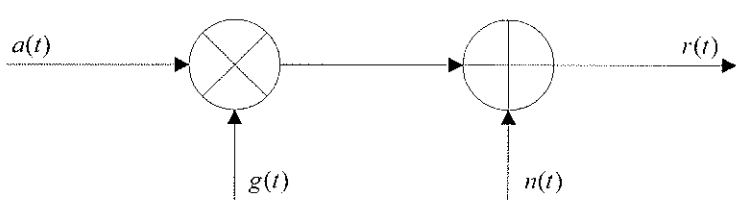

Figure 2.9: Model for the frequency non-selective channel

2.2.2. Demodulator and Decoder

There are two distinct components in a communication receiver. The first is the demodulator, which uses knowledge of the signal constellation to turn the received signal into a set of values. These values are called sufficient statistics, because they contain all

the relevant information in the received waveform. The second component of the receiver is the decoder. The role of the decoder is to transform the sufficient statistics into a symbol sequence.

Synchronization of symbol time intervals is assumed between the transmitter and receiver in this thesis. The optimum demodulator for the fading channel finds the correlation of the received signals to each possible sent signal. This correlation can be obtained by passing the received signal through a filter "matched" to the shaping filter used at the transmitter. One complex sufficient statistic for each transmitted symbol is obtained by sampling the output of the matched filter at the end of each symbol interval. This process is shown in Figure 2.10.

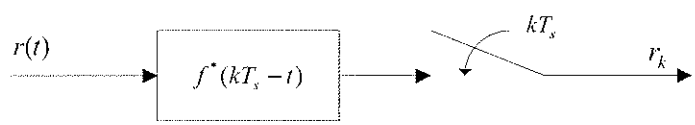

Figure 2.10: Matched filter demodulator

The kth sufficient statistic denoted as r_k can be expressed as

$$r_k = \int_{kT_s}^{(k+1)T_S} r(t) f^*(t - kT_S)dt \tag{2.15}$$

In the case of slow fading, the multiplicative gain is considered constant over each symbol interval, (i.e. $g(t) \approx g_k \quad t \in [kT_S, (k+1)T_S]$) and the received signal (2.14) can be rewritten as

$$r(t) = \sum_{k=0}^{N-1} g_k a_k f(t - kT_S) + n(t)\,. \tag{2.16}$$

When the shaping filter is designed to have unit energy, the output is

$$r_k = g_k a_k + n_k \tag{2.17}$$

where a_k represents the transmitted symbol, g_k is a multiplicative factor due to the fading over the time interval and n_k is complex Gaussian noise with variance $N_o/2$. The corresponding discrete block diagram of the channel is presented in Figure 2.11.

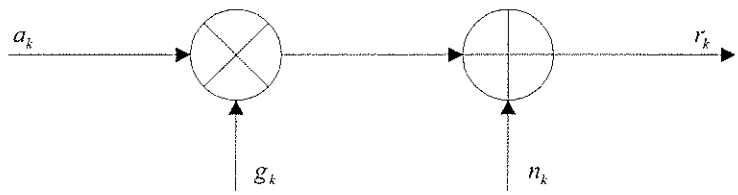

Figure 2.11: Discrete-time model of the communication system

The sufficient statistics obtained from the demodulator are processed by the decoder to estimate the original symbol sequence. The decoders utilized in this thesis are memory-based as mentioned in Chapter 1. Each criterion results in different corresponding decoding algorithms. Given the sufficient statistics there are two common decoding methodologies. One is to determine the most likely symbol sequence in its entirety, while the other is to determine the most likely symbol during each time interval. While the techniques differ greatly in complexity and concept, it is generally accepted that the overall performance of each is quite similar. *Maximum a posteriori* (MAP) sequence estimation is used in this thesis, and is described in further detail in Chapters 3 and 4.

When the shaping and matched filters are chosen correctly, the discrete model fully represents the channel. Consequently, computer simulations of the digital communication system need only to generate variables in the discrete time model and input them to a decoder. The methods used to simulate the channel and its discrete variables are described in the next section.

2.3. Channel Simulation

To test decoding algorithms for use on a fading channel, sufficient statistics need to be generated as input for the decoder. The sufficient statistics, as shown in (2.17), are a function of transmitted symbols, correlated fading samples, and white Gaussian samples. These variables have been simulated to test the decoding algorithms discussed in Chapter 3 and Chapter 4. The programs[1] used to obtain the results were written in the C programming language on a Unix terminal, or in the Matlab software environment on a Windows platform.

The modulation in this thesis is primarily *on-off keying* (OOK). OOK is a binary communication scheme that consists of simply turning a fixed amplitude signal on or off according to the value of each bit. This modulation scheme was chosen because the information is not contained exclusively in the phase or amplitude. As stated earlier, both the phase and envelope of the fading process are time varying in the Rayleigh channel environment. If the information is purely in the phase of the received signal, such as the case with *phase shift keying* (PSK) or *quadrature amplitude modulation* (QAM), receiver performance drops dramatically. This is due to the symmetry of the signal constellations combined with the phase uncertainty from the fading.

[1] All algorithms and routines used to obtain the results presented in this thesis were programmed by the author, unless stated otherwise. All of the code written for this thesis is available on compact disc and can be obtained from Dr. E. Shwedyk at the University of Manitoba.

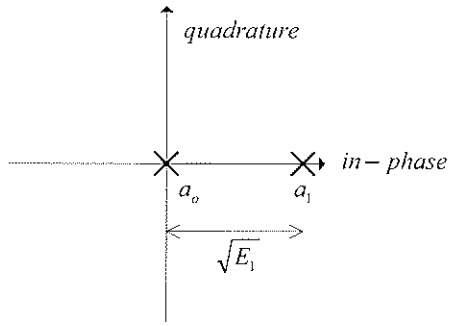

Figure 2.12: OOK signal constellation

The OOK signal constellation is shown in Figure 2.12, where E_1 is the transmitted energy when a "1" is sent. Although the "on" baseband signal is chosen to be real, the constellation is shown in two dimensions to signify that the sufficient statistics are complex. The sequence $\mathbf{a} = \{a_0, a_1, \ldots, a_{N-1}\}$ is generated randomly; each symbol is independent of each other and the other variables in the discrete-time model. Each time indexed element a_k in $\mathbf{a}$ can be either 0 or $\sqrt{E_1}$.

Since it is assumed that fading remains constant over each symbol interval T_S, the autocorrelation of the discrete fading process is a sampled version of the continuous autocorrelation:

$$R_g(kT_S) = (\sigma^2/2)J_o(2\pi k f_D T_S) \qquad k = \ldots, -1, 0, 1, \ldots. \tag{2.18}$$

When simulating the multiplicative fading, each real and imaginary component is required to be Gaussian distributed with an autocorrelation function of the above form.

It well known that passing a white Gaussian process through a filter will produce a colored Gaussian process with a power spectrum proportional to the squared magnitude of the filter transfer function. In Appendix A, the filter impulse response required to produce a Jakes-Clarke Doppler spectrum is derived

$$h(t) = \frac{\Gamma(3/4)}{\sqrt{2}} \left(\frac{f_D}{\pi |t|} \right)^{1/4} J_{1/4}(2\pi f_D |t|) \qquad t \neq 0,$$

$$\lim_{t \to 0} h(t) = \frac{1}{\sqrt{2}} \frac{f_D^{1/2} \Gamma(3/4)}{\Gamma(5/4)}. \tag{2.19}$$

It has been assumed that the filter has unity gain. $J_{1/4}(\cdot)$ is the ¼ order Bessel function of the first kind, and $\Gamma(\cdot)$ is the Gamma function.

The simulated complex fading $\mathbf{g} = \{g_0, g_1, \ldots, g_{N-1}\}$ is obtained by passing two streams of independent white Gaussian samples through identical filters. *Finite impulse response* (FIR) filters are used, which have an impulse response (2.19) sampled at a rate of T_S. The in-phase and quadrature components output by the two filters are combined as shown in Figure 2.13. It is important to note that the filters are theoretically infinite in length and non-causal. The filters used in this thesis are truncated at 2001 taps, which yields a sufficient representation of the filter. Non-causality can be disregarded since the fading is generated prior to the simulation. Several characteristics of this simulated fading, such as power spectrum and *level crossing rate* (LCR), were verified to confirm satisfactory filter implementation.

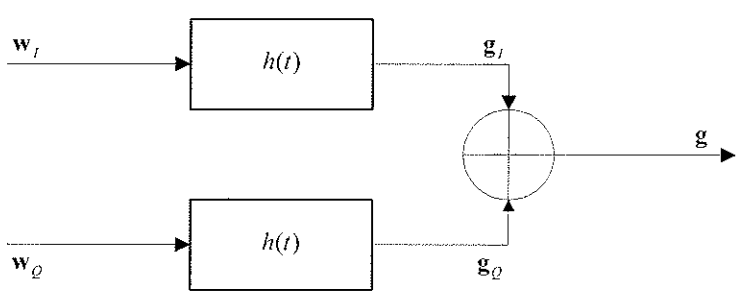

Figure 2.13: Block diagram of the fading simulator

The complex additive Gaussian noise sequence $\mathbf{n} = \{n_0, n_1, \ldots, n_{N-1}\}$ was generated using a Gaussian sample generator. The process is white with variance $N_o/2$, and the in-phase and quadrature components are independent.

The final step of this simulation is to control the SNR. The signal power component of the ratio is a combination of both the transmitted power and the power in the fading envelope. The noise power is strictly due to the additive Gaussian noise. The power contributed from the fading is expressed as an average power since the instantaneous fading power is a function of time. The power contributed by the transmitted signal is normalized in terms of average energy per bit E_b. The SNR in this thesis can now be defined as

$$\text{SNR} \equiv \frac{\text{average signal power per bit}}{\text{noise power}} = \frac{E_b \sigma^2}{N_o/2} \tag{2.20}$$

where σ^2 is the total average energy in the fading. By setting σ^2 and E_b to unity, the SNR can be controlled by simply adjusting the variance in the additive noise.

2.4. Summary

The concept of the transmitted signal arriving at the receiver via multiple paths was introduced in order to establish the *N*-ray multipath propagation model of the channel. With the assumption that no single ray is significantly stronger than another, it was shown that the channel impulse response could be viewed as a complex Gaussian random process. Further assumptions were made concerning equality of ray strength and uniform angle of arrival to develop the Jakes-Clarke model of the channel. The Jakes-Clarke model is distinguished by its channel autocorrelation function and Doppler spectra.

This chapter outlined the transmitter and receiver models used in typical communication systems. An appropriate choice of a filter pair (shaping filter at the modulator and matched filter at the demodulator) led to an equivalent discrete time model of the system. This thesis uses the discrete time model of the channel for purposes of analysis. This model employs sufficient statistics to equivalently represent the received analog signal.

The decoder, which is perhaps the most important element of the receiver, has yet to be discussed in detail. The sufficient statistics produced by the demodulator are used by the decoder to generate estimates of the transmitted symbols. These estimations are made using a predetermined algorithm that depends on the criterion and model of the channel considered by the receiver. Ideally, the receiver should take advantage of the memory in the fading process, but still manage to keep its complexity low. Although the real channel is Jakes-Clarke, the receivers used in this thesis will be based on channel models that readily adapt to decoding algorithms.

Chapter 3 will introduce a model of the channel with finite memory as opposed to the infinite memory present in the Jakes-Clarke model. Using a specific criterion, along with a finite memory model, allows for the development of two decoding algorithms later in the chapter. Chapter 4 will detail the Markov channel model, another finite memory model of the fading channel.

Chapter 3

Memory-Based Detection over Rayleigh Fading Channels

As mentioned in Chapter 1, memory based techniques will be used to map sufficient statistics into an estimated transmitted symbol sequence. Although there is no correlation between transmitted symbols or additive noise samples, the received signal has memory. This is due to the correlative nature of the fading process. A simple receiver could be implemented that makes decisions based solely on one sufficient statistic at a time. This method, however, does not take full advantage of the correlation in the received signal introduced by the channel. With memory present in the received signal, the optimal strategy is to make decisions based on the entire set of sufficient statistics or at least a larger portion of it. The next two chapters discuss methods used to make symbol decisions based on the entire received signal. Two receivers that implement an optimal decision-making strategy are presented later in this chapter.

3.1. Decoding Criterion

There are two fundamental criteria for making "optimal" decisions given the entire sequence of sufficient statistics. These are *maximum a posteriori* (MAP) symbol-by-symbol detection and maximum *a posteriori* sequence detection. These criteria will

be defined mathematically for a general communications problem. They will then be elaborated on in detail with regard to the specific channel investigated in this thesis.

Let the transmitted sequence be $\mathbf{a} = \{a_0, a_1, \ldots, a_{N-1}\}$, where N is the number of transmitted symbols. The set of sufficient statistics output by the demodulator is denoted as $\mathbf{r} = \{r_0, r_1, \ldots, r_{N-1}\}$, as described in Chapter 2. The output of the decoder, $\hat{\mathbf{a}} = \{\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{N-1}\}$, represents an estimate of the transmitted sequence. The term "estimation" generally refers to variables with continuous distributions, while the term "detection" generally refers to discrete variables. These two terms, however, will be interchanged loosely in this thesis and are to be interpreted in the context used.

The ultimate goal in a communications system is to minimize the probability of error. An error is defined as an incorrect decision. This event may be in reference to a decision at the sequence, symbol or bit level. In special cases, such as those with a memoryless channel or binary modulation, minimizing the probability of a specific type of error may also result in the minimization of the probability of another.

3.1.1. MAP Sequence Detection

Determining the MAP sequence estimate minimizes the probability of sequence error given the received set of sufficient statistics. Let the set of all possible sequences be $\mathbf{A}$, with cardinality M^N for uncoded M-ary modulation. The MAP symbol estimate of the transmitted sequence $\hat{\mathbf{a}}_{MAP}$ can be stated mathematically as

$$\hat{\mathbf{a}}_{MAP} \equiv \max_{\boldsymbol{\alpha} \in \mathbf{A}} \Pr(\mathbf{a} = \boldsymbol{\alpha} \mid \mathbf{r}) \tag{3.1}$$

where $\Pr(\cdot)$ represents a discrete probability mass. This equation can be read as "$\hat{\mathbf{a}}_{MAP}$ is the sequence $\boldsymbol{\alpha}$ in $\mathbf{A}$ that is more probable than any other, given the received vector." Using Bayes' rule, this statement can be rearranged as

$$\hat{\mathbf{a}}_{MAP} = \max_{\boldsymbol{\alpha} \in \mathbf{A}} \frac{p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}) \Pr(\mathbf{a} = \boldsymbol{\alpha})}{p_{\mathbf{r}}(\mathbf{r})} \tag{3.2}$$

where $p_x(\cdot)$ represents a continuous probability density in x. For clarity, the suffix is omitted when the argument gets large, and the variable of the density should be understood by context.

Note that the denominator is not a function of the estimated sequence and can be omitted for convenience when finding $\hat{\mathbf{a}}_{MAP}$. Furthermore, if each transmitted sequence has equal probability, then $\Pr(\mathbf{a} = \boldsymbol{\alpha})$ is equal to a constant $1/M^N$ and can also be disregarded. For this thesis the transmitted symbols are memoryless and of equal probability, therefore the estimate can be restated as

$$\hat{\mathbf{a}}_{MAP} = \max_{\boldsymbol{\alpha} \in \mathbf{A}} p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}). \tag{3.3}$$

The solution to this equation is also known as the *maximum likelihood sequence estimation* (MLSE) and is defined as

$$\hat{\mathbf{a}}_{MLSE} \equiv \max_{\boldsymbol{\alpha} \in \mathbf{A}} p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}). \tag{3.4}$$

Before proceeding further, additional knowledge is required of the communication channel, as well as the statistical dependence in the received signal. If there is no correlation in the received signal, the MLSE can be obtained symbol-by-symbol. If there is correlation, the solution is more difficult to realize. The traditional decoding methods used for a correlated fading channel are described in more detail in Section 3.2. Before specific receiver structures are developed, however, another decoding criterion will be introduced.

3.1.2. MAP Symbol-By-Symbol Detection

MAP symbol-by-symbol detection minimizes the probability of symbol error given the received set of sufficient statistics. Essentially, the estimate of each symbol can be found independently of all other symbol estimates. The set of all possible symbols is Λ, and has cardinality M for M-ary modulation. The MAP symbol estimate of the kth symbol is denoted as $\hat{a}_{k,MAP}$ and can be stated mathematically as

$$\hat{a}_{k,MAP} = \max_{\lambda \in \Lambda} \Pr(a_k = \lambda \mid \mathbf{r}) \qquad k = 0,1,\ldots,N-1 \quad . \tag{3.5}$$

This equation can be read as "$\hat{a}_{k,MAP}$ is the symbol λ in Λ that is more probable than any other a_k given the received vector." Note that the MAP symbol estimate in a binary communication system is also the MAP bit estimate.

It is well known that for a receiver that determines MAP symbol estimates or MAP sequence estimates, the receiver's complexity depends on the length of memory in the received signal. The received signal observed on the Jakes-Clarke channel actually has infinite memory, which is a result of the channel autocorrelation function (2.11) existing for all time. While the received signal technically has infinite memory, it is only natural to assume that the dependency between received samples decreases as the time distance increases. To practically implement a MAP receiver, the length of significant dependence in the received signal must be estimated. As mentioned in the Chapter 1, the estimate of the memory length is denoted as L. There are several techniques that have been suggested for determining an appropriate L for mobile wireless channels. Kong [30] chose L by determining a length that resulted in a relatively small memory loss due to the finite memory assumption. Traditionally, and perhaps more intuitively, a trial and error method is usually employed [19] [57] [35]. This research uses a trial and error method and justifies the choice of L according to receiver performance.

With a finite memory assumption, the complexity of receivers implementing the two proposed MAP criteria can be compared. A receiver with a complexity proportional to M^{2L+1} results when using the *Bahl-Cocke-Jelinek-Raviv* (BCJR) algorithm [5] to find the MAP symbol estimate. To determine the MAP sequence estimate the *Viterbi algorithm* (VA) [16] can be used, which has a complexity proportional to M^{L+1}. Past research has suggested that the resulting symbol error performance of both MAP sequence decoding and MAP symbol decoding are similar under high SNR conditions. These are the conditions of primary concern in a real environment. Since MAP sequence detection has a much lower complexity than MAP symbol-by-symbol detection, it is the criterion chosen for this thesis. The actual method used in this thesis to find the MAP estimate (and hence MLSE) will be elaborated on in the following sections and subsequent chapter.

3.2. MLSE Receivers

In this section, the MLSE stated in (3.4) will be expanded, since finite memory is now being assumed in the received signal. It will be shown that the MLSE can be found using optimal predictions of the fading process.

From the definition of MLSE, the term to be maximized in (3.4) can be rewritten as

$$p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a}=\boldsymbol{\alpha}) = p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(r_{N-1}, r_{N-2}, \ldots, r_0 \mid \alpha_{N-1}, \alpha_{N-2}, \ldots, \alpha_0)$$

$$= \prod_{k=0}^{N-1} p(r_k \mid r_{k-1}, r_{k-2}, \ldots, r_0, \alpha_{N-1}, \alpha_{N-2}, \ldots, \alpha_0)$$

$$= \prod_{k=0}^{N-1} p(r_i \mid r_{k-1}, r_{k-2}, \ldots, r_0, \alpha_k, \alpha_{k-1}, \ldots, \alpha_0)$$

$$= \prod_{k=0}^{N-1} p(r_k \mid r_{k-1}, r_{k-2}, \ldots, r_{k-L}, \alpha_k, \alpha_{k-1}, \ldots, \alpha_{k-L}). \tag{3.6}$$

The first step can be performed with successive applications of Bayes' rule, with the conditional probabilities only relying on past samples. The finite memory assumption allows the third step to be true.

Taking the natural logarithm of (3.6) results in a sum rather than a product

$$\ln\{p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha})\} = \sum_{k=0}^{N-1} \ln\{p(r_k \mid r_{k-1}, r_{k-2}, \ldots, r_{k-L}, \alpha_k, \alpha_{k-1}, \ldots, \alpha_{k-L})\}. \tag{3.7}$$

This form suggests that the Viterbi algorithm can be used to find the MLSE. Each state in the VA trellis is one of the possible M^{L+1} symbol combinations. The complexity of searching a VA trellis with M^{L+1} states agrees with the previously stated complexity of finding the MLSE.

For the Rayleigh channel considered in this thesis, the conditional probability (3.6) is Gaussian. By defining $\tilde{r}_k$ as the optimal conditional estimate of r_k and $\sigma^2_{r,k}$ as the conditional variance, (3.6) can be rewritten as

$$p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}) = \prod_{k=0}^{N-1} (2\pi\sigma^2_{r,k})^{-1/2} \exp\left(\frac{|r_k - \tilde{r}_k|^2}{2\sigma^2_{r,k}}\right). \tag{3.8}$$

It will be shown that optimal estimates of the received signal can be obtained from optimal estimates of the fading process. To obtain the required fading estimates, a model of the fading process with the corresponding finite memory must first be specified. Two techniques which provide fading estimates will be discussed in this section: Kalman filtering and prediction filtering.

3.2.1. Kalman Filtering

In general, discrete-time Kalman filters provide estimates for correlated processes observed after a linear transformation, and in the presence of additive noise [42]. The

Kalman filtering technique employed in this thesis associates the fading process with the correlated process, and the received sufficient statistics with the observed process.

Since the fading process is stationary, it can be represented as the output of a causal filter excited by a white noise process. The fading can then be modeled as an ARMA process, and predictions can be made given past values of the received signal. The general ARMA form of the fading process is

$$g_k = \sum_{i=1}^{L} c_i g_{k-i} + \sum_{j=0}^{W} d_j w_{k-j} \tag{3.9}$$

where c_i and d_j are the ARMA coefficients, $\{w_k\}$ is a zero mean white Gaussian process, and the assumed memory in the fading is L.

Using matrix notation, a vector of fading samples can be represented as

$$\mathbf{g}_k^{k-L+1} = \mathbf{C}\mathbf{g}_{k-1}^{k-L} + \mathbf{D}\mathbf{w}_k^{k-W} \tag{3.10}$$

where

$$\mathbf{C} \equiv \begin{bmatrix} c_1 & c_2 & \dots & c_{L-1} & c_L \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \tag{3.11}$$

and

$$\mathbf{D} \equiv \begin{bmatrix} d_1 & d_2 & \dots & d_{W-1} & d_W \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \tag{3.12}$$

Notationally, $\mathbf{x}_i^j$ denotes a vector $\{x_i, x_{i-1}, \dots, x_j\}^T$ where i must be greater than or equal to j. Under a specific transmitted sequence hypothesis $\boldsymbol{\alpha}$, the received sufficient statistics can be rewritten as a function of the fading vector

$$r_k = \mathbf{h}_k \mathbf{g}_{k-L+1}^{k} + n_k \,. \tag{3.13}$$

The vector $\mathbf{h}_k$ is defined as $\{\alpha_k, 0, \ldots, 0\}^T$ and corresponds with the kth symbol in a specific transmitted sequence hypothesis.

Summarizing the above, the state-space equations for the system are

$$\mathbf{g}_k^{k-L+1} = \mathbf{C}\mathbf{g}_{k-1}^{k-L} + \mathbf{D}\mathbf{w}_k^{k-W} \tag{3.14}$$

and

$$r_k = \mathbf{h}_k^T \mathbf{g}_k^{k-L+1} + n_k \,. \tag{3.15}$$

Using these state-space equations, Kalman filters can provide the maximum likelihood estimates of the fading vectors, denoted as $\tilde{\mathbf{g}}_k^{k-L+1}$, as well as their respective error covariance matrices $\mathbf{\Lambda}_k$. The maximum likelihood estimates are considered to be the optimal estimates. Borrowing the results from the method outlined in [42], the fading estimates and error covariance matrices are recursively provided by the Kalman equations

$$\tilde{\mathbf{g}}_{k+1}^{k-L} = (\mathbf{C} - \mathbf{K}_k \mathbf{h}_k^T)\tilde{\mathbf{g}}_k^{k-L+1} + \mathbf{K}_k r_k \tag{3.16}$$

$$\mathbf{\Lambda}_{k+1} = (\mathbf{C} - \mathbf{K}_k \mathbf{h}_k^T)\mathbf{\Lambda}_k \mathbf{C}^T + \mathbf{D}\mathbf{D}^T \tag{3.17}$$

$$\mathbf{K}_k \equiv \mathbf{C}\mathbf{\Lambda}_k \mathbf{h}_k (\mathbf{h}_k^T \mathbf{\Lambda}_k \mathbf{h}_k + N_0/2)^{-1} \tag{3.18}$$

Recall that $N_0/2$ is the variance of the additive channel noise.

Given the fading estimates and corresponding covariance matrices, the received signal estimates can be found given a transmitted sequence hypothesis. The conditional mean and variance of (3.8) are found to be (see Appendix B for detail)

$$\tilde{r}_k = \mathbf{h}_k^T \tilde{\mathbf{g}}_k^{k-L+1} \tag{3.19}$$

and

$$\sigma_{r,k}^2 = \mathbf{h}_k^T \mathbf{\Lambda}_k \mathbf{h}_k + N_0 / 2 . \tag{3.20}$$

The receiver structure using the fading and received signal estimates, and a Viterbi algorithm decoder is shown in Figure 3.1. Each branch of the receiver corresponds to a unique symbol vector of length L.

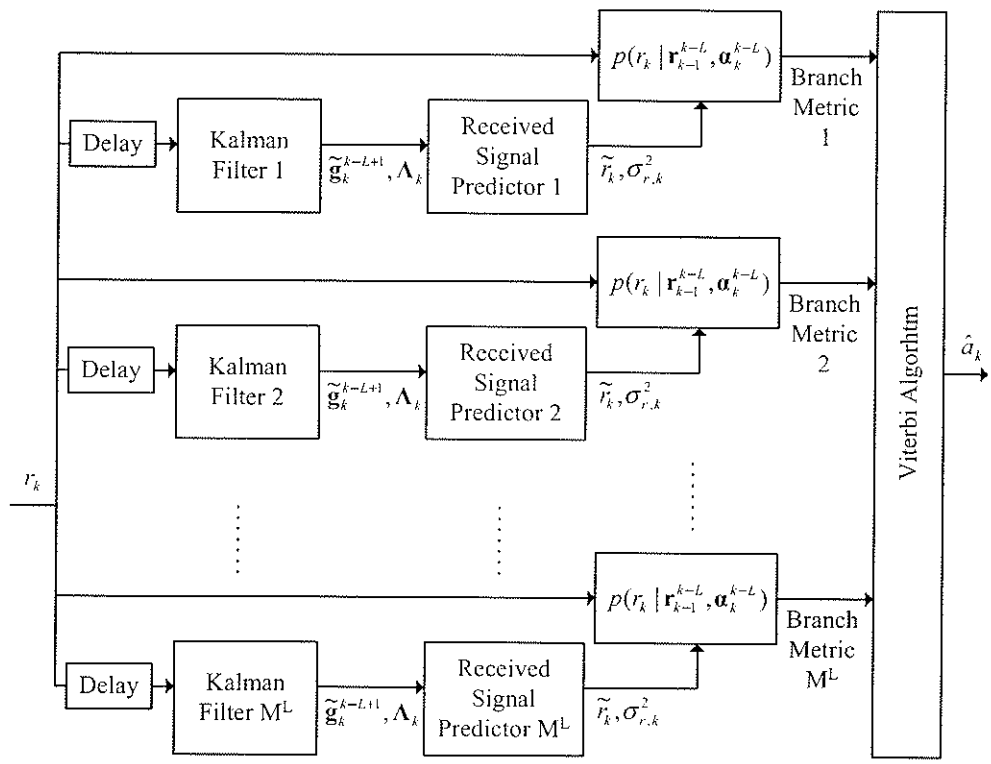

Figure 3.1: Kalman filtering combined with a Viterbi algorithm decoder

Although Kalman filters provide the optimal length-L fading vector and received signal estimates, they also prove to be complex due to the large amount of matrix multiplications and inversions. If the fading is assumed to be an *autoregressive* (AR) process (a special case of an ARMA process) the recursive matrix manipulations can be avoided. The next section discusses prediction filters, which can be used if an AR model of the fading process is assumed.

3.2.2. Prediction-filtering

A method that is similar to Kalman filtering, yet avoids the complex matrix manipulations, is prediction-filtering. By modeling the fading as an *autoregressive* (AR) process, one-step linear prediction can easily be performed [40]. When the fading is considered an AR process, it can be generated by an all-pole filter and represented by a modified version of (3.9)

$$g_k = \sum_{i=1}^{L} c_i g_{k-i} + d_0 w_k \tag{3.21}$$

Just as in (3.9), c_i and d_0 are the AR coefficients, $\{w_k\}$ is a zero mean white Gaussian process and the assumed memory in the fading is L.

Under a given transmitted sequence hypothesis $\boldsymbol{\alpha}$, as in (3.8), one can define the random variable

$$y_k \equiv \frac{r_k}{\alpha_k} = g_k + \frac{n_k}{\alpha_k}. \tag{3.22}$$

This is merely a modification of the received sufficient statistics. When g_k is considered to be an AR process, the *minimum mean squared error* (MMSE) estimate of g_k, denoted as $\tilde{g}_k$, can be calculated by a weighted summation of the modified sufficient statistics

$$\tilde{g}_k = \sum_{i=1}^{L} p_{k,i}(\boldsymbol{\alpha}) y_{k-i}. \tag{3.23}$$

The weights used in the summation are called prediction coefficients $\{p_{k,i}(\boldsymbol{\alpha})\}$ and are found by solving the set of Yule-Walker equations [40]

$$\begin{bmatrix} p_{k,1}(\boldsymbol{\alpha}) \\ p_{k,2}(\boldsymbol{\alpha}) \\ \vdots \\ p_{k,L}(\boldsymbol{\alpha}) \end{bmatrix} = \begin{bmatrix} |\alpha_{k-1}|^2 R_g(0)+\frac{N_0}{2} & |\alpha_{k-1}|^2 R_g(1) & \dots & |\alpha_{k-1}|^2 R_g(L-1) \\ |\alpha_{k-2}|^2 R_g(1) & |\alpha_{k-2}|^2 R_g(0)+\frac{N_0}{2} & \dots & |\alpha_{k-2}|^2 R_g(L-2) \\ & \vdots & \ddots & \vdots \\ |\alpha_{k-L}|^2 R_g(L-1) & |\alpha_{k-L}|^2 R_g(L-2) & \dots & |\alpha_{k-L}|^2 R_g(0)+\frac{N_0}{2} \end{bmatrix}^{-1} \begin{bmatrix} |\alpha_{k-1}|^2 R_g(1) \\ |\alpha_{k-2}|^2 R_g(2) \\ \vdots \\ |\alpha_{k-L}|^2 R_g(L) \end{bmatrix}. \tag{3.24}$$

The solution to the Yule-Walker equations requires the autocorrelation function of the discrete fading $R_g(k)$, as well as the strength of the additive noise $N_0/2$. From (3.24) it is apparent that there will be M^L sets of prediction coefficients to determine, corresponding to each possible length-L symbol vector. These can be calculated prior to processing the received signal.

As shown in Appendix C, the received signal estimate is determined from the fading estimate by

$$\tilde{r}_k = \alpha_k \tilde{g}_k = \alpha_k \sum_{m=1}^{L} p_{k,m}(\boldsymbol{\alpha}) \frac{r_{k-m}}{\alpha_{k-m}} \tag{3.25}$$

and the variance in the prediction is given by

$$\sigma_{r,k}^2 = |\alpha_k|^2 \left[R_g(0) - \sum_{m=1}^{L} p_{k,m}(\boldsymbol{\alpha}) R_g(m) \right] + \frac{N_0}{2}. \tag{3.26}$$

Division by zero is avoided in (3.25) when α_i is zero, since its respective prediction coefficient $p_{k,i}(\boldsymbol{\alpha})$ is also zero. In this case the term tends to zero, not infinity, so there is no contribution to the summation.

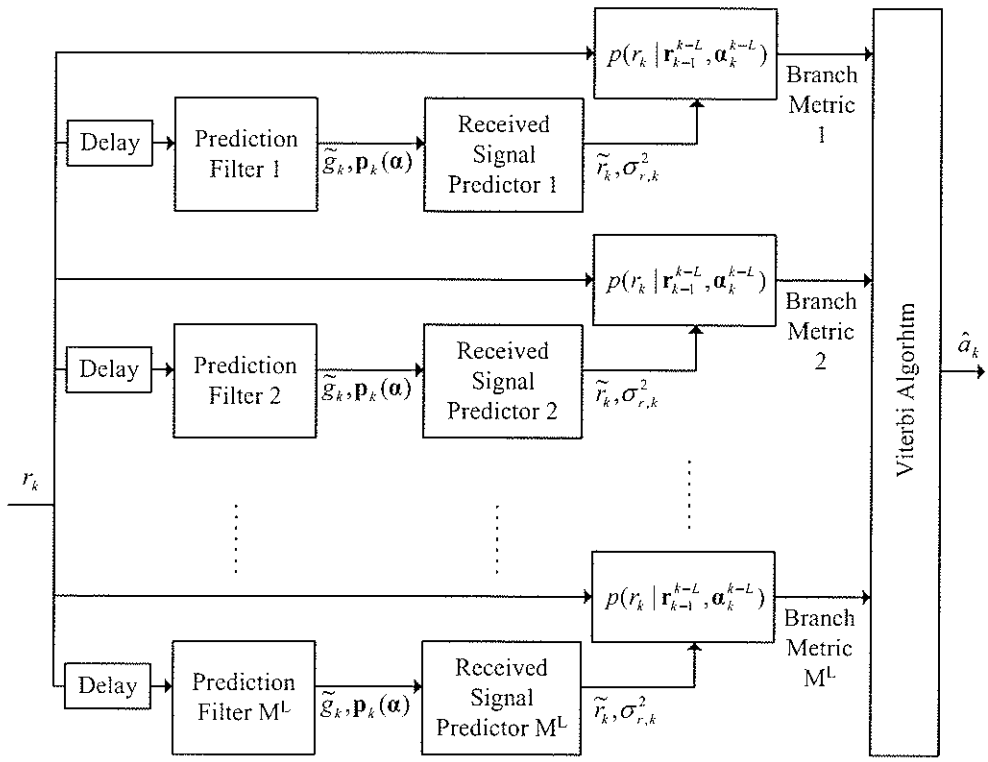

Figure 3.2: Prediction-filtering combined with Viterbi algorithm decoder

The receiver structure using the fading and received signal estimates is combined with a Viterbi algorithm decoder and shown in Figure 3.2. While the overall operation of the prediction filter decoder in Figure 3.2 and Kalman filter decoder in Figure 3.1 appear similar, the actual methods used to provide estimates differ greatly. The primary disadvantage of these receivers is the complexity of the Viterbi algorithm. The number of states used in the Viterbi algorithm is M^L, so obviously the complexity increases exponentially with the assumed memory in the channel. There are techniques that can be used to limit the number of states searched in the Viterbi trellis, such as adaptive prediction filters [35] or *per-survivor processing* (PSP) [42]. The purpose of this thesis is not to simplify known decoding algorithms, but to develop a new model of the channel to simplify the receiver. Therefore, reduced-state MLSE techniques will not be further investigated.

The structures of two well-known receivers for use on channels with memory have been discussed in this section. If the true fading is modeled as an ARMA process, Kalman filtering is a straightforward receiving technique that assists in determining the MLSE. Likewise, if the fading is considered to be an AR process, prediction-filtering will assist in finding the MLSE. As a true optimal receiver is not feasible to implement, due to the infinite memory in the Jakes-Clarke channel, the receivers discussed must be compared against each other to observe relative performance.

Given the order of the ARMA model, the Kalman filtering receiver requires the ARMA coefficients that best represent the channel. The method used to find the ARMA coefficients is discussed in the next section. The prediction filter, on the other hand, only requires an AR model order to be specified, as the autocorrelation function of the fading (2.11) is already known. The performance of receivers, using either Kalman or prediction filters, are compared for various channel memory lengths in the next section.

3.3. Results

The error performance of the previously discussed MLSE receivers was determined via simulation. The communication simulations were performed with an OOK modulation scheme on mobile channels, with fading rates $f_D T_S$ 0.01, 0.05 and 0.1. The fading rates of 0.1, 0.05 and 0.01 correspond to channels experiencing slow, medium and fast fading respectively. For a communication channel with a 2 GHz carrier transmitting 10,000 symbols per second, $f_D T_S$ of 0.01, 0.05 and 0.1 correspond to mobiles traveling at 540, 270, and 54 km/hour respectively. The goal of these simulations is to provide a reference for evaluating the Markov receivers to be introduced in the next section.

3.3.1. Kalman Filtering

When designing Kalman filters, ARMA coefficients that produce the fading process must first be determined. There is more than one method that can be used to find the coefficients, and the actual set of coefficients obtained by each method is not unique. Different methods require different information to generate coefficients, such as the autocorrelation function of the true process, the impulse response of a filter that would generate the true process, or even a sample sequence of the actual process. In addition, the number of autoregressive coefficients L and the number of moving average coefficients W must be specified.

Although there are many techniques that could have been used, the *Steiglitz-McBride iteration algorithm* (SMIA) was used in this thesis to find the ARMA coefficients. This algorithm requires the time domain impulse response of a filter that would generate the fading, given that the input is white. The SMIA minimizes the squared error between the input impulse response and the impulse response that corresponds to the generated coefficients. This algorithm was chosen because the ARMA coefficients generated are unique given $f_D T_S$, L and W, and not subject to the slight variations evident when generated with training sequences. Another reason for choosing the SMIA was because the required impulse response is already available from generating the fading process, as described in Section 2.3. Finally, the routine for the SMIA was available in the Matlab software environment this thesis used for simulation.

Using a trial and error method for determining L and W, various Kalman filtering receivers were constructed and implemented. From observations, it was concluded that there is no advantage to making W larger than L. In general, the performance of the receivers increased with L, and as expected, the length of time to complete the simulations increased exponentially. The performance of several Kalman filtering receivers is shown in Figures 3.3, 3.4 and 3.5 for channels with slow, medium

and fast fading respectively. The performance of a receiver equipped with knowledge of the fading process is also plotted. This serves as a lower bound on receiver performance for an uncoded OOK communication system.

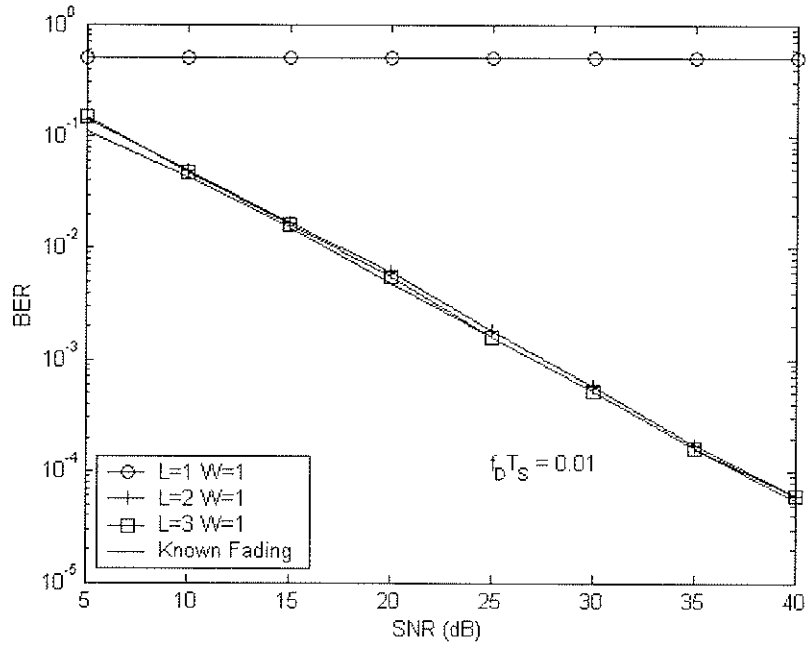

Figure 3.3: Error performance of Kalman filtering receivers $(f_D T_S = 0.01)$

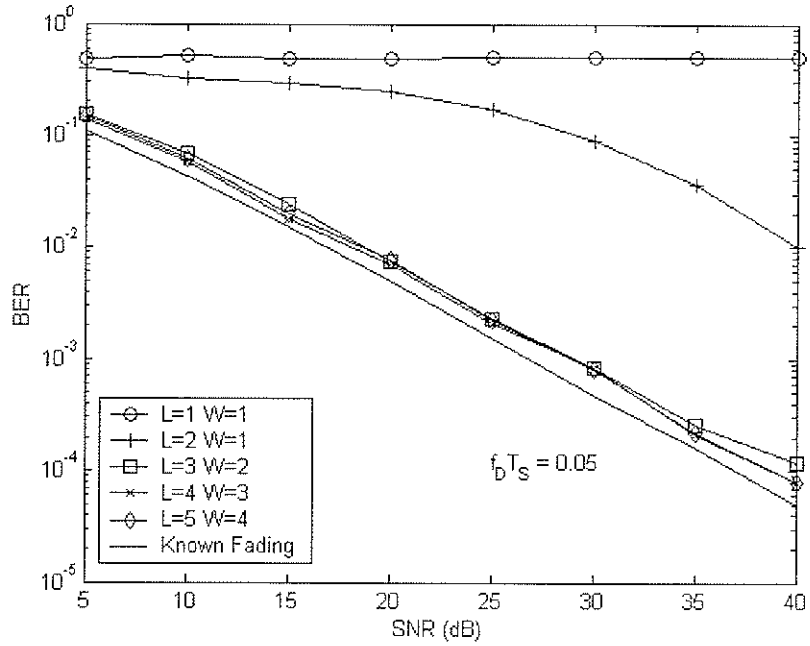

Figure 3.4: Error performance of Kalman filtering receivers $(f_D T_S = 0.05)$

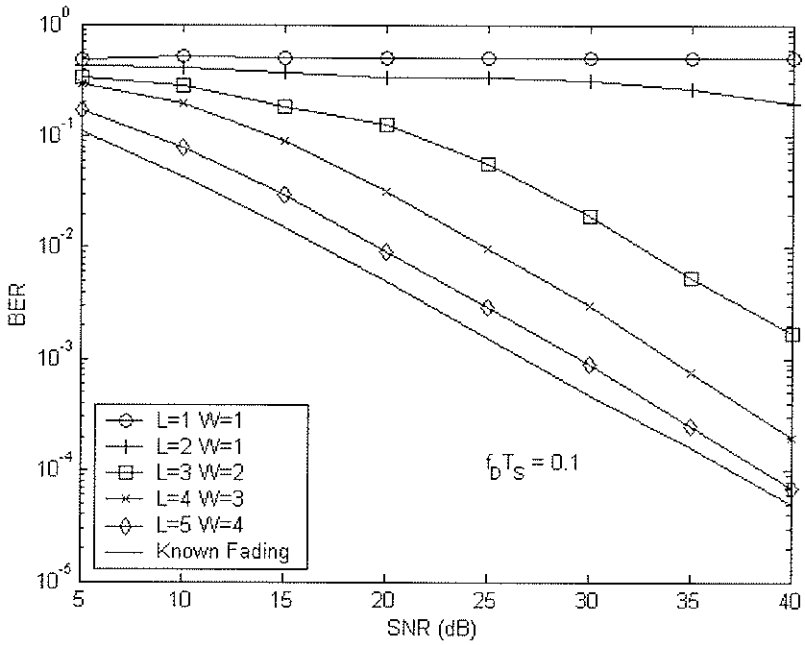

Figure 3.5: Error performance of Kalman filtering receivers $(f_D T_S = 0.1)$

As expected, the memory-based receivers performed differently on channels with different fading rates. The lower bound in all three figures is the same, however, since the correlation in the fading does not provide any new information for coherent receivers. For memory-based receivers, such as the Kalman filtering receiver, it is generally expected that better performance can be achieved when there is higher correlation in the channel. Higher correlation is observed when the fading rate is smaller.

The validity of the ARMA model representation of the fading is apparent from the above figures. For a channel with a fading rate of 0.01, Kalman filtering receivers with an L greater than 1 performed extremely well, since the lower bound appeared to have been met. On the medium fading channel, the performance of the Kalman receivers did not improve once L was greater than 3. This observation suggests that receiver performance equal to the lower bound is not always attainable. The fast fading channel requires L equal to 5 to achieve reasonable performance. Increasing L beyond 5 for the fast fading channel would be excessive, resulting in only marginal improvement in receiver performance due to the proximity to the lower bound.

3.3.2. Prediction-filtering

The only parameter to consider when performing prediction-filtering is the assumed memory length in the received signal. The influence of the assumed memory length on receiver performance was investigated through computer simulation. The results obtained are shown in Figures 3.6, 3.7 and 3.8 for channels with fading rates of 0.01, 0.05 and 0.1 respectively.

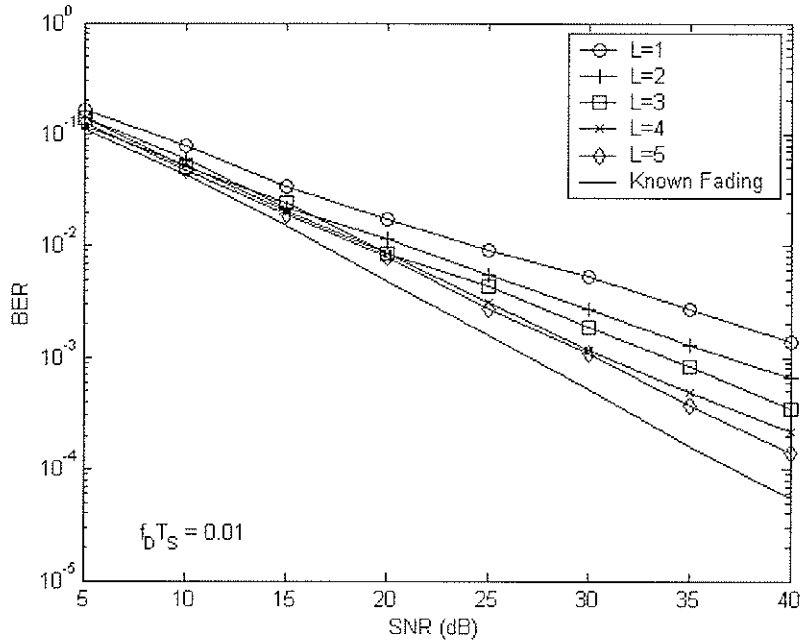

Figure 3.6: Error performance of prediction-filtering receivers $(f_D T_S = 0.01)$

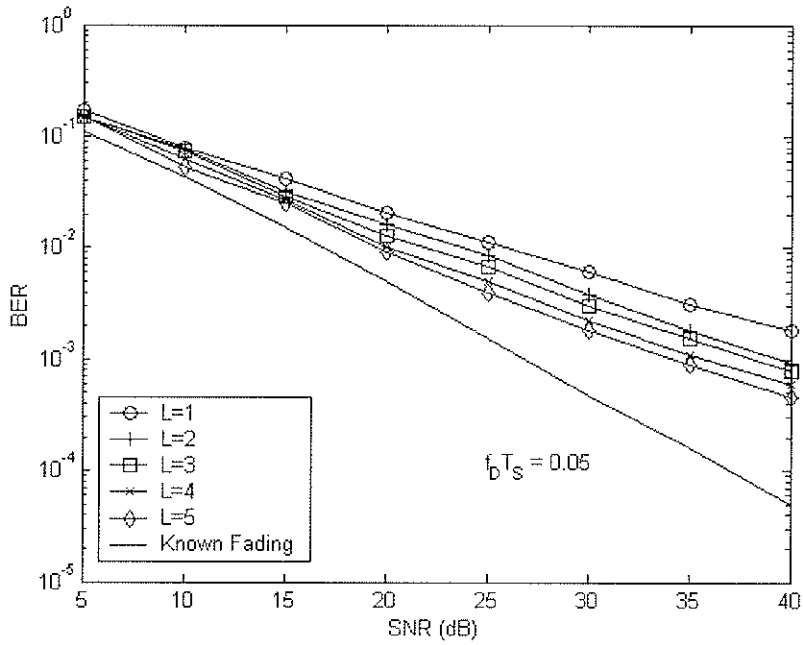

Figure 3.7: Error performance of prediction-filtering receivers $(f_D T_S = 0.05)$

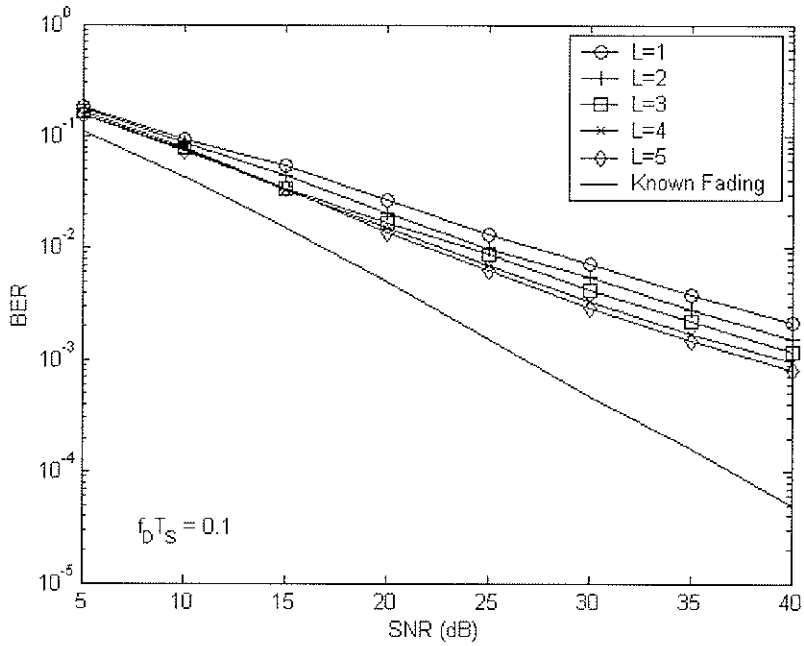

Figure 3.8: Error performance of prediction-filtering receivers $(f_D T_S = 0.1)$

For the slow fading channel, the performance of a prediction-filtering receiver with L equal to 5 is close to the lower bound given by the coherent receiver, but not as close as the Kalman filtering receiver with L equal to 5. Compared to the Kalman filtering receivers, the prediction-filtering receivers show poor overall performance on both medium and fast fading channels.

It is apparent from the above discussion that the Kalman filtering receivers generally perform better than the prediction-filtering receivers. There are two possible reasons for this observation. The first is that the Kalman filters provide the optimum length-L sequence estimates of the fading, while the prediction filters provide the optimum scalar estimate of the fading. The recursive nature of the Kalman filtering combined with the sequence estimates may provide better received signal estimates. The second reason is that Kalman filters use more knowledge of the fading when constructing predictions. While the Kalman filtering approach requires the entire impulse response of

a process-generating filter to construct the ARMA coefficients, the prediction-filtering receiver only uses L samples of the autocorrelation of the fading.

An observation can be made about the performance of the Kalman filtering and prediction-filtering receivers when increasing L. The increases in performance appear to be small but consistently incremental for the prediction-filtering receivers. The increases in performance for the Kalman filtering receivers, on the other hand, tend to vary considerably in magnitude and are not at all consistent.

3.4. Summary

Several criteria for memory-based receivers were introduced in this section, with MLSE being the criterion chosen for this thesis. The methods that find the MLSE are based on either an ARMA or an AR model of the fading process. The Kalman filtering method, which employs an ARMA model, obtains received signal estimates from fading vector estimates, using complex matrix manipulation. The prediction-filtering method, which employs an AR model, appears to be a more intuitive approach for finding the fading estimates and received signal estimates.

The main drawback to implementing either of these algorithms lies in the complexity of the Viterbi algorithm that processes the received signal estimates. The computational complexity of implementing the VA grows exponentially with the assumed memory length of the ARMA or AR model of the fading. Reference receivers should yield good performance, yet exhibit low complexity. Different reference receivers should be used for different fading rates, as the required memory used by the receiver depends on the correlation of the channel. The performance of the Kalman filtering and prediction-filtering receivers were compared for various L to determine which method to use. Kalman filtering was chosen for all three channels considered.

The reference receivers that will be used are shown in Table 3.1, along with their respective parameters. The performance of these receivers, as well as the lower bound provided by coherent receivers, will serve as the references for the next chapter.

Table 3.1: Reference receivers for channels with different fading rates

Channel	**Reference receiver**
Slow ($f_D T_S = 0.01$)	Kalman filtering - $L = 2$, $W = 1$
Medium ($f_D T_S = 0.05$)	Kalman filtering - $L = 3$, $W = 2$
Fast ($f_D T_S = 0.1$)	Kalman filtering - $L = 5$, $W = 4$

Chapter 4

Finite State Markov Channel

Chapter 2 discussed the physical environment of a mobile wireless communication system, and derived a mathematical model of the channel. The Jakes-Clarke channel was chosen as the true channel, but due to the infinite channel memory it was shown that optimum receivers could not be practically implemented. Chapter 3 introduced the ARMA and AR models of the fading channel. These models of the true channel enabled MLSE memory-based receivers to be constructed. An assumption of finite memory in the fading was made, and the complexity of these receivers grew exponentially with the assumed memory length. This chapter introduces an alternative finite memory model of the Jakes-Clarke channel to assist in MLSE memory-based receiver design. This model is the *finite state Markov channel* (FSMC) model.

Although much research has been conducted on modeling communication channels as FSMCs, little attention has been focused on the application of this knowledge in receiver design for mobile wireless channels. This chapter briefly summarizes some basic elements of FSMCs. It then proceeds to relate the Markov channel to the channel under investigation in this thesis. The receiver structures will be derived from the MLSE criterion and the FSMC assumption, and the performance of the MLSE FSMC receivers will be analyzed.

4.1. Introduction

Markov models are often used to describe stochastic processes. The most common application of Markov modeling is found in pattern recognition problems [41]. Markov models serve as simple comprehensive representations of discrete time processes. A finite state Markov model approximates a true process using a finite number of states, where each state represents a possible existence point. Typically each state represents different characteristics and has a different set of parameters from the other states. These parameters will be discussed in the next section. The Markov model can be observed in terms of a sequence of occupied states, or as the indirect output from a sequence of occupied states. If the state sequence is not directly observable, but an embedded state transition output is, the Markov model is commonly referred to as a hidden Markov model.

The primary reason Markov modeling is often employed is that there are many methods available for constructing Markov models from test data, without any prior knowledge of the source. This is not the reason, however, that Markov models are used in this thesis, since the Jakes-Clarke model already describes accurately the fading channel. The motivation for using Markov models is the potential for representing the true channel in a simpler manner than the ARMA or AR models do. If a Markov model can describe the correlation in a fading channel accurately, this knowledge can be used to design efficient receivers. To fully appreciate the relation between the Markov model and the Jakes-Clarke fading channel, a brief introduction to Markov processes and their respective parameters must first be presented.

A 1st order finite state Markov process with K states is completely described by three elements (a Dth order Markov process will be discussed in the next section in the context of the wireless channel):

i) state-space, $\mathbf{C} \equiv \{c_0, c_1, \ldots, c_{K-1}\}$,

ii) $K \times K$ transition probability matrix $\mathbf{P}$, and

iii) length-K steady state probability vector $\boldsymbol{\pi}$.

The Markov process represents a discrete time process. For the duration of any one time interval, the process can be in one and only one state. The set of probabilities for the process to transfer from any particular state to another is contained in the probability transition matrix $\mathbf{P}$. Each element $P_{i,j}$ of the matrix $\mathbf{P}$ is defined as

$$P_{ij} = P(c_j \text{ will be the next state } | \text{past state was } c_i) \tag{4.1}$$

where i and j denote specific state indices. An important aspect of Markov processes is that the transition probabilities depend only on the currently occupied state, and not on previous outputs or occupied states.

The elements of the steady-state probability vector $\boldsymbol{\pi}$ are the probabilities of being in any particular state, without any knowledge of outputs or transitions:

$$\pi_i \equiv \Pr(c_i). \tag{4.2}$$

A visual representation of a 3-state Markov process is shown in Figure 4.1. The state transitions are represented as arrows connecting states, and are labeled with arbitrarily assigned transition probabilities. The steady-state probabilities are primarily used for initializing a Markov process, and are omitted here for clarity.

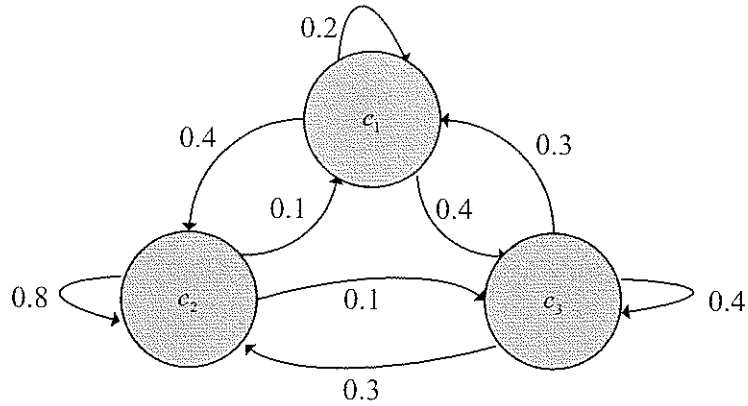

Figure 4.1: General example of a Markov process

4.2. Finite State Markov Channel (FSMC)

Markov processes were first used to describe a communication channel by Gilbert [20], and were later elaborated on by Elliott [15] in the 1960s. Their study of discrete channel models with memory led to a 2-state Markov channel. In the Gilbert-Elliot model, each state's defining characteristic is its probability of transmission error. One state is a "good state" G with a low probability of transmission error, and the other state is a "bad state" B with a high probability of transmission error. The state transitions reflect the pattern of burst errors observed on a communication channel. This model is shown in Figure 4.2.

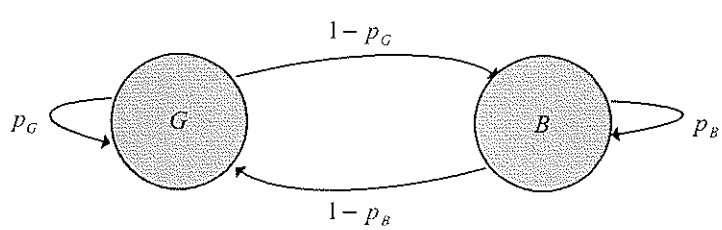

Figure 4.2: Gilbert-Elliott 2-state Markov channel

The Gilbert-Elliott channel model can be extended by increasing the number of states or varying the error characteristics associated with each state [28]. If the channel can be represented by a FSMC with a different probability of transmission error attributed to each state, decoding techniques are known that take advantage of the channel memory [51]. Recent research has taken a different approach by letting the Markov channel states represent a more fundamental channel characteristic, such as the fading envelope [50], SNR [2], or soft decision outputs [48]. The channel characteristic to be modeled as a Markov process in this thesis is the complex fading process $\{g_k\}$. Once a Markov model of the complex fading is obtained, it can be easily transformed into a Markov model of the fading envelope or SNR, as these are both functions of the complex fading. The development of the complex fading Markov model from the Jakes-Clarke channel is discussed in this section.

4.2.1. FSMC as a Model of the Rayleigh Fading Channel

Recall that the fading process g_k experienced on the mobile wireless channel is a correlated zero-mean complex Gaussian process. The in-phase and quadrature components are independent, and have identical statistical properties. Let either the in-phase or quadrature component of g be denoted as g_c without loss of generality:

$$g_c \equiv \mathrm{Re}\{g\} \text{ or } \mathrm{Im}\{g\} \tag{4.3}$$

If the context requires a time index k, the fading component will be denoted as $g_{c,k}$. Although it is possible to construct a FSMC based on the entire complex Gaussian process, this thesis will construct identical but independent Markov models for both the in-phase and quadrature components.

When extending the basic 1st order Markov process to a Dth order Markov process, the state-space $\mathbf{C}$ is redefined as $\{\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{K^D-1}\}$. The number of states is kept

as a function of K so that models of various order can later be compared by a normalized number of states. The corresponding transition probability matrix of a Dth order Markov process will be $K^D \times K^D$. Each state $\mathbf{c}_i$ will be a length-D vector of sequential g_c values

$$\mathbf{c}_i = \{c_{i,0}, c_{i,1}, \ldots, c_{i,D-1}\} \tag{4.4}$$

The specific elements of $\mathbf{c}_i$ do not change once the Markov model is constructed. The values are typically derived from statistical properties of the fading process. In this thesis, the Markov states are determined by quantization of higher order probability densities associated with g_c. Before proceeding further, these probability distributions must be stated.

Since the random variable g_c is Gaussian with an autocorrelation function given by (2.18), the length-D random vector $\mathbf{g}_{c,k}^{k-D+1} \equiv \{g_{c,k}, g_{c,k-1}, \ldots, g_{c,k-D+1}\}$ is D dimensional Gaussian. The corresponding probability distribution is

$$f_{\mathbf{g}_c}(\mathbf{g}_c) \equiv f_{\mathbf{g}_{c,k}^{k-D+1}}(\mathbf{g}_{c,k}^{k-D+1}) = (2\pi)^{-D/2} \, |\Sigma|^{-1/2} \exp\left\{-\mathbf{g}_c \Sigma^{-1} \mathbf{g}_c^{\,T} / 2\right\} \tag{4.5}$$

where the $D \times D$ covariance matrix Σ has elements $\Sigma_{i,j}$

$$\Sigma_{i,j} = [R_g((j-i)T_s)]_{D \times D}. \tag{4.6}$$

Recall that the fading process is stationary, therefore the density is independent of the time index. This enables $\mathbf{g}_{c,k}^{k-D+1}$ to be replaced by $\mathbf{g}_c$ for clarity, in the context of a Dth order model. The covariance matrix is symmetrical because the Jakes-Clarke autocorrelation function $R_g(m)$ is an even function in m.

The state-space of the Markov model of each fading component can be constructed by quantizing the D dimensional Gaussian density. Furthermore, the transition probabilities can be determined by finding the areas under regions of a $D+1$

dimensional Gaussian density. The theory to construct a FSMC from D and $D+1$ dimensional densities and the methods actually used are discussed next.

4.2.2. Markov Model Construction

The construction of a Dth order Markov model from the Jakes-Clarke model is a two step procedure. First the state-space is constructed from knowledge of the D dimensional probability distribution of $\mathbf{g}_c$. Recall that the length-D vector $\mathbf{g}_c$ can be either the real or imaginary component of the complex fading, without loss of generality. The state-space consists of K^D non-overlapping regions, and is determined by a quantization of the density $f_{\mathbf{g}_c}(\mathbf{g}_c)$. Let the region associated with state $\mathbf{c}_i$ from the quantization of $f_{\mathbf{g}_c}(\mathbf{g}_c)$ be R_i. The criteria and methods used to find R_i will be discussed later in this section. The steady state probabilities can be determined directly from this quantization

$$\pi_i = \Pr(\mathbf{c}_i) = \Pr(\mathbf{g}_c \in R_i) = \int_{R_i} f_{\mathbf{g}_c}(\mathbf{g}_c) d\mathbf{g}_c \ . \tag{4.7}$$

The second step is to determine the state transition probabilities:

$$P_{ij} = P(\mathbf{c}_j \text{ will be the next state} \mid \text{past state was } \mathbf{c}_i) = P(\mathbf{g}_{c,k} \in R_j \mid \mathbf{g}_{c,k-1} \in R_i) \ . \tag{4.8}$$

This can be rewritten using Bayes' rule

$$P_{ij} = \frac{P(\mathbf{g}_{c,k} \in R_j, \mathbf{g}_{c,k-1} \in R_i)}{P(\mathbf{g}_{c,k-1} \in R_i)} \ . \tag{4.9}$$

The denominator is identified as π_i, while the numerator can be manipulated with the realization that $\mathbf{g}_{c,k}$ and $\mathbf{g}_{c,k-1}$ share $D-1$ elements. The numerator is then the area under a $D+1$ dimensional Gaussian density

$$P(\mathbf{g}_{c,k} \in R_j, \mathbf{g}_{c,k-1} \in R_i) = \int_{Z_{i,j}} f_{\mathbf{g}'_{c,k}}(\mathbf{g}'_{c,k}) d\mathbf{g}'_{c,k} \tag{4.10}$$

where

$$\mathbf{g}'_{c,k} \equiv \mathbf{g}_{c,k}^{k-D} = \{g_{c,k}, g_{c,k-1}, \ldots, g_{c,k-D}\}, \tag{4.11}$$

and

$$f_{\mathbf{g}'_c}(\mathbf{g}'_c) = (2\pi)^{-(D+1)/2} \, |\mathbf{\Sigma}'|^{-1/2} \exp\left\{-\mathbf{g}'_c \mathbf{\Sigma}'^{-1} \mathbf{g}'^{T}_c / 2\right\}. \tag{4.12}$$

The entries $\Sigma'_{i,j}$ of the covariance matrix $\mathbf{\Sigma}'$ are

$$\Sigma'_{i,j} = [R_g((j-i)T_S)]_{D+1 \times D+1}. \tag{4.13}$$

The regions $Z_{i,j}$ are not readily determined, so in practice the transition probabilities must be calculated using training sequences with appropriate statistical properties.

Although the above analysis provides a theoretical understanding of the Markov model construction, it is not feasible to find actual Markov parameters this way. For the simple case when D is equal to 1, areas under regions of a 2-dimensional Gaussian density determine the transition probabilities. There are no closed form solutions for the areas however, even if the regions $Z_{i,j}$ are well defined. For this thesis, the state-space of a Markov model and its transition probabilities are determined with the assistance of the *Linde-Buzo-Gray* (LBG) algorithm, and Monte-Carlo techniques. Given training sequences of the fading process, the LBG algorithm determines the best quantization for the state-space, and the Monte-Carlo techniques determine approximate transition probabilities.

The LBG (or generalized Lloyd) algorithm is used to partition the distribution of fading vectors into a state-space. The LBG algorithm performs *vector quantization* (VQ) by repeatedly moving all reference vectors to the arithmetic means of their Voronoi sets [36]. The reference vectors are obtained from a source that exhibits appropriate statistical properties, and there are typically many more training vectors than states.

Early in this research, several state-spaces were constructed using the LBG algorithm and 1,000,000 training vectors. This number proved sufficient as the resultant state-spaces consistently approached the same values, and the arrangement of the state-spaces agreed with intuitive expectations. To assure that the number of training vectors used in this thesis was more than sufficient, 10,000,000 were used to construct each state-space.

The state-space of a Dth order Markov model is determined by quantizing the D dimensional probability distribution into K^D regions or states. The state-space must first be initialized with each of the K^D states being given a unique centroid. Each length-D training vector is then associated with the nearest centroid. A new centroid is then calculated for each state from its set of associated training vectors. The training vectors are once again assigned to the nearest centroid. This process is repeated until the state-space ceases to change significantly, or up to a predetermined number of iterations. The final values of the centroids are used as the state-space $\mathbf{C}$ of the Markov model. The step-by-step LBG algorithm is shown in Table 4.1.

Table 4.1: LBG algorithm used to determine a Markov state-space

Step 1	Initialize K^D unique centroids.
Step 2	Assign each length-D training vector to the nearest centroid.
Step 3	Calculate K^D new centroids from the training vectors associated with each state.
Step 4	If the state-space has settled or the maximum number of iterations has occurred – go to Step 5. Else – go to Step 2.
Step 5	Return current centroids as the Markov state-space $\mathbf{C}$.

Sample 1st and 2nd order state spaces are shown in Figure 4.3 and Figure 4.4 respectively. While both the targets and regions of the quantization are shown in Figure 4.3, only the targets are shown in Figure 4.4. It is understood that the quantization regions are determined by the nearest centroid.

Once the Markov state-space is obtained, the transition probabilities can be determined using the Monte-Carlo method, which approximates the probabilities using the statistics of training data. The first step is to assign each training vector to the state with the smallest Euclidian distance from it. Next, the sequence of training vectors (or equivalently the sequence of occupied states) is examined for state transitions and state occurrences. The ratio of specific state transitions to state occurrences determines the state transition probabilities.

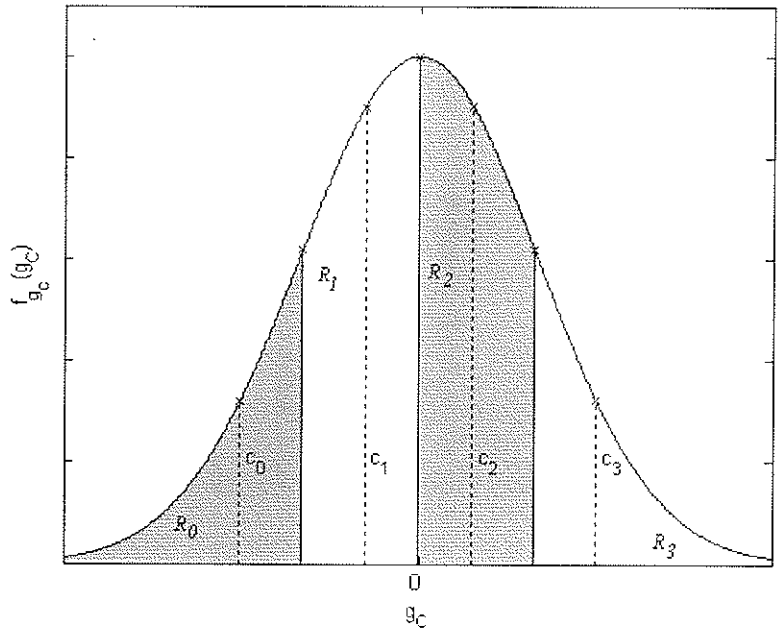

Figure 4.3: Quantization of 1st order density for 4-state 1st order state-space

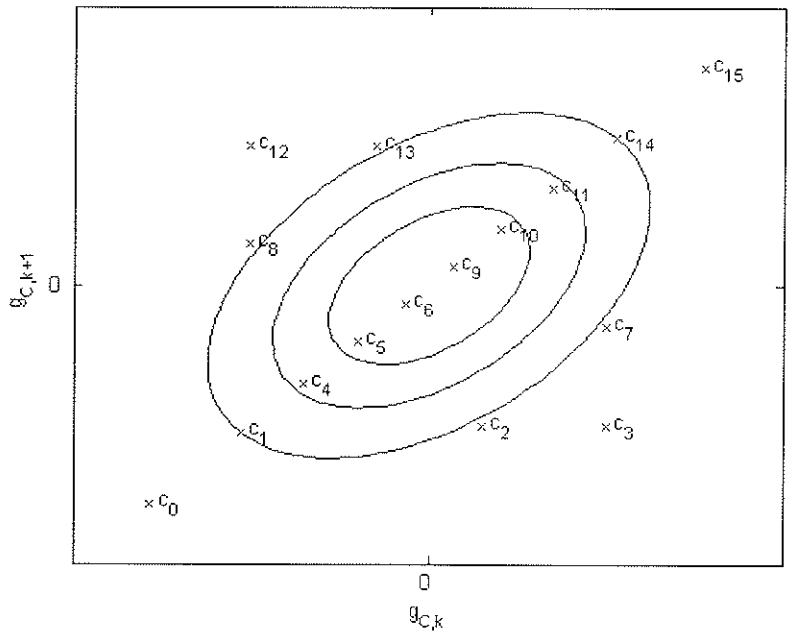

Figure 4.4: VQ of 2nd order density for 16-state 2nd order state-space

Although it is known that the fading process is complex Gaussian, the LBG algorithm and Monte-Carlo method do not require this information. These algorithms are attractive to use for any general time-varying channel, since only training data is required to construct a Markov model.

The FSMC as well as techniques for Markov model construction were presented in this section. The design of decoding algorithms for use on a FSMC is explained in the following section.

4.3. MAP Sequence Detection and Channel State Estimation

In Chapter 3, the MAP sequence detection criterion was stated, and the solution to the problem was obtained when the fading process was considered to be an ARMA or AR process. When considering a FSMC, finding the MAP sequence estimate proves to be more difficult. A different, but closely related criterion to the MAP sequence estimate, is joint MAP sequence detection and channel state estimation. Less computation is required for joint MAP decoding compared to MAP sequence decoding, and results will show that the joint MAP receivers compare well to the reference receivers. FSMC receivers using joint MAP decoding are simulated, and the resulting performance is compared to the performance of the reference receivers discussed in the previous chapter. This section will also discuss MAP channel state estimation, which is another possible criterion for FSMC receiver design.

As stated in the previous section, the in-phase and quadrature components of the fading have identical Markov representations. When deriving the estimation equations, however, the in-phase and quadrature Markov models must be distinct. In general, the Markov state-space of either component is denoted as $\mathbf{C}=\{\mathbf{c}_0,\mathbf{c}_1,\ldots,\mathbf{c}_{K^D-1}\}$ where $\mathbf{c}_i=\{c_{i,1},c_{i,1},\ldots,c_{i,D-1}\}$. Let a vector or the in-phase fading be denoted as

$$\mathbf{g}_{I,k}^{k-D+1} \equiv \mathrm{Re}\{\mathbf{g}_k^{k-D+1}\} = \mathrm{Re}\{g_k, g_{k-1}, \ldots, g_{k-D+1}\} \tag{4.14}$$

and similarly, a vector of the quadrature component of the fading

$$\mathbf{g}_{Q,k}^{k-D+1} \equiv \mathrm{Im}\{\mathbf{g}_k^{k-D+1}\} = \mathrm{Im}\{g_k, g_{k-1}, \ldots, g_{k-D+1}\}. \tag{4.15}$$

It is worthwhile to mention that since both components of the fading are represented by identical Markov models, $p\left(\mathbf{g}_{I,k}^{k-D+1} = \mathbf{c}_i\right)$ and $p\left(\mathbf{g}_{Q,k}^{k-D+1} = \mathbf{c}_i\right)$ are both valid probabilities.

4.3.1. MAP Sequence Detection

Let the transmitted sequence be denoted as $\mathbf{a} = \{a_0, a_1, \ldots, a_{N-1}\}$, the set of received sufficient statistics as $\mathbf{r} = \{r_0, r_1, \ldots, r_{N-1}\}$ and the symbol sequence estimate as $\hat{\mathbf{a}} = \{\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{N-1}\}$, where N is the total number of transmitted symbols. As mentioned in Chapter 3, the MAP symbol sequence estimate is also the MLSE for the communication system used in this thesis. The suffix on the continuous probability distribution will occasionally be omitted for clarity if it is of considerable length. The variable of the distribution should still be easily understood in its context. Restating (3.4), the MLSE is defined as

$$\hat{\mathbf{a}}_{MLSE} \equiv \max_{\boldsymbol{\alpha} \in \mathbf{A}} p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}) \tag{4.16}$$

The term to be maximized above can be rewritten as

$$\begin{aligned} p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}) &= p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(r_{N-1}, r_{N-2}, \ldots, r_0 \mid \alpha_{N-1}, \alpha_{N-2}, \ldots, \alpha_0) \\ &= \prod_{k=0}^{N-1} p(r_k \mid r_{k-1}, r_{k-2}, \ldots, r_0, \alpha_{N-1}, \alpha_{N-2}, \ldots, \alpha_0). \end{aligned} \tag{4.17}$$

The fading, transmitted signal, and additive noise are considered to be independent. Under the assumption that the channel is a FSMC, the probability inside the product can

be expanded as a double summation over both the in-phase and quadrature Markov state-spaces

$$p_{\mathbf{r}|\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r}\,|\,\mathbf{a}=\boldsymbol{\alpha})=\prod_{k=0}^{N-1}\sum_{i=0}^{K^D-1}\sum_{j=0}^{K^D-1}p\left(r_k,\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i,\mathbf{g}_{Q,k}^{k-D+1}=\mathbf{c}_j\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_k^0\right) \tag{4.18}$$

Using Bayes' rule, the probability inside the product and summations can be expanded

$$p\left(r_k,\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i,\mathbf{g}_{Q,k}^{k-D+1}=\mathbf{c}_j\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_k^0\right)$$

$$=p\left(r_k\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_k^0,\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i,\mathbf{g}_{Q,k}^{k-D+1}=\mathbf{c}_j\right)\Pr\left(\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i,\mathbf{g}_{Q,k}^{k-D+1}=\mathbf{c}_j\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_k^0\right) \tag{4.19}$$

Given the kth sample of complex fading and the kth transmitted symbol, the remaining prior information can be omitted from the conditional density of the received statistic. The conditional probability of the Markov processes can be separated into a product of two probabilities, as the in-phase and quadrature components are independent. Also, the prior information of α_k does not provide any information about the two new probabilities, and can therefore be omitted. With these simplifications, (4.19) can now be restated

$$p\left(r_k,\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i,\mathbf{g}_{Q,k}^{k-D+1}=\mathbf{c}_j\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_k^0\right)$$

$$=p\left(r_k\,|\,\alpha_k,g_{I,k}=c_{i,0},g_{Q,k}=c_{i,0}\right)\Pr\left(\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_{k-1}^0\right)\Pr\left(\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_{k-1}^0\right) \tag{4.20}$$

The conditional probability of the received statistic is recognized as Gaussian with mean $\alpha_k(g_{I,k}+jg_{Q,k})$, and variance $N_0/2$. The terms $\Pr\left(\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_{k-1}^0\right)$ and $\Pr\left(\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_{k-1}^0\right)$ are specific in-phase and quadrature Markov state probabilities, conditioned on the previous received signal samples under a specific transmitted hypothesis. With $P_{I,\alpha}(k,i)\equiv\Pr\left(\mathbf{g}_{I,k}^{k-D+1}=\mathbf{c}_i\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_{k-1}^0\right)$ and $P_{Q,\alpha}(k,j)\equiv\Pr\left(\mathbf{g}_{Q,k}^{k-D+1}=\mathbf{c}_j\,|\,\mathbf{r}_{k-1}^0,\boldsymbol{\alpha}_{k-1}^0\right)$ defined, (4.18) can be rewritten

$$p_{\mathbf{r},\mathbf{a}=\boldsymbol{\alpha}}(\mathbf{r}\,|\,\mathbf{a}=\boldsymbol{\alpha}) = \prod_{k=0}^{N-1}\sum_{i=0}^{K^D-1}\sum_{j=0}^{K^D-1}(2\pi)^{-1}(N_0/2)^{-1/2}\exp\left(-\frac{\left|r_k-\alpha_k(g_{I,k}+jg_{Q,k})\right|^2}{N_0}\right)P_{I,\alpha}(k,i)P_{I,\alpha}(k,j)\,. \tag{4.21}$$

Recursive solutions exist [23] to calculate $P_{I,\alpha}(k,i)$ and $P_{Q,\alpha}(k,j)$

$$P_{I,\alpha}(k,i) \equiv \frac{\sum_{m=0}^{K^D-1} p(r_{k-1}\,|\,g_{I,k-1}=c_{m,0},\alpha_{k-1})P_{I,\alpha}(k-1,m)P_{m,i}}{\sum_{m=0}^{K^D-1} p(r_{k-1}\,|\,g_{I,k-1}=c_{m,0},\alpha_{k-1})P_{I,\alpha}(k-1,m)} \tag{4.22}$$

$$P_{Q,\alpha}(k,j) \equiv \frac{\sum_{m=0}^{K^D-1} p(r_{k-1}\,|\,g_{Q,k-1}=c_{m,0},\alpha_{k-1})P_{I,\alpha}(k-1,m)P_{m,j}}{\sum_{m=0}^{K^D-1} p(r_{k-1}\,|\,g_{Q,k-1}=c_{m,0},\alpha_{k-1})P_{I,\alpha}(k-1,m)} \tag{4.23}$$

for i and j equal to $0,1,\ldots,K^D-1$. Recall from (4.8) that $P_{m,i}$ and $P_{m,j}$ represent Markov state transition probabilities. The conditional densities in (4.22) and (4.23) are Gaussian with variance $N_0/2$, and means equal to the conditional transmitted signal multiplied by the conditional complex fading.

Taking the natural logarithm of (4.21) provides an additive metric. Ideally, the Viterbi algorithm would then be applied to find the MLSE, however, it is apparent from (4.22) and (4.23) that each metric depends on the entire history of the received signal. Hence, the VA cannot be used to find the MLSE unless modified. Tree search algorithms [57] or reduced state Viterbi algorithms [17] that provide sub-optimal sequence estimates are not discussed in this thesis. The transmitted sequence estimates found in this thesis will be obtained using joint MAP sequence detection and channel estimation decoding.

4.3.2. Joint MAP Sequence Detection and Channel Estimation

Joint estimation is the process of obtaining the optimal estimate of a pair of variables at the same time. For certain communication channels and modulations, such as Rayleigh channels with constant amplitude modulation, the joint MAP sequence detection and channel estimation provide the ML sequence estimate and asymptotically the MAP channel estimate [30]. This justifies its use in this thesis. In general, joint MAP estimates are not necessarily equivalent to individual MAP estimates, but are still reliable.

Let the fading vector estimate obtained from the joint MAP decoding be denoted as $\hat{\mathbf{g}} = \{\hat{g}_0, \hat{g}_1, \ldots, \hat{g}_{N-1}\}$, and the set of all possible in-phase or quadrature Markov state sequences as $\mathbf{G}$ with cardinality K^{DN}. The complex fading sequence hypothesis, denoted as $\gamma_I + j\gamma_Q$, represents a specific sequence of Markov states. The joint MAP sequence and channel estimates is then defined

$$(\hat{\mathbf{a}}, \hat{\mathbf{g}})_{MAP} \equiv \max_{(\boldsymbol{\alpha} \in \mathbf{A}, \gamma_I \in \mathbf{G}_I, \gamma_Q \in \mathbf{G})} \Pr(\mathbf{a} = \boldsymbol{\alpha}, \mathbf{g} = \gamma_I + j\gamma_Q \mid \mathbf{r}) \tag{4.24}$$

Under the standard assumption that the fading and transmitted sequences are independent, the probability to be maximized can be rewritten as

$$\Pr(\mathbf{a} = \boldsymbol{\alpha}, \mathbf{g} = \gamma_I + j\gamma_Q \mid \mathbf{r}) = \frac{p(\mathbf{a} = \boldsymbol{\alpha}, \mathbf{g} = \gamma_I + j\gamma_Q, \mathbf{r})}{p_{\mathbf{r}}(\mathbf{r})}$$

$$= p(\mathbf{r} \mid \mathbf{a} = \boldsymbol{\alpha}, \mathbf{g} = \gamma_I + j\gamma_Q) \Pr(\mathbf{g} = \gamma_I + j\gamma_Q) \frac{\Pr(\mathbf{a} = \boldsymbol{\alpha})}{p_r(\mathbf{r})}. \tag{4.25}$$

The term $\Pr(\mathbf{a} = \boldsymbol{\alpha}) / p_{\mathbf{r}}(\mathbf{r})$ is a constant, independent of the symbol or channel hypothesis, and hence can be omitted when determining the joint MAP estimates. The joint MAP sequence and channel estimate (4.24) now reduces to

$$(\hat{\mathbf{a}},\hat{\mathbf{g}})_{MAP} = \max_{(\boldsymbol{\alpha}\in \mathbf{A},\boldsymbol{\gamma}_I \in \mathbf{G}_I,\boldsymbol{\gamma}_Q \in \mathbf{G})} p(\mathbf{r}\mid \mathbf{a}=\boldsymbol{\alpha},\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q)\Pr(\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q). \qquad (4.26)$$

The right hand side of the joint MAP estimates can now be expanded (as was done in the MAP sequence analysis) to yield

$$p(\mathbf{r}\mid \mathbf{a}=\boldsymbol{\alpha},\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q)\Pr(\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q)$$

$$= \prod_{k=0}^{N-1} p(r_k \mid \mathbf{r}_{k-1}^{0},\mathbf{a}=\boldsymbol{\alpha},\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q)\Pr(\mathbf{g}_k^{k-D+1}=\boldsymbol{\gamma}_{I,k} + j\boldsymbol{\gamma}_{Q,k} \mid \mathbf{g}_{k-1}^{k-D}=\boldsymbol{\gamma}_{I,k-1} + j\boldsymbol{\gamma}_{Q,k-1}) \qquad (4.27)$$

Given the transmitted symbol and the fading sample, both with the same time index, the conditional probability of the received signal does not dependant on the remaining prior information. The second term can be expanded into a product of two probabilities, since the in-phase and quadrature components of the fading are independent processes. Following from these observations, (4.27) simplifies to

$$p(\mathbf{r}\mid \mathbf{a}=\boldsymbol{\alpha},\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q)\Pr(\mathbf{g}=\boldsymbol{\gamma}_I + j\boldsymbol{\gamma}_Q)$$

$$= \prod_{k=0}^{N-1} p(r_k \mid a_k=\alpha_k, g_k=\gamma_{I,k,0} + j\gamma_{Q,k,0})\Pr(\boldsymbol{\gamma}_{I,k}\mid \boldsymbol{\gamma}_{I,k-1})\Pr(\boldsymbol{\gamma}_{Q,k}\mid \boldsymbol{\gamma}_{Q,k-1}). \qquad (4.28)$$

The conditional probability of the received statistic is Gaussian with mean $\alpha_k(\gamma_{I,k,0} + j\gamma_{Q,k,0})$ and variance $N_0/2$. The other two terms in the product are recognized as Markov transition probabilities, since the fading hypothesis is a sequence of Markov states. After substituting the first term, the final form of the joint MAP sequence detection and channel estimation becomes

$$(\hat{\mathbf{a}},\hat{\mathbf{g}})_{MAP} = \max_{(\boldsymbol{\alpha}\in \mathbf{A},\boldsymbol{\gamma}_I \in \mathbf{G}_I,\boldsymbol{\gamma}_Q \in \mathbf{G})}$$

$$\sum_{k=0}^{N-1}(2\pi)^{-1}\left(\frac{N_0}{2}\right)^{-1/2}\exp\left\{-\frac{|r_k-\alpha_k(\gamma_{I,k,0}+j\gamma_{Q,k,0})|^2}{N_0}\right\}\Pr(\boldsymbol{\gamma}_{I,k}\mid \boldsymbol{\gamma}_{I,k-1})\Pr(\boldsymbol{\gamma}_{Q,k}\mid \boldsymbol{\gamma}_{Q,k-1}). \qquad (4.29)$$

The exponential term can be interpreted as a simple distance metric, while the two Markov transition probability terms multiply the distance metric according to the probability of a specific fading sequence. Taking the natural logarithm of (4.29) provides an additive metric, and then the joint MAP estimates can be obtained by searching a VA trellis with MK^{2D} states. The cardinality of the signal set is M, the number of Markov states in the in-phase Markov model is K^D, and the number of states in the quadrature component is also K^D. The total number of VA states is a product of these three quantities. To complete the discussion of FSMC receivers, MAP channel state estimation is presented in the next section.

4.3.3. MAP Channel State Estimation

In some communication systems a channel estimate is necessary, while in others it can be used to improve the overall performance. Sub-optimal two-stage receiving strategies exist that make a channel estimate prior to any transmitted symbol decisions [30]. If a feedback channel is available, there are dynamic communication systems that will modify the transmission parameters, depending on channel quality [11]. These techniques provide the motivation for channel estimation. In general, when the channel fading is a continuous correlated process, the optimal channel estimate is difficult, if not impossible, to attain. For FSMCs, on the other hand, the optimal channel estimate is attainable since the channel fading can only occupy a finite number of levels or states.

With MAP as the optimal criterion in this thesis, the MAP channel estimate is stated as

$$\hat{\mathbf{g}}_{MAP} \equiv \max_{(\gamma_I \in \mathbf{G}_I, \gamma_Q \in \mathbf{G}_Q)} \Pr(\mathbf{g} = \gamma_I + j\gamma_Q \mid \mathbf{r}) \tag{4.30}$$

This can be rearranged

$$\hat{\mathbf{g}}_{MAP} = \max_{(\gamma_I \in \mathbf{G}_I, \gamma_Q \in \mathbf{G}_Q)} \frac{p(\mathbf{r} \mid \mathbf{g} = \gamma_I + j\gamma_Q)\Pr(\mathbf{g} = \gamma_I + j\gamma_Q)}{p_{\mathbf{r}}(\mathbf{r})} \tag{4.31}$$

The term $p_{\mathbf{r}}(\mathbf{r})$ is independent of the fading estimate and can be omitted, without loss of generality. The product to be maximized can be rewritten and expanded

$$p(\mathbf{r} \mid \mathbf{g} = \gamma_I + j\gamma_Q)\Pr(\mathbf{g} = \gamma_I + j\gamma_Q)$$

$$= \prod_{k=0}^{N-1} p(r_k \mid \mathbf{r}_{k-1}^0, \mathbf{g} = \gamma_I + j\gamma_Q)\Pr(\gamma_{I,k} \mid \gamma_{I,k-1})\Pr(\gamma_{Q,k} \mid \gamma_{Q,k-1})$$

$$= \prod_{k=0}^{N-1}\left(\sum_{a_k} p(r_k, a_k \mid \mathbf{r}_{k-1}^0, \mathbf{g} = \gamma_I + j\gamma_Q)\right)\Pr(\gamma_{I,k} \mid \gamma_{I,k-1})\Pr(\gamma_{Q,k} \mid \gamma_{Q,k-1})$$

$$= \prod_{k=0}^{N-1}\left(\sum_{a_k} p(r_k \mid \mathbf{r}_{k-1}^0, \mathbf{g} = \gamma_I + j\gamma_Q, a_k)\Pr(a_k)\right)\Pr(\gamma_{I,k} \mid \gamma_{I,k-1})\Pr(\gamma_{Q,k} \mid \gamma_{Q,k-1}) \tag{4.32}$$

The transmitted symbols are considered to be equally probable, so $\Pr(a_k)$ is equal to $1/M$ for M-ary modulation. The conditional probability inside the summation is Gaussian with mean $a_k(\gamma_{I,k,0} + j\gamma_{Q,k,0})$ and variance $N_0/2$. The final form of the MAP channel estimate can now be stated as:

$$\hat{\mathbf{g}}_{MAP} = \max_{(\gamma_I \in \mathbf{G}_I, \gamma_Q \in \mathbf{G}_Q)}$$

$$\frac{1}{M^N}\prod_{k=0}^{N-1}\left(\sum_{a_k}(2\pi)^{-1}\left(\frac{N_0}{2}\right)^{-1/2}\exp\left\{-\frac{|r_k - a_k(\gamma_{I,k,0} + j\gamma_{Q,k,0})|^2}{N_0}\right\}\right)\Pr(\gamma_{I,k} \mid \gamma_{I,k-1})\Pr(\gamma_{Q,k} \mid \gamma_{Q,k-1})$$

(4.33)

Although an additive metric is obtained by taking the natural logarithm of (4.33), it is worthwhile to note that the exponential remains since it is inside a summation. Searching a trellis with K^{2D} states using the Viterbi algorithm will provide the MAP channel estimate.

Three different MAP receiver criteria were discussed in this section. The MAP sequence estimate was found to be difficult to obtain directly, but a good sequence estimate was attainable from joint MAP sequence detection and channel estimation. In the next section joint MAP receivers are implemented for use on communication channels that have different fading rates. Both the joint MAP sequence and channel estimates are then used to compare the Markov receivers, and the error performance results are shown and discussed.

4.4. Simulation Results

Both 1st and 2nd order FSMCs were constructed to represent slow, medium and fast fading channels. The state-spaces of the 1st order Markov models ($D=1$) were determined using the LBG algorithm. The number of states in the 1st order Markov models was 2, 4, 8, 16 and 32. Although transition probabilities are different for channels with different fading rates, the state-spaces were identical since the 1st order distributions of the fading component were identical for all three channels.

The state spaces of the 2nd order ($D=2$) FSMCs were determined using two different methods. The LBG algorithm provided 2nd order state spaces via *vector quantization* (VQ). 2nd order state-spaces were also constructed from *Cartesian products* (CP) of two 1st order state-spaces. An example of the VQ state-space was shown in Figure 4.4, and an example of the CP of two 1st order state-spaces is now shown in Figure 4.5. The number of states used in the 2nd order models was 2^2, 4^2 and 8^2. The numbers are in exponential form so comparison can be made to the number of states in the 1st order models, normalized by model order. Higher numbers of Markov states were not considered for the 2nd order models, as implementation of these receivers proved to be too complex.

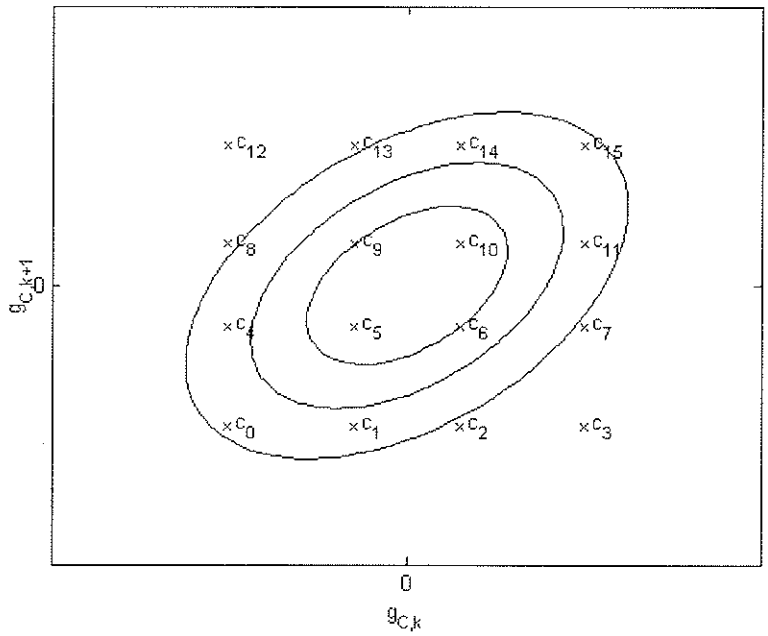

Figure 4.5: 2nd order state-space from CP of two 4-state 1st order state-spaces

The sequence and channel estimations were obtained from joint MAP sequence detection and channel estimation. The modulation scheme for the simulations was OOK, with performance averages obtained from test sequences of length 100,000. BER performances are investigated first in this section for the various Markov receivers, followed by channel estimate error analysis.

4.4.1. Sequence Detection Results

The error performance results of the 1st order, 2nd order CP, and 2nd order VQ Markov receivers for the slow fading channel are shown in Figure 4.6, Figure 4.7 and Figure 4.8 respectively. As mentioned in the previous section, the Kalman filtering receiver meets the lower bound of coherent detection on the slow fading channel. It is apparent from these figures that under certain conditions, the FSMC receivers implemented have the ability to perform comparably to the reference receiver.

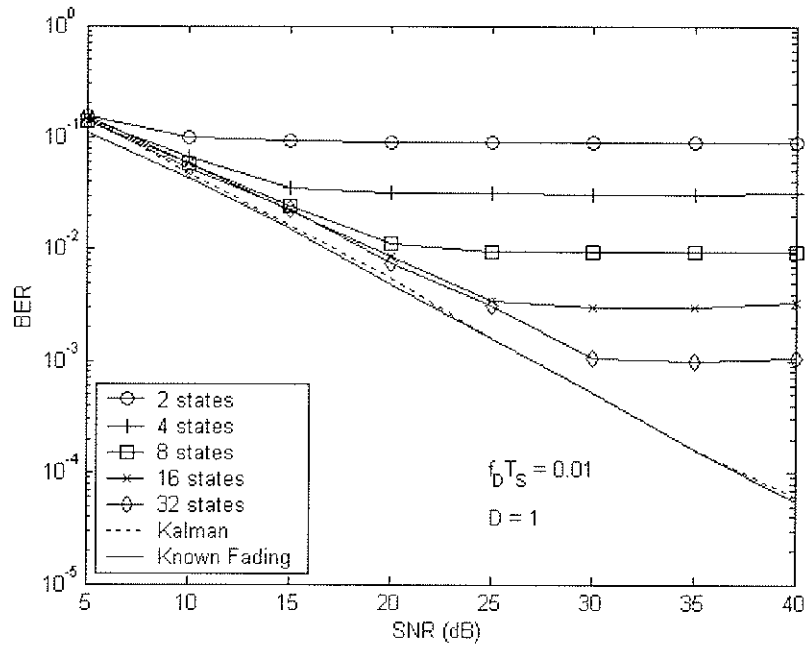

Figure 4.6: BERs of 1st order Markov receivers for slow fading channel

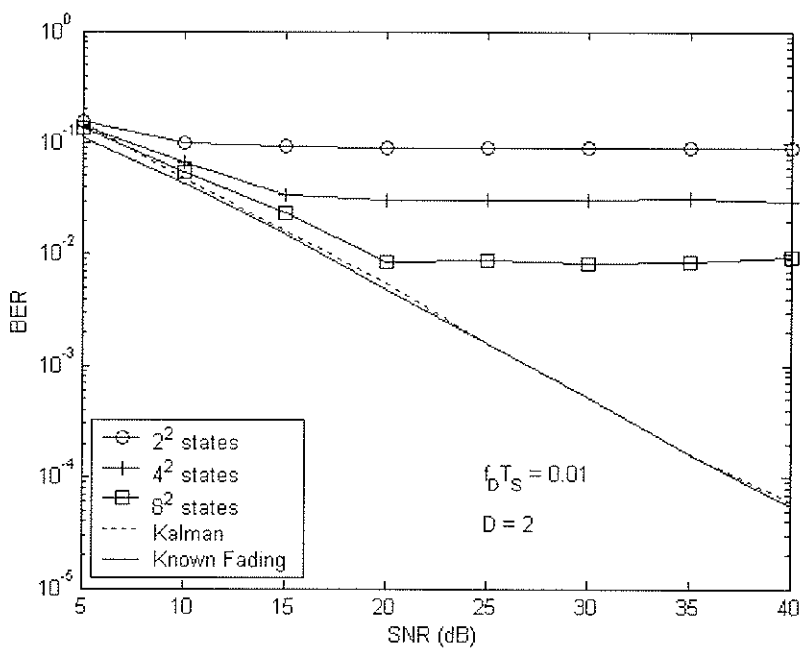

Figure 4.7: BERs of 2nd order CP Markov receivers for slow fading channel

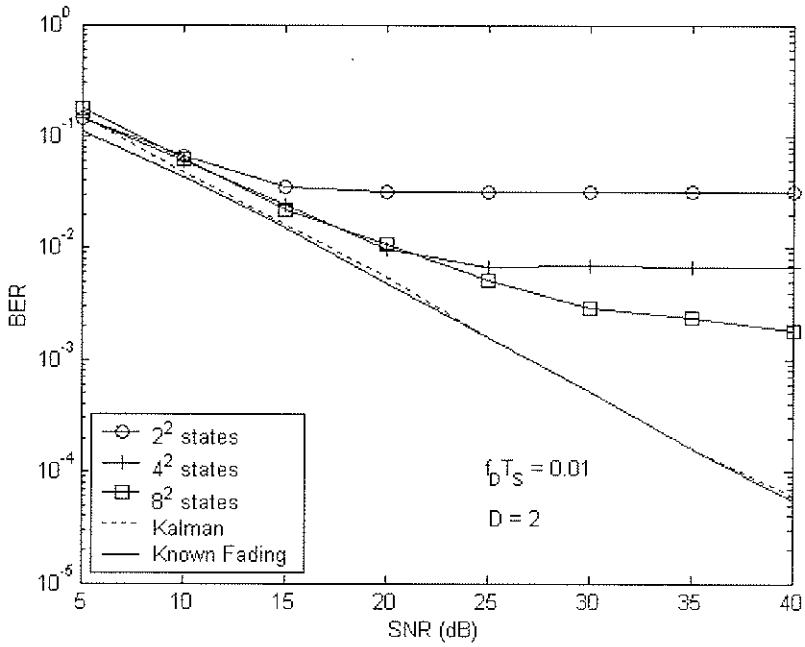

Figure 4.8: BERs of 2nd order VQ Markov receivers for slow fading channel

For higher SNR, the performance of Markov receivers is restricted by the number of states in the Markov model. For the 1st order receivers in Figure 4.6, the error floor appears to decrease linearly with the number of states in the Markov model.

Comparing the results of the 1st order Markov receivers in Figure 4.6 to the 2nd order CP Markov receivers in Figure 4.7 shows there is no advantage to the 2nd order CP receivers. A 1st order receiver with K states performs as well as a 2nd order CP receiver with K^2 states. The 2nd order CP receiver, however, is K^2 times (a factor of K from each fading component) more complex to implement because of the increase in number of VA trellis states.

When comparing the performance of 2nd order CP Markov receivers in Figure 4.7 with the performance of 2nd order VQ Markov receivers in Figure 4.8, it is apparent that for a given number of states, the VQ Markov receivers outperform the CP Markov receivers. The VQ Markov receivers do not, however, outperform the 1st order Markov

receivers. Comparing Figure 4.6 to Figure 4.8 the 32-state 1st order Markov receiver performs better than the 8^2-state 2nd order VQ Markov receiver. The 1st order receiver also has a smaller number of states and is hence less complex to implement. For slow fading channels, it appears that a 1st order Markov receiver will perform better than a 2nd order Markov receiver with the same complexity.

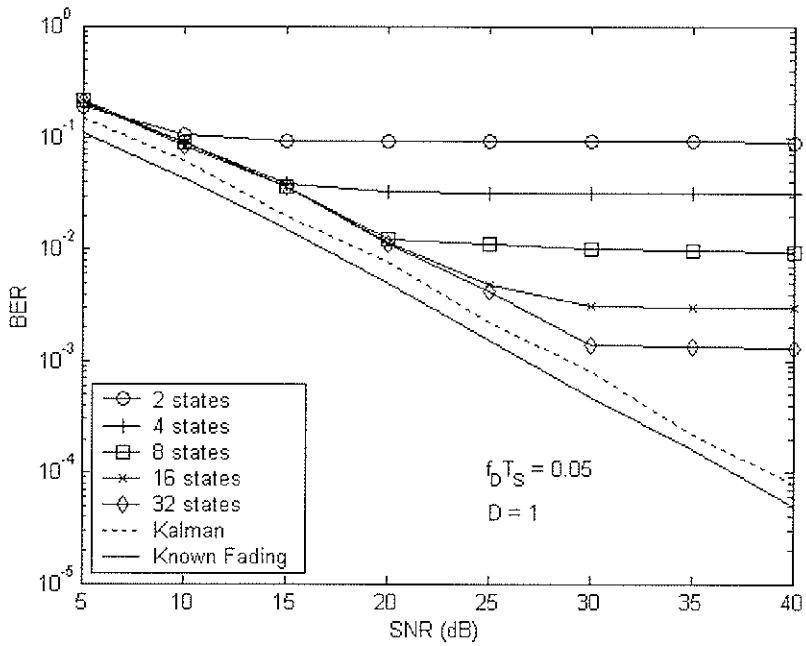

Figure 4.9: BERs of 1st order Markov receivers for medium fading channel

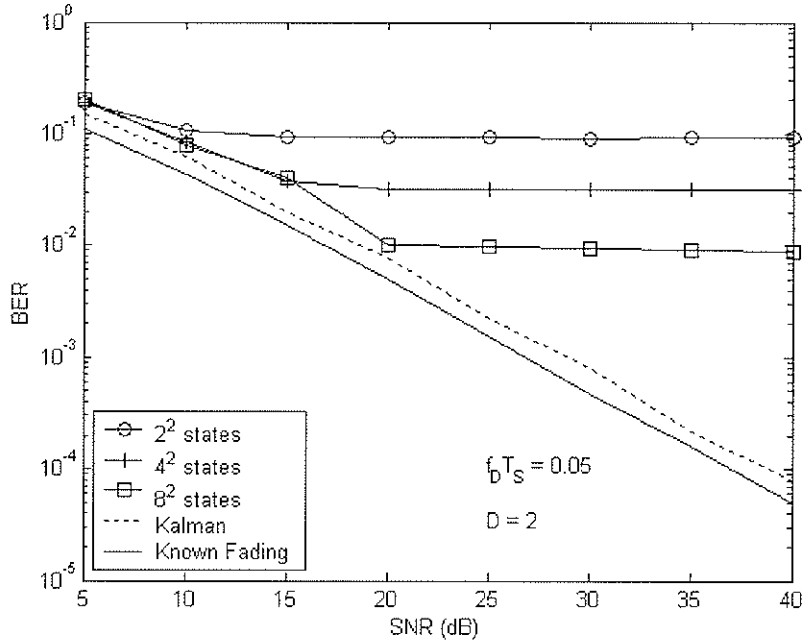

Figure 4.10: BERs of 2nd order CP Markov receivers for medium fading channel

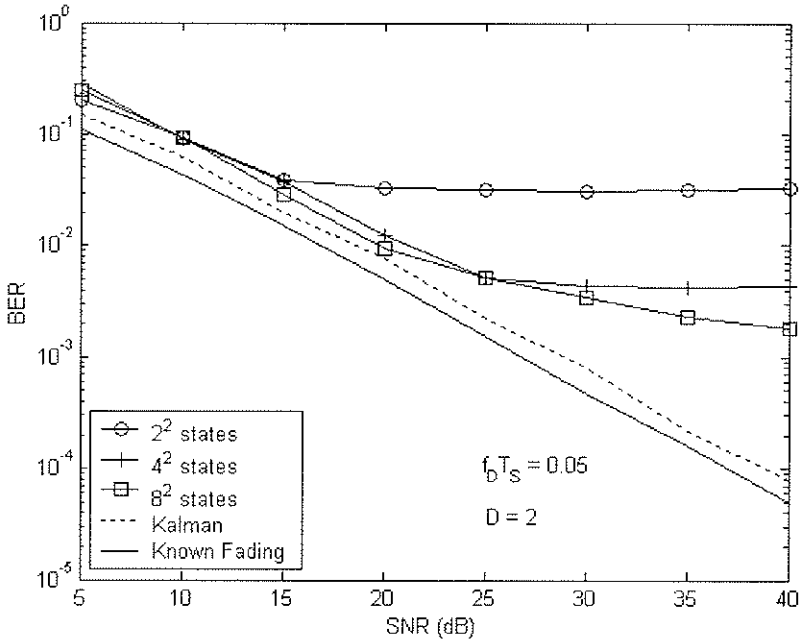

Figure 4.11: BERs of 2nd order VQ Markov receivers for medium fading channel

Figure 4.9, Figure 4.10 and Figure 4.11 show the error performance results of the 1st order, 2nd order CP, and 2nd order VQ Markov receivers respectively, for the medium fading channel. The Markov receivers exhibit the same trend as found with the Kalman filtering receivers, where receiver performance was not as close to the lower bound for the medium fading channel as for the slow fading channel. Since there is less correlation in the medium fading channel than in the slow fading channel, neither the Kalman filtering receiver performance nor the Markov receiver performance can approach the coherent detection lower bound.

The error floors observed for the slow fading channel are also evident for the medium fading channel. In fact, the error floors experienced on the medium fading channel are very close, if not identical, to the error floors of the slow fading channel. This suggests that for a specific number of Markov states, the error floor is not dependant on the fading rate of the channel.

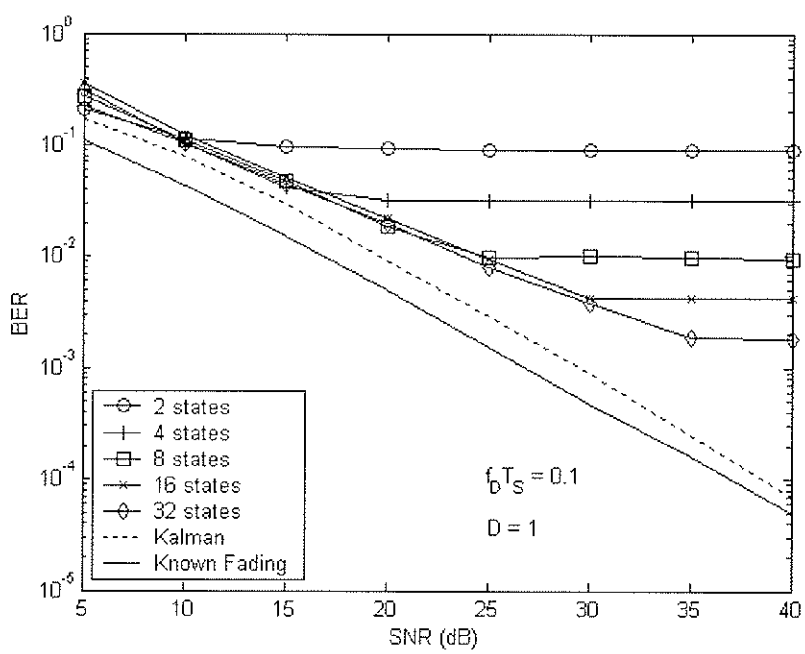

Figure 4.12: BERs of 1st order Markov receivers for fast fading channel

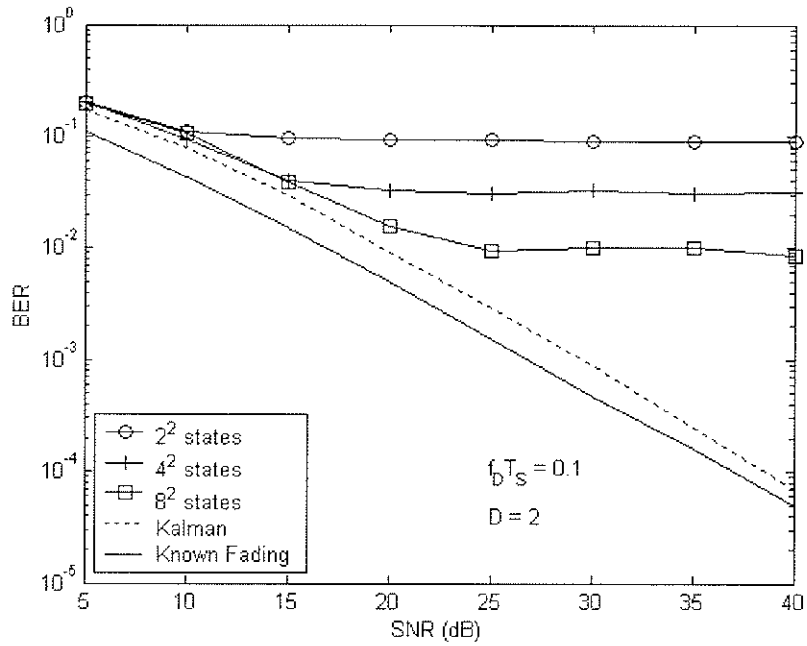

Figure 4.13: BERs of 2nd order CP Markov receivers for fast fading channel

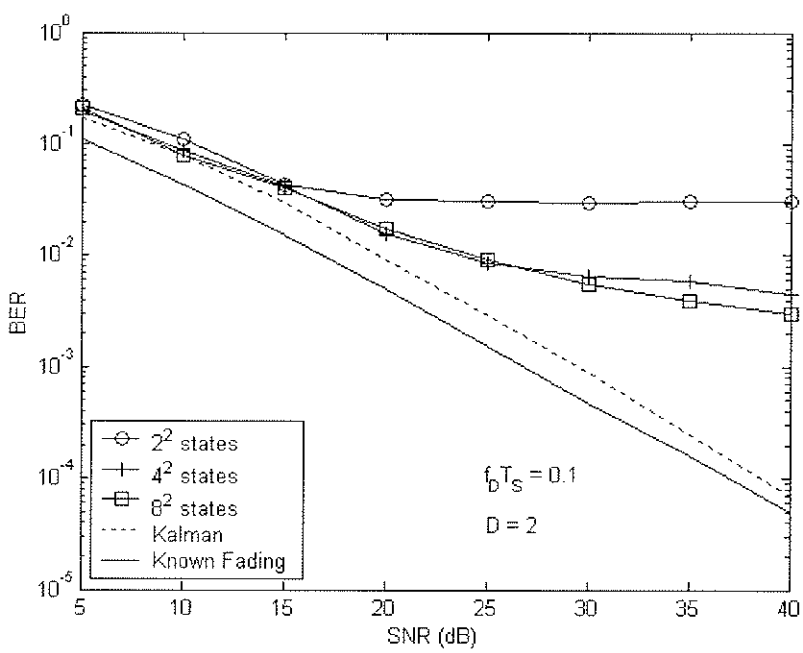

Figure 4.14: BERs of 2nd order VQ Markov receivers for fast fading channel

The error performance results of the 1st order, 2nd order CP, and 2nd order VQ Markov receivers for the fast fading channel are shown in Figure 4.12, Figure 4.13, and Figure 4.14 respectively. The same trends in Markov receiver performance for slow and medium fading channels can be observed in the Markov receiver performance for fast fading channels:

i) when higher correlation is evident in the channel, Markov receivers perform closer to the coherent detection lower bound,

ii) performance of both 1st and 2nd order Markov receivers is limited by the number of states in the Markov model,

iii) 2nd order CP Markov receivers perform the same as 1st order Markov receivers for a specific number of normalized states,

iv) 2nd order VQ Markov receivers perform better than the 2nd order CP Markov receivers with the same number of states, and

v) 1st order Markov receivers perform better than 2nd order VQ Markov receivers with the same complexity.

Since joint MAP sequence detection and channel estimation was performed, channel estimates are also available for receiver performance evaluation. The ability of the Markov receivers to produce accurate channel estimates is discussed next.

4.4.2. Channel Estimation Results

If the real channel was in fact a FSMC, a state error rate could have been calculated by simply comparing the true channel state sequence to the channel sequence estimate obtained via joint MAP detection. However, since the channel is not a FSMC, but in fact a Jakes-Clarke channel, the Markov fading values corresponding to the state sequence estimates must be compared to the real valued channel fading. This

performance measure is referred to as the *channel estimation error* (CEE). The CEE is normalized by the power in the fading.

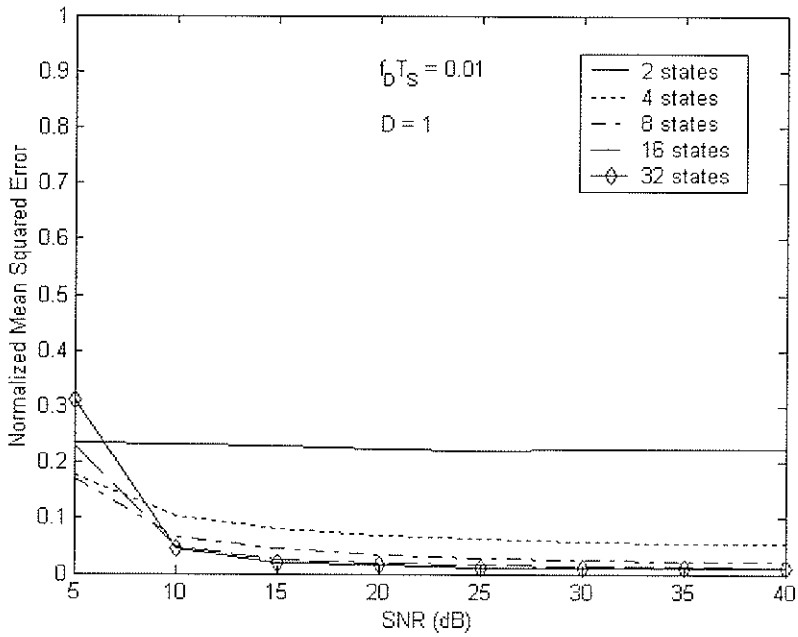

Figure 4.15: CEEs of 1st order Markov receivers for slow fading channel

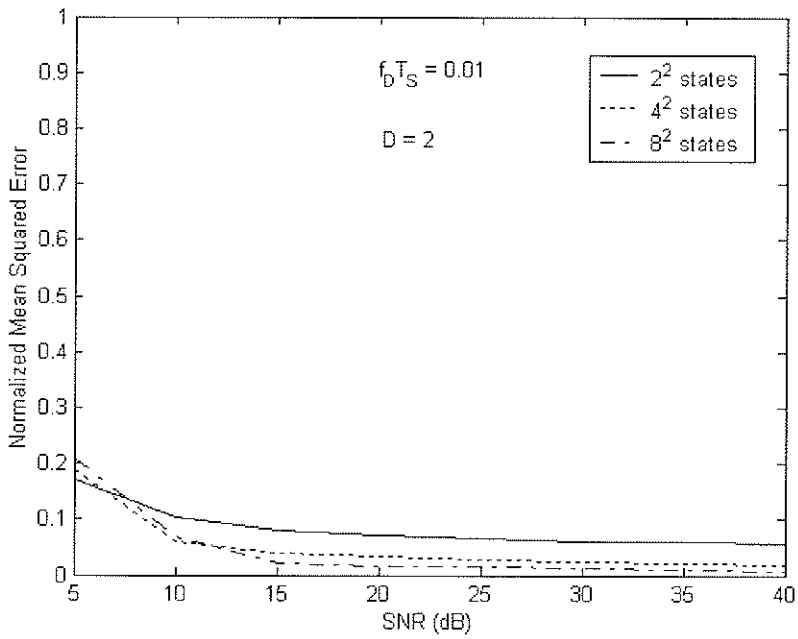

Figure 4.16: CEEs of 2nd order VQ Markov receivers for slow fading channel

The channel estimation performance of the 1st order and 2nd order VQ Markov receivers for the slow fading channel are shown in Figure 4.15 and Figure 4.16 respectively. For high SNR, the receivers with a higher number of states result in less CEE. For low SNR, though, an unexpected phenomenon is observed: better performance is attained from receivers with a smaller number of Markov states. This phenomenon can be explained by realizing that there are two contributors to the CEE: the state error from making an erroneous channel state estimate, and the quantization error related to the construction of the Markov state-space. When noise power is strong, the probability of making an incorrect state estimate is greater with a higher number of states. For low SNR, the state estimation error is the primary contributor to the CEE. For high SNR, the quantization error is the main contributor to the CEE.

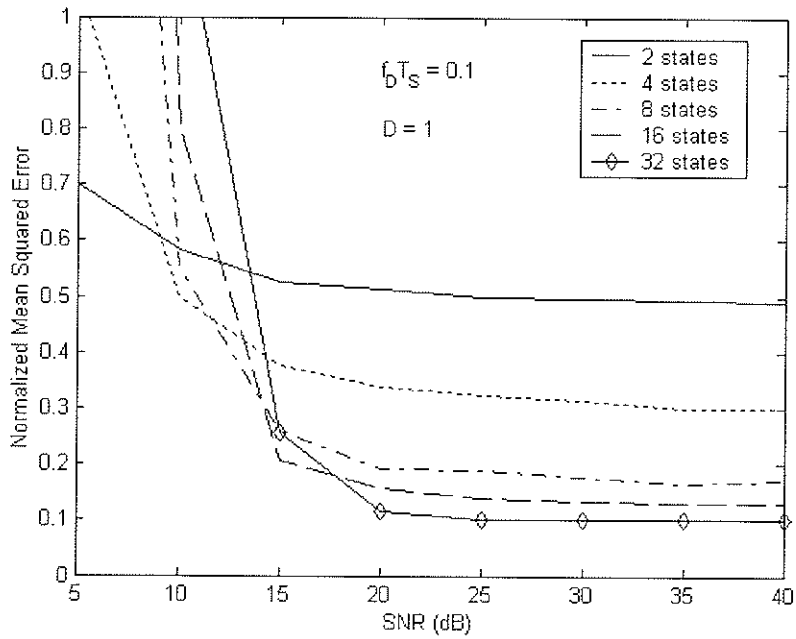

Figure 4.17: CEEs of 1st order Markov receivers for fast fading channel

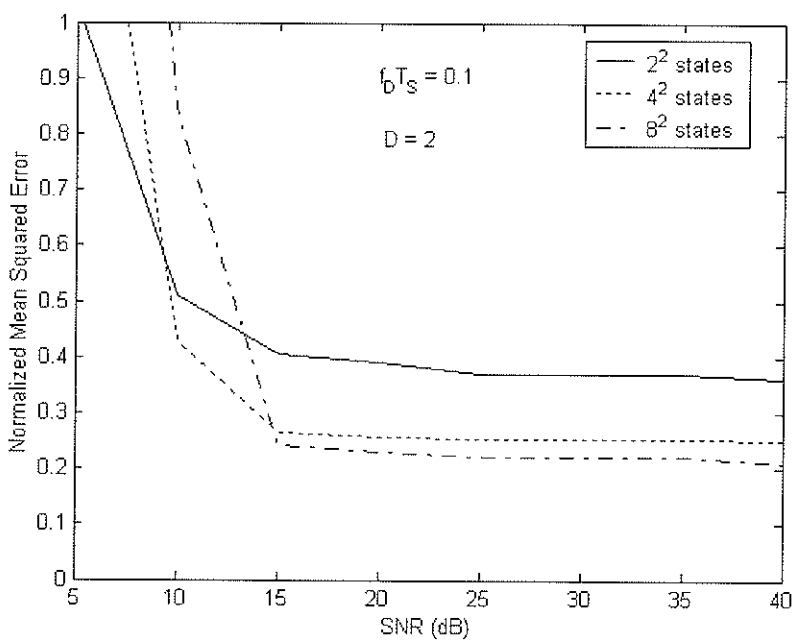

Figure 4.18: CEEs of 2nd order VQ Markov receivers for fast fading channel

Figure 4.17 and Figure 4.18 show the channel estimation error results of the 1st order and 2nd order VQ Markov receivers respectively, for a fast fading channel. There is considerably more CEE observed on the fast fading channel than was observed on the slow fading channel. For low SNR, higher CEE occurs when a Markov receiver has a greater number of states. This trend is more evident on the fast fading channel than on the slow fading channel. Remarkably, the channel estimates are so poor for low SNR, that the CEEs of the Markov receivers are occasionally larger than unity.

For high SNR on both the slow and fast fading channels, the 2nd order VQ Markov receivers have lower CEE than the 1st order Markov receivers with a normalized number of states. The channel estimate performances of 2nd order CP Markov receivers were also obtained and found to be the same as the 1st order Markov receivers. In addition, Markov receivers were implemented for the medium fading channel. The CEE observations corresponded with those for receivers on the slow and fast fading channels. These observations can be summarized as:

i) lower CEE can be achieved on channels exhibiting a higher correlation,

ii) 1st order Markov receivers perform as well as 2nd order CP Markov receivers with a normalized number of states,

iii) 2nd order VQ Markov receivers perform better than 1st order Markov receivers with a normalized number of states,

iv) Markov receivers with a smaller number of states achieve a lower CEE for low SNR, and

v) Markov receivers with a larger number of states achieve a lower CEE for high SNR.

For brevity, channel estimation performance of 2nd order CP Markov receivers as well as medium fading channel results are not included. This is justifiable since the primary

output of a communication receiver is the symbol estimates and not the channel estimates.

This section discussed performance of an OOK communication system on a channel experiencing a time-varying complex function that multiplies the transmitted signal. It was shown that the receiver performance was highly dependant on the receiver used as well as the channel fading rate. The next section briefly discusses receiver design for communication over a closely related communication system.

4.5. Quadrature Phase Shift Keying over Phase-compensated Channels

This thesis has focused on a communication channel that results in a complex time-varying multiplication of the transmitted signal as well as additive white noise. The complex multiplication is observed when there is no gain control or phase compensation device implemented at the receiver. If the receiver employs a phase compensation device, such as a PLL, the channel can be modeled as a time-varying gain on the transmitted signal plus additive white noise. This correlated time-varying gain is positive, real and Rayleigh distributed. There has been much documented research focusing on the development of Markov models of the Rayleigh distributed envelope [3] [49] [50].

The original plan for this research was not to develop Markov models of the complex fading process, but to develop Markov models of the Rayleigh distributed fading envelope. This plan was discarded, though, as the optimal receiver is trivial to implement if a proper modulation scheme is chosen. *Quadrature phase shift keying* (QPSK) is the specific modulation scheme that trivializes the design of an optimal receiver. The phase-compensated channel and QPSK are the focus of this section.

The QPSK modulation scheme consists of $M = 4$ signals with equal energy. The signal constellation is such that each signal has an antipodal partner, and zero correlation with the remaining two signals. A QPSK signal constellation is shown in Figure 4.19. This modulation scheme was initially chosen because the information is in the phase of the signals, and therefore less susceptible to the negative channel effects.

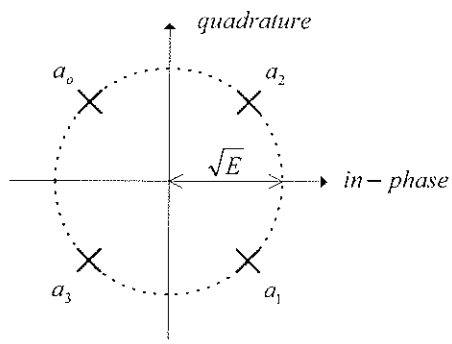

Figure 4.19: QPSK signal constellation

Only 1st order Markov models were constructed, and their state spaces were determined from the optimum quantization of a 1st order Rayleigh density. An example state-space with 4 states is shown in Figure 4.20. Markov models were constructed with 2, 4, 8 and 16 states. The Markov transition probabilities were determined using Monte-Carlo methods, and the Markov models were based on a slow fading channel. It will be shown, however, that receiver performance is not dependent on the fading rate of the channel.

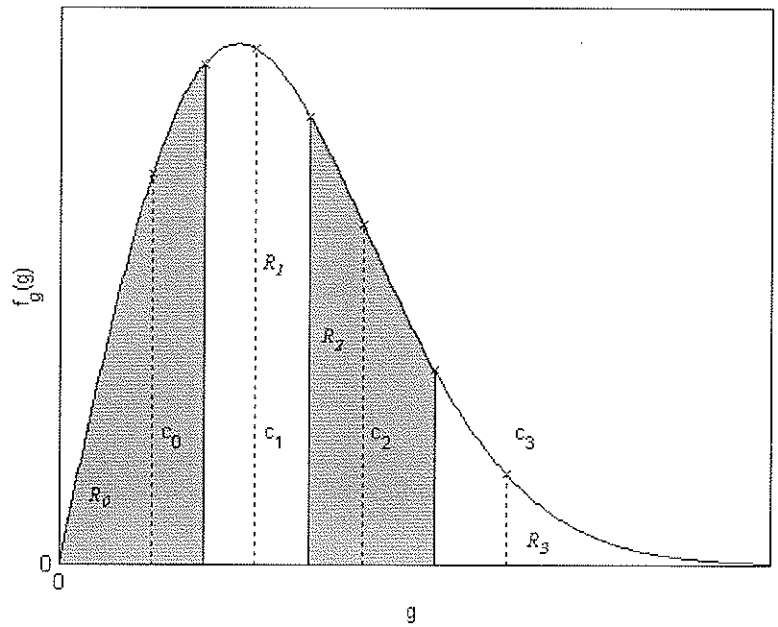

Figure 4.20: Quantization of 1st order density for 4-state 1st order state-space

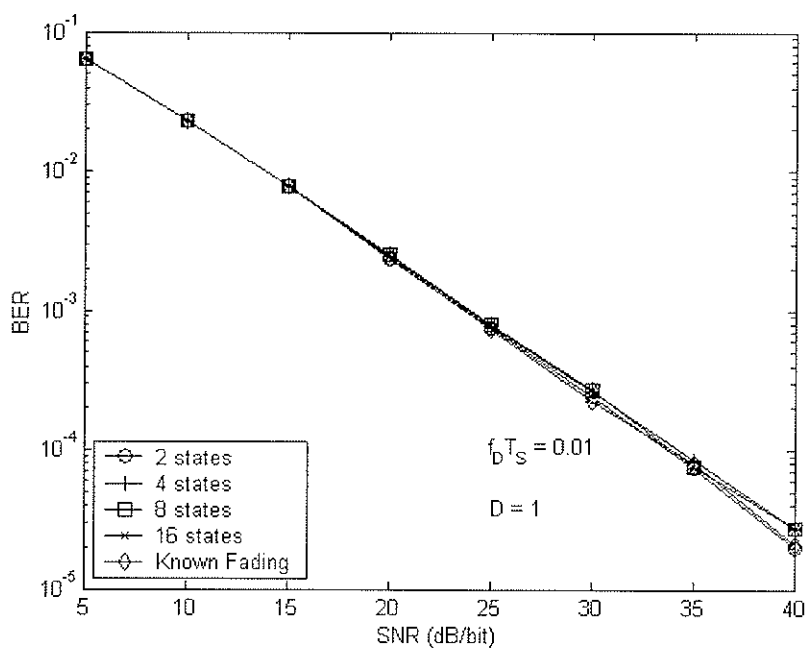

Figure 4.21: BERs of Markov receivers for phase-compensated channel

The equation used for joint MAP sequence detection and channel estimation on phase-compensated channels was derived by simply modifying the previous joint MAP equation from Section 4.3. The BER performance of the joint MAP Markov receivers is shown in Figure 4.21. It is immediately apparent that the performances of all of the Markov receivers are identical to the lower bound provided by the coherent receiver. Further investigation reveals that the actual signal estimates obtained by the Markov receivers are the same as the estimates obtained by the coherent receiver. There is a simple explanation for this. For any given time interval, if the fading value is known, the decision regions will be the same. The optimum strategy, therefore, is to choose the signal in the same quadrant as the received statistic. Decisions in this strategy do not depend on any other information. Since the memory in the received signal is not considered when determining the optimal sequence estimate, receiver performance does not depend on the fading rate of the channel.

Modeling the Rayleigh envelope as a Markov process does serve a purpose in receiver design. When QPSK is used, determining the optimal channel estimate requires that the memory in the received signal be considered. Determining the optimal sequence estimate, when using a modulation scheme that contains information in the constellation amplitudes, would also require the memory in the received signal. *Quadrature amplitude modulation* (QAM) is a scheme that contains information in the signal amplitudes if there are more than 4 signals in the signal constellation. 16 or more signals are often employed due to bandwidth constraints. If 64-QAM were used in a communication system with phase compensation, utilizing the memory of the channel would dramatically improve receiver performance.

4.6. Summary

The focus of this chapter was the finite state Markov channel. After a brief introduction to Markov models, the FSMC was related to the fading channel discussed in Chapter 2. It was shown that a FSMC could be theoretically constructed using probability distributions of the channel fading, or practically constructed using training sequences.

Three closely related receiver criteria were stated. Estimation methods that satisfy each criterion were derived for FSMCs. The criterion that was selected for receiver design was joint MAP sequence detection and channel estimation. 1st and 2nd order Markov receivers were designed for use on slow, medium and fast fading channels.

The performances of the Markov receivers were compared. They were also compared to the performances of the reference receivers chosen in Chapter 3. Comparisons were made for average sequence estimation error as well as average channel estimation error. It was determined that under certain conditions, the performance of Markov receivers was comparable to that of the reference receivers. It was determined that there was no benefit to be had in implementing a 2nd order Markov receiver over a 1st order Markov receiver, because of the corresponding increase in receiver complexity.

Digital communication of QPSK signals over a phase-compensated fading channel was the last consideration of the chapter. For this special case, it was concluded that there was no benefit to modeling the channel as a Markov process. The optimal receiving strategy was determined to be symbol-by-symbol detection.

A summary of this thesis, and some suggestions for further investigation are given in the next chapter.

Chapter 5

Conclusions and Suggestions for Further Study

This thesis focused on the design of simple receivers for efficient digital communication over mobile wireless channels. A mathematical model of the communication channel was developed from assumptions of the propagation environment. The received signal was modeled as the transmitted signal multiplied by a correlated complex time-function plus additive white noise. Optimum detection strategies were discussed and implemented which take advantage of the correlation in the received signal. Kalman filtering receivers, which perform maximum *a posteriori* sequence detection, were chosen as the reference receivers. The Markov receivers implemented for this thesis performed joint MAP sequence detection and channel estimation. A discussion of this research, and opportunities for future analysis are discussed in this final chapter.

5.1. Discussion

The primary goal of this research was to design efficient communication receivers by modeling the Jakes-Clarke channel as a finite state Markov channel. Previous research, as mentioned in Chapter 1, has offered differing conclusions on the Markov model order that best represents the true channel. The criterion used in this

thesis to judge the validity of a Markov model is the performance of a Markov receiver designed from FSMC parameters.

From the results in Chapter 4, it is evident that the 1st order Markov models represent the fading channel well, provided a sufficient number of states are used in the Markov model. The performance of a Markov receiver, however, is ultimately limited by the number of states in the Markov model. If the number of states is large enough, though, the Markov receivers perform comparably to the reference Kalman filtering receivers. For channels with a small fading rate, the performance of the Markov receivers approaches the lower performance bound given by the coherent receivers.

In comparing the performance of the 1st order Markov receivers to the 2nd order CP Markov receivers, it appears that increasing the order of Markov model does not effectively increase receiver performance. The drastic performance differences between the 2nd order CP Markov receivers and the 2nd order VQ Markov receivers show, however, that efficient state-space construction is important when designing a Markov receiver.

If the number of states in the VA trellis is the sole determinant of receiver complexity, comparing the efficiency of Markov receivers and Kalman filtering receivers is relatively straightforward. The number of states in the VA trellis, when using Markov receivers, is MK^{2D}, where M is the cardinality of the signal set, K^D is the number of states in the in-phase Markov representation, and K^D is the number of states in the quadrature Markov representation. When using Kalman prediction filtering, the number of states in the VA trellis is M^L, where L is the assumed memory length of the received signal. The number of states required in a Markov model can be determined, for certain SNR and BER specifications, by examining the receiver performances in Chapter 4. A sample of comparisons is given in Table 5.1.

Table 5.1: Comparing receiver complexity by number of VA states

Fading rate	**SNR (dB)**	**Tolerable BER**	**# of VA states for Kalman filtering receiver**	**# of VA states for Markov receiver**
Slow ($f_D T_S = 0.01$)	30	10^{-3}	$2^2 = 4$	$2 \times 32^2 = 2048$
Medium ($f_D T_S = 0.05$)	30	5×10^{-2}	$2^3 = 8$	$2 \times 16^2 = 512$
Fast ($f_D T_S = 0.1$)	25	10^{-2}	$2^5 = 32$	$2 \times 8^2 = 128$

These observations are somewhat misleading and do not immediately support the efficiency of Markov receivers. The flaw in this complexity analysis is that the mathematical operations required to arrive at the VA metrics are not considered. The metrics used by the Markov receivers are extremely simple to obtain and are acquired by straightforward scalar operations. The metrics associated with the Kalman filtering receivers, however, require complicated matrix manipulations. The computational cost for executing these algorithms, as well as the circuitry cost for implementing the receivers, should be the true basis when comparing receiver complexity.

In addition to BER performance, errors in the channel estimates were also compared. The main conclusion from comparing CEE results is that the estimation error has primarily two contributors: the error from choosing the wrong channel state, and the quantization error associated with the Markov model state-space. It was observed that for low SNR, the channel state errors were the dominant contributor to the CEE, and that for high SNR, the quantization error was the main contributor.

5.2. Suggestions for Further Study

When studying digital communications, changes in the assumptions and/or scope of previous research can often result in interesting prospects for further study. This section suggests some potential modifications to both the communication system and methods of analysis discussed in this thesis. These ideas could provide additional insight into digital communications and mobile wireless receiver design

In this thesis, an uncoded OOK modulation scheme was employed for simplicity, although there were several modulation techniques available with memory in the transmitted signal sequence. These modulation schemes with memory can take advantage of slowly varying phase and amplitude in the received signal. Differential modulation schemes [26], or the Kineplex system [14], are examples of these. Obviously, if modulation schemes with memory were utilized, the effectiveness of modeling the channel as a Markov process might not be as evident, but better receiver performance could be expected.

Although the frequency non-selective slow fading channel was chosen for this thesis, further investigation of different types of mobile wireless channels could prove interesting. Typically it is more difficult for receivers to perform well when communicating over channels that exhibit ISI, such as frequency selective or band-limited channels. If there is a clear path from the transmitter to the receiver, the channel is modeled as Rician, and better performance can be expected. Constructing Markov models of these channels and implementing Markov receivers should be simple, since Markov models can be generated from training sequences of channel data.

The Markov representation of the Jakes-Clarke channel in this thesis consisted of two independent Markov processes: a Markov model of the in-phase component of the

fading, and a Markov model of the quadrature component of the fading. It is widely known that joint quantization of two random variables provides a more efficient representation than that of the two variables quantized independently. This holds true, regardless of any correlation between the two variables. This idea could be extended to the Markov representation of the complex fading process. It should be expected that a Markov model of the fading with a complex state-space would be more efficient than the in-phase and quadrature Markov models employed in this thesis.

Using a method other than optimum density quantization when determining the Markov model state-spaces may be of significance. Intuitively, the Markov receiver estimates are more reliable when the envelope of the fading is large. It is also easier to track the fading when the envelope is large, because of the high instantaneous SNR. This suggests that for receiver design, fewer Markov states need be required to represent the envelope distribution corresponding to a "high SNR". It is this author's opinion that increasing the number of Markov states used to represent the "low SNR" of the envelope distribution, would result in a more efficient use of the Markov states.

The design of receivers depends on the fading rate of the channel, and hence the speed of the mobile. The assumption that the mobile speed is constant may not be a reasonable assumption. The versatility of the Markov receiver in performing on channels with dynamic fading rates is a valid problem. This comparison should be made against the robustness of the Kalman filtering or prediction-filtering receivers on the same channels.

Finally, extended analysis of the Markov receivers could be performed. Instead of using a Markov receiver on the Jakes-Clarke channel, it could be used on the actual FSMC for which it was designed. The (expected) decrease in Markov receiver performance when used for the FSMC, rather than the Jakes-Clarke channel, would prove interesting.

In summary, the suggestions for future research are:

i) performing further complexity analysis to include the cost of mathematical operations when determining VA metrics

ii) implementing modulation schemes with memory such as *differential phase shift keying* (DPSK),

iii) using a modified communication channel such as a frequency selective channel, Rician channel or phase-compensated Rayleigh channel,

iv) determining the Markov state-spaces by a joint quantization of the in-phase and quadrature components,

v) developing Markov state-spaces by using more states to represent "low SNR", and fewer states to represent "high SNR",

vi) testing the robustness of Markov receivers on channels with a different or variable fading rate, and

vii) investigating the (expected) degradation of Markov receiver performance when the true channel is a Jakes-Clarke channel instead of a finite state Markov channel.

Appendix A: Derivation of (2.19)

With the input to the filter having a flat power spectrum of strength one, and desiring to have an output power spectrum of (2.12), one can write

$$H(f)H^*(f) = S_D(f) = \frac{1}{2\pi f_D\sqrt{1-(f/f_D)^2}}, |f| \le f_D. \tag{A-1}$$

Assuming $H(f)$ is real

$$H(f) = \frac{1}{\sqrt{2\pi f_D}\left(1-(f/f_D)^2\right)^{1/4}}, |f| \le f_D. \tag{A-2}$$

Taking the inverse transform

$$h(t) = \int_{-\infty}^{\infty} H(f)\exp(j2\pi ft)df = \int_{-f_D}^{f_D} \frac{\exp(j2\pi ft)}{\sqrt{2\pi f_D}\left(1-(f/f_D)^2\right)^{1/4}}df \tag{A-3}$$

Substituting $\lambda = f/f_D$

$$h(t) = \sqrt{\frac{f_D}{2\pi}}\int_{-1}^{1}\frac{\exp(j2\pi f_D\lambda t)}{\left(1-\lambda^2\right)^{1/4}}d\lambda \tag{A-4}$$

$$= \sqrt{\frac{f_D}{2\pi}}\left[\int_{-1}^{1}\frac{\cos(2\pi f_D\lambda t)}{\left(1-\lambda^2\right)^{1/4}}d\lambda + j\int_{-1}^{1}\frac{\sin(2\pi f_D\lambda t)}{\left(1-\lambda^2\right)^{1/4}}d\lambda\right] \tag{A-5}$$

The second term is zero since the integral is over an odd function. The impulse response then becomes

$$h(t) = \sqrt{\frac{2f_D}{\pi}} \int_0^1 \frac{\cos(2\pi f_D \lambda t)}{\left(1-\lambda^2\right)^{1/4}} d\lambda \ . \tag{A-6}$$

Using equation 3.771-8 from [24] (page 427), with $u = 1$, $v = 1/4$ and $a = 2\pi f_D t$ the expression simplifies to

$$h(t) = \frac{\Gamma(3/4)}{\sqrt{2}} \left(\frac{f_D}{\pi |t|} \right)^{1/4} J_{1/4}(2\pi f_D |t|) \qquad \mathrm{t} \neq 0$$

$$\lim_{t \to 0} h(t) = \frac{1}{\sqrt{2}} \frac{f_D^{\ 1/2} \Gamma(3/4)}{\Gamma(5/4)} , \tag{A-7}$$

which corresponds with (2.19) as required.

Appendix B: Derivation of (3.19) and (3.20)

The conditional mean of the received signal $\tilde{r}_k$ and its respective variance can be obtained from the Kalman filter output. By definition

$$\tilde{r}_k \equiv \xi\{r_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \tag{A-8}$$

It then follows by substitution

$$\begin{aligned}
\tilde{r}_k &= \xi\{\mathbf{h}_k^T \mathbf{g}_k^{k-L+1} + n_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= \mathbf{h}_k^T \xi\{\mathbf{g}_k^{k-L+1} \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} + \xi\{n_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= \mathbf{h}_k^T \tilde{\mathbf{g}}_k^{k-L+1} + \xi\{n_k\} \\
&= \mathbf{h}_k^T \tilde{\mathbf{g}}_k^{k-L+1}
\end{aligned} \tag{A-9}$$

The fact that the noise is zero mean was used, and $\xi\{\mathbf{g}_k^{k-L+1} \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$ is recognized as the Kalman fading estimate $\tilde{\mathbf{g}}_k^{k-L+1}$. The conditional mean corresponds with (3.19) as required. To find the conditional variance of the received signal estimate, the definition is similarly introduced

$$\sigma_{r,k}^2 \equiv \xi\{(r_k - \tilde{r}_k)(r_k - \tilde{r}_k)^H \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \tag{A-10}$$

where $(\cdot)^H$ is the complex transpose operator. Performing substitutions for r_k and $\tilde{r}_k$

$$\sigma_{r,k}^2 = \xi\{(\mathbf{h}_k^T \mathbf{g}_k^{k-L+1} + n_k - \mathbf{h}_k^T \tilde{\mathbf{g}}_k^{k-L+1})(\mathbf{h}_k^T \mathbf{g}_k^{k-L+1} + n_k - \mathbf{h}_k^T \tilde{\mathbf{g}}_k^{k-L+1})^H \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$$

$$= \xi\{(\mathbf{h}_k^T (\mathbf{g}_k^{k-L+1} - \tilde{\mathbf{g}}_k^{k-L+1}) + n_k)(\mathbf{h}_k^T (\mathbf{g}_k^{k-L+1} - \tilde{\mathbf{g}}_k^{k-L+1}) + n_k)^H \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$$

$$= \mathbf{h}_k^T \xi\{(\mathbf{g}_k^{k-L+1} - \tilde{\mathbf{g}}_k^{k-L+1})(\mathbf{g}_k^{k-L+1} - \tilde{\mathbf{g}}_k^{k-L+1})^H \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}\mathbf{h}_k + \xi\{n_k n_k^* \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$$

$$= \mathbf{h}_k^T \boldsymbol{\Lambda}_k \mathbf{h}_k + \xi\{n_k n_k^*\}$$

$$= \mathbf{h}_k^T \boldsymbol{\Lambda}_k \mathbf{h}_k + N_0 / 2 \qquad \text{(A-11)}$$

The noise is uncorrelated with the fading and the fading estimate, and is also zero mean, therefore two terms disappear when multiplying out the second line. $\xi\{(\mathbf{g}_k^{k-L+1} - \tilde{\mathbf{g}}_k^{k-L+1})(\mathbf{g}_k^{k-L+1} - \tilde{\mathbf{g}}_k^{k-L+1})^H \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$ is recognized as the covariance matrix associated with the fading estimate $\boldsymbol{\Lambda}_k$. This result for the conditional variance of the received signal estimate corresponds with (3.20) as required.

Appendix C: Derivation of (3.25) and (3.26)

The solution to Yule–Walker equations (3.24) yields the prediction coefficients $\{p_{k,1}(\boldsymbol{\alpha}), p_{k,2}(\boldsymbol{\alpha}), \ldots, p_{k,L}(\boldsymbol{\alpha})\}$ to generate the fading estimate $\tilde{g}_k$ from the modified sufficient statistics $\{y_{k-1}, y_{k-2}, \ldots, y_{k-L}\}$

$$\tilde{g}_k = \sum_{m=1}^{L} p_{k,m}(\boldsymbol{\alpha}) y_{k-m} \tag{A-12}$$

By definition, the expected value of r_k given $\mathbf{r}_{k-1}^{k-L}$ and $\boldsymbol{\alpha}_k^{k-L}$ is

$$\tilde{r}_k \equiv \xi\{r_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \ . \tag{A-13}$$

It then follows by substitution

$$\begin{aligned}
\tilde{r}_k &= \xi\{\alpha_k g_k + n_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= \alpha_k \xi\{g_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} + \xi\{n_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= \alpha_k \tilde{g}_k + \xi\{n_k\} \\
&= \alpha_k \tilde{g}_k
\end{aligned} \tag{A-14}$$

The fact that the noise is zero mean was used, and $\xi\{g_k \mid \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$ is recognized as the fading prediction $\tilde{g}_k$ obtained using the Yule-Walker prediction coefficients. The

conditional mean corresponds with (3.25) as required. To find the variance of this prediction, the definition is stated

$$\sigma_{r,k}^2 \equiv \xi\{|r_k - \tilde{r}_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}. \tag{A-15}$$

Performing substitutions for r_k and $\tilde{r}_k$

$$\begin{aligned}
\sigma_{r,k}^2 &= \xi\{|\alpha_k g_k + n_k - \alpha_k \tilde{g}_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= \xi\{|\alpha_k (g_k - \tilde{g}_k) + n_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= |\alpha_k|^2 \, \xi\{|g_k - \tilde{g}_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} + \xi\{|n_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} \\
&= |\alpha_k|^2 \, \xi\{|g_k - \tilde{g}_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\} + \xi\{|n_k|^2\} \\
&= |\alpha_k|^2 \left[R_g(0) - \sum_{m=1}^{L} p_{k,m}(\boldsymbol{\alpha}) R_g(m) \right] + \frac{N_0}{2}
\end{aligned} \tag{A-16}$$

The knowledge that the fading and fading estimates are independent of the noise was used in the third step when multiplying out the square; also the noise is assumed to be zero mean. The term $\xi\{|g_k - \tilde{g}_k|^2 | \mathbf{r}_{k-1}^{k-L}, \boldsymbol{\alpha}_k^{k-L}\}$ in the fourth step is recognized as the variance in the error of the fading prediction supplied by the prediction filter. The error variance can be described in terms of the prediction coefficients $\{p_{k,1}(\boldsymbol{\alpha}), p_{k,2}(\boldsymbol{\alpha}), \ldots, p_{k,L}(\boldsymbol{\alpha})\}$ and autocorrelation of the fading $R_g(m)$ as written in (A-16). This result corresponds with (3.26) as required.

Bibliography

[1] B. D. O Anderson, J. B. Moore, Optimal Filtering, *Prentice-Hall*, Englewood Cliffs, NJ, 1979.

[2] F. Babich, G. Lombardi, "On verifying a first-order Markovian model for the mulit-threshold success-failure process for Rayleigh channel", *Proc. PIMRC '97*, Helsinki, Finland, pp. 12-16, Sept. 1997.

[3] F. Bacich, O. Kelly, G. Lomnardi, "Generalized Markov Modeling for Flat Fading", *IEEE Trans. on Commun.*, vol. 48, issue 4, pp. 547-551, April 2000.

[4] K.E. Baddour, N.C. Beaulieu, N.C., "Autoregressive models for fading channel simulation", *IEEE Global Telecommunications Conference, 2001. GLOBECOM '01*, pp. 1187-1192, November 2001.

[5] L.Bahl, J.Cocke, J.Raviv, F.Raviv, "Optimal Decoding of Linear Codes for minimising symbol error rate", *IEEE Transactions on Information Theory*, vol. IT-20, pp. 284-287, March 1974.

[6] R. Blahut, Theory Practice of Error Control Codes, *Addison-Wesley* , Reading, MA, 1983.

[7] J. K. Cavers, "The performance of phase-locked transparent tone-in-band with symmetric phase detection", *IEEE Trans. Commun.*, vol. 39 , pp. 1389-1399, Sept. 1991.

[8] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels", *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

[9] J. K. Cavers, M. Liao, "A comparison of pilot tone and pilot symbol techniques for digital mobile communications", *Globecom*, Orlando, FL, pp. 915-921, Dec. 1992.

[10] M. J. Chu, D. L. Goeckel, W. E. Stark, "On the design of Markov models for fading channels", *Proc. IEEE Vehicular Technology Conf. Fall*, Amsterdam, The Netherlands, Sept. 1999.

[11] A. P. Clark, S. F. Hau, "Adaptive adjustment of receiver for distorted digital signals", *IEE Proc.*, vol. 131, pp. 526-536, Aug. 1984.

[12] Q. Dai, E. Shwedyk, "Detection of bandlimited signals over frequency-selective Rayleigh fading channels", *IEEE Trans. Commun.*, vol. 42, Feb./Mar./Apr. 1994.

[13] Q. Dai, "Sequence Estimation for Rayleigh Fading Channels with Intersymbol Interference," *Ph. D Dissertation, University of Manitoba*, April 1993.

[14] M. Doelz, E. Heald, D. Martin, "Binary Date Transmission Techniques for Linear Systems," *Proc. IRE*, 656-661, May 1957.

[15] E. O. Elliot, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol 42, pp 1977-1997, Sept 1963.

[16] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE* vol. 61, pp268-278, Mar. 1973.

[17] G. J. Foschini; "A reduced state variant of maximum likelihood sequence detection attaining optimum performance for high signal-to-noise ratios", *IEEE Trans. Information Theory,* vol. IT-23, pp. 605-609, Sep 1977.

[18] R. G. Gallager, "Low-Density Parity-Check Codes", *MIT Press*, Cambridge, MA, 1962.

[19] M. J. Gertsman, J. H. Lodge, "Symbol-by-symbol MAP demodulation of CPM and PSK signals on Rayleigh flat-fading channels", *IEEE Trans. Commun.*, vol. 45, pp. 788-799, July 1997.

[20] E. N. Gilbert, "Capacity of a burst-noise channel", *Bell Syst. Tech. J.*, vol. 39, pp. 1253-1265, Sept. 1960.

[21] A. Goldsmith, "Joint source/channel coding for wireless channels", *Proc. IEEE Vehicular Technology Conf.*, pp. 614-618, July 1995.

[22] A. J. Goldsmith, P. Varaiya, "A decision-feedback maximum-likelihood decoder for finite state Markov channels", *Proc. 1994 IEEE Int. Conf. Commun.*, New Orleans, LA, pp. 1546-1550, 1994.

[23] A. J. Goldsmith, "Design and performance of high-speed communication systems over time-varying radio channels", *Ph. D Dissertation, U.C. Berkeley*, 1994.

[24] I. S. Gradshteyn I. M. Ryzhik. Table of Integrals, Series, and Products, *Academic Press*, San Diego, 1980.

[25] R. Haeb, H. Meyr, "A systematic approach to carrier recovery and detection of digitally phase modulated signals on fading channels", *IEEE Trans. Commun.*, vol. 37, pp. 748-754, July 1989.

[26] S. Haykin, Communication Systems, *John Wiley & Sons, Inc*, 1994.

[27] W. C. Jakes, Microwave Mobile Communications, *IEEE Press*, Piscataway, NJ, 1993.

[28] L. N. Kanal, A. Sastry, "Models for Channels with Memory and Their Applications to Error Control," *Proc. IEEE,* vol. 66, no. 6, July, 1978, pp. 724-744.

[29] S.A. Kassam, Q. Zhang, "Finite State Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1688–1692, Nov. 1999.

[30] H. Kong, "Receiver Structures for Wireless Mobile Channels," *Ph. D Dissertation, University of Manitoba*, Aug. 1997.

[31] H. Kong, E. Shwedyk, "A measure for the length of probabilistic dependence for Gaussian and lognormal distributed random processes", *Proc. Int. Symp. Information Theory*, Ulm, Germany, pp. 469-June 29–July 4 1997.

[32] H. Kong, E. Shwedyk, "Sequence detection and channel state estimation over finite state Markov channels" *IEEE Trans. on Vehicular Technology*, vol. 48, May 1999, pp. 833-839.

[33] H. Kong, E. Shwedyk, "On channel estimation and sequence detection of coded signals over frequency nonselective Rayleigh fading channels", *IEEE Trans. Veh. Technol.*, vol. 47, pp. 558-565, May 1998.

[34] H. Kong, E. Shwedyk, "Sequence estimation for frequency nonselective channels using a hidden Markov model", *Proc. IEEE 44th Veh. Technol. Conf.*, Stockholm, Sweden, pp. 1251-1253, June 1994.

[35] H. Kong, E. Shwedyk, "Adaptive MLSE of QAM signals over frequency nonselective Rayleigh fading channels", *Proc. Inst. Elect. Eng. Commun.*, vol. 145 pp 426-430 Dec 1998.

[36] Y. Linde, A. Buzo, R. M. Gray, "An algorithm for vector quantization design", *IEEE Trans. Commun.*, vol. 28, pp. 84-95, Jan. 1980

[37] D. Makrakis, P. Mathiopoulos, D. P. Bouras, "Optimal decoding of coded PSK and QAM signals in correlated fast fading channels and AWGN: a combined envelope, multiple differential, and coherent detection approach", *IEEE Trans. Commun.*, vol. 42, no.1, pp. 63-75, Jan. 1994.

[38] N. Nefedov, "Generative Markov Models for Discrete Channel Modelling," *Proc. IEEE PIMRC*, vol. 1, pp. 7-11, Sept. 1997.

[39] J. G. Proakis, Digital Communications, *McGraw-Hill*, New York, 1989.

[40] J. G. Proakis, Digital Signal Processing, *Prentice Hall*, New Jersey, 1996.

[41] L. R. Rabiner, B. H. Juang, "An introduction to hidden Markov models", *IEEE ASSP Magazine*, pp 4-16, January 1987.

[42] R. Raheli, A. Polydorus, C. K. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain environments", *IEEE Trans. Common.*, vol. 43, pp 354-364, Feb./Mar./Apr. 1995.

[43] L. L. Scharf, Statistical Signal Processing: Detection, Estimation, and Time Series Analysis, *Addison-Wesley*, Reading, MA, 1991.

[44] C. E. Shannon, "A mathematical theory of communication", *Bell Syst. Tech. J.*, vol. 27, pp. 379- 423, July 1948.

[45] J. P. Seymour, M. P. Fitz, "Near-optimal symbol-by-symbol detection schemes for flat rayleigh fading", *IEEE Trans. Commun.*, vol. 43 , pp. 1525-1533, Feb./Mar./Apr. 1995.

[46] K. Steiglitz, L.E. McBride, "A Technique for the Identification of Linear Systems," *IEEE Trans. Automatic Control*, pp. 461-464, vol. AC-10, 1965.

[47] G.L. Stuber, Principles of Mobile Communication, *Kluwer Academic Press*. 1996.

[48] F. Swarts, H. Ferreira, "Markov characterization of digital fading modile VHF channels", *IEEE Trans. Veh. Technol.,* vol. 43, pp. 977-985, Nov. 1994.

[49] C. Tan, N. Beaulieu, "On First-Order Markov Modeling for the Rayleigh Fading Channel", *IEEE Trans. on Commun.* Vol. 48, pp 2032-2040, Dec. 2000.

[50] W. Turin, R. van Nobelen, "Hidden Markov modeling of flat fading channels", *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1809-1817, Dec. 1998.

[51] W. Turin, "MAP symbol decoding in channels with error bursts", *IEEE Trans. Information Theory*, pp 1832-1838, vol. 47, issue 5, 2001.

[52] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets, Part II", *IEEE Commun. Mag.*, vol. 25, no.2, Feb. 1987.

[53] H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part I *John Wiley and Sons*, New York, 1968.

[54] H. S. Wang, P. Chang, "On Verifying the First Order Markovian Assumption for a Rayleigh Fading Channel Model," *IEEE Trans. on Vehicular Technology*, vol. 45, no. 2, pp. 353-357, May 1996.

[55] H. S. Wang, N. Moayeri, "Finite State Markov channel - A useful model for radio communication channels", *IEEE Trans. Veh. Technol.*, vol. 44, pp. 163-171, Feb. 1995.

[56] H. S. Wang, N. Moayeri, "Modeling, capacity, and joint source/channel coding for Rayleigh fading channels", *Proc. IEEE Vehicular Technology Conf.,* Secaucus, NJ, pp. 473-479, May 1993 .

[57] L. Wei, L. K. Rasmussen, "Near optimal tree-search detection schemes for bit synchronous multiuser CDMA systems over Gaussian and two-path Rayleigh fading channels", *IEEE Trans. Commun.*, vol. 45, pp. 691-700, June 1997.

[58] X. Yu, S. Pasupathy, "Innovations-based MLSE for Rayleigh fading channels", *IEEE Trans. Commun.*, vol. 43, pp. 1534-1544, Feb./Mar./Apr. 1995.